

Multiresolution Methods for Materials Modeling via Coarse-Graining

by

Ahmed E. Ismail

B.S., Yale University (1998)

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2005

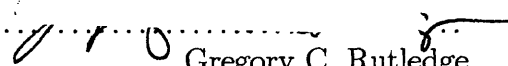
© Remove to use default copyright from the style

The author hereby grants to Massachusetts Institute of Technology permission to
reproduce and
to distribute copies of this thesis document in whole or in part.

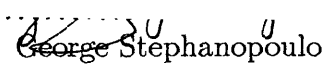
Signature of Author

Department of Chemical Engineering
29 October 2004

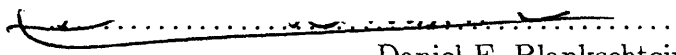
Certified by

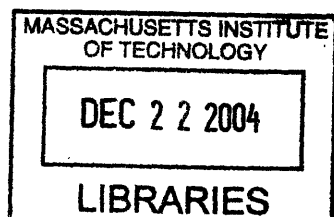

Gregory C. Rutledge
Professor of Chemical Engineering
Thesis Supervisor

Certified by


George Stephanopoulos
Arthur D. Little Professor of Chemical Engineering
Thesis Supervisor

Accepted by


Daniel E. Blankschtein
Professor of Chemical Engineering
Chairperson, Committee on Graduate Students



ARCHIVES

Multiresolution Methods for Materials Modeling via Coarse-Graining

by

Ahmed E. Ismail

Submitted to the Department of Chemical Engineering
on 29 October 2004, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Multiscale modeling of physical systems often requires the use of multiple types of simulations to bridge the various length scales that need to be considered: for example, a density-functional theory at the electronic scale will be combined with a molecular-dynamics simulation at the atomistic level, and with a finite-element method at the macroscopic level. An improvement to this scheme would be a method which is capable of consistently simulating a system at multiple levels of resolution without passing from one simulation type to another, so that different simulations can be studied at a common length scale by appropriate coarse-graining or refinement of a given model.

We introduce the wavelet transform as the basis for a new coarse-graining framework. A family of orthonormal basis, the wavelet transform separates data sets, such as spatial coordinates or signal strengths, into subsets representing local averages and local differences. The wavelet transform has several desirable properties for coarse-graining: it is hierarchical, compact, and has natural applications to approximating physical data sets. As a hierarchical method, it can be used to rescale a Hamiltonian to a desired length scale, and at the same time also rescales the particles of the system by creating “blocked” particles in the spirit of renormalization group (RG) calculations. The wavelet-accelerated Monte Carlo (WAMC) framework performs a Monte Carlo simulations on a small system which will be transformed into a block particle to obtain the probability distribution of the blocked particle; a Monte Carlo simulation is then performed on the resulting system of blocked particles. This method, which can be repeated as needed, can achieve significant speed-ups in computational time, while obtaining useful information about the thermodynamic behavior of the system.

We show how statistical mechanics can be formulated using the wavelet transform as a coarse-graining technique. For small systems in which exact enumerations of all states is possible, we illustrate how the method recovers reasonably good estimates for physical properties (errors no more than 10%) with several orders of magnitude fewer operations than are required for an exact enumeration. In addition, we illustrate that errors introduced by the wavelet transform vanish in the neighborhood of fixed points of systems as determined by RG theory.

Using scaling results from simulations at different length scales, we estimate the thermodynamic behavior of the original system without performing simulations on the full original system. In addition, we make the method adaptive by using fluctuation properties of the system to set criteria under which further coarse graining or refinement of the system is required. We demonstrate our method for the Ising universality class of problems.

We also examine the applicability of the WAMC framework to polymer chains. Polymers are quintessential examples of the need for simulations at multiple scales: at one end, we can study short chains using quantum chemistry methods; yet polymers can have relaxation times on the order of seconds or longer, and molecular weights of 10^6 or more. Even with modern computational resources, simulating behavior at long times or for long chains is still prohibitively expensive. While many approaches have been developed for studying such systems, many of these are specific to particular polymer chemistries, or fundamentally change the basic model of the system on an *ad hoc* basis. We also demonstrate how the WAMC framework can be adapted to study coarse-grained polymer chains represented as interacting lattice and off-lattice random walks.

These walks can incorporate many of the same interactions as traditional “off-lattice” polymer models: excluded volume, stiffness, and non-bonded pair interactions. Coarse-graining the chain using the wavelet transform leads to each segment of the chain being replaced by a bead located at the center of mass of the segment. Interactions along the contour—such as stiffness potentials—are directly handled as well and incorporated as an internal configuration energy. Non-bonded interactions, such as excluded volume and non-bonded pair interactions, must be handled differently; we discuss possible approaches for handling these terms hierarchically within the WAMC framework. We present the details of the implementation of this algorithm, its performance for basic thermodynamic properties, as well as its connections to other effective coarse-grained models such as freely-jointed and Gaussian chains.

In the development of our coarse-grained models, we have also discovered that the coarse-grained degrees of freedom—bond lengths, bond angles, and torsion angles—have distributions which are much more complicated than are typically employed in coarse-grained simulation techniques. This coupling of behavior is observed even when the model studied is as simple as a freely-jointed chain. In addition, Monte Carlo simulations have allowed us to establish the existence of numerical scaling laws for the overlap probabilities which we invert to determine the intra- and intermolecular potentials as a function of the number of repeat units as well as the ratio of repeat unit (or bead) size to the bond length. These results are compared to the results obtained from analytical derivations based on the freely-jointed chain which show qualitative agreement between the two approaches. Consequently, we can use the potentials determined by a simulation performed at one resolution of coarse-graining and derive from the scaling laws new potentials which describe the behavior at another resolution. This allows us to “tune” the WAMC algorithm to obtain results more efficiently than would be possible with algorithms that operate at fixed levels of coarse-grained resolution.

Thesis Supervisor: Gregory C. Rutledge
Title: Professor of Chemical Engineering

Thesis Supervisor: George Stephanopoulos
Title: Arthur D. Little Professor of Chemical Engineering

Acknowledgments

*Darest thou now, O Soul,
walk out with me toward the unknown region?*

—Walt Whitman, “Toward the Unknown Region”

Scientific discovery remains one of the last great mysteries of our time. Even with our increasingly interconnected, globalized society, we still have trouble answering fundamental questions of science. Even where we know in principle how to solve particular problems, the actual execution of such solutions is far beyond the realm of the currently possible, or at least the currently possible in the span of a single career. Moreover, the bar of what constitutes “novel” scientific research has been raised ever higher over time. What once sufficed for a graduate thesis—for example, the one-dimensional Ising model—now serves as a homework exercise in an undergraduate physics class. One can hardly imagine a topic like coarse-graining becoming a homework problem for some unsuspecting twenty-second-century engineering major, yet who knows what is possible a hundred years hence?

However, it is not time to look forward into a crystal ball, but instead a chance to take a look back. Yet, like most people, I find the challenge of appropriately acknowledging everyone who has had an important role to play most daunting. Even if the actual creative act of scientific research is solitary, the career of a researcher must be anything but. It is little surprise that

the list of people to be thanked will be quite long and almost surely quite incomplete as well. I hope anyone I've left off will accept my gratitude and thanks.

The natural place to start is with my thesis advisors, Professors Stephanopoulos and Rutledge. Their random suggestion of a project, proposed in the 10.990 seminar, seemed to this first-year the most interesting of all of the projects being offered. Since then, their support and encouragement in my endeavors—as well as patience with my exceedingly elliptical trains of thought—has been a tremendous benefit for me. [Of course, with the case of co-advisors comes the all-important question: which one to thank first? I decided the same way they decided who would be my “official” advisor—with a coin flip!]

The Stephanopoulos and Rutledge research groups also provided support, guidance, and convenient sounding boards for off-the-wall ideas and theories over the years. In particular, I'd like to thank Matthew Dyer, Orhan Karshgil, and Bill Schmitt in the Stephanopoulos group, and Numan Waheed, Mark Bathe, Marc Lavine, Pieter in 't Veld, and Shawn Mowry in the Rutledge group.

I would also like to thank the other members of my thesis committee, Professors Blankschtein and Trout, for their many useful comments and insights during thesis committee meetings, outside of them—particularly in the protracted process of finding a good post-doctoral appointment. They also gave me my first opportunity at being a teaching assistant—shepherding first-year graduate students in 2001 through 10.40. That opportunity was an enormous challenge, but simultaneously a most rewarding one. It also gave me the opportunity to come know an entire group of graduate students well—perhaps even a little too well! My more recent experience in 10.10, with Professors Stephanopoulos and Mohr, has been quite a different view of education—teaching undergraduates poses challenges quite unlike that of a graduate course, but is perhaps even the more rewarding in the long run. [I would be remiss not to also thank Henry and Kyle, without whose work as co-TA's would have made my job an impossible task.]

Of course, no university runs without its support staff, and the MIT Chemical Engineering Department has been blessed with a dedicated and competent cadre. Without the help of (in chronological order) Gabrielle, Liz, Jung, Tim, Greg, Susan, and Arline, a lot of minor bumps and hurdles could have become major administrative hassles. [Without Arline's bottomless supply of M&M's, I might also have been less hyper, but that's another story altogether.] The

Graduate Student Office has also been a major resource for our department, and its staff—Janet and Elaine, Suzanne and Jenn and Mary and Annie (and others)—are among the department’s other unsung heroes.

There have also been a number of other Course X graduate students who have been kindred spirits in this journey. The “power office” crew of 66-256, plus a few other brave souls from 66-264, made the first-year bonding experience an enjoyable one (even when frustration led us to generate the same results as an Excel spreadsheet by hand!). In later years, the Movie Night crew, plus other hooligans and miscreants from my 10.40 class (and more recent arrivals) have succeeded in injecting the occasional dose of levity and humor when desperately needed. My support circle also included a number of people from my days at Yale, plus from numerous random circumstances (Ben M., Nicola C., Ezra M., Mike H-L, Josh L., and Jay H., to name just a few).

Of course, the other passion that I’ve pursued in the rare hours off from engineering is singing, and that has also been a useful outlet and retreat for a busy engineer. Singing with the MIT Chamber Chorus and Concert Choir, with various summer groups, and as part of the Emerson Scholarship program has been a great joy for the last six-odd years. Special thanks go in particular to Pamela Wood and Margaret O’Keefe, who always challenged me to do what I never thought I could, yet always had in me. There aren’t many places where someone could go for a world-class education in both engineering and music; I’ve been privileged to find two such places in my educational career.

Research doesn’t happen in a vacuum; more importantly, it does not happen for free. I have been fortunate to have fellowship support for the vast majority of my education: the DOD National Defense Science and Engineering Graduate Fellowship (1998-2000), and the DOE Computational Sciences Graduate Fellowship (2001-2004). The latter program has been perhaps the most useful of my entire educational career, introducing me to a vast array of people with a similar interest in computational science, but with a vast array of different tools and different applications in mind. The CSGF fellows (and now alumni) are fast becoming a community whose impact will surely grow with time, and it is my good fortune to have been a part of this program for the last four years.

Finally, the unflagging, unstinting, and unending support of my parents has sustained me through this long, exciting, bewildering, perplexing, illuminating, and fascinating period of my life and career. Without their kindness and example, I wouldn't be who I am today. For that, they deserve above all my greatest love and admiration.

—Ahmed E. Ismail
Cambridge, MA
29 October 2004

P. S. At the time I had finished this thesis, who could have ever thought it possible that I would have to worry that a parade celebrating a Red Sox World Series championship would conflict with my defense date?

Contents

1	Introduction	23
1.1	Motivations	23
1.2	Coarse-graining	25
1.2.1	Substitution methods	25
1.2.2	Multiscale methods	27
1.2.3	Multiresolution methods	28
1.3	Thesis objectives and overview	29
2	Fundamental models and tools	31
2.1	Lattice theory	31
2.2	Physical models of polymer chains	33
2.2.1	Discrete and continuum “limiting” descriptions	33
2.2.2	Interacting chain	34
2.3	Wavelet transform fundamentals	36
2.3.1	The conceptual picture	36
2.3.2	Properties and examples of wavelet and scaling functions	38
2.3.3	Matrix formulation of the wavelet transform	40
2.3.4	Multidimensional wavelet transforms	42
2.4	Renormalization group principles	44
2.4.1	Renormalization group analysis of lattice systems	45
2.4.2	Renormalization group flows	46
2.4.3	Connections to our coarse-graining methodology	47

3	Multiresolution analysis in statistical mechanics. I. Using wavelets to calculate thermodynamic properties	49
3.1	Introduction	49
3.2	Wavelet analysis of lattice thermodynamics	51
3.2.1	Applying the wavelet transform to lattice Hamiltonians	51
3.2.2	Computing thermodynamic functions	52
3.3	Analysis	53
3.3.1	Weighting functions for wavelet-transformed statistics	54
3.3.2	Order parameter	55
3.3.3	Free-energy considerations	56
3.3.4	Entropy of a wavelet-transformed lattice system	59
3.3.5	Fluctuation properties	61
3.4	Similarities to Renormalization Group theories	64
3.5	Conclusions	65
4	Multiresolution analysis in statistical mechanics. II. The wavelet transform as a basis for Monte Carlo simulations on lattices	67
4.1	Introduction	67
4.2	The Wavelet-Accelerated Monte Carlo (WAMC) algorithm	68
4.3	Theoretical performance of WAMC versus traditional MC	71
4.4	Results	72
4.4.1	Order parameter	72
4.4.2	Internal energy	73
4.4.3	Fluctuation properties	74
4.4.4	Coarse-grained entropy	75
4.4.5	Scaling results	77
4.4.6	Decorrelation time	80
4.5	Analysis	83
4.5.1	Measured performance comparison	83
4.5.2	Comparison with renormalization group methods	84
4.5.3	Sources of error	85

4.5.4	Constructing an adaptive algorithm for MC using the wavelet transform	86
4.6	Conclusions	87
5	Topological coarse-graining of polymer chains. I. Freely-jointed chains	90
5.1	The freely-jointed chain model	91
5.2	Coarse-graining of polymer chains	92
5.2.1	Using wavelets to construct a coarse-grained model	92
5.2.2	“Equivalence” of on-lattice and off-lattice coarse-grained computations	93
5.3	Wavelet-accelerated Monte Carlo (WAMC) simulation of polymer chains	94
5.3.1	Hierarchical simulation approach	94
5.3.2	Coarse-grained simulation algorithm	96
5.4	Probability distributions for coarse-grained internal coordinates	97
5.4.1	Bond-length distribution	98
5.4.2	Bond-angle distribution	99
5.4.3	Torsion-angle distribution	102
5.5	Results	103
5.6	Conclusions	104
6	Topological coarse-graining of polymer chains. II. Self-avoiding chains	106
6.1	Self-avoiding chains	107
6.1.1	Definition	107
6.1.2	Simulating self-avoiding chains	107
6.1.3	Demonstration of $O(N)$ decorrelation time between independent configurations	109
6.2	Wavelet-accelerated Monte Carlo (WAMC) simulation of polymer chains	110
6.2.1	Coarse-grained simulation algorithm	110
6.2.2	Probability distributions for coarse-grained internal coordinates	112
6.3	Results	114
6.3.1	Atomistic simulations	114
6.3.2	WAMC simulations	116
6.3.3	Effects of additional interactions	121

6.3.4	Error tolerance and the $N \rightarrow \infty$ critical point	122
6.4	Performance	125
6.4.1	Decorrelation of samples	127
6.5	Geometric reverse-mapping of a polymer random walk	129
6.5.1	Description of the constraints	129
6.5.2	Implementation of the algorithm	131
6.5.3	Pathological cases	135
6.5.4	Handling the endpoints	137
6.5.5	Illustrative examples	138
6.5.6	Using the fine-grained representation	138
6.6	Conclusions	139
7	Coarse-grained potentials for polymer chains	142
7.1	Introduction	142
7.2	Definition of the problems to be examined	143
7.2.1	Transition to “soft” potentials	143
7.2.2	Overlap Probabilities	145
7.3	Examples of scaling behavior	146
7.4	Wavelet transform of the hard-sphere potential	152
7.5	Constructing coarse-grained potentials	153
7.6	Analytical solution: Freely-jointed chains	154
7.6.1	Invalidity of “maximum-term” approach	157
7.6.2	Asymptotic behavior of Equation (7.22)	161
7.7	The probability of overlap for two segments of a freely-jointed chain	162
7.8	Potentials derived from $P(\mathbf{u}, r_{test})$	167
7.9	Self-avoiding walk	171
7.10	Alternative approaches for computing new coarse-grained potentials	172
7.11	Conclusions	176
8	Coupling of variables as a consequence of coarse-graining	177
8.1	Discovering the coupling phenomenon	177

8.2	Should these results surprise us?	186
8.3	Information gain in coarse-grained systems	190
8.3.1	Polymer models	190
8.3.2	Ising lattices	192
8.4	When should we expect interdependence?	193
9	Future research directions and conclusions	196
9.1	Future research directions	196
9.2	Conclusions: Fundamental insights on coarse-graining	199
A	Center-of-mass bond statistics for a Gaussian chain	202
B	Converting internal coordinates	206
	Bibliography	208

List of Figures

2-1 (a) A sample signal $u(n)$. (b, c) The one-dimensional Haar scaling function $\phi(x)$ and wavelet (differencing) function $\psi(x)$. (d, e) The first-level scaling coefficients $s(n)$ and $\delta(n)$ produced from the original signal $u(n)$ using the Haar pair, following downsampling of the signal. 39

2-2 The two-dimensional orthogonal Haar wavelets. The coefficient in all cases is $1/2$, times the sign indicated in each quadrant. Note that the scaling function ϕ is built from the one-dimensional Haar scaling function, while the wavelet functions are built from the one-dimensional wavelet function. 43

2-3 Example of renormalization group flows, as applied to the two-dimensional Ising model. 47

3-1 (a) Average absolute magnetization $|m| = \sum_i |\sigma_i|$ of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dot-dashed and dotted lines, respectively). (b) Error in the average absolute magnetization of the 4×4 Ising model for one and two iterations of the two-dimensional Haar wavelet versus an exact enumeration. 57

3-2 Free energy A of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively). 58

3-3	Entropy of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively). The bottom of the y -axis corresponds to the zero-temperature limit of $S = \ln 2$	60
3-4	Heat capacity of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively).	62
3-5	Heat capacity of a 32×32 Ising model, illustrating the power-law divergence in the vicinity of the critical point.	63
4-1	Absolute average magnetization as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a $(4, 8)$ -WAMC simulation (center), and a $(8, 4)$ -WAMC simulation (right).	74
4-2	Internal energy as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a $(4, 8)$ -WAMC simulation (center), and a $(8, 4)$ -WAMC simulation (right).	75
4-3	Heat capacity as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a $(4, 8)$ -WAMC simulation (center), and a $(8, 4)$ -WAMC simulation (right).	76
4-4	Entropy of the 32×32 Ising model as a function of temperature, as computed by direct simulation (top) and using an $(8, 4)$ -WAMC simulation (bottom).	78
4-5	Distribution of the magnetization of states in an 8×8 subblock in the limit $T \rightarrow \infty$, as predicted by analytical arguments (line) and by direct simulation (histogram).	79
4-6	Graph showing variance in the estimate of energy as calculated using the method of Flyvbjerg and Petersen [60] for a 32×32 Ising model measured at its critical temperature.	81

4-7	Graph showing variance in the estimate of energy as calculated using the method of Flyvbjerg and Petersen [60] for a 32×32 Ising model in a (16, 2)-WAMC measured at the critical temperature determined from the simulation.	82
4-8	Phase diagram plotting average magnetization versus temperature and external field strength for a 64×64 ferromagnetic Ising lattice, computed using an (8, 8)-model via WAMC. The general features correspond to those that would be produced with standard MMC, but require less than 3 per cent of the computational time.	84
4-9	Phase diagram plotting average magnetization versus temperature for a 64×64 ferromagnetic Ising lattice, created using an adaptive WAMC algorithm, with refinement and coarsening criterion established using the change in heat capacity with respect to temperature $\Delta C_H/\Delta T$. The squares represent the simulation results, while the line reproduces Onsager's analytical result for the two-dimensional Ising model with zero external field. The simulation used in a given temperature region is shown on the plot.	88
5-1	Coarse-graining of a 16-site random walk in two dimensions using "center-of-mass aggregation" of adjacent points along the chain, as suggested by the wavelet transform method. The four sites which would be created after another iteration of this method are indicated on the graph; note that each is at a quarter-integer lattice point.	94
5-2	Coarse-graining of a three dimensional walk on a cubic lattice using the wavelet transform. (a) A 512-step self-avoiding random walk. (b)-(d) The same walk, after one, two, and three iterations of the Haar transform defined by (7.6) and (7.7).	95
5-3	Comparison of the probability distributions for the coarse-grained bond length (distance between adjacent center-of-mass beads) for freely-jointed chains for different chain lengths, with 32 atomistic beads per coarse-grained bead.	98
5-4	Ideal and observed coarse-grained cumulative distribution functions for a coarse-grained Gaussian random walk divided into segments of 32 beads each.	100

5-5	Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a freely-jointed chain with $N = 128$ and $N_e = 32$, as a function of r_1 , the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal Gaussian value of $r_1 = \sqrt{N_e} = \sqrt{32}$; the middle, dashed curve shows $\sqrt{N_e} < r_1 < \frac{3}{2}\sqrt{N_e}$; and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}$	101
5-6	CDF’s for the coarse-grained bond angle computed by simulation (solid) and using a “stretched Gaussian” approximation (dotted). The two curves are largely indistinguishable.	102
5-7	Probability distribution function for the torsion angle of a freely-jointed chain of length $N = 128$ considered as four beads of effective size $N_e = 32$	103
5-8	The mean square radius of gyration $\langle R_g^2 \rangle$ as a function of bead size N for both atomistic and WAMC representations of the freely-jointed chain.	105
6-1	Configurations of a 1024-step chain taken 1024 attempted pivots apart, showing the decorrelation between samples after $O(N)$ attempted pivots.	110
6-2	Autocorrelation function of the end-to-end distance for a self-avoiding walk of length $N = 1024$	111
6-3	Probability distributions for the distance between adjacent center-of-mass beads for self-avoiding chains (bottom), parametrized by total chain length.	113
6-4	Mean bond length connecting 32-mers as a function of the total length N for self-avoiding walks.	113
6-5	Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a self-avoiding chain with $N = 128$ and $N_e = 32$, as a function of r_1 , the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal value of $r_1 = \sqrt{N_e} = \sqrt{32}$ for the freely-jointed chain; the middle, dashed curve shows $\sqrt{N_e} < r_1 < \frac{3}{2}\sqrt{N_e}$; and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}$	115
6-6	Probability distribution function for the torsion angle of a chain of length $N = 128$ considered as four beads of effective size $N_e = 32$	116

6-7	Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (circles) and of coarse-grained simulations based on bead sizes of 32 and 64 (triangles and squares, respectively).	117
6-8	Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (diamonds) and of coarse-grained simulations based on bead sizes of 32 and 64 (squares and circles, respectively).	119
6-9	Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (top line) and of a four-scale coarse-grained simulations with bead sizes of 1, 128, 2048, and 65536 per stage. The observed Flory exponent is $\nu = 0.568$. Note that the end-to-end distance for the pivot algorithm is extrapolated from the behavior for $N \leq 16384$ using the correct scaling exponent $\nu \approx 0.577$	120
6-10	Demonstration of the end effect on the structural properties of an ideal Gaussian chain. The ratio $\langle R_g^2 \rangle_{cg} / \langle R_g^2 \rangle_{exact}$ was computed for different numbers of coarse-grained beads and for chains of length $N = 512$ and $N = 1024$. The comparison shows that the magnitude of the error is proportional to $1/N_b$, where N_b is the number of coarse-grained beads used.	121
6-11	Probability distribution for the distance between adjacent centers of mass of 16-bead segments for random walks of length $N = 1024$	123
6-12	Mean square end-to-end distance $\langle R_{256}^2 \rangle$ for a chain of length 256 as a function of the stiffness potential ϵ_s	124
6-13	Running time comparison for atomistic simulation (squares), optimized atomistic simulation (diamonds), and WAMC algorithm with $N_e = 32$ (crosses) and $N_e = 64$ (asterisks).	126
6-14	Running time per move of the pivot algorithm (upper curve) and of a four-scale coarse-grained algorithm with bead sizes of $N = (1, 128, 2048, 65536)$ per stage (lower curve).	128
6-15	The configuration of a coarse-grained polymer chain with $N_b = 256$ taken 256 steps apart.	129
6-16	Autocorrelation function for a the end-to-end distance of a self-avoiding walk studied using WAMC with $N_b = 256$ and $N_e = 64$	130

6-17	Reverse-mapping of a coarse-grained 64-mer to a fine-grained 128-mer. The coarse-grained chain is indicated by circles connected with dashed lines; the fine-grained chain is the solid curve.	139
6-18	Reverse-mapping of a coarse-grained 64-mer to a fine-grained 128-mer. The coarse-grained chain is indicated by circles connected with dashed lines; the fine-grained chain is the solid curve.	140
6-19	Contour length of the fine-grained 128-mer as a function of the parameter λ defined in (6.9) for various starting coarse-grained configurations.	141
7-1	(Top) Semi-log plot of the two-body correlation function $g(r_{ij})$ as a function of the interparticle separation distance r_{ij} . (Bottom) Semilog plot for small values of r_{ij} showing the sharp cutoff in $g(r_{ij})$ for $r_{ij} \leq 1$	144
7-2	Scaled probability of overlap for self-avoiding lattice chains of length $N = 2048$ with incorporated excluded-volume and stiffness interactions.	147
7-3	Scaled probability distribution for freely-jointed chains of varying lengths with equivalent numbers N_s of coarse-grained segments.	148
7-4	Overlap probability as a function of the test radius r_{test} for freely-jointed chains of length $N = 512$	149
7-5	Unscaled probability distribution $F_N(x, r)$ for finding a particle within a test volume of radius r a distance x away from the center of mass of a chain of length N	150
7-6	A scaled version of Figure 4-8, illustrating the $xN^{-1/2}$ scaling of the distance away from the center of mass and $r^3N^{-3/2}$ scaling of the probability.	151
7-7	Estimate for the true interatomic potential $u(r)$ used to improve the WAMC self-avoiding walk model. The potential is shown for coarse-grained beads representing 32 atomistic beads of diameter 1. $R_g \approx 7.1$ for the 32-mer.	155
7-8	The probability distribution function $F_{64}(\mathbf{x}, 1)$ as given by the true formula (7.22) [upper curve] and the approximation (7.25) [lower curve].	159
7-9	Estimate $P_{64}(\mathbf{x}, 1)$ for the probability distribution function for two-chain collisions, as given by the true formula (7.22) [upper curve] and the approximation (7.25) [lower curve].	160

7-10	Unscaled probability distribution $F_N(x, r)$ for finding a particle within a test volume of radius r a distance x away from the center of mass of a chain of length N . From top to bottom, $r_{test} = 1, 0.5, 0.25,$ and 0.125	162
7-11	A scaled version of Figure 7-10, illustrating the $xN^{-1/2}$ scaling of the distance away from the center of mass and $r^3N^{-3/2}$ scaling of the probability.	163
7-12	Dependence of overlap size on r_{test} : overlap between 32-mers with test-volume sizes $r_{test} = 1.0, r_{test} = 0.5,$ and $r_{test} = 0.25$ as a function of the separation of centers of mass, as obtained via direct simulation [solid curves] and approximation (7.32), with F_N given by (7.22) [dashed curves].	165
7-13	Dependence of overlap probability on r_{test} : overlap between 128-mers with test-volume size $r_{test} = 1.0$ and $r_{test} = 0.5$ as a function of the separation of centers of mass, as obtained via direct simulation and approximation (7.32).	166
7-14	Scaled potential showing the r_{test}^3 dependence of the potential for two freely-jointed chains of 32-mers in the limit $r_{test} \rightarrow 0$ using approximation (7.32).	167
7-15	Scaled potentials showing the r_{test}^3 dependence of the potential for overlap of two freely-jointed chains of 32-mers determined by direct simulation.	168
7-16	Scaling of the two-body potential created by approximation (7.32), parametrized by increasing bond length ($N = 32$ is the bottom curve; $N = 1024$ is the top curve), demonstrating convergence towards a fixed curve as $N \rightarrow \infty$	169
7-17	Scaling of the two-body potential created via direct simulation, parametrized by increasing bond length ($N = 32, 64,$ and 128 from bottom to top, respectively).	170
7-18	The coarse-grained overlap potential for two self-avoiding chain segments of length 32, parametrized as a function of the test-volume size r_{test} , demonstrating the r_{test}^3 -dependence of $U(r)$ as $r_{test} \rightarrow 0$	172
7-19	Collapse of coarse-grained overlap potentials for two self-avoiding chains with $N = 32, 64,$ or $128,$ and test-volume size $r_{test} = 0.5,$ showing an $N^{0.4} \approx N^{1-\nu}$ dependence.	173

7-20	Collapse of coarse-grained overlap potential for two self-avoiding chains with $N = 32, 64,$ or $128,$ and test-volume size $r_{test} = 0.125,$ showing an $N^{0.4} \approx N^{1-\nu}$ dependence for the range of the potential and an $N^{0.5}$ dependence on the magnitude.	174
8-1	Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (top line) and of coarse-grained simulations based on bead sizes of $32, 64,$ and 128 (lower curves). Note the excellent agreement of the different coarse-grained curves with each other. . .	178
8-2	Estimate for the true interatomic potential $u(r)$ used to improve the WAMC self-avoiding walk model.	179
8-3	A correlation plot of adjacent coarse-grained bond lengths in a 256-mer, showing decorrelation between adjacent bonds connecting beads of effective size $64.$. . .	180
8-4	A correlation plot of coarse-grained bond lengths and adjacent bond angles in a 256-mer, allegedly showing decorrelation between bonds connecting beads of effective size 64 and their neighboring bond angles.	181
8-5	Probability distribution function $p(\phi)$ (top) and cumulative distribution function (bottom) $P(\phi)$ for the coarse-grained bond angle between successive coarse-grained bonds, as computed for beads of size 32 in a freely-jointed chain model. The lower curve also shows the stretched Gaussian $p(\phi, 2.5)$ defined by (8.3). . .	184
8-6	Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a Gaussian chain with $N = 128$ and $N_e = 32,$ as a function of $r_1,$ the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal Gaussian value of $r_1 = \sqrt{N_e} = \sqrt{32};$ the middle, dashed curve shows $r_1 < \frac{3}{2}\sqrt{N_e};$ and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}.$	185
8-7	Probability distribution function for the torsion angle of a chain of length $N = 128$ considered as four beads of effective size $N_e = 32.$	187
8-8	The mean square radius of gyration $\langle R_g^2 \rangle$ as a function of bead size N for both atomistic and WAMC representations of the freely-jointed chain.	188

List of Tables

4.1	Performance comparison for MMC versus WAMC	83
5.1	Mean coarse-grained bond-lengths between beads representing 32-mers	99
5.2	Mean square radius of gyration using bond angle and torsion angle distributions	104
6.1	Mean end-to-end distance for atomistic versus WAMC algorithms	117
6.2	Radial statistics for self-avoiding walks with different potentials	122
8.1	Mean square radius of gyration using a single bond angle distribution	183
8.2	Mean square radius of gyration using multiple bond angle distributions	183
8.3	Mean square radius of gyration using bond angle and torsion angle distributions	186
8.4	Correlation parameter for different pairs of coarse-grained variables	190

Chapter 1

Introduction

1.1 Motivations

The rapid advances in available computational forces in recent years has improved our ability to study physical phenomena using numerical simulations. However, it is just as clear that the set of problems that we can currently address is a small subset of the problems of scientific and engineering interest. This mismatch between available resources and scientific needs becomes apparent when we note that atomistic techniques like molecular dynamics are currently capable of accessing time scales only on the order of nanoseconds and microseconds, while many chemical and physical processes—such as reactions and diffusion—occur on the scale of milliseconds and seconds. Similarly, unless one is working on a modern supercomputer, molecular dynamics simulations are usually limited to thousands of atoms, a substantial limitation when one considers that individual polymer chains can have degrees of polymerization on the order of hundreds of thousands, and that many processes of interest require dense melts containing hundreds of chains.

Numerous schemes, collectively known as *coarse-graining* methods, have been proposed and developed for dealing with the problem of “scale-bridging.” These techniques, however, generally have significant shortcomings that limit their application to specific fields or sometimes even to specific molecular chemistries. Consequently, in spite of the ability of these methods to address the systems for which they have been developed, a method which is capable of handling a variety of different molecular systems would be of great practical utility. With a unified method in hand,

we could address numerous kinds of systems with a clear, coherent methodology. In addition, the greater the amount of structure inherent in the method, the less it will be necessary to rely on *ad hoc* assumptions to construct our coarse-grained models.

In addition, we would like to understand better the coarse-graining process. As mentioned before, many different approaches have been developed, but many of the fundamentals underlying coarse-graining have remained largely unexplored. That is, researchers have shown how to coarse-grain physical systems, but they have not demonstrated what effects the coarse-graining method they have chosen has had on the physics of their new model. For example, questions such as the following have been left largely unanswered in the coarse-graining literature:

- What is the “optimal” amount of coarse graining to perform?
- How closely connected are the variables which survive the coarse-graining procedure?
- How do we determine when our coarse-graining has gone too far—or when we can coarse-grain further?
- Why favor one coarse-graining methodology over another?

Ideally, these questions could be addressed in sufficient detail that we would then have a “cook-book”: given the specifics of a particular problem, we would look up the specific coarse-graining technique and follow the appropriate “recipe.” However, the difficulty inherent in adequately answering these and other such basic questions makes the likelihood of codifying coarse-graining to such an extent unlikely.

The research described in this thesis develops a new, systematic and hierarchical coarse-graining methodology. To demonstrate the breadth of its applicability, we have applied it to two fundamental models of materials science: the Ising ferromagnet, and the polymer random walk. These models have been selected because of their general applicability to materials science, and for the richness of their physical behavior, as we shall see in greater detail below. In addition to comparing our simulation results to the well-established results for both of these models, we offer answers to the questions outlined above as they have arisen in the context of our research.

1.2 Coarse-graining

As mentioned above, the term *coarse-graining* has been applied to a wide variety of techniques, all of which address the problem of “bridging” time and length scales in molecular simulations. The general principle underlying these techniques is that we can replace expensive atomistic simulations with simpler simulations which, either singly or in series, will yield the same results as the more complicated atomistic model, but at much lower computational cost. We can divide the set of coarse-graining approaches broadly into three main subsets: substitution methods; handshaking methods, also known as multiscale modeling; and multiresolution coarse-graining.

1.2.1 Substitution methods

The substitution method is perhaps the most straightforward of the three approaches. As its name implies, the detailed atomistic model is replaced with a single, simpler model whose effective behavior is the same as the original model. The amount of coarse-graining can differ, as can the final model that is being computed. For example, within the realm of polymer physics, the range of possible substitutions runs from individual repeat groups all the way up to entire chains. We briefly outline some particularly notable approaches.

Kremer and co-workers [1, 75, 156, 157] have created an approach for polycarbonates which replaces a repeat unit with a united atom and have outlined a method for reverse-mapping this process, so that the original carbonate groups can be restored following a simulation of the coarse-grained representation. While similar to other substitution-based approaches [50], they have also developed a reverse-mapping procedure which allows them to convert their coarse-grained representation back to an atomic one. To date, however, the reverse mapping procedure has only been studied for the special case of bisphenol carbonates and their derivatives [10, 157].

Several other methods of interest for studying polymers map chains onto a lattice: two of these are the bond fluctuation method and the “high-coordination lattice” approach. In the former, a chain is replaced by a series of bonds which connect the center of plaquettes on a simple square or cubic lattice [26, 116, 117, 125, 132, 162]. Only certain bond lengths are permitted, thereby constraining the allowed geometries within a given chain. New configurations are created by adjusting which plaquettes are used to define the chain, according to the geometric bond

length constraints. Similarly, the “high-coordination lattice,” or “diamond lattice,” approach defines a chain as a series of second nearest neighbors on a tetrahedral lattice [43, 44, 76, 129]. Since the tetrahedral lattice has twelve nearest neighbors instead of the six found in a simple cubic lattice, many more possible configurations are available; these configurations are accessed by adjusting the position of bonds on the lattice or using a reptation-like approach. The resulting simulations have been used in conjunction with rotational isomeric state (RIS) theory in the study of melts.

The aforementioned schemes generally map one repeat unit onto one coarse-grained particle. However, it is also possible to map an entire chain onto a single coarse-grained particle, as is done in so-called ellipsoid models. These models replace a polymer chain with a single ellipsoid whose position is given by the center of mass of the original chain and whose size is determined by the eigenvalues of the chain’s radius of gyration tensor. Interpenetration of the ellipsoids representing different chains is permitted, so the hard-sphere potential describing non-bonded interactions is replaced by soft potentials. Different implementations of the ellipsoid method have been developed, including those of Kremer and co-workers [75, 119], Hansen and Louis [21, 23, 89, 98, 99], and Eurich and Maass [48]. While ellipsoid models lower the computational cost of simulations by reducing the number of degrees of freedom, they do so at the expense of having complicated three- (or six-) dimensional distributions which must be searched in order to generate a trial configuration. In addition, the coarse-grained potentials employed in these studies are surprisingly long-ranged: some of the proposed potentials do not decay to zero even at distances equal to three times the mean radius of gyration of a chain. The principal axes defining the soft ellipsoids are usually equal to or less than one mean radius of gyration, making it unlikely that two polymers whose centers of mass are separated by three mean radii will actually overlap.

Dissipative particle dynamics, which adds a random field to the dynamics of a system of interacting soft ellipsoids, may be viewed as a more general version of the ellipsoid model. The field of dissipative particle dynamics began with the work of Español [45–47] and Schlijper [135], and the later developments of Marsh, Flekkøy, and Coveney [32, 57, 58, 108–111]. Further extensions of the method include the introduction of electrostatic interactions [39], entanglements [123], and alternative methods for integrating the equations of motion, such as

operator splitting [138]. However, some serious fundamental issues make dissipative particle dynamics significantly more complicated than the soft ellipsoid methods described above. The most critical of these is that there need not exist a one-to-one correspondence between atomistic chains and dissipative particles. Several chains can be mapped to a single dissipative particle and, counterintuitively, a single chain can be mapped to several non-bonded dissipative particles. These non-bonded particles are then free to move about the system, effectively violating connectivity constraints. This makes it almost impossible to construct a reverse mapping to restore the atomistic chains, particularly when mass is redistributed between particles during the course of the simulation.

The last method we will discuss here is the *dynamic mean field* or *complex Langevin* approach, recently developed by Fredrickson and co-workers [4, 62, 63, 67, 68]. Based on the self-consistent field theory approach of Flerer *et al.* [56], the dynamic mean field approach replaces the simulation of atomistic chains with the self-consistent solution of partial differential equations for the density distribution of a polymer within a given domain. The method has been particularly useful in determining the morphology of block copolymers and similar problems in polymer physics. The relative efficiency of the method compared to a similar dense melt calculation, however, has not yet been discussed in the literature. On the other hand, since the mathematical problem is well-understood, any of the many approaches to solving PDE's can be reliably used here, with little effect on the results. The main connection of the dynamic mean field (and SCFT) methods with the method outlined in this thesis is that the detailed atomistic description of the chain is replaced with a density-based description. To an extent, this mimics the method of dealing with polymer chains in the WAMC approach, which replaces an atomistic chain with a coarser chain whose beads are located at the centers of mass (or effectively, the maxima of density) of segments of the chain.

1.2.2 Multiscale methods

The term “multiscale” has been used in the past few years to describe methods which employ multiple simulations that pass data back and forth between different length or time scales. For example, a multiscale approach might use molecular dynamics (an atomistic-scale method) to compute a diffusion coefficient which will be employed in a finite-element computation (a

macroscale method). Recent applications of these methods include Sheng *et al.*, who use molecular dynamics to compute silica lamina stiffness, which is in turn used in micromechanical simulations to determine the effects of particle aspect ratio and particle volume fraction on the mechanical properties of a polymer-clay mixture [139]. Other applications, such as chemical vapor deposition, use finite element methods on the scale of an entire nanochip, but molecular-level transport and reaction theory at the level of an individual feature [131]. Similar methods have been used for applications as diverse as crack propagation [25, 133], biopolymers [70], and turbulent flows around airfoils [149].

There are several advantages to multiscale modeling approaches: there already exist many well-designed codes at each different scale of interest, and the cost of “passing” parameters between models is relatively low. However, the price paid for the relative ease of linking multiple simulations together is that a gap remains in our understanding of the physics. This gap occurs whenever there exists a set of length or time scales which remains unexplored by any of the simulation methods being used. Consequently, we may have a detailed understanding of the physics at an atomistic level and at a macroscopic level, but we cannot tell if there are interesting behaviors at so-called *mesoscales* between the two levels.

1.2.3 Multiresolution methods

The last class of coarse-graining methods that we shall discuss here is the set of multiresolution methods, which bridge length scales within a single simulation by treating the amount of coarse-graining performed as an adjustable parameter. As a result, it is not known *a priori*, for example, how many repeat units will be combined into a bead, or how many lattice spins will be averaged into a single block spin, or how much discretization or how many basis functions will be applied to a quantum-mechanical description of a molecule. In principle, this level of flexibility makes coarse-graining both hierarchical and heterogeneous: the same coarse-graining method can be used several times in succession to yield progressively coarser representations of a given original structure, and different levels of resolution can be used in different parts of a single simulation.

Truly multiresolution methods have not been often employed to date. Most of the existing multiresolution methods have been based on the wavelet transform, which we introduce in

Chapter 2. As a basis set, wavelets have been applied to electronic structure calculations [7] and to the physics of elementary particles using quantum field theory [121,127,128]. They have also been applied to modeling of control systems [147] and to image processing [29,81,101].

The most similar use of wavelets completed prior to the work outlined in this thesis is found in the work of Best and Schäfer [14–16], who use wavelets to expand the field variable in Landau-Ginzburg theory of spin models. Since that time, the work of Katsoulakis *et al.* [86,87] appear to be the closest analogues to the present work. Their considerations, however, are principally studying the dynamics of coarse-grained processes. Like our work, they began by considering adsorption and desorption on an Ising lattice; however, they present their results only for the simplest, one-dimensional lattice.

1.3 Thesis objectives and overview

As mentioned above, the aim of this thesis has been to develop a systematic framework for multiresolution modelling. The design goals of this framework include adaptability, both to a variety of applications and in its use for a single application. We would like a framework that can handle hundreds of particles just as readily as hundreds of thousands of particles, and can address different parameter states by adjusting the amount of coarse-graining that is applied to our system.

We would also like our framework to be robust: its performance should be markedly superior to the corresponding detailed atomistic simulations on which they are based. This will allow the method to be a first choice as an analysis tool which can quickly predict the properties and behavior of a system under a wide range of conditions. Of course, this will be true only if the results of the coarse-grained simulation can accurately predict the behavior, or if the results can be easily correlated with results obtained from either atomistic simulations or experimental measurements. Thus, in addition to flexibility, we will attempt to demonstrate both the relative efficiency and accuracy of our method, and suggest how the performance of the method can be “tuned” to optimize efficiency and accuracy.

Chapter 2 introduces the fundamental models and tools needed to understand the work presented in the remainder of the thesis. Chapters 3 and 4 address this problem as applied to

Ising lattices, while Chapters 5 and 6 will present its application to polymer modelling. Two further chapters address interesting results that arose in the context of building the polymer model: scaling results for overlap probabilities in polymer chains in Chapter 7, and the increased coupling of variables inherent in most coarse-graining processes in Chapter 8. Finally, the principal contributions of this work, as well as avenues of future exploration, are discussed in Chapter 9.

Chapter 2

Fundamental models and tools

Before presenting the main work of this thesis, we briefly review the important models and theories which will be required to understand the later chapters. We begin by introducing the basic lattice and polymer models, the wavelet transform, and renormalization group theories and discuss their relevance to this work.

2.1 Lattice theory

The standard model for studying the thermodynamic behavior of lattice systems is the spin- $\frac{1}{2}$ Ising model [122], which contains both nearest-neighbor pairwise interactions as well as interactions between lattice sites and an external field. The Hamiltonian for this system is normally written in the form

$$-\beta\mathcal{H} = h \sum_i \sigma_i + J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \quad (2.1)$$

where h is the strength of the external field in the direction of the spins σ_i , and J is the strength of the interaction between pairs of nearest-neighbor spins on the lattice; these pairs are indicated by the subscript $\langle ij \rangle$ in the second summation in (2.1). The inverse temperature $\beta = (k_B T)^{-1}$; for convenience we let $k_B = 1$, so that temperature, external field, and nearest-neighbor interactions are all dimensionless quantities. The model can be further extended by the inclusion of a position-dependent external field h_i and nearest-neighbor interactions J_{ij} , or by the inclusion of pairwise interactions beyond nearest neighbors. The most general form of

(2.1) is given by

$$-\beta\mathcal{H} = \sum_i h_i \sigma_i + \sum_i \sum_j J_{ij} \sigma_i \sigma_j, \quad (2.2)$$

where in (2.2) h_i is the strength of the external field at lattice site i , and J_{ij} is the strength of the interaction between sites i and j .

The primary goal of most lattice theory models is the development of the relevant phase diagram. For Ising lattices, the relevant variables for constructing phase diagrams are usually the temperature T and the external field strength h . The most important quantity to determine is the ferromagnetic critical point (T_c, h_c) , above which it is impossible to observe spontaneous magnetization of the lattice. It is trivial to show that the one-dimensional Ising magnet has no spontaneous magnetization for finite temperatures [71]. The two-dimensional case can also be solved analytically; however, the calculations are significantly more challenging [73, 122, 124]. In both the one- and two-dimensional cases, the critical point is located at $h = 0$; application of an external field “breaks” the symmetry of the system, making the existence of a critical point effectively impossible [71]. For the three-dimensional Ising model, it has been necessary to rely on series solutions and numerical simulation techniques to estimate the critical temperature [93, 104, 126, 154]. Some recent results even suggest that an analytic solutions does not exist for the three-dimensional model, as it may belong to the class of “NP-complete” problems [31, 82].

In addition to the phase diagram, several other quantities are of interest for Ising lattices. For example, the average magnetization of the lattice is determined by

$$m = \frac{1}{N} \left\langle \sum_i \sigma_i \right\rangle, \quad (2.3)$$

where the normalization is taken to ensure that the magnetization remains finite even in the limit $N \rightarrow \infty$. A lattice in which all spins are aligned in the same direction has either $m = 1$ or $m = -1$; an antiferromagnet, in which the preferred alignment is every spin opposite its neighbor, would have $m = 0$. We also expect to see $m = 0$ in the high-temperature limit, since the thermal background energy is so high that spins can flip regardless of their preferred alignment.

The other important quantity frequently computed is the heat capacity,

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}. \quad (2.4)$$

where E is the total energy of the system. The heat capacity, which gives the amount of energy which must be added to or removed from the system to change its temperature, is an example of a fluctuation property: a nonnegative quantity that is finite except in the vicinity of a critical point, where it diverges according to a power law based on the “distance” $|T - T_c|$ away from the critical point.

2.2 Physical models of polymer chains

Current state-of-the-art simulations often simulate dense systems containing up to 10^4 polymer repeat units, but only for time scales on the order of nanoseconds. This is unfortunately several orders of magnitudes below what is needed to address problems of engineering interest: individual polymer chains can exist on length scales of 10^6 repeat units, and have relaxation times on the orders of milliseconds to seconds [10]. Thus, even with today’s computational power, we are still not capable, for example, of studying the relaxation of a melt of long polymer chains without massive simplification of the problem. Ideally, we would like to create an algorithm which can achieve large increases in computational efficiency with only small trade-offs in computational accuracy. However, most methods of “coarse-graining” the system to reduce the amount of information that must be computed during the course of a molecular simulation automatically introduces severe approximations into the structure of the system, leading to often significant errors in physical calculations of thermodynamic properties. We outline in Chapters 5 and 6 our work in the development of a hierarchical coarse-grained model for polymer chains that is based on the fine-grained representation of a lattice random walk yet can still reproduce the features found in off-lattice and continuum models of polymers.

2.2.1 Discrete and continuum “limiting” descriptions

There is a vast array of polymer chain models, including everything from discrete lattice models to continuous three-dimensional “thread” models. In addition, there are a host of other

models to describe polymer systems, including the so-called Flory-Huggins theory [59] and self-consistent field theory models [61, 63, 136], neither class of which explicitly represents chains in space.

The simplest example of a polymer walk is of course the Bernoulli random walk studied on a hypercubic lattice. Given a bead at lattice point \mathbf{x}_{i-1} , the next lattice point \mathbf{x}_i is selected with probability $(2d)^{-1}$ from any of the $2d$ neighboring lattice points, where d is the dimensionality of the lattice. In the Bernoulli walk, there are no constraints other than connectivity, so that lattice sites can be visited multiple times by one or more chains.

When the restriction of sitting on a hypercubic lattice is removed, the Bernoulli random walk becomes the freely-jointed chain [59], where there is no preferred orientation of bonds. Consequently, given a bead at point \mathbf{x}_{i-1} , the next point \mathbf{x}_i can be chosen according to the distribution

$$p(\mathbf{x}_i) = \frac{1}{4\pi a^2} \delta(|\mathbf{x}_i - \mathbf{x}_{i-1}| - a),$$

where a is the step size of the walk. Again, there is no self-avoidance requirement to prevent the chain from overlapping itself. A further extension of the freely-jointed chain, the Gaussian chain [161], permits individual bond lengths to vary according to the distribution

$$p(\mathbf{x}) = \left(\frac{3}{2\pi a^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{x}^2}{2a^2}\right), \quad (2.5)$$

where a is now the average bond length of the walk. We shall describe in Chapter 5 an algorithm that treats a coarse-grained chain like an on-lattice Gaussian chain model, as bond lengths are chosen according to a distribution whose limit approaches (2.5), but are then adjusted to fit on a finely-spaced hypercubic lattice.

We can make yet one more relaxation of the Gaussian chain by specifying a curve in space to represent the polymer. The resulting chain, sometimes called a ‘‘Gaussian thread,’’ is specified by $\mathbf{r}(\tau)$, where typically τ runs from 0 to 1 [42]. The thread description is particularly useful for self-consistent field theories and renormalization approaches [63, 64].

2.2.2 Interacting chain

So far, we have not specified any physical interactions which can affect the structure of the random walk. Consequently, while the walks we have described are analytically quite tractable, they do not accurately reflect the behavior of real polymer molecules.

For a single chain, we can categorize the various interactions as either between a particle and an external field, between two bonded particles, or between two or more non-bonded particles. For a given walk $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ on a (hypercubic) lattice, we can define the energy of the corresponding polymer chain as a sum of the different interactions:

$$U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N u_1(\mathbf{r}_i; \{\mathbf{r}\}) + \sum_{i=1}^N \sum_{j=i+2}^N u_2(\mathbf{r}_i, \mathbf{r}_j) + \dots, \quad (2.6)$$

where the one-body potential $u_1(\mathbf{r}_i; \{\mathbf{r}\})$ describes the internal energy corresponding to a bead at \mathbf{r}_i , while the two-body potential $u_2(\mathbf{r}_i, \mathbf{r}_j)$ incorporates both the self-avoidance condition as well as non-bonded interactions. Most two-body interactions are treated as functions only of the interparticle separation $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and therefore we write $u_2(\mathbf{r}_i, \mathbf{r}_j) = u_2(r_{ij})$, and define the potential as

$$u_2(r_{ij}) = \begin{cases} \infty, & r_{ij} < a \\ -\varepsilon_{nb}(r_{ij}), & a \leq r_{ij} \leq ca, \\ 0, & ca \leq r_{ij} \end{cases}, \quad (2.7)$$

where a represents the hard-core radius between particles, and c is a constant representing the “cutoff” of the potential (the distance beyond which two particles are assumed not to interact). In addition, it is generally assumed in (2.7) that particles i and j are non-bonded. The potential $-\varepsilon_{nb}(r_{ij})$ must be specified; for our purposes, we shall assume that it is a piecewise continuous function with $\varepsilon_{nb} > 0$ on $a \leq r_{ij} \leq ca$. If we define a to be the lattice spacing, then letting $1 < c < \sqrt{2}$ defines interactions between nearest-neighbor lattice sites, and $c < 2$ defines interactions between next-nearest-neighbors.

The internal energy $u_1(\mathbf{r}'; \{\mathbf{r}\})$ for our model is defined as

$$u_1(\mathbf{r}_i; \mathbf{r}_{i-1}, \mathbf{r}_{i+1}) = \begin{cases} \varepsilon_s, & \mathbf{r}_i - \mathbf{r}_{i-1} \neq \mathbf{r}_{i+1} - \mathbf{r}_i \\ 0, & \mathbf{r}_i - \mathbf{r}_{i-1} = \mathbf{r}_{i+1} - \mathbf{r}_i \end{cases} . \quad (2.8)$$

Equation (2.8) defines a *stiffness potential*—that is, an energy “penalty” paid for changing directions between $\mathbf{r}_i - \mathbf{r}_{i-1}$ and $\mathbf{r}_{i+1} - \mathbf{r}_i$; this is functionally similar to defining a persistent random walk, although its effects are observed in the acceptance or rejection of states, rather than in their selection.

In addition to the one- and two-body interactions already mentioned, we can also incorporate bond-angle and torsion-angle interactions between three and four consecutive bonded particles along a chain. Although the atomistic models we will consider do not explicitly incorporate these potentials, we shall see that there exists a bias in the distribution of the coarse-grained internal coordinates.

2.3 Wavelet transform fundamentals

2.3.1 The conceptual picture

The wavelet transform is a hierarchical method for decomposing a data set into averages and differences. Like the Fourier transform, it can be used to provide a decomposition in both real space and reciprocal space (\mathbf{k} -space), or time space and frequency space. Unlike the Fourier transform, however, it is capable of providing simultaneously localized transformations in both real and reciprocal space. A function localized in position space, such as a finite impulse function, cannot be represented by only a few terms of its Fourier series; many terms are required before good convergence is achieved. By contrast, in wavelet space, this same function can be almost completely described by just a handful of wavelet coefficients. Although the first wavelet was discovered almost a century ago by Haar [74], they have become an important computational technique only in the last decade, following the work of Mallat [106, 107], Daubechies [34], and others [30, 150, 151].

The wavelet transform, like any other transform, takes a mathematical object and transforms

it into another: we can represent its action by writing

$$\tilde{u} = \mathcal{W}[u]; \tag{2.9}$$

the specific form of \mathcal{W} depends both on the type of wavelet we have selected, and the object u which we wish to transform. All versions of the wavelet transform \mathcal{W} , however, are derived from the same source: a set of coefficients which define the transform. If u is a discrete data set, such as a signal sampled at regular intervals, then \mathcal{W} is usually represented as a matrix; while if u represents a continuous data set, such as the same signal measured at all times, then \mathcal{W} acts as an integral operator. While the matrix form of \mathcal{W} is often called a “filter bank” and the integral form a “wavelet transform,” we will not distinguish between them in what follows, as the theory developed here for discrete lattices and filter banks should carry over to continuous systems and wavelet transforms essentially unchanged.

Similar to the Fourier transform, the wavelet transform decomposes the object x into two separate components, as two different functions, a scaling function ϕ and a wavelet function ψ , both operate on x . However, the two functions separate its components not into cosines and sines, but into averages and differences, with a “wavelength” equal to the “window” over which the scaling and wavelet functions are nonzero. In another important distinction, the wavelet transform is *recursive*, so that it can be applied in succession to any set of averages which is produced using that wavelet transform, to produce another level of averages and another level of details.

To make the above concepts more mathematically precise, let us define u to be a discrete set of samples $u = (u(1), u(2), \dots, u(n))$. Then applying the scaling and wavelet functions ϕ and ψ to u create a set of averages $s(i)$ and a set of differences $\delta(i)$:

$$s(i) = \sum_{k=0}^{r-1} \phi(k) u(i+k), \tag{2.10}$$

$$\delta(i) = \sum_{k=0}^{r-1} \psi(k) u(i+k), \tag{2.11}$$

where r is a finite integer which defines the length scale, often referred to as the “size of the support,” over which ϕ and ψ are nonzero. The index i runs from 1 to n ; generally the data

set is padded with zeros to ensure that all sums in (2.10) and (2.11) are well-defined, although periodicity is sometimes used instead [148]. The coefficients $\phi(k)$ and $\psi(k)$ in (2.10) and (2.11) are related [34, 148], and are central in controlling the features of the wavelet transform. Note the wavelet transform is inherently redundant: for every sample $u(i)$ in the original set u , we now have two values, a local average $s(i)$ and a local difference $\delta(i)$. Since the new data are simply linear combinations of the original values, it is superfluous to retain both sets; at the same time, it is obvious that we cannot simply discard one set of data and recover all the original information using only the other data set. Instead, we choose to keep only the odd-numbered $s(i)$'s and $\delta(i)$'s, eliminating the even-numbered samples; this process is called *downsampling* [148]. Downsampling removes half of the $s(i)$ and half of the $\delta(i)$, regardless of the length r of the wavelet. Now we are left with n data: $s(1), s(3), \dots, s(n-1)$ and $\delta(1), \delta(3), \dots, \delta(n-1)$. These n data points can be stored as the level-one wavelet transform \tilde{u} of u , by assigning $s(1), s(3), \dots, s(n-1)$ to $\tilde{u}^{(1)}(1), \tilde{u}^{(1)}(2), \dots, \tilde{u}^{(1)}(n/2)$, and the corresponding $\delta(i)$'s as $\tilde{u}^{(1)}(n/2+1), \dots, \tilde{u}^{(1)}(n)$. [The superscript (1) denotes that the wavelet transform has been applied once to this data set.] We can either stop at this level of description, or continue by further decomposing the averages: then the new object $u^{(1)}$ to be transformed is $u^{(1)}(1) = \tilde{u}(1), \dots, u^{(1)}(n/2) = \tilde{u}(n/2)$, and so on. Note that although $\tilde{u}^{(1)}$ contains the averages $s(i)$'s and the differences $\delta(i)$'s obtained in the previous step, successive transforms only apply to averages obtained in the previous step. This process can be repeated until we have reduced our set of averages to a single point; no further averaging is possible. We assume henceforth that the data set $u^{(k)}$ has been sufficiently downsampled to retain only the minimum data set required.

2.3.2 Properties and examples of wavelet and scaling functions

Until this point, we have not introduced any specific wavelet or scaling functions. Before we do so, we note that the choice of a wavelet transform to apply to a given system usually hinges on the desired properties which one wishes to include in the transformed data. Three principal properties are almost universally required for filter banks and wavelet families [34, 148]:

1. *Perfect reconstruction*: No data is distorted by performing analysis followed by synthesis, so that the only permissible change is a delay in recovery of the original sample.

2. *Orthogonality*: Wavelets computed at different length scales or at different spatial locations are mutually orthogonal; thus fluctuations in the system are localized at the scales where they are most relevant.

3. *Compact support*: Properly designed wavelets are identically zero except for a finite interval, which means that exact results can be obtained using only a finite number of terms.

Other properties, such as orthonormality, symmetry in the functional form of the wavelet or a certain number of vanishing moments, can be taken into account when constructing the wavelet transform [152].

The two most commonly encountered selections are the Haar and Daubechies wavelets, named after their respective discoverers. The Haar pair is the oldest and simplest set of wavelets [74]: the coefficients of the scaling function are $\phi = (\phi(0), \phi(1)) = \frac{1}{\sqrt{2}}(1, 1)$, while the coefficients of the wavelet function are $\psi = (\psi(0), \psi(1)) = \frac{1}{\sqrt{2}}(-1, 1)$. No other wavelet can be described with two points, and therefore no other wavelet has a support as compact as the Haar wavelet. The scaling function ϕ simply averages the values stored at neighboring points, while ψ finds the difference between those values; the extra factor of $\sqrt{2}$ is incorporated to ensure orthonormality between overlapping $\phi(k)$ and $\psi(k)$. A simple example of the action of the Haar wavelet is shown in Fig. 2-1.

The Daubechies wavelets are a family of orthonormal functions explicitly designed to have orthogonality as well as vanishing higher-order moments [33]. Daubechies was able to show that the Haar wavelet is in fact the “first” member of the Daubechies family; that is, the Haar wavelet is the Daubechies wavelet with the shortest support. The second such member has four terms in its definition: the scaling function is defined by $\phi = (\phi(0), \phi(1), \phi(2), \phi(3)) = \frac{1}{4\sqrt{2}}(1 + \sqrt{3}, 3 + \sqrt{3}, 3 - \sqrt{3}, 1 - \sqrt{3})$. The wavelet function reverses the order of the coefficients and inverts the sign of every other component, which allows the orthonormality properties to be satisfied:

$$\psi = (\psi(0), \psi(1), \psi(2), \psi(3)) = (-\phi(3), \phi(2), -\phi(1), \phi(0)).$$

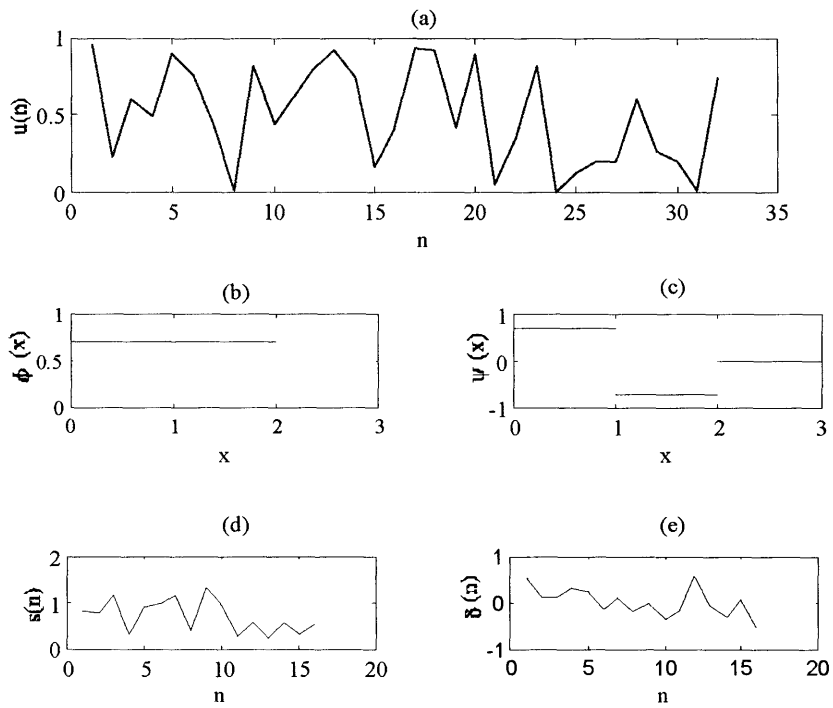


Figure 2-1: (a) A sample signal $u(n)$. (b, c) The one-dimensional Haar scaling function $\phi(x)$ and wavelet (differencing) function $\psi(x)$. (d, e) The first-level scaling coefficients $s(n)$ and $\delta(n)$ produced from the original signal $u(n)$ using the Haar pair, following downsampling of the signal.

We can see that the Haar wavelet obeys the same pattern as the Daubechies wavelet: $\psi_H = (-\phi(1), \phi(0))$. This pattern can be extended, using different coefficients but the same general sign rules, for wavelets with 6, 8, 10, ... coefficients. The resulting wavelet and scaling functions become increasingly smooth, and therefore are better suited for data sets in which there is only a gradual change in the data set with position—in thermodynamic systems, this would be more useful for, say, a spin- N Ising model than a spin- $\frac{1}{2}$ Ising model (presuming that $N \gg \frac{1}{2}$).

2.3.3 Matrix formulation of the wavelet transform

For discrete systems, a conceptually simple method of implementing the wavelet transform is to set up the transform as a matrix equation. The input u is converted into a column vector \mathbf{u} , so that the coefficients $s(i)$ and $\delta(i)$ are given by the dot product of \mathbf{u} with vectors $\mathbf{h}(i)$ and $\mathbf{l}(i)$, where the vectors are padded so that the first nonzero element is located at position i :

$$\begin{aligned}\mathbf{h}(i) &= (0, \dots, 0, \phi(0), \phi(1), \dots, \phi(r-1), 0, \dots, 0), \\ \mathbf{l}(i) &= (0, \dots, 0, \psi(0), \psi(1), \dots, \psi(r-1), 0, \dots, 0).\end{aligned}$$

The vectors \mathbf{s} and $\boldsymbol{\delta}$, which contain the wavelet-transformed coefficients of the decomposition, can be obtained by forming the matrices \mathbf{H} and \mathbf{L} and right-multiplying by the vector \mathbf{u} :

$$\mathbf{s} = \mathbf{H}\mathbf{u} \text{ and } \boldsymbol{\delta} = \mathbf{L}\mathbf{u},$$

where the rows of the matrices \mathbf{H} and \mathbf{L} are the vectors $(\mathbf{h}(1), \dots, \mathbf{h}(n))$ and $(\mathbf{l}(1), \dots, \mathbf{l}(n))$, respectively. As mentioned in the previous section, we do not need to keep all of the $s(i)$'s and $\delta(i)$'s in order to obtain perfect reconstruction of our signal; thus, we can obtain all the necessary coefficients in a single matrix multiplication by combining the relevant rows of \mathbf{H} and \mathbf{L} into a single matrix $\mathbf{W}^{(1)}$, whose rows are $(\mathbf{h}(1), \mathbf{h}(3), \dots, \mathbf{h}(n-1), \mathbf{l}(1), \mathbf{l}(3), \dots, \mathbf{l}(n-1))$. Thus,

the wavelet transformation (2.9) can be written as

$$\begin{bmatrix} s(1) \\ \vdots \\ s(n-1) \\ \delta(1) \\ \vdots \\ \delta(n-1) \end{bmatrix} = \begin{bmatrix} \mathbf{h}(1) \\ \vdots \\ \mathbf{h}(n-1) \\ \mathbf{l}(1) \\ \vdots \\ \mathbf{l}(n-1) \end{bmatrix} \begin{bmatrix} u(1) \\ u(2) \\ \vdots \\ u(n) \end{bmatrix}. \quad (2.12)$$

We will denote the product on the left-hand side of (2.12) as $\bar{\mathbf{u}}$.

As stated above, the wavelet process can be applied recursively: the set of averages

$$(s(1), s(3), \dots, s(n-1))$$

can be treated as a new data sample $\mathbf{u}^{(1)} = (u^{(1)}(1), \dots, u^{(1)}(n/2))$, and operated on by an $\frac{N}{2} \times \frac{N}{2}$ reduction of $\mathbf{W}^{(1)}$, which we denote $\mathbf{W}^{(2)}$, to produce a new set of $n/4$ averages

$$(s^{(2)}(1), s^{(2)}(3), \dots, s^{(2)}(n/2-1))$$

and corresponding new set of $n/4$ differences

$$(\delta^{(2)}(1), \delta^{(2)}(3), \dots, \delta^{(2)}(n/2-1)).$$

To reconstruct the original data set, we combine these $n/2$ values along with the $n/2$ differences $(\delta^{(1)}(1), \delta^{(1)}(3), \dots, \delta^{(1)}(n-1))$ obtained from applying $\mathbf{W}^{(1)}$. This process can be repeated as many times as desired, dividing the m non-downsampled averages $\mathbf{s}^{(k)}$ into $m/2$ averages $\mathbf{s}^{(k+1)}$ and $m/2$ differences $\delta^{(k+1)}$. However, since at each iteration the matrix $\mathbf{W}^{(k)}$ only operates on selected elements of the vector $\bar{\mathbf{u}}^{(k)} = \mathbf{W}^{(k-1)}\mathbf{u}^{(k-1)}$, computations for multiple levels can be performed at the same time. Thus, if we wish to apply the wavelet transform K

times, we can write this operation as an extended matrix product [69]:

$$\bar{\mathbf{u}}^{(K)} = \mathbf{W}\mathbf{u} = \prod_{k=1}^K \mathbf{Q}^{(k)}\mathbf{u}, \quad (2.13)$$

where the $\mathbf{Q}^{(k)}$ are a family of matrices of the form

$$\mathbf{Q}^{(k)} = \begin{bmatrix} \mathbf{W}^{(k)} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (2.14)$$

In (2.14), $\mathbf{Q}^{(k)}$ is always an $N \times N$ matrix, while the matrix $\mathbf{W}^{(k)}$ has size $(N/2^{k-1}) \times (N/2^{k-1})$.

To recover the original data sample \mathbf{u} following a wavelet transform \mathbf{W} , we can multiply $\bar{\mathbf{u}}$ by the inverse of \mathbf{W} . The matrix \mathbf{W} is unitary: that is, its inverse \mathbf{W}^{-1} is equal to its transpose \mathbf{W}^T . Consequently, if \mathbf{W} is known, all that is necessary to reverse the transformation is to left-multiply $\bar{\mathbf{u}}^{(k)}$ by the transpose \mathbf{W}^T . Moreover, the computation (2.12) of the wavelet transform can usually be carried out “in place” by manipulating local coordinates; in this manner, the computation is carried out even more rapidly than a standard multiplication, and without the increased storage costs associated with matrix multiplications [72, 155].

2.3.4 Multidimensional wavelet transforms

Since virtually all problems in lattice thermodynamics are in multiple dimensions, it is necessary to take the wavelet transform of a multivariate function or data set. Several methods have been developed to carry out such transformations; among them are Cohen and Daubechies’s *separable wavelets*, which form the multidimensional scaling and wavelet functions $\phi(x, y)$, $\psi_{xx}(x, y)$, $\psi_{xy}(x, y)$, and $\psi_{yy}(x, y)$ from products of the one-dimensional scaling and wavelet functions $\phi(x)$ and $\psi(x)$ [30]. A more general algorithm, the *lifting algorithm*, has been developed by Sweldens [35, 153]. It divides the wavelet transform into two steps: the first computes the wavelet coefficients $\delta(i)$; the second step uses the wavelet coefficients to speed up the calculation of the scaling coefficients. “Initialization” of the lifting algorithm requires the use of an appropriately selected basis function.

A convenient basis function for the multidimensional lifting transform is the generalized orthogonal Haar wavelets outlined by Sweldens [153]. An extension of the one-dimensional

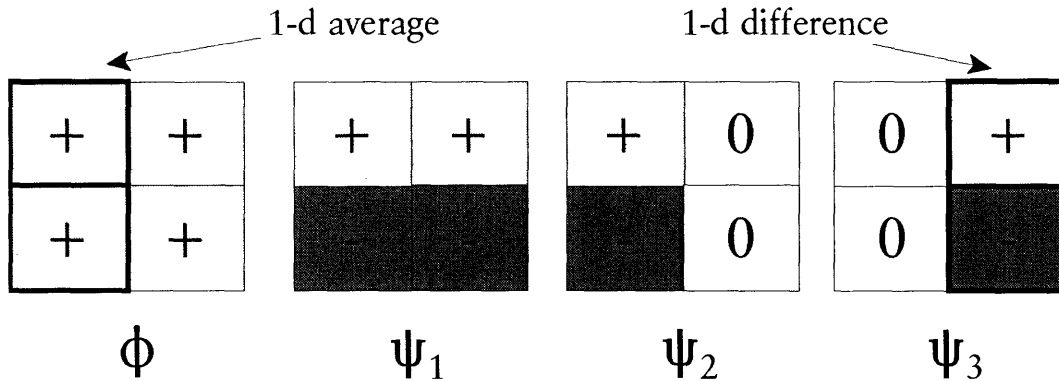


Figure 2-2: The two-dimensional orthogonal Haar wavelets. The coefficient in all cases is $1/2$, times the sign indicated in each quadrant. Note that the scaling function ϕ is built from the one-dimensional Haar scaling function, while the wavelet functions are built from the one-dimensional wavelet function.

wavelet transform, they can be created in any number of dimensions, and have the same basic orthonormality properties as the one-dimensional Haar functions, although the orthonormality constant becomes $2^{-d/2}$, where d is the dimensionality of the system. [The two-dimensional version is shown in Figure 2-2.] Moreover, the use of the Haar wavelets as a starting point for further iterations of the lifting algorithm allow the development of additional, “better” lifted wavelets with more desirable properties, such as smoothness.

It can further be seen that the “oversampling” problem which exists in one dimension will be magnified in multiple dimensions: since the number of wavelet functions which are produced from each data point increases by a factor of two with each additional dimension added, we must reduce the number of points maintained for each wavelet function by that same factor. Thus, in two dimensions, we keep only every fourth point; in three dimensions, every eighth point, and so on. The wavelet transform for multidimensional systems can thus still be written in the form of (2.13) and (2.14), after we have written the multidimensional data set in terms of a column vector. This can be accomplished by wrapping around the edges of the matrix in creating \mathbf{u} : for example, after inserting element $(1, N)$ of a two-dimensional data set into \mathbf{u} , we next store element $(2, 1)$, and so forth. The other significant difference in the structure of these equations is that the size of the submatrix $\mathbf{W}^{(k)}$ in (2.14) is now $N/2^{(k-1)d} \times N/2^{(k-1)d}$ instead of $N/2^{k-1} \times N/2^{k-1}$.

2.4 Renormalization group principles

Renormalization group (RG) theory, first developed by Wilson [160], has been applied primarily to systems in the thermodynamic limit near a critical point [6, 38, 55, 71], including glass transitions [137] and turbulent systems [49]. The general technique, however, is not restricted exclusively to systems in the vicinity of a critical point in the space of coupling parameters.

Although originally intended for use in quantum field theory, the range of applications of the renormalization group has been extended to include fields as diverse as polymer theory, hydrodynamics, and asymptotic analysis [6, 27, 38, 40, 55]. Other groups have applied renormalization group theory to nonlinear oscillators [27] and finite systems of ferromagnets [140–144]. RG theory can also be used in the study of systems approaching equilibrium, such as time-dependent diffusion and other self-similar phenomena from transport theory and fluid mechanics [9, 71].

The origins of RG theory lie in the effective field theories of Landau and Ginzburg, who introduced the notion of an order parameter which represents the *mesoscopic* scale of physical phenomena [55, 71]. At a fundamental level, the goal of renormalization methods is to identify the critical behaviors of a system as a function of the existing strengths of interaction, and to determine the conditions under which the system, when viewed at increasingly coarser scales, will evolve towards those critical points.

Perhaps the most significant feature—although some view it as a weakness—of the renormalization group transformations is that it is not a group, but in fact, only a semi-group [38, 55]. Consequently, the RG transformation does not possess a unique inverse transformation, and thus one generally cannot determine the unique initial state which resulted in a new renormalized state. This should not be surprising: virtually every type of coarse-graining method, including renormalization group theory, sacrifices detailed information about a thermodynamic system in order to carry out the analysis more efficiently.

One additional element of the renormalization group should also be noted here. While perturbation expansions may be required to carry out some steps in an RG analysis, the analysis itself is inherently non-perturbative: we do not require that there exists a small parameter be small to carry out the procedure.

2.4.1 Renormalization group analysis of lattice systems

The basic “recipe” for renormalization groups is a function mainly of the choice of approach: real-space versus momentum-space. As their names imply, real-space approaches work on the actual positions and particles in a given system, while momentum approaches are applied in Fourier space

For lattice systems, a typical real-space RG program entails several steps, which can be divided roughly as follows.

1. Identify the variables to be “blocked.” *Block variables* represent a set of variables in the original system, grouped together to function as a single unit at a higher scale. For example, in a two-dimensional rectangular Ising model, a block variable would be an $m \times n$ rectangle of spins. In RG theory, the block variables should have the same allowed values as the original variables. In the Ising model, this means that the block variables are allowed to take only the values -1 or $+1$.

2. Create the *projection operator*: the projection operator $\mathcal{P}(\{s_i\})$ is a probability assignment rule from original variables $\{s_i\}$ to block variables $\{s'_I\}$. Therefore, we require

$$(a) \mathcal{P}(\{s_i\}) \geq 0$$

$$(b) \mathcal{P}(\{s_i\}) \text{ preserves the } \textit{symmetries} \text{ and } \textit{interactions} \text{ of the original Hamiltonian}$$

$$(c) \sum_{\{s_i\}} \mathcal{P}(\{s_i\}) = 1 \text{ preserves the } \textit{value} \text{ of the original Hamiltonian}$$

3. Re-organize the Hamiltonian: the Hamiltonian of the system is written so that it is a function only of the new variables of the system. This normally entails rewriting the coupling constants in a more compact form, although the relationship between the original and renormalized coupling constants may be highly nonlinear [124]. As an example, consider the Hamiltonian of the one-dimensional Ising model,

$$\mathcal{H} = \sum_{i=1}^N \left\{ K + k_1 s_i s_{i+1} + \frac{1}{2} k_0 (s_i + s_{i+1}) \right\}, \quad (2.15)$$

where K is a constant which is initially set to zero. Upon renormalization, we obtain the

Hamiltonian

$$\mathcal{H}' = \sum_{i=1}^N \left\{ K' + k'_1 s_i s_{i+1} + \frac{1}{2} k'_0 (s_i + s_{i+1}) \right\},$$

where the new parameters K' , k'_0 , and k'_1 are related to K , k_0 , and k_1 by the relations

$$K' = \ln 2 + 2K + \frac{1}{4} \ln (\cosh (2k_1 + k_0) \cosh (2k_1 - k_0)) + \frac{1}{2} \ln \cosh k_0, \quad (2.16)$$

$$k'_1 = \frac{1}{4} \ln \left(\frac{\cosh (2k_1 + k_0) \cosh (2k_1 - k_0)}{\cosh^2 k_0} \right), \quad (2.17)$$

$$k'_0 = k_0 + \ln \sqrt{\frac{\cosh (2k_1 + k_0)}{\cosh (2k_1 - k_0)}}. \quad (2.18)$$

4. Identify the *fixed points* of the Hamiltonian. The fixed points represent the limit points in the space of coupling constants of multiple iterations of the RG transformation.
5. Construct the phase diagram or scaling laws for the system.

2.4.2 Renormalization group flows

To an extent, the principal goal of renormalization group analyses is to determine the *flows* through parameter space. For a given set of starting parameters \mathbf{K} , a renormalization transformation will map \mathbf{K} to a new point \mathbf{K}^* in parameter space. The only exception to this rule occurs at *fixed points*, which are unchanged by the renormalization transformation. For example, in the transformation specified by Equations (2.15) through (2.18), the fixed points correspond to those sets of parameters for which $K \rightarrow K'$, $k_0 \rightarrow k'_0$, and $k_1 \rightarrow k'_1$.

The three types of fixed points are attractive, mixed, and repulsive. For any given starting point in parameter space, an infinite number of iterations of the renormalization transformation must ultimately reach a fixed point. If the starting parameters \mathbf{K} is within the “well of attraction” of an attractive fixed point, it will eventually reach it. Similarly, for a mixed fixed point, some points will be ultimately mapped to it, while other points will move to an attractive fixed point. Finally, repulsive points are *never* mapped to by any other point in parameter space: it is possible to reach a repulsive fixed point only by starting at that point. Thermodynamic critical points are always repulsive fixed points.

An example of renormalization group flows is shown as Figure 2-3. The points designated S

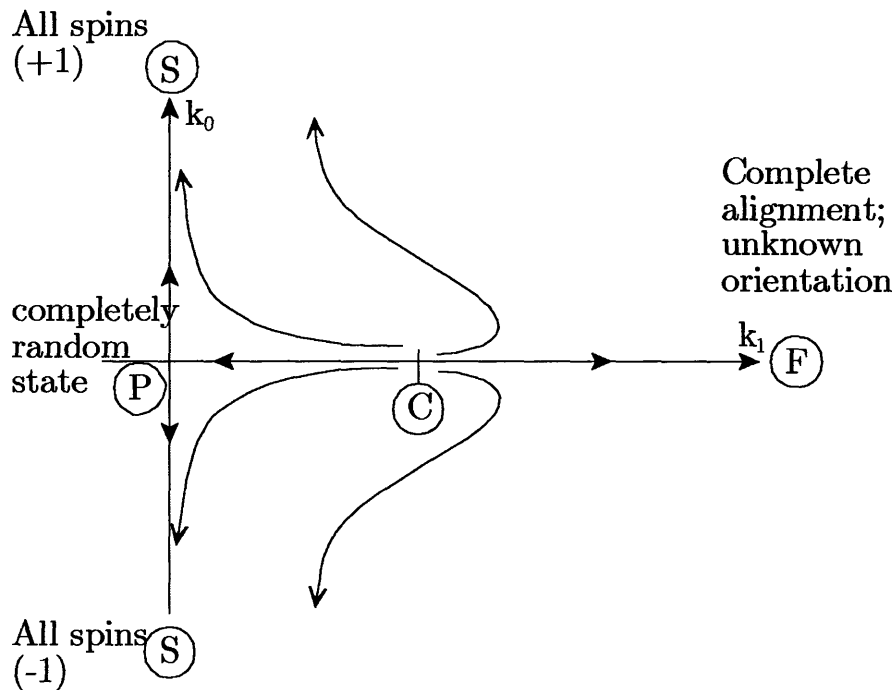


Figure 2-3: Example of renormalization group flows, as applied to the two-dimensional Ising model.

and F correspond to attractive fixed points: any point starting in their wells of attraction will yield a state of complete alignment, although for point F , corresponding to the low-temperature limit, the specific orientation obtained cannot be determined *a priori*. Point P corresponds to the high-temperature limit, in which all states are equally probable. The only point which is not mapped to by another point in parameter space is the ferromagnetic critical point, labelled C . Any other point in parameter space will tend to S , F , or P , although P can only be obtained along the line $k_0 = 0$ —that is, in the absence of an external magnetic field.

2.4.3 Connections to our coarse-graining methodology

The utility of the renormalization group method is that it can be used in conjunction with other numerical techniques. For example, the Monte Carlo renormalization group (MCRG) method, which has been successfully applied to both two-dimensional and three-dimensional

simple cubic lattice models [126, 154], uses iterations of Monte Carlo simulations to observe flows through parameter space; while density-matrix renormalization group (DMRG) has been developed for quantum chemical calculations [158, 159].

Our aim is to preserve the spirit of RG methods in our research, although we will not explicitly follow any particular RG method. For example, selecting the block variables and the projection operator are replaced by choosing an appropriate wavelet basis. Reorganization of the Hamiltonian using wavelets will proceed in a manner similar to the RG method, except we shall be more interested in the partition function than in the Hamiltonian itself. Finally, we are concerned with general calculations on lattices, rather than just the identification of fixed points and renormalization flows which are the hallmark of most RG and RG-like analyses [13, 71].

Chapter 3

Multiresolution analysis in statistical mechanics. I. Using wavelets to calculate thermodynamic properties

3.1 Introduction

Spin models are popular tools for theoretical calculations and for numerical simulations, as their universality classes allow a huge range of different systems—as varied as binary metal alloys, surface adsorption, and neural networks—to be modeled simultaneously. For example, even the “trivial” one-dimensional Ising model can be used to model the helix-coil transition in biopolymers; the deep connection between magnetic models and polymer chains allow us to predict scaling behavior and other properties across an even wider range of materials [64]. Lattice models are still widely used in modeling the thermodynamics of complex systems, because their regular structure simplifies the type and nature of interactions among components of the system. Moreover, the difficulty in obtaining analytical solutions of lattice systems, and the relative ease of computational simulations thereof, make them ideal test cases for new simulation algorithms.

Although simulations of lattice models are relatively straightforward to implement, they share the same drawbacks as off-lattice models. The chief drawback is that as the number of

particles grows large, the time required to sample the system accurately increases rapidly. A popular approach for addressing this problem is to *coarse-grain* the system: that is, we “rescale” the problem by increasing the basic size of a simulation element. For example, we might coarse-grain an atomic representation of a polymer chain into a “united atom” model, where a chain molecule is treated as if it consisted only of the backbone. More creative approaches redefine the problem to be addressed: for example, Mattice and coworkers have produced a method which maps a polymer chain atomistically onto a high-coordination lattice; this lattice is then used as the basis for a Monte Carlo simulation; the resulting configuration is then used to map back to continuous space to provide “evolution in time” [28, 43, 44, 129].

This chapter illustrates the use of the wavelet transform as a mathematical basis for performing thermodynamic computations of lattice models. The wavelet transform is an important tool in *multiresolution analysis*, which analyzes a system simultaneously at several length or frequency scales selected to reflect the actual physical processes underlying the observed behavior as closely as possible. The wavelet transform possesses a number of convenient properties, including orthogonality, compactness, and reconstruction; we will make these concepts more precise in Section 2.3.2 below. The orthogonal nature of most wavelet constructions makes them a logical choice for use in *ab initio* density-matrix quantum chemistry computations, in which the selection of an accurate basis set is crucial to the convergence and efficiency of the calculations [7, 17, 83]. Wavelet decompositions have been applied principally in electrical engineering, particularly in the field of signal processing. In this context, white noise and Markov processes have been studied using multiscale methods [100, 101]. To date, however, wavelet analysis does not seem to have been extensively applied to models in statistical mechanics. Huang uses wavelet analysis to observe the statistical distribution of multiplicity fluctuations in a lattice gas [79], while Gamero *et al.* employ wavelets to introduce their notion of *multiresolution entropy*, although their primary goal is dynamic signal analysis rather than statistical mechanics simulations [66], while O’Carroll attempts to establish a theoretical foundation connecting wavelets to the block renormalization group [120, 121]. A more in-depth review of the connection between wavelets and renormalization theory is provided in a recent monograph by Battle [12].

3.2 Wavelet analysis of lattice thermodynamics

3.2.1 Applying the wavelet transform to lattice Hamiltonians

While (2.1) and (2.2) are compact representations of the Hamiltonian of the system, the expansion of the lattice variables s_i and s_j as a sum of wavelet coefficients makes these equations impractical for applying the wavelet transformation. Since the system is described discretely, we want to use discrete wavelets, and therefore a matrix formulation of the Hamiltonian would be convenient. Using graph theory [31], this is readily accomplished: let the vectors $\mathbf{u} = (\sigma_1, \sigma_2, \dots, \sigma_N)$ and $\mathbf{h} = (h_1, \dots, h_N)$ denote the values of each of the N lattice variables in the system and the set of external-field interaction strengths, respectively, constructed in the row-wise manner described in the previous section. Furthermore, define the matrix \mathbf{J} such that element J_{ij} is the strength of the interaction between site i and site j . If these sites do not interact, then $J_{ij} = 0$. Then, the Hamiltonian (2.2) can be written in the form of a matrix equation:

$$-\beta\mathcal{H} = \mathbf{h}^T \mathbf{u} + \mathbf{u}^T \mathbf{J} \mathbf{u}, \quad (3.1)$$

where the superscript T denotes the transpose of the vector (or matrix) which precedes it.

The matrix \mathbf{W} which defines the wavelet transform satisfies by construction $\mathbf{W}^T \mathbf{W} = \mathbf{I}$, where \mathbf{I} is the identity matrix. Therefore, to apply the wavelet transform, we simply insert $\mathbf{W}^T \mathbf{W}$ between each pair of terms in (3.1), thereby obtaining

$$-\beta\mathcal{H} = (\mathbf{h}^T \mathbf{W}^T) (\mathbf{W} \mathbf{u}) + (\mathbf{u}^T \mathbf{W}^T) (\mathbf{W} \mathbf{J} \mathbf{W}^T) (\mathbf{W} \mathbf{u}). \quad (3.2)$$

Using the general matrix property that $\mathbf{B}^T \mathbf{A}^T = (\mathbf{A} \mathbf{B})^T$, (3.2) can be rewritten in terms of the wavelet-transformed vectors $\bar{\mathbf{h}}^{(K)} = \mathbf{W} \mathbf{h}$, $\bar{\mathbf{u}}^{(K)} = \mathbf{W} \mathbf{u}$, and the wavelet-transformed matrix $\bar{\mathbf{J}}^{(K)} = \mathbf{W} \mathbf{J} \mathbf{W}^T$:

$$-\beta\tilde{\mathcal{H}} = \left(\bar{\mathbf{h}}^{(K)} \right)^T \bar{\mathbf{u}}^{(K)} + \left(\bar{\mathbf{u}}^{(K)} \right)^T \bar{\mathbf{J}}^{(K)} \bar{\mathbf{u}}^{(K)}. \quad (3.3)$$

It is important to note that in writing equation (3.3), we have not made any explicit assumptions about the form of the matrix \mathbf{W} , other than to require it to be a matrix describing a wavelet transform. As a result, we can perform several levels of multiresolution simultaneously through appropriate preparation of the matrix \mathbf{W} in the manner outlined above. Using the inverse

wavelet transform $\mathbf{W}^{-1} = \mathbf{W}^T$ to recover an original configuration \mathbf{u} is a one-to-one mapping only if we are provided with n distinct data points, as contained in $\bar{\mathbf{u}}^{(K)}$.

Examining (3.3), we see that the result of applying the wavelet transform to a set of spins \mathbf{u} is to create a new representation $\bar{\mathbf{u}}^{(K)}$, which contains the same information about the spins as does the original state vector \mathbf{u} . However, the vector $\bar{\mathbf{u}}^{(K)}$ contains $n/2^{dK}$ averages, where d is the dimensionality of the lattice; these averages can be viewed as “block spins” in a sense similar to that of Kadanoff [71]. The remaining elements of $\bar{\mathbf{u}}^{(K)}$ contain the local differences in the spins; that is, they can be used to describe the specific set of spins which give rise to a particular block spin $\tilde{s}_i^{(K)}$.

3.2.2 Computing thermodynamic functions

The canonical partition function

$$Z = \sum_{\mathbf{u} \in \mathbb{S}} \exp(-\beta\mathcal{H}(\mathbf{u})), \quad (3.4)$$

where \mathbb{S} is the configuration space of the system, can be used to derive all the thermodynamic properties of a lattice system. Applying the wavelet transform \mathbf{W} to the state vector \mathbf{u} results in a new state vector $\bar{\mathbf{u}}^{(K)}$ belonging to the configuration space $\tilde{\mathbb{S}}^{(K)}$. Provided that \mathbf{W} satisfies the perfect reconstruction property, if the summation over $\mathbf{u} \in \mathbb{S}$ in (3.4) is replaced with the summation over $\bar{\mathbf{u}}^{(K)} \in \tilde{\mathbb{S}}^{(K)}$, the results will be identical. This result follows naturally, since perfect reconstruction necessarily implies that there is a unique state vector $\bar{\mathbf{u}}^{(K)} \in \tilde{\mathbb{S}}^{(K)}$ for each state vector $\mathbf{u} \in \mathbb{S}$, and by construction of the wavelet-transformed Hamiltonian (3.3), $-\beta\mathcal{H} = -\beta\tilde{\mathcal{H}}^{(K)}$.

Ensemble averages involving wavelet-transformed variables are in general no more complicated than computations involving the original variables. In general, the transformation of a function (or functional) of the original variables $f(\mathbf{u})$ or $f[u]$ will generally have the same characteristics after applying the wavelet transform to obtain $\tilde{f}^{(K)}(\bar{\mathbf{u}}^{(K)})$ or $\tilde{f}^{(K)}[\bar{u}^{(K)}]$. Moreover, the standard properties of ensemble averaging, such as linearity, also apply, which makes calculations of moments of the distribution particularly simple.

The above formulation can be applied to any system whose Hamiltonian can be written in

the form of equation (2.2), or as the sum of contributions, each of which is of that form. While the wavelet transform can be applied to any lattice system whose Hamiltonian is a function of the components of the state vector \mathbf{u} —or, indeed, to any Hamiltonian which is a functional of $\mathbf{u}(\mathbf{r})$ for continuous systems—in most of these cases, it is necessary to rely on the more cumbersome series expansions, or integral transforms in the case of continuous systems.

In addition, for Ising spin variables, the use of the wavelet transform poses an additional challenge. While it is entirely straightforward to describe the possible values an individual lattice site can take—for a spin- q Ising model, the allowed values of an individual spin σ are $-q, -q + 1, \dots, q - 1, q$ —the rules which determine whether an arbitrarily selected “transformed” state vector $\bar{\mathbf{u}}^{(K)}$ represents a real state vector \mathbf{u} are cumbersome to manipulate and to implement, and have proven a formidable challenge in prior research as well [15].

3.3 Analysis

In this section, we use two variants of the two-dimensional Ising model as a basis for our calculations: we look at 4×4 and 32×32 Ising lattices; the former is used for calculations when it is desired to perform calculations over all states explicitly, while the latter is illustrative of larger systems, for which exact calculations are intractable. The emphasis in this chapter will be on the use of wavelets to yield approximate answers in significantly faster time than is possible with a Monte Carlo simulation incorporating all degrees of freedom explicitly. The methodology by which the wavelet transform can be extended to Monte Carlo simulations of lattice systems is the focus of Chapter 4.

Before proceeding to the results of the calculations, we make note of the time required to execute the simulations. Each of Figures 3-1 through 3-4 plots the observed variables in the temperature range $T = 0.50$ to $T = 5.00$, with a step size of $\Delta T = 0.01$. The computations for the original problem, with $2^{16} = 65536$ states to consider explicitly, required more than 6.8 seconds per point to execute the required calculations on a 733 MHz Pentium III; the same calculations using one and two applications of the wavelet transform required just 0.061 and 0.026 seconds per point to consider $5^4 = 625$ and 17 states, respectively.

3.3.1 Weighting functions for wavelet-transformed statistics

Until now, we have worked with wavelet transforms which preserve the number of degrees of freedom between the original and transformed problems. This approach yields results for the wavelet-transformed system in exact agreement with those for the original system. However, as mentioned above, in such cases the transformed equations are usually harder to model than the original ones. Hence it is desirable to use the wavelet transform not only as a means of describing a lattice system, but also to derive an approximation scheme whereby estimates of thermodynamic properties can be made efficiently, while still offering error estimates that bound the true results.

The approach we adopt in this chapter is to ignore all local differences: that is, we assume that $\delta_i^{(k)} = 0$ for all values of i and k . As a secondary assumption, we assume that correlation functions which include wavelet coefficients are also equal to zero: that is, we assume $\langle \delta_i^{(k)} A(\cdot) \rangle = 0$ for any choice of property $A(\cdot)$. These extreme assumptions represent a “worst-case scenario” for the use of the wavelet transform method; any more accurate representation of the behavior of the wavelet coefficients will lead to similarly more accurate results in the calculations we present below. Note that under certain circumstances, these approximations are accurate: for Hamiltonians of the form (2.2) which do not exhibit interactions with an external field, there exist configurations with equal energies but opposite signs for $\delta_i^{(k)}$; consequently, when we take the average over all configurations, the ensemble average of $\delta_i^{(k)}$ will vanish.

Let us assume that we have applied the wavelet transform to our original Hamiltonian, $\mathcal{H}(\mathbf{s})$, and have obtained a new function $\tilde{\mathcal{H}}^{(K)}(\tilde{\mathbf{u}}^{(K)})$. If we then make our approximation that all the wavelet coefficients are negligible, then we can reduce $\tilde{\mathcal{H}}^{(K)}(\tilde{\mathbf{u}}^{(K)})$ to a new function $\tilde{\mathcal{H}}^{(K)}(\tilde{\mathbf{s}}^{(K)})$, where $\tilde{\mathbf{s}}^{(K)}$ represents the set of all averages $(s_1^{(K)}, \dots, s_m^{(K)})$ that were preserved by the wavelet transform. We know that the partition function of the system before we performed the wavelet transform is given by (3.4); after using the wavelet transform, we hypothesize that the partition function of the new system is given by

$$\tilde{Z} = \sum_{\tilde{\mathbf{s}}^{(K)} \in \tilde{\mathcal{S}}^{(K)}} w(\tilde{\mathbf{s}}^{(K)}) \exp\left(-\beta \tilde{\mathcal{H}}^{(K)}(\tilde{\mathbf{s}}^{(K)})\right), \quad (3.5)$$

where the $w(\tilde{\mathbf{s}}^{(K)})$ is the weight of configuration $\tilde{\mathbf{s}}^{(K)}$ in the configuration space $\tilde{\mathcal{S}}^{(K)}$; since

multiple configurations $\mathbf{u} \in \mathbb{S}$ can correspond to the same $\bar{\mathbf{s}}^{(K)}$, we cannot set $w(\bar{\mathbf{s}}^{(K)}) = 1$ as we did in the untransformed Ising model. Clearly, the new partition function (3.5) will be identical to the original partition function (3.4) if we define

$$w(\bar{\mathbf{s}}) = \sum_{\mathbf{u}:\mathbf{u}\rightarrow\bar{\mathbf{s}}^{(K)}} \exp\left(-\beta\left(\mathcal{H}(\mathbf{u}) - \tilde{\mathcal{H}}^{(K)}(\bar{\mathbf{s}}^{(K)})\right)\right), \quad (3.6)$$

where $\mathbf{u} : \mathbf{u} \rightarrow \bar{\mathbf{s}}^{(K)}$ denotes that the sum is performed over all configurations \mathbf{u} which have the same set of wavelet-transformed averages $\bar{\mathbf{s}}^{(K)}$. However, evaluating (3.6) to obtain the weighting functions is no more tractable than computing the original partition function. Thus, any computational efficiency to be gained is by finding an economical approximation for (3.6). One approach is to take $w(\bar{\mathbf{s}}^{(K)})$ to be equal to the number of states \mathbf{s} whose wavelet transform yields $\bar{\mathbf{s}}^{(K)}$, so that (3.5) is the standard form of the canonical partition function for a system with energy degeneracies. Let us call the number of such states \mathbf{s} with equivalent averages $\bar{\mathbf{s}}^{(K)}$ the degeneracy of state $\bar{\mathbf{s}}^{(K)}$, and denote this degeneracy as $g(\bar{\mathbf{s}}^{(K)})$. The restriction of $\bar{\mathbf{u}}^{(K)}$ to the averages $\bar{\mathbf{s}}^{(K)}$ prevents a unique reconstruction of \mathbf{u} , unless $g(\bar{\mathbf{s}}^{(K)}) = 1$.

3.3.2 Order parameter

A natural variable to compute is the order parameter, generically denoted η , which for lattice spin systems is the average magnetization,

$$m = \frac{1}{N} \left\langle \sum_i \sigma_i \right\rangle; \quad (3.7)$$

for other members of the Ising universality class, the order parameter can represent the overall density ρ or the difference between the densities of two phases [146]. While the computation of the order parameter is straightforward in simulations, its calculation can be made difficult in the case of zero external field because of symmetries in the configuration space: for every configuration with magnetization m_i , there exists another configuration with the same energy and magnetization $-m_i$. As a result, when all the states are combined using (3.7), we do not observe spontaneous magnetization, but instead find $m = 0$ at all temperatures. Thus, we consider the absolute value of the magnetization, $|m|$ in place of the magnetization m . The

results of this calculation for the 4×4 Ising model with the wavelet transform, setting all $\delta(i)$'s to zero, and without the wavelet transform are shown in Figure 3-1. We note that the error is essentially negligible for temperatures below $T = 1$, and decreases again for large values of T , where differences in energy levels become negligible and the average magnetization of a state is the only contributing factor to $|m|$.

3.3.3 Free-energy considerations

The Helmholtz free energy is just the logarithm of the partition function: $A = -k_B T \ln Z$; if we estimate the partition function using an expression such as (3.5), we naturally expect the approximated value A to differ from its true value. When we examine the behavior of the 4×4 Ising model under the wavelet transform, we find as expected that the two free energy surfaces are similar, although they are clearly not identical. In particular, the exact numerical values obtained from the two equations are not the same. However, since the assignment $A = 0$ is arbitrary in any system, we can choose to define $A = 0$ as either the maximum or the minimum free energy obtained in each system. Under these conditions, the energy scales for the exact calculation and the wavelet transform calculation are essentially identical, particularly at the so-called “fixed points” of the system—that is, for points which are not affected by renormalization transformations [104]. For the two-dimensional Ising model, these points are at $T = 0$, and at infinite (positive or negative) values of the external field interaction h .

This agreement at the fixed points should not be surprising: the fixed points correspond to unique configurations of the system, such that for the configurations $\bar{\mathbf{s}}^{(K)}$ which result from the wavelet transform of the fixed points \mathbf{u}^* , the degeneracy of the states is unity. Consequently, the approximation $w(\bar{\mathbf{s}}^{(K)}) = g(\bar{\mathbf{s}}^{(K)})$ is correct for the dominant configuration in the system, and therefore the behavior of the approximate partition function Z' is almost identical to that of the true partition function Z in the vicinity of the fixed points.

As can be seen in Figure 3-2, the free energy converges to the same values in the low-temperature limit, where only a few states which are essentially unaffected by the wavelet transform contribute to the partition function of the system. At high temperatures, the agreement is less exact, because of the approximations made in evaluating the Boltzmann factors of the block spin configurations.

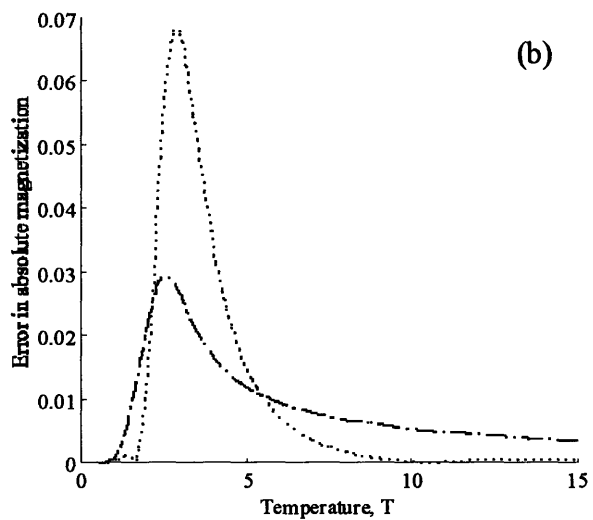
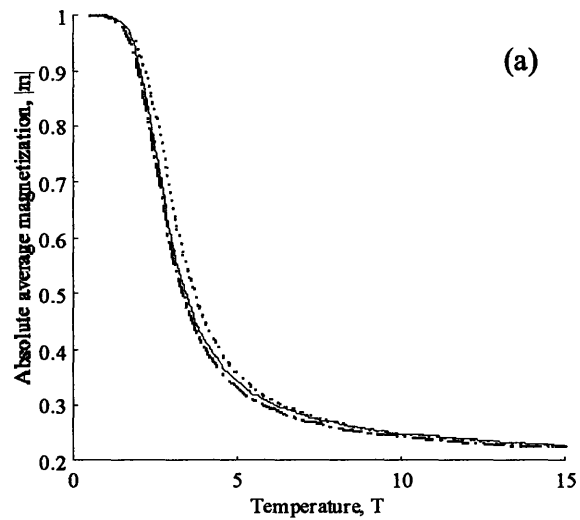


Figure 3-1: (a) Average absolute magnetization $|m| = \sum_i |\sigma_i|$ of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dot-dashed and dotted lines, respectively). (b) Error in the average absolute magnetization of the 4×4 Ising model for one and two iterations of the two-dimensional Haar wavelet versus an exact enumeration.

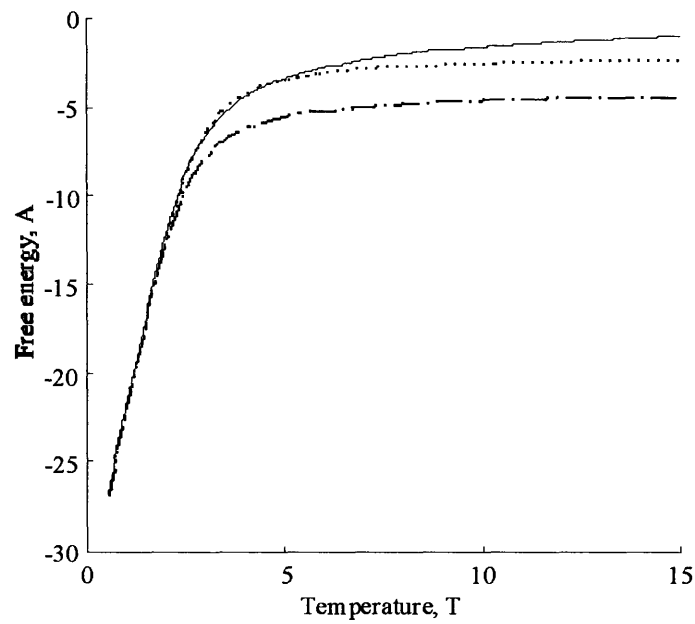


Figure 3-2: Free energy A of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively).

3.3.4 Entropy of a wavelet-transformed lattice system

It is relatively straightforward to show that coarse-graining a system with a countable phase space—that is, a phase space whose states can be completely enumerated—requires a correction factor to ensure satisfactory agreement with results from the fine-scale system. As an example, consider the entropy of a wavelet-transformed spin- $\frac{1}{2}$ Ising model. The original lattice has 2^{N_t} total configurations, and therefore in the high-temperature limit, the entropy approaches

$$S_{\max} = k_B N_t \ln 2. \quad (3.8)$$

If we ignore the weighting factor of the wavelet-transformed system by setting $w(\bar{\mathbf{s}}^{(K)}) = 1$ for all states $\bar{\mathbf{s}}^{(K)}$, the resulting system will have $\left(1 + \prod_{i=1}^{K-1} N^{(i)}\right)^{N^{(K)}}$, and the maximum possible entropy for this system is

$$S_{\max}^{(K)} = k_B N^{(K)} \ln \left(\prod_{i=1}^{K-1} N^{(i)} + 1 \right). \quad (3.9)$$

Comparing (3.9) to the original limit of $S_{\max} = k_B N_t \ln 2$, we must have $S_{\max} > S_{\max}^{(K)}$, since the configuration space is smaller, which indicates that our coarse-graining has effectively reduced the available entropy of the system. We can conclude from this that unless the volume of phase space is conserved by the coarse-graining procedure, some entropy will be lost at high temperatures, regardless of the accuracy of the coarse-graining. Our choice of the degeneracy $g(\bar{\mathbf{s}}^{(K)})$ for the weighting function $w(\bar{\mathbf{s}}^{(K)})$ preserves the volume of phase space in the limit of $T \rightarrow \infty$ for the systems under study here.

Looking at the specific case of the 4×4 Ising model, we compute the entropy as a function of temperature for zero external field for the exact problem, and for the wavelet-transformed problem using one and two iterations of the two-dimensional Haar wavelet. The results are shown in Figure 3-3. As before, errors vanish in the low-temperature limit, where only the lowest-energy states make a contribution to the partition function; since there are two such states, namely those with all spins up or all spins down, we find the $T \rightarrow 0$ limit of the entropy:

$$S(T \rightarrow 0) \approx - \sum_{i=1}^2 \frac{1}{2} \ln \frac{1}{2} = \ln 2.$$

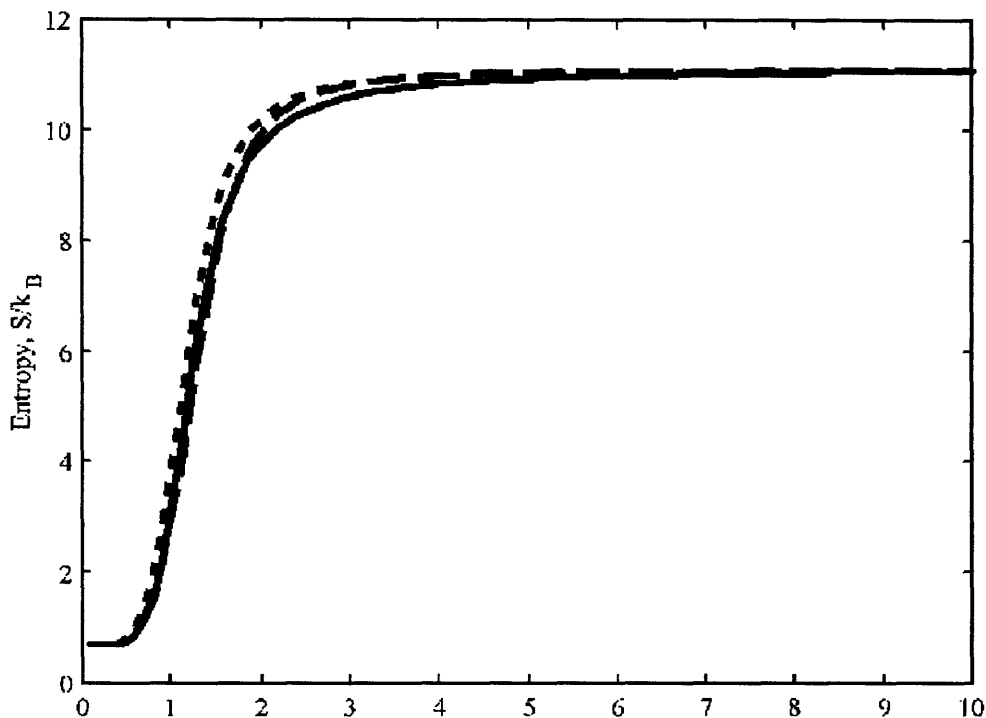


Figure 3-3: Entropy of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively). The bottom of the y -axis corresponds to the zero-temperature limit of $S = \ln 2$.

In the high-temperature limit, the entropies agree since, by construction, we have $\sum_{\bar{s} \in \tilde{\mathcal{S}}} w(\bar{s}) = 2^{16}$, and therefore, since the Boltzmann factor goes to unity for all states as $T \rightarrow \infty$, $Z \rightarrow \sum_{\bar{s} \in \tilde{\mathcal{S}}} w(\bar{s})$, and the entropy tends toward the value S_{\max} given by (3.8). Disagreement in the intermediate temperature regime is largely the result of grouping together states with different energies into a single transformed state with a single energy. In particular, the grouping of higher-energy states together at a lower energy increases the probability of those states, and therefore increases the total entropy of the system. This explains the seemingly anomalous result of coarse-graining increasing the entropy at intermediate temperatures observed in Figure 3-3.

For more general systems, we can consider the formal definition of the entropy for a continuous distribution,

$$S = -k_B \int_{\mathbf{u} \in \mathcal{S}} d\mathbf{u} p(\mathbf{u}) \ln p(\mathbf{u}), \quad (3.10)$$

where $p(\mathbf{u})$ is the probability of observing the configuration \mathbf{u} . For the original Ising model, each configuration \mathbf{u} is unique, and therefore the probability $p(\mathbf{u})$ is just the Boltzmann weight $Z^{-1} \exp(-\beta\mathcal{H}(\mathbf{u}))$, and thus the entropy is given by

$$S = k_B \ln Z + \frac{k_B}{Z} \int_{\mathbf{u} \in \mathcal{S}} d\mathbf{u} \beta \mathcal{H}(\mathbf{u}) \exp(-\beta \mathcal{H}(\mathbf{u})). \quad (3.11)$$

After performing the wavelet transform, and discarding finer-scale details, we must account for the weighting factor $w(\bar{\mathbf{s}})$ in our expression for the entropy. Consequently, the entropy for such a system is given by

$$\begin{aligned} S^{(K)} &= -k_B \int_{\bar{\mathbf{s}} \in \tilde{\mathcal{S}}} d\bar{\mathbf{s}} w(\bar{\mathbf{s}}) p(\bar{\mathbf{s}}) \ln p(\bar{\mathbf{s}}) \\ &= k_B \ln \tilde{Z} + \frac{k_B}{\tilde{Z}} \int_{\bar{\mathbf{s}} \in \tilde{\mathcal{S}}} d\bar{\mathbf{s}} w(\bar{\mathbf{s}}) \beta \tilde{\mathcal{H}}^{(K)}(\bar{\mathbf{s}}^{(K)}) e^{-\beta \tilde{\mathcal{H}}^{(K)}(\bar{\mathbf{s}}^{(K)})}, \end{aligned} \quad (3.12)$$

where the coarse-grained partition function \tilde{Z} is given by (3.5).

3.3.5 Fluctuation properties

For a given thermodynamic system, the constant-volume heat capacity C is defined as the fluctuation in the internal energy of the system:

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}. \quad (3.13)$$

It is well-known that in the vicinity of a continuous phase transition, the heat capacity C diverges. At the same time, it is also known that no system of finite size can have $C \rightarrow \infty$, since the energy, and hence the variance $\langle E^2 \rangle - \langle E \rangle^2$ in the energy of the system will also remain finite. However, we can still observe evidence of power-law divergence near the critical temperature. [19] For small finite systems, we do not observe evidence of a divergence in the heat capacity; instead, the heat capacity is a smooth function of temperature T . The wavelet transform largely preserves the behavior of the original model: we observe the same general functional form in both the original and the transformed systems, as seen in Figure 3-4. However, for large finite systems, we can still detect the characteristic power-law divergence in the vicinity of the critical point; such a divergence is shown for a 32×32 Ising model as Figure 3-5.

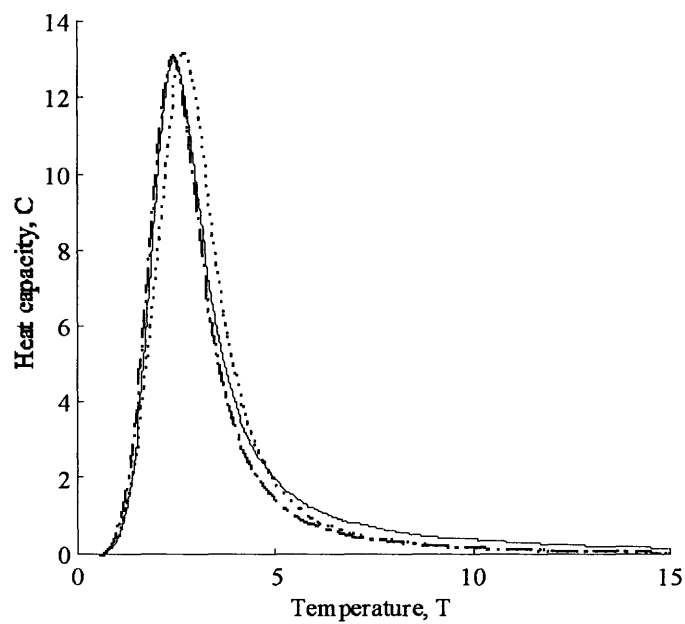


Figure 3-4: Heat capacity of the 4×4 Ising model at zero external field as a function of temperature, as computed by an exact enumeration using no wavelet transform (solid line), and using one and two iterations of the two-dimensional Haar wavelet (dashed and dot-dashed lines, respectively).

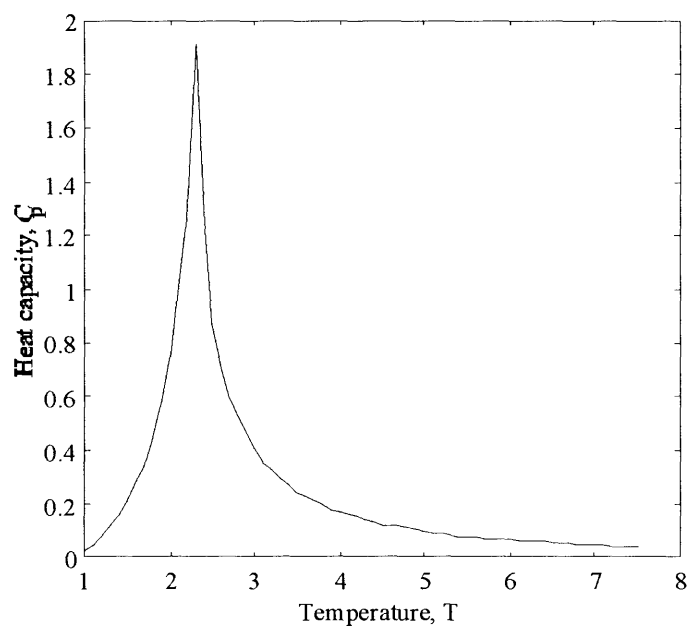


Figure 3-5: Heat capacity of a 32×32 Ising model, illustrating the power-law divergence in the vicinity of the critical point.

Applying the wavelet transform to a given thermodynamic system, we still expect to find evidence of a critical point; however, because of the mean-field like behavior of the wavelet transform, the critical point is found at a higher temperature in the wavelet-transformed system than in the original system. Computing the heat capacity for large lattices cannot be done by exhaustive enumeration; therefore, we save discussion of our numerical results for Chapter 4, and present a brief heuristic argument. As the size of a block spin increases, corresponding to a larger number of iterations of the wavelet transform, we expect that suppression of fluctuations in the system will lead to an increase in the phase transition temperature, while the actual numerical value of the heat capacity itself decreases. The decrease in the number of degrees of freedom in the system is accompanied by an increasing trend towards homogeneity of the system—the remaining configurations begin to look more and more similar to one another. As a result, the variance measured by $\langle E^2 \rangle - \langle E \rangle^2$ decreases with increasing amounts of coarse-graining. This trend is much less pronounced for small systems, as shown in Figure 3-4. When the number of degrees of freedom is small, the ability to perform an exact enumeration ensures that the only errors observed are those introduced by the wavelet approximation itself. Thus, since the free energy and entropy of small systems are accurately reproduced by the transform, only minor deviations in the phase transition temperature for small systems are expected.

3.4 Similarities to Renormalization Group theories

It should already be apparent that close affinities exist between the method described here and position-space renormalization methods. The construction of wavelet-based averages naturally corresponds to Kadanoff’s concept of a “block spin” transformation [71]: both the present method and Kadanoff’s approach rely on combining a region of contiguous spins into a single new “block spin.” However, with the approach outlined above, we do not seek to impose the requirement that the block spins must be restricted to have the same spin values as the original spins. For example, in a wavelet-transformed spin- $\frac{1}{2}$ Ising model, the block spins can take on values other than +1 and -1. In addition, we do not impose “majority rules” or other “tiebreakers” in the case where there are equal numbers of “up” and “down” spins contained in a single block.

A less apparent connection can also be established with Migdal’s “bond-moving” approximation [114, 115], or similarly with Wilson’s recursion method [160]. In particular, Migdal’s method can be compared to applying a separable wavelet transformation, in that the recursion is performed in only one spatial direction at a time. However, using the methods of either Migdal or Wilson, after performing the bond-moving or decimation transformation, the new interaction strengths J'_{ij} are determined by manipulating the resulting Hamiltonian and recasting it in the form of the original, leading at times to very complicated, nonlinear formulas for the J'_{ij} as a function of J_{ij} . By contrast, using wavelets, the values for the transformed coupling coefficients \tilde{J}_{ij} are obtained directly by transforming the interaction matrix \mathbf{J} , since $\bar{\mathbf{J}} = \mathbf{W}\mathbf{J}\mathbf{W}^T$. The trade-off that must be made for the algorithmic transparency of obtaining the new coupling constants via a matrix multiplication is that we cannot solve for the fixed points of the recursion relation, which makes the production of a renormalization “flow diagram” using the wavelet method a difficult problem.

The behavior of the wavelet transform and of the renormalization group differs in another, important manner. By the nature of its construction, the group of renormalization transformations is only a *semi-group*, inasmuch as reversing the mapping to move from a coarsened description to a more detailed description is impossible [38, 55]. Using the wavelet transform, a reverse mapping is theoretically possible: reversibility fails because of the approximations invoked to reduce the number of wavelet coefficients which are kept, rather than as an inherent limitation of the wavelet transform itself.

3.5 Conclusions

Using the wavelet transform as a basis for calculations of lattice systems yields an impressive reduction in the computational time required to obtain estimates of thermodynamic properties, at the price of modest errors in accuracy, given the relatively small size of the systems considered here, and the approximations introduced to simplify our calculations. The wavelet transform provides a systematic approach to coarse-grain systems to any level of resolution. The method produces correct results in the vicinity of fixed attractors, with decreasing accuracy as one approaches the critical point in parameter space.

In addition, the use of the wavelet transform on such systems is simple to implement: after selecting a particular set of wavelets, all of the operations can be reduced to matrix multiplications, as shown in Section 3.2.1. The computational requirements needed to implement the transformation are small: if the transformation cannot be accomplished in the explicit form of a matrix multiplication, it is possible to organize the calculation to be performed in-place [153]. Moreover, the amount of coarse-graining that can be achieved using the wavelet transform can be adjusted dynamically. These properties of the wavelet transform lead naturally to an extension of the method to systems for which simulations are required; the resulting algorithm is the focus of the following chapter.

Chapter 4

Multiresolution analysis in statistical mechanics. II. The wavelet transform as a basis for Monte Carlo simulations on lattices

4.1 Introduction

In the preceding chapter, we discussed wavelet-accelerated Monte Carlo (WAMC), a new method for multiscale simulations, and showed how to apply it as a topological tool to polymer simulations, the simplest case being the freely-jointed chain. The real power of the algorithm, however, lies in its ability to handle non-bonded interactions between particles, as is found in self-avoiding chains. Thus, in this chapter, we extend our development of the algorithm. We show that applying the wavelet transform to the hard-sphere potential justifies the use of soft potentials, and that, using the freely-jointed chain model, we can develop phenomenological scaling laws which can be used to estimate new coarse-grained interatomic potentials using previously computed data. We also demonstrate that our method reproduces expected scaling results for self-avoiding chains.

4.2 The Wavelet-Accelerated Monte Carlo (WAMC) algorithm

The principal difficulty of performing a wavelet transformation on a lattice system is in working with the discrete set of values that each spin is permitted to take, such as in a spin- q Ising model. Because the transformed variables are linear combinations of the original variables, the constraint that the spins of the individual sites on the original lattice must be drawn from the set $\{-q, -q + 1, \dots, q - 1, q\}$ quickly becomes a more complicated constraint on the transformed variables \tilde{s}_i and $\tilde{\delta}_i$. As the system size becomes large, the difficulty of rewriting the spin constraints proves so great that previous investigations of the use of wavelets in statistical field theory ignored Ising models altogether [15]. Consequently, we would like, if at all possible, to avoid computations involving original states after we have carried out the wavelet transformation. At this point, we take note of the application of wavelets to image compression, where the goal is to reduce the amount of information needed to reconstruct an image. We would like to apply this technique to lattice systems, and reduce the number of degrees of freedom which must be accounted for in our calculations.

We consider our system to be a d -dimensional regular lattice \mathcal{L} with side length l , so that the size of the lattice is $N = |\mathcal{L}| = l^d$, and we let a site σ_i on the lattice \mathcal{L} be characterized by a “spin” chosen from a finite set \mathcal{J} of values and by its physical location on the lattice. For the spin- $\frac{1}{2}$ Ising model, for example, the set \mathcal{J} is just $\{+\frac{1}{2}, -\frac{1}{2}\}$ (although for computational convenience this is usually treated as $\{+1, -1\}$, a convention which we follow below); similarly, for a lattice gas based on a spin-1 Ising model, $\mathcal{J} = \{0, 1, 2\}$ represents the allowed occupation numbers of each lattice site. We then assume that the only physical interactions that occur are either interactions with an external field h_i which can vary at each lattice site, or pairwise interactions with the bilinear form $U(\sigma_i, \sigma_j) = J_{ij}\sigma_i\sigma_j$, where J_{ij} is usually a function only of the spacing between sites i and j . Consequently, the Hamiltonian of the system can be written in the form

$$-\beta\mathcal{H} = \sum_i h_i\sigma_i + \sum_i \sum_j J_{ij}\sigma_i\sigma_j. \quad (4.1)$$

For the purposes of our simulations, however, we will find it more convenient to treat the set of spins $(\sigma_1, \dots, \sigma_N)$ and the external field (h_1, \dots, h_N) as vectors \mathbf{u} and \mathbf{h} , and the pairwise interaction strengths J_{ij} as a matrix \mathbf{J} . Then the Hamiltonian (4.1) can be written in matrix

form as

$$-\beta\mathcal{H} = \mathbf{h}^T \mathbf{u} + \mathbf{u}^T \mathbf{J} \mathbf{u}. \quad (4.2)$$

This formulation of the problem is similar in spirit to that of graph theory, where the pairwise potential J_{ij} is used to generate an adjacency list which specifies which edges interact [31]. Using (4.2) as the basis for a Monte Carlo simulation requires the calculation of the change of energy ΔE_{nm} from microstate \mathbf{u}_m to microstate \mathbf{u}_n :

$$\Delta E_{nm} = \mathbf{h}^T (\mathbf{u}_n - \mathbf{u}_m) + (\mathbf{u}_n - \mathbf{u}_m)^T \mathbf{J} (\mathbf{u}_n - \mathbf{u}_m). \quad (4.3)$$

If moves are restricted to changes of single spins, then only a single entry of $\mathbf{u}_n - \mathbf{u}_m$ is nonzero, and the calculation (4.3) reduces to a dot product, instead of a matrix multiplication.

As described in I, the action of the wavelet transform is to insert between each product in (4.2) or (4.3) the identity matrix in the form $\mathbf{I} = \mathbf{W}^T \mathbf{W}$, where \mathbf{W} is the wavelet transform which maps data from one scale to the next coarser scale, containing half as many data points. The resulting expressions rewrite the Hamiltonian in terms of wavelet-transformed averages and differences, with downsampling needed to reduce the number of variables from $2N$ to N . As before, the wavelet transform can be iterated by applying it to successive sets of averages, leading after K iterations to Hamiltonians of the form:

$$-\beta\tilde{H} = \left(\bar{\mathbf{h}}^{(K)}\right)^T \bar{\mathbf{u}}^{(K)} + \left(\bar{\mathbf{u}}^{(K)}\right)^T \bar{\mathbf{J}}^{(K)} \bar{\mathbf{u}}^{(K)}, \quad (4.4)$$

where in (4.4) the $\bar{\mathbf{u}}^{(K)}$ represent “block spins” whose values are determined by wavelet averaging over some well-defined region of the original system. The Hamiltonians (4.2) and (4.4) have the same formal structure, so that Monte Carlo simulations of the two systems are essentially identical. The only modifications needed to simulate a coarse-grained Hamiltonian are the ability to select new microstates $\bar{\mathbf{u}}_i^{(K)}$ which are generated through wavelet transformations of the original microstates \mathbf{u}_i , and the elimination of unwanted degrees of freedom from (4.4). It should be noted that in (4.4), the elements of $\bar{\mathbf{u}}^{(K)}$ are not restricted to the same values as in the original system, but are free to take on any value which is consistent with the wavelet transform applied to the system.

From above, we saw that for even Hamiltonians $\tilde{H}(\bar{\mathbf{x}})$, we should have that $\langle \delta \rangle_{\tilde{H}} = 0$ for any wavelet difference δ , where $\langle \cdot \rangle_{\tilde{H}}$ denotes the ensemble average weighted by the Hamiltonian \tilde{H} . As a “worst-case scenario” for our method, we shall assume not only that $\langle \delta \rangle_{\tilde{H}} = 0$, but also that any terms in the Hamiltonian (4.4) containing fluctuation terms can be neglected as well. This assumption allows us to reduce the size of $\bar{\mathbf{J}}^{(K)}$ from $N \times N$ to $2^{-Kd}N \times 2^{-Kd}N$, where d is the lattice dimensionality. Consequently, instead of performing calculations involving all of the original variables \mathbf{x} which describe the state of our system, we consider functions only of local averages of our original variables. However, we anticipate that this simplification of the interactions present in the system will have a significant impact on the thermodynamic behavior of the resulting system; we illustrate these effects below.

To generate the new microstates $\bar{\mathbf{u}}_i^{(K)}$, we need an estimate for the probability distribution $p(\tilde{u}_i^{(K)})$ which describes the individual sites in the coarse-grained lattice. Determining the correct distribution for a given $\tilde{u}_i^{(K)}$ would require a detailed simulation of the original system. An alternative, ignoring the effect of neighboring block spins, would be to perform an exact enumeration of the spins within a block, which is possible only for the smallest of block spins. Since we would like to apply this method to systems of arbitrary size, we want to avoid both of these options. Therefore, we simulate a sublattice with the same dimensions as $\tilde{u}_i^{(K)}$, ignoring physical interactions with the rest of the system by using either free or periodic boundary conditions. Using the standard Metropolis acceptance criterion, we compute distributions for the properties of the small lattice, such as the magnetization. Then, according to the matrix formulation described in the previous chapter, since the wavelet transform defines a single block spin $\tilde{u}_i^{(K)}$ as a linear function of the individual spins at level $K - 1$ which it replaces, we can use the linearity properties of probability distributions to convert the distribution of the properties directly into a distribution for the block spin $\tilde{u}_i^{(K)}$ [18]. Finally, using the distribution for the block spin $\tilde{u}_i^{(K)}$ as a starting point, we perform a Monte Carlo simulation on the system of block spins defined by the Hamiltonian (4.4).

Although (4.2) and (4.4) are structurally the same, we cannot impose a one-to-one correspondence between the states in the configuration space of (4.3) and the states in the configuration space of (4.4). Consequently, the thermodynamic information obtained from the two will not necessarily be identical; as we have shown in Chapter 3, there is under fairly broad conditions

a loss of entropy associated with the application of coarse-graining to a system. We can ensure that the detailed balance condition for the simulation based on (4.4) is satisfied for the new simulation by requiring

$$\frac{\alpha\left(\bar{\mathbf{u}}_m^{(K)} \rightarrow \bar{\mathbf{u}}_n^{(K)}\right)}{\alpha\left(\bar{\mathbf{u}}_n^{(K)} \rightarrow \bar{\mathbf{u}}_m^{(K)}\right)} = \frac{p\left(\bar{\mathbf{u}}_n^{(K)}\right)}{p\left(\bar{\mathbf{u}}_m^{(K)}\right)} e^{-\beta\left(\mathcal{H}\left(\bar{\mathbf{u}}_n^{(K)}\right) - \mathcal{H}\left(\bar{\mathbf{u}}_m^{(K)}\right)\right)},$$

where $\alpha(m \rightarrow n)$ is the probability of accepting a move from microstate m to microstate n , and $p(m)$ is the probability of selecting microstate m as determined from simulations on finer-grained lattices at lower scales.

4.3 Theoretical performance of WAMC versus traditional MC

The wavelet transform is a *hierarchical* method which can be applied iteratively to a system to obtain successively coarser descriptions of a system. To describe the operation of the wavelet transform on a lattice model, we need to introduce some notation based on the various length scales in the problem. In the original problem, the applicable length scales are the lattice spacing l , the correlation length ξ , and the total lattice size L . Applying the wavelet transform method once increases the lattice spacing by some factor a , so that the ratios of correlation length to lattice spacing and of system length to lattice spacing each decrease by a . If we apply the wavelet transform m times in succession, the corresponding factor becomes a^m .

We perform the simulation in a series of K stages, where the length scales at each stage are functions of the length scales at the previous stages. The initial simulation is performed on a sublattice of the original problem, with lattice size $L^{(1)} < L$, where the superscript denotes the first stage of the simulation. The lattice spacing of the first stage is the same as in the original problem, so we define $l^{(1)} = l$. At each subsequent stage of the simulation, the lattice spacing of the k th stage is defined by the recursive relation $l^{(k)} = L^{(k-1)}l^{(k-1)}$. Since $l^{(1)}$ is fixed to be the lattice spacing of the original lattice, the adjustable parameter in this relation is the lattice size $L^{(k)}$ of each stage. If we assume that the lattice is the same length in all dimensions at every stage, a single variable in stage k is a block variable representing the $(L^{(k-1)})^d$ variables simulated in stage $k - 1$.

Assuming that the lattice is the same length in all directions both in the original problem and at every stage in the wavelet-transformed problems, there are $N_t = L^d$ lattice variables in the original problem, and $N^{(k)} = (L^{(k)})^d$ lattice variables in the k^{th} stage of the wavelet-transformed problem. However, each variable in stage k is a block variable representing the average behavior of the $(L^{(k-1)})^d$ variables in a block at stage $k - 1$, so the number of *total* degrees of freedom represented at stage k is $N_t^{(k)} = \prod_{i=1}^k N^{(i)}$, where $N_t^{(K)} = N_t$. The number of *simulated* degrees of freedom is $N_s = N_t$ for traditional Metropolis Monte Carlo (MMC), but $N_s = \sum_{i=1}^K N^{(i)}$ for WAMC. Because the running time of Monte Carlo simulations is usually linear in the number of degrees of freedom being simulated, the advantage of coarse-graining the system using a wavelet transform becomes evident. For example, consider an “original problem” of simulating a cubic lattice with 256 Ising variables on a side. If we divide the original problem into two stages consisting of cubes of 16 Ising variables on a side, we reduce the original problem of analyzing $256^3 = 16,777,216$ variables to the simpler problem of analyzing $2(16^3) = 8192$ variables. Although it is more difficult to produce a trial configuration in a simulation of the wavelet-transformed problem than in a simulation of the original Ising lattice problem, this is more than offset by the reduction in the number of degrees of freedom being simulated.

4.4 Results

For the purposes of comparison, our “experimental” systems are two-dimensional Ising models of size 32×32 , where we have run both MMC simulations on the full lattice, and WAMC simulations at a variety of resolutions; we shall denote these resolutions using the notation (x, y) , where x indicates the length of the block size simulated in the first stage to estimate the probability distribution $p(\bar{\mathbf{u}}^{(K)})$ to be used in the second stage, and y denotes the number of blocks on a side of the lattice in the second stage of the simulation.

4.4.1 Order parameter

Usually, the property of greatest interest in a simulation of a lattice system is the order parameter η . For spin systems, η is generally taken to be the magnitude of the average magnetization, so that $\eta = \langle m \rangle$. [For XY and Heisenberg models, and other models where spins are oriented,

we generally consider only the magnitude of the average vector $\eta = \langle m \rangle = \langle |\mathbf{m}| \rangle$.] Generally, this is a very simple property to compute, since the value of the order parameter is constantly updated during the course of the simulation, and is thus always available.

For the 32×32 Ising model, the results of a MMC simulation, as well as (4, 8)- and (8, 4)-WAMC simulations are shown as Figure 4-1. The primary difference in the curves for the three cases is that as the coarse-graining process decreases the number of degrees of freedom in the final stage of the simulation, the location of the Curie temperature, indicating onset of spontaneous magnetization, increases and the steepness of the curve below the Curie temperature decreases. This result is consistent with our findings for average absolute magnetization $\langle |m| \rangle$ from analytical models, discussed in Chapter 3. In the present case, we note further that we achieve agreement between the different models not only in the low-temperature region, but also in the high-temperature regime $T \gg T_c$. The differences in the intermediate regime can be attributed largely to the difference in behavior that results from the use of the wavelet transform to move from the original Hamiltonian (4.2) to a coarse-grained Hamiltonian (4.4). Additionally, the increased noise in the WAMC results at intermediate and high temperatures arises because of the approximations used for the probability distributions $p(\mathbf{u}^{(K)})$ at the second stage of the simulation. The relative lack of noise in the MMC results stem in part from the fact that the Metropolis technique leads to non-ergodic sampling of phase space as temperature increases, as the simulation tends to cycle through a limited number of states [91].

4.4.2 Internal energy

Plotting the internal energy $\langle U \rangle$ as a function of the temperature, we obtain curves that follow the same general pattern outlined in Chapter 3. As illustrated in Figure 4-2, at low temperatures, the internal energy, as computed for the 32×32 model using standard MC as well as (4, 8)- and (8, 4)-WAMC simulations, is in exact agreement for all methods. This occurs because only a few microstates of the system, corresponding to states that have all spins aligned, are actually observed by the system, and the wavelet transform preserves the energy of these states exactly. All three eventually reach an average internal energy of zero, but exact agreement is only expected in the infinite-temperature limit, when the difference in energy levels between microstates becomes unimportant. For intermediate temperatures, as before, the disagreement

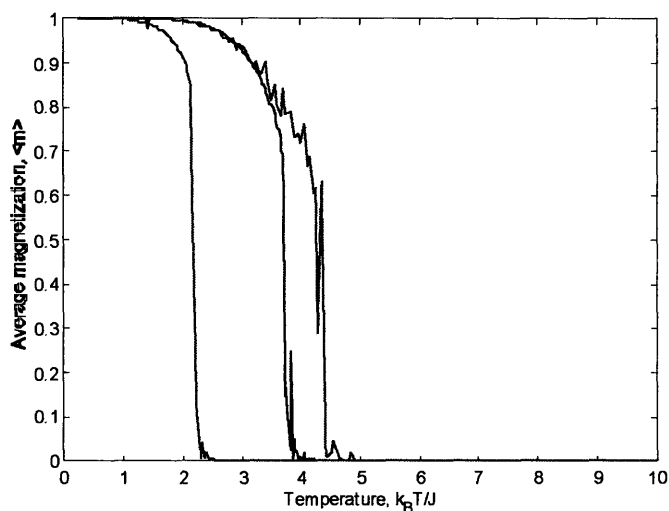


Figure 4-1: Absolute average magnetization as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a (4,8)-WAMC simulation (center), and a (8,4)-WAMC simulation (right).

is a result of the change in form of the Hamiltonian that results from neglecting local correlations. Also, we note that for WAMC the “noise” in the internal energy increases both with increasing proximity to the “observed” critical point of the system as well as with increasing coarse graining. The additional coarse graining yields a Hamiltonian with reduced numbers of energy levels, since the energy of a block spin is defined here to be a function only of its overall magnetization, and not of its internal magnetization fluctuations; the reduced number of discrete energy levels yields noisier data.

4.4.3 Fluctuation properties

Fluctuation properties are useful for locating critical points since in the vicinity of a critical point, the magnitude of fluctuation properties is known to diverge as $t^{-\alpha}$, where $t \equiv |T - T_c|/T_c$ [124]. Thus, a rapid increase in the value of a fluctuation property such as the heat capacity at constant external field, $C_H = (\langle E^2 \rangle - \langle E \rangle^2) / k_B T^2$, with respect to temperature can be used to estimate the critical temperature of a system. However, use of the wavelet transform leads to a decrease in the magnitude of the heat capacity, since the coarse-graining leads to smaller variances in the distribution of the energy $\langle U \rangle$. Consequently, the maximum value of the heat

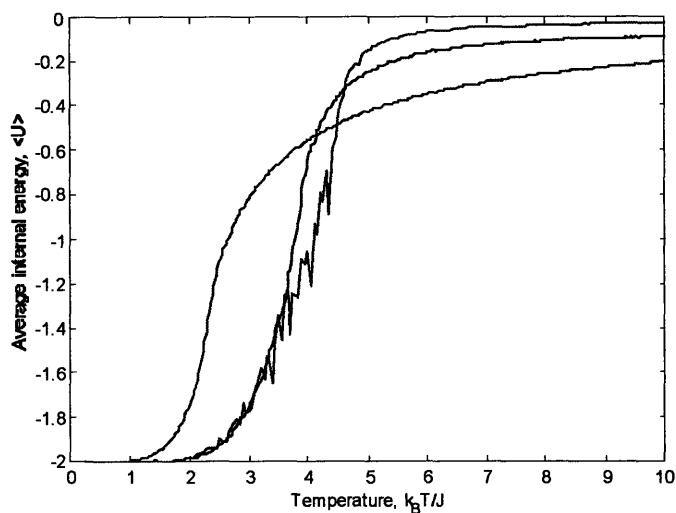


Figure 4-2: Internal energy as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a (4, 8)-WAMC simulation (center), and a (8, 4)-WAMC simulation (right).

capacity C_{\max} decreases as a function of the number of degrees maintained in the problem.

In Figure 7-1, the heat capacity is shown as a function of dimensionless temperature $k_B T/J$ for the same systems as for the order parameter and internal energy measured above. We see that the location of the maximum of the heat capacity does increase, as expected. Although the relative maximum of the heat capacity obtained from the (4, 8)- and (8, 4)-WAMC simulations appears to be identical, they differ by about 3 percent. Moreover, the actual value of the maximum is not as important as its existence and its location as a function of $k_B T/J$ and of the resolution of the model.

4.4.4 Coarse-grained entropy

In the previous chapter, we emphasized that in the low- and high-temperature limits, the entropy of the system should be preserved under the wavelet-transform method. However, looking at Figure 4-4, which plots the entropy of a 32×32 Ising model calculated via direct simulation and using an (8, 4)-WAMC simulation, we do not observe this behavior. Instead, we see that for a given temperature, the entropy predicted by the coarse-grained model is always lower than that determined by direct simulation.

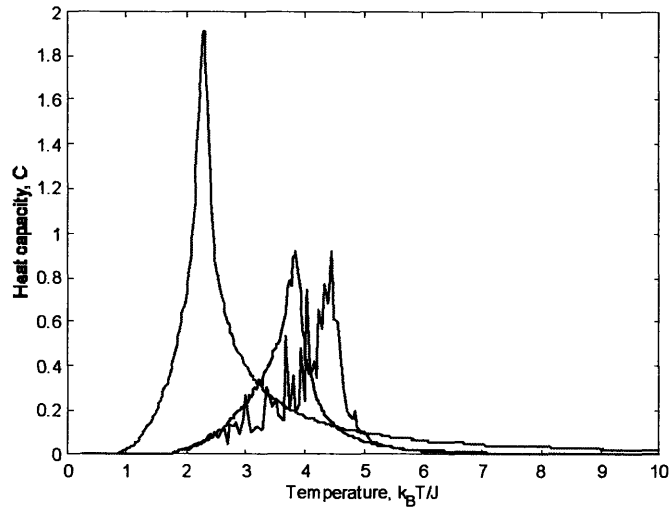


Figure 4-3: Heat capacity as a function of the dimensionless temperature $k_B T/J$ for the 32×32 Ising model computed using standard MC (left curve), a (4, 8)-WAMC simulation (center), and a (8, 4)-WAMC simulation (right).

Why does this not contradict our previous claim? We know that the problem does not occur at low temperatures, since the system is effectively trapped in those configurations with the lowest energies—the ones that have almost all spins up or all spins down. Similarly, the problem does not occur at high temperatures, as demonstrated by Figure 4-5, which plots the expected distribution of states in an 8×8 block as a function of its average magnetization. For very high temperatures, the agreement between the distribution of states yielded from a direct simulation of an 8×8 block versus the expected distribution is quite good. This is to be expected since for high temperatures, we can ignore the relative energies of the different configurations as the Boltzmann weight of each state goes to unity, and thus we are left with merely a count of the various energy states in the system. Thus, using Eq. (3.12), we can determine the entropy exactly at both low and high temperatures.

Therefore, we know that the problem is a result of the disagreement at intermediate temperatures. But why does inaccuracy at intermediate temperatures affect the entropy of the limits? The answer is that for systems too large to be enumerated explicitly, we cannot perform the calculation required to use Eq. (3.12). Except for very small systems, such as the 4×4

Using the model we considered in Section 3.3.4, there are simply too many states for an explicit categorization to be performed. Thus, we are forced to rely on thermodynamic integration to determine the entropy, calculating it using the heat capacity:

$$S(T) = S(T=0) + \int_0^T dT' \frac{C(T')}{T'}. \quad (4.5)$$

It is in fact (4.5) and not (3.12) that is used to derive the plots shown in Figure 4-4. As we explained in the previous section, the WAMC produces lower estimates for fluctuation properties than what is actually observed. Consequently, relying on results obtained at intermediate temperatures will throw off the answers predicted by thermodynamic integration, although other properties calculated by thermodynamic integration which do not require fluctuation properties in their definition, such as the free energy, do not suffer as much from this problem.

We could in principle also start at infinity and use the formula

$$S(T) = S(T=\infty) - \int_T^\infty dT' \frac{C(T')}{T'}; \quad (4.6)$$

however, this would create problems as $T \rightarrow 0$, since the coarse-grained results would not agree with the theoretical predictions of $S/Nk_B = \ln 2$. Alternatively, we could use both the high- and low-temperature limits, but this would mean that our predictions at intermediate temperatures would be invalid, especially at the jump discontinuity at the temperature where we would attempt to bridge the results produced by (4.5) and (4.6).

4.4.5 Scaling results

One application of the wavelet-accelerated MC method is to provide an upper bound for locating phase transitions. Running multiple simulations, at different levels of resolution, one can determine for each level of resolution the approximate phase transition temperature $T_p(N_s)$, where N_s is the number of degrees of freedom (here, block spins) in the given system. From these data, one can extrapolate a scaling relationship of the form

$$(T_p - T_p^*) \sim N_s^{-\gamma}, \quad (4.7)$$

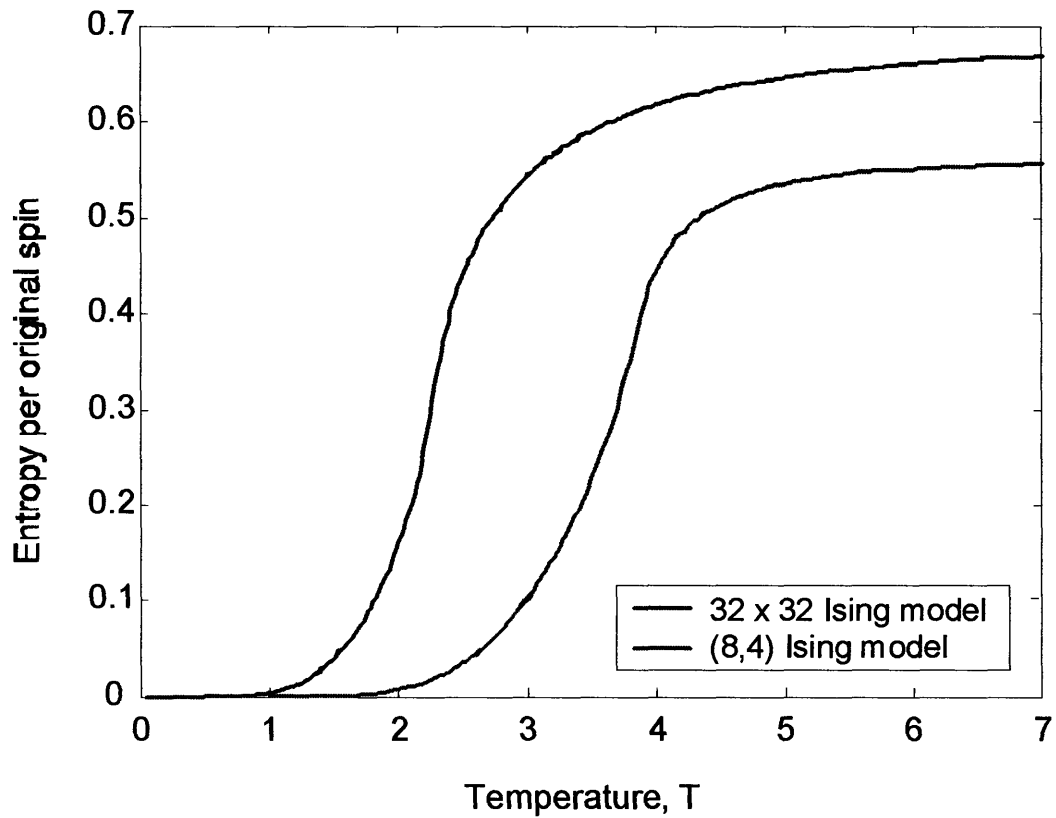


Figure 4-4: Entropy of the 32×32 Ising model as a function of temperature, as computed by direct simulation (top) and using an (8,4)-WAMC simulation (bottom).

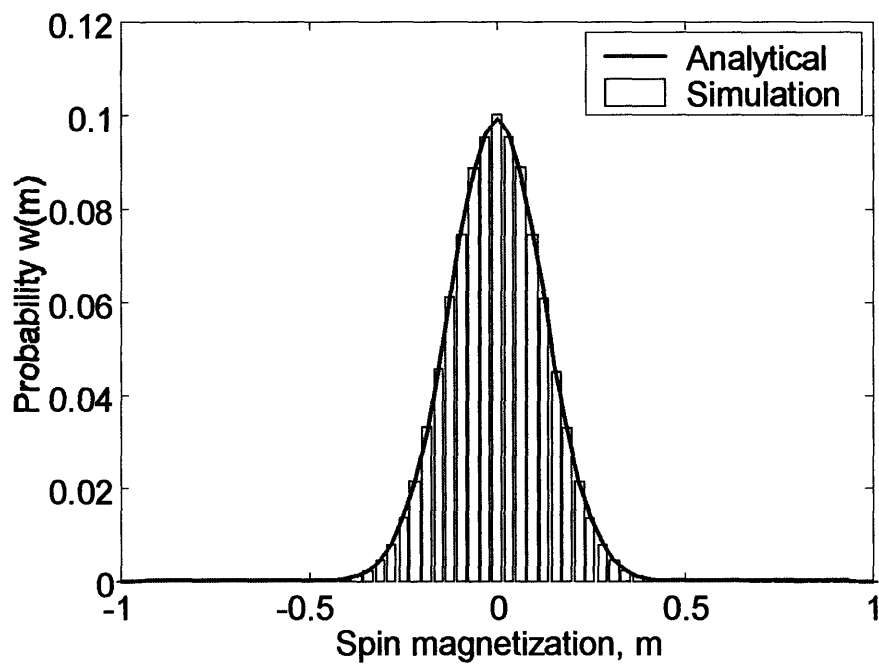


Figure 4-5: Distribution of the magnetization of states in an 8×8 subblock in the limit $T \rightarrow \infty$, as predicted by analytical arguments (line) and by direct simulation (histogram).

where γ is the corresponding “scaling” exponent, and T_p^* is the phase transition temperature for the untransformed model. The estimate obtained for the scaling exponent γ depends upon the technique used to calculate the transition temperature for a given system—for example, estimating divergence of the heat capacity versus the onset of spontaneous magnetization. Our simulations suggest that γ is typically between 0.20 and 0.25, with lower values obtained from divergence of heat capacity than from the onset of spontaneous magnetization.

As explained in I, using a relationship like (4.7) to estimate the phase transition temperature will usually lead to an overestimate of the phase transition temperature. This is a consequence of the underestimation of entropy that occurs through the reduction of the size of configuration space as a result of the wavelet transform. Estimates of $k_B T_p/J$, as determined by (4.7) for the two-dimensional Ising model considered here typically varied between about 2.7 and 2.9, which is an error of approximately 25 percent from the theoretical value of 2.27 provided by the Onsager solution, but only about 20 percent from the results determined by the traditional MC simulations, which gave $k_B T_p/J \approx 2.35$. Although these errors are somewhat sizable, it is useful to note that the total computation time required to obtain the estimate using scaling laws is at least an order of magnitude smaller than the computation time required to perform a direct simulation on the original system. Thus, if computational time is at a premium, an effective approach may be to use the wavelet transform method to provide an upper bound for T_p , and then perform a direct simulation for the parameter space with temperatures below T_p .

4.4.6 Decorrelation time

Another important measure to study is the time required for decorrelated samples. It is well known that in the vicinity of the critical point, traditional Monte Carlo algorithms experience so-called “critical slowing-down” [19]. The Monte Carlo aspect of the WAMC algorithm does not vary from traditional Metropolis Monte Carlo, so we expect that the performance of the two algorithms should be similar, when measured near their respective critical temperatures.

To compare the two methods, we generated $2^{24} = 16\,777\,216$ new configurations for the 32×32 Ising model at the critical point using traditional Metropolis Monte Carlo, as well as for the (4, 8)-, (8, 4)-, and (16, 2)-WAMC models. To determine the correlation time, we used the blocking technique of Flyvbjerg and Petersen [60]. The results are shown as Figures 4-6 and

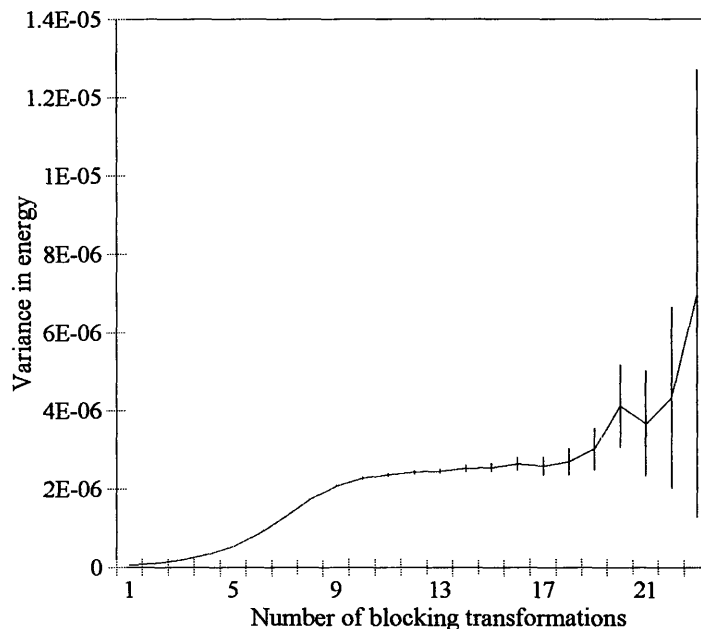


Figure 4-6: Graph showing variance in the estimate of energy as calculated using the method of Flyvbjerg and Petersen [60] for a 32×32 Ising model measured at its critical temperature.

4-7 for the MMC and (16,2)-WAMC models, respectively. The salient feature in the graph is the onset of a plateau in the value of the variance of the energy σ_E^2 ; according to the method of Flyvbjerg and Petersen, this indicates that configurations separated by a distance of 2^x steps are statistically independent, where x is the number of blocking transformations that have been performed. For the MMC model, we find that $x \approx 13$ or $x \approx 14$ provides a decent estimate; for the (16,2)-WAMC model, $x = 16$ is a good estimate for the index. [For the (4,8)- and (8,4)-models (not shown), $x = 15$ is a reliable estimate.] In each case, this indicates that between $2^{14} = 16384$ and $2^{16} = 65536$ steps are required between independent configurations. Thus, we conclude that there is no degradation of performance near the critical point of a WAMC simulation, relative to traditional MMC simulations.

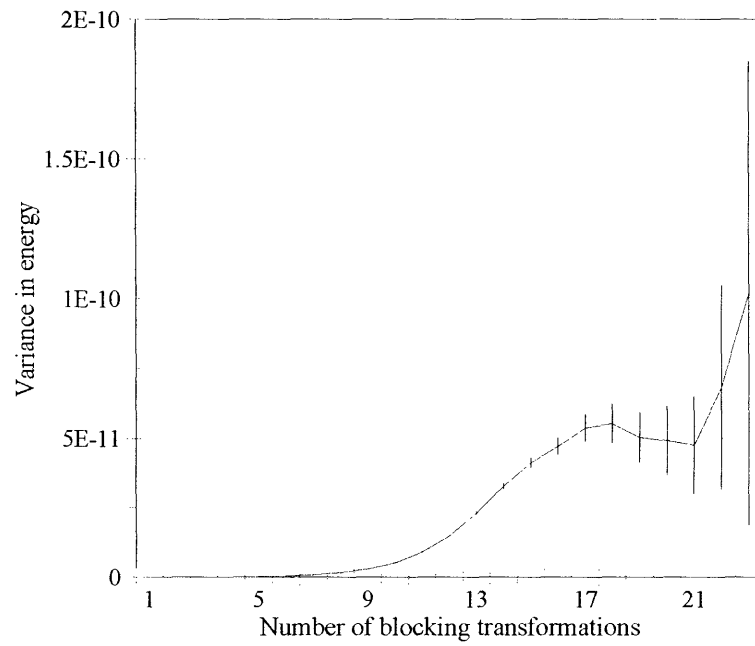


Figure 4-7: Graph showing variance in the estimate of energy as calculated using the method of Flyvbjerg and Petersen [60] for a 32×32 Ising model in a $(16, 2)$ -WAMC measured at the critical temperature determined from the simulation.

Table 4.1: Performance comparison for MMC versus WAMC

Simulation size	Time for 5×10^5 passes (s)
32×32 MMC	1824.22
(4, 8)-WAMC	64.5781
(8, 4)-WAMC	55.1094
(16, 2)-WAMC	144.516

4.5 Analysis

4.5.1 Measured performance comparison

In comparing the performance of the standard Monte Carlo algorithm to the wavelet-accelerated Monte Carlo algorithm, we performed 5×10^5 lattice passes on a 32×32 -lattice on a 733 MHz Pentium II: for the standard Monte Carlo algorithm, this meant that, on average, 5×10^5 attempts were made to flip each spin. For the WAMC algorithm, 5×10^5 attempts were made to flip a spin on a given level. The results are summarized in Table 4.1. We see that the (4, 8)- and (8, 4)-simulations, which have the smallest total number of lattice sites (80), perform the fastest; however, even the (16, 2)-simulation, which has one-fourth as many variables (260) as the 32×32 standard Monte Carlo simulation (1024), finishes in less than 8 per cent of the time required for the latter simulation. As the system size increases, the computational efficiency achieved by breaking down the system into multiple stages, all of relatively equal size, becomes even greater: for a 128×128 -lattice, the performance gain increases from a factor of approximately 25 to a factor of approximately 50, when we compare the (32, 4)- and (16, 8)-WAMC simulations to the standard MC model. However, for the (8, 16)-WAMC model, the complexity of assigning one of 65 possible values to each of 256 variables according to the correct probability distribution becomes comparable to that of the original problem, so that in fact standard MC runs in roughly a factor of 3 faster than the (8, 16)-model.

Similar results are also observed for calculating the phase diagram of a 64×64 -Ising lattice, which is shown in Figure 4-8 as a plot of average magnetization as a function of temperature and external field strength using an (8, 8)-model, for temperatures between $T = 0.5$ and $T = 5.0$, and for field strengths between $h = -1$ and $h = 1$. The phase diagram reproduces the essential features of the original two-dimensional ferromagnetic Ising lattice, such as the phase separation at $h = 0$, although the exact shape differs from the results obtained via a standard

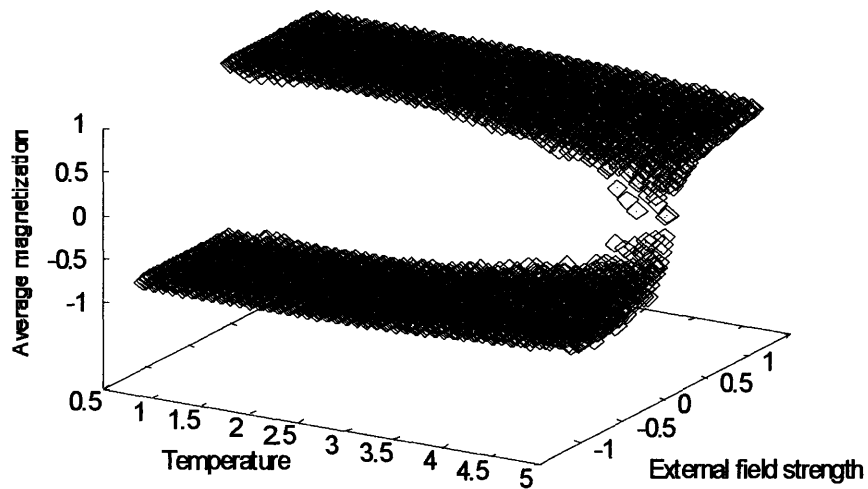


Figure 4-8: Phase diagram plotting average magnetization versus temperature and external field strength for a 64×64 ferromagnetic Ising lattice, computed using an $(8, 8)$ -model via WAMC. The general features correspond to those that would be produced with standard MMC, but require less than 3 per cent of the computational time.

Metropolis Monte Carlo simulations. However, the plot based on WAMC calculations is created approximately 40 times faster than would a comparable plot using standard MMC.

4.5.2 Comparison with renormalization group methods

Our observations also indicate that the accuracy of the wavelet-accelerated Monte Carlo simulations depends on the relative proximity to an “attractive fixed point” of the physical model in parameter space. Borrowed from renormalization group theory, these attractive fixed points represent the limiting behavior of the system under various conditions, such as the zero- and infinite-temperature limits and the limits of zero and infinite external field. As we approach these limiting cases, the approximations made in obtaining our wavelet-transformed Hamiltonian become increasingly less significant.

Combining these observations allows us to design an on-line fine-tuning algorithm for the coarse graining of our system: the further away from the critical point of the parameter space, the smaller the number of degrees of freedom $N^{(K)}$ necessary to simulate the system must be. Thus, if we keep track of changes in fluctuation properties such as the heat capacity or the magnetic susceptibility as we change the system parameters (h, T) , we can get an estimate of

our relative distance to the critical point. If we are sufficiently far away from the critical point, we can choose either to increase the number of stages K that we simulate, or we can look at more degrees of freedom at lower stages by increasing $N^{(1)}, N^{(2)}, \dots, N^{(K-1)}$. As we approach the critical point, we can either decrease the number of stages K or include more degrees of freedom at higher stages by reducing $N^{(1)}, \dots, N^{(K-1)}$.

4.5.3 Sources of error

In general, the source of our errors can be traced to the assumption that local fluctuation terms could be reasonably ignored in our coarse-grained Hamiltonian (4.4). This naive but otherwise useful assumption yields correct thermodynamic behavior when the overall physics of the system is particularly simple: in the low- and high-temperature regimes, for instance, when the number of observed microstates is small or when the differences between observed microstates is inconsequential. For more complicated behaviors, as found at intermediate temperatures and above all in the vicinity of a critical point, the use of this assumption has a drastic effect on both the phase space of the system, which in turn affects all thermodynamic properties of the system, including the internal energy and entropy of the system, as well as fluctuation properties of the system.

More complicated methods for dealing with fluctuation terms have significant drawbacks associated with them. Treating the fluctuation terms just like the block averages maintained in $\bar{\mathbf{u}}^{(K)}$ means that the wavelet-transformed Hamiltonian is no simpler than the original Hamiltonian, which affords few advantages in computational time. Likewise, other approaches, such as parametrizing the probability distribution for the elements of $\bar{\mathbf{u}}^{(K)}$ using a property like the energy E , introduce new functional dependencies which cannot be taken into account using the wavelet transform. Thus, we sacrifice one of the major advantages of the method—moving from one level to another is achieved exclusively through use of the wavelet transform. Thus, the most promising avenue for dealing with fluctuation terms is to develop a probability distribution for the fluctuation terms via the same approach used to determine probability distributions for the local averages. Then, using the probability distribution for the fluctuation terms, we can treat the discarded terms of the Hamiltonian as a noise term which can be used to restore some of the entropy that was lost as a result of the coarse-graining [see Chapter 3 for more details].

However, we have presented our results here to show what can be achieved under “worst-case” conditions, without the use of inverse coarse-graining methods.

Many coarse-graining techniques control errors by fitting the parameters of a new Hamiltonian to ensure agreement with some known structural information about the system, such as the radial distribution function [3, 10]. For lattice systems, this iterative approach is reflected in renormalization group theory, and notably the Wilson recursion method [160], which finds the fixed points of the system. As formulated, WAMC creates a coarse-grained Hamiltonian by truncating the Hamiltonian obtained after application of the wavelet transform. As an improvement to this, it should be possible to use the wavelet transform to determine which terms will appear in the Hamiltonian, and then determine the appropriate parameters to ensure the best fit for some desired property of the system using an iterative approach [102, 103, 134].

4.5.4 Constructing an adaptive algorithm for MC using the wavelet transform

As pointed out above, the wavelet transform method tends to produce overestimates for the critical point of the system; therefore, if we start with the high-temperature limit of our algorithm and slowly reduce the temperature in our simulation, we can observe the movement toward the critical point by watching various fluctuation parameters, such as the heat capacity $C_H = (\langle E^2 \rangle - \langle E \rangle^2) / k_B T^2$. Near the critical point, we expect to see a rapid increase in the value of C_H , consistent with the expected logarithmic divergence observed in the limit of finite-size systems [65, 124]. If we use the onset of this logarithmic divergence as an indicator, we can then “step down” and use a finer lattice including more degrees of freedom. This system will naturally better reflect the physics of our system, particularly in the vicinity of the critical point. We expect that very near the critical point, we will have to simulate the system at the original scale, since this will be effectively the only level which accurately represents the underlying behavior of the system. However, the region of parameter space where this is necessary is relatively small compared to the complete parameter space. This is especially true when we consider that as we proceed below the critical temperature of the system, the logarithmic divergence of C_H will also vanish. As a result, as we move increasingly far away from the critical point, we begin to approach the other fixed-point behaviors associated with

the low-temperature limits of the system. Since these are reasonably well-preserved using the wavelet transformation, we can safely return to increasingly coarse-grained descriptions of our system as the simulation proceeds past the critical point.

As an example, we compute the spontaneous magnetization curve for a 64×64 Ising lattice in the temperature range $0.5 \leq T \leq 10.0$, with $\Delta T = -0.05$, and choosing as our refinement criterion $\Delta C_H/\Delta T \leq -0.5$, until we reach the finest scale, corresponding to the original problem. We begin by coarse-graining the system to an $(8, 8)$ -model, where we find that the criterion is triggered only at $T = 5.1$; we then continue with a $(4, 16)$ -model, down to $T = 4.0$, at which point the refinement criterion is exceeded. Refining once more, we proceed with a $(2, 32)$ -model until $T = 3.4$, at which point the threshold is again crossed. Since the next refinement is the original problem, we proceed at this level of resolution until we have passed the critical point, so that $\Delta C_H/\Delta T$ is positive. As a coarsening criterion, we select for simplicity the opposite of the refinement criterion, $\Delta C_H/\Delta T \leq 0.5$. Using this criterion, we find that we coarsen the model to the $(2, 32)$ -, $(4, 16)$ - and $(8, 8)$ -models at temperatures of $T = 1.75$, $T = 1.65$, and $T = 1.55$, respectively. The rapid coarsening of the model results from the higher estimates of the critical point in the coarsened models. Since we are well past the critical point, we expect changes in the heat capacity as a function of temperature to be relatively small, and thus it is possible to obtain accurate results from a relatively coarse model. Computationally, the time required to create this diagram was only 28 per cent that required to perform a standard Metropolis Monte Carlo simulation with the same number of steps. Moreover, in the regions that were not simulated using MMC, the computation time required was just 8 per cent of the time required for MMC. The resulting plot of magnetization versus temperature, shown as Figure 4-9, compares favorably to the analytical solution of Onsager, which is also shown [122].

4.6 Conclusions

The WAMC algorithm can dramatically reduce the time required to calculate the thermodynamic behavior of a lattice system; the trade-off for these savings in time is in the accuracy of the results obtained, a general feature of coarse-graining techniques. The error that results is a function of position in parameter space: the results obtained are generally accurate in the

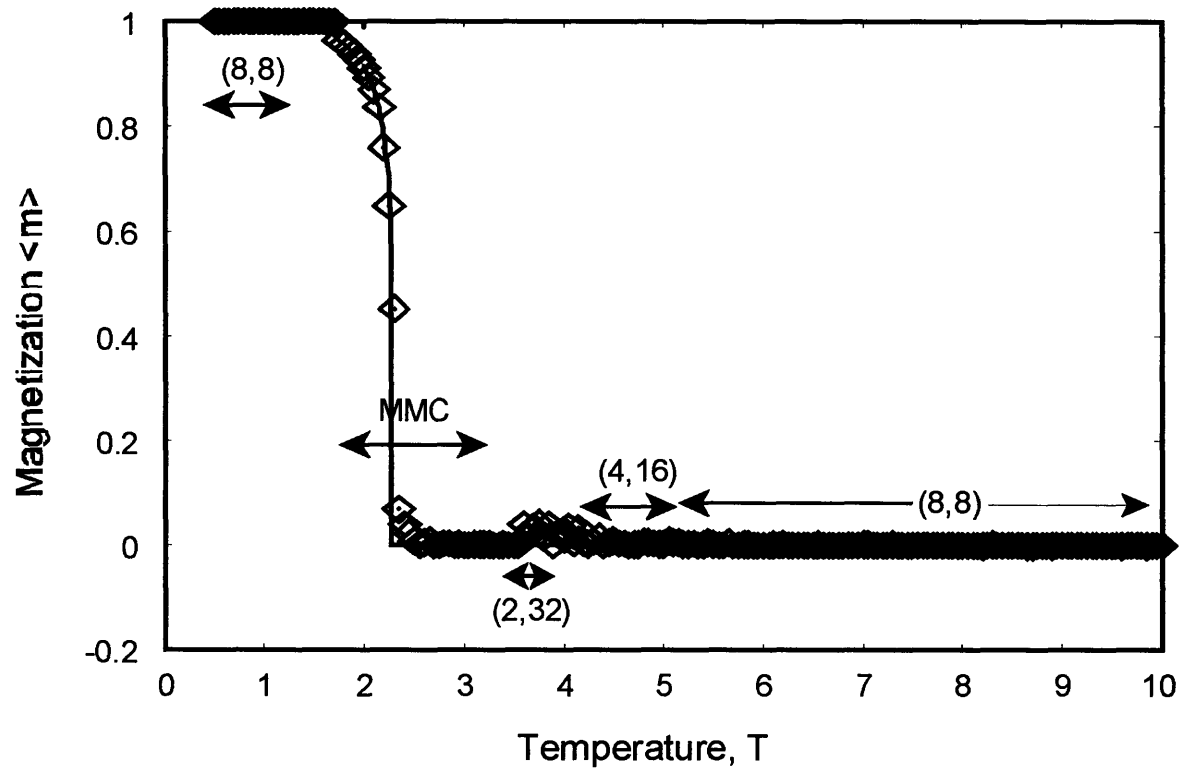


Figure 4-9: Phase diagram plotting average magnetization versus temperature for a 64×64 ferromagnetic Ising lattice, created using an adaptive WAMC algorithm, with refinement and coarsening criterion established using the change in heat capacity with respect to temperature $\Delta C_H / \Delta T$. The squares represent the simulation results, while the line reproduces Onsager's analytical result for the two-dimensional Ising model with zero external field. The simulation used in a given temperature region is shown on the plot.

vicinity of fixed attractors of the system, and decrease as one approaches critical points of the parameter space. Near critical points, deviations from results performed on the original lattice system are the result of coarse-graining the Hamiltonian by eliminating local fluctuation terms. Consequently, this suggests that a hierarchical simulation which uses fluctuation properties such as the heat capacity C_H to gauge proximity to critical points in “real time” would yield dramatic savings in the computation time of the behavior of a lattice system over a wide region of phase space, as regions of space close to fixed attractors would be simulated at a very coarse scale, with full-detailed simulations reserved only for regions of parameter space close to critical points.

Chapter 5

Topological coarse-graining of polymer chains. I. Freely-jointed chains

One method for simulating large polymer chains is to use a coarse-graining scheme which systematically reduces the number of degrees of freedom to improve the computational efficiency of a simulation. A number of coarse-graining schemes have been discussed in recent review articles by Baschnagel *et al.* [10] and by Müller-Plathe [118]. Several other theoretical methods based on renormalization group theory have also been developed to analyze polymer chains [38, 64]. More recently, polymer chains have also been treated as melts of interacting ellipsoids with soft potentials [75, 119] or as soft colloids [23, 97, 99]. To date, these coarse-graining approaches have always connected two specified length scales. In many cases, though, it is not clear in advance what level of coarse-graining is either appropriate or acceptable: different molecules may permit different amounts of coarse-graining. A scheme which can be adjusted according to the needs of a particular molecule during the course of a simulation would therefore be of enormous benefit, as it would have greater generality and flexibility than current coarse-graining approaches.

We employ the hierarchical methodology we developed in Chapters 3 and 4 for lattice spin systems, taking advantage of the self-similar structure of the polymer: at each stage of the

simulation, we sample some portion of the chain we want to study, and use data derived from a previous stage of the simulation to describe the behavior of the new, coarse-grained particles which represent an aggregation of particles in the finer-scale simulation. Unlike the lattice spin model we considered previously, our coarse-graining technique will be topological, as opposed to spatial: the “blocking” method now proceeds along the backbone of a polymer chain, rather than in a fixed region of space.

This chapter outlines the application of the wavelet transform to simulation of linear polymer chains, and presents results applicable to freely-jointed chains. The following chapter addresses the problem of excluded-volume interactions and presents results for self-avoiding chains.

5.1 The freely-jointed chain model

The freely-jointed chain model consists of a set of N beads joined together by a series of rigid bonds of length l whose orientation is drawn uniformly from the unit sphere:

$$p(r) = \delta(r - l), r \geq 0 \quad (5.1)$$

$$p(\phi) = \frac{1}{2} \sin \phi, 0 \leq \phi \leq \pi \quad (5.2)$$

$$p(\theta) = \frac{1}{2\pi}, 0 \leq \theta \leq 2\pi. \quad (5.3)$$

There are no interactions present between beads, so the excluded volume of the chain is zero. In the mathematical literature, the freely-jointed chain is referred to as a Gaussian random walk; we do not use that nomenclature here, to avoid confusion with the Gaussian polymer model, in which the rigid bonds are replaced with harmonic springs.

Simulating a freely-jointed chain is relatively simple, because there is no need to determine interaction energies for a particular configuration. Freely-jointed chain configurations can be generated via explicit construction, or using methods like the translate-jiggle algorithm [41] or the pivot algorithm [90] which have been developed for self-avoiding chains. Our atomistic simulations are performed via explicit construction; as discussed below, our coarse-grained simulation is based on a variant of the translate-jiggle algorithm.

5.2 Coarse-graining of polymer chains

5.2.1 Using wavelets to construct a coarse-grained model

A freely-jointed or self-avoiding polymer chain is an inherently topological structure: the order of beads $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ in the chain determines not only the position of the individual beads, but also their connectivity: bead 1 connects to bead 2, and so on. As a result, we can exploit this structure when constructing our coarse-grained representation. The resulting approach is closely related in spirit to renormalization group procedures, and in particular to the blocking technique of Kadanoff [71, 85].

We take as our input data the set of positions $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ of the beads in the polymer chain, and select the *unnormalized* Haar wavelet pair,

$$\mathbf{r}_n^{(k)} = \frac{1}{2} \left[\mathbf{r}_{2n-1}^{(k-1)} + \mathbf{r}_{2n}^{(k-1)} \right], \quad (5.4)$$

$$\mathbf{w}_n^{(k)} = \frac{1}{2} \left[\mathbf{r}_{2n-1}^{(k-1)} - \mathbf{r}_{2n}^{(k-1)} \right], \quad (5.5)$$

as our basis set. We have chosen to use the unnormalized form of the Haar wavelet to ensure that the physical dimensions of the system are not altered through the coarse-graining procedure.

Applying Eqs. (5.4) and (5.5) to $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ yields a set of averages $\left\{ \mathbf{r}_i^{(1)} \right\}_{i=1}^{N/2}$ and a set of differences $\left\{ \mathbf{w}_i^{(1)} \right\}_{i=1}^{N/2}$, where (1) denotes that Eqs. (5.4) and (5.5) have been applied once to our original data set. The effect of the averaging operator in (5.4) is to create a new coarse-grained bead $\mathbf{r}_n^{(1)}$ at the center of mass of the beads at \mathbf{r}_{2n-1} and \mathbf{r}_{2n} ; the differencing operator (5.5) returns the distance between the original particles' positions and their center of mass. We can then take the set of averages $\left\{ \mathbf{r}_i^{(1)} \right\}_{i=1}^{N/2}$ and iterate the wavelet transform once more, producing a set of averages $\left\{ \mathbf{r}_i^{(2)} \right\}_{i=1}^{N/4}$ and a set of differences $\left\{ \mathbf{w}_i^{(2)} \right\}_{i=1}^{N/4}$ which represent the behavior of the chain over twice the length scale as did the previous iteration. This process can be repeated as many times as desired, until we have a set of averages $\left\{ \mathbf{r}_i^{(K)} \right\}_{i=1}^{N/2^K}$ and K sets of differences, $\left\{ \left\{ \mathbf{w}_i^{(k)} \right\}_{i=1}^{N/2^k} \right\}_{k=1}^K$.

Applying the wavelet transform to our original data set re-expresses it in terms of a different set of variables, but does not reduce the total number of variables produced. In almost all cases, however, probability distributions for sums of random variables are more complicated than

the original variables. Therefore, to simplify our work further, our coarse-graining procedure eliminates the differencing variables $\{\mathbf{w}_i^{(K)}\}$ after each iteration of the wavelet transform. Elimination of the differencing variables is the fundamental step in which information about the underlying configuration is lost after coarse-graining.

5.2.2 “Equivalence” of on-lattice and off-lattice coarse-grained computations

There are some unusual features of our coarse-graining procedure, particularly for simulations performed on a lattice. Applying the wavelet transform to a topological structure increases the discretization of the lattice; this result has not previously been commented upon in the coarse-graining literature, including our work on lattice spin systems in Chapters 3 and 4.

To see how this increased discretization arises, define a to be the lattice spacing. In general, a is chosen to be equal to at least twice the radius of the bead to prevent beads sitting on neighboring lattice sites from overlapping. Our simulation starts on the lattice $\mathcal{L}^{(0)}$ of integer points, so that $a = 1$. The first set of averages created using (5.4) is found on the lattice $\mathcal{L}^{(1)}$ defined by the set of points $\mathbf{r} \in \mathcal{L}^{(0)}$, plus the set of midpoints of all pairs of points which can legally be bonded to one another. For the first set of averages, this corresponds to all neighboring pairs of points on $\mathcal{L}^{(0)}$. Consequently, the “effective” lattice spacing on $\mathcal{L}^{(1)}$ is $1/2$.

This process can be repeated, so that the lattice spacing for the lattice $\mathcal{L}^{(k)}$ obtained by k applications of (5.4) is 2^{-k} . As k increases, the lattice moves towards a continuum, as the lattice spacing decreases to zero. An example of this behavior in two dimensions is illustrated in Figure 5-1. A walk of sixteen segments is converted into a walk of eight segments, with the endpoint of each bead resting on a half-integer grid point. The next iteration would include quarter-integer grid points, and so on. For a given chain, this process can be repeated until a chain has been converted into a single point on a grid with spacing given by 2^{-K} , where K is the total number of iterations. At this point, no further coarse-graining of the walk is possible.

In addition to the lattice discretization increasing, the effect of the wavelet transform is to produce descriptions of the chain which are more “thread-like.” An example of this behavior is shown in Figure 5-2, which illustrates the action of the wavelet transform on a self-avoiding walk on the cubic lattice. The first iteration of the wavelet transform, illustrated in Figure 5-2 (b), essentially reproduces the atomistic structure of the walk in Figure 5-2 (a). The second and

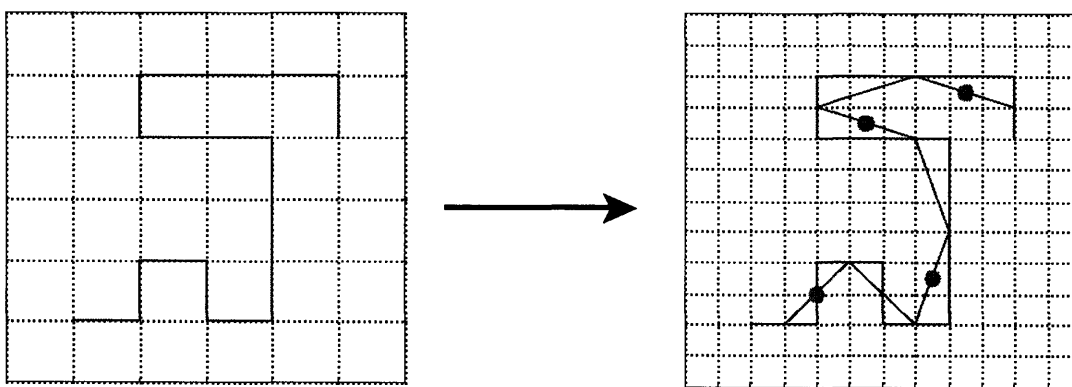


Figure 5-1: Coarse-graining of a 16-site random walk in two dimensions using “center-of-mass aggregation” of adjacent points along the chain, as suggested by the wavelet transform method. The four sites which would be created after another iteration of this method are indicated on the graph; note that each is at a quarter-integer lattice point.

third iterations largely preserve the major features of the original chain topology. However, the finer-scale details are lost as the amount of coarse-graining continues, so that after a number of coarse-graining iterations, the model will more closely resemble an off-lattice model, such as a Gaussian polymer model, than the original lattice walk.

5.3 Wavelet-accelerated Monte Carlo (WAMC) simulation of polymer chains

5.3.1 Hierarchical simulation approach

The basic principle of the WAMC algorithm is to “divide and conquer” a simulation. Instead of performing a single, fully atomistic simulation, one performs a series of simulations on chains whose degree of coarse-graining becomes progressively larger, as the probability distributions and other results obtained from one stage is passed on to the next, coarser stage of simulation.

Let us suppose that the goal is to model a chain of length N . A fully atomistic method would simulate all N beads in a single step. The WAMC method begins by performing a fully atomistic simulation on a chain containing $N_{b,1} \ll N$ beads. At the atomistic scale, no coarse-graining has yet been performed, so each bead has an “effective size” $N_{e,1} = 1$. [The

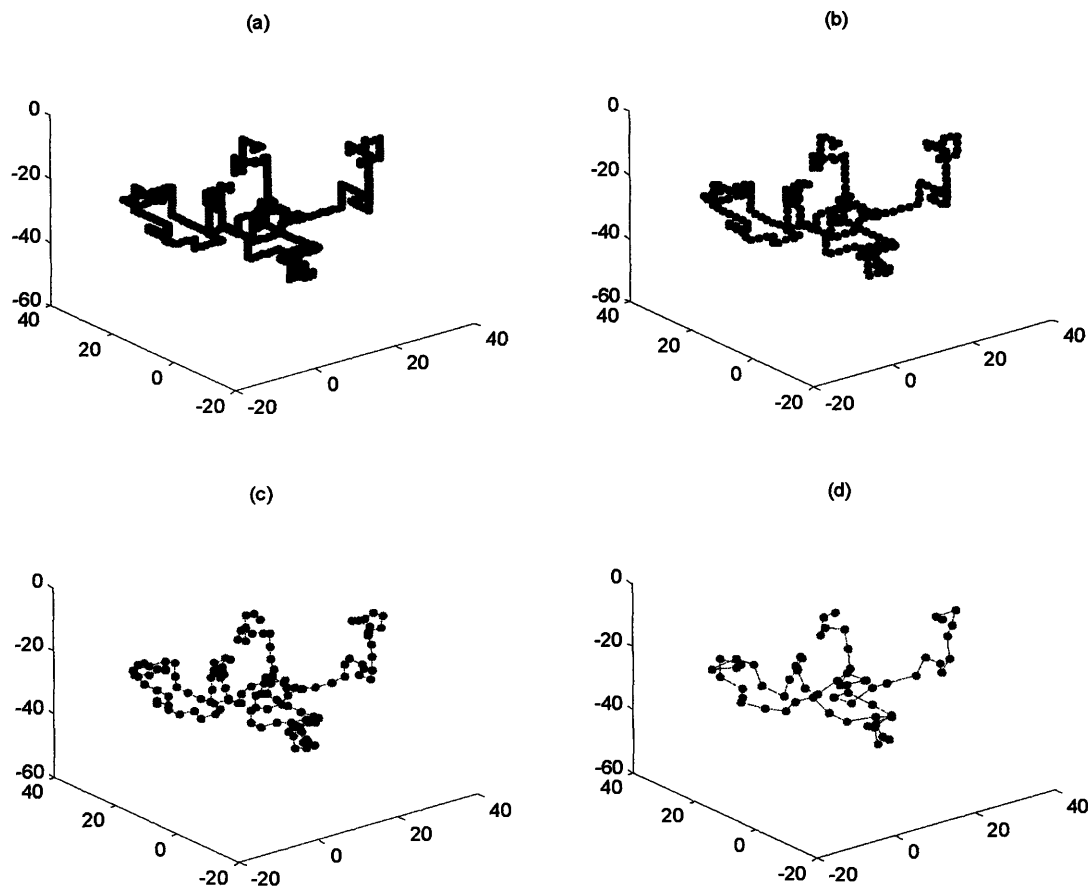


Figure 5-2: Coarse-graining of a three dimensional walk on a cubic lattice using the wavelet transform. (a) A 512-step self-avoiding random walk. (b)-(d) The same walk, after one, two, and three iterations of the Haar transform defined by (7.6) and (7.7).

numbers in the subscripts indicate the simulation stage.] For the simulations performed here, we used the pivot algorithm at the atomistic scale. After every $O(N_{b,1})$ steps, we apply the wavelet transform (5.4) K_1 times to determine the positions $\mathbf{r}^{(K_1)}$ of each new coarse-grained bead. These coarse-grained beads now have an effective size of $N_{e,2} = 2^{K_1}$ atomistic beads each. Using these results, we compute the distributions for the coarse-grained internal coordinates, as well as a coarse-grained inter-bead potential, if necessary. [See Chapter 7 for a further discussion of coarse-grained potentials.] An upper bound on the value of $N_{e,2}$ is placed by requiring at least four coarse-grained beads to determine a torsion angle, so that we must have $N_{e,2} \leq N_{b,1}/4$.

The second stage of the simulation consists of a chain of length $N_{b,2}$ beads. Since each coarse-grained bead has effective size $N_{e,2}$, the total effective chain length is therefore $N_{b,2} \times N_{e,2}$. If the total effective chain length is equal to the actual length N of the chain to be simulated, the simulation is terminated. Otherwise, one proceeds to a third stage of simulation by applying the wavelet transform (5.4) K_2 more times, to produce coarse-grained beads of effective size $N_{e,3} = 2^{K_2} N_{e,2}$. We then compute the new potential and distributions for the coarse-grained internal coordinates, and then simulate a chain of total size $N_{b,3} \times N_{e,3}$, and so on until the target chain size N is reached

5.3.2 Coarse-grained simulation algorithm

In applying the WAMC algorithm to a polymer chain, it is useful if every coarse-grained simulation stage uses the same basic method for sampling configuration space. Thus, we would like to develop a method capable of handling distributions in bond length, bond angle, and torsion angle, plus non-bonded potentials. We would also like our method to operate independently of the source of the distributions, so that its input data may be provided by calculation from experimental data, from an atomistic simulation, or from a previous iteration of the coarse-grained algorithm.

The coarse-grained Monte Carlo sampling algorithm employed here generates trial configurations using a variant of the “translate-jiggle” algorithm, developed by Dickman and Hall [41]. The “translate-jiggle” algorithm selects a bead along the chain, and perturbs it by a random displacement. The resulting bond, between beads k and $k + 1$, is then either stretched or com-

pressed to ensure that the previous bond length is conserved. The new displacement vector corresponding to the displacement is computed; then, to ensure connectivity constraints are satisfied, the remaining beads $k + 2, k + 3, \dots$, in the chain are then translated by the new displacement vector. The resulting chain is tested for self-avoidance and energy differences to determine if the proposed configuration should be accepted.

Our modifications are principally a relaxation of some of the constraints. Since we would like the bond lengths to be variable as well as the bond and torsion angles, bond lengths are not adjusted after selection. Instead, a bond length is selected at random from the provided distribution, and a new set of bond orientation information (ϕ_i, θ_i) is selected from the internal-coordinate distributions discussed above.

If the simulation is performed on-lattice, the bond is adjusted only so that both of its endpoints are mapped to the nearest lattice points consistent with the wavelet transform. Effectively, this means that for a coarse-grained simulation in which a bead represents 2^m beads on the atomistic lattice, the lattice spacing constant becomes 2^{-m} . Although similar to the approach used in the bond fluctuation method, [26] the number of allowed bonds is much greater, which makes the list-based approach of the bond-fluctuation approach impractical.

Since the allowed points on $\mathcal{L}^{(k)}$ consist of the midpoints of allowed bonds on $\mathcal{L}^{(k-1)}$, technically not all points will be valid lattice points, unless we take the limit $k \rightarrow \infty$. However, even if a discrete lattice has been used, the distribution of distances between nearest-neighbor bond segments rapidly approaches a continuous (Gaussian-like) distribution. Therefore, we relax this criterion in our simulations, and let the lattice $\mathcal{L}^{(k)}$ be the set of all lattice points with spacing 2^{-k} . This simplifies the problem of locating the nearest lattice point, and drastically increases the performance of the algorithm, with negligible effects on the accuracy of the simulation.

5.4 Probability distributions for coarse-grained internal coordinates

Our coarse-graining method replaces the description of an atomistic freely-jointed chain with a smaller number of coarse-grained beads whose positions are determined from center-of-mass averaging. In order for the results obtained from the coarse-grained model to be equivalent to the

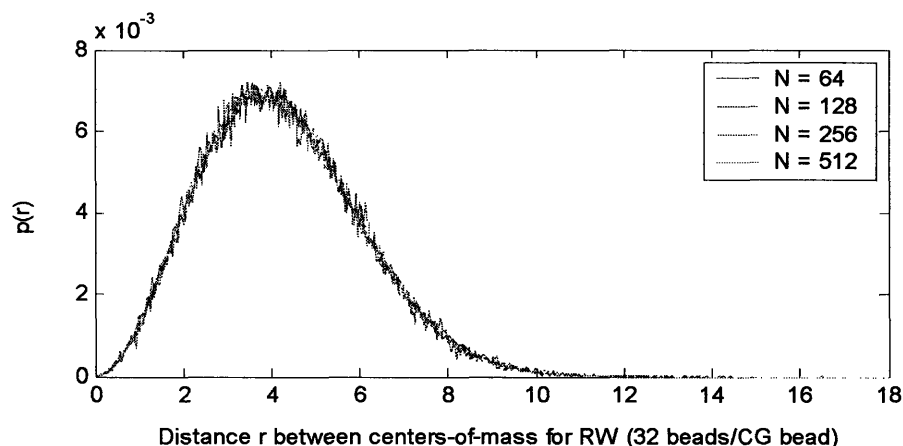


Figure 5-3: Comparison of the probability distributions for the coarse-grained bond length (distance between adjacent center-of-mass beads) for freely-jointed chains for different chain lengths, with 32 atomistic beads per coarse-grained bead.

original, the behavior of the coarse-grained degrees of freedom are generally more complicated than the uniform distribution given by Eqs. (5.1) through (5.3). The distributions of the coarse-grained internal coordinates necessary to achieve adequate agreement between the theoretical results and the coarse-grained model are discussed below.

5.4.1 Bond-length distribution

The coarse-grained bond length is given by the distance between two successive coarse-grained beads. On the atomistic scale, the coarse-grained bond length is the distance between the centers-of-mass of two consecutive segments of the chain. For a freely-jointed chain, the bond-length distribution at the coarse-grained scale should resemble the standard end-to-end distribution for a freely-jointed chain, although the scale of the bond-length distribution will be smaller than the end-to-end distribution. In Figure 5-3, we show the bond-length distributions for freely-jointed chains whose coarse-grained beads each represent 32 atomistic beads, simulated via the pivot algorithm for chain lengths ranging from $N = 64$ to $N = 512$. As shown in Table 5.1, the mean value of the bond-length distribution is essentially constant for freely-jointed chains: the distance between two segments should not depend on the number of beads in the chain, since there are no interactions between beads.

Table 5.1: Mean coarse-grained bond-lengths between beads representing 32-mers

N	$\langle r \rangle_{GRW}$
64	4.2534
128	4.2477
256	4.2452
512	4.2502

5.4.2 Bond-angle distribution

The coarse-grained bond angle is the angle formed between two adjacent coarse-grained bonds. From Eq. (5.2), we know that in the atomistic freely-jointed chain, the cumulative distribution function (CDF) of the bond angle distribution is given by

$$F(\phi) = \frac{1}{2} - \frac{\cos \phi}{2}. \quad (5.6)$$

The distribution for the atomistic freely-jointed chain is symmetric. While a single step has a symmetric bond-angle distribution, results for the coarse-grained chain exhibit a pronounced bias towards larger bond angles, as observed by Laso *et al.* [92]. We plot the coarse-grained CDF against the ideal CDF (5.6) in Figure 5-4. The distribution was determined using the pivot algorithm on a chain of length cN_e beads, combining N_e atomistic beads into a single coarse-grained bead, and then computing the $c - 2$ bond angles formed.

Even the distribution shown in Figure 5-4 is insufficient to yield adequate agreement between our coarse-grained simulations and the expected results for the Gaussian random walk. As suggested by Laso *et al.*, we parametrized the coarse-grained bond angle distribution by the lengths of the bonds forming the angle. Rather than use the product of the bond lengths as a parameter, we found sufficient agreement was obtained by using the longer of the two bond lengths r_1 as a parameter. Binning the results into three separate regimes, representing $r_1 < N_e^{1/2}$, $N_e^{1/2} < r_1 < \frac{3}{2}N_e^{1/2}$, and $r_1 > \frac{3}{2}N_e^{1/2}$, respectively, appears to be sufficient to model freely-jointed chains. The resulting CDF's are plotted in Figure 6-5, which clearly reveal the bias towards larger bond angles, particularly as the value of r_1 increases. In fact, the average bond angle for the latter case is almost $2\pi/3$, or roughly one-third more than the ideal result of $\pi/2$.

To simplify the handling of the bond-angle distribution, we noted that the CDF's shown in

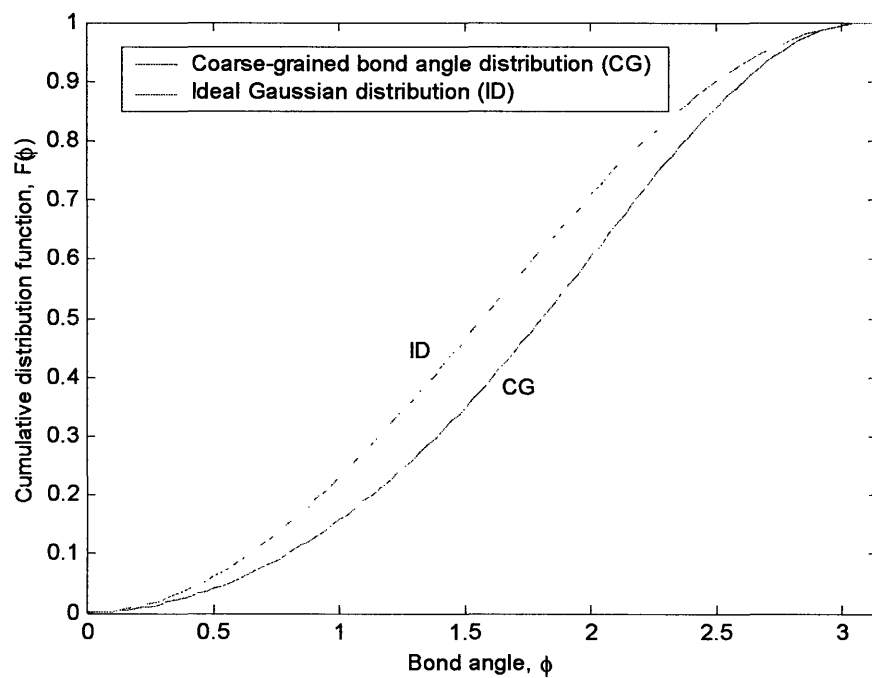


Figure 5-4: Ideal and observed coarse-grained cumulative distribution functions for a coarse-grained Gaussian random walk divided into segments of 32 beads each.

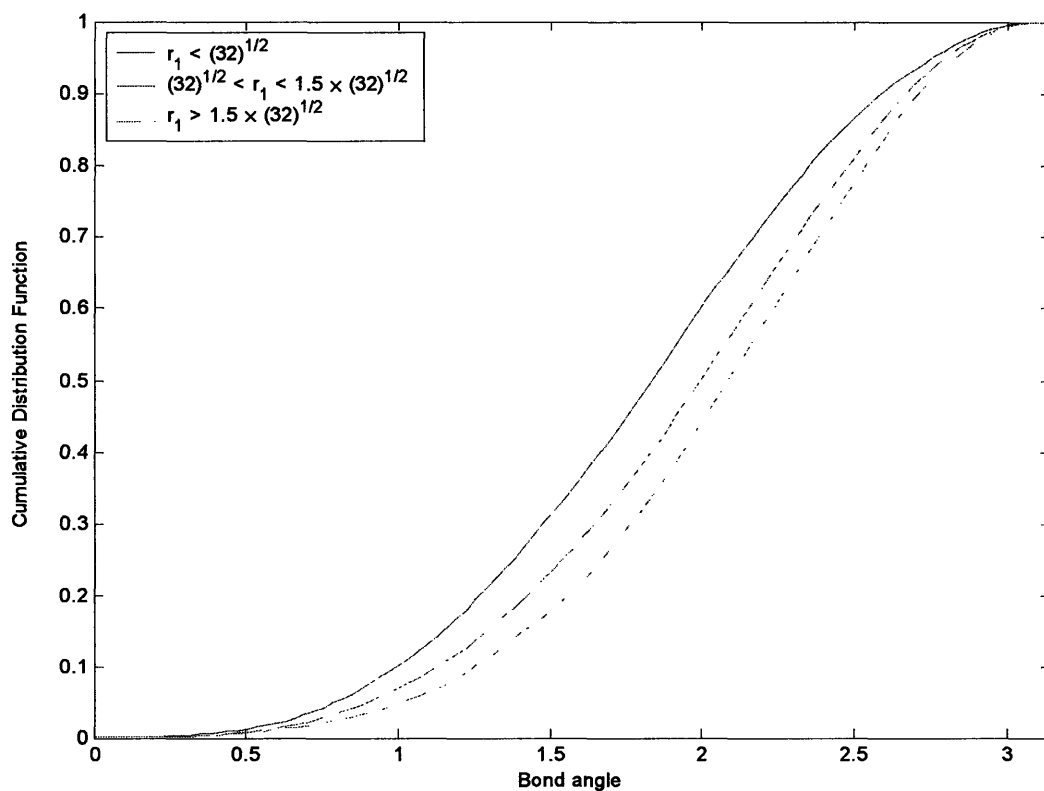


Figure 5-5: Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a freely-jointed chain with $N = 128$ and $N_e = 32$, as a function of r_1 , the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal Gaussian value of $r_1 = \sqrt{N_e} = \sqrt{32}$; the middle, dashed curve shows $\sqrt{N_e} < r_1 < \frac{3}{2}\sqrt{N_e}$; and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}$.

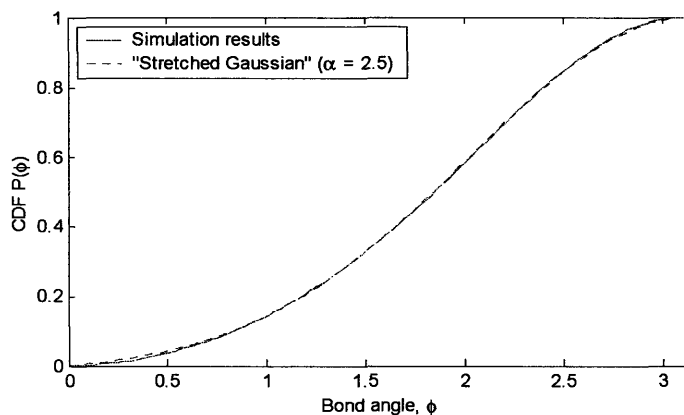


Figure 5-6: CDF's for the coarse-grained bond angle computed by simulation (solid) and using a “stretched Gaussian” approximation (dotted). The two curves are largely indistinguishable.

Figure 5-5 resemble “stretched” Gaussian exponential functions, so we fitted $p(\phi)$ to the form

$$p(\phi, \alpha) = \frac{\exp\left(-(\phi - \pi)^2 / \alpha\right) - \exp(-\pi^2 / \alpha)}{1 - \exp(-\pi^2 / \alpha)}, \quad (5.7)$$

where α is a fitted parameter. For the freely-jointed chain, excellent agreement between the unparametrized simulation results and the curve fit of Eq. (5.7) is obtained using $\alpha = 2.5$, as shown in Figure 5-6. The parametrized distributions shown in Figure 5-5, divided into regions $r_1 < \sqrt{32}$, $r_1 < \frac{3}{2}\sqrt{32}$, and $r_1 \geq \frac{3}{2}\sqrt{32}$, are fitted to the values $\alpha = 3.3$, 2.4, and 1.5, respectively.

5.4.3 Torsion-angle distribution

Finally, we considered the torsion-angle distribution, which was not reported by Laso *et al.* [92]. The coarse-grained torsion angle is defined as the analogue of the torsion angle in an atomistic chain. We determined the distribution for the coarse-grained torsion angle by simulating chains of length $4N_e$ beads, transformed N_e beads to a single coarse-grained bead, and determined the torsion angle of the resulting four-bead chain.

Like the bond-angle distribution, the torsion-angle distribution is asymmetric: *trans*-like configurations are favored over *cis*-like configurations. A sample probability distribution function for the torsion angle, plotted in terms of the cosine of the torsion angle, is shown in Figure

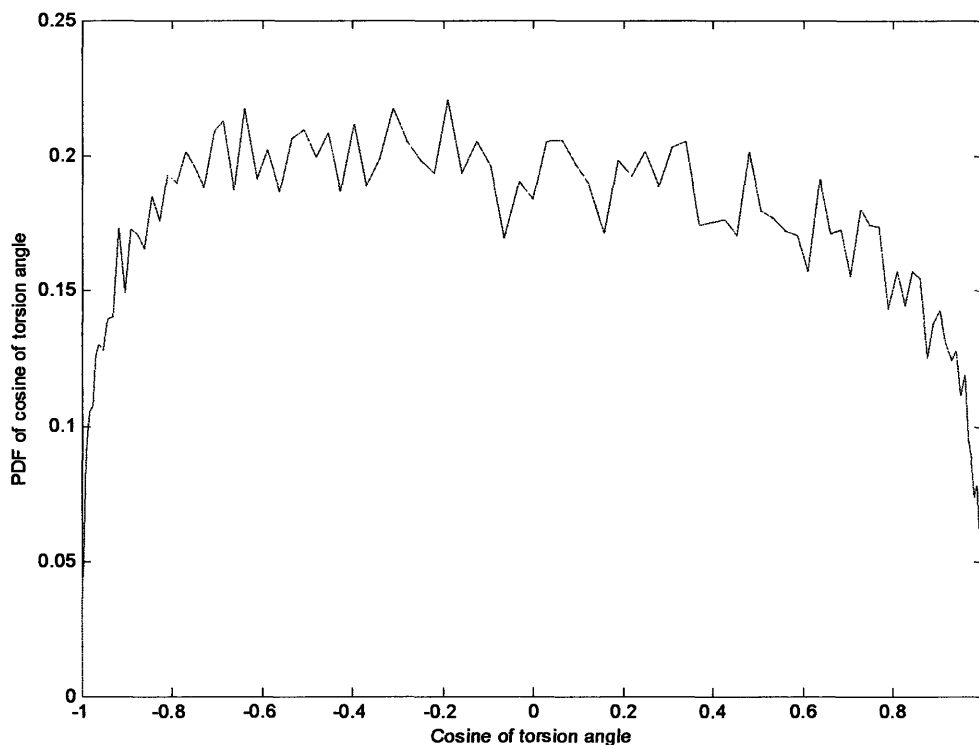


Figure 5-7: Probability distribution function for the torsion angle of a freely-jointed chain of length $N = 128$ considered as four beads of effective size $N_e = 32$.

5-7. Computing the mean value of the distributions reveal that the bias is actually quite small:

$$\langle \theta \rangle = 1.6177, \quad (5.8)$$

$$\langle \cos \theta \rangle = -0.0398. \quad (5.9)$$

However, as we shall see below, combining the biased torsion angle distribution specified by Figure 5-7, along with the coarse-grained bond length and bond-angle distributions outlined above, achieves excellent agreement with theoretical predictions for the freely-jointed chain.

5.5 Results

The well-known scaling regime for the radius of gyration for freely jointed chains, $\langle R_g^2 \rangle \sim N$, can be reproduced with excellent agreement via numerical simulation even for relatively small

Table 5.2: Mean square radius of gyration using bond angle and torsion angle distributions

N	$\langle R_g^2 \rangle$ (atomistic)	$\langle R_g^2 \rangle$ (WAMC)
$512 = (16 \times 32)$	84.75	82.74
$1024 = (32 \times 32)$	171.60	169.80
$2048 = (64 \times 32)$	342.70	342.10
$4096 = (128 \times 32)$	679.95	685.69
$8192 = (256 \times 32)$	1360.22	1375.09

values of N . For simulations performed with N between 16 and 32768, plotting $\langle R_g^2 \rangle$ as a function of N , and fitting to a power law $\langle R_g^2 \rangle \sim N^{2\nu}$ yields $\langle R_g^2 \rangle \sim N^{1.000}$, from which we obtain the Flory exponent

$$\nu = 0.500 \pm 0.002, \quad (5.10)$$

which agrees with the expected scaling for a freely-jointed chain.

We can repeat this calculation for our coarse-grained model using wavelet-accelerated Monte Carlo. Mapping 32 atomistic beads to each coarse-grained bead, and using the internal coordinate probability distributions given in Figures 5-3, 5-5, and 5-7, we simulated chains up to lengths equivalent to 32,768 atomistic beads and computed the mean square radius of gyration $\langle R_g^2 \rangle$. Results for short chains are given in Table 5.2, while additional results at larger values of N are shown in Figure 5-8. The scaling-law fit of the data,

$$\langle R_g^2 \rangle \sim N^{1.008 \pm 0.003}$$

is also shown in Figure 5-8; it is essentially indistinguishable from the atomistic scaling result $\langle R_g^2 \rangle \sim N$, except for very large values of N .

5.6 Conclusions

We have demonstrated that the WAMC algorithm can be readily adapted to study molecular systems in which the coarse-graining should be topological. The method recovers the expected scaling law $\langle R^2 \rangle \sim N^{2\nu}$ with correct estimates for ν for freely-jointed chains for a variety of chain lengths. Significantly, we find that intramolecular distributions of internal coordinates for polymer chains generally become both asymmetric and interdependent, in accord with

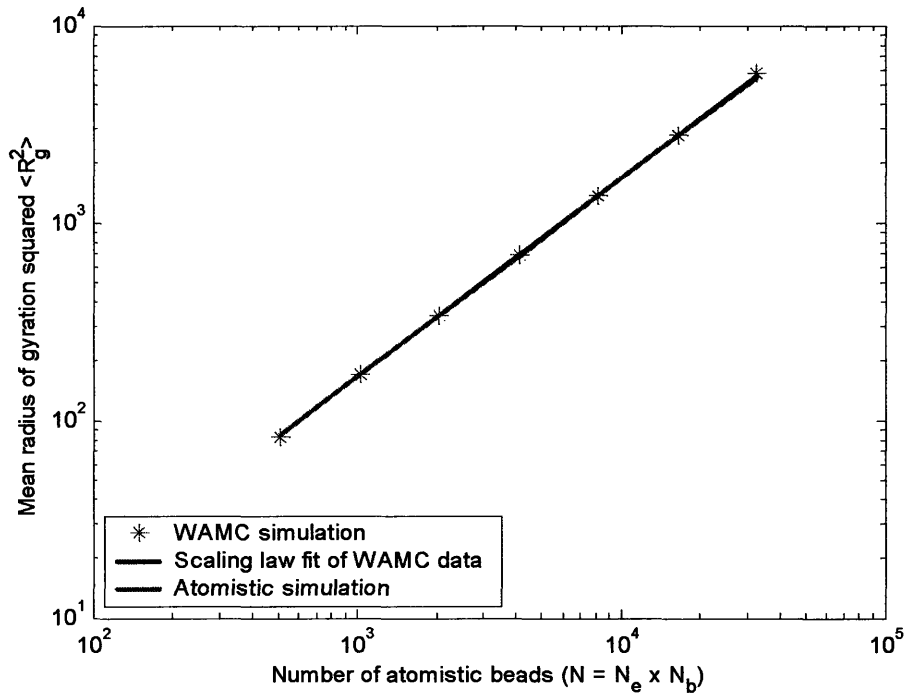


Figure 5-8: The mean square radius of gyration $\langle R_g^2 \rangle$ as a function of bead size N for both atomistic and WAMC representations of the freely-jointed chain.

the earlier results of Laso *et al.* [92] However, the real purpose of the WAMC method is to study chains with potential interactions. To that end, we discuss self-avoiding chains in the following chapter, and show how the presence of non-bonded interactions—namely, the hard-sphere potential—is handled by the WAMC algorithm, and how a phenomenological picture of interactions between particles can be used to develop scaling laws for coarse-grained potentials to reduce the computational cost.

Chapter 6

Topological coarse-graining of polymer chains. II. Self-avoiding chains

In the preceding chapter, we discussed wavelet-accelerated Monte Carlo (WAMC), a new method for multiscale simulations, and showed how to apply it as a topological tool to polymer simulations, the simplest case being the freely-jointed chain. The real power of the algorithm, however, lies in its ability to handle non-bonded interactions between particles, as is found in self-avoiding chains. Thus, in this chapter, we extend our development of the algorithm. We show that applying the wavelet transform to the hard-sphere potential justifies the use of soft potentials, and that, using the freely-jointed chain model, we can develop phenomenological scaling laws which can be used to estimate new coarse-grained interatomic potentials using previously computed data. We also demonstrate that our method reproduces expected scaling results for self-avoiding chains.

6.1 Self-avoiding chains

6.1.1 Definition

The self-avoiding chain is equivalent to a freely-jointed chain except for the inclusion of a “hard-sphere” potential that prohibits beads from overlapping:

$$u_2(\mathbf{r}_i, \mathbf{r}_j) = u_{HS}(r_{ij}) = \begin{cases} \infty, & r_{ij} < r_{HS} \\ 0, & r_{HS} \leq r_{ij} \end{cases}, \quad (6.1)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and r_{HS} is the cut-off for the hard-sphere potential. For chains defined on a lattice, r_{HS} is generally chosen to be equal to the lattice spacing a , thereby defining the maximum size of a particle on the chain and preventing multiple beads from occupying the same lattice site. We can incorporate other pairwise potentials, such as a Lennard-Jones potential, without difficulty.

Although the form of Eq. (6.1) is extremely simple, the consequences of introducing this potential are drastic. A freely-jointed chain can be analyzed analytically [80, 161]; including the hard-sphere potential Eq. (6.1) makes most of these analytical approaches intractable. Moreover, the difficulties carry over to numerical simulations. Generating a freely-jointed chain is trivial; generating a self-avoiding random walk is not, because one must ensure that $u_{HS}(r_{ij}) = 0$ for all pairs of beads i and j . Attempting to generate a large number of self-avoiding walks by explicit construction is impractical since the computational time required to verify that a walk of length N is self-avoiding can be as large as $O(N^2)$, and because we may need to generate many walks in order to find one that is self-avoiding.

6.1.2 Simulating self-avoiding chains

A well-known algorithm that circumvents the obstacles mentioned above is the *pivot algorithm* proposed by Lal [90]. The pivot algorithm generates a trial configuration from an existing configuration by selecting a pivot consisting of a rotation matrix $\Gamma_{\alpha\beta}(\theta)$ specifying a rotation of angle θ in the $\alpha\beta$ -plane and a random bead n in a chain of length N . The pivot $\Gamma_{\alpha\beta}(\theta)$ is then applied to all points either to the left or to the right of bead n , so that the new position

\mathbf{r}'_i of bead i is

$$\mathbf{r}'_i = \mathbf{r}_n + \Gamma_{\alpha\beta}(\theta)(\mathbf{r}_i - \mathbf{r}_n), \quad (6.2)$$

for all $1 \leq i \leq n$ or $n \geq i \geq N$. All possible rotations in each plane must be allowed although rotations can be constructed as the result of several other rotations, as is permitted in other Monte Carlo techniques. This allows the use a smaller set of rotation matrices which, when combined in series, yield all possible rotations.

For square and cubic lattices, all possible moves can be achieved by reflecting the chain across various planes. As a result, the complete set of symmetry operations is extremely compact: we need to specify just three rotation matrices in two dimensions and five rotation matrices in three dimensions. Consequently, all of the rotation matrices can easily be stored in memory, significantly speeding up the computations. In three dimensions, the five required matrices correspond to rotations of 90, 180, and 270 degrees in the xy -plane, and rotations of 90 and 270 degrees in the xz -plane:

$$\begin{aligned} \Gamma_{xy}\left(\frac{\pi}{2}\right) &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \Gamma_{xy}(\pi) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \Gamma_{xy}\left(\frac{3\pi}{2}\right) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \Gamma_{xz}\left(\frac{\pi}{2}\right) &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \text{ and } \Gamma_{xz}\left(\frac{3\pi}{2}\right) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (6.3)$$

For off-lattice simulations, we allow rotations in any of the xy -, xz -, or yz -planes, but restrict θ to small values; the cutoff is typically chosen to be ± 15 degrees.

While the pivot algorithm is straightforward and easy to implement, especially for lattice calculations, it languished in relative obscurity until a seminal study by Madras and Sokal in 1988 [105]. The principal criticism of the algorithm was its low rate of acceptance—typically orders of magnitude below the 50 per cent acceptance rate usually expected in stochastic molecular simulations [65,91]. Madras and Sokal have shown that while the acceptance rate is small, it decays only very slowly, approximately as $O(N^{-1/5})$ for large N . What had been neglected in previous analyses was that the acceptance of a pivot move generally results in a radical

change in the overall configuration, since on average one-fourth of the atoms in the chain will be rotated. Therefore, accepting even a handful of proposed moves can result in an independent configuration. Madras and Sokal were able to show that the pivot algorithm requires only $O(1)$ pivots—and $O(N)$ *attempts*—to achieve an independent configuration, as opposed to the reptation algorithm [38], which can require as many as $O(N^3)$ attempted moves to create an independent configuration. Moreover, given a self-avoiding walk as a starting configuration, the pivot algorithm is guaranteed to return a self-avoiding configuration, and in principle cannot be “trapped” in a particular configuration, a significant obstacle in the reptation algorithm [105].

The other obstacle for the pivot algorithm—its expected $O(N^2)$ running time per attempt—has also been addressed, at least for lattice systems. The $O(N^2)$ running time is obtained if we exhaustively test that $\mathbf{x}_i \neq \mathbf{x}_j$ for all pairs of i and j between 1 and N . While this test is trivial to implement, we can devise a much faster test for self-avoidance by observing that, if the distance on the lattice which must be traversed to get from \mathbf{x}_i to \mathbf{x}_j is equal to ra , then for a lattice spacing a , \mathbf{x}_i cannot overlap any \mathbf{x}_k for $j - r + 1 \leq k \leq j + r - 1$. Consequently, we can skip testing these points for self-avoidance for a given value of i . Looping this procedure over all values of i , from 1 to N , leads to a much more efficient algorithm, with a running time on the order of $O(N^{6/5})$. [In fact, Kennedy has shown that in limited cases, the algorithm can have a sub-linear running time [88].

6.1.3 Demonstration of $O(N)$ decorrelation time between independent configurations

Unlike explicit generation of freely-jointed random walks, where we create a new configuration on each iteration, the pivot algorithm does not produce independent configurations on successive moves, so we cannot include the results from every step in determining averages such as the root-mean-square end-to-end distance. Instead, we need to wait until essentially uncorrelated samples have been produced. For the pivot algorithm, it can be shown that this happens in $O(N)$ *attempted*—not *accepted*—pivots [105]. As a sample, we show two configurations of 1024-step walks separated by 1024 attempted pivot moves in Figure 6-1. The relative independence of the two configurations is made even clearer when one notes that the two sets of axes shown have different scales. A more rigorous demonstration of the decorrelation of the states is shown

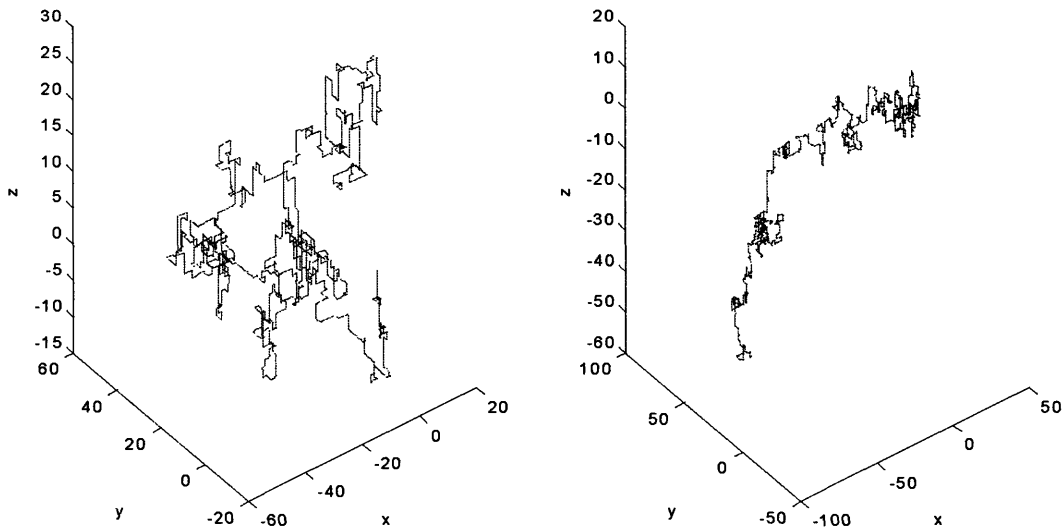


Figure 6-1: Configurations of a 1024-step chain taken 1024 attempted pivots apart, showing the decorrelation between samples after $O(N)$ attempted pivots.

in Figure 6-2, an autocorrelation plot of the root-mean-square end-to-end distance. As shown $\langle R(t) R(t + \tau) \rangle$ decays to zero for $\tau \geq 50$ moves, from which we can reasonably conclude that the decorrelation time is indeed $O(N)$.

6.2 Wavelet-accelerated Monte Carlo (WAMC) simulation of polymer chains

6.2.1 Coarse-grained simulation algorithm

We outlined the algorithm used for WAMC simulations in Section 5.3. Here, we briefly summarize the changes in the algorithm required to accommodate the hard-sphere potential Eq. (6.1). Since conformations can now have non-zero energies, after each trial conformation has been generated, it is necessary to determine whether or not the new configuration should be accepted. Just as in atomistic simulations, we use the Metropolis criterion to determine whether or not to accept the proposed conformation. However, unlike the atomistic self-avoiding walk,

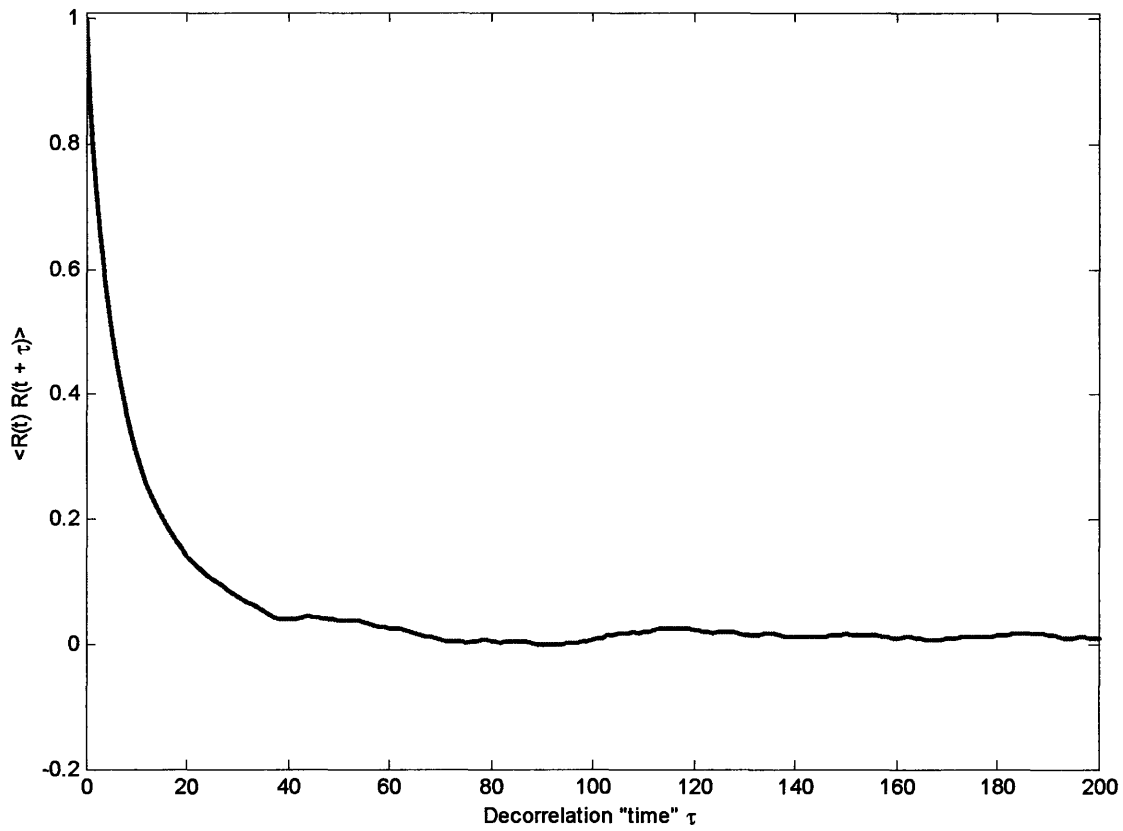


Figure 6-2: Autocorrelation function of the end-to-end distance for a self-avoiding walk of length $N = 1024$.

where the energy of a given configuration is either zero or infinity, the coarse-grained walk has a nontrivial energy function. The form of this potential must be determined before the beginning of each scale of the simulation; we discuss our method for computing coarse-grained potentials in Section 7.5 below.

In computing the non-bonded interactions, we do not recompute $\left| \mathbf{r}_i^{(0)} - \mathbf{r}_j^{(0)} \right|$ for every i and j . In accordance with the distance-checking improvements of Kennedy [88], if $r = \left| \mathbf{r}_i^{(0)} - \mathbf{r}_j^{(0)} \right| > r_{cut}$, where r_{cut} is the cutoff of our coarse-grained potential, we skip over consecutive beads until the total contour length of the skipped portion of the chain exceeds $r - r_{cut}$.

6.2.2 Probability distributions for coarse-grained internal coordinates

Bond-length distribution

As discussed in Section 5.4, instead of the simple distributions which describe bond orientations in the atomistic freely-jointed chain, we now require detailed information regarding the distribution of bond lengths, bond angles, and torsion angles. The methods for deriving these distributions are discussed in Chapter 5; we now present the necessary distributions for self-avoiding chains, and summarize the differences between these distributions and the corresponding distributions for freely-jointed chains.

In Figure 6-3, we show the bond-length distributions for self-avoiding chains whose coarse-grained beads each represent 32 atomistic beads, simulated using the pivot algorithm for chain lengths ranging from $N = 64$ to $N = 512$. As shown in Figure 6-4, the mean value of the bond-length distribution increases as a function of the chain length N , but converges to a limiting value as $N \rightarrow \infty$. Therefore, if we want to map N_e atomistic beads to a single coarse-grained bead, we only need to sample a walk of length $O(N_e)$ in order to obtain an appropriate bond-length distribution.

This behavior can be explained by noting that for a self-avoiding random walk, the presence of more beads in the chain leads to a greater excluded volume for the entire chain. This in turn requires that individual segments should be slightly more separated in a longer chain than in a shorter chain. However, we expect that for very long chains, this local stretching is no longer affected by increasing the size of the chain, since any new segments added will likely be far away from any given segment in the chain.

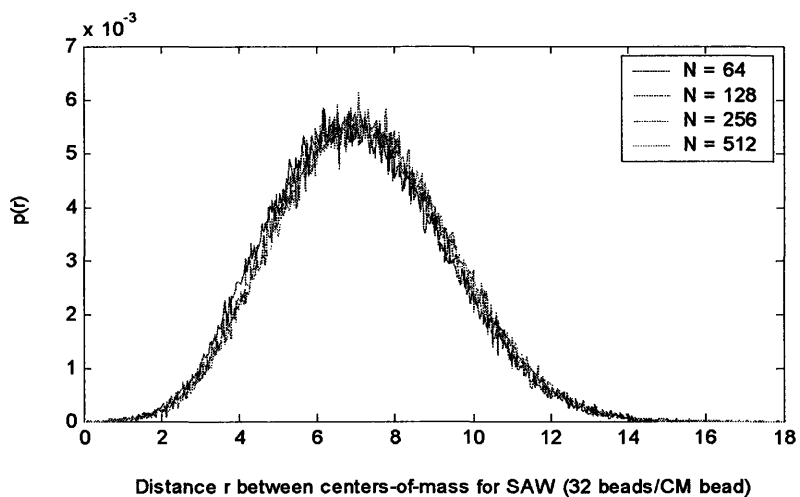


Figure 6-3: Probability distributions for the distance between adjacent center-of-mass beads for self-avoiding chains (bottom), parametrized by total chain length.

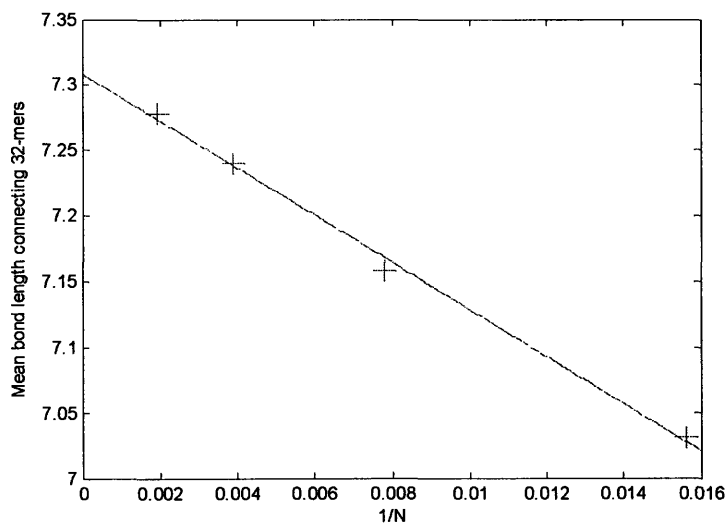


Figure 6-4: Mean bond length connecting 32-mers as a function of the total length N for self-avoiding walks.

Bond-angle and torsion-angle distributions

Our method also requires probability distributions for the coarse-grained bond and torsion angles. The bond-angle distribution is shown in Figure 6-5. As discussed in Section 5.4.2, there is a bias toward larger bond angles, as parametrized by the coarse-grained bond lengths forming the bond angle. The degree of bias observed is in fact greater than that for freely-jointed chains. In addition, there is a slight bias toward *trans*-like conformations over *cis*-like conformations in the torsion angles. Also, we fitted the cumulative distribution functions using stretched exponentials of the form

$$p(\phi, \alpha) = \frac{\exp\left(-(\phi - \pi)^2/\alpha\right) - \exp(-\pi^2/\alpha)}{1 - \exp(-\pi^2/\alpha)}, \quad (6.4)$$

where α is a fitted parameter. The parametrized distributions shown in Figure 6-5, divided into regions $r_1 < \sqrt{32}$, $r_1 < \frac{3}{2}\sqrt{32}$, and $r_1 \geq \frac{3}{2}\sqrt{32}$, are fitted to the values $\alpha = 1.6$, 1.1, and 0.8, respectively. These values are significantly smaller than the corresponding values for the freely-jointed chain, corresponding to the greater bias towards even larger bond angles for self-avoiding chains.

The torsion-angle distribution is shown in Figure 6-6. Again, as in the case of freely-jointed chains, the self-avoiding chain shows bias towards *trans*-like conformations over *cis*-conformations, with $\langle \theta \rangle = 1.6043$, and $\langle \cos \theta \rangle = -0.0335$.

6.3 Results

6.3.1 Atomistic simulations

For pivot algorithm simulations performed with N between 16 and 8192, plotting the resulting data of $\langle R_N^2 \rangle^{1/2}$ versus N yields $\langle R_N^2 \rangle \approx N^{1.194}$, from which we obtain the estimate

$$\nu = 0.597 \pm 0.006. \quad (6.5)$$

Thus, we obtain a reasonable estimate for ν , although the variance is significantly worse than that obtained from simulations of longer chains.

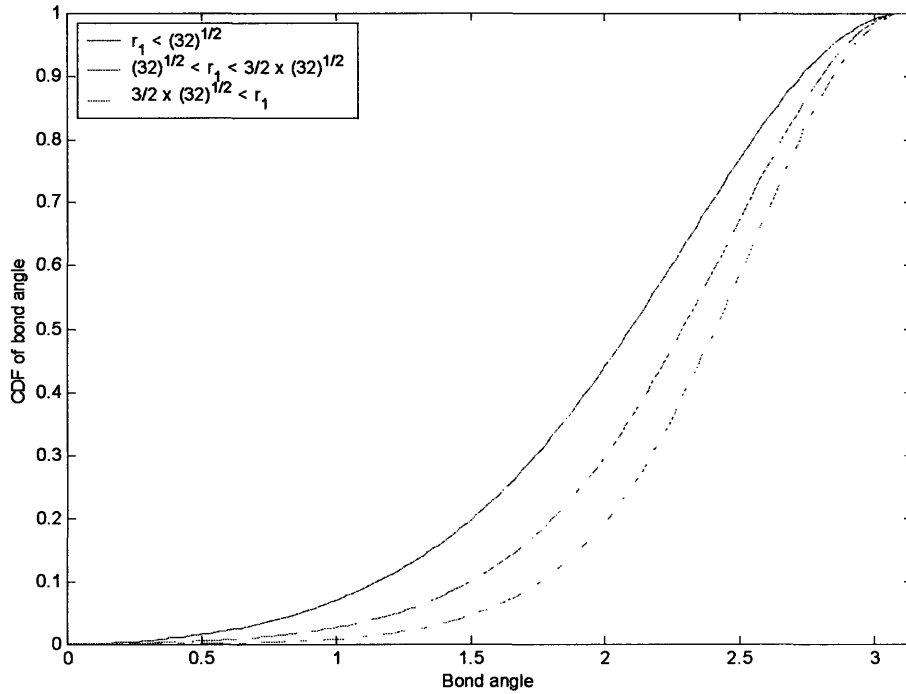


Figure 6-5: Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a self-avoiding chain with $N = 128$ and $N_e = 32$, as a function of r_1 , the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal value of $r_1 = \sqrt{N_e} = \sqrt{32}$ for the freely-jointed chain; the middle, dashed curve shows $\sqrt{N_e} < r_1 < \frac{3}{2}\sqrt{N_e}$; and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}$.

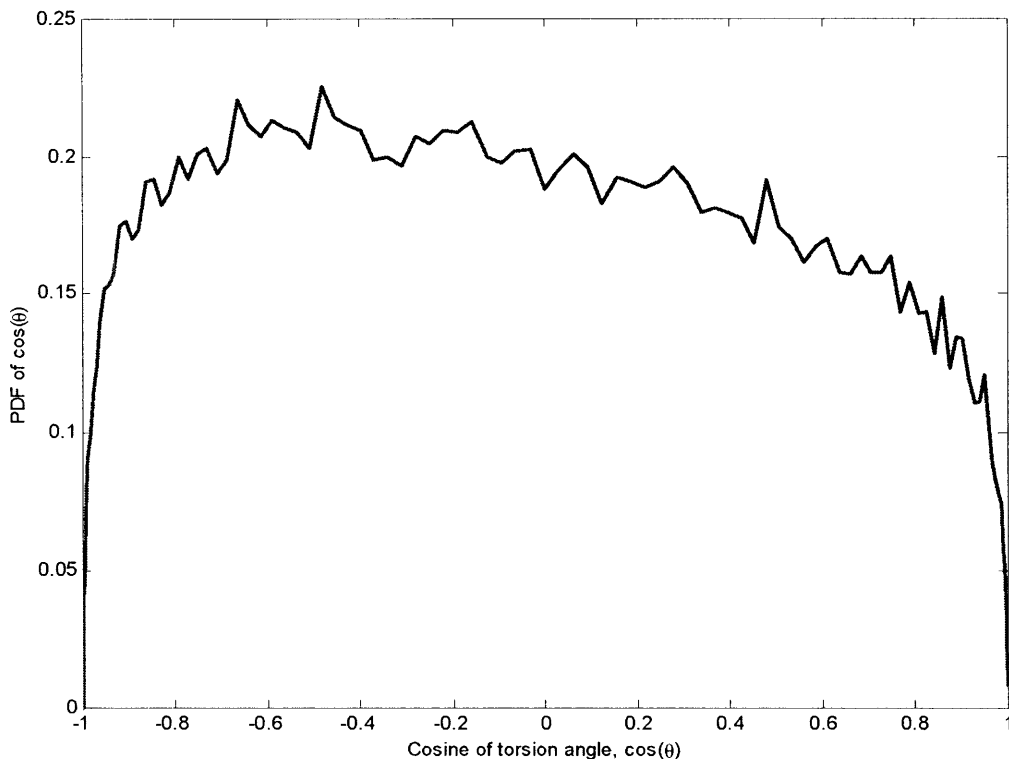


Figure 6-6: Probability distribution function for the torsion angle of a chain of length $N = 128$ considered as four beads of effective size $N_e = 32$.

6.3.2 WAMC simulations

A comparison of the results obtained for a detailed atomistic simulation and for two-scale coarse-grained simulations of self-avoiding random walks up to length $2^{15} = 32768$ beads are shown in Figure 6-8. Two different coarse-grained simulations are plotted, showing $N_{e,2} = 32$ and 64, respectively; the mean end-to-end distances of the three simulations show excellent agreement with one another, indicating that for relatively large values of $N_{e,2}$, there is little effect on the end-to-end distance. Additionally, the Flory exponent for the scaling of the root-mean-square end-to-end distance, $\langle R^2 \rangle^{1/2} \sim N^\nu$, is shown to be $\nu_{cg} \approx 0.578 \pm 0.006$ for $N_{e,2} = 32$ and $\nu = 0.581 \pm 0.006$ for $N_{e,2} = 64$, both of which are within one percent of the best available estimate $\nu \approx 0.577$ [80]; the corresponding data for values of N between 256 and 8192 are shown in Table 6.1.

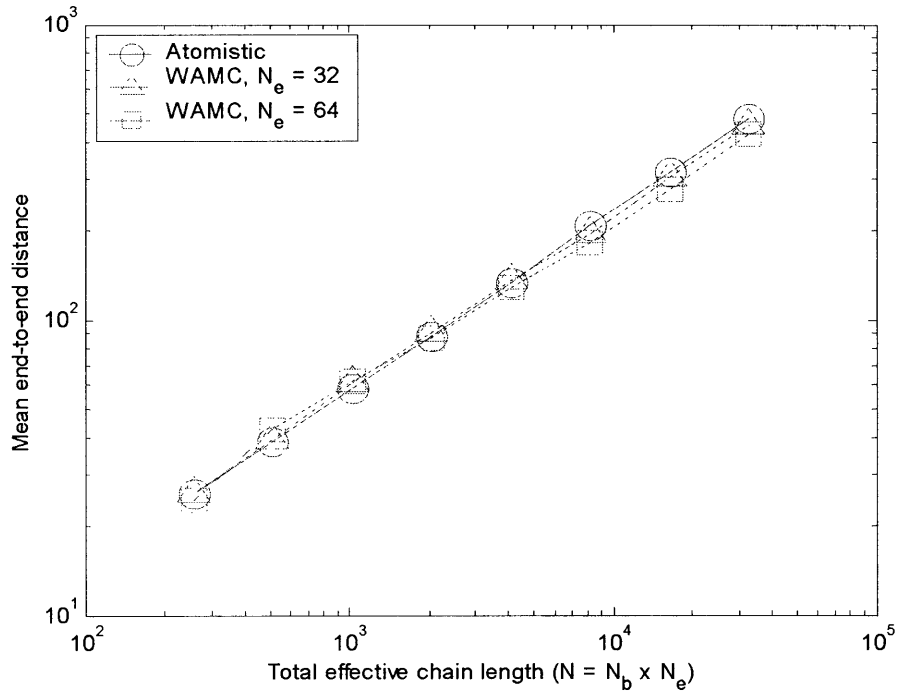


Figure 6-7: Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (circles) and of coarse-grained simulations based on bead sizes of 32 and 64 (triangles and squares, respectively).

Table 6.1: Mean end-to-end distance for atomistic versus WAMC algorithms

N	$\langle R \rangle$ (atomistic)	$\langle R \rangle$ ($N_e = 32$)	$\langle R \rangle$ ($N_e = 64$)
256	25.841	25.885	24.618
512	38.927	39.339	42.775
1024	58.482	61.357	62.207
2048	88.018	90.225	87.268
4096	133.973	135.228	127.651
8192	209.625	193.669	181.878

Results for a four-stage hierarchical simulation

To show the effectiveness of the hierarchical scheme, we performed a series of four-level simulations. The simulation started with $N_{b,1} = 512$ and $N_{e,1} = 1$; the second stage simulations had $N_{e,2} = 128$, but varying values for $N_{b,2}$. For the final two stages, we used $N_{e,3} = 2048$ and $N_{e,4} = 65536$. Results for the mean end-to-end distance for lengths up to $N = 2^{22} = 4194304$ are shown for the four-scale simulations as Figure 6-9. We find that the scaling of the end-to-end distance is $\nu = 0.568$ for the four-scale simulation, which still agrees with current estimates of ν . The numerical discrepancy increases with the number of scales in the simulation, by about a factor of 4.5. The increase in the numerical disagreement was further indication that mass-averaged bond-length distributions by themselves were insufficient to correctly model a coarse-grained polymer. We have not yet revisited the four-scale simulation after carrying out the analysis described in Chapter 8; performing these simulations at the present time should lead to significantly more accurate results from the simulation.

End effects

When gauging the accuracy of the coarse-grained algorithm it is important to recognize that the coarse-graining process will naturally affect the definition of what constitutes a “correct” answer for a given simulation. For example, there is a pronounced end effect in the size of a polymer as a function of the number of beads retained in the coarse-grained simulation. For a chain of length N , the observed radius of gyration squared, $\langle R_g^2 \rangle_{cg}$, for a coarse-grained simulation will be smaller than the value obtained from a simulation without coarse-graining, $\langle R_g^2 \rangle_{exact}$. The shortfall results because we have effectively omitted the “tails” of the chain, representing the distances between the beginning of the fine-grained chain and the first coarse-grained bead, and between the last coarse-grained bead and the end of the fine-grained chain, from our calculation of the radius of gyration.

A plot of this end effect is shown in Figure 6-10 for fine-grained chains of lengths $N = 512$ and $N = 1024$. The ratio between the coarse-grained and non-coarse-grained radii of gyration, $\langle R_g^2 \rangle_{cg} / \langle R_g^2 \rangle_{exact}$, is plotted versus the number of coarse-grained beads N_b . The solid curve

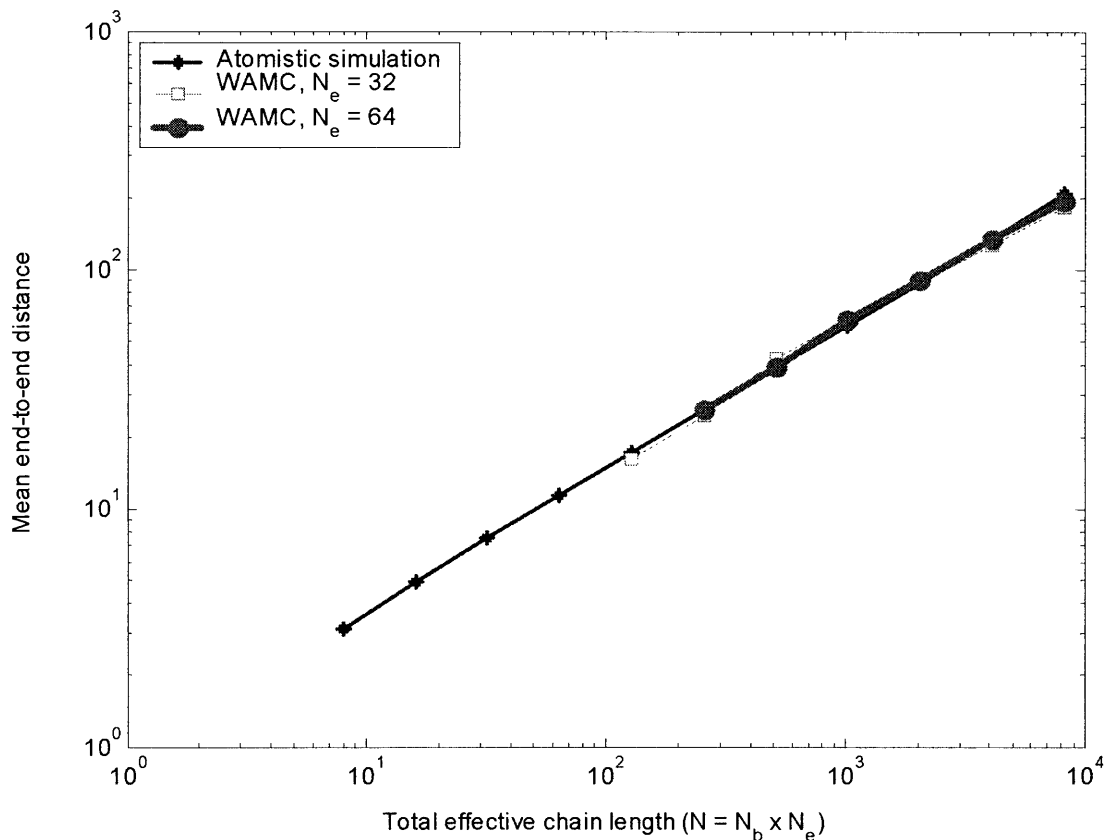


Figure 6-8: Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (diamonds) and of coarse-grained simulations based on bead sizes of 32 and 64 (squares and circles, respectively).

shows that the relative error is inversely proportional to N_b :

$$\frac{\langle R_g^2 \rangle_{cg}}{\langle R_g^2 \rangle_{exact}} \approx 1 - \frac{N_e}{N} = 1 - \frac{1}{N_b}. \quad (6.6)$$

We would expect comparable results for self-avoiding chain, although there may exist a scaling exponent related to the Flory exponent ν that is not observed in Eq. (6.6). Consequently, we again are faced with the challenge of balancing the needs for accuracy and performance: the bigger the chain, the smaller the resultant finite-size effects will be, but the slower the execution of the algorithm as well.

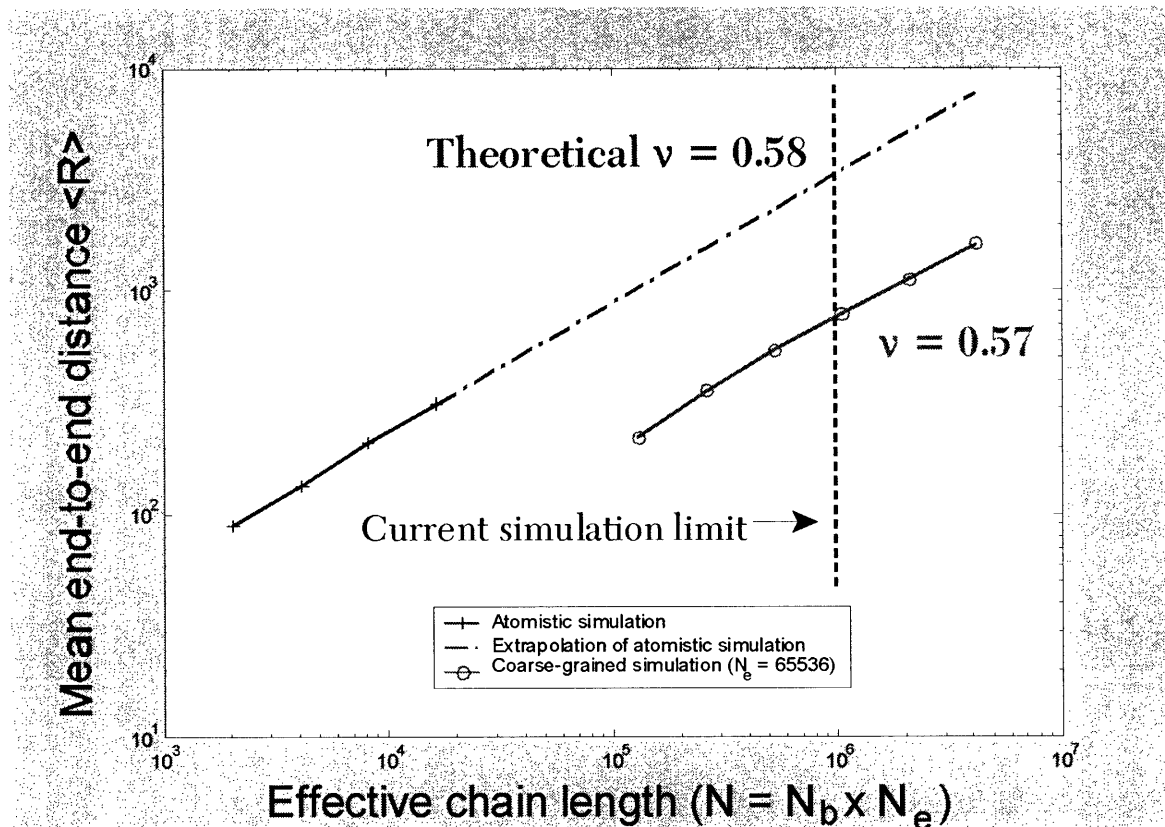


Figure 6-9: Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (top line) and of a four-scale coarse-grained simulations with bead sizes of 1, 128, 2048, and 65536 per stage. The observed Flory exponent is $\nu = 0.568$. Note that the end-to-end distance for the pivot algorithm is extrapolated from the behavior for $N \leq 16384$ using the correct scaling exponent $\nu \approx 0.577$.

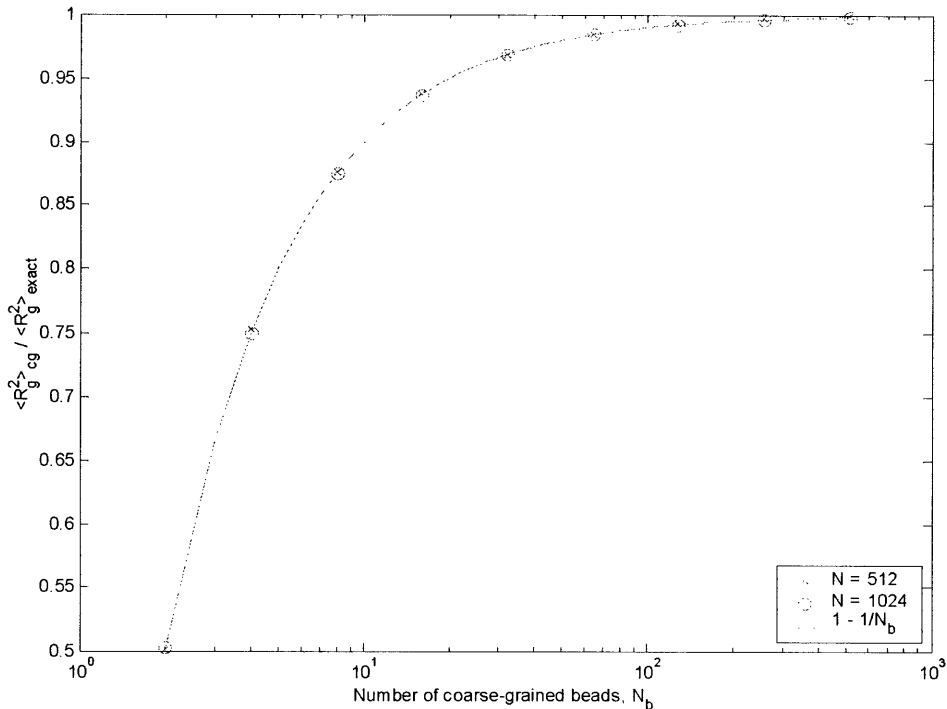


Figure 6-10: Demonstration of the end effect on the structural properties of an ideal Gaussian chain. The ratio $\langle R_g^2 \rangle_{cg} / \langle R_g^2 \rangle_{exact}$ was computed for different numbers of coarse-grained beads and for chains of length $N = 512$ and $N = 1024$. The comparison shows that the magnitude of the error is proportional to $1/N_b$, where N_b is the number of coarse-grained beads used.

6.3.3 Effects of additional interactions

Thus far, we have discussed only freely-jointed chains and self-avoiding chains. However, we want to consider the effects of additional interactions, such as stiffness potentials and non-bonded interactions, on the properties of the chain and of the individual bonds. Using the pivot algorithm, we simulated three different chains on-lattice, generating 50,000 independent configurations of length $N = 1024$: a self-avoiding walk with no other interactions, a self-avoiding walk with a stiffness potential of the form (2.8), with $\varepsilon_s = 1.0$, and a self-avoiding walk with stiffness interaction (2.8) with $\varepsilon_s = 1.0$ and non-bonded interaction (2.7) with strength $\varepsilon_{nb} = 1.0$. The distribution of bond lengths connecting coarse-grained segments of 16 beads each obtained from the different simulations is plotted in Figure 6-11. Note that there is essentially

Table 6.2: Radial statistics for self-avoiding walks with different potentials

	SAW	$\epsilon_s = 1.0$	$\epsilon_s = \epsilon_{nb} = 1.0$
$\langle r \rangle$ (16-mers)	4.8126	4.8129	4.8132
$\langle R^2 \rangle_{SAW}$	4116.98	4111.78	4108.07

no difference in the statistics for the distribution of a single bond; the mean bond distances and mean square end-to-end distances are essentially identical, as shown in Table 6.2. This suggests that the non-bonded interaction is of negligible effect for small values of ϵ_{nb} relative to $k_B T$.

Table 6.2 reveals that for small values of ϵ_s and ϵ_{nb} —that is, for values less than or equal to thermal energy $k_B T$ already present—the effect of these additional potentials on the behavior of the system is small. However, it is obvious that if the stiffness potential ϵ_s is large enough, the chain will prefer to exist in a rod-like conformation, for which the end-to-end distance and the distance between coarse-grained beads should grow essentially linearly with the number of beads N . As shown for a chain of length $N = 256$ in Figure 6-12, the mean square end-to-end distance $\langle R^2 \rangle$ is essentially constant for $\epsilon_s/k_B T < 5$, at which point it rapidly increases, and begins levelling off around $\epsilon_s/k_B T = 20$. A plateau is reached near this value because the chain is at maximal extension and cannot expand further. Consequently, we can suggest that a model without stiffness potential can be used to model systems for which $\epsilon_s \lesssim k_B T$.

6.3.4 Error tolerance and the $N \rightarrow \infty$ critical point

It is well known that the limit of $N \rightarrow \infty$ acts as a critical point in the study of polymer solutions [38, 64], much like the limit $T \rightarrow T_c$ for critical behavior in the Ising model. As discussed in Chapters 3 and 4, the WAMC algorithm acts much like a mean-field theory, in the sense that it tends to move the system away from the critical point as coarse-graining increases. Consequently, in the limit of very large polymer sizes, we would expect to see that the algorithm begins to break down as the number of stages of coarse-graining increases. However, we see little evidence of this in Figure 6-8 or in Table 6.1.

Perhaps the most likely explanation for why we were able to observe this phenomenon for the Ising model but not for the polymer model is the nature of the two critical points. The Ising critical point occurs at a finite temperature, which in principle we can reach with arbitrary

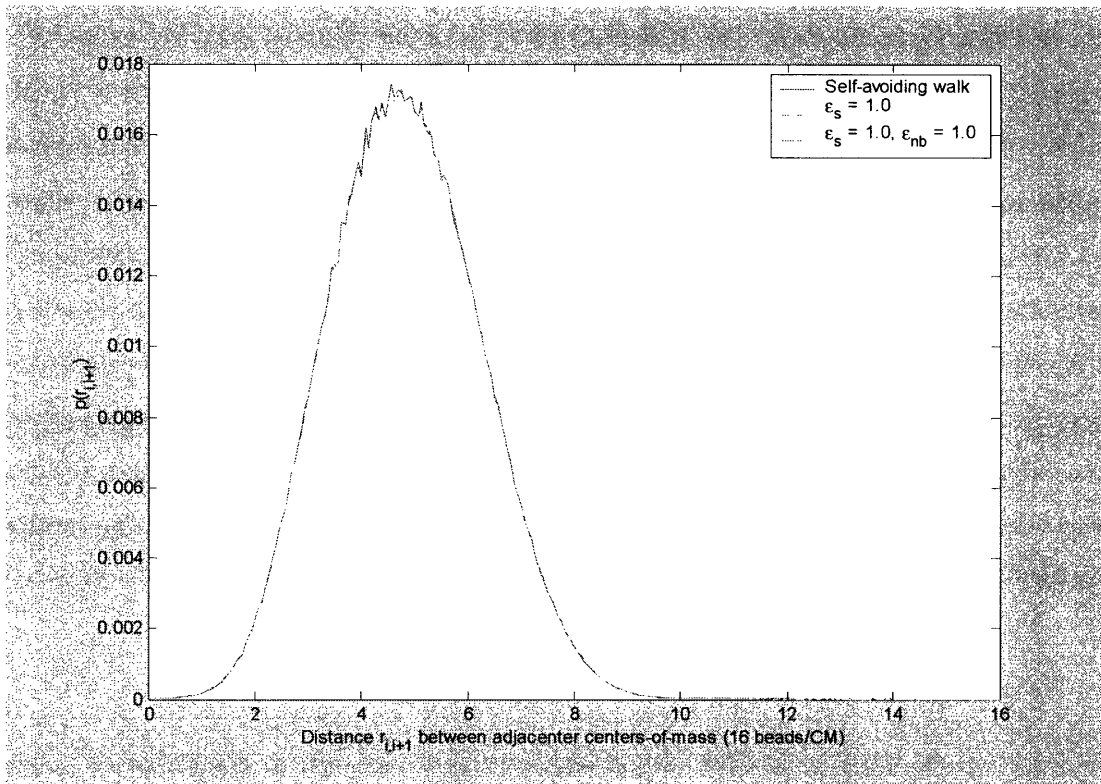


Figure 6-11: Probability distribution for the distance between adjacent centers of mass of 16-bead segments for random walks of length $N = 1024$.

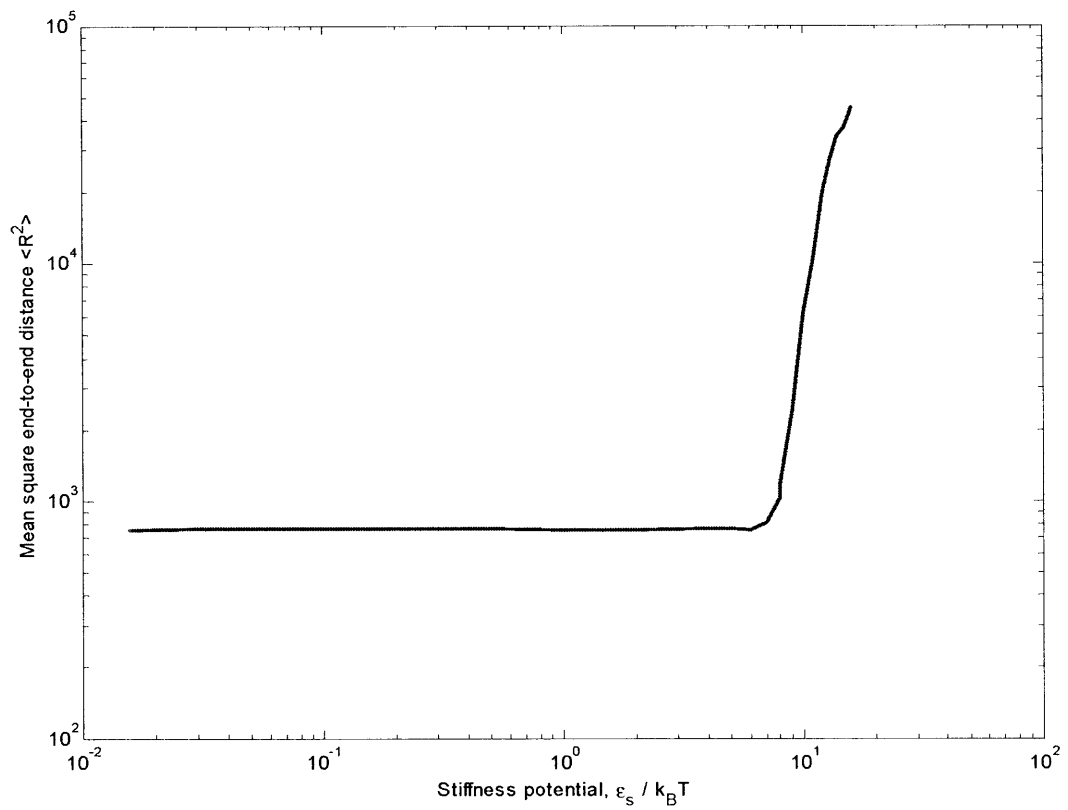


Figure 6-12: Mean square end-to-end distance $\langle R_{256}^2 \rangle$ for a chain of length 256 as a function of the stiffness potential ϵ_s .

precision by selecting our simulation temperature. For the polymer chain, in contrast, we have a singular limit: we cannot practically reach the polymer limit $N \rightarrow \infty$. This indicates that we may not have extended our results to chain sizes where we would observe increased error, or that the discrepancy occurs for chain sizes beyond the scope of interest.

Figure 6-8 does not show large errors in the predictions for the mean end-to-end distance $\langle R \rangle$ as a function of the level of coarse-graining. However, we do observe that while the scaling laws return nearly identical estimates for the Flory exponent for $N_e = 32$ and $N_e = 64$, there is a greater fluctuation around the scaling estimate for the coarse-grained simulations than for the corresponding atomistic simulation of the chain. This is somewhat unexpected, since in our Ising model simulations, we found that coarse-graining had a tendency to dampen fluctuations, not increase them. On the other hand, it may be that the nature of the coarse-graining for polymer chains is an increased “range” in the results: the overall predictions are correct, but a given result may be off by several percent or more. This is particularly true as the chain size increases; the largest deviations were observed at the largest chain lengths.

6.4 Performance

The paramount reason for employing coarse-graining is to increase the computational efficiency of a simulation. Therefore, let us consider the relative cost of our algorithm, as determined through clock measurements of performance. Generating 1.024×10^7 trial moves for an atomistic random walk of length $N = 1024$ required nearly 40 hours of computational time spread out over 20 1.4-GHz Pentium IV processors of a Beowulf cluster. By comparison, generating the equivalent number of trial moves using the WAMC algorithm for $N_b = 32$ coarse-grained beads, each representing $N_e = 32$ beads on the original chain, runs in approximately 25 minutes—a speed-up of two orders of magnitude, and comparable to other results based on wavelet transform averaging [148].

We compare the performance of the various algorithms over a variety of chain sizes in Figure 6-13. The topmost line in the graph corresponds to a detailed atomistic simulation using a WAMC-like algorithm; the running time of this algorithm is approximately $O(N^{7/4})$. If we optimize the atomistic algorithm by using the pivot algorithm, along with some of the

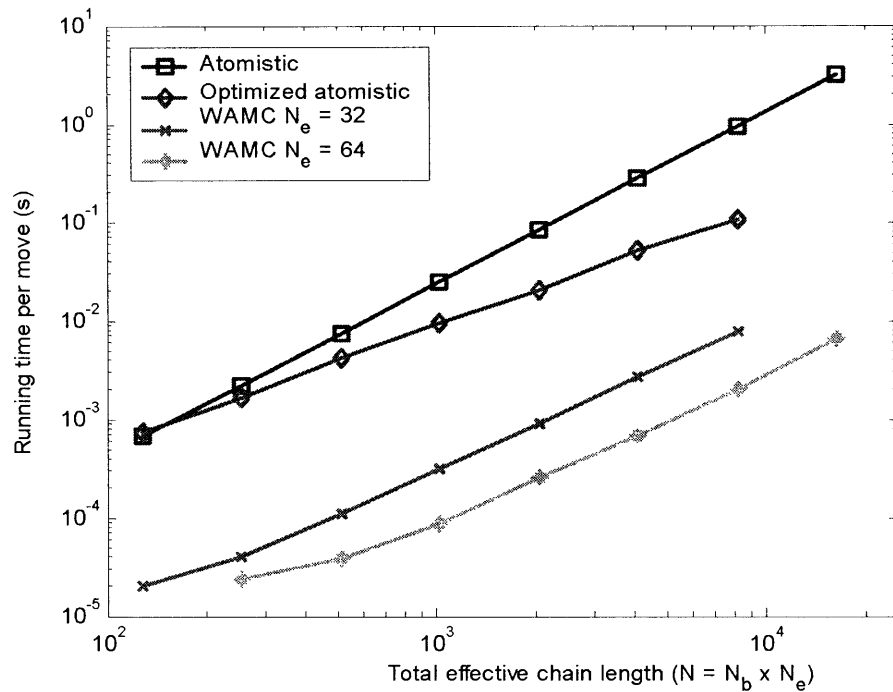


Figure 6-13: Running time comparison for atomistic simulation (squares), optimized atomistic simulation (diamonds), and WAMC algorithm with $N_e = 32$ (crosses) and $N_e = 64$ (asterisks).

additional improvements suggested by Kennedy [88], we obtain the “optimized atomistic” plot shown. The running time has been reduced to roughly $O(N^{6/5})$. Still further optimizations can reduce the running time to $O(N)$ or below; however, the resulting algorithm cannot be used in models with interparticle potentials that are continuous functions of the distance between particles.

The lower two lines plotted in Figure 6-13 represent the running time of the WAMC algorithm with $N_e = 32$ and $N_e = 64$ (crosses and asterisks, respectively). We note that both algorithms have a running time of $O(N_b^{7/5})$, which is in between the results for the atomistic and optimized atomistic algorithms. Because we have by definition $N_b \ll N$, we expect that the WAMC algorithm will be faster on a per-move basis than either atomistic algorithm, unless the associated prefactor is much larger for the WAMC algorithm than for the atomistic algorithms. However, since all of the methods are variations on the same basic algorithm, it is unlikely that the prefactor for WAMC will be larger than for the atomistic model. Therefore, we conclude

that the WAMC will generally be much faster than either atomistic technique.

One more interesting feature of the WAMC algorithm can be noted in Figure 6-13: the curves for $N_e = 32$ and $N_e = 64$ are essentially parallel to one another. This implies that the running time of the algorithm is independent of the effective bead size N_e . As we showed in Section 6.3 above, the accuracy of the simulation is essentially unaffected by the choice of N_e . Taken together, this means that in general, we can generally improve the performance of our algorithm by choosing N_e to be larger. However, as N_e increases, the running time of the previous stage of the simulation increases. We therefore conclude that optimizing the running time of the entire simulation requires us to balance the number of degrees of freedom in each level of the simulation as evenly as possible.

An example of the resulting improvement is shown in Figure 6-14, which shows the theoretical running time of the atomistic simulation as $N \rightarrow \infty$, compared to the observed running times for the four-scale simulation described above. At the largest chain size considered, the speed-up is more than six orders of magnitude: one *minute* of CPU time for the four-scale simulation with 64 beads of size 65536 corresponds to one *year* of simulation for an atomistic simulation of a chain of length 2^{22} beads.

6.4.1 Decorrelation of samples

Another important consideration in the algorithm is the number of simulation steps required to obtain independent configurations, as this governs the required length of the simulation. The decorrelation time for the coarse-grained simulation is linear, but with respect to N_b instead of N . A demonstration of this is shown as Figure 6-15, which illustrates the configuration of a coarse-grained polymer chain with $N_b = 256$ beads and effective size $N_e = 64$ taken 256 steps apart, which shows that our algorithm has an effectively linear decorrelation time.

To obtain z independent samples, the atomistic simulation based on the pivot algorithm requires, on average, $O(zN)$ steps [105]. The decorrelation of states can be confirmed after the simulation using the blocking technique of Flyvbjerg and Petersen [60], or during the simulation using the “on-the-fly” implementation of that method developed by Feldmann *et al.* [51, 52].

The decorrelation time for the coarse-grained simulation is linear, but with respect to N_b rather than N . A demonstration of this is shown as Figure 6-16, which plots the autocorrelation

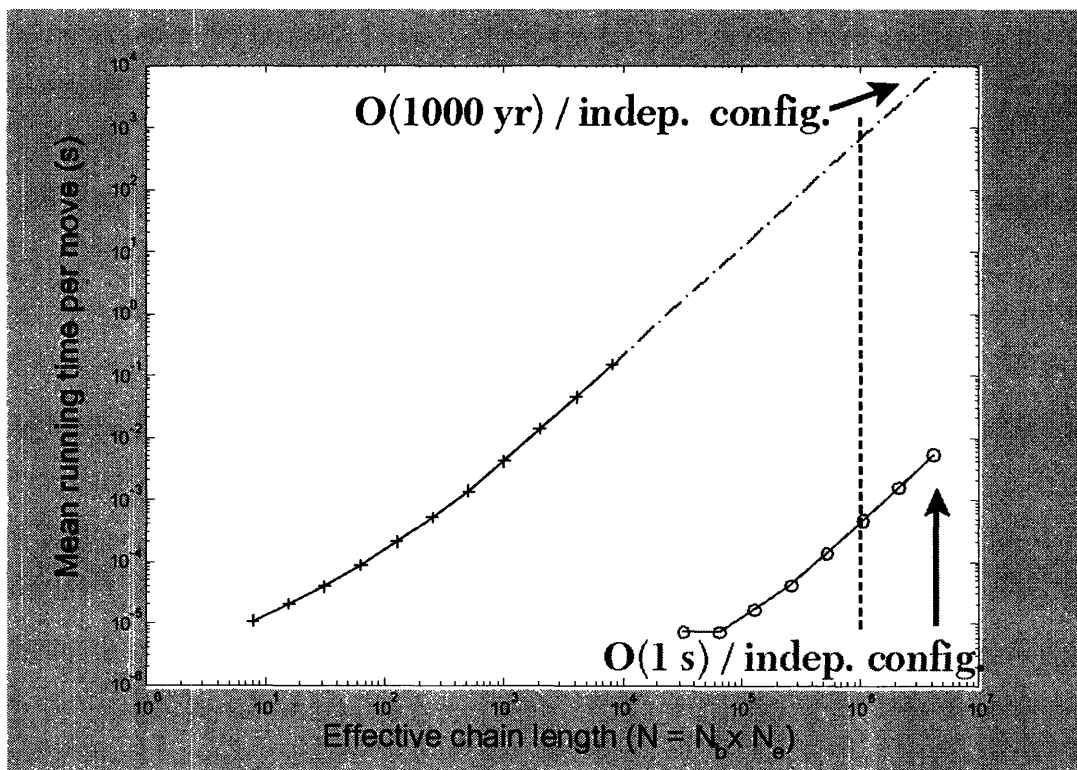


Figure 6-14: Running time per move of the pivot algorithm (upper curve) and of a four-scale coarse-grained algorithm with bead sizes of $N = (1, 128, 2048, 65536)$ per stage (lower curve).

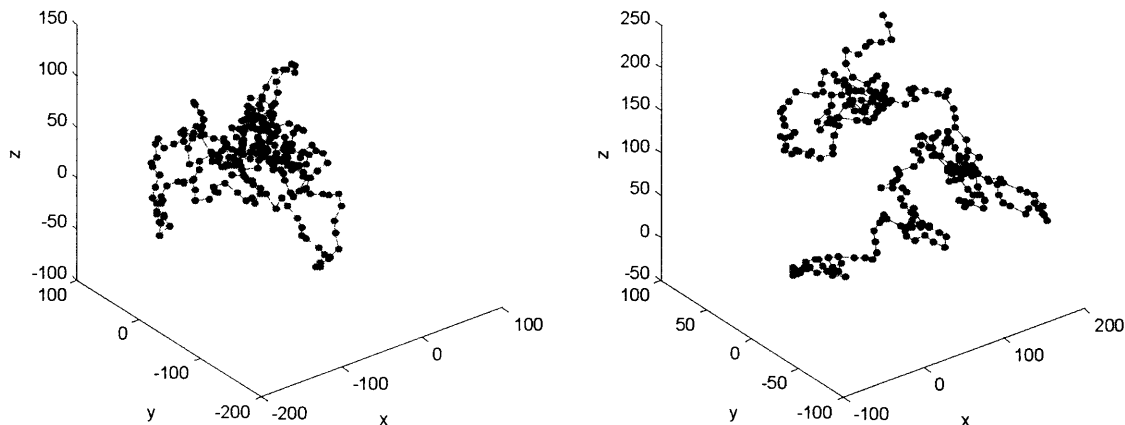


Figure 6-15: The configuration of a coarse-grained polymer chain with $N_b = 256$ taken 256 steps apart.

function of the end-to-end distance R for a coarse-grained polymer chain with $N_b = 256$ beads and effective size $N_e = 64$, which shows that our algorithm has a linear decorrelation time. As a consequence of this, we conclude that for a two-scale simulation, the total time required for z independent samples is $O(z(N_{b,1} + N_{b,2}))$, and since $N_{b,1} + N_{b,2} \ll N$, the number of simulation moves required also decreases through use of the hierarchical procedure. The overall efficiency increases with the number of stages in the simulation, as the total time required should scale as $O(z \sum_{i=1}^m N_{b,i})$.

6.5 Geometric reverse-mapping of a polymer random walk

6.5.1 Description of the constraints

We want to reverse-map a chain of length N to a chain of length $2N$. Since we would like to replace a bead on the coarse-grained chain with two beads on the corresponding fine-grained chain, we will need a total of six constraints for each pair of beads created. The definition of

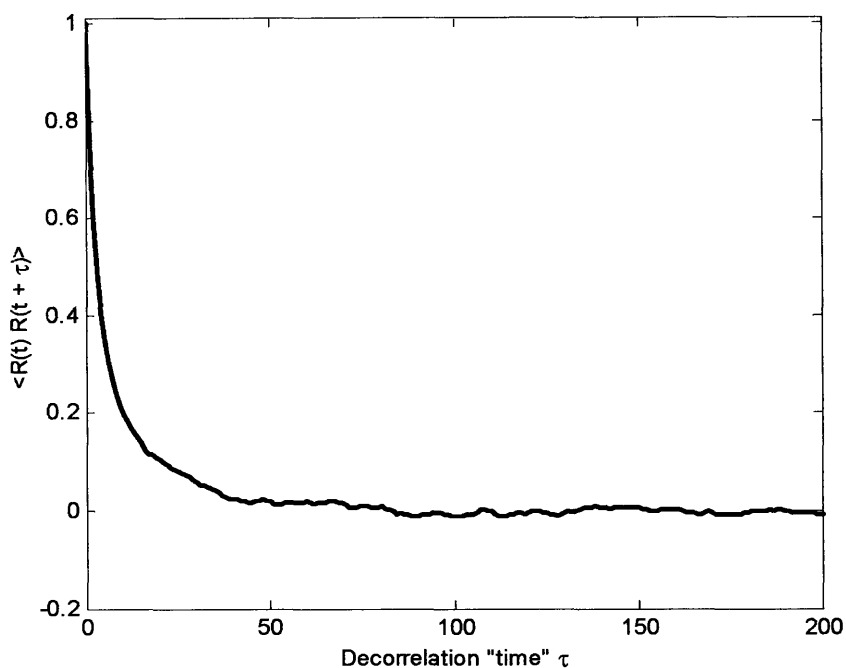


Figure 6-16: Autocorrelation function for a the end-to-end distance of a self-avoiding walk studied using WAMC with $N_b = 256$ and $N_e = 64$.

the wavelet transform, as applied to the polymer chain,

$$\mathbf{r}_{2i-1}^{(k-1)} + \mathbf{r}_{2i}^{(k-1)} = 2\mathbf{r}_i^{(k)}, \quad (6.7)$$

provides three pieces of information. [The superscript k denotes the coarse-grained chain and $k - 1$ the fine-grained chain throughout.] An additional constraint must obviously be the displacement of the two points relative to the center of mass—in other words, the bond length between the two new points to be constructed:

$$\left(\mathbf{r}_{2i}^{(k-1)} - \mathbf{r}_{2i-1}^{(k-1)} \right) \cdot \left(\mathbf{r}_{2i}^{(k-1)} - \mathbf{r}_{2i-1}^{(k-1)} \right) = l^2, \quad (6.8)$$

where l is the new bond length at the fine scale. The new bond length l can either be specified as a independent parameter, or determined as a geometric constraint. For the purposes of this presentation, we choose the latter approach, and require the distance l to be a multiple λ of the

distance spanned by the two bonds adjacent to the coarse-grained bead we wish to reverse-map:

$$l = \lambda \left| \mathbf{r}_{i+1}^{(k)} - \mathbf{r}_{i-1}^{(k)} \right|. \quad (6.9)$$

This is a purely matter of convenience; we could have chosen any other constraint similar to (6.9) for use here.

In addition to the bond-length constraint, we can create two other purely geometric constraints in selecting our reverse-mapping criterion. One of the two remaining criteria will be that the new vector must lie in the same plane as the two coarse-grained bonds on either side of the bead being replaced. That is, we want the new coordinates to satisfy

$$\left(\mathbf{r}_{2i}^{(k-1)} - \mathbf{r}_{2i-1}^{(k-1)} \right) \cdot \left(\left(\mathbf{r}_i^{(k)} - \mathbf{r}_{i-1}^{(k)} \right) \times \left(\mathbf{r}_{i+1}^{(k)} - \mathbf{r}_i^{(k)} \right) \right) = 0. \quad (6.10)$$

The other constraint that we want is for the new bond to be tangent to the two coarse-grained so that the angle between the new bond and either of the coarse-grained bonds are equal. Mathematically, we require

$$\frac{\left(\mathbf{r}_{2i}^{(k-1)} - \mathbf{r}_{2i-1}^{(k-1)} \right) \cdot \left(\mathbf{r}_i^{(k)} - \mathbf{r}_{i-1}^{(k)} \right)}{z \left| \mathbf{r}_i^{(k)} - \mathbf{r}_{i-1}^{(k)} \right|} = \frac{\left(\mathbf{r}_{2i}^{(k-1)} - \mathbf{r}_{2i-1}^{(k-1)} \right) \cdot \left(\mathbf{r}_{i+1}^{(k)} - \mathbf{r}_i^{(k)} \right)}{z \left| \mathbf{r}_{i+1}^{(k)} - \mathbf{r}_i^{(k)} \right|}. \quad (6.11)$$

6.5.2 Implementation of the algorithm

We would like to convert constraints (6.8), (6.10) and (6.11) into a form which can be programmed into an algorithm. Let us assume that we want to reverse-map point $\mathbf{r}_i^{(k)}$ on the coarse-grained scale into $\mathbf{r}_{2i-1}^{(k-1)}$ and $\mathbf{r}_{2i}^{(k-1)}$. We shall denote the components of the two beads at the fine scale as

$$\begin{aligned} \mathbf{r}_{2i-1}^{(k-1)} &= (x_1, y_1, z_1), \\ \mathbf{r}_{2i}^{(k-1)} &= (x_2, y_2, z_2), \end{aligned}$$

and the components of the coarse-grained bead $\mathbf{r}_i^{(k)}$ and its nearest neighbors at the coarse scale as follows

$$\mathbf{r}_{i-1}^{(k)} = (a_1, b_1, c_1)$$

$$\mathbf{r}_i^{(k)} = (a_2, b_2, c_2)$$

$$\mathbf{r}_{i+1}^{(k)} = (a_3, b_3, c_3)$$

and the new bond length as l . Then, we have from (6.7) above

$$x_1 + x_2 = 2a_2, \quad (6.12)$$

$$y_1 + y_2 = 2b_2, \quad (6.13)$$

$$z_1 + z_2 = 2c_2, \quad (6.14)$$

for the midpoint constraint. The coplanarity constraint is

$$A_1(x_2 - x_1) + A_2(y_2 - y_1) + A_3(z_2 - z_1) = 0, \quad (6.15)$$

where

$$A_1 = b_2c_3 - b_2c_1 - b_1c_3 - c_2b_3 + c_2b_1 + c_1b_3, \quad (6.16)$$

$$A_2 = c_2a_3 - c_2a_1 - c_1a_3 - a_2c_3 + a_2c_1 + a_1c_3, \quad (6.17)$$

$$A_3 = a_2b_3 - a_2b_1 - a_1b_3 - b_2a_3 + b_2a_1 + b_1a_3, \quad (6.18)$$

The equiangularity constraint is

$$\begin{aligned} & (R_2(a_1 - a_2) + R_1(a_3 - a_2))(x_1 - x_2) \\ & + (R_2(b_1 - b_2) + R_1(b_3 - b_2))(y_1 - y_2) \\ & + (R_2(c_1 - c_2) + R_1(c_3 - c_2))(z_1 - z_2) = 0, \quad (6.19) \end{aligned}$$

where

$$R_1 = \sqrt{(a_2 - a_1, b_2 - b_1, c_2 - c_1) \cdot (a_2 - a_1, b_2 - b_1, c_2 - c_1)}, \quad (6.20)$$

$$R_2 = \sqrt{(a_3 - a_2, b_3 - b_2, c_3 - c_2) \cdot (a_3 - a_2, b_3 - b_2, c_3 - c_2)}. \quad (6.21)$$

Finally, the bond length constraint is given by

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = l^2. \quad (6.22)$$

If the quadratic constraint in equation (6.22) were linear, we could solve the problem easily using Gaussian elimination using existing methods. However, since the constraint is non-linear, we will need to find an alternative approach. What we shall do is solve (6.12) through (6.19) in terms of one of the six unknowns, and then substitute the expressions for the other five variables into (6.22) to determine the sixth. We can then backsubstitute to determine the other variables. The first three equations, (6.12) through (6.14), are easily converted into

$$x_2 = 2a_2 - x_1, \quad (6.23)$$

$$y_2 = 2b_2 - y_1, \quad (6.24)$$

$$z_2 = 2c_2 - z_1, \quad (6.25)$$

inserting (6.23) through (6.25) into (6.15), (6.19), and (6.22) give

$$A_1 (a_2 - x_1) + A_2 (b_2 - y_1) + A_3 (c_2 - z_1) = 0, \quad (6.26)$$

$$\begin{aligned} & (R_2 (a_1 - a_2) + R_1 (a_3 - a_2)) (x_1 - a_2) \\ & + (R_2 (b_1 - b_2) + R_1 (b_3 - b_2)) (y_1 - b_2) \\ & + (R_2 (c_1 - c_2) + R_1 (c_3 - c_2)) (z_1 - c_2) = 0, \quad (6.27) \end{aligned}$$

and

$$(x_1 - a_2)^2 + (y_1 - b_2)^2 + (z_1 - c_2)^2 = \frac{l^2}{4}. \quad (6.28)$$

Eliminating z_1 from (6.26) gives

$$z_1 = c_2 + \frac{A_1}{A_3} (a_2 - x_1) + \frac{A_2}{A_3} (b_2 - y_1); \quad (6.29)$$

inserting this into (6.27) and (6.28) yields

$$\begin{aligned} & \left(R_2 (a_1 - a_2) + R_1 (a_3 - a_2) - \frac{A_1}{A_3} (R_2 (c_1 - c_2) + R_1 (c_3 - c_2)) \right) (x_1 - a_2) \\ & + \left(R_2 (b_1 - b_2) + R_1 (b_3 - b_2) - \frac{A_2}{A_3} (R_2 (c_1 - c_2) + R_1 (c_3 - c_2)) \right) (y_1 - b_2) = 0, \end{aligned} \quad (6.30)$$

and

$$(x_1 - a_2)^2 + (y_1 - b_2)^2 + \left(\frac{A_1}{A_3} (a_2 - x_1) + \frac{A_2}{A_3} (b_2 - y_1) \right)^2 = \frac{l^2}{4}. \quad (6.31)$$

We now need to eliminate one of the two remaining variables from (6.30) and (6.31). For convenience, we choose to rewrite these equations in the form

$$F_1 (x_1 - a_2) + F_2 (y_1 - b_2) = 0, \quad (6.32)$$

$$\left[\frac{A_1^2 + A_3^2}{A_3^2} \right] (x_1 - a_2)^2 + \left[\frac{A_2^2 + A_3^2}{A_3^2} \right] (y_1 - b_2)^2 + \frac{2A_1 A_2}{A_3^2} (x_1 - a_2) (y_1 - b_2) = \frac{l^2}{4}, \quad (6.33)$$

where

$$\begin{aligned} F_1 &= R_2 (a_1 - a_2) + R_1 (a_3 - a_2) - \frac{A_1}{A_3} (R_2 (c_1 - c_2) + R_1 (c_3 - c_2)), \\ F_2 &= R_2 (b_1 - b_2) + R_1 (b_3 - b_2) - \frac{A_2}{A_3} (R_2 (c_1 - c_2) + R_1 (c_3 - c_2)), \end{aligned}$$

Then, y_1 is given by

$$y_1 = \frac{F_1}{F_2} (a_2 - x_1) + b_2, \quad (6.34)$$

which makes (6.33)

$$G (x_1 - a_2)^2 = \frac{l^2}{4}, \quad (6.35)$$

where

$$G = \left(1 + \frac{A_1^2}{A_3^2}\right) + \left(1 + \frac{A_2^2}{A_3^2}\right) \left(\frac{F_1}{F_2}\right)^2 - \frac{2A_1A_2}{A_3^2} \left(\frac{F_1}{F_2}\right). \quad (6.36)$$

Solving (6.35) for x_1 yields two solutions:

$$x_1 = a_2 \pm \frac{1}{2}lG^{-1/2}. \quad (6.37)$$

After we have obtained either solution from (6.37), we can backsolve for the remaining solutions first using (6.34), then (6.29), and then finally determining the other point using (6.23) through (6.25).

One interesting feature should be noted: it does not matter which solution of (6.37) we use to determine the remaining variables; whichever root of (6.37) we do not assign to x_1 will in fact be x_2 . We can then use connectivity constraints on the coarse-grained chain to differentiate the two points.

6.5.3 Pathological cases

The above method only works if the points are not colinear in the plane $x = x_0$, where x_0 is some constant. In that case, we cannot use x as our variable of choice, and we must reformulate the problem in terms of y or z . This is because for points which are coplanar in x , $A_2 = A_3 = 0$, and thus we cannot calculate G in (6.36). Equations (6.23) through (6.28) remain the same, but now we need to eliminate x_1 from (6.26), instead of z_1 . Eliminating x_1 gives

$$x_1 = a_2 + \frac{A_2}{A_1}(b_2 - y_1) + \frac{A_3}{A_1}(c_2 - z_1); \quad (6.38)$$

however, since $A_2 = A_3 = 0$, (6.38) reduces to

$$x_1 = a_2. \quad (6.39)$$

Inserting (6.39) into (6.38) and (6.22) gives

$$F_3(y_1 - b_2) + F_4(z_1 - c_2) = 0, \quad (6.40)$$

and

$$(y_1 - b_2)^2 + (z_1 - c_2)^2 = \frac{l^2}{4}, \quad (6.41)$$

where in (6.40)

$$\begin{aligned} F_3 &= R_2(b_1 - b_2) + R_1(b_3 - b_2), \\ F_4 &= R_2(c_1 - c_2) + R_1(c_3 - c_2). \end{aligned}$$

Eliminating y_1 from (6.40) gives

$$y_1 = b_2 - \frac{F_4}{F_3}(z_1 - c_2); \quad (6.42)$$

combining (6.41) and (6.42) yields the quadratic equation

$$H_1(z_1 - c_2)^2 = \frac{l^2}{4}, \quad (6.43)$$

where

$$H_1 = 1 + \left(\frac{F_4}{F_3}\right)^2.$$

Solving (6.43) for z_1 leads to the solutions

$$z_1 = c_2 \pm \frac{1}{2}lH_1^{-1/2}.$$

We can now backsubstitute as before.

Alternatively, if the points are coplanar in y , we will need to use

$$y_1 = b_2,$$

from which we find

$$F_5(x_1 - a_2) + F_6(z_1 - c_2) = 0, \quad (6.44)$$

where

$$\begin{aligned} F_5 &= R_2 (a_1 - a_2) + R_1 (a_3 - a_2) \\ F_6 &= R_2 (c_1 - c_2) + R_1 (c_3 - c_2) \end{aligned}$$

which leads to the quadratic equation

$$H_2 (z_1 - c_2)^2 = \frac{l^2}{4}, \quad (6.45)$$

from which we obtain the solution

$$z_1 = c_2 \pm \frac{1}{2} l H_2^{-1/2},$$

where now we have

$$H_2 = 1 + \left(\frac{F_6}{F_5} \right)^2.$$

6.5.4 Handling the endpoints

The other “pathological” case are the end points of the chain, for which we do not have two different bonds to work with. Consequently, we cannot use the dot product and normality constraints as formulated above. We must find two alternate constraints with which to replace these points. To accomplish this, we will require that the segment connecting the second and third points from each end of the chain be parallel to the first and last coarse-grained bonds. This, coupled with the bond length and symmetry requirements should be sufficient information to locate the other beads.

Given $\mathbf{r}_2^{(k)} - \mathbf{r}_1^{(k)}$, and $\mathbf{r}_n^{(k)} - \mathbf{r}_{n-1}^{(k)}$, we can determine $\mathbf{r}_2^{(k-1)}$ and $\mathbf{r}_{2n-1}^{(k-1)}$ as

$$\mathbf{r}_2^{(k-1)} = \mathbf{r}_3^{(k-1)} - \frac{1}{3} \left(\mathbf{r}_2^{(k)} - \mathbf{r}_1^{(k)} \right), \quad (6.46)$$

$$\mathbf{r}_{2n-1}^{(k-1)} = \mathbf{r}_{2n-2}^{(k-1)} + \frac{1}{3} \left(\mathbf{r}_n^{(k)} - \mathbf{r}_{n-1}^{(k)} \right). \quad (6.47)$$

Then $\mathbf{r}_0^{(k-1)}$ and $\mathbf{r}_{2n}^{(k-1)}$ will be

$$\mathbf{r}_0^{(k-1)} = 2\mathbf{r}_1^{(k)} - \mathbf{r}_2^{(k-1)}, \quad (6.48)$$

$$\mathbf{r}_{2n}^{(k-1)} = 2\mathbf{r}_n^{(k)} - \mathbf{r}_{2n-1}^{(k-1)}. \quad (6.49)$$

Taken together, (6.46) through (6.49) are sufficient to specify uniquely the remaining four positions on the fine-grained chain. Thus, we can specify a representative reverse-mapped chain for a given coarse-grained chain using purely geometric constraints.

6.5.5 Illustrative examples

We illustrate two examples of this reverse-mapping process in Figures 6-17 and 6-18. In each case, a walk of 64 beads is converted into a walk with 128 beads, using a factor of $\lambda = 0.45$ in (6.9). We note in each case that the fine-grained walk adheres closely to the contours of the coarse-grained walk. This is because in selecting our constraints, we have kept in mind the convexity property of the wavelet transform: for a given curve, the coarse-grained walk created by averaging pairs of nearest-neighbor points will always be “convex”; that is, the coarse-grained curve will lie “inside” the fine-grained curve.

How to selecting the parameter λ in (6.9) is an open question. If we choose to make our criterion the shortest overall contour length of the fine-grained chain, then, as we can see from Figure 6-19, there is no general range of λ for which all chains are at their minimum. Instead, the values of λ for which the contour length is a minimum varies throughout the range, which suggests that we will have to average the results of many chains to determine an optimal value for the parameter λ . As a preliminary guess, however, one suspects that the optimal value for λ will lie in the range $0.4 \leq \lambda \leq 0.6$.

6.5.6 Using the fine-grained representation

It is also useful to ask what is the purpose of this reverse-mapping algorithm. If our goal is to produce a configuration of minimal energy, then our algorithm is unlikely to meet that aim, since we do not take energetic considerations into account. In addition, because of the large number of degrees of freedom, the probability that any configuration will be “the” configuration

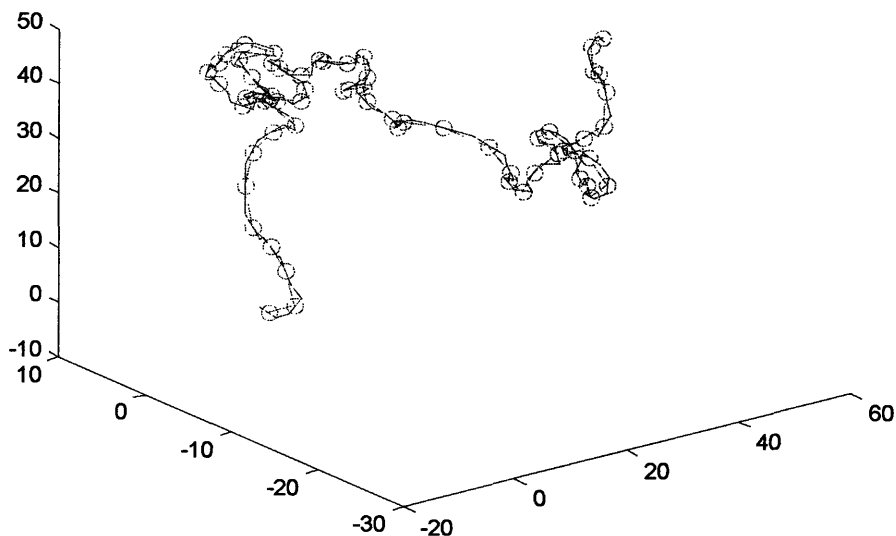


Figure 6-17: Reverse-mapping of a coarse-grained 64-mer to a fine-grained 128-mer. The coarse-grained chain is indicated by circles connected with dashed lines; the fine-grained chain is the solid curve.

of minimal energy is small. Using this geometric reverse-mapping as a starting point for an optimization process will likely produce a configuration which is a local minimum, but not the global minimum.

On the other hand, if we use this method to produce a configuration which is a starting point for further simulation, we will probably be on much safer footing, since we will have a configuration which is not too dissimilar from the coarse-grained configuration. This will be especially useful when we start having interactions between multiple chains in the semi-dilute or dense regimes.

6.6 Conclusions

We have demonstrated that the WAMC algorithm can be readily adapted to study molecular systems in which the coarse-graining should be topological. The method recovers the expected scaling law $\langle R^2 \rangle \sim N^{2\nu}$ with correct estimates for ν for both ideal and self-avoiding random

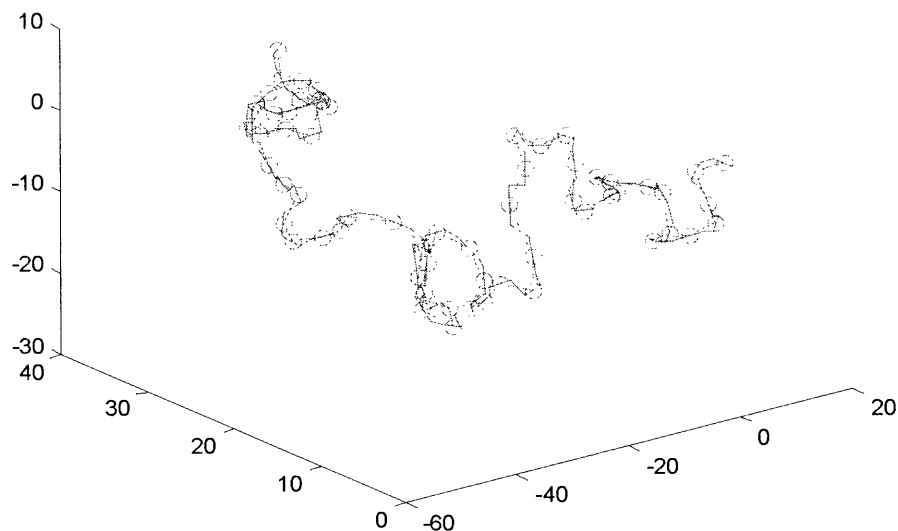


Figure 6-18: Reverse-mapping of a coarse-grained 64-mer to a fine-grained 128-mer. The coarse-grained chain is indicated by circles connected with dashed lines; the fine-grained chain is the solid curve.

walks, under different levels of coarse-graining and for a variety of chain lengths. In addition, we have illustrated that this algorithm achieves vast performance gains over “atomistic” simulations of polymers across chain lengths varying by several orders of magnitude. The ability to access conveniently large length scales within a single computational framework provides a flexible starting point from which we can develop simulations to study the behavior of semidilute and concentrated polymer solutions and polymer melts.

We have also shown that our method allows for the construction of coarse-grained probability distributions as well as potentials describing the interaction between coarse-grained particles. For freely-jointed chains, we have shown that very simple physical arguments can be used to approximate the two-body potentials observed for the overlap of chains. Similar phenomenological results can be derived for self-avoiding walks as well. Although the resulting analytical distributions tend to overestimate the potential, they show the correct qualitative behavior, suggesting that single-chain simulations may be sufficient to compute coarse-grained potentials.

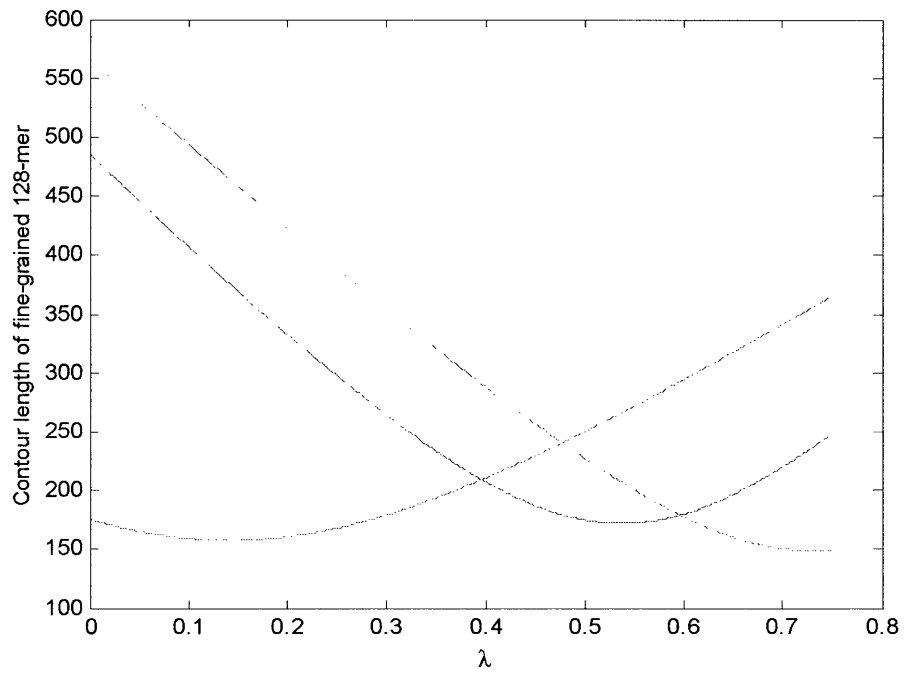


Figure 6-19: Contour length of the fine-grained 128-mer as a function of the parameter λ defined in (6.9) for various starting coarse-grained configurations.

Chapter 7

Coarse-grained potentials for polymer chains

7.1 Introduction

The work described in this chapter was undertaken as part of the investigation to achieve numerical agreement between our coarse-grained simulations and the results obtained from atomistic numerical simulations. Because the scaling results were essentially in agreement, it was suspected that some other parameter—such as the interparticle potential—was insufficient to describe the system.

For the coarse-grained simulations, our initial coarse-grained trials included as the interparticle potential the same hard-core repulsion that was assumed for the initial, atomistic simulations. However, the resulting values for the radius of gyration and end-to-end distance were significantly lower than the atomistic results. Following the work of Bolhuis *et al.* [20,22], we choose to compute the probability $F_N(\mathbf{x}, r)$ that a chain segment, containing N beads of radius r , will have a bead at a given distance $x = |\mathbf{x}|$ away from the center of mass. From the distribution $F_N(\mathbf{x}, r)$ we can compute a dimensionless potential $u(r)$ by taking the logarithm of $F_N(\mathbf{x}, r)$:

$$u(r) = -\ln F_N(\mathbf{x}, r). \quad (7.1)$$

The mean force potential that we obtain is structurally similar to those obtained by Bolhuis

et al., but the agreement is by no means exact. However, without additional details describing the algorithm they used, it is difficult to determine the source of the discrepancy.

Performing the simulations which would lead to an estimate of (7.1) led to another observation: there was a strong similarity among the various curves representing different numbers of beads, different bead sizes, and other parameters. This led us to look for scaling laws which would describe the probability distribution $F_N(\mathbf{x}, r)$. The results are presented below.

7.2 Definition of the problems to be examined

7.2.1 Transition to “soft” potentials

As stated above, the self-avoiding walk has a “hard-core” repulsive potential defined by (2.7). Thus, it is impossible for particles on the fine-grained chain to come closer to one another than the hard-core diameter a . However, when we consider the behavior of coarse-grained beads, we find that the hard-core potential begins to disappear, in the sense that coarse-grained beads can approach more closely together than the cutoff a . To do this, we construct the two-point correlation function $g^{(2)}(r_{ij})$, which tells us the probability of finding a particle a distance r_{ij} away from another, and, for a single chain molecule, satisfies the normalization criterion¹

$$\int_0^\infty dr g(r) = N_p - 1, \quad (7.2)$$

where ρ is the particle density of the system and N_p is the total number of particles. A sample plot of $g(r)$ for a chain of 512 beads is shown in Figure 7-1, we note that virtually all of the particles are found within a radius of 50 bond lengths from any other point on the chain, which is consistent with the $N^{3/5}$ scaling law ($512^{3/5} \approx 42.2$).

Although we know that a hard-core repulsion is inadequate, there is no general consensus on what model does correctly describe the physical interaction between the coarse-grained particles. Murat and Kremer [119], in their “soft ellipsoid algorithm” (further extended by Eurich and Maass [48]), generate the potential by determining the probability distribution $p(\mathbf{S})$ for the inertial matrix \mathbf{S} of the system, and then defining the force of the system as $F = -kT \ln p(\mathbf{S})$,

¹For a discrete distribution function, this means that the value of $g(r)$ for a given r will depend on the discretization Δr .

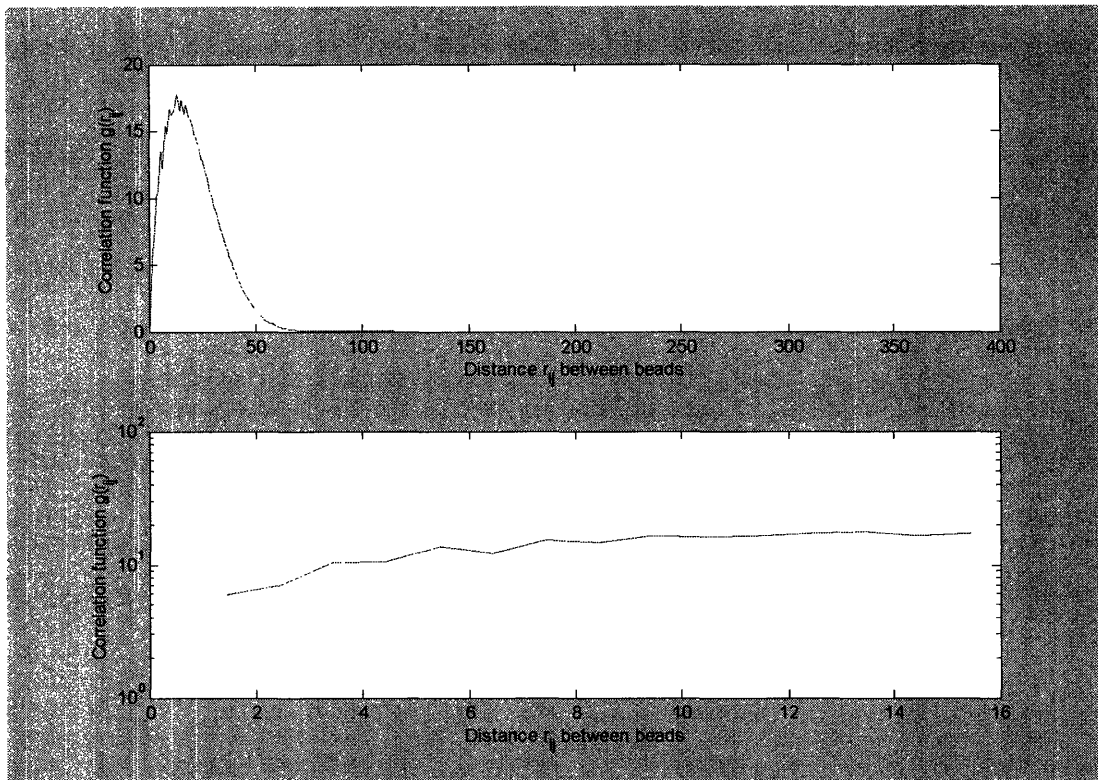


Figure 7-1: (Top) Semi-log plot of the two-body correlation function $g(r_{ij})$ as a function of the interparticle separation distance r_{ij} . (Bottom) Semilog plot for small values of r_{ij} showing the sharp cutoff in $g(r_{ij})$ for $r_{ij} \leq 1$.

and then generating the potential by integrating the force with respect to position. In the “soft colloid model” of Louis *et al.* [99], the potential is determined by solving the system of non-linear equations

$$\begin{aligned} g(r) &= \exp(-v(r) + g(r) - c(r) - 1), \\ c(r) &= g(r) - 1 - \rho \int dx c(x) (g(x) - 1), \end{aligned}$$

a method referred to as the “hypernetted chain approximation closure relation” [77]. However, each of these methods relies on mapping an entire chain into a single particle; to develop a hierarchical simulation method, we would like to determine a method capable of determining the potential at different levels of resolution of the chain.

7.2.2 Overlap Probabilities

We are interested in computing the probability of overlap, both between segments of a single chain and between segments on different chains. The methods used to determine the statistics of these types of overlap are different as a result of the different types of models we are considering.

One-chain overlap probability

The parameters of interest in this problem are the number of segments in the chain N_b , the number of fine-grain beads per segment N_e , and the test radius r_{test} , which gives the relative size of each bead. For a single chain, computing the overlap involves determining if the distance between bead i on segment a and bead j on segment b is less than r_{test} : that is, we check if

$$|\mathbf{r}_i - \mathbf{r}_j| < r_{test}, \tag{7.3}$$

for all $i \in [(a-1)N_e + 1, aN_e]$ and $j \in [(b-1)N_e + 1, bN_e]$ and $1 \leq a, b \leq N_b$. If (7.3) is false, then we proceed to the next configuration; if it is true for a given pair of a and b , then we assign weight 1 to the histogram $w(r_{ab})$, where

$$r_{ab} = |\mathbf{r}_a - \mathbf{r}_b|,$$

such that \mathbf{r}_a and \mathbf{r}_b are the centers of mass of each segment. The probability of overlap is then defined as

$$p(r) = \frac{w(r)}{n(r)}, \quad (7.4)$$

where $w(r)$ is the number of overlaps observed when the centers of mass are separated by distance r , and $n(r)$ is the total number of times the centers of mass were separated by distance r .

Two-chain overlap probability

We consider a slightly different problem here. Given two chains of length N , we perform a series of simulations on each chain; for instance, for two lattice walks, we perform a pivot simulation on each chain. After N moves, we compute the probability of overlap when their centers of mass are separated by a distance r . To determine the distance between the two chains, we move one chain along the vector joining their centers of mass such that the distance of the vector is r . We then test if any bead on the first chain intersects with any bead on the second chain:

$$r_{ab} = |\mathbf{r}_{i,a} - \mathbf{r}_{j,b}| < r_{test}, \quad (7.5)$$

where the subscripts a and b in (7.12) now denote the first and second chains, respectively. If (7.12) is satisfied for any i and j , then we increase $w(r_{ab})$ by one; otherwise we proceed to the next configuration to be tested. Again, the desired probability is (7.4).

7.3 Examples of scaling behavior

Before discussing the theory of overlap statistics, we first provide several illustrative examples of the scaling results obtained.

- To show that scaling behavior appears to be valid for a wide variety of models, we considered a lattice model with excluded-volume interactions, corresponding to a test volume of $r_{test} = 1$, plus stiffness and nearest-neighbor interactions given by $\varepsilon_{nb} = 1$ and $\varepsilon_s = 1$ in Eqs. (2.7) and (2.8), respectively. We show the scaled results for the probability of overlap in Figure 7-2 for chains of length $N = 2048$. It should be noted that although

the length scale has been scaled by $N_e^{-\nu}$, the probabilities have been scaled by N_e^ν , where $\nu \approx 0.58$ is the Flory exponent in three dimensions. This strongly suggests that there exists a close connection between the scaling parameters and the Flory exponent ν .

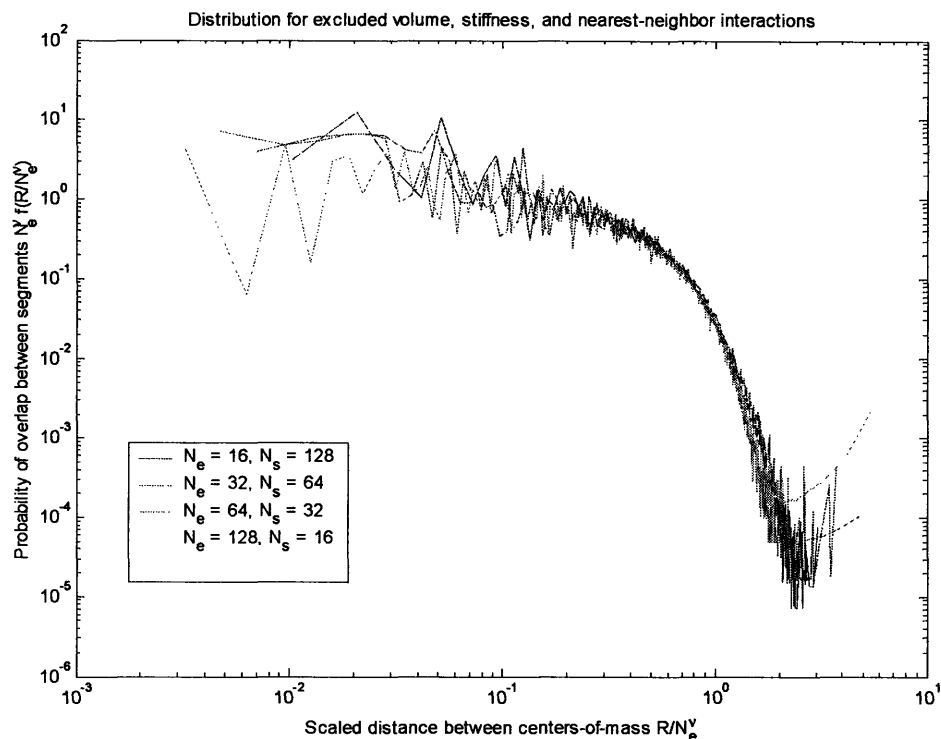


Figure 7-2: Scaled probability of overlap for self-avoiding lattice chains of length $N = 2048$ with incorporated excluded-volume and stiffness interactions.

- For freely-jointed chains of varying lengths, but with a constant number N_b of coarse-grained segments, we computed the probability distribution. The results, plotted in Figure 7-3, scales the distance by $N_e^{-0.58}$ and the probability distribution by $N_e^{0.6}$. The results are comparable for other nearby choices of the exponent, which again suggests that the “best” choice exponents are close to, if not exactly equal to, the Flory exponent ν of the particular model selected.
- For freely-jointed chains of length $N = 512$, we investigated the effect of bead size (or test volume radius r_{test}) on the overlap probability. The unscaled results are shown in Figure 7-4. The additional noise for the smallest bead sizes are principally the result of large bin

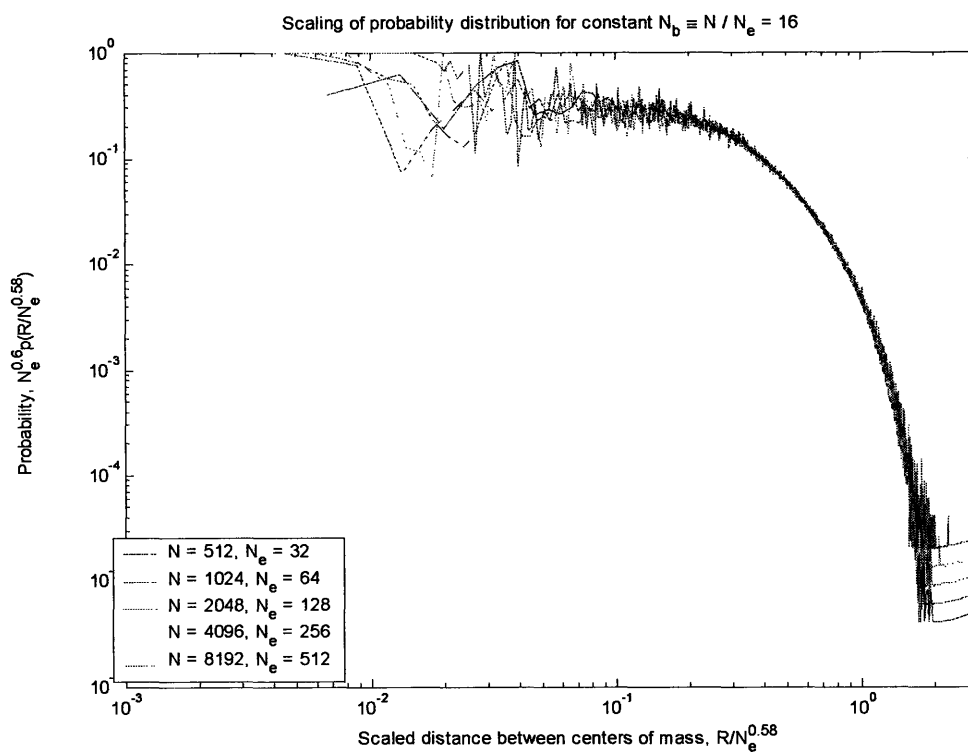


Figure 7-3: Scaled probability distribution for freely-jointed chains of varying lengths with equivalent numbers N_s of coarse-grained segments.

sizes for the distance in the probability distribution relative to the bead size. [The bins were chosen with $\Delta r = 0.02$.]

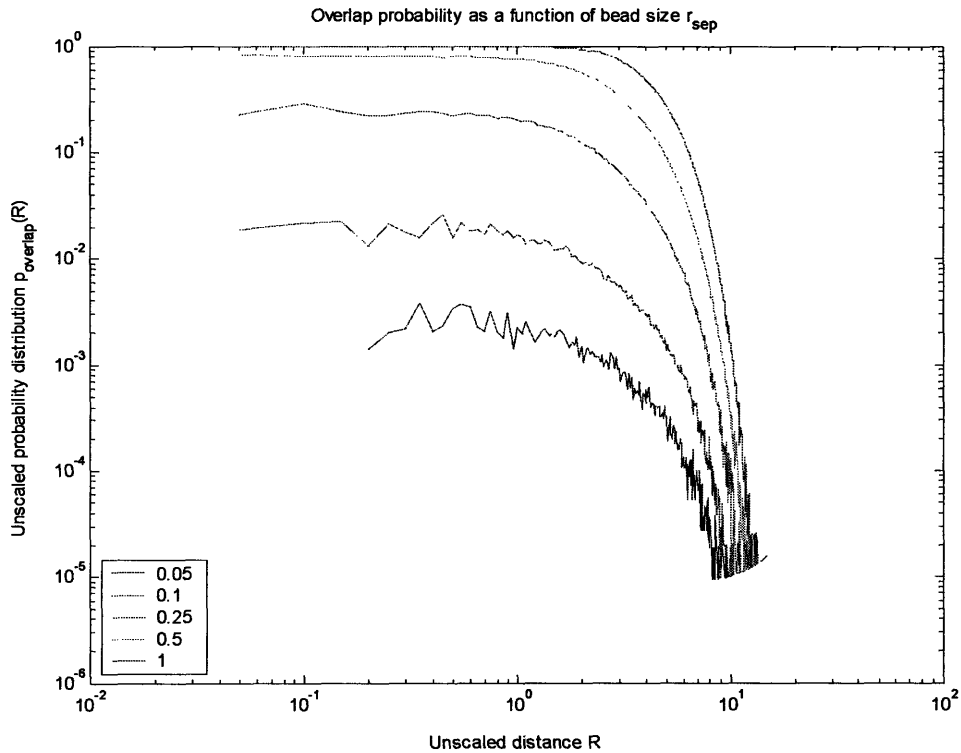


Figure 7-4: Overlap probability as a function of the test radius r_{test} for freely-jointed chains of length $N = 512$.

- For a single freely jointed chain, we determined analytically the probability $F_N(\mathbf{x}, r_{test})$ that there is a bead within a test sphere of radius r_{test} centered at \mathbf{x} , given the center of mass of the chain at the origin. The parameters here are the number of beads N in the chain, the distance x , the number of particles N , and the test radius r_{test} , which directly corresponds to the excluded volume and the “test volume.” The unscaled results, illustrating the general form of the scaling behavior, is shown as Figure 7-5. The corresponding scaled results are shown in Figure 7-6, which rescales the x -axis by $N^{-1/2}$, and the F_N -axis by $N^{3/2}r^{-3}$. The excellent agreement in the scaled results suggests that these parameters are important in the “universal function” underlying the scaling results; we show this more concretely below.

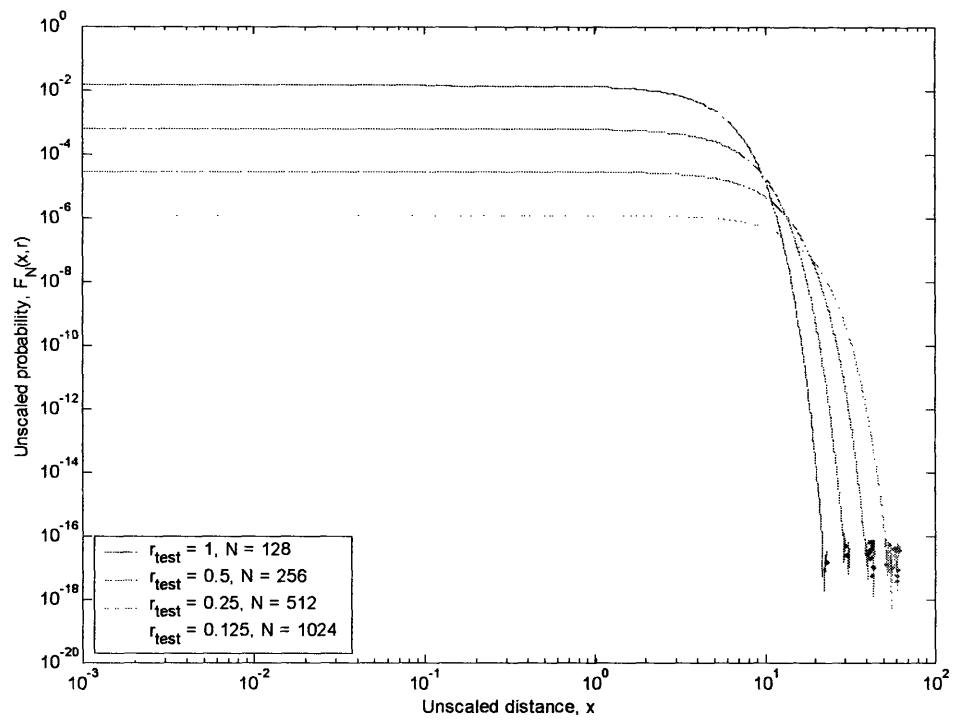


Figure 7-5: Unscaled probability distribution $F_N(x, r)$ for finding a particle within a test volume of radius r a distance x away from the center of mass of a chain of length N .

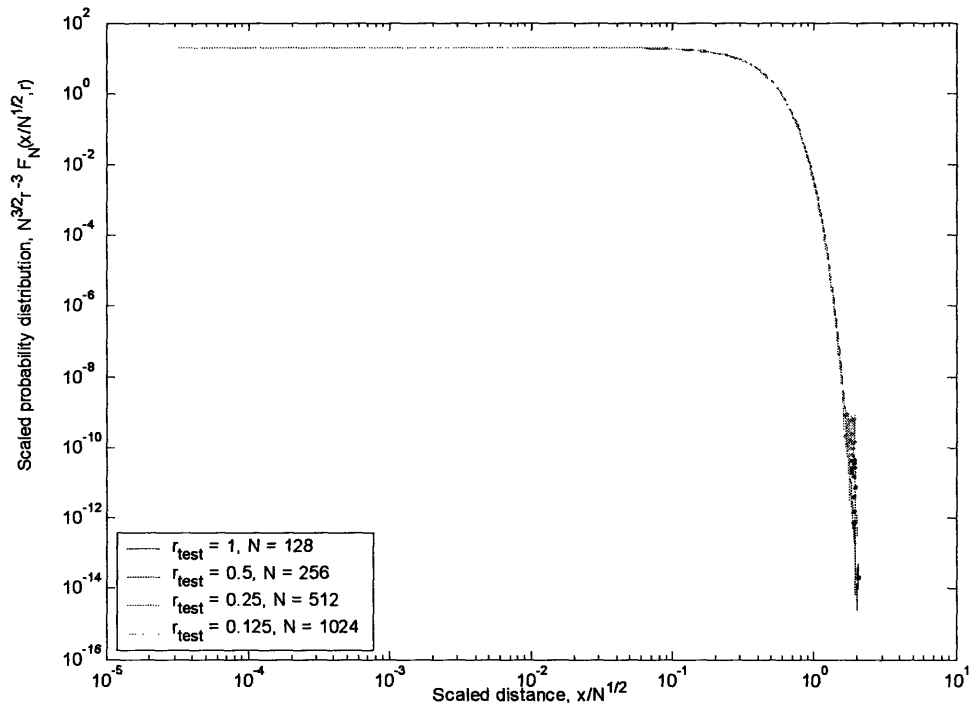


Figure 7-6: A scaled version of Figure 4-8, illustrating the $xN^{-1/2}$ scaling of the distance away from the center of mass and $r^3N^{-3/2}$ scaling of the probability.

7.4 Wavelet transform of the hard-sphere potential

In addition to probability distributions for the coarse-grained internal variables, we need potential functions describing their interaction. When we first developed the WAMC framework for magnetic systems on lattices, we were able to write the Hamiltonian of the system in matrix form. Because the lattice sites were fixed at the beginning of the simulation, the interactions were effectively position-independent. For a self-avoiding random walk, the positions of the individual atoms changes throughout the simulation, which prevents us from creating a matrix structure for the Hamiltonian.

Nonetheless, we can apply the wavelet transform to the hard-sphere potential Eq. (6.1). Because the hard-sphere potential is restricted to be either zero or infinity, we evaluate the arguments of the potential. Expanding $\mathbf{r}_i^{(0)}$ and $\mathbf{r}_j^{(0)}$ in terms of the scaling coefficients $\mathbf{r}^{(k)}$'s and the wavelet coefficients $\mathbf{w}^{(j)}$'s, $j = 1, \dots, k$ using the recursion formulas

$$\mathbf{r}_n^{(k)} = \frac{1}{2} \left[\mathbf{r}_{2n-1}^{(k-1)} + \mathbf{r}_{2n}^{(k-1)} \right], \quad (7.6)$$

$$\mathbf{w}_n^{(k)} = \frac{1}{2} \left[\mathbf{r}_{2n-1}^{(k-1)} - \mathbf{r}_{2n}^{(k-1)} \right], \quad (7.7)$$

leads to a coarse-grained potential which can be written in the form

$$u \left(\mathbf{r}_i^{(0)}, \mathbf{r}_j^{(0)} \right) = \begin{cases} \infty, & \left| \mathbf{r}_i^{(k)} - \mathbf{r}_j^{(k)} \right| < a + f \left(\mathbf{w}^{(j)} \right) \\ 0, & \left| \mathbf{r}_i^{(k)} - \mathbf{r}_j^{(k)} \right| \geq a - f \left(\mathbf{w}^{(j)} \right) \end{cases}, \quad (7.8)$$

where

$$f \left(\mathbf{w}^{(j)} \right) = \left| \sum_{k=1}^M \left(\chi_{m(i,k)} \mathbf{w}_{m(i,k)}^{(k)} - \chi_{m(j,k)} \mathbf{w}_{m(j,k)}^{(k)} \right) \right| \quad (7.9)$$

is a linear function of the wavelet coefficients corresponding to both sites $\mathbf{r}_i^{(0)}$ and $\mathbf{r}_j^{(0)}$. The index function $m(a, k)$ and the sign $\chi_{m(a,k)}$ are given by

$$m(a, k) = \left\lfloor a/2^k \right\rfloor, \quad (7.10)$$

$$\chi_{m(a,k)} = 2 \bmod (m(a, k) + 1, 2) - 1. \quad (7.11)$$

Rigorous use of (7.8) requires knowing all of the differencing terms $\mathbf{w}^{(j)}$ for the entire chain, and evaluating the arguments for all pairs of beads at the original scale. This limits its utility as a practical criterion for simulations. However, (7.8) highlights two competing trends that arise during coarse-graining. Because the magnitude of $f(\mathbf{w}^{(j)})$ can be larger than the lattice spacing a for a given configuration, it is possible that $a - f(\mathbf{w}^{(j)}) < 0$, which means that the zero-energy region can extend leftward to $r = 0$. However, according to (7.8), the hard-sphere exclusion region can extend past the lattice spacing a . Our potential must allow for both possibilities to be true. This suggests that the correct form of our potential will be a potential which remains finite-valued even as $r \rightarrow 0$. With such a potential, we pay an energy penalty for allowing particles to move too close together, without preventing such a possibility altogether. Thus, while unsuitable as a potential, (7.8) justifies the use of “soft” potentials, such as those used by Eurich and Maass [48] and Louis *et al.* [21, 23, 89, 98, 99].

7.5 Constructing coarse-grained potentials

To determine the correct coarse-grained potential, we use a variant of the method described by Bolhuis *et al.* [23]. We simulate two isolated chains, each of identical length N , using the pivot algorithm. After every N attempted pivots, we check for overlaps between the two chains. The respective centers-of-mass of each chain are separated by a random vector whose length increases from $r = 0$ to $r = 2N$ in increments of Δr , and whose orientation for each value of r is uniformly drawn from the unit sphere. At each r , we check to see if there are any excluded-volume interactions between the two chains: that is, we check if

$$r_{ab} = |\mathbf{r}_{i,a} - \mathbf{r}_{j,b}| < r_{test}, \quad (7.12)$$

where the subscripts a and b in (7.12) now denote the first and second chains, respectively, and r_{test} is the radius of the excluded-volume sphere. If (7.12) is satisfied for any i and j , then we increase the histogram $w(r)$ by one. The probability of overlap is then defined as $p(r) = w(r)/n(r)$, where $w(r)$ is the number of overlaps observed when the centers of mass are separated by distance r , and $n(r)$ is the total number of times the centers of mass were separated by distance r .

Repeating this process gives a probability distribution $p(r)$ for the overlap between any two component beads as a function of the distance r between coarse-grained beads. We then estimate the coarse-grained potential $u(r)$ by inversion of the Boltzmann factor:

$$u(r) = -k_B T \ln(1 - p(r)), \quad (7.13)$$

where the factor of $k_B T$ is included so that the relative effect of the potential is independent of the temperature. A sample potential produced by this method is shown in Figure 8-2, for coarse-grained beads representing 32 atomistic beads. We note that the potential is quite short-ranged: its magnitude approaches zero for distances larger than roughly one mean radius of gyration ($R_g \approx 7.1$ for the 32-mer), and remains finite-valued even as the distance between the centers-of-mass of the two segments goes to zero.

The basic form of the potential shown in Figure 8-2 was observed to remain unchanged regardless of whether freely-jointed or self-avoiding chain conformations were used to test for overlap, or for different number of particles in the chain. This suggests that we can derive scaling laws which would allow us to re-use potentials between different simulations without necessitating recalculations.

7.6 Analytical solution: Freely-jointed chains

In addition to a numerical procedure for obtaining coarse-grained potentials, we consider here analytical expressions for the coarse-grained potential which guide us identifying any scaling behavior which might be present. For freely-jointed chains, we can approximate the probability of overlap by considering the problem of a freely-jointed chain overlapping with a given “test volume” at a given radial distance away from the chain. This probability then leads to the distribution for overlap of an arbitrary chain with a specific polymer configuration. Using (7.13) would then allow us to determine the appropriate coarse-grained potential.

We now attempt to derive a rigorous result for the probability of overlap by considering the problem of a freely-jointed chain overlapping a given “test volume” at a specific location in space. We can then use this probability to compute the distribution for overlap of an arbitrary chain with a specific polymer configuration. For a Gaussian chain, Yamakawa [161] derived the

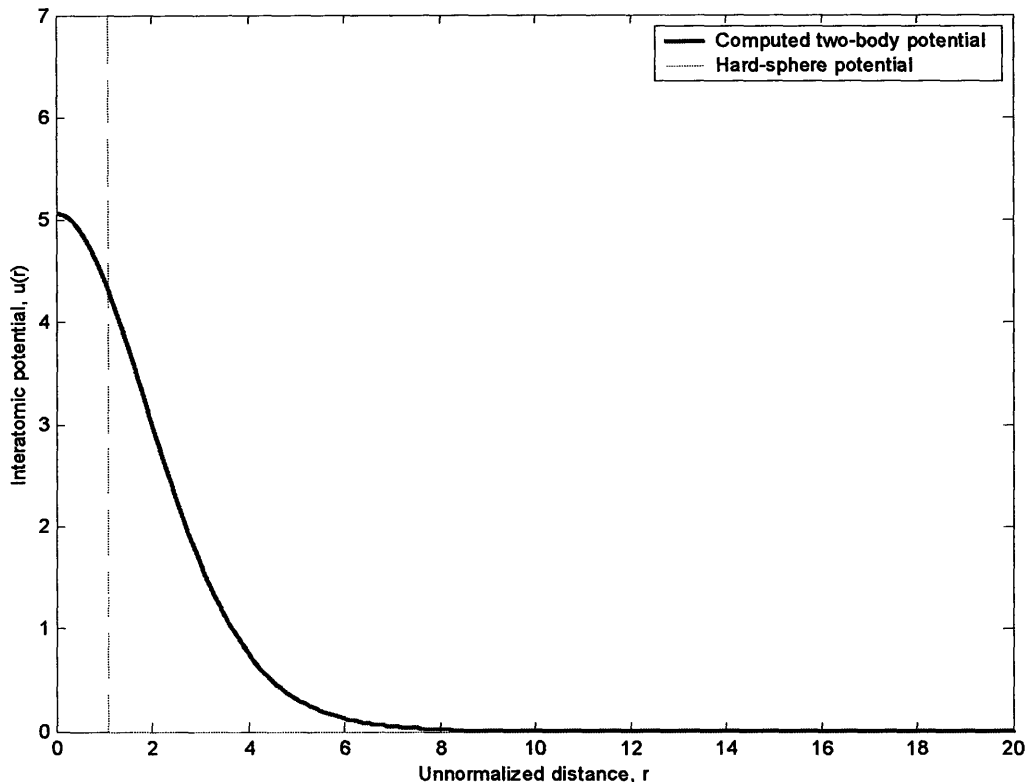


Figure 7-7: Estimate for the true interatomic potential $u(r)$ used to improve the WAMC self-avoiding walk model. The potential is shown for coarse-grained beads representing 32 atomistic beads of diameter 1. $R_g \approx 7.1$ for the 32-mer.

probability distribution $f_N(0, \mathbf{x})$ that any bead of a chain is at point \mathbf{x} , given that the center of mass is located at the origin:

$$f_N(0, \mathbf{x}) = \frac{1}{N} \sum_{j=1}^N \left(\frac{3}{2\pi \langle s_j^2 \rangle} \right)^{3/2} \exp\left(-\frac{3x^2}{2\sigma_j}\right), \quad (7.14)$$

where σ_j is defined as

$$\sigma_j = \langle s_j^2 \rangle = \frac{a^2}{3N} (N^2 - 3j(N - j)). \quad (7.15)$$

In (7.15), j is a free index, and a is the step size of the chain.

We would like to determine the probability that there is a particle within radius r of \mathbf{x} , given that the center of mass is at the origin. However, we can recast the problem as determining the

probability that there is a particle within radius r of the origin, given that the center of mass is located at \mathbf{x} . The quantity we need to determine is thus

$$F_N(\mathbf{x}, r) = \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta \int_0^r s^2 ds \frac{1}{N} \sum_{j=1}^N \left(\frac{3}{2\pi\sigma_j} \right)^{3/2} \exp\left(-\frac{3(\mathbf{x}-\mathbf{s}) \cdot (\mathbf{x}-\mathbf{s})}{2\sigma_j} \right). \quad (7.16)$$

The first integral in (7.16) is trivial; expanding the dot product inside the integrand and re-ordering the operations gives

$$F_N(\mathbf{x}, r) = \frac{2\pi}{N} \sum_{j=1}^N \left(\frac{3}{2\pi\sigma_j} \right)^{3/2} \int_0^r s^2 ds \int_0^\pi \sin\theta d\theta \exp\left(-\frac{3(x^2 + s^2 - 2xs \cos\theta)}{2\sigma_j} \right). \quad (7.17)$$

Taking the integral over θ in (7.17) first is simpler. Letting $u = \cos\theta$, so that $du = -\sin\theta d\theta$, (7.17) becomes

$$F_N(\mathbf{x}, r) = \frac{2\pi}{N} \sum_{j=1}^N \left(\frac{3}{2\pi\sigma_j} \right)^{3/2} \int_0^r s^2 ds \exp\left(-\frac{3(x^2 + s^2)}{2\sigma_j} \right) \int_{-1}^1 du \exp\left(\frac{3xsu}{\sigma_j} \right). \quad (7.18)$$

The inner integration is now straightforward, and gives

$$\begin{aligned} \int_{-1}^1 du \exp\left(\frac{3xsu}{\sigma_j} \right) &= \frac{\sigma_j}{3xs} \left(\exp\left(\frac{3xs}{\sigma_j} \right) - \exp\left(-\frac{3xs}{\sigma_j} \right) \right) \\ &= \frac{2\sigma_j}{3xs} \sinh\left(\frac{3xs}{\sigma_j} \right); \end{aligned} \quad (7.19)$$

inserting (7.19) into (7.18) yields

$$F_N(\mathbf{x}, r) = \frac{3}{N} \sqrt{\frac{3}{2\pi}} \sum_{j=1}^N \frac{1}{\sigma_j^{3/2}} \int_0^r s^2 ds \exp\left(-\frac{3(x^2 + s^2)}{2\sigma_j} \right) \left(\frac{2\sigma_j}{3xs} \sinh\left(\frac{3xs}{\sigma_j} \right) \right),$$

which can be rearranged as

$$F_N(\mathbf{x}, r) = \frac{1}{Nx} \sqrt{\frac{6}{\pi}} \sum_{j=1}^N \frac{e^{-\frac{3x^2}{2\sigma_j}}}{\sigma_j^{1/2}} \int_0^r s ds \exp\left(-\frac{3s^2}{2\sigma_j} \right) \sinh\left(\frac{3xs}{\sigma_j} \right). \quad (7.20)$$

Evaluating the integral over s in (7.20) is easier piecewise. The required integrals are

$$\int_0^c s ds \exp\left(\pm B_j x s - \frac{B_j s^2}{2}\right) = \frac{1}{B_j} \left(1 - e^{\pm B_j x c - \frac{1}{2} B_j c^2}\right) + x \sqrt{\frac{\pi}{2B_j}} e^{\frac{B_j x^2}{2}} \left(\pm \operatorname{erf}\left((c \mp x) \sqrt{\frac{B_j}{2}}\right) + \operatorname{erf}\left(\sqrt{\frac{B_j}{2}}\right)\right), \quad (7.21)$$

which makes the combined integral

$$\begin{aligned} \int_0^r s ds \exp\left(-\frac{3s^2}{2\langle s_j^2 \rangle}\right) \sinh\left(\frac{3xs}{\langle s_j^2 \rangle}\right) \\ = \frac{x}{2} \sqrt{\frac{\pi}{2B_j}} e^{\frac{B_j x^2}{2}} \left(\operatorname{erf}\left((c+x) \sqrt{\frac{B_j}{2}}\right) + \operatorname{erf}\left((c-x) \sqrt{\frac{B_j}{2}}\right)\right) \\ - \frac{1}{B_j} e^{-\frac{1}{2} B_j c^2} \sinh(B_j x c) \end{aligned}$$

Combining (7.20) and (7.21) gives

$$\begin{aligned} F_N(\mathbf{x}, r) = \frac{1}{2N} \sum_{j=1}^N \left(\operatorname{erf}\left((r+x) \sqrt{\frac{B_j}{2}}\right) + \operatorname{erf}\left((r-x) \sqrt{\frac{B_j}{2}}\right)\right) \\ - \frac{1}{Nx} \sum_{j=1}^N \sqrt{\frac{2}{\pi B_j}} e^{-\frac{1}{2} B_j (x^2 + r^2)} \sinh B_j x r, \quad (7.22) \end{aligned}$$

where

$$B_j = \frac{3}{\sigma_j} = \frac{9N}{a^2 (N^2 - 3j(N-j))}. \quad (7.23)$$

7.6.1 Invalidity of “maximum-term” approach

We previously performed the above analysis using a “maximum-term” approach to handle the $\langle s_j^2 \rangle$ term in (7.14) more efficiently. That is, we assumed that we could neglect differences in the value of $\langle s_j^2 \rangle$, so that (7.15) could be approximated by its mean value

$$\langle s^2 \rangle = \frac{1}{N} \sum_{j=1}^N \langle s_j^2 \rangle = \frac{1}{N} \sum_{j=1}^N \frac{a^2}{3N} (N^2 - 3j(N-j)) \approx \frac{Na^2}{6}. \quad (7.24)$$

Based on the approximation (7.24), we obtained the solution

$$F_N(\mathbf{x}, r) = \frac{1}{2} \left(\operatorname{erf} \left(\frac{3(x+r)}{a\sqrt{N}} \right) - \operatorname{erf} \left(\frac{3(x-r)}{a\sqrt{N}} \right) \right) - \frac{a}{3x} \sqrt{\frac{N}{\pi}} \exp \left(-\frac{9(r^2+x^2)}{a^2N} \right) \sinh \left(\frac{18xr}{a^2N} \right). \quad (7.25)$$

Unfortunately, the approximation (7.24) yields extremely inaccurate results, as the parameter $\langle s_j^2 \rangle$ appears in the denominator of the exponential function in (7.14), which suggests that the “maximum-term”-like approximation will be difficult to justify. To see this, consider Figure 7-8, which shows the difference between the two results (7.22) and (7.25). The consequence of the discrepancy is that there is a greater probability of particles being discovered close to the center of the molecule than would be expected from the maximum term approximation. In addition, this means that the probability of a collision between two different chain segments, which we can estimate as

$$P(\mathbf{u}, r_{test}) \approx 1 - (1 - F_N(\mathbf{u}, r_{test}))^N, \quad (7.26)$$

will be much higher than the approximation would suggest. For example, when $u = 0.01$, the more accurate equation (7.22) gives $F_{64}^t(0.01) = 0.0495$ while the approximation (7.25) yields $F_{64}^a(0.01) = 0.0365$. For these values of F_{64} , we use (7.26) to estimate the respective probabilities to be

$$P^t(0.01, 1.0) \approx 1 - (1 - 0.0539)^{64} = 0.961$$

$$P^a(0.01, 1.0) \approx 1 - (1 - 0.0365)^{64} = 0.907$$

The discrepancies noted in the values of P^t and P^a near the origin apply throughout the domain, as can be seen from Figure 7-9. Moreover, the fact that $P^t > P^a$ at all values of r indicates that using P^a to determine an interaction potential will yield a much softer potential than the more accurate formula P^t .

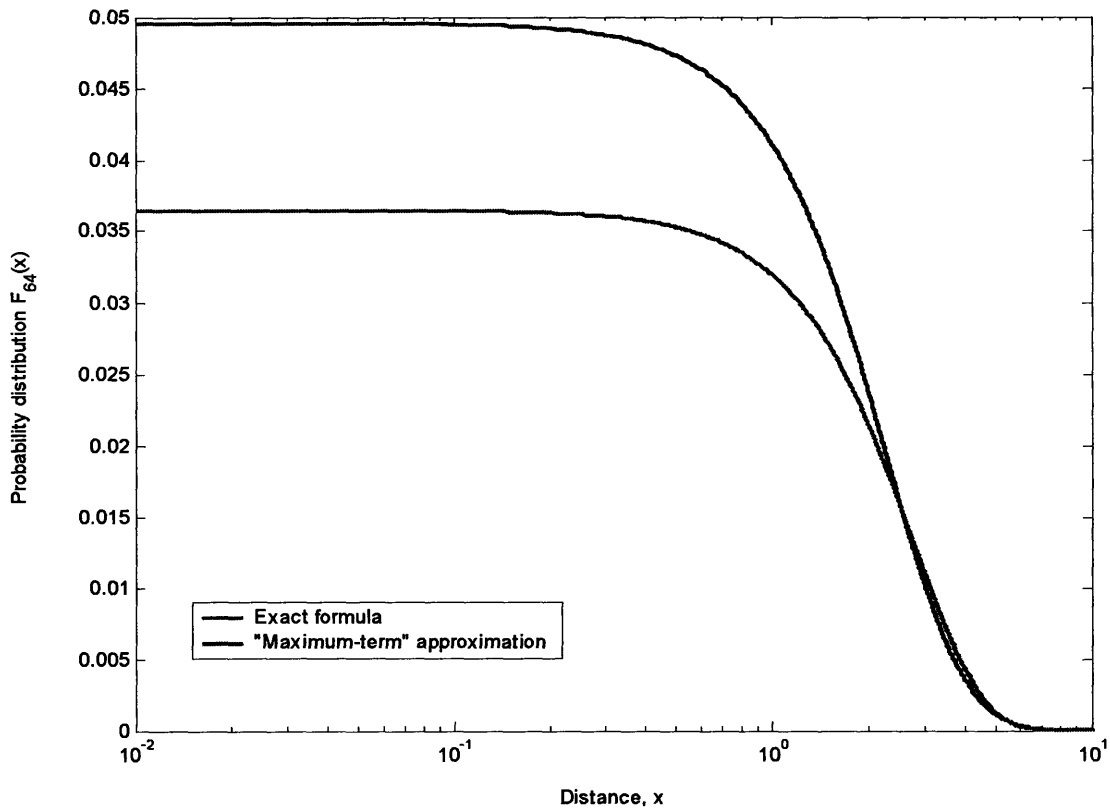


Figure 7-8: The probability distribution function $F_{64}(x, 1)$ as given by the true formula (7.22) [upper curve] and the approximation (7.25) [lower curve].

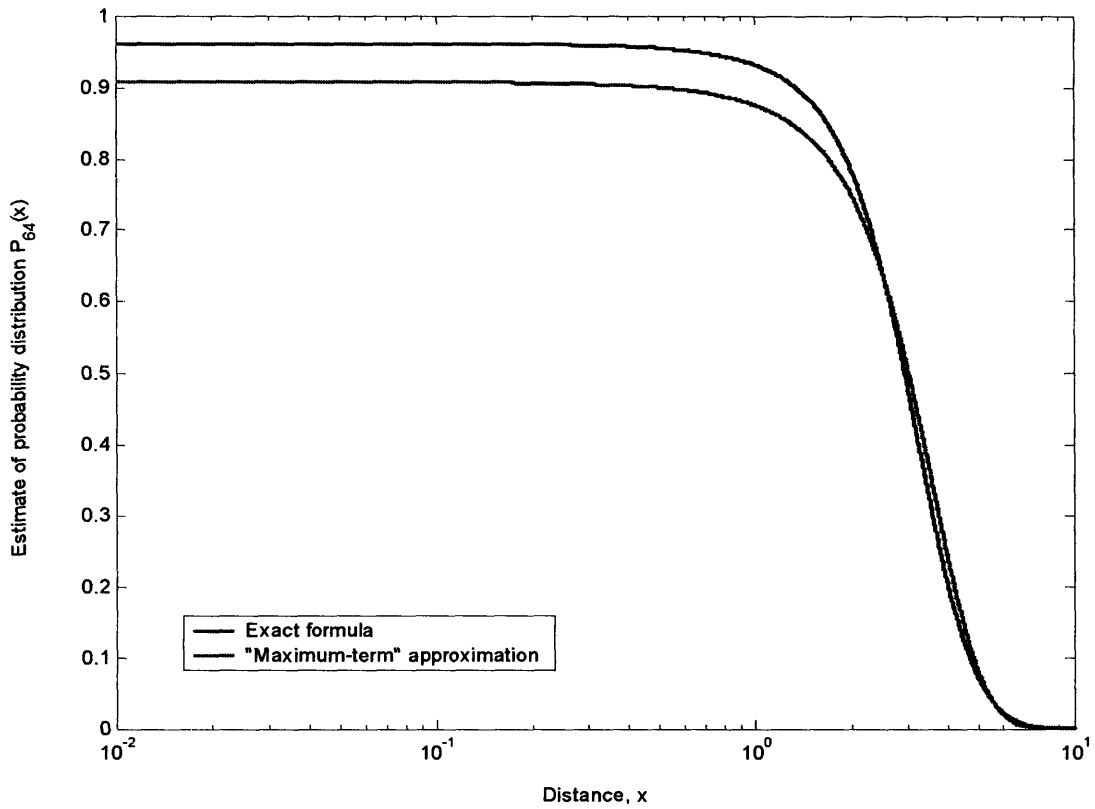


Figure 7-9: Estimate $P_{64}(x, 1)$ for the probability distribution function for two-chain collisions, as given by the true formula (7.22) [upper curve] and the approximation (7.25) [lower curve].

7.6.2 Asymptotic behavior of Equation (7.22)

To assist in the determination of scaling laws for the overlap between segments, we determine the limiting behavior of Eq. (7.22) as a function of x and the parameters r and N . For small values of x , expanding Eq. (7.22) as a power series in x gives to fourth order in x the formula

$$F_N(x \approx 0, r) \approx \frac{1}{N} \sum_{j=1}^N \left[\operatorname{erf} \left(r \sqrt{\frac{B_j}{2}} \right) - \sqrt{\frac{2B_j}{\pi}} \left(1 + \frac{1}{6} B_j^2 r^2 x^2 \right) r e^{-\frac{1}{2} B_j r^2} + O(x^4) \right]. \quad (7.27)$$

Equation (7.27) indicates that the probability distribution decays very slowly for test volumes located close to the center of mass of the segment. This result is easily observed in both the theoretical calculations and the simulation results. When x is large,

$$F_N(x \rightarrow \infty, r) = -\frac{1}{Nx} \sum_{j=1}^N \sqrt{\frac{2}{\pi B_j}} e^{-\frac{1}{2} B_j x^2}, \quad (7.28)$$

which suggests that we have Gaussian-like decay for $x \rightarrow \infty$. Similarly, if r is a small parameter, then the corresponding asymptotic expansion is

$$F_N(x, r \approx 0) \approx \frac{r^3 \sqrt{2}}{3N} \sum_{j=1}^N B_j^{3/2} \exp \left(-\frac{B_j x^2}{2} \right) + O(r^5). \quad (7.29)$$

Equation (7.29) estimates the probability distribution in the limit of vanishing test volumes, which corresponds to small bead sizes. Again, we observe from Figure 7-11 that there does exist an r^3 scaling in the overlap distribution, indicating (7.29) may be useful in determining the chain overlap probability.

The remaining case is the polymer scaling limit, $N \rightarrow \infty$. In this limit, the parameter B_j defined by Eq. (7.23) tends to zero as N^{-1} : thus, expanding Eq. (7.22) as a function of N leads to

$$F_{N \rightarrow \infty}(x, r) \approx \frac{r^3}{3N} \sqrt{\frac{2}{\pi}} \sum_{j=1}^N B_j^{3/2} - \frac{r^3}{30} \sqrt{\frac{2}{\pi}} \sum_{j=1}^N (3r^2 + 5x^2) B_j^{5/2} + O(B_j^{7/2}). \quad (7.30)$$

To leading order, F_N therefore scales as $N^{1/2}$. Comparing (7.29) to (7.30), we see that there are similar scaling results for both $N \rightarrow \infty$ and $r \rightarrow 0$, with identical results obtained when

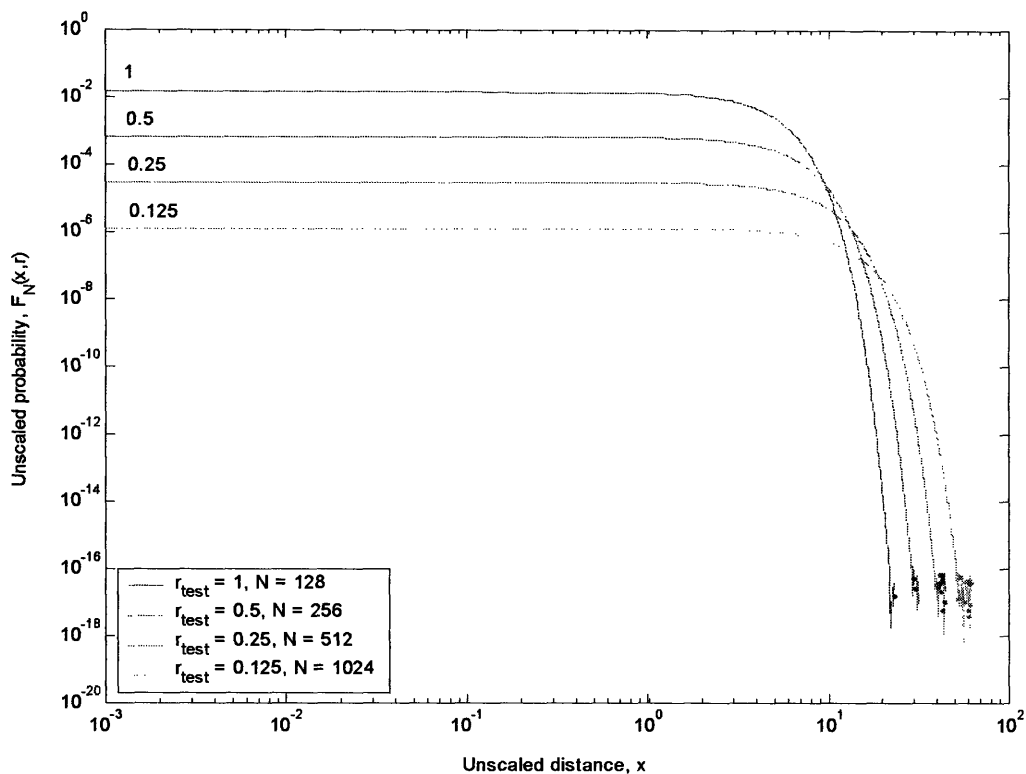


Figure 7-10: Unscaled probability distribution $F_N(x, r)$ for finding a particle within a test volume of radius r a distance x away from the center of mass of a chain of length N . From top to bottom, $r_{test} = 1, 0.5, 0.25,$ and 0.125 .

$r = 0$. Two features are common to both scaling laws: the distance scales as $xN^{-1/2}$, and the probability scales as $r^3N^{1/2}$, since $B_j \sim N^{-1}$.

Taking advantage of these insights provided by (7.22), we can collapse the data for different values of the parameters r and N if we rescale the x -axis by $N^{-1/2}$ and the y -axis by $r^{-3}N^{3/2}$. The unscaled data are shown in Figure 7-10, while the scaled plot is shown as Figure 7-11.

7.7 The probability of overlap for two segments of a freely-jointed chain

As mentioned above, given two independent freely-jointed chains, the probability that they will have an overlap, defined as any two component beads on different chains being separated by a

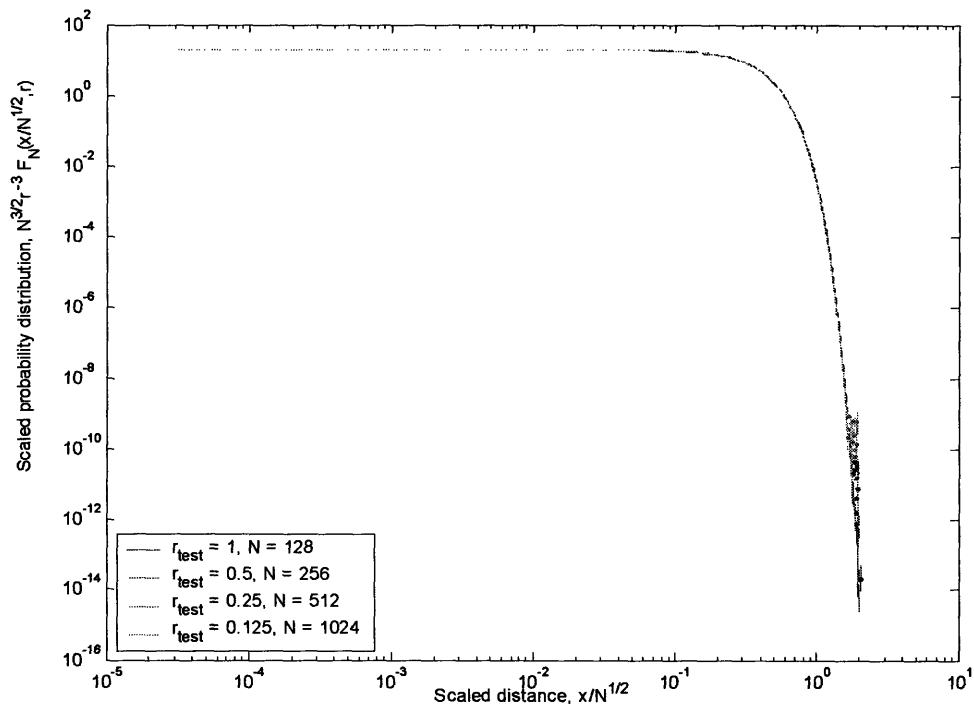


Figure 7-11: A scaled version of Figure 7-10, illustrating the $xN^{-1/2}$ scaling of the distance away from the center of mass and $r^3N^{-3/2}$ scaling of the probability.

distance less than or equal to r_{test} , is one minus the probability that they do not overlap. The exact probability should be given by

$$P_N(\mathbf{u}, r_{test}) \approx 1 - \prod_{i=1}^N (1 - F_N(\mathbf{u} + \Delta\mathbf{r}_i, r_{test})), \quad (7.31)$$

where $\Delta\mathbf{r}_i$ is the displacement of bead i in the test chain from the center of mass of the test chain. However, using Eq. (7.31) would require knowledge of the detailed atomistic chain, which is unavailable in the coarse-grained simulation. A rough estimate of the probability for a test volume of radius r_{test} which uses only information available at the coarse-grained scale is

$$P_N(\mathbf{u}, r_{test}) \approx 1 - (1 - F_N(\mathbf{u}, r_{test}))^N. \quad (7.32)$$

This is the probability we used above to determine the two-chain overlap probability. We note that this relatively simple formula for the overlap probability yields a reasonable estimate. Figure 7-12 shows the overlap probability for two chains of 32 beads each as a function of the separation of their centers of mass, x . The curves are qualitatively the same shape and the agreement between the simulation and the approximation (7.32) improves as $r_{test} \rightarrow 0$ and $x \rightarrow \infty$. Thus, although it is clear that (7.32) is not an exact formula for $P_N(\mathbf{u}, r_{test})$, we can approximate the behavior of $P_N(\mathbf{u}, r_{test})$ by calculating $F_N(\mathbf{u}, r_{test})$ for a number of different chains of varying size sufficient to develop the corresponding scaling behavior, and then use (7.32) to determine the probability for chains of the desired length N . We can then “invert” $P_N(\mathbf{u}, r_{test})$ via (7.13) to obtain the potential function.

We can also compare the results obtained holding r_{test} constant and varying N ; the results for $N = 128$ with $r_{test} = 1.0$ are shown in Figure 7-13. We note in these figures that the approximation (7.32) breaks down for large N , although we continue to observe the correct qualitative behavior for the overlap probability. This suggests that rather than trying to use an analytical law like (7.27), we are better off computing the underlying potential $F_N(\mathbf{u}, r_{test})$ for a given chain configuration, and then using (7.32) to estimate the potential.

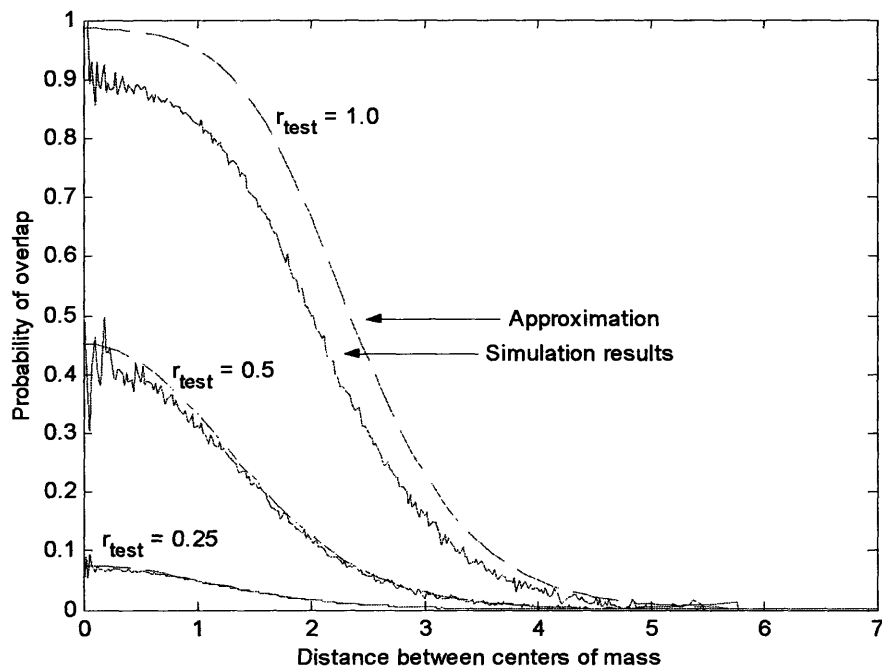


Figure 7-12: Dependence of overlap size on r_{test} : overlap between 32-mers with test-volume sizes $r_{test} = 1.0$, $r_{test} = 0.5$, and $r_{test} = 0.25$ as a function of the separation of centers of mass, as obtained via direct simulation [solid curves] and approximation (7.32), with F_N given by (7.22) [dashed curves].

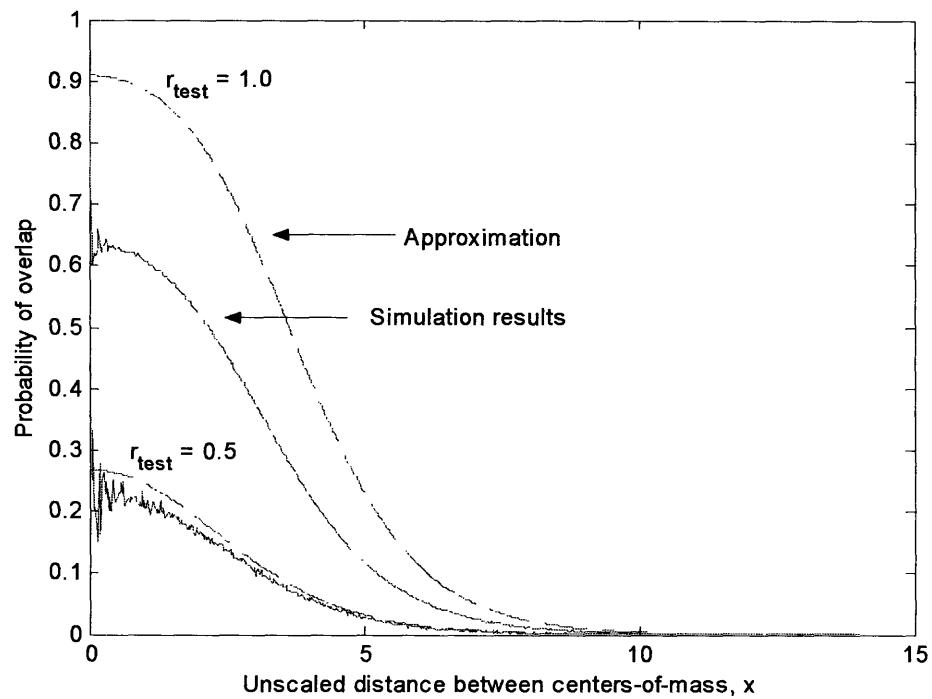


Figure 7-13: Dependence of overlap probability on r_{test} : overlap between 128-mers with test-volume size $r_{test} = 1.0$ and $r_{test} = 0.5$ as a function of the separation of centers of mass, as obtained via direct simulation and approximation (7.32).

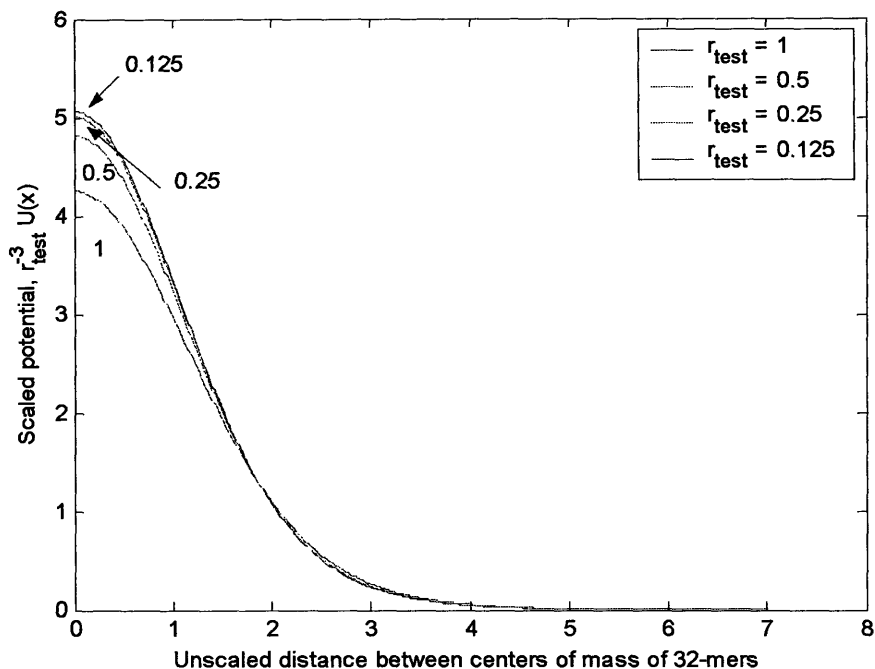


Figure 7-14: Scaled potential showing the r_{test}^3 dependence of the potential for two freely-jointed chains of 32-mers in the limit $r_{test} \rightarrow 0$ using approximation (7.32).

7.8 Potentials derived from $P(\mathbf{u}, r_{test})$

Using the probability distributions estimated by (7.32), we can construct approximate potentials for the polymer. For a freely-jointed chain, we obtain via the method described in Section 7.5 the potentials shown in Figure 7-14 for various values of r_{test} . The r_{test}^3 dependence postulated in (7.29) has been used to scale the results. We observe that the results converge in the limit of $r_{test} \rightarrow 0$ (Figure 7-15). Scaling the potentials for different chain lengths produced using (7.32) converge in the limit $N \rightarrow \infty$ when scaled by the factor $N^{1/2}$ (Figure 7-16). Convergence to a limiting curve as $N \rightarrow \infty$ appears to be much slower in the potentials computed from direct simulation than using the approximation (7.32). We do not get good agreement with the results from direct simulation, as seen in Figure 7-17. Consequently, (7.32) yields qualitatively useful but not quantitatively accurate results.

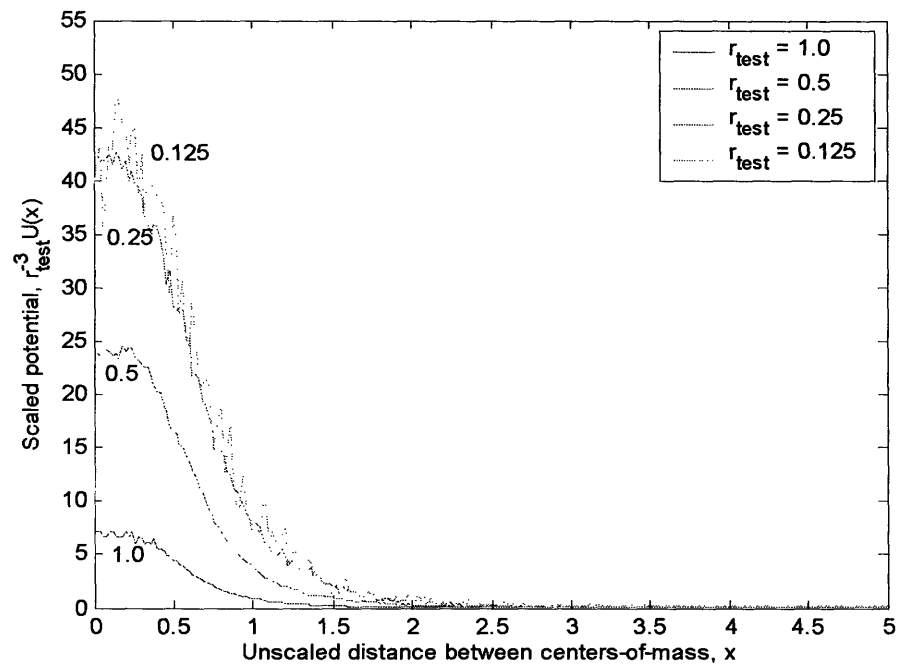


Figure 7-15: Scaled potentials showing the r_{test}^3 dependence of the potential for overlap of two freely-jointed chains of 32-mers determined by direct simulation.

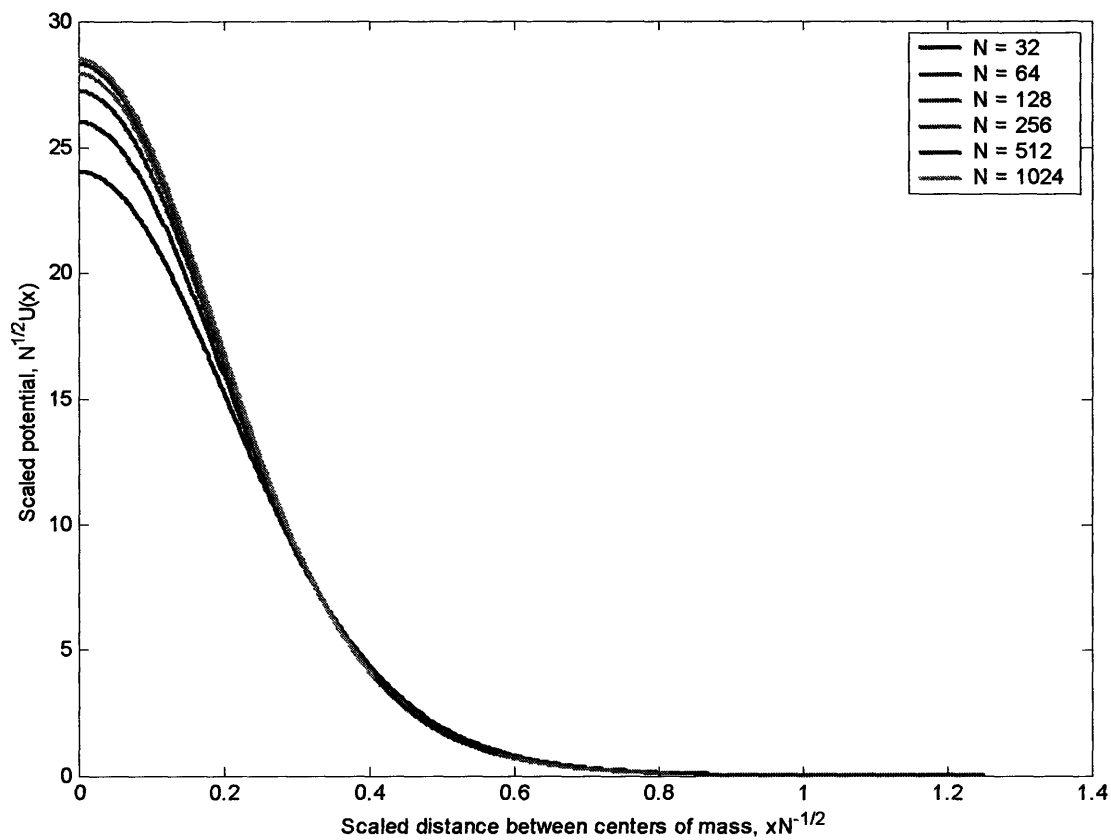


Figure 7-16: Scaling of the two-body potential created by approximation (7.32), parametrized by increasing bond length ($N = 32$ is the bottom curve; $N = 1024$ is the top curve), demonstrating convergence towards a fixed curve as $N \rightarrow \infty$.

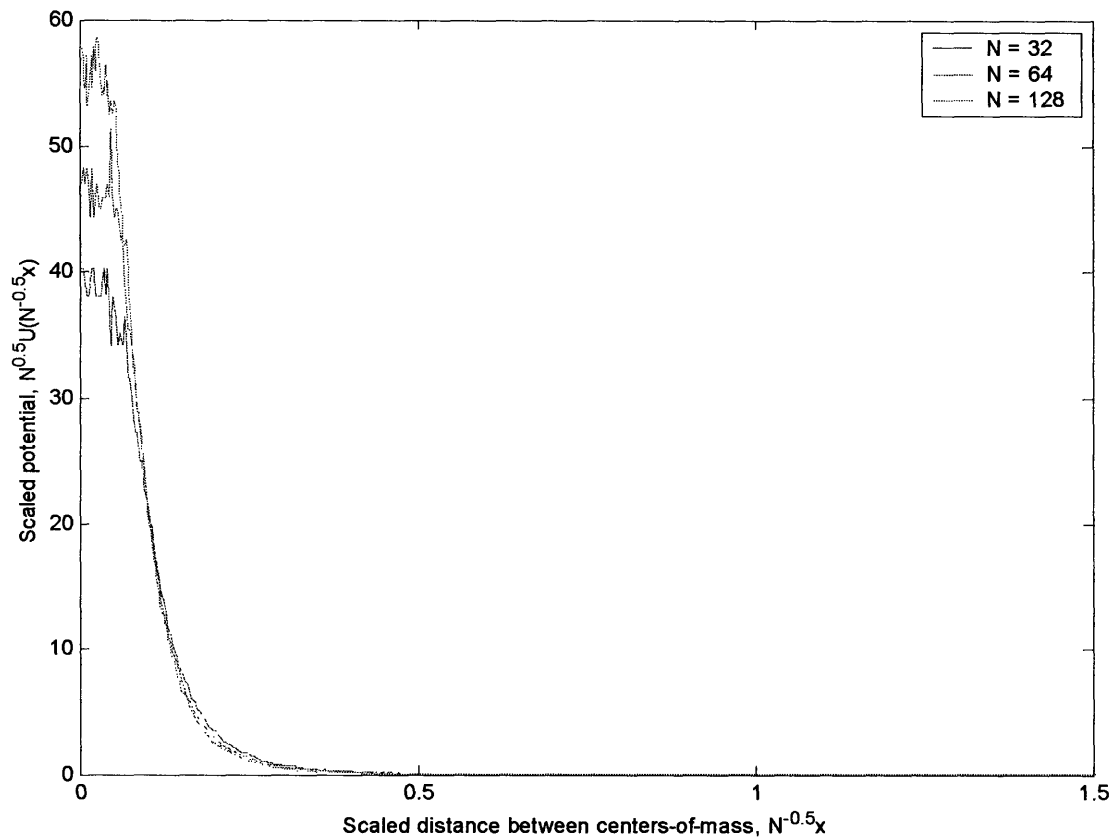


Figure 7-17: Scaling of the two-body potential created via direct simulation, parametrized by increasing bond length ($N = 32, 64,$ and 128 from bottom to top, respectively).

7.9 Self-avoiding walk

We have considered freely-jointed chains thus far for illustrative purposes and because they are analytically tractable. The practical value of coarse-grained intermolecular potentials arises when studying self-avoiding chains. It is possible to develop similar scaling laws for self-avoiding walks, as can be seen in Figures 7-18 through 7-20. These figures illustrate the dependence of the coarse-grained overlap potential on test volume size r_{test} and on the coarse-grained segment size N . Figure 7-18 shows the two-chain overlap potential for various test-volume sizes r . As the test-volume radius decreases, the scaled potential $r_{test}^{-3}U(x)$ converges to a limiting curve which reproduces the result obtained for freely-jointed chains above. This confirms that the radius of the test-volume (or, equivalently, the bead size) of the coarse-grained particle is an important parameter in our coarse-graining simulations.

Determining the scaling function for the chain segment length N , as shown in Figure 7-19, yields a much more complicated behavior than exhibited in the freely-jointed chain. For the self-avoiding walk, the range of the potential scales as $N^{0.4}$, whereas for the freely-jointed chain, the range scales as $N^{0.5}$.

There does not appear to be a universally applicable scaling for the magnitude of the potential for the freely-jointed chain. Instead, the scaling depends on the ratio of the test-volume radius r_{test} to the bond length of the chain. As shown in Figure 7-19, for $r_{test} = 0.5$, the magnitude is unchanged by adjusting the length of the chain; similar results persist for values up to $r_{test} = 1$. However, for smaller values of r_{test} , the magnitude of the potential collapses when scaled by $N^{-0.5}$, as shown in Figure 7-20 for $r_{test} = 0.125$. Thus, for small values of r_{test} , the self-avoiding and freely-jointed chains exhibit similar dependence in the magnitude of the coarse-grained potential. Such behavior is reasonable: when the test-volume size is large, the excluded volume of each chain will be too large for two chains to approach without some overlap. As the test-volume radius decreases relative to the bond length, the separation between the centers-of-mass of two chains can become small without beads on different chains overlapping. Thus we conclude that the ratio of test-volume size to bond length is an important parameter in constructing coarse-grained potentials.

From the results presented above, we can now outline a general approach for determining coarse-grained potentials. Since we have established scaling laws for the range and for the

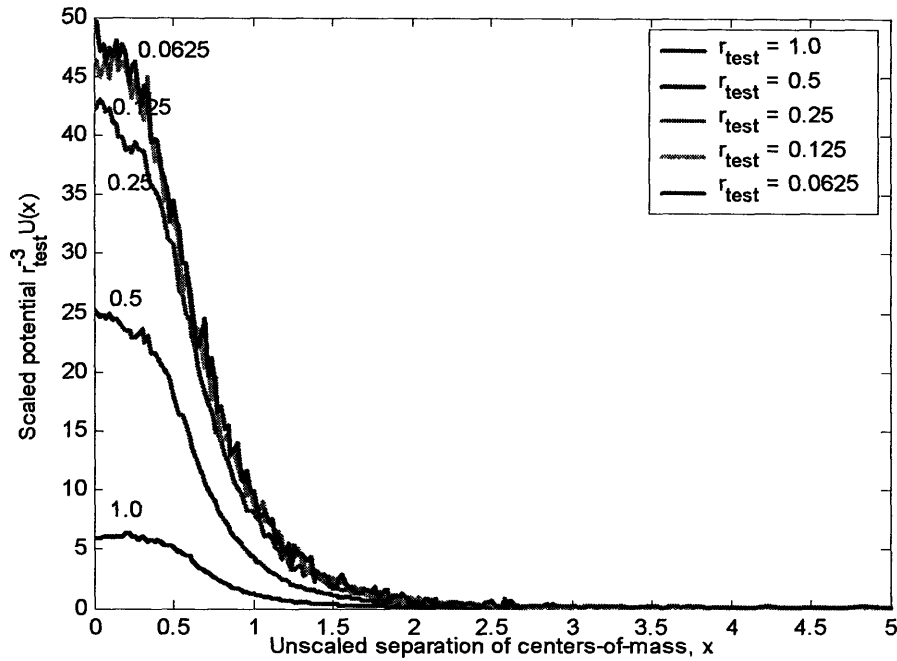


Figure 7-18: The coarse-grained overlap potential for two self-avoiding chain segments of length 32, parametrized as a function of the test-volume size r_{test} , demonstrating the r_{test}^3 -dependence of $U(r)$ as $r_{test} \rightarrow 0$.

magnitude of the potential, so we can compute potentials for various small system sizes and various values of r_{test} , and then use the phenomenological scaling laws outlined above to estimate what the correct potentials for the desired stage of coarse-graining are.

7.10 Alternative approaches for computing new coarse-grained potentials

The problem of determining a coarse-grained potential which accurately reproduces a given set of statistics is an example of “inverse Monte Carlo” methods. Simple methods such as inverting the radial distribution function $g(r)$ to obtain the potential of mean force

$$w(r) = -\ln g(r)$$

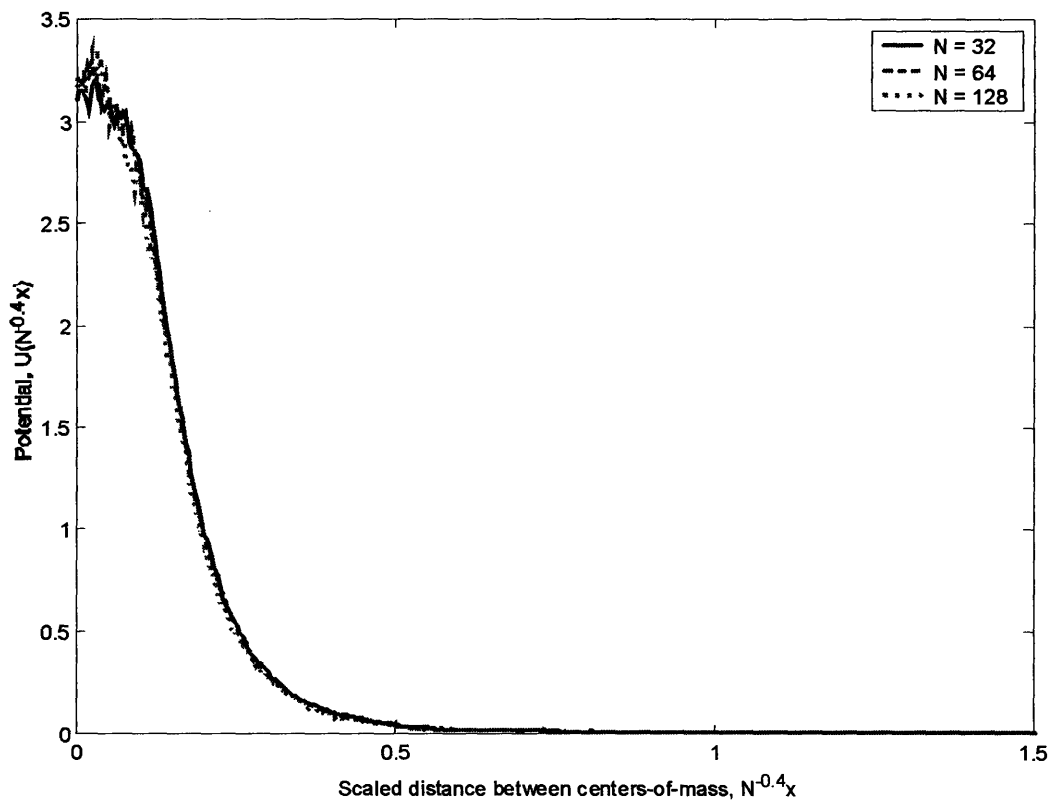


Figure 7-19: Collapse of coarse-grained overlap potentials for two self-avoiding chains with $N = 32, 64,$ or $128,$ and test-volume size $r_{test} = 0.5,$ showing an $N^{0.4} \approx N^{1-\nu}$ dependence.

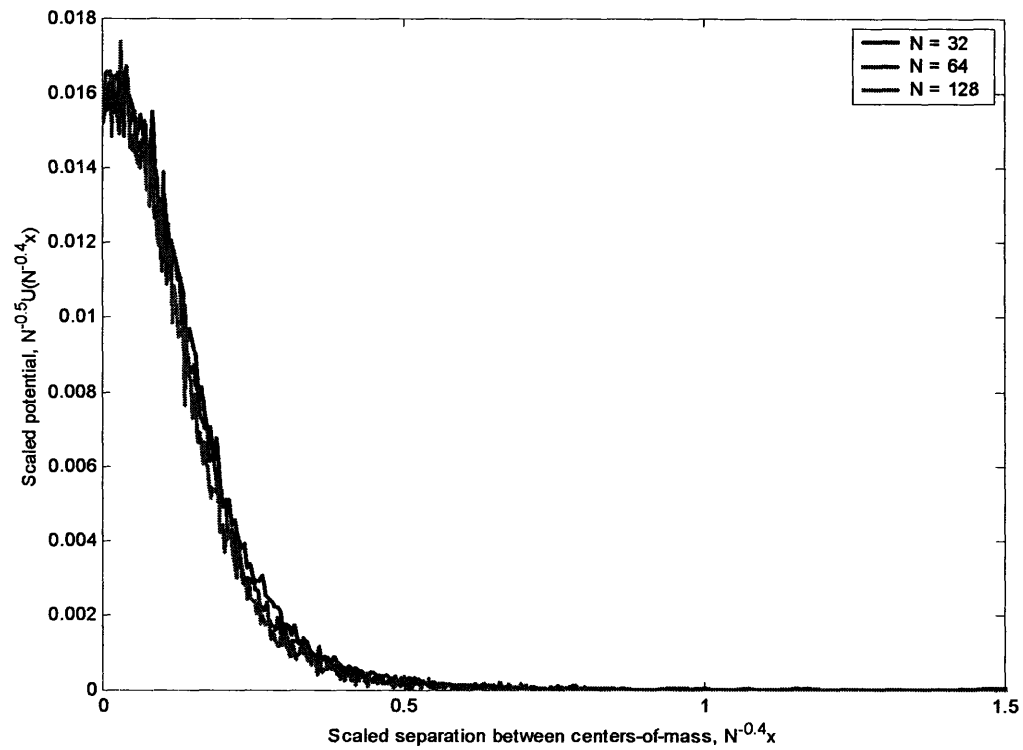


Figure 7-20: Collapse of coarse-grained overlap potential for two self-avoiding chains with $N = 32, 64,$ or $128,$ and test-volume size $r_{test} = 0.125,$ showing an $N^{0.4} \approx N^{1-\nu}$ dependence for the range of the potential and an $N^{0.5}$ dependence on the magnitude.

are well-known [113], but are often inadequate for systems such as chain molecules with connectivity constraints [118]. Instead, the usual approach for such problems is an iterative algorithm which biases potentials according to the difference between the computed distribution function and the “target” function. Such methods date back at least to the work of Reatto *et al.* [130], and a number of different implementations exist, developed by Soper [145], Lyubartsev and Laaksonen [102], and Rutledge [134], among others. While these methods are in theory general, most have been applied only to systems with relatively few degrees of freedom, and fewer still have been used to develop coarse-grained potentials. The closest work in this respect to the one proposed here is that of Akkermans and Briels [2, 3]. Unfortunately, there are some difficulties associated with this work, as several researchers have reported that thermodynamic inconsistencies can result from improper use of such coarse-grained potentials [2, 96].

The general outline of this method is as follows. Given a target radial distribution function, $g^*(r)$, we would like to determine a potential $u(r)$ which reproduces the desired target. For an ansatz $u_{(k)}(r)$, we generate as output a corresponding radial distribution function $g_{(k)}(r)$. The method states that the updated Hamiltonian can be found by solving the system

$$\Delta \langle 4\pi r^2 \mathbf{g} \rangle = \mathbf{A} \Delta \mathbf{u},$$

where the vectors \mathbf{g} and \mathbf{u} are appropriately discretized versions of $g(r)$ and $u(r)$, respectively, and the coefficients of the matrix \mathbf{A} are given by

$$\mathbf{A}_{ij} = \frac{\partial \langle 4\pi r^2 g_i \rangle}{\partial u_j} = -\beta \langle g_i g_j \rangle - \langle g_i \rangle \langle g_j \rangle.$$

The process is then repeated with the updated $u_{(k+1)}(r)$ until an appropriate error threshold is reached, at which point the algorithm terminates, and a long-time simulation can be run with the resulting coarse-grained potential u^* . The starting ansatz is often a square-well potential, such as (2.7), where $\varepsilon_{nb}(r_{ij})$ is a constant ε_{nb} . In subsequent iterations of the procedure, each discrete region of the potential $u(r)$ may be modified; however, in virtually all cases, the algorithm is able to converge in ten iterations or less [11, 118, 134].

Thus, we can envision implementing SGMC as a means of improving the performance of our simulation. Since the coarse-grained simulations tend to be faster than the numerical

simulations required to obtain the correct statistics, it makes more sense to spend more time at the coarse-grained scale than at the fine scale. Thus, we could run a short simulation on a small chain to obtain the distributions for $F_N(\mathbf{x}, r_{test})$, then use scaling laws and (7.32) to obtain a coarse-grained potential. This potential would then be used as the starting point in an SGMC simulation, and the results of successive iterations at the coarse-grained scale should then correct the potential to its “correct” shape.

7.11 Conclusions

For freely-jointed chains, we have shown that very simple physical arguments can be used to approximate the two-body potentials observed for the overlap of chains. Although the resulting analytical distributions tend to overestimate the potential, they show the correct qualitative behavior, suggesting that single-chain simulations may be sufficient to compute coarse-grained potentials. In addition, we have seen strong evidence for the existence of scaling laws for the probability distribution which generates the potential, for both freely-jointed and self-avoiding chains.

Chapter 8

Coupling of variables as a consequence of coarse-graining

8.1 Discovering the coupling phenomenon

Like the previous chapter, the results presented in this chapter were obtained in an attempt to explain the results obtained in Figure 8-1. The graph shows the results for the mean end-to-end distance $\langle R_N \rangle$ for a self-avoiding walk treated atomistically using the pivot algorithm, and as coarse-grained via WAMC, with effective bead sizes of $N_e = 32$, $N_e = 64$, and $N_e = 128$. As we have mentioned previously, there is excellent agreement between the different coarse-grained simulations, yet there was a substantial discrepancy between the set of coarse-grained simulations and the atomistic simulation. As a result, we believed that we did not have all the information necessary to describe the coarse-grained polymer model. The results in Figure 8-1 were generated from the distribution of bond lengths, and the assumption that the hard-sphere potential of radius $r_{hs} = 1$ used at the atomistic level was still valid at the coarse-grained scale. The first attempt at restoring detail into the model was to use the method described in the previous chapter to produce an estimate for the true potential $u(r)$. Using the potential shown in 8-2 caused an observable but ultimately insufficient adjustment in the results for $\langle R_N \rangle$; the three to four percent adjustment did not offset a factor of two difference.

At this point, it becomes clear that we are still missing some important data to resolve the problem. However, no such information is forthcoming from our original model. Further

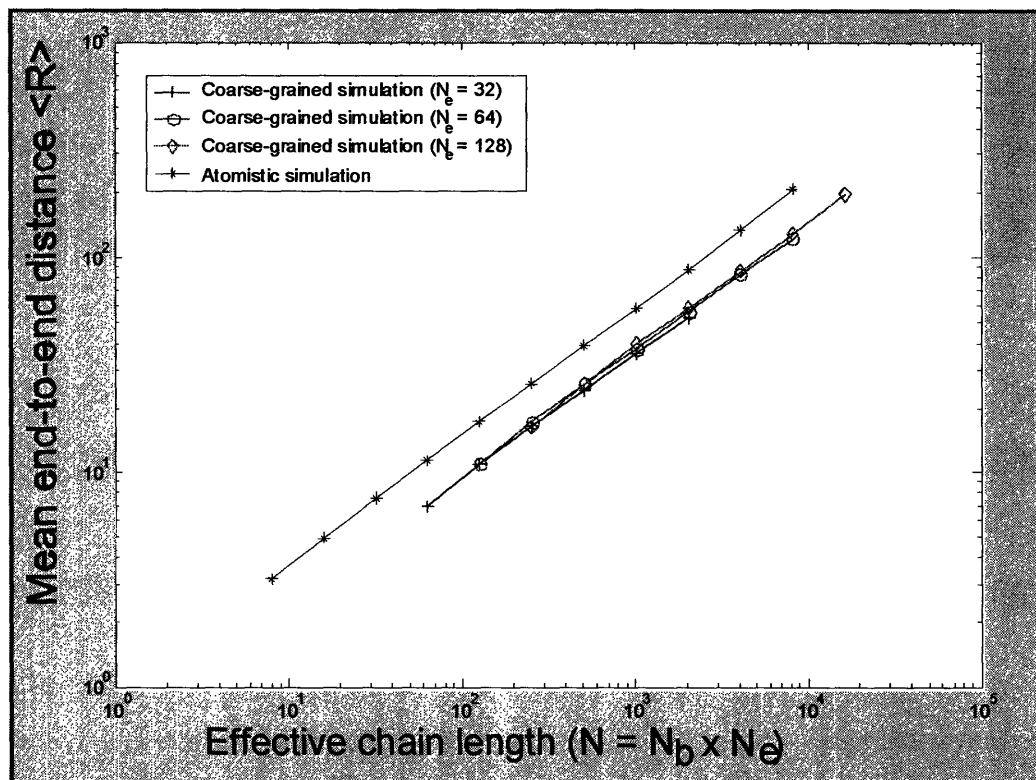


Figure 8-1: Mean end-to-end distance $\langle R \rangle$ of the pivot algorithm (top line) and of coarse-grained simulations based on bead sizes of 32, 64, and 128 (lower curves). Note the excellent agreement of the different coarse-grained curves with each other.

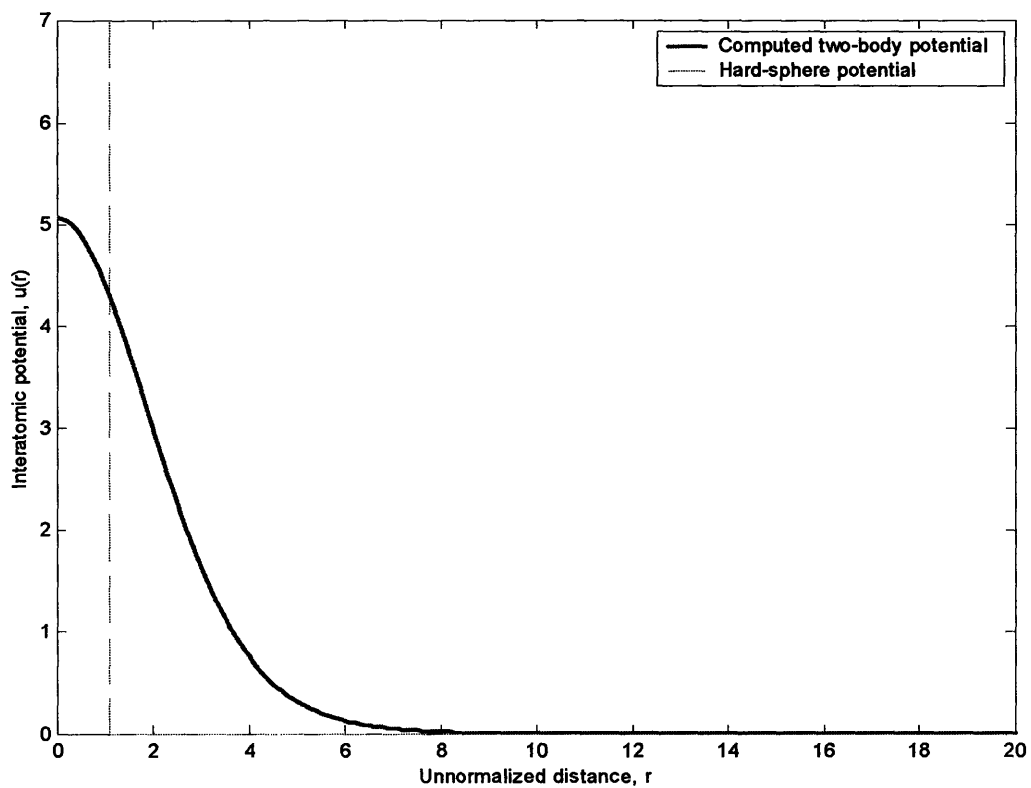


Figure 8-2: Estimate for the true interatomic potential $u(r)$ used to improve the WAMC self-avoiding walk model.

complicating the problem was data such as that shown in Figure 8-3. Figure 8-3 shows a scatter plot, in which each point indicates a pair of bond lengths, such that the length of bond i is the ordinate and the length of bond $i + 1$ is the abscissa. The solid curves underneath the plot represent contour lines of equal probability, under the assumption that the bond length probability distributions are independent of one another. The concentration of points at the center of the graph, and the relative “thinning out” observed as one moves away from the center implied that the bond length distributions were in fact independent of one another. Repeating the same plot for the bond length and bond angle distributions, as shown in Figure 8-4, shows that the assumption that bond lengths and bond angles are described by independent distributions appears to be reasonable.

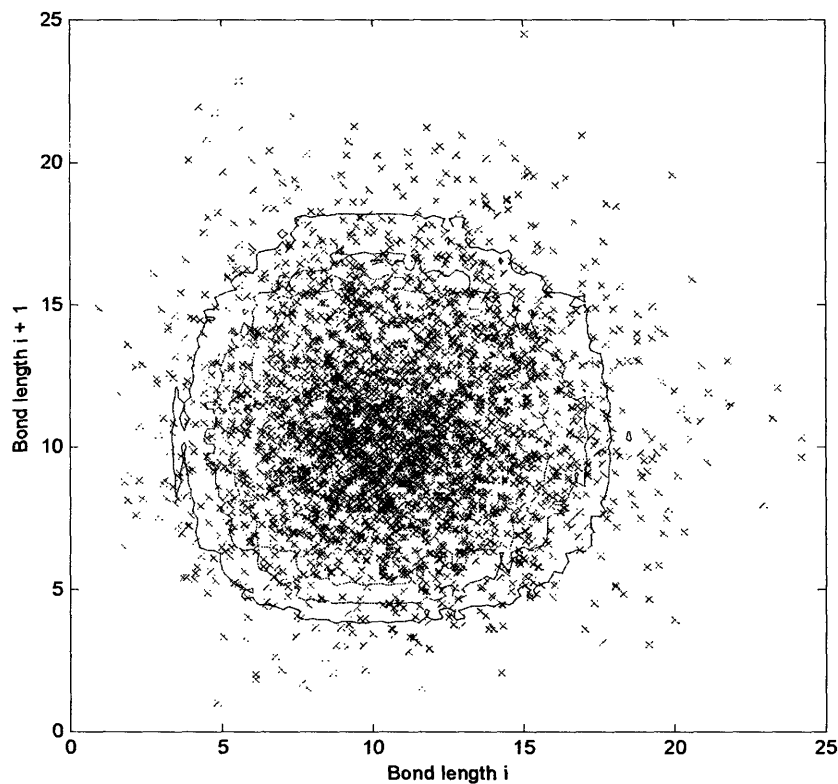


Figure 8-3: A correlation plot of adjacent coarse-grained bond lengths in a 256-mer, showing decorrelation between adjacent bonds connecting beads of effective size 64.

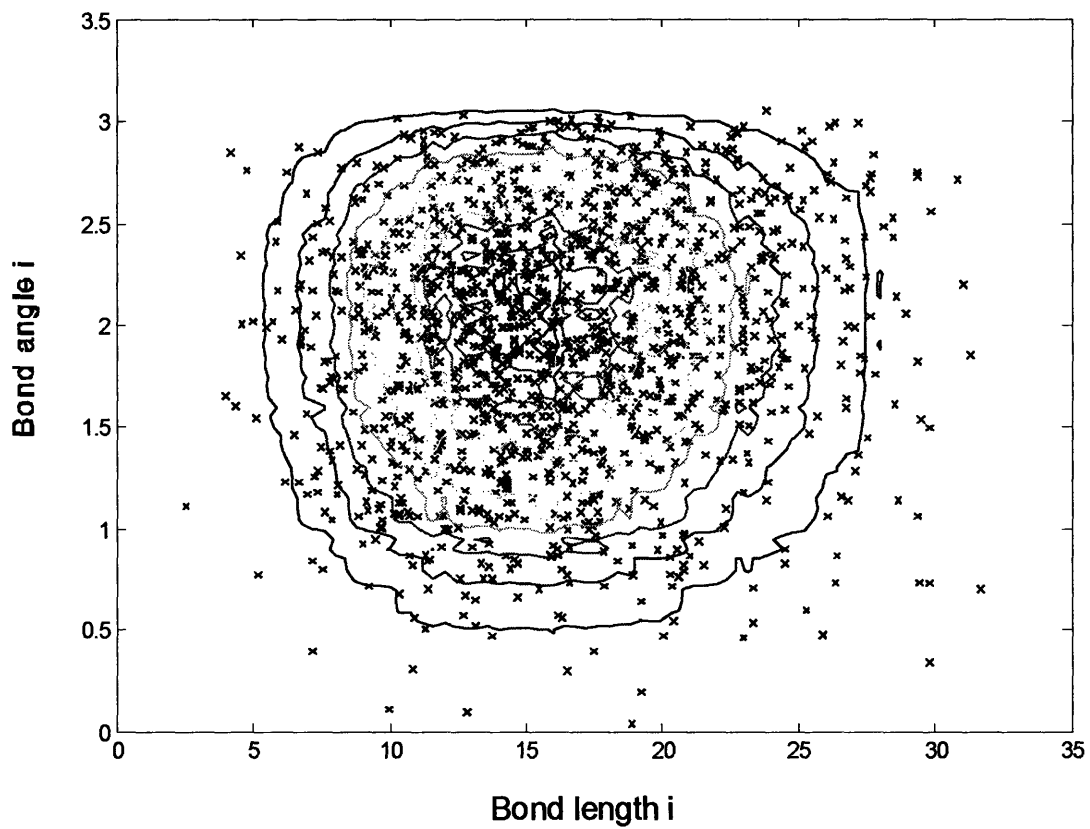


Figure 8-4: A correlation plot of coarse-grained bond lengths and adjacent bond angles in a 256-mer, allegedly showing decorrelation between bonds connecting beads of effective size 64 and their neighboring bond angles.

This led to an impasse: something appeared to be missing, but no leads to explain the situation remained. Consequently, an investigation of the freely jointed chain was undertaken, since it is conceptually simpler than the self-avoiding walk. However, initial results for the coarse-grained freely jointed chain model appeared no better than the self-avoiding walk, as significant errors remained.

Comparison of the coarse-grained freely-jointed chain model with that described in the work of Laso *et al.* [92] suggested that the assumption of independent bond lengths and bond angles was invalid for most coarse-grained systems, as the bond lengths are in fact interdependent: they claim that the covariance of adjacent bond vectors \mathbf{R}_i and \mathbf{R}_{i+1} is

$$\langle \mathbf{R}_{i-1,i} \mathbf{R}_{i,i+1} \rangle = -\frac{1-N^2}{18N} \langle l^2 \rangle \mathbf{E}, \quad (8.1)$$

where in 8.1 $\mathbf{R}_{i,i+1}$ is the vector connecting beads i and $i+1$, $\langle l^2 \rangle$ is the mean length of a bond at the atomistic scale, and \mathbf{E} is the three-dimensional unit tensor. Our derivation of this result, not supplied in the original paper, suggests that this result should actually be

$$\langle \mathbf{R}_{i-1,i} \mathbf{R}_{i,i+1} \rangle = \left(\frac{1}{3} + \frac{1}{18}N - \frac{1}{18N} \right) \langle l^2 \rangle \mathbf{E}, \quad (8.2)$$

which confirms that the correlation between neighboring bonds does not vanish. [See Appendix A for details on this derivation.] Moreover, they assert that the bond angles are dependent on the bond lengths, from which we conclude that our assumptions, based on Figures 8-3 and 8-4 were based on “false positives,” and not on the actual data.

To construct our coarse-grained model, it therefore became necessary to determine what additional information needed to be specified. For example, although (8.1) and (8.2) suggest that a conditional bond-length distribution is required, the results of Laso *et al.* [92] suggests that the actual effect of the correlation is negligible except for large bond lengths which comprise a negligible portion of the overall distribution—on the order of thousandths of one per cent of all bonds observed. Thus, we omitted the conditional bond distribution from our calculations.

On the other hand, we also realized that the bond-angle distribution of the coarse-grained walk was not, in fact, uniform. Instead, there is a pronounced bias in the distribution, favoring large bond angles over smaller ones, as shown in Figure 8-5. To simplify the work of using this

Table 8.1: Mean square radius of gyration using a single bond angle distribution

N	$\langle R_g^2 \rangle$ (atomistic)	$\langle R_g^2 \rangle$ (WAMC, single bond angle dist.)
[0.5ex] 512 = (16 × 32)	84.75	70.41
1024 = (32 × 32)	171.60	141.23

Table 8.2: Mean square radius of gyration using multiple bond angle distributions

N	$\langle R_g^2 \rangle$ (atomistic)	$\langle R_g^2 \rangle$ (WAMC, multiple bond angle dists.)
512 = (16 × 32)	84.75	74.15
1024 = (32 × 32)	171.60	150.10

angle distribution, we noted that the cumulative distribution function corresponding to $p(\phi)$ resembles a “stretched” Gaussian exponential, so we attempted to fit $p(\phi)$ to the form

$$p(\phi, \alpha) = \frac{\exp(-(\phi - \pi)^2 / \alpha) - \exp(-\pi^2 / \alpha)}{1 - \exp(-\pi^2 / \alpha)}, \quad (8.3)$$

where α is an adjustable parameter. For the freely-jointed chain, excellent agreement between the simulation results and the curve fit of (8.3) is obtained by selecting $\alpha = 2.5$, as shown in the lower graph in Figure 8-5.

Using the coarse-grained bond-length distribution and the approximation (8.3), both developed for effective bead sizes of 32, we then attempted to model a coarse-grained freely-jointed chain, computing its mean square radius of gyration, $\langle R_g^2 \rangle$. The results unfortunately still did not provide good quantitative agreement with the results of atomistic simulations, as is seen in Table 8.1.

The reasons for the discrepancy were at first unclear. At first, we considered the possibility that the bond angle depended more strongly upon the coarse-grained bond lengths than Figure 8-5 suggests. This is in fact true, as is shown in Figure 8-6, which parametrizes the bond angle according to the length r_1 “preceding” the bond angle. For the cases $r_1 < \sqrt{32}$, $r_1 < \frac{3}{2}\sqrt{32}$, and $r_1 \geq \frac{3}{2}\sqrt{32}$, we obtain α parameters of 3.3, 2.4, and 1.5, respectively. These parameters were then added to the simulation, which yielded improved values, but still not fully accurate values, as shown in Table 8.2.

Finally, we examined something not included in the work of Laso *et al.* [92]: the torsion-angle distribution. The asymmetry in the distribution is most clearly observed by plotting the

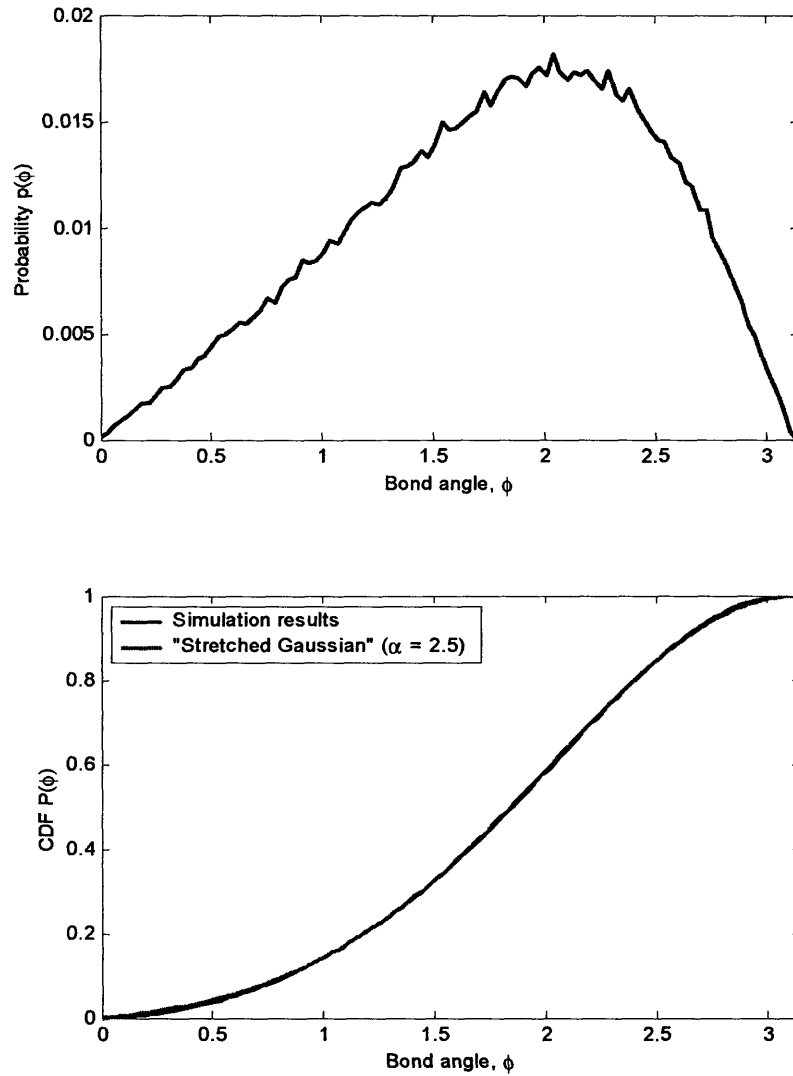


Figure 8-5: Probability distribution function $p(\phi)$ (top) and cumulative distribution function (bottom) $P(\phi)$ for the coarse-grained bond angle between successive coarse-grained bonds, as computed for beads of size 32 in a freely-jointed chain model. The lower curve also shows the stretched Gaussian $p(\phi, 2.5)$ defined by (8.3).

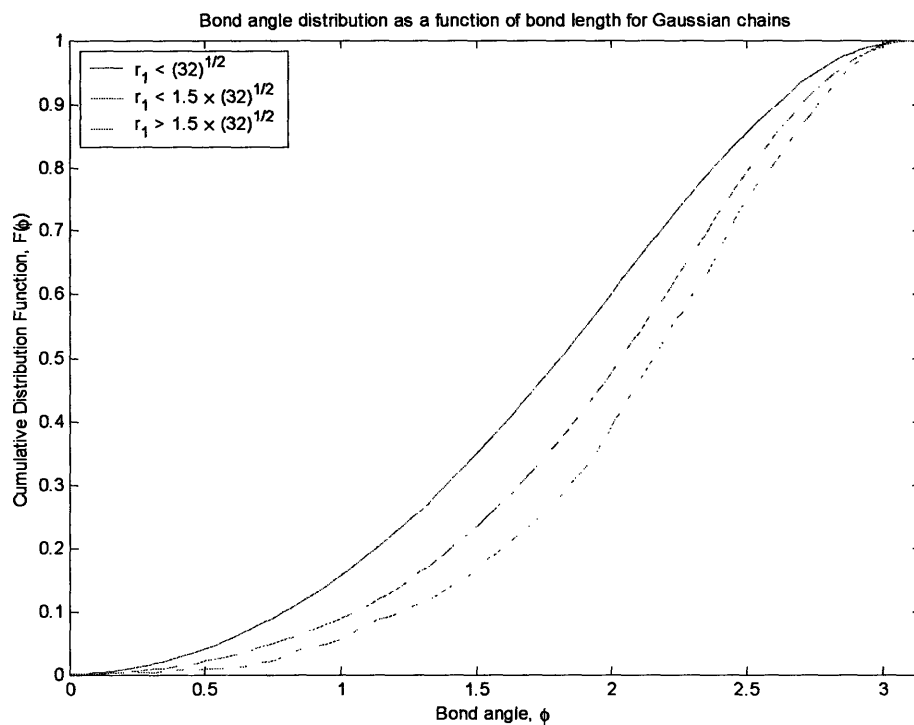


Figure 8-6: Cumulative distribution function $F(\phi)$ for the coarse-grained bond angle, in a Gaussian chain with $N = 128$ and $N_e = 32$, as a function of r_1 , the “first” bond length forming the angle. The top, solid curve shows bond lengths less than the ideal Gaussian value of $r_1 = \sqrt{N_e} = \sqrt{32}$; the middle, dashed curve shows $r_1 < \frac{3}{2}\sqrt{N_e}$; and the bottom, dash-dotted curve shows $r_1 \geq \frac{3}{2}\sqrt{N_e}$.

Table 8.3: Mean square radius of gyration using bond angle and torsion angle distributions

N	$\langle R_g^2 \rangle$ (atomistic)	$\langle R_g^2 \rangle$ (WAMC, bond and torsion angle dists.)
$512 = (16 \times 32)$	84.75	82.74
$1024 = (32 \times 32)$	171.60	169.80
$2048 = (64 \times 32)$	342.70	342.10
$4096 = (128 \times 32)$	679.95	685.69
$8192 = (256 \times 32)$	1360.22	1375.09

probability as a function of the cosine of the torsion angle ϕ , which is shown in Figure 8-7. This shows that there is a bias towards *trans* configurations instead of *cis* configurations, leading to a slight extension of the chain. Although the differences appear to be large, the bias turns out to be relatively small:

$$\begin{aligned}\langle \theta \rangle &= 1.6043, \\ \langle \cos \theta \rangle &= -0.0335.\end{aligned}$$

On the other hand, the effect of this change, coupled with the others, was significant enough to yield near-perfect agreement with the atomistic model, as is summarized in Table 8.3; more data is shown graphically in Figure 8-8, showing chains of lengths up to $N = 32768$. The scaling-law fit of the data,

$$\langle R_g^2 \rangle \sim N^{1.014 \pm 0.003}$$

is shown in Figure 8-8 as well; it is essentially indistinguishable from the atomistic result

$$\langle R_g^2 \rangle \sim N^{1.000 \pm 0.003},$$

except at the largest values of N .

8.2 Should these results surprise us?

In the previous section, we began with a freely-jointed chain model with no interactions, a monodisperse bond-length distribution and uniform bond-angle and torsion-angle distributions. The act of coarse-graining via the wavelet transform method produced a model which still had

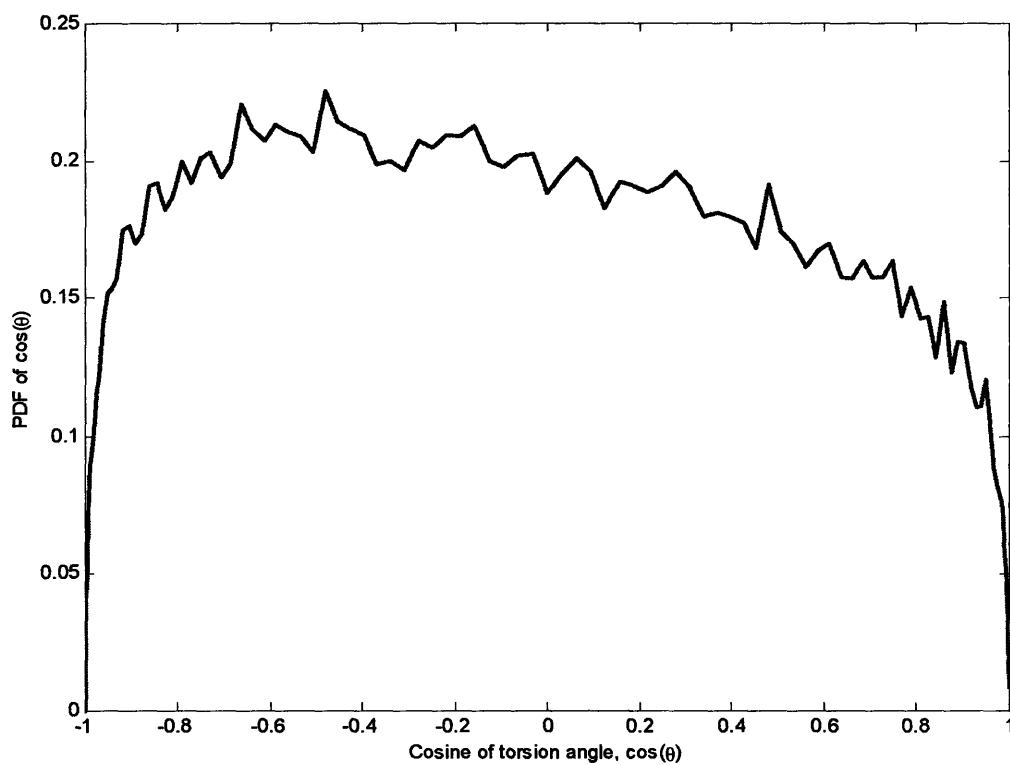


Figure 8-7: Probability distribution function for the torsion angle of a chain of length $N = 128$ considered as four beads of effective size $N_e = 32$.

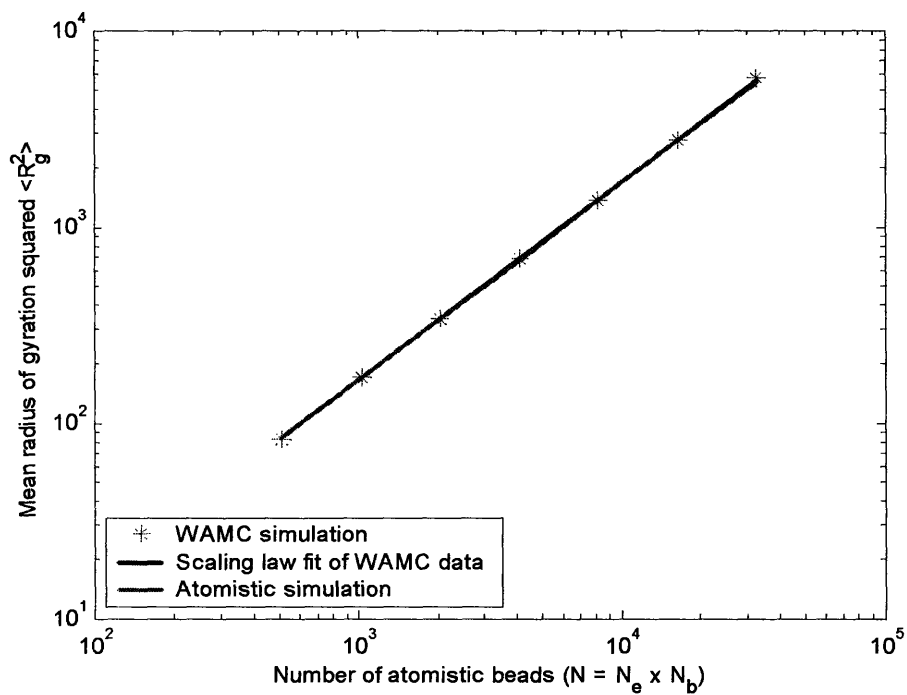


Figure 8-8: The mean square radius of gyration $\langle R_g^2 \rangle$ as a function of bead size N for both atomistic and WAMC representations of the freely-jointed chain.

no interactions, but in which bond angles were dependent upon bond lengths, and in which all the distributions were non-uniform. Is this result in fact surprising? Is there anything in the literature which would lead us to expect this behavior as a result of coarse-graining?

The most logical place to turn is the theory of critical phenomena and in particular the renormalization group method [19,85,104,146]. As explained in Section 2.4, that the one of the principal goals of using renormalization group principles to coarse-grain physical systems is to determine the flow diagrams in parameter space. That is, given a set of starting parameters \mathbf{K} , what new set of parameters \mathbf{K}^* correspond to \mathbf{K} after the coarse-graining procedure has been performed? The choice of parameters is arbitrary; however, we hope that the set we choose have well-defined (if not necessarily finite) values both before and after the coarse-graining procedure.

As previously noted, in the case of the Ising model, the wavelet transform coarse-graining method produces a self-interaction term that is not observed in the traditional Ising model. Similarly, renormalization-group analyses of the two-dimensional Ising model on the square lattice yields a more complicated model. Unlike the one-dimensional case, which only adds a constant term, the two-dimensional case adds both a constant term and a term describing interactions between *four* neighboring spins—those that meet at a common corner [124]. This introduces an increased level of correlation between the coarse-grained variables, not observed at the fine scale.

In the case of coarse-graining the freely-jointed chain, if we consider our parameter to be the lack of an interaction term, then we find that the wavelet transform method leaves the parameter unchanged—a result predicted by renormalization group theory. On the other hand, if we were to define a new parameter

$$\zeta_{\alpha\beta} = \frac{\langle\alpha\beta\rangle - \langle\alpha\rangle\langle\beta\rangle}{\langle\alpha\rangle\langle\beta\rangle},$$

which represents the coupling of parameters, we find that the coarse-graining operation has a very significant impact: for example, we cannot reasonably expect $\zeta_{r\phi}$ to be zero. In fact, we find that neighboring bond lengths r_i and r_{i+1} are correlated, as is the bond length r_i and the corresponding bond angle ϕ_i . Neighboring bond angles ϕ_i and ϕ_{i+1} show a much

Table 8.4: Correlation parameter for different pairs of coarse-grained variables

α	β	$\zeta_{\alpha\beta}$
r_i	r_{i+1}	0.00996
r_i	ϕ_i	0.02356
ϕ_i	ϕ_{i+1}	0.00407

smaller correlation, as can be seen in Table 8.4. These results confirm our previous hypotheses: neighboring bond angles can be chosen independently, and under some circumstances so can neighboring bond lengths.

However, we should expect that bond angles and bond lengths to be linked. The weak correlation between coarse-grained variables is actually the most difficult case to handle. If the correlation between the bond length and the bond angle was near-unity, we could easily devise a mapping which would convert a given bond length into a corresponding bond angle. If the correlation had been zero, we could simply choose the variables to be independent of one another. With the value non-negligible—it is significantly larger than the correlation either between adjacent bond angles or between adjacent bond lengths—we know that there is some structure in the data, but not a direct relationship between bond lengths and bond angles, thus leading us to expect a much more complicated behavior between the two.

8.3 Information gain in coarse-grained systems

8.3.1 Polymer models

Another notable feature of coarse-graining is the tendency of the coarse-grained variables to have greater information content [78] than any individual corresponding fine-grained variables which created it, but less than the total entropy of the entire set of corresponding fine-grained variables. The information inherent in a probability distribution is defined as

$$I = \frac{\langle \ln p \rangle}{\ln 2}; \quad (8.4)$$

that is, the information is the same as the entropy of the distribution times the constant $-1/(k_B \ln 2)$. The greater the value of I , the more we know about the system.

As an example, consider choosing a bond angle. If we were to assume a Gaussian bond angle, then the choice is arbitrary in the range $0 \leq \phi \leq \pi$. The information is then given by

$$I' = \frac{\int_0^\pi dx \frac{1}{\pi} \ln \frac{1}{\pi}}{\ln 2} = \frac{-\ln \pi}{\ln 2} \approx -1.651$$

In contrast, we have shown that (8.3) can be used to define the cumulative distribution function for the coarse-grained bond angle for a Gaussian chain. Using the derivative of (8.3),

$$p(\phi) = \frac{d}{d\phi} \left(\frac{\exp\left(-(\phi - \pi)^2/\alpha\right) - \exp(-\pi^2/\alpha)}{1 - \exp(-\pi^2/\alpha)} \right) = \frac{2(\pi - \phi) e^{-\frac{(\phi - \pi)^2}{\alpha}}}{\alpha \left(1 - e^{-\frac{\pi^2}{\alpha}}\right)}$$

with the parameter choice $\alpha = 2.5$ gives for the information content

$$I = \frac{\int_0^\pi d\phi p(\phi) \ln p(\phi)}{\ln 2} = \frac{-0.9777}{\ln 2} = -1.410,$$

which tells us that the coarse-grained bond angle distribution contains more information than the Gaussian bond angle distribution. Moreover, although I increases as $\alpha \rightarrow 0$, as expected—as the bond length increases, a smaller range of angles are available—even as $\alpha \rightarrow \infty$, corresponding to the smallest bond lengths, $I(\alpha) > I'$, so we will always have more information content in the coarse-grained bond angle than in the atomistic bond angles, even if we use multiple values of α to characterize the bond-angle distribution.

Similarly, the Gaussian torsion angle distribution, which is uniform over $0 \leq \theta \leq 2\pi$, has information content equal to

$$I = \frac{\int_0^{2\pi} dx \frac{1}{2\pi} \ln \frac{1}{2\pi}}{\ln 2} = \frac{-\ln 2\pi}{\ln 2} \approx -2.651.$$

If we use Figure 8-7 to estimate the information content of the torsion distribution, we obtain

$$I = \frac{1}{\ln 2} \sum_i (p(\theta_i) \ln p(\theta_i)) \Delta\theta \approx -1.324,$$

so we again find that the information content of the coarse-grained torsion angle is greater than that of the original torsion angle.

The only “exception” to this rule that we observe in the Gaussian case is the bond length distribution: the coarse-grained model will have $I < 0$, while the atomistic chain will have $I = 0$. While true, this is somewhat misleading, since there is no uncertainty in the bond length of the original model, and therefore we must obtain $I = 0$. However, we can compare the results for our bond-length distribution for distance between the centers-of-mass of 32-mers, for which we compute, using the data shown in Figure 6-3,

$$I = \frac{1}{\ln 2} \sum_i (p(\theta_i) \ln p(\theta_i)) \Delta r \approx -3.168$$

with the results for the Gaussian end-to-end distance of 32-mers,

$$\begin{aligned} I &= \frac{\int_0^\infty dr p(r) \ln p(r)}{\ln 2} \\ &= \frac{\int_0^\infty 4\pi r^2 \left(\frac{3}{2\pi(32)}\right)^{3/2} \exp\left(-\frac{3r^2}{2(32)}\right) \left(\frac{3}{2} \ln\left(\frac{3}{2\pi(32)}\right) - \frac{3r^2}{2(32)}\right) dr}{\ln 2} \\ &= -11.264, \end{aligned}$$

which indicates that much more information is contained in knowing the distribution of centers of mass of a chain than the end-to-end distance of coarse-grained segments.

A comparison of the three results is also instructive: the bond lengths and torsion angles have significantly large increases in information gain as a function of coarse-graining than does the bond angle. The differences are significant—a factor of 2 to 4 reduction for the bond lengths and torsion angles, but only about 15 to 20 per cent for the bond angle. Thus, we suspect that a small increase in information of a coarse-grained distribution is an indication that the distribution of the given variable is in fact dependent on one or more of the other degrees of freedom which are effectively independent of one another.

8.3.2 Ising lattices

We can obtain similar results for the Ising model. Let us consider the mapping of a square of size 8×8 to a single Ising spin. If we assumed that all block spins, were equally likely, the

information content of the block spin would be

$$I = \frac{1}{\ln 2} \sum_{j=0}^{64} \frac{1}{64} \ln \frac{1}{64} \approx -6.094;$$

by contrast, since we actually know the distribution of spins making up a block spin, we can repeat this calculation with the actual data, shown in Figure 4-5, to obtain

$$I = \frac{1}{\ln 2} \sum_{j=0}^{64} p \left(\frac{64-2j}{64} \right) \ln p \left(\frac{64-2j}{64} \right) \approx -4.047$$

in the high-temperature limit,

$$I = \frac{1}{\ln 2} \left(\frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{1}{2} \right) \approx -1$$

in the zero-temperature limit in the absence of magnetic field, and

$$I = 0,$$

as $T \rightarrow 0$ and $h \rightarrow \pm\infty$. Consequently, knowing the magnetization of the block spin provides much more information regarding the configuration of the block spin than the equal *a priori* assumption. This is particularly true in the vicinity of the attractive fixed points of the Ising model (shown in Figure 2-3).

8.4 When should we expect interdependence?

One of the reasons that using the wavelet transform as a coarse-graining mechanism leads to increased interdependence of the coarse-grained degrees of freedom is that it is an averaging process. Lattice spins are replaced with spins equal to the average spin of the block; beads are replaced with coarse-grained beads at their center of mass. Consequently, the interactions already inherent at the atomistic scale are carried over to the fine scale. In the case of lattice spins, we have interactions between the spins, which makes the distribution of expected states more difficult to predict; we use the behavior of a sublattice at a previous scale as a means of

approximating this behavior. Similarly, the coarse-grained bond length in the polymer model is a function of the positions corresponding to both coarse-grained beads in the chain.

We can therefore suggest the following principle as a rule for when interdependence increases.

Conjecture 1 *The complexity of the distributions of a variable in a coarse-grained system will increase beyond that of the distributions of the corresponding atomistic degrees of freedom which created it when:*

- 1. the quantity being investigated depends on interactions between two or more coarse-grained degrees of freedom, and*
- 2. the quantity has a small increase in information content after coarse-graining, or is weakly correlated with another degree of freedom.*

The first part of this principle should be intuitively obvious: The distribution of a variable changes when it is affected by another variable in the same system; this is merely rehashing the notion of conditional expectations. Thus, the only way to prevent more complicated distribution functions from appearing is to define coarse-grained variables that are non-interacting. However, our ability to define such coarse-grained variables is limited to a very small number of simple systems, or to coarse-grained length scales so large as to be meaningless. [For example, coarse-graining a semidilute polymer chain to a collection of particles, each of which is much smaller than the radius of gyration, would yield coarse-grained variables that are more or less non-interacting, but also non-informative.]

An example where coarse-graining does *not* lead to increased interdependence can be found in another model of the polymer random walk. If we choose to coarse-grain a freely-jointed chain by replacing a sequence of adjacent bonds by their end-to-end distance, we can create a model with independent coarse-grained bonds if we choose the segments to be longer than the persistence length of the original chain. However, in this case, the “atomistic” bond vectors which determine the coarse-grained bond length belong to only one coarse-grained bead, and therefore do not violate the principle set forth above. In the case of a lattice system, this would be akin to selecting a coarse-grained block size sufficiently large such that the spins are effectively decoupled. For such systems, the interactions between block spins are negligible in

comparison to the intra-block interactions, and thus knowing the distribution of spins suffices to determine the behavior of the overall system.

The second notion, that a small increase in information content or a weak correlation with another degree of freedom is related to the complexity of the coarse-grained probability distribution, is less obvious. However, upon reflection, it makes sense that these two factors would be correlated with increased complexity of behavior. If the information gain is small, it can mean that even though the distribution may be significantly altered from the previous fine-grained distribution, only a small amount of additional precise knowledge about its behavior has been gained—which may indicate that we have introduced a parametrized set of coarse-grained distributions, instead of a distribution valid under all conditions. Similarly, the presence of weak correlations can be a sign of increased dependence, since there is a clear sign of a link between variables, but not a strong enough correlation to infer a direct relationship between the given variables.

Chapter 9

Future research directions and conclusions

9.1 Future research directions

We have outlined in this thesis a framework, based on the principle of hierarchical organization, which is capable of simulating numerous different types of physical models. We can extend this work both within the context of the models we have studied, as well as branch out into studies of different systems. Some possible new extensions and applications of the WAMC framework are outlined below.

- Problems with defined length scales
 - The effects of confinement on the behavior of polymer systems has important ramifications for studies of separation and analysis of polymer, protein, and DNA molecules. It is well-known that a wall can be viewed as a colloidal particle of infinite radius, whose effects on the interaction of polymer molecules must be determined [22]. In the context of WAMC analysis, the presence of the confining walls or tube introduces a length scale into the problem which ultimately provides a bound on the amount of coarse-graining in a chain: it seems ill-advised to choose a coarse-graining scale in which the mean bond length is greater than the tube radius or the distance between walls, unless the behavior of the polymer is sufficiently rodlike to justify such

an approximation. Moreover, this would appear to be a better approach to modeling confined polymers than the soft colloid approach [20, 99], which creates a single particle with only nine degrees of freedom (a center of mass, and the orientation and principal axes of the ellipsoid defining the complete chain).

- Entangled melts are difficult to simulate, because of the complicated interactions where segments on different polymers intersect. However, we can determine the distance between cross-links of chains, and therefore derive an effective length scale for coarse-graining. The distance between cross-links should be greater than the coarse-grained effective bead size, to ensure that polymer chains between links are represented by multiple beads. This approach would be useful for simulating gels and other “networked” polymers.
- Heterogeneous coarse-graining would allow us to make our algorithm, which is multiresolution between stages of coarse-graining, a multiresolution algorithm within a single stage of the coarse-graining process. That is, we would have multiple resolutions of coarse-graining within the simulation system, or even within a single chain. For example, a block copolymer could have different blocks be assigned different effective bead sizes. Additionally, a semidilute homopolymer might be represented by a high level of coarse-graining when no polymer is nearby, but be represented by a finer-scaled description for those parts of the chain which are within the interaction range of another chain.

The main challenge in creating a heterogeneous simulation would be obtaining the statistics required to create the interaction potentials needed for the coarse-grained simulation. Different effective bead sizes—or different molecular structures—would require the sampling of all different pairs of possible effective bead sizes, which would effectively slow down the time required to perform the atomistic-level simulations. However, the increased information would allow us to focus computational time where it would be most useful in determining structural and thermodynamic properties.

- Another fruitful approach in molecular simulation has been hybrid Monte Carlo-molecular dynamics simulations, particularly for systems with surface reactions, where the reactive kinetics can be handled using Monte Carlo, while diffusion to and from the surface occurs

via molecular dynamics [24,84]. Other approaches integrate the two techniques by letting short molecular dynamics simulations generate new Monte Carlo configurations [5,65].

We can couple the WAMC framework to Molecular Dynamics simulations using several different approaches. One such method would be to create effective potentials for interactions between coarse-grained particles and between atomistic and coarse-grained particles. Then, instead of computing long-range potentials between an atom and a cluster of particles, we could cast the problem as an interaction between an atom and a coarse-grained particle, which would cut down on the time required to compute the intermolecular forces which is a central feature of MD simulations.

- Extensions to more complicated models
 - As mentioned above, the first obvious extension of the polymer modeling is to the semidilute and concentrated regimes, where interactions between polymers are important, allowing for the study of heterogeneous coarse-graining approaches.
 - Block copolymers could be described as the interaction of different types of atomistic beads—such as beads with different radii, or different interaction strengths [8,36,37,63,94,112].
 - Charged polymers should in principle also respond well to the WAMC approach. Coarse-graining techniques which only examine end-to-end distance to determine the location of coarse-grained beads will face serious obstacles in describing the effects of electrostatic interactions, which are strongly dependent on the location of the individual beads in the chain. Knowing the center of mass of a polymer segment provides much more information about the charge distribution than the end-to-end distance of the segment. However, the challenge introduced by electrostatic interactions is how to handle the additional intra-coarse-grained-bead interactions, as well as the interactions between the coarse-grained beads. In effect, we will need to determine either a new screening length for the coarse-grained particles, or a new effective charge that better reflects the electrostatic interactions between the coarse-grained beads.

- The extensions and further applications of the WAMC framework are not limited to polymer models. We can also consider the (restricted) primitive model [53, 54, 95], which is another approach for studying electrostatic interactions between ions. In this model, ions sit on discrete lattice sites, but have their standard Coulombic interactions. The model shows much the same behavior as more complicated models, but are good models for studying aggregation of charged particles, particularly when there is a size or charge discrepancy between the positive and negative charges.

9.2 Conclusions: Fundamental insights on coarse-graining

The principal products of this thesis have been algorithms for coarse-grained simulations of lattice models and polymer models based on the principle of wavelet transformations. We have demonstrated that the coarse-graining scheme is effective and efficient, compared to atomistic simulation techniques. In addition, the simulation algorithms can return accurate results and reproduces important cases, such as behavior near attractive fixed-points, essentially exactly.

However, equally valuable to the basic algorithms which we have produced is the insight gained about coarse-graining strategies. During the development of the WAMC algorithm and its model applications, we have uncovered a number of fundamental principles which have not yet garnered much attention in the literature on coarse-graining, despite their general applicability to most approaches.

- We have shown that coarse-graining is an adaptable process: there is no single “correct” way to coarse-grain a given system. Even after a particular coarse-graining mechanism has been selected, it is often necessary to make additional choices about the coarse-graining procedure. For example, if we choose to use the WAMC framework, how many levels of hierarchical simulation should we perform? What is the appropriate coarse-grained resolution to use? What data do we need to pass from one level to the next?
- In addition, the use of a hierarchical approach can be quite fruitful: the ability to break down complex simulations into the study of interactions of smaller, more easily manageable units has lead to a framework which can handle simulations of several different types (off-lattice or on-lattice, solid-state or condensed, positional or topological). Moreover,

the difficulty of the simulations can be reduced from running on a Beowulf cluster or supercomputer to running on a personal workstation, effectively removing improvements in and costs of hardware as an important factor in obtaining results from complex physical simulations.

- Coarse-graining intrinsically involves a trade-off between performance and the accuracy of the simulation both at the current scale and at later scales of a hierarchical coarse-graining scheme. Were we to use the $N^{-1/2}$ scaling law for the coarse-grained intramolecular potential and approximations for the coarse-grained variables we discussed in Chapters 7 and 8, we could save substantial time in performing the atomistic-level simulations. However, we have also shown that the behavior of this scaling law, while qualitatively correct, is quantitatively inaccurate, so the increased computational efficiency comes at the cost of greater error in our system.
- Moreover, we have a better understanding of how coarse-graining mechanisms achieve accurate results: like many other coarse-graining approaches, our WAMC framework is “brittle” at fine scales, and “resilient” at coarser scales. Quantitative accuracy of WAMC simulations depends largely on computing the behavior of the finest-scaled stage as accurately as possible; the more approximations introduced, the worse the overall performance of the simulation at coarser scales. However, further levels of coarse-graining do little to degrade performance of the algorithm: a crude four-scale simulation of a polymer using WAMC returns the same basic scaling behavior as a more accurate and detailed two-stage simulation, but with much faster execution times.
- As coarse-graining proceeds, there is an increase in the complexity of the remaining degrees of freedom in the simulation. Coarse-graining is widely perceived to produce a faster, simpler model than the original; thus such a claim appears counterintuitive. In reality, most coarse-graining techniques rely on “lumping” particles together. Since one particle now must act like a composite of many finer-scaled particles, this tends to yield models with degrees of freedom which can take many more allowed values than the original fine-grained models from which they were derived. For example, for the Ising lattice we showed that WAMC produces states which have a multiplicity of possible spins, instead

of the up-down dichotomy which must be maintained by renormalization approaches. In the same vein, for the polymer model, the descriptions of bond and torsion angles in even freely-jointed chains are much more complicated than the equal *a priori* probability distributions which are used for the atomistic chain.

- Finally, we have shown the utility of hierarchical and multiresolution methods. In principle, most “handshaking” coarse-graining techniques have been developed because it was difficult, if not impossible, to bridge the length scales inherent in different simulation methods, such as quantum chemistry and molecular-dynamics methods. However, a multiresolution method allows us to achieve coarse-graining scaleups of sufficient size that it may be feasible in the near future to create molecular-level simulations at length scales comparable with more mesoscale approaches such as self-consistent field theory and dissipative particle dynamics and macroscale approaches such as finite elements.

Appendix A

Center-of-mass bond statistics for a Gaussian chain

We would like to determine the correlation $\langle(-\mathbf{R}_{i-1,i})\mathbf{R}_{i,i+1}\rangle$ between adjacent bonds for an isotropic system in which self-avoidance constraints are ignored, the probability distribution for a single step of the chain is

$$p(\mathbf{x}) = \left(\frac{3}{2\pi x_0^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{x}^2}{2x_0^2}\right),$$

so that the chain obeys Gaussian statistics. It is readily shown that the end-to-end distance \mathbf{R} of a segment of N steps is given by [42]

$$P_N(\mathbf{R}) = \left(\frac{3}{2\pi N x_0^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{R}^2}{2N x_0^2}\right).$$

Let us define the distance between the first and last beads in a segment as \mathbf{r}_i , and the distance between the first bead and the center of mass as \mathbf{s}_i . Then the distance $\mathbf{R}_{i,i+1}$ between the centers of mass of the i th and $(i+1)$ th beads can be written as

$$\mathbf{R}_{i,i+1} = \mathbf{r}_i - \mathbf{s}_i + \mathbf{s}_{i+1}. \tag{A.1}$$

Since the walk is isotropic, we know that only the diagonal terms in $\langle \mathbf{r}_i \mathbf{r}_i \rangle$ will be nonzero, as all other terms vanish due to the periodicity of the trigonometric functions. Of the remaining terms, we know that all three will have the same value, given by a calculation involving any of the three. Thus, calculating $\langle r_{ix} r_{ix} \rangle$, we find that

$$\langle r_{ix} r_{ix} \rangle = \left(\frac{3}{2\pi N x_0^2} \right)^{3/2} \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \int_0^\infty t^2 dt r_{ix}^2 \exp\left(-\frac{3t^2}{2N x_0^2}\right) \quad (\text{A.2})$$

defining $r_{ix} = t \cos \phi \sin \theta$, we have from (A.2)

$$\begin{aligned} \langle r_{ix} r_{ix} \rangle &= \left(\frac{3}{2\pi N x_0^2} \right)^{3/2} \int_0^{2\pi} d\phi \cos^2 \phi \int_0^\pi \sin \theta d\theta (1 - \cos^2 \theta) \int_0^\infty t^2 dt t^2 \exp\left(-\frac{3t^2}{2N x_0^2}\right) \\ &= \left(\frac{3}{2\pi N x_0^2} \right)^{3/2} \left(\frac{4}{3}\pi \right) \int_0^\infty dt t^4 \exp\left(-\frac{3t^2}{2N x_0^2}\right). \end{aligned} \quad (\text{A.3})$$

To evaluate the Gaussian integral in (A.3), we can use the well-known formula

$$\int_0^\infty dx x^m \exp(-ax^n) = \frac{1}{na^{(m+1)/n}} \Gamma\left(\frac{m+1}{n}\right). \quad (\text{A.4})$$

Inserting $m = 4$, $n = 2$, and $a = 3/2N x_0^2$ in (A.4), we obtain from (A.3),

$$\langle r_{ix} r_{ix} \rangle = \left(\frac{3}{2\pi N x_0^2} \right)^{3/2} \left(\frac{4}{3}\pi \right) \frac{\Gamma\left(\frac{5}{2}\right)}{2 \left(\frac{3}{2N x_0^2}\right)^{5/2}} = \frac{1}{3} N x_0^2, \quad (\text{A.5})$$

which is the result we expected.¹ Consequently, the full correlation matrix is defined as

$$\langle \mathbf{r}_i \mathbf{r}_i \rangle = \frac{1}{3} N x_0^2 \mathbf{E},$$

where \mathbf{E} is the unit tensor in \mathbb{R}^3 .

Calculation of the correlation function $\langle \mathbf{s}_i \mathbf{s}_i \rangle$ of the center-of-mass function is slightly more complicated. However, the isotropic nature of the system does allow us to conclude that $\langle \mathbf{s}_i \mathbf{s}_i \rangle$ is again a diagonal matrix. Noting that the distance \mathbf{s} between the first bead and the center of

¹Obviously, if we had any other result for (??), this would imply that the system is either nonisotropic, or that $\langle r^2 \rangle \neq N r_0^2$. [We also could have simply proceeded backwards from the latter assumption and used the isotropic nature of the system to intuit (??).]

mass can be defined as

$$\mathbf{s} = \sum_{n=0}^{N-1} \left(\frac{n}{N} - 1 \right) \mathbf{x}_n,$$

where \mathbf{x}_n is the n th step in the segment, we have

$$\begin{aligned} \langle \mathbf{s}\mathbf{s} \rangle &= \left\langle \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left(\frac{j}{N} - 1 \right) \left(\frac{k}{N} - 1 \right) \mathbf{x}_j \mathbf{x}_k \right\rangle \\ &= \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left(\frac{j}{N} - 1 \right) \left(\frac{k}{N} - 1 \right) \langle \mathbf{x}_j \mathbf{x}_k \rangle \\ &= \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left(\frac{j}{N} - 1 \right)^2 \langle \mathbf{x}_j \mathbf{x}_j \rangle, \end{aligned} \quad (\text{A.6})$$

where (A.6) occurs because the steps are assumed independent. Since $\langle \mathbf{x}_j \mathbf{x}_j \rangle = \frac{1}{3} a^2 \mathbf{E}$, (A.6) becomes

$$\langle \mathbf{s}\mathbf{s} \rangle = \frac{1}{3} a^2 \mathbf{E} \sum_{j=0}^{N-1} \left(\frac{j}{N} - 1 \right)^2 = \frac{(N-1)(2N-1)}{18N} x_0^2 \mathbf{E}.$$

Similarly, the cross-correlation $\langle \mathbf{r}\mathbf{s} \rangle$ is

$$\langle \mathbf{r}\mathbf{s} \rangle = \left\langle \sum_{j=0}^{N-1} \mathbf{x}_j \sum_{k=0}^{N-1} \left(1 - \frac{k}{N} \right) \mathbf{x}_k \right\rangle = \sum_{k=0}^{N-1} \left(1 - \frac{k}{N} \right) \langle \mathbf{x}_k \mathbf{x}_k \rangle = \frac{N+1}{6} x_0^2 \mathbf{E}.$$

Therefore, from (A.1), we have

$$\begin{aligned} \langle \mathbf{R}_{i,i+1} \mathbf{R}_{i,i+1} \rangle &= \langle (\mathbf{r}_i - \mathbf{s}_i + \mathbf{s}_{i+1}) (\mathbf{r}_i - \mathbf{s}_i + \mathbf{s}_{i+1}) \rangle \\ &= \langle \mathbf{r}_i \mathbf{r}_i - 2\mathbf{r}_i \mathbf{s}_i + \mathbf{s}_i \mathbf{s}_i + \mathbf{s}_{i+1} \mathbf{s}_{i+1} \rangle \\ &= \frac{1}{3} N x_0^2 \mathbf{E} - 2 \frac{N+1}{6} x_0^2 \mathbf{E} + 2 \left(\frac{(N-1)(2N-1)}{18N} x_0^2 \mathbf{E} \right) \\ &= \frac{1}{9} \frac{2N^2 - 6N + 1}{N} x_0^2 \mathbf{E}, \end{aligned}$$

since the correlations between the different coarse-grained particles vanish. Similarly, for the

correlation $\langle \mathbf{R}_{i-1,i} \mathbf{R}_{i,i+1} \rangle$, we have

$$\begin{aligned}
 \langle -(\mathbf{R}_{i-1,i}) \mathbf{R}_{i,i+1} \rangle &= \langle (-\mathbf{r}_{i-1} + \mathbf{s}_{i-1} - \mathbf{s}_i)(\mathbf{r}_i - \mathbf{s}_i + \mathbf{s}_{i+1}) \rangle \\
 &= \langle \mathbf{s}_i \mathbf{s}_i - \mathbf{r}_i \mathbf{s}_i \rangle \\
 &= \left(-\frac{1}{3} - \frac{1}{18}N + \frac{1}{18N} \right) x_0^2 \mathbf{E},
 \end{aligned}$$

which demonstrates that adjacent center-of-mass bonds *do not* have an independent distribution, and therefore require some sort of bias, either in the sampling of states or in the form of a potential, to uncouple the bond lengths.

Appendix B

Converting internal coordinates

Given a chain described by the displacement vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, bond length r , bond angle ϕ , and torsion angle θ , we would like to determine the displacement vector \mathbf{x}_{n+1} corresponding to (r, ϕ, θ) . The radial component of \mathbf{x}_{n+1} should be in the direction of \mathbf{x}_n ; thus, we choose as the first unit vector $\mathbf{e}_1 = \mathbf{x}_n / |\mathbf{x}_n|$. The remaining two unit vectors should be orthogonal to \mathbf{e}_1 . One such vector can be found by taking the cross product of \mathbf{x}_{n-1} and \mathbf{x}_n , so that

$$\mathbf{e}_2 = \frac{\mathbf{x}_{n-1} \times \mathbf{x}_n}{|\mathbf{x}_{n-1} \times \mathbf{x}_n|}.$$

The third vector needs to be orthogonal to both \mathbf{e}_1 and \mathbf{e}_2 , so we choose

$$\mathbf{e}_3 = \frac{\mathbf{e}_2 \times \mathbf{e}_1}{|\mathbf{e}_2 \times \mathbf{e}_1|}.$$

Given (r, ϕ, θ) , we can determine their relative contributions in the \mathbf{e}_1 , \mathbf{e}_2 , and \mathbf{e}_3 directions as

$$\begin{aligned} u_1 &= r \cos(\pi - \phi), \\ u_2 &= r \sin(\pi - \phi) \cos(\pi - \theta), \\ u_3 &= r \sin(\pi - \phi) \sin(\pi - \theta). \end{aligned}$$

Thus, in standard Cartesian components, we can write \mathbf{x}_{n+1} as

$$\mathbf{x}_{n+1} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \mathbf{e}_3^T \end{bmatrix} \mathbf{u}, \quad (\text{B.1})$$

where in (B.1), \mathbf{e}_i^T is the transpose of vector \mathbf{e}_i , and \mathbf{u} is the vector $\mathbf{u} = (u_1, u_2, u_3)$.

Bibliography

- [1] Abrams, C. F. and K. Kremer. “Combined coarse-grained and atomistic simulation of liquid bisphenol A-polycarbonate: Liquid packing and intramolecular structure,” *Macromolecules*, *36*:260–267 (2003).
- [2] Akkermans, Reinier L. C. and W. J. Briels. “Coarse-Grained Interactions in Polymer Melts: A Variational Approach,” *J. Chem. Phys.*, *115*:6210–6219 (2001).
- [3] Akkermans, Reinier L. C. and W. J. Briels. “A Structure-Based Coarse-Grained Model for Polymer Melts,” *J. Chem. Phys.*, *114*:1020–1031 (2001).
- [4] Alexander-Katz, A., et al. “Field-Theoretic Simulations of Confined Polymer Solutions,” *J. Chem. Phys.*, *118*:9030–9036 (2003).
- [5] Allen, M. P. and D. J. Tildesley. *Computer Simulation of Liquids*. London: Oxford University Press, 1989.
- [6] Amit, Daniel J. *Field Theory, the Renormalization Group, and Critical Phenomena*. New Jersey: World Scientific, 1978.
- [7] Arias, T. A. “Multiresolution Analysis of Electronic Structure: Semicardinal and Wavelet Bases,” *Rev. Mod. Phys.*, *71*:267–311 (1999).
- [8] Bahiana, M. and Y. Oono. “Cell Dynamical System Approach to Block Copolymers,” *Phys. Rev. A*, *41*:6763–6781 (1990).
- [9] Barenblatt, G. I. *Scaling, Self-Similarity, and Intermediate Asymptotics*. New York: Cambridge University Press, 1996.

- [10] Baschnagel, J., et al. “Bridging the Gap Between Atomistic and Coarse-Grained Models of Polymers: Status and Perspectives,” *Adv. Poly. Sci.*, 152:41–156 (2000).
- [11] Bathe, Mark and G. C. Rutledge. “An Inverse Monte Carlo Procedure for Conformation Determination of Macromolecules,” *J. Comp. Chem.*, 24:876–890 (2003).
- [12] Battle, Guy. *Wavelets and Renormalization*. Singapore: World Scientific, 1999.
- [13] Bazant, Martin Z. “The Largest Cluster in Subcritical Percolation.” <http://xxx.lanl.gov/abs/cond-mat/9905191>, 1999.
- [14] Best, Christoph. “Wavelet-Induced Renormalization Group for the Landau-Ginzburg Model.” <http://xxx.lanl.gov/abs/hep-lat/9909151>, 1999.
- [15] Best, Christoph, et al. “Wavelets as a Variational Basis of the XY Model,” *Nucl. Phys. B: Proc. Suppl.*, 34:780–782 (1994).
- [16] Best, Christoph and Andreas Schäfer. “Variational Description of Statistical Field Theories Using Daubechies’ Wavelets.” <http://xxx.lanl.gov/abs/hep-lat/9402012>, 1994.
- [17] Beylkin, Gregory, et al. “Fast Spectral Projection Algorithms for Density-Matrix Computations,” *J. Comput. Phys.*, 152:32–54 (1999).
- [18] Billingsley, Patrick. *Probability and Measure*. New York: Wiley Interscience, 1995.
- [19] Binney, J. J., et al. *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*. Oxford: Oxford University Press, 1993.
- [20] Bolhuis, P. G. and A. A. Louis. “How to Derive and Parametrize Effective Potentials in Colloid-Polymer Mixtures,” *Macromolecules*, 35:1860–1869 (2002).
- [21] Bolhuis, P. G., et al. “Many-Body Interactions and Correlations in Coarse-Grained Descriptions of Polymer Solutions,” *Phys. Rev. E*, 64:021801 (2001).
- [22] Bolhuis, P. G., et al. “Influence of polymer-excluded volume on the phase-behavior of colloid-polymer mixtures,” *Phys. Rev. Lett.*, 89:128302 (2002).

- [23] Bolhuis, P. G., et al. “Accurate Effective Pair Potentials for Polymer Solutions,” *J. Chem. Phys.*, *114*:4296–4311 (2001).
- [24] Brennan, J. K. and B. M. Rice. “Molecular Simulation of Shocked Materials Using the Reactive Monte Carlo Method,” *Phys. Rev. E*, *66*:021105 (2002).
- [25] Broughton, Jeremy Q., et al. “Concurrent Coupling of Length Scales: Methodology and Application,” *Phys. Rev. B*, *60*:2391–2402 (1999).
- [26] Carmesin, I. and Kurt Kremer. “The Bond Fluctuation Method: A New Effective Algorithm for the Dynamics of Polymers in All Spatial Dimensions,” *Macromolecules*, *21*:2819–2823 (1988).
- [27] Chen, Lin-Yuan, et al. “Renormalization Group Theory for Global Asymptotic Analysis,” *Phys. Rev. Lett.*, *73*:1311–1315 (1994).
- [28] Cho, Junhan and Wayne L. Mattice. “Estimation of Long-Range Interaction in Coarse-Grained Rotational Isomeric State Polyethylene Chains on a High Coordination Lattice,” *Macromolecules*, *30*:637–644 (1997).
- [29] Chou, Kenneth Chien-Ko. *A Stochastic Modeling Approach to Multiscale Signal Processing*. PhD dissertation, Massachusetts Institute of Technology, 1991.
- [30] Cohen, Albert and Ingrid Daubechies. “Non-Separable Bidimensional Wavelet Bases,” *Rev. Mat. Ibero-amer.*, *9*:51–137 (1993).
- [31] Cormen, Thomas H., et al. *Introduction to Algorithms*. Cambridge, MA: McGraw Hill-MIT Press, 1990.
- [32] Coveney, Peter V. and Keir E. Novik. “Computer Simulations of Domain Growth and Phase Separation in Two-Dimensional Immiscible Fluids Using Dissipative Particle Dynamics,” *Phys. Rev. E*, *54*:5134–5141 (1996).
- [33] Daubechies, I. “Orthonormal Bases of Compactly Supported Wavelets,” *Commun. Pure Appl. Math.*, *41*:909–996 (1988).

- [34] Daubechies, Ingrid. *Ten Lectures on Wavelets*, 61. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia: SIAM, 1992.
- [35] Daubechies, Ingrid and Wim Sweldens. "Factoring Wavelet Transforms Into Lifting Steps," *J. Fourier Anal. Appl.*, 4:247–269 (1998).
- [36] David, Edwin F. and Kenneth S. Schweizer. "Integral-Equation Theory of Block Copolymer Liquids. I. General Formalism and Analytic Predictions for Symmetrical Copolymers," *J. Chem. Phys.*, 100:7767–7783 (1994).
- [37] David, Edwin F. and Kenneth S. Schweizer. "Integral Equation Theory of Block Copolymer Liquids. II. Numerical Results for Finite Hard-Core Diameter Chains," *J. Chem. Phys.*, 100:7784–7793 (1994).
- [38] de Gennes, Pierre-Gilles. *Scaling Concepts in Polymer Physics*. New York: Cornell University Press, 1979.
- [39] De Groot, R. "Electrostatic Interactions in Dissipative Particle Dynamics-Simulation of Polyelectrolytes and Anionic Surfactants," *J. Chem. Phys.*, 118:11265–11277 (2003).
- [40] Des Cloizeaux, Jacques and Gérard Jannink. *Polymers in Solution: Their Modelling and Structure*. New York: Oxford-Clarendon, 1990.
- [41] Dickman, R. and C. K. Hall. "Equation of State for Chain Molecules: Continuous-Space Analog of Flory Theory," *J. Chem. Phys.*, 85:4108–4115 (1986).
- [42] Doi, M. and S. F. Edwards. *The Theory of Polymer Dynamics*. London: Clarendon Press, 1988.
- [43] Doruker, Pemra and Wayne L. Mattice. "Reverse Mapping of Coarse-Grained Polyethylene Chains from the Second Nearest Neighbor Diamond Lattice to an Atomistic Model in Continuous Space," *Macromolecules*, 30:5520–5526 (1997).
- [44] Doruker, Pemra and Wayne L. Mattice. "A Second Generation of Mapping/Reverse Mapping of Coarse-Grained and Fully Atomistic Models of Polymer Melts," *Macromol. Theory Simul.*, 8:463–478 (1999).

- [45] Español, Pep. “Hydrodynamics from Dissipative Particle Dynamics,” *Phys. Rev. E*, 52:1734–1742 (1995).
- [46] Español, Pep. “Dissipative Particle Dynamics for a Harmonic Chain: A First-Principles Derivation,” *Phys. Rev. E*, 53:1572–1578 (1996).
- [47] Español, Pep. “Fluid Particle Model,” *Phys. Rev. E*, 57:2930–2948 (1998).
- [48] Eurich, F. and P. Maass. “Soft Ellipsoid Model for Gaussian Polymer Chains,” *J. Chem. Phys.*, 114:7655–7668 (2001).
- [49] Eyink, Gregory and Nigel Goldenfeld. “Analogies Between Scaling in Turbulence, Field Theory, and Critical Phenomena,” *Phys. Rev. E*, 50:4679–4683 (1994).
- [50] Faller, R., et al. “Local Structure and Dynamics of Trans-Polyisoprene Oligomers,” *Macromolecules*, 34:1436–1448 (2001).
- [51] Feldmann, M. T., et al. “Efficient Algorithm for "On-the-Fly" Error Analysis of Local or Distributed Serially-Correlated Data.” Submitted to *J. Chem. Phys.*, 2003.
- [52] Feldmann, Michael Todd. *Quantum Monte Carlo: Quest to Bigger, Faster, and Cheaper*. PhD dissertation, California Institute of Technology, 2002.
- [53] Fisher, M. E. “The Story of Coulombic Criticality,” *J. Stat. Phys.*, 75:1–36 (1994).
- [54] Fisher, M. E. “The Nature of Criticality in Ionic Fluids,” *J. Phys. Cond. Mat.*, 8:9103–9109 (1996).
- [55] Fisher, Michael E. “Renormalization Group Theory: Its Basis and Formulation in Statistical Physics,” *Rev. Mod. Phys.*, 70:653–681 (1998).
- [56] Fleer, G. J., et al. *Polymers at Interfaces*. London: Chapman and Hall, 1993.
- [57] Flekkøy, Eirik G. and Peter V. Coveney. “From Molecular Dynamics to Dissipative Particle Dynamics,” *Phys. Rev. Lett.*, 83:1775–1778 (1999).
- [58] Flekkøy, Eirik G., et al. “Foundations of Dissipative Particle Dynamics,” *Phys. Rev. E*, 62:2140–2157 (2000).

- [59] Flory, Paul J. *Principles of Polymer Chemistry*. Ithaca, NY: Cornell University Press, 1953.
- [60] Flyvbjerg, H. and H. G. Petersen. “Error Estimates on Averages of Correlated Data,” *J. Chem. Phys.*, *91*:461–466 (1989).
- [61] Fraaije, J. G. E. M. “Dynamic Density Functional Theory for Microphase Separation Kinetics of Block Copolymer Melts,” *J. Chem. Phys.*, *99*:9202–9212 (1993).
- [62] Fredrickson, Glenn H. “Dynamics and Rheology of Inhomogeneous Polymeric Fluids: A Complex Langevin Approach,” *J. Chem. Phys.*, *117*:6810–6820 (2002).
- [63] Fredrickson, Glenn H., et al. “Field-Theoretic Computer Simulation Methods for Polymers and Complex Fluids,” *Macromolecules*, *35*:16–39 (2002).
- [64] Freed, Karl F. *Renormalization Group Theory of Macromolecules*. New York: Wiley, 1987.
- [65] Frenkel, Daan and Berend Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. San Diego: Academic Press, 1996.
- [66] Gamero, L. G., et al. “Wavelet Analysis and Nonlinear Dynamics in a Nonextensive Setting,” *Physica A*, *246*:487–509 (1997).
- [67] Ganesan, V. and G. H. Frederickson. “Field-Theoretic Polymer Simulations,” *Europhys. Lett.*, *55*:814–820 (2001).
- [68] Ganesan, V. and V. Pryamitsyn. “Dynamical Mean-Field Theory for Inhomogeneous Polymeric Systems,” *J. Chem. Phys.*, *118*:4345–4348 (2003).
- [69] Gines, D., et al. “LU Factorization of Non-Standard Forms and Direct Multiresolution Solvers,” *Appl. Comput. Harmon. A.*, *5*:156–201 (1998).
- [70] Goddard, W. A., et al. “Strategies for Multiscale Modeling and Simulation of Organic Materials: Polymers and Biopolymers,” *Comput. Theor. Polym. Sci.*, *11*:329–343 (2001).
- [71] Goldenfeld, Nigel. *Lectures on Phase Transitions and the Renormalization Group*. Reading, MA: Addison-Wesley, 1992.

- [72] Golub, Gene H. and Charles F. Van Loan. *Matrix Computations*. Baltimore: Johns Hopkins University Press, 1996.
- [73] Green, H. S. and C. A. Hurst. *Order-Disorder Phenomena*. New York: Wiley Interscience, 1964.
- [74] Haar, A. “Zur Theorie der Orthogonalen Funktionen-Systeme,” *Math. Ann.*, 69:331–371 (1910).
- [75] Hahn, Oliver, et al. “Simulation of Polymer Melts: From Spherical to Ellipsoidal Beads,” *Macromol. Theor. Simul.*, 10:288–303 (2001).
- [76] Haliloglu, T., et al. “Simulations of Rotational Isomeric State Models for Poly(Propylene) Melts on a High Coordination Lattice,” *Macromol. Theory Simul.*, 7:613–617 (1998).
- [77] Hansen, J.-P. and I. R. McDonald. *Theory of Simple Liquids*. New York: Academic Press, 1986.
- [78] Honerkamp, Josef. *Statistical Physics: An Advanced Approach with Applications*. Berlin: Springer, 1998.
- [79] Huang, Ding-Wei. “Wavelet Analysis in Multiplicity Fluctuations,” *Phys. Rev. D*, 56:3961–3969 (1997).
- [80] Hughes, Barry D. *Random Walks and Random Environments*. New York: Oxford-Clarendon Press, 1995.
- [81] Irving, William W. *Multiscale Stochastic Realization and Model Identification with Applications to Large-Scale Estimation Problems*. PhD dissertation, Massachusetts Institute of Technology, 1995.
- [82] Istrail, Sorin. “Statistical Mechanics, Three-Dimensionality, and NP-Completeness: I. Universality of Intractability for the Partition Function of the Ising Model Across Non-Planar Surfaces.” *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*. 87–96. Portland, OR: ACM Press, 2000.

- [83] Johnson, Bruce R., et al. "Quadrature Integration for Orthogonal Wavelet Systems," *J. Chem. Phys.*, 110:8309–8317 (1999).
- [84] Johnson, J. K., et al. "Reactive Canonical Monte-Carlo: A New Simulation Technique for Reacting or Associating Fluids," *Mol. Phys.*, 81:717–733 (1994).
- [85] Kadanoff, Leo P. *Statistical Physics: Statics, Dynamics, and Renormalization*. Singapore: World Scientific, 2000.
- [86] Katsoulakis, Markos A., et al. "Coarse-Grained Stochastic Processes and Monte Carlo Simulations in Lattice Systems," *J. Comp. Phys.*, 186:250–278 (2003).
- [87] Katsoulakis, Markos A. and D. G. Vlachos. "Coarse-Grained Stochastic Processes and Kinetic Monte Carlo Simulators for the Diffusion of Interacting Particles," *J. Chem. Phys.*, 119:9412–9427 (2003).
- [88] Kennedy, Tom. "A Faster Implementation of the Pivot Algorithm for Self-Avoiding Walks," *J. Stat. Phys.*, 106:407–429 (2002).
- [89] Krakoviack, V., et al. "Relating Monomer to Centre-of-Mass Distribution Functions in Polymer Solutions," *Europhys. Lett.*, 58:53–59 (2002).
- [90] Lal, Moti. "'Monte Carlo' Computer Simulation of Chain Molecules. I.," *Molec. Phys.*, 17:57–64 (1969).
- [91] Landau, David P. and Kurt Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge: Cambridge University Press, 2000.
- [92] Laso, M., et al. "Bond-Length and Bond-Angle Distributions in Coarse-Grained Polymer Chains," *J. Chem. Phys.*, 95:2178–2182 (1991).
- [93] Lavis, David A. and George M. Bell. *Statistical Mechanics of Lattice Systems 1: Closed-Form and Exact Solutions*. Berlin: Springer, 1999.
- [94] Leibler, Ludwik. "Theory of Microphase Separations in Block Copolymers," *Macromolecules*, 13:1602–1617 (1980).

- [95] Levin, Y. and M. E. Fisher. “Criticality in the Hard-Sphere Ionic Fluid,” *Physica A*, 225:164–220 (1996).
- [96] Louis, A. A. “Beware of Density-Dependent Pair Potentials,” *J. Phys.: Cond. Mat.*, 14:9187–9206 (2002).
- [97] Louis, A. A., et al. “Coarse-Graining Polymers as Soft Colloids,” *Physica A*, 306:251–261 (2002).
- [98] Louis, A. A., et al. “Mean-Field Fluid Behavior of the Gaussian Core Model,” *Phys. Rev. E*, 62:7961–7972 (2000).
- [99] Louis, A. A., et al. “Can Polymer Coils Be Modeled as “Soft Colloids”?” *Phys. Rev. Lett.*, 85:2522–2525 (2000).
- [100] Luetzgen, Mark R. *Image Processing with Multiscale Stochastic Models*. PhD dissertation, Massachusetts Institute of Technology, 1993.
- [101] Luetzgen, Mark R., et al. “Multiscale Representations of Markov Random Fields,” *IEEE Trans. Sign. Process.*, 41:3377–3396 (1993).
- [102] Lyubartsev, Alexander P. and Aatto Laaksonen. “Calculation of Effective Interaction Potentials from Radial Distribution Functions: A Reverse Monte Carlo Approach,” *Phys. Rev. E*, 52:3730–3737 (1995).
- [103] Lyubartsev, Alexander P. and Aatto Laaksonen. “Osmotic and Activity Coefficients from Effective Potentials for Hydrated Ions,” *Phys. Rev. E*, 55:5689–5696 (1997).
- [104] Ma, Shang-Keng. *Modern Theory of Critical Phenomena*. Reading, MA: Benjamin/Cummings, 1976.
- [105] Madras, N. and A. D. Sokal. “The Pivot Algorithm—a Highly Efficient Monte-Carlo Method for the Self-Avoiding Walk,” *J. Stat. Phys.*, 50:109–186 (1988).
- [106] Mallat, Stephane G. “Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$,” *Trans. Amer. Math. Soc.*, 315:69–87 (1989).

- [107] Mallat, Stephane G. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11:674–693 (1989).
- [108] Marsh, C. A., et al. “Fokker-Planck-Boltzmann Equation for Dissipative Particle Dynamics,” *Europhys. Lett.*, 38:411–415 (1997).
- [109] Marsh, C. A., et al. “Static and Dynamic Properties of Dissipative Particle Dynamics,” *Phys. Rev. E*, 56:1676–1691 (1997).
- [110] Marsh, C. A. and P. V. Coveney. “Detailed Balance and H -Theorems for Dissipative Particle Dynamics,” *J. Phys. A: Math. Gen.*, 31:6561–6568 (1998).
- [111] Marsh, C. A. and J. M. Yeomans. “Dissipative Particle Dynamics: The Equilibrium for Finite Time Steps,” *Europhys. Lett.*, 37:511–516 (1997).
- [112] Matsen, M. W. and M. Schick. “Stable and Unstable Phases of a Diblock Copolymer Melt,” *Phys. Rev. Lett.*, 72:2660–2663 (1994).
- [113] McQuarrie, Donald A. *Statistical Mechanics*. New York: HarperCollins, 1976.
- [114] Migdal, A. A. “Phase Transitions in Gauge and Spin-Lattice Systems,” *Sov. Phys.-JETP*, 42:743–746 (1976).
- [115] Migdal, A. A. “Recursion Equations in Gauge Field Theories,” *Sov. Phys.-JETP*, 42:413–418 (1976).
- [116] Müller, M., et al. “Intra- and Interchain Correlations in Semidilute Polymer Solutions: Monte Carlo Simulations and Renormalization Group Results,” *Macromolecules*, 33:4568–4580 (2000).
- [117] Müller, Marcus. “Miscibility Behavior and Single Chain Properties in Polymer Blends: A Bond Fluctuation Model Study,” *Macromol. Theory Simul.*, 8:343–374 (1999).
- [118] Müller-Plathe, Florian. “Scale-Hopping in Computer Simulations of Polymers,” *Soft Mat.*, 1:1–31 (2003).

- [119] Murat, Michael and Kurt Kremer. “From Many Monomers to Many Polymers: Soft Ellipsoid Model for Polymer Melts and Mixtures,” *J. Chem. Phys.*, *108*:4340–4348 (1998).
- [120] O’Carroll, Michael. “Lattice and Continuum Wavelets and the Block Renormalization Group,” *J. Stat. Phys.*, *71*:415–423 (1993).
- [121] O’Carroll, Michael. “Multiscale Representation of Generating and Correlation Functions for Some Models of Statistical Mechanics and Quantum Field Theory,” *J. Stat. Phys.*, *73*:945–958 (1993).
- [122] Onsager, Lars. “Crystal Statistics 1: A Two-Dimensional Model with an Order-Disorder Transition,” *Phys. Rev.*, *65*:117–149 (1944).
- [123] Pan, G. and C. W. Manke. “Developments Toward Simulation of Entangled Polymer Melts by Dissipative Particle Dynamics (DPD),” *Int. J. Mod. Phys. B*, *17*:231–235 (2003).
- [124] Pathria, R. K. *Statistical Mechanics*. Woburn, MA: Butterworth-Heinemann, 1996.
- [125] Paul, Wolfgang, et al. “Dynamics of Polymer Solutions and Melts. Reptation Predictions and Scaling of Relaxation Times,” *J. Chem. Phys.*, *95*:7726–7740 (1991).
- [126] Pawley, G. S., et al. “Monte Carlo Renormalization-Group Calculations of Critical Behavior in the Simple-Cubic Ising Model,” *Phys. Rev. B*, *29*:4030–4040 (1984).
- [127] Pereira, E. and A. Procacci. “Block Renormalization Group Approach for Correlation Functions of Interacting Fermions,” *Lett. Math. Phys.*, *42*:261–270 (1997).
- [128] Pereira, E., et al. “Multiscale Formalism for Correlation Functions of Fermions. Infrared Analysis of the Tridimensional Gross-Neveu Model,” *J. Stat. Phys.*, *95*:665–692 (1999).
- [129] Rapold, Roland F. and Wayne L. Mattice. “Introduction of Short and Long Range Energies to Simulate Real Chains on the 2nd Lattice,” *Macromolecules*, *29*:2457–2466 (1996).
- [130] Reatto, L., et al. “Iterative Predictor-Corrector Method for Extraction of the Pair Interaction from Structural Data for Dense Classical Liquids,” *Phys. Rev. A*, *33*:3451–3465 (1986).

- [131] Rodgers, Seth T. and Klavs F. Jensen. “Multiscale Modeling of Chemical Vapor Deposition,” *J. Appl. Phys.*, *83*:524–530 (1998).
- [132] Rouault, Y., et al. “Concentration Profile Near the Surface of Polymer Mixtures: A Monte Carlo Study,” *Polymer*, *37*:297–304 (1996).
- [133] Rudd, R. E. and J. Q. Broughton. “Concurrent Coupling of Length Scales in Solid State Systems,” *Phys. Stat. Sol. B*, *217*:251–291 (2000).
- [134] Rutledge, Gregory C. “Modeling Experimental Data in a Monte Carlo Simulation,” *Phys. Rev. E*, *63*:021111 (2001).
- [135] Schlijper, A. G., et al. “Computer Simulation of Dilute Polymer Solutions with the Dissipative Particle Dynamics Method,” *J. Rheol.*, *39*:567–579 (1995).
- [136] Schmid, F. “Self-Consistent-Field Theories for Complex Fluids,” *J. Phys.: Cond. Mat.*, *10*:8105–8138 (1998).
- [137] Sethna, James P., et al. “Scaling Theory for the Glass Transition,” *Phys. Rev. B*, *44*:4943–4959 (1991).
- [138] Shardlow, T. “Splitting for Dissipative Particle Dynamics,” *SIAM J. Sci. Comp.*, *24*:1267–1282 (2003).
- [139] Sheng, N., et al. “Multiscale Micromechanical Modeling of Polymer/Clay Nanocomposites and the Effective Clay Particle,” *Polymer*, *45*:487–506 (2004).
- [140] Singh, Surjit and R. K. Pathria. “Spin-Spin Correlations in Finite Systems: Scaling Hypothesis and Corrections to Bulk Behavior,” *Phys. Rev. B*, *33*:672–674 (1986).
- [141] Singh, Surjit and R. K. Pathria. “Spin-Spin Correlations in Finite Systems with $O(n)$ Symmetry: Scaling Hypothesis and Corrections to Bulk Behavior,” *Phys. Rev. B*, *36*:3769–3781 (1987).
- [142] Singh, Surjit and R. K. Pathria. “Finite-Size Scaling of $O(n)$ Models in Higher Dimensions,” *Phys. Rev. B*, *38*:2740–2748 (1988).

- [143] Singh, Surjit and R. K. Pathria. “Phase Transitions in Finite Systems with $O(n)$ Symmetry: Magnetization and Susceptibility in the Presence of an External Field,” *Phys. Rev. B*, *37*:7806–7814 (1988).
- [144] Singh, Surjit and R. K. Pathria. “Finite-Size Scaling of $O(n)$ Models with Long-Range Interactions,” *Phys. Rev. B*, *40*:9238–9248 (1989).
- [145] Soper, A. K. “Empirical Potential Monte Carlo Simulation of Fluid Structure,” *Chem. Phys.*, *202*:295–306 (1996).
- [146] Stanley, H. Eugene. *Introduction to Phase Transitions and Critical Phenomena*. Oxford: Clarendon Press-Oxford, 1971.
- [147] Stephanopoulos, George, et al. “Multi-Scale Aspects in Model-Predictive Control,” *J. Proc. Cont.*, *10*:275–282 (2000).
- [148] Strang, Gilbert and Truong Nguyen. *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
- [149] Succi, S., et al. “Applying the Lattice Boltzmann Equation to Multiscale Fluid Problems,” *Comp. Sci. Eng.*, *3*:26–37 (2001).
- [150] Sweldens, W. and R. Piessens. “Asymptotic Error Expansions of Wavelet Approximations of Smooth Functions II,” *Numer. Math.*, *68*:377–401 (1994).
- [151] Sweldens, W. and R. Piessens. “Quadrature Formulae and Asymptotic Error Expansions for Wavelet Approximations of Smooth Functions,” *SIAM J. Numer. Anal.*, *31*:1240–1264 (1994).
- [152] Sweldens, W. and P. Schröder. “Building your own wavelets at home.” *Wavelets in Computer Graphics* 15–87, New Orleans: ACM SIGGRAPH Course notes, 1996.
- [153] Sweldens, Wim. “The Lifting Scheme: A Construction of Second Generation Wavelets,” *SIAM J. Math. Anal.*, *29*:511–546 (1997).
- [154] Swendsen, Robert H. “Monte Carlo Renormalization Group,” *Phys. Rev. Lett.*, *42*:859–861 (1979).

- [155] Trefethen, Lloyd N. and David Bau III. *Numerical Linear Algebra*. Philadelphia: SIAM, 1997.
- [156] Tschöp, W., et al. "Simulation of Polymer Melts. I. Coarse-Graining Procedure for Polycarbonates," *Acta Polymer.*, 49:61–74 (1998).
- [157] Tschöp, W., et al. "Simulation of Polymer Melts. II. From Coarse-Grained Models Back to Atomistic Description," *Acta Polymer.*, 49:75–79 (1998).
- [158] White, Steven R. "Density-Matrix Algorithms for Quantum Renormalization Groups," *Phys. Rev. B*, 48:10345–10356 (1993).
- [159] White, Steven R. and Richard L. Martin. "Ab initio Quantum Chemistry Using the Density Matrix Renormalization Group," *J. Chem. Phys.*, 110:4127–4130 (1999).
- [160] Wilson, Kenneth J. and J. Kogut. "The Renormalization Group and the ϵ Expansion," *Phys. Rep.*, 12:75–200 (1974).
- [161] Yamakawa, Hiromi. *Modern Theory of Polymer Solutions*. New York: Harper & Row, 1971.
- [162] Yong, C. W. and P. G. Higgs. "Chain Orientation in Polymer Networks: Computer Simulations Using the Bond Fluctuation Model," *Macromolecules*, 32:5062–5071 (1999).