



HAL
open science

Phase Transitions in Long-range Spin Models: The Power of Generalized Ensembles

Sylvain Reynal

► **To cite this version:**

Sylvain Reynal. Phase Transitions in Long-range Spin Models: The Power of Generalized Ensembles. Data Analysis, Statistics and Probability [physics.data-an]. Université de Cergy Pontoise, 2005. English. NNT: . tel-00010256

HAL Id: tel-00010256

<https://theses.hal.science/tel-00010256>

Submitted on 23 Sep 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE CERGY-PONTOISE

THÈSE

présentée
pour obtenir

le grade de Docteur en Sciences
Spécialité Physique Théorique

par

SYLVAIN REYNAL

Phase Transitions in Long-range Spin Models:
The Power of Generalized Ensembles

le 13 Juin 2005

devant la commission d'examen composée de :

Président :	Prof. David P. Landau	Center for Simulation Physics, The University of Georgia, Athens
Rapporteurs :	Dr. Werner Krauth	DR CNRS, Laboratoire de Physique Statistique, Ecole Normale Supérieure, Paris
	Prof. Olivier Martin	Laboratoire de Physique Théorique et Modèles Statistiques Université Paris Sud, Orsay
Examineur :	Dr. Pascal Viot	DR CNRS, Laboratoire de Physique Théorique des Liquides Université P. et M. Curie, Paris
Directeur de thèse :	Prof. Hung-The Diep	Laboratoire de Physique Théorique et Modélisation Université de Cergy-Pontoise, Cergy-Pontoise

A Stéphanie,

Short abstract/Résumé

English version

This thesis uses generalized ensembles Monte Carlo methods to explore the critical behavior of spin chains with algebraically decaying interactions. The first part of this thesis investigates the phase diagram of a long-range Potts chain using a multicanonical algorithm. A new method based on spinodal points is proposed to detect the order of phase transitions. The boundary between first- and second-order transitions is located with unprecedented accuracy using this method, and a new, unusual finite-size effect is observed. The second part of this thesis formulates a new, versatile multicanonical method that includes cluster updates, considerably extending the range of attainable lattice sizes. The method is shown to be far more accurate than standard multicanonical methods. It is applied to the investigation of finite-size effects at first-order transitions, where strong evidence suggests that the mixed-phase configuration has a fractal dimension depending on the decay parameter of the interaction. Finally, a long-range Ising chain with bimodal random fields is studied. The existence of a tricritical point for slowly decaying interactions is demonstrated.

Version française

Dans cette thèse, nous explorons le comportement critiques de chaînes de spins gouvernées par des interactions à décroissance algébrique. Dans une première partie, nous étudions le diagramme de phase d'une chaîne de Potts en utilisant un algorithme multicanonique. Nous proposons une nouvelle méthode de détection de l'ordre des transitions de phase exploitant les points spinodaux. A l'aide de cette méthode, nous localisons la ligne séparant les transitions du premier et du second ordre avec une précision sans précédent, et mettons en évidence un effet de taille finie inhabituel. Dans une deuxième partie, nous introduisons une nouvelle méthode multicanonique intégrant un algorithme de mise-à-jour collective des spins. Cette méthode, extrêmement souple, étend considérablement l'intervalle de tailles simulables, et s'avère bien plus précise que les méthodes multicanoniques usuelles. Nous appliquons cette méthode à l'étude d'effets de taille finie dans le cadre de transitions du premier ordre : les résultats suggèrent fortement que les configurations correspondant à des phases en coexistence sont caractérisées par une dimen-

Short abstract

sion fractale dépendent du taux de décroissance de l'interaction. Dans un dernier chapitre, nous étudions une chaîne d'Ising régie par des interactions à longue portée en présence de champs aléatoires à distribution bimodale, et prouvons l'existence d'un point tricritique pour des interactions à décroissance lente.

Abstract

This thesis addresses the fundamental question that pertains to spin models governed by algebraically decaying interactions: the influence of the interaction range on their critical behavior. Two of the most challenging issues when investigating these models by means of Monte Carlo methods are (i) the onset of a diversity of phase transitions as the decay parameter is varied, and (ii) a large algorithmic complexity stemming from the huge number of interactions between spins. The purpose of this thesis is to show that methods operating in generalized ensembles present, to date, the most efficient and versatile means of meeting these challenges.

We begin this thesis with a review of the considerable literature that the study of long-range models has produced during the last three decades. Following an overview of the wide variety of Monte Carlo methods available to date, the second part of the thesis presents a broad study of the phase diagram of the q -state Potts chain by means of a local-update version of Berg's multicanonical method. Several improvements tailored to long-range models are proposed. The utility of the method for the simulation of medium-sized long-range models is demonstrated by extensive tests of performance and accuracy over a large range of decay parameters.

Then, a new method for the detection of the order of a phase transition is developed: by following the position of spinodal points, the boundary separating first- and second-order transitions is located with unprecedented accuracy. On the line of inverse square interactions, a very unusual finite-size effect is found: while the transition seems to be first order at finite lattice size, it becomes a continuous transition in the thermodynamic limit. This settles a long-standing controversy surrounding the nature of the phase transition on this line, and suggests that finite-size scaling at first-order transitions is in general highly atypical in long-range models. Beyond this line, the behavior of the model is shown to turn into a short-range one within a narrow window, lending clear support to a long-suggested renormalization group scenario.

With the aim of scrutinizing finite-size scaling effects, the second part of this thesis introduces a novel multicanonical method that drastically extends the range of attainable lattice sizes. The method builds on the ability of cluster algorithms to rapidly reduce temporal correlations, and relies upon the microcanonical temperature to bridge the gap between multicanonical ensembles and random-cluster representations. Owing to its straightforward formulation, it can be used in a va-

Abstract

riety of multicanonical implementations, including the Wang-Landau algorithm and the Transition Matrix method. Systematic tests demonstrate the superiority of this method over standard multicanonical methods: the numerical accuracy is drastically improved, and the reduction of the algorithm complexity to that of a short-range model of the same linear size ensures remarkable scalability.

Application to long-range chains at very large sizes ($\sim 2^{16}$ spins) delivers the best numerical estimates to date. The investigation of finite-size effects shows important finite-size corrections, and suggests that the mixed-phase interface has a fractal dimension depending on the decay parameter of the interaction. An estimate of the correlation length is provided, that is consistent with this picture. The benefits of the method extend far beyond the realm of long-range models: the estimate of the surface tension for the two-dimensional nearest-neighbor Potts model (up to 256×256 spins) matches the exact value to an accuracy never attained with Monte Carlo methods, although modest statistics were used.

A later chapter centers on extensions to disordered models. A long-range Ising chain with bimodal random fields is studied, and the onset of a discontinuous transition is demonstrated at large fields for slowly decaying interactions. This is consistent with a similar scenario reported in the nearest-neighbor version of this model. The limitations of the method for this class of model are examined; improvements and possible combinations with other generalized ensembles are suggested.

Some of the work presented in this thesis has been published or submitted for publication. The references are:

- S. Reynal and H. T. Diep, Reexamination of the long-range Potts model: A multicanonical approach, *Phys. Rev. E* **69**, 026109 (2004), 17 pages.
- S. Reynal and H. T. Diep, Q-state Potts model with power-law decaying interactions: Along the tricritical line, *J. Appl. Phys.* **95**, 6968 (2004), 4 pages.
- S. Reynal and H. T. Diep, Hybrid multicanonical cluster algorithm for efficient simulations of long-range spin models, *Comp. Phys. Comm.* **169**, 243 (2005), 4 pages.
- S. Reynal and H. T. Diep, Simulation of spin models in multicanonical ensemble with collective updates, submitted to *Phys. Rev. Lett.* (cond-mat/0409529).
- S. Reynal and H. T. Diep, Fast flat-histogram method for generalized spin models, submitted to *Phys. Rev. E.* (cond-mat/0504367).

Remerciements/Acknowledgements

Commencer une thèse et la mener à terme est toujours un défi. Le faire à un âge légèrement plus avancé qu'il est d'usage, et en parallèle avec un service d'enseignement de professeur agrégé l'est sans doute bien plus encore. Il est donc évident que, ni le démarrage, ni la conclusion de cette thèse n'aurait pu voir le jour sans l'intervention, le soutien, l'aide, l'appui et l'amitié de nombreuses personnes.

Cette thèse a été dirigée par Hung-The Diep. Compte-tenu de mon parcours atypique, la confiance que m'a accordée Hung *a priori* fût un élément déterminant dans l'amorce de la thèse. Je l'en remercie profondément, comme je lui suis infiniment reconnaissant de m'avoir accordé, pendant ces trois années, une entière autonomie. Son regard critique sur mon travail et son enthousiasme fûrent une aide précieuse durant les inévitables périodes de doute. Notre complicité a fait le reste.

Je remercie très chaleureusement David P. Landau d'avoir accepté de présider mon jury ; sa relecture très attentive du manuscrit et ses nombreuses remarques et conseils fûrent extrêmement enrichissantes. Je suis également très reconnaissant à Pascal Viot d'avoir consenti à juger mon travail. J'exprime enfin ma profonde gratitude à Werner Krauth et Olivier C. Martin pour avoir endossé la lourde tâche de rapporteur.

La décharge partielle d'enseignement dont j'ai bénéficié de la part de l'Ecole Nationale Supérieure de l'Electronique et de ses Applications (au titre du décret 2000-552 du 16 juin 2000), durant les deux premières années de thèse, fût la garantie d'une réelle tranquillité d'esprit. J'exprime ma gratitude envers les collègues qui m'ont fait confiance lors des commissions d'audition, et qui ont compris le sens de ma démarche consistant à effectuer ma thèse dans un laboratoire de recherche extérieur à l'école, en particulier Christian Ngô, qui fût mon rapporteur externe, Didier Demigny, Inbar Fijalkov, Daniel Pasquet, Pierre Pouvil et Arnaud Revel. J'ai une pensée particulière pour Philippe Gaussier, sans qui je n'aurais peut-être jamais songé à aller explorer l'autre rive scientifique de l'Oise.

Les calculs numériques intensifs ont tenu une place centrale dans ce travail de thèse. Ils ont été réalisés sur le calculateur de l'équipe E.T.I.S, équipe commune ENSEA/Université de Cergy-Pontoise associée au CNRS (UMR 8051), et au centre de calcul de l'Université de Cergy-Pontoise. Mes vifs remerciements vont à Laurent Protois, Jean-Christophe Baccon et Yann Costes. Mon obstination à développer en langage Java leur a donné, je n'en doute pas, bien du fil à retordre.

Cette thèse a été effectuée au Laboratoire de Physique Théorique et Modélisa-

Remerciements/Acknowledgements

tion de l'Université de Cergy-Pontoise, UMR 8089 associée au CNRS. Ce fût un plaisir permanent de travailler dans une atmosphère aussi chaleureuse et amicale ; la personnalité, la modestie, les encouragements et le sens de l'humour de chacun des membres du laboratoire y fûrent essentielles. J'ai une pensée amicale et complice pour Damien P. Foster¹, Claire Pinet, Geneviève Rollet et Laura Hernandez, Jean Avan, Philippe Lecheminant, Christophe Oguey, Guy Trambly, Thierry Huillet, Tuong Truong et Rossen Dandoloff. Mes remerciements vont aussi à Sylvie Villemin, dont le pragmatisme et l'efficacité ont su me faire oublier qu'il existe une administration dans la fonction publique.

A aucun instant durant ces trois ans je n'ai eu le sentiment que la qualité de mon enseignement a pâti de mon activité de recherche : cela tient pour beaucoup au soutien inconditionnel et à l'engagement de mes collègues à l'E.N.S.E.A., qui ont tout mis en oeuvre pour compenser mon investissement parfois chancelant. Je pense tout particulièrement à Bruno Darracq, Emanuelle Bourdel, Laurent Monchal et Frédéric Pépin. Je remercie Pascal Goureau de m'avoir concocté un emploi du temps sur mesure². J'ai également une pensée émue pour mon ancien collègue Bertrand David, qui a su me donner l'espoir de tenter, moi aussi, l'aventure. Mijn hartelijke dank tot Sabine Laroche voor haar voortdurend steun, geduld en humor.

This work has also benefited from the invaluable aid of some senior scientists outside the lab. I am particularly indebted to Robert H. Swendsen for suggesting me a way of combining the transition matrix method with one of the methods I introduce in this thesis. I have been delighted to discuss parallels between long-range systems and spin systems on fractal lattices with Pascal Monceau. I wish to thank Thierry Dauxois for suggesting me several references regarding the interpretation of the unusual finite-size effect observed at $\sigma = 1.0$. I have also greatly appreciated the stimulating discussions with Simon Trebst on the optimized ensemble, and with Alexander Hartman regarding disordered models.

My wholehearted thanks go to Kari Foster for her attentive reading and English correction of several parts of the manuscript.

By the time I started this thesis, I was still involved in several collective projects which I had set up long before. I am grateful to the developers of the jPicEdt opensource project for having efficiently taken over, and to my dear friends Eva Pfannes and Sylvain Hartenberg from Ooze-Rotterdam for nicely accepting that I cut down my commitment to our collaboration.

Mes pensées vont enfin à ma petite famille et à mes proches amis qui ont accepté (de force) de me voir si rarement durant ces dernières années, Jérôme et Anne-Laure en particulier, and (now back in the big apple) Christina³ and Stephane.

¹Complice, y compris dans les projets pédagogiques les plus insensés...

²Il me dira, bien entendu, qu'il ne l'a pas fait exprès.

³Special thanks go to you again for proofreading an article published in Phys. Rev. E.

Table of Contents

Introduction and motivations	1
Long-range spin models	1
Monte Carlo simulations in generalized ensembles	3
Organization of the Thesis	5
I Models and methods: review	9
1 Spin models with long-range interactions: review and pending questions	11
1.1 Long-range vs short-range interaction: semantic issues	12
1.2 From Van der Waals forces to neural networks	13
1.3 Spin models governed by long-range interactions: raising the effective dimension	15
1.3.1 Results for ferromagnetic O_n models	16
1.3.2 Antiferromagnetic, frustrated and disordered models	17
1.4 Long-Range Potts model: conjectures and controversies	20
2 From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?	25
2.1 Markov chains and the Metropolis algorithm	26
2.1.1 The Metropolis algorithm	27
2.1.2 Detailed balance for moves that are not micro-reversible	28
2.1.3 Random numbers generators	29
2.2 Dynamic characteristics and statistical efficiency	30
2.2.1 Autocorrelation times	30
2.2.2 Statistical efficiency	32

Table of Contents

2.2.3	Dynamic exponents for correlation times and critical slowing down	33
2.2.4	Scaling of relaxation times and the non-linear dynamic exponent	35
2.3	Cluster algorithms	36
2.3.1	Wolff algorithm	37
2.3.2	Luijten-Blöte algorithm: efficient cluster construction . . .	41
2.3.3	Swendsen-Wang algorithm	44
2.3.4	Extensions to other models	48
2.3.5	Niedermayer’s construct, improved estimators and cluster distributions	49
2.4	Reweighting methods: single- and multi-histograms	51
2.5	Simulation in generalized ensembles	52
2.5.1	From rare-events sampling to flat-histogram algorithms . .	52
2.5.2	Supercritical slowing down and surface tension at first-order transitions	55
2.5.3	Multicanonical ensemble	56
2.5.4	Wang-Landau’s random walker	60
2.5.5	Transition matrices and the Broad Histogram Equation . .	62
2.5.6	Other generalized ensembles	64
2.6	Outlook	65
II A multicanonical study of the long-range Potts model		67
3 A multicanonical algorithm for long-range spin models		69
3.1	Canonical vs multicanonical ensembles and long-range models . .	70
3.2	Multicanonical weights	71
3.3	Recursion scheme	74
3.3.1	Accidental vs non-accidental histogram entries	74
3.3.2	Optimal histogram bin width	75
3.3.3	Unequal spacing of energy levels	75
3.3.4	Convergence criterion	77
3.4	Reweighting procedure	77
3.4.1	Free energies	78
3.5	Predictor choice	78
3.6	Algorithm performance	82

Table of Contents

3.6.1	Tunneling times	82
3.6.2	Autocorrelation times	82
3.6.3	Results	83
3.6.4	Discussion and conclusion	84
4	From mean-field to the <i>tricritical</i> line	87
4.1	Model, existing results and unsettled controversies	88
4.1.1	Mean-field theory	90
4.1.2	Finite range scaling and cluster mean-field approaches	93
4.1.3	Results from real-space renormalization	96
4.2	Estimates of transition temperatures from multicanonical simulations	97
4.2.1	Periodic boundary conditions	98
4.2.2	Methodology	99
4.2.3	Results	104
4.3	Spinodals: a novel approach to assess the order of phase transitions	107
4.3.1	Outline of the method	108
4.3.2	Obtaining reliable information from (variational) free energies	109
4.3.3	Finite size scaling for spinodal temperatures	110
4.3.4	A digression: glean information from transition matrices	111
4.3.5	Application to the phase diagram of the long-range Potts chain	114
4.3.6	Asymptotic $q \rightarrow \infty$ behavior and the $1/r^2$ line	118
4.3.7	Unexpected finite-size scaling behavior at $\sigma = 1.0$: case closed?	122
4.3.8	Outlook: boundary effects and fractal geometry	124
5	From long-range to short-range behavior	129
5.1	Universality and critical properties for long-range Hamiltonians: conjectures	130
5.1.1	Ginzburg-Landau functional	130
5.1.2	Upper critical dimension	131
5.1.3	Momentum-shell renormalization	131
5.1.4	Gaussian model with long-range interactions	132
5.1.5	Perturbative expansion in $\epsilon = 2\sigma - D$	132
5.2	Critical exponents for $q > 2$: multicanonical simulations	133
5.2.1	Objectives	134

Table of Contents

5.2.2	Critical exponents: method	135
5.2.3	Discussion	136
5.3	Long-range to short-range crossover: three decades of controversies	136
5.4	Crossover in the three-state Potts chain: multicanonical simulations vs exact transfer matrix	141
5.4.1	Exact results for the short-range chain: transfer matrix derivation	142
5.4.2	Pertinent crossover indicators and discussion	143
Conclusion of Part II		147
 III Breathing clusters: a novel approach to simulate long-range models in generalized ensembles		 149
 6 Limitations of standard multicanonical methods, and beyond		 151
6.1	Non-optimal dynamic scaling	152
6.2	Algorithm complexity and long-range models	153
6.3	The multibond algorithm and other concurrent approaches	155
6.4	A novel approach: breathing clusters and the microcanonical temperature	157
 7 Fast Flat-Histogram Method for Generalized Spin Models		 159
7.1	Introduction	159
7.2	A method to embed cluster updates in a flat histogram algorithm	163
7.3	Optimization for long-range models	166
7.3.1	Computing the lattice energy through FFT acceleration	166
7.3.2	Efficient cluster construction for long-range interactions decaying with the distance	168
7.4	Numerical tests of algorithm performance	170
7.4.1	Phase space exploration and mean acceptance rates	171
7.4.2	Dynamic properties	173
7.4.3	Overall CPU demand for LR models	179
7.5	Two-dimensional NN Potts model: comparison with exact results	180
7.6	LR Potts chain: error estimates and tests of finite-size scaling	180
7.6.1	Statistical error of the density of states	181
7.6.2	Tests of finite-size scaling: transition temperatures and surface tensions	183

Table of Contents

7.7	Combination with the transition matrix method	188
7.7.1	Efficient estimation of $\beta(E)$ and bootstrapping	188
7.7.2	Efficient predictors for the Wang-Landau algorithm	191
7.8	Conclusion	193
8	Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder	195
8.1	Model and existing results	196
8.2	Gaussian vs bimodal distributions: what's the point?	198
8.3	Mean-field theory	199
8.4	Dimensional argument in the present of long-range interactions	204
8.5	Method	206
8.5.1	Algorithm	206
8.5.2	Influence of periodic boundary conditions	210
8.6	Phase diagram of the long-range RFI model	211
8.7	Discussion and perspectives	215
8.8	Conclusion	217
	General conclusion and developments	221
	Bibliography	223
A	Mean-field theory for the long-range Potts chain	249
B	Transfer matrix derivation of critical exponents for the short-range Potts chain	253
C	Error calculation from Monte Carlo data	257

List of Figures

2.1	Parking lot test over $N = 10^4$ numbers $\{i_n\}$ for the Mersenne Twister generator: the graph is built from points of coordinates (i_n, i_{n+10}) , and displays no stripe that would be characteristic of hidden correlations.	31
2.2	Construction of a Wolff cluster for a nearest-neighbor model, from state (a) and from state (b), where configurations (a) and (b) are related by a cluster flip. The solid lines correspond to activated (or satisfied) bonds ($C_a^i = C_b^i = 7$), and the dotted lines to non-satisfied bonds. In (a), there are two non-satisfied <i>edge</i> -bonds ($C_a^e = 2$, see explanation in text).	38
2.3	Construction of a Wolff cluster for a one-dimensional model with long-range interactions (this applies to any D -dimensional lattice, since the underlying graph is a complete graph in any case). Solid lines denote active bonds, dotted lines represent non-satisfied <i>edge</i> -bonds, while the dashed line corresponds to a non-satisfied inner bond which as such does not explicitly enter the <i>proposed-update</i> probability.	40

List of Figures

- 2.4 Construction of a Wolff cluster for a one-dimensional model with long-range interactions using cumulative probabilities $C_{j_\alpha}(j_{\alpha+1})$. The cluster seed, i.e., the “current” spin to which bonds are added is denoted as 0. The first spin *provisionally* added to the right is at relative position $j_1^R = 2$, i.e., the random number r is such that $C_0(1) \leq r < C_0(2)$. This spin is *actually* not added to the cluster (dashed line) since it does not have the correct sign (i.e., $f(\sigma_0, \sigma_2) = 0$, see explanation in text). A second random number such that $C_2(7) \leq r < C_2(8)$ yields $j_2^R = 8$. The corresponding spin has the correct sign, so that it can be added to the cluster (solid line). Alternatively, the random number r may be transformed to $r' = r[1 - C_0(2)] + C_0(2)$ and j_2^R obtained from $C_0(7) \leq r' < C_0(8)$. Then, the same procedure is performed for spins having negative relative indices: this yields $j_1 = 2$, i.e., $j_1^L = -2$. Finally, one proceeds further with the addition of spins to, e.g., spin $j_1^L = -2$, which becomes the new “current” spin 0 (dotted line). 42
- 2.5 Tree-like (left) and flattened (right) representations of a complete graph \mathcal{G} for a two-state Potts model (i.e., the Ising model) with long-range J_{ij} interactions and $L = 6$ spins. Solid and dotted lines correspond to active ($b = 1$) and inactive ($b = 0$) bonds, respectively. The corresponding bond configuration engenders three connected components, each being associated with a different cluster in the Swendsen-Wang algorithm: $\{0, 3\}$, $\{1, 2, 5\}$ and $\{4\}$ 45
- 2.6 An illustration of a bond update in the flattened representation of the complete graph \mathcal{G} , for the Swendsen-Wang algorithm applied to a two-state Potts model with long-range J_{ij} interactions. The vertical dashed arrow shows the order in which bonds are considered for activation, starting from an empty bond set (i.e., all bonds inactive). The bond between spins 2 and 4 may be left inactive with probability $e^{-\beta J_{24}}$ (a) or activated with probability $1 - e^{-\beta J_{24}}$ (b), see Eq. (2.10). The procedure continues until every bond has been considered for activation. 47

List of Figures

2.7	Histogram of the energy filled with the samples produced by a canonical simulation at inverse temperature β (solid line), and what the same histogram should <i>look like</i> at inverse temperature β_0 . Both histograms overlap only poorly, and the resulting accuracy of reweighted averages at β_0 is thus very low.	53
2.8	A schematic illustration of the surface tension phenomenon that results from phase coexistence at a first-order transition in, e.g., a two-dimensional Potts model with $q > 4$. The two types of hatches represent the ordered and disordered phases. The dimension of the interface is $D - 1$ for a D -dimensional model with nearest-neighbor interactions. The size of a droplet (indicated with arrows) is roughly given by the correlation length of the corresponding phase.	55
2.9	Canonical distribution of the energy at a first-order transition for various lattice sizes L (data from multicanonical simulations of a three-state long-range Potts chain with $1/r^{1.5}$ interactions, see Chap. 7).	56
3.1	Flat energy distribution engendered by the multicanonical algorithm (multicanonical histogram) vs the reweighted histogram at the “equal heights” transition temperature (denoted as “canonical” histogram in the figure). The strong suppressing of mixed-phase configurations that would occur in a canonical simulation, is clearly visible: the dash line illustrates (in a very sketchy manner) what would be the dynamics of a canonical simulation in this case. The model parameters are: $q = 5, \sigma = 0.3, L = 400$. E/L designates the energy per spin.	72
3.2	Lowest energy levels for $q = 5, \sigma = 0.5, N = 400$, computed by sorting energy samples from a long simulation run. Each level is drawn as a horizontal line.	76
3.3	Illustration of the reweighting procedure for (variational) free energies. The graph sketches the variational free energy with respect to the energy, i.e. the logarithm of the reweighted histogram of the energy (up to a change of sign), for three distinct temperatures: T_1 and T_2 correspond to limits of metastability; T_c is one possible definition of the (finite-size) transition temperature.	79

List of Figures

3.4	Dots indicate the initial guesses $\bar{H}_{mu}(E)$ that were fed into the iteration scheme at $L = 400$, $q = 5$, and $\sigma = 0.3(\diamond), 0.5(+), 0.9(\square)$. Each initial guess is computed using Eq. (3.6), i.e., by scaling a true estimate obtained at $L = 200$. Solid lines show true estimates $H_{mu}(E)$ as obtained after the whole iteration scheme at $L = 400$ converged. The straight dashed line sketches the original Hamiltonian, i.e., $H_{mu}(E) = E$	81
3.5	Energy histogram as computed after indicated runs, for $q = 5, \sigma = 0.9, L = 400$ spins, using Eq. (3.6) to compute the initial effective Hamiltonian $\bar{H}_{mu}(E)$ from a previous run at $L = 200$ spins. Labeling on y axis indicates normalized probabilities.	81
3.6	Integrated autocorrelation time τ vs lattice size L for $q = 7$ and $\sigma = 0.2, 0.4, 0.6, 0.8$. Dynamic exponents computed from a fit to L^z are $z = 1.09(1), 1.15(1), 1.38(1), 1.55(1)$, respectively.	83
3.7	Tunneling time τ_{tun} vs lattice size L for $q = 7$ and $\sigma = 0.2, 0.4, 0.6$, and 0.8 . Dynamic exponents computed from a fit to L^z are $z = 1.25(1), 1.30(2), 1.37(1), 1.53(1)$, respectively.	84
4.1	Free energy per spin in the mean-field approximation for the nine-state Potts model in zero external field, and with $1/r^{3/2}$ interactions. T_1 and T_2 refer to the temperatures of metastability, while T_c designates the transition temperature corresponding to the two minima having the same magnitude. ΔF is the surface tension (the <i>reduced</i> surface tension is usually defined from $\Delta F/L^{D-1}$ for nearest-neighbor models).	92
4.2	Spinodal temperatures T_1 and T_2 , along with the transition temperature kT_c , as a function of q in the mean-field approximation. Shown are quantities normalized with respect to $\zeta(1 + \sigma)$	93
4.3	Transfer matrix method for finite-range interactions: the zig-zag scheme brings the problem back to a conventional transfer matrix problem, with each column interacting with its nearest-neighbor only.	94
4.4	Sketchy illustration of the two-gaussian phenomenological theory of finite-size scaling at a first-order transition. The dashed curve represents the <i>true</i> distribution.	101

List of Figures

4.5	Dotted curve: reweighted histogram of the energy for the three-state Potts model with $1/r^{1.2}$ interactions ($L = 1000$), at a temperature corresponding to the equal-weights condition ($kT = 3.9599(2)$). The superimposed solid curves indicate a non-linear least-square fit to the two-gaussian theory.	102
4.6	Critical couplings for $q = 3, 5, 7, 9$, from top to bottom (solid lines). Mean-field predictions are shown for comparison (dotted lines).	105
4.7	Free energy $F_m(kT, m)$ for $q = 9, \sigma = 0.3, L = 400$ (solid line), along with the mean-field prediction (dashed line) as given by Eq. (4.4). 107	107
4.8	Graphs of $F_e(kT, E) = -\ln N_{rw}(kT, E)$ for $q = 5, \sigma = 0.3, N = 400$, at four characteristic temperatures: T_1, T_2, T_{eqh} , and $T_{eqw} = T_c$ denote the temperatures of the two metastable states, and the temperature of equal peak heights, and that of equal peak weights, respectively. E/L stands for the energy per spin.	109
4.9	Average energy per spin for $q = 3, \sigma = 0.2, L = 400$, computed over both phases (solid line), ordered phase only (lower dashed line), and disordered phase only (upper dashed line). Vertical dotted lines indicate the four characteristic temperatures: from left to right, lower limit of metastability (kT_1), transition temperatures (equal heights kT_{eqh} , then equal weights kT_{eqw}), and upper limit of metastability (kT_2).	110
4.10	An example of calculation of spinodal points from the information provided by transition matrices. Model parameters are $q = 6, \sigma = 0.9, L = 256$, and the simulation was performed with the breathing cluster method presented in Chap. 6. Dotted and dashed lines refer to the first and the second spinodal points, respectively (the error on the spinodal temperatures lies around 5×10^{-5} and is therefore not indicated on the graph). The inset shows the free energy $F_e(kT, E)$ at the second spinodal temperature kT_2 , before filtering was applied, and the large vertical error bars indicate the error on the spinodal temperatures obtained from $F_e(kT, E)$	113

List of Figures

4.11 Spinodal curve for $0.3 \leq \sigma \leq 0.8$ ($q = 5$). The limits of metastability T_1 and T_2 (i.e., defining spinodal temperatures) are indicated as triangles and diamonds, respectively. The transition temperature T_c is reminded as dotted line. Errors are smaller than the size of symbols.	115
4.12 Difference between temperatures of metastability $dkT_m = kT_2 - kT_1$ vs $1/\sigma$ for $q = 7$ (circles connected by solid lines). Errors are smaller than the size of symbols. Mean-field prediction are shown for comparison (dashed line).	115
4.13 T_2/T_c vs $1/\sigma$ for $q = 3, 5, 7, 9$. Solid lines indicate polynomial fits. Dotted lines are guides to the eyes. Error bars are smaller than the size of dots, except where explicitly indicated.	116
4.14 Phase diagram computed using finite-size scaling properties of spinodal points, for $\sigma < 1.0$. Dotted lines are guides to the eyes. . . .	117
4.15 Phase diagram of Fig. 4.14, along with the disputed line of inverse square interactions. The open circle shows where the $\sigma_c(q)$ line was claimed to terminate in [19]. According to Cardy's scenario [72], on the contrary, the <i>whole</i> line is the locus of topological transitions; grayed disks refer to a numerical study by Luijten and Messingfeld [235] reporting the onset of a topological transition for $q = 2$ and $q = 3$	121
4.16 Linear fit of finite size temperatures vs $1/L$ for $q = 9, \sigma = 1.0$. Dotted, solid, and dashed lines correspond to kT_1 , kT_{eqh} , and kT_2 , respectively. Error bars are smaller than the size of symbols, except where explicitly indicated. In the limit $L \rightarrow \infty$, the difference between temperatures of metastability tends to 0.0012. Within error bars, the transition is thus clearly not of the first order.	122
4.17 Metastability temperature differences $kT_2(L) - kT_1(L)$ on the line $\sigma = 1.0$ for $q = 8$ and $q = 9$, along with the corresponding linear fits (solid lines). For the sake of clarity, the size of the error for each lattice size is shown as a bar; the error shown corresponds to the largest of both errors computed for $q = 8$ and $q = 9$	123
4.18 Reduced surface tension $F_m(kT, m)/L$ for $q = 9$ and various values of σ (top), and a detailed view for $\sigma = 1.0$ (bottom).	125

List of Figures

4.19	Extrapolated value of the reduced surface tension $\Delta F_m(L)/L$ in the thermodynamic limit for $q = 9$. A linear scaling law of the form $\Delta F_m(L)/L \sim a + b/L$ was assumed.	126
5.1	The graph shows the peaks of the cross-cumulants of the magnetization, V_1^{min} and V_2^{min} with respect to the lattice size L on a log-log scale, for $L = 50, 100, 150, 200, 400$. The model parameters are $q = 5$ and $\sigma = 0.9$. The power-law fit yields $1/\nu = 0.6675(1)$ and $1/\nu = 0.669(2)$ for V_1 and V_2 , respectively. Errors were computed from a bootstrap method; they are smaller than the size of symbols.	137
5.2	Maximum of the susceptibility χ^{max} plotted against the lattice size L , for $L = 50, 100, 150, 200, 400$, and the following model parameter: $q = 5, \sigma = 0.9$. A fit of the data points to $\chi^{max} \propto L^{\gamma/\nu}$ yielded $\gamma/\nu = 0.940(4)$. Errors are smaller than the size of symbols.	137
5.3	Specific heat for various lattice sizes and $\sigma = 1.0, 1.1, 1.2, 1.7$, along with the exact nearest-neighbor case, from right to left. Data for other values of σ have been omitted in order to preserve the clarity of the figure. Solid, dashed, dotted and long-dashed styles refer to $L = 50, 100, 200$, and 400 respectively, except for the short-range case where they refer to $L = 5, 10, 100, 200$	144
5.4	Maximum of the specific heat vs inverse lattice size for $\sigma = 1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0, 3.0$, and 4.0 , from top to bottom. The solid line is a reminder for the (exact) short-range case in the thermodynamic limit. Other lines are guides to the eyes.	145
5.5	Magnetization vs kT for $\sigma = 1.0, 1.1, 1.2$, and 1.7 from right to left. Solid, dashed, dotted and long-dashed styles refer to $L = 50, 100, 200$, and 400 respectively.	146
7.1	Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts chain with $\sigma = 0.7$, and $L = 1024$ spins. The dashed line shows the estimated inverse microcanonical temperature. The vertical dotted lines indicate the position of the histogram peaks corresponding to the ordered and disordered phases.	171

List of Figures

7.2	Tunneling times for the long-range Potts chain with $q = 3$, $\sigma = 0.4$ (dashed lines) and 0.6 (dotted lines), and $q = 6$, $\sigma = 0.7$ (solid lines). Triangles refer to the SSU implementation, while squares indicates estimates for our method (CU). Dynamic exponents z were determined from a fit to the power law $\tau_e \sim L^z$	173
7.3	Effective autocorrelation time τ_{eff} for $q = 6$, $\sigma = 0.9$ and $L = 512$ with (a) cluster updates (b) single-spin updates. The effective transition temperature defined from the peak of the specific heat is $T_c(C_v) = 0.7163(2)$	175
7.4	Fit of effective autocorrelation times τ_{eff} to the power law $\tau_{\text{eff}} \propto L^\alpha$ for the six-state Potts chain ($\sigma = 0.9$ and $L = 512$) with (a) cluster updates (b) single-spin updates.	176
7.5	Plot of $\chi^2(t)/r$ for the six-state Potts chain ($\sigma = 0.9$, $L = 512$) using (a) cluster updates and (b) single-spin updates. The regression was carried out over a histogram containing 20 bins populated from 1000 runs, all starting in ground state configuration but with distinct random seeds.	178
7.6	Fit of equilibrium times to the power law $\tau_{eq} \propto L^\alpha$ for the six-state Potts chain ($\sigma = 0.9$). (a) cluster updates and (b) single-spin updates.	178
7.7	CPU time per MCS and per spin for the long-range Potts chain. Triangles indicate typical CPU times for the local-update algorithm (SSU), irrespective of q and σ . Squares refer to our algorithm (CU) with LR specific optimizations included; for $q = 3$ and $q = 6$, estimates were determined by averaging over the indicated σ values.	179
7.8	Statistical error on the density of states of the six-state Potts chain for two distinct modification factors $\ln f$ of the Wang-Landau algorithm. The statistical errors were obtained from 12 independent runs. The parameters of the model are $\sigma = 0.9$ and $L = 512$. (a) and (b) correspond to our method and the local-update algorithm respectively.	182
7.9	Graph of the specific heat for the six-state Potts chain ($\sigma = 0.9$ and $L = 512$) obtained directly from the final estimate of the density of states with (a) our method and (b) the local-update algorithm. The inset shows the relative error $\text{err}(C_v)/C_v$	182
7.10	Specific heat for the three-state Potts chain with $\sigma = 0.5$ as obtained with our method.	183

List of Figures

7.11	Peak of the cumulant of the energy $U_L = \langle E^4 \rangle / \langle E^2 \rangle^2$ as a function of the lattice size for the three-state LR Potts chain with $\sigma = 0.2$ and $\sigma = 0.5$. The inset shows a magnification near the origin for $\sigma = 0.5$	185
7.12	Graph of the free energy $F(T, E) = -\ln N(T, E)$ for the three-state LR Potts chain with $\sigma = 0.5$. $N(T, E)$ is the reweighted histogram at a transition temperature defined by equal peak heights. For the four lattice sizes shown here, lattice configurations corresponding to phase coexistence are suppressed by a factor ranging from 0.1 to 10^{-6} with respect to pure phase configurations; for the three largest sizes, we note that a canonical simulation is clearly intractable. . .	186
7.13	Fit of the interface free energy ΔF to a power law of the lattice size for the three-state LR Potts chain with $\sigma = 0.5$. All estimates of ΔF were obtained with our method.	186
7.14	Fit of the interface free energy ΔF to a power law of the lattice size for the six-state LR Potts chain with $\sigma = 0.5, 0.7,$ and 0.9 . All estimates of ΔF were obtained with our method.	187
7.15	Microcanonical inverse temperature $\beta(E) = dS(E)/dE$ computed from the estimated density of states using a spline interpolation, for the three-state long-range chain with $\sigma = 0.4, 0.5,$ and 0.6 from bottom to top.	189
7.16	Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts chain with $\sigma = 0.5,$ and $L = 512$ spins (strong first-order regime) for three different estimates of $\beta(E)$. (a) best estimate, as given by the ultimate iteration of the Wang-Landau algorithm; (b) $\beta(E)$ scaled by 0.9; (c) $\beta(E)$ scaled by 1.1. . .	190
7.17	Symbols show the microcanonical inverse temperature $\beta(E)$ computed from the transition matrix accumulated over 2000 MCS, for the six-state LR model ($\sigma = 0.5$) containing 512 spins. The estimate obtained from an interpolation scheme after the ultimate iteration is shown as a solid line for comparison.	190

List of Figures

7.18	The graph shows the number of MCS needed to reach the ground-state (dashed horizontal line) of the six-state Potts chain ($\sigma = 0.5$ and $L = 512$) for an initially unknown density of states, using three distinct schemes: (a) and (b) predictor based on $\beta(E)$, local- and collective-update algorithms respectively; (c) no predictor ($S(E) = 0, \forall E$), local-update algorithm.	192
8.1	Free energy $F(kT, m, h)$ at $T = 0$ for several values of the field amplitude. The ferromagnetic phase (F) is stable up to $h/h_c = 1$ (this defines the critical field amplitude). The spin-glass phase (SG) is metastable at small field amplitudes. This feature is a peculiarity of the bimodal distribution: for a trimodal distribution, this phase becomes unstable at sufficiently low fields.	201
8.2	Free energy $F(kT, m, h)$ for $h/h_c = 0.3$. The metastable state at $m = 0$ corresponds to the spin glass phase, with spins aligned with their local field, and vanishes above $T_a = 0.205 < T_c$. Temperatures are normalized with respect to the transition temperature at zero field.	202
8.3	Free energy $F(kT, m, h)$ for $h/h_c = 0.9$. The metastable state at $m = 0$ exists from $T = 0$ up to the transition temperature, where it becomes stable. Temperatures are normalized with respect to the transition temperature at zero field.	203
8.4	Illustration of Imry-Ma dimensional argument for a long-range chain with a bimodal distribution of fields: the stability of an elementary excitation consisting of R contiguous spins is monitored against the decay parameter of the long-range interaction.	206
8.5	Total CPU time needed to perform a Monte Carlo simulation delivering the <i>same</i> number of uncorrelated measurements, with respect to the lattice size. Single-spin update: dashed line; cluster-update: solid line. Units on the vertical axis are arbitrary.	210
8.6	Graph of the critical field $h_c(L)$ with respect to the lattice size for $\sigma = 0.1$ and two choices of periodic boundary conditions: infinite image periodic boundary conditions (IIPBC), and first-image periodic boundary conditions (FIPBC).	211
8.7	Fit of finite-size temperatures $T_{c,\chi}(L)$ to the power law $T_{c,\chi}(L) - T_{c,\chi}(\infty) \propto L^{-1/\chi}$, for $\sigma = 0.1$ and $h = 1.0$	212

List of Figures

8.8 Phase diagram of the random-field Ising chain with $1/r^{1+\sigma}$ interactions, for $\sigma = 0.1$ and 0.4 . Open symbols indicate a first-order transition, while filled symbols refer to a second-order transition (lines are guides to the eyes and errors are smaller than the size of symbols except were explicitly shown). The thick, dotted and solid lines show the mean-field predictions in the first- and second- order regime, respectively. The thick dashed line shows the boundary between both regime in the mean-field approximation. 213

8.9 Fit of the temperatures of metastability $T_1(L)$ and $T_2(L)$ to a power law of the lattice size, for $\sigma = 0.1$ and $h = 10.0$. In spite of the strong dispersion of the data points, the transition is clearly of the first order. 214

8.10 Top: Joint distribution of the (normalized) random-field part of the interaction (i.e., $(1/L) \sum_i \sigma_i h_i$) with respect to the total energy per spin E/L , for $\sigma = 0.1$, $h = 10$ and $L = 512$. Bottom: Reweighted histogram of the energy at the equal-height transition temperature, for the same sample. 216

8.11 Joint distribution of the (normalized) random-field part of the interaction with respect to the total energy per spin E/L , for $\sigma = 0.1$, $h = 10.2$ and $L = 256$ 218

List of Tables

4.1	Estimates of the critical temperature in the first- and second-order regimes (the latter is indicated by an asterisk): MF, mean-field predictions; χ , using peaks of the susceptibility; $U^{(4)}$ using crossing points of Binder cumulants of the magnetization; eqw,eqh, using the free energy, where T_c corresponds to equal peak weights and heights, respectively; Ref. [131], Monte Carlo study based on multihistogramming and the Luijten-Blöte cluster algorithm ($q = 3$) and a standard metropolis algorithm ($q = 5$), where (a) refers to $1/K_e(\Delta F)$, and (b) to $1/K_e(U^{(4)})$; Ref. [251]), cluster mean-field method combined with an extrapolation technique based on the VBS (Vanden Broeck and Schwartz) algorithm; Ref. [130]), transfer matrix method combined with finite-range scaling.	106
5.1	Critical exponents in the second-order regime $\sigma > \sigma_c(q)$, and $q = 3, 4, 5$. Shown for comparison are results from Ref. [130] (transfer matrix method), from Ref. [325] (largest cluster probability, with sizes up to $L = 20000$ spins), and from Ref. [19] using a Monte Carlo single-histogram method.	138

List of Tables

5.2	Estimates of the critical couplings in the second-order regime. MF, mean-field predictions; χ , using peaks of the susceptibility; $U^{(4)}$ using crossing points of Binder cumulants of the magnetization; Ref. [325], Monte Carlo study relying on the largest cluster probability, with sizes up to $L = 20000$ spins, Ref. [131], Monte Carlo study based on histogramming and the Luijten-Blöte cluster algorithm ($q = 3$) and a standard metropolis algorithm ($q = 5$); Ref. [251]), cluster mean-field method combined with an extrapolation technique based on the VBS (Vanden Broeck and Schwartz) algorithm; Ref. [130]), transfer matrix method combined with finite-range scaling.	138
7.1	Dynamic exponents z for the q -state Potts chain with power-law decaying interactions (a) and its two-dimensional counterpart with nearest-neighbor interactions (b). $z(SSU)$ and $z(CU)$ refer to single-spin and cluster updates respectively, while $z_{\mu Bo}$ and $z_{\mu Clus}$ make reference to the multibond method [183] and Rummukainen's multi-microcanonical cluster method [291] applied to the NN model. . .	173
7.2	Effective autocorrelation times at the transition temperature defined from the location of the peak of the specific heat, for the six-state LR Potts chain with $\sigma = 0.9$	176
7.3	Equilibrium times for the six-state LR Potts chain with $\sigma = 0.9$ obtained by monitoring the graph of $\chi^2(t)/r$	178
7.4	Estimates of peaks of the specific heat C_V and the susceptibility χ , and corresponding effective transition temperatures for the three-state LR Potts chain with $\sigma = 0.5$. Error calculations were carried out by means of the jackknife method applied to a single production run. The number of MCS per production run is the same for both methods, yet varies between 10^6 and 10^7 from the smaller to the larger lattice sizes.	184

List of Tables

- 8.1 Effective autocorrelation times at the transition temperature for $\sigma = 0.4$ and $h/h_c = 0.3$, and $L = 128, 256$ and 512 . (SSU) refers to the single-spin update algorithm. (I) indicate estimates obtained with algorithm I for several scaling factors s applied to the microcanonical temperature $\beta(E)$ entering bond probabilities. (II) refers to the ghost-spin algorithm. A and B indicate the acceptance rate and the fraction of the lattice occupied by the largest cluster, respectively. . 209

Introduction and motivations

This thesis narrates two distinct, and yet tightly related stories: the investigation, with the concepts and tools of statistical physics, of the critical behavior of a class of spin models governed by long-range interactions, and the development of new methods dedicated to their investigation. The former focuses on the q -state Potts model, and, to a lesser extent, on the Random Field Ising model. The latter revolves around *generalized ensembles* methods. This introduction is aimed at outlining the scientific motivations behind this thesis, and in particular, at explaining — although partially⁴ — (i) why long-range spin models, and specifically the two models mentioned above, are appealing from the viewpoint of equilibrium statistical physics, and (ii) why Monte Carlo simulations in generalized ensembles might be an efficient approach to the exploration of these models.

Long-range spin models

The model-systems investigated in this thesis are discrete spin models living on a lattice. For instance, the q -state Potts model I will be considering from Chap. 4 to Chap. 7 is described by the following Hamiltonian, $H[\{\sigma_i\}] = -\sum_{i \neq j} J_{ij} \delta(\sigma_i, \sigma_j)$, where the σ_i 's represent discrete spin variables taking on integer values between 1 and q , and the sum runs over every pair of spins. J_{ij} is a long-range coupling constant of the form $|i - j|^{-D-\sigma}$, i.e., decaying algebraically with the interspin distance, that favors ferromagnetic ordering at low temperature. The random-field Ising chain that will be studied in Chap. 8 is governed by a similar Hamiltonian, $H[\{\sigma_i\}] = -\sum_{i \neq j} J_{ij} \sigma_i \sigma_j - \sum_l h_l \sigma_l$, where $\sigma_i = \pm 1$, and the second term represents the interaction between spins and external random fields $\{h_i\}$. The equilibrium thermodynamic properties of these *classical* spin models⁵ at a given temperature T are described by the usual Boltzmann distribution, the probability of occurrence of a given configuration of spins $[\{\sigma_i\}]$ being proportional to $e^{-H[\{\sigma_i\}]/kT}$. An approximate treatment through mean-field theory will be provided for both models in Chap. 4 and Chap. 8, respectively.

Though seemingly simple in their construction, these models are in fact en-

⁴Chap. 1 and 2 provide more detailed material in this respect.

⁵Quantum fluctuations are assumed to be negligible with respect to thermal fluctuations, and only thermal phase transitions are considered in this work.

Introduction and motivations

dowed with a built-in complexity that stems from the fundamental role played by the decay parameter σ in shaping their critical behavior, akin to an effective dimension. Their phase diagram is in general richer than that of their nearest-neighbor counterpart: as the effective dimension is varied, from the geometric dimension (where the behavior is that of a nearest-neighbor model) to infinity (where the behavior is mean-field-like), they go through a variety of universality classes [225]. In this respect, they may be regarded as a powerful paradigm for studying the influence of dimensionality in the physics of critical phenomena: crossovers between universality classes and tests of renormalization group ϵ -expansions, for instance, have been extensively studied in the past [40].

Yet long-range models are also of fundamental interest in the microscopic modeling of a variety of systems, from the Kondo effect in the physics of high temperature superconductivity [11] to the fast-growing technological field of ultrathin magnetic films [45, 66], from autoassociative memories [268] to small-world networks [193]. This means that investigating these models is not merely a matter of academic interest.

Inside the realm of long-range spin models, the q -state Potts model is perhaps the most exciting, owing to its parameter space comprising three parameters: the temperature, the number of states q , and the decay parameter σ . In fact, its appealing character might be already expected from the rich behavior of its nearest-neighbor counterpart. This last model undergoes a phase transition that turns from a continuous into a first-order one as the number of states q is increased beyond a threshold value $q_c(d)$ depending on the dimensionality of the lattice. Therefore, the connection between the decay parameter and the effective dimension suggests that a similar behavior might occur in the long-range version.

The Potts model is in fact one of the less studied long-range models, with only a rough estimate of its phase diagram available; it is also one of the most intriguing, with numerous controversies and conjectures surrounding it, e.g., the nature of the phase transition on the line of inverse square interactions (topological vs first-order), the exact position of the boundary separating first- and second-order transitions, or the crossover from the long-range to the short-range regimes. Finally, and as opposed to its nearest-neighbor counterpart, no rigorous treatment (e.g., based on duality) has been proposed thus far. It is one of the aims of this thesis to investigate this model from a new perspective.

The Random Field Ising chain is the second model that is studied in this thesis, though to a far lesser extent than the previous model. Here again, the parameter space is three-dimensional, owing to the presence of disorder. While very substantial literature has been produced on the *nearest-neighbor* version of this model, its long-range version has been much less investigated [52, 343]. One of the most intriguing questions regarding the nearest-neighbor model with a *bimodal* field distribution is the possible onset of a first-order transition for sufficiently large fields [2]. This has been predicted in the mean-field case, yet the issue seems to be somewhat challenging in the three-dimensional nearest-neighbor model [154]: one might thus expect that in a long-range version of this model with interactions de-

Introduction and motivations

caying at a sufficiently slow pace (i.e., low decay parameter σ), the onset of such a transition, if at all, may be more clearly visible. The aim of Chap. 8 is to examine this feature.

Overall, and where *both* models are concerned, the fundamental issues examined in this thesis revolve around the two following questions:

- In what respect is the critical behavior of long-range models, and in particular the order of their phase transitions, influenced by the decay parameter of the interaction? What is the nature of the crossover effects that occur between the regimes exhibited?
- To what extent may a long-range model serve as a guide in the investigation of its nearest-neighbor counterpart, where the effect of dimensionality is concerned?

A significant part of this thesis has been devoted to obtaining a comprehensive view of the current state of knowledge regarding long-range models, and to keep up with the steady production of new results. This is presented in Chap. 1. A background in condensed matter physics is assumed there. The phase diagram of the q -state Potts chain ($D = 1$) is investigated in Chap. 4 and 5. An atypical finite-size effect will be found out on the line of inverse square interactions ($\sigma = 1$), that will trigger a reexamination at much larger lattice sizes in Chap. 7. Finally, the phase diagram of the long-range Random Field Ising chain is explored in Chap. 8.

Monte Carlo simulations in generalized ensembles

There are, broadly speaking, two means of studying the thermodynamic properties of a model from the viewpoint of equilibrium statistical physics: *analytically*, using rigorous methods, e.g., contour expansions or inequalities on correlation functions, or approximation-based methods, e.g., series expansions, transfer matrices⁶ or renormalization group methods; or *numerically*, by means of Monte Carlo (i.e., stochastic) or Molecular Dynamics (i.e., deterministic) methods. The latter have turned into a powerful tool to investigate complex models where no rigorous treatment is available. Monte Carlo methods, and specifically, their *generalized ensembles* flavor, play a central role in this thesis, and will be reviewed in Chap. 2. Other methods, specifically analytical methods dedicated to long-range models, are outlined in the core of the manuscript when deemed necessary.

Basically, Monte Carlo methods (inside and outside the realm of physics) are a tool to compute a multi-dimensional integral using random sampling; in statistical physics, the role of the multi-dimensional integral is played by the *partition function* of the model and the statistical moments that are generated from it by differentiation. The most common implementation of this method in statistical physics is *importance sampling* Monte Carlo, whereby a Markov chain generates a

⁶Methods based on transfer matrices may be exact or approximate.

Introduction and motivations

sequence of (properly distributed) configurations of the system that is a tiny subset of all microscopic configurations. The first implementation of this method was introduced by Metropolis et al. [246] in 1953: configurations are sampled according to the Boltzmann distribution $e^{-E/kT}$ (this method is therefore sometimes referred to as “canonical sampling”), and statistical averages of thermodynamic quantities at temperature T are directly computed from the sequence of configurations.

The achievements of this algorithm span more than five decades, and it has long been regarded as the paradigm for Monte Carlo simulations in statistical physics. Fifteen years of Monte Carlo simulations of long-range spin models based on this algorithm (or close variants thereof) attest that, in this field also, strategies did not depart from custom. Yet this method faces some severe drawbacks, in particular when simulating systems at first-order transitions or with complicated energy landscapes: both behaviors are present in the models explored in this thesis, and suggest that the issue be met from a different perspective using new algorithms.

New methods that seem promising candidates in this respect are methods operating in *generalized ensembles* [25], and in particular, in the *multicanonical ensemble*. Introduced in the early 90’s, these rely on *non-Boltzmann* sampling, and simulate the model over a large range of energy with no explicit reference to the temperature. They offer several benefits with respect to the Metropolis algorithm: statistical averages of thermodynamic quantities can be obtained from a single simulation, by resorting to a so-called *reweighting* procedure; the presence of *local minima* in the free energy landscape, which leads to repetitive dynamics and thermalization problems when using the Metropolis algorithm, is efficiently tackled in generalized ensembles. Therefore, investigating long-range spin models by means of these new methods seems an exciting way of meeting the various controversies that surround these models.

A substantial part of this thesis was devoted to learning how to use the various Monte Carlo methods that have been made available to date, including (i) the multicanonical method, the Wang-Landau algorithm, the transition matrix method and methods operating in other generalized ensembles, and (ii) a large class of cluster algorithms (Chap. 2). The former, and specifically the multicanonical method, form the core of the numerical methods used in this thesis; the latter will turn out to be an essential ingredient of the breathing cluster method developed in Chap. 7.

On the *methods* side, one of my first items of business (Chap. 3) will be to test the usability of these methods in the context of long-range models, e.g., their performance, reliability and ease of implementation. In this regard, the diversity of phase transitions exhibited by the q -state Potts model will make it a perfect test case. The multicanonical method — in its pristine formulation by Berg [30] — will play a central role here, and the phase diagram of the long-range Potts chain (Part II) will be investigated by means of an improved version of this method tailored to long-range models.

In spite of its efficiency at medium lattice sizes, the multicanonical method will turn out to exhibit severe scalability issues. This pertains, among other things, to the large *algorithmic complexity* brought about by the long-range potential (with

Introduction and motivations

respect to nearest-neighbor models): since each particle interacts with each other, a huge number of interparticle interactions must be taken into account at each Monte Carlo step. The need to investigate finite-size effects at large lattice sizes will prompt the development of a new method (Chap. 4), that considerably extends the range of attainable lattice sizes by lumping together the benefits of generalized ensembles and those of cluster algorithms.

Beyond the design of new simulation methods (or the improvement of existing methods), this thesis will also introduce a new method to detect the order of phase transitions. This will prove particularly stringent when investigating the position of the boundary separating first- and second-order transitions. There, it will be shown that traditional techniques fail, owing to the weakening of the first-order transition as the boundary is approached. The proposed method will efficiently capitalize on the information provided by the multicanonical method.

From the viewpoint of the methods, this thesis is therefore aimed at answering the two following questions:

- To what extent do generalized ensembles provide better, more efficient and reliable methods for the investigation of phase transitions in long-range spin models? As a corollary, can they help in settling (some of) the various controversies that surround them?
- In what respect can long-range models constitute an ideal case for the testing of new numerical methods?

Organization of the Thesis

Both the *models* and the *methods* are central to this thesis, and will thus be treated on an equal footing. As a result, both stories are markedly interwoven, and will be presented according to this picture.

Part I of this thesis (Chap. 1 and 2) is devoted to a review of the models and the numerical methods that will be considered in the remainder of this work. It is aimed at providing the necessary background to understand the work presented in Part II and III, yet also at shedding deeper light on the motivations behind this thesis.

Models governed by long-range interactions are reviewed in Chap. 1: applications of these models to various fields of physics are considered, but the emphasis is markedly on lattice spin models, including — but not restricted to — ferromagnetic models. In particular, disordered and frustrated models are also considered, inasmuch as several pending questions regarding these models are common to most long-range spin models. The key point in this review will be the notion of *effective dimension*, and the role played by the decay parameter of the long-range interaction in shaping the critical properties of the models. As far as the long-range Potts model is concerned, this chapter is intentionally succinct: most results will be re-examined in the second part of this thesis, and further detailed explanations, when

Introduction and motivations

needed, will be provided in the corresponding chapters.

Chap. 2 is dedicated to stochastic (Monte Carlo) methods. This Chapter concentrates not only on simulations in generalized ensembles (including the multicanonical ensemble), but also on cluster algorithms: most methods presented in this chapter will be referred to, at some point or another, in this thesis, either for the mere purpose of comparisons, or because they are directly used in numerical simulations, or because they comprise a large proportion of the ingredients of a novel method.

Part II (Chap. 3 to 5) concerns itself with the multicanonical simulation of a one-dimensional Potts model with power-law decaying interactions.

Chap. 3 centers on the implementation of the multicanonical algorithm in the context of long-range models: it comprises an introduction to the method, a second part dedicated to the improvements that must be carried out in order to make the multicanonical recursion scheme suitable for long-range models, and a study of the performance of the algorithms in terms of dynamic exponents.

Chap. 4 focuses on the location of the boundary separating the first- and the second-order regime of the model, and on the estimation of critical couplings in a large part of the phase diagram. Results from other methods, e.g., renormalization group or transfer matrices, are extensively reviewed there. A novel method to detect the order of phase transitions is introduced. The results obtained with this method, combined with a careful finite-size scaling analysis, shed new light on a long-standing controversy concerning the asymptotic behavior of the boundary.

Chap. 5 specializes in the critical behavior of the model in the second-order regime, and the crossover from the long-range to the short-range regime. A detailed review of controversial results is provided, and a novel approach is presented to investigate the onset of the short-range regime. An intermediate conclusion is provided at the end of this part.

Part III (Chap. 6 to 8) puts forward a novel method to simulate long-range models in the multicanonical ensemble, which is able to considerably extend the range of attainable lattice sizes.

In Chap. 6, the limitations of standard multicanonical methods are discussed, first from a general perspective, and then in the context of long-range models. Existing schemes aimed at improving the performance of multicanonical methods are reviewed, and I discuss the main impediments to their extension to long-range models. A later section highlights the main issues that must be overcome in order to improve the multicanonical method in this context, and outlines an original, innovative strategy.

Chap. 7 describes the methods in extensive detail. It comprises an article that was submitted to Phys. Rev. E. and is currently being refereed. Two optimization schemes dedicated to long-range models are presented. Extensive tests of the methods are performed, which demonstrate the superiority of the method with respect to standard implementations in terms of accuracy and scalability. The method is used to test several finite-size effects in long-range models.

An extension of the method to disordered models is presented in Chap 8. A

Introduction and motivations

long-range Ising chain with bimodal random fields is investigated, and a phase diagram is provided. Possible improvements of the method in this context are discussed.

Developments and perspectives are discussed at the end of each chapter. They are summarized in the conclusion, both from the perspective of the models and the methods.

Part I

Models and methods: review

Chapter 1

Spin models with long-range interactions: review and pending questions

Long-range spatial interactions are ubiquitous in nature. At all length scales, they contribute to shape our Universe, from the electromagnetic force at the microscopic level to the gravitational force in astrophysical structures. While the gravitational force makes its genuine long-range nature felt everywhere, this is no longer the case of the electromagnetic force. The existence of electric charges and currents of opposite sign gives rise, among other things, to screening effects and multipolar interactions, that tend to reduce its range. In some cases, the reduction is so marked that the interaction becomes effectively short ranged, i.e., restricted to “neighboring” particles.

Looking back on more than a century of condensed matter physics, one realizes that *effective* short-range interactions have in fact played a greater role in shaping this field than have long-range interactions. The pivotal role played by the nearest-neighbor Ising model in various subfields of physics undoubtedly accounts for this. In recent decades, however, long-range interactions have started drawing unabated interest in microscopic modeling, spanning an ever increasing number of systems. The goal of the present chapter is to give an overview of these systems, with an emphasis on the topic of *spin models* which is central to this thesis.

The distinction between short and long ranged force probably deserves a more rigorous definition than the sketchy account given above. This will be my first item of business in this chapter. Then, I will give an overview of the various fields where long-range interactions hold a prominent role. Moving in to the more specific subject of long-range spin models, I will review the main results that have been made hitherto available, from ferromagnetic chains to long-range spin glasses to Lévy flights. One of the purposes of that part will be to highlight the many similarities regarding the (pending) questions that surround seemingly very different models.

Finally, the long-range Potts model will be singled out as a possible paradigm

of phase transitions in long-range spin systems, owing to its particularly rich phase diagram and the abundance of pending questions and controversies that remain attached to it. These will be given special attention, inasmuch as they constitute one of the central motivations of this thesis.

1.1 Long-range vs short-range interaction: semantic issues

Locating the frontier between long- and short-range interaction in a unique fashion is not trivial, as many (sometimes controversial) definitions have coexisted in the literature of the past three decades. Since the early work of Ruelle [290] on ferromagnetic chains with interactions decaying as a power law of the interparticle distance r , the accepted lore in the framework of critical phenomena has been to distinguish between long ranged and short ranged interactions from the behavior of the moments of the interaction potential, i.e., the integral $\int V(\vec{r})r^n d^D r$ defining the n th moment.

In the one-dimensional Ising model, for instance, there is no phase transition at finite temperature whenever the first moment is finite [290, 107], which suggests that the interaction is short ranged here for $\sigma > 1$. Looking back on the early work of Fisher, Ma and Nickel on the critical behavior of long-range ferromagnetic spin systems [115], the suggested situation is that, on a D -dimensional lattice, any interaction between two spins separated by an interparticle distance r is said to be *short ranged* if it assumes, either an exponential decay $e^{-r/b}$, or a power-law decay $1/r^{D+\sigma}$ provided $\sigma > 2$. This definition relies on the critical behavior of the model, inasmuch as a critical behavior is said to be long ranged if the corresponding critical exponents match those of the same model with nearest-neighbor interactions. The situation is in fact dramatically more intricate, as is genuinely certified by more than two decades of vivid debate [292, 142, 232]¹; and for a large class of models², the case is far from being closed.

Dantchev and Rudnick [91] recently reexamined the issue under an enlarged perspective encompassing short-range, long-range, and *subleading* long-range interactions. An interaction is said to be of the long-range type if the moment of n th order of the interaction potential $J(r)$ diverges for a sufficiently large n . This in effect rules out *without ambiguity* interactions that follow an exponential decay, e^{-r/r_0} . In Fourier space, the expansion of $J(q)$ gives rise to so-called anomalies, i.e., it can no longer be expressed in terms of powers of q^2 only. In this respect, a criterion for short-range interactions is a finite second moment for $J(r)$. This means that the q -dependent term in the Ginzburg-Landau functional may be written as $r + v_2 k^2 + R(k)$, with $R(k)$ asymptotically smaller than the first two terms. Systems with *subleading* long-range interactions (for instance, Van der Waals interactions) belong to this class: here, $R(k)$ contains a term of the form k^σ , with

¹The whole picture is described at length in Sec. 5.3.

²With a negative Fisher exponent.

1.2. From Van der Waals forces to neural networks

$\sigma > 2$ (σ non-even); in other words, all moments are finite up to the σ th order (or the largest integer smaller than σ).

Looking at the issue from a totally different perspective, physicists working in the field of non-extensive thermodynamics and Tsallis's generalized entropy [322, 68, 294, 89, 92] connect the long-range character of an interaction to its non-integrability, i.e., the interaction is long ranged as soon as the zeroth-order moment diverges, which implies that the thermodynamic limit becomes ill-defined, and an appropriate redefinition of thermodynamic variables is required [68, 65, 312, 89, 136]. Incidentally, the contrasting definitions emanating from these two (sub)communities have raised gentle disputes at times, e.g., when both communities went to investigating the crossover between extensive (i.e., short-range vs long-range) and non-extensive systems (i.e., long-range vs non-integrable) [331, 332].

In this thesis, I will exclusively consider long-range interactions in the sense of integrable interactions, where the thermodynamic limit always exists. I will also merely concentrate on power-law decaying interactions, i.e., of the form $1/r^{D+\sigma}$, focusing on the change of qualitative and quantitative behavior as σ is varied. In this respect, it is worth underlining that σ is not strictly speaking a *range* parameter, since a power-law interaction has no intrinsic length-scale. This parameter will thus be termed *decay parameter* in this work.

As a last point in this section, it is also important to mention that, whenever they are investigated by means of numerical methods — as will mostly be the case in this thesis — long-range models exhibit a distinctive feature with respect to their nearest-neighbor counterparts. Indeed, most numerical methods — specifically Monte Carlo methods — impose that systems be investigated at finite geometry: in a long-range model, every particle sees the boundaries, and as will be witnessed in several parts of this work, this feature will give rise to interesting, unusual finite-size effects that do not show up in nearest-neighbor models.

1.2 From Van der Waals forces to neural networks

This section mostly centers on microscopic models in condensed matter physics and related fields, since this is where integrable interactions have had their longest record of promising applications. In particular, applications to astrophysical structures will be left out, as the gravitational force is indeed non-integrable. It should be mentioned, however, that this field has been extensively studied during recent decades, and is still a matter of intense debate [203].

Microscopic models with long-range interactions decaying as a power law, i.e., as $1/r^{D+\sigma}$, have aroused renewed interest during the last decade. Beyond their relevance to the understanding of fundamental issues in the physics of critical phenomena, which I will review in the next section, they have also started playing a seminal role in the modeling of a large class of physical, chemical or biological models where electrostatic interactions, polarization, or van der Waals forces play a central role. Thus long-range interactions do not simply represent a case of

Chapter 1. Spin models with long-range interactions: review and pending questions

academic interest.

Applications to condensed matter include systems undergoing phase separation, e.g., liquid-liquid phase separation in a highly ionic system [278], phase segregation in a lattice gas model binary alloy governed by Kawasaki exchange dynamics [125], and naturally systems governed by Van der Waals interactions. As regards critical phenomena, the Van der Waals interaction leads to subleading long-range behavior, i.e., short-range-like, yet giving rise to non-trivial effects in terms of critical behavior; this has been extensively studied in [91, 90]. The Casimir effect, initially discovered in the context of quantum electrodynamics — where it stems from vacuum fluctuations of the electromagnetic field — has also been observed in fluids composed of neutral particles near the critical point, with a decay exponent being far below that of van der Waals interactions [60].

In the physics of lattice spin models, long-range Hamiltonians appear in connection with a variety of exciting contemporary problems, in particular the Kondo effect, which is central to the investigation of high temperature superconductivity, e.g., in Josephson junctions [77, 78]. An equivalence between the (classical) q -state Potts chain with inverse square interactions ($1/r^2$) and the Kondo problem was proposed by Anderson [11], who showed that the path-integral representation of the spin- $\frac{1}{2}$ Kondo problem could be mapped onto a classical Ising chain, with successive spins on the chain representing the time history of the single impurity along imaginary time. Classical Ising models with long-range, RKKY³ interactions of the form $1/r^D \cos(k_0 r + \phi)$ also appear in a quite natural way in the study of spin glasses [121]. More recently, models with competing short-range ferromagnetic interactions and long-range antiferromagnetic dipolar forces have surfaced in the fast-growing technological field of ultrathin magnetic films [45, 66]. These models are of fundamental importance to the understanding of the geometric and dynamic properties of magnetic domains, e.g., in metal-on-metal thin films [239], which exhibit complex patterns (for instance, “striped”, i.e., lamellar phases characterized by modulated patterns [329]). They are therefore vital to the development of new data storage devices [8] (see, for instance, [99] for a recent review), e.g., permanent random-access memories. In higher dimensions, these models are also of interest in the modeling of a variety of related problems in soft-matter physics (see references in [138]).

In the fringe of the customary physicist’s turf, lattice models with long-range interactions also turned out to play an important role in the modeling of the brain [10], where neurons far apart interact through an action potential that decays slowly along the axon, in the context of pattern recognition (as an alternative to multilayer neural networks) [302, 135] and auto-associative memories with dilute long-range connections [268], in neural network learning processes [123], and in the modeling of visual perception and orientation [313]. Ising models on small-world networks with interactions decaying as a power law of the euclidian distance between nodes were also studied in [193].

³Ruderman-Kittel-Kasuya-Yosida.

1.3. Spin models governed by long-range interactions: raising the effective dimension

Power-law decaying interactions also appeared in self-organized criticality: the Bak-Sneppen model with $1/r^{D+\sigma}$ interactions, which is claimed to represent a simple, yet realistic model of coevolutionary species, with the “distance” between species decaying as a power law, was investigated in [126], and also in [12] in connection with Tsallis’s generalized thermodynamics. A model of spreading epidemics with a long-range probability of infection, in which the spreading agents perform Lévy flights⁴, was studied by Hinrichsen and Howard [157] by means of anomalous directed percolation. The equivalence between Lévy flights and long-range ferromagnetic spin models was also explored in [32, 352] in the context of nonequilibrium steady states, by means of a kinetic Ising model in which a standard (thermal) spin-flip dynamics competes with random Lévy-flight spin-exchanges. Long-range forces also emerged in non-linear wave theory: an extension of the Klein-Gordon equation whose solutions are long-range interacting solitons, was proposed by Guerrero and González [140].

Finally, long-range interactions have also attracted much attention in the framework of nonextensive thermodynamics and Tsallis generalized statistics, where a possible equivalence with short-range models is under consideration [322]. Nonetheless, most of the work dedicated to Tsallis’s statistics centers on *non-integrable* long-range interactions, and as stated above, will not be considered here.

1.3 Spin models governed by long-range interactions: raising the effective dimension

I now turn to a review of the equilibrium properties of long-range spin models, in particular models governed by algebraically decaying interactions $1/r^{D+\sigma}$. Some of the results reviewed here have been the subject of intense, long-running debate, and are reexamined in several parts of this thesis. More detailed explanations will be given when needed. Therefore, this paragraph only aims to give a succinct description of the most “distinguished” results, in a way that may help the reader get a quick grasp of the subject.

Since the seminal work of Joyce on the spherical model [195], these models — beyond their relevance in microscopic modeling — have also turned into an ideal testing ground for the physics of critical phenomena: the key point here is that the decay parameter influences the universal properties of the model in much the same way as the dimension does in a nearest-neighbor model. In a sense, the decay parameter σ plays the role of an effective dimension which, in the framework of the field-theoretic renormalization group, allows the validity of ϵ -expansion to be checked in a very flexible way by means of numerical methods, since the ϵ parameter can now be *continuously* varied (as against, e.g., $\epsilon = 4 - D$ in ϕ^4 theories). Long-range models are therefore a powerful paradigm for studying the dependence of critical properties on dimensionality, e.g., in systems above their upper critical

⁴A random walk that is steered by a so-called “superdiffusive” motion, and in which the lengths of the (random) steps are distributed according to a power law.

dimension [228]. In the context of the long-range Ising chain, this was extensively investigated by Luijten et al. [40] by means of Monte Carlo methods, with a particular focus on crossover exponents describing the crossover from the mean-field to the non-mean-field regime [39, 41]. On the numerical side, the initial spark which ignited the investigation of crossover functions between universality classes may certainly be traced back to the work of Luijten and Blöte on the nearest-neighbor Ising model in $D \geq 4$ [228, 229] or with constant interactions of variable (finite) range [230, 233].

1.3.1 Results for ferromagnetic O_n models

Rigorous studies of the Ising chain ($n = 1$) with $1/r^{D+\sigma}$ ($D = 1$) interactions can be traced back to the seminal works of Ruelle [290] and Dyson [107, 108], which (i) ruled out long-range order at finite temperature for $\sigma > 1$, and (ii) proved the existence of a phase transition at a non-zero critical temperature for $0 < \sigma < 1$, respectively⁵. This rigorous result, which is in strong contrast to the nearest-neighbor case where no phase transition exists at finite temperature, was later extended to $\sigma = 1.0$ [300, 124]. Its generalization to D -dimensional models was proposed in [6, 55], showing that $D_l = \sigma$ is the lower critical dimension of the model.

For one- and two-dimensional Heisenberg models, similar rigorous results were obtained by Mermin and Wagner [245], ruling out long-range ferro- and antiferromagnetic order whenever the second-order moment of the interaction is finite, i.e., $\sigma > 2$ where power-law decaying interactions are concerned. In one dimension, Simon and Sokal [300] proved that no long-range order at finite temperature exists on the $\sigma = 1$ line, in contrast with Ising chains (although essential singularities may show up irrespective of the number of components of the order parameter, yet with different behavior depending on whether $n = 1$ or $n > 1$ [208]). Heisenberg models were reexamined by means of spin wave theory in [74, 75], showing long-range order at finite temperature for $0 < \sigma < D$ in one and two dimensions, and all $\sigma > 0$ (i.e., in the extensive regime) in higher dimensions. The two-dimensional XY model with dipolar ($1/r$) interactions was also investigated very recently [240] by means of renormalization theory, and shown to display a phase transition similar to the Kosterlitz-Thouless transition [207].

Early works based on Wilson's version of the renormalization group were initiated by Fisher, Ma and Nickel [115] on general O_n models⁶. This work concerned itself with the universality classes of the model, with two regimes identified: for $0 < \sigma < 0.5$, the critical exponents are classical (i.e., mean-field-like, with the exception of $\nu = 1/\sigma$), whereas the region $0.5 < \sigma < 1$ displays non-trivial critical exponents; naturally, these results assume that the phase transition is continuous. These surmises were partially confirmed for the Ising chain by the early series ex-

⁵Ruelle's result is in fact more general and makes no assumption about the exact form of the interaction: there is no phase transition if the zeroth- and first-order moments of the interaction are finite. See a related discussion in Sec. 1.1.

⁶An outline of the derivation is provided in Sec. 5.1.

1.3. Spin models governed by long-range interactions: raising the effective dimension

pansions of Nagle and Bonner [254], and by later (extensive) investigations based on transfer matrices [323] or Monte Carlo simulations [228, 231, 225, 213, 43]. Overall, these results confirmed that long-range models in effect go through a variety of universality classes as σ is varied within the non-classical range, thus exhibiting rich critical behavior. I will turn to them again in greater detail in Chap. 5, in the context of the long-range to short-range crossover.

In [115], Fisher and coworkers also addressed the crossover from the long-range to the short-range regime. The argued value of $\sigma_{co} = 2$ as the crossover decay exponent raised several inconsistencies, however, in particular with regard to the one-dimensional case where no phase transition exists for this value. A later work by Sak [292] on the same issue arguing that $\sigma_{co} = 2 - \eta_{SR}$, where η_{SR} is the Fisher exponent of the nearest-neighbor model, only added fuel to the row, which then turned into a long-standing controversy still unresolved in some respects (see Sec. 5.4 for more on this).

The $\sigma = 1$ case corresponding to the line of inverse square interactions, and for which possible connections with the Kondo problem were mentioned in the previous section, was first elucidated by Kosterlitz [208] by means of an expansion in $\epsilon = 1 - \sigma$, and later revisited by Imbrie [171]. There it was claimed that this line may be the locus of an infinite-order phase transition similar in many respects to the Kosterlitz-Thouless transition occurring in the two-dimensional XY model. The result was confirmed by Cardy [72] for a discrete q -state model, by mapping the model to a gas of kinks interacting logarithmically, yet some authors also argued that the line might not correspond to the onset of a Kosterlitz-Thouless-like transition for some values of q [19]. Controversies regarding inverse square interactions will be discussed in much more detail in Sec. 4.3.6, in the context of the q -state Potts chain.

1.3.2 Antiferromagnetic, frustrated and disordered models

The discussion below is aimed at giving a cursory look at some prominent results, including a handful of conjectures, that emerged in the field of long-range models with antiferromagnetic, competing or random interactions. Several features, as will be witnessed, are shared by a large class of long-range models, including the q -state Potts chain which I investigate in this thesis: here again, the decay parameter of the interaction plays a central role, and most debates revolve around the change in critical behavior as this parameter is varied.

Models with oscillatory long-range interactions, i.e., of the form $r^{-D-\sigma} \cos(k_0 r + \phi)$ where $k_0 = 0$ corresponds to ferromagnetic couplings, were rigorously investigated by Bruno [58]. The aim was to generalize the rigorous result of Mermin and Wagner [245], which rules out long-range order in Heisenberg models whenever the second-order moment of the interaction is finite. One of the main rigorous results of [58] in one dimension is that long-range order cannot exist in O_2 and O_3 chains (yet the result does not apply to O_1 models) with long-range oscillatory interactions of the kind given above and $0 < |k_0| \leq \pi$, whenever $\sigma \geq 0$. This, in

Chapter 1. Spin models with long-range interactions: review and pending questions

particular, forbids long-range order for systems governed by true RKKY interactions ($\sigma = 0$ here). In two dimensions, the result is somewhat more intricate, since it involves the phase of the oscillatory potential.

On the side of purely antiferromagnetic (i.e., staggered) interactions, which do not induce any frustration, most works targeted the effect of *quantum* fluctuations on the destruction of the Néel order at $T = 0$. By considering Heisenberg chains with pair interactions of the form $V_{ij} = -(-1)^{|i-j|}/|i-j|^{1+\sigma} \vec{S}_i \cdot \vec{S}_j$, where the \vec{S}_i are quantum spins whose coordinates satisfy the usual commutation relations, and $\vec{S}^2 = s(s+1)$, Parreira et al. proved on rigorous grounds that the ground state exhibits Néel order, stable against quantum fluctuations for $\sigma < 2$ and sufficiently large value of s . The same model was studied in [361] by means of spin wave theory, and led to the conclusion that Néel order is also stable at finite temperature for $\sigma < 1$. As a side-note, quantum fluctuations were also studied in a variety of long-range quantum models, including a quantum flavor of the classical XY model (the so-called “quantum rotator” model) [106], a quantum spin glass [104, 105], and diluted models [105]. As quantum phase transitions are somewhat orthogonal to the core subject of this thesis, I will not comment on them further.

While frustration in short-range models is usually induced by a particular choice of the lattice geometry, long-range models can become frustrated under the sole influence of the long-range interaction. Two classes of long-range frustrated models have been considered in the recent past: completely frustrated models with a single, frustrating long-range interaction, and weakly frustrated models where frustration is induced by competing short- and (weak) long-range interactions. In the context of long-range interactions, *complete* frustration can be induced by “repulsive” interactions between *all* spins. This can be carried out by taking a pair interaction of the form $V_{ij} = +1/|i-j|^{\sigma+D} \vec{S}_i \cdot \vec{S}_j$. Completely frustrated O_2 and O_3 models were explored by Romano [288] using spin-wave theory and Monte Carlo simulations. The system ground state corresponds to a staggered configuration, i.e., with spins pointing alternatively up and down. Spin wave theory predicts no long range order at finite temperature, although Monte Carlo simulations suggest residual order at finite temperature that wanes off whenever lattice size is increased.

Much more attention has been given to weakly-frustrated models (as opposed to complete frustration), however. As mentioned in the previous section, two-dimensional Ising models with competing short- and long-range interactions are indeed of fundamental interest for the understanding of the physical properties of magnetic ultrathin films, while their three-dimensional counterparts are of relevance for the modeling of a variety of systems in soft-matter physics (see references in Sec. 1.2 above). In addition, the phase diagram of these models is much more interesting than that of completely frustrated models, because weak frustration suppresses the usual long-range ferromagnetic order of the short-range Ising model and produces complex mesophases characterized by lamellar patterns. The model Hamiltonian comprises two parts: an exchange interaction yielding *ferromagnetic* coupling between nearest-neighbor spins, and a weak, long-range, either dipolar or Coulombian *antiferromagnetic* interaction. The Coulomb frustrated Ising model in

1.3. Spin models governed by long-range interactions: raising the effective dimension

$D = 3$ was explored in [329, 137, 138], while the two-dimensional version with dipolar, $1/r^3$ interactions is reviewed in [66]. Note that the first model corresponds to the non-extensive regime, which requires that the total magnetization be fixed to zero for the thermodynamic limit to be properly defined [138].

For both models, the ground state of the system corresponds to a staggered configuration for large frustration (as in the completely frustrated model studied by Romano [288]), whereas at low frustration it exhibits anti-aligned stripes whose width grows exponentially with the relative strength of the (short-range) ferromagnetic interaction with respect to the long-range part [239]. As for the dipolar version in two dimensions, it has been suggested by means of Monte Carlo simulations, that the presence of many striped phases geometrically close to the ground state generate a complex structure of interwoven metastable states in the vicinity of the ground state energy [127], a feature which on the numerical side is known to induce long thermalization time. Incidentally, the transition between striped phases of increasing width was reported to be of the first order.

The phase diagram of the Coulomb frustrated model (in $D = 3$) is even more elaborate: at zero temperature, lamellar phases turn into so-called *tubular* phases and then orthorhombic phases as the frustration is increased, a feature originating in successive translational symmetry breaking in the second and third direction [137]. As temperature is increased, and for low frustration, lamellar phases repeatedly transform into more and more complex modulated phases, with the phase diagram showing an infinite-tree-like structure, until an infinite number of incommensurate modulated phases eventually appears at the order-disorder transition temperature. Noteworthy enough, and as mentioned by the authors [138], the intricate geometry of these phases renders numerical investigation particularly demanding, as (i) the presence of long-range interactions imposes a very low limit on the maximum attainable size, and (ii) finite geometry restrains the number of modulated phases that can be observed. The algorithm used in this study is an interesting extension of the geometric cluster algorithm, and is briefly detailed in Sec. 2.3.

As a last point in this section, I briefly turn to disordered systems governed by long-range interactions. As far as random-field models are concerned, a more detailed discussion in the context of bimodal fields will be provided in Chap. 8.

The random-field Ising model was investigated by Bray [52] in the context of the long- to short-range crossover (the point is reexamined in Sec. 5.3), showing a crossover to take place at $\sigma = 2 - \eta_{SR}$, where η_{SR} is the Fisher exponent of the nearest-neighbor model, in compliance with the result previously obtained by Sak [292]. One of the most interesting results was established in [343] using a renormalization-group approach applied to an Ising chain with uncorrelated, gaussian distributed random fields. By generalizing the Imry-Ma dimensional argument to long-range models [173], i.e., by analyzing the stability of the ordered state against an arbitrarily weak random field, a critical value of $\sigma = 1/2$ was reported [343], above which no phase transition occurs at finite temperature. An expansion was then carried out around $\sigma = 1/2$, a value which seems to play the same role for the random-field model as $\sigma = 1$ for the pure long-range model. As

in the short-range version of this model, the critical behavior is governed by a fixed point at zero-temperature and fixed (non-zero) randomness, yet the initially gaussian distribution is shown to become strongly non-gaussian under renormalization. It was then suggested by the authors in [343] that this may signal a first-order transition driven by fluctuations. This point will be reexamined in Chap. 8 in the case of a long-range Ising chain with bimodal random fields.

As regards spin glasses, few studies of the Ising chain with random, power-law interactions $1/r^{1+\sigma}$ have been made available in the last fifteen years. Kotliar explored the case of gaussian-distributed random couplings using a kink-gas model and the standard replica symmetry-breaking theory [209], and found a phase transition⁷ to occur at finite temperature for $-1/2 < \sigma < 0$, with critical exponents taking on their mean-field value for $\sigma < -1/3$. By pondering on possible equivalence with short-range models, it was also suggested that the lower critical dimension of the corresponding short-range spin glass may be 3. Noteworthy enough, Sak's scenario seems to hold also for spin glasses, provided in Sak's relation the η_{SR} exponent is replaced by the (domain-wall) stiffness exponent θ_{SR} of the short-range spin glass, which means that long-range behavior becomes dominant whenever $\theta_{LR} > \theta_{SR}$ ($= -1$ for $D = 1$, which yields $\sigma_{co} = 1$ for the crossover value, since $\theta = 1 - \sigma$ [116, 53], see also Sec. 5.3). More recently, extensive Monte Carlo simulations [200] were conducted on the same model for linear sizes up to 256 spins⁸, with emphasis on the values of the stiffness exponent(s) θ , still a matter of debate [215, 163, 271]. The issue revolves around the question of whether the stiffness exponent, which characterizes the scaling of the energy of elementary excitations with respect to size, is the same for domain walls at zero temperature and for droplet-like excitations. Results confirm the replica-symmetry-breaking picture with regard to the shape of the overlap parameter, and droplet theory is ruled out by monitoring the behavior of the minimum of the overlap parameter distribution (i.e. at its zero value). The stiffness exponent obtained for droplet-like excitation is ~ 0 , in compliance with both the replica-symmetry-breaking theory of Parisi and the trivial-non-trivial scheme [215, 271].

1.4 Long-Range Potts model: conjectures and controversies

The long-range flavor of the q -state Potts model will play a central role in this thesis, and as such deserves particular attention in this review. Yet several controversies related to this model are somewhat intricate, and I deemed it more useful to describe them at length in the core of this manuscript rather than here.

Owing partly to a higher ground state degeneracy, this model reveals a phase

⁷It turns out that, owing to the gaussian distribution of random couplings, extensivity is ensured even for $-0.5 < \sigma < 0$ (in $D = 1$). This would not be the case for a ferromagnetic model.

⁸Note that an original scheme for the implementation of periodic boundary conditions is used in this work, see Sec. 4.2.1.

1.4. Long-Range Potts model: conjectures and controversies

diagram markedly richer than that of the Ising chain, for instance, and as such may perhaps be regarded as a *paradigm* of model systems. This diagram is reminiscent, though only incompletely, of the phase diagram of its nearest-neighbor counterpart, with the order-disorder transition turning from a continuous to a first-order one as the number of states q is increased beyond a threshold value $q_c(d)$ depending on the dimensionality of the model. For instance, $q_c(2) = 4$ and $q_c(4) = 2$ [4, 17], and there is strong evidence in favor of $q_c(3) \sim 3$ [350].

It turns out that one of the most recurrently intriguing questions in the long-range case is the possible existence of a so-called “tricritical” line $\sigma(q)$ separating a first- and a second-order regime (and, as a corollary, the location thereof). There are other intriguing questions though: the behavior on the $\sigma = 1$ boundary, where a dispute recently emerged over the topological nature of the transition and the range of values of q where it may set in; the location of the long- to short-range crossover, which in the case of one-dimensional models, *may* take place also on the line $\sigma = 1$; the crossover from the extensive to the non-extensive regime, in connection with Tsallis’s conjecture, which aspires to unify these two regimes in a single framework. These questions are briefly reviewed below.

Boundary between first- and second-order regimes On the numerical side, a clue for the existence of a “tricritical” line was first provided by Glumac and Uzelac [131] using Monte Carlo simulation based on the Luijten-Blöte cluster algorithm (see Sec. 2.3.2). The investigation performed in this work targets the three-state Potts model, and suggests that there is indeed a value of σ_c below which the transition is of the first order. This *qualitative* picture was later reinforced in [19] for $q = 3, 5, 7, 9$, and in [213] for $q = 3$, both relying on Monte Carlo studies, and simultaneously in [132] using an approach based on the random-cluster representation of the model, which made it possible to handle non-integer values of q . Even so, the exact location of the tricritical line separating both regions is still fairly controversial. While for $q = 3$, σ_c was claimed to lie between 0.6 and 0.7 [131], Krech and Luijten pointed out that $\sigma = 0.7$ still belongs to the first-order regime, and that the second-order regime may set in for $\sigma = 0.75$ only [213]. The situation with $q = 5$ turns out to be even worse, with numerical estimates available only within fairly large ranges: a lower boundary value of 0.8 was reported in [131], whereas $0.7 < \sigma_c(5) < 1.0$ according to [19]. Overall, the only reliable assertion is that $\sigma_c(q)$ increases with q .

On the analytical side, studies specifically dedicated to the Potts model have been equally scarce. Whether based on real-space procedures [64], or Wilson’s momentum-shell method [279, 315], these studies essentially target the critical behavior of the model (irrespective of the order of the transition), and remain inconclusive, where distinguishing between first- and second-order transitions is concerned. The first enlightening step in this direction emerged very recently in a work by Biskup, Chayes, and Crawford [42], that attempts to draw very general conclusions for models with interactions that (i) satisfy the reflection-positivity

Chapter 1. Spin models with long-range interactions: review and pending questions

condition, and (ii) are sufficiently “spread out” (in practice, with an exponential or a power-law decay). In one-dimension, the authors prove that, whenever mean-field theory signals a (thermally-driven) first-order transition, then the system has a similar transition for $\sigma \in [0, 1]$. This result would imply that, for $q \geq 3$, the long-range Potts chain would exhibit a transition already for $\sigma \leq 1$. Several remarks are in order:

- this fixes an upper bound to $\sigma_c(q)$ which lies above that given in some numerical studies [213];
- and yet the derivation of Biskup et al. also means that, irrespective of the exact value of $\sigma_c(q)$, $\sigma_c(q) > 1$ is ruled out.
- reflection-positivity conditions pose strong constraints on the decay parameter, which in practice reduces the validity of the derivation to $\sigma \sim 1$;

To recap, no utterly convincing *quantitative* picture has been made available with regard to the location of the line $\sigma_c(q)$, and a refinement of the phase diagram seems compulsory in this respect.

The line of inverse square interactions The behavior on the line of inverse square interactions raises another set of thorny questions. The work by Cardy [72] on the $\sigma = 1$ boundary shows the onset, for all q , of a Kosterlitz-Thouless-like transition exhibiting singularities of the essential type. While this confirms the previous result of Kosterlitz [208], it was also argued in [19] that the phase transition changes from a second-order to a first-order one for $q \geq 9$. While the recent work of Luijten and Messingfeld on the Ising chain [235] lends further support to Cardy’s assertion, the upper bound obtained by Biskup et al. is *once again* inconsistent with the result of Cardy, and it seems that we are back to square one.

Overall, the controversy still appears unsettled. *How can these contradictory scenarios be reconciled, if at all?* One of my *objectives* in this thesis will be to “broker” a scenario that gives a convincing picture of the phenomena that occur on this borderline. With this in mind, a prerequisite might very well be the inspection of the asymptotic behavior of $\sigma_c(q)$ as $q \rightarrow \infty$. In this respect, it is interesting to note — in anticipation of Chap. 4 — that the rigorous result of Biskup et al., while no precise estimation of the phase diagram is given, does not contradict the scenario that will be suggested in Chap. 4, according to which $\sigma_c(q) \rightarrow 1$ asymptotically as $q \rightarrow \infty$.

The long-range to short-range crossover Here again, few analyses concerned themselves with the q -state Potts model, and yet — as mentioned in an early section of this chapter — the location of the crossover has been the matter of debate. Renormalization group analyses resorted to a continuum version of the Potts model [279, 315], and did not clearly settle the controversy. In the one-dimensional case,

1.4. Long-Range Potts model: conjectures and controversies

no numerical simulations have addressed the issue thus far, and the situation remains unclear. If Sak's scenario proves reliable for this model as well, then the crossover should take place at $\sigma = 1$, where an infinite-order transition may *also* occur. From the viewpoint of numerical simulations, settling the controversy might be challenging, owing to the presence of essential singularities.

Critical couplings and the mean-field regime Most numerical investigations that concerned themselves with the Potts model focused on the estimation of critical couplings (irrespective, in terms of the estimators utilized, of the order of the transition — with the exception of [213]). These include the transfer matrix approach of Glumac and Uzelac [130], the coherent anomaly method [251], and Monte Carlo simulations [131, 132, 19, 325, 213]. The last, however, mostly addressed the case $q = 3$, and led to numerical estimates of critical temperatures (and, in some respects, critical exponents) showing severe discrepancies. In addition, the behavior in the vicinity of the non-extensive regime ($\sigma \rightarrow 0$) is inconsistent with Tsallis's conjecture $1/T_c \sim \sigma$, except in a cluster mean-field approach provided by Monroe [251]. In this respect, it is important to emphasize that the periodic boundary conditions implemented in these numerical studies sometimes differ markedly from one work to the other: *to what extent do they influence the finite-size scaling behavior? Do the observed discrepancies presage the failure of Tsallis's conjecture?*

Outlook As is all too apparent from the review given above, the long-range Potts model is still a fertile ground for the investigation of critical phenomena. Paradoxical though this may seem, specific studies are markedly less numerous and in any case rather recent, whilst significant emphasis has long been placed on the investigation of O_n models. Every part of the phase diagram of this model is pregnant with intriguing, indeed unsettled questions: these are the main motivations behind the second part of this thesis. Prior to going over this subject, however, I will review the methods that will be used in Part II to scrutinize these questions.

Chapter 2

From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

In this thesis, I will mostly use Monte Carlo methods to investigate the physics of long-range spin models; I will resort to analytical methods on some rare occasions as a guide (Sec. 4.1.1 and 5.4.1 in particular). Since the seminal work of Metropolis et al. [246] on the eponymous algorithm, more than five decades have passed: a wealth of methods of astonishing diversity have been born, all more or less inheriting from the Metropolis algorithm, yet adding to it an ever increased level of complexity. Unraveling ever more complex physical phenomena imposes that new algorithms be continuously designed, because the efficiency of numerical approaches is not simply a matter of available computer resources. Even if these have exponentially soared during the last three decades, there are *still* physical problems that just cannot be tackled with the Metropolis algorithm.

This chapter is thus devoted exclusively to Monte Carlo methods, with a marked emphasis on:

- methods that were specifically tailored to long-range models, e.g., dedicated cluster algorithms;
- methods that were *not* specifically tailored to long-range models, and yet might help in investigating a particular class of problems encountered in long-range models: methods operating in *generalized ensembles* are examples of these.

However, this chapter is not aimed at giving a comprehensive review of *all* Monte Carlo methods, or at providing extensive detail on the theory of finite-size scaling and other data analysis schemes. With regards to these, more detailed material may be found, e.g., in [216] or [259], where progress in the field is described up to 2001.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

The layout of this chapter is as follows. First, I will briefly review some fundamentals, i.e., Markov chains and detailed balance, the Metropolis algorithm, and error calculation. The Monte Carlo method is primarily a *stochastic* method, and the amount of statistical error on generated data hinges on the choice of the underlying algorithm. I will describe a class of algorithms in widespread use, cluster algorithms, that lead to a drastic lowering of this error. All the algorithms I will have presented up to that point operate in the canonical ensemble, i.e., engender a Boltzmann distribution. In a variety of situations, including the investigation of first-order phase transitions or the physics of disordered systems, these methods behave poorly and become of little use. I will describe a category of algorithms that efficiently tackle such situations by operating in other ensembles: generalized ensembles algorithms. The multicanonical method will be singled out as the paradigm of simulations in generalized ensembles; it will also be my method of choice for the investigation of long-range systems, and as such will be given particular attention in this chapter.

2.1 Markov chains and the Metropolis algorithm

Irrespective of the particular scientific field considered, the Monte Carlo is essentially a method that computes (multidimensional) integrals in an approximate fashion. It does so by relying on a stochastic scheme, i.e., by drawing *random* numbers; hence its name. .

In statistical physics, the multidimensional integral to be (stochastically) computed is the partition function. In a numerical implementation, the integral is actually a discrete sum over all the configurations of the system, $Z = \sum_{[\sigma]} e^{-E/kT}$, and the Monte Carlo method aims to produce an estimate of it by means of a stochastic sampling scheme. In short, we have two sampling schemes at our disposal.

“Simple sampling” samples the partition function by generating configurations at random. Since there are astronomical numbers of configurations of high energy in comparison with those of lower energy, this sampling scheme is particularly inefficient, because in the temperature range we are usually interested in, a vast majority of these configurations do not contribute significantly to the partition function.

“Importance sampling”, on the contrary, is a *biased* sampling scheme, which concentrates much of its effort on the configurations that are dominant in the expression of Z . It does so by generating configurations that are distributed according to the Boltzmann distribution at a given temperature: thus this sampling scheme (at least in its initial formulation termed the Metropolis algorithm [246]) is essentially a single-temperature method.

Importance sampling relies on a Markov chain of configurations (or states, or microstates) to engender the proper distribution: starting from a given configuration σ_i , the chain chooses a new configuration σ_f according to a transition proba-

2.1. Markov chains and the Metropolis algorithm

bility $W(i \rightarrow f)$ which must satisfy the so-called *detailed balance equation*

$$w_i W(i \rightarrow f) = w_f W(f \rightarrow i), \quad (2.1)$$

where w_i and w_f refer to the probability of occurrence (often termed the *weight*) of states i and f in the engendered distribution. For quite some time, only Boltzmann distributions have been considered, whereby $w_i = e^{-E_i/kT}$ (where E_i is the energy of state i). Section 2.5 will show that other, sometimes wiser choices are possible.

The previous equation can be derived from many sources. One may, for example, start from the *master equation*

$$\frac{dw_f(t)}{dt} = \sum_i [w_i(t)W(i \rightarrow f) - w_f(t)W(f \rightarrow i)], \quad (2.2)$$

which governs the dynamics of the weight $w_f(t)$. There is actually one such equation for each configurational weight, with the sum rule $\sum_i w_i(t) = 1$ holding at all times. At equilibrium, the rate of change of any weight is down to zero, whereby detailed balance is readily recovered: a Markov chain thus ideally generates an equilibrium distribution. As we will witness below, however, things are not *that* simple.

2.1.1 The Metropolis algorithm

Replacing the weights in Eq. (2.1) by their actual Boltzmann value, one obtains $e^{(E_f - E_i)/kT} = W(f \rightarrow i)/W(i \rightarrow f)$, and transition probabilities must be designed which satisfy this equation. Metropolis's prescription, known as the *Metropolis acceptance rate* [246], reads

$$W(i \rightarrow f) = \min\left(1, \frac{w_f}{w_i}\right) = \min\left(1, e^{(E_f - E_i)/kT}\right), \quad (2.3)$$

It is straightforward to show that this acceptance rate is in fact valid for any distribution w . Basically, for a discrete spin model, the Metropolis algorithm works as follows:

1. pick a spin at random in the lattice (this operation is what will be referred to later on as “a priori” choice [211], or “proposed update”);
2. flip the spin (i.e., *attempt* a move *update*);
3. compute the energy change $E_f - E_i$, and then the transition probability $W(i \rightarrow f)$ (this is also termed the *acceptance rate*);
4. draw a number r at random between 0 and 1 and,
 - accept the attempted move if $r < W(i \rightarrow f)$, i.e., the next configuration in the Markov chain is the state σ_f ;

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

- otherwise reject the move, i.e., the next configuration in the Markov chain is the same as the current configuration, σ_i .

The previous set of operations is termed a *Monte Carlo step*. A Monte Carlo *sweep* (MCS) corresponds to N spins being updated on a lattice containing N spins.

Other choices of acceptance rate are possible, e.g., the *heat-bath* algorithm,

$$W(i \rightarrow f) = \frac{e^{-E_f/kT}}{\sum_j e^{-E_j/kT}}.$$

As opposed to the Metropolis transition rate, the acceptance rate depends only on the final state σ_f . This may speed up the dynamics for (nearest-neighbor) Potts model, and this is also the acceptance rate generally used to update bond configurations in the Swendsen-Wang cluster algorithm (Sec. 2.3.3).

Once a Markov chain of sufficient length has been generated, thermodynamic averages can be readily computed from direct averages, e.g. $\langle E \rangle = 1/N \sum_i E_i$ where $\{E_1, E_2, \dots\}$ denotes a set of measurements taken along the Markov chain (as we will see below, not every configuration generated should generally enter the average).

2.1.2 Detailed balance for moves that are not micro-reversible

In some situations (and particularly those reviewed below, and in Part III), it is more efficient to implement *move updates* that are not micro-reversible, i.e., for which the proposed update is not symmetric when going from i to f and back. We need to generalize Eq. 2.3 for this purpose. Proceeding along the line of [210], for instance, the transition probability is split into two terms, $W(i \rightarrow f) = P(i \rightarrow f)A(i \rightarrow f)$, where

- $P(i \rightarrow f)$ denotes a *proposed-update probability*, i.e., associated with the *choice* of a particular move among the set of every possible move;
- and $A(i \rightarrow f)$ is simply the *acceptance probability* for the proposed update.

It is trivial to show that

$$A(i \rightarrow f) = \min \left\{ 1, \frac{P(f \rightarrow i) w(E_f)}{P(i \rightarrow f) w(E_i)} \right\} \quad (2.4)$$

is a valid acceptance probability which satisfies the detailed balance equation. This will be referred to in the rest of this thesis as the *generalized Metropolis acceptance rate*. For a reversible move like the one implemented in the original Metropolis algorithm, Eq. (2.3) is recovered, of course, seeing that $P(i \rightarrow f) = P(f \rightarrow i) = 1/N$ where N is the number of spins in the lattice.

2.1.3 Random numbers generators

Random numbers are a fundamental ingredients of Monte Carlo methods, and the reliability of the results obtained from a Monte Carlo simulation depends heavily on the quality of the random numbers used by the Markov chain. Monte Carlo simulations generally rely on a *pseudo* random numbers generator (PRNG), i.e., a deterministic algorithm that generates random numbers in a way that emulates *true* random numbers¹, usually according to a uniform distribution. It is not the goal of the present work to go into the archanes of random numbers generators, and I will hereafter only provide some information regarding the generators I use in the simulations performed in this thesis.

The quality of a PRNG depends on several parameters: the distribution of the random numbers it generates must be as close as possible to the uniform distribution (or any other distribution it is assumed to engender), correlations between successive random numbers must be as low as possible, and it should be reasonably efficient since a great deal of random numbers are usually required at each Monte Carlo sweep. It is also important that the range of random values allows us to reach all the sites of a (large) lattice, which implies that the words must have a width of at least 32 bits.

The general scheme used by a PRNG is to generate a sequence $\{i_n\}$ of pseudo-random numbers from the recurrence equation $i_n = f(i_{n-1}, i_{n-2}, \dots, i_{n-k})$. The algorithm is initialized with an array of k numbers that is computed from a seed, and different seeds will produce distinct sequences. A feature common to these generators is that they will produce cycles: a good generator must offer a large period, preferably far larger than the total amount of random numbers that is required by the simulation (of the order of $L^D \times M$, where L^D is the volume of the lattice, and M is the number of MCS in the simulation). Otherwise, systematic errors may arise [270].

Several classes of generators have been devised in the last decades, that are based on the previous recurrence equation. Linear congruential generators (LCG) rely on the equation $i_n = (ai_{n-1} + c) \bmod m$, with a , c and m set to some “magic” values in order to achieve the longest cycle. Improved LCG’s combine a linear congruential method with a shuffling scheme, i.e., the numbers generated by the LCG are shuffled using a distinct generator. This reduces the correlation between numbers, and increases the cycle. More recently, a new class of generators termed “shift register” generators has surfaced: being based solely on logical shift and XOR operations (both are available in assembly language), they are usually faster than LCG’s [205]. These generators have been subsequently generalized to *lagged Fibonacci* generators (also termed *generalized feedback shift-register*), where the XOR operation is replaced by additions or multiplications, and offer very long periods (of the order of 10^{43} for the RANMAR generator). In this respect, it has recently been shown [270] that all these generators can cause systematic errors

¹For instance, those generated by a hardware device, such as the amplification of the Schottky noise a diode

using the Swendsen-Algorithm (see Sec. 2.3), if the lattice size is a multiple of a very large powers of 2 and one random number is used per bond. This stems from the synchronization between the periodicity of bond updates and the large, yet finite periodicity of the random number generator. Ossola and Sokal [270] proposed a work-around that consists in breaking the synchronization by updating bonds either in a random order, or in an aperiodic manner.

The *Mersenne-Twister* generator [243], the generator used in this thesis, uses a twisted generalized feedback shift-register algorithm which has a Mersenne prime period of $2^{19937} - 1$, or about 10^{6000} . This period is orders of magnitude larger than shift-register algorithms of the previous generation (for instance, the RANLUX generator [176, 236], which throws away numbers to destroy correlations, has a period of 10^{171}), and makes it one of the strongest uniform generator to date. It generates numbers in batches of 624 numbers at a time, thereby efficiently utilizing the caching mechanism of modern CPU's, and is extremely fast since only 32-bit additions and logical bit operations are used. It also passed many empirical tests, including the diehard test of Marsaglia [242], and the k -distribution test (every k -dimensional vector made up of successive numbers appears the same number of times over the period length, for each $k \leq 623$). In order to check for the presence of possible hidden correlations, I performed several *parking lot* tests [216]: a plane is filled with points of coordinates (i_n, i_{n+k}) (for several values of k), and hidden correlations will show up through striped patterns. For instance, the absence of such patterns for $k = 10$ is illustrated in Fig. 2.1. In order to check for possible systematic errors, I also performed some tests by simulating the two-dimensional short-range Potts model ($q = 2$ and $q = 3$, up to 128×128 spins), using the Swendsen-Wang cluster algorithm and two distinct generators (Mersenne-Twister with the default parameters given in [243], and RANLUX with the largest luxury level), and did not find discrepancies between the generators.

All these generators are available from the package `cern.jet.random.engine` of the Java COLT library released by the CERN. As most Java objects, they can store and retrieve their state through the serialization mechanism introduced in Java 1.1, which makes it possible to stop and restart a simulation at any point. As far as efficiency is concerned, the Mersenne-Twister generator is only slightly slower than the RANLUX and the RANMAR generators (approximately 1.8 to 1 when RANLUX is set to the highest luxury level), although its initialization time is markedly longer. With regards to the benefit brought about by the very large period length offered by this generator, this overhead is negligible.

2.2 Dynamic characteristics and statistical efficiency

2.2.1 Autocorrelation times

Markov chains do usually not engender statistically independent configurations. The amount of correlation between successive values of a given observable $A(\sigma)$

2.2. Dynamic characteristics and statistical efficiency

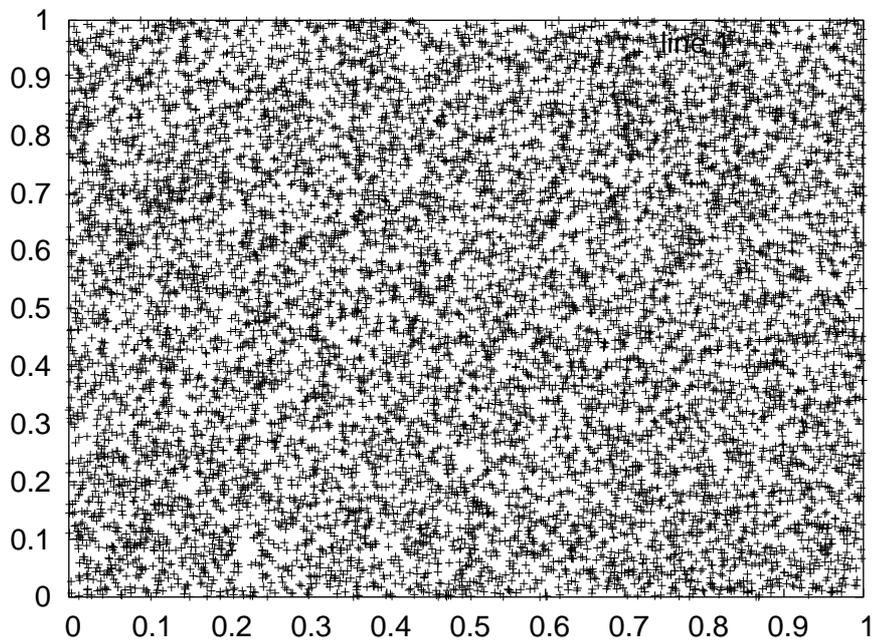


Figure 2.1: Parking lot test over $N = 10^4$ numbers $\{i_n\}$ for the Mersenne Twister generator: the graph is built from points of coordinates (i_n, i_{n+10}) , and displays no stripe that would be characteristic of hidden correlations.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

along a chain of length M is measured by a time-displaced correlation function

$$\chi_A(t) = \sum_{n=1}^{M-t} \{A_n A_{n+t} - \langle A \rangle^2\} \quad (2.5)$$

where A_n is the n th record of the value of the observable A , and $\langle A \rangle$ is the expectation value of A at equilibrium (see thermalization, Sec. 2.2.4). This function tells us to what extent measurements of A separated by t MCS are correlated. A convenient (though somewhat approximate) indicator of the average time-scale (if any) between independent measurements is the *autocorrelation time*. This can be obtained, either from the long-time exponentially decaying behavior $\chi(t) \propto e^{-t/\tau}$, or by integration, i.e., $\tau = \sum_{t=0}^M \chi(t) / \chi(0)$. Both definitions lead to the so-called *exponential* and *integrated* autocorrelation times τ_e and τ_i , respectively. These generally do not exactly coincide, owing to the fact that the exponential behavior is a simplified view of the real correlation mechanism, i.e., there are actually as many correlation times as there are states in the systems (one for each master equation 2.2), and what τ represents is actually the *largest* correlation time (whereby all other modes have died off in the long-time limit, see, for instance, [216]). As a rule of thumb, measurements should be taken at intervals of 2τ along the chain to form a sequence of *independent* measurements (see Appendix C).

From a practical point of view, $\chi(t)$ may be efficiently estimated by relying on a Fast Fourier Transform (FFT) algorithm [259], which reduces the burden from $O(M^2)$ to $O(M \ln M)$ operations. Denoting as $\tilde{A}(\omega)$ the Fourier transform of the sequence $\{A_i\}$, we have $\tilde{\chi}(\omega) = |\tilde{A}(\omega) - \tilde{A}(0)|^2$ (i.e., the zero-mode is set to zero, or, in other words, $\tilde{\chi}(\omega)$ is the discrete Fourier transform of $A_i - \langle A \rangle$). From there on, $\chi(t)$ is obtained by taking the inverse Fourier transform of $\tilde{\chi}(\omega)$.

2.2.2 Statistical efficiency

There is an innate source of error in every Monte Carlo simulation, that stems from the random nature of the Markov chain. Specifically, the source of error in the estimation of $\langle A \rangle$ by means of the *estimator* $\bar{A} = 1/M \sum_i A_i$, is the fluctuation of A_i from one step of the chain to the other. For a set of measurements associated with an autocorrelation time τ , the squared statistical error ϵ_A^2 on the estimated mean \bar{A} is given² by [252]

$$\epsilon_A^2 = \frac{1 + 2\tau}{M - 1} (\bar{A}^2 - \bar{A}^2) \sim \frac{2\tau}{M} (\bar{A}^2 - \bar{A}^2) \quad (2.6)$$

where $\bar{A}^2 - \bar{A}^2$ is the variance of A within a single chain (or Monte Carlo run), M designates the *total* number of measurements, and the second form applies when $\tau \gg 1$ and $M \gg 1$. For $\tau = 0$, i.e., perfectly uncorrelated measurements, this reduces to the well-known formula of elementary statistics. For $\tau > 0$, the number

²i.e., it is the variance of the mean from one Markov chain to the other.

2.2. Dynamic characteristics and statistical efficiency

of independent measurements is reduced by a factor $\sim 2\tau$ with respect to the total number of samples: this reduction factor is called the *statistical efficiency* of the simulation, the lower the better.

Equation 2.6 yields another way to estimate autocorrelation times: by computing the error on \bar{A} from multiple, independent runs, or by means of a blocking or resampling method (Appendix C); inverting the previous equation then allows one to retrieve τ . The corresponding time is termed the *effective* autocorrelation time, and is practically the sole autocorrelation time that makes sense in the context of multicanonical simulations [179]. This point will be reexamined in Sec. 7.4. There are other ways to estimate τ , for example, by combining the blocking method (Appendix C) with a renormalization group transform [120]³, or by explicitly analyzing autocorrelation functions [85].

2.2.3 Dynamic exponents for correlation times and critical slowing down

Critical and supercritical slowing down effects As we have seen in the previous paragraph, autocorrelation times are directly linked to the statistical efficiency of a Monte Carlo simulation. One of the core issues in the numerical study of phase transitions is related to the fact that autocorrelation times generally become very large in the vicinity of the transition temperature, and as a result make the simulation less accurate (or alternatively, require longer simulation runs). This behavior has two distinct origins, depending on whether we are investigating second-order or discontinuous phase transitions, and is respectively termed “critical slowing down” and “supercritical slowing down”. Overall, it is an innate (dynamic) feature of the model under investigation, in the sense that it is a direct consequence of the physics being studied, not of the algorithms in hand (although wise algorithm designs can significantly reduce it). Critical slowing down has incidentally been observed in many physical systems [158]. Supercritical slowing down is probably even more common, since its physical significance is directly related to the very existence of supercooled or superheated states in systems at first-order transitions.

The amount of critical slowing down depends on the actual algorithm being used, and it is the goal of a large class of algorithms introduced in the last fifteen years, and called “cluster algorithms”, to reduce this effect as much as possible. As for *super* critical slowing down, the benefit of cluster algorithms is less clear-cut [134], and other strategies must be used. Simulations in *generalized ensembles* represent such a strategy, and are considered in deeper detail in the next section.

Dynamic exponent and critical fluctuations Common to both types of transition is the fact that, beyond (and somehow correlated with) the sharp increase of correlation times at the transition, autocorrelation times also increase with the lattice size. In the context of second order transitions, a convenient indicator that gives

³This works only for a certain class of Hamiltonians, however.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

a way of quantifying this effect is the dynamic exponent z . As far as simulations of finite size systems are concerned, this exponent controls the behavior of autocorrelation times with respect to the system size, in much the same way as critical exponents control the divergence of (among other things) response functions at a critical transition. Here, it is perhaps interesting to mention that dynamic exponents can be associated with so-called *dynamic universality classes* in the same fashion as critical exponents determine static universality classes; this was established in the context of dynamic critical phenomena [158].

Critical slowing down is here largely reminiscent of the diverging fluctuations that show up at a critical point, where in systems of finite size, fluctuations are actually bounded by the large but finite correlation length. At a critical phase transition, large domains of predominantly up- or down- pointing spins tend to form as the temperature approaches the critical temperature from above. Inside these domains, we can identify one or more subsets of highly correlated spins — termed *clusters* — which, as they flip, generate large fluctuations in the energy and the magnetization. Since these fluctuations are linked to response functions through the fluctuation-dissipation theorem, we expect divergences in these functions also. Critical slowing down results from the fact that, in very much the same way as the characteristic length scale diverges at the critical temperature, the characteristic *time scale* also diverges. The determining role played by the spins belonging to the spanning cluster was clearly demonstrated, because these spins are perfectly correlated in the thermodynamic limit [282]. Note that, in the vicinity of a critical transition, the correlation length ξ behaves as $\xi \sim |t|^{-\nu}$, where $t = (T - T_c)/T_c$ is the reduced temperature, and $t \sim 0$ at the transition. In some cases, we may witness a logarithmic behavior $\tau \sim \ln |t|$ instead, yet this does not represent the majority of cases. To describe the divergence of the autocorrelation time, a new exponent z is defined so that $\tau \sim |t|^{-\nu z}$ in the vicinity of the transition, where τ is usually measured in MCS. Now, we also have that $\xi \sim |t|^{-\nu}$, thus $\tau \sim \xi^z$. Whence for finite-size systems and sufficiently close to the critical temperature, we obtain $\tau \sim L^z$ because the divergence is cut off in the region for which $\xi > L$.

Dynamic exponents as a powerful indicator of scalability Loosely speaking, z thus tells us what we may expect in terms of simulation accuracy whenever, for example, we double the size of the lattice. Since autocorrelation times determine, in the first place, the amount of CPU resources needed to reach a given accuracy, the knowledge of the dynamic exponent is thus crucial to a proper estimation of the scalability of a given algorithm. It must be mentioned that, as opposed to critical exponents, z is not a universal quantity, since its hinges heavily on what algorithm is being used. For instance, cluster algorithms generally yield lower z exponents than their Metropolis counterpart [153, 309], and are thus more efficient in this respect. To be specific, one of the most precise estimates of z in the case of the two-dimensional Ising model simulated with a Metropolis algorithm is $z = 2.1665(12)$ [264], while cluster algorithms yield $z \sim 0.25$ for the same model [83].

2.2. Dynamic characteristics and statistical efficiency

In terms of total CPU load, z clearly enters the overall scaling of CPU resources with the system size. For Metropolis implementations and models with finite-range interactions, we must perform the order of L^D operations per MCS (i.e., updating the whole lattice), hence the total CPU load behaves as L^{z+D} . For models with long-range interactions, the situation looks even worse, because L^D operations are needed to compute the energy change when a single spin is updated, and consequently L^{2D} operations are now required to update the whole lattice. This means that the total CPU load needed to reach a given accuracy scales as L^{2D+z} . Whether short- or long-range models are concerned, this shows that the Metropolis is certainly not the best algorithm when investigating critical phase transitions.

2.2.4 Scaling of relaxation times and the non-linear dynamic exponent

Autocorrelation times are not the only dynamic quantities that become very large at a critical transition. Relaxation times⁴, which inform us about the rate at which the Markov chain reaches the equilibrium distribution, are also known to exhibit the same scaling behavior (sometimes termed “critical relaxation”). Since what we aim at sampling in the first place is the equilibrium distribution, relaxation times thus clearly impinge on the accuracy of the simulation and, as a result, on the scalability of the algorithm. It turns out that algorithms that can reduce critical slowing down, e.g., cluster algorithms, are — on the same grounds — also better candidates with regard to relaxation behavior.

The approach to equilibrium of a given sampled quantity A can be described by the following non-linear relaxation function [216]

$$\Phi_A(t) = \frac{\langle A(t) - A(\infty) \rangle}{\langle A(0) - A(\infty) \rangle} \quad (2.7)$$

where $\langle \cdot \rangle$ stands for ensemble averages, i.e., averages over multiple, separate runs. Similar to (time-displaced) autocorrelation functions, this relaxation function displays a rough exponential decay with an associated (exponential) characteristic time scale τ , i.e., $\Phi_A(t) \sim e^{-t/\tau}$; τ may also be defined as an integrated relaxation (or equilibrium) time from the equation $\tau = \int_0^\infty \Phi_A(t) dt$. In practice, one may perform a single simulation, and then estimate τ by replacing $A(\infty)$ with an average over a subset of samples such that $t > \tau$, that is, in a self-consistent fashion. As a side note, let me mention that, as an exception to the exponential decay just mentioned, glassy systems and diluted models display a relaxation function that must be described by a more complex relation.

Relaxation times exhibit a scaling behavior similar to that experienced by autocorrelation times, yet with a different (non-linear) dynamic exponent z_{nl} . It has been established, however, that the dynamic exponent z and its non-linear counterpart z_{nl} are not independent. Instead, they are related to each other and to other

⁴Also termed *equilibration times*, *thermalization times*, and sometimes even *equilibrium times* [139]

(static) exponents through an equation similar in spirit to the hyperscaling relation [119] $z = z_{nl} + \beta/\nu$, where both dynamic exponents are computed from measurements of the magnetization, and β and ν refer to the usual critical exponents.

The non-linear dynamic exponent defined above retains its significance for simulations based on cluster algorithms; these, as already mentioned, lead to clear improvements over the Metropolis algorithm also in this area. However, and as will be observed in Chap. 7, non-linear dynamic exponents defined in this fashion become ill-defined where simulations in the multicanonical ensemble are concerned: first, the underlying dynamics of the Markov chain is all too different from the quite realistic dynamics engendered by canonical algorithms for the exponential decay behavior to remain valid; second, what we are interested in is no longer an average over raw samples, but a *reweighted* average. In this respect, I will show that it becomes necessary to define relaxation times with respect to the *broad* equilibrium distribution rather than to single averages, and that it is under this assumption that power-law scaling behaviors may be recovered.

2.3 Cluster algorithms

How do single-spin updates cope with critical slowing down? We have just seen in the previous paragraph that correlations between successive states generated by the Markov chain become large at a critical point, and thus severely endanger the scalability of the Metropolis algorithm. Indeed, for a given desired accuracy, this would imply running the simulation for a much longer time. It is above all in the critical region that the Metropolis algorithm is less accurate, because update moves based on single-spin flips are no longer physically relevant in this region. As I have mentioned, large clusters of correlated spins tend to form, with all spins lined up in the same direction, or, for Potts models, belonging to the same subphase. However, single-spin updates cannot flip a large domain, because in order to achieve this, they would have to flip a great deal of correlated spins one at a time, and this is associated with a tremendously low acceptance probability. Indeed, trying to flip a down-pointing spin in a sea of up-pointing spins costs an energy interface, e.g., of $4J$ for the two-dimensional Potts model, which leads to an acceptance probability of $e^{-4J/kT}$, i.e., around two per cent for the three-state version of this model at its transition temperature $T_c = J/\ln[1 + \sqrt{3}]$. For a cluster comprising N spins, the joint probability of shrinking the whole cluster is approximately given by $e^{-4NJ/kT}$, a quantity which scales in a quite unfortunate way indeed with the system size. Reworded differently, critical slowing down is particularly stark in the context of Metropolis algorithms owing to the drastically low pace at which the Markov chain explores the whole phase space when two successive states differ by at most one spin value. As will be witnessed in Chap. 7, multicanonical algorithms also experience critical slowing down: the wide phase space explored by the chain encompasses a “critical” energy range where correlations become large and — for the same reason as in the canonical case — single-

2.3. Cluster algorithms

spin updates deliver non-optimal dynamic performance.

Flipping clusters in one shot A solution to the previous problem consists in flipping groups of correlated spins in a single Monte Carlo step. Algorithms resorting to this principle are dubbed cluster-flipping algorithms, or simply cluster algorithms. Broadly speaking, such algorithms first inspect the lattice in search for groups of similarly oriented *and* correlated spins, and then flip them in one shot. These algorithms nearly entirely remove the critical slowing down at second-order phase transitions, and have thus proved extremely popular in the last fifteen years. This is, however, at the expense of an increased complexity, which should be pitted against the benefits brought in by the algorithm in terms of accuracy. As regards long-range models, it will become clear (see section 2.3.2) that naively transposing to long-range models cluster algorithms designed for short-range models, might tip the balance in favor of the Metropolis algorithm, or to say the least, make the issue unclear.

2.3.1 Wolff algorithm

Wolff algorithm for models with nearest-neighbor interactions The simplest strategy is to pick a spin at random, and then to look for neighboring spins that point in the same direction, and eventually flip the entire cluster. Since we want to restrict the group of spins to correlated spins only (i.e., to a *cluster*, which is a subset of a *domain*), and because these correlations depend on the temperature, there must be a temperature-dependent way to limit the addition of spins to the group. In other words, the cluster size must depend on the temperature. At high temperatures, for example, correlations between spins are scarce, which means that small clusters must be flipped. On the contrary, below the critical temperature T_c , the tendency of the system is to create clusters of spins spanning the entire lattice, and the algorithm should update large groups of spins in a row. The solution to this problem is to add a spin to the group with a probability depending on the temperature, in such a way that this probability increases with the inverse temperature. This is the rationale behind Wolff's cluster algorithm [347], otherwise termed the single-cluster algorithm, and to some extent, behind the related *generalized cluster algorithm* introduced by Niedermayer [261].

To set the stage, let us see how the algorithm works for a Potts model with nearest-neighbor interactions, i.e., governed by the following Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j)$$

where the sum runs over pairs of nearest-neighbor spins. I shall consider the extension to long-range Potts models in a later section, because such an extension is not trivial and deserves special scrutiny. As a first step, the algorithm chooses a cluster seed at random, that is, one spin among $N = L^D$ possible spins on the lattice, and assigns it a new value drawn at random; in what follows, the same value

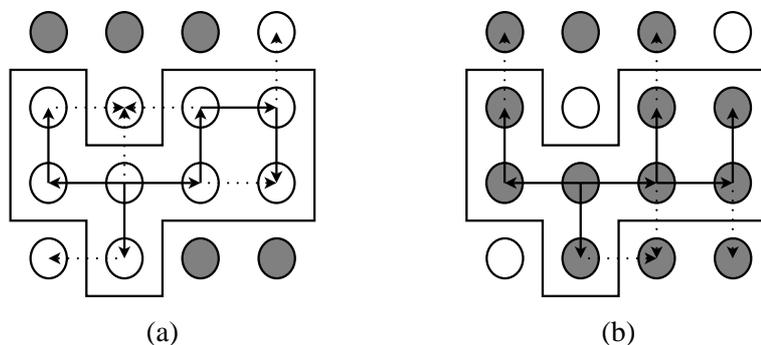


Figure 2.2: Construction of a Wolff cluster for a nearest-neighbor model, from state (a) and from state (b), where configurations (a) and (b) are related by a cluster flip. The solid lines correspond to activated (or satisfied) bonds ($C_a^i = C_b^i = 7$), and the dotted lines to non-satisfied bonds. In (a), there are two non-satisfied *edge*-bonds ($C_a^e = 2$, see explanation in text).

will be assigned to each spin being successfully added to the cluster. Then, each nearest neighbor of the seed is considered in turn for addition to the cluster, with a probability $p = 1 - e^{-\beta J}$, where β is the inverse temperature and J stands for the coupling constant of the model. From the viewpoint of practical implementation, each spin *actually* added to the cluster is concurrently pushed on top of a stack, for instance a first-in-first-out buffer. Once every nearest neighbor has been considered for addition, a new “seed” is popped from the stack and the algorithm proceeds with further spin addition until the stack is empty.

The cluster construction process is illustrated in Fig. 2.2 in the case of the two-dimensional Ising model. Whenever a spin is actually added to the cluster, we say that the corresponding *bond* has been *activated*, whereas in the contrary the bond is said *inactive*. In the following, I will denote as \mathcal{B} the set of active bonds. As can be witnessed in the figure, \mathcal{B} is actually a subset of the set of spin interactions belonging to the cluster. This simply means that some spins were added to the cluster after a first rejected attempt, although through a different path.

Detailed balance and cluster flip acceptance rate As has to be the case for any valid Monte Carlo algorithm, the Wolff cluster algorithm must satisfy detailed balance, or equivalently, the acceptance probability associated with the cluster flip must be constructed from Eq. (2.4). The proof is as follows. The true acceptance probability actually comprises two parts: a proposed-update probability, $P(a \rightarrow b)$ and an acceptance probability for the proposed update, $A_{\text{flip}}(a \rightarrow b)$. The proposed-update probability is the probability to choose a new state b from an existing state a , or equivalently the probability to construct a given cluster from a

2.3. Cluster algorithms

seed drawn at random. This probability reads

$$A(a \rightarrow b) = \prod_{l \in B} p_l(S) \prod_{l \notin B} (1 - p_l(S))$$

where l designates a bond, S is the associated spin pair, and $p_l(S)$ is zero whenever both spins differ, $1 - e^{-\beta J}$ otherwise. In Wolff's single-cluster approach, the first product runs over all pair of interacting spins until the construction stops, hence the second product runs over adjacent bonds to the cluster only. For a given cluster configuration constructed from state a , I now write C_a^i for the number of satisfied bonds, and C_a^e for the number of non-satisfied *edge*-bonds among pair of spins having identical values. The proposed-update probability may then be rewritten as:

$$A(a \rightarrow b) = p^{C_a^i} (1 - p)^{C_a^e}$$

I actually omitted a leading factor n/Nq , where n is the number of spins in the cluster, N is the number of spins in the lattice and q is the number of states of the model: this represents the probability to pick a particular seed inside the cluster, yet this term cancels out with that of the reverse move. Let me now consider the reverse move: as illustrated in Fig. 2.2, the proposed-update probability to select a from an initial configuration b writes

$$A(b \rightarrow a) = p^{C_b^i} (1 - p)^{C_b^e}$$

It is obvious that $C_a^i = C_b^i = C^i$ (although the distribution of active bonds over available bonds may differ in both configuration), and thus the ratio of proposed-update probabilities can be written as

$$\frac{A(b \rightarrow a)}{A(a \rightarrow b)} = (1 - p)^{C_b^e - C_a^e}$$

On the other hand, $J(C_a^e - C_b^e)$ represents exactly the energy change when going from a to b . Plugging the last equation into Eq. (2.4), we are led to the conclusion that cluster flips are always accepted. In other word, it is the choice of $1 - e^{-\beta J}$ as the bond activation probability which ensures that an appropriate distribution of cluster sizes is engendered, so that in the end the equilibrium distribution for the energy matches the Boltzmann weight.

It is crucial to note here that this *perfect* acceptance rate does no longer hold in the presence of an external field (or random-fields), except when the external field is an integral part of the cluster construction, as is the case in the ghost-spin algorithm.

Generalization to long-range models At first glance, the generalization of the above algorithm to long-range Hamiltonians is straightforward. Considering a long-range Potts model with interactions J_{ij} depending on the spin positions i and j , the construction is illustrated in Fig. 2.3, and the only change is that:

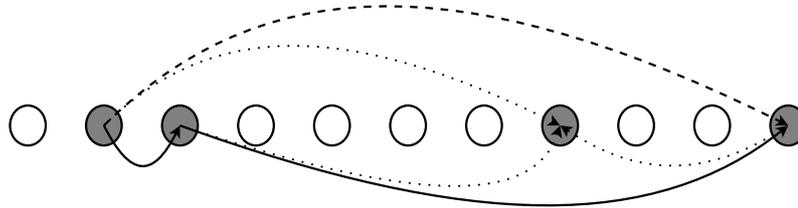


Figure 2.3: Construction of a Wolff cluster for a one-dimensional model with long-range interactions (this applies to any D -dimensional lattice, since the underlying graph is a complete graph in any case). Solid lines denote active bonds, dotted lines represent non-satisfied *edge*-bonds, while the dashed line corresponds to a non-satisfied inner bond which as such does not explicitly enter the *proposed-update* probability.

- Bond activation is considered for any pair of spins, irrespective of their being neighbors. The underlying graph is now a *complete* graph.
- Bond probabilities now depend on the coupling constant through $p_{ij} = 1 - e^{-\beta J_{ij}}$.
- A given cluster is no longer a block of contiguous spins with identical value, and may now span the entire lattice. In this regard, thermal clusters differ drastically from geometrical clusters in long-range models (in the figure, three geometrical clusters can be identified, although two of them belong to the same thermal cluster).

Algorithm complexity of the cluster construction for long-range models From a practical point of view, a seed is drawn at random, and then each spin *in the lattice* is considered for addition to the cluster. The construction process then goes on along the same line as with nearest-neighbor models, until the stack is empty, except that for each new seed, $\sim N$ bonds have to be checked. Thus it becomes clear that the cluster construction takes the order of N^2 operations for a lattice containing N spins. This represents an algorithm complexity substantially larger than that of nearest-neighbor cluster algorithms, and is incidentally reminiscent of the algorithm complexity associated with the computation of the energy in the long-range model.

The issue becomes particularly stringent for interactions which decay with the distance between spin, which is the case generally considered. Here, the probability of adding a bond between two spins falls off quite rapidly as the distance be-

2.3. Cluster algorithms

tween them increases. A significant amount of time during the construction of the cluster is thus wasted because an overwhelming number of bonds are considered for activation which have only a negligible probability to be activated. Even in the case of interactions decaying as $1/|i - j|^{1+\sigma}$ with σ close to 0, does the bond count never exceed a few percent of the whole number of available bonds. With regard to this, switching from a single-spin update algorithm to a cluster algorithm might well be an ill-fated choice, as the gain in terms of autocorrelation time is spoiled by the exceedingly time-consuming construction of the cluster. A work-around to this issue is considered in the next section.

2.3.2 Luijten-Blöte algorithm: efficient cluster construction

In this section, I describe an efficient construction method introduced by Luijten and Blöte [227, 226, 234], which works for Hamiltonian with any number of interactions between spins, and whose demand in terms of computer resources scales roughly as the volume of the lattice. The sole constraint of the method is that interactions should be invariant by translation along the lattice⁵, e.g., $J_{ij} = J(i - j)$ for a one-dimensional lattice. In Part III, I will show how this method can be straightforwardly combined with a multicanonical algorithm, and the material exposed below is therefore quite detailed.

The rationale behind Luijten-Blöte's method is to rely on *cumulative* probabilities, in the spirit of the N-fold-way algorithm [50, 212]. Instead of considering each spin in turn for addition to a given cluster and thereby experiencing numerous rejected attempts, it is the *index* of the next spin to be added which is drawn at random, either from a look-up table, or, as was proposed by Luijten in the context of Ising chains, by direct calculation.

First of all, the probability to add a bond is split up into two parts: (i) a provisional probability p_l depending on the distance $l = |i - j|$ between spins, and (ii) a factor $f(\sigma_i, \sigma_j)$ controlled by the spin values, e.g., a Kronecker delta symbol in the case of a Potts model, or a function of the angle between spins in the case of an XY model. If 0 designates the index of the current spin to which we are adding bonds (i.e., spin indices are considered to be relative to the current spin, see Fig. 2.4), then the provisional probability of skipping $k - 1$ spins and binding the current spin with a spin at position $k > 0$ is given by

$$P_0(k) = \prod_{m=1}^{k-1} (1 - p_m) p_k.$$

From there on, one builds a table of cumulative probabilities

$$C_0(j_1) = \sum_{k=1}^{j_1} P_0(k),$$

⁵Although the method works *in theory* also for non-invariant interactions, yet with a far higher demand in terms of memory resources, see the issue discussed farther regarding look-up tables.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

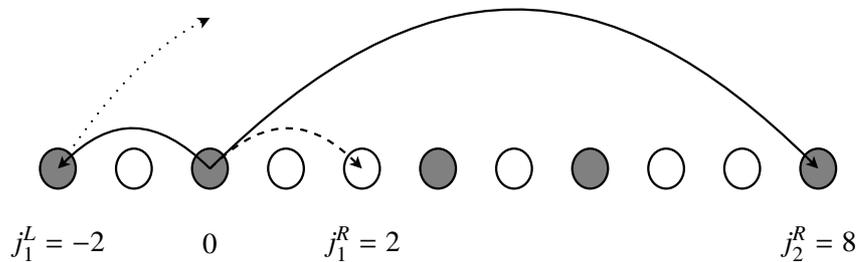


Figure 2.4: Construction of a Wolff cluster for a one-dimensional model with long-range interactions using cumulative probabilities $C_{j_\alpha}(j_{\alpha+1})$. The cluster seed, i.e., the “current” spin to which bonds are added is denoted as 0. The first spin *provisionally* added to the right is at relative position $j_1^R = 2$, i.e., the random number r is such that $C_0(1) \leq r < C_0(2)$. This spin is *actually* not added to the cluster (dashed line) since it does not have the correct sign (i.e., $f(\sigma_0, \sigma_2) = 0$, see explanation in text). A second random number such that $C_2(7) \leq r < C_2(8)$ yields $j_2^R = 8$. The corresponding spin has the correct sign, so that it can be added to the cluster (solid line). Alternatively, the random number r may be transformed to $r' = r[1 - C_0(2)] + C_0(2)$ and j_2^R obtained from $C_0(7) \leq r' < C_0(8)$. Then, the same procedure is performed for spins having negative relative indices: this yields $j_1 = 2$, i.e., $j_1^L = -2$. Finally, one proceeds further with the addition of spins to, e.g., spin $j_1^L = -2$, which becomes the new “current” spin 0 (dotted line).

2.3. Cluster algorithms

for all $j_1 > 0$, and $C_0(0) = 0$, so that the index j_1 of the spin to be bound with the current spin 0 is obtained by first drawing a random number $0 \leq r < 1$ and then reading out j_1 from the table. In other words, j_1 is such that $C_0(j_1 - 1) \leq r < C_0(j_1)$. Last, a bond is activated between spins 0 and j_1 with a probability $f(\sigma_0, \sigma_{j_1})$.

One proceeds further with the computation of the index $j_2 > j_1$ of the next spin to be bound with current spin 0. The corresponding provisional probability thus becomes

$$P_{j_1}(k) = \prod_{m=j_1+1}^{k-1} (1 - p_m)p_k,$$

and the cumulative probabilities read

$$C_{j_1}(j_2) = \sum_{k=j_1+1}^{j_2} P_{j_1}(k).$$

The same procedure is repeated for $\{j_3, j_4, \dots\}$ until an index $j_\alpha > L$ is drawn⁶, in which case we jump to the next current spin, which in a one-dimensional model is the nearest-neighbor of the previous current spin.

The efficiency of the algorithm is greatly improved by using two formulas which make it easier to compute cumulative probabilities. First, it can be shown that the cumulative probability $C_0(j)$ may efficiently be computed from

$$C_0(j) = 1 - \exp \left[-\beta \sum_{k=1}^j J(|k|) \right],$$

Second, the cumulative probabilities $C_{j_\alpha}(j_{\alpha+1})$ can be straightforwardly derived from the $C_0(j)$ coefficients through the relation

$$C_{j_\alpha}(j_{\alpha+1}) = \frac{C_0(j_{\alpha+1}) - C_0(j_\alpha)}{1 - C_0(j_\alpha)}.$$

It follows from the last relation that, instead of building a look-up table for each $C_{j_\alpha}(j_{\alpha+1})$, one may as well draw a random number $0 \leq r < 1$, transform it to $r' = r[1 - C_0(j_\alpha)] + C_0(j_\alpha)$, and choose the next spin to be added from the relation $C_0(j_{\alpha+1} - 1) \leq r' < C_0(j_{\alpha+1})$. In practice, one thus simply needs to compute a single look-up table $\sum_{k=1}^j J(|k|)$ for $0 < j < L$ at the beginning of the simulation, from where the $C_0(j)$ coefficients are immediately derived.

In the case of $1/r^\alpha$ interactions, Luijten proposed an efficient way to get around the (somewhat tedious) task of looking up a precomputed table⁷, which consists

⁶This conditions implies that infinite-image periodic boundary conditions are included in the coupling constant, see Sec. 4.2.1; otherwise, the actual spin index is $j_\alpha \bmod L$, i.e., periodic boundary conditions are taken into account in the course of the cluster construction. The former approach is more efficient in one-dimension, because an exact calculation of the “renormalized” coupling is possible. The latter, however, proves useful in higher dimensions.

⁷There are efficient binary-search algorithms for this purpose, yet even so this takes computer resources.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

in replacing the discrete sum $\sum_{k=1}^j J(|k|)$ by an integral and carrying out a direct calculation of the spin index. It was safely argued that this does not change the *universal* properties of the model, yet only non-universal quantities like critical couplings. This entails, however, that this kind of optimization is intractable for models experiencing first-order transitions.

2.3.3 Swendsen-Wang algorithm

As opposed to the single-cluster algorithm, Swendsen-Wang's algorithm [309] looks for *all* clusters, including those reduced to a single site, and then flips them all independently at random, that is, assigning a new value to each cluster. Historically, this so-called multi-cluster algorithm was introduced before Wolff's algorithm, and the latter algorithm may actually be viewed as a particular case of the multi-cluster algorithm. The single-cluster algorithm, however, was reported to work better in higher dimensions, at least for nearest-neighbor interactions [282]; for the Ising model in $D = 4$, for instance, Coddington and Baillie [83] reported $z = 0.25(1)$ for Wolff's algorithm vs $z = 0.86$ for Swendsen-Wang's algorithm. The Swendsen-Wang algorithm can be best envisioned in the context of bond-percolation, which I review hereafter.

Fortuijn-Kasteleyn random-cluster mapping for long-range models In the early 70's, Fortuijn and Kasteleyn [122] showed that spin models with a \mathbb{Z}_n symmetry can be mapped onto a bond-percolation model, the so-called *random-cluster model*. In its first flavor, termed the *spin-bond* representation, this model is made up of thermally fluctuating bonds and spins, whereas in its second version, an appropriate integration over spin degrees of freedom yields a *pure bond* representation. Both representations form the basis core of multi-cluster algorithms, though they are definitely not on an equal footing where concrete implementation is concerned.

The random-cluster mapping was initially contrived with nearest-neighbor models in view. However, since taking long-range interactions into account does not pose any major difficulties with respect to the original line of arguments, I will readily consider in the following a long-range potential. To be specific, I will take a generalized *ferromagnetic* Potts model with a coupling constant J_{ij} depending on positions i and j on the lattice ⁸, i.e., with an Hamiltonian reading $H = -\sum_{i<j} J_{ij}\delta_{\sigma_i,\sigma_j}$. At a given inverse temperature β , the partition function of the model reads

$$Z = \sum_{[\sigma]} e^{\beta \sum_{i<j} J_{ij}\delta_{\sigma_i,\sigma_j}}$$

where the sum runs over all spin configurations and $J_{ij} > 0, \forall(i, j) \in \mathbb{Z}^2$. The key idea consists in reexpressing Z in terms of a trace over bonds, where each bond lives on an edge of the *complete* graph \mathcal{G} engendered by the set of all possible

⁸Note that, as opposed to the Luijten-Blöte cluster construction, translational invariance is not a prerequisite here.

2.3. Cluster algorithms

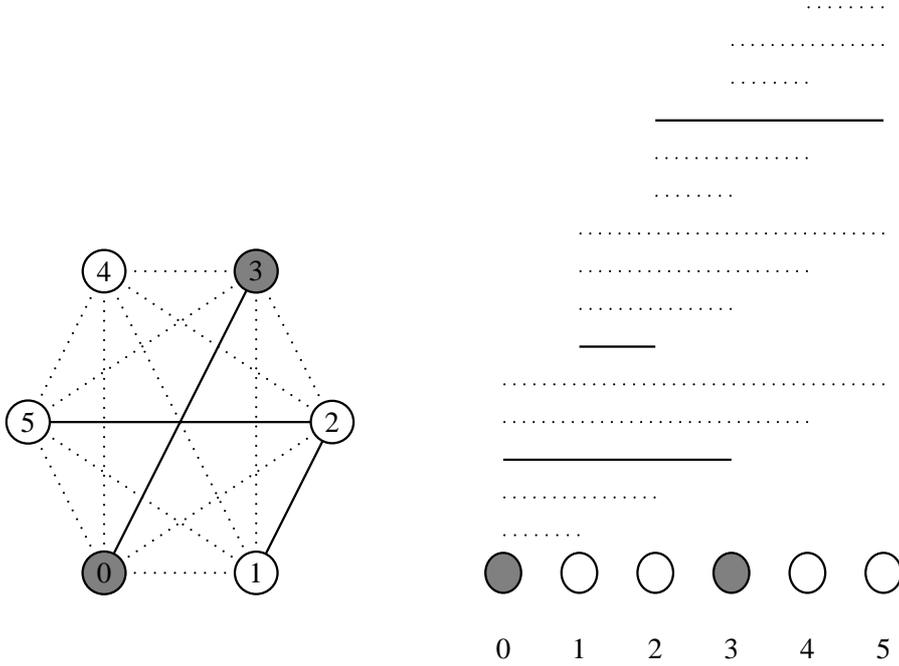


Figure 2.5: Tree-like (left) and flattened (right) representations of a complete graph \mathcal{G} for a two-state Potts model (i.e., the Ising model) with long-range J_{ij} interactions and $L = 6$ spins. Solid and dotted lines correspond to active ($b = 1$) and inactive ($b = 0$) bonds, respectively. The corresponding bond configuration engenders three connected components, each being associated with a different cluster in the Swendsen-Wang algorithm: $\{0, 3\}$, $\{1, 2, 5\}$ and $\{4\}$.

interactions, that is, such that the graph vertices are in a one-to-one correspondence with the lattice cells.⁹

Let b_{ij} be a binary variable associated with the edge joining vertices i and j , such that $b_{ij} = 1$ if a bond is active on this edge, and $b_{ij} = 0$ otherwise. Let also $p_{ij} = e^{\beta J_{ij}} - 1$. Then the partition function may be reexpressed as

$$Z_{FK} = \sum_{[\sigma]} \sum_{[b]} \prod_{i < j} (p_{ij} \delta_{\sigma_i, \sigma_j} \delta_{b_{ij}, 1} + \delta_{b_{ij}, 0}) \quad (2.8)$$

where the second sum runs over all *bond* configurations. In order to show that both expressions of the partition match, it is enough to limit ourselves, to set the stage, to a graph reduced to a single bond, say p_{01} : the trace over bonds yields $p_{01} \delta_{\sigma_0, \sigma_1} + 1$, that is, $e^{\beta J_{01} \delta_{\sigma_0, \sigma_1}} = e^{-\beta H}$. Coming back to the general case, it is clear that, since the argument of the product in Z_{FK} depends on a single bond, the same line of argument as above can be carried out by first swapping the product

⁹The engendered graph is complete if and only if $J_{ij} \neq 0, \forall i, j$. This is the case for power-law decaying interactions, yet not for interactions restricted to nearest neighbors or for medium-range interactions.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

over pairs and the sum over bond configurations, and then tracing over each bond independently. Clearly, there is no requirement as for the explicit form of J_{ij} in the argument above.

The representation of the partition function in terms of both traces over spin *and* bond configuration underlies the spin-bond representation of a Potts model with long-range interactions. This is the basis of the Swendsen-Wang algorithm presented in the next section. This will also be an important ingredient of a novel Monte Carlo method, the breathing cluster method. This method represents a salient part of this thesis, and is described in detail in Part III.

By integrating over spin degrees of freedom, one may derive another interesting representation of the partition function in terms of trace over bonds only,

$$Z_{RC} = \sum_{[b]} \prod_{i<j} p_{ij}^{b_{ij}} q^{N_c}, \quad (2.9)$$

where q is the number of states of the model and N_c stands for the number of connected subgraphs (i.e., the number of clusters in Swendsen-Wang's parlance). This last expression may be readily obtained by first swapping the traces over bonds and spins, and then observing that, for a given bond configuration, the trace over spin configurations amounts to assigning to each connected component one spin state among q distinct values in turn. Indeed, it should be noted that spin configurations that are incompatible with a given bond configuration do not contribute to the trace.

Application to the Swendsen-Wang algorithm Swendsen-Wang's algorithm consists in simulating the system in an extended phase space comprising the original spin degrees of freedom *and* the bond degrees of freedom. The weight of a given $[b, \sigma]$ configuration in this extended phase space is simply given by

$$w([b, \sigma]) = \prod_{i<j} (p_{ij} \delta_{\sigma_i, \sigma_j} \delta_{b_{ij}, 1} + \delta_{b_{ij}, 0}).$$

The original implementation of the authors puts these degrees of freedom on an equal footing, that is, each Monte Carlo step comprises two stages: first, *bonds* are updated from the current spin configuration, and then, *spins* are updated under the constraint imposed by the current bond configuration. The second stage is rather straightforward, and actually amounts to assigning a new value to each cluster (i.e., connected subgraph) separately. The associated acceptance rate is equal to one, since the weight of a given $[b, \sigma]$ configuration and the weight of the same configuration with any cluster being assigned a new value, are identical. As for the bond configuration update, we have two schemes at our disposal.

First, we may rely on the Metropolis acceptance rate, Eq. (2.4), i.e., the probability to activate an initially open bond is given by $\mathcal{P}(b_{ij} : 0 \rightarrow 1) = \min(1, p_{ij})$, whereas the reverse move probability is $\mathcal{P}(b_{ij} : 1 \rightarrow 0) = \min(1, p_{ij}^{-1})$. For a nearest-neighbor model where $p_{ij} = e^{\beta} - 1$, where the (i, j) edge is restricted to

2.3. Cluster algorithms

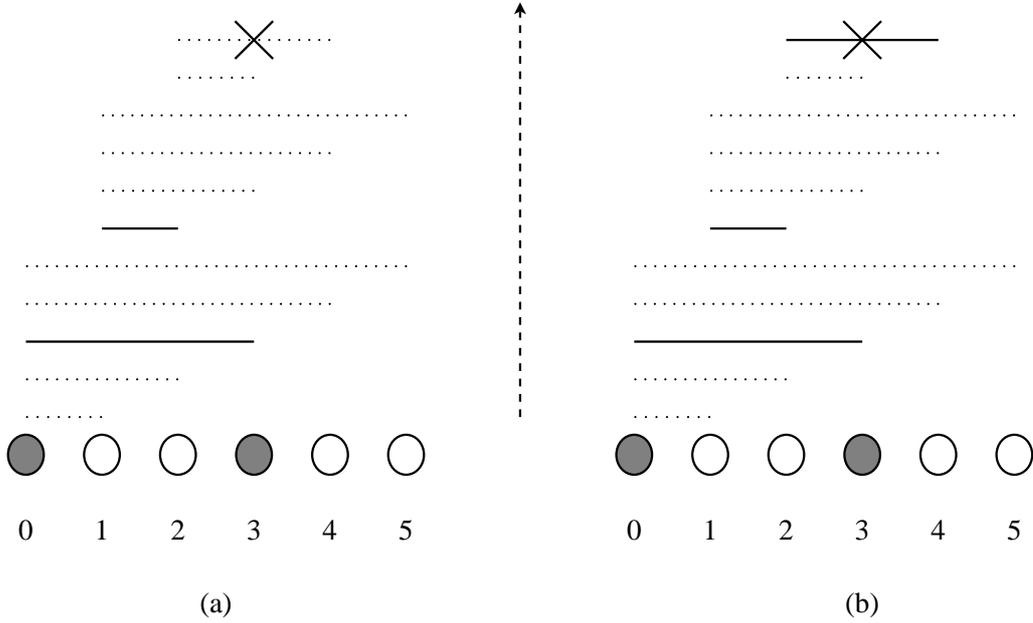


Figure 2.6: An illustration of a bond update in the flattened representation of the complete graph \mathcal{G} , for the Swendsen-Wang algorithm applied to a two-state Potts model with long-range J_{ij} interactions. The vertical dashed arrow shows the order in which bonds are considered for activation, starting from an empty bond set (i.e., all bonds inactive). The bond between spins 2 and 4 may be left inactive with probability $e^{-\beta J_{24}}$ (a) or activated with probability $1 - e^{-\beta J_{24}}$ (b), see Eq. (2.10). The procedure continues until every bond has been considered for activation.

nearest neighbors, bonds are thus always activated whenever $\beta > \ln 2$, and conversely, deactivated for $\beta < \ln 2$. The smaller probability associated with the reverse move clearly yields the correct weighting. By construction, this approach implies that bond configurations be kept in memory between each spin update.

Another scheme, which proves particularly useful whenever one prefers to rebuild the cluster *from scratch* after every spin update, is the heat-bath algorithm depicted in Fig. 2.6. The implementation is particularly straightforward here, because we have only two possible final states: $b_{ij} = 0$ (state A), and $b_{ij} = 1$ (state B). In the heat-bath algorithm, the probability associated with each final state is independent of the initial state, and is simply made proportional to the weight of the final state, i.e., $\mathcal{P}(b_{ij} = 1) \propto p_{ij}$ and $\mathcal{P}(b_{ij} = 0) \propto 1$. After proper normalization, this yields

$$\mathcal{P}(b_{ij} = 1) = \frac{p_{ij}}{1 + p_{ij}} = 1 - e^{-\beta J_{ij}} \quad (2.10)$$

which is analog to Wolff's bond probability.

Long-range interactions In the case of power-law decaying interactions, the activation probability rapidly becomes negligible as the distance increases, and the same optimization scheme as in the Luijten-Blöte algorithm must be performed. The main change with respect to Luijten-Blöte’s formulation is that the construction of the cluster now consists in (i) choosing a “current” spin among $L-1$ possible spins *in turn*, e.g. starting from the leftmost one, and then (ii) activating bonds between the current spin and other spins located to its right (in one dimension) by drawing a random number, scaling it, and selecting the bond indices from a look-up table containing the $C_0(j)$ coefficients at energy E . Once each spin has been considered as a “current” spin, a cluster multiple labeling technique can eventually be used to identify every set of spins actually belonging to the same cluster [161]. In dimensions greater than one, care must be taken to consider every pair of spins only *once*: in two dimensions, for instance, the look-up table must exclusively contain indices of spins located in a half-plane $\theta \in]-\pi, \pi]$.

2.3.4 Extensions to other models

Cluster algorithms have been generalized to a broad class of models, yet with varied success. In his seminal paper [347], Wolff considered the extension to O_n models with continuous symmetry breaking, e.g., the XY model, by implementing a so-called *embedding trick*, whereby XY spins are embedded in Ising variables through a projection onto a direction chosen at random at each Monte Carlo step. A further generalization to ϕ^4 field-theoretic models was presented by Brower [57], and to σ -nonlinear models by Carracciolo et al. [70]. Antiferromagnetic models can also be handled by cluster methods, provided bonds are activated between spin of opposite signs. The antiferromagnetic Potts model, in particular, was considered in [340], using a “freezing” mechanism, whereby for a q -state model, $q-2$ phases are (temporarily) frozen and the cluster is built from the remaining two phases.

Cluster algorithms in their initial formulation were claimed to work rather poorly for disordered systems, and there have been numerous attempts in the last decade at designing novel cluster algorithms that are (more or less) specifically tailored to this class of systems. The random-field Ising model, for instance, was considered in [258]: here, the “pinning” of large domains due to the random fields is circumvented by flipping clusters of limited radius, with a distribution of radius following a power law. Extensions of the multi-cluster algorithm to simple frustrated models were considered in [197, 198, 84], although these were reported to work only poorly for spin glasses. In this regard, a promising direction turned out to be *hybridation*, e.g., a combination of cluster algorithms with methods working in expanded or generalized ensembles. The two-replica cluster method [238, 283] and further extensions thereof [162, 194] is an example in this class, which combines the replica exchange method [308] with a cluster algorithm that activates bonds between two distinct replicas as a function of their mutual overlap. The method bears resemblance to simulated tempering (Sec. 2.5.6), yet what is swapped is the magnetization instead of the temperature.

2.3. Cluster algorithms

Cluster algorithms have also been applied to non-magnetic systems, e.g., to hard sphere mixtures [103] where “clutches” of spheres are grown then swapped through a geometric transform (e.g., a reflection) in a cluster-flip like fashion. The algorithm was extended to nearest-neighbor spin systems by Herringa and Blöte [155], whereby two clusters of opposite sign are grown then swapped through a geometric transform, instead of being flipped through a symmetry operation¹⁰. In the presence of long-range interactions competing with short-range ferromagnetic couplings, a generalization was proposed in [138], whereby clusters are grown as if the model were governed by the short-range part only, yet with a reduced bond probability that in effect limits the cluster size (for this purpose, the author makes use of a so-called effective temperature that is tweaked in order to yield the highest acceptance ratio); the presence of long-range interactions is subsequently taken care of by accepting the cluster-swap with a probability lower than one and depending on the long-range part.

Finally, it is worth mentioning the invaded cluster algorithm [317, 267], which is able to find out the critical temperature by tuning the cluster probability (through a feedback mechanism) until cluster percolate (this implies that the algorithm works for models where the percolation threshold and the critical point coincide).

2.3.5 Niedermayer’s construct, improved estimators and cluster distributions

Approximately at the same period where Wolff introduced his single-cluster algorithm, Niedermayer presented a very general formalism for cluster algorithms, the general cluster updating method [261, 262], which in a sense unifies Swendsen-Wang’s and Wolff’s formulations. Niedermayer showed that, provided a global (either discrete or continuous) symmetry underlies the model Hamiltonian, the original system of spin variables can be mapped onto a new system of “macro-spin” variables; each macro-spin represents a cluster, and interacts with the other macro-spins through an interaction which is directly related to the bond probability, in a way that ensures detailed balance. The key point is that there is some degree of freedom in the choice of the bond probability p which was not present in Wolff’s original formulation, and which allows the average cluster sizes to be altered. The prescription, if one requires that every macro-spin flip (through the underlying symmetry operation) be *always* successful, writes

$$p_{ij}(\sigma_i, \sigma_j) = 1 - e^{E_{ij}(\sigma_i, \sigma_j) - Q_{ij}},$$

where $E_{ij}(\sigma_i, \sigma_j)$ is the interaction between spins σ_i and σ_j , and Q_{ij} must be chosen greater than the maximum of E_{ij} over all possible symmetry operations involved in the macro-spins flipping; for the Ising model, for instance, one obtains $Q_{ij} \geq J_{ij}$, where Wolff’s algorithm is recovered for $Q_{ij} = J_{ij}$, and larger values

¹⁰This is akin to the two-replica cluster algorithm [238], yet with exchange taking place *inside* the same system.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

increase the average cluster size by explicitly allowing bonds between spins of opposite sign. The previous condition on Q_{ij} ensures that flipping macro-spins is always associated with a unit probability; other choices (which would lead, e.g., to a larger bond probability for identical spins, but to a null probability for spins of opposite sign) are perfectly possible, yet they impose that a correction factor enters the Metropolis acceptance rate [222]. A correction term is also needed in the presence of an external field, e.g. in the random-field Ising model, because the field breaks the symmetry of the Hamiltonian (and it is no longer possible to find an “optimum” value of Q which yields a unit flipping probability).

Niedermayer’s formulation also makes it straightforward to construct *improved estimators* from the macro-spin variables¹¹, namely, estimators with the same mean but *reduced* variance. For a q -state Potts model, the two-point correlation function¹² is $\langle q\delta_{\sigma_i,\sigma_j} - 1 \rangle / (q - 1)$, and $\langle \delta_{\sigma_i,\sigma_j} \rangle$ may be (better) estimated from

$$\sum_{\mathcal{B}} \mathcal{P}(\mathcal{B}) \langle \delta_{\sigma_i,\sigma_j} \rangle_{\mathcal{B}},$$

where \mathcal{B} is a given bond distribution yielding particular macro-spin variables, and occurring with a probability $\mathcal{P}(\mathcal{B})$, and the macro-spin correlation function $\langle \delta_{\sigma_i,\sigma_j} \rangle_{\mathcal{B}}$ is an average over the q^{N_c} possible macro-spins configurations (for a q -state model). By symmetry, the latter quantity reduces to 1 if both spins belong to the same cluster, and $1/q$ otherwise. In practice, one just has to implement a data structure that memorizes at each Monte Carlo step and for each (i, j) pair whether the corresponding spins belong to the same cluster or not.

Other improved estimators can be constructed from the *largest cluster distribution*. Denoting as l the size of the largest cluster in a given bond configuration, and $P_L(l)$ the distribution of the largest cluster of size l , one may define improved estimators for the moments of the order parameter from the moments of $P_L(l)$, i.e., $\langle l^n \rangle = \sum_l l^n P_L(l)$ is an improved estimator for $\langle m^n \rangle$. In particular, the largest cluster distribution has an associated (modified) Binder cumulant $B_L = \langle l^4 \rangle / \langle l^2 \rangle^2$ such that $B_L(t) = f(L^{1/\nu}t)$, where $t = (T - T_c)/T_c$ is the reduced deviation to the critical temperature and ν is the critical exponent of the correlation length. Since B_L is invariant under a renormalization group transform [35], the critical temperature can be located by monitoring the crossing point of these cumulants at increasing lattice sizes.

In two recent Monte-Carlo studies focusing the q -state long-range Potts models [326, 325], improved estimators were ingeniously exploited to yield information regarding the Kosterlitz-Thouless-like transition that occurs for $1/r^2$ decaying interactions, a transition which is notably difficult to investigate owing to the presence of essential singularities in the correlation length and the susceptibility.

¹¹Any cluster construction actually allows one to construct these estimators, yet this was first put into words by Niedermayer.

¹²Janke and Kappler use another normalization factor [181, 184].

2.4. Reweighting methods: single- and multi-histograms

Finally, it is worthy of mention that the entire cluster size distribution (as opposed to the *largest* cluster distribution) may provide useful insight into the critical properties of a model. Here, a connection with fractal geometry was recently proposed [13, 190, 188, 189]. This distribution was also suggested to be an efficient way of estimating the correlation length of the ordered phase in a first-order transition [277], by monitoring the cluster size above which the power-law behavior of the distribution¹³ breaks down.

2.4 Reweighting methods: single- and multi-histograms

As already stated in the introduction, canonical simulations rely on a Markov chain weighted by a Boltzmann factor $e^{-E/kT}$, and thermodynamic averages can be straightforwardly estimated by averaging over the engendered sequence of states. Naturally, this implies that these averages are obtained at a *single* temperature point. Reweighting methods aim at getting around this limitation by providing a way to obtain averages at other temperature points, and as *continuous* functions of the temperature. In this respect, reweighting methods are not specific to canonical simulations, and simulations in generalized ensembles (see next section) actually constitute their favorite playground. Their *realistic* application to the study of critical phenomena first appeared, however, in the context of histogram methods [113, 114], which I briefly review hereafter¹⁴.

The idea behind the single-histogram method [113] is that the sampled distribution of the energy provides in fact much more than mere thermodynamic averages: it also offers a way to estimate the density of states, yet as we shall see, in a narrow energy window. Indeed, the energy distribution (which is estimated through an *histogram* of the energy; hence the name of the method) is given by $N(E) \propto n(E)e^{-\beta_0 E}$, where β_0 is the simulation temperature. Inverting this relation yields (an estimate of) the density of states $n(E) \propto N(E)e^{\beta_0 E}$. From there on, any moment of the energy may (in theory) be computed at any temperature point $kT = 1/\beta$ through the reweighting equation

$$\langle E^n \rangle(\beta) = \frac{\sum_E E^n N(E) e^{\beta_0 E - \beta E}}{\sum_E N(E) e^{\beta_0 E - \beta E}}, \quad (2.11)$$

where the denominator is, as a by-product, an estimate of the partition function Z , and the free energy may be obtained (up to an additive constant $\propto T$) from $F = -kT \ln Z$. This equation can be generalized to any thermodynamic variable A whose microcanonical average $\langle A \rangle_E$ is known (which just entails recording A during the simulation),

$$\langle A \rangle(\beta) = \frac{\sum_E \langle A \rangle_E N(E) e^{\beta_0 E - \beta E}}{\sum_E N(E) e^{\beta_0 E - \beta E}}. \quad (2.12)$$

¹³This power-law behavior can be thought of as stemming from the scale invariance “felt” by droplets when their size is smaller than the finite correlation length.

¹⁴See, also, [180] for an extensive review that also encompasses applications to multicanonical and tempering methods.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

In terms of sums over measurements, this can be reexpressed as

$$\langle A \rangle (\beta) = \frac{\sum_i A_i e^{\beta_0 E_i - \beta E_i}}{\sum_E N(E) e^{\beta_0 E - \beta E}}, \quad (2.13)$$

where A_i and E_i refer to the i th measurement of A and E , respectively.

In the context of long-range models, the method was used to study the q -state Potts model [19]. While it is ideal for locating peaks of response functions at a critical point, it suffers from flawed reweighting far from $1/kT_0$, because the histogram has too few entries far from the energy region sampled by the (usually gaussian) canonical distribution. This point will be further discussed in the next section in the context of generalized ensembles simulations introduced in the early 90's, and which represent very efficient techniques to get around the issue.

In the meantime, an extension of the previous single-histogram method was proposed, that aims to improve the reliability of reweighting over a larger temperature range. The *multi-histogram* method [114] relies on a set of simulations at distinct yet nearby temperatures points, so that energy distribution overlap and the resulting “compound” histogram spans a larger energy range. The reweighting equation now involves all histogram data, and in order to minimize errors on the estimation of the density of states, histograms should appear in the equation with a weight that is proportional to the number of *independent* measurements from which they were filled, or alternatively, to the inverse of the variance of the autocorrelation time of the corresponding data bag (if, say, raw, non-independent data are used). The procedure proposed by Ferrenberg and Swendsen is in fact slightly more involved: the previous rule is used only as a starting guess, and weights are refined in a self-consistent manner until the minimum error is reached. Even so, while the multi-histogram was shown to improve the accuracy of reweighted averages in a dramatic way with respect to single histogramming, it will be seen below that it is no longer the best method where efficiently sampling a large energy window is the goal.

2.5 Simulation in generalized ensembles

2.5.1 From rare-events sampling to flat-histogram algorithms

As was argued in the previous section, the canonical distribution $w(E) = e^{-\beta E}$ is definitely not the best choice where estimating the average of certain operators is concerned, including (but not limited to) measurements of free energy differences. This can be traced back to the fact that, in general, configurations that dominate in reweighted thermodynamic averages are generated with negligible probability. To be specific, if we consider the reweighted average of an operator A at a given inverse temperature β_0 and from a set of measurements $\{A_i\}$ generated from a canonical distribution, i.e.,

$$\langle A \rangle (\beta_0) = \frac{\sum_i A_i w^{-1}(E_i) e^{-\beta_0 E_i}}{\sum_i w^{-1}(E_i) e^{-\beta_0 E_i}},$$

2.5. Simulation in generalized ensembles

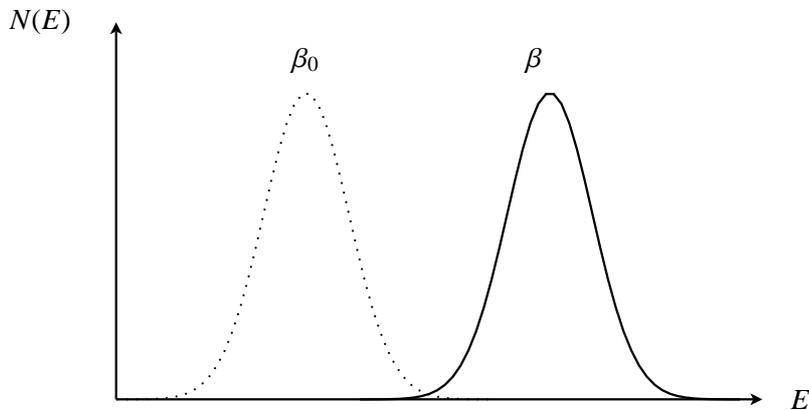


Figure 2.7: Histogram of the energy filled with the samples produced by a canonical simulation at inverse temperature β (solid line), and what the same histogram should *look like* at inverse temperature β_0 . Both histograms overlap only poorly, and the resulting accuracy of reweighted averages at β_0 is thus very low.

then the measurements that contribute predominantly to the sum in the numerator are actually those associated with an energy $E_i \sim \langle E \rangle(\beta_0)$. If β differs too much from β_0 , then the tail of the Boltzmann distribution is very unlikely to produce the appropriate samples, at the expense of poor accuracy (see Fig. 2.7). This is all the more stringent at large lattice sizes, that the Boltzmann distribution becomes narrower, seeing that its width scales with the inverse of the specific heat. Multi-histogram methods [114] are of little help for that matter, because the increasing narrowness of the energy window that is sampled requires that more and more simulations are performed at ever closer temperatures¹⁵.

As mentioned in the introduction of this chapter, the Metropolis algorithm may, in theory, sample any distribution. By choosing a distribution that gives equal importance to each energy windows that enter the sum in the reweighted averages, one may clearly increase the accuracy of the averages over a much wider range of temperature. This is the key idea underlying what was first coined as "non-Boltzmann sampling" (in the context of umbrella sampling) [318], and later disseminated to the community as "generalized ensembles Monte Carlo"¹⁶. Generalized ensembles methods currently cover a variety of algorithms, from transition matrices [341] to multicanonical sampling [25] and variants thereof [335] to the recently introduced optimized ensemble method [319, 351]. I will give a survey of these algorithms in subsequent parts of this section, yet before moving in to more technical detail, I

¹⁵The same pitfall occurs in simulated tempering, although the last algorithm was designed with a different purpose in mind.

¹⁶One may also encounter the terminology "extended/expanded ensembles Monte Carlo" now and often, see for instance [166], although this usage is generally restricted to methods related to simulated tempering, i.e., with an enlarged configuration space (hence the name) containing additional dynamic variables like the temperature.

Chapter 2. *From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?*

would like to address another thorny issue that is tightly related to the concept of generalized ensembles: rare events sampling.

In recent past, as a matter of fact, the attention of the community gradually shifted from the problem of sampling *broad* energy windows to the subject of sampling rare events. This subject slowly surfaced as a result of unabated interest for a variety of physical systems exhibiting rugged free energy landscapes, including heteropolymers and protein folding [147], random field models or spin glasses [26, 29, 156, 27], yet also (and in spite of their apparent simplicity as against the previous models) models exhibiting a discontinuous phase transition [31]. Canonical methods face some severe drawbacks here as well, this time owing to a so-called *supercritical slowing down* effect [31] which manifests itself through highly repetitive dynamics¹⁷.

When the free energy landscapes comprises metastable states, the dynamics produced by the Markov chain can easily be trapped in a local minimum of the free energy, especially at low temperature where the correspondingly strong rejection ratio makes it difficult for the system to climb up even small energy barriers “uphill”. This leads to quasi-ergodicity breaking, and as a consequence unreliable statistics. What generalized ensembles bring about here is a more efficient dynamics, whereby the Markov chain is able to wade across free energy barriers in a nearly transparent way, and thus can sample the phase space far more efficiently.

From a practical point of view, this is carried out by feeding the Markov chain with an appropriate distribution $w(E)$, in such a way that unlikely events corresponding to local maxima in the free energy are firmly enhanced. It turns out that this is precisely what generalized ensembles algorithms do when they lead the dynamics to explore a much wider energy window than in canonical simulations: they assign equal importance to a large set of macrostates with distinct energies, and yet some of these macrostates *may* contain the infrequent events we wish to sample. In a large class of algorithms termed “flat histogram algorithms”, for instance, sampling a wide energy range is carried out by setting $w(E)$ to the inverse of the density of states (or, since it is generally unknown, to an estimate of it obtained in an appropriate way), so that the resulting energy histogram is indeed (approximately) flat over some range of energy values. As a result, the corresponding energy levels are sampled with equal frequency, including those which have an associated low density of states and represent occasional events.

¹⁷Highly repetitive dynamics may also be efficiently overcome, for some models, by continuous time algorithms — also termed N-fold-way, or faster-than-the-clock algorithms — [50, 212]. Basically, such algorithms compute the number of Monte Carlo steps one has to skip before a move update is accepted; this clearly turns out to be very efficient at low acceptance rates. Combination with multicanonical ensembles simulations were also considered [297, 310].

2.5. Simulation in generalized ensembles

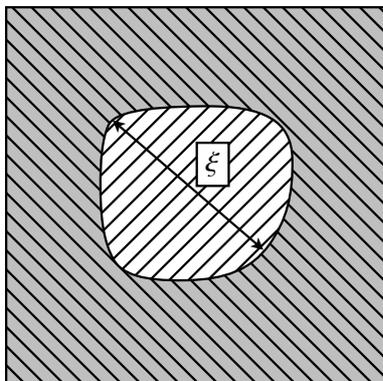


Figure 2.8: A schematic illustration of the surface tension phenomenon that results from phase coexistence at a first-order transition in, e.g., a two-dimensional Potts model with $q > 4$. The two types of hatches represent the ordered and disordered phases. The dimension of the interface is $D - 1$ for a D -dimensional model with nearest-neighbor interactions. The size of a droplet (indicated with arrows) is roughly given by the correlation length of the corresponding phase.

2.5.2 Supercritical slowing down and surface tension at first-order transitions

At a first-order phase transition, the presence of two stable phases in coexistence manifests itself through a double-peak structure in the histogram of the energy (or, alternatively, of the order parameter). When the model is simulated by means of a Boltzmann weight, configurations near or at the maximum of the free energy are strongly suppressed. This suppression is termed supercritical slowing down and was first investigated in the context of the multicanonical method [31]. This behavior is linked to the excess free energy ΔF that is associated with the interface between the two coexisting phases (Fig. 2.8), in such a way that mixed phase configurations are reduced by a factor $\sim e^{\Delta F}$ with respect to pure phase states. As illustrated in Fig. 2.9, this surface tension increases with the lattice size. For models with nearest-neighbor interactions, it is actually expected to grow like L^{D-1} (up to finite-size corrections of higher order [220, 219]) where L and D are the linear size and the dimension of the lattice, because the dimension of the interface is reduced by one with respect to the lattice dimension. This was proven on rigorous grounds in [46] for the two-dimensional Potts model using arguments based on complete wetting. As for models with long-range interactions, the situation is less clear; although some authors claimed that one should expect the same sort of behavior as in short-range models [131], yet with stronger finite-size corrections, I give marked evidence in Sec. 7.6 that the interface actually has a non-integer dimension lying between $D - 1$ and D .

Metropolis algorithms do not cope well with this effect: the reason is that, when going from one phase to the other, the Markov chain has to go through a

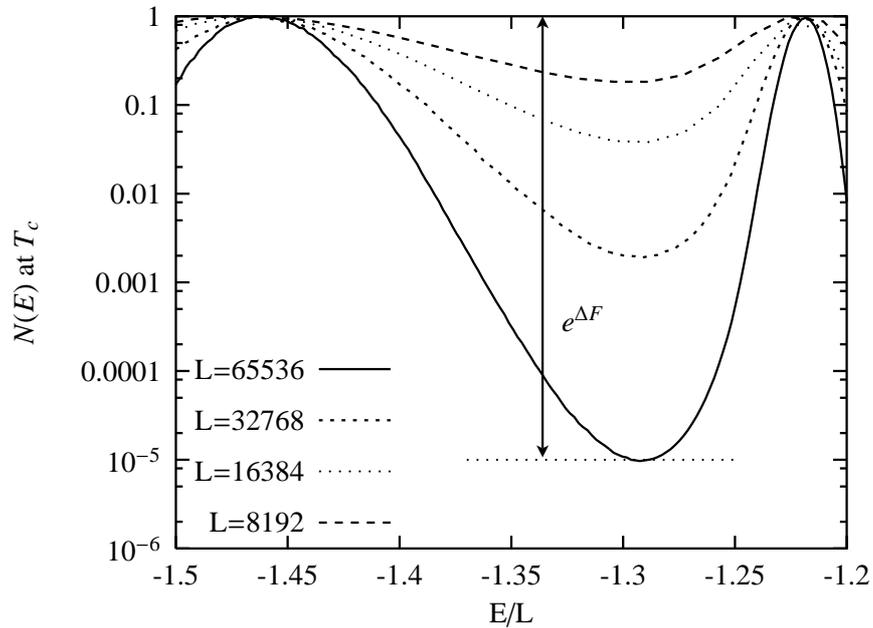


Figure 2.9: Canonical distribution of the energy at a first-order transition for various lattice sizes L (data from multicanonical simulations of a three-state long-range Potts chain with $1/r^{1.5}$ interactions, see Chap. 7).

series of mixed-phase configurations associated with large rejection ratios. Single-spin updates work all the more inefficiently in this respect that they do this at a drastically low pace by updating one spin at a time. This means that the dynamics might stay an exceedingly long time in one of the two phases before the algorithm gets a chance to switch to the other one. Incidentally, cluster algorithms work only slightly better here; it was shown in [134] that their relaxation time grows in general exponentially with the system size, a signature of supercritical slowing down.

Irrespective of the underlying move update, canonical simulations of first-order transitions are severely hampered by the suppression of mixed-phase configurations, since this suppression grows exponentially with the system size. With the exception of weak-first order transitions, i.e. with a large correlation length and a correlatively low surface tension (see [46], where it is shown that the reduced surface tension, i.e., $\Delta F/L$, is given by the inverse of the correlation length of the ordered phase), one must rely on other schemes in order to obtain reliable statistics.

2.5.3 Multicanonical ensemble

From the overlapping distribution method to umbrella sampling The multicanonical method stems indirectly from the idea of the “overlapping distribution method” [22], which later inspired the *umbrella sampling* method [318]. All these methods more or less revolve around enlarging the sampled energy windows (or

2.5. Simulation in generalized ensembles

any other parameter of physical interest), and were initially designed with measurement of free energy differences in mind. In the overlapping distribution method, two identical systems with their Hamiltonian differing in some respects are considered at the same temperature; one simulates one of the two systems while accumulating a histogram of the difference between the energy of the system being simulated, and the energy of the second system in the same microscopical configuration. The free energy difference between the two systems is then reconstructed from this histogram¹⁸. This method suffers from poor accuracy in the estimation of the free energy, however, because the physically relevant configurations of the second system are only poorly sampled by the simulation of the first system (this is reminiscent of the issue encountered with single histogramming). Umbrella sampling [318] aspires to alleviate this problem by altering the weight of the Markov chain in such a way that both distributions overlap appreciably.

Multicanonical weights Simulations in the multicanonical ensemble [30, 31] (also [25] for a concise review targeting comparison with other generalized ensembles) extends this idea by setting the weight of the Markov chain to the inverse of the density of states, i.e., $w(E) \propto 1/n(E)$, or equivalently, $w(E) \propto e^{-S(E)}$ where $S(E)$ is the microcanonical entropy. This particular choice leads to a histogram of the energy that is ideally flat over the whole energy axis. In practice, $w(E)$ is often truncated to the energy range of interest for the problem in hand, e.g., the interval spanning the two peaks of the energy histogram at a first-order transition. There may be some situation, however, where one seeks to explore energy states in the vicinity of the ground states, e.g. investigations of spin glass ground states or native protein conformations. In any case, imposing a flat energy histogram means that the dynamics in effect performs a *random walk* in the energy space. It turns out, however, that single-spin updates do not lead to an ideal random walk, i.e., one drifting along the energy axis at a pace given by the square root of the number of Monte Carlo steps. In particular, it was suggested in [24] that Berg's recursion scheme creates an additional slowing down, so that the distance covered by the random walker along the energy axis actually scales as $N^{2.5}$ instead of N^2 , where N is the number of single-spin updates (in units of lattice sweeps, this would amount to $N^{2.5/2}$ and N respectively). It is the goal of several improved schemes introduced in the last ten years to optimize the random walk dynamics in the sense mentioned above, either by combining simulation in the multicanonical ensembles with collective updates [183, 291, 353, 285], or by moving to a different ensemble [319]. A substantial part of this thesis is devoted to the development of a new method which, among other things, efficiently tackles this issue; this is presented in Part III.

¹⁸In some respect, it seems to me that this very ancient and bright idea might have inspired the recently introduced Adaptive Integration Method [111]

Multicanonical recursions Still and all, one of the sticking points of multicanonical simulations is the estimation of the density of states: naturally, it is initially unknown, for if we knew it, a major part of the problem would have been solved in the first place. A variety of schemes of varied efficiency have thus been devised in the last decade in this purpose. The first scheme was made available by Berg in [23]: it is an iterative scheme which aims to estimate the density of states in a recursive way by accumulating a histogram of the energy and modifying the current estimate of the density of states from it. Starting from an initial canonical guess $w(E) = e^{-\beta_0 E}$, the algorithm progressively refines the estimate by performing several iterations, updating $w(E)$ from the histogram of the energy obtained in the previous iteration, until the histogram has eventually become sufficiently flat. Once accurately estimated, the resulting density of states is fed into an ultimate, long (production) run during which measurements are recorded. A reweighting scheme similar to that explained in Sec. 2.4 is then used to obtain averages of thermodynamic data at any temperature below $1/\beta_0$. I will comment more on this method in Chap. 3, with special focus on (i) the improvements that must be performed when applying the method to long-range models; (ii) the dynamic characteristics of the algorithm. Another approach proposed by Wang and Landau a few years ago is discussed in a subsequent paragraph.

Entropic sampling This method was formulated by Lee [218] short after Berg's multicanonical recursions. It more or less boils down to a rewording of Berg's method in terms of the microcanonical entropy, the latter quantity being related to the density of states through $S(E) = \ln n(E)$. The microcanonical entropy is estimated in a recursive way, starting from $S(E) = 0, \forall E$. Then a histogram of the energy $H(E)$ is accumulated during an iteration run, and $S(E)$ is eventually updated according to $S(E) \leftarrow S(E) + H(E)$, except at empty bins where $S(E)$ is left unchanged. As in Berg's recursion, several iterations must be performed until the histogram becomes flat to a good approximation. It is perhaps interesting to notice that this method *might* contain the germ¹⁹ of Wang-Landau's method [335], where $S(E)$ is updated in nearly the same way (up to an "attenuation" factor that gradually dies off to zero as the simulation goes on), but at each MCS rather than at the end of each iteration run.

Applications and extensions of the multicanonical method The multicanonical method of Berg has been applied to a variety of situations since the early work of Berg [25], including first-order transitions driven by the temperature [31, 30] or by a magnetic field [28], spin glasses [27, 29] and the helix-coil transition in proteins [145]. The first application of the multicanonical method was the study of the temperature-driven transition in the two-dimensional ten-state Potts model [31], i.e., in its strong first-order regime. The authors focused here on the estimation of the surface tension of this model, and obtained numerical values that were

¹⁹That is not to say that the method *really* inspired the authors, though.

2.5. Simulation in generalized ensembles

found to be in very good agreement (within three per cent) with the exact values obtained slightly later by Borgs and Janke [46].

The multimagnetic method [28] represents an interesting extension of the method to random walks in the space of the order parameter, i.e., with the correspondence $E \leftrightarrow M$ and $\beta \leftrightarrow H$ where M denotes the magnetization and H the external field. This approach implies that the model Hamiltonian contains a term that linearly couples the order parameter with some external source, whereby averages of thermodynamic data are obtained as continuous functions of the external field. The method was shown to completely suppress the supercritical slowing down effect observed in the field-driven first-order transition in the two-dimensional Ising model below its critical temperature.

A related approach is the so-called *multioverlap* extension [29] that was successfully applied to the simulation of the three-dimensional $\pm J$ spin glass — the Edward-Anderson model [110]—, one of the simplest and most investigated approximates to realistic spin glasses. In this model, the free energy landscape displays a complex tree-like structure comprising numerous energy basins — in particular at low energy —, owing to the presence of disorder, and frustration that might arise from it. This feature is actually shared by a large class of disordered systems, including random-field models. The presence of local minima manifests itself through energy barriers in the space of the order parameter, i.e., the Parisi overlap parameter (this represents the overlap — or, broadly speaking, the resemblance — between two replicas of the same system, where a “system” refers to a set of random couplings). The purpose of the multioverlap method is thus to produce a random walk in the space of this overlap parameter, so that the corresponding barriers can be effectively overcome. The same approach was reconsidered recently in the framework of the Wang-Landau algorithm [334], where it was observed that the algorithm, as a by-product, could very efficiently find the ground states of the model.

The multicanonical ensemble can easily be extended to multi-parameter Hamiltonians, provided they are made up of a sum of linear couplings²⁰ involving a pair of intensive and extensive variables, e.g. $\beta \times E$ or $M \times H$. In this case, the multicanonical weight is set to, e.g., $w(E, M) = 1/n(E, M)$ where $n(E, M)$ is a *joint* density of states, so that a random walk is in effect performed on a hypersurface underlaid by the (intensive) variables axes. This was considered, for instance, in the context of lattice polymers [167], protein folding [82] and frustrated systems [168], and also in combination with the replica-exchange method (REM) [303]. Noteworthy enough, Guerra and Muñoz recently suggested [139] that multi-dimensional random walks might experience stronger relaxation effects than their unidimensional counterparts, as a result of the higher dimension of the hypersurface explored. The multibond method [183], which simulates spin models in their bond representation and yields a random walk in the space of the average

²⁰Although slightly different in its spirit, the Adaptive Integration Method [111] alleviates this limitation by directly computing the free energy with respect to some physically relevant parameter.

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

bond number, may also be viewed as a multi-dimension random walk as soon as next-nearest neighbor interactions are taken into account. Although this approach was suggested in [353] as a possible extension of the method in the context of the Wang-Landau algorithm, no numeric study based on this extension has been made available thus far, to the best of my knowledge.

Finally, let me mention that simulations in the multicanonical ensemble were also hybridized with Langevin algorithms and molecular dynamics [152]. Here, the potential energy entering the Hamiltonian is modified (in a way involving, here again, the microcanonical entropy) so that the resulting dynamics is in effect a random walk in the space of the potential energy.

2.5.4 Wang-Landau's random walker

As already mentioned above, one of the biggest challenges in multicanonical simulations is the estimation of the density of states that is required to perform the multicanonical production run (if any). In this respect, Berg's multicanonical recursion scheme [23] suffers from two deficiencies of distinct magnitude: it is poorly scalable, i.e., its performance in terms of convergence to the true value wanes at large lattice sizes; perhaps less importantly, it may have been deemed difficult to implement.

Meeting the problem from a slightly different angle, Wang and Landau recently proposed a so-called "acceleration method" [335, 334] for the estimation of the density of states, which — as it turns out — swiftly caught on in the community, and prompted a flurry of papers concentrating on improvements [7, 93, 297, 304, 355] or generalizations [320, 353, 267] of the method. Indeed, the scheme is simple, somewhat more straightforward to implement than Berg's one (although one may delude oneself into thinking that the scheme is childlike on the grounds that the equation does not look as sophisticated as Berg's one: the reality looks somewhat harsher, see implementation issues in Chap. 7).

Wang and Landau's method updates the multicanonical weight $w(E) = e^{-S(E)}$ in *real-time* during the course of the simulation. Each time an energy level E is visited, the microcanonical entropy is updated according to $S(E) \leftarrow S(E) + \ln f$, where $\ln f > 0$ is a modification factor that is kept constant during a given iteration. Thus energy levels which correspond to occasional events are *indirectly* enhanced because those that are visited most often will rapidly gain a higher $S(E)$, or equivalently a lower weight. Since modifying the weight of the Markov chain during a simulation is known to break detailed balance, the amount by which $S(E)$ is modified during a given iteration is decreased from one iteration to the other until it reaches a negligible value, so that detailed balance is restored in the last step of the iteration scheme. Initially [334], the suggested prescription to switch from one iteration to the other was to monitor the histogram "flatness" until it reaches an acceptably low value, in much the same way as in Berg's recursions. Histogram flatness can be estimated in many ways, i.e., one may simply want to compute the percentage of histogram bins that departs from the baseline by more than a given

2.5. Simulation in generalized ensembles

amount, or one may estimate it from the standard deviation of the *logarithm* of the histogram (this scheme is used in Chap. 3 and I suggest that this is a more natural choice, on the grounds that the quantity entering the reweighting equation is the logarithm of the histogram entries). Incidentally, estimation of histogram flatness is a feature shared by all estimators working in an iterative manner.

The convergence of the algorithm towards the flat energy distribution is somewhat tedious to prove on rigorous grounds [362], yet the intuitive picture is that, as soon as the energy distribution has become flat, each energy level occurs with the same frequency and thus –for a Markov chain of infinite length– the effect is just to translate the whole curve $S(E)$ vertically by a global amount. If the energy distribution is flat in the last step, we also have that $S(E)$ is an estimator for the density of state, with a relative uncertainty which ideally amounts to $\sqrt{\ln f}$ [362]. The situation is actually somewhat more intricate, because other parameters impinge on the global uncertainty, including the number of entries in the histogram at the end of each iteration, and correlations between successive measurements. In addition, the maximum accuracy affordable with the method was also reported to be limited by *construction*, irrespective of the number of MCS performed as a whole [358, 339], yet in the meantime it was also suggested that subtle choices of parameters may greatly help in taming several sources of error [362]. These points will be discussed in detail in Chap. 6.

Another improvement over Berg’s scheme that was proposed by Wang and Landau, consists in performing independent random walks in distinct energy ranges, possibly in a parallel way, and then to “stitch” the separate estimates together. This was suggested to markedly cut down the total amount of computer time required [335, 353], on the grounds that each random walk in a given energy interval of width ΔE takes a time $\propto \Delta E^2$ to correctly explore the available phase space, whereby dividing the energy range by N yields an improvement of $\sim N^2$. Here again, things are not that simple, and it was argued again by Zhou and Bhatt [362] that starting from a good initial guess of $S(E)$ might be as efficient as performing multiple random walks from an initially “flat” $S(E)$. Some insight into the issue is provided in Chap. 6, where I show that relying on an efficient predictor for $S(E)$ during the first iterations indeed drastically reduces the burden in terms of computer load.

The “stitching” operation may also give rise to systematic errors at right edges, owing to the fact that updates which move the system outside the permitted range of energy are systematically rejected, and thus increase $S(E)$ at right edges by an unwanted excessive amount [297]. This must be taken into account one way or the other, otherwise the seams may show: the easiest way to skirt the problem is to leave $S(E)$ unchanged whenever a move update attempts to take the system outside the energy range; another trick, which I put to use in Chap. 7, is to make the energy ranges slightly overlap *and* to ensure a “soft” rejection at right edges by continuing $S(E)$ alongside its tangent (hence the algorithm behaves locally as a canonical one).

Another pitfall which was apparently left unnoticed in recent literature is re-

Chapter 2. From Metropolis to the Wang-Landau algorithm and beyond: what can we learn from generalized ensembles methods?

lated to the time needed by the Markov chain to thermalize, which — whenever, say, the simulation kicks off from a disordered state— artificially enhances $S(E)$ in the upper energy range. Proper estimation of thermalization times is thus required, so that the corresponding events in the Markov chain may be safely skipped. The problem, as expected, does surface not only on start-up, yet *also* each time the system is re-thermalized. Indeed, it is necessary to “jar” the system now and often in the course of an iteration — which entails re-thermalizing right afterwards —, for the following reasons:

- first, this lowers the statistical error on $S(E)$, inasmuch as the procedure is equivalent to performing several, independent runs (with the same fixed $\ln f$), and averaging $S(E)$ over these [362];
- and then, this helps getting rid of potential ergodicity breaking for those energy ranges that lie on the lower side, i.e., near the ground state; here, the system may stick to the same energy basin for a long time (for much the same reason as occurs in canonical simulations at low temperature), and shaking the system is an efficient way of thrusting it into another basin [334].

Again, these points are considered in greater detail in Chap. 6 in the context of the breathing cluster method.

2.5.5 Transition matrices and the Broad Histogram Equation

That one may rely on transition probabilities between energy levels to estimate the density of states was considered as early as in 1995 by Smith and Bruce [301]. The idea was revisited sometime later by Wang, Tay and Swendsen [341] in the context of the *transition matrix* method [339, 338, 337], and separately by De Oliveira et al. [98, 96, 97], although the initial formulation of the latter was deemed somewhat flawed [95, 336] (see, also, extension to other models [253, 224] and comparison with Berg’s multicanonical method [223]).

A common feature to both methods is the notion of *potential moves*, whereby *every* potential move conveys its own piece of information to the simulation, not just accepted moves. This is expressed through the so-called Broad Histogram Equation,

$$n(E)T_{\infty}(E \rightarrow E') = n(E')T_{\infty}(E' \rightarrow E) \quad (2.14)$$

where $T_{\infty}(E \rightarrow E')$ is the (infinite temperature) transition matrix element between energy levels E and E' (also denoted as $\langle N(\sigma, E' - E) \rangle_E$ in [98]). This quantity is defined as

$$T_{\infty}(E \rightarrow E') = \frac{1}{n(E)} \sum_{E(\sigma)=E} \sum_{E(\sigma')=E'} P(\sigma \rightarrow \sigma')$$

where $S(\sigma \rightarrow \sigma')$ is the *proposed-update* probability²¹ to reach the state σ' from

²¹See Sec. 2.1.

2.5. Simulation in generalized ensembles

the state σ through a *single spin flip*²². In other words, this matrix element represents the microcanonical average at energy E of the number of potential single-spin moves from a state σ of energy E to a state σ' of energy E' . From a practical point of view, this average is estimated by accumulating a double-entry histogram $h(E, \Delta E)$ containing the number of potential moves from E to $E + \Delta E$ each time the energy level E is visited. This implies that, at each Monte Carlo step, a non-negligible amount of time will be spent computing this number by moving spins one by one and monitoring the change in energy²³. It is interesting to mention at this point that an extension of the Broad Histogram equation to other variables, e.g., the number of bonds in the graph representation of the Potts model, was recently proposed [354]; it was argued by the authors, in particular, that the equation may actually be generalized to any number of variables, i.e., in the same spirit as the multicanonical method applied to multiparameter Hamiltonians (see Sec. 2.5.3).

Although the definition of $T_\infty(E \rightarrow E')$ is very general, the Broad Histogram Equation applies only if move updates satisfy a so-called microreversibility hypothesis, e.g., are local (see previous footnote). Yet provided this is the case, the density of states can be readily estimated by computing, e.g., the quantity $\ln n(E')/n(E) = S(E') - S(E)$ from $T_\infty(E \rightarrow E')$ and $T_\infty(E' \rightarrow E)$, and integrating over E to yield $S(E)$ over the whole energy axis. Other, more reliable schemes based on least-square optimization were proposed in [339]. For the purpose of multicanonical simulations, i.e., yielding a flat energy distribution, one is only interested in the ratio $n(E')/n(E)$ which enters the Metropolis acceptance rate of Eq. (2.4), so that integrating is not necessary at this stage. Additionally, a working estimate of $T_\infty(E \rightarrow E')$ is already available in the first steps of a simulation; it is simply estimated from the number of potential moves from the current state at energy E , and seeing that for sufficiently close energy levels E and $E + \Delta E$, one may also approximate $T_\infty(E + \Delta E \rightarrow E)$ by $T_\infty(E \rightarrow E - \Delta E)$, the acceptance ratio can be well approximated even in the very first stage of the simulation.

Finally, a useful quantity related to the transition matrix method is the so-called *TTT* identity, which directly stems from the *histogram detailed balance* derived in [339], and yields useful information regarding the degree of detailed balance violation in a given simulation. For single-spin updates *only*, this degree reads

$$\nu = \left| 1 - \frac{T_\infty(E \rightarrow E')T_\infty(E' \rightarrow E'')T_\infty(E'' \rightarrow E)}{T_\infty(E \rightarrow E'')T_\infty(E'' \rightarrow E')T_\infty(E' \rightarrow E)} \right|$$

and $\nu = 0$ for a perfect simulation with zero detailed-balance violation. A similar equation also holds for cluster updates, yet in a clearly more complex form since

²²Wang's initial formulation relied on this acceptance rate to derive a detailed balance equation in the energy space, yet it was established that, provided all moves satisfy a micro-reversibility hypothesis, the corresponding (finite temperature) transition matrix $T(E \rightarrow E')$ can be factorized into a term involving the usual acceptance rate — e.g., Metropolis-like — and an infinite temperature transition matrix $T_\infty(E \rightarrow E')$. The formulation does not apply as is to cluster updates, however.

²³For nearest-neighbor models, this can be readily calculated from the local environment of each spin using “classes” as in the heat-bath algorithm. The task is much more demanding, however, for long-range models.

cluster updates do not satisfy the micro-reversibility condition.

Transition matrices generally yield more accurate estimates of the density of state, although at a higher cost in terms of computer resources. As we will witness in several parts of this thesis, the method can be efficiently and straightforwardly combined with other algorithms to improve their accuracy: this will be considered first in Sec. 4.3 in the context of the spinodal method to estimate the order of a phase transition, and then in Chap. 6 in combination with the breathing cluster method.

2.5.6 Other generalized ensembles

Simulated tempering Although simulated tempering methods do not, strictly speaking, operate in a generalized ensemble, it is worth mentioning them here, because they represent a class of methods that has proven to efficiently overcome quasi-ergodicity breaking — in particular in the investigation of spin glasses [162, 194, 221] and protein folding [146] —, in spite of their being somewhat limited in some other respects. Simulated tempering actually exists in many flavors [237], including a parallel version termed *parallel tempering* [241]. The idea is to engender a random walk in the temperature, yet over a discrete set of variables — and this is where the method differs in an essential way from generalized ensembles methods²⁴. This is carried out by performing several simulations (possibly in parallel) of the same system, yet at different temperatures, and exchanging the temperatures (or equivalently, the configurations) between the two systems now and often. Obviously, this exchange operation is subject to some acceptance probability, in order for detailed balance to be satisfied. The method can efficiently overcome quasi-ergodicity breaking, because the high-temperature simulation assists the system in crossing free energy barriers which it would not have been able to cross in the low-temperature phase. Thermodynamic averages are eventually computed in much the same way as with the multi-histogram method [114].

1/k ensemble As opposed to simulated tempering, the 1/k ensemble [156] clearly belongs to the family of generalized ensembles. It differs from the multicanonical ensemble, however, in that it does not engender a flat energy histogram, yet rather a random walk in the space of the microcanonical entropy. Since (i) the microcanonical entropy $S(E)$ is a monotonous increasing function of the energy, and (ii) the microcanonical temperature $\beta(E)$ decreases with increasing energy (except at a first-order transitions in the region of phase coexistence), this ensemble favors low-energy states and thus targets mainly optimization problems and searches for ground states. It can be established that the appropriate weight for this ensemble can be easily related to the multicanonical weight $1/n(E)$, and thus it can be indirectly estimated by relying upon a multicanonical iteration scheme.

²⁴In fact, the energy axis is also a discrete set in any numerical implementations, yet *every* energy level — at least inside a given range — needs to be sampled, whereas temperatures can be spaced with more freedom in simulated tempering.

2.6. Outlook

Generalized ensemble with Tsallis weights A somewhat related approach was formulated by Hansmann and Okamoto [148] with the protein-folding problem in view. This ensemble is somehow related to Tsallis statistics, in that the weight used by the author mimics the generalized distribution proposed by Tsallis [321]. The Markov weight is made of two parts: a low-energy part given by a Boltzmann weight, and a high-energy part following a power-law behavior [151]. The latter choice is shown to ensure that free energy barriers can be efficiently overcome, while simplifying the estimation of the optimal distribution [149]. This approach was also reconsidered in the context of molecular dynamics [150], along the same line of arguments as those used in [152] for the multicanonical-MD hybrid algorithm.

Optimized ensemble: increasing diffusion currents A complementary approach based on the maximization of diffusion rates was very freshly introduced in [319, 351]: in this so-called *optimal ensemble*, the engendered histogram is still broad, but no longer flat. Rather, it is peaked around the critical region²⁵, where the *diffusivity* coefficient of the random walker is claimed by the authors to experience its lower value (akin to critical slowing down; this was already alluded to by Guerra and Muñoz in the context of equilibrating properties of multicanonical simulations [139]). The so-called *optimal weight* $w(E)$ of the new ensemble is computed through a feedback procedure, whereby $w(E)$ is increased in the region where the diffusivity is smaller. This method seems very promising, with applications to a simple frustrated model proving very convincing.

2.6 Outlook

In this chapter, two classes of Monte Carlo algorithms have been reviewed: cluster algorithms rapidly reduce temporal correlations by flipping blocks of spins in a collective way, resulting in an increased accuracy in the estimation of thermodynamic averages; algorithms operating in generalized ensembles efficiently overcome repetitive dynamics originating from rugged free energy landscapes in disordered models or surface tensions at discontinuous transitions. The next part of this thesis investigates a long-range Potts chain by relying on the last approach. Yet as was mentioned above, there have been several attempts in the recent past at combining both approaches in order to deliver better dynamic performance. These methods, among which the multibond method [183], are reviewed in Part III (Chap. 6), where I will describe the main impediments to their usability in the context of long-range models. These will be the main motivations behind the development of a new method (Chap. 7) that will indeed efficiently combine both approaches using an innovative strategy.

²⁵This is the energy interval that would be sampled by a canonical simulation at the critical temperature.

Part II

A multicanonical study of the long-range Potts model

Chapter 3

A multicanonical algorithm for long-range spin models

In this chapter, I will present the multicanonical method that will be used during the first part of this work, Chap. 4 and 5. The purpose of this chapter is threefold:

- it may be regarded as an introduction to the method in its pristine formulation (Berg, [31]), in particular as far as implementation details are concerned;
- it introduces several improvements over Berg's initial recursion scheme [30, 31], that make the algorithm suitable for long-range models; in this respect, it will be shown that the presence of unequally spaced energy levels in these models requires a modification of the recursion equation (with regard to the last point, long-range *discrete* spin models are somewhat peculiar indeed); the choice of an efficient predictor and a reliable convergence criterion for the recursion scheme is also discussed;
- it contains a study of the dynamic characteristics of the algorithm in the context of long-range spin models, in particular in terms of dynamic exponents.

The model being investigated in this thesis part is a long-range generalization of the q -state Potts model, i.e., with interactions $r^{-\sigma-D}$ decaying as a power law of the interparticle distance. More detailed material regarding the *model* is exposed in the next two chapters, which are dedicated to its numerical investigation. On the contrary, this chapter is exclusively devoted to the *method*, so that it may be skipped by a reader only interested in the results regarding the critical behavior of the model.

The main features of the multicanonical method are presented in Sec. 3.1: it is explained why this approach is *a priori* suitable for the investigation of the long-range Potts chain. Multicanonical weighting is presented in Sec. 3.2; the particular role played by the microcanonical temperature in the dynamics is singled out. The recursion scheme used to determine the multicanonical weights is described in Sec. 3.3: taking guidance from the initial formulation of Berg and Neuhaus, I

propose several improvements regarding the convergence of the algorithm. The reweighting procedure is discussed in Sec. 3.4, with emphasis on the estimation of free energies. In Sec. 3.5, the choice of an efficient predictor is discussed, that increases the robustness of the algorithm in terms of convergence to the best estimate of the density of states. Finally, the performance of the algorithm is measured in Sec. 3.6 for the long-range Potts chain with power-law decaying interactions, by relying on the estimation of various characteristic times of the simulation. Dynamic exponents are compared with several other existing schemes.

3.1 Canonical vs multicanonical ensembles and long-range models

As was seen in Chap. 2, the Metropolis algorithm (hereafter denoted as belonging to the class of *canonical* algorithms, i.e., operating in the canonical ensemble) has long been considered the paradigm for Monte Carlo simulations in statistical physics. However, the method faces some severe drawbacks in situations where the sequence of states created by the Markovian chain leads to very repetitive dynamics, i.e., dramatically low acceptance rates. On the dynamic side, this implies exponentially diverging autocorrelation times (or tunneling times, to be defined below), and fewer independent samples for a given simulation time. As a result, reaching a given accuracy requires that the system be simulated over exceedingly long runs in order to obtain good statistics and reliable estimates of thermodynamic averages.

Repetitive dynamics are encountered in canonical simulations of first-order phase transitions (the so-called supercritical slowing down [31]). Here, tunneling time between coexisting phases grows exponentially with the system size, due to the increasingly high free energy barrier to be overcome. This is one of the reasons that prompted me to focus on the class of multicanonical algorithms to investigate the phase diagram of the long-range Potts chain. As will be seen in the next two chapters, a multicanonical approach is indeed an appropriate choice for the determination of the location of the boundary separating the first- and the second-order regimes of this model, and for the estimation of its critical couplings in the first-order regime. The present purpose is therefore to adapt the scheme initially formulated for nearest-neighbor models so as to make it suitable for a large class of long-range models.

In order to circumvent repetitive dynamics, multicanonical methods perform a random walk in the energy space. It is important to note that this scheme does not make any assumption about the particular move update utilized (this may be a single-spin flip or a collective update, as we shall consider in Chap. 6). This random walk engenders in turn a *flat* energy distribution. The benefits of the algorithm are in fact twofold:

- This results in the algorithm quickly sampling a much wider phase space

3.2. Multicanonical weights

than in the canonical case, by allowing the system to cross any free energy barrier.

- This allows the density of states to be computed over the whole energy axis, thus extending the reliability of reweighting procedures over a much wider range of temperature than in the case of standard histogram methods. An example of such calculation is illustrated in Fig. 3.4.

As a corollary, and as opposed to multihistogramming [114], a single run is needed to cover the energy range of interest¹. Once a reliable estimate of the density of states has been obtained, it is then straightforward to compute thermodynamic functions otherwise hardly within reach of canonical simulations, e.g., entropies and free energies. The last functions will be extensively used in several parts of this thesis, e.g., to accurately determine the order of the phase transitions exhibited by the long-range Potts chain.

The first point mentioned above, that is, much wider and efficient phase sampling, is illustrated in Fig. 3.1: here, the *flat* histogram engendered by the random walk is depicted along with the *canonical* histogram, i.e., the histogram that would — ideally — be obtained from a canonical simulation at the transition temperature. This canonical histogram was actually obtained by a reweighting procedure (see Sec. 3.4 below): a canonical simulation would have hardly produced the *correct* energy distribution, owing to the presence of the free-energy barrier (this roughly corresponds to $E/L = -1.5$ in the figure). As is depicted in the figure, the dynamics of an equivalent “canonical” walker would be fairly inefficient, with the walker “bumping” into the free-energy barrier a huge number of times before having a chance to go across the barrier and reach the ordered phase.

3.2 Multicanonical weights

The rationale behind the multicanonical algorithm is the generation of a Markovian chain of states $\{\sigma_i\}$, whose weights $w_{mu}(E(\sigma_i))$ are adjusted so that one eventually gets a flat energy histogram. Denoting as $P_{mu}(E)$ the energy distribution and as $n(E)$ the density of states, we want

$$P_{mu}(E) \propto n(E)w_{mu}(E) = \text{const.}$$

Since $n(E)$ usually increases drastically with energy, low-energy states are thus sampled much more often than high-energy ones. One of the *core* issues in the multicanonical method is to determine these weights, since these are not known *prior* to starting the simulation.

¹This is a somewhat idealized picture though, because many iterative runs are needed to obtain the correct weights. One usually speaks of a single *production* run, i.e., the run from which thermodynamic averages are produced.

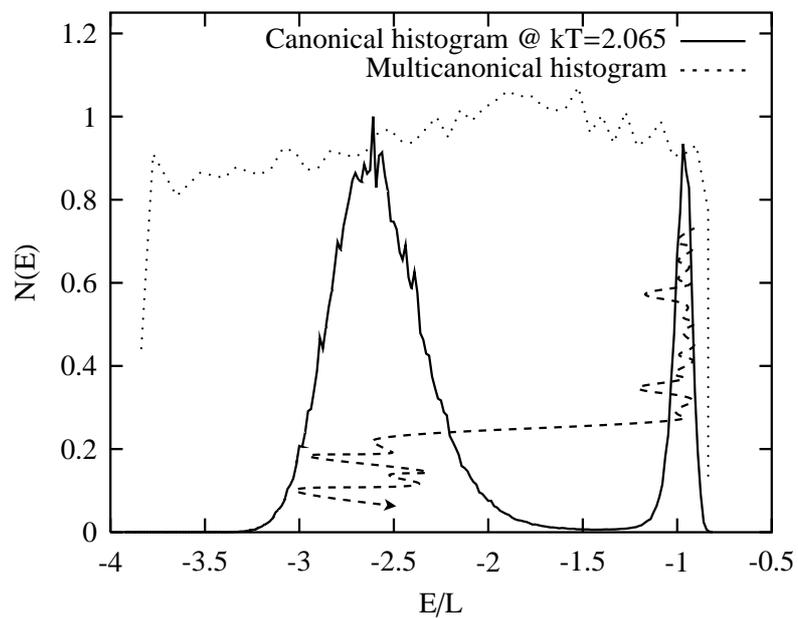


Figure 3.1: Flat energy distribution engendered by the multicanonical algorithm (multicanonical histogram) vs the reweighted histogram at the “equal heights” transition temperature (denoted as “canonical” histogram in the figure). The strong suppressing of mixed-phase configurations that would occur in a canonical simulation, is clearly visible: the dash line illustrates (in a very sketchy manner) what would be the dynamics of a canonical simulation in this case. The model parameters are: $q = 5$, $\sigma = 0.3$, $L = 400$. E/L designates the energy per spin.

3.2. Multicanonical weights

The following outline takes guidance from Berg's recursion scheme [23]. The weights $w_{mu}(E)$ are computed through an iterative procedure, starting from an initial canonical simulation at inverse temperature β_0 . β_0 indirectly sets the energy below which the energy histogram is to be flat, i.e. $E_{max} = \langle E \rangle_{\beta_0}$ (in Fig. 3.1, this roughly corresponds to $E/L \sim -0.8$ for the energy per spin). Thus, $kT_0 = 1/\beta_0$ must be chosen high enough to ensure that the final energy histogram spans a suitably large energy range upward, e.g., reaches the energy of the disordered phase in the case of a first-order transition, and extends even further away if one wants to observe with satisfying accuracy the free energy plateaus signaling the limit of metastability (see Fig. 3.3, and also the application of metastability plateaus to the detection of the order of phase transitions in Chap. 4).

For convenience, one may define an effective Hamiltonian $H_{mu}(E)$, so that

$$w_{mu}(E, \beta_0) = e^{-\beta_0 H_{mu}(E)} = e^{-S(E)}$$

where $S(E)$ plays the role of a microcanonical entropy (and is an estimate for it if the recursion scheme converges to the true value). The second formulation will be relied on in Chap. 6, in the context of the breathing cluster method. Here, I follow the initial formulation of Berg and Neuhaus. Multicanonical simulations can thus be envisioned as a canonical simulation at inverse temperature β_0 with the usual Boltzmann weight, provided the original Hamiltonian is replaced by an effective Hamiltonian to be determined iteratively.

As a side note, a cluster implementation in the framework of the multicanonical algorithm is clearly less straightforward, since this effective Hamiltonian has fundamentally a global nature, whereas canonical simulations explicitly preserve the locality of the original Hamiltonian: this point will be reexamined in detail in Chap. 6.

Denoting $H_{mu}^\infty(E)$ as the true estimate of the effective Hamiltonian, one may thus write

$$n(E) \propto e^{\beta_0 H_{mu}^\infty(E)}.$$

The microcanonical inverse temperature $\beta(E)$ may be easily related to $H_{mu}^\infty(E)$, as one has (assuming $k = 1$)

$$\beta(E) = \frac{d \ln n(E)}{dE} = \beta_0 \frac{dH_{mu}^\infty(E)}{dE}$$

Since the dynamics of the Markovian chain is governed by the transition rate

$$W(a \rightarrow b) = \min(1, \exp\{\beta_0 [H_{mu}(E_a) - H_{mu}(E_b)]\}),$$

one may write, for two states infinitely close in energy, i.e., whenever $E_b = E_a + \delta E$,

$$W(a \rightarrow b) = \min(1, \exp[-\beta(E_a)\delta E]).$$

Hence it is the *microcanonical temperature* which is the relevant quantity where the dynamics (e.g., the acceptance rate) of the multicanonical algorithm is concerned. This observation will be main thrust of the breathing cluster method developed in Chap. 7.

3.3 Recursion scheme

The effective Hamiltonian is initially set to $H_{mu}^0(E) = E$, or equivalently $\beta^0(E) = \beta_0$, as this indeed corresponds to a canonical simulation at temperature $1/\beta_0$. At step i , a simulation is performed using a Boltzmann weight with effective Hamiltonian $H_{mu}^i(E)$; then an energy histogram $N^i(E)$ is eventually computed using independent samples. Incidentally, taking truly independent samples proves useful during the late stages of the iteration scheme only, where the aim is then to refine a nearly flat histogram. During early iteration steps, histograms may be computed using nonindependent samples without significantly affecting the convergence.

I now denote E_{min}^i as the lowest energy level that was reached throughout the previous runs, including step i : this is the energy level below which $H_{mu}^{i+1}(E)$ will have to be predicted, since no histogram data are available inside this energy range. Issues regarding adequate predictor choice will be considered later on in this chapter. The rules for updating H_{mu}^{i+1} at step $i + 1$ from H_{mu}^i at step i are based on the following equations. For $E \geq E_{max}$, $H_{mu}^{i+1}(E) = E$, i.e., the dynamics is canonical-like at inverse temperature β_0 for all iteration steps. For $E_{min}^i \leq E < E_{max}$,

$$\beta^{i+1}(E) = \beta^i(E) + \frac{\hat{g}_0^i}{\delta E} \ln \frac{N^i(E + \delta E)}{N^i(E)}, \quad (3.1)$$

where

$$\hat{g}_0^i = \frac{g_0^i}{\sum_{k=1}^i g_0^k}$$

and g_0^k is a *raw* inverse damping factor proportional to the reliability of the k th histogram. It has been established in [23], following an error calculation argument, that

$$g_0 = \frac{N(E)N(E + \delta E)}{N(E) + N(E + \delta E)}$$

provides an estimator proportional to the inverse of the variance of $\beta^{i+1}(E)$.

Once $\beta^{i+1}(E)$ is known, $H_{mu}^{i+1}(E)$ is derived by a mere integration scheme starting from the initial condition $H_{mu}(E_{max}) = E_{max}$. Finally, for $E < E_{min}^i$, $H_{mu}^{i+1}(E)$ will have to be computed using a suitably chosen predictor, until at last E_{min}^i becomes equal to the ground state energy. A cubic spline is then fitted to $H_{mu}(E)$ at every bin center, and this curve is used to compute acceptance probabilities during the next run. It can be seen that Eq. (3.1) leads to a steady state whenever $N(E)$ is constant over the energy range of interest.

3.3.1 Accidental vs non-accidental histogram entries

Writing a recursion equation involving $\beta(E)$ instead of $H_{mu}(E)$, together with the inclusion of a damping factor, allows one to handle the situation where some bins have null entries, a case which otherwise leads to a fairly spiky graph for $H_{mu}(E)$ and inconsistent dynamics. *Accidental* null entries at energy values E or $E + \delta E$ will

3.3. Recursion scheme

simply leave $\beta(E)$ unchanged, and the corresponding parts of $H_{mu}(E)$ thus move as a block. Since acceptance rates hinge on the microcanonical temperature, this in effect drastically reduces bias on the dynamics.

Considering a small set of histogram bins that are copiously filled for the first time during a given iteration run (e.g., high-energy bins during the early iteration runs whenever one starts with a canonical simulation), one realizes that the related *cumulative* inverse damping factor first soars and produces a great amount of change in $\beta(E)$ in the couple of runs that follow, and then decays progressively to zero as these bins continue to be filled. By taking into account all the data that have been sampled up to step i , this modified recursion both clearly stabilizes the algorithm and reduces relative errors due to poor histogram sampling.

3.3.2 Optimal histogram bin width

Choosing the most appropriate value of the histogram bin width results from a trade-off between resolution and computation time. A higher resolution on the one hand guarantees good histogram flatness, and is especially crucial at low energy levels, where the density of states displays a rugged graph. On the other hand, I impose a fixed number of independent samples per histogram bin, so as to give the histogram variance an acceptably low value. This means that a low δE imposes that more simulation steps be performed per iteration. The present approach is thus to choose a fairly high δE , e.g., one yielding around 20 bins, during the early stages of the iteration process in order to obtain a rough picture of the density of states, and then to progressively reduce δE once the ground state has been reached. As will become clear in Sec. 4.3, the ultimate value of δE deeply affects the attainable precision on the computation of spinodal points, since the latter is based on a precise location of free energy plateaus, and this indeed entails having enough bins belonging to a given plateau. As a rule of thumb, the best compromise is then to obtain between 100 and 300 histogram bins in the final stage, with the number of bins increasing as the σ value corresponding to the second-order regime is approached.

3.3.3 Unequal spacing of energy levels

The unequal spacing of energy levels in long-range spin models deserves specific attention. As witnessed in Fig. 3.2, large energy gaps separate isolated energy levels or tiny groups thereof in the vicinity of the ground state, whereas the distribution gradually turns into a near continuum above $E \sim -1025$. Setting a low δE value leads in turn to *nonaccidental* null entries in those bins located inside energy gaps, whereby $\beta(E)$ never gets updated at isolated energy levels and g_0 is always zero. Since the graph of the density of states looks indeed fairly wrinkled near the ground state, and the dynamics there is noticeably sensitive to even the smallest departure of $H_{mu}(E)$ from the ideal line, one would then observe a sharp steady peak in the lowest part of the energy histogram, which the present recursion

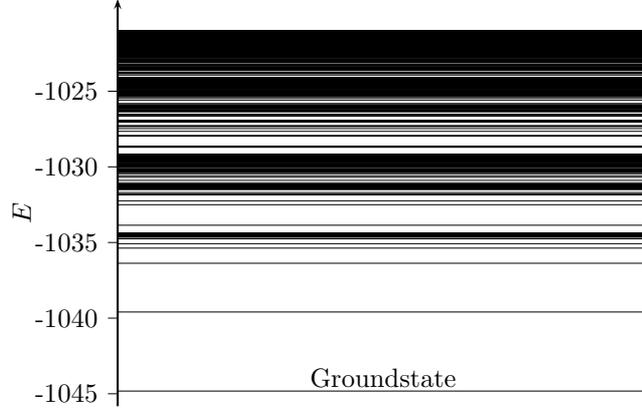


Figure 3.2: Lowest energy levels for $q = 5$, $\sigma = 0.5$, $N = 400$, computed by sorting energy samples from a long simulation run. Each level is drawn as a horizontal line.

would not be able to suppress. One could trivially think of working this out by implementing variable-width bins that would span energy gaps. This is, however, impracticable since the distribution of energy levels is not known prior to starting the iteration process (for this is precisely what one intends to compute with the density of states). To circumvent this limitation, I have modified the previous recursion so that null entries are always skipped, however accidental or nonaccidental they may be. Denoting by E_a and E_b , with $E_a < E_b$, the centers of histogram bins located on each side of a set of contiguous empty bins, one has

$$\beta^{i+1}(E_a) = \beta^i(E_a) + \frac{\hat{g}_0^i(E_a)}{E_b - E_a} \ln \frac{N^i(E_b)}{N^i(E_a)}, \quad (3.2)$$

where $\beta(E_a) = \beta_0 \{H_{mu}(E_b) - H_{mu}(E_a)\}$ and I now impose

$$g_0(E_a) = \frac{N(E_a)N(E_b)}{N(E_a) + N(E_b)};$$

hence g_0 can never be zero.

In order to avoid losing details of the shape of $H_{mu}(E)$ for $E_a < E < E_b$ that were possibly collected during previous runs, $H_{mu}(E)$ is updated through a linear difference scheme,

$$\delta H_{mu}(E) = \frac{\delta H_{mu}(E_b) - \delta H_{mu}(E_a)}{E_b - E_a} (E - E_a) + \delta H_{mu}(E_a),$$

where $\delta H_{mu}(E) = H_{mu}^{i+1}(E) - H_{mu}^i(E)$. While this has obviously no effect where nonaccidental null entries are concerned, this favors quicker convergence during the early runs where the inadequate shape of $H_{mu}(E)$ is more likely to produce empty bins.

3.4. Reweighting procedure

3.3.4 Convergence criterion

The iteration process stops whenever the energy histogram has become suitably flat over the energy range of interest: between the ground state energy and E_{max} for the present purpose. This property is evaluated by computing the standard deviation of histogram entries, as well as the same quantity for the logarithm of histogram entries restricted to nonempty bins. The latter seems to be a better indicator since it is sensitive to both poorly populated bins and histogram peaks, whereas the former increases only with rather spiky histograms. Furthermore, it is this quantity which enters the reweighting formula for the estimation of free energies.

The degree of convergence of the algorithm is eventually estimated by computing the mean square distance between $H_{mu}^i(E)$ and $H_{mu}^{i+1}(E)$ after the ground state has been reached. Then, a threshold value is computed for each indicator by trial and error, based on a couple of short runs for various lattice sizes and bin widths.

3.4 Reweighting procedure

Once $H_{mu}(E)$ has been satisfactorily computed, a long production run is performed using this effective Hamiltonian in place of the original one, and then estimates of thermodynamic quantities of interest at inverse temperature β are computed using a reweighting scheme, i.e., formally,

$$\langle A \rangle_\beta = \frac{\sum_E \langle A \rangle_E n(E) e^{-\beta E}}{Z(\beta)},$$

where $\langle A \rangle_E$ is the microcanonical average of A at energy E , and the partition function is given by $Z = \sum_E n(E) e^{-\beta E}$.

The best estimate for the density of states $n(E)$ is provided by

$$n(E) \propto N(E) e^{\beta_0 H_{mu}(E)},$$

where $N(E)$ stands for the number of bin entries at energy E computed from the production run. In order to avoid numerical overflows, as well as to suppress bias resulting from possibly strong variance on microcanonical averages, it is more appropriate to compute $\langle A \rangle_\beta$ from a sum running over samples instead of energy bins, i.e., $\langle A \rangle_\beta = \sum_i A_i w(E_i) / \sum_i w(E_i)$, where $w(E_i) = e^{\beta_0 H_{mu}(E_i) - \beta E_i - K}$. K is then determined so as to avoid both numerator and denominator overflows. However, the downside of this approach is a marked memory overhead for large sets of data.

Provided the histogram sampled during the production run is flat to a good approximation, the maximum in $e^{\beta_0 H_{mu}(E) - \beta E}$ is reached whenever $dH_{mu}(E)/dE \sim \beta/\beta_0$, which yields the energy value at which K is to be computed. In addition, since the reweighting scheme involves an exponential contribution of $H_{mu}(E)$, the resulting curve $e^{\beta_0 H_{mu}(E) - \beta E}$ is strongly peaked around the maximum. Therefore, it is clear that only histogram points in the vicinity of this maximum contribute to $\langle A \rangle_\beta$. In effect, the existence of two distinct maxima, or equivalently of two

energy values for which $\beta(E)$ has the same value, coincides with the occurrence of a first-order phase transition²

3.4.1 Free energies

Following the same reweighting procedure, one computes “partial” (or variational) free energy functions, i.e., $F(\beta, m)$ where m is the order parameter, and reweighted histograms of the energy, i.e., $N_{rw}(\beta, E)$. The partial partition function is straightforwardly derived from a sum over samples having the prescribed order parameter,

$$Z(\beta, m) = \sum_i e^{\beta_0 H_{mu}(E_i) - \beta E_i} \delta_{m, m_i}, \quad (3.3)$$

which then yields $F(\beta, m) = -\ln Z(\beta, m)/\beta$. Similarly, a reweighted histogram of the energy is obtained from

$$N_{rw}(\beta, E) = N(E) e^{\beta_0 H_{mu}(E) - \beta E}, \quad (3.4)$$

from where on a (variational) free energy function with respect to the energy may be derived,

$$F_e(\beta, E) = -\ln N_{rw}(\beta, E). \quad (3.5)$$

This is illustrated in Fig. 3.3 for the five-state long-range Potts chain with $1/r^{1.3}$ interactions and $L = 400$ spins.

3.5 Predictor choice

I now address some issues related to the choice of an efficient predictor for $E < E_{min}$. At small lattice sizes, the algorithm is initially fed with an effective Hamiltonian $H_{mu}(E) = E$, and the goal is then to find an appropriate trade-off between speeding up the convergence of E_{min}^i toward the ground state and avoiding algorithm instability. While the former demands that $H_{mu}^i(E)$ have a sufficiently high slope below E_{min}^i , the latter still requires that the algorithm remain ergodic to a suitable extent.

The present implementation relies on a first-order predictor, $H_{mu}(E) = a + bE$, and continuity is imposed on $H_{mu}(E)$ at E_{min} . The simplest approach is then to choose a predictor slope so that continuity on $H'_{mu}(E)$ is enforced at $E = E_{min}$, i.e., $b = \beta(E_{min})/\beta_0$. While E_{min} reaches the ground state rather quickly using this predictor, the dynamics often gets locked in very low energy levels due to the particularly steep slope of $H_{mu}(E)$ in the vicinity of the ground state. The time needed by the iteration scheme to get over from this deadlock and obtain a flat histogram thus becomes prohibitive. On the other hand, choosing $b = 1$ leads to the smoothest yet slowest convergence, and avoids deadlock issues. An efficient

²This observation will be relied on in Sec. 4.3.4, where it will be demonstrated that the transition matrix method [337] can produce useful information with respect to the order of phase transitions.

3.5. Predictor choice

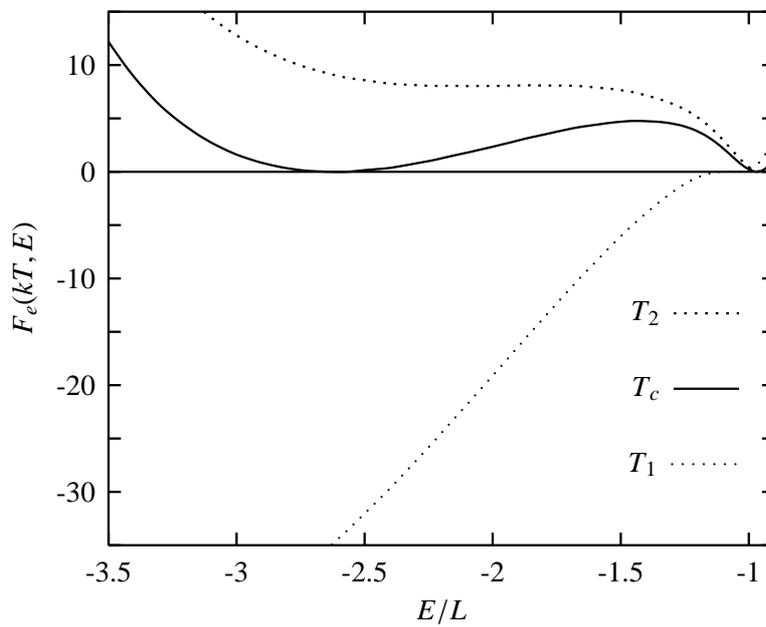


Figure 3.3: Illustration of the reweighting procedure for (variational) free energies. The graph sketches the variational free energy with respect to the energy, i.e. the logarithm of the reweighted histogram of the energy (up to a change of sign), for three distinct temperatures: T_1 and T_2 correspond to limits of metastability; T_c is one possible definition of the (finite-size) transition temperature.

compromise is thus to ensure a “weak” continuity at E_{min} , i.e., by computing the slope of the predictor using a least-square scheme based, e.g., on the first ten per cent of points above E_{min} .

Reusing predictors - finite size scaling At large lattice sizes, where reaching the ground state energy can become time consuming, one has to resort to a “scaling trick” whereby $H_{mu}(E)$ is initially guessed from the density of states obtained at a smaller lattice size. This approach was initially mentioned by Berg and Neuhaus [31], and claimed to work perfectly within the framework of a study of the two-dimensional ten-state Potts model with *nearest-neighbor* interactions, where the energy is additive to a perfect extent. The presence of long-range interactions, however, slightly worsens the case, especially at low σ . The scaled density of states is computed as follows. Let me consider, for the sake of simplicity, two systems Σ and $\bar{\Sigma}$ with respective lattice sizes $L = N$ and $\bar{L} = 2N$. The latter is split up into two subsystems Σ_1 and Σ_2 of equal size L . Since $H_{mu}(E) = kT_0 \ln n(E)$, where $n(E)$ stands for the density of states, one has to compute $\bar{n}(E)$ for $\bar{\Sigma}$ as a function of $n(E)$ for Σ . Neglecting the interaction between subsystems Σ_1 and Σ_2 , and denoting by E_1 the energy of Σ_1 , the density of states for $\bar{\Sigma}$ just reads $\bar{n}(E) \simeq \sum_{E_1} n(E_1)n(E - E_1)$, which yields

$$\begin{aligned} \beta_0 \bar{H}_{mu}(E) &\simeq \ln \sum_{E_1} e^{\beta_0 [H_{mu}(E_1) + H_{mu}(E - E_1)]} \\ &\sim \ln \frac{1}{\delta E} \int dE_1 e^{\beta_0 [H_{mu}(E_1) + H_{mu}(E - E_1)]}, \end{aligned}$$

where δE is the energy histogram bin width. Providing that $n(E)$ is a monotonic and rapidly increasing function of E , one may use a saddle-point approximation to evaluate the former sum. The maximum of $H_{mu}(E_1) + H_{mu}(E - E_1)$ is reached whenever $E_1 = E/2$. Whence

$$\bar{H}_{mu}(E) \simeq 2H_{mu}\left(\frac{E}{2}\right) + kT_0 \ln \frac{\sqrt{\pi/|H''_{mu}(E/2)|}}{\delta E} \quad (3.6)$$

This expression may be readily extended to lattice sizes that are any multiple of the original size. Figure 3.4 sketches results obtained for the following model parameters³: $q = 5$ and $\sigma = 0.3, 0.5$, and 0.9 . A series of iteration runs is first conducted with $L = 200$ spins in order to obtain an estimate of $H_{mu}(E)$ for this lattice size, then this estimate was scaled using Eq. (3.6), and then used as the initial guess $\bar{H}_{mu}(E)$ for the next series of iteration runs at $L = 400$. Equation (3.6) yields a very acceptable guess for $\sigma = 0.9$, and the two curves are hardly distinguishable from each other. As illustrated in Fig. 3.5, the energy histogram becomes nearly flat within five iterations. For $\sigma = 0.3$ and 0.5 , the agreement remains quite satisfying, yet the initial guess falls slightly below the true estimate at low energy levels, and the lowest-energy bins are exceedingly enhanced during the first iteration runs.

³The Hamiltonian of the long-range Potts chain is defined in Eq. 4.1.

3.5. Predictor choice

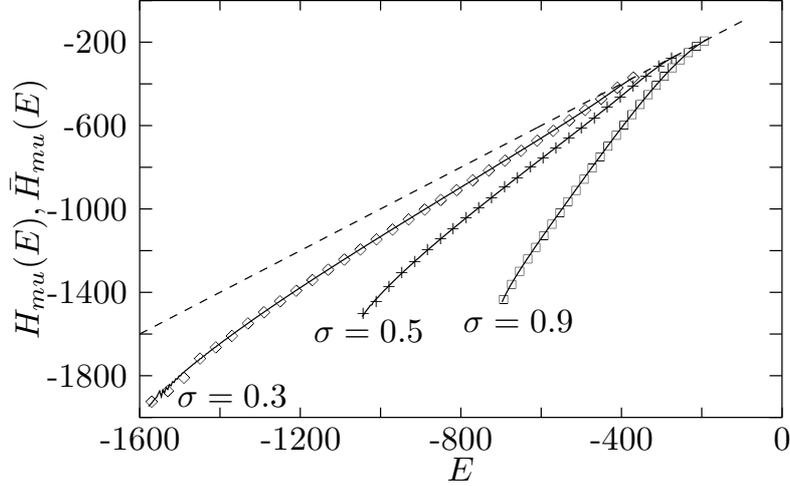


Figure 3.4: Dots indicate the initial guesses $\bar{H}_{\mu}(E)$ that were fed into the iteration scheme at $L = 400$, $q = 5$, and $\sigma = 0.3(\diamond)$, $0.5(+)$, $0.9(\square)$. Each initial guess is computed using Eq. (3.6), i.e., by scaling a true estimate obtained at $L = 200$. Solid lines show true estimates $H_{\mu}(E)$ as obtained after the whole iteration scheme at $L = 400$ converged. The straight dashed line sketches the original Hamiltonian, i.e., $H_{\mu}(E) = E$.

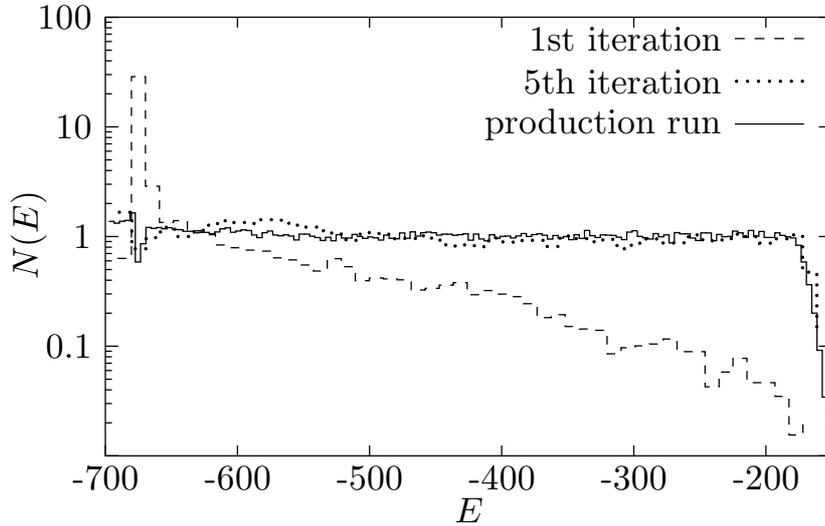


Figure 3.5: Energy histogram as computed after indicated runs, for $q = 5$, $\sigma = 0.9$, $L = 400$ spins, using Eq. (3.6) to compute the initial effective Hamiltonian $\bar{H}_{\mu}(E)$ from a previous run at $L = 200$ spins. Labeling on y axis indicates normalized probabilities.

More iteration runs are thus required to obtain a perfectly flat histogram as σ is decreased, and the benefit of this approach in effect becomes negligible for $\sigma \leq 0.3$. Indeed, the algorithm then spends a great number of iteration steps being trapped in low energy levels, seeking to rectify the shape of the density of states in this energy region until convergence is obtained: starting from an initial canonical effective Hamiltonian actually yields better performance. Since for systems with long-range interactions, the computation time behaves at least like L^2 , using this “scaling trick” thus greatly reduces the time needed for proper convergence, at least for $\sigma > 0.3$; this partially makes up for the lack of a hybrid multicanonical-cluster algorithm dedicated to long-range models⁴.

3.6 Algorithm performance

In order to measure the performance of the present implementation, a dynamic exponent z is computed, which is defined as the scaling exponent of a relevant characteristic time τ of the simulation, i.e., $\tau \propto L^z$, where L is the lattice size: while at second-order transitions, the integrated autocorrelation time represents such a relevant time (see Chap. 2), at first-order transitions *tunneling* times through the energy barrier (τ_{tun}) prove to be a more meaningful indicator [183].

3.6.1 Tunneling times

Tunneling times are defined as one quarter of the average number of Monte Carlo steps per spin (MCS) needed to travel from one peak of the reweighted energy histogram ($N_{rw}(\beta, E)$) to the other and back, with the temperature being set to the transition temperature. Tunneling times are expected to grow exponentially with L for canonical algorithms, and to scale as a power law of L for multicanonical algorithms [31] (see Sec. 2.2.3). In both cases, it appears that the chosen characteristic time is a good indicator of how quickly the demands in CPU time should grow with increasing lattice size: for second-order transitions, this is the time needed to generate truly independent samples, while for first-order transitions, this tells us at what rate the dynamics spreads out over the energy barrier and thus to what extent samples get efficiently picked from the two phases in coexistence.

3.6.2 Autocorrelation times

The integrated autocorrelation time is computed by using the time-displaced correlation function presented in Sec. 2.2.1. This function displays an exponential-like short-time behavior $\sim e^{-t/\tau}$; τ is then derived from a simple integration scheme. Since the latter function makes sense within equilibrium only, n thermalization steps are first discarded, where n is obtained by using the nonlinear relaxation

⁴Such an hybrid algorithm will be introduced in Chap. 6

3.6. Algorithm performance

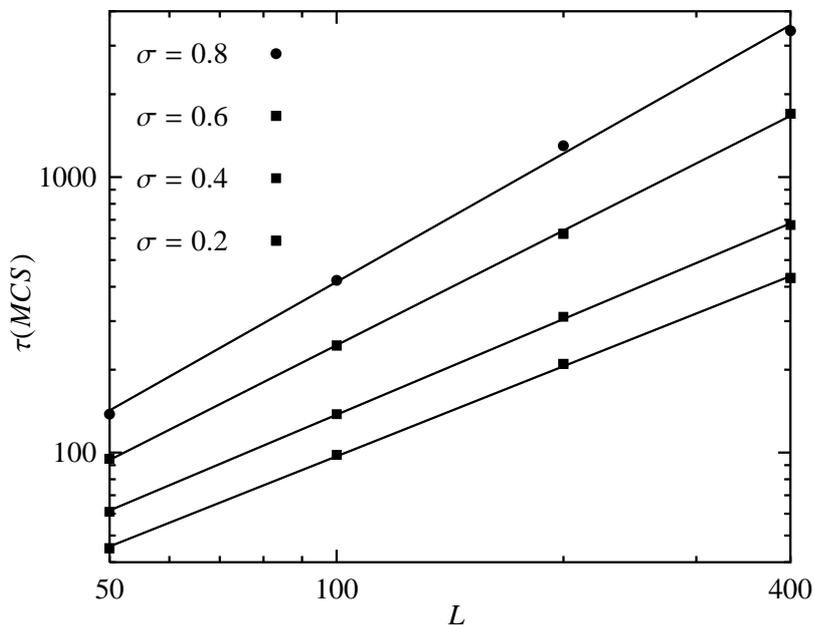


Figure 3.6: Integrated autocorrelation time τ vs lattice size L for $q = 7$ and $\sigma = 0.2, 0.4, 0.6, 0.8$. Dynamic exponents computed from a fit to L^z are $z = 1.09(1), 1.15(1), 1.38(1), 1.55(1)$, respectively.

function that describes the approach to equilibrium [216] and averaging over several dry runs. An interesting point regarding multicanonical simulations is that, since they are random walks in energy space, “thermalization” (although this term is no longer appropriate as far as generalized ensembles algorithms are concerned) always occurs rather rapidly; simulations based on a nearly flat histogram show that a value of 1000 MCS is indeed appropriate on average⁵.

3.6.3 Results

Results for $q = 7$ and σ lying between 0.2 and 0.8 are shown in Fig. 3.6 for integrated autocorrelation times, and in Fig. 3.7 for tunneling times. The slight dispersion in the power-law fits arises from the fact that simulations at larger sizes are conducted with a higher number of MCS between measurements in order to reduce memory overhead. Where computing tunneling times is concerned, this results in some tunneling events being possibly skipped and the average tunneling time being overestimated. Both figures show, however, that a power-law behavior is perfectly plausible. In the case of first-order transitions, the reduction in simulation costs is thus drastic in comparison with standard canonical algorithms.

⁵In this regard, a much more efficient scheme based on a χ^2 regression will be presented in Chap. 7.

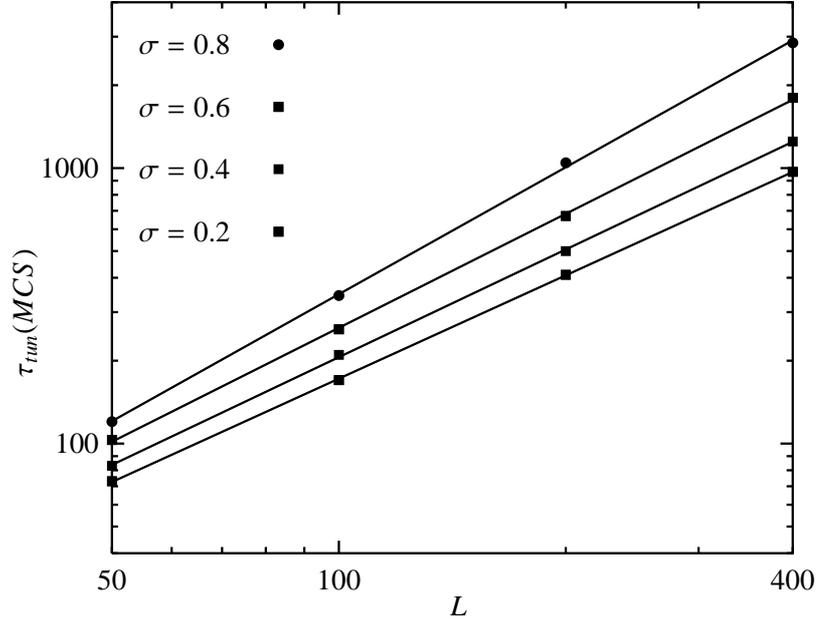


Figure 3.7: Tunneling time τ_{tun} vs lattice size L for $q = 7$ and $\sigma = 0.2, 0.4, 0.6,$ and 0.8 . Dynamic exponents computed from a fit to L^z are $z = 1.25(1), 1.30(2), 1.37(1), 1.53(1)$, respectively.

3.6.4 Discussion and conclusion

For both indicators, an average z slightly above 1.0 is obtained for $\sigma = 0.2$, yet z increases smoothly with decreasing range of interaction. This may be accounted for by the fact that spatial and time correlations grow as one departs from the mean-field regime and approaches the short-range regime. As for tunneling times, the prefactor turns out to be slightly higher near the mean-field regime, and z increases at a lower rate with increasing σ than is the case for autocorrelation times.

Since there are no other numerical studies of long-range spin models based on multicanonical algorithms⁶, comparison is limited to estimates obtained for nearest-neighbor models. For the three-state Potts model, canonical simulations using local updates led to $z = 2.7$ [344], while Swendsen and Wang obtained $z \sim 0.6$ using the eponymous cluster algorithm [309]. For further comparison, the Metropolis algorithm applied to the (nearest-neighbor) Ising model in $d = 2$ and $d = 3$ yielded a value of z slightly above 2 [342], whereas Wolff's cluster algorithm led to $z \sim 0.27$ [348].

To conclude, and in view of these results, two remarks are in order:

- first, the dynamic exponents of the long-range Potts chain are moderate, even at large decay parameter; it turns out, however, that prefactors increase rapidly with σ ; with regard to this, the fact that first-order transitions occur

⁶To the best of my knowledge.

3.6. Algorithm performance

for a value of σ *below* a given threshold makes the method an appropriate choice for the investigation of the model.

- owing to the first-order nature of the transitions, a comparison with a canonical algorithm (even using Luijten-Blöte cluster algorithm, Sec. 2.3.2) would not make sense here (except perhaps at large σ where the transition is weakly first order); as can be witnessed in Fig. 3.1, the suppression of mixed-phase configurations is quite large already for $L = 400$ spins, so that the method seems, here again, an adequate choice as against a Metropolis-based algorithm;
- In view of the marked reduction in dynamic exponents brought about by cluster methods, one may ponder on the benefit of these methods in the context of multicanonical simulations. As mentioned in this chapter, however, combining both schemes is not straightforward: this will be the purpose of the second part of this thesis. In this regard, and in anticipation of Chap. 7, it is perhaps worthy of mention that the breathing cluster method developed in the second part of this thesis will *indeed* lead to a clear reduction of both dynamic exponents and prefactors for long-range models.

Chapter 4

From mean-field to the *tricritical* line

As was discussed in Chap 2, one of the main problems of *canonical*¹ Monte Carlo simulations of phase transitions is the need to perform a large number of simulations at distinct temperatures in order to obtain *reliable* reweighted averages of thermodynamic data over a sufficiently large range of temperature. A second problem surfaces in the context of first-order phase transitions, where supercritical slowing down leads to very repetitive dynamics and unreliable statistics. The multicanonical method presented in the previous chapter represents an efficient way to tackle both problems “at one whack”.

In the present chapter, I will show how one can take advantage of this method to investigate the phase diagram of a Potts chain governed by power-law decaying interactions. The investigations will center around:

- the strong first-order regime, i.e., at low decay parameter,
- the tricritical² line separating the first- and second-order regimes.

The chapter begins with a review of the prominent features of the nearest-neighbor model, as well as some important mean-field results. Previous investigations of the phase diagram of the long-range model are described in Sec. 4.1.2 and 4.1.3. Results from previous Monte Carlo methods are reviewed in Sec. 4.2.3. The critical behavior of the model in the second-order regime is examined in the next chapter.

¹In the canonical ensemble.

²The term “tricritical” might have something of a misnomer, since in the Ginzburg-Landau functional of the model, the decay parameter σ is at the same time a conjugated — or non-ordering — field *and* an exponent of the wave-vector k (see Sec. 5.1.1). I will stick to this terminology on *pragmatic* grounds, however, owing to the fact that (i) it evokes the picture — usually associated with this term — of a point separating a first- and a second-order line of transitions, and (ii) there are other “boundaries” in the phase diagram of the model investigated here, and the present terminology might help reducing confusion.

Chapter 4. From mean-field to the tricritical line

In Sec. 4.2, it will be demonstrated that, in spite of the modest lattice sizes investigated, the multicanonical method yields very precise estimates of transition temperatures, in far better agreement with mean-field predictions whenever $\sigma \rightarrow 0$ than in previous Monte Carlo studies (Sec. 4.2.3). This pertains to the superiority of the method at first-order transitions, yet also to the implementation of infinite-image periodic boundary conditions (Sec. 4.2.1). Indicators based on free energies will also prove to play an important role in a reliable estimation of transition temperatures.

As regards the second point, i.e., the location of the tricritical line, the weakening of the first-order transition as the line is approached presents a particular challenge. Several other approaches are reviewed (Sec. 4.3), most of them inconclusive in this respect. With this in mind, a novel method is introduced for the detection of the order of phase transitions, which relies on the location of spinodal points. The multicanonical method makes it possible to reweight over a large range of temperatures, and to reliably determine the position of metastable states: its versatility in this respect will thus prove essential here. I will illustrate the method for $q = 3, 5, 7, 9$, and prove that the position of the tricritical line can be determined to an unprecedented two-digit precision.

Finally, the controversial nature of the phase transition on the $\sigma = 1$ line is considered in Sec. 4.3.6. I will point out that the asymptotic behavior of the tricritical line is crucial in this respect, since it might confirm or rule out one or more scenarios suggested in recent time. By carrying out a careful finite-size scaling analysis, a new, unusual finite-size effect is found out, with a first-order transition gradually waning into a second-order one in the thermodynamic limit (Sec. 4.3.7). This behavior, I claim, clearly washes out some previous claims, and lends support to one of the suggested scenarios on this line: $\sigma = 1$ is a line of topological phase transitions [72]. A discussion of this atypical finite-size effect is provided: it is suggested that several causes enter this effect, including — but not restricted to — boundaries effect.

4.1 Model, existing results and unsettled controversies

Throughout this chapter and the next chapter, I will consider a one-dimensional ferromagnetic Potts model incorporating power-law decaying interactions. This model is derived from a generalized q -state Potts Hamiltonian, i.e.,

$$H = -\frac{1}{2} \sum_{i \neq j} J_{ij} \delta_{\sigma_i, \sigma_j} - \sum_i h_i \delta_{\sigma_i, \sigma_0}, \quad (4.1)$$

where the Potts spin σ_i at site i can take on the values $1, \dots, q$, and the first sum runs over all pairs of sites. h_i is an external aligning field favoring condensation in state σ_0 , yet numerical investigations in what follows are conducted with no external magnetic field. Incorporation of power-law interactions is carried out by

4.1. Model, existing results and unsettled controversies

setting

$$J_{ij} = J(|i - j|) = \frac{1}{|i - j|^{D+\sigma}}, \quad (4.2)$$

where $D = 1$ throughout this work except when otherwise specified, and σ is an adjustable parameter that — broadly speaking — “alters” the interaction range. For $\sigma \leq 0$, the interaction is no longer integrable; this corresponds to rather different — though interesting — physics³, and I shall not consider this case in the present work. Rather, I will concentrate on the so-called “integrable” regime⁴, where the phase diagram of the model exhibits not less appealing behavior, including first- and second- order transitions, Kosterlitz-Thouless like transitions, and intriguing crossover effects.

It is perhaps interesting to recall that, as σ falls off to -1 , this model tends⁵ to the exact mean-field regime where all interactions have equal strength, so that one *might* expect to recover, for a sufficiently low value of σ , the features of the mean-field model⁶. Conversely, the limiting case $\sigma \rightarrow \infty$ corresponds to a pure short-range model with nearest-neighbor interactions, and the absence of phase transition at finite temperature for $\sigma > 1$ has been rigorously proven [107, 108, 300, 124]. The issue related to the crossover between the long-range and the short-range regime has been the matter of long-running debate, and is reexamined in Chap. 5.

The critical behavior of this model is studied by way of the following order parameter,

$$m = \frac{q \max_n \rho_n - 1}{q - 1}, \quad (4.3)$$

(hereafter termed “magnetization”) where $n = 1, \dots, q$, and ρ_n is the density of Potts spins in phase n . The last quantity varies between $1/q$ at infinite temperature and 1 in the ground state, so that $m = 1$ in the ordered phase, and 0 otherwise.

The Potts model can be mapped onto an O_{q-1} -like vector model (though with stronger constraints on the degrees of freedom). This offers the benefit of transforming the Kronecker delta symbol into a sometimes more tractable dot product, e.g., for field-theoretic renormalization purpose, or as will be witnessed in Chap. 7, whenever the algorithm efficiency requests so (see [279, 333] for two possible formulations, and also Sec. 7.3.1 where I introduce a slightly different mapping, better suited for numerical evaluation). Another useful representation is that of Mittag and Stephen [247], where the set of spin states $1 \dots q$ is mapped onto the set $1, \omega, \omega^2, \dots, \omega^{q-1}$, where $\omega = e^{2i\pi/q}$ is a q th root of unity.

³See references given in Chap. 1 for connections with Tsallis’s thermodynamics

⁴See a note on a semantic row in this respect in Sec. 1.1.

⁵After proper normalization of the coupling constant so that the thermodynamic limit is not ill-defined

⁶In this regard, it has been suggested in a recent paper that, if mean-field theory predicts a first-order transition for any long-range interaction that satisfies the so-called reflection-positivity condition, then — as a sufficient and necessary condition — there is a first-order transition for some sufficiently spread out interactions [42].

Nearest-neighbor model The nearest-neighbor Potts has been extensively studied since its “discovery” by Domb and Potts in 1952 (and independently by Kihara in 1954). Its critical properties and its applications to microscopic modeling in physics were extensively reviewed in [350]. For instance, the limit $q = 0$ and $q = 1$ describe the resistor-network problem and percolation, respectively. Non-integer values are also of interest in the modeling of dilute spin glasses ($q = 1/2$) or branched polymers ($0 \leq q \leq 1$). Using duality arguments, Baxter obtained the critical temperature of the two-dimensional model, $T_c = 1/\ln(1 + \sqrt{q})$, $q > 2$ [17]. The critical temperature is not known exactly in three dimensions. The model is also expected to undergo a first-order transition for sufficiently high values of q . For instance, it was proven that this is the case for $q \geq 4$ in two dimensions, $q > 3$ in three dimensions (not known exactly), and $q \geq 2$ in four dimensions, [4, 17, 350], so that the threshold value q_c above which the transition changes from second- to first-order is expected to decrease continuously all the way up to $D \rightarrow \infty$.

4.1.1 Mean-field theory

Mean-field theory for the Potts model was initially formulated by Kihara et al. using Bragg-Williams theory (see, e.g., [350]). The most prominent result of mean-field theory is that the model undergoes a first-order transition for $q > 2$.

Mean-field behavior can also be obtained quite straightforwardly by using the variational mean-field method, which relies on the minimization of the free energy expressed as a functional of a trial density matrix, whereby the Boltzmann density matrix is recovered at the minimum. The variational method is very general⁷, and its mean-field formulation simply consists in factorizing the trial density matrix in terms of independent one-site density matrices, as in standard molecular-field theory. This is roughly equivalent to neglecting fluctuations to second-order in Bragg-William’s theory, although a treatment of the Potts model with $q > 2$ is somewhat less tractable with the latter approach⁸.

A detailed calculation is presented in Appendix A. I will here only review the main results. In particular, I will focus on the location of spinodal points: these form the core of a novel method to detect the order of phase transitions, which I introduce in Sec 4.3. In the presence of power-law decaying interactions, the mean-field free energy per spin is expressed with respect to the order parameter (up to an additive constant) as

$$\begin{aligned} \frac{qf(m)}{q-1} = & -hm - \zeta(1 + \sigma)m^2 + kT\{(1 - m)\ln(1 - m) \\ & + \frac{1 + m(q - 1)}{q - 1} \ln[1 + m(q - 1)]\}, \end{aligned} \quad (4.4)$$

where $\zeta(1 + \sigma)$ is the Riemann zeta function and h an external aligning field (see Eq. A.2). The graph of $f(m)$ is depicted in Fig. 4.1 for $q = 9$ and $\sigma = 0.5$. Equilib-

⁷See, for instance, [76] for an introduction to the method.

⁸This requires mapping the model to an O_{q-1} model.

4.1. Model, existing results and unsettled controversies

rium values of the magnetization are located at minima of the free energy. At the transition temperature T_c , two minima are clearly visible, which in the thermodynamic limit signals a first-order transition associated with a jump in the magnetization and a latent heat. This is always true whenever $q > 2$, where the transition temperature is given by

$$kT_c = \zeta(1 + \sigma) \frac{q - 2}{(q - 1) \ln(q - 1)}. \quad (4.5)$$

For $q = 2$ (the Ising model), a second-order transition occurs (i.e., with critical fluctuations induced by $F''(0) = 0$ at the transition⁹), and

$$kT_c = \zeta(1 + \sigma). \quad (4.6)$$

which is also recovered by taking the limit $q \rightarrow 2$ in Eq. (4.5).

The zeta function expands as $1/\sigma$ around $\sigma = 0$, thus transition temperatures are expected to vary as $1/\sigma$ in the vicinity of the mean-field regime. Conversely, $\lim_{\sigma \rightarrow \infty} \zeta(1 + \sigma) = 1 > 0$, which confirms — as expected — that mean-field theory is a very poor approximation in the short-range regime.

Metastable states are defined from the condition that either one of the two minima disappears, i.e., both the first and the second derivative of the free energy must be zero (this is indicated with horizontal arrows in Fig. 4.1). This yields

$$kT_1 = \frac{2\zeta(1 + \sigma)}{q}$$

corresponding to the extremum at $m = 0$ becoming unstable. The temperature kT_2 at which the second extremum vanishes is unfortunately not available in closed form (see Appendix A). Thus there are two temperatures of metastability T_1 and T_2 on each side of T_c , that physically correspond to the limit of stability of each subphase. These are depicted in Fig. 4.2 for $2 \leq q \leq 10$. Inside the interval $[T_1, T_2]$, there exists two values of the order parameter corresponding to a null curvature of the free energy. This is known to induce a long-range (i.e., low wave number) instability that can trigger a phase transition through the spinodal decomposition mechanism [141]. In the following, I may denote T_1 and T_2 as *spinodal temperatures*, and the corresponding values for the order parameter at the inflexion point as *spinodal points*. Alternatively, one may also express the free energy as a function of the mean-field energy $E = -\zeta(1 + \sigma)m^2$ and obtain the same spinodal temperatures from the joint condition $f'(E) = f''(E) = 0$. Obviously, the two expressions of f do not have the same shape.

An important point that will be invoked later in this chapter is that the width of the spinodal curve $T_2 - T_1$ shrinks to zero as $q \rightarrow 2$. This indeed accounts for the second-order nature of the transition at $q = 2$, since in this limit the two minima merge into a single large minimum responsible for the well-known divergence of fluctuations at a continuous transition.

⁹For the same reason, spurious critical behavior may also occur at weakly first-order transitions in the vicinity of a local minimum of the free energy; this may show up in renormalization group scenarios.

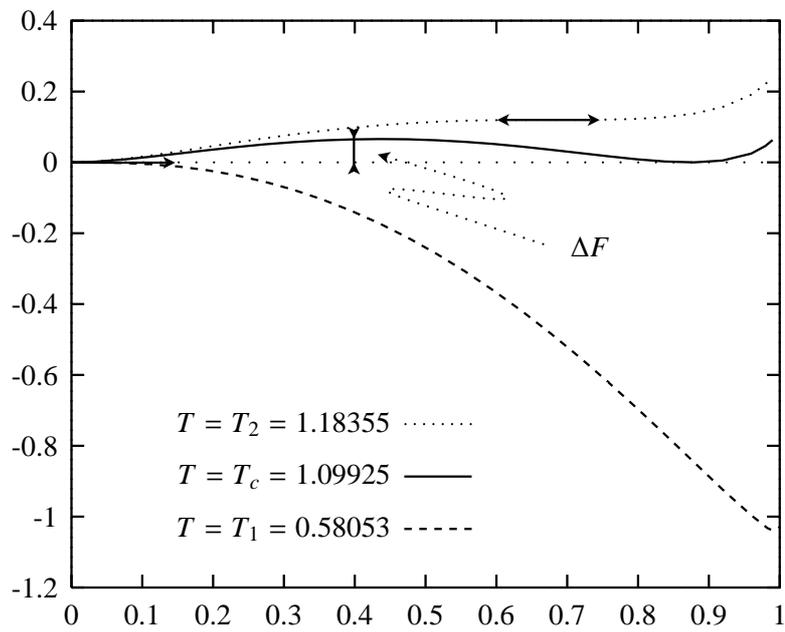


Figure 4.1: Free energy per spin in the mean-field approximation for the nine-state Potts model in zero external field, and with $1/r^{3/2}$ interactions. T_1 and T_2 refer to the temperatures of metastability, while T_c designates the transition temperature corresponding to the two minima having the same magnitude. ΔF is the surface tension (the *reduced* surface tension is usually defined from $\Delta F/L^{D-1}$ for nearest-neighbor models).

4.1. Model, existing results and unsettled controversies

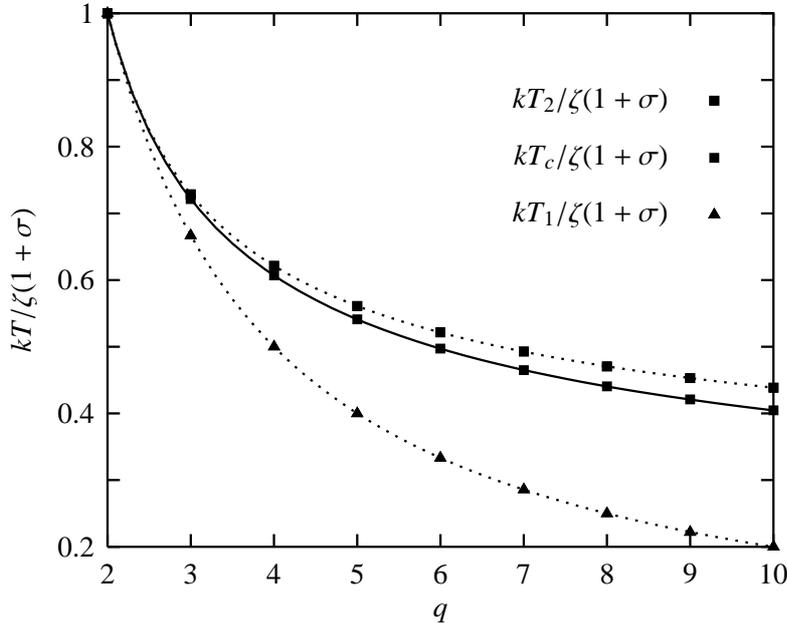


Figure 4.2: Spinodal temperatures T_1 and T_2 , along with the transition temperature kT_c , as a function of q in the mean-field approximation. Shown are quantities normalized with respect to $\zeta(1 + \sigma)$.

4.1.2 Finite range scaling and cluster mean-field approaches

Apart from Monte Carlo simulations (which are reviewed in Sec. 4.2.3 below), long-range chains were studied by means of two distinct numerical approaches: finite range scaling (FRS) combined with an extension of the transfer matrix method to long-range interactions [130], and the cluster-mean field method [251]. Both methods mainly targeted the estimation of critical temperatures, although the former also addressed possible crossover issues by relying on the finite-range scaling behavior of the correlation length exponent. However, none of these methods concerned itself with the order of the transition. I hereafter briefly review them in some detail; Section 4.2.3 contains a comparison of the numerical results that were obtained with these methods [130, 251] with the estimates I obtain with the multicanonical method; Section 5.3 in the next chapter deals more specifically with crossover issues.

Finite range scaling combined with transfer matrices

This method was initially introduced by Glumac and Uzelac [323, 129] in the context of Ising chains with power-law decaying interactions, and extended to the Potts model in an ensuing study [130]. More recently, it was applied to the $S \geq 1/2$ spin chain [16] and to a two-sited (i.e., AB model) Ising chain [15]. The method targets the estimation of critical exponents, transition temperatures and correlation lengths

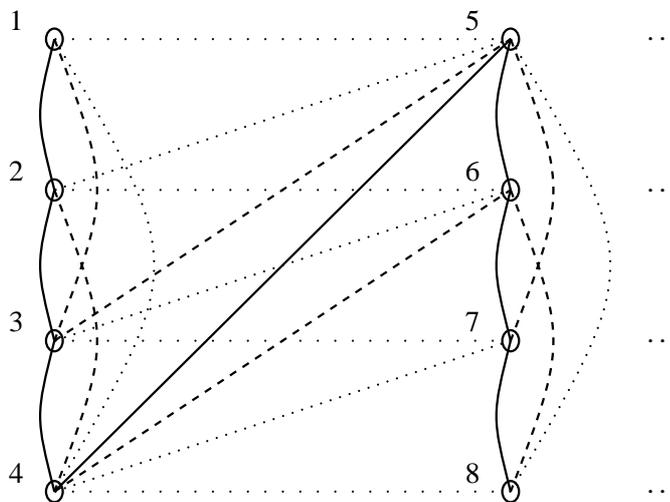


Figure 4.3: Transfer matrix method for finite-range interactions: the zig-zag scheme brings the problem back to a conventional transfer matrix problem, with each column interacting with its nearest-neighbor only.

in the critical regime; yet it has remained rather inconclusive where distinguishing between first- and second-order transitions is concerned.

The method comprises two essential ingredients. First, it generalizes to long-range chains the transfer matrix method in its variable strip-width flavor, e.g., the (numerical) approach used for two-dimensional Potts models with nearest-neighbor interactions. However, the technique cannot be used as is with infinite range interactions, as will be witnessed below, and prerequisite is to truncate the power-law decaying interaction below some finite cut-off M . It was suggested [129] that response functions should behave as $\chi_M(t) \sim \chi_\infty(t)f(M/\xi_\infty)$, where ξ_∞ is the infinite cut-off correlation length, and $f(u)$ denotes a dimensionless finite-range scaling function that prevents critical divergences at finite cut-off¹⁰. Several numerical calculations are performed at increasingly large cut-offs, and (this the second ingredient) a finite-range scaling scheme is carried out with respect to M : for two distinct cut-offs M and M' , the correlation length scales according to $\xi_{M'}(t) = (M'/M)\xi_M[(M'/M)^{1/\nu}t]$. Critical exponents are thus estimated from an extrapolation to the infinite cut-off limit.

In order to transform the finite-range problem into a nearest-neighbor-like problem that may become tractable with conventional transfer matrix methods, spins are grouped by blocks of M subsequent spins, where each block represents a matrix column; the chain is treated in the thermodynamic limit, and the strip count increases with the interaction cut-off. Interactions between spins are handled in a zig-zag fashion, as illustrated in Fig. 4.3 for $M = 4$. This makes it possible to restrict the interactions to neighboring columns only, so that the partition function

¹⁰It was claimed by the authors that this scaling behavior should hold also in the mean-field regime, contrary to what occurs for finite-size scaling.

4.1. Model, existing results and unsettled controversies

of the system is simply given by the trace of the n th power of the transfer matrix over q^M possible states of the column vectors. The correlation length is obtained in the standard way from the two largest eigenvalues λ_1 and λ_2 of the transfer matrix, i.e., $\xi_M = M/\ln(\lambda_1/\lambda_2)$, and the correlation length exponent follows from a finite-range scaling extrapolation based on $1/\nu = \ln[\xi'_{M'}/\xi'_M]/\ln(M'/M)$. Here, ξ' refers to the derivative of the correlation length with respect to the critical temperature, estimated at the critical point. Interestingly enough, it was proven that the long-range nature of interactions allows one to decompose the transfer matrix into much simpler matrices, so that the complexity of the calculation could be drastically reduced. The method was also successfully extended to the q -state Potts with a non-integer q value [130], by relying on a bond representation of this model, and tracing over bond configurations instead of spin configurations.

In [130], Glumac and Uzelac applied this method to the q -state Potts chain with interactions decaying as $1/r^{1+\sigma}$, with σ ranging from 0.1 to 1.0 and q chosen between 1 and 64 for the integer case, and 1/16 and 1/2 for the non-integer case. Estimates of critical temperatures are claimed to support a precision of one per cent, which is clearly less accurate than the coherent anomaly method [251]. Since the transfer matrix is a $q^M \times q^M$ matrix, however, extrapolations are in practice restricted to modest cut-off values, i.e., around 20 for q not too large. It should be mentioned, however, that the results obtained recently by Barati and Ramazani [16] for the $S \geq 1/2$ Ising chain are somewhat more accurate, which suggests that the choice of a different extrapolation procedure might help in this respect; yet the last approach also relied on an extrapolation based on the exactly known value of the correlation length exponent in the classical regime, and this was not the case in [130].

In spite of the limitations of the method in terms of transfer matrix sizes, the authors in [130] were able to shed partial light on some controversies: by relying on the non-monotonic behavior of the critical temperature with respect to σ , and observing that the correlation length exponent ν increases as $\sigma \rightarrow 1$, they argued that the change of regime from short-range to long-range behavior (more on this in Sec. 5.3) should take place at $\sigma = 1.0$, but overall failed to (i) reproduce the essential singularity at $\sigma = 1.0$ predicted in [208, 72] (see Sec. 4.3.6 below), and (ii) bring clear evidence as for the order of the transition, with a notable inconsistency between the behavior of ν at large q (suggesting a change of order) and the absence of crossing of the largest and the third largest eigenvalues of the transfer matrix [169].

Cluster mean-field

A second approach, the so-called *cluster mean-field* method was developed by Monroe and applied to the Ising and q -state Potts chains [250, 251]. This method is a close relative of the coherent anomaly method (CAM) proposed by Suzuki [305] in the mid 80's. It is also somewhat reminiscent of recursive methods based on Bethe lattice approximations [248, 249], and might be traced back to Kadanoff's

decimation scheme in its spirit.

In this approach, interactions between spins in the same cluster are handled exactly, while interactions between a given spin and spins belonging to other clusters are treated in a mean-field manner. The resulting n -site Hamiltonian (where n is the cluster size) involves a mean field term, which can be obtained self-consistently. By construction, the (improved) mean-field nature of the approximation results in the method yielding upper bounds of critical couplings. By carrying out several calculations at increasingly large clusters (i.e., containing between 1 and 9 sites), and making use of an extrapolation scheme (the Vanden-Broeck-Schwartz transformations, which is generalization of the Padé approximant method), Monroe was able to obtain very precise estimates of transition temperatures supporting a four-digit precision, at least for interaction decaying slowly enough with the interparticle distance (a feature which is indeed consistent with the mean-field approximation that underlies the method). It turns out that, for $q > 2$, these are the most precise estimates to date along with those I will obtain with the breathing cluster method (see Sec. 7.6 for a deepened comparison). In the Ising case, cluster mean-field and Monte Carlo calculations performed by Luijten and Blöte [231] yield similarly accurate results.

4.1.3 Results from real-space renormalization

Several results for the Ising and Potts chains with power-law decaying interactions were obtained by means of a real-space renormalization group approach [63, 64]. This method is based on three ingredients: (i) Kadanoff's blocking method; (ii) a non-linear transform using a (modified) majority rule, the so-called *equally probable tie-breaking*, whereby subgroups of similar spins representing the majority in a given block are equally weighted in case of conflict; (iii) the cumulant expansion method of Niemeijer and van Leeuwen¹¹, whereby the Hamiltonian is first split up into a term H_0 involving intra-block interactions and a term V corresponding to the interactions between spins in distinct blocks, and then the renormalized Hamiltonian is obtained from a cumulant expansion of the average (i.e., intra-block partial traces) of e^{-V} with respect to the Hamiltonian H_0 ; in [64], an expansion to first order is considered.

In the context of long-range interactions, the generalization of Kadanoff's construct introduced in [63, 64] consists in building blocks of spins of size b , and extrapolating the results to the limit $b \rightarrow \infty$. Since the terms in the cumulant expansion behave as $1/b^{1+\sigma}$, the authors expect this extrapolation to give good results in the long-range case.

From their recurrence equation (Eq. (17) in [64]), the authors obtain, for all $q \geq 2$ and in the limit $b \rightarrow \infty$, a trivial unstable fixed-point at infinite temperature for $\sigma \rightarrow 0$ and at zero temperature for $\sigma > 1$ (recovering the exact result in [108]), and a non-trivial unstable fixed-point at finite temperature between these two limiting

¹¹A very explanatory introduction can be found in [21] in the context of the renormalization of the Ising model on the triangular lattice.

4.2. Estimates of transition temperatures from multicanonical simulations

values. From the analysis of the change of curvature of $T_c(\sigma)$, they assess a non-zero transition temperature $T_c = \pi^2/12$ on the line of inverse square interactions, and suggest that this temperature is independent of the number of states of the model (thus yielding a discontinuity in T_c that should be independent of q). While the authors claim that this assumption is supported by previous works in the Ising case, it contradicts the results obtained in [130], whereby $1/T_c$ increases with q . Along the same lines, the authors obtain a correlation length exponent $\nu^{-1} \sim 2(1 - \sigma)$ as $\sigma \rightarrow 1$ for all $q \geq 2$, a result which reproduces the expected divergence on the line of inverse square interactions, yet contrasts with the expression obtained by Kosterlitz using a momentum-shell renormalization method [208]. It is mentioned by the authors, however, that their extrapolation procedure $b \rightarrow \infty$ should not be expected to yield reliable results near $\sigma \sim 1$.

In the region $\sigma \rightarrow 0$, i.e., when the model approaches the regime of non-integrable interactions, the authors report (under the assumption of a continuous transition) the following scaling behavior for the critical temperature, $T_c \propto 1/\sigma$, and claim that this behavior underpins Tsallis's conjecture [322] proposed in the context of non-extensive thermodynamic. As will be witnessed later in this chapter (Sec. 4.2), the estimates I obtained with a multicanonical approach lend clear support to this conjecture, yet *irrespective* of the order of the transition.

Incidentally, and as the authors hasten to recall, these results are based on the assumption of a continuous phase transition. Yet in the large q limit, mean-field theory was proven to become exact [276], and a first-order transition should therefore be expected for either a sufficiently large value of q or a small enough power-law exponent σ . This behavior is indeed consistent with the observation of a first-order transition for $\sigma \leq 0.6$ in the three-state model, that was reported in a later numerical work based on standard Monte Carlo simulations [131, 132]. It is also consistent with the phase diagram I obtained for this model using a multicanonical method (see Sec. 4.3.5). It is worth stressing here that real-space renormalization methods based on such majority rules hardly enable one to distinguish between first- and second-order transition, and a generalization to a lattice-gas model incorporating vacancies [263] should indeed be used for this purpose. To the best of my knowledge, such a generalization has never been performed in the context of long-range spin models.

4.2 Estimates of transition temperatures from multicanonical simulations

In this section, I present an extensive investigation of the transition temperatures of the long-range Potts chain. Multicanonical simulations are performed for a large range of decay parameters and values of q , using the algorithm presented in the previous chapter. To set the stage, I will briefly discuss the influence of periodic boundary conditions in the presence of a long-range interaction. I will then describe the method employed for the estimation of transition temperatures. A last

paragraph provides a discussion regarding the comparison of my results with mean-field predictions and other methods.

4.2.1 Periodic boundary conditions

In numerical simulations, systems of *finite* geometry are simulated, and it is vital that efficient boundary conditions be implemented, since these generally influence the rate at which thermodynamic quantities reach the infinite-size limit as size is increased. Periodic boundary conditions are the most widespread choice, inasmuch as the (initially broken) translational invariance of the system is restored, and convergence toward the thermodynamic limit is improved. The general prescription for spin models with nearest-neighbor ferromagnetic interactions¹² is to use *first-image* conditions, i.e., the lattice is simply wrapped around itself (from a topological viewpoint, this yields a D -dimensional torus), so that each spin feels the same local environment, irrespective of its position on the lattice.

In the presence of long-range interactions, the choice is no longer unique:

- one may impose some cut-off on the interaction range, so that first-image conditions may be used as in any model with a finite-range interaction; in a previous Monte Carlo study of the q -state Potts chain [19], for instance, the cut-off was set to $L/2$, where L is the size of the chain;
- or one may use so-called *infinite-image* boundary conditions, whereby the lattice is wrapped an *infinite* number of times around itself, and each spin interacts in effect with an infinite number of replicas of the original lattice. For a one-dimensional lattice of size L , this consists in connecting the two end-points of the chain in a ring-like fashion, and computing the interaction between two given spins at a relative position r by walking along the ring an infinite number of times and adding up every interaction at distance $r + nL$, $n \in \mathbb{Z}$.

A comparative study of these two choices in terms of finite-size effects was carried out in [67]. For models exhibiting a phase transition at finite temperature, it was established that infinite-image conditions are much more efficient than single-image conditions in wiping out finite-size effects, especially at medium lattice sizes. In any case, using the former conditions is no more costly in terms of computational resource than first-image conditions: the long-range coupling needs just be “renormalized” in order to take all possible interactions into account, and from the viewpoint of numerical implementation, only interactions between spins of the *original* lattice must then be considered.

For a one-dimensional lattice and power-law decaying interactions, the “bare” coupling constant $J(r) = 1/r^{1+\sigma}$ has to be replaced by $\tilde{J}(r) = \sum_{n=-\infty}^{+\infty} J(r + nL)$. For

¹²And generally, any finite-range interaction whose support is smaller than the lattice linear size.

4.2. Estimates of transition temperatures from multicanonical simulations

the purpose of numerical evaluation, this sum may be reexpressed as

$$\tilde{J}(r) = \frac{1}{r^{1+\sigma}} + \frac{1}{L^{1+\sigma}} \left[\zeta\left(1 + \sigma, 1 + \frac{r}{L}\right) + \zeta\left(1 + \sigma, 1 - \frac{r}{L}\right) \right], \quad (4.7)$$

where $\zeta(s, \alpha)$ denotes the generalized Hurwitz-Riemann zeta function¹³. The self-energy corresponding to the interaction of each spin with itself in each replica is obviously omitted since it is just an additive constant to the total energy.

Slightly distinct conditions were used in [200] in the context of a long-range spin glass, whereby instead of considering the topological distance along the chain, the authors made use of the euclidian distance: for the one-dimensional chain, this amounts to setting the distance r entering the coupling constant $J(r)$ to $r = \sin \pi|i - j|/L$ where $|i - j|$ is the distance *along* the chain. In this case, infinite-image conditions just make no sense.

4.2.2 Methodology

The multicanonical simulations presented hereafter were performed for $q \in [3, 9]$, using for each value of q an appropriate set of σ parameters between 0.3 and 0.9, so that a variety of transition strengths could be observed. As will be observed in Sec. 4.3.5, though, the shape of the line separating the first- and the second-order regimes is such that the second-order regime is reduced to a tiny portion of the phase diagram (except perhaps at $q = 3$). This means that most values of σ investigated here correspond to first-order transition temperatures. Critical exponents in the second-order regime will be investigated in the next chapter in the context of the long-range to short-range crossover.

Once the density of states has been determined using the iteration process described in the previous chapter, a production run is performed for each lattice size between $L = 50$ and $L = 400$. The number of MC sweeps needed for each production run is computed so as to yield approximately 5×10^4 truly independent samples. Owing to the combination of rapidly soaring autocorrelation times — especially at large σ — and a computer load increasing with the square of the lattice size, I had to refrain from investigating larger lattice sizes, i.e., above $L \sim 400$. As regards the estimation of transition temperatures, this entails extrapolating infinite-size temperatures from linear sizes that all lie within one decade, e.g. approximately [40 – 400], and in this respect it was rather good surprise that the precision of the estimates could attain three digits. As is discussed in the next section (Sec. 4.3.5), this limitation creates a serious hindrance, however, when discriminating between first- and second- order transition is concerned, and there it will prove mandatory to meet the challenge from a different angle.

¹³This function is available from many software packages, e.g., the GNU Scientific Library (GSL) (<http://www.gnu.org/software/gsl/>). Recently, a port to the Java language was also made available (<http://sourceforge.net/projects/gsl-java>).

Finite size scaling at first-order transitions

At second-order transitions in finite geometry, the correlation length is bounded by the finite lattice size, so that response functions which diverge in the thermodynamic limit, actually experience a finite maximum: this maximum scales as a power law of the lattice size, with an exponent depending solely on the critical exponents of the model (see Sec. 5.2). At first-order transition, a similar behavior takes place, although it is now the *dimension* of the lattice which plays the role of a critical exponent. In finite geometry, the discontinuity, e.g., in the energy (connected to the latent heat), is rounded proportionally to the inverse lattice size, so that peaks of response functions are subject to finite-size scaling in much the same way as in continuous transitions.

Several theories of finite-size scaling at discontinuous transitions have been proposed (see, for instance, the seminal works of Fisher [117, 118] based on renormalization group arguments). In the following, I will mostly rely on the *double-gaussian* theory of finite-size scaling — the most useful from the viewpoint of numerical simulations — initially introduced by Binder and Landau [38] in the context of the field-driven transition in the Ising model below T_c , and extended to thermally-driven transition [79]. I will also briefly review the results yielded by the more rigorous treatment provided later by Borgs and Kotecký [48] for the nearest-neighbor model.

The double-gaussian theory is purely phenomenological, very general, and applies *a priori* to any first-order transition exhibited by a model with *short-range* interactions (the case of long-range interactions will be considered later on). At a first-order transition, the free energy displays two minima of equal magnitude corresponding to the ordered and the disordered phases, respectively. In the *vicinity* of each minimum, the phenomenological theory of Challa and Landau consists in approximating the free energy by its Taylor expansion limited to the second order; as a result, the histogram of the energy (or the magnetization in the case of a field-driven transition [38]) is approximated by the sum of two gaussian curves. Such an approximation is useful for deriving finite-size scaling behaviors for thermodynamic quantities related to the energy, e.g., the specific heat or the Binder cumulant; it fails, however, to describe the scaling behavior of the surface tension, since a two-gaussian scheme alone is clearly not suitable to describe mixed-phase configurations.

For a temperature-driven transition, as is the case for the q -state Potts model, the gaussian curves are centered at $E_o + C_o(T - T_c)$ and $E_d + C_d(T - T_c)$ respectively, where T_c denotes the temperature at *equal weights* [48, 49], i.e., when the area under both gaussian curves are the same (this amounts to imposing that the mean energy be the arithmetic mean of the energy of the ordered and disordered phases). C_o and C_d refer to the specific heat (in the thermodynamic limit) of the ordered and disordered phases, and the relations given above for the position of the curve imply that they are assumed constant in the vicinity of the transition. The width of each gaussian curve grows with the square root of the corresponding specific heat (per

4.2. Estimates of transition temperatures from multicanonical simulations

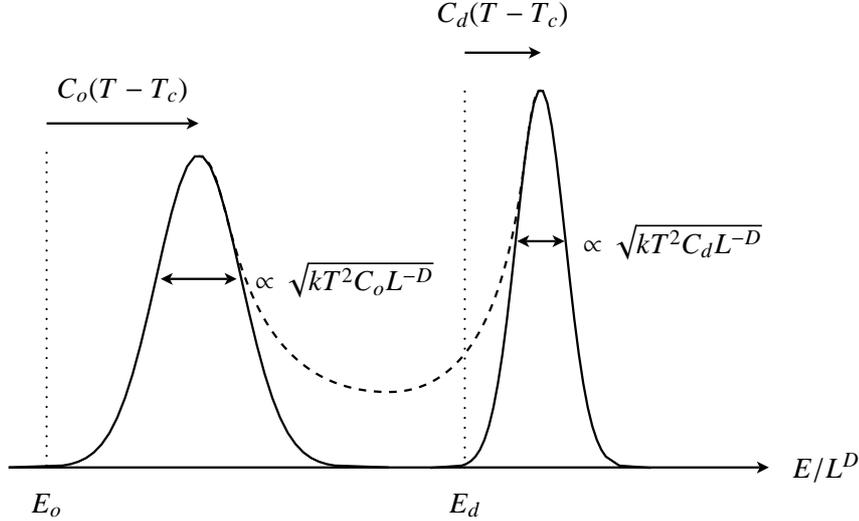


Figure 4.4: Sketchy illustration of the two-gaussian phenomenological theory of finite-size scaling at a first-order transition. The dashed curve represents the *true* distribution.

spin). Finally, E_o and E_d refer to the location of the peaks in the thermodynamic limit, so that $E_d - E_o$ is the latent heat associated with the transition. A sketchy illustration is given in Fig. 4.4.

In Fig. 4.5, it is illustrated how the parameters may be retrieved through a non-linear least-square fit to a reweighted histogram of the energy. This is hardly a very convenient way to estimate specific heats, however, as non-linear fits are notoriously difficult to cope with, especially when one wishes to automate the process through a piece of software. As a result, reweighted averages of the second moment of the energy usually yield far more reliable results; see, for instance, Fig. 4.9, where the mean energy and the specific heat are estimated separately for each subphase by reweighting over a subset of samples whose energy belong to the appropriate subphase. Additionally, it is clear from the figure that, as expected, the surface tension is not well described by the theory.

Still and all, the theory provides essentially what we are seeking, i.e., scaling laws for the statistical moments of the energy and the associated finite-size transition temperatures. By computing these moments with respect to the two-gaussian approximation of the distribution of the energy, one obtains that all quantities scale with the volume of the system [79], i.e., for instance,

$$T_c(L) \sim T_c(\infty) + aL^{-D},$$

$$\langle E \rangle(L) \sim \frac{E_d - E_o}{2} + b(T - T_c(L))L^D,$$

and

$$C_v^{\max}(L) \sim \frac{C_o + C_d}{2} + cL^D$$

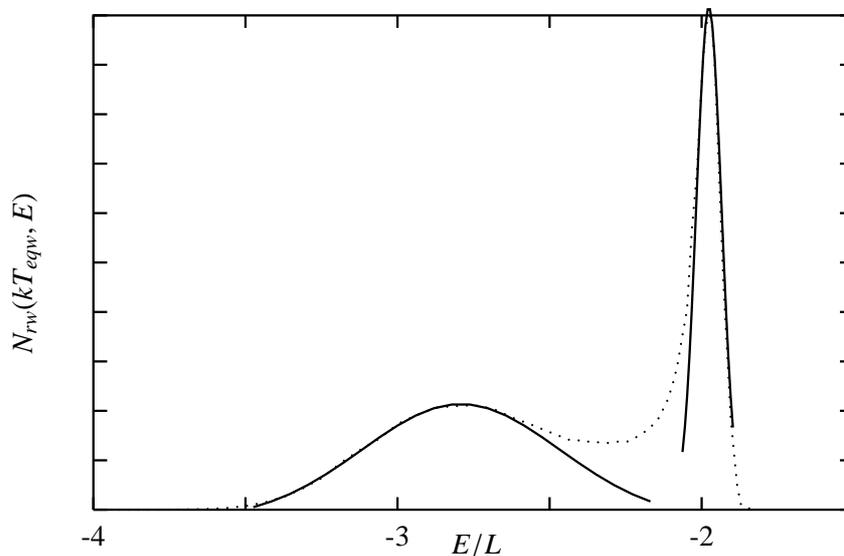


Figure 4.5: Dotted curve: reweighted histogram of the energy for the three-state Potts model with $1/r^{1.2}$ interactions ($L = 1000$), at a temperature corresponding to the equal-weights condition ($kT = 3.9599(2)$). The superimposed solid curves indicate a non-linear least-square fit to the two-gaussian theory.

where the coefficients a , b and c depend, among other things, on the latent heat, so that this quantity may as well be estimated from the finite-size dependence of the three quantities above¹⁴. The last relation shows that, in the limit $L \rightarrow \infty$, one has $C_v/L^D \neq 0$ in contrast with a second-order transition; this criterion was used in [131] as a possible discriminator for the order of the transition.

For the q -state Potts model with nearest-neighbor interactions, Borgs and Kotecký developed a more rigorous theory of these finite-size effects [49], by expressing the partition function of the model (in a periodic box) in terms of contours through a Fortuin-Kastelein random-cluster mapping¹⁵. The derivation is valid only for large q , but the authors suggest that it should actually hold for the whole first-order regime (this was confirmed numerically in [191] for the three-dimensional model). It is not the goal of the present work to dwell upon the technical details involved in the derivation, and I will simply consider the most prominent results: there exist estimators for the (pseudo-)transition temperature which deviate from the thermodynamic limit by terms falling off *exponentially* with the lattice size; in particular, these include (i) the crossing point of the mean energy density vs the temperature, for two distinct lattice sizes, and (ii) the crossing point of the ratios W_o/W_d at distinct sizes, where W_o (W_d) designates the area under the peak of the

¹⁴Another way to compute the latent heat that relies on Borgs and Kotecký theory (see below) was proposed by Janke in [177].

¹⁵See Sec. 2.3.3.

4.2. Estimates of transition temperatures from multicanonical simulations

ordered (disordered) phase, and $W_o/W_d \rightarrow q$ in the thermodynamic limit¹⁶. These two criteria require simulating the lattice at two distinct sizes. On the contrary, the so-called *number-of-phases* criterion, i.e., $W_o/W_d = q$, directly yields an estimate of the transition temperature from a single lattice size, yet its validity is, as it turns out, restricted to the short-range flavor of the q -state Potts model. Interestingly, as far as the long-range model is concerned, it was suggested by Krech and Luijten [213] that an effective number of states q_{eff} may be defined from the asymptotic behavior of the ratio W_o/W_d as $L \rightarrow \infty$, with $q_{eff} \sim 1.67$ reported for $q = 3$ (since, however, q_{eff} is initially unknown, the criterion cannot be used to locate $T_c(\infty)$).

An extension of the two previous theories to long-range interactions is not trivial. As regards the long-range Potts chain in its first-order regime, the recipe suggested thus far [131] is to consider that the volume term should be simply set to L (i.e., with finite-size scaling involving corrections in $1/L$), yet (in the same article) it is also claimed that the interface tension should *also* scale proportional to L . My approach in the following will be to consider two possible schemes: $1/L$ corrections *explicitly* allowing for second-order terms, and $1/L^x$ corrections with an unknown exponent, yet *without* correction terms. As will be discussed in Chap. 7, however, the second definition, along with a σ -dependent exponent, is by far the most reliable (and I will also argue that is is the most physically pertinent definition).

Estimators

In order to estimate the transition temperatures, I relied on (i) free energies, and (ii) moments of the magnetization. The specific heat exhibits a less pronounced peak, and is usually less suited for this purpose. The (variational) free energy of the magnetization, hereunder denoted as $F_m(kT, m)$, is computed from the reweighting equation, Eq. (3.3), which amounts to computing the distribution of the first (microcanonical) moment of the order parameter, and reweighting at temperature T . As regards the free energy associated with the energy distribution, the scheme is even simpler, since this is just (minus) the logarithm of the reweighted histogram of the energy, i.e., $F_e(kT, E) = -\ln N_{rw}(kT, E)$. Here and in the following, $N_{rw}(kT, E)$ designates the reweighted histogram of the energy and is given by

$$N_{rw}(kT, E) = N(E)e^{S(E)-\beta E},$$

where $e^{S(E)}$ is the estimated density of states (i.e. $e^{-S(E)}$ is the multicanonical weight). $N(E)$ denotes the *raw* (expectedly flat) histogram of the energy accumulated during the production run, yet filled with *independent* samples taken every two tunneling times. In practice, this means that the total number of MC sweeps may reach 10^7 , indeed 10^8 in some cases (especially when close to $\sigma = 1.0$). It is important to note that both free energy functions play the same role, i.e., that

¹⁶In [47], it is also suggested to use as a *phase separator* the mean energy of the system at the temperature where the specific heat is maximum.

of a thermodynamic potential, although with respect to a different parameter. As far as estimating transition temperatures is concerned, choosing one or the other is really a matter of attainable precision. F_m turned out to yield more reliable results at large q owing to its higher symmetry (too sharp wells make the estimation of the equal-height condition cumbersome).

Two conditions for the estimation of T_c were used:

- the equal-height condition is defined as the temperature where histogram peaks have the same height, i.e., where minima of $F_e(kT, E)$ (or $F_m(kT, m)$) have the same value;
- the equal-weight condition, introduced by Lee and Kosterlitz in [219]: here, the transition temperature $T_c(L)$ is obtained by imposing that the number of bin entries in $N_{rw}(E)$ be the same below and above the energy corresponding to the maximum of $F_e(kT, E)$ (i.e., the mixed-phase configuration). This is equivalent to the condition that the average energy be the arithmetic mean of the energy of each phase, yet this condition cannot always be used in practice, especially at large q value where the free energy is strongly asymmetric.

Both definition should yield the same result in the infinite-size limit [79]. In [131], it must be noted, only the first condition was invoked.

For the sake of completeness, I also compute transition temperatures by relying on two other estimators: the magnetic susceptibility, and a Binder cumulant of the magnetization. From the fluctuation-dissipation theorem, the susceptibility is readily obtained from the second moment of the magnetization. Binder cumulants of the magnetization are defined as

$$U^{(4)} = 1 - \frac{\langle m^4 \rangle}{3 \langle m^2 \rangle^2}.$$

Owing to its invariance under a renormalization group transform, this cumulant is a very convenient estimator of the critical point, because the associated convergence is really high. In the thermodynamic limit, cumulants cross at a critical fixed point $U_*^{(4)}$ defining the true critical temperature. In practice (i.e. at finite-lattice size), I found the crossing points to drift smoothly over the range of lattice sizes. Following Binder [35], I assumed a power law of the form L^w for the crossing point of $[U^{(4)}(L), U^{(4)}(L')]$, with an unknown exponent w [35].

4.2.3 Results

Results for $q = 3, 5, 7, 9$ are summarized in Table 4.1 and sketched in Fig. 4.6. In the second-order regime (indicated by an asterisk in the table), I computed the critical temperatures from χ and $U^{(4)}$ only, since free energies exhibit no double-peak structure here (see Sec. 4.3 for details on how the order of the transition was ascertained). As expected according to finite-size scaling theory, both definitions

4.2. Estimates of transition temperatures from multicanonical simulations

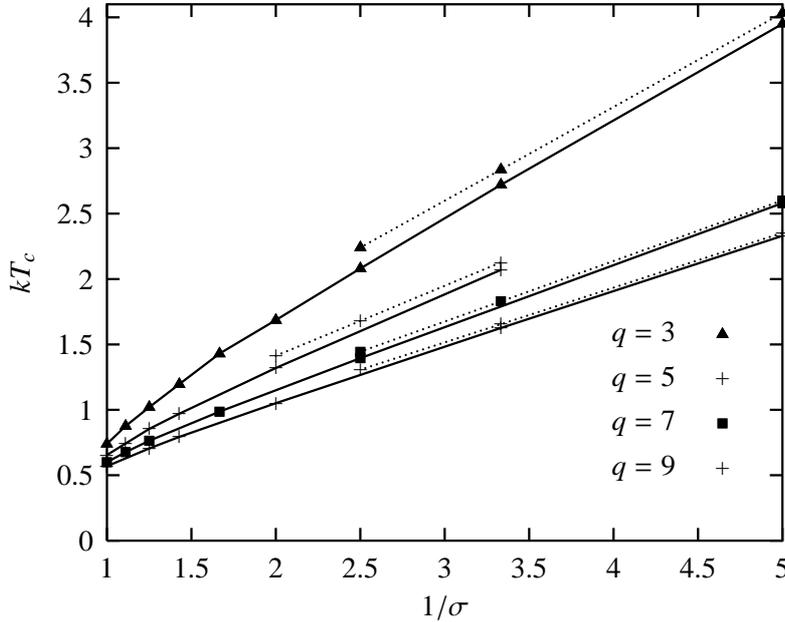


Figure 4.6: Critical couplings for $q = 3, 5, 7, 9$, from top to bottom (solid lines). Mean-field predictions are shown for comparison (dotted lines).

of the transition temperature, i.e., using equal peak weights vs equal peak heights, lead within error bars to the same estimates at infinite lattice size. Other quantities $T_c(\chi)$ and $T_c(U^{(4)})$ yield very similar results, with a discrepancy never exceeding 1%.

Comparison with exact mean-field predictions

For all values of q , transition temperatures progressively depart from the mean-field line as σ is increased. For $q = 5$, for instance, the ratio between $T_c(\chi)$ and the mean-field value ranges from 97.3% at $\sigma = 0.3$ to 83.9% at $\sigma = 0.8$. For a given range of interaction, the adequacy of mean-field results is also markedly improved at large q , as in the nearest-neighbor model; this is consistent with the large- q expansion of Pearce and Griffiths [276] showing that mean-field theory becomes exact in this limit. In anticipation of the next section, and as illustrated in Fig. 4.7 for $q = 9, \sigma = 0.3$, and $L = 400$, this agreement also holds, even at finite lattice sizes, for the shape of the free energy $F_m(kT, m)$ and the position of metastability plateaus related to spinodal points.

Comparison with other methods

For $q = 3$ and $q = 5$, several results have been made available from previous Monte Carlo simulations. Results obtained in [131] using either Luijten-Blöte's cluster algorithm ($q = 3$) or a standard Metropolis algorithm ($q = 5$) are in fairly good

q	σ	T_c (MF)	$T_c(\chi)$	$T_c(U^{(4)})$	$T_c(\text{eqh})$	$T_c(\text{eqw})$	T_c [131]	T_c [251]	T_c [130]
3	0.2	4.034	3.97(1)	3.98(1)	3.94(1)	3.97(1)	3.70 ^a		3.7023
	0.3	2.836	2.72(1)	2.72(1)	2.71(1)	2.71(1)	2.70 ^a	2.71669	2.5893
	0.4	2.240	2.086(4)	2.089(6)	2.075(5)	2.074(4)	2.08 ^a		2.0247
	0.5	1.884	1.691(3)	1.685(3)	1.686(4)	1.684(2)	1.70 ^a	1.68542	1.6631
	0.6	1.649	1.44(1)	1.43(1)	1.43(1)	1.43(1)	1.41 ^a		1.4000
	0.7	1.482	1.196(3)	1.19(1)	1.18(1)		1.19 ^b	1.1968	1.1942
	0.8*	1.358	1.019(4)	1.03(1)			1.01 ^b		1.0231
	0.9*	1.262	0.876	0.875			0.88 ^b	0.8785	0.874
5	0.3	2.127	2.07(1)	2.07(1)	2.072(6)	2.070(4)	2.033 ^a	2.06900	1.736
	0.5	1.413	1.321(3)	1.319(4)	1.319(3)	1.319(2)	1.297 ^a	1.31638	1.245
	0.7	1.111	0.973(1)	0.973(2)	0.970(3)	0.970(2)	0.981 ^a	0.96963	0.956
	0.8	1.018	0.854(1)	0.853(1)	0.857(1)	0.857(1)			0.844
	0.9*	0.947	0.743(2)	0.739(4)				0.74673	0.745
7	0.2	2.600	2.58(1)	2.58(2)	2.578(2)	2.577(1)			
	0.4	1.444	1.395(5)	1.394(4)	1.394(1)	1.393(1)			
	0.6	1.063	0.986(2)	0.985(3)	0.984(1)	0.986(1)			
	0.8	0.875	0.764(1)	0.763(1)	0.764(1)	0.764(1)			
	0.9	0.814	0.677(1)	0.676(1)					
9	0.2	2.353	2.33(1)	2.33(1)	2.33(1)	2.32(1)			
	0.3	1.655	1.626(3)	1.625(4)	1.627(3)	1.626(1)			
	0.5	1.099	1.052(2)	1.051(2)	1.050(3)	1.052(1)			
	0.7	0.864	0.793(2)	0.792(2)	0.794(2)	0.794(1)			
	0.8	0.792	0.705(2)	0.704(1)	0.704(1)	0.704(1)			

Table 4.1: Estimates of the critical temperature in the first- and second-order regimes (the latter is indicated by an asterisk): MF, mean-field predictions; χ , using peaks of the susceptibility; $U^{(4)}$ using crossing points of Binder cumulants of the magnetization; eqw, eqh, using the free energy, where T_c corresponds to equal peak weights and heights, respectively; Ref. [131], Monte Carlo study based on multihistogramming and the Luijten-Blöte cluster algorithm ($q = 3$) and a standard metropolis algorithm ($q = 5$), where (a) refers to $1/K_e(\Delta F)$, and (b) to $1/K_e(U^{(4)})$; Ref. [251]), cluster mean-field method combined with an extrapolation technique based on the VBS (Vanden Broeck and Schwartz) algorithm; Ref. [130]), transfer matrix method combined with finite-range scaling.

4.3. Spinodals: a novel approach to assess the order of phase transitions

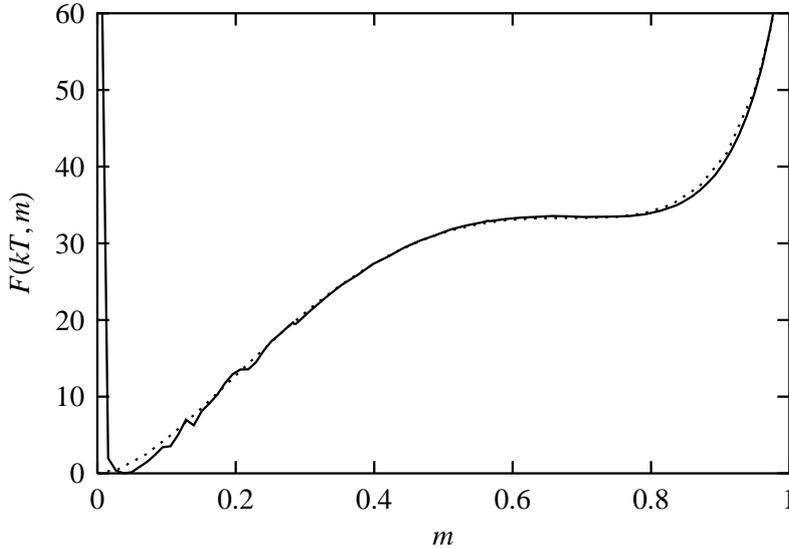


Figure 4.7: Free energy $F_m(kT, m)$ for $q = 9, \sigma = 0.3, L = 400$ (solid line), along with the mean-field prediction (dashed line) as given by Eq. (4.4).

agreement with multicanonical ones. The discrepancy does not exceed 1%, except in the case $\sigma = 0.2$, where multicanonical estimates lie much closer to the mean-field prediction. This might certainly be (partially) attributed to the use of different periodic boundary conditions, since I found Infinite Image Periodic Boundary Conditions to *drastically* speed up the convergence to the thermodynamic limit, as against the first image conditions used in [131, 19].

I further compare these estimates with those obtained by Monroe with the cluster mean-field method [251], and by Glumac and Uzelac using a transfer matrix approach [130]. As illustrated in Table 4.1, results obtained using the cluster mean-field approach combined with the Vanden-Broeck-Schwartz extrapolation algorithm (Sec. 4.1.2) yield a very satisfying match, with a deviation as low as 0.1% on average over the whole range of σ values. The discrepancy with estimates obtained using the transfer matrix method is slightly higher and amounts to 2% on average, except for low values of σ where the agreement of my results with the mean-field prediction is, here again, far better. Since implementing Infinite Image Periodic Boundary Conditions in the finite-range scaling approach is simply intractable, this undoubtedly accounts for the discrepancy.

4.3 Spinodals: a novel approach to assess the order of phase transitions

As already stated in the introduction of this chapter, a precise determination of the tricritical value $\sigma_c(q)$ is a real challenge, due to the weakening of the first-order transition as σ_c is approached from below. This makes traditional indicators, e.g.,

latent heats or energy jumps¹⁷, fairly inefficient, since observing clear jumps in the vicinity of the tricritical value entails simulating huge lattices.

Glumac and Uzelac in [131] used three less traditional estimators: the interface free energy, the specific heat, and the reduced fourth-order Binder cumulant, which all turned out to be less sensitive to this weakening: in particular, the last quantity defined as $U_L = \langle E^4 \rangle / \langle E^2 \rangle$ is expected to reach a nontrivial constant $U_\infty \neq 1$ as $L \rightarrow \infty$ at a first-order transition only [79]; by extrapolating to the thermodynamic limit from measures taken at different lattice sizes, they found σ_c to fall between 0.6 and 0.7 for the three-state model. Still and all, this approach imposes simulating fairly large lattices (around $L = 3000$) for the extrapolation procedure to be reliable, let alone the fact that Binder cumulants may experience uncontrollable crossover effects [219].

Another crucial point, which will be developed in Sec. 4.3.5, is the fact that all these indicators do not allow the location of the tricritical line to be determined accurately, because they tend *smoothly* to their second-order value as $\sigma \rightarrow \sigma_c(q)$. As will be seen in the remainder of this chapter, this is not the case of the method developed hereafter.

4.3.1 Outline of the method

The present approach is based on the location of spinodal points, which may be accurately determined already for medium lattice sizes. In marked contrast to multihistogram techniques, the multicanonical method indeed allows one to obtain free energy functions (or, equally, reweighted histograms of the energy) over a range of temperature which extends fairly far away from the transition temperature, with remarkably modest numerical effort.

The basis of this method relies on the fact that the temperature difference between both spinodal points will tend to zero as σ_c is approached, since there are no metastable states in the case of continuous transitions. Stated differently, the conditions under which metastability occurs, i.e., both first and second derivatives of the free energy are zero, are met only at the critical point for a continuous transition: hence metastable states merge into a single large minimum as the first-order transition turns into a second-order one. Such behavior has indeed been widely observed, e.g., for liquid-vapor transitions near the critical point, and is borne out by the mean-field calculation of Sec. 4.1.1.

For a given set of (q, σ) parameters, the location of spinodal points is determined by first computing the (variational) free energy function of the order parameter [$F_m(kT, m)$, see Eq. 3.3] over a large temperature range. Alternatively, one can make use of a similar function of the energy, i.e., $F_e(kT, E) = -\ln N_{rw}(kT, E)$, where $N_{rw}(kT, E)$ designates the reweighted histogram of the energy. While the latter function plays the same role as the free energy of the magnetization, it turns out to yield a higher precision at low q , as we will witness in a moment. The

¹⁷Energy jumps tend to zero as $\sigma \rightarrow \sigma_c^-$.

4.3. Spinodals: a novel approach to assess the order of phase transitions

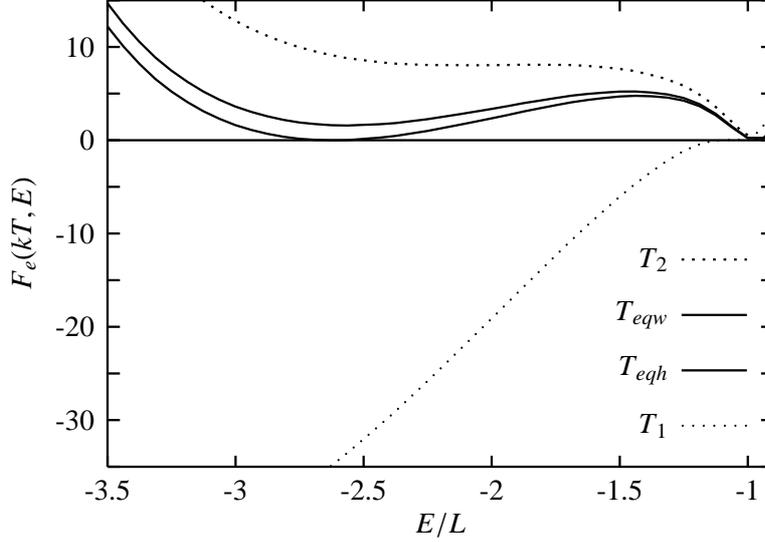


Figure 4.8: Graphs of $F_e(kT, E) = -\ln N_{rw}(kT, E)$ for $q = 5, \sigma = 0.3, N = 400$, at four characteristic temperatures: T_1, T_2, T_{eqh} , and $T_{eqw} = T_c$ denote the temperatures of the two metastable states, and the temperature of equal peak heights, and that of equal peak weights, respectively. E/L stands for the energy per spin.

limit of metastability at finite lattice size is then defined by the joint condition $dF_e/dE = d^2F_e/dE^2 = 0$, or alternatively $dF/dm = d^2F/dm^2 = 0$: for a first-order transition, this condition is met at two temperatures T_1 and T_2 which satisfy the inequality $T_1 < T_c < T_2$, where T_c denotes the transition temperature.

4.3.2 Obtaining reliable information from (variational) free energies

The previous free energy functions usually have rather rugged graphs, owing to statistical fluctuations that occur on the corresponding histograms, and before any reliable estimation can be carried out, rapid oscillations must be filtered out by means of a linear smoothing filter. The order of this filter is computed so that we end up with at most three extrema over the whole temperature range of interest. By continuously varying kT within this range, one determines the temperature of each metastable state by monitoring the change in the number of minima (see Fig. 4.8). An alternative approach that builds on the transition matrix method [341] is introduced in the next section.

Graphs of the free energy $F_e(kT, E)$ in Fig. 4.8 illustrate that the peak and the plateau corresponding to the disordered phase are much narrower than those of the ordered phase. As a result, the precision in the determination of the temperature T_1 of the lowest metastable state is substantially lower than that of the upper metastable state (T_2). This asymmetry increases with increasing q (irrespective of the power-law decay parameter), and in effect precludes the use of reweighted histograms for the estimation of spinodal points at $q > 7$.

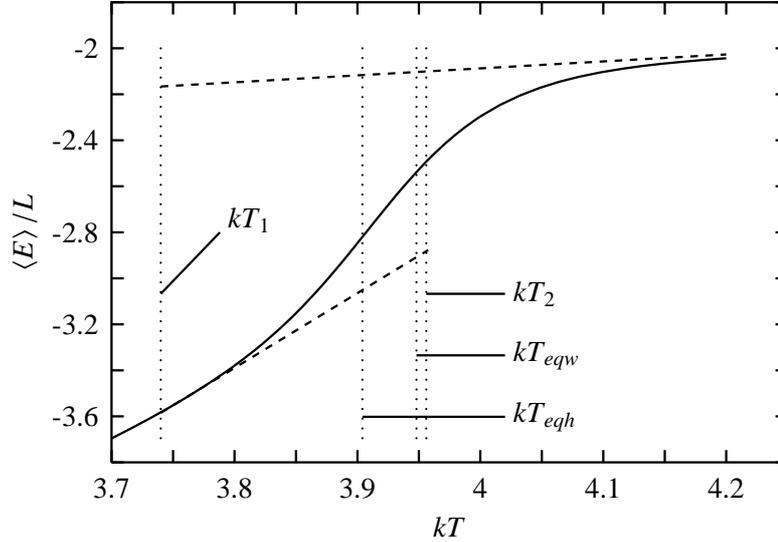


Figure 4.9: Average energy per spin for $q = 3, \sigma = 0.2, L = 400$, computed over both phases (solid line), ordered phase only (lower dashed line), and disordered phase only (upper dashed line). Vertical dotted lines indicate the four characteristic temperatures: from left to right, lower limit of metastability (kT_1), transition temperatures (equal heights kT_{eqh} , then equal weights kT_{eqw}), and upper limit of metastability (kT_2).

For $q = 9$, one has thus to rely on the extrema of the “other” free energy, $F_m(kT, m)$, since this function then becomes nearly symmetric and displays two peaks that are well separated. Incidentally, the asymmetric shape of $F_e(kT, E)$ can be accounted for by the fact that specific heats have a different magnitude in the disordered and ordered phases, since this thermodynamic quantity is simply proportional to the standard deviation of the associated gaussian peak [79]. This may be readily confirmed by reweighting thermodynamic averages over a single subphase at a time, once the maximum of $F_e(kT, E)$ which separates the two phases has been located. Figure 4.9 illustrates how this procedure was applied to the computation of the mean energy per spin of each subphase for $q = 3, \sigma = 0.2$, and $L = 400$ spins. A simple visual inspection allows one to assess a much lower specific heat for the disordered phase than for the ordered phase.

4.3.3 Finite size scaling for spinodal temperatures

At finite lattice size, transition temperatures are known to experience a shift proportional to the distance to the thermodynamic limit. As can be easily suspected, the same feature applies to spinodal temperatures as well. Assuming that the finite-size scaling theory developed in [79] for first-order transitions is also valid (i) in the case of long-range Hamiltonians, and (ii) not only for transition temperatures, but also for spinodal temperatures, I therefore compute temperatures at infinite lattice

4.3. Spinodals: a novel approach to assess the order of phase transitions

size by assuming power-law corrections in $1/L$. I take it for granted that temperatures defining the limit of metastability do indeed obey the same scaling behavior, although the phenomenological theory proposed in [79] does not explicitly handle them. The inclusion of a second-order term proves necessary in order to obtain satisfying fits, due to the presence of small lattice sizes in the set of data. Yet, interestingly enough, fitting finite-size temperatures to a power law of the form $T(L) = T(\infty) + aL^b$ yields very similar extrapolated values, with discrepancies smaller than 0.1%, i.e., within the range of uncertainty. In addition, $F_e(kT, E)$ and $F_m(kT, m)$ lead to distinct finite-size shifts, with the latter function easily allowing one to drop second-order corrections without much affecting the final result, yet I did not find any convincing explanation for this feature.

4.3.4 A digression: glean information from transition matrices

Anticipating the introduction of a new multicanonical algorithm in Chap. 6, I would like to briefly discuss at this point a possible combination of the spinodal method presented above with the transition matrix approach¹⁸. As will be witnessed in Chap. 6, the algorithm needs a vital ingredient to work: the *microcanonical temperature* $\beta(E)$. The transition matrix method [339] will prove an efficient way of obtaining a reliable estimate of $\beta(E)$ (although at the expense of increased computation cost). I show hereafter that the microcanonical temperature, once determined by means of the transition matrix method, can also readily yield estimates of spinodal temperatures.

The point of departure is the so-called Broad Histogram Equation [339, 98] which reads

$$\frac{T_\infty(E \rightarrow E')}{T_\infty(E' \rightarrow E)} = \frac{n(E')}{n(E)}, \quad (4.8)$$

where $n(E) = e^{S(E)}$ is the density of states, and $T_\infty(E \rightarrow E')$ is the transition matrix element between energy levels E and E' . The last quantity is estimated by accumulating a double-entry histogram $h(E, \Delta E)$ containing the number of potential moves¹⁹ from E to $E + \Delta E$ each time the energy level E is visited. In terms of the microcanonical entropy $S(E)$, the previous equation writes

$$S(E') - S(E) = \ln \frac{T_\infty(E \rightarrow E')}{T_\infty(E' \rightarrow E)}. \quad (4.9)$$

I will now denote as $\gamma(E, E')$ the *numerical estimate* of the ratio $T_\infty(E \rightarrow E')/T_\infty(E' \rightarrow E)$. Writing $\epsilon = E' - E$, and performing an expansion on both sides to second order in ϵ , and then identifying order by order in ϵ , we thus have

$$\beta(E) = \frac{dS(E)}{dE} \hat{=} \frac{d \ln \gamma(E)}{dE} \quad (4.10)$$

¹⁸The method was described in Sec. 2.5.5.

¹⁹Actually, the microreversibility hypothesis that underlies the Broad Histogram equation imposes that only single-spin updates be used for the estimation of the number of potential moves.

and

$$\frac{d^2 S(E)}{dE^2} \hat{=} \frac{d^2 \ln \gamma(E)}{dE^2} \quad (4.11)$$

Alternatively, one may fetch the two previous quantities by performing a polynomial fit of $\ln \gamma(E, E')$ with respect to $\epsilon = E' - E$, i.e.

$$\ln \gamma(E, E') = \beta(E)\epsilon + \frac{1}{2} \frac{d\beta(E)}{dE} \epsilon^2 + O(\epsilon^3)$$

For long-range potentials, this is all the more appropriate that the discretization of the energy axis can usually be chosen very small, so that the polynomial fit can be carried out over many points.

Now, the main thrust of the argument somewhat takes after Maxwell's construction: the free energy $F_e(kT, E)$, which defines the locus of spinodal points, is obtained through the reweighting equation $F_e(kT, E) = E - kTS(E)$ ²⁰(Sec. 3.4). At a given temperature kT , equilibrium values for the energy are thus obtained from $dF_e(kT, E)/dE = 0$ and $d^2 F_e(kT, E)/dE^2 > 0$, i.e.²¹,

$$\beta(E) = 1/kT \text{ and } \frac{d\beta(E)}{dE} < 0.$$

While inflexion points are given by $d^2 F_e(kT, E)/dE^2 = 0$; whence spinodal points follow from

$$\frac{d\beta(E)}{dE} = 0 \text{ and } kT = 1/\beta(E)$$

An example of such calculation is depicted in Fig. 4.10 for the following model parameters: $q = 6, \sigma = 0.9, L = 256$ (i.e., these correspond to a weakly first-order transition). The following temperature were obtained: $kT_1 = 0.72525(5)$ and $kT_2 = 0.72224(4)$. The error is approximately two orders of magnitude smaller than when using a reweighted histogram, as illustrated in the figure.

To wind up, spinodal points are obtained directly from numerical estimates of transition matrices, and it is no longer necessary to resort to energy histograms, nor to a production run, because $\beta(E)$ is estimated *directly* in the course of the iteration procedure. It can be argued that $\beta(E)$ and its derivative might as well be obtained from the estimated density of states, by relying on a discrete differentiation scheme. This is not always practicable, however, because (i) in the multicanonical approach, the need to perform a production run stems from the fact that the estimated density of states *may not* yield a perfectly flat histogram, and the production run in a sense *corrects* for this, and (ii) a discrete differentiation scheme produces a lot of noise (although this may be compensated for by applying a low-pass filter, or by relying to a spline interpolation). Additionally, transition matrices are *by construction* more reliable in the previous respect, inasmuch as they contain much more information than the density of states (for one thing because *every* potential move

²⁰ $k = 1$ is implicit in the definition of $S(E)$ invoked here.

²¹In fact, this conditions along with the negative curvature condition yields the most probable value, which coincides with the mean value in the thermodynamic limit only.

4.3. Spinodals: a novel approach to assess the order of phase transitions

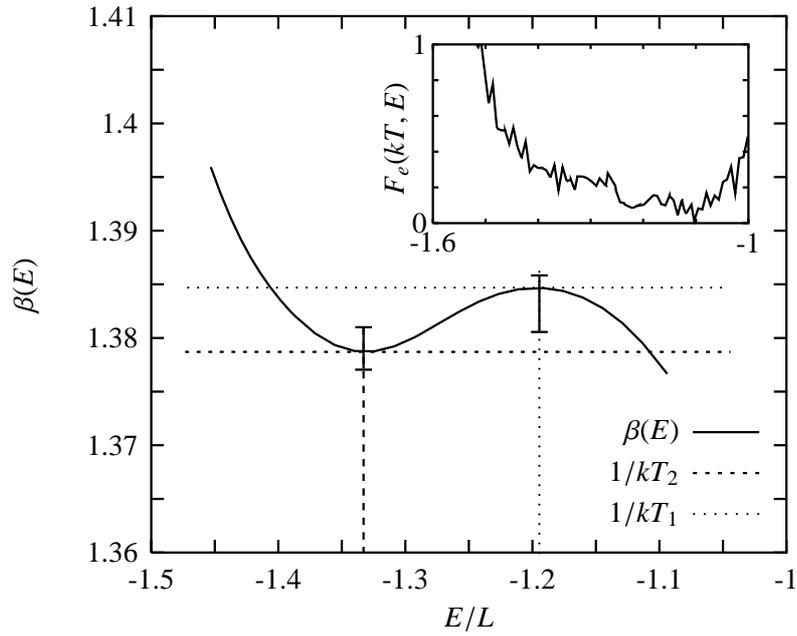


Figure 4.10: An example of calculation of spinodal points from the information provided by transition matrices. Model parameters are $q = 6, \sigma = 0.9, L = 256$, and the simulation was performed with the breathing cluster method presented in Chap. 6. Dotted and dashed lines refer to the first and the second spinodal points, respectively (the error on the spinodal temperatures lies around 5×10^{-5} and is therefore not indicated on the graph). The inset shows the free energy $F_e(kT, E)$ at the second spinodal temperature kT_2 , before filtering was applied, and the large vertical error bars indicate the error on the spinodal temperatures obtained from $F_e(kT, E)$.

brings its own piece of information, not just accepted moves), and these information are effectively utilized in the estimation of $\beta(E)$ ²²; this was already realized by the authors of the method themselves, and some variants of the multicanonical method incidentally capitalized on this, see, for instance, [299, 358].

4.3.5 Application to the phase diagram of the long-range Potts chain

The present section is devoted to the application of the spinodal method to the investigation of the phase diagram of the long-range Potts chain. Here, I concentrate on the location of the so-called “tricritical” line $\sigma_c(q)$ where the transition changes from first to second order. The outline is as follows: first, spinodal temperatures T_1 and T_2 are determined from finite-size scaling for several σ values in the first-order regime; then it is demonstrated that, by relying on the ratio T_2/T_c , where T_2 is the upper spinodal temperature and T_c the transition temperature, one can devise an efficient method to determine the location of the border line with markedly higher precision than with other indicators.

Spinodals As can be viewed in Fig. 4.11, spinodal points merge slightly above $\sigma \sim 0.8$ for $q = 5$, and according to the spinodal method developed above, this indeed signals a change of the nature of the transition. A plot of $dkT_m = kT_2 - kT_1$ against $1/\sigma$ indicates that, for all values of q , the points fit quite well on a straight line for low enough σ , and the slope of this line tends toward that of the mean-field curve. The $q = 7$ case is sketched in Fig. 4.12, where it is clear that the point at $\sigma = 0.6$ ($1/\sigma \sim 1.67$) marks the border between the linear and nonlinear behavior, illustrating the weakening of the first-order transition as σ_c is approached.

Since both transition and spinodal temperatures appear to scale as $1/\sigma$ in the vicinity of the mean-field regime, it is more appropriate to work with the ratios T_1/T_c and T_2/T_c , for the scaling factors will then cancel out neatly, *except* when approaching $\sigma_c(q)$. As previously mentioned, the second ratio, which is sketched in Fig. 4.13, offers a higher precision owing to a larger free energy plateau in the ordered phase. As σ falls off to the mean-field regime, this ratio tends, within error bars, to the value predicted by mean-field theory, i.e., $T_2/T_c = 1.01, 1.037, 1.059, 1.077$ for $q = 3, 5, 7, 9$, respectively.

On the leftmost side of the graph in the figure, a sudden sharp decrease of T_2/T_c can be witnessed as $\sigma \rightarrow \sigma_c$. This suggests a quite interesting way of determining $\sigma_c(q)$ without much ambiguity, as opposed to, e.g., methods using the interfacial free energy or Binder cumulants. In view of the shape of the graph near σ_c , the most reliable method is a fit of data points to a polynomial. The lowest error was reached with a polynomial of degree 2 for $q = 5, 7, 9$, and of degree 3 for $q = 3$.

²²There is one pitfall, however, because it is practically impossible to estimate the confidence interval on $\beta(E)$ other than by performing multiple independent simulations. This is not the case when one has performed a production run, where standard jackknife methods can be used.

4.3. Spinodals: a novel approach to assess the order of phase transitions

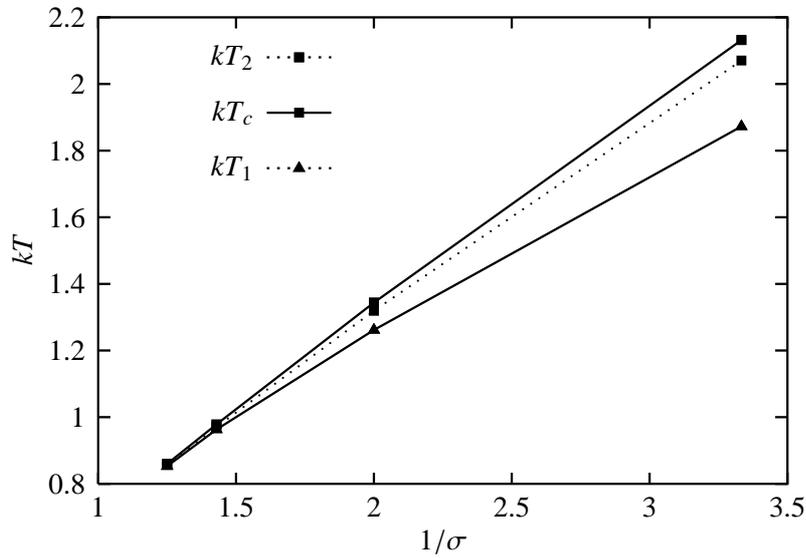


Figure 4.11: Spinodal curve for $0.3 \leq \sigma \leq 0.8$ ($q = 5$). The limits of metastability T_1 and T_2 (i.e., defining spinodal temperatures) are indicated as triangles and diamonds, respectively. The transition temperature T_c is reminded as dotted line. Errors are smaller than the size of symbols.

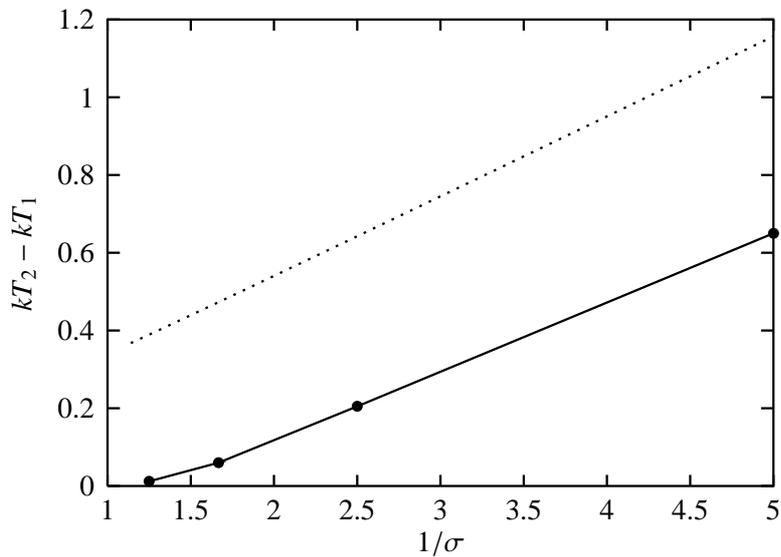


Figure 4.12: Difference between temperatures of metastability $dkT_m = kT_2 - kT_1$ vs $1/\sigma$ for $q = 7$ (circles connected by solid lines). Errors are smaller than the size of symbols. Mean-field prediction are shown for comparison (dashed line).

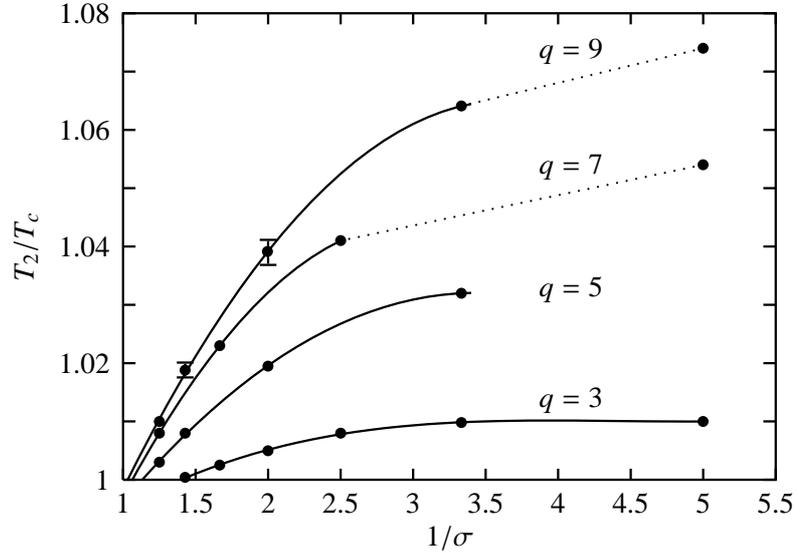


Figure 4.13: T_2/T_c vs $1/\sigma$ for $q = 3, 5, 7, 9$. Solid lines indicate polynomial fits. Dotted lines are guides to the eyes. Error bars are smaller than the size of dots, except where explicitly indicated.

The following numerical estimates are obtained:

q	σ_c
3	0.72(1)
5	0.88(2)
7	0.94(2)
9	0.965(20)

The graph of $\sigma_c(q)$ is sketched in Fig. 4.14, along with two other estimations obtained from Monte Carlo simulations and the Luijten-Blöte cluster algorithm:

- $\sigma \in [0.6, 0.7]$ for $q = 3$ and $\sigma > 0.8$ for $q = 5$ [131]; the authors resorted to three distinct indicators, i.e., Binder cumulants of the energy, peaks of the specific heat, and surface tensions (Lee's criterion [219]);
- $\sigma \in [0.7, 0.8]$ for $q = 3$ [213] (it was suggested by the authors, however, that $\sigma = 0.75$ may already belong to the second-order regime); this estimation is based on the standard double-peak structure of the histogram, yet the authors investigated sizes up to $L = 2^{19}$ spins.

Discussion The crucial point that ensures the superiority of the above method as against other approaches is that the quantities $T_2 - T_1$ and T_2/T_c do not reach 0 in the same fashion. This is *apparently* so because the first-order coefficient in the expansion of $T_2/T_c - 1$ in terms of $\epsilon = \sigma - \sigma_c$ has a large value, which is not the

4.3. Spinodals: a novel approach to assess the order of phase transitions

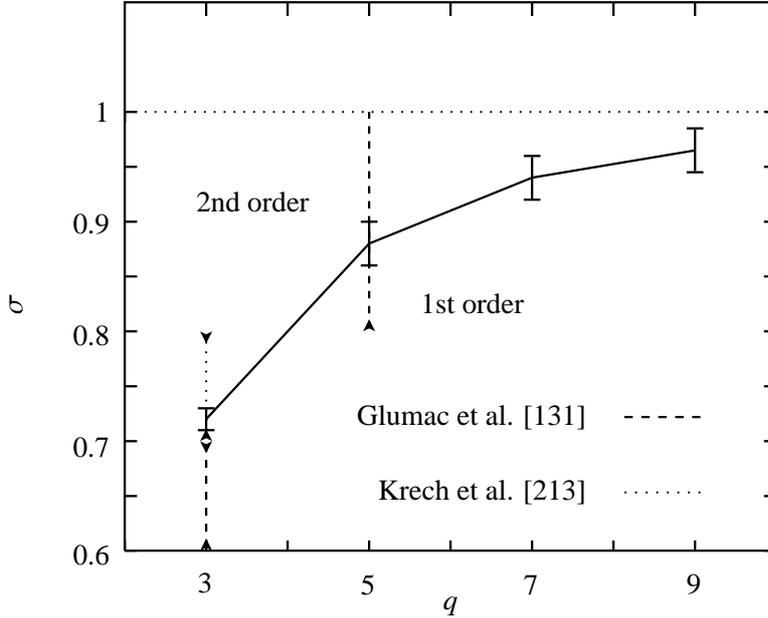


Figure 4.14: Phase diagram computed using finite-size scaling properties of spinodal points, for $\sigma < 1.0$. Dotted lines are guides to the eyes.

case for $T_2 - T_1$. This means that the error on the determination of the value of σ for which $T_2/T_c \rightarrow 1$, is smaller²³ than that associated with the condition that $T_2 - T_1 \rightarrow 0$.

I deem it important to stress, at this point, that the previous discussion does not *outlaw* a null first-order coefficient: in the very vicinity of σ_c , the graph of T_2/T_1 may very well be smoother than expected, but this cannot be ascertained from available numerical data; if this is the case, however, then (i) higher order terms might rapidly compensate for this behavior, so that the graph looks nearly linear at a coarser grain, and (ii) overall, the error on the estimation of σ_c is probably very low, leading to an underestimation of $\sigma_c(q)$ (as will be discussed in Sec. 4.3.7, there is compelling evidence indeed that $\sigma_c(q)$ is bounded by 1.0).

As a side note, this suggests that, from an analytical point of view, it would be quite interesting indeed to derive such an expansion, and to check for the value of the first order coefficient. Whether the real-space renormalization group approach of Cannas and Magalhães [64] would be suited for this purpose, *once* dilution has been included in the model²⁴, is an appealing question indeed.

Comparison with other indicators Noteworthy enough, Lee's criterion [219] suffers from the same shortcoming as does the quantity $T_2 - T_1$. In [131], this criterion was used (along with the Binder cumulant of the energy and the specific heat)

²³Or equally, T_1/T_c , even if the ratio T_2/T_c is endowed with a lower uncertainty.

²⁴See Sec. 4.1.3 for a discussion on this issue

to discuss the change of regime for the three-state model. While the behavior of the reduced surface tension $\Delta F/L$ as a function of σ and L *qualitatively* demonstrates the weakening of the transition as σ is increased, more quantitative information can hardly be drawn from the graph of $\Delta F/L$ vs σ . For the sake of completeness, I performed the calculation for $q = 9$ and $\sigma = 1.0$: the value of $\Delta F_m/L$ extrapolated in the thermodynamic limit is depicted in Fig. 4.18; it is blatant that, although $\Delta F_m/L$ tends smoothly to zero as $\sigma \rightarrow 1$, thus signaling a change of regime in this region, the shape of the curve as $\sigma \rightarrow 1$ does not allow one to draw convincing conclusion regarding the location of σ_c .

The specific heat does not represent a better estimator for that matter (this leads to a lower bound of 0.65 for $\sigma_c(3)$). Incidentally, it is by relying on Binder cumulants of the energy extrapolated to the thermodynamic limit that the authors in [131] were able to suggest a change of regime in the interval [0.6, 0.7], and yet the graph of the extrapolated value (Fig. 3 in the article) is extremely smooth as well in the interval [0.65, 0.8] so that a more precise estimation seems in fact out of reach (let alone the strong crossover effects that Binder cumulants may experience [34]; I shall further examine this point in Sec. 7.6.

4.3.6 Asymptotic $q \rightarrow \infty$ behavior and the $1/r^2$ line

In view of the phase diagram just obtained (Fig. 4.14), one may readily ponder on the asymptotic behavior of the $\sigma_c(q)$ line in the limit $q \rightarrow \infty$. Indeed, this behavior lies at the core of a controversy surrounding the nature of the phase transition on the line of inverse square interactions. Before reexamining the controversy in the light of the previous results, I will therefore provide a blow-by-blow account of the current state of the battleground.

Topological phase transition in models with inverse square interactions For one-dimensional models, the line of inverse square interaction is special, since it was rigorously proven that no phase transition at finite temperature can occur for $\sigma > 1$ [55, 107, 124, 290, 316]. This marginal case, as it turns out, may not merely be an “upper critical range” of interaction: it has been claimed [11, 33, 72, 208] that it is also the locus of a topological phase transition, similar in several respects to the Kosterlitz-Thouless transition governed by topological defects that XY models exhibit in two dimensions [207]. This class of transitions of “infinite order” also covers phase transitions in superfluid helium and superconductors, and melting transitions in two dimensions [257].

Indeed, it has been found that the correlation function displays a power law decay with an exponent depending on the temperature²⁵ [33, 171], leading to a line of critical points below a critical temperature T_c . At T_c , however, correlations die off logarithmically with the distance (specifically, as $1/\ln r$, [33, 235]), yet for T

²⁵Though with an exponent $1 + 4\sqrt{(T_c - T)/T}$, so that correlations drop off more rapidly with decreasing temperature than in the XY model.

4.3. Spinodals: a novel approach to assess the order of phase transitions

above T_c , it was established from renormalization group arguments that the correlation length follows an essential singular behavior again like the XY model, i.e., of the form e^{bt^α} where t is the reduced deviation from T_c (recalling that $\alpha = -1/2$ for the XY model). Kosterlitz also reported a scaling of the form $\exp[(n-1)/2\pi^2 T]$ for the correlation length of the O_n model with $n > 1$ [208]. Noteworthy enough, $T_c = 0$ in the latter case.

Early Monte Carlo simulation [33] suggested that the specific heat increases above T_c with a finite-size scaling behavior that rules out a second-order transition, although it was pointed out by the authors that the saturation — which is characteristic of the XY model — was not reached for the sizes investigated in their numerical study; a recent investigation based on much larger lattice sizes confirmed the saturation at a temperature slightly above the critical temperature [235].

Cardy [72] extended the investigation to q -state Potts models by means of a kink-gas representation. Here, kinks represent elementary excitations of the system (akin to domain walls in nearest-neighbor models), that is to say, with spins lined-up in the same direction over large domains separated by walls, and their density is adjustable through a chemical potential (which emerges from a re-expression of the Hamiltonian in terms of kinks). It was established that these excitations interact logarithmically with the distance, and in this respect they can indeed be viewed as “topological defects” resembling (anti)vortex in XY models, since vortex/anti-vortex pairs interact in the same fashion (the same behavior is found in the two-dimensional Coulomb gas, where the electrostatic potential decays logarithmically [143, 144]). Renormalization of the kink-gas Hamiltonian yields an essential divergence for the correlation length, $\ln \xi \sim bt^{-\tilde{\nu}}$ as T_c is approached from above, with an exponent $\tilde{\nu}$ depending on the number of states of the model.

Recently, an extensive Monte Carlo investigation based on the Luijten-Blöte cluster algorithm was carried out for the Ising and three-state Potts chains [235]. The authors focus on the scaled order parameter $\langle m^2 \rangle / kT$, and show that this quantity undergoes a universal jump at the transition temperature T_c (together with a superposed “singularity” which vanishes at T_c and is governed by the same exponents as the correlation length). Lower bounds for this jump were given on rigorous grounds in [5], but its universal character for systems with inverse square interactions was first suggested in [235]. From the finite-size scaling behavior of the scaled order parameter, and by performing simulations up to 4×10^5 spins, the authors were able to estimate the magnitude of the critical jump to lie at 0.49(2), the transition temperature around 0.764 (in units of the Potts model) and the leading critical exponent of the superposed singularity near $\tilde{\nu} = 0.54$, the latter being in very good agreement with the value of $1/2$ found by Cardy (i.e. $2/(q+2)$ with $q = 2$ for the Ising chain) [72]. Similar agreement is observed for the three-state Potts chain, with for instance $T_c = 0.7089(2)$. As a by-product, this study thus rules out, at least for $q = 2$ and $q = 3$, the invariance of the critical temperature with respect to the number of states q , as suggested from real-space renormalization [64] (Sec. 4.1.3). These results were backed up by another numerical study of the same model based on the largest cluster distribution [325] (where also the

Luijten-Blöte cluster algorithm was used, yet in its Swendsen-Wang flavor). The authors propose to determine the critical exponent of the superposed singularity from the quantity $L(\partial T/\partial L)_{\bar{\chi}} \sim -(T - T_c)^{1-\bar{\nu}}$ [289], where $\bar{\chi}$ denotes a reduced susceptibility χ/L^σ , and χ is computed through an improved estimator²⁶ based on the largest cluster distribution.

So, where does the tricritical line *really* terminate? Overall, the previous findings (mostly based on renormalization techniques, yet also including a handful of Monte Carlo simulations) give a picture of models with inverse square interactions that seems to mimic the vertex-unbinding transition of the XY model. However, several numerical works contradict this picture to some extent. Both the transfer matrix method of Glumac and Uzelac [130] and the real-space renormalization approach of Cannas and Magalhães [64] failed to reproduce the essential singularity at $\sigma = 1.0$. Incidentally, the issue was not addressed in the work of Monroe [251], although the coherent anomaly method, being essentially a *mean-field* method, is certainly not the best method in this respect.

Most importantly, a recent work by Bayong, Diep and Dotsenko [19] on the q -state Potts model cast real doubt upon the issue, by arguing that the line $\sigma_c(q)$ crosses the line of inverse square interactions at $q \sim 8$, i.e., at $\sigma = 1$, the transition is of the first order for $q > 8$. The method used by the authors is basically a single-histogram Monte Carlo technique [113] with a truncated long-range potential, i.e., periodic boundary conditions are limited to the first image only (see Sec. 4.2.1 for an extensive discussion of this topic). The first-order nature of the transition was asserted by the authors on the grounds that the histogram of the energy displayed a clearly visible double-peak structure above $q = 8$. From a purely numerical viewpoint, the result does not contradict previous numerical studies [235, 325], in so far as the last studies demonstrated the topological nature of the transition for $q = 2$ and $q = 3$ only. However, the assumption that the “tricritical” line crosses the $\sigma = 1.0$ line raises a handful of markedly intriguing questions:

- First, this would imply that Cardy’s scenario [72] is somewhat flawed, since above $q = 8$, the transition can definitely not be *at the same time* of the first order and of infinite order; this entails considering that Cardy’s development based on a diluted kink-gas might not be valid at large q ; it is appealing to note at this point that several models, which normally exhibit a Kosterlitz-Thouless phase transition, have been found to undergo a first-order transition for sufficiently non-linear interactions between spins [100, 101, 328], so that the scenario mentioned above may not be deemed totally atypical;
- Although it is not stated in the article of Bayong et al. whether $\sigma = 1.0$ is indeed a “termination” point or if the tricritical line gets across it, the former hypothesis would give the point ($\sigma = 1.0, q \sim 8$) a very prominent status in the field of critical phenomena, i.e., a point terminating a line of topological

²⁶See Sec. 2.3.5.

4.3. Spinodals: a novel approach to assess the order of phase transitions

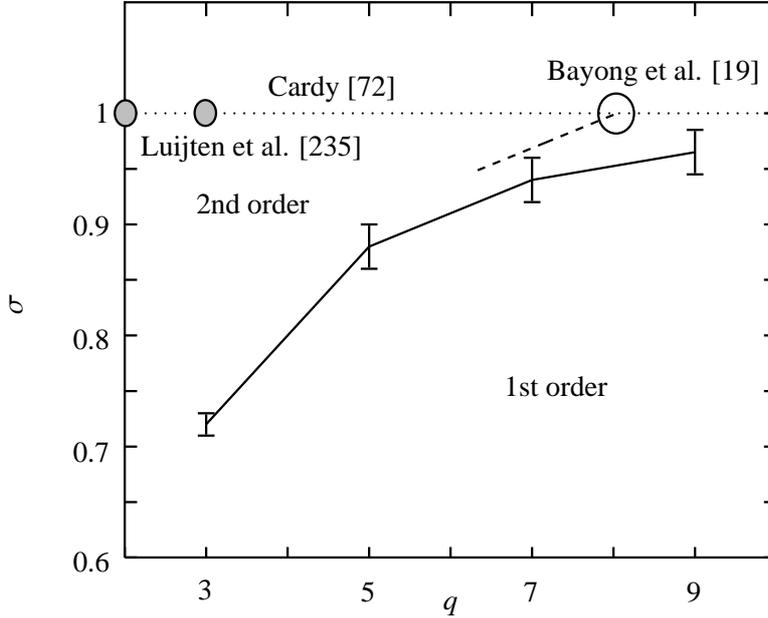


Figure 4.15: Phase diagram of Fig. 4.14, along with the disputed line of inverse square interactions. The open circle shows where the $\sigma_c(q)$ line was claimed to terminate in [19]. According to Cardy’s scenario [72], on the contrary, the *whole* line is the locus of topological transitions; grayed disks refer to a numerical study by Luijten and Messingfeld [235] reporting the onset of a topological transition for $q = 2$ and $q = 3$.

transitions, a region of second-order transitions, and a domain of discontinuous transitions; conversely, the latter hypothesis would prompt a reexamination of the (as yet debate-prone) long- to short-range crossover scenario (see Sec. 5.4 where the controversy is described at length, and revisited with multicanonical simulations).

The phase diagram obtained in the previous section is reminded in Fig. 4.15 along with the controversial areas. In what respect can this phase diagram shed light on this controversy? Clearly, if Cardy’s scenario proves reliable, then we must have $\lim_{q \rightarrow \infty} \sigma_c(q) \leq 1$. Although a numerical simulation may hardly bring about *rigorous* evidence with regard to an asymptotic behavior, the shape of the $\sigma_c(q)$ line in Fig. 4.14 lends strong support to such a behavior. Still, the double-peak structure at $\sigma = 1.0$ and large q remains intriguing and must be resolved in some way or another for the whole picture to be convincing. The purpose of the next section is to scrutinize again into this issue, yet with a markedly different approach involving (i) the spinodal method developed in Sec. 4.3, and (ii) a finite-size scaling analysis.

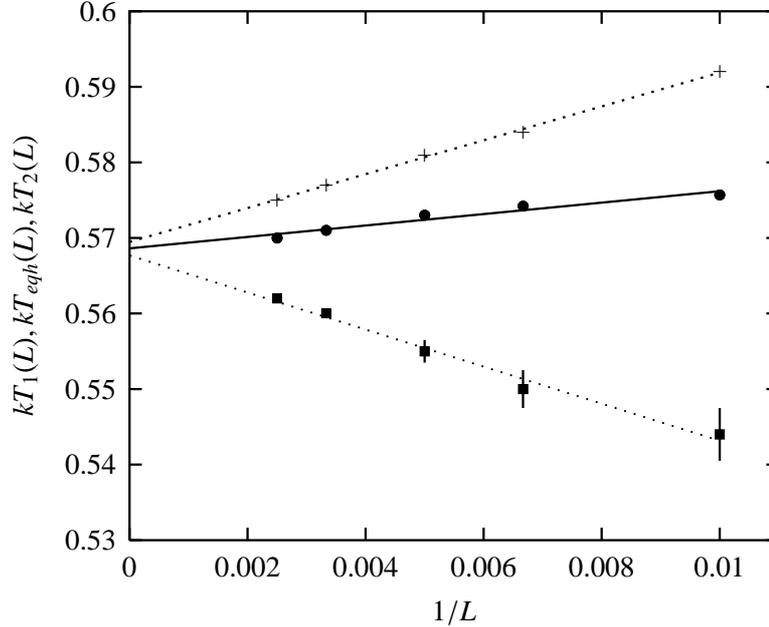


Figure 4.16: Linear fit of finite size temperatures vs $1/L$ for $q = 9, \sigma = 1.0$. Dotted, solid, and dashed lines correspond to kT_1 , kT_{eqh} , and kT_2 , respectively. Error bars are smaller than the size of symbols, except where explicitly indicated. In the limit $L \rightarrow \infty$, the difference between temperatures of metastability tends to 0.0012. Within error bars, the transition is thus clearly not of the first order.

4.3.7 Unexpected finite-size scaling behavior at $\sigma = 1.0$: case closed?

To set the stage, I will briefly inspect the case $q = 9, \sigma = 1.0$, that is, just above the change of regime claimed in [19]. I performed a series of simulations at $L = 50, 100, 150, 200, 300$, and 400 for this set of model parameters, and then I computed the corresponding (finite-size) spinodal temperatures $T_1(L)$ and $T_2(L)$ by relying on the free energy $F_m(kT, m)$. First, a naive analysis based on the shape of the free energy at a given lattice size might be markedly misleading, because the histogram indeed displays a double-peak structure *already* at sizes as low as $L = 100$.

As may be noticed in Fig. 4.16, a striking feature of this limiting case is the existence of metastable states at all finite lattice sizes, with a first-order character strongly enhanced at low sizes, despite the fact that finite-size scaling yields $T_2 - T_1 = 0$ in the thermodynamic limit. It turns out that, as opposed to the claim in [19], the transition is clearly not of the first order in the thermodynamic limit. This *unusual* (for reasons explained below) finite-size scaling behavior is also confirmed, though in a less marked way, for $q = 6, 7$, and 8^{27} . The latter case is

²⁷After the completion of this thesis, I recently reexamined the situation at $q = 12$ by means of the breathing cluster method and in the two-dimensional Potts model with medium-range interactions, where I witnessed the same finite-size effect.

4.3. Spinodals: a novel approach to assess the order of phase transitions

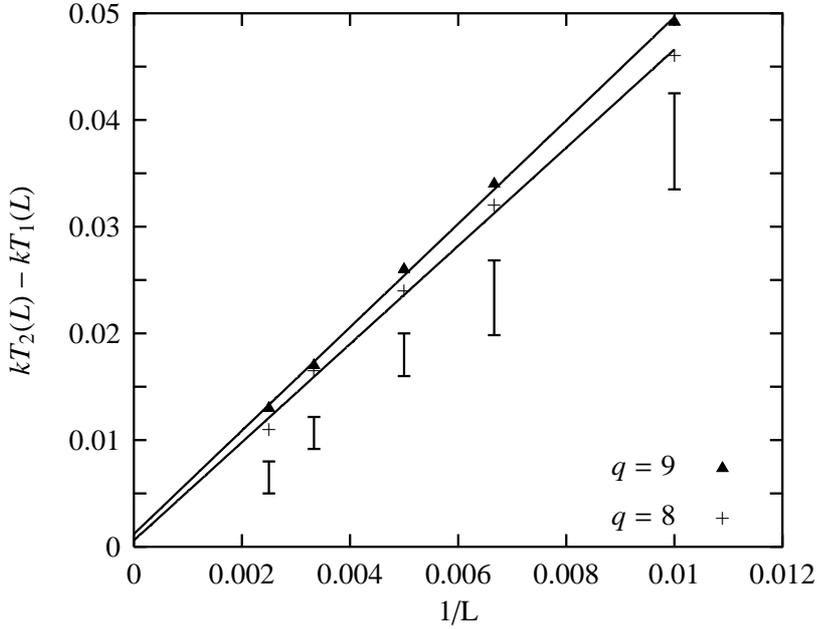


Figure 4.17: Metastability temperature differences $kT_2(L) - kT_1(L)$ on the line $\sigma = 1.0$ for $q = 8$ and $q = 9$, along with the corresponding linear fits (solid lines). For the sake of clarity, the size of the error for each lattice size is shown as a bar; the error shown corresponds to the largest of both errors computed for $q = 8$ and $q = 9$.

depicted in Fig. 4.17 along with $q = 9$. It is apparent that the strength of the first-order transition at small sizes increases with q . Conversely, for $q < 6$, a precise location of metastable states become impracticable at the medium lattice sizes that are under investigation, so that it seems *at first blush* that the distance between the line of inverse square interactions and the line $\sigma_c(q)$ plays a prominent role in the issue.

To wind up, there are two remarkable points in the issue:

- First, simply relying on the double-peak structure of the histogram (or on the temperature difference $T_2 - T_1$) at the transition temperature is clearly misleading; it is crucial that a finite-size scaling analysis should be carried out. This indeed rules out a first-order transition in the thermodynamic limit.
- Second, the observed finite-size effect is somewhat atypical, as against the usual picture where transitions turn from second- into first-order as the lattice size is increased. This point is discussed in greater detail hereafter.

Comparison with Lee's criterion

Before going over a more detailed interpretation of the previous effect, however, it is perhaps interesting to compare the prediction given by the spinodal method

with those yielded by more *established* indicators. Lee's criterion [219] relies on the behavior of the reduced surface tension $\Delta F/L$ in the thermodynamic limit. Figure 4.18 depicts a plot of the reduced "surface" tension with respect to the lattice size for $q = 9$ and various values of σ between 0.2 and 1.0. First, it is noticed that the tension indeed decreases as $\sigma \rightarrow 1$, confirming the change of regime near $\sigma = 1$. An extrapolation to the thermodynamic limit is depicted in Fig. 4.19, where a scaling law of the form $\Delta F_m/L \sim a + b/L$ was assumed, i.e., by considering that the long-range nature of the potential implies that the dimension of the interface is the same as that of the lattice, and restricting the fit to first-order corrections. As witnessed in Fig. 4.18 (top)), two regimes may be easily distinguished:

- For $\sigma = 1.0$, $\Delta F(L)/L$ behaves as $1/L$ and, according to Lee's criterion, this signals a *continuous* transition. It should be noted, however, that this scaling behavior is different from the one observed for "true" continuous transitions, e.g., for $q = 3$ and $\sigma = 0.9$, where I found $\Delta F(L) = 0$ at all sizes within error bars;
- For lower values of σ , the above quantity now behaves as $a + b/L$, and the transition is of the first order. This is in agreement with the results obtained using spinodal curves. Contrary to the case of first-order transitions in short-range models, however, the tension scales in first approximation as L^D rather than L^{D-1} . The figure also clearly indicates that correction terms are non-negligible (i.e. b is non-null).

As will be discussed in Chap. 7, the "surface" tension does actually scale very nicely as L^α , where α is a non-integer exponent that will be interpreted as being reminiscent of the fractal structure of the interface. It is somewhat awkward to ascertain such scaling behavior from the small set of data points available here, but this intriguing scaling behavior will be clearly visible when carrying investigation over several orders of magnitude in terms of linear size, as will done in Sec. 7.6

4.3.8 Outlook: boundary effects and fractal geometry

At first glance, this unusual finite-size effect substantially contradicts the expected picture, whereby at first-order transitions, the correlation length is finite and roughly independent of the lattice size (provided $\xi < L$), and is roughly connected to the size of clusters. As a result, first-order transitions appear as if they were continuous until the lattice size overtakes the correlation length. With regard to short-range models, this has been the standard scenario thus far: the behavior observed on the line of inverse square interaction suggests, however, that this scenario may be somewhat pondered about in the case of long-range models. This is met from three distinct perspectives in what follows.

The crucial observation that will underly the arguments laid down hereafter is that, in a long-range system, each particle sees the boundaries, irrespective of the lattice size, although as size is increased, particles in the bulk become more and more "blind" in this respect.

4.3. Spinodals: a novel approach to assess the order of phase transitions

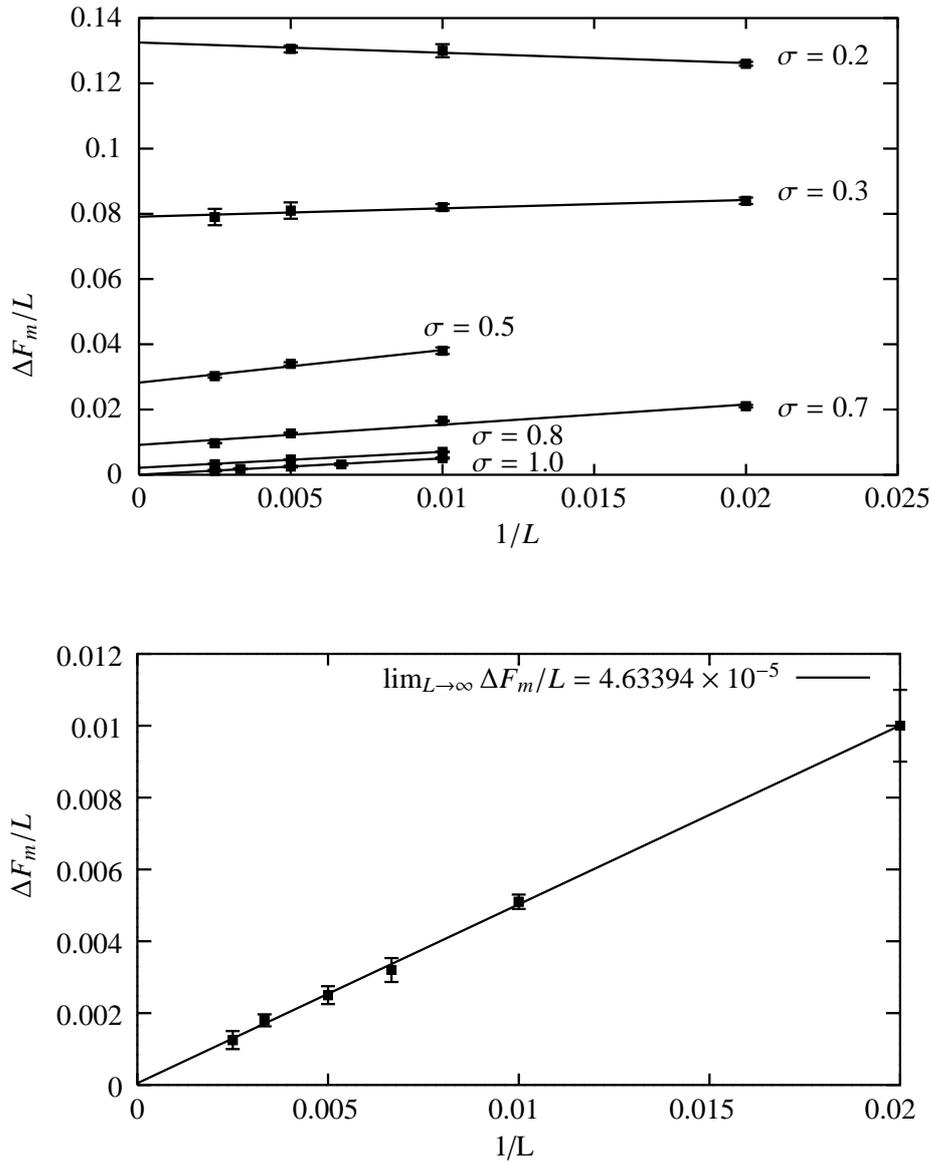


Figure 4.18: Reduced surface tension $F_m(kT, m)/L$ for $q = 9$ and various values of σ (top), and a detailed view for $\sigma = 1.0$ (bottom).

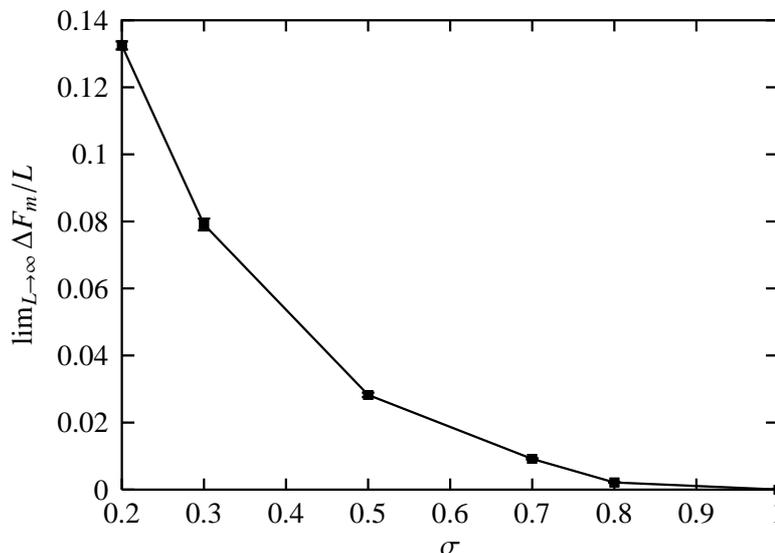


Figure 4.19: Extrapolated value of the reduced surface tension $\Delta F_m(L)/L$ in the thermodynamic limit for $q = 9$. A linear scaling law of the form $\Delta F_m(L)/L \sim a + b/L$ was assumed.

Truncation of the long-range potential One way to examine the issue is to consider the influence of the truncation of the long-range potential in finite geometry. In this regard, it is crucial to note that this truncation occurs *irrespective* of the periodic boundary conditions utilized: relying on infinite-image periodic boundary conditions simply speeds up the convergence towards the thermodynamic limit.

On a lattice of size L with periodic boundary conditions, the largest allowed distance between any two spins is $L/2$, and this also corresponds to the smallest interacting potential affordable on a given lattice. It is manifest that these spins experience a stronger interacting potential whenever L is small, and hence the whole array of spins may be rigidly tied to an adequate extent for an order-disorder transition to occur through metastability. When increasing the lattice size, on the contrary, spins being a distance $L/2$ apart now experience weaker interaction, and this results in a softening of the transition. Still and all, whether this softening might be sufficient to yield a change of nature of the transition at some (either finite or infinite) lattice size, so that the transition may be continuous in the thermodynamic limit, is an unsettled question; this assumption is borne out at least for $q = 9$ and $\sigma = 1.0$, as witnessed by the present results. Alternatively, one may say that the truncation of the long-range potential at small lattice size artificially shifts the model toward the mean-field regime, since the interacting potential now varies smoothly over the available distance of interaction. This “shift” should occur at all lattice sizes, although it will make its influence felt preferably at small sizes, and at lower values of the decay parameter. The fact that the “shift” is particularly visible here, on the line of inverse square interactions, pertains to the fact that this line lies very close to the change of regime. Overall, this feature certainly accounts — at

4.3. Spinodals: a novel approach to assess the order of phase transitions

least partially — for non-trivial finite size scaling.

Correlation lengths The usual physical meaning attributed to the correlation length in the case of short-range models, i.e., loosely speaking the average size of a cluster of contiguous spins having the same value, may no longer hold in the case of long-range potentials: since all the spins of the lattice, however distant they may be, are tied together through an interacting potential, there is basically no need of a long-range order for two distant spins to already have slightly correlated fluctuations. At continuous transitions, it has been shown that the correlation function comprises two parts: a long-range part reminiscent of the power-law decaying interaction, and a *true* correlation function [231] (see, also, Sec. 5.1). In the context of first-order transitions, this means that either clusters may extend well beyond the size permitted by the value of the *true* correlation length (i.e., once the long-range potential has been removed), or the correlation length itself may become infinite in the thermodynamic limit. This behavior has indeed already been reported in models of DNA thermal denaturation [314] as well as in the context of wetting [280].

Finite Size Scaling for long-range models revisited As was mentioned above, it is clear that finite-size effects must play a non-trivial role in the observed *unusual* behavior.

The key issue here, however, does not revolve around the finite-size scaling behavior of critical couplings, but that of free energies, surface tensions and metastable temperatures. In the previous section, it was shown that Lee’s criterion underpins this unusual finite-size scaling: this suggests that the thermodynamic quantities that are influenced by this effect are connected to the concepts of domain walls, surface tensions, and correlation lengths (in the nearest-neighbor Potts model, we know that the last two are connected [46]). To what extent are these quantities influenced by the boundaries in the case of long-range models is an intriguing question: in other words, what we lack for is a theory of finite-size scaling of *surface tensions* and metastable states at first-order transitions in long-range models.

In view of this, one may readily ponder on the topology of the mixed-phase configuration in long-range models, and its influence on finite-size scaling: owing to the long-range potential, the geometry of the interface is probably far less trivial than it is in nearest-neighbor models, and might exhibit *holes*, viz. a fractal (or “spongy”) structure.

Chapter 5

From long-range to short-range behavior

The present chapter focuses on the critical behavior of the long-range Potts chain *beyond* the boundary $\sigma_c(q)$. The investigations concentrate on the second-order regime of the model and on the crossover from the long-range to the short-range regime.

To set the stage, the main results from functional renormalization theory are reviewed in Sec. 5.1, and an outline of the derivation of a perturbative expansion around $\sigma = 1/2$ is given. It is important to underline, however, that these results correspond either to the O_n model, or to a continuum version of the Potts model with $q < 3$, while $q \geq 3$ in the model investigated here.

With this in mind, critical exponents are estimated for $q > 2$ using the multicanonical method outlined in Chap. 3. The focal point is the comparison of these exponents with (i) theoretical predictions and (ii) previous numerical studies (Sec. 5.2). The aim is to evaluate to what extent critical exponents for the Potts chain with $q \geq 3$ depart from those obtained from functional theory. A second objective is to estimate the precision of the multicanonical algorithm at modest lattice sizes, in particular as regards the peaks of response functions (which are obtained from a reweighting procedure), since these enter the estimation of critical exponents through finite-size scaling relations.

The second part of this chapter is devoted to the long- to short-range crossover. First, an extensive account of the controversies surrounding this issue is given in Sec. 5.3. Then, a novel approach is introduced to investigate the onset of the short-range regime: an exact transfer matrix is carried out for the *pure* short-range model, and compared with the temperature dependence of some carefully chosen thermodynamic quantities (Sec. 5.4). The multicanonical method will turn out to play a pivotal role in this approach.

5.1 Universality and critical properties for long-range Hamiltonians: conjectures

As already stated in Chap. 1, the interaction range affects the universal critical properties of a long-range spin models in much the same way as the dimension for their nearest-neighbor counterparts. As for interactions decaying as a power law of the interparticle distance, this role is played by the decay parameter σ , and in this respect, Ising chains represent perfect candidates since their behavior is critical along the whole σ line [40] (with even essential singularities inside some intervals). Although it has been demonstrated in the previous chapter that a substantial part of the phase diagram of the long-range Potts model corresponds to discontinuous transitions, i.e., the transition is first order for $\sigma < \sigma_c(q)$, the situation is not less interesting here for $\sigma > \sigma_c(q)$ for the three following reasons: first, there is still a small region under the tricritical line $\sigma_c(q)$ where the model exhibits a critical behavior, and estimates of critical exponents reported in several works show significant discrepancies in this respect; second, several rigorous results have long been made available in the case of O_n models, and it is tempting to pit them against models with a \mathbb{Z}_q symmetry; third, critical exponents are connected in some way to the long-range to short-range crossover problem, a long-disputed issue. The last point is considered in greater detail in a subsequent section of this chapter. The first two points are discussed hereafter, and to set the stage I will begin with a discussion of available results from functional theory regarding critical exponents for $1/r^{D+\sigma}$ interactions.

5.1.1 Ginzburg-Landau functional

In Fourier space, the corresponding Ginzburg-Landau functional admits, in the small- k limit and for an n -component field $\vec{\phi}(k)$, the following expansion,

$$H = \int d^D k (r + v_2 k^2 + v_\sigma k^\sigma) \vec{\phi}(k) \cdot \vec{\phi}(-k) \quad (5.1)$$

$$+ \int d^D k_1 d^D k_2 d^D k_3 \frac{u}{6} (\vec{\phi}(k_1) \cdot \vec{\phi}(k_2)) (\vec{\phi}(k_3) \cdot \vec{\phi}(-k_1 - k_2 - k_3))$$

where the mass term r varies linearly with the temperature¹ near T_c . The only difference with standard ϕ^4 theory resides in the presence of a k^σ term in the gaussian part of the Hamiltonian, which is non-analytical for $\sigma \notin \mathbb{N}$. At $\sigma = 2$, logarithmic terms must enter the previous expression, i.e. k^σ is replaced by $k^2 \ln k$. From a dimensional argument, it is easy to see that the k^2 term (which is reminiscent of the “gradient” term in short-range models) is irrelevant for $\sigma < 2$, i.e., vanishes under successive renormalization transforms. Therefore, the (bare) inverse propagator

¹That is, $r \propto t$, where t is the reduced temperature deviation with respect to the critical temperature of the gaussian model.

5.1. Universality and critical properties for long-range Hamiltonians: conjectures

actually reads $\sim (r + k^\sigma)$ in this region² It is interesting to note that, for $\sigma > 2$, it is now the k^σ term which is irrelevant and the model belongs to the short-range universality class. However, the (now subleading) long-range interaction gives rise to non-trivial finite-size scaling effects [80, 91, 90].

5.1.2 Upper critical dimension

The upper critical dimension can be readily estimated by relying on the Ginzburg criterion (see for instance [76]), which sets the limit of validity of the mean-field approximation from the ratio of field fluctuations to field averages. Taking guidance from standard mean-field theory applied to short-range functionals, the correlation function in the mean-field approximation writes

$$G(k) = \frac{kT}{v_\sigma} \frac{\xi^\sigma}{1 + (k\xi)^\sigma}$$

where the correlation length is given by

$$\xi = \left(\frac{v_\sigma}{r + u \langle \phi \rangle^2} \right)^{1/\sigma}$$

and $\langle \phi \rangle$ is the order parameter in this approximation, with $\langle \phi \rangle = -3r/u$ in the ordered phase. The Ginzburg criterion reads $\xi^{-D} G(k=0) \ll \langle \phi \rangle^2$, i.e. (in the ordered phase), $r^{2-D/\sigma} \gg 1$; whence $D_c = 2\sigma$ is the upper critical dimension. Alternatively, this result can be established from renormalization theory³, and has also been proven on rigorous grounds in [6].

5.1.3 Momentum-shell renormalization

The first calculation of the critical exponents for a general O_n model with long-range interactions $1/r^{D+\sigma}$ can be traced back to the seminal works by Fisher, Ma and Nickel [115] and Suzuki, Yamazaki and Igarashi [306] in the early 70's. Both are based on the momentum-shell renormalization group method introduced by Wilson [345, 346]. Basically, Wilson's approach is a Fourier-space version of Kadanoff's decimation scheme, where a renormalization transform of scaling b consists in integrating over field fluctuations of large momenta $k > \Lambda/b$, where the integral cut-off is defined by $\Lambda = 2\pi/a$ with a the (current) lattice spacing. Formalizing this a little, I write $\phi(k) = \phi^<(k) + \phi^>(k)$, i.e., the original field is split up into low- and high-momentum terms having supports $[0, \Lambda/b[$ and $[\Lambda/b, \Lambda[$ respectively. The renormalized Hamiltonian H' is formally obtained through a functional

²This would indeed correspond to a $(m^2 + k^\sigma)^{-1}$ bare propagator in field theory parlance, and thus $m^2 + k^\sigma + \Sigma(k)$ is the inverse two-point correlation function where Σ denotes the proper energy as usual. Apart from the non-analyticity of k^σ , which has non-trivial consequences, the same line of arguments as in ϕ^4 theory applies.

³This corresponds to the gaussian fixed point becoming unstable.

integration over $\phi^>(k)$,

$$e^{-H'[\phi'^<(k')]} = \int \mathcal{D}\phi^>(k) e^{-H[\phi^<(k)+\phi^>(k)]}$$

where primes indicate rescaled quantities: $k' = kb$ (this follows from $\xi \rightarrow \xi/b$ in real-space) and $\phi'(k') = \zeta\phi(k)$, where ζ is the *field rescaling factor*. The latter is generally fixed by imposing that the amplitudes of correlation functions be left unchanged under rescaling at the critical point, i.e., $\zeta^2 = b^{D-\eta-2}$. The exponent η enters the expression of the correlation function in Fourier-space (at criticality) through $G(k) \sim k^{\eta-2}$, and is termed the *correlation function exponent* or *Fisher's exponent*.

5.1.4 Gaussian model with long-range interactions

For a gaussian model with a $r + k^\sigma$ inverse propagator ($v_\sigma = 0$), the derivation is as straightforward as in the standard gaussian model (i.e., with $r+k^2$), since the Hamiltonian H can be separated into a low-momentum part $H^<$ and high-momentum part $H^>$, and the integration over higher momenta trivially yields the following recursion relations,

$$\begin{aligned} r' &= b^{2-\eta}r \\ v'_\sigma &= v_\sigma b^{2-\eta-\sigma} \end{aligned}$$

After imposing that v_σ be kept constant (i.e., ζ is imposed by the most stable fixed point), one obtains $\eta = 2-\sigma$ and $\nu = 1/\sigma$; the standard gaussian model is recovered when $\sigma = 2$, i.e., $\eta = 0$ and $\nu = 1/2$. The last results follow from the scaling behavior of the correlation length, $\xi \sim |r|^{-\nu}$; whence $r'/r = b^{1/\nu}$.

5.1.5 Perturbative expansion in $\epsilon = 2\sigma - D$

When $v_\sigma \neq 0$, the derivation is less trivial, since the Hamiltonian is no longer separable into low- and high-momentum terms, and a (diagrammatic) perturbative expansion in v_σ must be carried out. This expansion shows in particular that the upper critical dimension of the model is $D_c = 2\sigma$ (whereby $D_c = 4$ is recovered for the short-range ϕ^4 theory⁴). It is not the aim of this thesis to describe the method in detail; numerous writings have been published on the subject, see, for instance, [76, 21] for an introductory course, and [174, 363] for more specialized monographies. Therefore, my goal in the following will be to discuss the results more than their derivation.

As in conventional field-theoretic renormalization applied to short-range models, the derivation in [115] is an ϵ -expansion around the upper critical dimension $D_c = 2\sigma$, i.e., $\epsilon = 2\sigma - D$. It is valid for any D -dimensional O_n model provided

⁴And yet, as will be seen below, this that not entail that $\sigma = 2$ is the boundary between the short- and long-range regimes.

5.2. Critical exponents for $q > 2$: multicanonical simulations

$\sigma \neq 2$; I recall that the last case induces a different form for the Ginzburg-Landau functional due to the presence of a logarithmic term in the gaussian part. The spherical model can be recovered in the limit $n \rightarrow \infty$. Renormalization group equations yield two fixed points: the standard gaussian fixed point ($r^* = u^* = 0$), and a non-classical (Heisenberg-like) fixed point of order ϵ ($u^* = \epsilon v_\sigma^2 / (n + 8)$, unstable direction). The main findings of [115] are as follows:

- $\sigma < D/2$ defines the so-called “classical” regime where the gaussian fixed point is stable (i.e., mean-field like where critical exponents are concerned), where the correlation function exponent $\eta = 2 - \sigma$, the correlation length exponent $\nu = 1/\sigma$ and the susceptibility exponent $\gamma = 1$;
- on the boundary $\sigma = D/2$, logarithmic corrections to scaling apply;
- for $D/2 < \sigma < 2$, the “non-classical” region (where the gaussian fixed point loses its stability), η is still given by $\eta = 2 - \sigma$ (with no correction to order ϵ^2 and ϵ^3 at least), i.e., η retains its classical value, whereas γ (after reexpressing the ϵ -expansion in terms of $\Delta\sigma = \sigma - D/2$) is given to first order in $\Delta\sigma$ by $1 + \frac{4}{D} \frac{n+2}{n+8} \Delta\sigma$; the expression of ν follows from the standard scaling relation $\nu = \gamma / (2 - \eta) = \gamma / \sigma$;
- for $\sigma > 2$, the exponents take their short-range value. Note that, in one-dimension, this picture is somewhat incomplete, since the absence of a phase transition at finite temperature was rigorously proven for $\sigma > 1$ [107]; this is discussed in greater detail in Sec. 5.3; another interesting remark is the fact that, at least for $D > 4$, the system is in its mean-field regime whatever σ (although this condition is sufficient, yet not necessary).

From the expression of η , the correlation function thus behaves according to $G(r) \sim 1/r^{D-\sigma}$ (to be compared with $J(r) = 1/r^{D+\sigma}$ for the interaction), which means that the correlation function dies off more slowly than the interaction, as expected from the intuitive picture. This behavior was confirmed at short distance by means of simulations in the classical regime [231].

As a last point in this section, I would like to mention how startling it seems at first that more than two decades have separated the seminal derivation of these results through analytical perturbation methods and their verification by means of Monte Carlo studies (see, for instance, [230], which — historically — perhaps represents the first really extensive study in this respect). This should certainly be seen as another evidence of the challenge that long-range interactions represent from the viewpoint of numerical simulations.

5.2 Critical exponents for $q > 2$: multicanonical simulations

This section reports results obtained by multicanonical simulations regarding the critical properties of the q -state Potts chain in the *continuous* regime, i.e., for

$\sigma_c(q) < \sigma \leq 1.0$. The $\sigma > 1.0$ case will be considered in a subsequent part of this chapter in connection with the crossover from long- to short-range behavior.

As far as the $\sigma = 1$ boundary is concerned, a few remarks are in order. As mentioned in Sec. 4.3.6, the line of inverse square interactions, $\sigma = 1$, has been claimed by some authors to be the locus of a Kosterlitz-Thouless-like transition exhibiting *essential* singularities. The work of Cardy [72] is enlightening in this respect, since the q -state Potts model should exhibit this kind of transition for all q . Although several other authors also reported results in contradiction with this assumption, claiming power-law singularities and a finite correlation length exponent ν [19, 64, 129, 130], a recent Monte Carlo study tipped the balance in favor of Cardy's scenario, at least for $q = 2$ and $q = 3$. A work by Glumac and Uzelac [325] on the three-state model, based on the largest cluster probability, also proved consistent with this scenario, with exponents of the essential singularity found very close to the values predicted by Cardy. This scenario is also supported by the findings of the previous chapter (Sec. 4.3.7), where I suggest that $\sigma_c(q) \rightarrow 1$ asymptotically in the $q \rightarrow \infty$ limit. It is not the purpose of the present work, however, to (try to) resolve the issue on the line of inverse square interactions for $q > 3$. The presence of essential singularities renders the endeavor particularly challenging, for this imposes performing exceedingly long runs and to investigate huge lattice sizes, and a multicanonical method with single-spin updates is certainly not the most suited approach in this respect.

It is perhaps interesting to mention, however, that I found finite-size scaling fits on this line to be particularly poor; in particular for those regarding the estimate of the correlation length exponent $1/\nu$, log-log fits depart strongly from the straight line expected from a power-law fit. In addition, autocorrelation times diverge in a startlingly rapid manner with the lattice size at $\sigma = 1$. Although these observations do not yield any *quantitative* information concerning essential singularities, it lends at least some support to the onset of a topological phase transition for $q > 3$.

I will concentrate in the following on the estimates of critical exponents for $q = 3, 4, 5$, excluding the correlation length exponent at $\sigma = 1$ (although exponent ratios γ/ν and β/ν will be considered, as they are still well-defined in the case of essential singularities). Higher values of q are not investigated, for the region corresponding to a continuous transition is then too narrow.

5.2.1 Objectives

In the previous section, the estimates of critical exponents from functional theory were reviewed. For the O_n model, an interesting finding in the non-classical regime is $\eta = 2 - \sigma$, which should hold up to order ϵ^3 ($\epsilon = 2\sigma - D$) for all n .

Since critical exponents are otherwise related by standard scaling relations, e.g., $\gamma/\nu = 2 - \eta$, one also expects the simple relation $\gamma/\nu = \sigma$. This can easily be checked in numerical simulations, since γ/ν is directly related to the finite-size scaling behavior of the susceptibility through $\chi^{max} \sim L^{\gamma/\nu}$, where χ^{max} designates the peak of the susceptibility.

5.2. Critical exponents for $q > 2$: multicanonical simulations

Because the previous relation was derived, however, for model with an n -component order parameter and a continuous symmetry (for $n > 1$), it is interesting to check if the same relation holds also for a discrete spin model with a \mathbb{Z}_q symmetry, $q > 2$, and in particular, to monitor the validity of $\gamma/\nu = \sigma$ for several values of q . To the best of my knowledge, this has never been investigated numerically for the q -state Potts chain, $q > 2$; on the analytical side, the renormalization group derivations by Priest and Lubensky [279] and Theumann and Gusmão [315] target $q < 3$ (non-integer), and particularly the percolation limit $q \rightarrow 1$.

A similarly interesting question naturally arises for the correlation length exponent, since the derivation for the O_n models indicates that it depends on n to order ϵ . This was already investigated in the context of real-space renormalization [64] (see also Sec. 4.1.3 where an extensive description of the method is given), where it emerged that $1/\nu$ should not depend on the number of states of the model. A work by Glumac and Uzelac based on transfer matrices [130] showed, however, that this might not be true, with $1/\nu$ increasing with q for a given decay parameter. This was confirmed by the same authors using Monte Carlo simulations and the largest cluster probability [325]⁵. Both issues are looked upon for $q > 2$ in the remainder of this section.

5.2.2 Critical exponents: method

As mentioned in [231], “standard” finite-size scaling theory is valid for long-range systems provided the effective upper critical dimension $d^* = 2\sigma$ is greater than the geometrical dimension $d = 1$, i.e., $\sigma > 0.5$ (this is similar to the short-range case, where standard scaling relations hold below the upper critical dimension, see [228]). Thus it is assumed here that “standard” finite-size scaling equations are also valid for $q \geq 3$.

The critical exponent ν is determined using n th-order cumulants of the magnetization, i.e.,

$$V_n = \frac{d \ln \langle m^n \rangle}{d\beta},$$

which have minima obeying the scaling law $V_n^{min} \propto L^{1/\nu}$ [112]. The approach consists in computing two numerical estimates of ν by fitting reweighted averages of V_1^{min} and V_2^{min} to a power law of the lattice size, and then to average over both values. Other critical exponents, i.e., β and γ , are computed using similar scaling laws, i.e., $M(T_c(\infty)) \propto L^{-\beta/\nu}$, and $\chi^{max} \propto L^{\gamma/\nu}$.

Figures 5.1 and 5.2 depict a power-law fit of peaks of V_1 , V_2 and χ against the lattice size obtained for $q = 5, \sigma = 0.9$. Points lie neatly on a straight line when using a log-log scale, and give the following estimates: $1/\nu_1 = 0.668(2)$, $1/\nu_2 = 0.669(2)$, $\gamma/\nu = 0.940(4)$. Error bars were computed using a bootstrap procedure, see Appendix C.

⁵See Sec. 2.3.5 for more details on this approach.

Once ν has been estimated, finite-size temperatures $T_c(L)$ defined from peaks of the magnetic susceptibility are fitted to a power law of the form

$$T_c(L) = T_c(\infty) + \lambda L^{-1/\nu},$$

and yield an estimate of the infinite-size critical temperature. With regard to critical couplings obtained from Binder cumulants of the magnetization, I follow the same procedure as in the first-order regime (Sec. 4.2.2).

Finally, the critical exponent β of the magnetization is determined by performing a fit of the magnetization estimated at the infinite-size critical temperature, $M(T_c(\infty))$, to a power law of the lattice size. In order to improve the accuracy of the estimation, the (reweighting) temperature at which M is to be sampled is slowly varied around $T_c(\infty)$ until the best fit is obtained. In the example considered above, this leads to $\beta/\nu = 0.103(2)$.

5.2.3 Discussion

Results for several pairs of (q, σ) values are summarized in Table 5.1 for the critical exponents, and Table 5.2 for the critical couplings. On the $\sigma = 1.0$ boundary, only exponent ratios are shown. It can be seen that these estimates match fairly well with those obtained from a previous Monte Carlo study based on the single-histogram method and sizes up to $L = 900$ spins [19], and that the discrepancy with results obtained from a transfer matrix approach in [130] never exceeds 8%.

As opposed to the conjecture made in [64], the exponent ν does clearly depend on q . This feature was also recently reported for $q = 2$ and $q = 3$ in a numerical study exploring the largest cluster probability [326, 325]. It is interesting to note that the “criticality” of the model decreases with increasing q for a given range of interaction, in compliance with the convex shape of the tricritical line $\sigma_c(q)$ (Fig. 4.14) obtained in Chap. 4.

If the relation $\sigma = 2 - \eta$ derived in [115] for the O_n model, and in [315] for the continuum version of the Potts model (yet *provided* $q < 3$), holds also for $q \geq 3$, the simple behavior $\gamma/\nu = 2 - \eta = \sigma$ should be observed in the second-order regime. As depicted in the sixth column of Table 5.1, the qualitative behavior follows the conjecture, yet clearly $\sigma < 2 - \eta$, and the discrepancy is remarkably higher for $q = 5$ than for $q = 3$. Moreover, while γ/ν appears to tend to 1 as $\sigma \rightarrow 1$, it is unclear whether this ratio varies linearly with σ , considering the small number of data points available.

5.3 Long-range to short-range crossover: three decades of controversies

As already mentioned in this chapter, one should *naively* expect a short-range behavior whenever $\sigma > 2$, for the k^σ term has then the same exponent as the k^2 term in Eq. 5.1, and the latter corresponds to the squared gradient term (in real space) that

5.3. Long-range to short-range crossover: three decades of controversies

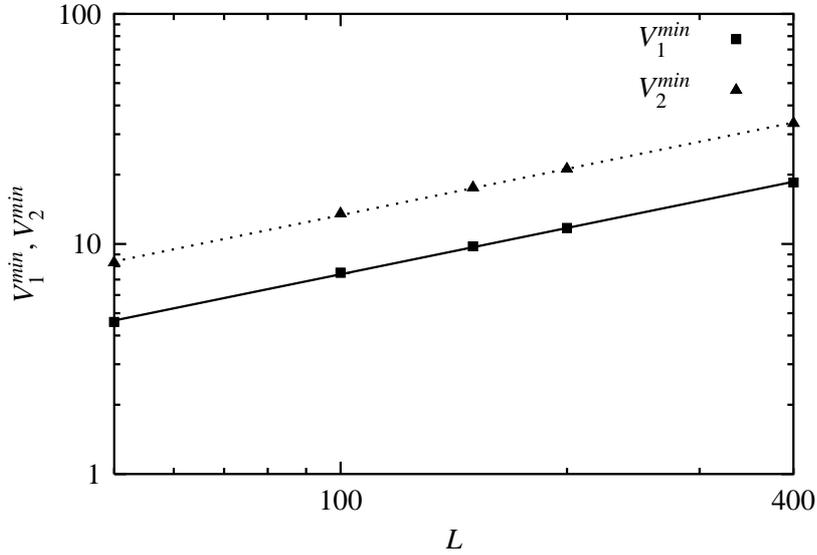


Figure 5.1: The graph shows the peaks of the cross-cumulants of the magnetization, V_1^{min} and V_2^{min} with respect to the lattice size L on a log-log scale, for $L = 50, 100, 150, 200, 400$. The model parameters are $q = 5$ and $\sigma = 0.9$. The power-law fit yields $1/\nu = 0.6675(1)$ and $1/\nu = 0.669(2)$ for V_1 and V_2 , respectively. Errors were computed from a bootstrap method; they are smaller than the size of symbols.

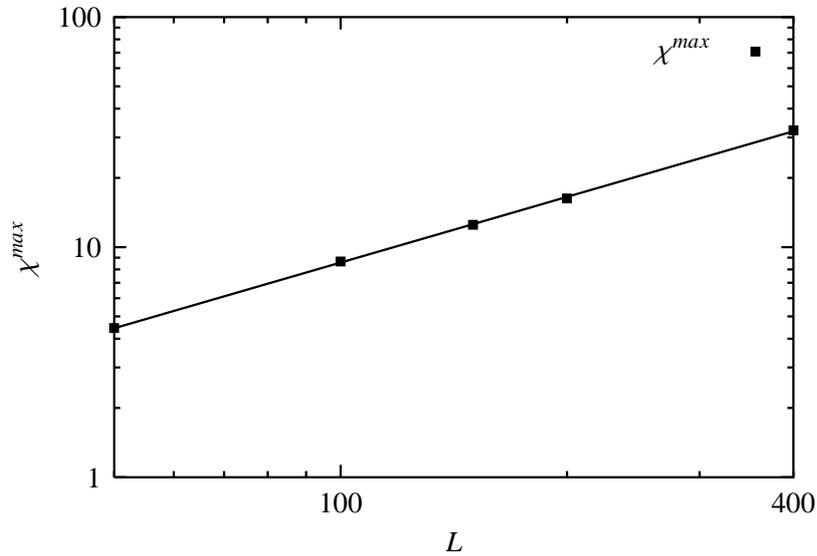


Figure 5.2: Maximum of the susceptibility χ^{max} plotted against the lattice size L , for $L = 50, 100, 150, 200, 400$, and the following model parameter: $q = 5, \sigma = 0.9$. A fit of the data points to $\chi^{max} \propto L^{\gamma/\nu}$ yielded $\gamma/\nu = 0.940(4)$. Errors are smaller than the size of symbols.

Chapter 5. From long-range to short-range behavior

q	σ	ν^{-1}	ν^{-1} [325]	ν^{-1} [130]	γ/ν	β/ν
3	0.8	0.624(6)	0.61(2)	0.574	0.842(5)	0.101(5)
	0.9	0.54(1)	0.48(1)	0.491	0.908(5)	0.053(5)
	1.0				0.96(1)	0.025(8)
4	0.8	0.71(1)		0.67	0.882(3)	0.122(4)
	0.9	0.610(5)		0.56	0.920(4)	0.050(3)
	1.0				0.96(1)	0.022(9)
5	0.9	0.668(2)		0.62	0.940(4)	0.103(2)
	1.0				0.97(1)	0.04(1)
	1.0 [19]				0.966	0.017

Table 5.1: Critical exponents in the second-order regime $\sigma > \sigma_c(q)$, and $q = 3, 4, 5$. Shown for comparison are results from Ref. [130] (transfer matrix method), from Ref. [325] (largest cluster probability, with sizes up to $L = 20000$ spins), and from Ref. [19] using a Monte Carlo single-histogram method.

q	σ	T_c (MF)	$T_c(\chi)$	$T_c(U^{(4)})$	T_c [325]	T_c [131]	T_c [251]	T_c [130]
3	0.8	1.358	1.019(4)	1.03(1)	1.026(1)	1.01		1.0231
	0.9	1.262	0.876	0.875	0.8735(10)	0.88	0.8785	0.874
5	0.9	0.947	0.743(2)	0.739(4)		0.74673	0.745	

Table 5.2: Estimates of the critical couplings in the second-order regime. MF, mean-field predictions; χ , using peaks of the susceptibility; $U^{(4)}$ using crossing points of Binder cumulants of the magnetization; Ref. [325], Monte Carlo study relying on the largest cluster probability, with sizes up to $L = 20000$ spins, Ref. [131], Monte Carlo study based on histogramming and the Luijten-Blöte cluster algorithm ($q = 3$) and a standard metropolis algorithm ($q = 5$); Ref. [251]), cluster mean-field method combined with an extrapolation technique based on the VBS (Vanden Broeck and Schwartz) algorithm; Ref. [130]), transfer matrix method combined with finite-range scaling.

5.3. Long-range to short-range crossover: three decades of controversies

stems from small- k expansions in short-range models. Yet the situation, as it turns out, is not *that* trivial, and has raised much controversy indeed during the last two decades. A first contradiction in this approach emerges for one-dimensional models, where a phase transition at finite temperature occurs for $\sigma \leq 1$ only [108, 124], so that a crossover to short-range behavior can definitely not take place at $\sigma = 2$. A second issue is related to the η critical exponent, since we have $\eta_{LR} = 2 - \sigma$, and thus $\eta_{LR} = 0$ at $\sigma = 2$, whereas the short-range exponent $\eta_{SR} \neq 0$, with for instance $\eta_{SR} = 1$ for all values of q in the one-dimensional Potts chain (See appendix B) and $\eta_{SR} = 1/4$ for the two-dimensional Ising model. This would imply a discontinuity in η at $\sigma = 2$. From a very general viewpoint, there are at least two possible scenarios:

- either $\sigma = 2$ determines the locus of the crossover, and there is a jump in η ; following thermodynamic stability arguments, one must simply ensure that $\eta \leq 2 + \sigma$, so that this scenario, though somewhat unusual in the theory of critical phenomena, is perfectly allowed; still, one-dimensional models would not fit into this scheme;
- or the crossover occurs at another value of σ , i.e., $\sigma_{co} = 2 - \eta_{SR}$, in which case there is no longer a discontinuity; this would imply $\sigma_{co} = 1$ for the Potts chain, in accordance with Dyson's prediction of the absence of a phase transition for $\sigma > 1$; for model with negative Fisher exponent (e.g., Yang-Lee Hamiltonians with imaginary couplings), however, this is in contradiction with the onset of the short-range universality class at $\sigma > 2$.

The best way to try to treat this problem from an analytical viewpoint is to consider changes in universality classes, i.e., to be specific, to monitor the variation of critical exponents with σ and to look for a change of regime at some crossover value σ_{co} . In the language of Wilson's renormalization theory, this question reduces (although not trivially) to the investigation of the competition between long- and short-range fixed points. Numerous works based on renormalized perturbation theory or on Wilson's recursion relations have been performed in this respect [292, 356, 357, 142, 315, 160, 159, 349], including specific investigations in finite geometry [91, 80]; the main picture that seems to emerge today is the prominent role played by the sign of Fisher's exponent in the short-range regime η_{SR} with respect to the locus of the crossover.

The second scenario mentioned above, along which $\sigma_{co} = 2 - \eta_{SR}$, was first proposed by Sak [292], on the grounds that higher order terms in the ϵ -expansion (Eq. 5.1), which were not taken into account in [115], influence the competition between long-range and short-range fixed points. To be specific, this means that the full functional $v_2 k^2 + v_\sigma k^\sigma$ must be considered in recursion relations, with the short-range part k^2 explicitly taken into account. By carrying out an expansion to order ϵ^2 (where $\epsilon = 2\sigma - D$), Sak established that at the (Heisenberg-like) fixed point, the v_2 term is not zero to order ϵ^2 (whereas it is zero to order ϵ). In other words, starting from $v_2 = 0$, the short-range term increases under renormalization

(because the recursion relation for v_2 involves v_σ to order ϵ^2), and starts to compete with the long-range term. On the grounds that the choice of the field rescaling factor ζ must depend on the sign of $\eta - \eta_{SR}$, Sak established that this leads to a special region of weak long-range interactions for $\sigma > 2 - \eta_{SR}$ where $\eta = \eta_{SR}$, which corresponds to η and ϵ^2 having the same magnitude. In this region, the fixed point is given by $v_\sigma^* = 0$, and the universality class is short-range. On the other hand, if σ is sufficiently small (i.e., below $2 - \eta_{SR}$), then the long-range term wins out over the short-range one under renormalization and the v_2 term can be safely ignored. The scenario of Yamazaki [356] draws more or less the same conclusions. These findings were later reassessed to all orders of the expansion by means of renormalized perturbation theory [160], and rederived for the random-field Ising model with long-range interactions [52]. However, Van Enter [327] pointed out that this scenario may be somewhat flawed with regard to models with rotational symmetry breaking in their short-range regime (i.e., XY models in $D \geq 3$ dimensions in the case considered in [327]), since symmetry breaking should be destabilized by the presence of an irrelevant long-range perturbation, which according to the above scenario corresponds to $2 - \eta < \sigma < 2$.

Another expansion parameter, $\epsilon = 3\sigma - D$, was considered by Priest and Lubensky [279] in the context of a (short-range) continuum version of the Ashkin-Teller-Potts model. Here the authors suggested in their conclusion, as an extension of their results to power-law decaying interactions, that the long-range fixed point is again stable with respect to the short-range one as long as $\sigma < 2 - \eta_{SR}$. They also pointed out that the case $\eta_{SR} < 0$, which along the previous lines of arguments should yield inconsistent behavior with a crossover taking place at $\sigma > 2$, might actually be straightforwardly resolved by observing that the long-range fixed point no longer exists for $\sigma > 2$, implying that long-range behavior is dominant up to $\sigma = 2$ only. The same conclusion appeared again in [315] (from an expansion to second order in ϵ). Systems with $\eta_{SR} < 0$ (corresponding to ϕ^3 theories, or ϕ^4 theories with $D \geq 4$) were then reconsidered recently by Janssen [192], who claimed that the boundary is given by $2 - \sigma_{co} = \eta_{SR}$ *also* when $\eta_{SR} < 0$, owing to the pivotal role played by nonquadratic terms of higher-order in the Ginzburg-Landau functional (in particular the $k^2 \ln k$ term that, incidentally, shows up *also* when expanding the Fourier transform of the interaction exactly at $\sigma = 2$).

Finally, it is worth mentioning the contribution to this debate by Wragg and Gehring [349], who proposed an extension of the variational method of Takahashi [311] to an Ising model with power-law decaying interactions (i.e. the long-range part of the interaction is considered as a perturbation of the nearest-neighbor potential, and the minimization of the free-energy is carried out with respect to the last potential). In particular, the long-distance behavior of the correlation functions was carefully examined, and it was established in this respect that the long-range contribution to the correlation function is dominated by the short-range contribution for $\sigma > 2 - \eta_{SR}$ (including the one-dimensional case). No mention was made here, however, of the influence of the sign of η_{SR} .

5.4. Crossover in the three-state Potts chain: multicanonical simulations vs exact transfer matrix

For systems with negative Fisher exponents, therefore, the situation seems still rather unsettled, and numerical investigations on the D -dimensional long-range Ising model with $D > 4$ would probably be enlightening in this respect. In two dimensions ($\eta_{SR} > 0$), however, compelling evidence in favor of Sak's scenario was produced in a recent numerical study by Luijten [232]; it was claimed that the modified Binder cumulant $Q = \langle m^2 \rangle^2 / \langle m^4 \rangle$ takes its short-range value around $\sigma = 1.75$, and similarly that η exhibits a (clearly visible) kink at the same value of σ , in accordance with $\sigma_{co} = 2 - \eta_{SR} = 7/4$ for the two-dimensional Ising model. In one dimension, on the contrary, the transfer matrix approach of Glumac and Uzelac was quite unconvincing in this respect [130].

Still, the situation remains unclear in the one-dimensional case, where only approximate renormalization group results are available, and no numerical simulation has addressed the issue thus far. It is interesting to note here that the value $\sigma_{co} = 1$ predicted by Sak's scenario is also the locus of the Kosterlitz-Thouless transition suggested in [208, 72] (see, also, Sec. 4.3.6 for more material on this issue), so that the phase diagram of the long-range Potts chain would in effect contain a region of essential singularities for all $\sigma \geq 1$, seeing that the nearest-neighbor model exhibits this kind of divergences as well. From the viewpoint of numerical simulations, the endeavor in one dimension is thus markedly challenging, since the presence of essential singularities *a priori* requires covering a large range in linear system size for finite-size scaling to be reliable. It is the goal of the next section to investigate these questions by means of a multicanonical approach; I will show, in particular, that by relying on properly chosen indicators, it is perfectly possible to carry off the investigation without having to resort to huge simulation sizes.

5.4 Crossover in the three-state Potts chain: multicanonical simulations vs exact transfer matrix

In this section, I investigate the crossover from short-range to long-range behavior in the three-state Potts chain with power-law decaying interactions. I make use of the multicanonical method outlined in Chap. 3, and concentrate on σ above the boundary value $\sigma_{co} = 1$, where a crossover is expected to take place along the line of Sak's scenario [292] (recalling that $\gamma/\nu = 1$ for all values of q in the short-range case, $\eta_{SR} = 1$, and this indeed leads to $\sigma_{co} = 1$). The choice of $q = 3$ as the particular value under investigation here is merely a matter of coherence with the rest of this work, since the investigations (in particular in Chap. 4) targetted the phase diagram for $q > 2$. Investigating the Ising chain would not have made sense in this respect.

As described above, the definition initially proposed by Sak in [292] on theoretical grounds, as well as the exact location of σ_{co} within the interval $[1.0, 2.0]$, is still somewhat controversial. Signs that the crossover takes place at $\sigma \sim 1$ *might* be seen from the dependence of γ/ν on σ , as shown in Table 5.1: γ/ν indeed appears to reach its short-range value as $\sigma \rightarrow 1^-$, yet this ratio proves no longer reliable

above the boundary value, as we will witness below, and it will become necessary to meet the problem from a completely new angle.

One of the key points in the present approach, indeed, is that I will be able to glean crucial information over the change of regime from a careful comparison of the temperature-dependence of (some carefully chosen) reweighted thermodynamic quantities with the exactly known results in the nearest-neighbor case. This is where the multicanonical method turns out to play a pivotal role, because it yields reweighted averages over a large range of temperature and makes the comparison pertinent. My first item of business will be to derive exact results in the short-range case. This is most easily done using transfer matrices.

5.4.1 Exact results for the short-range chain: transfer matrix derivation

I first briefly review some exact results concerning the q -state Potts chain with nearest-neighbor interactions. Detailed calculations can be found in Appendix B. For $q = 3$, the transfer matrix is a 3×3 matrix having three eigenvalues, which in zero external field read $\lambda_1 = 3 \cosh(\beta/2) - \sinh(\beta/2)$, $\lambda_2 = \lambda_3 = 2 \sinh(\beta/2)$, where $\beta = 1/kT$. By retaining the largest eigenvalue λ_1 only, and taking the limit $L \rightarrow \infty$, one successively obtains the free energy per spin

$$F(\beta) = -\frac{\ln(2 + e^\beta)}{\beta}$$

and the specific heat

$$C_v(\beta) = \frac{2\beta^2}{(\sinh \beta/2 - 3 \cosh \beta/2)^2}$$

From there on, the correlation length is then computed using the standard formula [279] $\xi = 1/\ln(\lambda_1/\lambda_2)$, which then yields

$$\xi(\beta) = \left[\ln \frac{3 \coth \beta/2 - 1}{2} \right]^{-1}$$

Finally, the magnetic susceptibility is obtained using the fluctuation-dissipation relation, which gives

$$\chi(\beta) = \frac{8}{27}\beta(1 + 2e^\beta)$$

It is then straightforward to prove that, at the zero-temperature critical point, the ratio γ/ν is given by $\lim_{\beta \rightarrow \infty} \ln \chi(\beta)/\ln \xi(\beta) = \gamma/\nu = 1$.

Numerical signs for exponential divergences In numerical simulations, the correlation length is bounded by the finite lattice size. In the previous calculation, this is equivalent to evaluating the ratio γ/ν at a finite (i.e., positive) temperature, as

5.4. Crossover in the three-state Potts chain: multicanonical simulations vs exact transfer matrix

given by Eq. (B.1). This yields, however, a substantially overestimated result: for instance, I obtain $\gamma/\nu \sim 1.3$ instead of $\gamma/\nu = 1$ for $L = 400$.

The important point is that this picture is perfectly supported by the simulations, with for instance, $\gamma/\nu = 1.02(1)$, $1.14(1)$, and $1.23(1)$ for $\sigma = 1.1$, 1.5 , and 4.0 , respectively. Since the last two values are clearly overestimated (within error bars), this in effect indicates the presence of exponential divergences characteristic of the short-range regime. As a by-product, this also means drastically slow convergence of finite-size scaling.

I would like to mention at this point that for $\sigma = 1.1$, the topological transition at $\sigma = 1.0$ may very well make its influence (moderately) felt, owing to the fact that the finite geometry induces a truncation of the long-range potential at $L/2$ (this point was already addressed in Sec. 4.3.7 in the context of the unusual finite-size effect observed on the line of inverse square interaction). This point is looked upon in greater detail below by means of other indicators.

5.4.2 Pertinent crossover indicators and discussion

Specific heat Investigating the shape of the specific heat turns out to provide the most tractable approach at medium lattice sizes where distinguishing between the short-range and the long-range regime is concerned. This investigation, as will be witnessed, corroborates the previous analysis.

In the thermodynamic limit, the specific heat $C_v(\beta)$ of the nearest-neighbor model admits a maximum $C_v^{max} = 0.7618$ at $kT_m = 0.3767$. It is enlightening to examine the nonmonotonic behavior of this maximum at finite L : this can be carried out by computing $F(\beta, L)$ and then $C_v(\beta, L)$ while retaining all three eigenvalues. Since the calculation is fairly involved, and the final result admits no simple expression, I shall hereafter simply refer to the corresponding curve sketched in Fig. 5.3. Detailed calculations are reported in Appendix B, Eq. (B.2).

When L is increased, the peak of the specific heat first increases to a maximum, and then graphs of C_v collapse and merge gently as the thermodynamic limit is approached. Whenever it is witnessed in graphs obtained from simulation data, this feature thus signals a short-range-like behavior.

Simulations were performed for $1.0 \leq \sigma \leq 4.0$ for various lattice sizes between $L = 50$ and $L = 400$, and I fixed the initial canonical temperature to $kT_0 = 1.0$ so that the maximum of C_v would be clearly visible within the whole range $\sigma \geq 1.0$. C_v was computed using the fluctuation-dissipation relation $C_v = (\langle E^2 \rangle - \langle E \rangle^2)/(kT^2L)$.

As appears obvious from a glance at Fig. 5.3, the cases $\sigma = 1.0$ and $\sigma = 1.1$, on the one hand, and $\sigma \geq 1.2$, on the other hand, display fairly distinct qualitative behaviors:

- For $\sigma = 1.0$, the specific heat reaches its maximum monotonically, at least for the lattice sizes that were investigated. The slowing down in the increase rate as $1/L \rightarrow 0$ allows one to assess a finite maximum in the thermodynamic

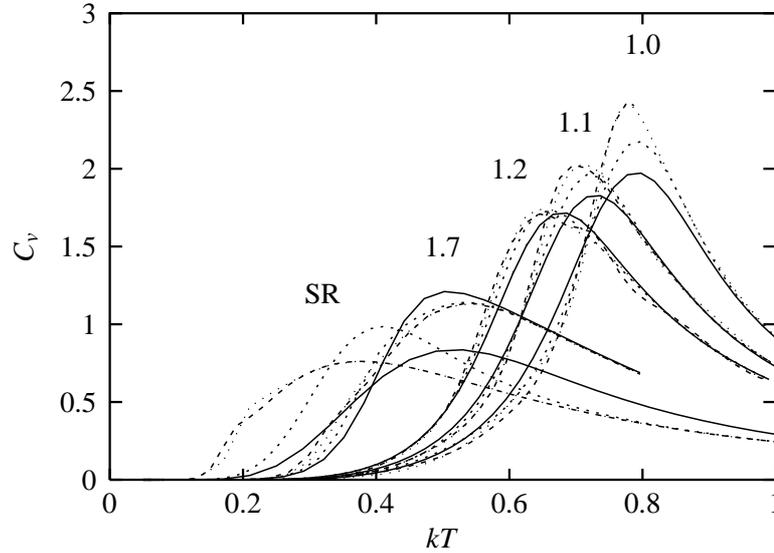


Figure 5.3: Specific heat for various lattice sizes and $\sigma = 1.0, 1.1, 1.2, 1.7$, along with the exact nearest-neighbor case, from right to left. Data for other values of σ have been omitted in order to preserve the clarity of the figure. Solid, dashed, dotted and long-dashed styles refer to $L = 50, 100, 200$, and 400 respectively, except for the short-range case where they refer to $L = 5, 10, 100, 200$.

limit, and this shows that C_v is a nondivergent quantity. Since this is the expected behavior if the line of inverse square interactions is indeed a line of topological transitions (Sec. 4.3.6), this feature brings support to Cardy's scenario. The same behavior is observed for $\sigma = 1.1$, although in a smoother way.

- On the contrary, the qualitative behavior is clearly different for $\sigma \geq 1.2$, where the maximum of C_v first decreases with increasing lattice size, and then quickly reaches a plateau reminiscent of the exact short-range behavior investigated above. While this plateau only slowly reaches the exact short-range value as $\sigma \rightarrow 4.0$ (see Fig. 5.4), it can be concluded, however, that the behavior is already short-range-like.

First, what we observe may boil down to the same finite-size effect as was observed on the $1/r^2$ line at high values of q (Sec. 4.3.7): there, I suggested that the truncation of the long-range potential resulting from the finite lattice size (whatever the periodic boundary conditions) artificially shifted the decay parameter toward the mean-field regime; here, this would correspond to bringing the model closer to the line of topological transitions at $\sigma = 1$, although it is difficult to say in a more quantitative way what is the exact amount of “drift” that the Hamiltonian experiences. Overall, since the truncation is all the more important that small sizes are considered, the whole picture presented here is certainly consistent with the

5.4. Crossover in the three-state Potts chain: multicanonical simulations vs exact transfer matrix

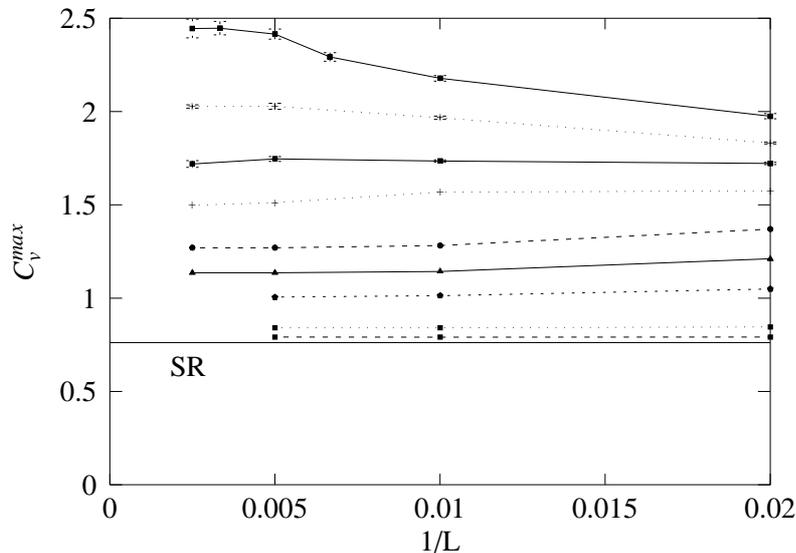


Figure 5.4: Maximum of the specific heat vs inverse lattice size for $\sigma = 1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0, 3.0,$ and 4.0 , from top to bottom. The solid line is a reminder for the (exact) short-range case in the thermodynamic limit. Other lines are guides to the eyes.

fact that the sizes investigated in the present work are rather modest (in comparison with the sizes attainable with the breathing cluster method introduced in Chap. 6).

Second, the renormalization scenario of Sak (and related works, see Sec. 5.3) implies that the crossover actually results from a competition between short- and long-range fixed points. This competition, as seems manifest to me, may not resolve instantly whenever σ crosses the $\sigma = 1.0$ line, and may thus smear the boundary over some finite region. Incidentally, the same kind of effect was observed by Luijten and Blöte in the two-dimensional long-range Ising model [232]. Additionally, it is hard to tell to what extent the subleading short-range interaction (when $\sigma > 1$) influences the finite-size scaling behavior in the interval $[1.0, 1.2]$. Here, it would be interesting to make use of exact finite-size scaling relations in the spirit of the work of Chamati and Dantchev [80], and then to perform extensive checks at large lattice sizes.

Magnetization and Binder cumulant The previous picture can be confirmed by monitoring the behavior of critical temperatures with respect to σ . The magnetization is depicted in Fig. 5.5: graphs merge slightly above $kT = 0$, whenever $\sigma \geq 1.2$, which means that there is no transition at finite temperature. While for $\sigma = 1.1$ there remains some ambiguity due to statistical errors, for $\sigma = 1.0$ the curves now clearly intersect around $kT \sim 0.7$; this, at least, confirms that the behavior is no longer short-range. Incidentally, this temperature is consistent with the value of $0.7089(2)$ obtained by Luijten and Messinfield in [235], yet by monitor-

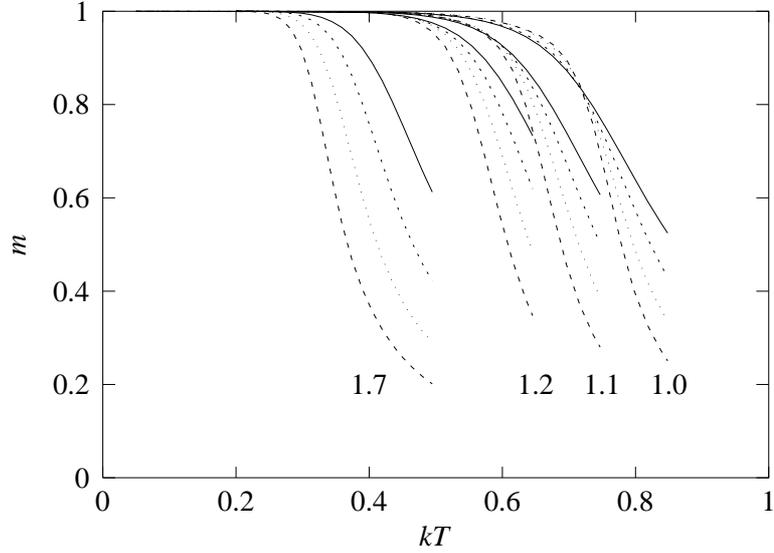


Figure 5.5: Magnetization vs kT for $\sigma = 1.0, 1.1, 1.2,$ and 1.7 from right to left. Solid, dashed, dotted and long-dashed styles refer to $L = 50, 100, 200,$ and 400 respectively.

ing the finite-size scaling behavior of the universal critical jump $\langle m^2 \rangle / kT$ over far larger lattice sizes.

I also computed critical temperatures from the crossing points of Binder cumulants of the magnetization. This yields $\beta_c = 3.3, 6.5,$ and 19 for $\sigma = 1.1, 1.3,$ and 1.5 . Concerning $\sigma = 1.7$ and $\sigma = 2.0$, cumulants no longer cross except at $kT = 0$ within statistical error (the latter case giving β_c between 150 and 200 , yet with excessive dispersion). While the crossover appears to take place in the very vicinity of the $\sigma = 1.0$ boundary, the critical temperature actually dies off quite slowly to 0 as σ increases.

Overall, these results lend support to Sak's scenario for $\sigma > 1.0$: a crossover from long- to short-range behavior occurs whenever $\sigma_{co} = 2 - \eta_{SR}$. Nonetheless, it is worth stressing that I found this crossover to occur within the finite, yet narrow range $1.0 < \sigma < 1.2$, and the pure short-range case to be reached in the limit $\sigma \rightarrow \infty$ only. This is particularly apparent from the behavior of the specific heat.

Whether this results from the truncation of the long-range potential, from non-trivial finite size effects due to the subleading short-range part of the interaction, from a competition between fixed-points that does not resolve instantly as σ crosses the $1/r^2$ line, or (to all probabilities) from a combination thereof, is still an open question.

Conclusion of Part II

In the second part of this thesis, the phase diagram of a long-range Potts chain was extensively studied using a generalized-ensemble algorithm operating in the multicanonical ensemble. In Chap. 3, a single-spin update version of this algorithm was presented, and several improvements tailored to long-range models were proposed. An improved recursion scheme was proposed that is able to efficiently tackle the instability raised by the presence of unequally spaced low energy levels, a peculiarity of long-range spin models. The choice of an efficient predictor and a reliable convergence criterion was discussed, and shown to yield much quicker convergence than with Berg's original algorithm. The utility of a multicanonical approach for the simulation of medium-sized long-range models was demonstrated by tests of performance and accuracy over a large range of decay parameters: the method efficiently circumvents the slowing down traditionally experienced at first-order transitions, delivers an accurate estimate of the density of states over a large energy range, and allows one to monitor thermodynamic quantities over a large range of temperature with strikingly modest numerical effort.

In Chap. 4, the multicanonical algorithm was applied to the estimation of the transition temperatures of a q -state Potts chain with power-law decaying interactions. This study significantly extends the range of available estimates; in the first-order regime of the model, the agreement with mean-field predictions is exceptionally good. In particular, results are perfectly consistent with Tsallis's conjecture $T_c \sim 1/\sigma$ in the limit $\sigma \rightarrow 0$.

In order to locate the boundary separating the first- and second-order regimes, a new method was proposed that detects the order of the phase transition by following the position of spinodal temperatures as the decay parameter of the interaction is varied. The applicability of the method is not restricted to long-range models, and it may represent an efficient, alternative way to other methods, e.g., Lee-Kosterlitz criterion or Binder cumulants of the energy. The multicanonical method was shown to play a pivotal role in this method, owing to its ability to produce accurate estimates of the position of metastable states, as opposed to methods based on multihistogramming. An efficient combination with the transition matrix method was also proposed, yielding even more precise estimates, yet at the expense of an overhead in computation time.

The application of the method to the q -state long-range Potts chain yielded highly precise estimates of the boundary value $\sigma_c(q)$, the accuracy attaining two

Conclusion of Part II

digits. In particular, the value $\sigma_c(3) = 0.72(1)$ is perfectly consistent with the lower bound of 0.7 proposed by Krech and Luijten [213]: however, the precision is markedly higher by an order of magnitude, although the simulations were performed on lattices having fewer than 400 spins. On the line of inverse square interactions, a new, unusual finite-size effects was observed. A detailed finite-size scaling analysis showed that, while the transition seems to be first order at finite lattice size, it becomes continuous in the thermodynamic limit. First, this offers convincing numerical evidence supporting $\sigma_c(q) < 1.0$ for all q , and settles a long-standing controversy surrounding the nature of the phase transition on the $\sigma = 1.0$ line. This result is one of the most surprising findings of this thesis part: it suggests that finite-size scaling at first-order transitions is highly atypical in long-range models. I proposed several interpretations: first, every particle “sees” the boundaries in a long-range system, and the truncation of the long-range potential may artificially pull the decay parameter of the interaction towards the mean-field regime; second, the usual physical meaning attributed to the correlation length in the case of nearest-neighbor models should be markedly challenged; third, the geometry of the interface separating phases in coexistence is certainly far less trivial in long-range models than in their nearest-neighbor counterpart. This suggests several interesting directions for future works, in particular the development of a theory of finite-size scaling at first-order transition for long-range models, that takes metastable states explicitly into account.

Finally, Chap. 5 explored the second-order regime of the model, and the crossover from the long- to the short-range regimes. In the second-order regime, the relation $\eta = 2 - \sigma$, conjectured to be exact for $q = 2$, was shown to yield an increasingly high discrepancy when q is increased, and its validity may just be reinforced in the vicinity of $\sigma = 1.0$. By comparing the behavior of several reweighted thermodynamic quantities with a transfer matrix calculation carried out in the pure short-range case, the crossover between the long-range and short-range regimes was demonstrated to occur inside a narrow window $1.0 < \sigma < 1.2$. This lends strong support to a long-suggested renormalization group scenario. An interpretation of the finite narrowness of the window inside which the crossover takes place was provided in terms of the competition between short-range and long-range fixed points.

To sum up, several controversies were reexamined and successfully settled in the second part of this thesis. Yet this also raised new, exciting questions regarding the finite-size scaling behavior of long-range models at discontinuous transitions. In order to investigate this questions, far larger linear sizes must be attained: this is out of reach of the present multicanonical implementation. It is the goal of the next part of this thesis to introduce a new multicanonical method that efficiently tackles this issue.

Part III

Breathing clusters: a novel approach to simulate long-range models in generalized ensembles

Chapter 6

Limitations of standard multicanonical methods, and beyond

In this and the following chapter, I introduce a new Monte Carlo method, the breathing cluster method, that lumps together a cluster-update scheme and a multicanonical approach. The purpose of this method is twofold:

- first, it is aimed at improving the statistical efficiency of the Markov chain by reducing correlations between successive measurements;
- second, it addresses the problem of scalability in simulations of long-range models, by drastically reducing the algorithm complexity associated with the computation of the lattice energy.

In the following chapter, I will show that this method draws within computation reach one of the challenging issues that was suggested at the end of Chap. 4: the investigation of finite-size effects at first-order transitions in long-range spin models. Yet prior to describing the method, and addressing the last issue, I will examine the main limitations of standard multicanonical methods (i.e., relying on *single-spin* updates), in a general context to begin with (Sec. 6.1), and then in the context of long-range models (Sec. 6.2), and then I will describe the ingredients that are required in order to carry out significant improvements over the existing single-spin-flip scheme. Concurrent approaches will be described in Sec. 6.3, with an emphasis on their limitation as regards long-range models. Finally, the main outlines of the method will be given in Sec. 6.4. The method is then described in deeper detail in the next chapter.

When implemented through single-spin updates (Chap. 3), the multicanonical method is subject to two serious drawbacks. One drawback is tightly related to the dynamic characteristics of move updates based on single-spin flips, and is somewhat independent of the particular model Hamiltonian under study; the other one is

specific to long-range models, and stems from the huge number of interactions to be taken into account when one has to compute the energy of a long-range model.

6.1 Non-optimal dynamic scaling

Let me first examine the dynamic aspect. It was shown by Berg that supercritical slowing down is totally suppressed by a simulation with a multicanonical weight. Specifically, tunneling times are reduced from an exponential-law to a power-law $\tau_e \sim L^z$ of the lattice size, as a result of the ability of the multicanonical weight to assist the Markov chain in jumping across free energy barriers. However, the dynamic exponent z is still substantially higher than the ideal value $z \sim D$ that should be expected from a perfect random walk dynamics [183, 93].

This ideal value can be obtained from the following line of argument. For the sake of simplicity, I consider a Potts model with nearest-neighbor interaction and an interaction constant $J = 1$, on a D -dimensional lattice containing $N = L^D$ spins. Assuming that each single-spin update ideally changes the lattice energy by an amount $\pm\Delta E$, and considering that a lattice sweep (that is, one Monte Carlo sweep) consists of N single-spin updates, one may conclude that the random walk has drifted by about $\sqrt{N}\Delta E$ along the energy axis after one MC sweep. After N lattice sweeps, the drift thus amounts to $N\Delta E$. Up to a multiplicative constant, this is the energy interval between the energy peaks E_o and E_d associated with the ordered and disordered phases (up to yet another constant, this may be the interval between the ground state and the upper energy level). As a consequence, a perfect random walker requires N lattice sweeps to drift from one peak of the energy histogram to the other peak. This means that tunneling times should ideally scale as $N = L^D$, and $z_{ideal} = D$.

For long-range models, it is relevant to mention that the distance $E_d - E_o$ the random walker must travel does not scale linearly with the number of spins. This feature is especially apparent for interactions decaying as a power law, where $E_d - E_o$ grows all the more faster with increasing lattice size that σ comes closer to 0. As a result, the power law $\tau_e \sim L^z$ yields dynamic exponents which are underestimated with respect to the value given by a power law of the form $\tau_e \sim (E_d - E_o)^z$. As a result, the ideal value of z might be lower than D in this case. Where the performance in terms of CPU load is concerned (and in particular if one is interested in how tunneling times grow with the size of the system), I think that the traditional definition $\tau_e \sim L^z$ is more meaningful: the crucial point is the reduction in z that may be brought about when using other move updates.

It is crucial at this point to mention that all indicators of statistical efficiency are subject to the same non-optimal scaling behavior, e.g., effective autocorrelation times [178] and equilibrium times [139]. This is very similar to the behavior observed in canonical simulations, and on exactly the same grounds, it is thus the scalability of the algorithm as a whole which is at stake. I will scrutinize more deeply into this issue in Sec. 7.4.2, and will show that the breathing cluster method

6.2. Algorithm complexity and long-range models

engenders an optimal random walk also with respect to the indicators just mentioned.

It is equally important to mention that correlations between successive measurements not only have an impact on the statistical efficiency of multicanonical production runs, yet also represent a source of systematic error regarding the estimation of the density of states [362].

One of the possible ways to get around these limitations is clearly to consider a combination with a cluster algorithm. As was seen in Chap. 2, these algorithms offer fast decorrelating capabilities, lead to drastically reduced dynamic exponents, and thus improved accuracy for a *given* run length; this pertains to their ability to get rid of spatial correlations by updating spins in a collective way. Yet, when expressed in terms of computer load, the situation is less clear-cut, for cluster algorithms may also pose huge demand in this respect. Overall, the former observation clearly acts as a spur to the development of multicanonical method including cluster updates; the latter, however, implies that care must be taken as regards the efficiency of the implementation, so that the gain in autocorrelation time is not scuppered by the computational effort required to construct the clusters. This is vital indeed in the context of long-range models. In this respect, the major breakthrough that was recently initiated by the introduction of a cluster method whose algorithm complexity scales roughly like the size of the lattice [227] (see, also, Sec. 2.3.2), will play a pivotal role in the design of such a method.

6.2 Algorithm complexity and long-range models

The core issue in the simulation of long-range systems resides in the need to take a huge number of interactions into account when computing the system energy. For potentials restricted to two-body interactions, this task requires the order of N^2 operations, where N is the number of particles in the system.

However, computing the energy of the system is a crucial ingredient in a multicanonical simulation, for the random walk is controlled by a Markov weight which hinges on the energy. On the contrary, canonical cluster algorithms do not require such a step¹. In the context of multicanonical simulations, the multimagnetic ensemble [28] probably represents the only exception to the rule.

When single-spin updates are used, each attempted spin flip actually requires the order of N operations, since it is generally not necessary to recompute the total lattice energy from scratch, yet this also means that a complete lattice sweep requires N^2 operations as whole. Overall, this represented the main impediment to the scalability of the multicanonical implementation used in Chap. 3, where I was in practice limited to lattice sizes of around several hundred spins.

These observations suggest that, in order to improve the scalability of multicanonical algorithms (irrespective of the dynamic characteristics of the algorithm),

¹Except if an estimate of the energy is explicitly needed.

one must find a way to reduce the algorithmic complexity stemming from the computation of the energy. For potentials that are invariants by translation, it turns out, however, that there is a very efficient way of reducing this complexity, which consists in computing the *total* lattice energy through a Fast Fourier Transform (FFT) algorithm, since the complexity of these algorithms scales roughly as $N \ln N$. This approach, which pertains to the use of the convolution theorem (incidentally in widespread use, for nearly the same reasons, in the implementation of fast digital filters in signal processing), was already suggested by Krech et al. [213], yet in the context of *canonical* simulations. In the case of interest here, this would reduce the number of operations required to compute the *total* energy of the lattice to that of a nearest-neighbor model (up to a logarithmic factor).

However, meeting the problem from this new perspective requires in turn that we are able to flip several spins at once: if this is not the case, that is, if we have to compute the energy after each single-spin flip, the total burden for a whole lattice sweep will now amount to roughly $N^2 \ln N$, which proves even worse than our genuine single-spin flip algorithm (which, recalling, scales as N^2). There is, here again, a perfect way to lump all these requirements together in a single method, for *cluster updates* provide what we are seeking: they “know” how to update several spins at once in the most efficient way. If an efficient cluster algorithm can be devised, then the benefit of such a method would be twofold:

- with respect to the computation of the energy, this would cut down the algorithm complexity to that of a short-range model;
- on the dynamic side, one might expect — as a by-product — a gain with regard to the dynamic exponents.

Noteworthy enough, the same $O(N^2)$ complexity *a priori* arises when cluster algorithms are implemented in the context of long-range models, seeing that roughly N^2 bonds have to be considered for activation, and (to crown it all) a vast majority of these bonds only have a negligible probability to be activated. In the framework of *canonical* simulations of long-range models, the problem was elegantly solved by Luijten and Blöte [227], who introduced an efficient algorithm that is able to build clusters with a CPU load that is roughly independent of the number of interactions per spin.

To recap, we have two ingredients in hand which, once combined, may cut down the burden of computing the energy to that of a short-range model. It turns out, however, that there is yet another item of business to be carried out, and this one is perhaps the central subject of this thesis part: how can we *technically* include cluster-updates in a multicanonical algorithm? Broadly speaking, the issue amounts to plugging two schemes into one another, which are seemingly incompatible:

- on the one hand, cluster algorithms rely on a (canonical) *temperature* in order to create blocks of spins with the correct bond probability, i.e., one that

6.3. The multibond algorithm and other concurrent approaches

ensures not only that detailed balance holds, yet also that the efficiency of the algorithm is maximized (this implies, in general, that the critical point and the percolation threshold coincide);

- on the other hand, multicanonical algorithms rely on the *density of states* to meet their goals, namely, producing a flat histogram; the multicanonical weight $w(E)$ has a fundamentally *non-local* nature, that loses track of the symmetries of the model Hamiltonian; and yet these symmetries are crucial to the *technical* implementation of a cluster algorithm.

It is the purpose of the next chapter to develop such a program, in an efficient way (in particular where *ease of implementation* is concerned) and with wide applicability in mind (i.e., we would not like to restrict the method to a *peculiar* multicanonical scheme).

Prior to going over this subject, however, I will briefly review a method proposed in the mid-90's by Janke and Kappler [182]: the *multibond* method combines — from a clearly different perspective — cluster updates and generalized ensembles weighting; I will explain why this approach is not suited for long-range models.

6.3 The multibond algorithm and other concurrent approaches

As regards short-range spin models, the issue presented in the previous paragraph was met by the multibond algorithm [182, 183, 178] (or variants thereof, [353], including a combination with the transition matrix method [354]) and, although with a somewhat different approach, by the hybrid daemon-cluster approach [291]. The common denominator to both methods is the fact that they combine a *generalized ensemble* weighting of the Markov chain with a global, i.e., cluster-like update scheme, and in effect yield nearly ideal tunneling-time dynamic exponents. These two methods, however, do not lend themselves to an efficient implementation in the case of long-range interactions, since in particular they preclude the use of the previously mentioned optimizations dedicated to long-range models.

The hybrid daemon-cluster algorithm [291] comprises two sub-algorithms: a microcanonical cluster [87, 86], which simulates the system at a constant energy using demons (these act as energy buffers, allowing the energy to slightly fluctuate); and a so-called multicanonical demon refresh, which moves the system from one energy level to another by “refreshing” the demons. One the main impediment to the application of this method to long-range model is the very presence of a long-range potential: the hybrid approach in [291] relies explicitly on the presence of *equally spaced* (discrete) energy levels to connect the microcanonical daemon and the multicanonical heat bath, while energy levels are unequally spaced in long-range models (Chap. 3). In addition, assuming that a generalization

of this approach to long-range models *might* be tractable, no microcanonical cluster algorithm exists for these potentials, and in this respect an extension of Creutz's algorithm [86] is not trivial. Overall, the method introduced in the next chapter will prove much more straightforward in its implementation, by relying on a single update scheme.

The multibond method simulates the (nearest-neighbor) Potts model in a so-called multibondic ensemble, where the role of the energy variable E is taken over by that of the number of active bonds B . To be specific, the method resorts to a *modified* partition function of the model in its spin-bond representation (see Sec. 2.3.3),

$$Z_{FK} = \sum_{[\sigma]} \sum_{[b]} \prod_{\langle i,j \rangle} (p \delta_{\sigma_i, \sigma_j} \delta_{b_{ij}, 1} + \delta_{b_{ij}, 0}) w_{mubo}(B)$$

where $\langle i, j \rangle$ means that only pairs of nearest-neighbor spins are considered, $B = \sum_{\langle i,j \rangle} b_{ij}$ is the total number of active bonds, $p = e^{\beta c} - 1$ is the bond probability (which, in the case of a nearest-neighbor model, is unique), and $w_{mubo}(B)$ is a so-called multibond weight which aims to enhance rare events. In this context, rare events correspond to *spin-bond* configurations which are strongly suppressed, e.g., by a free energy barrier. The prescription of the authors is to set β_c to the transition temperature (which, incidentally, requires that this last quantity be known in advance, even to a moderate accuracy).

The rationale behind this approach is the equivalence between the average energy and the average number of active bonds, i.e., $-\langle E \rangle = (1 + 1/p) \langle B \rangle$. This can be readily derived by differentiating the (logarithm of the) partition function of the model in its pure-bond representation, i.e., Z_{RC} (Eq. 2.9), with respect to β , and equating the result with $-\langle E \rangle$. Then it is assumed that E and B roughly satisfy the same proportionality relation as $\langle E \rangle$ and $\langle B \rangle$, i.e., up to thermal fluctuations in the number of bonds for a given energy (and vice-versa). From there on, one possible implementation of the method consists in estimating the multicanonical weight $w(E)$, and then obtaining $w_{mubo}(B)$ from $-E \sim (1 + 1/p)B$. Of course, this requires knowing $w(E)$ beforehand, but the authors in [182] argue that this can be carried out by rescaling $w(E)$ from an estimate obtained at a lower size². This also requires that the proportionality relation between E and B holds to a sufficient extent, so that a *flat* histogram in the *bond* landscape indeed yields an equally flat histogram in the energy landscape. In this respect, it is shown that this is the case, at least for the sizes investigated by the author.

A variant was recently proposed by Yamaguchi and Kawashima [353] based on the Wang-Landau algorithm, whereby the method directly estimates the density of states in the bond space, $n(B)$, and a correspondence between $n(B)$ and $n(E)$ is derived analytically, so that thermodynamic averages (with respect to E) may be obtained in the first place. This correspondence is intractable for long-range potentials. Turning back to the original multibond approach, an extension to long-range spin models would involve, however, mapping the spin model to a spin-bond representation, with bond probabilities p_{ij} now depending on the interaction strength

²For long-range models, this *may* be tractable for σ large enough.

6.4. A novel approach: breathing clusters and the microcanonical temperature

J_{ij} between spins i and j : the proportionality relation between the average energy $\langle E \rangle$ and the average number of bonds $\langle B \rangle$ does no longer hold. Instead, it has to be replaced with a proportionality relation involving $\langle E \rangle$ and a linear combination of $\langle b_{ij} \rangle$ coefficients, where $b_{ij} = 1$ whenever the bond between spins i and j is active, and $b_{ij} = 0$ otherwise. From a technical point of view, this approach is somewhat intractable as soon as the number of interactions between spins becomes large. As a result, it does not prove efficient in the long-range case either.

In addition, there are two further impediments to the usability of the method in the context of long-range models. First, analyzing first-order transitions hinges on computing energy-related quantities, e.g., specific heats, Binder cumulants or interfacial free energies, and we are still left with the $O(N^2)$ issue mentioned in the previous section. Second, there is an additional $O(N^2)$ cost related to the cluster building operations: there are N^2 possible bonds to test for activation, and furthermore most of these bonds only have a negligible probability of being activated. The mechanism used in the multibondic method to build the cluster step by step, i.e., broadly speaking, by computing the probability to activate a new bond from the *current* number of active bonds, renders the implementation of Luijten-Blöte cluster algorithm clearly intractable (in particular where resorting to cumulative probabilities is concerned).

Finally, let me mention another approach based on multigrid methods [133]. These methods update blocks of spins on different length scales according to a given sequence (rather straightforward in the *unigrid* approach; possibly involving a renormalization scheme in the *true* multigrid approach). They have proven remarkably efficient at reducing dynamic exponents to a nearly zero value [199]. A combination with multicanonical weighting was presented by Janke and Sauer in [186, 178] for the simulation of the ϕ^4 field theory, and indeed multigrids methods have focused (thus far) on systems with continuous degrees of freedom, e.g., lattice field theories. In this respect, a connection with the Path Integral Monte Carlo was also considered [187]. These methods will thus not be considered here.

6.4 A novel approach: breathing clusters and the microcanonical temperature

The key point of the method introduced in the next chapter is the control of the cluster construction process by a so-called instantaneous temperature which depends on the lattice energy. This quantity is estimated in a such way that, at each energy level visited by the multicanonical random walk, clusters are built as if the system were *canonically* simulated at this temperature: I show that the microcanonical temperature is the best candidate in this respect, in that it guarantees the highest decorrelating performance.

The choice of the microcanonical temperature as the temperature that drives the cluster construction results in cluster bond probabilities changing continuously as the available energy range is walked along by the random-walker: small clusters are

Chapter 6. Limitations of standard multicanonical methods, and beyond

built in the upper energy range (where the microcanonical temperature is large and the bond probability is low), and conversely large clusters are grown in the lower energy range. From a physical point of view, the average cluster size oscillates at the rhythm of the random walker, evoking the picture of “breathing” clusters.

With regards to long-range models, and more generally spin models having any number of interaction, I will explain how the method straightforwardly combines with the efficient Luijten-Blöte cluster construction algorithm, and yield a reduction of the algorithm complexity to that of a short-range model with the same number of spins.

This new method is general and versatile in the sense that it can host a variety of modern multicanonical algorithms dedicated to the estimation of the density of states, whether iterative or not. In particular, I will consider its combination with the recently proposed Wang-Landau method, which has proven to markedly overcome the drawbacks encountered with Berg’s approach in terms of robustness and scalability. I will also address its integration into the transition matrix method, which turns out to provide, at the expense of a slight overhead, a very efficient way of estimating the information required to properly steer the cluster construction.

The superiority of the method over concurrent multicanonical schemes will be illustrated on q -state long-range Potts chains with q ranging from 3 to 12. I will investigate chains containing up to 2^{16} spins, otherwise largely out of reach of other multicanonical methods, and show that this draws precise tests of finite-size scaling within computation range (Sec. 7.6). These tests will be enlightening as regards the finite-size scaling behavior of long-range models at first-order phase transitions (this was one of the developments suggested at the end of Chap. 4).

The emphasis will be clearly given to indicators of performance, including indicators that were introduced only very recently, e.g., the estimation of equilibration times from χ^2 regression [139]. I feel strongly, indeed, that a new Monte Carlo method may demonstrate its superiority only after being intensively pitted against the large variety of performance indicators available to date, notwithstanding its ease of implementation. I will show that, in terms of computer load, the method already outperforms conventional local-update algorithms for sizes above a hundred spins in one-dimensional models

The next chapter contains a version of an article that was recently submitted, and describes the method and the results in a very detailed way.

Chapter 7

Fast Flat-Histogram Method for Generalized Spin Models

With minor changes, this chapter reproduces a preprint submitted for publication in Phys. Rev. E. In Sec.7.6.2, in particular, a paragraph dedicated to the finite-size scaling behavior of the surface tension was added.

[cond-mat/0504367]

Abstract

We present a Monte Carlo method that efficiently computes the density of states for spin models having any number of interaction per spin. By combining a random-walk in the energy space with collective updates controlled by the microcanonical temperature, our method yields dynamic exponents close to their ideal random-walk values, reduced equilibrium times, and very low statistical error on the density of states. The method can host any density of states estimation scheme, including the Wang-Landau algorithm and the transition matrix method. Our approach proves remarkably powerful in the numerical study of models governed by long-range interactions, where it is shown to reduce the algorithm complexity to that of a short-range model with the same number of spins. We apply the method to the q -state Potts chains ($3 \leq q \leq 12$) with power-law decaying interactions in their first-order regime; we find that conventional local-update algorithms are outperformed already for sizes above a few hundred spins. By considering chains containing up to 2^{16} spins, which we simulated in fairly reasonable time, we obtain estimates of transition temperatures correct to five-figure accuracy. Finally, we propose several efficient schemes aimed at estimating the microcanonical temperature.

7.1 Introduction

Long-range spin models have drawn increasing interest in the last decade, both in the microscopic modeling of a variety of systems ranging from model alloys [125] to spin glasses [121] to neural networks [10], and as powerful laboratory frame to investigate fundamental issues in the physics of critical phenomena. These include, e.g., the effect of dimensionality [225], the crossover from short-range to

long-range behavior [315, 232, 284], mean-field driven phase transitions [42], and possible connections with Tsallis's non-extensive thermodynamics [68, 332, 67]. Monte Carlo (MC) methods have now gained a prominent role in the investigation of phase transitions in these models [227, 324, 19, 39, 41, 40, 326]. In particular, a major breakthrough was recently initiated by the introduction of a (canonical) cluster algorithm able to overcome the algorithm complexity inherent to long-range (LR) models, namely, the need to take a huge number of interactions into account at each Monte Carlo step (MCS) [227]. In a recent article, we proposed a generalization of this algorithm to simulations in the multicanonical ensemble [285]. It is the goal of the present work to introduce a general and versatile method aimed at embedding any cluster update scheme in a flat histogram algorithm, with special emphasis given to LR spin models.

Whether short- or long-range interactions are considered, canonical MC simulations of long-range spin models suffer indeed from severe shortcomings, the use of cluster updates notwithstanding. First and foremost, models exhibiting first order phase transitions or complicated energy landscapes experience supercritical slowing down [31]: the time needed for the dynamics to tunnel through free energy barriers grows exponentially with the lattice size, leading to quasi ergodicity breaking and unreliable statistics. Second, the computation of free energies and related thermodynamic quantities is highly involved, and a precise determination of the order of the transition is often intractable. In practice, these shortcomings preclude the use of canonical MC algorithms at first-order transitions except at modest lattice sizes and in the case of weakly first-order transitions.

An efficient approach aimed at overcoming this limitation is the simulation in generalized ensembles [166, 25], in particular its multicanonical flavor initially proposed by Berg [30, 31], reconsidered in the context of transition matrix dynamics [341, 301] and recently revisited by Wang and Landau [335, 334]. The key-idea here is to artificially enhance rare events corresponding to local maxima in the free energy, by feeding the Markov chain with an appropriate distribution $w(E)$. In the *multicanonical ensemble*, $w(E)$ is set to the inverse of the density of states, so that the resulting dynamics is a random walk in the energy space that yields a flat histogram of the energy. Other ensembles have been proposed in the last decade, including the $1/k$ ensemble, which enhances low-energy states [156], and very recently, the optimal ensemble, which aims to optimize the distribution $w(E)$ with respect to the local diffusivity of the random walker, so that tunneling times are minimized [319, 351]. While still broad, histograms engendered by these last ensembles are no longer flat; in the optimal ensemble for instance, the histogram is slightly peaked around the critical region, so that the larger time spent by the random walker inside the critical region makes up for the lower diffusivity in this region.

When implemented through local (i.e., single-spin) updates [284], simulations in the multicanonical ensemble suffer, however, from two serious hurdles. First, while tunneling times — measured in Monte Carlo steps (MCS) — are reduced from an exponential to a power law $\tau \sim L^z$ of the lattice size, the dynamic ex-

7.1. Introduction

ponents z are still substantially higher than the ideal value $z \sim D$ that should be expected from the dynamics of a random walker [183, 93]. This observation, as we will witness in this article, applies equally well to effective autocorrelation times and to equilibrium times; this represents a serious hindrance in terms of scalability, in particular whenever a higher precision is desired and large amounts of decorrelated data need to be gathered. In this respect, it is important to mention that correlations between successive measurements do not only have an impact on the statistical efficiency of multicanonical production runs, yet also represent a source of systematic error regarding the estimation of the density of states [362]. A second impediment to the scalability of local-update implementations specifically relates to long-range models. Here, the very presence of long-range interactions makes the computation of the energy — an essential ingredient of multicanonical methods — a very time consuming operation, namely, one associated with an $O(L^{2D})$ algorithm complexity. As a result, the demand on CPU time needed to generate perfectly decorrelated statistics grows as L^{z+2D} , with $z > D$.

In this article, we present a Monte Carlo method which successfully tackles these issues by performing simulations in the multicanonical ensemble using collective updates. Our method combines the fast-decorrelating capabilities of cluster algorithms with the versatility of flat-histogram methods in an efficient and straightforward way, and with wide applicability in view. In particular, it can be readily combined with any iteration scheme dedicated to the estimation of the density of states, e.g., Wang-Landau’s method [335] or transition matrix algorithms [341]. Additionally, while our method is presented here in the context of long-range spin models, where it gives drastic improvements over commonly used methods, it is perfectly applicable to any class of models for which a cluster algorithm exists in the canonical ensemble.

Noteworthy enough, embedding a collective update scheme in a multicanonical algorithm is not straightforward, however, due to the fundamentally *non-local* nature of the multicanonical weight $w(E)$. Indeed, cluster algorithms depend heavily upon particular symmetries of the model Hamiltonian, which $w(E)$ does not keep track of; in particular, there is no longer a canonical temperature. With simulations of spin models with nearest-neighbors interactions in view, several attempts have been made at combining cluster updates with multicanonical methods in some way or another during the last decade: the multibond algorithm [182, 183, 178, 73] or variants thereof targeting Wang-Landau’s algorithm [353, 351] simulate the model in its spin-bond representation; Rummukainen’s hybrid-like two-step algorithm lumps together a microcanonical cluster algorithm and a multicanonical daemon refresh [291]. As opposed to these, however, our method relies on a cluster-building process which simply depends on the microcanonical temperature of the current configuration — a quantity that may be readily derived from the estimated density of states — in order to determine appropriate bond probabilities. In particular, it does not require prior knowledge of the transition temperature, as is the case in the multibond method. We further show that our approach makes it particularly straightforward to incorporate two optimization schemes dedicated to LR models

[227, 213], which cut down the algorithm complexity from $O(L^{2D})$ to $O(L^D \ln L^D)$. As a result, the total demand on CPU time with respect to uncorrelated data is reduced to approximately $L^{2D} \ln L^D$, since cluster updates also lower z to around D ; where LR models are concerned, the benefit of cluster updates is thus clearly twofold. Let us also mention that, as a by-product, using cluster updates provides improved estimators for the statistical moments of the order parameter [325] and for spin-spin correlation functions; for instance, the last quantity can be better estimated by counting the fraction of time two given sites belong to the same cluster [261, 262]. Further interesting information, including information connected with fractal geometry, may also be gleaned from cluster statistics [13, 189].

Overall, the sharp reduction of the computer load brought about by our method allowed us to study q -state Potts chains with $1/r^{1+\sigma}$ interactions containing up to 2^{16} spins in a few days on a modern Intel-based computer. It must be noted that, with standard multicanonical methods based on single-spin updates, such huge sizes are simply intractable, since the largest size of 2^{16} investigated in this work would demand several months of computation. As regards dynamic performance, we obtain a substantial reduction in the dynamic exponent, from e.g., $z \sim 1.35(3)$ to $z \sim 1.05(1)$ for $q = 6$ and $\sigma = 0.7$. We also show that our method produces faster equilibration, lower effective autocorrelation times, and — where implementations based on the Wang-Landau algorithm are concerned — lower statistical errors on the estimate of the density of states, e.g., of nearly an order of magnitude for $q = 6$, $\sigma = 0.9$ and $L = 512$ spins. As a result, we obtain estimates of transition temperatures that have a noticeably higher precision than those obtained using local updates [284] or standard canonical methods [324, 19]. Finally, in order to check that our method did not produce systematic errors, we performed several simulations of the two-dimensional seven- and ten-state Potts models with nearest-neighbor interactions and sizes up to $L = 256 \times 256$. We obtain dynamic exponents close to the ideal random-walk value $z \sim 2$. Although computed from rather modest statistics, our estimate of the interfacial free energy for the largest size reaches a precision of four digits. In this respect, our method compares perfectly well with other methods operating in the multicanonical ensemble, and represents an alternative way for short-range spin models.

The layout of this article is as follows. In Sec. 7.2, after briefly reviewing some prominent features of multicanonical methods, we explain how we combine a multicanonical weighting with collective updates, with special emphasis given to the detailed balance equation. Section 7.3 addresses optimizations dedicated to long-range models. Numerical results regarding the dynamic characteristics of our method are presented in Sec. 7.4. In Sec. 7.5, we compare our results for the two-dimensional Potts model with nearest-neighbor interactions, with exactly known results, and section 7.6 is devoted to the investigation of the precision of our method in the context of the long-range Potts chain with power-law decaying interactions. Overall, we pay particular attention to comparison with other algorithms operating in the multicanonical ensemble, especially in terms of tunneling rates, dynamic exponents and estimates of thermodynamic averages. Finally, we discuss several

7.2. A method to embed cluster updates in a flat histogram algorithm

procedures aimed at estimating the microcanonical temperature, and in particular, how we can efficiently combine our method with the transition matrix approach.

7.2 A method to embed cluster updates in a flat histogram algorithm

Monte Carlo simulations are based on the generation of a Markov chain of configurations $\{\sigma_i\}$, where each configuration is assigned a weight $w[E(\sigma_i)]$ corresponding to the probability distribution one wishes to sample. In canonical simulations, i.e., carried out at a fixed inverse temperature β , one chooses a Boltzmann weight $w[E(\sigma_i)] = e^{-\beta E(\sigma_i)}$, thus thermodynamic averages are straightforwardly obtained by computing the appropriate moments of the data accumulated at the given temperature. On the other hand, reweighting methods based on multihistogramming [114] are hampered at large lattice sizes by the narrowness of the energy window that is sampled, let alone additional supercritical slowing down. In the multicanonical ensemble, one allows the dynamics to jump across free energy barriers and, from a more general standpoint, to sample wide energy windows, by producing a flat energy distribution over the energy range of interest for the problem at hand. This is formally carried out by setting $w(E) = e^{-S(E)} \propto 1/n(E)$, where $n(E)$ is the density of states and $S(E)$ is the microcanonical entropy. This in effect leads to $N(E) \propto n(E)w(E) = \text{const.}$ for the number of visits to energy E . Since the density of states is a priori unknown, $w(E)$ is estimated using an iterative procedure initially fed from, e.g., a canonical guess $w(E) = e^{-\beta_0 E}$ at a carefully chosen inverse temperature β_0 , a flat guess $w(E) = 1$, or — whenever feasible — a properly scaled estimate obtained at a smaller lattice size. Thermodynamic quantities that depend solely on the energy, like the specific heat or Binder cumulants, can then be computed directly from the estimated density of states. Other quantities, e.g., those depending on the order parameter, are obtained through a reweighting procedure based on data gathered during an additional production run.

Historically, Berg's recursion scheme [23, 24] was the first iteration procedure specifically dedicated to multicanonical simulations. It consists in accumulating histogram entries of the energy during each iteration run, and updating $w(E)$ from the histogram of the energy obtained in a previous iteration run, until eventually the histogram becomes flat up to a given tolerance. Entropic sampling [218] more or less boils down to the same key principle. Both methods suffer, however, from poor scalability. Looking at this issue from a slightly different angle, the recently proposed Wang-Landau acceleration method [335, 334] updates $w(E)$ in real-time during the course of the simulation, performing independent random walks in distinct energy ranges. Since modifying the weight of the Markov chain during a simulation breaks detailed balance, the amount by which $w(E)$ is modified during a given iteration is decreased from one iteration to the other until it reaches a negligible value. As a result, detailed balance is approximately restored in the last step of the iteration scheme. In this regard, an original approach aimed at reducing

the statistical error on the estimate of the density of states was recently proposed by Yan and de Pablo [358], whereby the density of states is obtained by integrating an instantaneous temperature computed from configurational information or from a so-called multimicrocanonical ensemble. Finally, a large class of iteration schemes have been proposed that are based on matrices of transition probabilities [301, 81, 341, 337, 339] or a combination thereof with Wang-Landau's algorithm [299, 304]. Here, the density of states is computed through a so-called broad histogram equation involving infinite temperature transition matrices, where transition matrices keep track of the microcanonical average of the number of potential moves from one energy levels to another (Sec. 7.7 gives more details on how our method can efficiently capitalize on transition matrices). Historically, procedures based on transition matrices were termed *flat histogram methods* in order to distinguish them from Berg's multicanonical method, although both approaches in effect yield a flat, broad histogram. To sum up, the main benefit of multicanonical methods is twofold: first, a wide energy range is sampled, irrespective of the presence of free energy barriers; second, the methods yield a direct estimate of the density of states.

A local-update implementation of a multicanonical algorithm may consist in updating a single spin at a time and accepting the attempted move from state a to state b with a probability given by $W(a \rightarrow b) = \min[1, e^{S(E_a)-S(E_b)}]$. We now show that the microcanonical temperature $\beta(E)$ defined as $dS(E)/dE$ is a relevant quantity for the acceptance rate of this process. Denoting $E_b = E_a + \epsilon$, we expand the probability $W(a \rightarrow b)$ for small ϵ , and obtain $W(a \rightarrow b) \sim \min[1, e^{-\beta(E_a)\epsilon}]$. This shows that, for small enough energy changes, the dynamics is equivalent to that of a canonical simulation at an inverse temperature $\beta(E)$. Our departure point for a collective-update implementation in the multicanonical ensemble is thus to build clusters of spins with the same bond probabilities as would be given by a canonical simulation at inverse temperature $\beta(E)$.

Although our algorithm may be equally well applied to other spin models, e.g., models incorporating disorder or exhibiting a continuous symmetry, we now consider, for the sake of clarity, a generalized ferromagnetic spin model with a \mathbb{Z}_q symmetry, whose Hamiltonian reads $H = -\sum_{i<j} J_{ij}\delta_{\sigma_i,\sigma_j}$. Here $J_{ij} > 0$ and the σ_i variables can take on integer values between 1 and q . Taking guidance from Swendsen-Wang's cluster algorithm [309], we start from an empty bond set, consider each pair of spins $\{\sigma_i, \sigma_j\}$ in turn, and activate a bond between them with a bond probability given by $\pi_{ij}(E_a) = \delta_{\sigma_i,\sigma_j} [1 - e^{-J_{ij}\beta(E_a)}]$, where E_a is the current lattice energy and $\beta(E_a)$ the inverse microcanonical temperature at energy E_a . Efficient ways of estimating $\beta(E)$ are considered later on in Sec. 7.7. Then, we identify clusters of connected spins using, e.g., a multiple-labeling scheme [161], draw a new spin value at random for each cluster, and accept the attempted move with an acceptance probability $A(a \rightarrow b)$ which ensures that detailed balance is satisfied. The derivation of this probability may be carried out in the following way. First, the total acceptance probability $W(a \rightarrow b)$, i.e., the quantity that enters

7.2. A method to embed cluster updates in a flat histogram algorithm

detailed balance in such a way that $e^{-S(E_a)}W(a \rightarrow b) = e^{-S(E_b)}W(b \rightarrow a)$, is split into two terms $P(a \rightarrow b)$ and $A_{\text{flip}}(a \rightarrow b)$ representing a *proposed update probability* and an *acceptance probability* for the proposed update, respectively. It is straightforward to show that the choice $A_{\text{flip}}(a \rightarrow b) = \min \left[1, \frac{P(b \rightarrow a)}{P(a \rightarrow b)} e^{S(E_a) - S(E_b)} \right]$ satisfies the detailed balance equation. Let us denote \mathcal{B} the set of active bonds over the complete graph \mathcal{G} engendered by all possible interactions: the *proposed update probability* is given by the probability to construct a given set \mathcal{B} from an empty bond set, i.e.,

$$P(a \rightarrow b) = \prod_{b_{ij} \in \mathcal{B}} \pi_{ij}(E_a) \prod_{b_{ij} \in \mathcal{G} \setminus \mathcal{B}} [1 - \pi_{ij}(E_a)].$$

After simplification, we obtain for $\frac{P(b \rightarrow a)}{P(a \rightarrow b)}$,

$$e^{\beta(E_b)E_b - \beta(E_a)E_a} \prod_{b_{ij} \in \mathcal{B}} \frac{e^{J_{ij}\beta(E_b)} - 1}{e^{J_{ij}\beta(E_a)} - 1};$$

whence

$$A_{\text{flip}}(a \rightarrow b) = \min \left[1, \frac{e^{\alpha(E_a)}}{e^{\alpha(E_b)}} \prod_{b_{ij} \in \mathcal{B}} \frac{p_{ij}(E_b)}{p_{ij}(E_a)} \right], \quad (7.1)$$

where $\alpha(E) = S(E) - \beta(E)E$ and $p_{ij}(E) = e^{J_{ij}\beta(E)} - 1$. This expression can be further simplified if we consider long-range models whose coupling constant depends only on the distance between spins, i.e., $J_{ij} = J(l)$, where $l = \text{dist}(i, j)$. We have for $A_{\text{flip}}(a \rightarrow b)$:

$$A_{\text{flip}}(a \rightarrow b) = \min \left[1, \frac{e^{\alpha(E_a)}}{e^{\alpha(E_b)}} \prod_{l>0} \left[\frac{p_l(E_b)}{p_l(E_a)} \right]^{B(l)} \right], \quad (7.2)$$

where $B(l)$ stands for the number of bonds of length l . It is worthy of mention that, if one looks at this equation from the standpoint of canonical simulations at inverse temperature β_0 , we have $w(E) = e^{-\beta_0 E}$; whence $\beta(E) = \beta_0$ and $\alpha(E)$ does no longer depend on E . As a result, the acceptance rate $A_{\text{flip}}(a \rightarrow b)$ is equal to 1 and we are back to the original Swendsen-Wang algorithm.

It is also crucial to underline that it is the microcanonical temperature, i.e., the lattice energy in the first place, which entirely governs the construction of clusters; indeed, for a given lattice configuration at energy E , bonds are placed as if the model were simulated at its microcanonical temperature using a Swendsen-Wang algorithm. As a result, cluster bond probabilities change continuously as the lattice configuration walks along the available energy range of the random walk, so that, e.g., small clusters are built in the upper energy range and conversely large clusters in the lower energy range.

7.3 Optimization for long-range models

7.3.1 Computing the lattice energy through FFT acceleration

As is apparent in Eq. (7.1), determining the acceptance rate of a cluster flip demands that we compute the energy of the new (attempted) lattice configuration, which for long-range models is an $O(L^{2D})$ operation. This is similar to the local-update case, where performing one MC step, i.e., updating L^D spins subsequently, takes a CPU time proportional to the square of the number of spins, seeing that L^D operations are needed after each single spin update to compute the new partial energy between the updated spin and the rest of the lattice. Recently, Krech and Luijten proposed an algorithm that is able to compute the energy of a model governed by translation invariant interactions in $O(L^D \ln L^D)$ operations [213]. This method leans on the convolution theorem and the Fast Fourier Transform (FFT), for which numerous efficient radix-based implementations are available. As a result, updating the lattice configuration *globally* rather than a single spin at a time allows us to cut the $O(L^{2D})$ complexity down to an $O(L^D \ln L^D)$ one. A crucial point to be noted here is that this reduction is absolutely intractable with single-spin updates, owing to the very reason that the energy would have to be computed again after each single-spin update; this requires L^D operations, and an FFT algorithm would output no gain at all.

Let us assume that we can write down the model Hamiltonian as a sum of dot products, i.e., $H = -\frac{1}{2} \sum_{i \neq j} J_{ij} \vec{S}(i) \cdot \vec{S}(j)$, with J_{ij} invariant by translation. This is straightforwardly done when $q = 2$, since in this case the dot product reduces to a product of scalar Ising spins. As we will witness in a moment, the presence of a delta Kronecker symbol in the Hamiltonian whenever $q > 2$ requires, however, a minor transformation of the Hamiltonian. For simplicity, we consider hereafter a one-dimensional lattice with an interaction $J(l)$ depending on the distance l between spins. The line argument is similar in higher dimensions, with the sole exception that multidimensional Fourier transforms are then performed. The Discrete Fourier Transform (DFT) of the spin sequence $\{\vec{S}(l)\}_{l=1 \dots L}$ reads

$$\tilde{\vec{S}}(k) = \sum_{l=0}^{l=L-1} \vec{S}(l) e^{-i2\pi kl/L},$$

and reciprocally,

$$\vec{S}(l) = \frac{1}{L} \sum_{k=0}^{k=L-1} \tilde{\vec{S}}(k) e^{i2\pi kl/L}.$$

Similarly, we define the DFT of the sequence of coupling constants $\{J(l)\}$ as

$$\tilde{J}(k) = \sum_{l=0}^{l=L-1} J_{pbc}(l) e^{-i2\pi kl/L},$$

where $J_{pbc}(l)$ incorporates the effect of Infinite Image Periodic Boundary Conditions (IIPBC) [67], that is, $J_{pbc}(l) = \sum_{m=-\infty}^{+\infty} J(l + mL)$; for algebraically decaying

7.3. Optimization for long-range models

interactions, this sum can be exactly expressed in terms of Hurwitz functions [284]. We diagonalize the original Hamiltonian H by rewriting it in terms of the $\tilde{J}(k)$ and $\tilde{S}(k)$,

$$H = -\frac{1}{2L} \sum_{k=0}^{k=L-1} \tilde{J}(k) \tilde{S}(k) \cdot \tilde{S}(-k),$$

where it should be emphasized that $\tilde{S}(-k)$ and $\tilde{S}(k)$ are complex conjugates, since the original vectors $\vec{S}(l)$ have real coordinates. By relying on an FFT radix-2 algorithm, the task of computing the lattice energy is consequently reduced to $O(L \ln L)$ operations.

For $q > 2$, the Kronecker delta symbol in the Hamiltonian unfortunately rules out the previous diagonalization. One way to resolve this issue is to map the q -state Potts model onto a $(q - 1)$ -dimensional vector model, so that the Kronecker delta function in the original Hamiltonian is turned into a dot product. We define a one-to-one mapping between each Potts spin $\sigma = 1 \dots q$ and a unit-length vector $\vec{S}^{(\sigma)}$ belonging to a $(q - 1)$ -dimensional hypersphere, so that $\vec{S}^{(\sigma)} \cdot \vec{S}^{(\sigma')} = \frac{q\delta_{\sigma,\sigma'} - 1}{q-1}$. It is straightforward to prove that $\sum_{\sigma} \vec{S}^{(\sigma)} = 0$, and that

$$H = \frac{q-1}{2q} \sum_{i \neq j} J(i-j) \vec{S}^{(\sigma_i)} \cdot \vec{S}^{(\sigma_j)} + \frac{1}{q} \sum_{i < j} J(i-j).$$

In the case of the three-state model, this transformation is equivalent to mapping Potts variables onto the complex plane, i.e., $\sigma \rightarrow S^{(\sigma)} = e^{i2\pi(\sigma-1)/3}$, and writing the dot product $\vec{S}^{(\sigma_i)} \cdot \vec{S}^{(\sigma_j)}$ as $\text{Re}\{S^{(\sigma_i)} S^{(\sigma_j)*}\}$. In this case, the term $\tilde{S}(k) \cdot \tilde{S}(-k)$ becomes $|S(k)|^2$, where $S(k)$ is the DFT of the sequence of (complex) variables $\{S^{(\sigma)}\}$. This reduces by one the number of $O(L)$ operations required, since computing a dot product is no longer required.

For $q > 3$, spin vectors on the $(q - 1)$ -dimensional hypersphere may be determined by using hyperspherical coordinates in $D = q - 1$ dimensions, i.e., for the i th vector $\vec{S}^{(i)}$ (with $1 \leq i \leq q$),

$$\begin{aligned} x_1^{(i)} &= \sin \theta_1^{(i)} \sin \theta_2^{(i)} \dots \sin \theta_{q-3}^{(i)} \sin \theta_{q-2}^{(i)} \\ x_2^{(i)} &= \sin \theta_1^{(i)} \sin \theta_2^{(i)} \dots \sin \theta_{q-3}^{(i)} \cos \theta_{q-2}^{(i)} \\ x_3^{(i)} &= \sin \theta_1^{(i)} \sin \theta_2^{(i)} \dots \cos \theta_{q-3}^{(i)} \\ &\dots \\ x_{q-2}^{(i)} &= \sin \theta_1^{(i)} \cos \theta_2^{(i)} \\ x_{q-1}^{(i)} &= \cos \theta_1^{(i)} \end{aligned}$$

We initially set $\theta_i^{(i)} = 0$ for $1 \leq i \leq q - 2$, $\theta_j^{(i)} = \alpha_j$ for $j < i \leq q$ and $1 \leq j \leq q - 3$, and $\theta_{q-2}^{(q-1)} = -\theta_{q-2}^{(q)} = \alpha_{q-2}$. There remains $q - 2$ angles α_j to be determined from $q - 2$ equations $\vec{S}^{(i)} \cdot \vec{S}^{(i+1)} = -1/(q - 1)$ with $1 \leq i \leq q - 2$, from where we obtain

Chapter 7. Fast Flat-Histogram Method for Generalized Spin Models

$\alpha_1 = \arccos \frac{-1}{q-1}$, $\cos \alpha_{j+1} = \frac{\cos \alpha_j}{1 + \cos \alpha_j}$, and thus by induction $\cos \alpha_j = \frac{-1}{q-j}$. After a bit of algebra, we find $\vec{S}^{(1)} = (0, \dots, 0, 1)$, and

$$\vec{S}^{(i)} = \left(\underbrace{0, \dots, 0}_{q-1-i \text{ terms}}, \sqrt{\frac{q(q-i)}{(q-1)(q-i+1)}}, \{x_{q-1-i+j}^{(i)}\}_{1 < j < i}, \frac{-1}{q-1} \right)$$

for $1 < i < q$, where the $(q-1-i+j)$ th coordinate reads

$$x_{q-1-i+j}^{(i)} = -\sqrt{\frac{q}{(q-1)(q-1-i+j)(q-i+j)}}.$$

$\vec{S}^{(q)}$ and $\vec{S}^{(q-1)}$ differ only in the sign of their first coordinate x_1 . Once these vectors have been computed for a given q , which may be done on start-up, determining the lattice energy requires, first computing the DFT $\tilde{S}_j(k)$ of each sequence of coordinates $\{\vec{S}^{(l)}\}_{l=1, \dots, L}$, and then evaluating the double sum $\sum_{k=0}^{L-1} \sum_{j=1}^q \tilde{J}(k) |\tilde{S}_j(k)|^2$. As a result, the whole operation is associated with a $O(qL \ln L)$ complexity — or in general $O(qL^D \ln L^D)$ —, provided the implementation relies on a FFT radix algorithm. As a by-product, it should be noted that once the Fourier components have been computed, it is straightforward to derive the Fourier transform of the spin-spin correlation functions at any inverse temperature β from $\tilde{g}_\beta(k) = 1/L \langle \sum_{j=1}^q |\tilde{S}_j(k)|^2 \rangle_\beta$, where the mean value is obtained from a reweighting procedure. At large lattice sizes, the requirement that L Fourier components be stored at each MCS may constitute a significant challenge in terms of computer memory; in this case, a practical work-around consists in computing microcanonical averages for each energy level visited during the simulation, and then to perform the reweighting procedure directly from these microcanonical averages. In the case of long-range interactions, careful attention must be paid, however, to the influence of the discretization of the energy axis in terms of systematic error.

7.3.2 Efficient cluster construction for long-range interactions decaying with the distance

For long-range spin models, building a new cluster at each MCS takes the order of L^{2D} operations, since $L^D(L^D - 1)/2$ pairs of spins are considered in turn for bond activation. When interactions decay with distance, the probability of adding a bond between two spins falls off quite rapidly as the distance between them increases. A significant amount of time during the construction of the cluster is thus wasted because an overwhelming number of bonds are considered for activation which have only a negligible probability to be activated. Even in the case of interactions decaying as $1/|i-j|^{1+\sigma}$ with σ close to 0, does the bond count never exceed a few percent

7.3. Optimization for long-range models

of the whole number of available bonds. In this respect, switching from a local- to a global- update scheme might well be an ill-fated choice as the gain in terms of autocorrelation time is spoiled by the exceedingly time consuming construction of the cluster. However, an efficient construction method was proposed by Luijten and Blöte in the recent past [227], with an efficiency that is independent of the number of interactions per spin, and a CPU demand that scales roughly as L^D . The rationale behind this method is to use cumulative probabilities, whereby instead of considering each spin in turn for addition to a given cluster, it is the index of the next spin to be added which is drawn at random. We now give a sketchy outline of the method in the context of long-range chains. Extensive details may otherwise be found in [227, 226]. First of all, the probability to add a bond is split up into two parts, namely, (i) a provisional probability $\pi_l(E)$ (hereafter simply denoted π_l) depending on the distance $l = |i - j|$ between spins and on the lattice energy E , and (ii) a factor $f(\sigma_i, \sigma_j)$ controlled by the spin values, e.g., a Kronecker delta symbol in the case of a Potts model. If 0 designates the index of the current spin to which we are adding bonds (i.e., spin indices are considered to be relative to the current spin), then the provisional probability of skipping $k - 1$ spins and bonding the current spin with a spin at position $k > 0$ is given by $P_0(k) = \prod_{m=1}^{k-1} (1 - \pi_m)\pi_k$. From there, one builds a table of cumulative probabilities $C_0(j_1) = \sum_{k=1}^{j_1} P_0(k)$ for all $j_1 > 0$, so that the index j_1 of the spin to be bound with current spin 0 is obtained by first drawing a random number $0 < r < 1$ and then reading out j_1 from the table, i.e., j_1 is such that $C_0(j_1 - 1) < r < C_0(j_1)$. Standard binary-search algorithms may be used for this purpose. Last, a bond is activated between spins 0 and j_1 with a probability $f(\sigma_0, \sigma_{j_1})$, and we proceed further with the computation of the index $j_2 > j_1$ of the next spin to be bound with current spin 0. The corresponding provisional probability thus becomes $P_{j_1}(k) = \prod_{m=j_1+1}^{k-1} (1 - \pi_m)\pi_k$, and the cumulative probabilities read $C_{j_1}(j_2) = \sum_{k=j_1+1}^{j_2} P_{j_1}(k)$. The same procedure is repeated for $\{j_3, j_4, \dots\}$ until we draw a $j_\alpha > L$, in which case we jump to the next current spin, which in a one-dimensional model is the nearest-neighbor of the previous current spin. In addition, there are two formulas which make it easier to compute cumulative probabilities: first, one can show that $C_0(j) = 1 - \exp[-\beta(E) \sum_{k=1}^j J(k)]$, where E is the energy of the current configuration, and second, the cumulative probabilities $C_{j_\alpha}(j_{\alpha+1})$ can be straightforwardly derived from the $C_0(j)$ coefficients through the relation $C_{j_\alpha}(j_{\alpha+1}) = \frac{C_0(j_{\alpha+1}) - C_0(j_\alpha)}{1 - C_0(j_\alpha)}$. It follows from the last relation that, instead of building a look-up table for each $C_{j_\alpha}(j_{\alpha+1})$, we may as well draw a random number $0 < r < 1$, transform it to $r' = r[1 - C_0(j_\alpha)] + C_0(j_\alpha)$, and choose the next spin to be added from the relation $C_0(j_{\alpha+1} - 1) < r' < C_0(j_{\alpha+1})$. In practice, we thus simply need to compute a single look-up table filled with $\sum_{k=1}^j J(k)$ for each j at the beginning of the simulation, from where we will derive the $C_0(j)$ coefficients at each new MCS corresponding to a lattice configuration with a given energy E . This last task requires the order of L^D operations. To sum up, the construction of each cluster thus consists in choosing a “current” spin among $L - 1$ possible spins in turn, e.g., starting from the leftmost one, and then activating bonds between the

current spin and other spins located to its right by drawing a random number, scaling it, and selecting the bond indices from a look-up table containing the $C_0(j)$ coefficients at energy E . Once each spin has been considered as a current spin, a cluster multiple labeling technique can eventually be used to identify every set of spins actually belonging to the same cluster [161].

7.4 Numerical tests of algorithm performance

In this section, we address the performance of our algorithm in terms of dynamic behavior. Since our work focuses mainly on long-range spin models, we decided to perform intensive numerical tests on the one-dimensional q -state Potts chain with LR interactions $1/|i - j|^{1+\sigma}$ decaying as a power law of the distance between spins. The rich phase diagram of this model, and the fact that several numerical studies have been carried out on this model in the recent past, makes it a perfect test-case. For the sake of comparison with other numerical methods, and in order to ensure that our algorithm did not produce systematic errors, we also performed several tests on the two-dimensional model with nearest-neighbor (NN) interactions, for which exact results are known (see [46, 59]; also references in [179]). Both models are known to exhibit a first-order transition for an appropriate set of parameters, namely, $q > 4$ for the NN model [350], and $\sigma < \sigma_c(q)$ for the LR one, with for instance $\sigma(3) = 0.72(1)$ [284]. We chose a set of parameters that would allow us to observe both weak and strong first-order transitions, and concentrated on several indicators of performance, reliability, and scalability: these include tunneling, equilibrium and effective autocorrelation times, and mean acceptance rates. These indicators inform us about the efficiency with which the Markov chain reaches the equilibrium distribution and explores the phase space. They also tell us at what rate successive measurements decorrelate from each other, or equivalently what amount of resources is needed to obtain reliable statistics. Overall, they are therefore good indicators of whether CPU resources are efficiently utilized or not. As regards scalability, we also computed the dynamic exponents associated with tunneling and equilibrium times; these indicate how fast needs in CPU time grow with the lattice size.

All densities of states were calculated by means of the Wang-Landau algorithm, whereby, starting from an initial guess of the density of states $n(E)$, we update $n(E)$ after each visit to energy level E according to the rule $\ln n(E) \leftarrow \ln n(E) + \ln f$, where $\ln f$ is hereafter termed *Wang-Landau modification factor*. In the case of LR models, the unequal spacing of energy levels and the existence of energy gaps in the vicinity of the ground state required that we introduced a few changes over the original version. In particular, using an interpolator for $\ln n(E)$ turned out to be mandatory in order to compensate for the finite width of histogram bins — as would also be required for models having a continuous symmetry; indeed, we observed that using large bins tends to strongly reduce the acceptance rate if no interpolator is used. Bezier splines provide good interpolators, although a linear interpolation

7.4. Numerical tests of algorithm performance

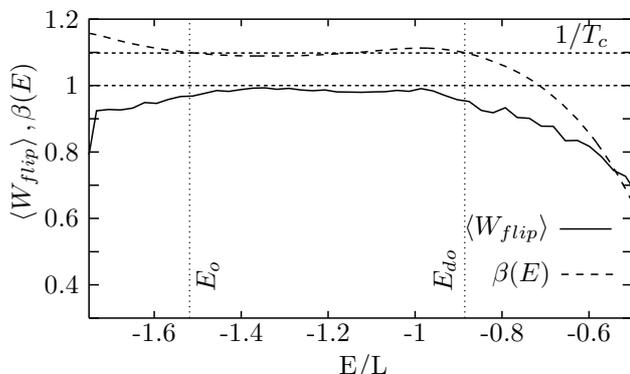


Figure 7.1: Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts chain with $\sigma = 0.7$, and $L = 1024$ spins. The dashed line shows the estimated inverse microcanonical temperature. The vertical dotted lines indicate the position of the histogram peaks corresponding to the ordered and disordered phases.

with a slope given by the microcanonical temperature $\beta(E)$ also proved to be particularly efficient whenever this last quantity was made available by other means, e.g., the transition matrix.

For small and medium lattice sizes, we systematically performed all simulations twice, first with standard single-spin updates (SSU) and then with our method embedding cluster updates (CU). We give an estimate of the error on the density of states obtained from both types of update scheme. For the largest lattice sizes we studied, however, the SSU implementation simply turned out to be impracticable, due to either exceedingly high tunneling times, and — for LR models — excessive CPU demands, and we present results for the CU algorithm only.

7.4.1 Phase space exploration and mean acceptance rates

As opposed to the (canonical) Swendsen-Wang cluster algorithm, the acceptance rate of our algorithm — Eq. (7.1) — is not trivially equal to unity. Still, it is tightly related to the efficiency with which the Markov chain wanders about the phase space, since a low acceptance rate would lead to very repetitive dynamics. In view of this, it is instructive to compute an approximate analytical expression of this acceptance rate when the initial and the final energies E_a and E_b differ only by a small amount. Writing $E_b = E_a + \epsilon$, and carrying out a series development to first order in ϵ , one obtains $W_{flip} = \min(1, 1 + \Delta(E_a)dE)$, where

$$\Delta(E_a) = \beta'(E_a) \left(\sum_{b_{ij} \in \mathcal{B}} J_{ij} \frac{1 + p_{ij}(E_a)}{p_{ij}(E_a)} - |E_a| \right),$$

with the same notation as in Sec. 7.2. We wish to obtain an estimate of the first statistical moments of $\Delta(E)$. We hereafter consider the case of a model with nearest-

Chapter 7. Fast Flat-Histogram Method for Generalized Spin Models

neighbor interactions ($J = 1$), for which we can carry out an exact derivation. The last expression simplifies to

$$\Delta(E_a) = \beta'(E_a) \left(B \frac{1+p}{p} - |E_a| \right),$$

where B stands for the total number of bonds and $p = p(E_a)$. From the distribution of bond counts at a given energy E_a ,

$$P(B) = \binom{|E_a|}{B} \frac{p^B}{(1+p)^{|E_a|}},$$

we can derive the average bond count, $\langle B \rangle = |E_a| \frac{p}{1+p}$. This allows us to rewrite $\Delta(E_a)$ as

$$\Delta(E_a) = \beta'(E_a) \frac{1+p}{p} (B - \langle B \rangle);$$

hence $\langle \Delta(E_a) \rangle = 0$. The variance of $\Delta(E_a)$ is thus proportional to the variance of the bond count distribution, i.e., $\langle B^2 \rangle - \langle B \rangle^2 = |E_a| \frac{p}{(1+p)^2}$, which yields

$$\sqrt{\langle \Delta(E_a)^2 \rangle} = \delta \Delta(E_a) = |\beta'(E_a)| \sqrt{\frac{|E_a|}{\exp \beta(E_a) - 1}}$$

For a given $\epsilon > 0$, one half of all attempted cluster flips thus leads to an acceptance rate which is lower than 1, the other half saturating at unity. Assuming a gaussian distribution for $\Delta(E)$, with the standard deviation computed above (which is valid for large enough lattice sizes), the mean acceptance rate is readily obtained from the mean value of a gaussian distribution centered at unity and truncated above 1, which yields

$$\langle W_{\text{flip}} \rangle (E_a) = 1 - \frac{\delta \Delta(E_a)}{2 \sqrt{2\pi}} \epsilon$$

In the case of interactions depending on the distance l between spins, one may observe that the average energy is related to the average number of bonds of length l by $\langle E \rangle = \sum_{l>0} J(l) \frac{1+p_l}{p_l} \langle B(l) \rangle$, which shows that $\langle \Delta(E) \rangle = 0$ also in this case.

At a first-order transition, $\beta(E)$ varies smoothly between the energy peaks of the ordered and disordered phases, which ensures that $\Delta(E)$ remains small. The mean acceptance rate for the six-state LR Potts chain with $\sigma = 0.7$ and $L = 1024$ spins is sketched in Fig. 7.1. While the acceptance rate is close to 1 inside the range of phase coexistence, the variance of $\Delta(E)$ increases when E lies outside the range of phase coexistence, and therefore leads to a reduction in the acceptance rate. We observe that this diminution is less marked at low-energy levels, for the energy cost associated with flipping a small number of big clusters is lower than that associated with randomly updating a great deal of small clusters, and $E_b - E_a$ is consequently lower in the last case. It is worth stressing, however, that the energy range of interest in the analysis of first-order phase transitions spans an interval

7.4. Numerical tests of algorithm performance

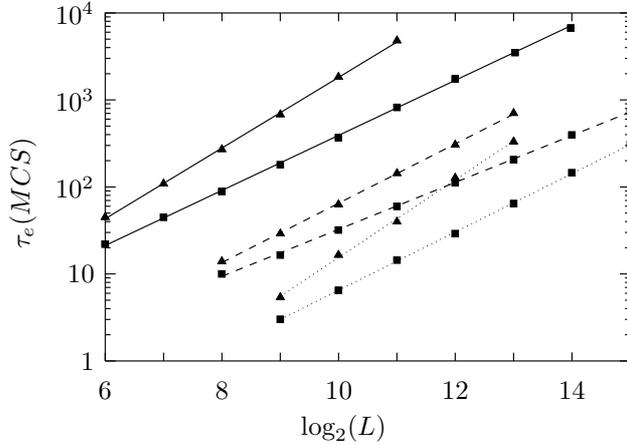


Figure 7.2: Tunneling times for the long-range Potts chain with $q = 3$, $\sigma = 0.4$ (dashed lines) and 0.6 (dotted lines), and $q = 6$, $\sigma = 0.7$ (solid lines). Triangles refer to the SSU implementation, while squares indicates estimates for our method (CU). Dynamic exponents z were determined from a fit to the power law $\tau_e \sim L^z$.

which is only moderately larger than that corresponding to phase coexistence, the only requirement being that metastability plateaus [284] and histogram peaks must be clearly visible. As a result, the fact that the mean acceptance rate for cluster flips remains well above 90% inside this range of energy represents already an improvement of a factor 3 with respect to the standard multicanonical approach, where we obtained acceptance rates oscillating around 30%.

7.4.2 Dynamic properties

Where performance measurements at first-order transitions are concerned, tunneling times have thus far been regarded as one of the most meaningful measurement

Table 7.1: Dynamic exponents z for the q -state Potts chain with power-law decaying interactions (a) and its two-dimensional counterpart with nearest-neighbor interactions (b). $z(SSU)$ and $z(CU)$ refer to single-spin and cluster updates respectively, while z_{muBo} and z_{muClus} make reference to the multibond method [183] and Rummukainen's multi-microcanonical cluster method [291] applied to the NN model.

q	σ	$z(SSU)$	$z(CU)$	z_{muBo}	z_{muClus}
6^a	0.7	1.35(3)	1.05(1)		
3^a	0.6	1.48(2)	1.11(1)		
3^a	0.4	1.13(2)	0.89(1)		
7^b		2.60(4)	1.82(2)	1.84	1.82(3)
10^b		2.87(4)	2.23(1)	2.1	

parameters [186, 179, 183]. They are defined as one half of the average number of MCS needed for the walk to travel from one peak of the energy histogram to the other – where peaks are defined with respect to the finite-size transition temperature – and turn out to represent a fairly good indicator of the interval between roughly independent samples.

Results for the LR chain with $q = 3$ and 6 are shown in Fig. 7.2. Dynamic exponents z were determined from a fit to the power law $\tau_e \sim L^z$, and are summarized in Table 7.1. We can witness a substantial reduction for both the LR and the NN models, with exponents close to and sometimes even below the ideal random-walk value $z = D$. As regards the NN model, our values compare extremely well with those obtained with the multibond method [183] and with Rummukainen’s hybrid-like two-step algorithm [291], although these approaches and ours differ markedly in the way clusters are constructed.

It should be mentioned that the distance $E_d - E_o$ the random walker must travel, i.e., the energy gap between the peaks of the histogram, does not scale linearly with the number of spins. This feature is especially apparent for long-range interactions, where $E_d - E_o$ grows all the more faster with increasing lattice size that σ comes closer to 0. As a result, the power law $\tau_e \sim L^z$ yields dynamic exponents which are underestimated with respect to the value given by a power law of the form $\tau_e \sim (E_d - E_o)^z$ (up to a dimensional factor 2 for the NN model). For instance, we would obtain $z = 1.40(3)$ instead of $z = 1.35(3)$ for $q = 6$ and $\sigma = 0.7$, and $z = 1.10(1)$ instead of $z = 1.05(1)$. Where the performance in terms of CPU demands is concerned (and in particular if one is interested in how it grows with the size of the system), we think, however, that the traditional definition $\tau_e \sim L^z$ is more meaningful.

While tunneling times represent a practical way to estimate the efficiency with which the random walker drifts along the energy landscape, they are subject to two limitations. First, they cannot be properly defined in the case of second-order phase transitions, since the histogram of the energy does no longer exhibit two peaks. Second, there is no direct connection between tunneling times and autocorrelation times, which makes it difficult to estimate the optimum interval between measurements that will yield perfectly uncorrelated data, and thus minimum statistical error on estimates of thermodynamic data. It is worth mentioning here that computing integrated autocorrelation times naively from the set of measurements, i.e., just as is usually done in the canonical case, simply makes no sense when simulating in the multicanonical ensemble, because the quantities we are interested in are, in the first place, reweighted averages of thermodynamic data [179].

Therefore, alternate definitions have been proposed, which try to circumvent these limitations. One approach is to compute the so-called *round-trip times* [7], which are computed from the number of MCS needed to get across the whole energy axis, that is, from the ground state to the upper energy level. Although round-trip times may be determined for any order of phase transition, they present unfortunately no more connection with statistical errors than do tunneling times. On the contrary, multicanonical *effective autocorrelation times*, which were first

7.4. Numerical tests of algorithm performance

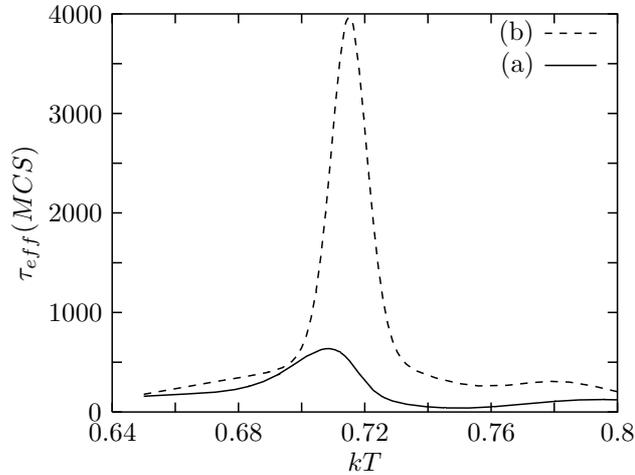


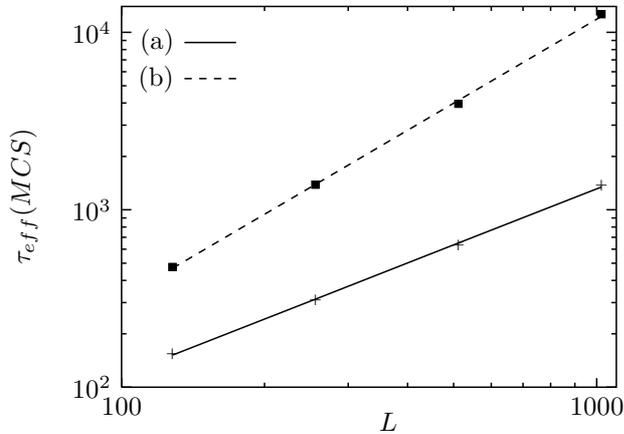
Figure 7.3: Effective autocorrelation time τ_{eff} for $q = 6$, $\sigma = 0.9$ and $L = 512$ with (a) cluster updates (b) single-spin updates. The effective transition temperature defined from the peak of the specific heat is $T_c(C_v) = 0.7163(2)$.

introduced in the framework of the multibond algorithm [183], offer a direct comparison with exponential or integrated autocorrelation times traditionally used in canonical simulations. Mimicking the canonical case, the effective autocorrelation time τ_{eff} can be defined for any thermodynamic variable θ by inverting the standard error formula $\epsilon_\theta^2 = \sigma_\theta^2 2\tau_{\text{eff}}/N$, where N stands for the total number of (possibly correlated) measurements, σ_θ^2 denotes the variance of the (reweighted) thermodynamic variable θ , e.g., $\langle E^2 \rangle - \langle E \rangle^2$, and ϵ_θ^2 is the squared statistical error on the same variable. The error may be estimated either from resampling or (jackknife) blocking procedures, or by performing multiple independent runs. Since both the variance and the error depend on the reweighting temperature, the previous definition obviously yields an effective autocorrelation time which also depends on the temperature.

We now discuss our results for effective autocorrelation times obtained for the six-state LR Potts chain with $\sigma = 0.9$ and $128 \leq L \leq 1024$ spins. For this value of σ , the model exhibits a very weak first-order transitions with no clearly visible histogram peaks for sizes below $L \sim 2000$. The choice of medium lattice sizes was dictated by the fact that we computed the error from multiple independent runs (around 20 runs of 10^6 MCS each), which we found a more reliable way of computing the statistical error than using a blocking procedure. Figure 7.3 shows the dependence of τ_{eff} on the temperature for $L = 512$. For both algorithms, τ_{eff} exhibits a peak in the vicinity of the effective transition temperature $T_c(C_v) = 0.7163(2)$. As expected, the reduction brought about by cluster updates in terms of correlation between measurements is marked, especially in the transition region, where single-spin update lead to a critical slowing down similar to the one encountered in canonical simulations. This behavior is consistent with the

Table 7.2: Effective autocorrelation times at the transition temperature defined from the location of the peak of the specific heat, for the six-state LR Potts chain with $\sigma = 0.9$

L	$\tau_{\text{eff}}(SSU)$	$\tau_{\text{eff}}(CU)$
128	475	155
256	1390	310
512	3960	635
1024	12700	1370
z	1.6(1)	1.0(1)


 Figure 7.4: Fit of effective autocorrelation times τ_{eff} to the power law $\tau_{\text{eff}} \propto L^\alpha$ for the six-state Potts chain ($\sigma = 0.9$ and $L = 512$) with (a) cluster updates (b) single-spin updates.

very general observation reported recently in [319] in the framework of the optimal ensemble, and also in [139] in the context of equilibration time for multicanonical algorithms (see also the next paragraph for more details on this issue), whereby the random walker diffuses at a slower pace in the critical region. In this regard, cluster updates optimize the diffusive current of the random walker in the critical region in much the same way as do the optimal ensemble weighting proposed in [319], yet with a different strategy: in the latter, the error is reduced by allowing the walker to spend more time in the critical region than in the rest of the energy axis; in our approach, it is the decorrelating capability of the move update itself which reduces the statistical error in the transition region. As is well known, however, cluster updates are especially efficient at the percolating threshold, and the reduction in terms of correlation is large *because* bond probabilities are governed by the microcanonical temperature. This interpretation is clearly underpinned by our investigation of the effect of poor estimates of $\beta(E)$ on tunneling times, presented later in Sec. 7.7.

Finally, we focus on the scaling behavior of autocorrelation times. Table 7.2 reports our results for L ranging from 128 to 1024 spins, where τ_{eff} is evaluated at

7.4. Numerical tests of algorithm performance

the effective transition temperature determined from the peak of the specific heat. Our method gives smaller autocorrelation times already for $L = 128$ spins. From these values, we also determined the associated scaling exponents by a fit to the power law $\tau_{\text{eff}} \propto L^z$ (Fig. 7.4), and obtained a highly satisfying value of $z \sim 1.0(1)$ with cluster updates.

We conclude the discussion on the dynamic characteristics of our algorithm with an investigation of equilibrating properties. As opposed to canonical simulation, estimating equilibrium times has been much less common in the context of multicanonical simulations; the non-linear relaxation function, while very informative when the equilibrium distribution is driven by a Boltzmann weight [216], is of limited use indeed if the engendered distribution is flat. Recently, however, an efficient procedure aimed at estimating equilibrium times for any equilibrium distribution was proposed by Guerra and Muñoz [139]. This procedure relies on a χ^2 regression with respect to the (expected) flat equilibrium distribution $\mathcal{P}(E)$. Starting from the same initial lattice configuration, n Markov processes are run with distinct random seeds, and at each MC step t , a histogram of the energy $V_t(E)$ is filled with the value of the energy of each process. Asymptotically, $V_t(E)$ should approximate the expected flat distribution $\mathcal{P}(E) \propto n(E)w(E)$. In order to estimate the equilibrium time in a more quantitative way, a $\chi^2(t)$ deviation of $V_t(E)$ with respect to the flat distribution is carried out at each MC step t , i.e., $\chi^2(t) = \sum_E (V_t(E) - n\mathcal{P}(E))^2 / (n\mathcal{P}(E))$, where the sum runs over histogram bins. For large n , and provided equilibrium has been reached, the distribution of $\chi^2(t)$ over m experiments obeys a χ^2 law with a number r of degrees of freedom given by the number of histogram bins minus one, that is, with a mean equal to r and a standard deviation given by $\sqrt{2r/m}$. Due to the intensive demand in CPU required by this procedure, we restricted our estimation of equilibrium times to the single case $q = 6$ and $\sigma = 0.9$. We performed $n = 1000$ Markov processes for sizes between $L = 128$ and $L = 512$, and estimated the equilibrium time from a single experiment (that is, $m = 1$) by simply monitoring the time needed for $\chi^2(t)/r$ to reach unity and then stay within the interval $[1 - 2\sigma/r, 1 + 2\sigma/r]$. As illustrated in Fig. 7.5, relying on a single experiment leads to quite large error bars, yet this is sufficient for our purpose. From the graphs of $\chi^2(t)$ we read $\tau_{eq} = 4500 \pm 500$ MCS and $\tau_{eq} = 23000 \pm 2000$ MCS for the cluster- and single-spin updates respectively; in spite of the large uncertainty, the reduction in terms of equilibrium time brought about by our method is clearly visible. Results for other lattice sizes are summarized in Table 7.3. A fit to the power law $\tau_{eq} \propto L_{eq}^z$ (see Fig. 7.6) yields the scaling exponents $z_{eq} = 1.96(5)$ and $z_{eq} = 1.21(3)$ for the single-spin and the cluster updates respectively. Here again, we think that lower diffusion currents in the latest case account for the higher pace at which the random walker reaches the equilibrium distribution.

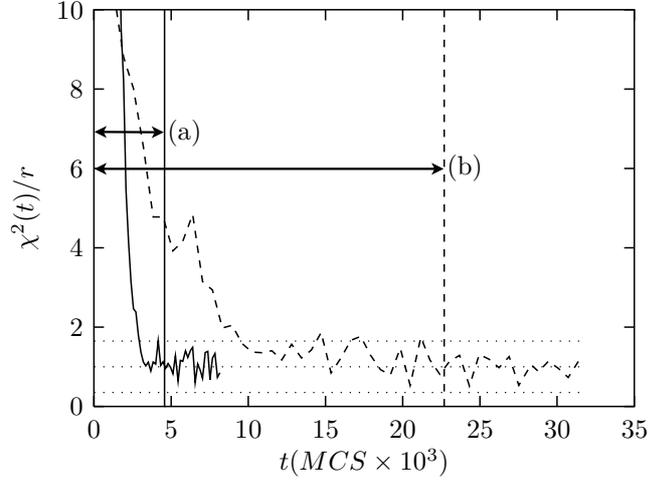


Figure 7.5: Plot of $\chi^2(t)/r$ for the six-state Potts chain ($\sigma = 0.9$, $L = 512$) using (a) cluster updates and (b) single-spin updates. The regression was carried out over a histogram containing 20 bins populated from 1000 runs, all starting in ground state configuration but with distinct random seeds.

Table 7.3: Equilibrium times for the six-state LR Potts chain with $\sigma = 0.9$ obtained by monitoring the graph of $\chi^2(t)/r$.

L	$\tau_{eq}(SSU)$	$\tau_{eq}(CU)$
128	1700(100)	800(120)
256	6000(750)	2000(200)
512	23000(2000)	4500(500)
1024	101000(8000)	10000(800)

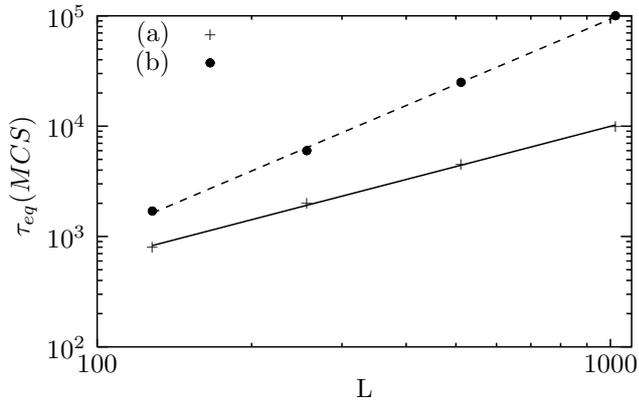


Figure 7.6: Fit of equilibrium times to the power law $\tau_{eq} \propto L^\alpha$ for the six-state Potts chain ($\sigma = 0.9$). (a) cluster updates and (b) single-spin updates.

7.4. Numerical tests of algorithm performance

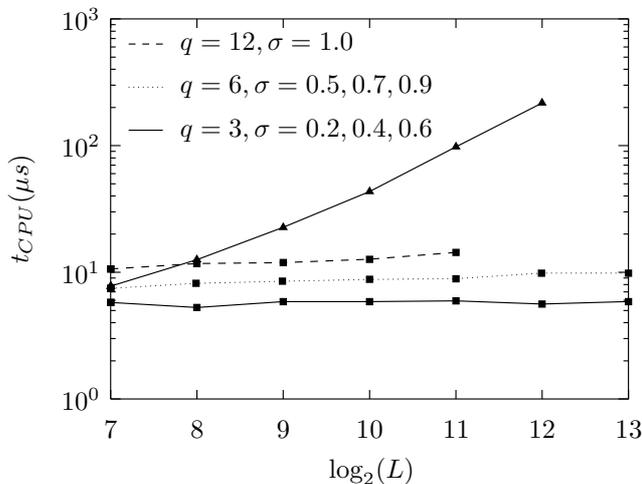


Figure 7.7: CPU time per MCS and per spin for the long-range Potts chain. Triangles indicate typical CPU times for the local-update algorithm (SSU), irrespective of q and σ . Squares refer to our algorithm (CU) with LR specific optimizations included; for $q = 3$ and $q = 6$, estimates were determined by averaging over the indicated σ values.

7.4.3 Overall CPU demand for LR models

We now discuss CPU demand in the case of LR models, and concentrate on the gain in CPU resources brought about by the optimization schemes proposed in Sec. 7.3. Assuming a decently efficient algorithm implementation, this indicator yields a rough account of the real algorithm complexity, although it should be mentioned that it is usually an elaborate task to estimate this quantity rigorously, partly because its value hinges heavily on a variety of implementation, CPU architecture and compiler dependent properties. We decided to measure CPU times over a series of one-hour long simulation runs on a handful of distinct CPU architectures, including Intel Pentium and Xeon at 2.4 and 3.2GHz. Figure 7.7 sketches averages of the CPU (user) time per MCS and per spin, where small fluctuations might be attributed to the effect of varying CPU cache sizes amidst our clutch of CPU's. While for the local-update implementation the demand in CPU per spin grows linearly with the number of spin, it is roughly constant over a fairly large range of lattice sizes in the case of our cluster-update algorithm. Moreover, our method already outperforms the local-update scheme starting from several hundreds spins, with nonetheless an increased footprint for higher q values which is accounted for by the correspondingly higher number of FFT's to be computed. This, however, clearly demonstrates the breakthrough that our method brings about for the study of long-range models, paving the way for precise tests of finite-size scaling.

7.5 Two-dimensional NN Potts model: comparison with exact results

In order to check that our algorithm did not produce systematic errors, we computed transition temperatures and interface tensions between coexisting phases for the two-dimensional q -state Potts model ($q = 7, 10$) with nearest-neighbor (NN) interactions and helical boundary conditions. Results regarding the dynamic characteristics of our algorithm for this model were reported in Sec. 7.4; we will concentrate here on precision matters. For $q = 10$, we obtained $T_c(L) = 0.70699(5), 0.70491(5), 0.70300(2), 0.70278(1), 0.70164(1), 0.701328(4)$ and $0.701249(2)$ for $L = 16, 20, 30, 32, 64, 128$, and 256 , where T_c was determined from the location of peaks of the specific heat. C_v was computed directly from the estimated density of states, and then refined from an additional production run of length 10^7 MCS. The error was estimated by means of the jackknife method. Following standard FSS theory at first-order phase transitions, we collapsed $C_v(T)/L^2$ vs $(T - T_c)L^2$ over the five highest lattice sizes and found an infinite size temperature $T_c(\infty) = 0.701236(3)$ in very good agreement with the exact value $0.70123157\dots$. The same procedure applied to $q = 7$ and $L = 64, 128$, and 256 yielded $T_c(\infty) = 0.773059(1)$ which again matches perfectly the exact value $0.7730589\dots$. We estimated the interface tension Σ from the histogram of the energy, reweighted at a temperature where energy peaks have the same height, namely, $2\Sigma = -L^{-1} \ln P_{min}$. Here, P_{min} designates the minimum of the histogram between the two energy peaks and the peak heights are normalized to unity. We computed Σ directly from the density of states, and estimated the error from the additional production run. In this regard, it should be noted that estimating interface tensions directly from the density of states generally yields values that lie below those computed from histograms collected during production runs. Our algorithm allowed us to determine Σ with a four-digit precision for sizes up to $L = 256$ and nonetheless rather modest statistics. For the seven-state model, we obtained $2\Sigma = 0.0336(6), 0.0294(1), 0.02631(8)$, and $0.02384(9)$ for $L = 32, 64, 128$, and 256 ; a linear fit of the form $\Sigma \sim \Sigma(\infty) + c/L$ [220] performed over the three largest sizes (i.e., for L above the disordered phase correlation length $\xi \sim 48$ [59]) yielded the infinite size value $0.02230(11)$, still above the exact value 0.020792 , yet closer to it than estimates reported in several previous studies [291, 183, 179].

7.6 LR Potts chain: error estimates and tests of finite-size scaling

In this section, we discuss the precision of our results for the q -state Potts chain with algebraically decaying interactions, i.e., $J(r) = 1/r^{1+\sigma}$. Our purpose is twofold. First, we estimate the error in the density of states $n(E)$ obtained from the Wang-Landau algorithm, so that we can obtain a better insight into the benefit of our method with regard to the iterative calculation of $n(E)$. Second, we determine con-

7.6. LR Potts chain: error estimates and tests of finite-size scaling

fidence intervals on reweighted averages computed from an additional production run. Since computing thermodynamic quantities from a production run does not require that the histogram be perfectly flat, nor that the estimate of the density of states be perfectly accurate, this amounts to estimating the gain in precision brought about by lower autocorrelation times.

7.6.1 Statistical error of the density of states

In order to compare the error on the density of states produced by the single-spin update implementation and our method, we performed for each method a series of 12 independent simulations with the Wang-Landau algorithm, all starting with the same initial guess of the density of states. The model parameters were set to $q = 6$, $\sigma = 0.9$ and $L = 512$. This choice of parameters guarantees that, in spite of the modest lattice size we consider, autocorrelation times differ by a sufficient amount between the single-spin updates method and our method, so that the benefit may be clearly interpreted in terms of decorrelating capabilities. The initial guess of $S(E)$ was scaled up from an estimate obtained at $L = 256$, and the updating factor of the Wang-Landau algorithm was initially set to $\ln f = 5$. We did not make use of all improvements to the original Wang-Landau algorithm, as proposed by Zhou and Bhatt in [362], since these would have partly overshadowed the gain produced solely by lower autocorrelation times. Indeed, we mainly focused on the systematic error (rather than the whole statistical error) that may show up during the first iterations. It was established in [362] that this systematic error results from the combination of a large $\ln f$ coefficient with the presence of strong correlations between adjacent binning. We thus simply relied on the original histogram flatness criterion to switch from one iteration to another, and divided $\ln f$ by the same amount (namely, 5) after each iteration which passed the flatness check. We found out, however, that using the criterion in [362] instead, that is, averaging $n(E)$ on multiple independent runs after each iteration, and switching to the next iteration only after a given number of entries was recorded in the histogram (see Eq. (12) in [362]), led to markedly lower statistical errors. As illustrated in Fig. 7.8, the statistical error on the density of states is clearly improved by our method. In particular, cluster updates lead to a spread of the error over the whole energy axis. In this respect, and as already mentioned in Sec. 7.4, the lower diffusion rates associated with collective updates in the critical energy region offer a clear benefit. As expected from the arguments of Zhou and Bhatt, the reduction is also more marked for $\ln f = 0.04$ than for $\ln f = 10^{-7}$, and the systematic error brought about by correlation between successive binning is indeed partly tamed by a lower Wang-Landau modification factor. Finally, we show in Fig. 7.9 the resulting statistical error on the specific heat, since thermodynamic averages are the relevant quantities in the first place. C_v was computed directly from the estimated density of states $n(E)$, i.e., according to the formula $C_v(kT) = (\langle E^2 \rangle_{kT} - \langle E \rangle_{kT}^2) / (L kT^2)$, where $\langle E^n \rangle = (\sum_E E^n n(E) e^{-E/kT}) / (\sum_E n(E) e^{-E/kT})$. For long-range models, energy levels are not equally spaced, and it should be noted that too large histogram bins may

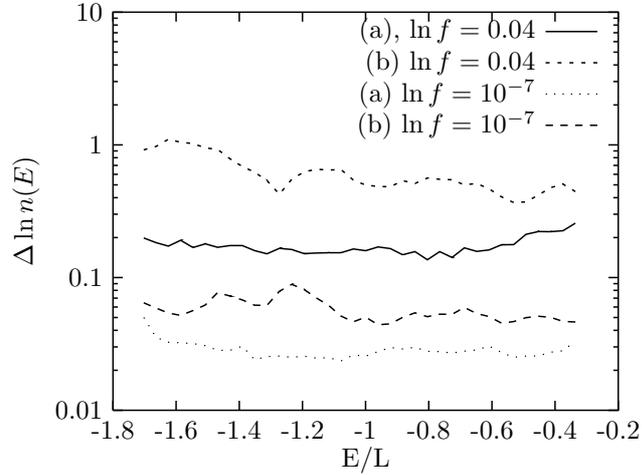


Figure 7.8: Statistical error on the density of states of the six-state Potts chain for two distinct modification factors $\ln f$ of the Wang-Landau algorithm. The statistical errors were obtained from 12 independent runs. The parameters of the model are $\sigma = 0.9$ and $L = 512$. (a) and (b) correspond to our method and the local-update algorithm respectively.

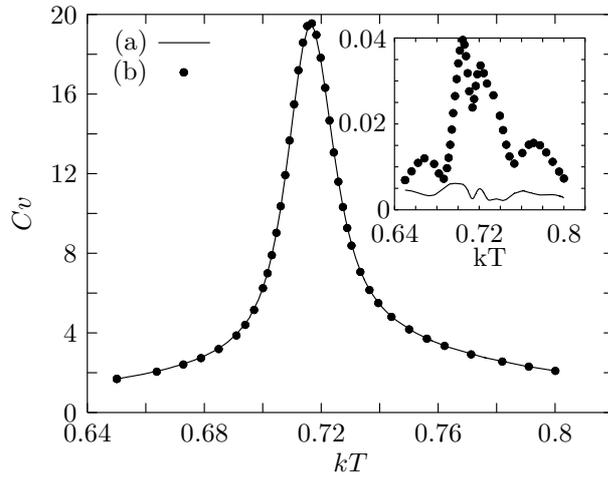


Figure 7.9: Graph of the specific heat for the six-state Potts chain ($\sigma = 0.9$ and $L = 512$) obtained directly from the final estimate of the density of states with (a) our method and (b) the local-update algorithm. The inset shows the relative error $\text{err}(C_v)/C_v$.

7.6. LR Potts chain: error estimates and tests of finite-size scaling

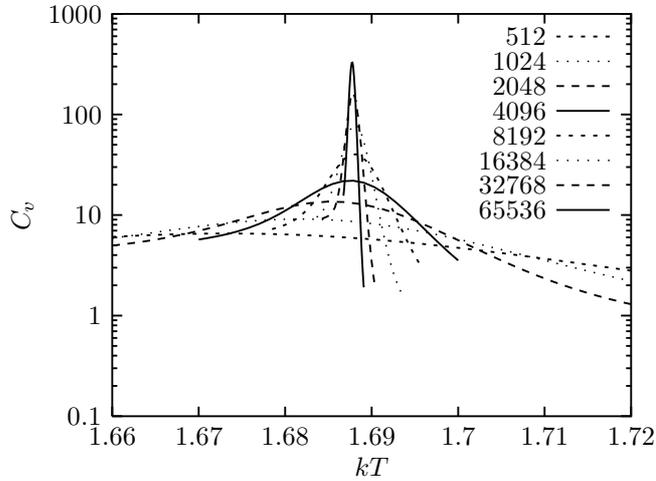


Figure 7.10: Specific heat for the three-state Potts chain with $\sigma = 0.5$ as obtained with our method.

cause a systematic deviation on the averages as well. We paid attention to this by comparing our results for several bin widths, and made sure that the systematic deviation engendered was always lower than the statistical error itself. As depicted in the inset of Fig. 7.9, the accuracy on the estimation of C_v is larger by nearly an order of magnitude at the transition temperature. Incidentally, we observe that this is comparable to the gain in terms of autocorrelation times, as already presented in Fig. 7.3.

7.6.2 Tests of finite-size scaling: transition temperatures and surface tensions

We now discuss some of our results for the three-state Potts chain, for which we performed extensive simulations for sizes ranging from $L = 128$ to $L = 65536$. As opposed to higher values of q , there exists indeed a large set of numerical studies for $q = 3$, so that comparison with previous estimates is easier. Table 7.4 reports our values for transition temperatures and peaks of response functions for $\sigma = 0.5$. Both C_v and χ were computed from a production run whose length varied between 10^6 and 10^7 MCS depending on the lattice size, and error bars were computed by means of the jackknife blocking method. We performed these production runs twice, first using single-spin updates, and then using our method, yet in both cases with the same estimate of the density of states. Figure 7.10 shows the graph of C_v as obtained with cluster updates. We mention that, for $L > 4096$, the local-update implementation was simply intractable as a result of excessive computation times. For all sizes, our results match within error bars for both methods, and it should be noticed that, for the two largest sizes, we obtain estimates of transition temperatures accurate up to the fifth digit. A fit of $T_c(L)$, as given by the location of the

Table 7.4: Estimates of peaks of the specific heat C_v and the susceptibility χ , and corresponding effective transition temperatures for the three-state LR Potts chain with $\sigma = 0.5$. Error calculations were carried out by means of the jackknife method applied to a single production run. The number of MCS per production run is the same for both methods, yet varies between 10^6 and 10^7 from the smaller to the larger lattice sizes.

L	$T_c(C_v)$ (CU)	(SSU)	C_v^{max} (CU)	(SSU)
128	1.6450(18)	1.645(3)	3.55(2)	3.55(3)
256	1.6607(2)	1.6607(13)	4.86(2)	4.88(5)
512	1.6741(9)	1.675(1)	6.54(3)	6.47(6)
1024	1.6815(2)	1.6815(17)	9.14(8)	9.10(24)
2048	1.6856(3)	1.685(1)	13.63(15)	13.73(80)
4096	1.68742(9)	1.6875(10)	22.21(34)	21.9(2.2)
8192	1.68801(7)		40.28(44)	
16384	1.688031(34)		79.46(35)	
32768	1.687851(12)		164.1(4)	
65536	1.687749(09)		332.8(6)	
L	$T_c(\chi)$ (CU)	(SSU)	χ^{max} (CU)	(SSU)
128	1.6793(14)	1.679(4)	3.44(3)	3.46(3)
256	1.6837(2)	1.6837(15)	6.09(3)	6.13(5)
512	1.6864(8)	1.6877(15)	10.81(7)	10.86(13)
1024	1.6882(3)	1.688(2)	19.73(28)	19.8(5)
2048	1.6887(2)	1.6882(15)	37.6(5)	37.5(1.8)
4096	1.68869(9)	1.6887(11)	75.4(1.2)	74.6(6.7)
8192	1.68842(7)		165.2(1.8)	
16384	1.688148(35)		369.8(1.1)	
32768	1.687870(20)		827(2)	
65536	1.687773(16)		1754(3)	

7.6. LR Potts chain: error estimates and tests of finite-size scaling

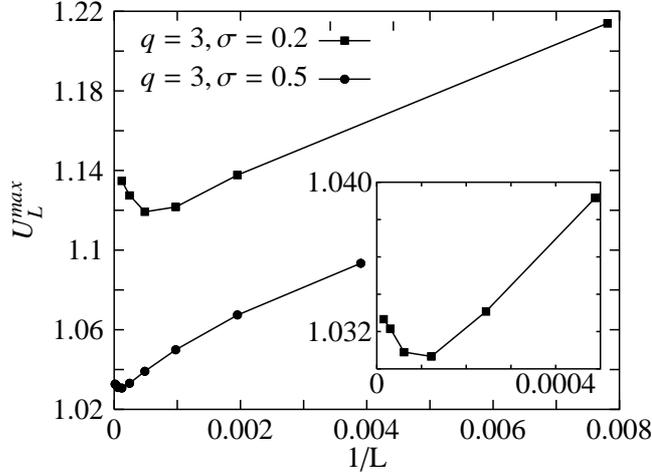


Figure 7.11: Peak of the cumulant of the energy $U_L = \langle E^4 \rangle / \langle E^2 \rangle^2$ as a function of the lattice size for the three-state LR Potts chain with $\sigma = 0.2$ and $\sigma = 0.5$. The inset shows a magnification near the origin for $\sigma = 0.5$.

peaks of C_v , to the power law $T_c(L) = T_c(\infty) + a/L$ yielded $T_c(\infty) = 1.68764(1)$ for our method, and $T_c(\infty) = 1.6888(8)$ for the local-update implementation. The same fit performed on $T_c(\chi)$ gave $T_c(\infty) = 1.68765(2)$ and $T_c(\infty) = 1.6892(6)$ respectively. These values compare very well with each other. However, our infinite size transition temperature is slightly larger than the best estimate determined so far (to the best of our knowledge) with a numerical approach, namely, the value of $T_c = 1.68542$ obtained in [251] with the cluster mean-field method. An important effect we noticed is the presence of a crossover around $L = 32768$ for $T_c(C_v)$, where the finite-size transition temperature starts to decrease slightly. This means that we had to restrict our fits to the largest lattice sizes. We think that this crossover may be attributed to the large correlation length at $\sigma = 0.5$. In Fig. 7.11, we observe indeed that the lattice size at which the crossover occurs is the same as the size where the peak of the reduced Binder cumulant of the energy, namely, $U_L = \langle E^4 \rangle / \langle E^2 \rangle^2$, experiences a minimum. For $\sigma = 0.2$, the same effect is witnessed by our results, with the change of slope of U_L taking place around $L = 2048$, and a change of behavior for $T_c(C_v)$ occurring near $L = 4096$. We also note in passing that, for $\sigma = 0.5$, relying on the Binder cumulant to assess the first-order nature of the transition requires simulating the system up to sizes that are far beyond the capabilities of single-spin update implementations. In particular, carrying a power-law fit of U_L restricted to sizes below $L \sim 3000$ would yield underestimated values. Our results in Fig. 7.11 show that the infinite size value lies around 1.033, and thus that the transition is stronger than suggested for instance in [131].

FSS behavior of the surface tension and evidence for fractal dimension Although a precise determination of correlation lengths for the long-range Potts chain

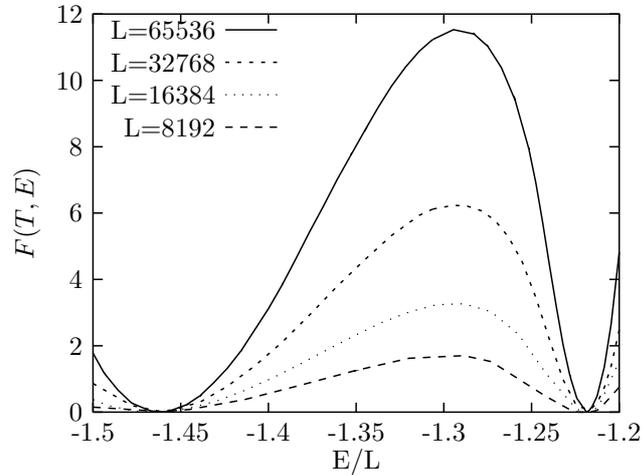


Figure 7.12: Graph of the free energy $F(T, E) = -\ln N(T, E)$ for the three-state LR Potts chain with $\sigma = 0.5$. $N(T, E)$ is the reweighted histogram at a transition temperature defined by equal peak heights. For the four lattice sizes shown here, lattice configurations corresponding to phase coexistence are suppressed by a factor ranging from 0.1 to 10^{-6} with respect to pure phase configurations; for the three largest sizes, we note that a canonical simulation is clearly intractable.

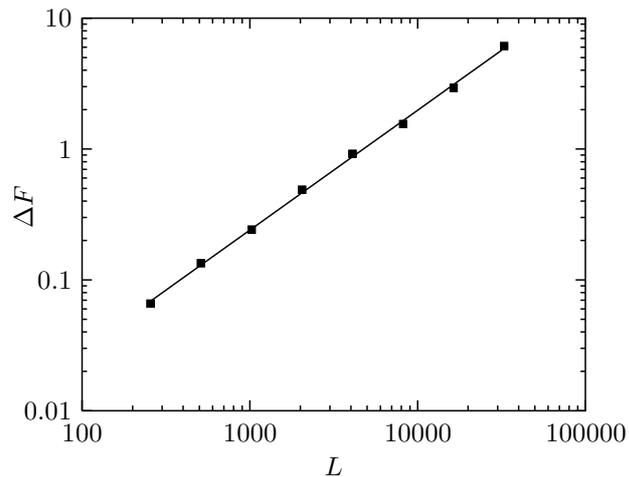


Figure 7.13: Fit of the interface free energy ΔF to a power law of the lattice size for the three-state LR Potts chain with $\sigma = 0.5$. All estimates of ΔF were obtained with our method.

7.6. LR Potts chain: error estimates and tests of finite-size scaling

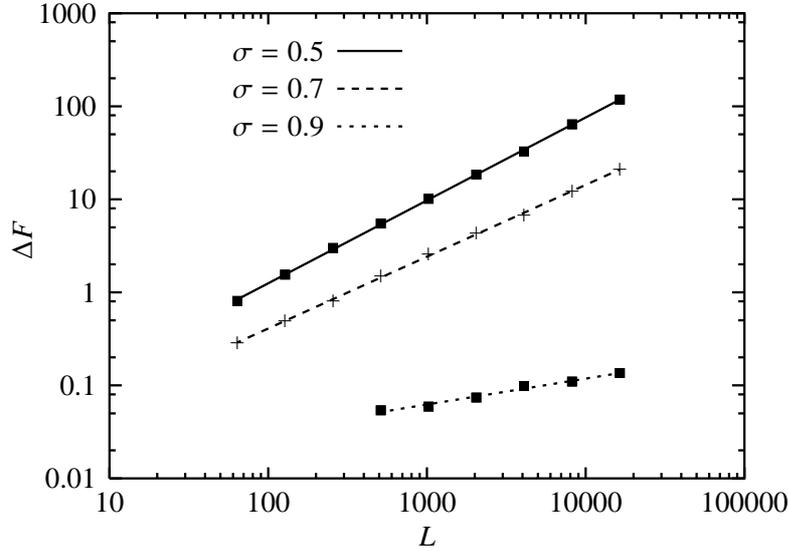


Figure 7.14: Fit of the interface free energy ΔF to a power law of the lattice size for the six-state LR Potts chain with $\sigma = 0.5, 0.7,$ and 0.9 . All estimates of ΔF were obtained with our method.

is beyond the scope of this work, we tried to obtain a rough estimate of it from the finite-size behavior of the interface free energy ΔF (akin to a “surface tension”, although this term is certainly no longer appropriate for long-range models). First, we computed a reweighted histogram $N(T, E)$ of the energy at the transition temperature T_{eqh} where both peaks of the histogram have equal height (see Fig 7.12). Then we measured ΔF from $\ln P_{max} - \ln P_{min}$, where P_{max} and P_{min} stand for the height of the peaks, and the minimum of the histogram between the two energy peaks, respectively.

By fitting the interface free energy to the power law $\Delta F \propto L^\alpha$, we obtain a very good fit for sizes ranging from $L = 256$ to $L = 65536$, yielding $\alpha = 0.91(2)$ and $\Delta F/L^\alpha = 0.0004$ in the thermodynamic limit. This is illustrated in Fig. 7.13. In view of the expected behavior for short-range models, namely, ΔF scales to leading order as a power of the lattice size with an exponent given by the dimension of the interface [219], this suggests that the effective dimension of the interface lies between 0 and 1 for long-range chains, evoking a fractal geometry with an Hausdorff dimension given by this exponent. This assumption is also supported by the fact that the fits of $\Delta F/L$ in [131] exhibit important finite-size corrections, while our fit with a non-integer exponent is exceptionally good over more than *two orders* of magnitude in linear size, and does not suggest such corrections.

Fig. 7.14 shows the same fits performed for the six-state model. The corresponding exponents are $\alpha = 0.89(1), 0.77(1),$ and $0.28(2)$ for $\sigma = 0.5, 0.7$ and 0.9 respectively. This behavior suggests that the fractal dimension of the interface, if any, increases and tends to the geometric dimension $D = 1$ as $\sigma \rightarrow 0$. In addition, the fact that exponents for $\sigma = 0.5$ and $q=3$ and 5 match within error bars seems to

indicate that the dimension of the interface does not depend on q , yet only on the decay parameter of the interaction.

Finally, if we mimic the large q arguments proposed in [46] for the nearest-neighbor Potts model, namely, that the correlation length ξ of the disordered phase is given by $(\Delta F/L)^{-1}$, we obtain after changing the unit exponent to α , an estimate of $\xi \sim 2500$ for the LR chain at $\sigma = 0.5$. In this respect, we would like to mention that: (i) the topology of the interface between the ordered and disordered phases is certainly far more complex than in short-range models, and (ii) we make use of Infinite Image Periodic Boundary Conditions. Therefore, this estimate should be taken as very rough one, since for instance the factor 2 in the definition of the interface tension (see Eq. (7) in [46]) might be questionable in LR models. Nonetheless, our estimate seems at least consistent with the fact that the change of slope of U_L sets in for sizes slightly above this size, i.e., $L \sim 5000$.

7.7 Combination with the transition matrix method

In this section, we examine how our method can be efficiently combined with the transition matrix method [341]. We show in particular that transition matrices represent a very efficient way of estimating the microcanonical temperature $\beta(E)$ used to compute cluster bond probabilities when nothing is known initially about the density of states. We also discuss how the estimated $\beta(E)$ can then be used as an efficient predictor to speed up the convergence towards the ground state during the early iterations of the Wang-Landau algorithm.

7.7.1 Efficient estimation of $\beta(E)$ and bootstrapping

As seems manifest from the scheme presented in Sec. 7.2, one of the basic requirements of our algorithm is to have an estimate of $\beta(E)$ at our disposal over the whole energy axis in order to compute cluster bond probabilities. One rather simple way of estimating $\beta(E)$ is to compute it from the current estimate of the density of states $n(E)$ using a finite-difference scheme, i.e., in real-time in the course of the iteration scheme. This is the most tractable approach if one decides to rely solely on Wang-Landau's algorithm to estimate $n(E)$. During early iterations, however, the estimate of $n(E)$ is somewhat rough and it is necessary to resort to a spline interpolation in order to obtain a sufficiently smooth estimate of $\beta(E)$. Since the unequal spacing of energy levels in long-range models renders an interpolation scheme for $n(E)$ absolutely mandatory [284], $\beta(E)$ is already available to us for free. Figure 7.15 shows estimates obtained with this approach for the three-state long-range chain with various interaction ranges, computed after ten iteration steps of 10000 measurements each. We note in passing that the presence of a clearly visible minimum in the three cases results from the first-order nature of the transition. For sufficiently short-range interactions, and when no random disorder is present, the microcanonical entropy $S(E)$ scales quite gently with the lattice size, and it is also

7.7. Combination with the transition matrix method

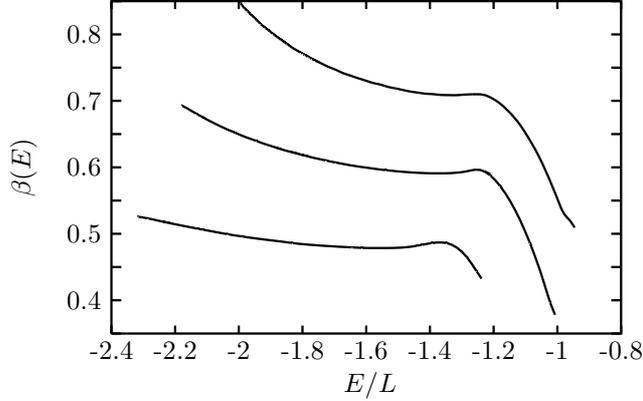


Figure 7.15: Microcanonical inverse temperature $\beta(E) = dS(E)/dE$ computed from the estimated density of states using a spline interpolation, for the three-state long-range chain with $\sigma = 0.4, 0.5,$ and 0.6 from bottom to top.

perfectly feasible to use the value of $\beta(E)$ obtained at a smaller lattice size as an initial guess.

In any case, it is crucial for the performance of our algorithm that we should compute $\beta(E)$ to sufficient accuracy. Indeed, we have found that any departure from the ideal line results in poorer performance, as illustrated in Fig. 7.16. The curve (a) in the figure shows the mean acceptance rate as a function of the energy for an estimate of $\beta(E)$ obtained after the ultimate Wang-Landau iteration and a modification factor $\ln f = 10^{-7}$. Curves (b) and (c) show the same quantity for microcanonical temperatures that were under- and overestimated by 10%. The poor estimate of $\beta(E)$ causes a marked decrease of the acceptance rate in the transition region (around $E/L \sim -1.5$), from around 100% to nearly 40%. Tunneling times clearly experience a corresponding increase, from 243 for the best estimate, to 737 and 1150 for the under- and overestimated temperatures, respectively. This can be easily explained, if one considers that the efficiency of cluster updates reaches a maximum at the percolation threshold. Any departure of the estimate of $\beta(E)$ from the ideal line results in a shift between the temperature at which clusters percolate (which depends on $\beta(E)$) and the *effective* temperature of the system (which is given by $dS(E)/dE$). This behavior has been observed in the context of canonical simulations of disordered systems, e.g., the Random Field Ising model [258], where the presence of randomness depresses the critical temperature. In this case, using the (canonical) simulation temperature to compute the bond probabilities simply results in a growing shift between the critical temperature and the percolation threshold as the randomness is increased.

In view of the previously mentioned requirements on the estimation of $\beta(E)$, it is clear that, if one does not have a reliable guess of $\beta(E)$ at hand before the simulation starts, an efficient scheme must be devised in order to compute $\beta(E)$ in

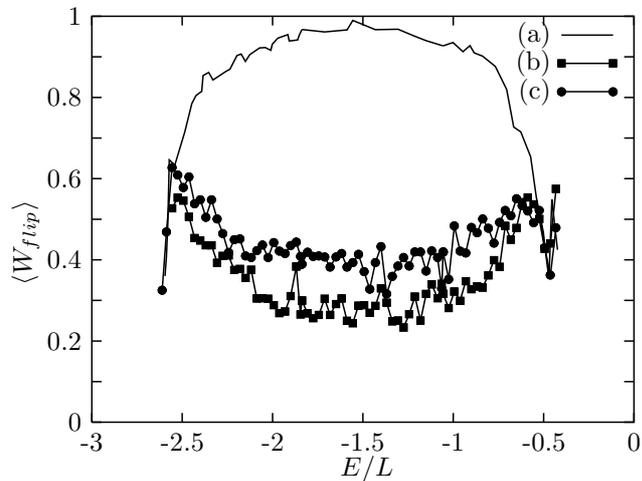


Figure 7.16: Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts chain with $\sigma = 0.5$, and $L = 512$ spins (strong first-order regime) for three different estimates of $\beta(E)$. (a) best estimate, as given by the ultimate iteration of the Wang-Landau algorithm; (b) $\beta(E)$ scaled by 0.9; (c) $\beta(E)$ scaled by 1.1.

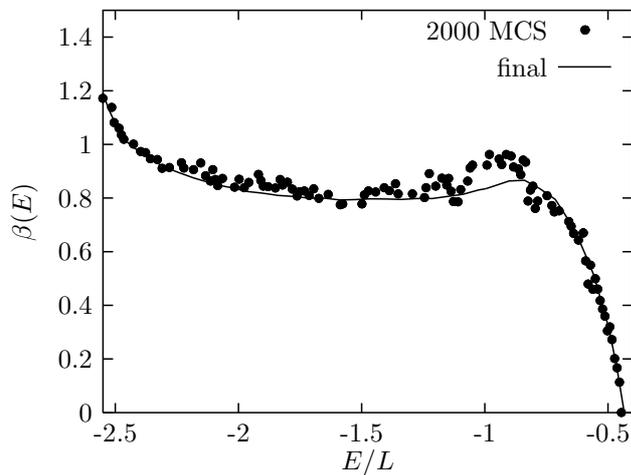


Figure 7.17: Symbols show the microcanonical inverse temperature $\beta(E)$ computed from the transition matrix accumulated over 2000 MCS, for the six-state LR model ($\sigma = 0.5$) containing 512 spins. The estimate obtained from an interpolation scheme after the ultimate iteration is shown as a solid line for comparison.

7.7. Combination with the transition matrix method

the early stage of the Wang-Landau algorithm. This is vital at this stage, because the exceedingly noisy estimate of the density of states makes it more likely to obtain under- or over-estimated values for $\beta(E)$. An efficient approach in this regard relies on transition matrices [341, 307]. This method produces highly precise estimates of $\beta(E)$, although it has an inherently higher cost in terms of computer load. The starting point is the Broad Histogram equation [339, 98]:

$$n(E)T_{\infty}(E \rightarrow E') = n(E')T_{\infty}(E' \rightarrow E),$$

where $T_{\infty}(E \rightarrow E')$ is the transition matrix element between energy levels E and E' (also denoted as $\langle N(\sigma, E' - E) \rangle_E$ in [98]). This quantity contains the micro-canonical average at energy E of the number of potential single-spin moves from a state σ of energy E to a state σ' of energy E' . It is estimated by accumulating a double-entry histogram $h(E, \Delta E)$ containing the number of potential moves from E to $E + \Delta E$ each time the energy level E is visited. Long-range interactions lead to energy levels which are irregularly spaced, with in particular a few gaps in the vicinity of the ground state [284], and it is necessary to choose an axis bin small enough to minimize discretization errors, and at the same time sufficiently large to contain at least a handful of entries. In this case, $T_{\infty}(E \rightarrow E')$ varies sufficiently smoothly for the following approximation scheme to be valid:

$$\beta(E) = \frac{1}{\Delta E} \ln \frac{T_{\infty}(E \rightarrow E + \Delta E)}{T_{\infty}(E \rightarrow E - \Delta E)},$$

where the actual estimate is obtained by weighted-averaging over several values of ΔE . As illustrated in Fig. 7.17 for the six-state LR chain, the estimation of $\beta(E)$ from the transition matrix elements is reliable already after 2000 MCS, which roughly corresponds to 50 round-trips between the upper and the lower energy range. For long-range models, each estimation of the number of potential moves requires the order of L^{2D} operations (as opposed to L^D for nearest-neighbor interactions). However, we have shown in Sec. 7.3 that a single cluster update can demand as little as $O(L^D \ln L^D)$ operations when long-range specific optimizations are carried out. This means that estimation schemes based on transition matrices partly scupper the benefits of the last optimizations, and should therefore be employed only as a bootstrap procedure when nothing is known yet about the micro-canonical temperature. Conversely, models with nearest-neighbor interactions do not undergo such a drawback, and make the transition matrix approach a perfectly transparent one from the viewpoint of algorithm complexity.

7.7.2 Efficient predictors for the Wang-Landau algorithm

Finally, we discuss how $\beta(E)$ can be used as an efficient predictor during the early stage of the Wang-Landau algorithm when nothing is known about the density of states. In the original implementation of this algorithm, we start with $S(E) = 0$ for all energy levels, and simply increment $S(E)$ by the modification factor $\ln f$

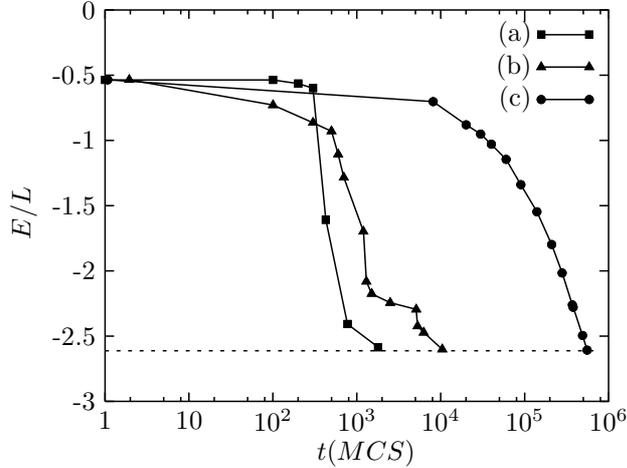


Figure 7.18: The graph shows the number of MCS needed to reach the ground-state (dashed horizontal line) of the six-state Potts chain ($\sigma = 0.5$ and $L = 512$) for an initially unknown density of states, using three distinct schemes: (a) and (b) predictor based on $\beta(E)$, local- and collective-update algorithms respectively; (c) no predictor ($S(E) = 0, \forall E$), local-update algorithm.

each time the corresponding energy level is visited. One of the main drawbacks of this approach is that the Markov chain tends to wander around a fairly long time in the upper energy range, until eventually enough visits have been recorded in the histogram for the system to start exploring low-energy levels. This point has already been mentioned in [362], where it was suggested that starting with a good initial guess of $S(E)$ was more efficient in terms of the number of histogram entries required to reach the final estimate, than performing a multi-range run with no initial guess at all. To circumvent this drawback when no initial guess is available, we therefore propose to use $\beta(E)$ to predict $S(E)$ for energy levels that are visited for the first time, and thus for which $S(E)$ is not available (i.e., it is set to $S(E) = 0$ in the original implementation of the Wang-Landau algorithm). A linear prediction scheme turned out to sufficiently efficient for our purpose. As illustrated in Fig. 7.18, using a predictor brings about a gain of three orders of magnitude in the time needed to reach the ground state. Our method and the single-spin update method lead similar performance, with nonetheless a slightly better behavior when cluster updates are used. We note that the Markov chain stays initially somewhat longer in the upper energy range when cluster updates are used, since a good estimate of $\beta(E)$ is needed to build the clusters with the correct bond probabilities. We think that this approach would prove particularly useful when the characteristics of the model makes it impossible to obtain an initial guess of $S(E)$ from simulations at smaller lattice sizes, e.g., in the presence of disorder or when the long-range interaction experience a slow decay.

7.8 Conclusion

In conclusion, we have developed a new Monte Carlo method which combines in an efficient and straightforward way the benefits of flat histogram algorithms with the ability of cluster algorithms to rapidly reduce temporal correlations. It is suited for spin models with any number of interaction between spins. Our formulation is versatile, and the method can be applied to a variety of density of states estimation schemes, including the Wang-Landau algorithm, Berg's recursion scheme or the transition matrix method. We have shown that using the microcanonical temperature to compute cluster bond probabilities leads to a drastic reduction in effective autocorrelation times, tunneling times and equilibration times. In the context of the Wang-Landau implementation, the reduced correlation between successive binning of the energy histogram yields a lower error in the estimation of the density of states, and as a result more reliable estimates of thermodynamic averages. Several schemes for the estimation of the microcanonical temperature were proposed, among which an efficient procedure which harnesses the power of the transition matrix method, and allows us to bootstrap the algorithm even if nothing is known initially about the density of states. Finally, we carefully examined the precision of our method in the case of spin models with power-law decaying interactions. Here, our method proves all the more powerful that it is able to reduce the algorithm complexity to that of a short-range model having the same number of spins. This allowed us to study several finite-size effects at large lattice sizes, otherwise largely out of reach of conventional local-update implementations. In particular, we found out that the interface free energy scales perfectly well with a power of the lattice size, yet with a non-integer exponent which lies between 0 and 1. This, we think, is accounted for by the complex topology of the phases in coexistence in long-range models. A more detailed study, including a deeper insight into the topological properties of the generated clusters and the estimation of correlation lengths at large lattice sizes, would be very promising. We think that our method clearly draws this challenge within computation range.

Chapter 8

Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

In this Chapter, a preliminary investigation of the critical behavior of the random field Ising (RFI) model with power-law decaying is proposed. This model, in its nearest-neighbor flavor, has been the subject of intense scrutiny during the last thirty years, with numerous controversies still surrounding its critical behavior. On the contrary, its long-range counterpart has been much less studied, and for exactly the same reasons as for other long-range models, the ability to vary the effective dimension of the model — from mean-field behavior to short-range behavior — makes it a perfect laboratory framework that might help shedding new light on these controversies.

This model would probably deserve a whole thesis on its own, and it is clearly not the purpose of the present chapter to investigate its behavior in detail. Rather, my goal in the following will be to capitalize on the methods developed in the previous parts in order to scrutinize one specific question regarding this model: the possible existence of a tricritical point separating first- and second-order transitions when random fields are generated by a bimodal distribution. The question has been mooted several times for the nearest-neighbor model, mean-field theory predicts such a behavior, and it is obviously exciting to check if the long-range model also exhibits this behavior, at least when the interaction is sufficiently close to the mean-field regime.

To what extent does such a study make sense with the remainder of this thesis? First, disordered models exhibit complicated energy landscapes, and as such they may be regarded as systems of choice for Monte Carlo methods operating in generalized ensembles. In this respect, I will naturally examine the efficiency of single-spin updates vs collective updates in the context of disordered systems.

Second, the detection of the order of the phase transition in the Potts chain has been central to the first part of this thesis: a dedicated method based on spinodal points has been introduced in this purpose in Chap. 4, and it is therefore appealing to check the efficiency of this method on a different model where a first-order transition is expected to occur.

The layout of this Chapter is as follows. The main results and points of controversy regarding the nearest-neighbor are recalled in Sec. 8.1. In particular, the influence of the shape of the field distribution on the phase diagram is reviewed in Sec. 8.2. Then, a derivation of mean-field results using the replica formalism is provided in Sec. 8.3: these results are expected to become exact for interactions that decay sufficiently slowly¹, and will thus serve as a guide. The upper critical range of the model is calculated in Sec. 8.4 by relying on a generalization to power-law decaying interactions of the dimensional argument of Imry and Ma. The algorithms are described in Sec. 8.5. Results, including a preliminary estimation of the phase diagram for two decay parameters, are presented in Sec. 8.6. Consistency with results obtained for the nearest-neighbor model, and possible improvements of the method, are discussed in Sec. 8.7.

8.1 Model and existing results

The random-field Ising model is a generalization of the Ising model in a uniform external field to the case of non-uniform, randomly distributed fields. The model Hamiltonian, in its most generic formulation, writes

$$H = -\frac{1}{2} \sum_{i \neq j} J_{ij} \sigma_i \sigma_j - \sum_l \sigma_l h_l$$

where $\sigma_i = \pm 1$ are usual Ising variables and J_{ij} is a generalized coupling constant. The set $\{h_l\}$ represents external fields coupling linearly with the spin variables, and randomly distributed according to some given law of probability, e.g., gaussian, or multimodal (e.g., made up of Dirac “functions”). As far as modeling of *real* systems is concerned, random fields are assumed to evolve on a time scale orders of magnitude larger than the usual time scale associated with thermal fluctuations of the spin variables, so that they represent a *quenched* disorder. In particular, thermodynamic averages are computed by assuming constant random fields: this is to be distinguished from the notion of *configurational* averages, i.e., averages over several realizations of disorder. An important point in this respect is the *self-averaging* property of the free energy with respect to configurational averages, whereby the free energy (per spin) for a given set of quenched fields tends towards its configurational average in the thermodynamic limit, so that averaging over disorder realizations — if this is technically feasible — proves an efficient way of obtaining

¹Although the exactness of the mean-field regime for sufficiently slowly decaying interactions has been proven for non-disordered models only [65, 68, 196, 331].

8.1. Model and existing results

the critical properties for *any* given set of quenched fields. This clearly plays a crucial role also in numerical simulations.

In its nearest-neighbor formulation, i.e., with J_{ij} taking non-null values for interactions between neighboring spins only, the RFI model has been the subject of unabated interest during the last three decades, particularly for fields distributed according to a gaussian law. Beyond its importance in the microscopic modeling of a variety of magnetic or ferroelectric materials, e.g., diluted antiferromagnets (see [20] for a review addressing experimental realizations), it has also become — along with the Edwards-Anderson model of spin glasses [110] — a paradigm for the investigation of critical phenomena in disordered systems. There are still several controversial issues surrounding the critical behavior of this model. One is the value of the lower critical dimension: the answer oscillates between $D_l = 2$ and $D_l = 3$ depending on the line of arguments invoked [3, 56, 71, 170, 173, 274], although strong evidence tips the balance in favor of $D_l = 2$. Another thorny question is the possible existence of a tricritical point for some distributions of random fields [2, 94, 54, 69, 154, 164, 238, 266, 275]. For a gaussian distribution of fields, it was shown by Schneider [295] that mean-field theory (which becomes exact for the infinite-range model) predicts a second-order phase transition at finite temperature with T_c decreasing with increasing field strength. On the contrary, it was argued on the grounds of renormalization-group arguments [2] that multimodal distributions lead to a first-order transition at sufficiently large fields, provided the distribution has a minimum at zero field (which is the case for the bimodal distribution). Finally, the need for a third, independent critical exponent entering a so-called modified hyperscaling relation — with respect to the pure Ising model — and the question of which universality class the model belongs has also been the matter of intense debate [69, 258, 260, 266, 275]. More detailed material on these questions may be found, for instance, in two reviews by Nattermann et al. [256, 255].

Long-range interactions The model studied hereafter is a generalization of a one-dimensional, nearest-neighbor RFI model to pair interactions decaying as a power law of the distance between spins, i.e., with the following Hamiltonian²

$$H = - \sum_{i < j} \frac{1}{|i - j|^{1+\sigma}} \sigma_i \sigma_j - \sum_l \sigma_l h_l \quad (8.1)$$

I will restrict the present study to a *bimodal* distribution of fields, i.e.,

$$\mathcal{P}(h_l) = \frac{1}{2} [\delta(h_l - h) + \delta(h_l + h)], \quad (8.2)$$

and assume *uncorrelated* fields, i.e., such that $[h_i h_j] = \delta_{ij}$ where the brackets refer to a configurational average, i.e., an average over quenched disorder.

²A factor $\frac{1}{2}$ has been included in numerical simulations for the sake of comparison, in the zero field limit, with the temperatures of the Potts model with $q = 2$.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

As opposed to the nearest-neighbor model, specific studies of the long-range model have been scarce: critical exponents have been calculated by means of renormalization group methods in [52, 343] (Bray also studied the problem of long- to short-range crossover [52]), and a rigorous work by Klein and Massoumian [206] recently addressed the long-distance behavior of correlation functions by means of high temperature/field expansions. Vojta investigated a spherical version of this model [330] using the replica method, with special attention given to the lower critical dimension.

Since the decay parameter of the interaction influences the effective dimension of long-range models, it is clear, however, that the controversies surrounding the nearest-neighbor model may be worthy of reexamination in the case of power-law decaying interactions. In particular, since mean-field behavior becomes exact for $\sigma \rightarrow -1$, one may expect some of the predictions of mean-field theory, e.g., the existence of a tricritical point for bimodal distributions, to show up for a sufficiently low decay parameter. On the other side of the phase diagram, i.e., for rapidly decaying interactions, the ability to continuously vary the effective dimension may prove a very efficient way to compute the upper critical range (i.e., the analog of the lower critical dimension), and to check for the validity of the various arguments invoked to estimate it.

In the following, I will only address the nature of the phase transition in the context of long-range interactions and with bimodal distributions. In addition, since no study, whether analytical or based on numeric simulations, has been made available — to the best of my knowledge — on this subject in the past, the emphasis in what follows will be given to a qualitative estimation of the phase diagram rather than on precise estimates of critical couplings or critical exponents.

8.2 Gaussian vs bimodal distributions: what's the point?

As already mentioned in the introduction, mean-field theory predicts that the transition remains continuous for all field strengths for a gaussian distribution of fields [295, 244]. Conversely, the existence of a tricritical point has been argued when fields are engendered by a bimodal distribution [244], or more generally, by any symmetric distribution with a minimum at zero field [2]: the transition turns from second- to first-order at sufficiently large field amplitude. With an eye to future use for comparison with numerical results, mean-field results will be derived in greater detail in Sec. 8.3 using a different line of argument based on the replica method.

Predictions based on mean-field approximations [201, 298] or combination thereof with renormalization-group schemes [94], confirm that, for a trimodal distribution of the form $\mathcal{P}(h_i) = p\delta(h_i) + \frac{1}{2}(1-p)[\delta(h_i - h) + \delta(h_i + h)]$, there is a tricritical point whenever $p \leq 1/4$, and a first-order transition is ruled out for $p > 1/4$ [69]. The latter case is thus expected to exhibit the same behavior as the gaussian distribution, yet the low value of the magnetization exponent β reported in some studies of the latter distribution [266] may be indeed misleading in this re-

8.3. Mean-field theory

spect, since this would suggest either a weak first-order transition, or a very sharp second-order transition. This behavior was also reported in Monte Carlo studies of the bimodal case [360, 287]: while the transition may indeed be first-order, it may also have unusual attributes, namely, no latent heat but a jump in the magnetization [287, 286].

By monitoring the behavior of the critical exponent of the connected and disconnected correlation functions, Young and Nauenberg [360] suggested the onset of a first-order transition driven by *fluctuations* for a bimodal distribution (a feature which is in contrast with the mean-field mechanism, yet was also mentioned in [343] for the long-range model with a gaussian distribution). In a recent Monte Carlo study, a tricritical point at large field was also asserted from the presence of an hysteresis loop in the field-cooling/field-heating cycle [154]. However, a high-temperature expansion predicted a fluctuation-driven first-order phase transition *also* for a gaussian distribution (provided $D < 4$) [164], a feature that was subsequently borne out by field-theoretic renormalization-group derivations [54]; the onset of a first-order transition was also confirmed for this distribution by a Monte Carlo study based on the replica-exchange algorithm [238], although for some, but not all realizations of disorder. Finally, Gofman et al. predicted the onset of a first-order transition for *both* types of distribution, so that *overall* the issue is still somewhat unclear as far as the influence of the shape of the distribution is concerned.

8.3 Mean-field theory

In this Section, I give a brief outline of the derivation of the free energy for the RFI model with a bimodal field distribution, in the mean-field approximation. The point of departure is the infinite-range Hamiltonian, i.e., $H = -\frac{J}{2N} \sum_{i \neq j} S_i S_j - \sum_l h_l S_l$, where J will be replaced by $\zeta(1 + \sigma)$ in the last step in order to recover the free energy of a long-range chain with $1/r^{1+\sigma}$ interactions in the mean-field approximation (see Chap. 4). I will take guidance from the derivation carried out in [265] for the Sherrington-Kirkpatrick (infinite-range) spin glass.

The free energy F for a fixed set of quenched fields reads $F(\{h_i\}) = -kT \ln Z(\{h_i\})$, which we rewrite as $F(\{h_i\}) = -kT \lim_{n \rightarrow 0} \frac{Z(\{h_i\})^n - 1}{n}$ by resorting to the usual replica trick ([110], see also, for instance, [265]): Z^n actually stands for $Z_\alpha Z_\beta \dots Z_\omega$, where each partition function corresponds to a replica of the original system with its own set of spin variables $\{S_i^\alpha\}_{i=1 \dots N}$. The configurational average then reads

$$[F] = -kT [\ln Z] = \lim_{n \rightarrow 0} \frac{[Z^n] - 1}{n},$$

and the derivation of the mean-field free energy thus involves first computing $\prod_{\alpha=1}^n Z_\alpha$, averaging over the distribution of random fields, taking the limit $n \rightarrow 0$ and eventually taking the thermodynamic limit $N \rightarrow \infty$. The “replicated” partition

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

function Z^n writes (with $\beta = 1/kT$)

$$Z^n = \text{Tr} \exp \left\{ \frac{\beta J}{2N} \sum_{\alpha} \sum_{i \neq j} S_i^{\alpha} S_j^{\alpha} + \beta \sum_{\alpha} \sum_l h_l S_l^{\alpha} \right\}$$

where the trace operation is over all sets of replica's variables $\{\{S_i^{\alpha}\}_{i=1 \dots N}\}_{\alpha=1 \dots n}$. The configurational average over fields generated by a bimodal distribution (Eq. 8.2) yields

$$[Z^n] = \text{Tr} \exp \left\{ \frac{\beta J}{2N} \sum_{\alpha} \sum_{i \neq j} S_i^{\alpha} S_j^{\alpha} + \sum_i \ln \cosh \beta h \sum_{\alpha} S_i^{\alpha} \right\}.$$

In order to carry out the trace independently over each spin variable, one first “linearizes” the $S_i^{\alpha} S_j^{\alpha}$ term by resorting to an (inverse) gaussian integral (akin to a Hubbard-Stratanovitch transform) over auxiliary variables m_{α} . It can be shown [265] that these variables m_{α} also represent the order parameter of each replica α . Whence, after tracing over all spin variables,

$$[Z^n] = \left(\frac{\beta J N}{2\pi} \right)^{n/2} \int \prod_{\alpha} dm_{\alpha} e^{-\beta N J m_{\alpha}^2 / 2 + N \ln L}$$

where

$$L = \frac{1}{2} \prod_{\alpha} 2 \cosh \beta (J m_{\alpha} + h) + \frac{1}{2} \prod_{\alpha} 2 \cosh \beta (J m_{\alpha} - h)$$

Equilibrium solutions correspond to the maxima of $[Z^n]$. Since the argument of the exponential function in the integral above is proportional to N , it is possible to evaluate $[F]$ through a saddle-point approximation, i.e., $[Z^n] \propto e^{-\beta [F]_0}$, where $[F]_0$ is the argument of the exponential function evaluated at its minimum. Assuming replica symmetry [272], i.e., $m_{\alpha} = m$ for all α , where m now denotes the order parameter of the “original” system, and considering that $n \sim 0$, equilibrium solutions are given by the minima of $\beta N J m^2 / 2 - N \ln L$ (where L is now a function of m only), which we identify with the free energy of the system (up to a factor n stemming from the presence of n identical replica in the expression of $[Z^n]$). Whence the free energy per spin (and per replica) is given, up to an irrelevant additive constant, by

$$f(m) = \frac{J m^2}{2} - \frac{kT}{2} \left\{ \log \cosh \frac{Jm - h}{kT} + \log \cosh \frac{Jm + h}{kT} \right\} \quad (8.3)$$

where $J = 2\zeta(1 + \sigma)$. The order parameter m is a solution of the equation of state $df(m)/dm = 0$, i.e., in implicit form,

$$m = \frac{1}{2} \left\{ \tanh \frac{Jm + h}{kT} + \tanh \frac{Jm - h}{kT} \right\}$$

The free energy is depicted in Fig. 8.1 for several field strengths at zero temperature. Two minima are exhibited, which correspond to a so-called spin glass (SG)

8.3. Mean-field theory

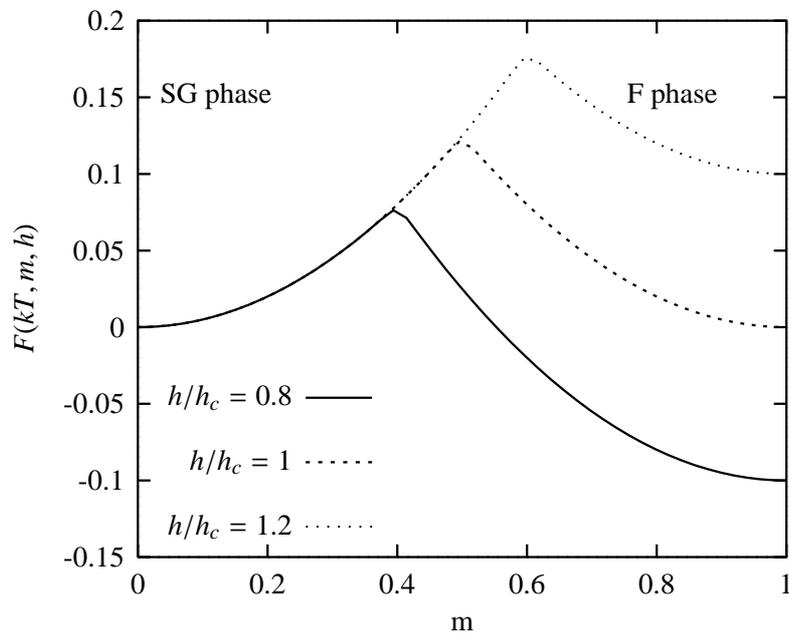


Figure 8.1: Free energy $F(kT, m, h)$ at $T = 0$ for several values of the field amplitude. The ferromagnetic phase (F) is stable up to $h/h_c = 1$ (this defines the critical field amplitude). The spin-glass phase (SG) is metastable at small field amplitudes. This feature is a peculiarity of the bimodal distribution: for a trimodal distribution, this phase becomes unstable at sufficiently low fields.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

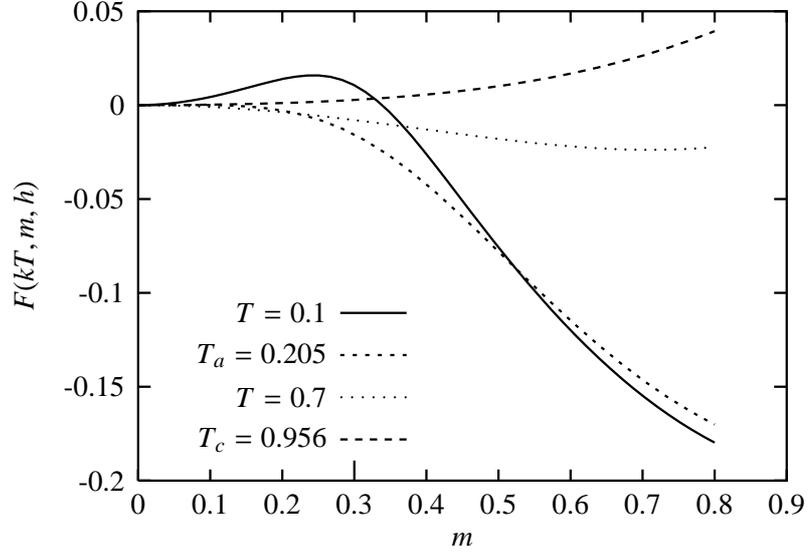


Figure 8.2: Free energy $F(kT, m, h)$ for $h/h_c = 0.3$. The metastable state at $m = 0$ corresponds to the spin glass phase, with spins aligned with their local field, and vanishes above $T_a = 0.205 < T_c$. Temperatures are normalized with respect to the transition temperature at zero field.

phase — or *independent* spin phase — and a ferromagnetic phase (F) respectively. In the SG phase, spins align with their local field independently of their neighbors, which leads to $m = 0$ since the distribution of fields is symmetric. This phase is metastable for h below a critical strength $h_c = \zeta(1 + \sigma)$, ; for $h > h_c$, it is the ferromagnetic phase which becomes metastable.

Graphs of $f(m)$ at several temperatures are shown in Fig. 8.2 and 8.3 for $h/h_c = 0.3$ and $h/h_c = 0.9$, respectively. These two field strengths define two qualitatively distinct regimes: for $h/h_c = 0.3$, the metastable state at $m = 0$ vanishes for a temperature T_a lower than the transition temperature, and the transition is continuous, whereas for $h/h_c = 0.9$ the metastable state exists up to the transition temperature and the transition is first order.

The value of the field at which this change of regime occurs can be computed by performing a Landau expansion of $f(m)$, and monitoring the change of sign in the second derivative of $f(m)$. Expanding $f(m)$ to fourth order in m yields

$$f(m) = am^2 + bm^4 + O(m^5)$$

with

$$a = \frac{J}{2kT} \left\{ -J + kT + J \tanh^2 \frac{h}{kT} \right\}$$

and

$$b = \frac{J^4}{12kT^3} \frac{\cosh(2h/kT) - 2}{\sin^4(h/kT)}$$

8.3. Mean-field theory

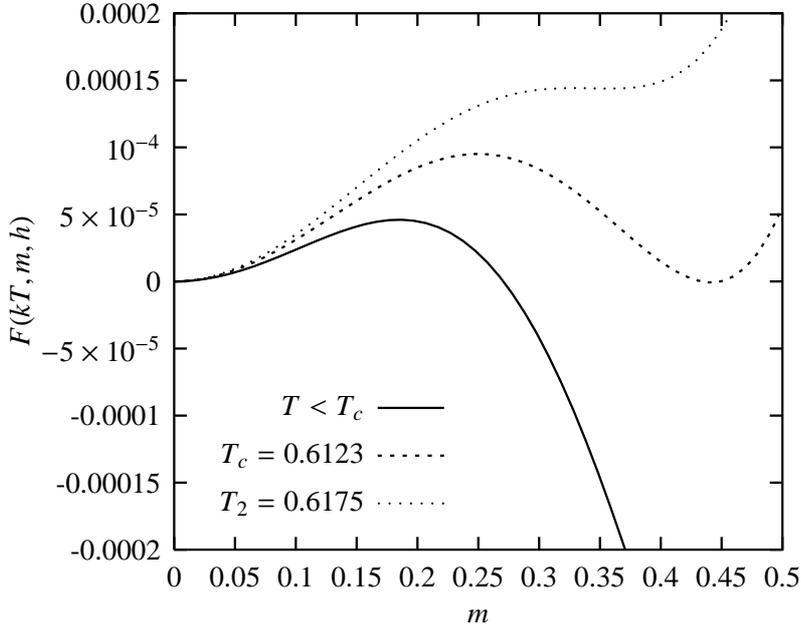


Figure 8.3: Free energy $F(kT, m, h)$ for $h/h_c = 0.9$. The metastable state at $m = 0$ exists from $T = 0$ up to the transition temperature, where it becomes stable. Temperatures are normalized with respect to the transition temperature at zero field.

In the low- m limit, the transition temperature T_c is defined by $a = 0$ (provided $b > 0$), i.e., by the implicit equation $kT_c/J = 1 - \tanh^2(h/kT_c)$. This yields a good estimate of T_c in the second-order regime only; in the first-order regime, I will rely on the exact expression of the free energy, Eq. 8.3.

The metastable state at $m = 0$ (SG phase) disappears above a temperature kT_a such that $F''(m = 0)$ changes sign (it is positive below T_a). This corresponds to the condition that

$$F''(m = 0, kT_a, h) = J - \frac{J^2}{kT \sin^2(h/kT)} = 0.$$

This equation has two solutions for $h/h_c < \kappa = 0.8955$: the first temperature corresponds to the vanishing of the metastable state (T_a in Fig. 8.2), while the second solution corresponds to the transition temperature T_c . In this case, the transition is continuous since the SG phase has lost its stability *before* the transition occurs. For h above the previous value, the equation has no solution, which means that the metastable state exists at all temperatures up to the transition temperature, where it eventually becomes stable. As a result, the transition is of first order, with a coexistence between the SG and the ferromagnetic (F) phases; the ferromagnetic phase becomes unstable at a metastability temperature $T_2 > T_c$ corresponding to the fading away of the associated local minimum.

At $h/h_c = \kappa$, therefore, there is a tricritical point where the transition changes from second to first order. The associated mean-field transition temperature is

$T_c/\zeta(1 + \sigma) = 0.6238$ (Landau expansion leads to $T_c/\zeta(1 + \sigma) = 0.5949$). It is important to note at this point that for $h/h_c > \kappa$, the histogram of the energy shows two spinodal points contrary to $f(m)$ which is associated with a single temperature of metastability T_2 .

8.4 Dimensional argument in the present of long-range interactions

Although the present chapter does not explicitly center on the estimation of the upper critical range σ^* of the model (i.e., there is no phase transition at finite temperature for $\sigma > \sigma^*$, irrespective of the field amplitude), it is nonetheless interesting to calculate it since this sets the range of σ values that should be *a priori* explored in numerical investigations. In addition, the shape of the phase diagram given below (Fig. 8.8) may certainly shed some light on this question, at least qualitatively to begin with.

The most straightforward way to calculate σ^* is by generalizing the dimensional argument proposed by Imry and Ma [173] for the nearest-neighbor model. This argument relies on the stability of the ferromagnetic phase at $T = 0$ against localized excitations made up of isolated domains of, e.g., up-pointing spins in a sea of down-pointing spins. The main lines of the original argument are as follows. As a domain \mathcal{D} of size R is flipped, the energy of the system changes as a result of (i) the onset of a domain wall located on the domain boundary, and (ii) a change in the interaction of the spins belonging to the domain with their local fields. In a nearest-neighbor model of dimension D , the energy change reads

$$\Delta E = \Delta E_{DW} + \Delta E_Z \sim R^{D-1} + \sum_{i \in \mathcal{D}} h_i$$

where the first term denotes the increase in energy due to a domain wall whose length behaves as R^{D-1} , and the second term is a *volume* term ($\sim R^D$) corresponding to the change in Zeeman energy of the domain. Whether fields are distributed according to a gaussian or a bimodal distribution, the contribution of the volume term is zero on average, and one must therefore consider the second moment of the distribution: in the large R limit, and provided fields are not correlated, the central-limit theorem states that the quantity $\sum_{i \in \mathcal{D}} h_i$ converges to a gaussian law with a standard deviation given by $h \sqrt{R^D}$. As a result, the change in energy reads $\Delta E \sim R^{D-1} \pm R^{D/2}$, where the minus sign corresponds to a fluctuation of the random field inside the domain \mathcal{D} that would lead to a negative change in the Zeeman energy; hence to a competition with the domain wall energy (it is always possible to find such a domain in the thermodynamic limit). The ferromagnetic phase remains stable as long as $\Delta E > 0$, i.e., $D > 2$; in this case, only *small* domains may show up (up to a size R_M depending on the field amplitude), that are indeed unable to destabilize the whole ordered phase. This dimensional argument shows that the lower critical dimension of the nearest-neighbor model is $D = 2$. It is

8.4. Dimensional argument in the present of long-range interactions

worth to briefly mention at this point that this argument has been the subject of much criticism, in particular as regards the shape of the domain boundary. It was argued by Binder [37], for instance, that the domain boundary has in fact a much more complex, possibly fractal structure, owing to the tendency of the boundary to “steer” around spins pinned by a large local field — in the case of a gaussian distribution of fields — in order to minimize the total energy of the interface.

Turning to the long-range model, it is clear that the change in Zeeman energy is still given by $hR^{D/2}$. The main change with respect to the nearest-neighbor model is thus related to the domain-wall energy. The argument given hereafter assumes that the domain \mathcal{D} is compact, i.e., made up of *contiguous* spins; in the presence of long-range interactions, this assumption is markedly questionable, in particular in view of the fractal structure of the interface argued in Chap. 7 in the case of first-order transitions (let alone the argument of Binder [37] regarding the fractal structure of the domain wall that was already argued for the nearest-neighbor model and might also hold here). For the sake of simplicity, however, I will stick to this hypothesis in what follows; I will also restrict the discussion to $D = 1$, i.e., the model under investigation in this chapter. Considering a domain of width R (see Fig. 8.4), the change in domain-wall energy is given by

$$\Delta E_{DW} \sim R \sum_{r=1}^{\infty} \frac{1}{r^{1+\sigma}} - \sum_{r=1}^{R-1} (R-r) \frac{1}{r^{1+\sigma}}$$

where the first sum corresponds to the change in energy when R spins are flipped, and the second sum compensates for interactions between spins inside the domain. By transforming sums into integrals, one obtains

$$\Delta E_{DW} \sim \frac{R^{1-\sigma}}{\sigma(1-\sigma)}$$

whereby the upper critical range is given by $\sigma^* = 1/2$ (this value was also reported in [343]). So far, it has been assumed that $R = N$, where N is the number of spins inside the domain. However, if the domain has indeed a fractal structure with a Hausdorff dimension lower than 1, then $N < R$, E_Z would still behave as $N^{1/2}$, yet one might expect that ΔE_{DW} scales more rapidly with N (because the second sum in ΔE_{DW} would grow more slowly), so that one would end up with a larger σ^* . However, the domains also tend to modify their shape (at fixed N) in order to minimize their energy with respect to the random fields, and this would conversely *depress* σ^* ; this is similar to the *domain-roughening* argument and, to some extent, the *domain-within-domain* argument invoked in the nearest-neighbor case [37, 214]. It is interesting to note at this point that the dimensional-reduction argument [3, 273, 359], whereby $D \rightarrow D - 2$ in the presence of random fields, would yield $\sigma^* = 0$, since $D^* = 2\sigma$ (see Sec. 5.1) and $\sigma^* = 1$ for the pure Ising model.

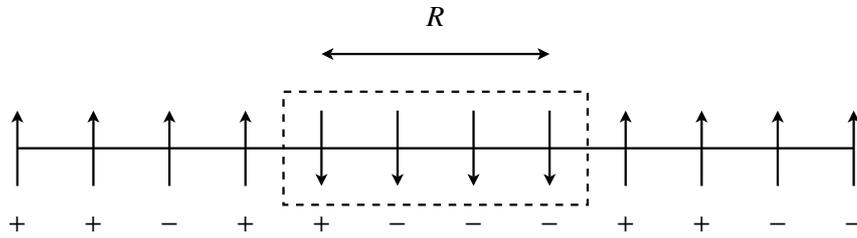


Figure 8.4: Illustration of Imry-Ma dimensional argument for a long-range chain with a bimodal distribution of fields: the stability of an elementary excitation consisting of R contiguous spins is monitored against the decay parameter of the long-range interaction.

8.5 Method

8.5.1 Algorithm

The main objective of the numerical study presented in the next paragraph is the investigation of the order of the phase transition for a long-range RFI model with a *bimodal* distribution of fields. As in the remainder of this thesis, the model is studied numerically by means of a Monte Carlo method operating in the multicanonical ensemble: as far as the RFI model is concerned, it is expected to deal better with the complex free energy landscape exhibited by the model — in particular, the presence of a metastable state at all temperatures $T < T_c$ whenever $h/h_c > \kappa$ — than canonical methods. The model is studied for two decay parameters, $\sigma = 0.1$ and 0.4 , and field amplitudes ranging from $h = 0$ to $h = h_c$, by means of the two following algorithms:

- a single-spin update version of the Wang-Landau algorithm (Sec. 2.5.4), adapted to long-range models (i.e., following the prescriptions presented in Chap. 3 regarding the presence of unequally spaced energy levels), and combined with the transition matrix method; the last method is used to estimate the temperatures of metastability (if at all), see Sec. 4.3.4 and Sec. 7.7.
- a cluster-update version along the line of the breathing cluster method presented in Chap. 7.

As far as the single-spin update algorithm is concerned, the implementation does not differ markedly from that presented in Chap. 7. The only difference amounts to the use of Zhou and Bhatt's prescription [362] instead of the original prescription [334] for estimating the minimum number of histogram entries M per iteration, and for decreasing the modification factor. These prescriptions were claimed — and indeed observed in Chap. 7 — to yield improved accuracy on the estimate of the density of states. The modification factor $\gamma = \ln f$ is initially set to around 20, and divided by a number oscillating between 5 and 10 at the end of each iteration: this choice is clearly empirical, and values are obtained by trial and

8.5. Method

error, based on a couple of short runs for various realizations of disorder. The minimum number of histogram entries per bin is set³ to $1/\sqrt{\gamma}$, and the transition matrix method is used during the first iteration to efficiently predict a (rough) estimate of the density of states (see Sec. 7.7.2).

As regards the cluster-update version, the goal of the present study is clearly modest, and I simply considered a straightforward generalization of the breathing cluster method introduced in the previous chapter. Devising an efficient cluster algorithm specifically tailored to disordered models and which may easily combine with the breathing cluster method is beyond the scope of the present work. Two cluster schemes may be considered:

- clusters can be constructed as if the model were a pure ferromagnetic model (algorithm I)
- or random fields may be included in the cluster construction process as well, i.e., by considering random fields as discrete ± 1 spins and a coupling constant given by the strength of the field; in this case, clusters “pinned” by at least one random field are not flipped (algorithm II).

In the first scheme, the construction of the cluster thus proceeds along the same lines as given in Chap. 7 for the Potts chain, and clusters are (virtually) flipped with an adjustable probability p_{flip} (see below). However, since random fields are not included in the random-cluster representation of the model, the acceptance rate for the cluster-flips from state a to state b (Eq. 7.2) must be modified as follows:

$$A_{flip}(a \rightarrow b) = \min \left\{ 1, \frac{e^{\alpha(E_a)} e^{-\beta(E_b)H_{RF}(b)}}{e^{\alpha(E_b)} e^{-\beta(E_a)H_{RF}(a)}} \prod_{l>0} \left[\frac{p_l(E_b)}{p_l(E_a)} \right]^{B(l)} \right\}, \quad (8.4)$$

H_{RF} denotes the random-field part of the Hamiltonian, i.e., $H_{RF} = -\sum_l \sigma_l h_l$. This term is easily evaluated in both states a and b by computing the number of spins that were *virtually* flipped and the total field in each cluster: both operations are $O(N)$.

This construction scheme is the multicanonical analog of the (canonical) cluster algorithm of Dotsenko [102]. The pitfall is clear: as was already alluded to in [258] in the context of canonical simulations, the presence of disorder depresses the transition temperature below that of the pure model. As a result, clusters constructed using the inverse microcanonical temperature $\beta(E) = dS(E)/dE$ are markedly too large for the same reason as in the canonical implementation: for instance, considering the system at a given energy E corresponding to a microcanonical temperature $T(E) \sim T_c(h)$, with $T_c(h)$ denoting the transition temperature of the disordered model at field amplitude h , the *pure* model at the same temperature is already in its ordered phase because $T_c(h) < T_c(0)$. Since clusters percolate at

³The “original” prescription in [362], i.e., $M \sim 1/\gamma$, turns out to be exceedingly demanding, especially in view of numerous simulations that need to be performed for disordered models.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

the critical temperature for the pure model, this can no longer be the case for the disordered model, and the offset between the optimal cluster size and that given by the microcanonical temperature obviously increases with the field strength. Yet there is more to be fathomed out in the context multicanonical simulations: the acceptance rate — hence, in some respect, the autocorrelation time — of a multicanonical update scheme is optimal if the proposal-move scheme picks rare events and non-rare events with equal chance, i.e., according to the random walker picture, picks left- and rightwards moves (on the energy axis) with equal weight. This is clearly what the cluster-update scheme of the breathing cluster method does for a ferromagnetic model: turning back to the RFI model, if clusters are too large, low energy states will be picked more often than high energy ones, and this clearly reduces the acceptance rate. Similarly, the acceptance rate is also reduced if clusters are too small (this can be easily explained by considering clusters of unit size, where we are back to simple Monte Carlo sampling).

One way to get around the issue is to decrease the bond probability, i.e., to decrease the inverse temperature $\beta(E)$ that governs the cluster construction, since this lowers the average cluster size⁴. However, this must be done empirically — in the absence to date of any *smarter* scheme —, by scaling $\beta(E)$ globally until the minimum *effective* autocorrelation time is reached⁵. Let alone the fact that scaling $\beta(E)$ globally does certainly not yield the *optimal* bond probability at *all* energy levels, this approach is clearly tedious: it turns out that the procedure is efficient at modest lattice sizes and low field strengths (up to $h/h_c \sim 0.3$), because the efficiency of the algorithm is only moderately affected by a value of $\beta(E)$ departing from the ideal value; this becomes clearly cumbersome, however, above $L \sim 400$. At medium field, the performance is illustrated in Table 8.1, where effective autocorrelation times (at the transition temperature) are indicated for several scaling values: as can be seen, the maximum attainable acceptance rate for $L = 256$ hardly exceeds 10%, which clearly shows that scaling $\beta(E)$ is not enough to obtain an optimal move-proposal scheme, at least at medium and large field strengths.

At medium and large lattice sizes (and medium field strengths), algorithm II proves more efficient. Here, bonds are activated not only between spins, but also between spins and random fields (akin to “ghost-spins”), using a bond probability given by

$$p_{ij}(E) = \delta_{\sigma_i, \sigma_j} \left(1 - e^{-\beta(E)|J_{ij}|}\right)$$

for the former type of bonds, and

$$p_i(E) = \delta_{\sigma_i, h_i/h} \left(1 - e^{-\beta(E)h}\right)$$

⁴Another similar method consists in shifting the function $\beta(E)$ toward the left of the energy axis.

⁵The same procedure was proposed, although in the context of a canonical cluster algorithm, in a numerical study of the frustrated Coulomb model [138].

8.5. Method

Table 8.1: Effective autocorrelation times at the transition temperature for $\sigma = 0.4$ and $h/h_c = 0.3$, and $L = 128, 256$ and 512 . (SSU) refers to the single-spin update algorithm. (I) indicate estimates obtained with algorithm I for several scaling factors s applied to the microcanonical temperature $\beta(E)$ entering bond probabilities. (II) refers to the ghost-spin algorithm. A and B indicate the acceptance rate and the fraction of the lattice occupied by the largest cluster, respectively.

L		SSU	I: s=1.1	s=1.0	s=0.9	s=0.8	II:
128	τ_{eff}	20	32	21	16	40	30
	A(%)	35	12	15	14	7	58
	B(%)		68	63	57	47	64
256	τ_{eff}	55	800	670	590	1000	90
	A(%)	35	8	10	8.5	5	55
	B(%)		56	52	48	41	60
512	τ_{eff}	150					205

for the latter⁶. In this case, the acceptance rate is simply given by Eq. 7.2, yet taking bonds between spins and random fields into account, i.e.

$$A_{\text{flip}}(a \rightarrow b) = \min \left\{ 1, \frac{e^{\alpha(E_a)}}{e^{\alpha(E_b)}} \prod_{l>0} \left[\frac{p_l(E_b)}{p_l(E_a)} \right]^{B(l)} \left[\frac{q_b}{q_a} \right]^C \right\} \quad (8.5)$$

where $q_a = e^{\beta(E_a)h} - 1$, and C is the number of bonds that have been activated between spins and random fields. Clusters that have at least one active bond with a random field are said “frozen”, and are not flipped.

It turns out that constructing clusters by taking the disorder into account yields substantially larger acceptance rates than with algorithm I: clusters are still constructed using the microcanonical temperature $\beta(E)$, yet large clusters are never flipped because they most often end up pinned by a random field. To be specific, the probability that a cluster of size B be frozen is given by $1 - e^{-\beta(E)H}$, where $H = \sum_{i \in \mathcal{B}} h_i$ is the total field felt by the cluster, and this quantity increases exponentially with the square root of the cluster size. As a result, only small clusters are flipped, which is exactly what we were seeking when decreasing $\beta(E)$ in algorithm I. As shown in Table 8.1, however, the scaling is not markedly better than with a single-spin update (SSU) implementation for the fields considered: the dynamic exponents amount to $z = 1.4$ and $z = 1.39$ for the SSU and the cluster-update algorithms, respectively. Overall, and although cluster updates turn out to be more efficient at lower field amplitudes, I found the balance to be clearly tipped in favor of the SSU algorithm at larger fields. Nonetheless, it should be remarked that canonical algorithms behave poorly in this region, so that the power law behavior obtained for both move updates represents in any case a clear improvement.

⁶These probabilities take into account a factor $\frac{1}{2}$ that has been included in the definition of the Hamiltonian, i.e., pair interactions are actually Kronecker delta functions.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

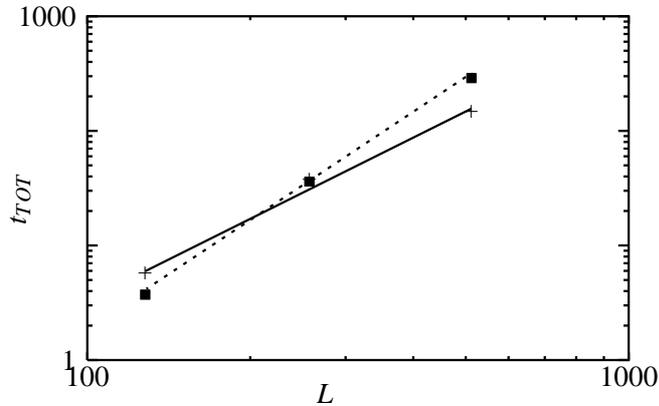


Figure 8.5: Total CPU time needed to perform a Monte Carlo simulation delivering the *same* number of uncorrelated measurements, with respect to the lattice size. Single-spin update: dashed line; cluster-update: solid line. Units on the vertical axis are arbitrary.

While dynamic characteristics are more or less similar for both move updates, it must be reminded, however, that the benefits of cluster algorithms are twofold when simulating long-range models: there is also an associated reduction in the number of operations required to compute the energy, from N^2 to $N \ln N$. In view of the prefactors, the cluster-update version becomes more efficient above $L \sim 200$ (see Fig. 8.5), although the boundary clearly depends on the type of Fast Fourier Transform implementation used to compute the energy. To sum up, cluster updates represent an efficient update scheme up to medium field strengths, yet essentially as a result of the reduction of the algorithm complexity rather than in terms of dynamic behavior.

8.5.2 Influence of periodic boundary conditions

An appropriate choice of periodic boundary conditions proves crucial to investigate the model at large fields. Indeed, as the amplitude of the random field increases, the transition temperature dies off to zero; on the line $T = 0$, there is a critical value h_c above which the ferromagnetic ground state becomes unstable, and spins reduce their energy by aligning with their local random field. While in models with nearest-neighbor interactions, the critical value h_c does not depend on the lattice size L , but merely on the coordination of the lattice, this is no longer the case for long-range models: the (long-range) ferromagnetic part of the interaction that keep all spins lined up in the same direction, increases with the lattice size, and as a result the smallest field amplitude needed to flip a spin pointing in a direction opposed to its local random field increases as well.

Formalizing this a little, the critical field amplitude h_c for a chain of size L is given by $h_c(L) = \sum_{r=1}^L J(r)$, with $\lim_{L \rightarrow \infty} h_c(L) = \zeta(1 + \sigma)$. If first-image periodic boundary conditions (FIPBC) are being used, then $J(r) = 1/r^{1+\sigma}$ (truncated to

8.6. Phase diagram of the long-range RFI model

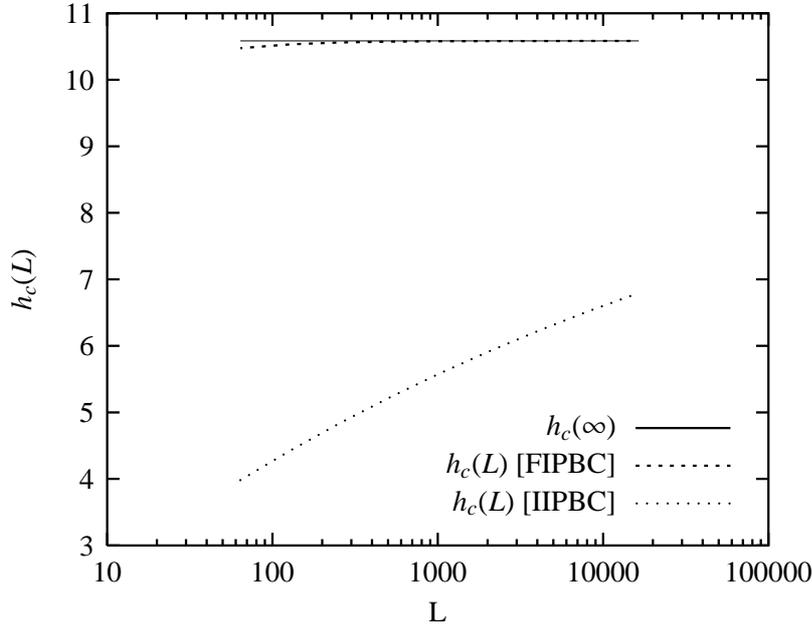


Figure 8.6: Graph of the critical field $h_c(L)$ with respect to the lattice size for $\sigma = 0.1$ and two choices of periodic boundary conditions: infinite image periodic boundary conditions (IIPBC), and first-image periodic boundary conditions (FIPBC).

$r < L/2$) and $h_c(L)$ increases at a drastically low pace with the lattice size; this is clearly visible in Fig. 8.6 for $\sigma = 0.1$. As a result, investigating the region near $T = 0$ imposes that chains of huge sizes should be simulated lest no phase transition occurs. In this respect, using infinite image periodic boundary conditions (IIPBC), as described in Sec. 4.2.1 dramatically improves the convergence. In this case, each spin interacts with an infinite number of periodic replica of the chain, and $J(r)$ is replaced by an effective interaction $\tilde{J}(r)$ given by Eq. 4.7. As can be seen in Fig. 8.6, $h_c(L)$ converges much more rapidly to the infinite-size value, so that investigations of the region ($h \sim h_c, T \sim 0$) are clearly feasible for sizes below $L \sim 1000$.

8.6 Phase diagram of the long-range RFI model

Simulations were carried out for $\sigma = 0.1$ and $\sigma = 0.4$, i.e., near the expected onset of the mean-field regime, and in the vicinity of the upper critical range, respectively. In order to ensure that every realization of disorder exhibits a phase transition, the zero total field constraint was enforced, $\sum_{i=1}^L h_i = 0$. Lattice sizes between $L = 64$ and $L = 1024$ were investigated, and between 20 and 50 realizations of disorder (i.e., samples) were generated, depending on the lattice size; at large fields, I had to rely on more samples owing to a larger dispersion in the data

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

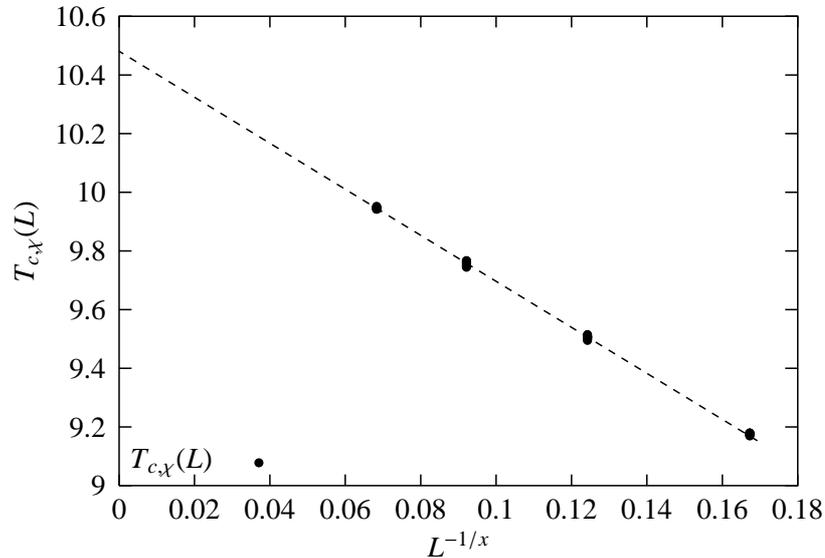


Figure 8.7: Fit of finite-size temperatures $T_{c,\chi}(L)$ to the power law $T_{c,\chi}(L) - T_{c,\chi}(\infty) \propto L^{-1/x}$, for $\sigma = 0.1$ and $h = 1.0$.

points — as expected. Algorithm II was used for fields $h/h_c < 0.3$, and the single-spin update algorithm at larger fields. Equilibrium time was checked by relying on the χ^2 -regression scheme described in Sec. 7.4.

Transition temperatures In order to determine the transition temperatures, I rely on the scaling behavior of peaks of the specific heat, the susceptibility, and cross-cumulants V_1 and V_2 of the magnetization (see Chap. 4), and estimate the infinite size temperature by performing a fit of *all* data points (i.e., for each realization of disorder) to a power law of the form $T_c(L) - T_c(\infty) \propto L^{-1/x}$, where x is adjusted so as to yield the lowest error: it turns out that, in the second-order regime, setting $x = \nu$, where ν is the correlation length exponent given by the finite-size scaling behavior of V_1 and V_2 , leads to slightly underestimated temperatures and poorer fits. A fit to the peaks of the susceptibility is illustrated in Fig. 8.7 for $\sigma = 0.1$ and $h/h_c = 0.0945$: the sample-to-sample dispersion is small; this is no longer the case at large fields, and the fact that fits are performed over four to five lattice sizes, and with a rather modest number of realizations of disorder, produces quite large error bars. The corresponding estimates of transition temperatures are depicted in Fig. 8.8.

Order of the phase transition The method used to detect the order of the phase transition is the spinodal method presented in Sec. 4.3. The microcanonical temperature $\beta(E)$ is computed using the transition matrix method, and temperatures of metastability, if any, are determined from the extrema of $\beta(E)$ in the critical region. This is depicted in Fig. 8.9 for $\sigma = 0.1$ and $h/h_c = 0.945$ ($h_c = \zeta(1 + \sigma)$): while

8.6. Phase diagram of the long-range RFI model

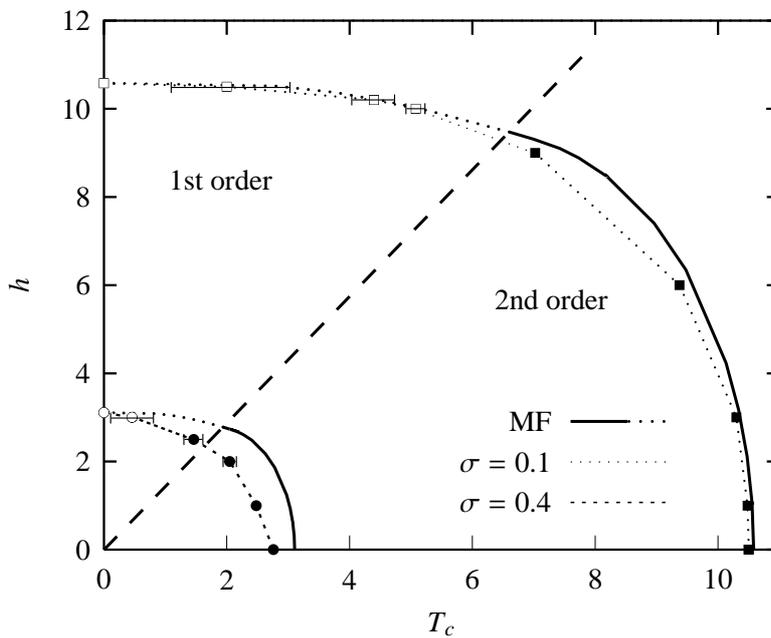


Figure 8.8: Phase diagram of the random-field Ising chain with $1/r^{1+\sigma}$ interactions, for $\sigma = 0.1$ and 0.4 . Open symbols indicate a first-order transition, while filled symbols refer to a second-order transition (lines are guides to the eyes and errors are smaller than the size of symbols except were explicitly shown). The thick, dotted and solid lines show the mean-field predictions in the first- and second-order regime, respectively. The thick dashed line shows the boundary between both regime in the mean-field approximation.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

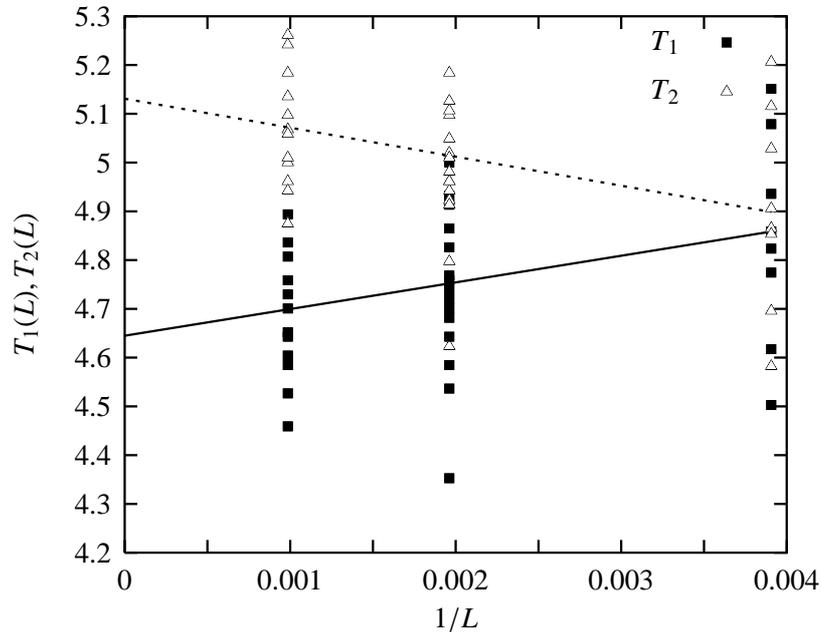


Figure 8.9: Fit of the temperatures of metastability $T_1(L)$ and $T_2(L)$ to a power law of the lattice size, for $\sigma = 0.1$ and $h = 10.0$. In spite of the strong dispersion of the data points, the transition is clearly of the first order.

the sample-to-sample variation is large, the finite-size scaling behavior of the two temperature of metastability clearly indicates that the transition is first-order. The nature of the phase transition for the field amplitudes investigated in the present work is indicated in Fig. 8.8 as open and filled symbols for first- and second-order transitions, respectively. In view of the modest lattice sizes investigated, the onset of a first-order transition⁷ cannot be definitely ruled out, even if the transition is marked as being continuous in the phase diagram: this is particularly true regarding the points that lie on the boundary between the two regimes.

The coexistence between the two phases is illustrated in Fig 8.10: the bottom part of the figure depicts the two-peak structure characteristic of phase coexistence, while the top part shows the joint distribution of the random-field part of the interaction, E_{RF} , with respect to the lattice energy. In the upper energy range, both E_{RF} and the order parameter (not shown) are close to zero, showing that spins behave paramagnetically. As the critical energy range is approach, the random-field part of the interaction increases, signaling a tendency of spins to align with their local random field (the independent spin phase): this tendency is reminiscent of the metastable state existing at $m = 0$ in the mean-field approximation. Simultaneously, near $E/L \sim -14.3$, the ferromagnetic phase develops, characterized by a large magnetization and a low value of E_{RF} . The interface tension that is clearly visible on the reweighted histogram of the energy near $E/L \sim -14.3$ corresponds

⁷In particular if it happens to be driven by fluctuations

8.7. Discussion and perspectives

to the coexistence between these two phases.

8.7 Discussion and perspectives

Phase diagram Mean-field results are indicated along with the other estimates in the phase diagram shown in Fig. 8.8. In particular, the thick dashed line is a reminder of the boundary between first- and second-order transitions as obtained from the mean-field approximation.

For $\sigma = 0.1$, estimates of transition temperatures match well the mean-field estimates, although the discrepancy is slightly larger at medium field amplitudes; the same behavior is observed in a slightly more marked way for $\sigma = 0.4$, where the curvature of $T_c(h)$ is smaller than for $\sigma = 0.1$. This behavior may very well signal the approach of the upper critical range, since at $\sigma^* = 0.5$, $T_c(h = 0) \sim 2.182$ and $h_c = 2.612$, so that the curve $T_c(h)$ must in any case terminate at these two points: one may expect that the curvature does not change abruptly right at the upper critical range, but rather gradually changes sign. This clearly suggests to further investigate the behavior of the model for $0.4 < \sigma < 0.5$.

As regards the position of the tricritical point, the large error bars do not allow to draw more than mere qualitative conclusions. For $\sigma = 0.1$, the position of the boundary is consistent with the mean-field prediction; for $\sigma = 0.4$, however, the continuous nature of the transition for $h = 2.5$ suggests that the tricritical point shifts towards the low temperature region. This is in contrast with the findings of Khurana et al. [202] obtained from a high-temperature expansion of the nearest-neighbor model, showing that, while the tricritical point persists in lower dimensions, it also appears at weaker fields (with respect to the critical field amplitude): this means that the boundary is pulled towards the large temperature region, a feature also suggested by Monte Carlo Renormalization Group studies [281]. Although a fluctuation-driven first-order transition cannot be totally ruled out for $\sigma = 0.4$ and $h = 2.5$, seeing that the investigation were restricted to $L < 1000$, this could also imply that the nearest-neighbor and the long-range models do not behave *quantitatively* the same way in this respect. Finally, and as regards the adequacy with mean-field results, it would be very appealing to check (i) the value of σ at which a classical critical behavior sets in, if at all, and (ii) the exactness of the mean-field behavior for $\sigma < 0$.

Dynamics of the algorithm at large fields and possible improvements For the two largest fields considered for $\sigma = 0.1$, namely, $h = 10$ and $h = 10.5$, the transition is first-order except for a few samples which do not exhibit a transition at all: the magnetization remains close to zero, with a smooth, light bump instead of the usual jump observed for other samples. The same behavior occurs for $\sigma = 0.4$ and $h = 3$. Since $h < h_c$ and the zero total field condition ensures that there must be a phase transition, such a behavior must be attributed to the poor phase space exploration performed by the algorithm at low energy for some realizations of disorder.

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

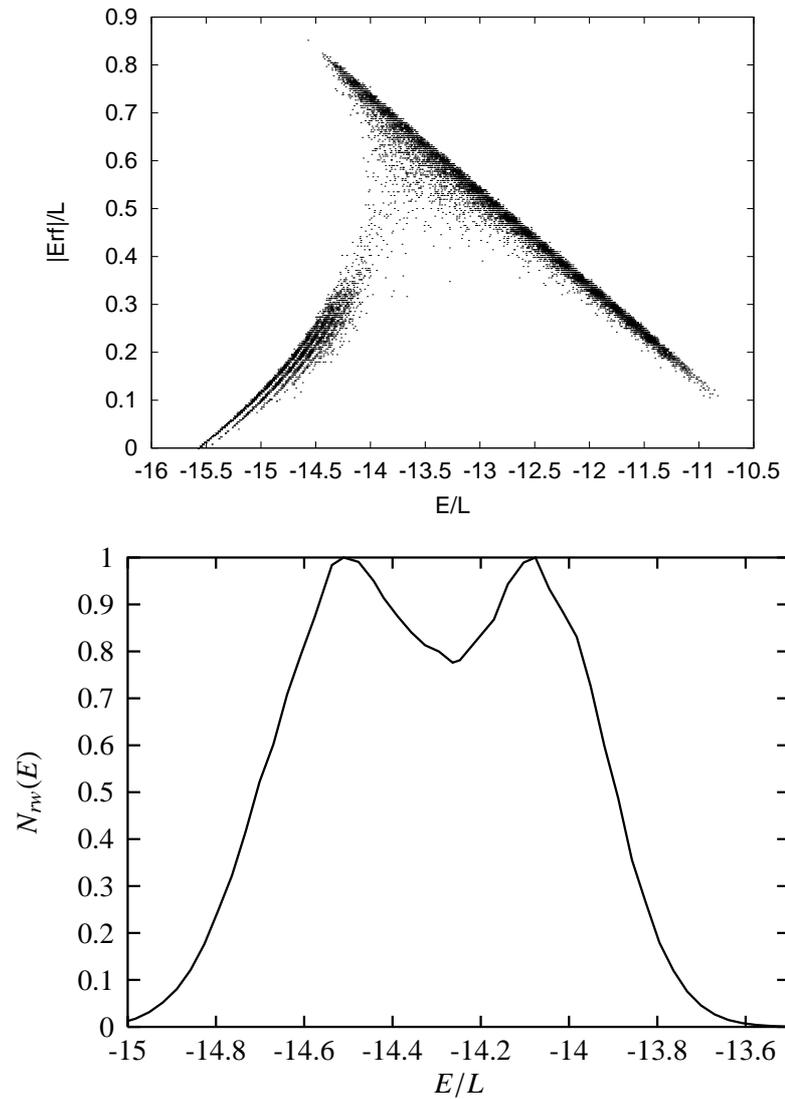


Figure 8.10: Top: Joint distribution of the (normalized) random-field part of the interaction (i.e., $(1/L) \sum_l \sigma_l h_l$) with respect to the total energy per spin E/L , for $\sigma = 0.1$, $h = 10$ and $L = 512$. Bottom: Reweighted histogram of the energy at the equal-height transition temperature, for the same sample.

8.8. Conclusion

For field amplitudes very close to the critical amplitude h_c , the two phases coexist down to the vicinity of the ground state (see Fig. 8.11), and the algorithm sometimes fails to sample the ferromagnetic phase efficiently: as the random walker travels down the energy axis, the system remains in the independent-spin phase although the ferromagnetic phase is more stable. In this respect, a multicanonical weighting is indeed of little help, since the two phases have the very close energies; the behavior is a sort of critical slowing down, yet in the space of the E_{RF} variable. The “faulty” behavior must therefore be attributed (i) to the type of move updates utilized (although both types of move fail in this respect), and (ii) the fact that the path “connecting” the two phases (around $E/L \sim -14.5$ in the figure) is not sufficiently given weight. This suggests several possible improvements:

- First, one may rely on the recently proposed optimal ensemble [319, 351] to increase the weight in the region connecting the two phases; this indeed would increase the diffusion current of the random walker in this region; yet an adequate optimal weight remains to be found that takes into account the particular dynamics engendered by cluster updates: these are necessary to reach larger sizes owing to the recurrent algorithm complexity problem. In this respect, optimizing $\beta(E)$ as well as the weight $w(E)$, in the spirit of the optimal ensemble algorithm, would certainly be very promising.
- Second, and perhaps in connection with the previous point, one may perform a bidimensional random-walk, where the first axis would correspond to the ferromagnetic energy and the second axis to the random-field interaction; as a by-product, and provided the phase space is efficiently sampled, this would make it possible to obtain reweighted averages for any field strength. As far as cluster updates are concerned, it remains to be seen how an appropriate bond probability may be devised, e.g., since the presence of two dimensions in the density of states makes this point highly non-trivial. Along the same line, an extension of the transition matrix method to a bidimensional space may also be of interest, in the same spirit as the extension proposed by Yamaguchi et al. [354].

8.8 Conclusion

In this chapter, multicanonical algorithms were used to investigate the phase diagram of a long-range random-field Ising chain with a bimodal distribution of fields. The prominent result of this study is the existence of a tricritical point for the two decay parameters investigated ($\sigma = 0.1$ and 0.4), with a first-order transition occurring for field amplitudes larger than a threshold value depending on the decay parameter σ . This is in *qualitative* agreement with mean-field predictions and with the behavior reported in a few studies of the nearest-neighbor model in three dimensions. While at $\sigma = 0.1$, the transition line $T_c(h)$ separating ordered and disordered phases is in very good *quantitative* agreement with mean-field predictions,

Chapter 8. Long-range Ising chain with bimodal random fields: first-order transitions induced by disorder

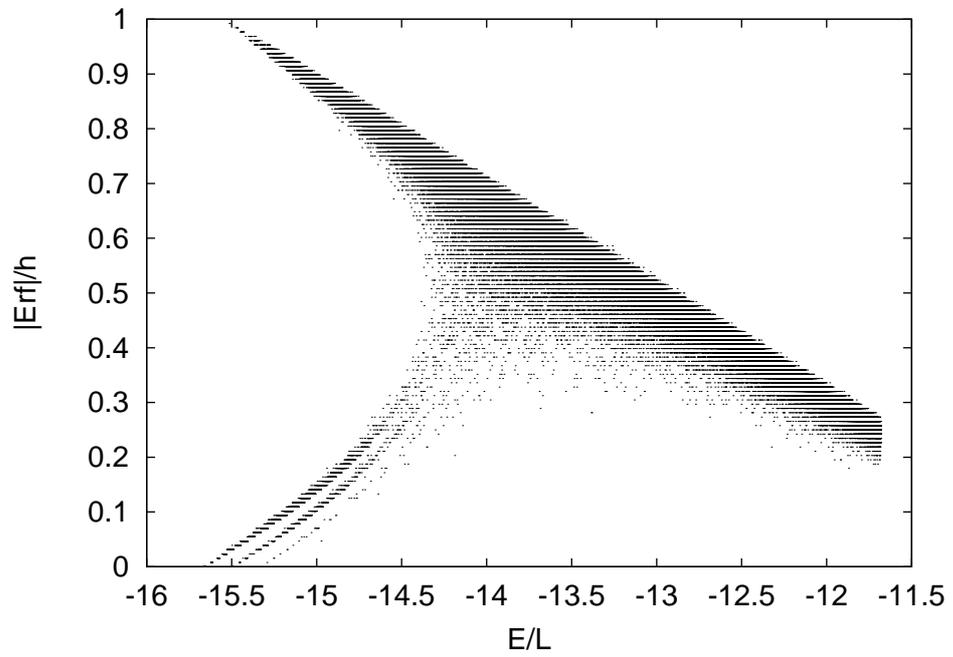


Figure 8.11: Joint distribution of the (normalized) random-field part of the interaction with respect to the total energy per spin E/L , for $\sigma = 0.1$, $h = 10.2$ and $L = 256$.

8.8. Conclusion

this is no longer the case for $\sigma = 0.4$: the area under this line is markedly reduced, and the curvature is weaker. The weakening of the curvature suggests the approach of the critical upper range $\sigma^* = 0.5$ obtained from a generalization of Imry-Ma argument.

Two multicanonical algorithms were used in this study: a single-spin update version of the Wang-Landau algorithm adapted to long-range models, and an extension of the breathing cluster method to disordered models. Both algorithms efficiently explore the phase space in spite of the rugged energy landscape of the model: this represents a major improvements over canonical methods, that are known to produce very repetitive dynamics — especially at large field strenghts — for this class of models. However, the benefit brought about by cluster updates only amounts to the reduction of the algorithm complexity connected with the computation of the energy: no marked gain was observed regarding the dynamic characteristics, and prefactors favor the cluster-update version only for sizes above 500 spins. In view of this, a combination of the breathing cluster method with a cluster-update scheme specifically tailored to disordered models, with the optimized ensemble [319, 351], or with a two-dimensional random walk, would be worthy of investigation.

General conclusion and developments

The essential message that has guided this work is perhaps central to computational physics: the interplay between models and methods is a two-way trade.

For a given model, it is crucial that the most appropriate method should be chosen to investigate its physical behavior. If such a method does not exist, then an existing one may be adapted (Berg's multicanonical recursions, Chap. 3), or a new method devised from scratch (the "spinodal" method, Chap. 4), or made up by combining existing methods, if necessary by creating a bridge between them (clusters, Chap. 7). It turns out that long-range models, because they are highly demanding in many respects, may be regarded as a very efficient spur to the development of new methods. Conversely, they may also be considered an ideal testbed for existing *general* methods: the ability of the multicanonical method to tackle a variety of phase transitions with equal ease (Chap. 4 and 5) is clearly promising in this respect.

On the other hand, methods that are initially devised to resolve a specific issue for a given model may prove efficient at tackling other models from a new perspective: the detection of the order of phase transitions (Chap. 4) and the Breathing Cluster method (Chap. 7, yielding better estimates of the surface tension of the nearest-neighbor Potts model) are examples of this. Here, a promising line of research is the investigation of the Potts models defined on a fractal lattice⁸, where the Hausdorff dimension — among other fractal parameters — is expected to influence the order of the phase transition in a way similar to the decay parameter in long-range models. Extensions to disordered models (Chap. 8), however, show that the interplay between models and methods is an unfinished story: the Breathing Cluster method must be improved at large fields. A possible way is the combination with the optimized ensemble. Since there is no objection in principle to using a different temperature than the microcanonical temperature for the bond probabilities, optimizing this temperature to obtain the best performance would be

⁸Collaboration with P. Monceau, Pôle Matière et Systèmes Complexes, UMR 7057 CNRS/Paris 7.

an interesting way to improve the method.

The second message conveyed by this thesis is perhaps of fundamental relevance to the physics of critical phenomena: the ability to continuously vary the *effective* dimension of long-range models by means of the decay parameter makes them perfect tools for the investigation of crossover effects. The q -state Potts model is enlightening in this respect, exhibiting numerous crossovers, from the non-extensive to the extensive regime (Sec. 4.2), from the first- to the second-order regime (Sec. 4.3), from continuous to topological transitions (Sec. 4.3.6), and from long-range to short-range behavior (Chap. 5). The crossover from the first- to the second-order regime that was investigated in the Random Field Ising model (Chap. 8) shows that long-range models may also serve as a guide in the investigation of their short-range counterpart.

However, the connection between both classes of model seems to be merely qualitative. First, no quantitative relation has yet been found between the effective dimension and the decay parameter that may enable a direct transposition to short-range models of the results obtained in long-range models. Second, each particle in a long-range model interacts with the entire system, including the boundaries. This leads to startling, unusual finite-size effects (Sec. 4.3.7), suggesting that fractal geometry may also play an important role in shaping the behavior of these models at a discontinuous transition (Chap. 7). This evokes several exciting developments: if indeed the mixed-phase interface is characterized by a fractal structure, what are its Hausdorff dimension, its lacunarity and its connectivity? The investigation of the cluster distribution and the estimation of correlation lengths would certainly be useful to glean more information in this respect; this is perfectly within reach of the method introduced in Chap. 7. On the analytical side, this also suggests the need to develop a rigorous theory of finite-size scaling at first-order transitions for the long-range Potts model: with this in mind, the two-dimensional Potts model with long, but *finite*, range interactions is currently being investigated, both numerically and analytically (by means of a contour-based method)⁹. Finally, beyond the equilibrium properties of these models, out-of-equilibrium properties, and particularly nucleation, might be worthy of deep scrutiny, as mixed-phase interfaces play a crucial role in this mechanism.

⁹Collaboration with T. Gobron, LPTM.

Bibliography

- [1] S. Abe, S. Martínez, F. Pennini, and A. Plastino. *Nonextensive thermodynamics relations*. Phys. Lett. A **281**, 126 (2001).
- [2] A. Aharony. *Tricritical points in systems with random fields*. Phys. Rev. B **18**, 3318 (1978).
- [3] A. Aharony, Y. Imry, and S. K. Ma. *Lowering of dimensionality in phase transitions with random fields*. Phys. Rev. Lett. **37**, 1364 (1976).
- [4] A. Aharony and E. Pytte. *First- and second-order transitions in the Potts model near four dimensions*. Phys. Rev. B **23**, 362 (1981).
- [5] M. Aizenman, J. T. Chayes, L. Chayes, and C. M. Newman. *The phase boundary in dilute and random Ising and Potts ferromagnets*. J. Stat. Phys. **50**, 1 (1988).
- [6] M. Aizenman and R. Fernández. *Critical exponents for long-range interactions*. Lett. Math. Phys. **16**, 39 (1988).
- [7] S. Alder, S. Trebst, A. K. Hartmann, and M. Troyer. *Dynamics of the Wang-Landau algorithm and complexity of rare events for the three-dimensional bimodal Ising spin glass*. J. Stat. Mech. **7**, P07008 (2004)..
- [8] R. Allenspach and A. Bischof. *Magnetization direction switching in Fe/Cu(100) epitaxial films: temperature and thickness dependence*. Phys. Rev. Lett. **69**, 3385 (1992).
- [9] M. P. Almeida. *Thermodynamical entropy (and its additivity) within generalized thermodynamics*. Physica A **325**, 426 (2003).
- [10] D. J. Amit. *Modeling brain functions* (Cambridge University Press, Cambridge, England, 1989).
- [11] P. W. Anderson and G. Yuval. *Some numerical results on the Kondo problem and the inverse square one-dimensional Ising model*. J. Phys. C **4**, 607 (1971)..

Bibliography

- [12] C. Anteneodo and C. Tsallis. *Breakdown of exponential sensitivity to initial conditions: role of the range of interactions*. Phys. Rev. Lett. **80**, 5313 (1998).
- [13] J. Asikainen, A. Aharony, B. B. Mandelbrot, E. Rausch, and J.-P. Hov. *Fractal geometry of critical Potts clusters*. Eur. Phys. J. B **34**, 479 (2003).
- [14] F. Baldovin, E. Brigatti, and C. Tsallis. *Quasi-stationary states in low-dimensional hamiltonian systems*. Phys. Lett. A **320**, 254 (2004).
- [15] M. Barati and A. Ramazani. *Magnetic phase diagram of an alternating Ising chain with long-range interactions*. Phys. Rev. B **64**, 024407 (2001).
- [16] M. Barati and A. Ramazani. *Classical phase diagram of $s \geq \frac{1}{2}$ Ising chains with long-range interactions: finite-range scaling*. Phys. Rev. B **65**, 012406 (2002).
- [17] R. J. Baxter. *Potts model at the critical temperature*. J. Phys. C **6**, L445 (1973)..
- [18] E. Bayong and H. T. Diep. *Effect of long-range interactions on the critical behavior of the continuous Ising model*. Phys. Rev. B **59**, 11919 (1999).
- [19] E. Bayong, H. T. Diep, and V. Dotsenko. *Potts model with long-range interactions in one dimension*. Phys. Rev. Lett. **83**, 14 (1999).
- [20] D. P. Belanger. *Experiments on the random field Ising model*. In *Spin glasses and random fields* (World Scientific, Singapore, 1997).
- [21] M. Le Bellac, *Des Phénomènes Critiques aux Champs de Jauge* (InterEditions/Editions du CNRS, 1992).
- [22] C. H. Bennett. *Efficient estimation of free energy differences from Monte Carlo data*. J. Comput. Phys. **22**, 245 (1976).
- [23] B. A. Berg. *Multicanonical recursions*. J. Stat. Phys. **82**, 323 (1996).
- [24] B. A. Berg. *Introduction to multicanonical Monte Carlo simulations*. Fields Inst. Commun. **26**, 1 (2000).
- [25] B. A. Berg. *Generalized ensemble simulations for complex systems*. Comp. Phys. Commun. **147**, 52 (2002).
- [26] B. A. Berg, A. Billoire, and W. Janke. *Numerical study of the two-replica overlap of the 3d Edwards-Anderson Ising spin glass*. Physica A **321**, 49 (2003).
- [27] B. A. Berg and T. Celik. *New approach to spin-glass simulations*. Phys. Rev. Lett. **69**, 2292 (1992).

Bibliography

- [28] B. A. Berg and U. Hansmann. *Simulation of an ensemble with varying magnetic field: a numerical determination of the order-order interface tension in the $d = 2$ Ising model*. Phys. Rev. B **47**, 497 (1993).
- [29] B. A. Berg and W. Janke. *Multioverlap simulations of the 3d Edwards-Anderson Ising spin glass*. Phys. Rev. Lett. **80**, 4771 (1998).
- [30] B. A. Berg and T. Neuhaus. *Multicanonical algorithms for first order phase transitions*. Phys. Lett. B **267**, 249 (1991).
- [31] B. A. Berg and T. Neuhaus. *Multicanonical ensemble: a new approach to simulate first-order phase transitions*. Phys. Rev. Lett. **68**, 9 (1992).
- [32] B. Bergersen and Z. Rácz. *Dynamical generation of long-range interactions: random Lévy flights in the kinetic Ising and spherical models*. Phys. Rev. Lett. **67**, 3047 (1991).
- [33] J. Bhattacharjee, S. Chakravarty, J. L. Richardson, and D. J. Scalapino. *Some properties of a one-dimensional Ising chain with an inverse-square interaction*. Phys. Rev. B **24**, 3862 (1981).
- [34] A. Billoire, R. Lacaze, and A. Morel. *A numerical study of finite-size scaling for first-order phase transitions*. Nucl. Phys. B **370**, 773 (1992).
- [35] K. Binder. *Critical properties from Monte Carlo coarse graining and renormalization*. Phys. Rev. Lett. **47**, 693 (1981).
- [36] K. Binder. *Monte Carlo calculation of the surface tension for two- and three-dimensional lattice-gas models*. Phys. Rev. A **25**, 1699 (1982).
- [37] K. Binder. *Random-field induced interface widths in Ising systems*. Z. Phys. B **50**, 343 (1983).
- [38] K. Binder and D. P. Landau. *Finite-size scaling at first-order phase transitions*. Phys. Rev. B **30**, 1477 (1984).
- [39] K. Binder and E. Luijten. *Monte Carlo tests of theoretical predictions for critical phenomena: still a problem?* Comp. Phys. Commun. **127**, 126 (2000).
- [40] K. Binder and E. Luijten. *Monte Carlo tests of renormalization-group predictions for critical phenomena in Ising models*. Phys. Rep. **344**, 179 (2001).
- [41] K. Binder, E. Luijten, M. Müller, N. B. Wilding, and H. W. J. Blöte. *Monte Carlo investigations of phase transitions: status and perspectives*. Physica A **281**, 112 (2000).
- [42] M. Biskup, L. Chayes, and N. Crawford. *Mean-field driven first-order phase transitions in systems with long-range interactions*. math-ph/0501067.

Bibliography

- [43] H. W. J. Blöte, J. R. Heringa, and E. Luijten. *Cluster Monte Carlo: extending the range*. Comp. Phys. Commun. **147**, 58 (2002).
- [44] B. M. Boghosian. *Thermodynamic description of the relaxation of two-dimensional turbulence using Tsallis statistics*. Phys. Rev. E **53**, 4754 (1996).
- [45] I. Booth, A. B. MacIsaac, J. P. Whitehead, and K. De’Bell. *Domain structures in ultrathin magnetic films*. Phys. Rev. Lett. **75**, 950 (1995).
- [46] C. Borgs and W. Janke. *An explicit formula for the interface tension of the 2d Potts model*. J. Phys. I (Paris) **2**, 2011 (1992)..
- [47] C. Borgs and W. Janke. *New method to determine first-order transition points from finite-size data*. Phys. Rev. Lett. **68**, 1738 (1992).
- [48] C. Borgs and R. Kotecky. *A rigorous theory of finite-size scaling at first-order transitions*. J. Stat. Phys. **61**, 79 (1990).
- [49] C. Borgs, R. Kotecký, and S. Miracle-Solé. *Finite-size scaling for Potts models*. J. Stat. Phys. **62**, 529 (1991).
- [50] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. *A new algorithm for Monte Carlo simulation of Ising spin systems*. J. Comput. Phys. **17**, 10 (1975).
- [51] R. Botet, M. Poszajczak, and J. A. González. *Phase transitions in nonextensive spin systems*. Phys. Rev. E **65**, 015103(R) (2002)..
- [52] A. J. Bray. *Long-range random-field models: scaling theory and $1/n$ expansion*. J. Phys. C **19**, 6225 (1986)..
- [53] A. J. Bray, M. A. Moore, and A. P. Young. *Lower critical dimension of metallic vector spin-glasses*. Phys. Rev. Lett. **56**, 2641 (1986).
- [54] E. Brézin and C. De Dominicis. *New phenomena in the random field Ising model*. Europhys. Lett. **44**, 13 (1998).
- [55] E. Brézin, J. Zinn-Justin, and J. C. Le Guillou. *Critical properties near σ dimensions for long-range interactions*. J. Phys. A **9**, L119 (1976)..
- [56] J. Bricmont and A. Kupiainen. *Lower critical dimension for the random-field Ising model*. Phys. Rev. Lett. **59**, 1829 (1987).
- [57] R. C. Brower and P. Tamayo. *Embedded dynamics for ϕ^4 theory*. Phys. Rev. Lett. **62**, 1087 (1989).
- [58] P. Bruno. *Absence of spontaneous magnetic order at nonzero temperature in one- and two-dimensional Heisenberg and XY systems with long-range interactions*. Phys. Rev. Lett. **87**, 137203 (2001).

Bibliography

- [59] E. Buddenoir and S. Wallon. *The correlation length of the Potts model at the first-order transition point*. J. Phys. A **26**, 3045 (1993).
- [60] T. W. Burkhardt and E. Eisenriegler. *Casimir interaction of spheres in a fluid at the critical point*. Phys. Rev. Lett. **74**, 3189 (1995).
- [61] A. Campa, A. Giansanti, and D. Moroni. *Canonical solution of a system of long-range interacting rotators on a lattice*. Phys. Rev. E **62**, 303 (2000).
- [62] A. Campa, A. Giansanti, D. Moroni, and C. Tsallis. *Classical spin systems with long-range interactions: universal reduction of mixing*. Phys. Lett. A **286**, 251 (2001).
- [63] S. A. Cannas. *One-dimensional Ising model with long-range interactions: a renormalization-group treatment*. Phys. Rev. B **52**, 3034 (1995).
- [64] S. A. Cannas and A. C. N. de Magalhães. *The one-dimensional Potts model with long-range interactions: a renormalization group approach*. J. Phys. A **30**, 3345 (1997).
- [65] S. A. Cannas, A. C. N. de Magalhães, and F. A. Tamarit. *Evidence of exactness of the mean field theory in the nonextensive regime of long-range spin models*. Phys. Rev. B **61**, 11521 (2000).
- [66] S. A. Cannas, P. M. Gleiser, and F. A. Tamarit. *Two dimensional Ising model with long-range competing interactions*. cond-mat/0502403.
- [67] S. A. Cannas, C. M. Lapilli, and D. A. Stariolo. *Testing boundary conditions efficiency in simulations of long-range interacting magnetic models*. Int. J. Mod. Phys. C **15**, 115 (2004).
- [68] S. A. Cannas and F. A. Tamarit. *Long-range interactions and nonextensivity in ferromagnetic spin models*. Phys. Rev. B **54**, R12661 (1996)..
- [69] M. S. Cao and J. Machta. *Migdal-Kadanoff study of the random-field Ising model*. Phys. Rev. B **48**, 3177 (1993).
- [70] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal. *Wolff-type embedding algorithms for general nonlinear σ -models*. Nucl.Phys. B **403**, 475 (1993).
- [71] J. Cardy. *Nonperturbative effects in a scalar supersymmetric theory*. Phys. Lett. B **125**, 470 (1983).
- [72] J. L. Cardy. *One-dimensional models with $1/r^2$ interactions*. J. Phys. A **14**, 1407 (1981).
- [73] M. S. Carroll, W. Janke, and S. Kappler. *Dynamical behavior of the multi-bondic and multicanonic algorithm in the 3d q-state Potts model*. J. Stat. Phys. **90**, 1277 (1998).

Bibliography

- [74] A. Cavallo, F. Cosenza, and L. De Cesare. *Classical Heisenberg ferromagnetic chain with long-range interactions: a spectral density approach*. Phys. Rev. B **66**, 174439 (2002).
- [75] A. Cavallo, F. Cosenza, and L. De Cesare. *Thermodynamic properties of a classical d -dimensional spin- s Heisenberg ferromagnet with long-range interactions via the spectral density method*. Physica A **332**, 301 (2004).
- [76] P. M. Chaikin and T. C. Lubensky. *Principles of condensed matter physics* (Cambridge University Press, Cambridge, England, 1995).
- [77] S. Chakravarty. *Quantum fluctuations in the tunneling between superconductors*. Phys. Rev. Lett. **49**, 681 (1982).
- [78] S. Chakravarty and P. W. Anderson. *Interlayer josephson tunneling and breakdown of fermi liquid theory*. Phys. Rev. Lett. **72**, 3859 (1994).
- [79] M. S. S. Challa, D. P. Landau, and K. Binder. *Finite-size effects at temperature-driven first-order transitions*. Phys. Rev. B **34**, 1841 (1986).
- [80] H. Chamati and D. M. Dantchev. *Renormalization group treatment of the scaling properties of finite systems with subleading long-range interaction*. Eur. Phys. J. B **26**, 89 (2002)..
- [81] H. S. Chan and K. A. Dill. *Energy landscapes and the collapse dynamics of homopolymers*. J. Chem. Phys. **99**, 2116 (1993).
- [82] G. Chikenji, M. Kikuchi, and Y. Iba. *Multi-self-overlap ensemble for protein folding: ground state search and thermodynamics*. Phys. Rev. Lett. **83**, 1886 (1999).
- [83] P. D. Coddington and C. F. Baillie. *Empirical relations between static and dynamic exponents for Ising model cluster algorithms*. Phys. Rev. Lett. **68**, 962 (1992).
- [84] P. D. Coddington and L. Han. *Generalized cluster algorithms for frustrated spin models*. Phys. Rev. B **50**, 3058 (1994).
- [85] ALPHA Collaboration and U. Wolff. *Monte Carlo errors with less errors*. Comp. Phys. Commun. **156**, 143 (2004).
- [86] M. Creutz. *Microcanonical Monte Carlo simulation*. Phys. Rev. Lett. **50**, 1411 (1983).
- [87] M. Creutz. *Microcanonical cluster Monte Carlo simulation*. Phys. Rev. Lett. **69**, 1002 (1992).
- [88] S. Curilef. *A long-range ferromagnetic spin model with periodic boundary conditions*. Phys. Lett. A **299**, 366 (2002).

Bibliography

- [89] S. Curilef and C. Tsallis. *Critical temperature and nonextensivity in long-range interacting Lennard-Jones-like fluids*. Phys. Lett. A **264**, 270 (1999).
- [90] D. Dantchev. *Two-point correlation function in systems with van der Waals type interaction*. Eur. Phys. J. B **23**, 211 (2001).
- [91] D. Dantchev and J. Rudnick. *Subleading long-range interactions and violations of finite size scaling*. Eur. Phys. J. B **21**, 251 (2001).
- [92] T. Dauxois, S. Ruffo, E. Arimondo, and M. Wilkens, editors. *Dynamics and thermodynamics of systems with long-range interactions* (Springer, 2002).
- [93] P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith. *Performance limitations of flat-histogram methods*. Phys. Rev. Lett. **92**, 097201 (2004).
- [94] A. S. de Arruda, W. Figueiredo, and R. M. Sebastianes. *Field distributions and tricritical points in a random-field Ising model: mean-field renormalization-group approach*. Phys. Rev. B **39**, 4409 (1989).
- [95] P. M. C. de Oliveira. *Broad histogram relation is exact*. Eur. Phys. J. B **6**, 111 (1998).
- [96] P. M. C. de Oliveira. *Broad histogram: an overview*. Braz. J. Phys. **30**, 195 (2000).
- [97] P. M. C. de Oliveira. *Broad histogram: tests for a simple and efficient microcanonical simulator*. Braz. J. Phys. **30**, 766 (2000).
- [98] P. M. C. de Oliveira, T. J. P. Penna, and H. J. Herrmann. *Broad histogram method*. Braz. J. Phys. **26**, 677 (1996).
- [99] K. De'Bell, A. B. MacIsaac, and J. P. Whitehead. *Dipolar effects in magnetic thin films and quasi-two-dimensional systems*. Rev. Mod. Phys. **72**, 225 (2000).
- [100] E. Domany, D. Mukamel, and A. Schwimmer. *Phase diagram of the Z(5) model on the square lattice*. J. Phys. A **13**, 1311 (1980).
- [101] E. Domany, M. Schick, and R. H. Swendsen. *First-order transition in an XY model with nearest-neighbour interactions*. Phys. Rev. Lett. **52**, 1535 (1984).
- [102] V. I. S. Dotsenko, W. Selke, and A. L. Talapov. *Cluster Monte Carlo algorithms for random Ising models*. Physica A **170**, 278 (1991).
- [103] C. Dress and W. Krauth. *Cluster algorithm for hard spheres and related systems*. J. Phys. A **28**, L597 (1995)..

Bibliography

- [104] A. Dutta. *Quantum spin glass with long-range random interactions*. Phys. Rev. B **65**, 224427 (2002).
- [105] A. Dutta. *Effect of long-range interactions on the pure and random quantum Ising transitions*. Physica A **318**, 63 (2003).
- [106] A. Dutta and J. K. Bhattacharjee. *Phase transitions in the quantum Ising and rotor models with a long-range interaction*. Phys. Rev. B **64**, 184106 (2001).
- [107] F. J. Dyson. *Existence of a phase-transition in a one-dimensional Ising ferromagnet*. Commun. Math. Phys. **12**, 91 (1969).
- [108] F. J. Dyson. *Non-existence of spontaneous magnetization in a one-dimensional Ising ferromagnet*. Commun. Math. Phys. **12**, 212 (1969).
- [109] F. J. Dyson. *An Ising ferromagnet with discontinuous long-range order*. Commun. Math. Phys. **21**, 269 (1971).
- [110] S. F. Edwards and P. W. Anderson. *Theory of spin glasses*. J. Phys. F **5**, 965 (1975)..
- [111] M. Fasnacht, R. H. Swendsen, and J. M. Rosenberg. *Adaptive integration method for Monte Carlo simulations*. Phys. Rev. E **69**, 056704 (2004).
- [112] A. M. Ferrenberg and D. P. Landau. *Critical behavior of the three-dimensional Ising model: a high-resolution Monte Carlo study*. Phys. Rev. B **44**, 5081 (1991).
- [113] A. M. Ferrenberg and R. H. Swendsen. *New Monte Carlo technique for studying phase transitions*. Phys. Rev. Lett. **61**, 2635 (1988).
- [114] A. M. Ferrenberg and R. H. Swendsen. *Optimized Monte Carlo data analysis*. Phys. Rev. Lett. **63**, 1195 (1989).
- [115] M. E. Fischer, S. Ma, and B. G. Nickel. *Critical exponents for long-range interactions*. Phys. Rev. Lett. **29**, 917 (1972).
- [116] D. S. Fisher and D. A. Huse. *Equilibrium behavior of the spin-glass ordered phase*. Phys. Rev. B **38**, 386 (1988).
- [117] M. E. Fisher and A. N. Berker. *Scaling for first-order phase transitions in thermodynamic and finite systems*. Phys. Rev. B **26**, 2507 (1982).
- [118] M. E. Fisher and V. Privman. *First-order transitions breaking $O(n)$ symmetry: finite-size scaling*. Phys. Rev. B **32**, 447 (1985).
- [119] M. E. Fisher and Z. Rácz. *Scaling theory of nonlinear critical relaxation*. Phys. Rev. B **13**, 5039 (1976).

Bibliography

- [120] H. Flyvbjerg. *Error estimates on averages of correlated data*. In *Advances in computer simulation, lectures held at the Eötvös summer school in Budapest, Hungary, 1996*, edited by J. Kertész & I. Kondor (Springer, Berlin, 1998).
- [121] P. J. Ford. *Spin glasses*. *Contemp. Phys.* **23**, 141 (1982).
- [122] C. M. Fortuin and P. W. Kasteleyn. *On the random-cluster model i: introduction and relation to other models*. *Physica* **57**, 536 (1972).
- [123] L. Franco and S. A. Cannas. *Non-glassy ground state in a long-range anti-ferromagnetic frustrated model in the hypercubic cell*. *Physica A* **332**, 337 (2004).
- [124] J. Fröhlich and T. Spencer. *The phase transition in the one-dimensional Ising model with $1/r^2$ interaction energy*. *Commun. Math. Phys.* **84**, 87 (1982).
- [125] G. Giacomin and J. L. Lebowitz. *Exact macroscopic description of phase segregation in model alloys with long range interactions*. *Phys. Rev. Lett.* **76**, 1094 (1996).
- [126] P. M. Gleiser, F. A. Tamarit, and S. A. Cannas. *Self-organized criticality in a model of biological evolution with long-range interactions*. *Physica A* **275**, 272 (2000).
- [127] P. M. Gleiser, F. A. Tamarit, and S. A. Cannas. *Metastable states in a two-dimensional Ising model with dipolar interactions*. *Physica D* **168**, 73 (2002)..
- [128] F. Gliozzi. *Simulation of Potts models with real q and no critical slowing down*. *Phys.Rev. E* **66**, 016115 (2002).
- [129] Z. Glumac and K. Uzelac. *Finite-range scaling study of the 1d long-range Ising model*. *J. Phys. A* **22**, 4439 (1989).
- [130] Z. Glumac and K. Uzelac. *Critical behaviour of the 1d q -state Potts model with long-range interactions*. *J. Phys. A* **26**, 5267 (1993).
- [131] Z. Glumac and K. Uzelac. *First-order transition in the 1d three-state Potts model with LR interactions*. *Phys. Rev. E* **58**, 4372 (1998).
- [132] Z. Glumac and K. Uzelac. *Determination of the order of phase transitions in Potts model by the graph-weight approach*. *Physica A* **271**, 147 (1999).
- [133] J. Goodman and A. D. Sokal. *Multigrid Monte Carlo method for lattice field theories*. *Phys. Rev. Lett.* **56**, 1015 (1986).
- [134] V. K. Gore and M. R. Jerrum. *The Swendsen-Wang process does not always mix rapidly*. *J. Stat. Phys.* **97**, 67 (1999).

Bibliography

- [135] A. Grabowski, R. A. Kosinski, and A. Krawiecki. *Dynamics of a one-dimensional neural network with a "small world" topology of synaptic connections*. Physica A **341**, 702 (2004).
- [136] J. Raúl Grigera. *Extensive and non-extensive thermodynamics. a molecular dynamic test*. Phys. Lett. A **217**, 47 (1996).
- [137] M. Grousson, G. Tarjus, and P. Viot. *Phase diagram of an Ising model with long-range frustrating interactions: a theoretical analysis*. Phys. Rev. E **62**, 7781 (2000).
- [138] M. Grousson, G. Tarjus, and P. Viot. *Monte Carlo study of the three-dimensional Coulomb frustrated Ising ferromagnet*. Phys. Rev. E **64**, 036109 (2001).
- [139] M. L. Guerra and J. D. Mu noz. *Equilibrium times for the multicanonical method*. Int. J. Mod. Phys. C **15**, 471 (2004)..
- [140] L. E. Guerrero and J. A. Gonzáles. *Long-range interacting solitons: pattern formation and nonextensive thermostatics*. patt-sol/9905010.
- [141] J. D. Gunton, M. San Miguel, and P. S. Sahni. In *Phase transitions and critical phenomena*, Vol. 8, edited by C. Domb and J. L. Lebowitz (Academic Press, New York, 1989).
- [142] M. A. Gusmao and W. K. Theumann. *Validity of the long-range expansion in the n-vector model*. Phys. Rev. B **28**, 6545 (1983).
- [143] J. P. Hansen and P. Viot. *Pair correlations and internal energy of the two-dimensional Coulomb gas*. Phys. Lett. A **95**, 155 (1983).
- [144] J. P. Hansen and P. Viot. *Two-body correlations and pair formation in the two-dimensional Coulomb gas*. J. Stat. Phys. **38**, 823 (1985).
- [145] U. H. E. Hansmann. *On the protein folding problem*. Nucl. Phys. B (Proc. Suppl.) **47**, 188 (1996)..
- [146] U. H. E. Hansmann. *Parallel tempering algorithm for conformational studies of biological molecules*. Chem. Phys. Lett. **281**, 140 (1997).
- [147] U. H. E. Hansmann. *Generalized ensembles: a new way of simulating proteins*. Physica A **254**, 15 (1998).
- [148] U. H. E. Hansmann. *Generalized ensemble technique and protein folding simulations*. Comp. Phys. Commun. **147**, 604 (2002).
- [149] U. H. E. Hansmann. *New algorithms and the physics of proteins*. Physica A **321**, 152 (2003).

Bibliography

- [150] U. H. E. Hansmann, F. Eisenmenger, and Y. Okamoto. *Stochastic dynamics simulation in a new generalized ensemble*. Chem. Phys. Lett. **297**, 374 (1998).
- [151] U. H. E. Hansmann and Y. Okamoto. *Generalized-ensemble Monte Carlo method for systems with rough energy landscape*. Phys. Rev. E **56**, 2228 (1997).
- [152] U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger. *Molecular dynamics, langevin and hybrid Monte Carlo simulations in a multicanonical ensemble*. Chem. Phys. Lett. **259**, 321 (1996).
- [153] D. W. Heermann and A. N. Burkitt. *System size dependence of the auto-correlation time for the Swendsen-Wang Ising model*. Physica A **162**, 210 (1990).
- [154] L. Hernández and H. T. Diep. *Existence of a tricritical point at finite field in the three-dimensional random-field Ising model*. Phys. Rev. B **55**, 14080 (1997).
- [155] J. R. Heringa and H. W. J. Blöte. *Geometric cluster Monte Carlo simulations*. Phys. Rev. E **57**, 4976 (1998).
- [156] B. Hesselbo and R. B. Stinchcombe. *Monte Carlo simulation and global optimization without parameters*. Phys. Rev. Lett. **74**, 2151 (1995).
- [157] H. Hinrichsen and M. Howard. *A model for anomalous directed percolation*. Eur. Phys. J. B **7**, 635 (1999).
- [158] P. Hohenberg and B. Halperin. *Theory of dynamic critical phenomena*. Rev. Mod. Phys. **49**, 435 (1977).
- [159] J. Honkonen. *Critical behaviour of the long-range $(\phi^2)^2$ model in the short-range limit*. J. Phys. A **23**, 825 (1990).
- [160] J. Honkonen and M. Y. Nalimov. *Crossover between field theories with short-range and long-range exchange or correlations*. J. Phys. A **22**, 751 (1989).
- [161] J. Hoshen and R. Kopelman. *Percolation and cluster distribution. i. cluster multiple labeling technique and critical concentration algorithm*. Phys. Rev. B **14**, 3438 (1974).
- [162] J. Houdayer. *A cluster Monte Carlo algorithm for 2-dimensional spin glasses*. Eur. Phys. J. B **22**, 479 (2001).
- [163] J. Houdayer and O. C. Martin. *A geometrical picture for finite-dimensional spin glasses*. Europhys. Lett. **49**, 794 (2000).

Bibliography

- [164] A. Houghton, A. Khurana, and F. J. Seco. *Fluctuation-driven first-order phase transition, below four dimensions, in the random-field Ising model with a gaussian random-field distribution*. Phys. Rev. Lett. **55**, 856 (1985).
- [165] X.-P. Huang and C. F. Driscoll. *Relaxation of 2d turbulence to a metaequilibrium near the minimum enstrophy state*. Phys. Rev. Lett. **72**, 2187 (1994).
- [166] Y. Iba. *Extended ensemble Monte Carlo*. Int. J. Mod. Phys. C **12**, 623 (2001).
- [167] Y. Iba, G. Chikenji, and M. Kikuchi. *Simulation of lattice polymers with multi-self-overlap ensemble*. J. Phys. Soc. Japan **67**, 3327 (1998).
- [168] Y. Iba and H. Takahashi. *Exploration of multi-dimensional density of states by multicanonical Monte Carlo algorithm*. cond-mat/0407396.
- [169] F. Iglói and J. Sólyom. *First-order transition for the (1+1)-dimensional $q=4$ Potts model from finite lattice extrapolation*. J. Phys. C **16**, 2833 (1983)..
- [170] J. Z. Imbrie. *Lower critical dimension of the random-field Ising model*. Phys. Rev. Lett. **53**, 1747 (1984).
- [171] J. Z. Imbrie and C. M. Newman. *An intermediate phase with slow decay of correlations in one dimensional $1/|x - y|^2$ percolation, Ising and Potts models*. Commun. Math. Phys. **118**, 303 (1988).
- [172] Y. Imry. *Finite-size rounding of a first-order phase transition*. Phys. Rev. B **21**, 2042 (1980).
- [173] Y. Imry and S. Ma. *Random-field instability of the ordered state of continuous symmetry*. Phys. Rev. Lett. **35**, 1399 (1975).
- [174] Y. M. Ivanchenko and A. A. Lisiansky, *Physics of critical fluctuations* (Springer, Berlin, 1999).
- [175] D. Ivaneyko, J. Ilnytskyi, B. Berche, and Yu. Holovatch. *Criticality of the random-site Ising model: Metropolis, Swendsen-Wang and Wolff Monte Carlo algorithms*. cond-mat/0501291.
- [176] F. James, *RANLUX: A Fortran implementation of the high-quality pseudorandom number generator of Lüscher*. Comput. Phys. Commun. **79**, 111 (1994).
- [177] W. Janke. *Accurate first-order transition points from finite-size data without power-law corrections*. Phys. Rev. B **47**, 14757 (1993).
- [178] W. Janke. *Multicanonical Monte Carlo simulations*. Physica A **254**, 164 (1998).

Bibliography

- [179] W. Janke. *Statistical analysis of simulations: data correlations and error estimation*. In *Quantum Simulations of Complex Many-Body Systems: from Theory to Algorithms*, NIC Series - Vol. 10, Proceedings of the Euro Winter School, 2002, edited by J. Grotendorst, D. Marx, and A. Muramatsu (John Von Neumann Institute for Computing, Jülich, 2002), pp. 423–445
- [180] W. Janke. *Histograms and all that*. In *Computer Simulations of Surfaces and Interfaces*, NATO Science Series, II. Mathematics, Physics and Chemistry - Vol. 114, Proceedings of the NATO Advanced Study Institute, Albena, Bulgaria, 9 - 20 September 2002, edited by B. Dünweg, D. P. Landau, and A. I. Milchev (Kluwer, Dordrecht, 2003), pp. 137–157.
- [181] W. Janke and S. Kappler. *Correlation function at β_t in the disordered phase of 2d Potts models*. Phys. Lett. A **197**, 227 (1995).
- [182] W. Janke and S. Kappler. *Multibondic cluster algorithm*. Nucl. Phys. B (Proc. Suppl.) **42**, 876 (1995)..
- [183] W. Janke and S. Kappler. *Multibondic cluster algorithm for Monte Carlo simulations of first-order phase transitions*. Phys. Rev. Lett. **74**, 212 (1995).
- [184] W. Janke and S. Kappler. *Ordered vs disordered: correlation lengths of 2d Potts models at β_t* . Nucl. Phys. B (Proc. Suppl.) **42**, 770 (1995)..
- [185] W. Janke and R. Kenna. *The strength of first and second order phase transitions from partition function zeroes*. J. Stat. Phys. **102**, 1211 (2001).
- [186] W. Janke and T. Sauer. *Multicanonical multigrid Monte Carlo method*. Phys. Rev. E **49**, 3475 (1994).
- [187] W. Janke and T. Sauer. *Multigrid method versus staging algorithm for pimc simulations*. Chem. Phys. Lett. **263**, 488 (1996).
- [188] W. Janke and A. M. J. Schakel. *Fractal structure of high-temperature graphs of $O(n)$ models in two dimensions*. cond-mat/0502062.
- [189] W. Janke and A. M. J. Schakel. *Fractal structure of spin clusters and domain walls in the two-dimensional Ising model*. Phys. Rev. E **71**, 036703 (2005).
- [190] W. Janke and A. M. J. Schakel. *Geometrical phase transitions*. Comp. Phys. Commun., In press, Available online 7 April 2005.
- [191] W. Janke and R. Villanova. *Three-dimensional 3-state Potts model revisited with new techniques*. Nucl. Phys. B **489**, 679 (1997).
- [192] H. K. Janssen. *Influence of long-range interactions on the critical behaviour of systems with negative fisher-exponent*. Phys. Rev. E **58**, 2673 (1998).

Bibliography

- [193] D. Jeong, H. Hong, B. J. Kim, and M. Y. Choi. *Phase transition in the Ising model on a small-world network with distance-dependent interactions*. Phys. Rev. E **68**, 027101 (2003).
- [194] T. Jörg, *Cluster Monte Carlo algorithms for diluted spin glasses*. In *Progress of Theoretical Physics Supplement*, Proceedings of “Statistical Physics of Disordered Systems and Its Applications”, Hayama (Japan), July 2004.
- [195] G. S. Joyce. *Spherical model with long-range ferromagnetic interactions*. Phys. Rev. **146**, 349 (1966).
- [196] P. Jund, S. G. Kim, and C. Tsallis. *Crossover from extensive to nonextensive behavior driven by long-range interactions*. Phys. Rev. B **52**, 50 (1995).
- [197] D. Kandel, R. Ben-Av, and E. Domany. *Cluster dynamics for fully frustrated systems*. Phys. Rev. Lett. **65**, 941 (1990).
- [198] D. Kandel, R. Ben-Av, and E. Domany. *Cluster Monte Carlo dynamics for the fully frustrated Ising model*. Phys. Rev. B **45**, 4700 (1992).
- [199] D. Kandel, E. Domany, and A. Brandt. *Simulations without critical slowing down: Ising and three-state Potts models*. Phys. Rev. B **40**, 330 (1989).
- [200] H. G. Katzgraber and A. P. Young. *Monte Carlo studies of the one-dimensional Ising spin glass with power-law interactions*. Phys. Rev. B **67**, 134410 (2003).
- [201] M. Kaufman, P. E. Klunzinger, and A. Khurana. *Multicritical points in an Ising random-field model*. Phys. Rev. B **34**, 4766 (1986).
- [202] A. Khurana, F. J. Seco, and A. Houghton. *Ising model in a quenched random field: critical exponents in three dimensions from high-temperature series*. Phys. Rev. Lett. **54**, 357 (1985).
- [203] M. K.-H. Kiessling. *On the equilibrium statistical mechanics of isothermal classical self-gravitating matter*. J. Stat. Phys. **55**, 203 (1989).
- [204] T. Kihara, Y. Midzuno, and T. Shizume. J. Phys. Soc. Jpn. **9**, 681 (1954).
- [205] S. Kirkpatrick and E. P. Stoll, *A very fast shift-register sequence random number generator*. J. Comput. Phys. **40**, 517 (1981).
- [206] A. Klein and S. Masoolean. *Taming griffiths’ singularities in long range random Ising models*. Commun. Math. Phys. **189**, 497 (1997)..
- [207] J. M. Kosterlitz. *The critical properties of the two-dimensional XY model*. J. Phys. C **7**, 1046 (1974)..
- [208] J. M. Kosterlitz. *Phase transitions in long-range ferromagnetic chains*. Phys. Rev. Lett. **37**, 1577 (1976).

Bibliography

- [209] G. Kotliar, P. W. Anderson, and D. L. Stein. *One-dimensional spin-glass model with long-range random interactions*. Phys. Rev. B **27**, 602 (1983).
- [210] W. Krauth, *Introduction to Monte Carlo algorithms*. In *Advances in computer simulation, lectures held at the Eötvös summer school in Budapest, Hungary, 1996*, edited by J. Kertész and I. Kondor (Springer, Berlin, 1998).
- [211] W. Krauth, *Cluster Monte Carlo algorithms*. In *New Optimization Algorithms in Physics*, edited by A. K. Hartmann and H. Rieger (Wiley-VCh, 2004).
- [212] W. Krauth and O. Pluchery. *A rapid dynamical Monte Carlo algorithm for glassy systems*. J. Phys. A **27**, L715 (1994)..
- [213] M. Krech and E. Luijten. *Optimized energy calculation in lattice systems with long-range interactions*. Phys. Rev. E **61**, 2058 (2000).
- [214] U. Krey. *On the lower critical dimension of spin systems in random fields*. J. Phys. C **18**, 1455 (1985).
- [215] F. Krzakala and O. C. Martin. *Spin and link overlaps in three-dimensional spin glasses*. Phys. Rev. Lett. **85**, 3013 (2000).
- [216] D. P. Landau and K. Binder. *A guide to Monte Carlo simulations in statistical physics* (Cambridge University Press, Cambridge, England, 2000).
- [217] V. Latora, A. Rapisarda, and C. Tsallis. *Non-gaussian equilibrium in a long-range hamiltonian system*. Phys. Rev. E **64**, 056134 (2001).
- [218] J. Lee. *New Monte Carlo algorithm: entropic sampling*. Phys. Rev. Lett. **71**, 211 (1993).
- [219] J. Lee and J. M. Kosterlitz. *New numerical method to study phase transitions*. Phys. Rev. Lett. **65**, 137 (1990).
- [220] J. Lee and J. M. Kosterlitz. *Finite-size scaling and Monte Carlo simulations of first-order phase transitions*. Phys. Rev. B **43**, 3265 (1991).
- [221] L. Leuzzi. *Critical behaviour and ultrametricity of Ising spin-glass with long-range interactions*. J. Phys. A **32**, 1417 (1999)..
- [222] S. Liang. *Application of cluster algorithms to spin glasses*. Phys. Rev. Lett. **69**, 2145 (1992).
- [223] A. R. Lima. *A comparison between broad histogram and multicanonical methods*. J. Stat. Phys. **99**, 691 (2000).
- [224] A. R. Lima, P. M. C. de Oliveira, and T. J. P. Penna. *Broad histogram method for multiparametric hamiltonians*. cond-mat/9912152.

Bibliography

- [225] E. Luijten. *Test of renormalization predictions for universal finite-size scaling functions*. Phys. Rev. E **60**, 7558 (1999).
- [226] E. Luijten, *Monte Carlo simulation of spin models with long-range interactions*. In *Computer simulation studies in condensed-matter physics*, Vol. 12 (Springer, Heidelberg, 2000), pp. 86–99.
- [227] E. Luijten and H. W. J. Blöte. *Monte Carlo method for spin models with long-range interactions*. Int. J. Mod. Phys. C **6**, 359 (1995).
- [228] E. Luijten and H. W. J. Blöte. *Finite-size scaling and universality above the upper critical dimensionality*. Phys. Rev. Lett. **76**, 1557 (1996).
- [229] E. Luijten and H. W. J. Blöte. *Finite-size scaling and universality above the upper critical dimensionality [erratum to phys. rev. lett. 76, 1557 (1996)]*. Phys. Rev. Lett. **76**, 3662 (1996).
- [230] E. Luijten and H. W. J. Blöte. *Medium-range interactions and crossover to classical critical behavior*. Phys. Rev. E **54**, 4626 (1996).
- [231] E. Luijten and H. W. J. Blöte. *Classical critical behaviour of spin models with LR interactions*. Phys. Rev. B **56**, 8945 (1997).
- [232] E. Luijten and H. W. J. Blöte. *Boundary between long-range and short-range critical behavior in systems with algebraic interactions*. Phys. Rev. Lett. **89**, 025703 (2002).
- [233] E. Luijten, H. W. J. Blöte, and K. Binder. *Non monotonic crossover of the effective susceptibility exponent*. Phys. Rev. Lett. **79**, 561 (1997).
- [234] E. Luijten and J. Liu. *Cluster algorithms: beyond suppression of critical slowing down*. In *The Monte Carlo method in the physical sciences: celebrating the 50th anniversary of the Metropolis algorithm.*, Vol. 690, Los Alamos National Laboratory, June 9-11 2003, edited by J. E. Gubernatis. (AIP Conference Proceedings, American Institute of Physics, Melville, 2003) pp. 225–231.
- [235] E. Luijten and H. Messingfeld. *Criticality in one dimension with inverse square-law potentials*. Phys. Rev. Lett. **86**, 5305 (2001).
- [236] M. Lüscher. *A portable high-quality random number generator for lattice field theory simulations*. Comput. Phys. Comm. **79**, 100 (1994).
- [237] A. P. Lyubartsev, A. A. Martinovsky, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. *New approach to Monte Carlo calculation of the free energy: method of expanded ensembles*. J. Chem. Phys. **96**, 1776 (1992).

Bibliography

- [238] J. Machta, M. Newman, and L. Chayes. *Replica-exchange algorithm and results for the three-dimensional random field Ising model*. Phys. Rev. E **62**, 8782 (2000).
- [239] A. B. MacIsaac, J. P. Whitehead, M. C. Robinson, and K. De'Bell. *Striped phases in two-dimensional dipolar ferromagnets*. Phys. Rev. B **51**, 16033 (1995).
- [240] P. G. Maier and F. Schwabl. *Novel phase transition in two-dimensional XY-models with long-range interaction*. Cond. Mat. Phys. **8**, 103 (2005).
- [241] E. Marinari and G. Parisi. *Simulated tempering: a new Monte Carlo scheme*. Europhys. Lett. **19**, 451 (1992).
- [242] G. Marsaglia. *A current view of random numbers*. In *Computer Science and Statistics, Proceedings of the Sixteenth Symposium on The Interface* (North-Holland, Amsterdam, 1985), pp. 3-10.
- [243] M. Matsumoto and T. Nishimura, *Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator*, ACM Transactions on Modeling and Computer Simulation **8**, 3 (1998).
- [244] D. C. Mattis. *Tricritical point in random-field Ising model*. Phys. Rev. Lett. **55**, 3009 (1985).
- [245] N. D. Mermin and H. Wagner. *Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models*. Phys. Rev. Lett. **17**, 1133 (1966).
- [246] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, and E. Teller. *Equation of state calculations by fast computing machines*. J. Chem. Phys. **21**, 1087 (1953).
- [247] L. Mittag and M. J. Stephen. *Mean-field theory of the many component Potts model*. J. Phys. A **7**, L109 (1974)..
- [248] J. L. Monroe. *Bethe lattice approximation of long-range interaction Ising models*. Phys. Lett. A **171**, 427 (1992).
- [249] J. L. Monroe. *A new criterion for the location of phase transitions for spin systems on recursive lattices*. Phys. Lett. A **188**, 80 (1994).
- [250] J. L. Monroe. *One-dimensional Ising models with long-range interactions*. J. Phys. A **31**, 9809 (1998)..
- [251] J. L. Monroe. *Long-range one-dimensional Potts models: a cluster mean-field and extrapolation approach*. J. Phys. A **32**, 7803 (1999).

Bibliography

- [252] H. Müller-Krumbhaar and K. Binder. *Dynamical properties of the Monte Carlo method in statistical mechanics*. J. Stat. Phys. **8**, 1 (1973).
- [253] J. D. Munoz and H. J. Herrmann. *Broad histogram method for continuous systems: the XY-model*. cond-mat/9810024.
- [254] J. F. Nagle and J. C. Bonner. *Numerical studies of the Ising chain with long-range ferromagnetic interactions*. J. Phys. C **3**, 352 (1970)..
- [255] T. Nattermann. *Theory of the random field Ising model*. In *Spin glasses and random fields* (World Scientific, Singapore, 1997), pp. 31.
- [256] T. Nattermann and P. Rujan. *Random field and other systems dominated by disorder fluctuations*. Int. J. Mod. Phys. B **11**, 1597 (1989).
- [257] D. R. Nelson. *Study of melting in two dimensions*. Phys. Rev. B **18**, 2318 (1978).
- [258] M. E. J. Newman and G. T. Barkema. *Monte Carlo study of the random-field Ising model*. Phys. Rev. E **53**, 393 (1996).
- [259] M. E. J. Newman and G. T. Barkema. *Monte Carlo methods in statistical physics* (Oxford University Press, Oxford, England, 2001).
- [260] M. E. J. Newman, B. W. Roberts, G. T. Barkema, and J. P. Sethna. *Real-space renormalization group for the random-field Ising model*. Phys. Rev. B **48**, 16533 (1993).
- [261] F. Niedermayer. *General cluster updating method for Monte Carlo simulations*. Phys. Rev. Lett. **61**, 2026 (1988).
- [262] F. Niedermayer, *Cluster algorithms*. In *Advances in computer simulation, lectures held at the Eötvös summer school in Budapest, Hungary, 1996*, edited by J. Kertész & I. Kondor (Springer, Berlin, 1998).
- [263] B. Nienhuis, A. N. Berker, E. K. Riedel, and M. Schick. *First- and second-order phase transitions in Potts models: renormalization-group solution*. Phys. Rev. Lett. **43**, 737 (1979).
- [264] M. P. Nightingale and H. W. J. Blöte. *Dynamic exponent of the two-dimensional Ising model and Monte Carlo computation of the subdominant eigenvalue of the stochastic matrix*. Phys. Rev. Lett. **76**, 4548 (1996).
- [265] H. Nishimori. *Statistical physics of spin glasses and information processing* (Oxford University Press, Oxford, England, 2001).
- [266] Andrew T. Ogielski. *Integer optimization and zero-temperature fixed point in Ising random-field systems*. Phys. Rev. Lett. **57**, 1251 (1986).

Bibliography

- [267] Y. Okabe, Y. Tomita, and C. Yamaguchi. *Application of new Monte Carlo algorithms to random spin systems*. Comp. Phys. Commun. **146**, 63 (2002).
- [268] D. O’Kane and A. Treves. *Short- and long-range connections in autoassociative memory*. J. Phys. A **25**, 5055 (1992).
- [269] J. Oppenheim. *Thermodynamics with long-range interactions: from Ising models to black holes*. Phys. Rev. E **68**, 016108 (2003).
- [270] G. Ossola and A. D. Sokal. *Systematic errors due to linear congruential random-number generators with the Swendsen-Wang algorithm: A warning*. Phys. Rev. E **70**, 027701 (2004).
- [271] M. Palassini and A. P. Young. *Nature of the spin glass state*. Phys. Rev. Lett. **85**, 3017 (2000).
- [272] G. Parisi. *Magnetic properties of spin glasses in a new mean field theory*. J. Phys. A **13**, 1887 (1980).
- [273] G. Parisi and N. Sourlas. *Random magnetic fields, supersymmetry, and negative dimensions*. Phys. Rev. Lett. **43**, 744 (1979).
- [274] G. Parisi and N. Sourlas. *Supersymmetric field theories and stochastic differential equations*. Nucl. Phys. B **206**, 321 (1982).
- [275] Y. S. Parmar and J. K. Bhattacharjee. *Renormalization group for the random-field Ising model*. Phys. Rev. B **46**, 1216 (1992).
- [276] P. A. Pearce and R. B. Griffiths. *Potts model in the many-component limit*. J. Phys. A **13**, 2143 (1980)..
- [277] P. Peczak and D. P. Landau. *Monte Carlo study of finite-size effects at a weakly first-order phase transition*. Phys. Rev. B **39**, 11932 (1989).
- [278] K. S. Pitzer, M. C. P. de Lima, and D. R. Schreiber. *Critical point and phase separation for an ionic system*. J. Phys. Chem. **89**, 1854 (1985).
- [279] R. G. Priest and T. C. Lubensky. *Critical properties of two tensor models with application to the percolation problem*. Phys. Rev. B **13**, 4159 (1976).
- [280] V. Privman and N. M. Švrakić. *Wetting phenomena with long-range forces: exact results for the solid-on-solid model with the $1/r$ substrate potential*. Phys. Rev. B **37**, 5974 (1988).
- [281] E. B. Rasmussen, M. A. Novotny, and D. P. Landau. *MCRG study of d -dimensional random field Ising models*. J. Appl. Phys. **53**, 1925 (1982).
- [282] T. S. Ray, P. Tamayo, and W. Klein. *Mean-field study of the Swendsen-Wang dynamics*. Phys. Rev. A **39**, 5949 (1989).

Bibliography

- [283] O. Redner, J. Machta, and L. Chayes. *Graphical representations and cluster algorithms for critical points with fields*. Phys. Rev. E **58**, 2749 (1998).
- [284] S. Reynal and H. T. Diep. *Reexamination of the long-range Potts model: a multicanonical approach*. Phys. Rev. E **69**, 026109 (2004).
- [285] S. Reynal and H. T. Diep. *Hybrid multicanonical cluster algorithm for efficient simulations of long-range spin models*. Comp. Phys. Commun., In press, Available online 8 April 2005.
- [286] H. Rieger. *Critical behavior of the three-dimensional random-field Ising model: two-exponent scaling and discontinuous transition?* Phys. Rev. B **52**, 6659 (1995).
- [287] H. Rieger and A. P. Young. *Critical exponents of the three dimensional random field Ising model*. J. Phys. A **26**, 5279 (1993).
- [288] S. Romano. *One-dimensional-lattice spin models with long-range antiferromagnetic interactions*. Phys. Rev. B **42**, 6739 (1990).
- [289] H. H. Roomany and H. W. Wyld. *Finite-lattice approach to the $O(2)$ and $O(3)$ models in $1 + 1$ dimensions and the $(2+1)$ -dimensional Ising model*. Phys. Rev. D **21**, 3341 (1980).
- [290] D. Ruelle. *Statistical mechanics of a one-dimensional lattice gas*. Commun. Math. Phys. **9**, 267 (1968).
- [291] K. Rummukainen. *Multicanonical cluster algorithm*. Nucl. Phys. B (Proc. Suppl.) **390**, 621 (1993)..
- [292] J. Sak. *Recursion relations and fixed points for ferromagnets with long-range interactions*. Phys. Rev. B **8**, 281 (1973).
- [293] R. Salazar and R. Toral. *Scaling laws for a system with long-range interactions within Tsallis statistics*. Phys. Rev. Lett. **83**, 4233 (1999).
- [294] L. C. Sampaio, M. P. de Albuquerque, and F. S. de Menezes. *Nonextensivity and Tsallis statistics in magnetic systems*. Phys. Rev. B **55**, 5611 (1997).
- [295] T. Schneider and E. Pytte. *Random-field instability of the ferromagnetic state*. Phys. Rev. B **15**, 1519 (1977).
- [296] L. S. Schulman. *Long range percolation in one dimension*. J. Phys. A **16**, L639 (1983)..
- [297] B. J. Schulz, K. Binder, and M. Mueller. *Flat histogram method of Wang-Landau and n-fold way*. Int. J. Mod. Phys. C **13**, 477 (2002).
- [298] R. M. Sebastianes and V. K. Saxena. *Phase diagram of the random-field Ising model with a trimodal distribution*. Phys. Rev. B **35**, 2058 (1987).

Bibliography

- [299] M. S. Shell, P. G. Debenedetti, and A. Z. Panagiotopoulos. *An improved Monte Carlo method for direct calculation of the density of states*. J. Chem. Phys. **119**, 9406 (2003)..
- [300] B. Simon and A. Sokal. *Rigorous entropy-energy arguments*. J. Stat. Phys. **25**, 679 (1981).
- [301] G. R. Smith and A. D. Bruce. *A study of the multi-canonical Monte Carlo method*. J. Phys. A **28**, 6623 (1995).
- [302] B. Stosik and I. P. Fittipaldi. *Pattern recognition via Ising model with long range interactions*. Physica A **242**, 323 (1997).
- [303] Y. Sugita, A. Kitao, and Y. Okamoto. *Multidimensional replica-exchange method for free-energy calculations*. J. Chem. Phys. **113**, 6042 (2000).
- [304] T. Surungan, Y. Okabe, and Y. Tomita. *Study of the fully frustrated clock model using the Wang-Landau algorithm*. J. Phys. A **37**, 4219 (2004).
- [305] M. Suzuki, X. Hu, M. Katori, A. Lipowsky, N. Hatano, K. Minami, and Y. Nonomura. *Coherent Anomaly Method: mean field, fluctuations and systematics* (World Scientific, Singapore, 1995).
- [306] M. Suzuki, Y. Yamazaki, and G. Igarashi. *Wilson-type expansions of critical exponents for long-range interactions*. Phys. Lett. A **42**, 313 (1972).
- [307] R. H. Swendsen. Private communication, 2004.
- [308] R. H. Swendsen and J.-S. Wang. *Replica Monte Carlo simulation of spin glasses*. Phys. Rev. Lett. **57**, 2607 (1986).
- [309] R. H. Swendsen and J.-S. Wang. *Nonuniversal critical dynamics in Monte Carlo simulations*. Phys. Rev. Lett. **58**, 86 (1987).
- [310] R. H. Swendsen, J.-S. Wang, S.-T. Li, B. Diggs, C. Genovese, and J. B. Kadane. *Transition matrix Monte Carlo*. Int. J. Mod. Phys. C **10**, 1563 (1999).
- [311] M. Takahashi. *Generalization of mean-field approximations by the Feynman inequality and application to long-range Ising chain*. J. Phys. Soc. Japan **50**, 1854 (1981).
- [312] F. Tamarit and C. Anteneodo. *Rotators with long-range interactions: connection with the mean-field approximation*. Phys. Rev. Lett. **84**, 208 (1999).
- [313] A. F. Teich and N. Qian. *Learning and adaptation in a recurrent model of VI orientation selectivity*. J. Neurophysiol. **13**, 452 (2002)..

Bibliography

- [314] N. Theodorakopoulos, T. Dauxois, and M. Peyrard. *Order of the phase transition in models of DNA thermal denaturation*. Phys. Rev. Lett. **85**, 6 (2000).
- [315] W. K. Theumann and M. A. Gusmao. *Critical exponents for ϕ^3 -field models with long-range interactions*. Phys. Rev. B **31**, 379 (1985).
- [316] D. J. Thouless. *Long-range order in one-dimensional Ising systems*. Phys. Rev. **187**, 732 (1969).
- [317] Y. Tomita and Y. Okabe. *Probability-changing cluster algorithm for Potts models*. Phys. Rev. Lett. **86**, 572 (2001).
- [318] J. M. Torrie and J. P. Valleau. *Monte Carlo free energy estimates using non-boltzmann sampling: application to the sub-critical Lennard-Jones fluid*. Chem. Phys. Lett. **28**, 578 (1974).
- [319] S. Trebst, D. A. Huse, and M. Troyer. *Optimizing the ensemble for equilibration in broad-histogram Monte Carlo simulations*. Phys. Rev. E **70**, 046701 (2004).
- [320] M. Troyer, S. Wessel, and F. Alet. *Flat histogram methods for quantum systems: algorithms to overcome tunneling problems and calculate the free energy*. Phys. Rev. Lett. **90**, 120201 (2003).
- [321] C. Tsallis. *Non-extensive thermostatistics: brief review and comments*. Physica A **221**, 277 (1995).
- [322] C. Tsallis. *Nonextensive thermostatistics and fractals*. Fractals **3**, 541 (1995).
- [323] K. Uzelac and Z. Glumac. *Finite-range scaling for a one-dimensional system with long-range interactions*. J. Phys. A **21**, L421 (1988)..
- [324] K. Uzelac and Z. Glumac. *First-order phase transition in 1d Potts model with long-range interactions*. Fiz. B **6**, 133 (1997).
- [325] K. Uzelac and Z. Glumac. *The critical behaviour of the long-range Potts chain from the largest cluster probability distribution*. Physica A **314**, 448 (2002).
- [326] K. Uzelac, Z. Glumac, and A. Anicic. *Critical behavior of the long-range Ising chain from the largest-cluster probability distribution*. Phys. Rev. E **63**, 037101 (2001).
- [327] A. C. D. van Enter. *Instability of phase diagrams for a class of “irrelevant” perturbations*. Phys. Rev. B **26**, 1336 (1982).

Bibliography

- [328] A. C. D. van Enter and S. B. Shlosman. *First-order transitions for n -vector models in two and more dimensions: rigorous proof*. Phys. Rev. Lett. **89**, 285702 (2002).
- [329] P. Viot and G. Tarjus. *Microphase separation and modulated phases in a Coulomb frustrated Ising ferromagnet*. Europhys. Lett. **44**, 423 (1998).
- [330] T. Vojta. *Spherical random-field systems with long-range interactions: general results and application to the Coulomb glass*. J. Phys. A **26**, 2883 (1993).
- [331] B. P. Vollmayr-Lee and E. Luijten. *Comment on “scaling laws for a system with long-range interactions within Tsallis statistics”*. Phys. Rev. Lett. **85**, 470 (2000).
- [332] B. P. Vollmayr-Lee and E. Luijten. *A Kac-potential treatment of nonintegrable interactions*. Phys. Rev. E **63**, 031108 (2001).
- [333] D. J. Wallace and A. P. Young. *Spin anisotropy and crossover in the Potts model*. Phys. Rev. B **17**, 2384 (1978).
- [334] F. Wang and D. P. Landau. *Determining the density of states for classical statistical models: a random walk algorithm to produce a flat histogram*. Phys. Rev. E **64**, 056101 (2001).
- [335] F. Wang and D. P. Landau. *Efficient, multiple-range random walk algorithm to calculate the density of states*. Phys. Rev. Lett. **86**, 2050 (2001).
- [336] J.-S. Wang. *Is the broad histogram random walk dynamics correct?* Eur. Phys. J. B **8**, 287 (1999).
- [337] J.-S. Wang. *Flat histogram Monte Carlo method*. Physica A **281**, 147 (2000).
- [338] J.-S. Wang and L. W. Lee. *Monte Carlo algorithms based on the number of potential moves*. Comp. Phys. Commun. **127**, 131 (2000).
- [339] J.-S. Wang and R. H. Swendsen. *Transition matrix Monte Carlo method*. J. Stat. Phys. **106**, 245 (2002).
- [340] J.-S. Wang, Robert H. Swendsen, and R. Kotecký. *Antiferromagnetic Potts models*. Phys. Rev. Lett. **63**, 109 (1989).
- [341] J.-S. Wang, T. K. Tay, and R. H. Swendsen. *Transition matrix Monte Carlo reweighting and dynamics*. Phys. Rev. Lett. **82**, 476 (1999).
- [342] S. Wansleben and D. P. Landau. *Monte Carlo investigation of critical dynamics in the three-dimensional Ising model*. Phys. Rev. B **43**, 6006 (1991).

Bibliography

- [343] P. O. Weir, N. Read, and J. M. Kosterlitz. *Renormalization-group treatment of the long-ranged one-dimensional Ising model with random fields*. Phys. Rev. B **36**, 5760 (1987).
- [344] J. K. Williams. *Monte Carlo estimate of the dynamical critical exponent of the 2d kinetic Ising model*. J. Phys. A **18**, 49 (1985).
- [345] K. G. Wilson. *Renormalization group and critical phenomena. i. renormalization group and the Kadanoff scaling picture*. Phys. Rev. B **4**, 3174 (1971).
- [346] K. G. Wilson. *Renormalization group and critical phenomena. ii. phase-space cell analysis of critical behavior*. Phys. Rev. B **4**, 3184 (1971).
- [347] U. Wolff. *Collective Monte Carlo updating for spin systems*. Phys. Rev. Lett. **62**, 361 (1989).
- [348] U. Wolff. *Comparison between cluster Monte Carlo algorithms in the Ising model*. Phys. Lett. B **228**, 379 (1989).
- [349] M. J. Wragg and G. A. Gehring. *The Ising model with LR ferromagnetic interactions*. J. Phys. A **23**, 2157 (1990).
- [350] F. Y. Wu. *The Potts model*. Rev. Mod. Phys. **54**, 235 (1982).
- [351] Y. Wu, M. Koerner, L. Colonna-Romano, S. Trebst, H. Gould, J. Machta, and M. Troyer. *Overcoming the critical slowing down of flat-histogram Monte Carlo simulations: cluster updates and optimized broad-histogram ensembles*. cond-mat/0412076.
- [352] H. J. Xu, B. Bergersen, and Z. Rácz. *Long-range interactions generated by random Lévy flights: spin-flip and spin-exchange kinetic Ising model in two dimensions*. Phys. Rev. E **47**, 1520 (1993).
- [353] C. Yamaguchi and N. Kawashima. *Combination of improved multibondic method and the Wang-Landau method*. Phys. Rev. E **65**, 056710 (2003).
- [354] C. Yamaguchi, N. Kawashima, and Y. Okabe. *Broad histogram relation for the bond number and its applications*. Phys.Rev.E **66**, 036704 (2002).
- [355] C. Yamaguchi and Y. Okabe. *Three-dimensional antiferromagnetic q-state Potts models: application of the Wang-Landau algorithm*. J. Phys. A **34**, 8781 (2001).
- [356] Y. Yamazaki. *Critical exponent η of isotropic spin systems with long and short-range interactions*. Phys. Lett. A **61**, 207 (1977).
- [357] Y. Yamazaki. *Comments on the critical behavior of isotropic spin systems with long- and short-range interactions*. Physica A **92**, 446 (1978).

Bibliography

- [358] Q. Yan and J. J. de Pablo. *Fast calculation of the density of states of a fluid by Monte Carlo simulations*. Phys. Rev. Lett. **90**, 035701 (2003).
- [359] A. P. Young. *On the lowering of dimensionality in phase transitions with random fields*. J. Phys. C **10**, L257 (1977)..
- [360] A. P. Young and M. Nauenberg. *Quasicritical behavior and first-order transition in the $d=3$ random-field Ising model*. Phys. Rev. Lett. **54**, 2429 (1985).
- [361] E. Yusuf, A. Joshi, and K. Yang. *Spin waves in antiferromagnetic spin chains with long range interactions*. Phys. Rev. B **69**, 144412 (2004).
- [362] C. Zhou and R. N. Bhatt. *Understanding and improving the Wang-Landau algorithm*. cond-mat/0306711.
- [363] J. Zinn-Justin. *Quantum field theory and critical phenomena* (Oxford University Press, Oxford, England, 2001).

Appendix A

Mean-field theory for the long-range Potts chain

The model is described by the following Hamiltonian,

$$H = -\frac{1}{2} \sum_{i \neq j} J_{ij} \delta_{\sigma_i, \sigma_j} - \sum_i h_i \delta_{\sigma_i, \sigma_0},$$

where $J_{ij} = |i - j|^{-\alpha}$, the spin σ_i at site i can take on the values $1, \dots, q$, the first sum runs over all pairs of sites on a chain of size L , and h_i is an external aligning field favoring condensation in state σ_0 .

The derivation in what follows is based on the variational mean-field method¹, which relies on the minimization of the following functional

$$F[\rho] = U[\rho] - TS[\rho] = \text{Tr} \rho H + kT \text{Tr} \rho \ln \rho \quad (\text{A.1})$$

with respect to a trial distribution ρ depending on the spin configuration $[\sigma]$. The trace operation means a sum over all spin configurations, and the dependence of H and ρ on the spin configuration is implied in the notation above. By relying on Lagrange multipliers, it is easy to show that $F[\rho]$ reaches a minimum whenever $\rho = e^{-H/kT} / Z$, i.e., in the case of a Boltzmann distribution. This minimum yields the free energy of the system.

The mean-field approximation consists in expressing the trial distribution ρ as a product of one-site distributions ρ_i which depend only on the spin variable at site i , i.e.,

$$\rho[\sigma] = \prod_{i=1}^L \rho_i(\sigma_i)$$

The trace operation may be rearranged as a sum involving traces on single spin variables, namely,

$$F[\rho] = kT \sum_{i=1}^L \text{Tr}_i \rho_i \ln \rho_i - \sum_{i=1}^L h_i \text{Tr}_i \rho_i \delta_{\sigma_i, \sigma_0} - \frac{1}{2} \sum_{i \neq j} \text{Tr}_i \rho_i \text{Tr}_j \rho_j J_{ij} \delta_{\sigma_i, \sigma_j}.$$

¹See, for instance, [76].

Appendix A. Mean-field theory for the long-range Potts chain

where $\text{Tr}_i = \sum_{\sigma_i=1\dots q}$. The local order parameter is given by

$$m_i = \left\langle \frac{q\delta_{\sigma_i, \sigma_0} - 1}{q - 1} \right\rangle_{\rho_i},$$

where the average is weighted by the trial density matrix ρ_i . It is straightforward to check that $m_i = 1$ whenever the spin σ_i has “condensated” in phase σ_0 , while $m_i = 0$ if it fluctuates at random (whenever $m_i = m, \forall i$, the usual *global* order parameter used, e.g., in numerical simulations, is thus recovered).

Seeing that all states but state σ_0 are equivalent, the constraint $\text{Tr} \rho_i = 1$ yields

$$\rho_i = \frac{1 - m_i}{q} + m_i \delta_{\sigma_i, \sigma_0}.$$

which leads to a re-parametrization $F[\rho]$ in terms of the m_i . The entropic term in $F[\rho]$ writes

$$\begin{aligned} \text{Tr}_i \rho_i \ln \rho_i &= \sum_{\sigma_i=1\dots q} \rho_i(\sigma_i) \ln \rho_i(\sigma_i) \\ &= \rho_i(\sigma_0) \ln \rho_i(\sigma_0) + (q - 1) \rho_i(\sigma_1) \ln \rho_i(\sigma_1) \\ &= \frac{1}{q} \left[(1 - q)(m_i - 1) \ln \frac{1 - m_i}{q} + (1 + (q - 1)m_i) \ln \frac{1 + (q - 1)m_i}{q} \right] \end{aligned}$$

The external field term reduces to

$$\text{Tr}_i \rho_i \delta_{\sigma_i, \sigma_0} = \rho_i(\sigma_0) = \frac{1 + (q - 1)m_i}{q},$$

while the “coupling” term writes

$$\begin{aligned} \text{Tr}_i \rho_i \text{Tr}_j \rho_j J_{ij} \delta_{\sigma_i, \sigma_j} &= \sum_{\sigma_i=1\dots q} \sum_{\sigma_j=1\dots q} \rho_i(\sigma_i) \rho_j(\sigma_j) J_{ij} \delta_{\sigma_i, \sigma_j} \\ &= \rho_i(\sigma_0) \rho_j(\sigma_0) J_{ij} + (q - 1) \rho_i(\sigma_1) \rho_j(\sigma_1) J_{ij} \\ &= \frac{J_{ij}}{q} [1 + (q - 1)m_i m_j] \end{aligned}$$

The mean-field free energy thus writes

$$\begin{aligned} F[\{m_i\}] &= kT \sum_{i=1}^L \frac{1}{q} \left[(1 - q)(m_i - 1) \ln \frac{1 - m_i}{q} + (1 + (q - 1)m_i) \ln \frac{1 + (q - 1)m_i}{q} \right] \\ &\quad - \sum_{i=1}^L h_i \frac{1 + (q - 1)m_i}{q} - \frac{1}{2} \sum_{j \neq i} \frac{J_{ij}}{q} [1 + (q - 1)m_i m_j]. \end{aligned}$$

Considering a uniform external field $h_i = h$, we have $m_i = m$ for all sites. In the thermodynamic limit, this yields for the free energy per spin,

$$\begin{aligned} f(m) = \frac{F(m)}{L} &= kT \frac{1}{q} \left[(1 - q)(m - 1) \ln \frac{1 - m}{q} + (1 + (q - 1)m) \ln \frac{1 + (q - 1)m}{q} \right] \\ &\quad - h \frac{1 + (q - 1)m}{q} - \frac{1 + (q - 1)m^2}{q} \sum_{r>0} \frac{1}{r^\alpha} \end{aligned}$$

Dropping terms which are constant in m , so that $f(0) = 0$, this leads to

$$\frac{qf(m)}{q-1} = kT\{(1-m)\ln(1-m) + \frac{1+m(q-1)}{q-1}\ln[1+m(q-1)]\} - hm - \zeta(\alpha)m^2, \quad (\text{A.2})$$

where $\zeta(\alpha) = \sum_{r=1}^{\infty} \frac{1}{r^\alpha}$ is the Riemann zeta function. The equilibrium order parameter is obtained from the equation of state

$$\frac{q}{q-1} \frac{df(m)}{dm} = kT \ln \frac{1+m(q-1)}{1-m} - 2m\zeta(1+\sigma) - h = 0$$

In the following, we are interested in the zero-field solutions. Equilibrium values of the order parameter are located at minima of the free energy, and it can be seen from the Landau-expansion

$$\frac{qf(m)}{q-1} = \left(\frac{kTq}{2} - \zeta(1+\sigma) \right) m^2 - \frac{kT(q-2)q}{6} m^3 + O(m^4)$$

that $m = 0$ is a stable minimum for $kT \geq 2\zeta(1+\sigma)/q$. For $q = 2$, there is no third-order term: a second-order transition occurs at

$$kT_c = \zeta(1+\sigma).$$

For $q \geq 3$, the negative coefficient in the third-order term creates a second minimum, which corresponds to a first-order transition. At the transition temperature, the free energy has the same value at both minima. The exact transition temperature kT_c may be computed by simultaneously solving $f(m) = f'(m) = 0$ and yields

$$kT_c = \zeta(1+\sigma) \frac{q-2}{(q-1)\ln(q-1)}.$$

Metastable states (i.e., spinodal points) are computed by jointly solving $f'(m) = f''(m) = 0$, giving temperature points at which either one of the two minima vanishes,

$$\frac{1}{m} \ln \frac{1+m(q-1)}{1-m} = \frac{q}{(1-m)(1+m(q-1))} = \frac{2\zeta(1+\sigma)}{kT}$$

These equations possess one trivial solution, namely, $kT_1 = 2\zeta(1+\sigma)/q$ corresponding to the extrema at $m_1 = 0$ becoming unstable, and a nontrivial solution kT_2 which may be obtained numerically by solving the following equation,

$$\frac{KqS-2}{2q-1} = \ln \left(S \sqrt{\frac{Kq}{2}} \right), \quad (\text{A.3})$$

where $S = 1 + \sqrt{1+2(1-q)/(Kq)}$ and $K = \zeta(1+\sigma)/kT_2$. The corresponding value of the order parameter at which the metastable state vanishes is

$$m_2 = \frac{q-2 + \sqrt{q(2+(K-2)q)/K}}{2(q-1)}.$$

Appendix B

Transfer matrix derivation of critical exponents for the short-range Potts chain

The model is defined on a one-dimensional lattice of length L , and is governed by the following Hamiltonian,

$$H = - \sum_{i=1}^L J \delta_{\sigma_i, \sigma_{i+1}} - \sum_i h_i \delta_{\sigma_i, \sigma_0},$$

The spin σ_i at site i can take on the values $1, \dots, q$, h_i is an external aligning field favoring condensation in state σ_0 , and periodic boundary conditions are implied, i.e., $\sigma_{L+1} = \sigma_1$. The transfer matrix approach consists in expressing $e^{-\beta H}$ as a product of transfer matrices coupling two neighboring spins, and evaluating $Z = \text{Tr} e^{-\beta H}$ at inverse temperature β from the trace of this product.

In order to give a symmetric role to the spins σ_i and σ_{i+1} , I now write $-\beta H$ as

$$-\beta H = \sum_i K (2\delta_{\sigma_i, \sigma_{i+1}} - 1) + L (\delta_{\sigma_i, \sigma_0} + \delta_{\sigma_{i+1}, \sigma_0} - 1) = \sum_i \alpha(\sigma_i, \sigma_{i+1})$$

where

$$K = \beta J / 2, \quad L = \beta h / 2,$$

and the energy has been shifted by an irrelevant constant. The transfer matrix T is defined in such a way that, in the basis $\{\sigma = 1 \dots q\}$, the matrix element $\langle \sigma | T | \sigma' \rangle = e^{\alpha(\sigma, \sigma')}$. Whence

$$\begin{aligned} e^{-\beta H} &= \prod_i e^{\alpha(\sigma_i, \sigma_{i+1})} = \prod_i \langle \sigma_i | T | \sigma_{i+1} \rangle \\ &= \langle \sigma_1 | T | \sigma_2 \rangle \langle \sigma_2 | T | \sigma_3 \rangle \dots \langle \sigma_{L-1} | T | \sigma_L \rangle \langle \sigma_L | T | \sigma_1 \rangle, \end{aligned}$$

Appendix B. Transfer matrix derivation of critical exponents for the short-range Potts chain

whereby the partition function Z at finite lattice size and inverse temperature β is given by

$$\begin{aligned} Z &= \text{Tr} e^{-\beta H} \\ &= \sum_{\sigma_1=1\dots q} \langle \sigma_1 | T^L | \sigma_1 \rangle \\ &= \sum_{i=1\dots q} \lambda_i^L \end{aligned}$$

where $\{\lambda_i\}$ is the spectrum of eigenvalues of T . In the thermodynamic limit,

$$\lim_{L \rightarrow \infty} Z(\beta, L) \lambda_1^{-L} = 1$$

where λ_1 is the largest eigenvalue (provided it is not degenerate); whence the free energy per spin

$$f = -kT \ln \lambda_1.$$

For $q = 3$, T writes

$$T = \begin{pmatrix} e^{K+L} & e^{-K} & e^{-K} \\ e^{-K} & e^{K-L} & e^{-K-L} \\ e^{-K} & e^{-K-L} & e^{K-L} \end{pmatrix}$$

Its eigenvalues are in descending order

$$\lambda_1 = \frac{e^{-K-L} \left(1 + e^{2K} + e^{2(K+L)} + \sqrt{e^{4(K+L)} + (1 + e^{2K})^2 - 2e^{2L}(-4 + e^{2K} + e^{4K})} \right)}{2}$$

$$\lambda_2 = \frac{e^{-K-L} \left(1 + e^{2K} + e^{2(K+L)} - \sqrt{e^{4(K+L)} + (1 + e^{2K})^2 - 2e^{2L}(-4 + e^{2K} + e^{4K})} \right)}{2}$$

and

$$\lambda_3 = 2e^{-L} \sinh K$$

In the thermodynamic limit, the free energy per spin is thus given by

$$f = \frac{K + L + \log(2) - \log(1 + e^{2K} + e^{2(K+L)} + \sqrt{e^{4(K+L)} + (1 + e^{2K})^2 - 2e^{2L}(-4 + e^{2K} + e^{4K})})}{K}.$$

The magnetic susceptibility in zero field is obtained from $\chi = -d^2 f / dh^2$ at $L = \beta h = 0$,

$$\chi = \frac{8(1 + 2e^{2K})K}{27}$$

and exhibits an essential singularity at $kT = 0$ (in particular, $\nu \rightarrow \infty$). The inverse correlation length is given by the standard formula $\xi^{-1} = \ln(\lambda_1/\lambda_2)$,

$$\xi^{-1} = \ln \frac{3 \coth K - 1}{2}$$

From $\chi \propto \xi^{\gamma/\nu}$, we obtain

$$\gamma/\nu = \lim_{K \rightarrow \infty} \frac{\ln \chi}{\ln \xi} = 1$$

In Chap. 5, we are also interested in finite size quantities. In order to take the finite lattice size into account, we first express the susceptibility as a function of the correlation length:

$$\chi(\xi) = \frac{8}{9} \coth^{-1} \left(\frac{1 + 2e^{1/\xi}}{3} \right) \coth \left(\frac{1}{2\xi} \right) \quad (\text{B.1})$$

At finite lattice size L , the saturation of the correlation length $\xi \sim L$ when kT is sufficiently close to zero leads to the susceptibility saturating at $\chi^{\max} = \chi(L)$. The ratio γ/ν computed from $d \ln \chi^{\max} / d \ln L$, i.e., by fitting the peaks of the susceptibility to the power law $\chi^{\max} \propto L^{\gamma/\nu}$ over nearby lattice sizes, is always overestimated with respect to the infinite size value:

$$\frac{d \ln \chi^{\max}}{d \ln L} = \frac{e^{1/L} \left(3 \left(1 + e^{1/L} \right) + 4 \left(2 + e^{1/L} \right) \coth^{-1} \left(\frac{1+2e^{1/L}}{3} \right) \right)}{2 \left(1 + e^{1/L} \right) \left(-2 + e^{1/L} + e^{\frac{2}{L}} \right) L \coth^{-1} \left(\frac{1+2e^{1/L}}{3} \right)}$$

This is even worse if one considers $\ln \chi^{\max} / \ln L$ instead.

A second quantity that is needed in Chap. 5 is the expression of the specific heat C_v . This entails explicitly taking all eigenvalues into account in the expression of Z . In what follows, I set $h = 0$. The partition function admits a simplified expression in this case,

$$Z_L = (3 \cosh(K) - \sinh(K))^L + 2^{1+N} \sinh(K)^L$$

The mean energy per spin is given by $\langle E \rangle / L = -L^{-1} d \ln Z_L / d\beta$. Recalling that $K = \beta J / 2$, this leads to

$$\frac{\langle E \rangle}{L} = -\coth(K) + \frac{3 / \sinh(K)}{3 \cosh(K) - \sinh(K) + 2^{1+L} (3 \cosh(K) - \sinh(K))^{1-L} \sinh(K)^L}$$

whence the specific heat for a chain of length L writes

$$C_v = \frac{K^2}{\sinh^2 K} \left(-1 - \frac{9L (3 \cosh K - \sinh K)^{2(L-1)}}{\left((3 \cosh K - \sinh K)^L + 2^{1+L} \sinh K^L \right)^2} + \frac{9 (L + \cosh(2K)) - 3 \sinh(2K)}{2^{1+L} (3 \cosh K - \sinh K)^{2-L} \sinh K^L + (3 \cosh K - \sinh K)^2} \right) \quad (\text{B.2})$$

Appendix C

Error calculation from Monte Carlo data

Autocorrelation time and statistical error

We consider Ω simulation runs, each containing N measurements $\{A_i\}$ of the observable A . We are interested in the error $\text{err}(\bar{A})$ on the estimator \bar{A} of the quantity $\langle A \rangle$, where

$$\bar{A} = \frac{1}{N} \sum_{i=1}^N A_i$$

for a given run. The squared error is the variance of \bar{A} ,

$$\text{err}^2(\bar{A}) = \langle (\bar{A} - \langle A \rangle)^2 \rangle = \left\langle \left[\frac{1}{N} \sum_{i=1}^N A_i - \langle A \rangle \right]^2 \right\rangle$$

where $\langle \cdot \rangle$ refers to an average over Ω runs, and $\langle A \rangle$ is estimated by $\langle \bar{A} \rangle$. $\text{err}^2(\bar{A})$ can be reexpressed in term of the characteristic time τ_A of the time-displaced autocorrelation function $\Phi_A(t)$, where the latter is defined as

$$\Phi_A(t) = \frac{\langle A_0 A_t \rangle - \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2}$$

and the former is obtained from $\tau_A = \sum_0^N \Phi_A(t)$ (Sec. 2.2.1 gives more details on how to estimate τ_A). Expanding the previous definition of $\langle \delta A^2 \rangle$, and assuming time invariance, i.e., $\langle A_i A_{i+j} \rangle = \langle A_0 A_j \rangle$, $\forall i$,

Appendix C. Error calculation from Monte Carlo data

$$\begin{aligned}
 \text{err}^2(\bar{A}) &= \frac{1}{N^2} \sum_{i=1}^N \langle (A_i - \langle A^2 \rangle) \rangle + \frac{2}{N^2} \sum_{i=1}^N \sum_{j=1}^{N-i} \langle A_0 A_j \rangle \\
 &= \frac{1}{N} [\langle A^2 \rangle - \langle A^2 \rangle] \left[1 + 2\tau_A \left(1 - \frac{\tau_A}{N} \right) \right] \\
 &\sim \frac{2\tau_A}{N} [\langle A^2 \rangle - \langle A^2 \rangle]
 \end{aligned}$$

where the last line assumes that $1 \ll \tau_A \ll N$, which is usually the case because (i) one wants to have a substantial amount of independent samples, and (ii) measurements are carried out at intervals sufficiently small with respect to τ_A , so that τ_A can be reliably estimated. The last equation also shows that the number of independent samples from where thermodynamic averages must be computed is given by

$$N_{ind} = \frac{N}{2\tau_A}$$

so that the usual formula of elementary statistics is recovered, i.e.,

$$\text{err}^2(\bar{A}) \sim \frac{\langle A^2 \rangle - \langle A \rangle^2}{N_{ind}}$$

(biased and unbiased estimators coincide whenever $N_{ind} \rightarrow \infty$). This equation yields a convenient way of computing τ_A and N_{ind} , by separately estimating $\langle A^2 \rangle - \langle A \rangle^2$ and $\text{err}^2(\bar{A})$. The former is readily obtained from the set of measurements, where $\text{err}^2(\bar{A})$ may be estimated, either by performing multiple, independent runs, or by relying on a procedure based on a single run. Three methods are presented hereafter: the bootstrap, the jackknife, and the blocking/binning methods.

The bootstrap

This is a resampling method, whereby several estimators of $\langle A \rangle$ are computed on resampled sets of measurements, and the bias on the original estimator is computed from the deviation of these “resampling” estimators. The bootstrap method relies on a Random Sample With Replacement (RSWR) scheme, i.e., a fresh set of N measurements is built from the original set by explicitly allowing each measurement in the original set to be picked more than once (usually twice). As far as implementations in Java are concerned, several random samplers are available in the package `cern.jet.random.sampling` of the COLT library¹.

Denoting as \bar{A}_0 the estimator computed from the original set of measurements, and as \bar{A}_i that computed from the i th resampled set (where $i = 1 \dots p$ and p is the

¹<http://hoschek.home.cern.ch/hoschek/colt/>

number of resampled sets), the bootstrapped estimator is given by

$$\bar{A}_B = \frac{1}{p} \sum_{i=1}^p \bar{A}_i$$

The bias of the estimator \bar{A} is $\bar{A}_B - \bar{A}_0$, whereby the *unbiased* estimator is given by

$$\bar{A}_u = \bar{A}_0 - (\bar{A}_B - \bar{A}_0) = 2\bar{A}_0 - \bar{A}_B \quad (\text{C.1})$$

The error, i.e., the **confidence interval** on \bar{A} is eventually obtained from the variance

$$\text{err}^2(A) = \frac{1}{N-1} \sum_{i=1}^p (\bar{A}_i - \bar{A}_B)^2$$

where N is the number of measurements! For instance, a 95% confidence interval corresponds to $\pm 1.96\sigma$. For this confidence interval to be reliable, at least 100 to 200 bootstrapped estimators have to be computed. This is the most accurate method to estimate errors on statistical averages, yet this method is also markedly demanding in terms of CPU and memory consumption when performed over large sets of measurements.

The Jackknife

This is also a resampling method, yet now resampled sets are constructed by simply deleting one or more measurements from the original set (hence the name “jackknifed” sets). Denoting as $\{\bar{A}_i\}$, $i = 1 \dots J$, the set of jackknifed estimators, and \bar{A}_J the mean of these estimators, the error on \bar{A} is given by

$$\begin{aligned} \text{err}^2(A) &= \frac{J-1}{J} \sum_{i=1}^J (\bar{A}_i - \bar{A}_J)^2 \\ &\sim \sum_{i=1}^J (\bar{A}_i - \bar{A}_J)^2 \end{aligned}$$

provided $J \gg 1$. Usually one takes $J \sim N$, where N is the number of measurements, so that this method suffers from the same issues as the bootstrap method when large sets are considered.

The blocking method

The original set of measurements is split up into B contiguous blocks of length k , i.e., with $N = kB$. Denoting as $\{\bar{A}_i\}$, $i = 1 \dots B$, the set of block estimators, as \bar{A}_B

Appendix C. Error calculation from Monte Carlo data

the mean of these estimators, and as $\text{err}^2(A)_B$ their variance, the error on \bar{A} is given by

$$\begin{aligned}\text{err}^2(A) &= \frac{\text{err}^2(A)_B}{B} \\ &= \frac{1}{B(B-1)} \sum_{i=1}^B (\bar{A}_i - \bar{A}_B)^2\end{aligned}$$

This method is efficient in terms of CPU load (the number of operations is constant whatever the number of blocks), yet it suffers from a stark dependence on the number of blocks. Several estimations must therefore be carried out at distinct block lengths, until a plateau is observed.