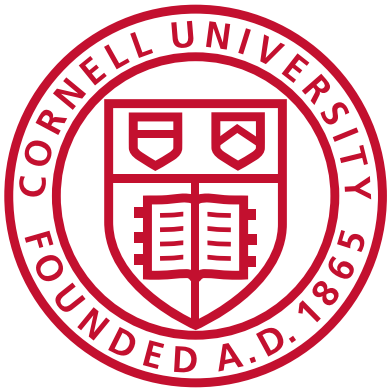

NOVEL CRITICAL PHENOMENA

A dissertation presented to the faculty of the Graduate School of Cornell University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Jaron Patrick Kent-Dobias

August 2020

Novel Critical Phenomena

Jaron Patrick Kent-Dobias, Ph.D.
Cornell University 2020

The renormalization group is nearly fifty years old, and has proven to be a comprehensive description of equilibrium critical phenomena. Fifty years may seem a long time, but after fifty years of quantum mechanics physicists were only just coming to terms with the significance of entanglement. New frontiers abound, but dark corners in our understanding remain. This thesis, in various ways, pokes around some of those dark corners.

One burgeoning frontier is in the critical phenomena of nonequilibrium systems. Chapter 1 describes the skeleton of a new scaling theory for quasibrittle fracture as a departure from the percolation fixed point along an invariant surface that becomes trivial in the thermodynamic limit. After reviewing the set of numeric tools needed to simulate a popular model for quasibrittle fracture across a wide scale of disorders, we compare simulation results to several versions of the scaling theory. Qualities of the numeric model at critical stress and after rupture that are singular at the percolation fixed point seem well-described by the simplest versions of that theory, while dynamic qualities that are nonsingular at the percolation fixed point appear to be governed by different scaling. At the end we reflect on possible explanations for this discrepancy, and paint a path towards distinguishing them.

If the first chapter demonstrates the importance of numeric experiments for assessing the predictions of scaling theories, the next two offer some new tools for performing certain numeric experiments. Chapter 2 describes an extension to cluster Monte Carlo algorithms that permit their efficient use

in the presence of arbitrary on-site potentials. These algorithms were vital for making precision numeric measurements near the critical points of lattice models, and our extension offers new opportunities for study. We give two examples: determining the relevance of symmetry-breaking fields on the 2D XY model, and constraining the form of universal scaling functions in the metastable state. Chapter 3 describes how our extension works equally well for the application of inhomogeneous potentials in geometric cluster flip algorithms for the simulation of colloidal or atomistic systems. We detail the form of the extension and speculate on possible uses, from the simulation of novel phase transition boundaries in gravitational fields to improving the equilibration of computer glasses.

After building some useful tools, we return to a dark corner. Chapter 4 examines the influence that an obscure and subtle singularity near lines of abrupt phase transitions has on the universal scaling at the critical points that terminate them. We show that a simple ansatz for the scaling functions in the metastable state, across the abrupt transition, gives a close prediction for the scaling functions and their derivatives in the stable state, at least for the 2D Ising model. We discuss techniques for improving the simplest prediction perturbatively and speculate on why the description is so much worse in 3D.

Chapter 5 addresses another subtle question: what makes two critical points the same? Specifically, it examines the breadth of phenomena, equilibrium or not, whose critical behavior is *infinite order* and whose critical singularities are not power laws, but stretched exponentials. We question whether every transition claimed to be in the Berezinskiĭ–Kosterlitz–Thouless universality class actually is, and go hunting for corrections to scaling in a

model of growing networks that would prove otherwise.

Finally, we return to more familiar waters. Chapter 6 uses functional mean field theory to explain anomalous singular behavior near the critical point of URu_2Si_2. Data from experiments that measure components of the modulus tensor across an array of temperatures reveal behavior in one mode of the modulus that appears to decay like a power law above the critical point but is not singular at it. We show that this can be explained by the critical point of an order parameter with the same symmetry of that mode, but which becomes modulated, not uniform, in the low-temperature phase. We reflect on other evidence for this idea and on future experiments that could support or falsify it.

BIOGRAPHICAL SKETCH

I was born on 20 June 1992 in Walnut Creek, California, though I remained there for less than two months. After brief stints in New Orleans and Redmond, my family settled in Sammamish, Washington, at the time a fledgling suburb on the outskirts of the Seattle metropolitan area consisting entirely of single-family homes, primary schools, churches, and strip malls. In this respect it is largely unchanged, though it has now has grown into the 19th most populous city in Washington. Having gotten in relatively early, I grew up in a more spread-out neighborhood of somewhat unique houses on a property with a small private forest. My parents still live there.

Growing up in that environment, though full of the organized activities of Tae Kwon Do and the Boy Scouts, was ultimately isolating, with an activation energy and a car required to go most anywhere or do most anything outside the home. I only infrequently met that activation energy, and spent much of my free time at home reading. By middle school part of that reading attention was turned towards popular scientific nonfiction. My parents encouraged me to maintain better grades by enrolling in study hall, and rather than doing homework, I would just read those books for an hour every day, eventually going through most of the middle school library's offerings. Internally, I imagined scientific life as Doc-from-*Back-to-the-Future*-equese and full of those

lightening lamps from electronics catalogs, for some reason. Though I took apart a lot of electronics, I never managed to put any back together.

By high school that vague science interest had coalesced on physics. I vividly remember finding a copy of Feynman's QED, a popular introduction to path integrals, in a used bookstore and becoming obsessed with the concepts it laid out. Through the rest of primary school I was on a quantum physics and cosmology bender, fully into the ideology and fetishization of physics as "fundamental."

I (thankfully) got into the only college I applied to: Harvey Mudd College, a tiny liberal arts college east of Los Angeles with a specific focus on math, engineering, and the sciences, always intending to study physics. I got a wonderful training there, including mentorship from Professors Sharon Gerbode and Andrew Bernoff, who introduced me to a frontier of physics far from the "fundamental" but just as challenging: that of the soft and squishy. This interest led me to Professor Gerbode's alma mater, and an old collaborator of hers.

At Cornell I have had the opportunity to become involved in a wide variety of research topics, both with my advisor and a rich array of collaborators. I have been able to teach classes in innovative and exciting ways with the guidance of experts in physics education. I have been involved a unionization struggle, which though illegally suppressed by the graduate school taught me a lot about power and made me many rich friendships. I met and married my wife. I adopted a cat that is now near death. A lot has happened, and I am richer for it.

This document is dedicated to Penny.

ACKNOWLEDGEMENTS

Extensive use was made of GNU Parallel in nearly every chapter of this work.[1]

Thanks to Adam Stones, who suggested the idea to extend the work of Chapter 2 into that of Chapter 3 in a conversation after my talk at StatPhys27.

Kacey Acquilano and Deb Hatfield were very helpful in managing my research and teaching appointments. Special thanks to Karl Hill for helping me print an awful number of things.

Thanks to John Miner for trusting me with a cool assistantship my first semester as a graduate student.

Jonathan Fuller was extremely helpful and patient in managing my conference reimbursements despite the increasingly intricate trips I arranged.

Barry Robinson has been an essential help in setting up and maintaining the research computers that so much of this work relies on.

Thanks to Natasha Holmes, Emily Smith, and Z. Yasemin Kalender for trusting me with and guiding me in research-based laboratory instruction, which has enriched both my teaching and my philosophy of science.

Thanks to Archishman Raju and David Hathcock, who were both valuable collaborators on the content of Chapter 5. Thanks to Michael Matty, who pushed me to finally resolve a dangling question from my candidacy exam and was so much fun to work with on Chapter 6.

Thanks to my wife Emily, who supported me through much of my degree and especially so during my A and B exams.

My extended committee, Professors Brad Ramshaw and Veit Elser, have both been helpful mentors and proffered great collaborations.

Of course my advisor, Jim Sethna, has been an invaluable resource, mentor, and inspiration. I can only hope that I've picked up a little of his knack for knowing what questions to ask.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

SCALING IN QUASIBRITTLE FRACTURE

Everything we build will eventually fail, but we would like to know when and how. Here I will discuss some ideas we have about how the way concrete fails might depend on a structure's size. Concrete is an example of a *quasibrittle* material. This means that it is brittle (does not tolerate much plastic deformation) but disordered (full of inhomogeneities). Though it is one of the most common building materials, we don't have a firm idea about how to take its failure statistics at small, laboratory scales and extrapolate them to the scale of buildings and dams. This is a problem because concrete is often used to construct buildings and dams.

The trouble seems to be wrapped up in the way that concrete yields under strain. In plastic materials like metal and—well—plastic, complicated microscopic processes mediate plastic yielding: individual and collective movement of dislocations, and polymer breaking, rejoining, and untangling, respectively. Despite the complexity, these phenomena behave well under mesoscopic and ultimately macroscopic coarse graining, and yielding is well-described by theories with continuum elastic fields.[2,3] In concrete, yielding largely progresses through a process called *microfracture*: when 'stretched,' a sample opens up

myriad tiny cracks across many length scales which collectively conspire to reduce the bulk modulus and admit more strain.[4] This process does not behave well under coarse graining, and continuum theories of concrete yielding preform poorly outside of the systems they were tuned to.

This leads to material behavior whose size effects are difficult to predict; strength tests on a concrete sample in the lab have little bearing on the strength of concrete in a dam.[5] The reason for this lies in the fracture procedure, which is heavily affected by microfracture. At the tip of a perfect crack, stress diverges. In ordinary brittle materials, this divergent stress leads to immediate breaking,* while in plastic materials like metals a small region—usually on the order of micrometers—around the crack tip is plastically deformed, absorbing energy and blunting the tip. In quasibrittle materials, however, stress at the crack tip is relieved by microcracking: the formation of myriad small cracks in the vicinity of the larger one, with stress being distributed across all. This microcracking is a highly correlated phenomena, and for materials like concrete can take place on the order of feet.[4] Since the microfracture that blunts the progression of cracks though the material appears to span length scales up to the size of the structure, neglecting the finite system size in our analysis just won't do.

Numerous scaling theories have been proposed for disordered fracture, some at fixed disorder and others while the disorder is varied, that seek to explain both the structures that result and the avalanche dynamics that characterize the lead up to fracture, often with great success.[6–16] A few have attempted to trace the behavior of disordered fracture to infinite disorder,

*This is why you see so few partial cracks in glass.

where percolation physics is thought to describe the crumbling failure that occurs.[17–22] At this point, it has been well-established that common infinite-disorder numeric models do reproduce percolation physics in this limit, and have a crossover governed by the percolation fixed point as disorder is decreased.[17,21] There has been less success connecting the critical avalanche dynamics seen at intermediate disorder to this infinite disorder limit. The main prior work making this connection relies heavily on corrections to scaling.[17]

In this chapter, I examine the question of this infinite-disorder to intermediate disorder crossover in quasibrittle systems, mostly by numeric experiments on fuse networks. First, I will introduce those networks and the details of how they are simulated. Then, I'll lay out the minimum pieces of a scaling theory that has fracture cross over from a percolation critical point, and make some guesses at some underconstrained details. We'll step through the simulation results and reflect on their support—or lack thereof—for the simplest crossover models. Finally, I'll discuss some ideas on resolving seeming inconsistencies with a scaling picture, and look towards using these results to describe the structure and dynamics of growing cracks in these materials.

§1. Resistive networks

We study fracture using a simple analogue to elasticity: DC circuits. Stress on a bond is represented by current, and strain at a node is represented by voltage. The scalar-valued objects current and voltage are far more wieldy than their tensor-valued elastic counterparts, and their governing equations match the elastic equations in the limit that strain is purely compressive,

e.g., maintains the relative positions of all elements up to a global length scale. Some features are lost in this translation—for instance, we'll talk of percolation on our networks, where the relevant process for an elastic system is *rigidity percolation*—but we believe the same scaling laws should apply. The details of our numeric model follow.

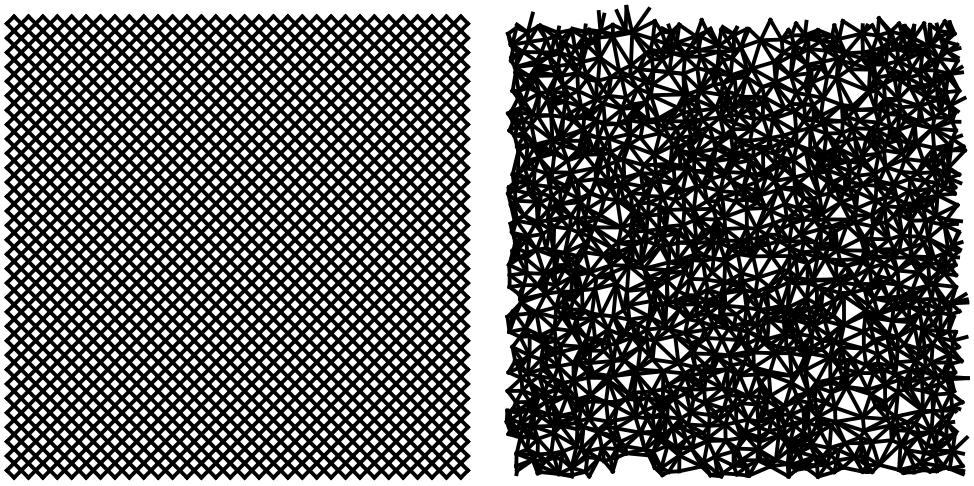

Figure 1.1: Resistive networks that we use. (Left) a square network tilted so as to ensure that bonds have symmetric exposure to current (right) a network comprised of the faces of voronoi cells for uniform random points.

A *resistive network* G is a simple connected undirected graph $G = (V, E)$. Each edge $e \in E$ represents a resistor of unit resistance. Assuming $|V| = N$, let $\{x_1, \cdots, x_N\} = V$ and adopt the natural ordering $x_i < x_j$ if $i < j$. In our work we strictly deal with two-dimensional resistive networks, which further requires that G be planar.* Kirchoff's law requires that the net current flowing from each node be zero, which (using Ohm's law) means that the voltages $u(x)$

*Or at least planar *locally*. We will be modelling fracture on a torus, after all.

for each $x \in V$ obey

$$0 = \sum_{\{y|\{x,y\}\in E\}} u(x) - u(y) \tag{1.1}$$

where the sum is over neighboring sites. We can represent this requirement as a matrix equation. Consider the adjacency function $A(G) : V \times V \to \mathbb{Z}_2$ defined for $x, y \in V$ by

$$(A(G))(x,y) = \begin{cases} 1 & \text{if } \{x,y\} \in E \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

This gives 1 if the sites are connected by an edge, and if they are not. Consider further the degree function $D(G) : V \to \mathbb{N}$ defined for $x \in V$ by

$$(D(G))(x) = \sum_{y\in V} \begin{cases} 1 & \{x,y\} \in E \\ 0 & \text{otherwise} \end{cases} \tag{1.3}$$

which gives the degree of each vertex. If we define the function $L(G) : V \times V \to \mathbb{Z}$ by $L(G)(x,y) = \delta_{xy}(D(G))(x) - (A(G))(x,y)$, then (1.1) is equivalent to the requirement that

$$\begin{aligned} 0 &= (D(G))(x)u(x) - \sum_{y\in V}(A(G))(x,y)u(y) \\ &= \sum_{y\in V}(L(G))(x,y)u(y) \end{aligned} \tag{1.4}$$

which is equivalent to the matrix equation $[L(G)][u] = 0$ for the matrix with components

$$[L(G)]_{ij} = (L(G))(x_i, x_j) \tag{1.5}$$

and the vector with components $[u]_i = u(x_i)$. $L(G)$ is called the *Laplacian* of G, and $[L(G)]$ is called the *Laplacian matrix.*[23] The name is not spurious; in fact, the Laplacian matrix of a planar network is the natural discretization of the Laplacian operator ∇^2 for continuous space. Its relevance here is similar to its relevance for the diffusion equation: it is the simplest isotropic linear operator which produces a conservation law.

Boundary conditions

Though we have reduced determining vertex voltages to a linear problem, that linear problem is trivial as currently posed. First, the zero vector is a solution. Since we have required that G be connected, $[L(G)]$ has precisely one zero eigenvalue, which corresponds to an eigenvector $[u] = (1, \ldots, 1)$.[24] All solutions to the problem posed above, then, have $u(x)$ constant for all $x \in V$, and constant voltage implies zero current. This is a complex way of seeing an obvious physical fact: isolated networks of resistors do not spontaneously carry currents. We will need to modify the problem by adding a boundary condition of some kind. Given a resistive network and a boundary condition x of boundary type B, we want to find the current across each edge. The graph $G^* = (V^*, E^*)$ and the injective function $f^* : E \to E^*$ represent a transformed version of G for which the desired boundary has been accounted for.* G^* resembles G, but depending on the boundary conditions desired G^* can equal G, contain G, or neither. In all cases, $V^* \subseteq V$.

No matter the boundary conditions, the same general technique is used. G^* is derived from G based on the boundary type desired. A boundary condi-

*This seems very cryptic now, but will become clearer given the coming examples.

tion $x : V^* \to \mathbb{R}$, is chosen, along with an additional function $M : V^* \times V^* \to \mathbb{R}$ that helps fix the boundaries in certain cases. A *voltage function* $u : V^* \to \mathbb{R}$ is computed by solving the linear system $[L(G^*) + M(G^*)][u] = [x]$. A *current function* $i : E \to \mathbb{R}$ is then computed from u by the linear transformation $m : (V^* \to \mathbb{R}) \to (E \to \mathbb{R})$ defined for $\{x, y\} \in E$ and $\{x^*, y^*\} = f^*(\{x, y\}) \in E^*$ by

$$i(\{x, y\}) = (m(u))(\{x, y\}) = \begin{cases} u(x^*) - u(y^*) & x^* > y^* \\ u(y^*) - u(x^*) & x^* < y^* \end{cases} \tag{1.6}$$

which is the realization of Ohm's law. The resulting values $i(\{x, y\})$ give the current (in the direction $x \to y$ for $x < y$) on the resistor represented by $\{x, y\}$.

The simplest boundaries to adopt are current boundaries. In this case, one defines a *current source function* $I : V^* \to \mathbb{R}$ which gives the net current flowing out of each vertex. Now Kirchoff's law gives

$$I(x) = \sum_{\{y | \{x,y\} \in E^*\}} u(x) - u(y) \tag{1.7}$$

for $x \in V^*$, or $[L(G^*)][u] = [I]$, which yields nontrivial solutions for any I for which $\sum_{x \in V^*} I(x) = 0$. Mathematically, this is because the sum over vertices of the adjacency function equals the degree, so

$$\begin{aligned} \sum_{x \in V^*} I(x) &= \sum_{x \in V^*} \sum_{y \in V^*} (L(G^*))(x, y) u(y) \\ &= \sum_{y \in V^*} u(y) [(D(G^*))(y) - \sum_{x \in V^*} (A(G^*))(x, y)] = 0 \end{aligned} \tag{1.8}$$

but physically this is the statement that the amount of current entering the network must equal that exiting. If one is content selecting point sources within the network as the boundary, one can take $G^* = G$ and $M = 0$ and be done. For our purposes, we will tweak this slightly so as to apply virtual bus-bars to our network. Let $V_t, V_b \subset V$ be the vertices on the top and bottom boundaries of our network, respectively. Then we define $V^* = V \cup \{x_t, x_b\}$ and

$$E^* = E \cup \{\{y, x_t\} \mid y \in V_t\} \cup \{\{y, x_b\} \mid y \in V_b\} \tag{1.9}$$

Let $I(x) = 0$ for all $x \in V$ and $I(x_t) = 1$, $I(x_b) = -1$. We have added two fictitious vertices to our network, connected them to the network's top and bottom, and made them a source and sink, respectively. In order to produce more numeric stability and make voltage solutions unique, we will also require $u(x_t) = 0$ by setting $M(x, y) = \delta_{x_t x}\delta_{x_t y}$, as

$$\begin{aligned} 0 &= I(x_t) - \sum_{y \in V^*} \big[(L(G^*))(x_t, y) + M(x_t, y)\big]u(y) \\ &= u(x_t) + I(x_t) - \sum_{y \in V^*} (L(G^*))(x_t, y)u(y) = u(x_t) \end{aligned} \tag{1.10}$$

This removes the zero eigenvalue from $[L(G^*) + M]$ and therefore the degree of freedom represented by constant shifts in u. In order for the presence of the boundary to minimally effect the resulting currents, we often set the resistance of the extra edges in $E^* - E$ to near-zero values, which is equivalent to multiplying all associated terms in the Laplacian by a large constant.

Now suppose we want to fix the *voltage* on a subset of vertices $V_1 \subset V$

to equal U and the voltage on another subset $V_0 \subset V$ to equal zero. These vertices will be removed from the linear problem because they no longer need to be solved for, but how do we account for them in the solution for the free vertices? If we define

$$\begin{gathered} E_0(x) = \{\{x,y\} \in E \mid x \in V_0\} \quad E_1(x) = \{\{x,y\} \in E \mid x \in V_1\} \\ E_0 = \bigcup_{x \in V} E_0(x) \qquad\qquad E_1 = \bigcup_{x \in V} E_1(x) \end{gathered} \tag{1.11}$$

and $E^* = E - (E_0 \cup E_1)$, then for a vertex $x \in V^* = V - (V_0 \cup V_1)$, we have

$$\begin{aligned} 0 &= \sum_{\{x,y\} \in E} (u(x) - u(y)) \\ &= \sum_{\{x,y\} \in E^*} (u(x) - u(y)) + \sum_{\{x,y\} \in E_0} (u(x) - u(y)) + \sum_{\{x,y\} \in E_1} (u(x) - u(y)) \\ &= \sum_{y \in V^*} L_{G^*}(x,y) u(y) + u(x)\Big(\sum_{\{x,y\} \in E_0} 1 + \sum_{\{x,y\} \in E_1} 1 \Big) - \sum_{\{x,y\} \in E_1} U \end{aligned} \tag{1.12}$$

If we write $M(x,y) = \delta_{xy}(|E_0(x)| + |E_1(x)|)$, then we have

$$U|E_1(x)| = \sum_{y \in V^*} (L_{G^*}(x,y) + M(x,y)) y \tag{1.13}$$

Thus the boundary condition x is given by $x(v) = U|E_1(v)|$ for all $v \in V^*$.

If G is locally planar but globally a tiling of a torus,* we can still compute currents on it assuming that voltage around one axis of the torus increases

*I was introduced to the possibility of toroidal boundaries in an unpublished document sent to me by either Stefano Zapperi or Vincenzo Vitelli (I can't find the relevant email), but I couldn't make heads or tails of that description, so I eventually worked this out myself.

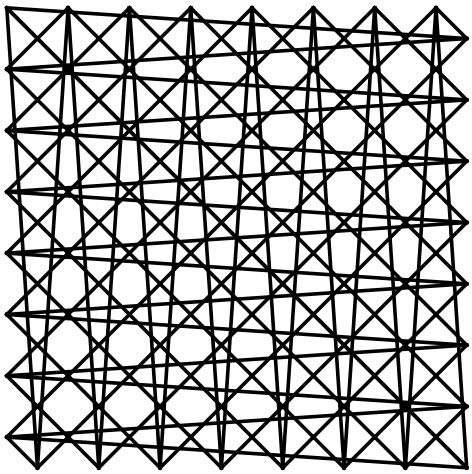

Figure 1.2: While all figures in this paper seem to show planar networks, they are actually just visually appealing representations of networks on a torus. This is a square network with the wrapping edges explicitly shown—not very comprehensible!

by ΔU with each traversal.* This would require the voltage function to be multi-valued, and actually violate Kirchoff's laws. We will get around this by cutting the torus. First, cut the torus along one axis. Let the set of cut edges be E_c, and let V_1 and V_2 be the sets of vertices represented in E_c on each side of the cut. Consider new sets V_1^* and V_2^* with $|V_1^*| = |V_1|$ and $|V_2^*| = |V_2|$, and let $f_1 : V_1 \to V_1^*$ and $f_2 : V_2 \to V_2^*$ be bijections. Then define the graph $G_f = (V_f, E_f)$ where $V_f = V \cup V_1^* \cup V_2^*$,

$$\begin{aligned} E_1 &= \big\{\{x, f_1(y)\} \mid \{x, y\} \in E_c, y \in V_1\big\} \\ E_2 &= \big\{\{x, f_2(y)\} \mid \{x, y\} \in E_c, y \in V_2\big\} \end{aligned} \tag{1.14}$$

*This is equivalent to there being some net magnetic flux threaded through one of the principal axes of the torus.

and

$$E_f = (E - E_{\mathsf{c}}) \cup E_1 \cup E_2 \tag{1.15}$$

Kirchoff's laws will faithfully reproduce the torus current on this graph provided we ensure that $u(f_1(y)) = u(y) + \Delta U$ and $u(f_2(y)) = u(y) - \Delta U$. Let $E' = E - E_{\mathsf{c}}$. We have, for some $x \in V$,

$$\begin{aligned}
0 &= \sum_{\{x,y\}\in E_f} (u(x) - u(y)) \\
&= \sum_{\{x,y\}\in E'} (u(x) - u(y)) + \sum_{\{x,y\}\in E_1} (u(x) - u(y)) + \sum_{\{x,y\}\in E_2} (u(x) - u(y)) \\
&= \sum_{y\in V} L_{(V,E')}(x,y)u(y) + \sum_{\{x,y\}\in E_1} (u(x) - u(f_1^{-1}(y)) + \Delta U) \\
&\quad + \sum_{\{x,y\}\in E_2} (u(x) - u(f_2^{-1}(y)) - \Delta U) \\
&= \sum_{y\in V} L_{(V,E')}(x,y)u(y) + \sum_{\{x,y\}\in E_{\mathsf{c}}} (u(x) - u(y)) \\
&\quad + \sum_{\{x,y\}\in E_1} \Delta U - \sum_{\{x,y\}\in E_2} \Delta U \\
&= \sum_{y\in V} L_G(x,y) + |E_1(x)|\Delta U - |E_2(x)|\Delta U
\end{aligned} \tag{1.16}$$

Remarkably, the linear problem can be solved in terms of the Laplacian of the original torus, with specially added sources and sinks! We let $x(v) = (|E_2(v)| - |E_1(v)|)\Delta U$. We can now also remove a final trivial degree of freedom by setting $u(x_1) = 0$ as we did for the current boundaries.

§2. Fuse networks

A *fuse network* $F = (G, t)$ is a resistive network $G = (V, E)$ with a function $t : E \to (0, 1]$ that corresponds to the breaking current of each fuse.[25–27] Our networks will take thresholds distributed by $p(t \mid \beta) = \beta t^{\beta-1}$, which is very peaked around $t = 1$ for large β and very peaked around $t = 0$ (spanning many orders of magnitude) for small β. A fuse network is fractured deterministically under the following procedure. The current function $i : E \to \mathbb{R}$ is computed on G using whatever boundaries are desired. The edge $e \in E$ which maximizes $|i(e)|/t(e)$ is found. A new graph $G' = (V, E - \{e\})$ and a new fuse network $F' = (G', t')$ (for the natural restriction of the domain of t) are then created. This simulates the breaking of elements in the network as the voltage (or current) is increased adiabatically. Since we are simulating the fracture of tori, this process can be thought of as physically like increasing the amount of flux threaded through one of the principal axes. When this procedure is iterated, it results in a chain $(F, F', F'', \ldots, F^{\prime\cdots\prime})$ of networks, the final one of which is *fractured*. A network is fractured when

- (in the case of current boundaries) the source and sink belong to different connected clusters, or
- (in the case of voltage or toroidal boundaries) the current function $i(e) = 0$ for all $e \in E$.

The distinction in the definition of "fractured" is present only because in the current case the linear problem used to find the current function has no solution, and so it cannot be solved for $i(e) = 0$. This process of breaking networks in sequence can be done very efficiently. Because, when a fuse is

broken, only four elements in two rows and columns of the Laplacian are affected, one can solve for the currents of the new graph by using a rank-one transformation to the Laplacian's Cholesky decomposition (used by the numeric solver) rather than refactor the new Laplacian each time, a very expensive operation.[28] We use the CHOLMOD library for this.[29]

Trimming the backbone: lollipops and bow ties

Simulation of the fuse network may seem straightforward enough, but a complexity arises in practice. As fuses are broken, the distribution of currents on the remaining fuses are changed, and sometimes carry no current at all. The numeric problem of determining the currents cannot distinguish between a fuse which carries no current and one that carries a current at or less than the level of precision in the linear problem.* Figure 1.3 shows the distribution of currents for large systems with small β. The hump at small currents is not physical—it represents bonds that should be identically zero. For this system size those bonds cannot be distinguished from ones carrying a physical current, since the distributions overlap.

Though the inability to distinguish between fuses with small current and those with none may seem a problem of no more consequence than the inability

*A careful reader might notice another problem that could arise from the indistinguishably of small and zero current: how can you ensure that the torus is broken by confirming that $i(e) = 0$ if we can't be sure $i(e)$ is zero or just near zero? An inconvenient way would be to search for a cycle of the correct signature over and over until we find it, confirming current can no longer be carried. A better way is to mark the edges along the cut in the torus, and after every bond breaks sum up their values (careful to preserve the sign of the currents across the cut). This gives the total current on the system, which for $\Delta U = 1$ is also the modulus. The *minimum* modulus is given by a network with all the fuses lined up in series, which would be $1/N$. This lower bound on the modulus, or total current, is much greater than the numeric jitter in our systems, and so therefore it can be reliably used to determine if the system has ruptured or not.

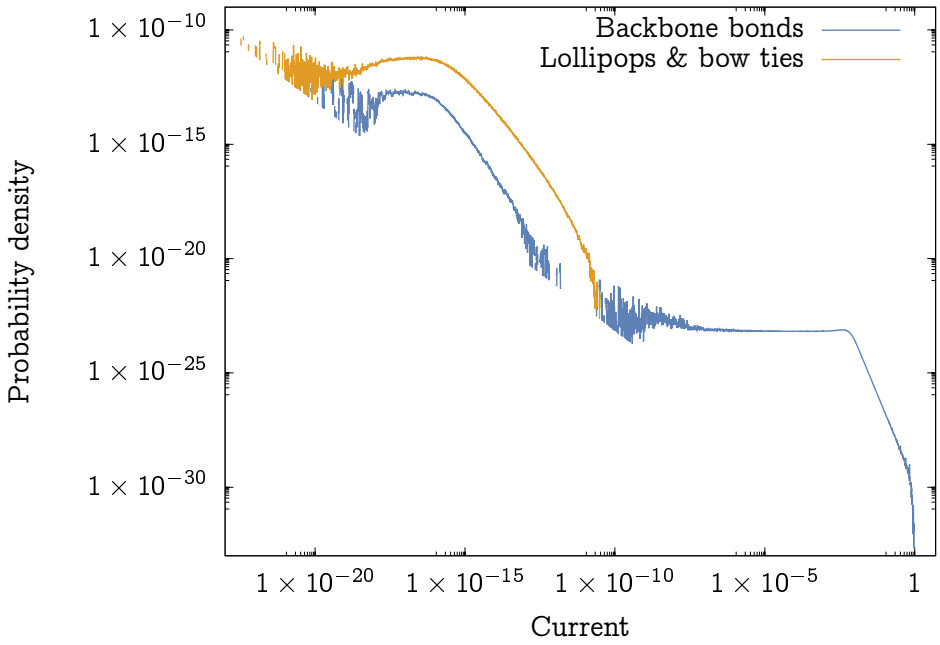

Figure 1.3: The distribution of currents on all fuses measured over the course of many fuse network fractures for a voronoi lattice with $L = 128$ and $\beta = 0.0001$. The blue line depicts the current on bonds that have (by the methods detailed below) been identified as backbone bones, and the orange line depicts the current on bonds that have been identified as not in the backbone.

to distinguish between fuses with different small currents, the former often leads to a complete breakdown of the numeric problem. The trick we use to efficiently update the set of currents after each fuse is broken relies on modifying the Cholesky decomposition of the Laplacian matrix of the network, which itself relies on the positive-definite nature of that matrix. Breaking certain identically-zero fuses renders that matrix merely positive semi-definite.

Why should this be true? Here is an example. If two collections of fuses are joined only by one fuse, no current can be carried on that fuse, because

there cannot be a net one-way flow from one bunch to the other. This is a bond whose current must be identically zero. If that bond is removed, the two collections are disconnected. Because they are disconnected, the potential on each can be set arbitrarily with the same outcome. This is responsible for a zero eigenvalue—the difference from an Ohm's law solution is unchanged by such a shift. These bonds and the collections of bonds they connect to we will call *lollipops* because of their resemblance. Another class of identically-zero bonds is formed when a collection of bonds is connected to the backbone by only a single vertex—we will call these *bow ties*.

Unless we would like to destroy our chances at analyzing a network as often as such a bond is broken, we must prevent lollipops from being broken, and once we are doing this, making the same analysis for bow ties is not too much more work. Many "burning" strategies for this have been proposed,[30] but we shall make use of an idea first articulated by Roux and Hansen, suitably generalized for fuse networks spanning a torus.[31] We will first describe the idea on a simply connected network, where the application is more simple.

A group of fuses that is connected to the rest of the network by only one bond—like that pictured in the left panel of Figure 1.4—can be identified by looking at the connectivity of the network's dual. If the dual to a bond (highlighted gray in the figure) has both ends in the same connected cluster on the dual network, then the bond is the stem of a lollipop. This is because a continuous path may be drawn from one end of the dual bond to the other, and the bonds enclosed by that path must be connected to the rest of the network only by the stem in question. After identifying a stem this way, the rest of the lollipop may be found by summing the currents on the pieces of

the network on each side of the stem and looking for the one with zero net current.

A similar logic is used to identify a bow tie. Now, instead of looking for bonds on the lattice whose dual has both ends in the same connected cluster, one searches for sites—like that pictured in the right panel of Figure 1.4—that have two adjacent faces that lie in the same cluster. That condition isn't quite sufficient to merit a bow tie, though: a vertex along line of fuses has two connected faces on each side of it, but that doesn't mean that the vertex has a dangling tie flap on one side. A *sufficient* condition (at least on a simply connected network) is that the site has two adjacent faces that lie in the same cluster *and are not themselves adjacent.* Then, just as in the case of the lollipop, a connected path can be made between the faces, and the bonds enclosed must be connected to the network only by the vertex in question. The flap(s)* are then identified by traversing sections of the network bounded by the vertex in question and computing the net current they carry.

The story for the torus is considerably more complicated. Figure 1.5 shows neat counterexamples to the carefully constructed lollipop and bow tie conditions we listed above. The left (right) panel depicts a path within a connected cluster on the dual network connecting one side of a dual bond (one face adjacent to a vertex) to the other side of the dual bond (another face adjacent to the same vertex but not adjacent to the previous face). Concluding that the offending bond (site) was the stem (knot) of a lollipop (bow tie) would be

*The only unique bow tie geometry on the square lattice is shown in Figure 1.4, and so bow ties can only have one flap on those lattices. This is because every vertex has a degree four, and so there is only one way to have two adjacent faces that are not themselves adjacent. On the voronoi lattice, where the vertex degree is often considerably higher, more exotic bow ties are permitted, including sometimes with multiple flaps.

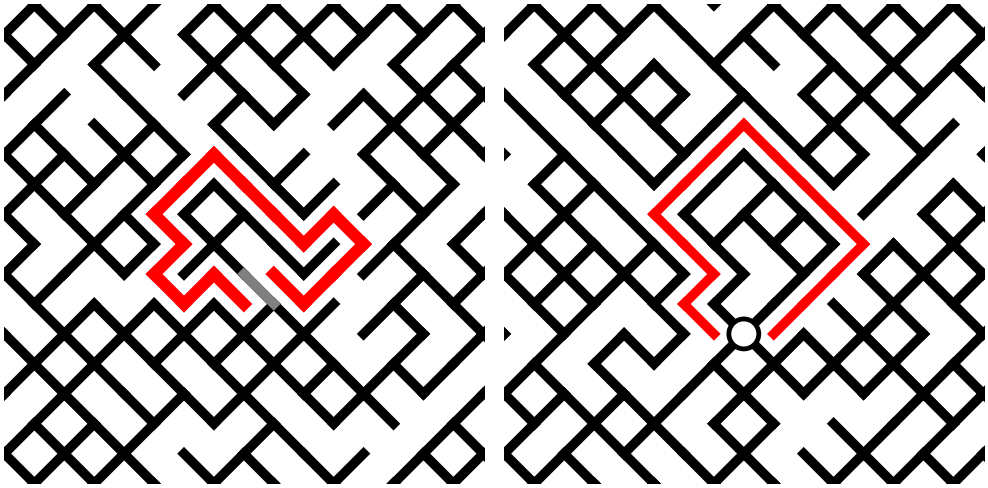

Figure 1.4: Basic collections of bonds that carry no current and the cycles that identify them (red). Left: A lollipop and its stem (highlighted edge). Right: A bow tie and its knot (highlighted vertex).

inaccurate, and severely so: these features actually carry the entirety of the current on the network, being the only elements remaining that connect top to bottom!

We must therefore add more nuance to the identification of these features. We must also exclude bonds and sites that meet the previous criteria but where the path that connects their corresponding dual sites has an odd topological signature on the torus. Such a path on the dual violates the principle that it encloses a collection of bonds connected to the network at only one point: rather, it encloses *every* bond in the network!

This combination of criteria has the potential for blowing up our efficiency. Every time a bond is broken, we would naïvely have to:

1. Iterate over every bond still in the backbone and ask if the ends of its dual lie in the same cluster (efficient), and if they do perform a depth

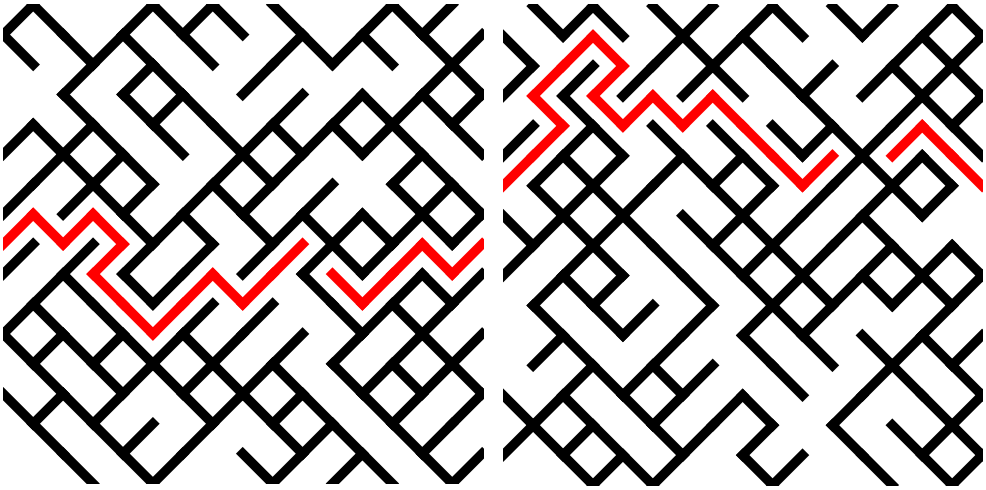

Figure 1.5: Dual vertices belonging to the same cluster as (Left) another on the same bond (Right) another around the same vertex, but that don't imply no current over the bond or site in question.

first search to find the path(s)* on the cluster that connect the two ends and ask if its topological signature is odd (inefficient).

2. Iterate over every site still in the backbone and ask if any of its non-adjacent faces lie in the same cluster (efficient), and if they do perform a depth first search to find the path(s) on the cluster that connect the two ends and ask if its topological signature is odd (inefficient).

The great inefficiency that could be involved on the torus is: once a candidate lollipop stem or bow tie knot is found that actually has a spanning path and should not be removed, we might retry removing that bond or site over and over again, yielding many unnecessary depth first searches. We can save a lot of time by remembering what dual-lattice pairs of sites we've tried before and

*Path*s*? Yes! Once the network has been broken along one axis of the torus, the two ends of a lollipop are connected by both an ordinary enclosing path *and* a path that spans the torus (and the composition of the two!). We have to make sure we've found *every* path, because if we stop after finding a spanning one then we risk missing an ordinary loop.

found a spanning path, and then immediately skip those candidates. There is one wrinkle in this picture, and it arises when the network has broken along one axis (or diagonally!) but still conducts along the other. Upon breaking along the first direction, candidate bonds or sites that previously only had a cycle that spanned the torus along the direction that is now broken can now have a couple more cycles, one of which is an ordinary cycle of the type that signifies a dangling cluster.

Therefore, the following method allows the removal of dangling lollipops and bow ties without unnecessary path searches. Every time a fuse is broken,

1. Iterate over every bond still in the backbone. If the ends of its dual lie in the same cluster and the network isn't broken along the x (y) axis, check if the pair of ends has been tried before and had a cycle with an odd signature in the x (y) direction. If it hasn't (or the network is broken along the direction that it has been tried), perform a depth first search to find all cycles connecting those sites. If *any* cycle is found with only even signatures, identify the bond as a lollipop stem and process the two connected pieces of the backbone it separates, removing the one without current. If the only cycle found has an odd signature in some direction, add the pair of sites to the dictionary of pairs that have been tried with a signature along that direction.
2. Iterate over every site still in the backbone. For every pair of adjacent faces not adjacent to each other that lie in the same cluster and the network isn't broken along the x (y) axis, check if the pair of faces has been tried before and had a cycle with an odd signature in the x (y) direction. If it hasn't (or the network is broken along the direction that it

has been tried), perform a depth first search to find all cycles connecting those sites. If *any* cycle is found with only even signatures, identify the site as a bow tie knot and process each of the connected pieces of the backbone it separates, removing every one without current. If the only cycle found has an odd signature in some direction, add the pair of sites to the dictionary of pairs that have been tried with a signature along that direction.

This suffices to successfully remove many of the offending zero-current bonds. However, as emphasized by the remaining lump of probability for currents below 10^{-10} in the curve corresponding to only purported backbone bonds in Figure 1.3, we didn't get them all...

One class of bonds we necessarily miss are those that carry zero current because of symmetry protection. Two examples of these situations are shown in Figure 1.6. They arise where two sites that join one (or several) bonds have, because of the lattice geometry surrounding them, exactly the same voltage. Though certain small configurations are by far the most common, these symmetry-protected lattice animals can get quite large. There is no good way to find and remove them from the backbone, because even if all offending subgraphs could be identified, finding them within the network is a subgraph isomorphism problem, which is NP-complete.* This isn't very concerning, since we don't expect these rare bonds are relevant to any sort of structure or avalanche dynamics. They would be completely removed if an infinitesimal amount of randomness were added to the resistances of the fuses, or if we

*A linear-time subgraph isomorphism algorithm exists for *planar* graphs, but it isn't clear whether this can be extended to toroidal graphs.[32]

incorporated dynamic jitter in the bonds. One might worry that the presence of these bonds is indicative of something fundamentally wrong with the model, since reality does not feature elements with exactly the same modulus—could lacking any resistive disorder change the universality, as it does in quantum percolation?[33] Even if this were the case, numeric jitter causes their rupture regardless, and we are powerless to stop it. If one were concerned that the preservation of the bonds would change something fundamental in the fracture process, that numeric jitter provides same effect that a small random spread in the bond resistances would.

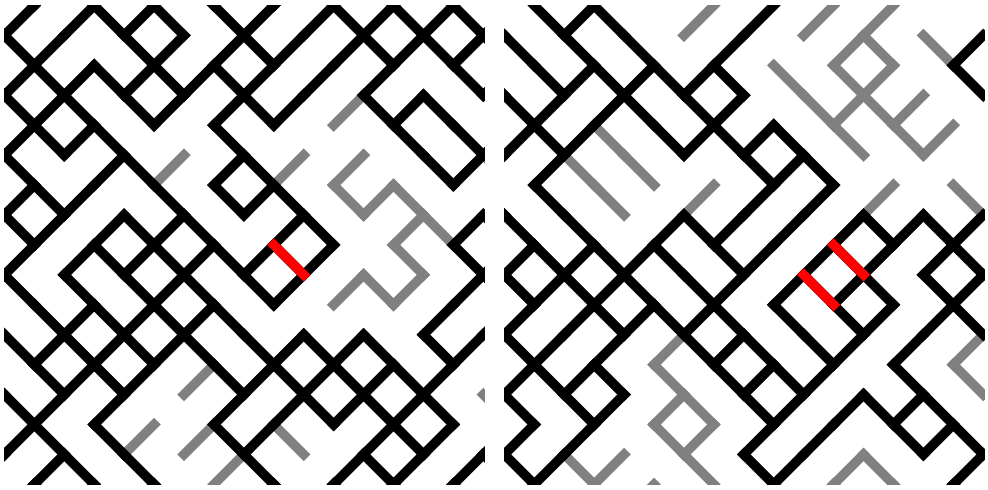

Figure 1.6: Edges that carry no current because they are protected by symmetry (red).

There is a final class of structures which have identically zero current over them, and they are *topological bow ties and symmetry-protected clusters*. These rarities appear sometimes in very disordered systems, and involve a loop of bonds that spans the torus and connects back to the current-carrying

backbone in a way that puts both ends of the loop at the same voltage. Figure 1.7 shows an example of each variety, with the left depicting a topological symmetry-protected loop, while the right depicts a topological bowie.* The same considerations that apply to standard symmetry-protected clusters applies to the topological ones, and we neither can identify them easily nor do we expect them to affect anything important. The bow ties can be identified more easily, if the same site is found to have two different nonadjacent pairs of nonadjacent adjacent faces that are connected by loops that span the torus. Right now we don't find and remove these, since they are both very rare and not conceptually very different from the symmetry-protected cases that we aren't concerned with anyway.

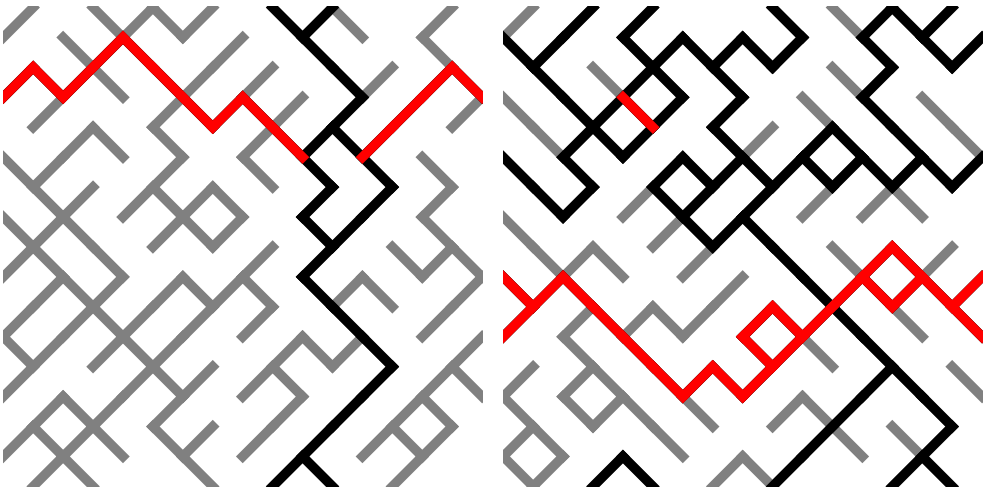

Figure 1.7: Networks with novel bond configurations. Backbone bonds are black, bonds identified as not current-carrying are grey, and bonds that cannot be identified using the described methods are red. Left: A spanning collection of bonds that connect to sites with the same voltage. Right: A spanning collection of bonds that connect to the same site.

*Why bow tie? Imagine the loop popping out of the page.

On avalanche measurement: 'stress' versus 'strain'

For a given network, one can find the conductivity Λ of the network (the analogue to elastic modulus) by

- (in the case of current boundaries) dividing the applied current (usually unit current) by the voltage difference between the source and sink, or
- (in the case of voltage boundaries) identifying a minimal set of edges which cuts the system, summing the currents on these edges *in the direction they cross the cut*, and dividing it by the applied voltage (usually unit voltage).

For the chain of fuse networks defined above, the conductivity should be a monotonically decreasing function of time, or fuses broken, because removing a resistor from a network of resistors cannot increase its conductivity. One network in the chain will be known as the *critical fuse network*. If e is the edge broken in each network and I is the current flowing across the network, then the critical network is that for which $It(e)/|i(e)|$ is maximized. In the case of current boundaries $I = 1$, and in the case of voltage boundaries $I = \Lambda$. The quantity $It(e)/|i(e)|$ can be thought of as the current on the network *in the units of the fuse threshold*—put another way, if we were to actually simulate slowly ramping up the boundary conditions until the literal result of our linear problem equals the fuse value, this would be the value of the applied current. This has other significance: if the value of $It(e)/|i(e)|$ for a network in the chain is greater than that of n networks following it, then the breaking of the fuse in the first network can be thought of as having caused the breaking of those in the next n networks—an avalanche. Avalanches occur in

all matter of critical phenomena, and specifically in fracture and other forms of yielding when the failure of one region triggers the successive and recursive failure of others. Another way of viewing the critical network, then, is the network prior to the final avalanche that breaks the system.

The procedure for breaking bonds described above does not care whether the system is broken with current or voltage boundaries: the choice of the next fuse is a function of only local properties. The procedure for determining *avalanches* does, however, care about the boundary conditions. Avalanches are determined by finding the current or voltage that would have had to been applied to the system for that bond to break, then compute when that necessary external current or voltage backtracks.

The numeric problem described above gives the numeric current i_{num} for each fuse in the network. The next fuse that will break is that whose threshold would be first reached by uniformly scaling all numeric currents from zero. The current i on that fuse when it breaks is its threshold value, and the ratios of real to numeric currents

$$\frac{i}{i_{\text{num}}} = \frac{I}{I_{\text{num}}} = \frac{U}{U_{\text{num}}} \tag{1.17}$$

all are equal. The numeric voltage is always $U_{\text{num}} = 1$, defined by the boundary conditions detailed above. The numeric current $I_{\text{num}} = \Lambda U_{\text{num}} = \Lambda$ (and therefore the modulus Λ) can be measured by summing the currents flowing over bonds whose duals form a closed path around the torus. It follows that $U = i/i_{\text{num}}$ and $I = \Lambda U = \Lambda \times (i/i_{\text{num}})$; the two differ by a factor of the modulus Λ.

Since fuses are only broken, Λ is a monotonically decreasing function of time. This has a direct consequence on the avalanches: those defined by fracture at external voltage are fully contained by those defined by fracture at external current. If I_n is the external current after the nth fuse has burned and it ends a current avalanche, $I_m < I_n$ for all $m < n$. It therefore follows that, since monotonicity implies that $\Lambda_m \geq \Lambda_n$ for any $m < n$,

$$U_m = \frac{I_m}{\Lambda_m} < \frac{I_n}{\Lambda_m} \leq \frac{I_n}{\Lambda_n} = U_n \tag{1.18}$$

whence $U_m < U_n$ for all $m < n$ and U_n must terminate a voltage avalanche. The opposite is not true, and in fact current avalanches may and do often contain multiple voltage avalanches. This is an example of a no-passing principle.[34–36] An example where a large current-based avalanche contains several voltage-based avalanches can be seen in Fig. 1.8.

The difference between these two types of avalanches is more than just academic. Numeric measurements indicate that the two types of avalanches appear to scale differently, at least up to the system sizes that are accessible to us right now. There are compelling reasons for this. If we think of the fracture event as an abrupt phase transition, as it has been thought of historically,[37] then trading stress for strain can be thought of as trading control parameters from, e.g., one like external field to one like magnetization. In a constant-magnetization Ising model, the transition is washed out with a coexistence region, and study of the abrupt transition itself is more conveniently done at constant field. In the fracture problem, the stress (current) is the canonical control parameter,[37,38] and we expect that it is more natural for describing the

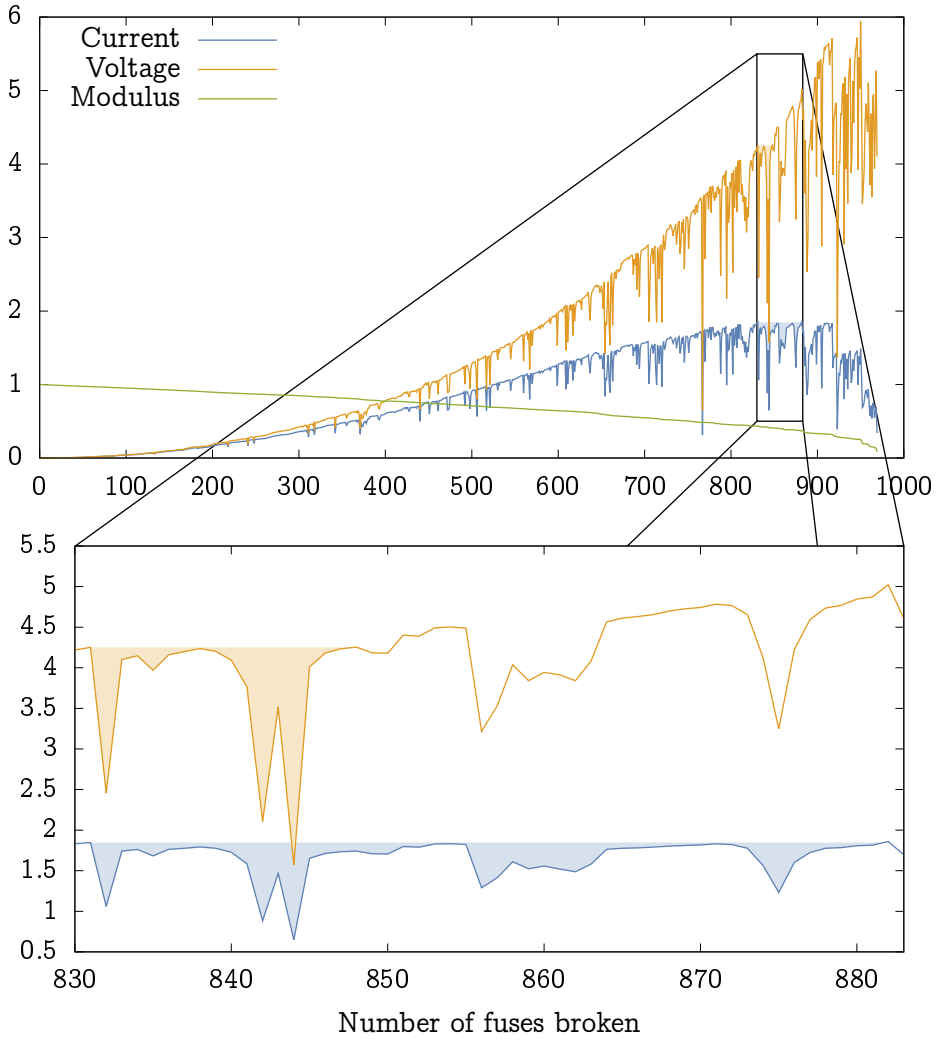

Figure 1.8: Voltage, current, and modulus versus fuse number in the fracture of a 64×64 square lattice with $\beta = 0.5$. An avalanche in both current and voltage starting at fuse 831 is highlighted. The height of the filled regions show the current or voltage that began that avalanche and that a bond must break at or above in order to end it. (Top) Data over the entire fracture. (Bottom) Data only over the course of the avalanches starting at fuse 831.

transition. An intuition around this can be developed: in the thermodynamic limit, even a perfectly brittle system held at nonzero constant *strain* will not propagate a crack forever, since eventually the lowering modulus and widening crack will separate the two sides enough to meet the imposed strain—partial cracks can be held at constant strain. However, this is not true for constant *stress*, since as the modulus lowers and the sides of the system separate the force applied does not subside. We therefore expect that studies at constant strain will be dogged by an analogy to coexistence phenomena, where the 'critical strain' happens at a place with the growing crack already present.

§3. Percolation and the $\beta \to 0$ limit

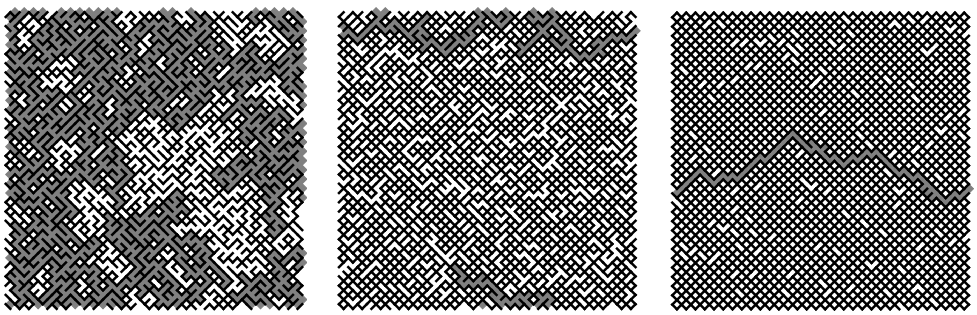

Figure 1.9: Broken networks for three choices of β (from left to right): 0.03, 0.5, and 3. The fracture surface is highlighted in each.

We're going to talk a lot about percolation, and here's why. Small β corresponds to high disorder, where material strengths vary across orders of magnitude, while large β corresponds to very little disorder. In the limit of infinite disorder, or zero β, the fuse network problem reduces to percolation.[39] This can be seen by the following argument: for any β, the expected value of

the nth strongest fuse is

$$\begin{aligned}\langle x_n \,|\, N,\beta\rangle &= \int_0^1 x_n P(x_n \mid N,\beta)\,dx_n \\ &= \int_0^1 x_n \Big(\int_0^{x_n} \beta x_<^{\beta-1}\,dx_<\Big)^{n-1} \beta x_n^{\beta-1}\Big(\int_{x_n}^1 \beta x_>^{\beta-1}\,dx_>\Big)^{N-n} dx_n \\ &= \frac{\Gamma(N+1)\Gamma(n+\frac{1}{\beta})}{\Gamma(N+1+\frac{1}{\beta})\Gamma(n)} = (\tfrac{n}{N})^{1/\beta} + \mathcal{O}(N^{-1})\end{aligned} \tag{1.19}$$

For sufficiently small β, the ratio $((n+1)/n)^{1/\beta}$ of any two subsequent microscopic strengths dwarfs the ratios of stress in those regions, and fuses break simply in the order of their strengths. Since strengths are independent, this corresponds to breaking regions in random order, with the system fractured when a contiguous broken region spans the system.* This is simply a percolation problem.

A review of percolation physics[40]

We will therefore build our scaling theory around the notion that $\beta = 0$, $L = \infty$ is a critical point in the percolation universality class. It's worth taking a moment to review percolation physics, as we will rely on it heavily. In a (bond) percolation process, a lattice has its connections depleted at random, where the probability that a given connection is present is p. When p is reduced through a critical bond density p_c, the lattice transitions from connected to disconnected via a continuous phase transition.

*One could also study the case where the system is considered fractured when its rigidity goes to zero. Our scaling analysis would be unchanged, with the percolation exponents substituted for those of rigidity percolation, though our numerics would need to be changed entirely.

At the transition, the regions that remain connected by bonds have sizes that span all length scales, and in particular are distributed with a power law $n_s \sim s^{-\tau}$. One infinite cluster remains but is so internally depleted as to be fractal, with fractal dimension d_f. The probability that two sites separated by a distance r belong to the same cluster is power-law distributed as well, with $g(r) \sim r^{-2+(d-\eta)}$ with anomalous dimension η related to the fractal dimension by $\eta = 2 + d - 2d_f$. The portion of the infinite cluster that would support a current (excluding lollipops and bow ties) grows with a different fractal dimension d_b, and the length of the shortest path between two sites in the same cluster grows as a power $d_{\min}$ of the Euclidean distance between them.

As the transition is approached from above or below, the distribution of non-infinite clusters inherits a cutoff $s_{\max} \sim \Delta p^{-1/\sigma}$, so that $n_s \sim s^{-\tau} f(s\Delta p^{1/\sigma})$ for some function f. The correlation length of growing clusters (or shrinking voids) grows as $\xi \sim \Delta p^{-\nu}$. Since the cutoff in the cluster size distribution can be thought of as the size of the growing infinite cluster and its size is given by its extent to the power of its dimension, $s_{\max} \sim \xi^{d_f}$ implying $d_f \nu = 1/\sigma$.[41]

We will be working with two-dimensional simulations, and so we will be making frequent use of the two-dimensional values of these exponents. Most referenced here are known from exact solutions, and listed in Table 1.1. Also present is the first singular correction to scaling ω, which provides corrections of the form $L^{-\omega}$. Checking that the exponent relations given above work in two dimensions is left as an exercise of the reader.

The renormalization group analysis of percolation is a textbook example

Exponent	Description	2D value
ν	Correlation length divergence	$4/3$
τ	Cluster size distribution tail	$187/91$
σ	Cluster size distribution cutoff	$36/91$
d_f	Fractal dimension of infinite cluster	$91/48$
η	Anomalous dimension in correlation function	$5/24$
ω	First singular correction to scaling	$3/2$ [42]
d_b	Fractal dimension of conducting backbone	1.64336 [43]
$d_{\min}$	Fractal dimension of shortest paths	1.13077 [44]

Table 1.1: Percolation exponents and their two-dimensional values.

of its use: a nice, ordinary, hyperbolic fixed point.* The distance Δp from critical depletion is proportional to the scaling field u_p with largest scaling dimension y_p, and if we take u_s to be the scaling field with the next largest scaling dimension y_s we have

$$\begin{aligned} \frac{du_p}{d\ell} &= y_p u_p \\ \frac{du_s}{d\ell} &= -y_s u_s \\ \frac{dL}{d\ell} &= -L \end{aligned} \tag{1.20}$$

where L is the system size. The correlation length is a function of these scaling fields that leaves the combination $L^{-1}\xi$ invariant under the flow, which gives

$$\begin{aligned} 0 = \frac{d(L^{-1}\xi)}{d\ell} &= \frac{1}{L}\frac{d\xi}{d\ell} - \frac{1}{L^2}\frac{dL}{d\ell}\xi \\ &= \frac{1}{L}\left(\frac{du_p}{d\ell}\frac{d\xi}{du_p} + \frac{du_s}{d\ell}\frac{d\xi}{du_s} + \frac{dL}{d\ell}\frac{d\xi}{dL}\right) + \frac{1}{L}\xi \\ &= \frac{1}{L}\left(y_p u_p \frac{d\xi}{du_p} - y_s u_s \frac{d\xi}{du_s} - L\frac{d\xi}{dL}\right) + \frac{1}{L}\xi \end{aligned} \tag{1.21}$$

*What does hyperbolic mean? Check out Chapter 5!

which in turn can be solved to yield

$$\xi = u_p^{-1/y_p} f(u_p L^{y_p}, u_s L^{-y_s}) \tag{1.22}$$

for some function f. Near the critical point, the scaling fields can be expanded with $u_p = \Delta p + O(\Delta p^2, u_s^2, \Delta p u_s)$, and we have

$$\xi \simeq \Delta p^{-1/y_p} f(\Delta p L^{y_p}, u_s L^{-y_s}) \tag{1.23}$$

Given this form of the correlation length, we can immediately infer that $\nu = 1/y_p$, and that $\omega = y_u$. Functional dependence like this is how finite size effects and corrections to scaling principally affect critical behavior.

§4. Modelling fracture: depletion with correlated avalanches

How do we work our fracture problem into this picture? At $\beta = 0$ we should see perfect percolation, with fuses breaking in a (mostly) uncorrelated way (mostly) in order of strength until the system breaks, exhibiting all the nice scaling properties discussed above at fracture.* With nonzero β, several new phenomena begin to emerge:

*The "mostly"s here are in reference to the fact that—because of the lollipops and bow ties discussed in the previous section—bonds do not actually get broken in the order of their thresholds, and in fact whole connected regions of the system are removed from the damage process at intermediate p because they belong to one of these features. It has been argued[39] that these differences do not affect the universality of percolation on the basis that any two points that would belong to the same cluster if not separated by a lollipop stem in fact *do* belong to the same cluster, by the definition of a lollipop detailed above. This argument preserves the scaling of the backbone, but is not bulletproof in regard to the cluster size distribution, since it neglects the effect of myriad bonds within lollipops that would have been broken but weren't. We find, as has been found previously, numeric evidence that certain quantities, like the fractal dimensions of the infinite cluster and the minimum path, are different in the fuse problem.[45]

- The depletion events will become increasingly correlated.
- The network will be depleted to a point increasingly far from p_c at the point it is broken.
- The network becomes broken (its conductivity drops to zero) in an increasingly discontinuous way.
- In the infinite system limit, the first depletion event immediately ruptures the system.

Proposed scaling theories of these combined phenomena can be quite complex.[46]

Sufficiently long range correlations in percolation are known to be relevant in the renormalization group sense,[47–49] and so it is plausible that these correlations in the depletion process will significantly change the critical behavior. In particular, a Harris-style criterion concludes that correlations that fall off like $r^{-\alpha}$ are relevant if $\alpha\nu - 2 < 0$, or $\alpha < 2/\nu = 3/2$ in two dimensions. Given that the effective force field of broken fuses goes like $r^{-1/2}$, it is plausible we would see correlations of sufficiently long range! It is also known that the correlation length exponent for the resulting long-range percolation fixed point is given by $\tilde{\nu} = 2/a$.

Another important feature of the fracture problem is the lack of a sensible thermodynamic limit. This is because, as the size of a sample is increased, the weakest flaw present within it becomes weaker. An infinite system where *any* avalanches are permitted will inevitably have a rare, extremely weak region so large as to trigger a spanning avalanche.[10] Therefore, any average of the properties of a system at critical stress, or averaged over the lead up to critical stress, must have finite size integrated into the description in a central way,

since the critical stress is zero in the thermodynamic limit.

In addition to the primary scaling field u_p, we imagine a second relevant scaling field u_d, with associated scaling dimension y_d. The surface along which the rupture happens is thought to define a first order transition.[37] If this is the case, there must exist a fixed point associated with the abrupt transition with a relevant direction that acts as a separatrix for the two 'phases,' broken and unbroken.[50–52] A point inside the basin of attraction of this fixed point cannot be brought outside of it, since that would violate the meaning of a basin of attraction, and therefore the basin of attraction must be an invariant surface of the RG. Since the basin of attraction for the first-order fixed point is *also* the transition surface, this means that the values of u_{pf} and u_{df} at critical stress, or the fracture transition, obey a relationship like

$$0 = F(u_{df} u_{pf}^{-y_d/y_p}, u_{df} L^{y_d}) \tag{1.24}$$

for some function F. This definition of the critical surface is naturally invariant under the RG, because both of the arguments to F are invariants of the RG. Assuming the surface is single-valued in the correct arguments, we can equivalently write the relationship as

$$u_{df} u_{pf}^{-y_d/y_p} = F(u_{df} L^{y_d}) \tag{1.25}$$

In the infinite-size limit, the system does not become broken if u_d is zero, but instantly breaks for infinitesimal u_d. We therefore conclude that the transition

surface must run along the line $u_d = 0$ at $L^{-1} = 0$, and

$$\lim_{x \to \infty} F(x) = 0 \tag{1.26}$$

For nonzero L^{-1}, F is likewise nonzero and the transition surface pulls away from $u_d = 0$. How it does this is the largest outstanding question in this theory, and I will try to address it with our existing data as best as possible.

How should we expect to relate our physical parameter, β, to these abstract scaling fields? Since for $\beta = 0$ we have no fracture transition and simply see ordinary percolation, we expect that

$$u_d = \beta^\alpha + \cdots \tag{1.27}$$

where α is some positive constant. While at the moment our data is only at the fracture transition, we do not vary Δp or its analogue ΔI from their transition values $\Delta p_f = p_f - p_c$ or $\Delta I_f = I_f - I_c = I_f$, but imagine that

$$u_p = \Delta p + \cdots \tag{1.28}$$

Model A: Power law approach

For the purposes of a scaling theory, the simplest answer is to slap a power law on it. This would give $F(x) = Ax^{-z}$ for same constant z. The value of z is constrained by the limiting consideration discussed above, and since

$$u_{df} \propto u_{pf}^{y_d/(1+z)y_p} L^{-y_p z/(z+1)} \tag{1.29}$$

we have simply $z > 0$. This relates the finite-size invariants $u_p L^{y_p}$ and $u_d L^{y_d}$ along the critical surface by

$$(u_{df} L^{y_d})^{1+z} = A(u_{pf} L^{y_p})^{y_d/y_p} \tag{1.30}$$

For small β, this gives

$$\Delta p_f \simeq u_{pf} \propto u_{df}^{(z+1)y_p/y_d} L^{z y_p} \simeq (\beta L^{z y_d/\alpha(z+1)})^{\alpha(z+1)y_p/y_d} \tag{1.31}$$

Model B: Logarithmic approach

In the science of fracture, it is known that the critical stress of a system often scales like an inverse logarithm of its size.[10] That therefore suggests to us that, at fixed u_d (which we will heuristically think of as fixing the properties, e.g., heterogeneity, of the system) the transition point u_{pf} should approach infinity as some logarithm of L^{-1}. Why infinity and not zero? Recall that $u_p \sim \Delta p$, the distance from fully percolated, and the larger u_p is the *fewer* bonds we have broken. Since we expect infinite systems to break upon the first depletion event, the critical u_p should diverge with system size.

A simple form of F that is consistent with this limit is given by

$$F(x) = \frac{A}{\log(B + x^z)} \tag{1.32}$$

This gives, at finite size,

$$u_{pf} = \tilde{A} u_{df}^{y_p/y_d} \log(B + (u_{df} L^{y_d})^z)^{y_p/y_d} \tag{1.33}$$

which has a logarithmic singularity as $L \to \infty$, and where $\tilde{A} = A^{-y_p/y_d}$. We must check that this ansatz for the surface preserves the correct behavior in u_{df} as $L \to \infty$. Unfortunately this form is not analytically invertible, so we cannot write u_{df} as an explicit function of u_{pf} and L. We will make an assumption about the dependence of u_{df} on L for fixed u_{pf} then check if the solution self-consistently satisfies the assumption. Namely, we will suppose that L^{y_d} grows much faster than u_{df} shrinks in the limit of large L for any fixed u_{pf}. This would imply that $B + u_{df}L^{y_d} \simeq u_{df}L^{y_d}$, and therefore that

$$u_{df} \simeq \frac{\tilde{A}u_{pf}^{y_d/y_p}}{zW\left[\tilde{A}(u_{pf}L^{y_p})^{y_d/y_p}/z\right]} \tag{1.34}$$

where W is the principal branch of the Lambert W function. At large argument, $W(x) = \log x - \log\log x + \mathcal{O}(1)$, and

$$u_{df} \simeq \frac{\tilde{A}u_{pf}^{y_d/y_p}}{z\log\left[\tilde{A}(u_{pf}L^{y_p})^{y_d/y_p}/z\right]} \tag{1.35}$$

which both vanishes as $L \to \infty$ for any fixed u_{pf} and does so much faster than any positive power of L grows.

What does this form for the critical surface imply for the relationship between the invariants $u_{pf}L^{y_p}$ and $u_{df}L^{y_d}$ along it? Equation (1.33) implies

$$u_{pf}L^{y_p} = \tilde{A}(u_{df}L^{y_d})^{y_p/y_d}\log(B + (u_{df}L^{y_d})^z)^{y_p/y_d} \tag{1.36}$$

For small $u_{df}L^{y_d}$, we find

$$u_{pf}L^{y_p} \simeq \tilde{A}\log(B)(u_{df}L^{y_d})^{y_p/y_d} \tag{1.37}$$

which implies $u_{pf} u_{df}^{-y_p/y_d} = \tilde{A} \log(B)$, e.g., a line of constant flow for the percolation fixed point in the limit $L \to \infty$. For large $u_{df} L^{y_d}$, we find

$$u_{pf} L^{y_p} \simeq \tilde{A} z (u_{df} L^{y_d})^{y_p/y_d} \log(u_{df} L^{y_d})^{y_p/y_d} \tag{1.38}$$

which implies $u_{pf} u_{df}^{-y_p/y_d} = \tilde{A} z \log(u_{df} L^{y_d})^{y_p/y_d}$, or that the invariant combination grows logarithmically with fixed L along the fracture transition surface with u_{df}. In the limit of small β, we should find

$$\Delta p_f \simeq u_{pf} \propto u_{df}^{y_p/y_d} \simeq \beta^{\alpha y_p/y_d} \tag{1.39}$$

Notice that this predicts no rescaling of β with L.

Model C: Power law approach with crossover

At this point we should pause and allow a little more complexity into our thinking, lest we construct a power-law-scaling cargo cult. Besides the constraint that it diverge sufficiently quickly with infinite L, very little practically constrains the function F, and *any* function will result in a critical surface that is invariant under the renormalization group flow, since it is built from pieces that are each invariant under the renormalization group flow. Though it's reasonable to assume that things will approach their limits with some sort of power law, the middle is fair play.

The simplest surface function that has power law tails but does something different in the middle might be something like

$$F(z) = A x^{-z} (B + x^m)^{-(y-z)/m} \tag{1.40}$$

which goes as x^{-z} for small x and x^{-y} for large, with a crossover whose width is tuned by m and whose position is tuned by B. The same considerations on z discussed in Model A apply here for y, which sets the large-argument behavior. This relates the finite-size invariants $u_p L^{y_p}$ and $u_d L^{y_d}$ along the critical surface by

$$u_{pf} L^{y_p} = \tilde{A} (u_d L^{y_d})^{(1+z)y_p/y_d} (B + (u_d L^{y_d})^m)^{(y-z)y_p/y_d m} \tag{1.41}$$

Since we suspect the critical surface meets the origin, we must have $z > -1$. For small β, this gives

$$\Delta p_f \simeq u_{pf} \propto u_{df}^{(1+z)y_p/y_d} L^{z y_p} \simeq (\beta L^{z y_d/\alpha(1+z)})^{\alpha(1+z)y_p/y_d} \tag{1.42}$$

which is exactly the same as in Model B.

§5. Numeric comparison with the critical damage

To investigate the plausibility of this picture, we will start by looking at data giving p_f as a function of β from fuse network simulations. Inspired by the considerations above, we fit

$$p_f = \frac{p_c}{1 + a x^\gamma + b x^{2\gamma} + c x^{3\gamma}} \tag{1.43}$$

where $x = s\beta L^q$ and where p_c and s are taken to be lattice dependent (with $s_{\text{sq}} = 1$). The form of the function means that

$$\Delta p_f = p_c - p_f = a p_c x^\gamma + \mathcal{O}(x^{2\gamma}) \tag{1.44}$$

p_c^{sq}	p_c^{vo}	a	b	c	γ	q	s_{sq}	s_{vo}
0.437	0.604	1.18	0.103	0.00709	1.19	0.144646	1	1.21832

Table 1.2: The parameters and their values for the fit to the damage at critical stress.

which allows for the departure of Δp_f from the critical point to be a nonlinear power law of β. Unfortunately, the p_cs from ordinary percolation on e.g., voronoi lattices,[53] are not relevant here, because we leave so many bonds in lollipops and bow ties as to dramatically modify the proportion broken at the critical point.

The resulting fit is shown in Figure 1.10, in which data from both square and voronoi lattices at many system sizes is superposed. The fit parameters are shown in 1.2. The result fits the data well until the edge of the expected scaling region. The difference between the fit and p_c does poorly in the limit of small β, where properly taking into account the finite size scaling of the effective p_c would be necessary to improve the agreement between data in the collapse. Our value of p_c^{sq} is consistent with that measured in other studies.[39]

Notice that γ is very nearly one, and q is quite small, which could suggest the presence of a logarithm, though the logarithmic Model B does *not* predict logarithmic corrections to Δp_f as a function of β for small β, only for large β. Therefore we will take it seriously as an exponent for now. In the framework of Model A or Model C, these exponents are $q = zy_d/\alpha(z+1)$ and $\gamma = \alpha(z+1)y_p/y_d$. Without a serious consideration of the uncertainty analysis, this immediately suggests $z \simeq 0.23$ and $y_d/\alpha \simeq 0.78$.

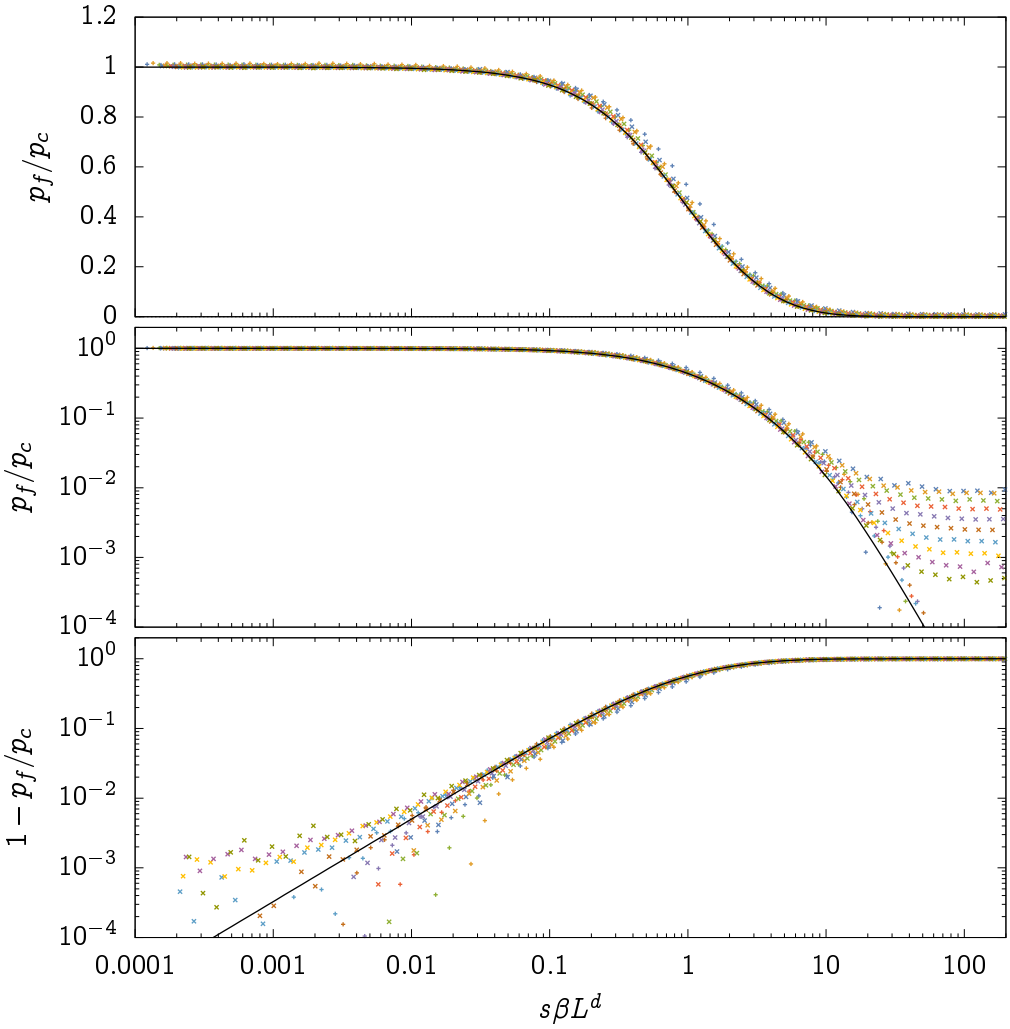

Figure 1.10: (Top) The probability p_f that a fuse is unburned at failure as a function of the fit variable $x = s\beta L^d$ for both square and voronoi data, along with a fit in black. (Middle) Same, but on a logarithmic scale. (Bottom) The deviation $\Delta p_f = p_f - p_c$.

Figure 1.11: A 64×64 square lattice with $\beta = 0.316$, held at critical stress. Connected clusters on the dual lattice are highlighted with individual colors.

§6. Dual lattice clusters

Besides the asymptotic value of Δp along the failure surface, the other properties that can be measured in this numeric system and then directly compared to percolation are properties of the *clusters*. In our case, any interesting analysis will have to be made of the clusters of sites on the dual lattice connected by dual bonds corresponding to burnt fuses. Trying to do the opposite, and analyze the cluster size distribution of connected components on the principal network, is not very interesting because, as we discussed in a previous section, we go to great pains to makes sure that there is only ever *one* connected component on the principal network! Therefore that size distribution at fracture, say, would be exactly the analysis of the damage done above. The amount of damage done at percolation does not have interesting scaling properties beyond its mean.

Cluster size

Recall that the distribution of clusters in percolation has a power law form $s^{-\tau}$ that is cut off at some $s_{\max} \sim u_p^{-1/\sigma}$. Let us take this scaling seriously in our extended analysis, and therefore consider

$$n_s = s^{-\tau} \mathcal{N}(s u_p^{1/\sigma}, u_p L^{y_p}, u_d L^{y_d}) \tag{1.45}$$

Fitting the entire distribution is challenging, and it helps to reduce our mental (and computational) load by instead fitting *moments* of that distribution. These will have scaling forms given by

$$\begin{aligned}
\langle s^n \rangle &= \int_1^\infty s^n n_s \, ds \\
&= \int_1^\infty s^{n-\tau} \mathcal{N}(s u_p^{1/\sigma}, u_p L^{y_p}, u_d L^{y_d}) \, ds \\
&= \int_{u_p^{1/\sigma}}^\infty u_p^{-1/\sigma} (y u_p^{-1/\sigma})^{n-\tau} \mathcal{N}(y, u_p L^{y_p}, u_d L^{y_d}) u_p^{-1/\sigma} \, dy \\
&= u_p^{-(n+1-\tau)/\sigma} \left[\mathcal{N}_n(u_p L^{y_p}, u_d L^{y_d}) + \mathcal{O}(u_p^{(n+1-\tau)/\sigma}) \right] \\
&= L^{d_f(n+1-\tau)} \left[\tilde{\mathcal{N}}_n(u_p L^{y_p}, u_d L^{y_d}) + \mathcal{O}(L^{-d_f(n+1-\tau)}) \right]
\end{aligned} \tag{1.46}$$

where we have used the exponent relation $d_f = 1/\sigma\nu = y_p/\sigma$. The first term is leading for small u_p, i.e., in the vicinity of the critical point, if $n+1-\tau > 1$. Since we expect from percolation that $\tau = 187/91$, this means that we can expect the moments to be well-described by the truncation above for $n = 2$ and up.

We can take more from our knowledge of the percolation problem than simply the exponents—there's a reason we call them *universal* scaling func-

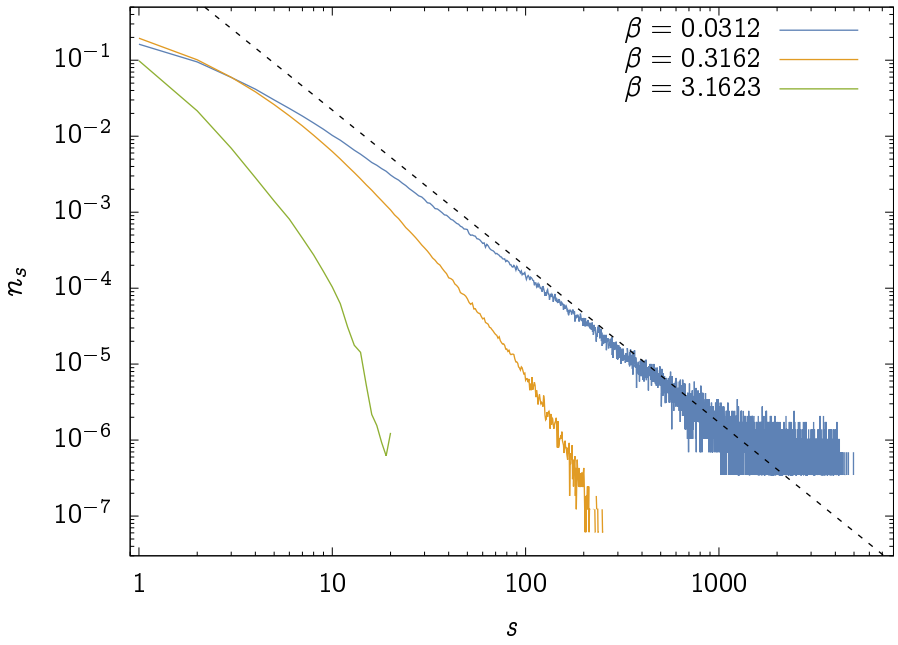

Figure 1.12: Example cluster size distributions for a 128×128 square lattice fuse network for several disorder parameters β. The dashed line shows $s^{-\tau}$ for $\tau = 187/91$, from 2D percolation.

tions, after all! In particular,

$$\mathcal{P}_n(u_p L^{y_p}) = \tilde{\mathcal{N}}_n(u_p L^{y_p}, 0) \tag{1.47}$$

should match (up to constant scale factors) the scaling functions found for the moments of the cluster size distributions in ordinary bond percolation. Therefore, if we can fit the values of $\mathcal{P}_n$ we can write

$$\tilde{\mathcal{N}}_n(x, y) = \mathcal{P}_n(x) + y\tilde{\mathcal{N}}^{(1)}(x) + \mathcal{O}(y^2) \tag{1.48}$$

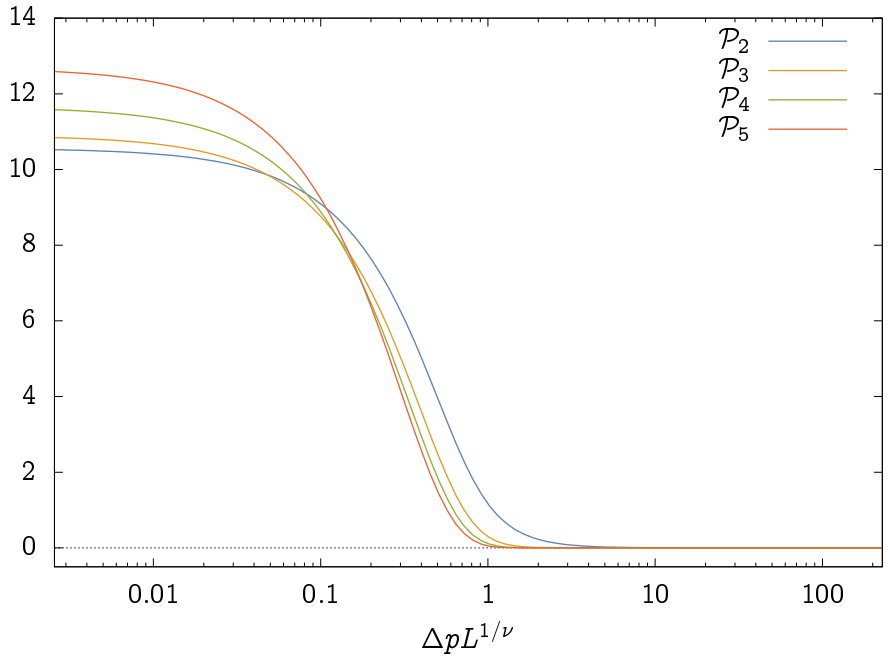

Figure 1.13: Percolation scaling functions measured by use of direct simulation on a 4096×4096 toroidal square lattice.

which significantly constrains our fits to fracture data. Percolation cluster statistics have been observed in fracture problems before.[17,21,54]

We have simulated standard bond percolation on $L \times L$ toroidal square lattices for L up to 4096 and measured the cluster size distribution as a function of Δp using the Newman–Ziff algorithm.[55] These were then fit to the largest system size—in the region where the previous several system sizes collapsed well together—to functions of the form

$$A\left[1 + \left(\sum_{i=0}^{N} b_i x^{\nu d_f (n+1-\tau)-i} + \sum_{i=1}^{N} c_i x^i\right)^m\right]^{-1/m} \tag{1.49}$$

for sufficient N that the agreement was close. The results are shown in 1.13.

As a first-order attempt to compare the data from fracture to the scaling functions of our theory, we will perform a fit focusing on small β, so that the next terms in the series for the scaling function are relatively small. A fit in this regime then takes the form

$$\langle s^n \rangle = a_n L^{-d_f(n+1-\tau)} \mathcal{P}_n(b_n x^\zeta), \tag{1.50}$$

where $x = s\beta L^r$ and s and the a_n depend on the lattice. We again take $s_{\text{sq}} = 1$. The resulting fits are shown in Figure 1.14 for both square and voronoi lattices, along with the percolation scaling function we fit earlier. This fit fixes the percolation exponents τ and d_f. We measure $r \simeq 0.87$ and $\zeta \simeq 0.96$. Slightly better agreement in the small β limit is obtained if d_f is allowed to be fit as well, which might imply that the internal structure inside lollipops and bow ties that is held in place throughout the fracture process rather than continuing to percolate is relevant. When it is allowed to vary, it takes the value $d_f \simeq 1.869$, which is to be compared with the percolation value $d_f = 91/48 \simeq 1.896$. Our fit value for the infinite cluster fractal dimension is consistent with that found by other researchers on a similar model.[45] The same fits with d_f allowed to vary are seen in Figure 1.15—these result in $r \simeq 0.85$ and $\zeta \simeq 0.96$. The parameters from this latter fit, which we believe is more trustworthy than the first,[45] can be found in Table 1.3.

Even this basic fit, without matching more than the percolation scaling function, allows us to connect back with our scaling theory. It predicts that

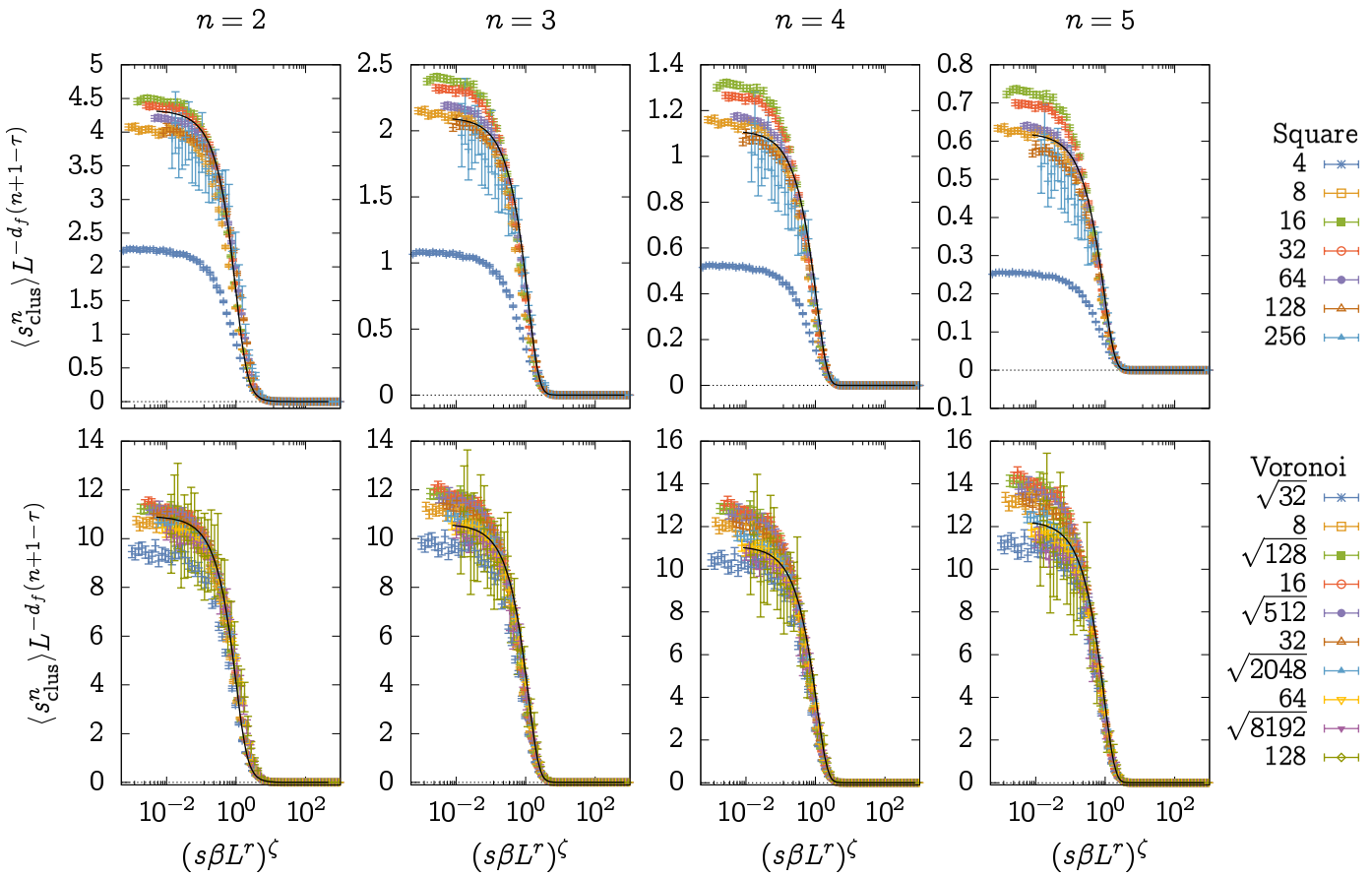

Figure 1.14: Cluster size distributions measured in fuse networks at critical stress in (Top row) square and (Bottom row) voronoi lattices, along with (black line) fits to associated scaling function from percolation. In these fits, d_f is fixed to its percolation value.

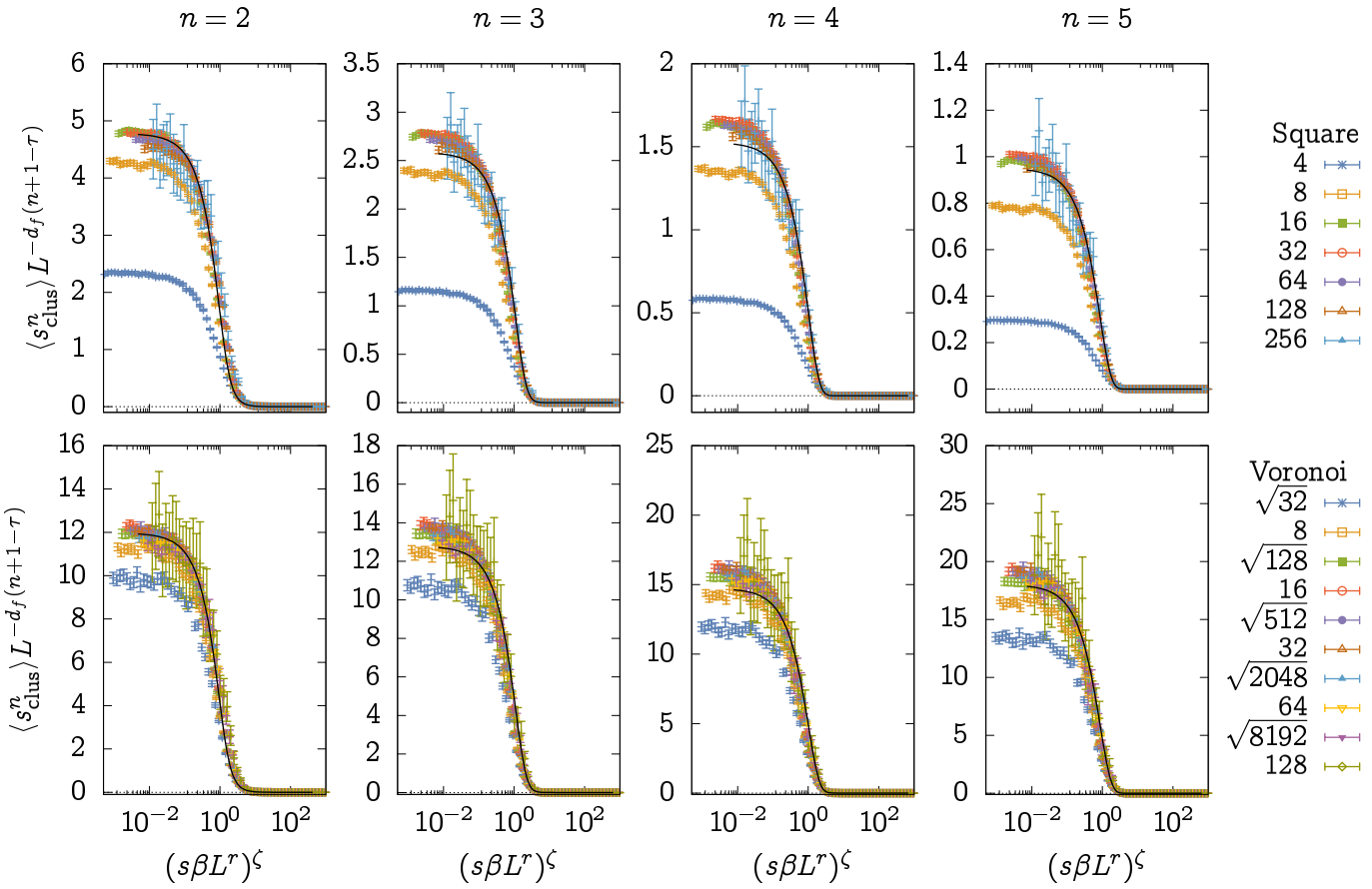

Figure 1.15: Cluster size distributions measured in fuse networks at critical stress in (Top row) square and (Bottom row) voronoi lattices, along with (black line) fits to associated scaling function from percolation. In these fits, d_f is allowed to vary from its percolation value.

ζ	r	d_f	b_2	b_3	b_4	b_5
0.96	0.85	1.87	0.52	0.35	0.33	0.35

s_{sq}	a_2^{sq}	a_3^{sq}	a_4^{sq}	a_5^{sq}
1	0.45	0.24	0.13	0.075

s_{vo}	a_2^{vo}	a_3^{vo}	a_4^{vo}	a_5^{vo}
1.65	1.1	1.2	1.3	1.4

Table 1.3: Parameters resulting from fitting (1.50) to fuse network data for $L > 16$ and $\beta < 0.5$. No corrections to scaling are present, and including them would likely lead to a better fit.

a function of the invariant $u_p L^{y_p}$ should depend on β like

$$u_{pf} L^{y_p} \propto (u_{df} L^{y_d})^{(1+z)y_p/y_d} \simeq (\beta L^{y_d/\alpha})^{\alpha(1+z)y_p/y_d} \tag{1.51}$$

which means that the parameters in our fit correspond to with the exponents in our theory like $r = y_d/\alpha$ and $\zeta = \alpha(1+z)y_p/y_d$. This fit gives $y_d/\alpha \simeq 0.85$ and $z \simeq 0.1$. The former prediction agrees with our fit from the previous section, while the latter is about half its previous value. The fact that z is so small once again suggests that a power law may not be the right fit in this limit, but a different scaling theory than Model B would have to incorporate it.

Here we have only fit the small-β dependence of the functions, explicitly ignoring the far tails. Taking a look at those tails by plotting on a logarithmic scale as in Figure 1.16 reveals some of the intermediate and large β structure. At very large β, the square and voronoi lattice models cross over into different, nonuniversal behavior, with the square lattice becoming unstable to

every avalanche, and with the baked-in disorder of the voronoi cutting off the decreasing disorder in β. This behavior has nothing to do with critical phenomena, and we don't expect to be able to describe it with a scaling theory. There does appear to be an intermediate regime, however, which looks similar across both lattices, in which the fracture data pulls away from the percolation scaling function with a slightly shallower slope and in a way that depends on the system size. Intermediate disorder is known to have clusters that are not described by percolation physics.[38,56] This behavior is not explained by Model A alone, but may be accommodated by Model C, or by the addition of singular corrections to scaling in Model A.

Cluster geometry

Besides having sizes, the clusters that form on the network has a *spatial extent*, and we can measure their two-point correlation function as well. More precisely, we measure $g(r)$, which is the probability that, starting on a site at the origin, a site displaced by r belongs to the same cluster. Recall that in percolation, $g(r) \sim r^{d-2-\eta}$ and is cut off by the correlation length $\xi \sim u_p^{-\nu}$. Taking these behaviors to heart again, we look for a scaling form

$$g(x_\parallel, x_\perp) = x_\parallel^{-\eta} \mathcal{G}(x_\parallel u_p^\nu, x_\parallel x_\perp^{-1}, u_p L^{u_p}, u_d L^{u_d}) \tag{1.52}$$

Here and throughout, the parallel direction refers to the direction parallel to the applied stress, and the perpendicular direction refers to the direction of crack propagation. This predicts that its moments will likely have scaling

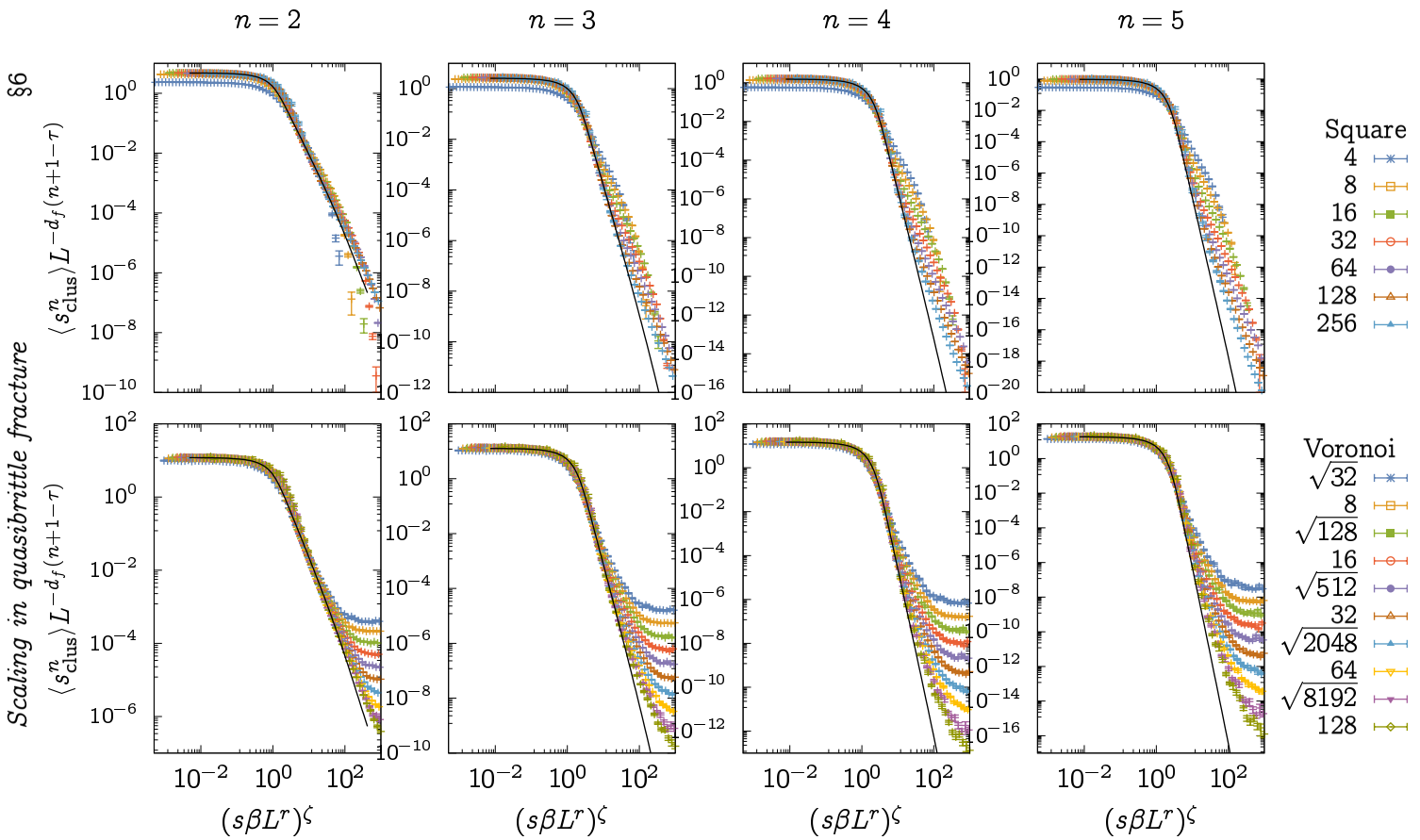

Figure 1.16: Cluster size distributions measured in fuse networks at critical stress in (Top row) square and (Bottom row) voronoi lattices, along with (black line) fits to associated scaling function from percolation, now plotted on a logarithmic scale.

forms, now of the form

$$\begin{aligned}\langle x_{\parallel}^{n}\rangle &= \int x_{\parallel}^{n-\eta}\mathcal{G}(x_{\parallel}u_p^{\nu}, x_{\perp}x_{\parallel}^{-1}, u_pL^{u_p}, u_dL^{u_d})\,dx_{\parallel} \\ &= \int u_p^{-\nu(n-\eta)}\mathcal{G}(v, vu_p^{-\nu}x_{\perp}^{-1}, u_pL^{u_p}, u_dL^{u_d})u_p^{-\nu}\,dv \\ &= u_p^{-\nu(n+1-\eta)}\mathcal{G}_n(x_{\perp}u_p^{\nu}, u_pL^{u_p}, u_dL^{u_d}) \\ &= L^{n+1-\eta}\tilde{\mathcal{G}}_n(x_{\perp}u_p^{\nu}, u_pL^{u_p}, u_dL^{u_d})\end{aligned} \tag{1.53}$$

with a similar expression for the moments in $x_{\perp}$. In particular, using these moments we can estimate the correlation length directly, using

$$\xi = \frac{\langle x^2\rangle}{\langle s^2\rangle} = L^{2+1-\eta-df(2+1-\tau)}\frac{\tilde{\mathcal{G}}_2}{\tilde{\mathcal{N}}_2} = LC(u_pL^{y_p}, u_dL^{y_d}) \tag{1.54}$$

where we used the hyperscaling relationships $\eta = 2+d-2d_f$ and $\tau = d/d_f+1$. Also of interest are the aspect ratio of clusters, given by the ratio of second moments and scaling as L^0. While the lattices we study in this thesis are isotropic save the voltage applied along one direction but not the other, in fracture surfaces are thought to be strongly anisotropic, and so more nontrivial scaling of aspect ratios is expected as $L_{\parallel}$ and $L_{\perp}$ are varied separately.

Figure 1.17 shows the second-moment correlation lengths in directions both parallel and perpendicular to the applied stress as a function of the scaling variable we fit in the previous subsection. For small β the correlation length is larger than the system size, and the scaling behaves as expected. For intermediate β, the larger values of the ratio of correlation length to system size for smaller systems indicate that the correlation length becomes smaller than the system size.

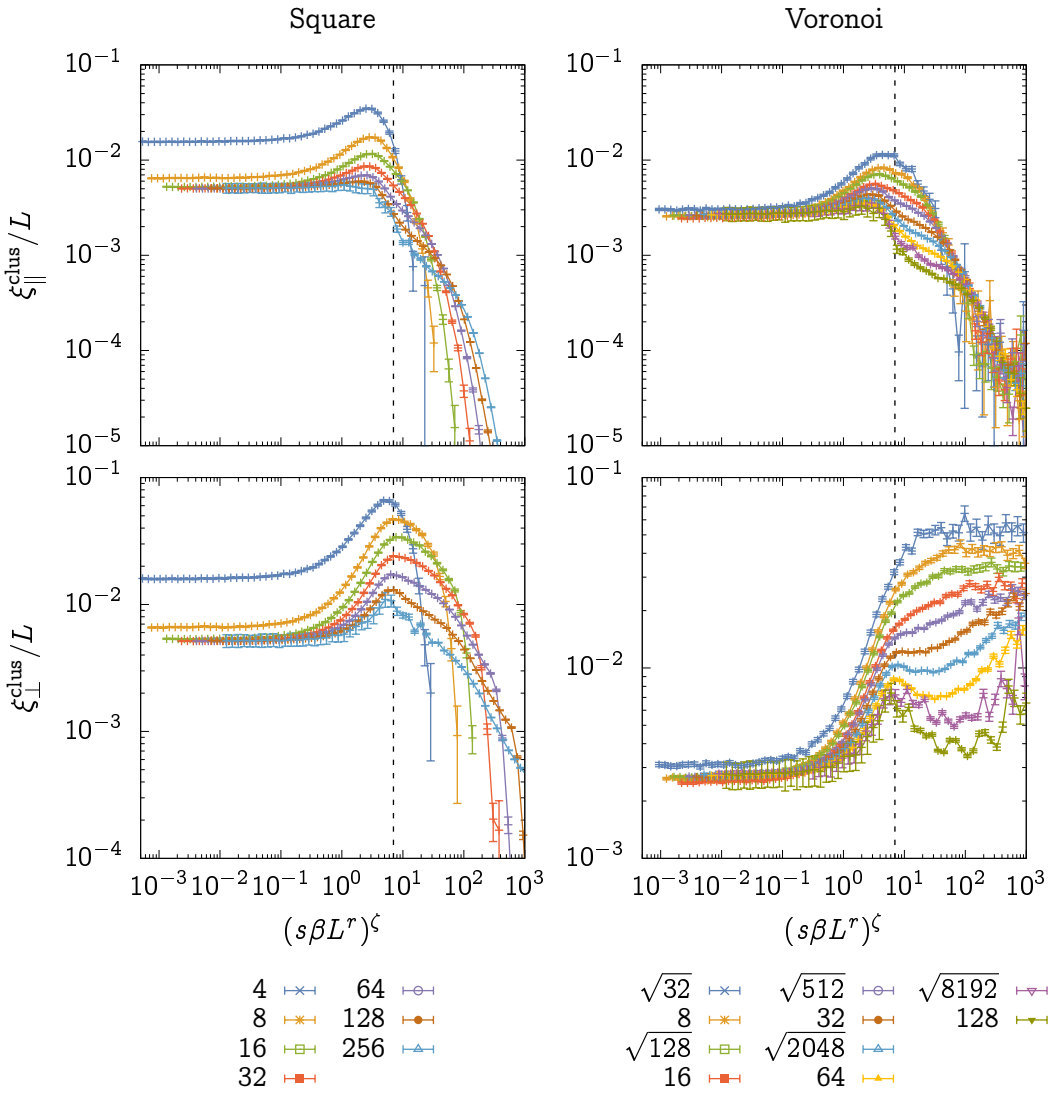

Figure 1.17: The second-moment correlation lengths of clusters at critical stress as a function of the scaling variable fit for the cluster size distribution in §6. The parallel direction is the direction along which stress is applied, and the perpendicular direction is the direction along which the crack propagates.

The correlation length along the direction stress is applied develops a maximum near the crossover regime, but that gives way for larger systems and has an apparently universal tail that is *not* described by our single scaling variable. The correlation length along the direction of crack propagation also develops of maximum at around the same point (emphasized with a dashed line on all plots), but unlike the correlation length parallel to the applied stress this appears to sharpen as the peak becomes smaller. For larger systems, the location of the peak in $\xi_{\perp}^{\text{clus}}$ appears to coincide with the formation of an inflection point in $\xi_{\parallel}^{\text{clus}}$. Right now our scaling theory does not address why this should be the case.

Figure 1.18 shows the aspect ratio of the connected clusters at the critical stress, again as a function of our fit scaling invariant. In both lattice geometries it is very well described by our fit scaling variable alone up to around the same place where the peak in $\xi_{\perp}^{\text{clus}}$ occurs, again drawn as a vertical dashed line, at which point the curves fall away in a way that at least initially seems independent of the lattice.* There may be the start of a power law decay after the second crossover, but our system sizes are not able to resolve this more than a decade.

Notice that while the aspect ratios are near one for very small β, they asymptote on a value less than one that does not approach it with increasing L, at around 0.98. This may seem troubling, since percolation is definitely an isotropic phenomena—how does our limit have a baked-in anisotropy if it is meant to replicate percolation? Though our lattice-breaking procedure is

*Far outside the field of view, the square lattice curves drop to zero while the voronoi lattice curves flatten.

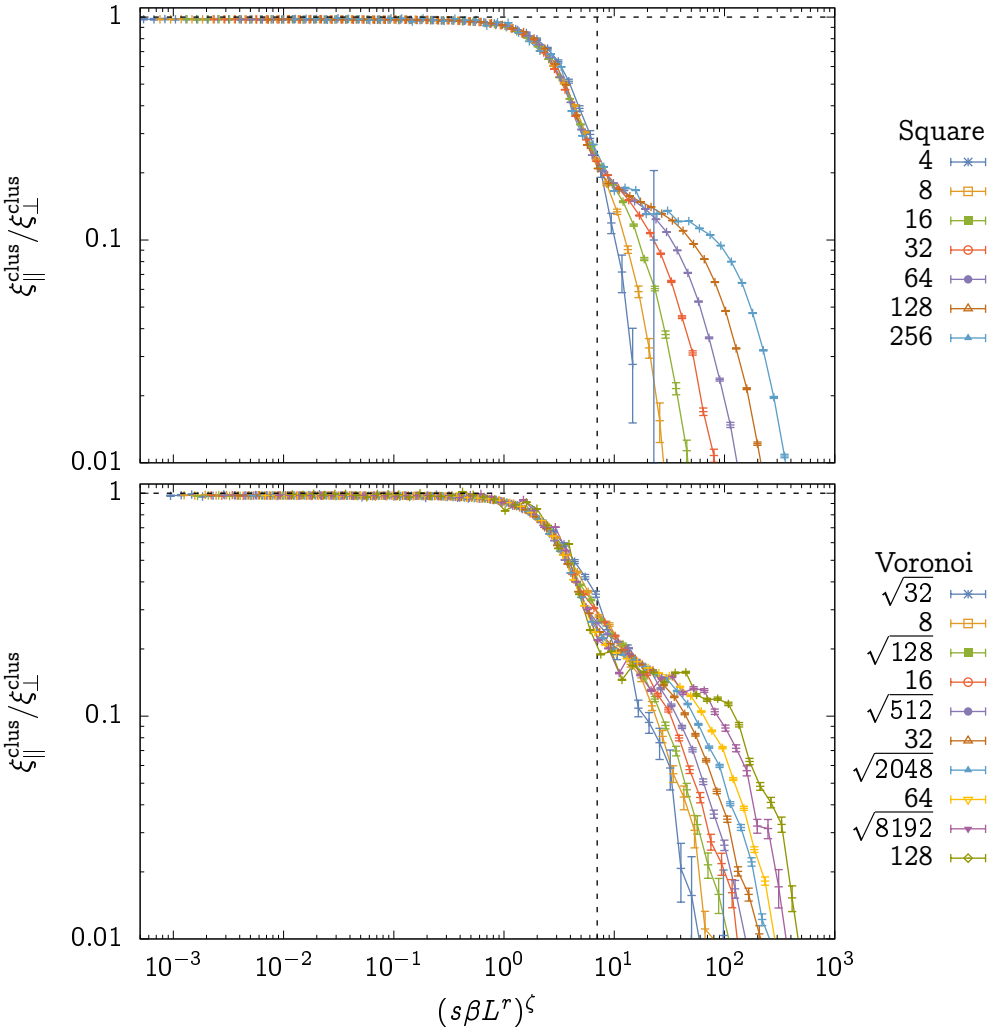

Figure 1.18: Aspect ratios of connected clusters at the critical stress as a function of the scaling variable we fit to the cluster size distributions in §6. Note that the parallel direction is the direction of applied stress, and the perpendicular direction is the direction of crack propagation.

different from percolation even in the infinite-disorder limit, the identification of lollipops and bow ties is independent of the direction of applied stress, and so this is not the source of our problem. Rather, it has to do with the definition of critical stress: the point at which the lattice stops conducting *along the direction of applied stress*. Here is the asymmetry: if we percolate along the parallel direction first, we don't stop breaking fuses, while if we percolate along the perpendicular direction first we are done. This leads, even in the limit of infinite disorder, to clusters that are more stretched along the perpendicular direction, since we must guarantee that a spanning cluster exists along that direction but not the other. This sort of behavior makes our small deviation from the percolation fractal dimension more plausible, since it suggests that, despite sharing many properties with percolation exactly, the bonds left unbroken by the fuse problem that would be broken by percolation lead to a fundamentally different network connectivity. We will see this very starkly when we examine the minimal crack surface in §9.

§7. The spanning cluster

At percolation there is a cluster of broken bonds whose duals span the system. At the percolation fixed point, this infinite cluster has a mass that grows like L^{d_f} with the system size. We can identify and measure the size and correlations in this infinite cluster in our fuse networks after they have broken.

The average size of the spanning cluster is shown plotted in Figure 1.20 as a function of the scaling variable we fit in §6. Here the agreement is good far out in the scaling variable, with a behavior that seems to converge to a

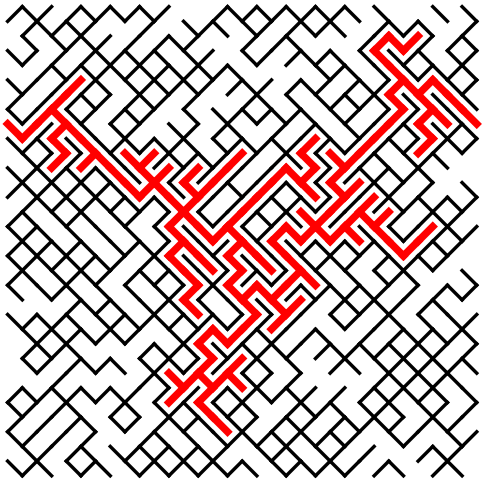

Figure 1.19: A 64×64 square lattice with $\beta = 0.316$, after rupture. The spanning cluster on the dual lattice is highlighted in red.

universal scaling function as L grows. The identical crossovers for the two lattices are indicative of a trivial regime, where the system is broken by a clean crack whose size grows proportional to L.

Since the distribution of spanning clusters is peaked heavily around its mean and not power-law distributed like the clusters before fracture, we now expect

$$\xi^{\text{span}} \simeq \frac{\langle x^2 \rangle}{\langle s^2 \rangle} \propto \frac{L^{2+1-\eta}}{L^{2d_f}} = L^{-1} \tag{1.55}$$

though this may prove to be different if we cannot trust percolation d_f or η to describe our spanning cluster, which recall is underpopulated due to not breaking bonds contained in lollipops and bow ties. If η were unchanged but d_f took the value we fit earlier, we would see $\xi \sim L^{0.946}$. Instead, the data we plot in Figure 1.21 agrees best at small β for $\xi \sim L^{0.9}$. The correlation length $\xi_{\parallel}^{\text{span}}$ along the direction that current is applied has a apparently

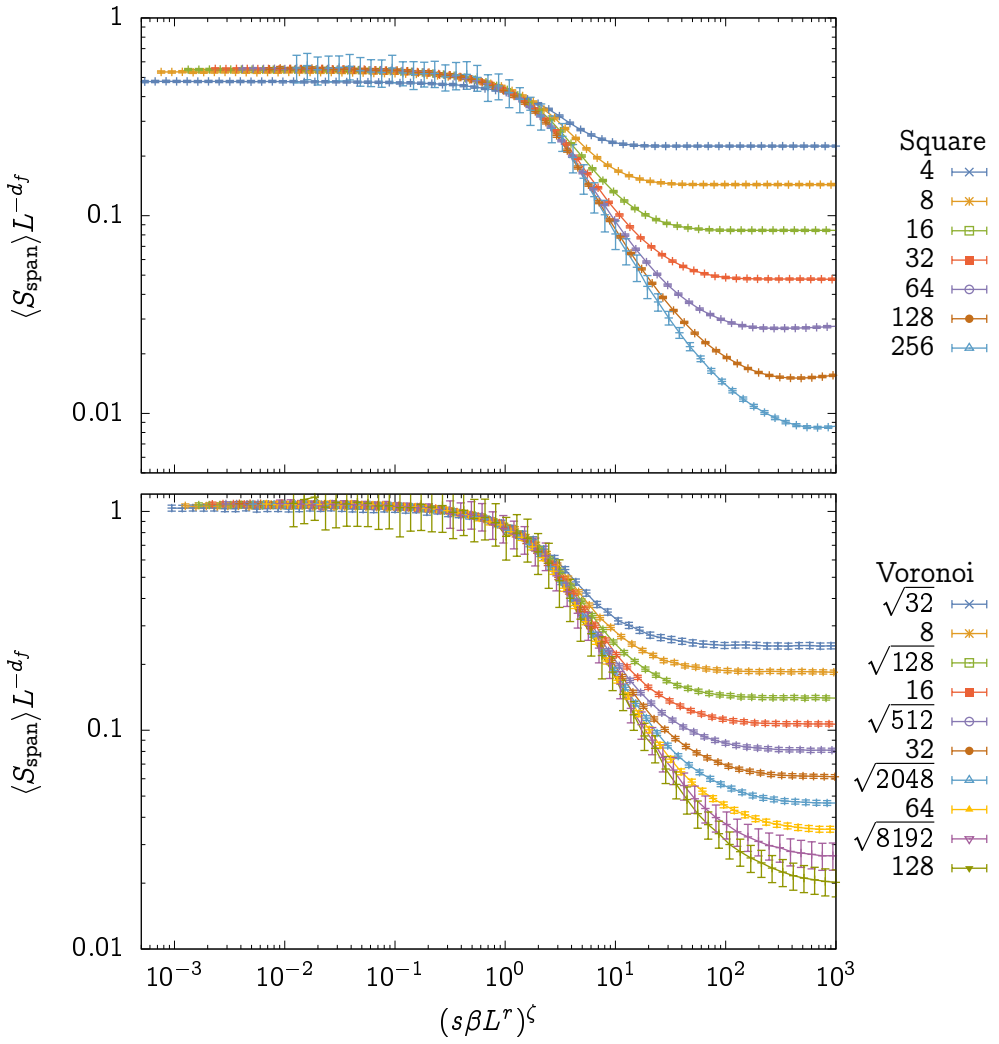

Figure 1.20: The average size of the spanning cluster rescaled by the fractal dimension as a function of the scale invariant we fit to the cluster size distributions in §6.

universal peak around the same place that the correlation length in the perpendicular direction did for the clusters at critical stress, shown again by the dashed vertical line. The tail of $\xi_{\parallel}^{\text{span}}$ appears to be a power law, apparently maintained in the voronoi case for several decades. This length, which can be considered as a sort of "crack width" should be compared with results in depleted elastic networks that show monotonic increase with disorder.[57] The correlation length $\xi_{\perp}^{\text{span}}$ along the direction of crack propagation simply grows monotonically from a minimum value before nonuniversal lattice effects take over.

Figure 1.22 shows the aspect ratio of the spanning cluster as a function of our scaling invariant. Since the individual correlation lengths both were well-described by simple functions of our scale invariant, it is not surprising that the aspect ratio of the spanning cluster likewise is well described as a function of the scale invariant alone. This, like the clusters, does not approach isotropy in the small β limit, as more clearly seen in this figure than for the clusters in Figure 1.18 since we have forgone the logarithmic scale.

§8. The conducting backbone

At critical stress, when the network is about to be ruptured, there are a great many bonds that don't carry any current. We spoke at length about the technical reasons why in a previous section, but now we will reckon with exactly what this does by looking at the scaling of the conducting backbone, which is the set of bonds at critical stress that still carry a current. At percolation we have already seen that the size of spanning clusters grows with

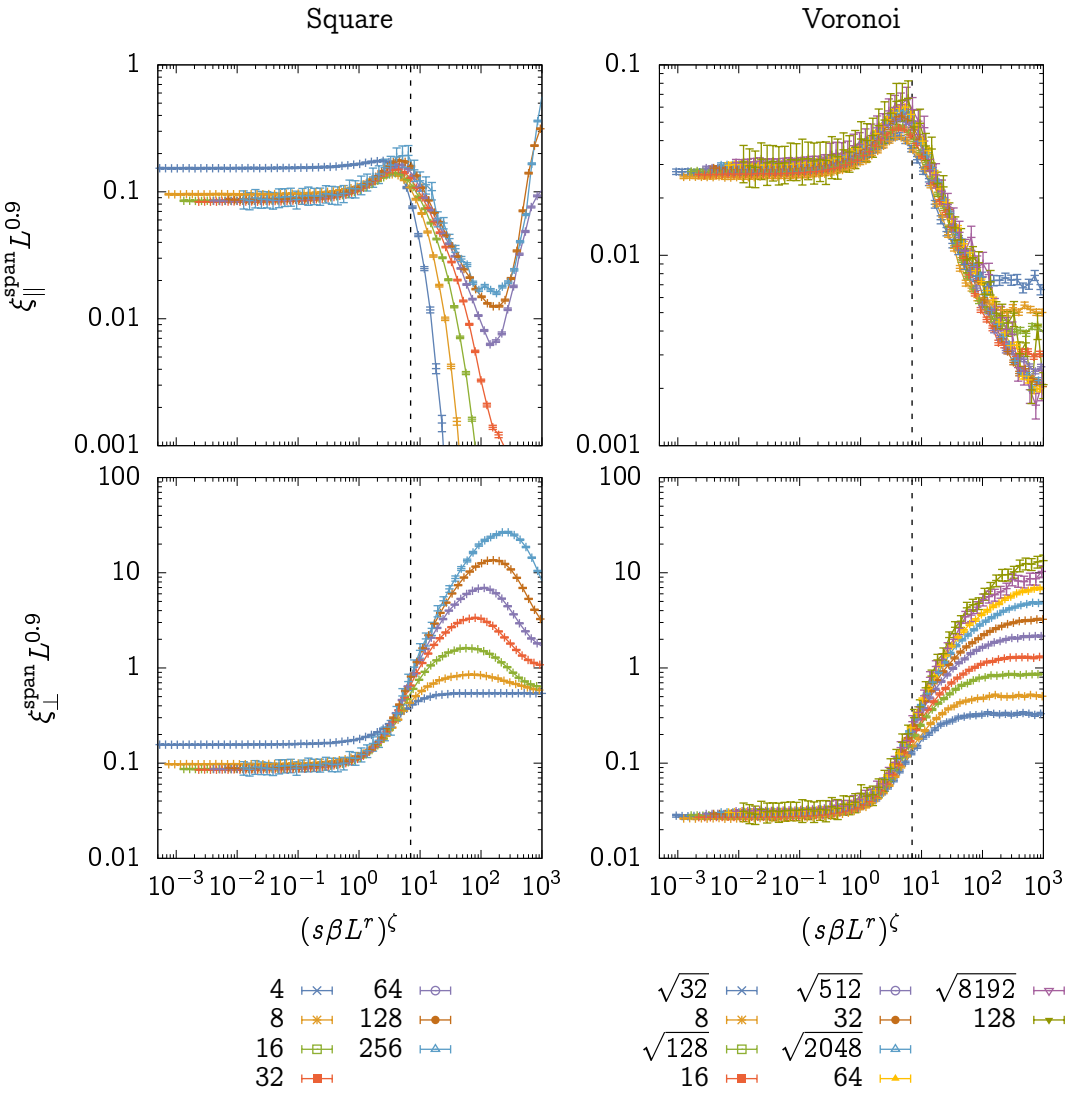

Figure 1.21: Correlation lengths for the spanning cluster after the system is ruptured. As a function of the invariant fit in §6. The parallel direction is the direction of applied current, while the perpendicular direction is the direction of crack propagation.

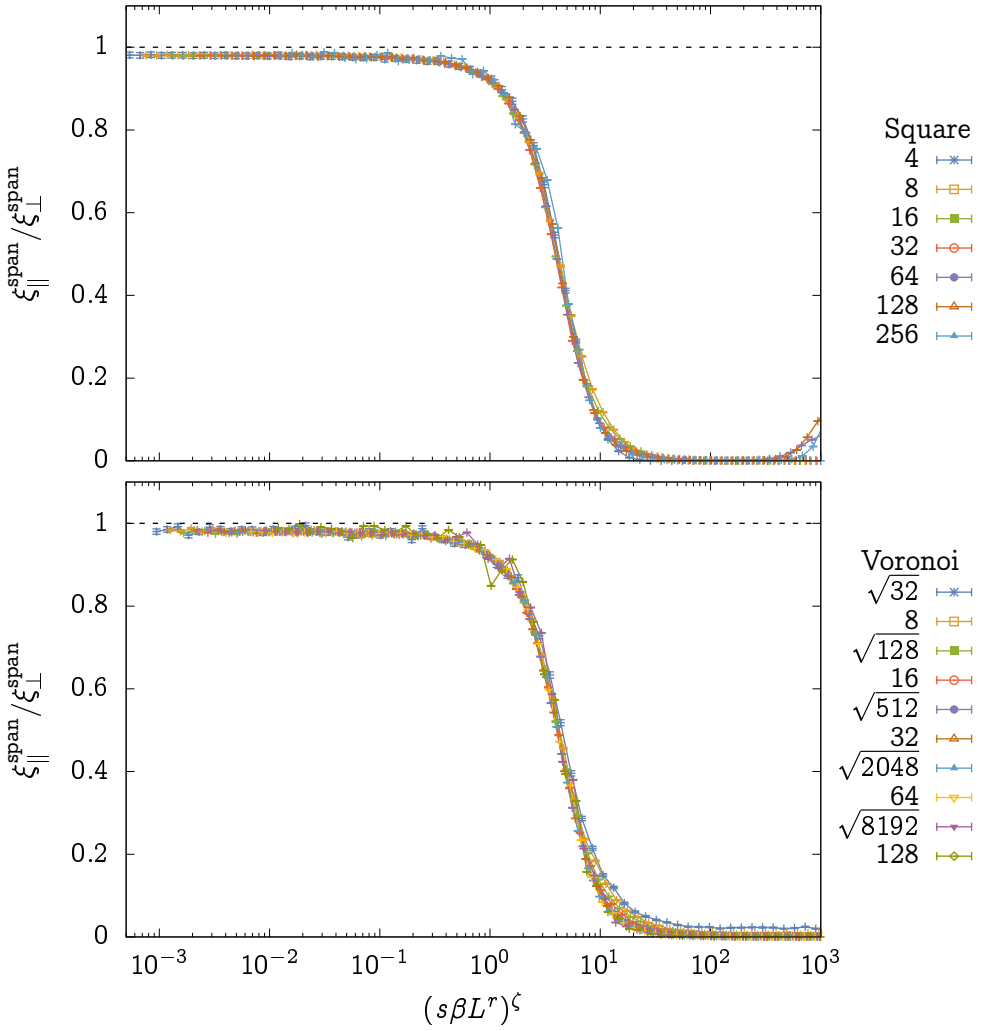

Figure 1.22: The aspect ratio of the spanning cluster after rupture as a function of our scale invariant fit in §6. Note that the parallel direction is the direction that current is applied, while the perpendicular direction is the direction of crack propagation.

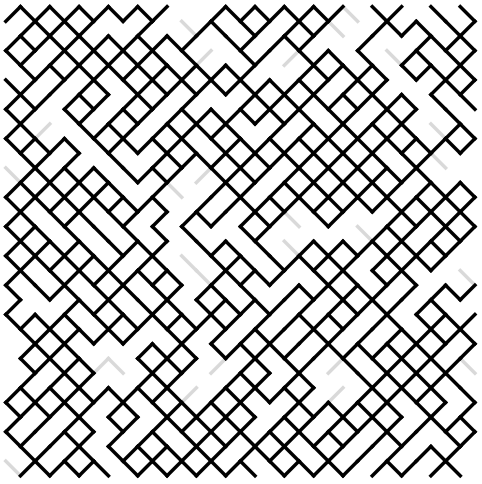

Figure 1.23: A 64×64 square lattice with $\beta = 0.316$, held at critical stress. Bonds that belong to the conducting backbone are black, while those that do not are light gray.

a fractal dimension; it turns out that there are so many dead ends in the percolated network, those remaining grow with a *different*, smaller fractal dimension d_b. The best estimates have $d_b \simeq 1.64336$.[43]

Unlike the case of other properties, we have no reason to suspect our modifications to straight percolation will influence the size scaling of the backbone, since every single bond we would remove in percolation but don't is a bond that carries no current, and therefore isn't part of the backbone anyway. Figure 1.24 shows the average size of the conducting backbone at the critical stress as a function of our invariant fit in §6, rescaled by the percolation prediction, and the agreement is good.

Since $\xi^{d_b} \sim \langle S \rangle$, we might expect $\xi \sim L^{1/d_b} = L^{0.6085}$. Figure 1.25 shows the correlation length data as a function of the scaling invariant, rescaled by that factor with reasonable agreement. More interesting is the aspect ratio, shown in Figure 1.26, and which starts at an anisotropic value of about 0.9 near percolation before peaking at intermediate β before becoming isotropic

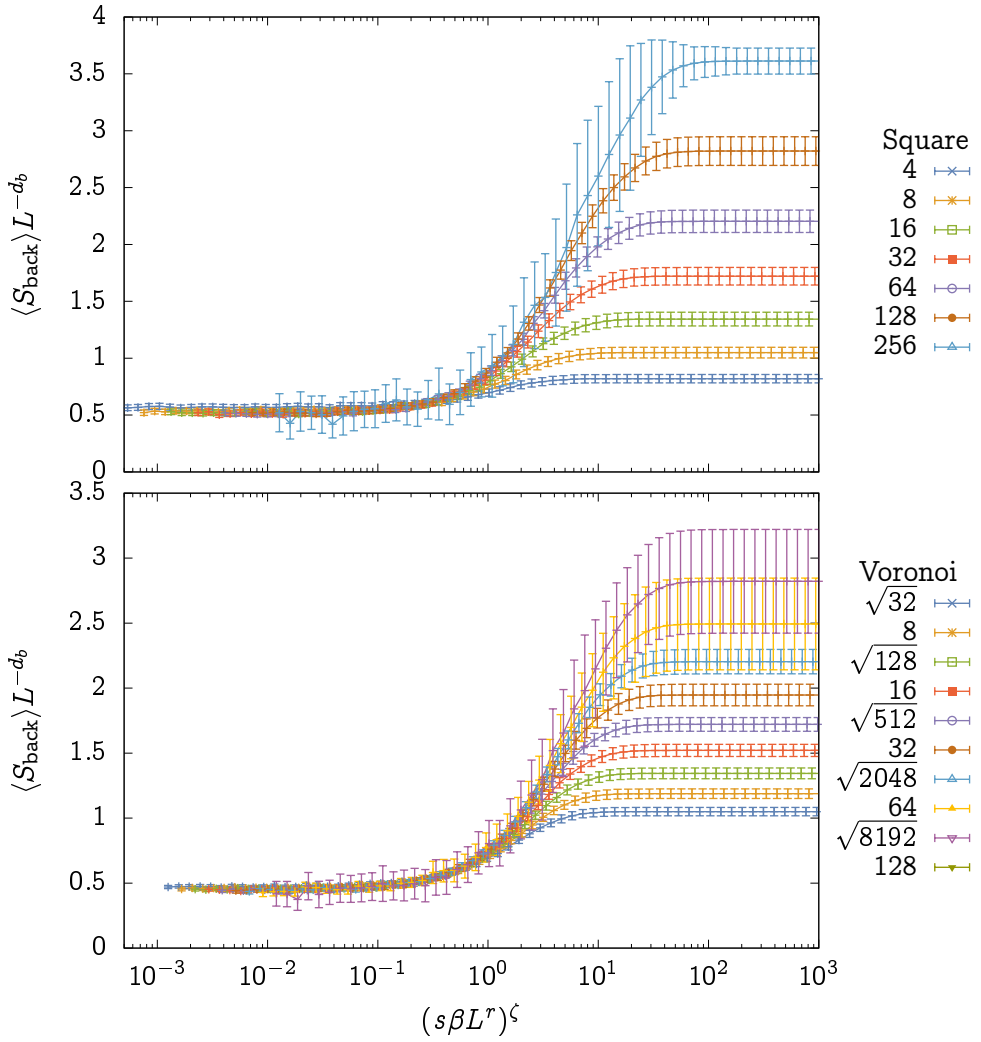

Figure 1.24: Average size of the conducting backbone at critical stress as a function of the scaling variable we fit for small β in §6.

in the limit of no damage at critical stress.

The properties of the distributions of currents on the conducting backbone are quite rich, and beyond the scope of the current study, but they are worth touching on because that distribution should affect the structure of avalanches in the near-percolation limit. At percolation, the distribution of currents is *multifractal*, meaning that unlike the power-law distributed cluster size distribution, for instance, moments of the current distribution are not linear in the moment number.[58–81] This can be interpreted as having assigned a different fractal dimension to different characteristic current magnitudes, with the linchpin fuses that carry most of the current having a very small dimension, while the fuses with middling current have one that more closely resembles that of the backbone itself. Multifractal scaling has been used to measure structure in concrete.[82]

§9. The crack surface

The minimal crack surface is defined by the unique path on the dual lattice that spans the torus in the direction opposite the applied current. An example is shown in Figure 1.27. When we want to draw something that one would intuitively describe as "the crack," this is it. In the percolation limit it is forced to crawl around a fractal maze and therefore its length grows more quickly than linear with L. In percolation, the analogue to this minimal crack is the minimal path, which can be defined as the shortest path that spans, e.g., the torus. In 2D its fractal dimension has been measured to be $d_{\min} \simeq 1.13077$.[44]

Figure 1.28 shows the average size of the minimal crack measured in our

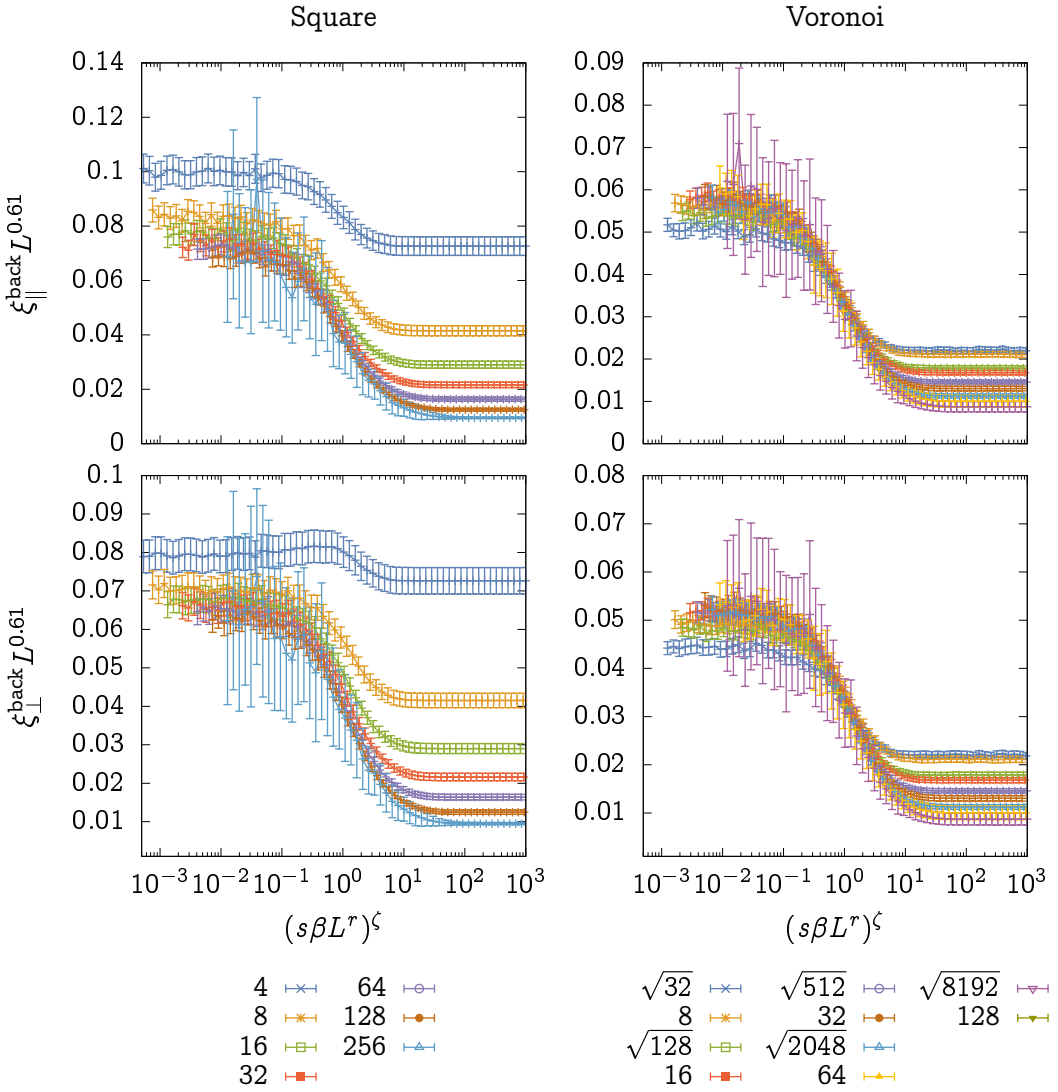

Figure 1.25: Correlation lengths for the conducting backbone at critical stress as a function of the scale invariant fit in §6. Recall that the parallel direction is the direction current is applied and the perpendicular direction is the direction of crack propagation.

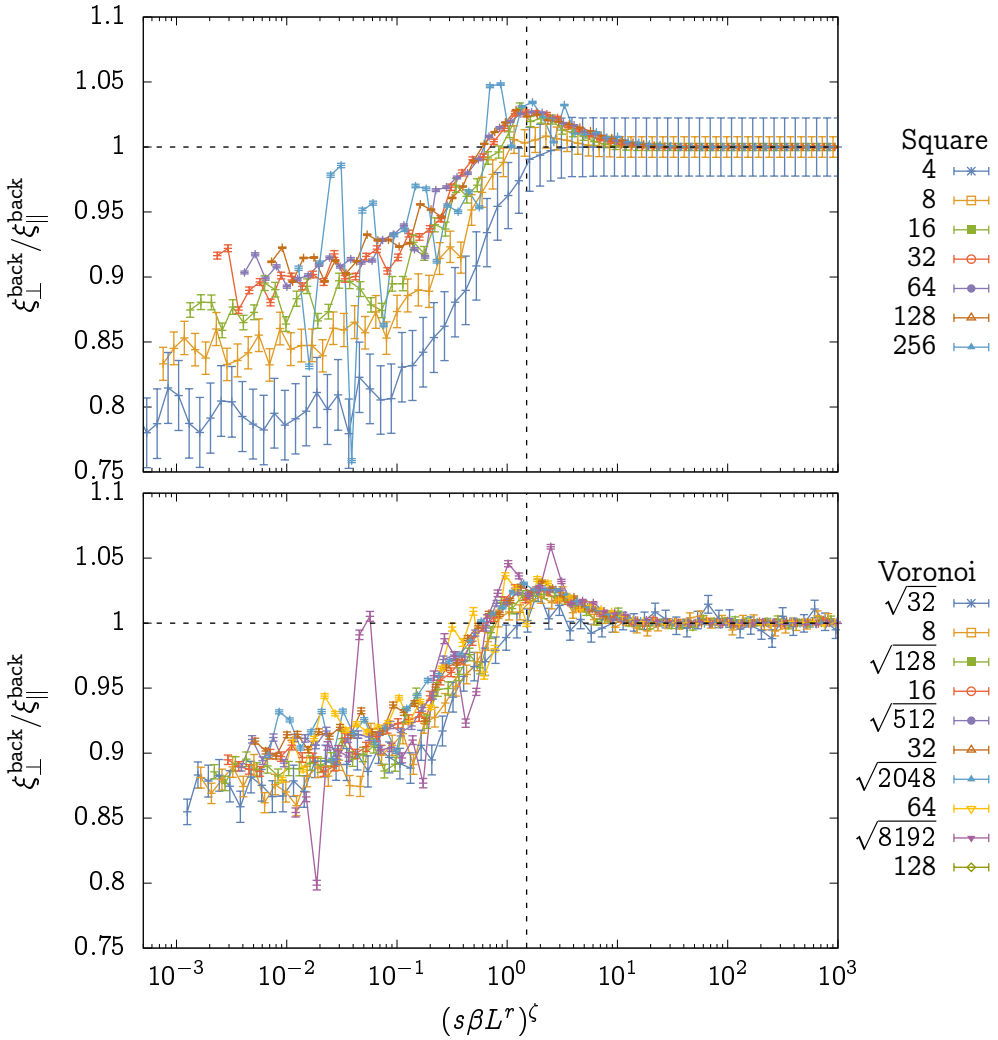

Figure 1.26: Aspect ratio of the conducting backbone as a function of the scale invariant fit in §6.

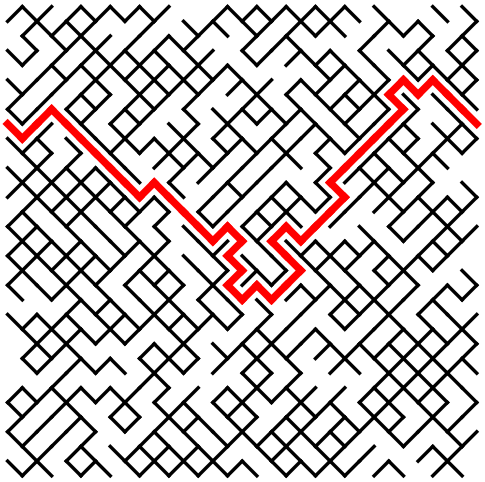

Figure 1.27: A 64×64 square lattice with $\beta = 0.316$, held at critical stress. The set of bonds that make up the crack surface on the dual lattice are highlighted in red.

simulations, rescaled by $d_{\min} \simeq 1.23$, substantially larger than that measured for 2D percolation but consistent with that measured by researchers on a similar model.[45] Here, the discrepancy between fuse network structure and percolation structure is more clear. Not only is the minimal crack surface not necessarily the shortest path (since it must span against the direction of applied current), but it isn't even close. The unbroken stems of lollipops dramatically decrease the paths that the crack surface can take on the fractured system, to the extent that in actual percolation there are many paths that span in a given direction, but in this problem there is *exactly one*. It is not surprising, then, that we have to travel farther on average. Besides bucking the expectations set by percolation, the size of the minimal crack appears to scale reasonably well with our scaling variable fit in §6.

Figure 1.29 shows the correlation lengths for the minimal crack surface. These are all described well by our single fit scaling variable, with those along the direction of applied current crossing over into a decreasing power law

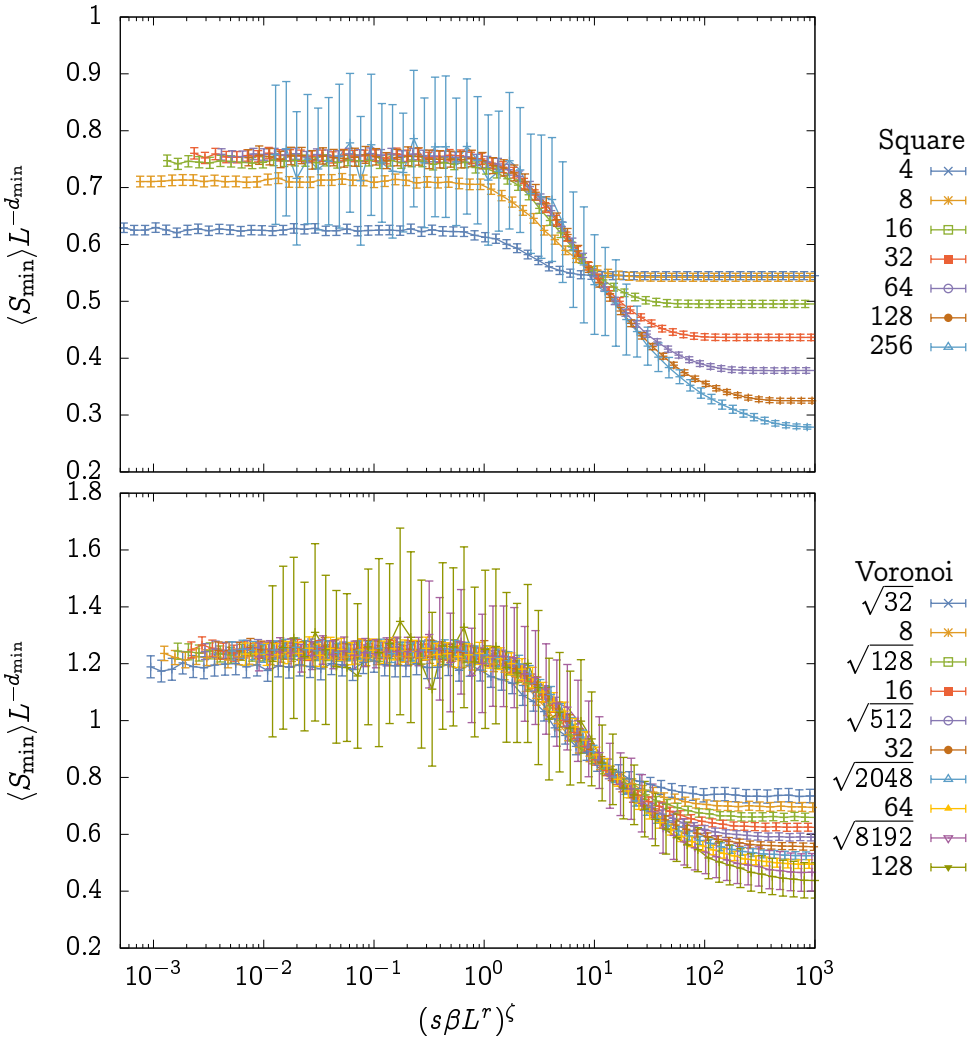

Figure 1.28: Average size of the minimal crack surface rescaled by $L^{d_{\min}}$ for $d_{\min} \simeq 1.23$ as a function of the scale invariant we fit in §6.

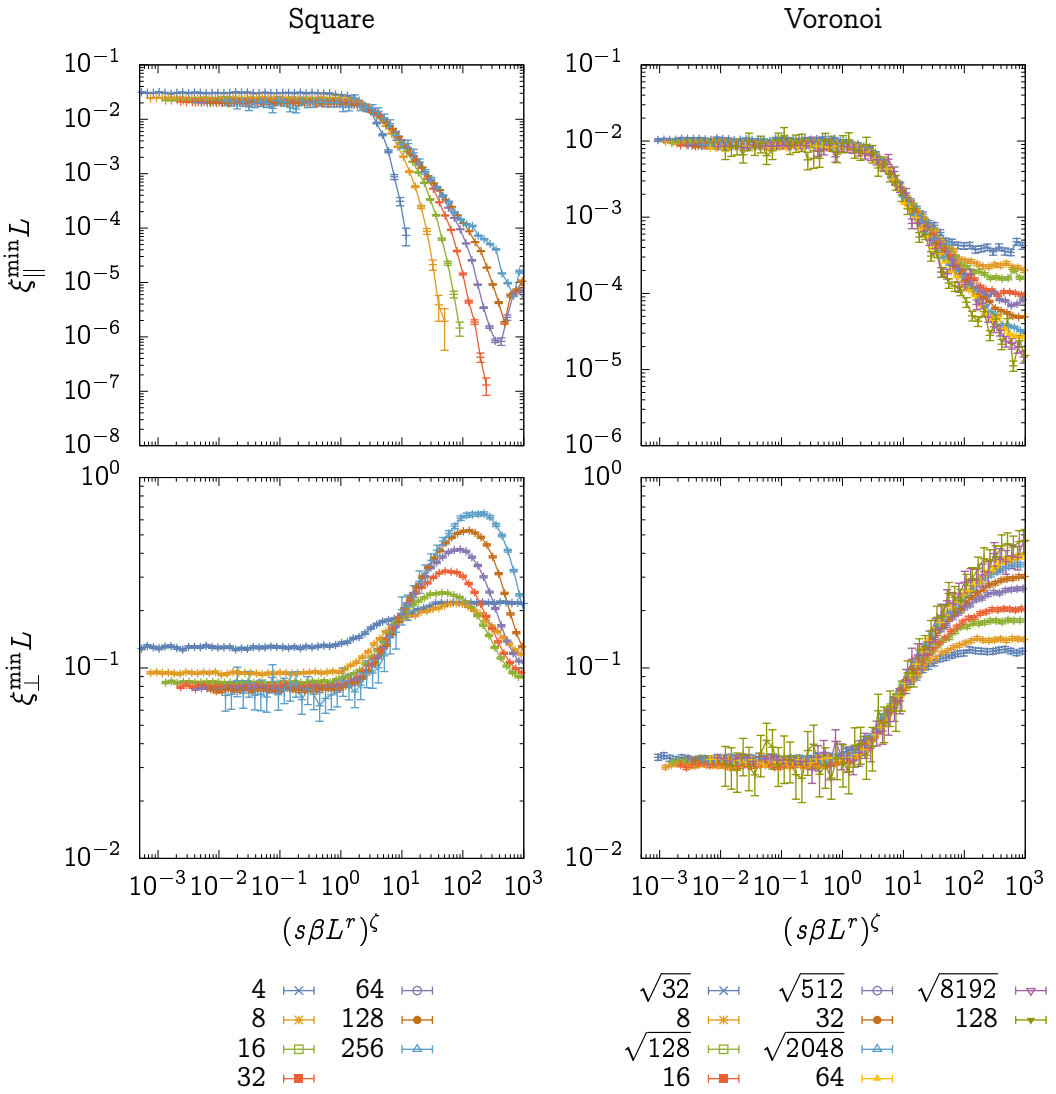

Figure 1.29: Correlation lengths for sites in the minimal crack surface as a function of the scale invariant fit to cluster size distributions in §6. Recall that the parallel direction is the direction current is applied in, while the perpendicular direction is the direction of crack propagation.

while those along the direction of crack propagation crossing over into an increasing power law. Its aspect ratio, shown in Figure 1.30, likewise has very clean behavior, with a simple power-law crossover from a value that seems to coincide on the square and voronoi lattices, around 0.26.

Besides being an abstract quantity that percolation researchers are interested in, the minimal crack surface has the possibility for direct connection to experiment: in the case it happens to be single-valued in $x_\parallel$, it is the surface one would see if one separated the two halves of a fractured object. The quantity that people are interested in in this context is the *height–height correlation function*, defined by

$$h(x_\perp)^2 = \langle (x_\parallel(0) - x_\parallel(x_\perp))^2 \rangle = \int x_\parallel^2 g_{\min}(x_\parallel, x_\perp)\, dx \tag{1.56}$$

Normally, this is expected to scale like $h(x_\perp) \sim x_\perp^\zeta$ for some positive exponent ζ that describes the self-affine fractal nature of the crack surface.[21,83–96] Because we have a minimal surface that can wrap back on itself ($x_\parallel(x_\perp)$ is not single-valued), the $h(x_\perp)$ we get by integrating the correlation functions for the minimal surface do not go to zero at zero $x_\perp$. Figure 1.31 shows some examples of these for an $L = 256$ square lattice for several β. For large β there is relatively good power law dependence up to a point, while this is quickly swamped by a constant floor that results from how increasingly agnostic the crack surface is about moving vertically. They all turn over at large $x_\perp$ due to the fact we are on a torus, and these functions must therefore be periodic. This strong system-size dependent feature makes it difficult to compare the height–height correlation functions directly across β and L.

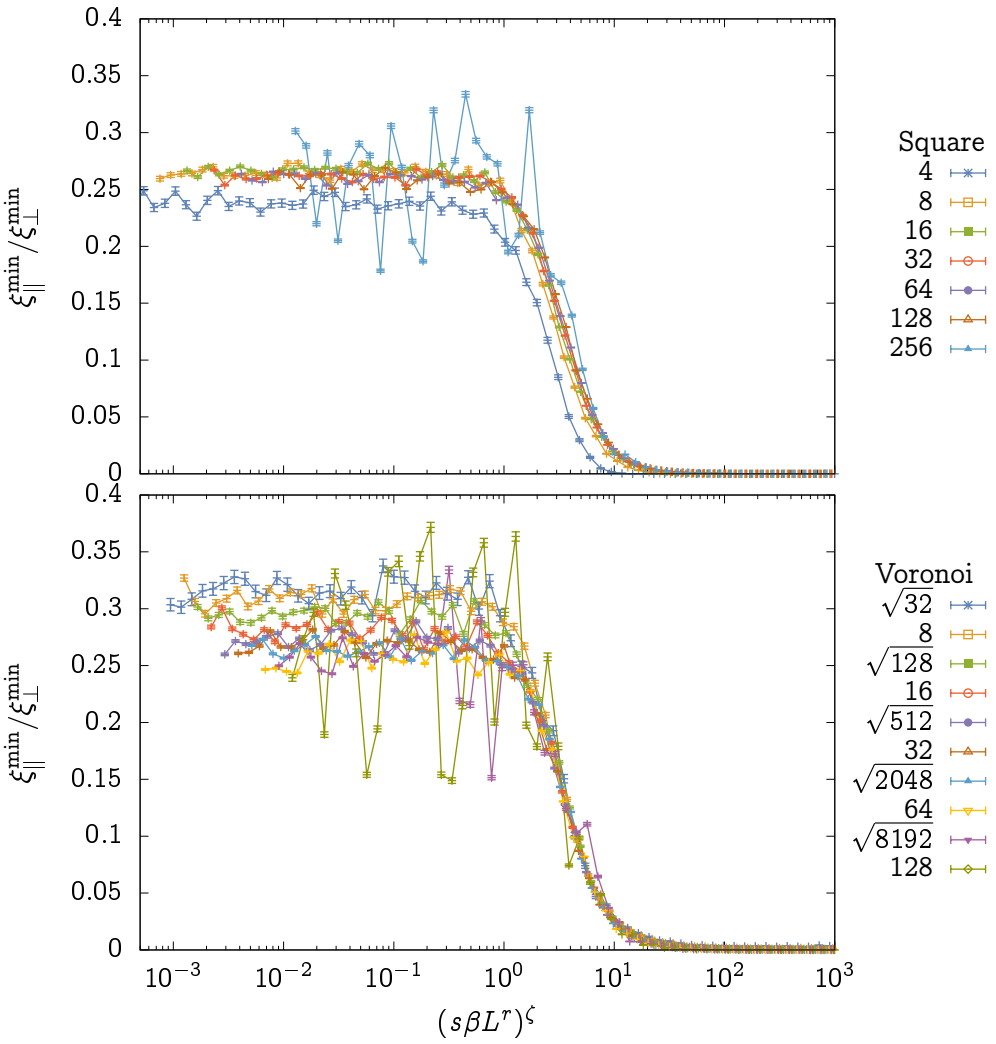

Figure 1.30: The aspect ratio of the minimal crack surface as a function of our scaling invariant fit in §6. Recall that the parallel direction is the direction that current is applied in, while the perpendicular direction is the direction of crack propagation.

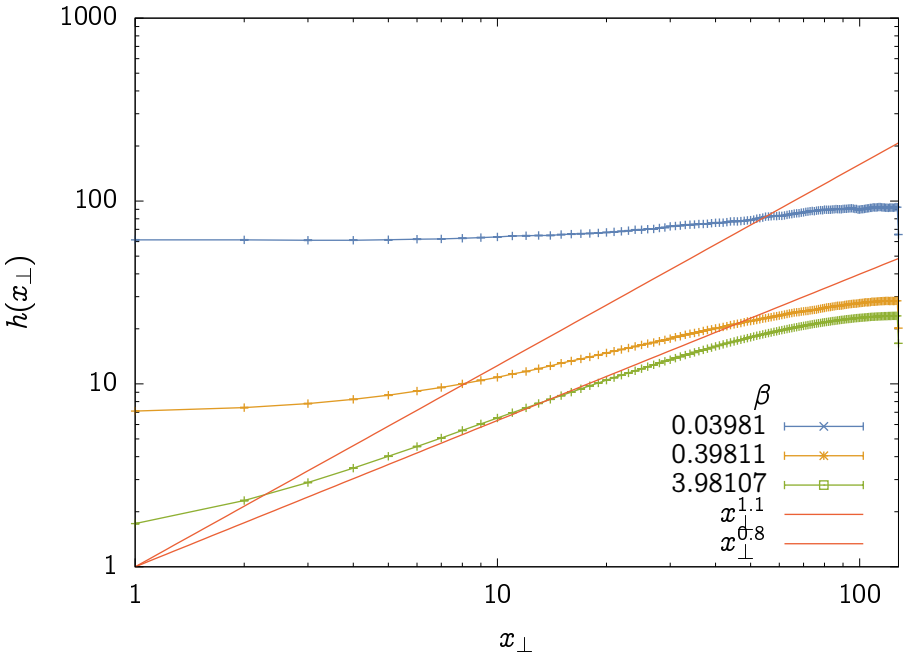

Figure 1.31: Example height-height correlation functions for an $L = 256$ square lattice for several values of β.

If h goes as $h(x_\perp) \sim x_\perp^\zeta$, we might expect its first moment to scale as

$$\begin{aligned} h_1 &= \int x_\perp h(x_\perp \mid L, u_p, u_d)\, dx_\perp = \int x_\perp^{1+\zeta} \mathcal{H}(x_\perp L^{-1}, u_p L^{y_p}, u_d L^{y_d})\, dx_\perp \\ &= L^{2+\zeta} \int u^{1+\zeta} \mathcal{H}(u, u_p L^{y_p}, u_d L^{y_d}) L\, du \end{aligned} \tag{1.57}$$

and therefore with $L^{2+\zeta}$. Figure 1.32 shows data depicting h_1 as a function of our scaling invariant and rescaled to collapse in small β by $L^{3.1}$, which would seem to imply that $\zeta \simeq 1.1$, which is also plotted as a line in Figure 1.31. However, that region does not have $h \sim x_\perp^\zeta$, and so our scaling consideration isn't expected to hold. In the large-β limit, we get good agreement for $\zeta \simeq 0.8$, but the moments act like a function of a different scaling variable βL^t for

$t \simeq 0.4$. The ζ resulting in this limit is roughly consistent with what we see in Figure 1.31 for $\beta > 1$.

§10. Precursor avalanches

We are now finished reviewing the *structural* properties of the network at or after the critical stress, and will now turn to *dynamical* properties. Here we will have very little to compare directly with percolation, since the dynamics of percolation are "break one bond at a time until you're done." However, mean-field models of fracture will guide our way. These may be expected to work reasonably well, since the range of the interaction between rupturing fuses is so long. First, we will look at the distribution of avalanches leading up to the rupture. Here we venture away from the rupture surface for the first time, averaging together *all* avalanches before the final one.

Avalanches are the groups of bonds that break in response to the failure of a single fuse, and they are found in our simulations by looking for backtracking in the external current necessary to break successive fuses. In fracture, they empirically have power-law tails.[97,98] The democratic fiber bundle model[99–103] is a mean-field theory of fracture that provides some predictions for the way that the precursor avalanches are distributed. It can be connected directly to a mean-field theory of failure in continuous solids,[37] and predicts that the avalanche sizes should be distributed by

$$n_s \sim s^{-\tau} f(s(I_f - I)^{\kappa}) \tag{1.58}$$

where I is the applied current and I_f is the current that causes fracture. Com-

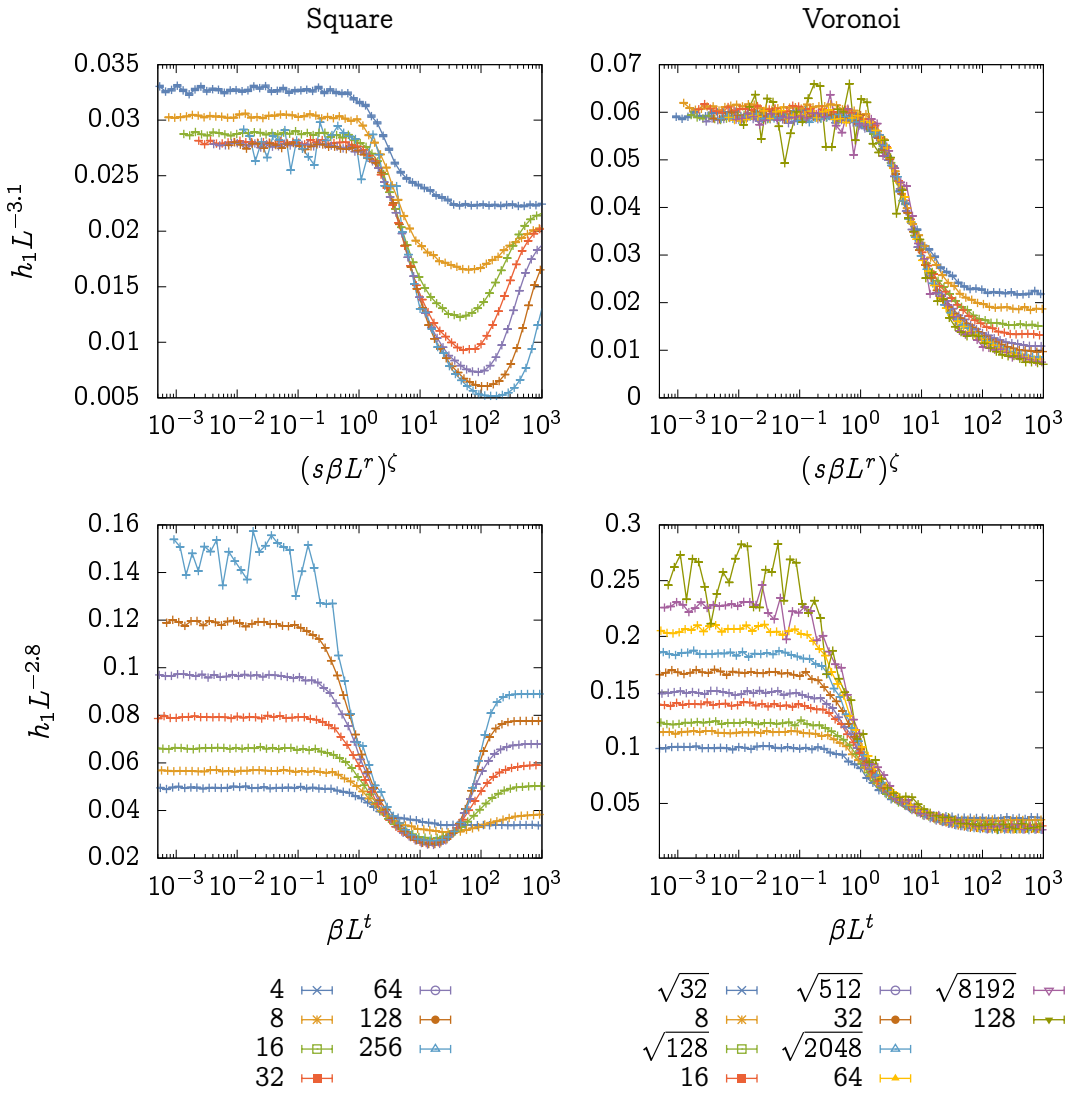

Figure 1.32: First moment of the height–height correlation function as a function of two different scaling variables, and rescaled two different ways. (Top) As a function of the scaling variable fit to the cluster size distributions for small β in §6. (Bottom) As a function of a scaling variable fit by eye with $t \simeq 0.4$.

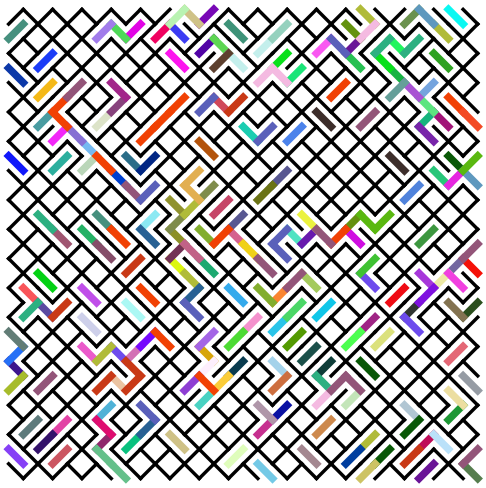

Figure 1.33: A 64×64 square lattice with $\beta = 0.316$, held at critical stress. Dual bonds whose partners were broken in the same avalanche are highlighted with the same color.

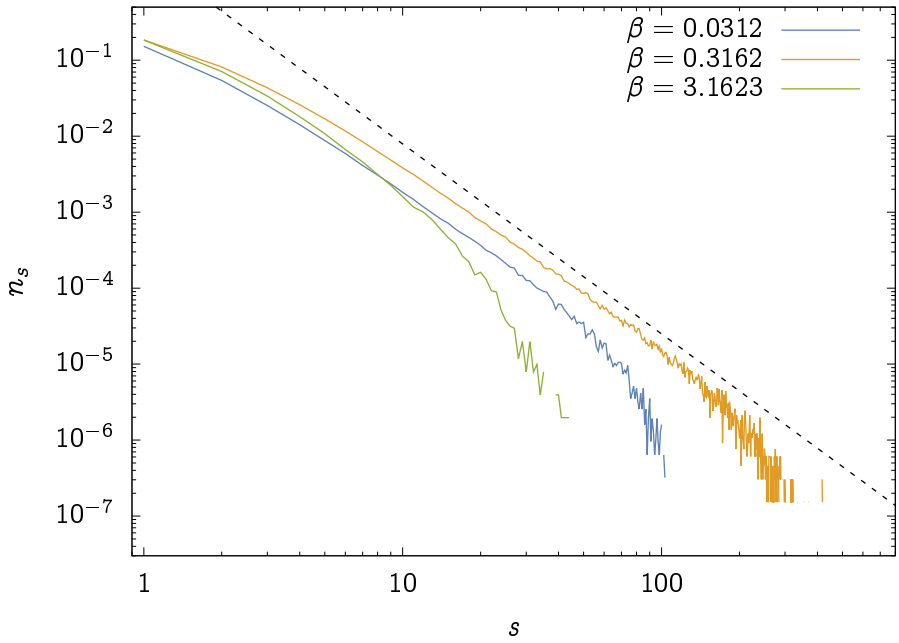

Figure 1.34: Example avalanche size distributions averaged over all precursors for a 128×128 square lattice fuse network for several disorder parameters β. The dashed line shows $s^{-\tau-1/\kappa}$ for $\tau + 1/\kappa = 5/2$, from mean field fiber bundles.

parison has been made between these avalanches and fluctuations observed on approach to a spinodal, where metastable states become unstable.[56] The mean field prediction gives $\tau = 3/2$ and $\kappa = 1$.

To make a connection with our scaling idea, we need a correspondence between I and our scaling variables. Since increasing the applied current is what advances the network depletion at fixed disorder, it is therefore reasonable to associate $I_f - I$ with $(u_p - u_{pf})^w$ for some power w, which likewise gives a proximity in the unbroken network from rupture. We take as a guess for the form of the avalanche size distribution

$$n_s = s^{-\tau}\mathcal{N}(s(u_p - u_{pf})^{w\kappa}, u_p L^{y_p}, u_d L^{y_d}) \tag{1.59}$$

It follows that the distribution averaged over all precursors is given by

$$\begin{aligned}
\tilde{n}_s &= \int_0^{I_f} n_s \, dI \\
&= \int_0^{I_f} s^{-\tau}\mathcal{N}(s(I_f - I)^{\kappa}, \left[u_{pf} + (I_f - I)^{1/w}\right]L^{y_p}, u_d L^{y_d}) \, dI \\
&= \int_{-I_f}^{0} s^{-\tau}\mathcal{N}(s\Delta I^{\kappa}, \left[u_{pf} + \Delta I^{1/w}\right]L^{y_p}, u_d L^{y_d}) \, d\Delta I \\
&= \int_{-I_f s^{1/\kappa}}^{0} s^{-\tau}\mathcal{N}(y^{\kappa}, \left[u_{pf} + y^{1/w} s^{-1/w\kappa}\right]L^{y_p}, u_d L^{y_d}) s^{-1/\kappa} \, dy \\
&= s^{-\tau-1/\kappa}\tilde{\mathcal{N}}(u_{pf} s^{1/\kappa w}, u_{pf} L^{y_p}, u_d L^{y_d})
\end{aligned} \tag{1.60}$$

Therefore, moments of the avalanche size distribution along the rupture sur-

face should be given by something of the form

$$\begin{aligned}\langle s^n \rangle &= \int_1^\infty s^n \tilde{n}_s \, ds = \int_1^\infty s^{n-\tau} \tilde{\mathcal{N}}(u_{pf} s^{1/\kappa w}, u_{pf} L^{y_p}, u_d L^{y_d}) \, ds \\ &= \int_{u_{pf}^{\kappa w}}^\infty u_{pf}^{-\kappa w(n-\tau)} y^{n-\tau} \tilde{\mathcal{N}}(y^{1/\kappa w}, u_{pf} L^{y_p}, u_d L^{y_d}) u_{pf}^{-\kappa w} \, dy \\ &= u_{pf}^{-\kappa w(n+1-\tau)} \left[\mathcal{N}_n(u_{pf} L^{y_p}, u_d L^{y_d}) + \mathcal{O}(u_{pf}^{\kappa w}) \right] \\ &= L^{y_p \kappa w(n+1-\tau)} \left[\tilde{\mathcal{N}}_n(u_{pf} L^{y_p}, u_d L^{y_d}) + (u_{pf} L^{y_p})^{-\kappa w(n+1-\tau)} \mathcal{O}(u_{pf}^{\kappa w}) \right]\end{aligned} \tag{1.61}$$

This predicts, up to eventual corrections, linear scaling in the moment n as a function of the scaling variables we have seen before, with eventual corrections.

Plots of moments $n = 2, \ldots, 5$ are shown in Figure 1.35 scaled by the scaling invariant fit to the cluster size distributions in §6, along with $\tau = 5/2$ and $w\kappa y_p \simeq 1.3$. Though this rough fit for the size scaling appears to perform well in terms of, e.g., the heights of the peaks, the distributions are *not* well-described by a single function of the scaling invariant we have been using. This data also appears inconsistent with a model that would attempt to explain this poor description using an irrelevant singular correction to scaling, which should decay like $L^{-\omega}$. The horizontal separation of, e.g., the peaks of these curves appears to not converge as this would predict, but (especially for larger n) remain constant, suggesting we have scaled by the wrong factor of L.

When rescaled by the same empirical invariant βL^t used in the lower plots of Figure 1.32 in §9 with $t \simeq 0.4$, the curves collapse well. The peaks appear to line up at around $s\beta L^t \simeq 2$, shown as a dashed line on the plot This suggests that the scaling of precursor avalanches may be governed by a

different length scale than governs most structural aspects for small β, either because the abrupt transition surface has experienced the sort of crossover behavior predicted by Models B or C, or as a result of proximity to the abrupt transition fixed point. It also appears that it is the *same* length scale that governs the formation of cracks with power law height–height correlations for large β.

The second-moment correlation lengths give another angle on the picture of precursor avalanches. These are shown, rescaled, in Figure 1.37. It is hard to reason much about the scaling of these, except to say that the precursor avalanches are *always* very long-ranged, even in the limit of zero β where they become small. Look, for instance, at the avalanche examples in Figure 1.33, where very few avalanches affect just nearby bonds. The correlation length along the direction of applied current, perpendicular to crack propagation, monotonically decreases with β, while the length along the direction of crack propagation has a shallow local minimum at around $(s\beta L^r)^\zeta \simeq 4$, which is shown is a dashed line in the plot. This local minimum, unlike the peaks of the size distributions, appears to scale with the scaling invariant fit from the cluster size distribution in §6.

The aspect ratio of these precursor avalanches is shown in Figure 1.38, rescaled by the empirical invariant $s\beta L^t$ introduced in §9. A dashed line at $s\beta L^t \simeq 2$ is shown at the same place as in Figure 1.36 for the avalanche size distribution moment peaks. The aspect ratio appears to have some asymptotic value for small β, and a local peak, governed by the cluster invariant $(s\beta L^r)^\zeta$, around the same time as similar peaks in the aspect ratio of static structural properties, but then a decay afterwards that appears again to be governed by

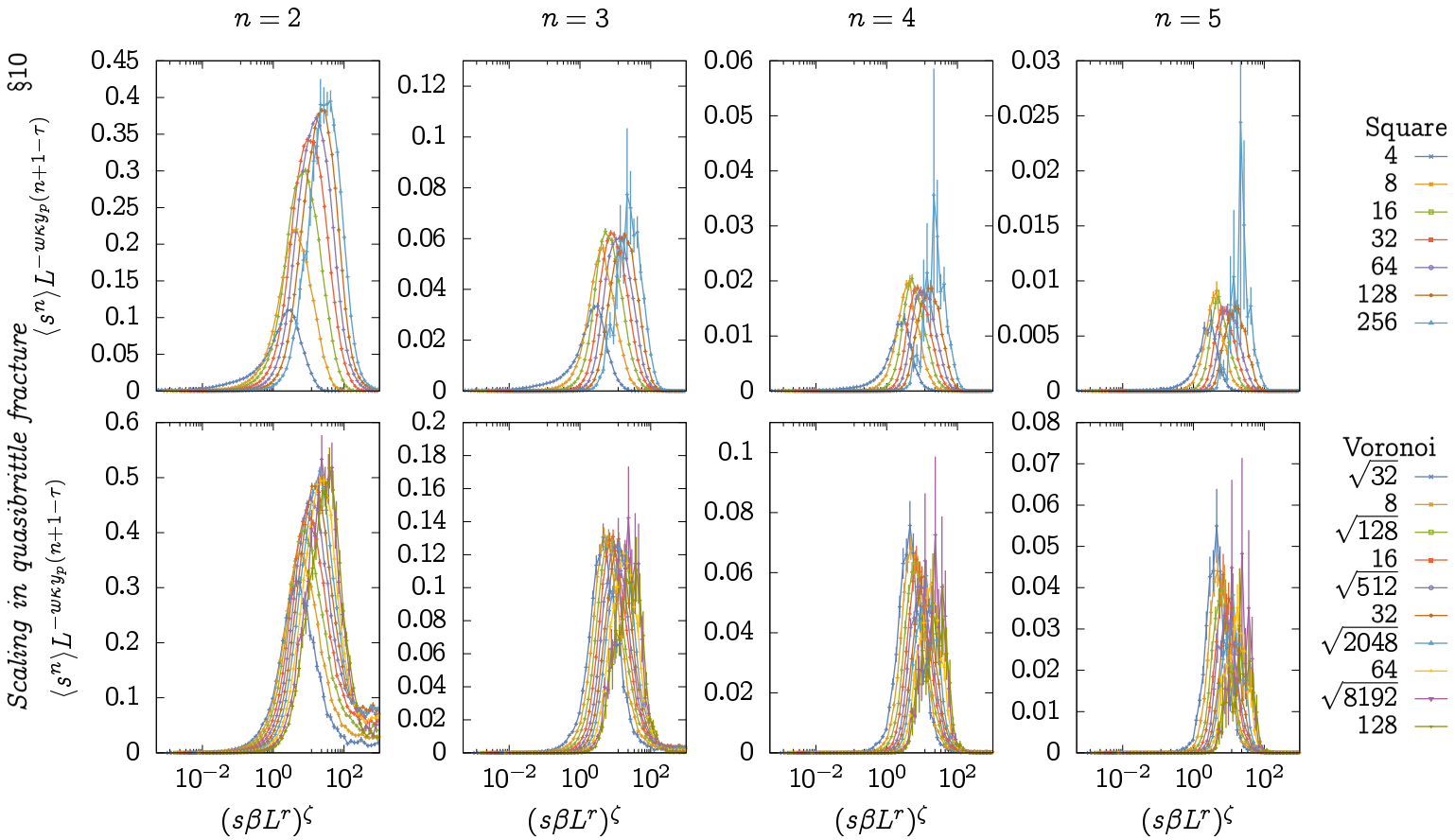

Figure 1.35: Rescaled moments of the precursor avalanche size distribution as a function of the scale invariant fit to the cluster size distribution for small β in §6.

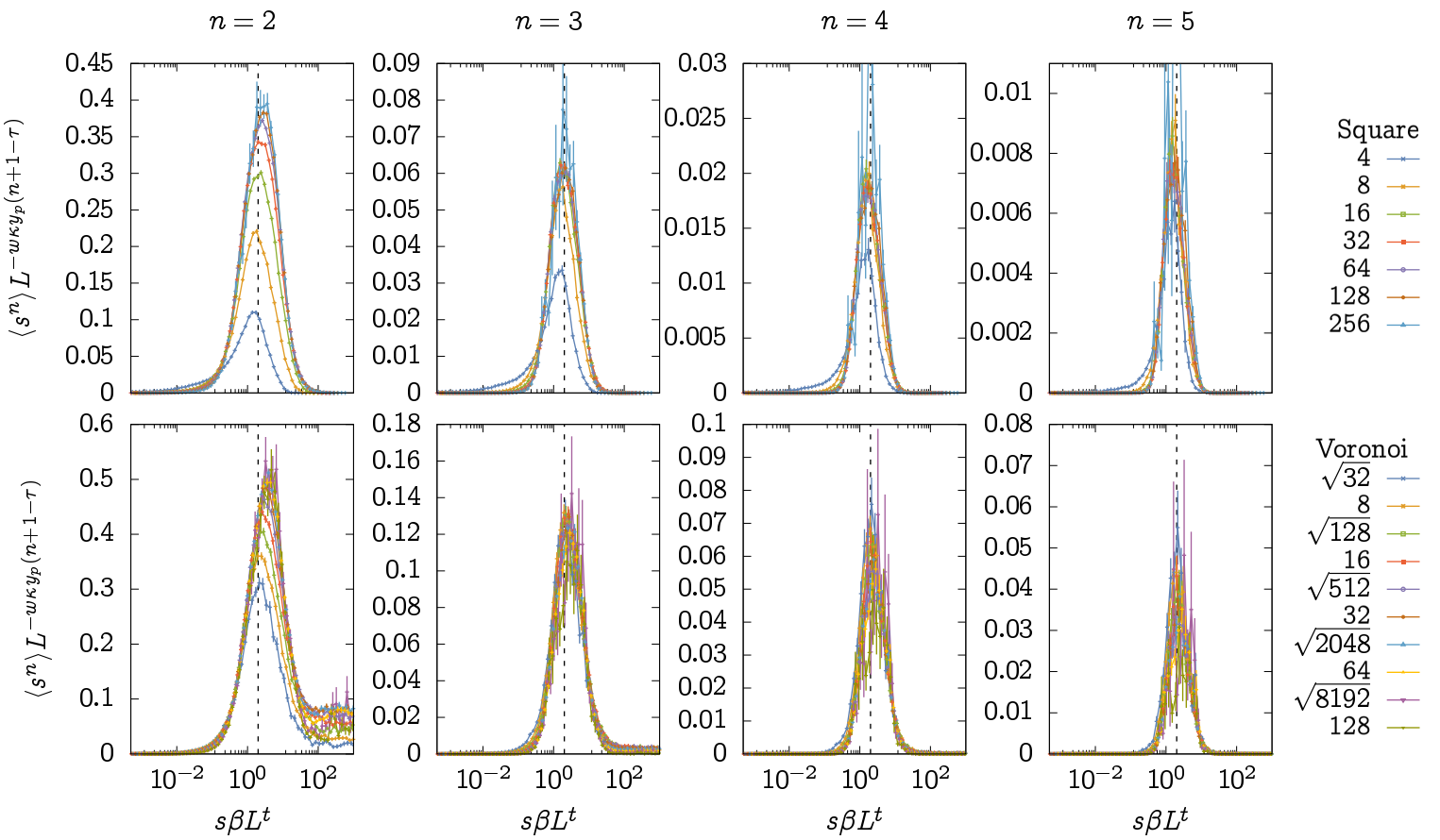

Figure 1.36: Rescaled moments of the precursor avalanche size distribution as a function of the scale invariant first introduced in §9.

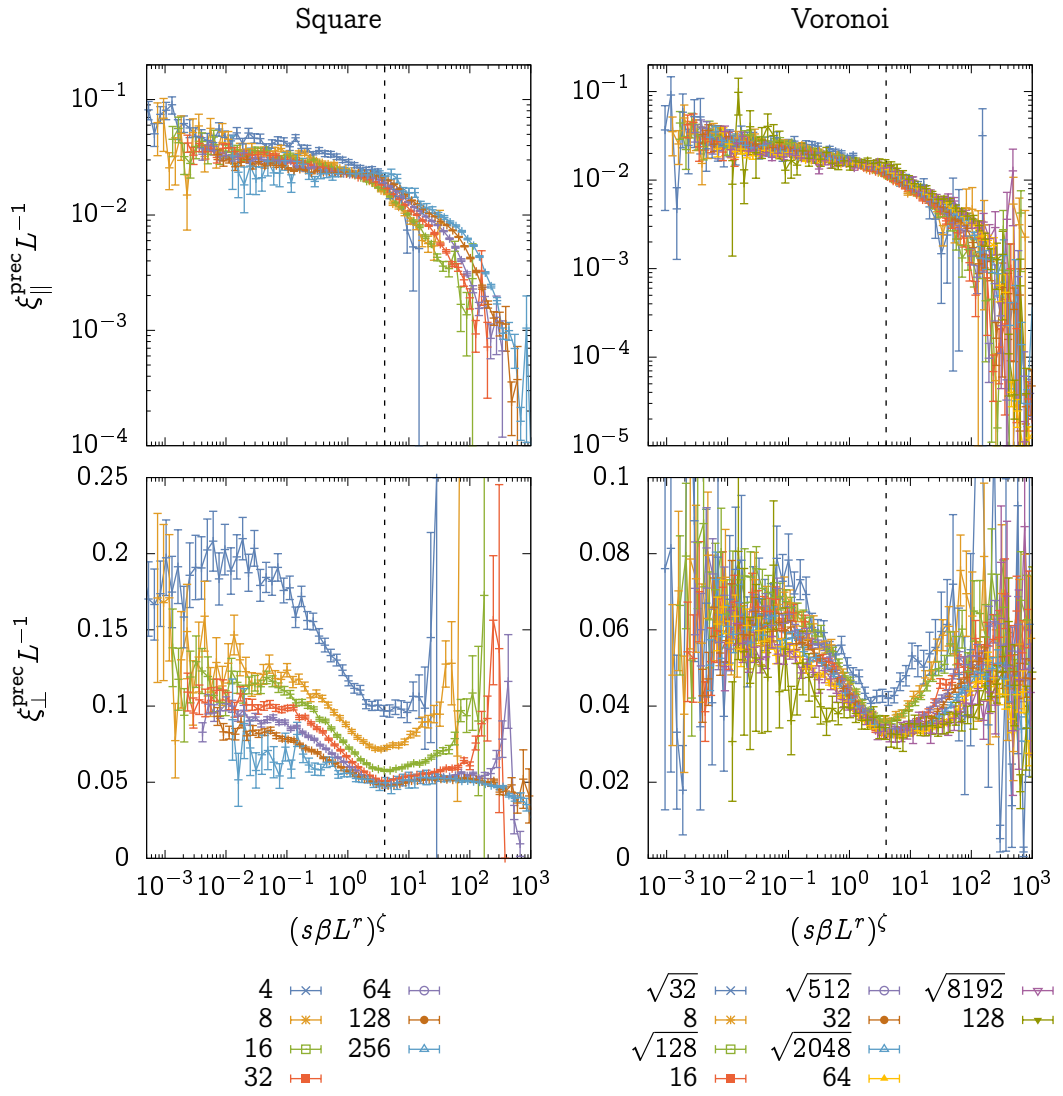

Figure 1.37: Correlation lengths for precursor avalanches as a function of the effective scale invariant fit in §6. Recall that the parallel direction is the direction of applied stress, while the perpendicular direction is the direction of crack propagation.

βL^t. In this viewing, the phenomenology seems to look like a steady growth of the aspect ratio with *decreasing* β towards isotropy that is cut off at finite size by the onset of percolation phenomenology.

Besides the sizes of the precursor avalanches, we can also look at the distribution of the modulus drops that they arrive with. Every time an avalanche occurs and some number of bonds break, the modulus (or more literally the conductivity) of the fuse network goes down by some amount. These drops are more experimentally accessible than the avalanche sizes themselves. Examples of the distributions we see in fuse networks are shown in Figure 1.39 for several values of β. It appears that the distributions have a β-independent tail for large β, with a β-dependent beak and tail towards zero.

Right now we do not have any theory for the scaling of this quantity, but one could likely be formed from an ansatz between the average avalanche size at a given u_p to the average modulus drop that results. In any case, the data in Figure 1.40 collapses reasonably well with the empirical invariant βL^t and a linear dependence of n, as is usually expected for a power-law distributed quantity. The dashed line in that figure shows again the same position that was shown in Figure 1.36 of the avalanche sizes scaled with the same horizontal axis. The two peak in the same place, which is not surprising. Once again, this scaling suggests a universal form that is a function of βL^t but cutoff for small β by the percolation phenomena, in which there is a smallest possible modulus drop that decreases with the size of the system.

If we ever find ourselves wanting for properties of these avalanches to measure, their size is not their only feature thought to be universal: the *shape* of avalanches is also thought to be universal, and is easily measurable

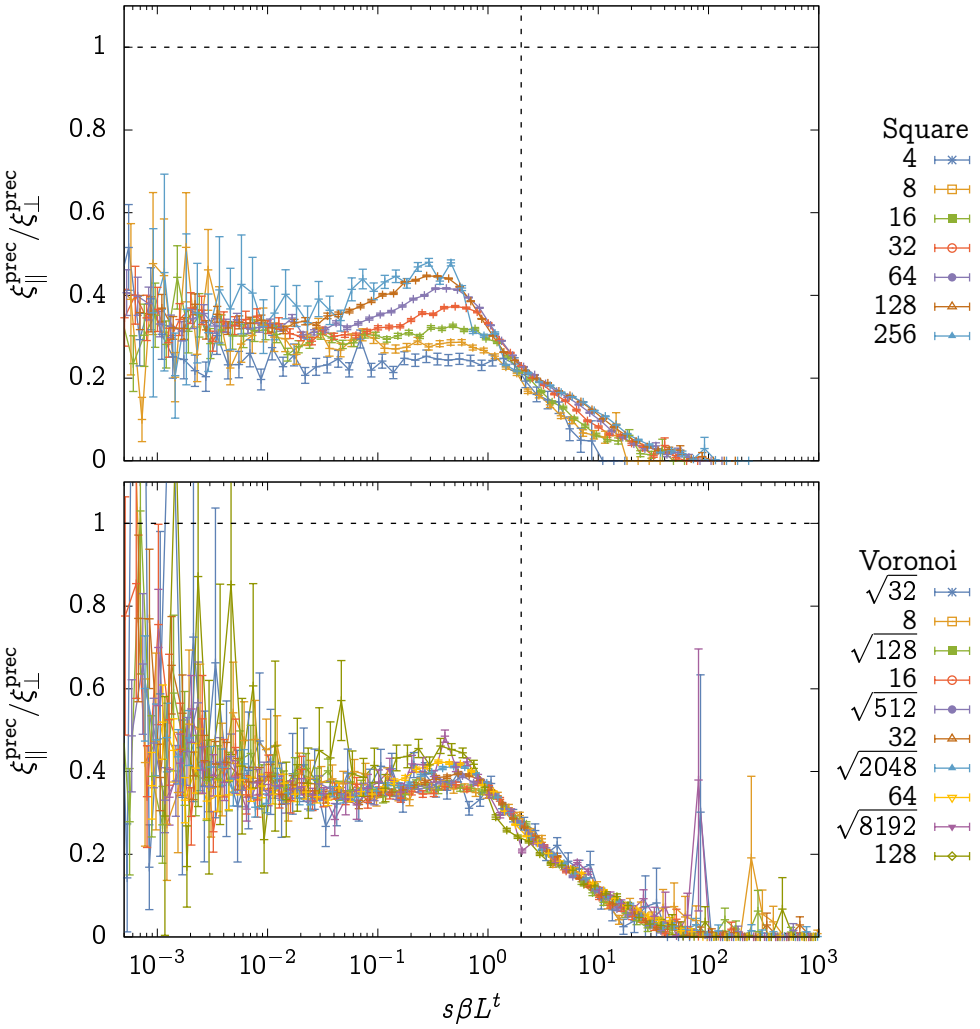

Figure 1.38: Aspect ratios of the precursor avalanches as a function of the invariant first introduced in §9. Recall that the parallel direction is the direction that stress is applied, while the perpendicular direction is that direction of crack propagation.

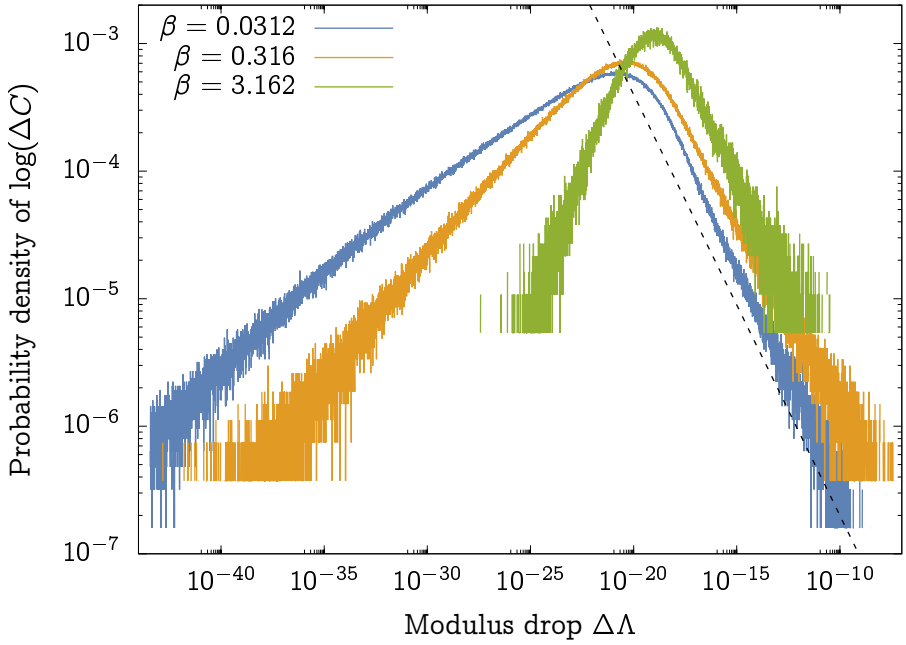

Figure 1.39: Examples of modulus drop distributions for a 2048-vertex voronoi lattice for various β. The dashed line shows $\Delta\Lambda^{-0.33}$, with no theoretical basis.

in the context of our simulations.[104–106]

§11. The final avalanche

When a little more than critical stress is applied, the system ruptures in a violent event with a large final avalanche. Unlike the precursor avalanches, the distribution of its size does not have a power-law tail, and instead has a well-defined typical size. This average size is plotted in Figure 1.42 as a function of the scale invariant fit in §6 for the cluster size distribution. For small β, the size of the final avalanche grows with a power law and is independent of

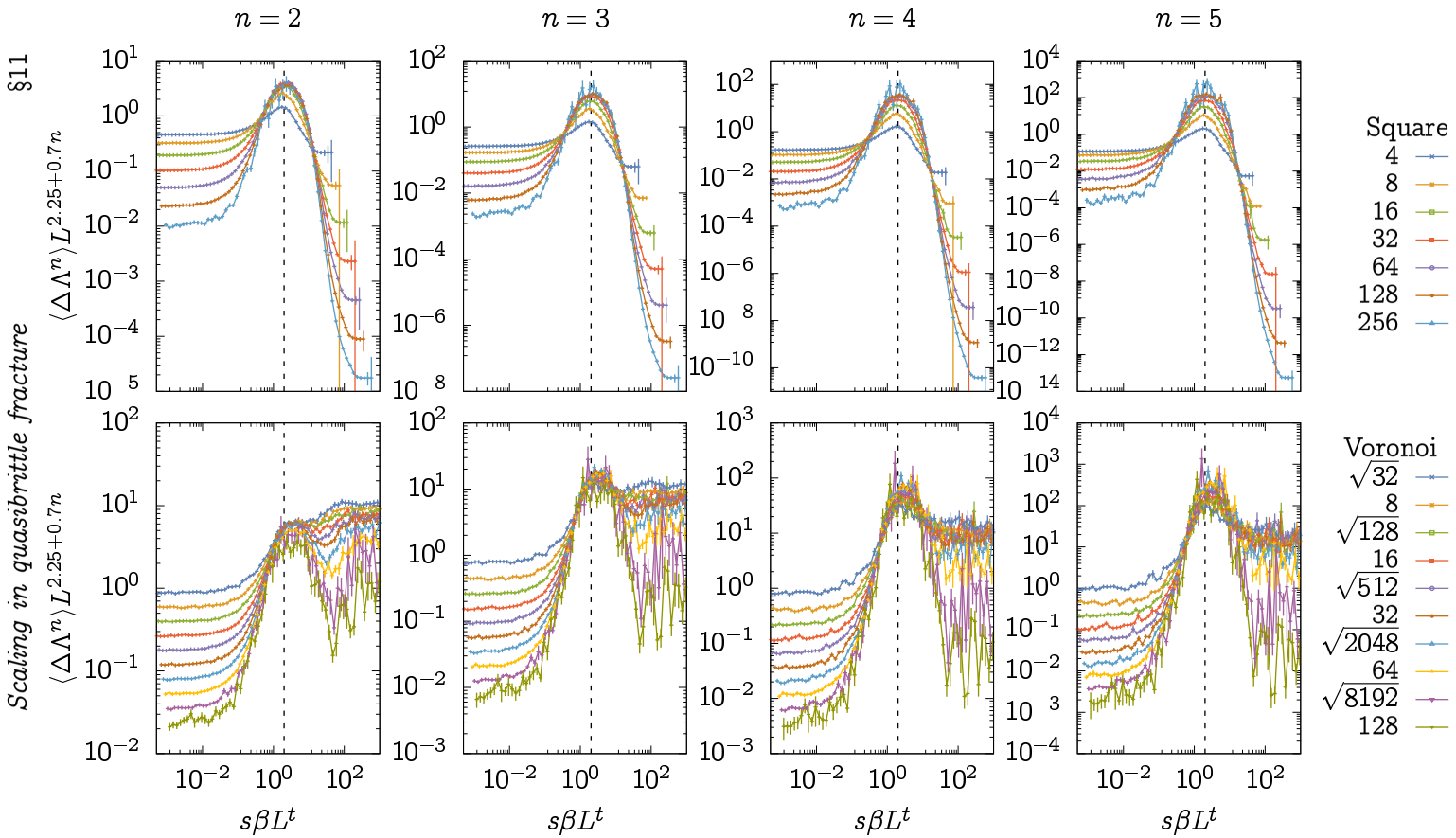

Figure 1.40: Rescaled moments of the precursor avalanche modulus drop distribution, rescaled by a heuristic factor linear in the moment n and as a function of the invariant first introduced in §9.

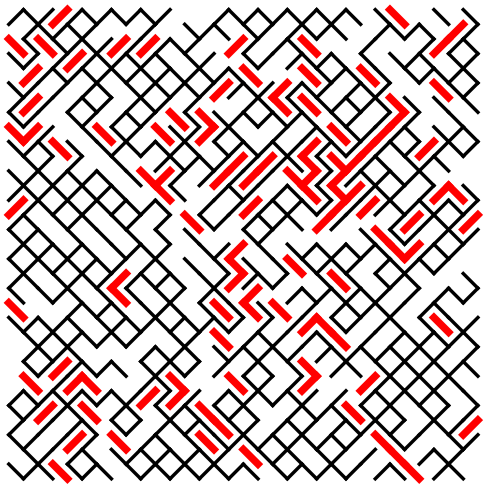

Figure 1.41: A 64×64 square lattice with $\beta = 0.316$, after rupture. Dual bonds whose partners were broken in the final avalanche are highlighted in red.

L. This then begins to grow faster, and has an L dependent peak for large β. One might wonder how the final avalanche at some intermediate β is larger than that at very large β, where the final avalanche is the only avalanche and is therefore system spanning. The answer is that the final avalanche in the intermediate β region is both system spanning and *wide*: it breaks many bonds to its sides as it moves. Look, for instance, at Figure 1.41, which shows the bonds broken by the final avalanche in a system at intermediate β. While some are systematically along the final crack surface, the damage is extremely diffuse.

This is seen strikingly in Figure 1.43, where the same average sizes are shown as a function of the invariant βL^t and rescaled by $L^{-1.5}$, and the peaks match well for sufficiently large system size. This suggests that the final avalanche grows with a fractal dimension ~ 1.5, which while subextensive is much faster than linear. As this final crack propagates through the system, it tears a wide path in this regime.

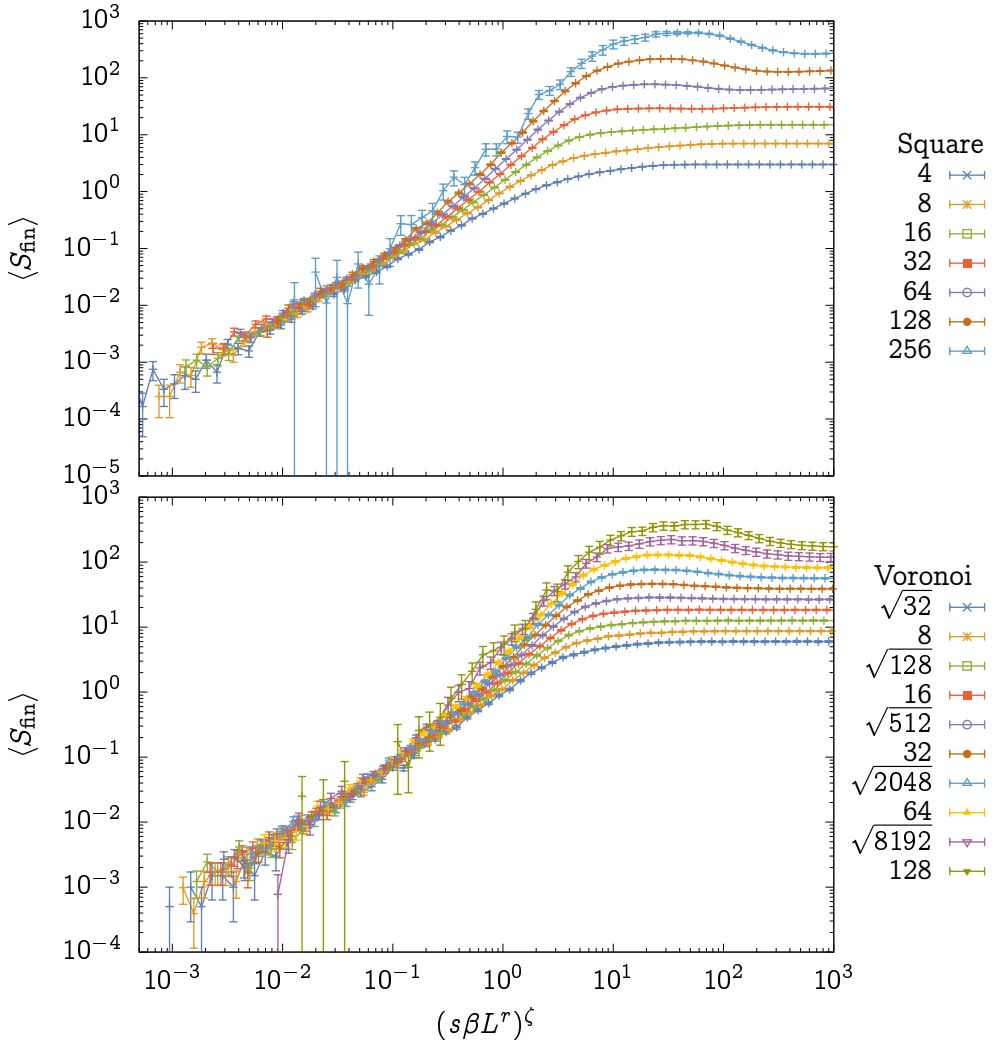

Figure 1.42: Average size of the final avalanche as a function of the scale invariant fit from the cluster size distribution in §6.

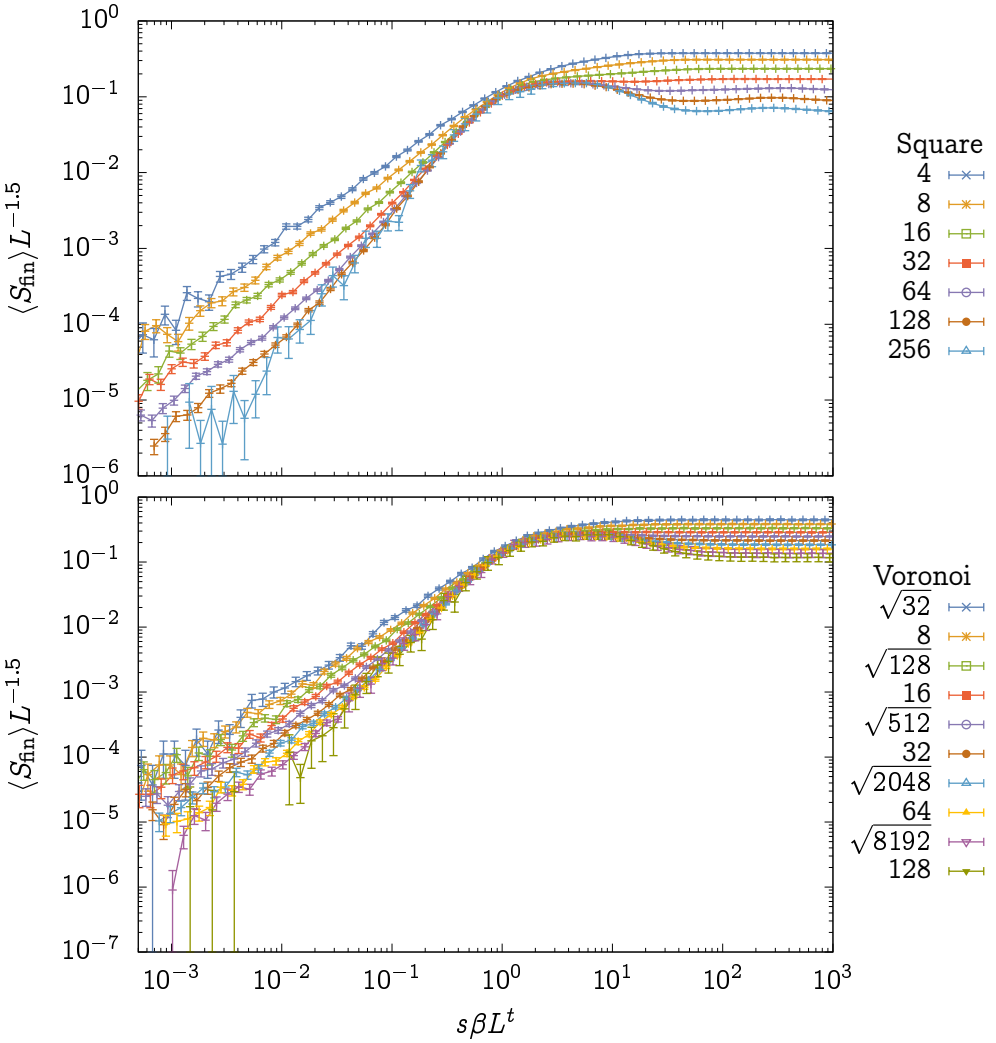

Figure 1.43: Average size of the final avalanche as a function of the scale invariant fit from the empirical invariant fit to the avalanche size distribution first introduced in §9.

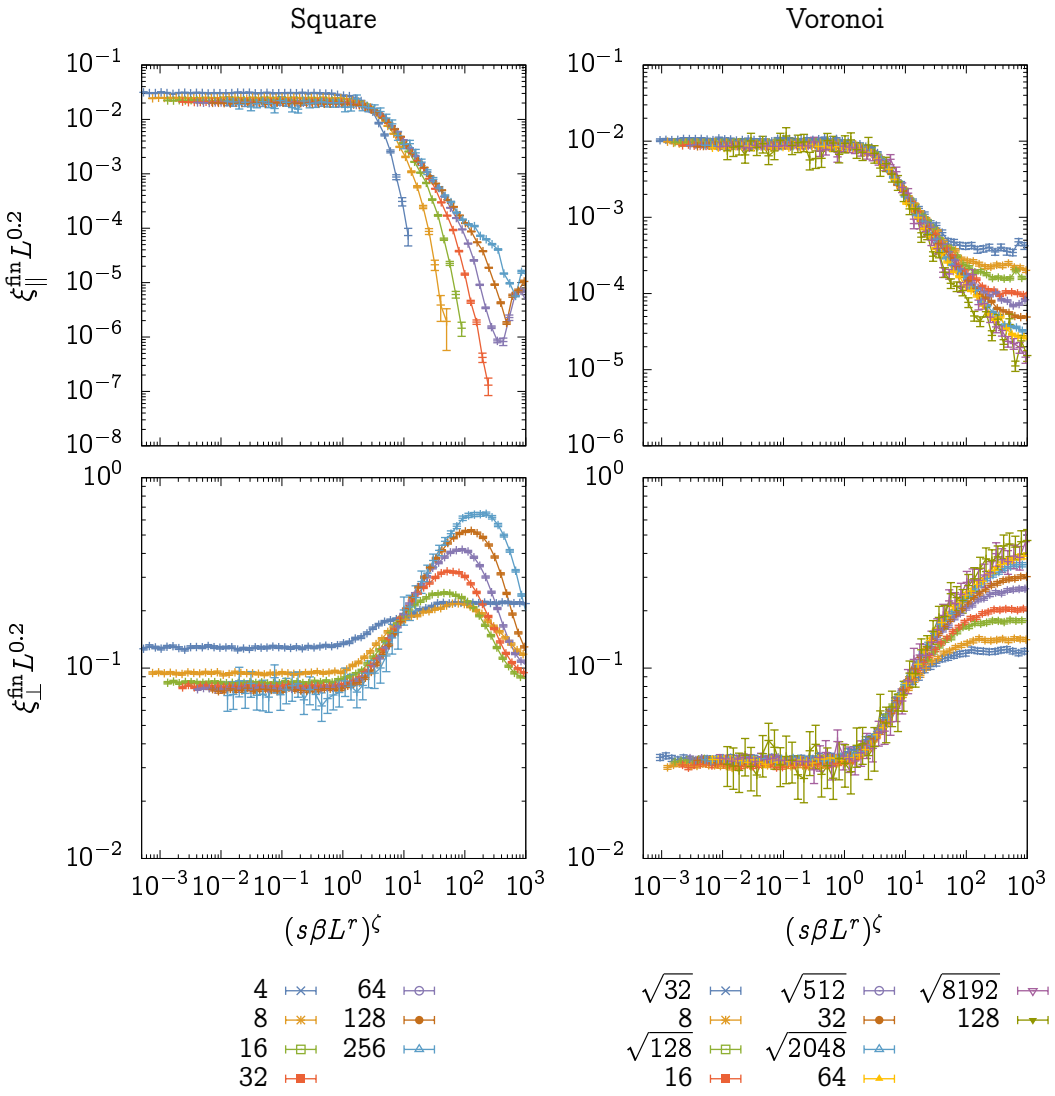

Figure 1.44: Correlation lengths for the final avalanche as a function of the invariant fit to the cluster size distribution in §6.

Interestingly, the wide avalanches we conclude must be happening in the large-β limit do not show up in the second-moment correlation lengths, shown in Figure 1.44. The "width," as measured this way, monotonically decreases with β from a maximum in the percolation limit with a crossover to a power law for intermediate β, with no local peak. Likewise, the "length" of the avalanche grows monotonically from a percolation value until cut off by the size of the lattice.* Both length scales are described well as functions of the invariant βL^r from the cluster size distribution fit in §6 alone.

The aspect ratio of the final avalanche shows more interesting behavior, having a peak at intermediate β that increases steadily with system size, suggesting an approach towards isotropy in the thermodynamic limit. This is shown in Figure 1.45, where this is plotted as a function of the invariant fit to the cluster size distribution in §6. A dashed line is drawn near its peak, which seems to be stably around $(s\beta L^r)^\zeta \simeq 2$. When plotted as a function of the invariant βL^r introduced in §9, the aspect ratio appears to have a universal form for *large* β which, like that for the precursor avalanches, seems to be headed towards isotropy at small β but is cut off by the emergence of the percolation phenomenology.

Finally, we can also examine the modulus drop that happens over the final avalanche. When scaled with the invariant fit to the cluster size distribution in §6, as in Figure 1.47, these are well-described by a function of that invariant up to a crossover. This isn't surprising, since the modulus drop over the final avalanche is precisely the modulus of the conducting backbone at crit-

*Or, in the case of the square lattice, is cut off by the strange, perfectly diagonal cracks that form and therefore no longer tend to run straight across the system.

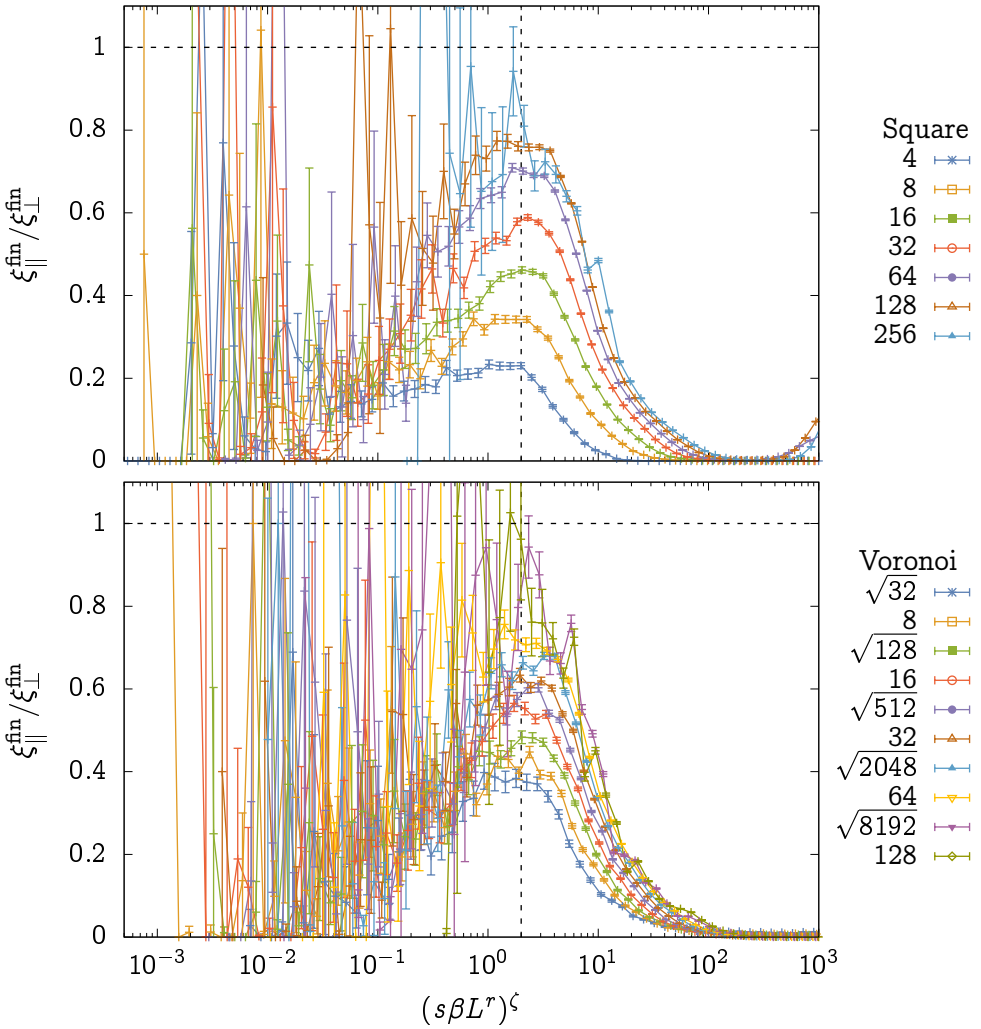

Figure 1.45: Aspect ratio of the final avalanche as a function of the invariant fit to the cluster size distribution in §6.

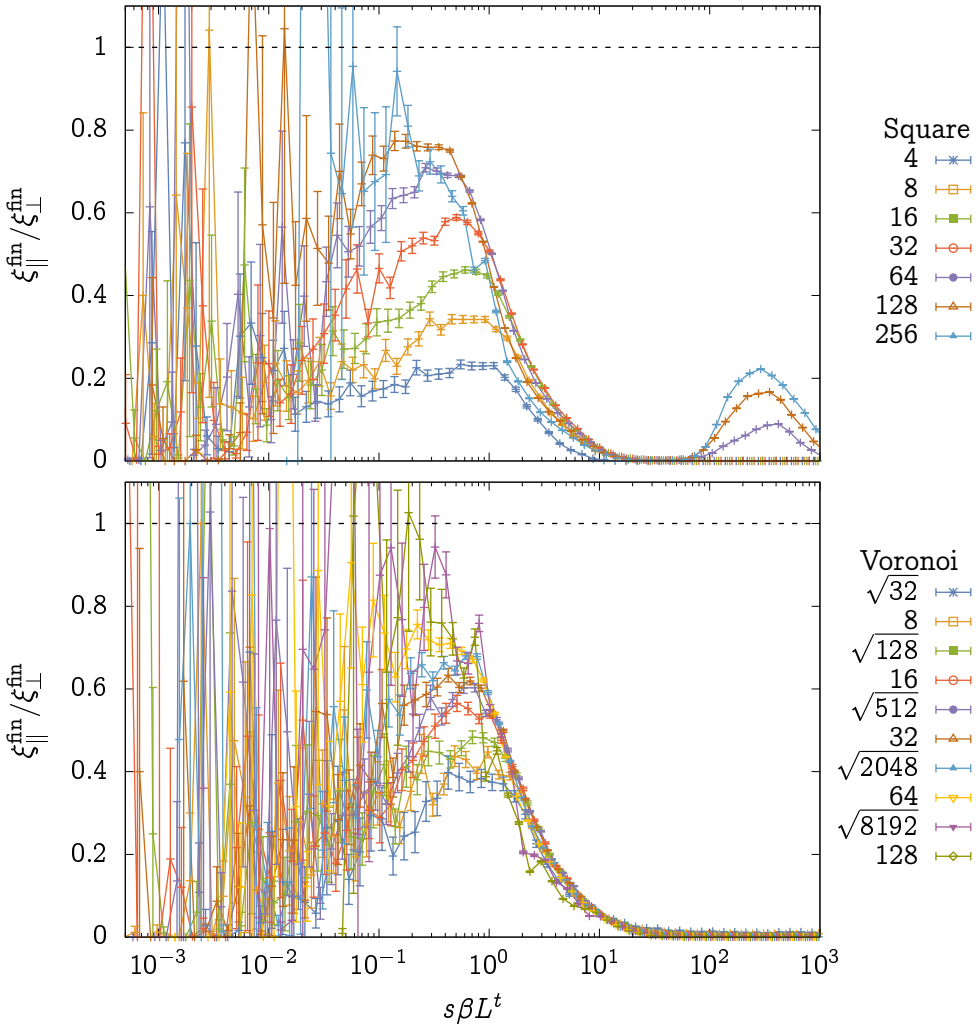

Figure 1.46: Aspect ratio of the final avalanche as a function of the invariant first introduced in §9.

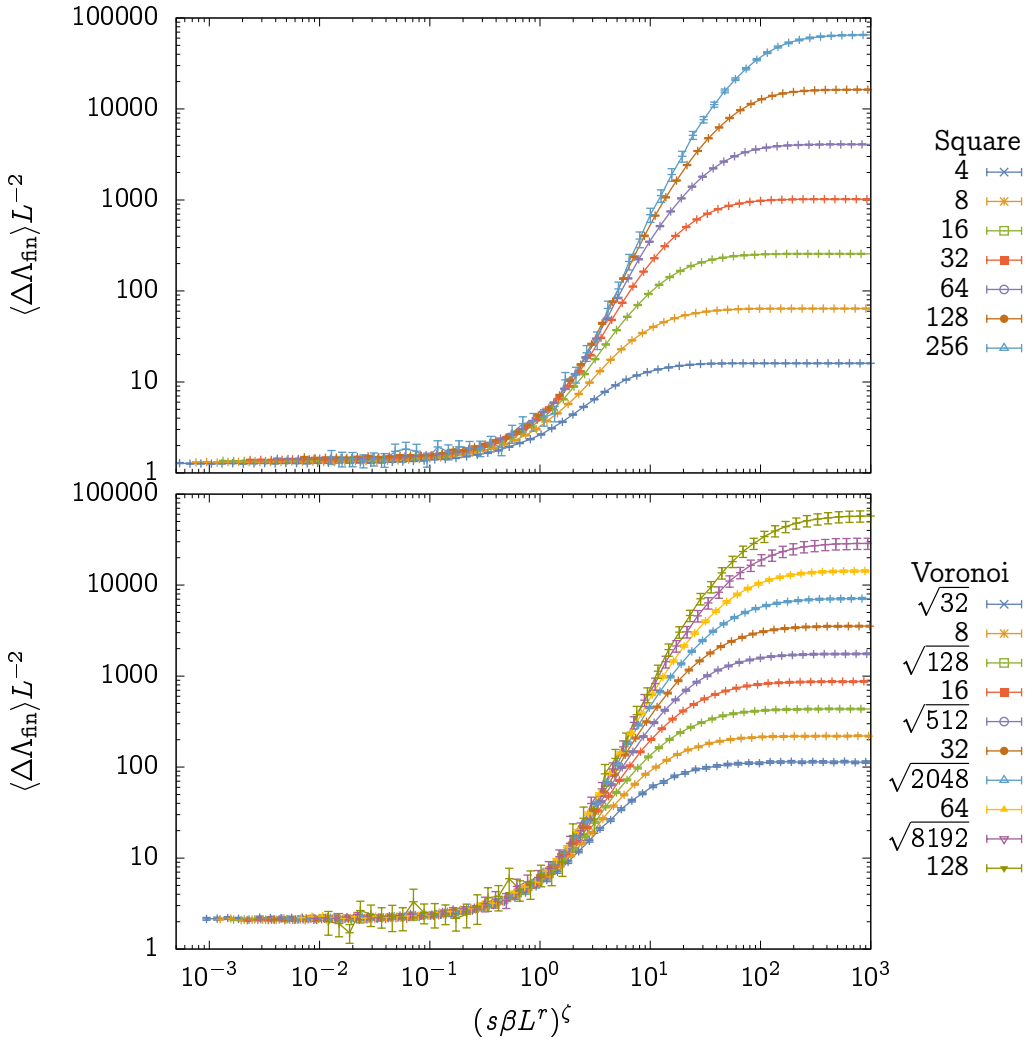

Figure 1.47: Modulus drop over the final avalanche as a function of the invariant fit to the cluster size distribution in §6.

ical stress, and therefore should be expected to scale with other percolation properties.

§12. Conclusions & steps forward

We have seen the skeleton of a scaling theory that can describe the crossover of an abrupt transition line from coinciding with a scaling field in the thermodynamic limit to acting as a nontrivial function of the relevant scaling fields at finite size. We have also seen a menagerie of properties of a numeric system purported to belong to the universality class described by the scaling theory. Some of those properties, namely most every static structural property associated with quantities that have analogues in the percolation problem, are very well described by a simple scaling picture like our Model A, in which we simply leave the percolation fixed point along a weird direction. Other properties, namely most every dynamic property associated with the form of the individual depletion steps in the limit of small disorder, are poorly described by such a picture, and instead seem to be governed by scaling with a different relationship between β and L than that predicted near the percolation fixed point.

How can we reconcile these phenomena? Models B and C are two attempts, but neither quite succeeds. Model B is clearly too simple a model of the fracture surface and doesn't predict the correct finite size scaling in the small β limit that is so clearly seen in the cluster size and p_{cf} data of §6 and §5. If a logarithmic singularity is present it likely needs an associated power law, which makes the analysis we use in the Model B subsection much more complicated.* Model C and perhaps more complex models like it offer some flexibility, as they predict that the scaling invariants along the fracture surface

*There is no commonly named function for the solutions to $y = x^{\alpha} \log^{\beta} x$.

should experience a crossover as one departs from the percolation fixed point. Since this prediction gives a crossover in the scaling invariants, any function of the scaling invariants should see the same crossover to dependence on βL^t instead of βL^r for large β. However, since the structural properties associated with percolation don't clearly exhibit such crossovers, it is difficult to imagine it could be responsible for the inconsistency.

How else to square this circle? Perhaps, we have been too focused on where the flow is coming from, and haven't thought enough about where it is going to. As mentioned while constructing the scaling theory, any abrupt transition is the result of a fixed point that acts like a separatrix for the two phases. This fixed point has its own properties, and its own scaling fields, which are not necessarily proportional to the scaling fields of the unstable critical fixed point we are flowing from. Take the example of the Ising model: below a critical fixed point at $T = T_c$, flow is separated either towards infinite or zero field by a separatrix at $T = 0$. Near the critical point, the external field has a flow equation

$$\frac{dH}{d\ell} = \frac{\beta\delta}{\nu} H + \cdots \tag{1.62}$$

which is linearized by the scaling field u_H with

$$\frac{du_H}{d\ell} = \frac{\beta\delta}{\nu} u_H \tag{1.63}$$

Near the low-temperature fixed point the external field has a flow equation

$$\frac{dH}{d\ell} = DH + \cdots \tag{1.64}$$

and the flow is linearized by the scaling field u'_H with

$$\frac{du'_H}{d\ell} = Du'_H \tag{1.65}$$

Clearly, u'_H and u_H cannot coincide, and while both are near-identity functions of H near their respective fixed points, they are not at all when you venture away.

Continuing the Ising metaphor, we now consider measuring a property of the system associated with the critical point, e.g., the correlation length ξ. This has a finite-size singular form

$$\xi = Lf(u_H u_t^{-\beta\delta}, u_H L^{\beta\delta/\nu}) \tag{1.66}$$

that describes its singular part for *all* T and H, up to irrelevant corrections to scaling. On the other hand, a property of the system associated with the zero-temperature fixed point, like the facet length—defined by the average length scale along which an interface is perfectly straight, and which diverges at zero temperature—will also have a singular scaling form, like

$$\xi_f = Lg(u'_H u_T^{\phi}, u'_H L^D) \tag{1.67}$$

Near the critical point, we don't expect this facet length to do anything inter-

esting, since the lattice that defines it is literally irrelevant. Its singular scaling from should be largely independent of the critical scaling. Therefore, it would be folly to attempt to develop a scaling theory where the facet length depends on the critical point invariant $u_H u_t^{-\beta\delta}$, since—while it might be technically true—the only region in which the facet length experiences singular scaling is a region where u_H is no longer linear in H and the connection between our control parameters and our scaling invariants is lost. This is epitomized at intermediate temperatures, where the facet length and the correlation length are both nonzero and noninfinite, and where each should have a reasonable description in terms of the fixed points they are respectively singular at. One would expect different finite-size scaling in T for each!

The traditional crossover picture, in which the scaling variables of the unstable fixed point *do* well-describe the behavior crossing over to the stable one, works because one is usually describing quantities that are singular at both fixed points, e.g., the susceptibility or correlation length, and so therefore must also be singular along the entire line connecting them.[107,108] But if a quantity *isn't* singular at both fixed points, and it mustn't necessarily be singular along the line connecting them, then the argument falls apart, because the behavior of the quantity as a function of the scale invariants of the fixed point that it is not singular at can be explained by analytic corrections to its behavior as a function of the scale invariants of the fixed point it is singular at. Analytic corrections aren't a part of universal scaling functions!

This metaphor is a long example that is meant to reflect on our thinking about avalanches in this problem. Near percolation, what avalanches exist are a sort of lattice effect and their behavior is not singular. Far enough along

the fracture surface away from percolation, however, they become system-spanning and are power-law distributed, clear singular behavior. Are the avalanches and their properties a kind of facet-length to the percolation problem? If this were the case, then what we are measuring with, e.g., exponents like t are in fact universal properties of the *fracture* fixed point that the critical surface is flowing towards. This is consistent with the fact that we would see different finite-size scaling for those properties in the same region we see percolation-like finite-size scaling for structural properties. Unlike the facet length, the intermediate β properties of the avalanches appear to be lattice-independent, which suggests there is a universal fracture fixed point with a basin of attraction that our fuse networks come into the vicinity of, before our large-β limit sees behavior in both lattices cross over to nonuniversal behavior.

One might have expected to see a crossover from isotropic percolation to something self-affine, in which the crack is described by a fractal dimension that is self similar under different rescales parallel and perpendicular to its direction. This is well-understood to be the coarse-grained geometry of cracks as review in §9. In that case we may have seen something similar to the crossover from isotropic to directed percolation,[109–111] which is even known to have associated avalanche phenomena.[112] The treatment of truly self-affine scaling cannot be done with ordinary finite-size methods, since correlation functions along different directions scale with different exponents.[113,114]

This is all, right now, speculation, and more work needs to be done fitting various models of the critical surface to rule it out as a possibility entirely. Now that we have a better sense of the space of theories and what looks interesting or is unexplained, taking more fine-grained data through the fracture process

as a function of the external current I or Δp (as in [21]) would provide a lot more constraint on the form of possible scaling theories, since we would then have two knobs to play with and could directly measure, e.g., I as a function of Δp and the shape of the fracture transition surface near the critical point. This work is far from over.

We are also looking to begin extending these ideas from describing the static spatial properties of the completed crack to describing the dynamic spatial properties of the growing crack. Of the data we have seen, those describing the final avalanche come closest, but likely are strongly governed by finite-size and boundary effects.* To this end, we are developing an coarse-grained model of the fracture process in high disorder, inspired by this data, to look into what conditions are required for cracks in these materials to propagate versus stand still and grow fat, similar to [115] but for quasibrittle systems that suffer from diffuse damage that modifies the local modulus. That work has only barely begun.

*A torus has no boundary, but it does let the crack interact with its other end as it propagates...

CHAPTER II

CLUSTER MONTE CARLO WITH ON-SITE POTENTIALS

Lattice models are important tools for studying emergent and critical phenomena. Besides describing the form of crystals, admitting exact solutions or perturbative expansions, the discreteness of a lattice makes them amenable to simulation. Simulations of critical phenomena are challenging, however, because they must contend with *critical slowing down*.[117] As length scales diverge near a critical point, dynamic scaling dictates that so too must time scales. In nature this slowness can make fluctuations observable; on a computer it can make them interminable. The autocorrelation time τ is predicted to diverge with system size L like L^z, and 2D Ising models have $z = 4 - \eta = \frac{15}{4} = 3.75$ with conserved dynamics[118] and $z \simeq 2.1767$ without.[119] With the length of a computational time step increasing like L^D,[†] this can mean the difference between enough system-size range to measure something or not.

[*] Portions of the material of this chapter were published in J Kent-Dobias and JP Sethna, "Cluster representations and the Wolff algorithm in arbitrary external fields", Physical Review E **98**, 10.1103/physreve.98.063306 (2018).

[†] After all, while we must update each site in our system one by one, nature is the ultimate parallelizer and handily updates every site simultaneously! The system size dependent timescales of computer simulations therefore differ from those of natural systems by a factor of the number of sites.

Thankfully there are methods for alleviating this. One of the most useful is the cluster method, which is the basis for both the celebrated Swendsen–Wang[120] and Wolff algorithms.[121] They use unphysical nonlocal updates[122] whose clusters undergo a percolation transition at the critical point of the system.[123] Moving beyond the constraint of realistic local dynamics results in relatively small dynamic exponents for many spin systems,[119,124–126] including the Ising, O(n),[121] Potts,[120,127] and other models.[128–132] For the 2D Ising model they give $z \simeq 0.3$,[119] a remarkable improvement in speed. Unfortunately these methods have limited applicability; in their original form they cannot even deal with the application of a simple magnetic field. These algorithms rely on the natural invariance of the systems in question under symmetry transformations on their spins.

Some success has been made in extending these algorithms to systems in certain external fields by adding a 'ghost site'[133] that returns global rotation invariance to spin Hamiltonians at the cost of an extra degree of freedom, allowing the method to be used in a subcategory of interesting fields.[134–136] Static fields have also been applied by including a separate metropolis or heat bath update step after cluster formation,[137–139] and other categories of fields have been applied using replica methods.[140–142] Monte Carlo techniques that involve cluster updates at fixed magnetization have been used to examine quantities at fixed field by later integrating measured thermodynamic functions.[143,144]

Here, I will discuss a useful generalization of these methods applicable to a broad class of lattice models with both pair interactions and arbitrary on-site potentials.[116] We show that the scaling of correlation time near the critical

point of several models suggests that the 'ghost' approach is a natural one, e.g., that it extends the celebrated scaling of dynamics in these algorithms at zero field to various non-symmetric perturbations. We also show, by a redefinition of the spin–spin coupling in a generic class of spin systems, *arbitrary* external fields can be treated using cluster methods. Rather than the introduction of a 'ghost spin,' our representation relies on introducing a 'ghost transformation,' an extra degree of freedom residing on a 'ghost' site coupled to all other sites that takes its values from the collection of spin symmetry transformations of the base model rather than resemble the base spins themselves.

We provide an open-source implementation of this method in the form of a C++ library, available at `https://git.kent-dobias.com/wolff/`. Use of this library will be described briefly within, but extensive documentation is also available at `https://doc.kent-dobias.com/wolff/`.

We will first briefly review the formal basis of cluster methods, which rests in a duality between the Ising model and a strange model on graphs. We will then see how this duality can be extended to one between many more complicated models and ensembles of strange graph models. We will then introduce the Wolff algorithm for a general model, and see the small conceptual tweak that allows its use with arbitrary on-site potentials. We'll then see this borne out in several canonical models, and gauge its efficiency in those cases. Finally, we'll see some direct applications: measuring the relevance of lattice anisotropies on the two-dimensional XY model, and directly measuring the value of scaling functions in the metastable state of the Ising model.

§13. Duality in Markov-chain Monte Carlo

Equilibrium statistical models are usually expressed in the form of an unnormalized probability distribution f over phase space Ω parameterized by values x. For instance, the Boltzmann weight for a system with Hamiltonian H and parameter given by inverse temperature β is

$$f(s \mid \beta) = e^{-\beta H(s)} \tag{2.1}$$

The properties of such a model are summed up in its partition function

$$Z(x) = \mathrm{Tr}_s\, f(s \mid x) \tag{2.2}$$

which is also the normalization of f, so that

$$P(s \mid x) = \frac{f(s \mid x)}{Z(x)} \tag{2.3}$$

is a proper probability distribution on Ω. Most properties of interest can be expressed as various derivatives of the partition function Z, and so its form and dependence on the model parameters is in many ways defining of a model.

A *duality* exists between two models f_a and f_b on phase spaces Ω_a and Ω_b when there exists a joint model f_{ab} on $\Omega_a \times \Omega_b$ with

$$\begin{aligned} f_a(s_a \mid x) &= \mathrm{Tr}_{s_b}\, f_{ab}(s_a, s_b \mid x) \\ f_b(s_b \mid x) &= \mathrm{Tr}_{s_a}\, f_{ab}(s_a, s_b \mid x) \end{aligned} \tag{2.4}$$

[†]Many of the ideas in this section are inspired by HG Evertz, HM Erkinger, and W von der Linden, "New cluster method for the Ising model", in *Springer proceedings in physics* (Springer Berlin Heidelberg, 2002), pp. 123–133.

This implies a close connection between the partition functions of the two models; as a function of the parameters their partition functions must have the same singularities, since

$$Z_a(x) = \operatorname{Tr}_a f_a(s_a \mid x) = \operatorname{Tr}_a \operatorname{Tr}_b f_{ab}(s_a, s_b \mid x) = \operatorname{Tr}_b f_b(s_b \mid x) = Z_b(x) \quad (2.5)$$

Posing the duality here as being characterized by a joint model may seem unusual, since ordinarily it is the correspondence between the two partition functions that is of interest, and any temporary expressions of both models' variables in a joint trace is just a means to that end. Here, though, the fact that both models have a joint expression is something we will exploit, and in particular it induces conditional probabilities

$$P(s_a \mid s_b, x) = \frac{f_{ab}(s_a, s_b \mid x)}{f_b(s_b \mid x)} \qquad P(s_b \mid s_a, x) = \frac{f_{ab}(s_a, s_b \mid x)}{f_a(s_a \mid x)} \qquad (2.6)$$

that—if they end up being simply expressed functions of the microstates—will allow us to switch back and forth rapidly between the two representations.[145]

Fortuin–Kasteleyn

Cluster methods are based on a representation of the Ising model developed by Fortuin and Kasteleyn,[146–148] and the connection between them was elucidated by Edwards and Sokal.[149] The Ising model in the absence of a field, which we shall see many more times in this thesis, is given on a graph

$G = (V, E)$ by the unnormalized distribution

$$f_{\text{Ising}}(s \mid \beta) = \exp\Big[\beta \sum_{\{i,j\}\in E} J_{ij} s_i s_j\Big] \tag{2.7}$$

over sets s of spins that take values of one or minus one and reside on the vertices of the graph. The Fortuin–Kasteleyn random graph model is given on the same graph by the distribution

$$f_{\text{FK}}(b \mid p) = 2^{N_c(b)} \prod_{\{i,j\}\in E} (1 - p_{ij})^{\frac{1}{2}} \left(\frac{p_{ij}}{1 - p_{ij}}\right)^{b_{ij}} \tag{2.8}$$

over sets b of binary values (zero or one) that reside in the edges of the graph, and where $N_c(b)$ gives the total number of connected clusters of active (value one) bonds defined by the configuration b.

We will show that the joint model defined by the unnormalized distribution

$$f_{\text{Ising/FK}}(s, b \mid \beta) = \prod_{\{i,j\}\in E} (1-p_{ij})^{-\frac{1}{2}} \left[\delta_{b_{ij},0}(1-p_{ij}) + \delta_{b_{ij},1}\delta_{J_{ij}s_is_j>0} p_{ij}\right] \tag{2.9}$$

has the property (2.4) if the model parameters J and p are taken to relate to each other by

$$p_{ij} = \begin{cases} 1 - e^{-2\beta J_{ij}} & J_{ij} > 0 \\ 1 - e^{+2\beta J_{ij}} & J_{ij} < 0 \end{cases} \tag{2.10}$$

It is sufficient to show that the partial trace over each configuration subspace is zero. First, tracing over the bond degrees of freedom yields the Ising model:

$$\begin{aligned}\mathrm{Tr}_b\, f_{\mathrm{Ising/FK}}(s,b\mid\beta) &= \mathrm{Tr}_b \prod_{\{i,j\}\in E}\left[\delta_{b_{ij},0}e^{-\beta J_{ij}} + \delta_{b_{ij},1}\delta_{s_i,s_j}(e^{\beta J_{ij}} - e^{-\beta J_{ij}})\right] \\ &= \prod_{\{i,j\}\in E}\sum_{b_{ij}=\{0,1\}}\left[\delta_{b_{ij},0}e^{-\beta J_{ij}} + \delta_{b_{ij},1}\delta_{s_i,s_j}(e^{\beta J_{ij}} - e^{-\beta J_{ij}})\right] \\ &= \prod_{\{i,j\}\in E}\left[e^{-\beta J_{ij}} + \delta_{s_i,s_j}(e^{\beta J_{ij}} - e^{-\beta J_{ij}})\right] = f_{\mathrm{Ising}}(s\mid\beta)\end{aligned} \tag{2.11}$$

since the identity

$$e^{\beta J_{ij}s_is_j} = e^{-\beta J_{ij}} + \delta_{s_i,s_j}(e^{\beta J_{ij}} - e^{-\beta J_{ij}}) \tag{2.12}$$

holds for $s_i, s_j = \pm 1$. The other direction is a little more complicated. For a given configuration b, the product is only nonzero if for every $\{i,j\} \in E$ with $b_{ij} = 1$, $s_i = s_j$. This means that a connected cluster of sites joined together by bonds with $b_{ij} = 1$ throughout must have $s_i = s_j$ through the entire cluster they form in order to have a nonzero product. Therefore, there are two distinct states per such cluster of sites connected by activated bonds (all up and all down in each cluster) in each contributing term in the trace, or $2^{N_c(b)}$. This leaves

$$\begin{aligned}\mathrm{Tr}_s\, f_{\mathrm{Ising/FK}}(s,b\mid\beta) &= 2^{N_c(b)}\prod_{\{i,j\}\in E}(1-p_{ij})^{-\frac{1}{2}}(1-p_{ij})^{1-b_{ij}}p_{ij}^{b_{ij}} \\ &= 2^{N_c(b)}\prod_{\{i,j\}\in E}(1-p_{ij})^{\frac{1}{2}}\left(\frac{p_{ij}}{1-p_{ij}}\right)^{b_{ij}}\end{aligned} \tag{2.13}$$

The models are therefore dual.

This is a neat and ancient exercise, but what does it have to do with lattice models and Monte Carlo? The useful insight comes when one writes down the joint distributions that result. From examination of the form of the joint model, one can see that

$$P(s \mid b, x) = \prod_{C \in \mathcal{C}(b)} \begin{cases} \frac{1}{2} & J_{ij} s_i s_j > 0 \text{ for all } b_{ij} \in C \\ 0 & \text{otherwise} \end{cases} \tag{2.14}$$

where $\mathcal{C}(b)$ is the set of clusters associated with a configuration b, and

$$P(b \mid s, x) = \prod_{ij} \begin{cases} p_{ij} & b_{ij} = 1 \text{ and } J_{ij} s_i s_j > 0 \\ 0 & b_{ij} = 1 \text{ and } J_{ij} s_i s_j < 0 \\ 1 - p_{ij} & b_{ij} = 0 \text{ and } J_{ij} s_i s_j > 0 \\ 1 & b_{ij} = 0 \text{ and } J_{ij} s_i s_j < 0 \end{cases} \tag{2.15}$$

Because both marginal distributions factor into distributions on individual clusters and bonds, respectively, and both distributions are nontrivial, i.e., they aren't delta functions on a single configuration, we now have our Monte Carlo strategy: take a configuration of one model, then sample a configuration of the other model, then repeat back and forth. A configuration of bonds is sampled from one of sites by setting b_{ij} to one with probability p_{ij} if $J_{ij} s_i s_j > 0$ and to zero otherwise. A configuration of sites is sampled from one of bonds by setting s_i up or down with equal probability, so long as it is set that way along with every other site in the cluster connected by activated bonds it belongs to.

Adding a field: a preview

Notice that the procedure followed to arrive at a cluster representation of the Ising model is independent of the structure of the lattice, or even the presence of a lattice: the coupling J can be *anything*, so long as it is a pair coupling between objects with two states, symmetric under the inversion of both. We will exploit this generality to add an external field by defining a strange lattice.

First, the more general form of our classic Ising model also allows for the addition of a field:

$$f_{\text{Ising}}(s \mid \beta, h) = \exp\left[\beta \sum_{\{i,j\}\in E} J_{ij} s_i s_j + \sum_i h_i s_i\right] \tag{2.16}$$

Though this seems to have spoiled our construction in the previous subsection, we can recover it at the cost of adding another site with a spin s_0 to the lattice. We have

$$\begin{aligned} \tilde{f}_{\text{Ising}}(s_0, s \mid \beta, h) &= \exp\left[\beta \sum_{ij} J_{ij} s_i s_j + h s_0 \sum_i s_i\right] \\ &= \exp\left[\beta \sum_{ij} \tilde{J}_{ij} s_i s_j\right] \end{aligned} \tag{2.17}$$

for $\tilde{J}_{ij} = J_{ij}$ for $i, j \geq 1$ and $\tilde{J}_{0i} = \tilde{J}_{i0} = h$. This new model has two important features: it has the form of (2.7) and expectation values of the distribution (2.16) can be computed. Indeed, any observable A of the original model can be written as an observable $\tilde{A}(s_0, s) = A(s_0 s)$ of the new model such that $\langle \tilde{A} \rangle = \langle A \rangle$, as we shall see in detail later. This representation and

its connection to the Ising model in a field was noticed in the original paper introducing the random cluster model,[146] and later drawn out in the context of cluster methods.[133,135]

General lattice models

The specific connection between the Ising and random-cluster models is useful far beyond the context of either model. This is because the Ising model is so simple: it describes any theory of objects under a pairwise coupling that can take one of two states, and that has a global symmetry when everyone is brought from one state to the other. Up to isomorphism, $\mathbb{Z}/2\mathbb{Z}$ is the only symmetry group with two elements, and any element with an order-two orbit inside another symmetry group is isomorphic to it. This means that the Ising model can be embedded inside other models that contain this subgroup, and clusters can be used to rapidly sample new states exactly as above.

This embedding scheme also opens a new degree of freedom in the construction of a cluster algorithm: the distribution of embeddings one uses. As long as this distribution satisfies some loose conditions covered later, a wide variety of cluster-forming choices lead to detailed balance. This freedom will be especially important in the next chapter.

Since the formal correspondence between models and dual ensembles of Fortuin–Kasteleyn graph models becomes complicated and tedious with diminishing conceptual returns (though some write it out[150]), we will move onto an alternative framing for the rest of this chapter which focuses on the way the method arises in practice, dynamically in cluster construction. The new feature introduced here is the extension of the field-as-lattice-site above to these

general models by introducing a field-as-lattice-site with a sufficiently large configuration space to capture the composition of flips over many different Ising embeddings.

§14. Clusters without a field

We will pose the problem in a general way, but several specific examples can be found in Table 2.1 for concreteness. Let $G = (V, E)$ be a graph, where the set of vertices $V = \{1, \dots, N\}$ enumerates the sites of a lattice and the set of edges E contains pairs of neighboring sites. Let R be a group acting on a set X, with the action of group elements $r \in R$ on elements $s \in X$ denoted $r \cdot s$. X is the set of states accessible by each spin, and R is the *symmetry group* of X. The set X must admit a measure μ that is invariant under the action of R, e.g., for any $A \subseteq X$ and $r \in R$, $\mu(r \cdot A) = \mu(A)$. This trait is shared by the counting measure on any discrete set, or by any group acting by isometries on a Riemannian manifold, such as $\mathrm{O}(n)$ on S^{n-1} in the $\mathrm{O}(n)$ models.[150] Finally, a subset R_2 of elements in R of order two must act transitively on X. This property, while apparently obscure, is shared by any symmetric space[151] or by any transitive, finitely generated isometry group. In fact, all the examples listed here have spin spaces with natural metrics whose symmetry group is their set of isometries. We put one spin at each site of the lattice described by G, so that the state of the entire system is described by elements $\mathbf{s} \in X \times \cdots \times X = X^N$.

The Hamiltonian of this system is a function $\mathcal{H} : X^N \to \mathbb{R}$ defined by

$$\mathcal{H}(\mathbf{s}) = -\sum_{\{i,j\}\in E} Z(s_i, s_j) - \sum_{i\in V} B(s_i), \tag{2.18}$$

where $Z : X \times X \to \mathbb{R}$ couples adjacent spins and $B : X \to \mathbb{R}$ is an external field. Z must be symmetric in its arguments and invariant under the action of any element of R applied to the entire lattice, that is, for any $r \in R$ and $s, t \in X$, $Z(r \cdot s, r \cdot t) = Z(s, t)$. One may also allow Z to also be a function of edge—for modelling random-bond, long-range, or anisotropic interactions—or allow B to be a function of site—for applying arbitrary boundary conditions or modelling random fields. The formal results of this chapter (that the algorithm obeys detailed balance and ergodicity) hold equally well for these cases, but we will drop the additional index notation for clarity. Some extensions, like adding strong random fields or bonds, ultimately prove inefficient,[140,152] though random, frustrated bonds have proven amenable to a cluster approach.[153] The approach should also work fine for randomness added perturbatively to nonrandom models.

Implementation of a model in the provided library is as simple as defining a class that represents an element of the state space X, with default constructor (and destructor, if necessary), and a class that represents an element of the group R, with default constructor and member functions that define the action and inverse action of the class on both states and group elements. Specific details may be found at `https://doc.kent-dobias.com/wolff/models.html`.

The goal of statistical mechanics is to compute expectation values of observables $A : X^N \to \mathbb{R}$. Assuming the ergodic hypothesis holds (for systems

Model	Spins (X)	Symmetry (R)	Action ($g \cdot s$)	Coupling ($Z(s,t)$)	Common Field ($B(s)$)
Ising	$\{-1, 1\}$	$\mathbb{Z}/2\mathbb{Z}$	$0 \cdot s \mapsto s$ $1 \cdot s \mapsto -s$	st	Hs
$\mathrm{O}(n)$	S^{n-1}	$\mathrm{O}(n)$	$M \cdot s \mapsto Ms$	$s^{\mathrm{T}}t$	$H^{\mathrm{T}}s$
Potts	$\{1, \dots, q\}$	S_n	$(i_1, \dots, i_q) \cdot s \mapsto i_s$	$\delta(s,t)$	$\sum_m H_m \delta(m,s)$
Clock	$\mathbb{Z}/q\mathbb{Z}$	D_n	$r_m \cdot s \mapsto m + s$ $s_m \cdot s \mapsto -m - s$	$\cos(2\pi\frac{s-t}{q})$	$\sum_m H_m \cos(2\pi\frac{s-m}{q})$
DGM	$\mathbb{Z}$	$D_{\inf}$	$r_m \cdot s \mapsto m + s$ $s_m \cdot s \mapsto -m - s$	$(s-t)^2$	Hs^2

Table 2.1: Several examples of spin systems and the symmetry groups that act on them. Common choices for the spin–spin coupling in these systems and their external fields are also given. Other fields are possible, of course: for instance, some are interested in modulated fields $H\cos(2\pi k\theta(s))$ for integer k and $\theta(s)$ giving the angle of s to some axis applied to the $\mathrm{O}(2)$ model [154]. All models listed here have example implementations in the provided C++ library.

with broken-symmetry states, it does not and these phase-space integrals need to be restricted to a subset), the expected value $\langle A \rangle$ of an observable A is its average over every state $\mathbf{s}$ in the configuration space X^N weighted by the Boltzmann probability of that state appearing, or

$$\langle A \rangle = \frac{\int_{X^N} A(\mathbf{s}) e^{-\beta \mathcal{H}(\mathbf{s})} \, d\mu(\mathbf{s})}{\int_{X^N} e^{-\beta \mathcal{H}(\mathbf{s})} \, d\mu(\mathbf{s})}, \tag{2.19}$$

where for $Y_1 \times \cdots \times Y_N = Y \subseteq X^N$ the product measure $\mu(Y) = \mu(Y_1) \cdots \mu(Y_N)$ is the simple extension of the measure on X to a measure on X^N. These values are estimated using Monte Carlo techniques by constructing a finite sequence of states $\{\mathbf{s}_1, \ldots, \mathbf{s}_M\}$ such that

$$\langle A \rangle \simeq \frac{1}{M} \sum_{i=1}^{M} A(\mathbf{s}_i). \tag{2.20}$$

Sufficient conditions for this average to converge to $\langle A \rangle$ as $M \to \infty$ are that the process that selects $\mathbf{s}_{i+1}$ given the previous states be Markovian (only depends on $\mathbf{s}_i$), ergodic (any state can be accessed), and obey detailed balance (the ratio of probabilities that $\mathbf{s}'$ follows $\mathbf{s}$ and vice versa is equal to the ratio of weights for $\mathbf{s}$ and $\mathbf{s}'$ in the ensemble).

Measurements of observables during Monte Carlo in the provided library are made by the use of hooks, which are member functions of a measurement class that are run at designated points during the algorithm's execution and are provided arbitrary information about the internal state of all relevant objects. A detailed description of these hooks can be found at `https://doc.kent-dobias.com/wolff/measurement.html`.

While any of several related cluster algorithms can be described for this system, we will focus on the Wolff algorithm [121]. In the absence of an external field, e.g., $B(s) = 0$, the Wolff algorithm proceeds in the following way.

Algorithm 1 Wolff

1. Pick a random site m_0 and add it to the stack.
2. Select a transformation $r \in R_2$ distributed by $f(r \mid m_0, \mathbf{s})$. Often f is taken as uniform on R_2, but it is sufficient for preserving detailed balance that f be any function of the seed site m_0 and $Z(s, r \cdot s)$ for all $s \in \mathbf{s}$. The flexibility offered by the choice of distribution will be useful in situations where the set of spin states is infinite.
3. While the stack isn't empty,
   a) pop site m from the stack.
   b) If site m isn't marked,
      i. mark the site.
      ii. For every j such that $\{m, j\} \in E$, add site j to the stack with probability

$$p_r(s_m, s_j) = \min\{0, 1 - e^{\beta(Z(r \cdot s_m, s_j) - Z(s_m, s_j))}\}. \qquad (2.21)$$

      iii. Take $s_m \mapsto r \cdot s_m$.

When the stack is exhausted, a cluster of connected spins will have been transformed by the action of r. In order for this algorithm to be useful, it must satisfy ergodicity and detailed balance. Ergodicity is satisfied since we have ensured that R_2 acts transitively on X, e.g., for any $s, t \in X$ there exists $r \in R_2$ such that $r \cdot s = t$. Since there is a nonzero probability that only one spin is transformed and that spin can be transformed into any state, ergodicity follows. The probability $P(\mathbf{s} \to \mathbf{s}')$ that the configuration $\mathbf{s}$ is brought to $\mathbf{s}'$ by the flipping of a cluster formed by accepting transformations of spins via bonds $C \subseteq E$ and rejecting transformations via bonds $\partial C \subset E$ is related to

the probability of the reverse process $P(\mathbf{s}' \to \mathbf{s})$ by

$$\begin{aligned}\frac{P(\mathbf{s} \to \mathbf{s}')}{P(\mathbf{s}' \to \mathbf{s})} &= \frac{f(r \mid m_0, \mathbf{s})}{f(r^{-1} \mid m_0, \mathbf{s}')} \prod_{\{i,j\} \in C} \frac{p_r(s_i, s_j)}{p_{r^{-1}}(s'_i, s'_j)} \prod_{\{i,j\} \in \partial C} \frac{1 - p_r(s_i, s_j)}{1 - p_{r^{-1}}(s'_i, s'_j)} \\ &= \prod_{\{i,j\} \in \partial C} e^{\beta(Z(s'_i, s'_j) - Z(s_i, s_j))} = \frac{e^{-\beta \mathcal{H}(\mathbf{s}')}}{e^{-\beta \mathcal{H}(\mathbf{s})}},\end{aligned} \tag{2.22}$$

whence detailed balance is also satisfied, using $r = r^{-1}$ and $Z(r \cdot s', s') = Z(r \cdot s, s)$.

The Wolff algorithm is well known to be efficient in sampling many spin models near and away from criticality, including the Ising, Potts, and O(n) models. In general, its efficiency will depend on the system at hand, e.g., the structure of the configurations X and group R. A detailed discussion of this dependence for a class of configuration spaces with continuous symmetry groups can be found in [150, 155].

This algorithm can be run on a system using the provided library. To construct a system, you must provide a graph representing the lattice, a temperature, the spin coupling function Z, and the field coupling function B. Once constructed, cluster flips as described in Alg. 1 can be performed by directly providing seed sites m_0 and transformations r, or many in sequence by providing a function that generates random (appropriately distributed to preserve detailed balance) transformations r. The construction and use of Wolff systems is described at `https://doc.kent-dobias.com/wolff/system.html`.

§15. Adding the field

This algorithm relies on the fact that the coupling Z depends only on relative orientation of the spins—global reorientations do not affect the Hamiltonian. The external field B breaks this symmetry. Fortunately it can be restored. Define a new graph $\tilde{G} = (\tilde{V}, \tilde{E})$, where $\tilde{V} = \{0, 1, \dots, N\}$ adds the new 'ghost' site 0 which is connected by

$$\tilde{E} = E \cup \big\{\{0, i\} \mid i \in V\big\} \tag{2.23}$$

to all other sites. Instead of assigning the ghost site a spin whose value comes from X, we assign it values in the symmetry group $s_0 \in R$, so that the configuration space of the new model is $R \times X^N$. We introduce the Hamiltonian $\tilde{\mathcal{H}} : R \times X^N \to \mathbb{R}$ defined by

$$\begin{aligned} \tilde{\mathcal{H}}(s_0, \mathbf{s}) &= -\sum_{\{i,j\}\in E} Z(s_i, s_j) - \sum_{i \in V} B(s_0^{-1} \cdot s_i) \\ &= -\sum_{\{i,j\}\in \tilde{E}} \tilde{Z}(s_i, s_j), \end{aligned} \tag{2.24}$$

where the new coupling $\tilde{Z} : (R \cup X) \times (R \cup X) \to \mathbb{R}$ is defined for $s, t \in R \cup X$ by

$$\tilde{Z}(s,t) = \begin{cases} Z(s,t) & \text{if } s, t \in X \\ B(s^{-1} \cdot t) & \text{if } s \in R \\ B(t^{-1} \cdot s) & \text{if } t \in R. \end{cases} \tag{2.25}$$

The modified coupling is invariant under the action of group elements: for any $r, s_0 \in R$ and $s \in X$,

$$\begin{aligned} \tilde{Z}(rs_0, r \cdot s) &= B((rs_0)^{-1} \cdot (r \cdot s)) \\ &= B(s_0^{-1} \cdot s) = \tilde{Z}(s_0, s) \end{aligned} \tag{2.26}$$

The invariance of $\tilde{Z}$ to global transformations given other arguments follows from the invariance properties of Z.

We have produced a system incorporating the field function B whose Hamiltonian is invariant under global rotations, but how does it relate to our old system, whose properties we actually want to measure? If $A : X^N \to \mathbb{R}$ is an observable of the original system, we construct an observable $\tilde{A} : R \times X^N \to \mathbb{R}$ of the new system defined by

$$\tilde{A}(s_0, \mathbf{s}) = A(s_0^{-1} \cdot \mathbf{s}) \tag{2.27}$$

whose expectation value in the new system equals that of the original observable in the old system. First, note that $\tilde{\mathcal{H}}(1, \mathbf{s}) = \mathcal{H}(\mathbf{s})$. Since the Hamiltonian is invariant under global rotations, it follows that for any $g \in R$, $\tilde{\mathcal{H}}(g, g \cdot \mathbf{s}) = \mathcal{H}(\mathbf{s})$. Using the invariance properties of the measure on X and

introducing a measure ρ on R, it follows that

$$\begin{aligned}\langle \tilde{A} \rangle &= \frac{\int_R \int_{X^N} \tilde{A}(s_0, \mathbf{s}) e^{-\beta \tilde{\mathcal{H}}(s_0, \mathbf{s})} \, d\mu(\mathbf{s}) \, d\rho(s_0)}{\int_R \int_{X^N} e^{-\beta \tilde{\mathcal{H}}(s_0, \mathbf{s})} \, d\mu(\mathbf{s}) \, d\rho(s_0)} \\ &= \frac{\int_R \int_{X^N} A(s_0^{-1} \cdot \mathbf{s}) e^{-\beta \tilde{\mathcal{H}}(s_0, \mathbf{s})} \, d\mu(\mathbf{s}) \, d\rho(s_0)}{\int_R \int_{X^N} e^{-\beta \tilde{\mathcal{H}}(s_0, \mathbf{s})} \, d\mu(\mathbf{s}) \, d\rho(s_0)} \\ &= \frac{\int_R \int_{X^N} A(\mathbf{s}') e^{-\beta \tilde{\mathcal{H}}(s_0, s_0 \cdot \mathbf{s}')} d\mu(s_0 \cdot \mathbf{s}') \, d\rho(s_0)}{\int_R \int_{X^N} e^{-\beta \tilde{\mathcal{H}}(s_0, s_0 \cdot \mathbf{s}')} d\mu(s_0 \cdot \mathbf{s}') \, d\rho(s_0)} \\ &= \frac{\int_R d\rho(s_0)}{\int_R d\rho(s_0)} \frac{\int_{X^N} A(\mathbf{s}') e^{-\beta \mathcal{H}(\mathbf{s}')} d\mu(\mathbf{s}')}{\int_{X^N} e^{-\beta \mathcal{H}(\mathbf{s}')} d\mu(\mathbf{s}')} = \langle A \rangle.\end{aligned} \tag{2.28}$$

Using this equivalence, spin systems in a field may be treated in the following way.

1. Add a site to your lattice adjacent to every other site.
2. Initialize a 'spin' at that site whose value is a representation of a member of the symmetry group of your ordinary spins.
3. Carry out the ordinary Wolff cluster-flip procedure on this new lattice, substituting $\tilde{Z}$ as defined in (2.25) for Z.

Ensemble averages of observables A can then be estimated by sampling the value of $\tilde{A}$ on the new system. In contrast with the simpler ghost spin representation, this form of the Hamiltonian might be considered the 'ghost transformation' representation.

One of the celebrated features of the cluster representation of the Ising and associated models are the improved estimators of various quantities in the base model, found by measuring conjugate properties of the clusters themselves.[156] What of these quantities survive this translation? As is noted in the formative construction of the cluster representation for the Ising and Potts models, all estimators involving correlators between spins are preserved, including corre-

lators with the ghost site.[146] Where a previous improved estimator exists, we expect this representation to extend it to finite field, all other features of the algorithm held constant. For instance, the average cluster size in the Wolff algorithm is often said to be an estimator for the magnetic susceptibility in the Ising, Potts, and (with clusters weighted by the components of their spins along the reflection direction)[157] O(n) models, but really what it estimates is the averaged squared magnetization, which corresponds to the susceptibility when the average magnetization is zero. At finite field the latter thing is no longer true, but the correspondence between cluster size and the squared magnetization continues to hold (see (2.33) and Figure 2.3 below).

§16. Examples

Several specific examples from Table 2.1 are described in the following.

The Ising model

In the Ising model spins are drawn from the set $\{1, -1\}$. Its symmetry group is C_2, the cyclic group on two elements, which can be conveniently represented by a multiplicative group with elements $\{1, -1\}$, exactly the same as the spins themselves. The only nontrivial element is of order two, and is selected every time in the algorithm. Since the symmetry group and the spins are described by the same elements, performing the algorithm on the Ising model in a field is fully described by just using the 'ghost spin' representation. This algorithm or algorithms based on the same decomposition of the Hamiltonian have been applied by several researchers.[134–136] I have implemented this algorithm in

an existing interactive Ising simulator at `https://mattbierbaum.github.io/ising.js`.

The $\mathrm{O}(n)$ models

In the $\mathrm{O}(n)$ model spins are described by vectors on the $(n-1)$-sphere S^{n-1}. Its symmetry group is $\mathrm{O}(n)$, $n \times n$ orthogonal matrices, which act on the spins by matrix multiplication. The elements of $\mathrm{O}(n)$ of order two are reflections about hyperplanes through the origin and π rotations about any axis through the origin. Since the former generate the entire group, reflections alone suffice to provide ergodicity. Sampling those reflections uniformly works well at criticality. The 'ghost spin' version of the algorithm has been used to apply a simple vector field to the $\mathrm{O}(3)$ model.[158] Other fields of interest include $(n+1)$-dimensional spherical harmonics[154] and cubic fields,[159,160] which can be applied with the new method. The method is quickly generalized to spins whose symmetry groups are other compact Lie groups.[150,155]

At low temperature or high external vector field selecting reflections uniformly becomes inefficient because the excitations of the model are spin waves, in which the magnetization only differs by a small amount between neighboring spins. Under these conditions, most choices of reflection plane will cause a change in energy so great that the whole system is always flipped, resulting in many correlated samples. To ameliorate this, one can draw reflections from a distribution that depends on how the seed spin is transformed, taking advantage of the freedom to choose the function f in Alg. 1. We implement this in the following way. Say that the state of the seed of the cluster is s. Generate a vector t taken uniformly from the space of unit vectors orthogonal to s. Let

the plane of reflection be that whose normal is $n = s + \zeta t$, where ζ is drawn from a normal distribution of mean zero and variance σ. It follows that the tangent of the angle between s and the plane of reflection is also distributed normally with zero mean and variance σ. Since the distribution of reflection planes only depends on the angle between s and the plane, and since that angle is invariant under the reflection, this choice preserves detailed balance.

The choice of σ can be inspired by mean field theory. At high field or low temperature, spins are likely to both align with the field and each other and the model is asymptotically equal to a simple Gaussian one, in which in the limit of large L the expected square angle between neighbors is

$$\langle \theta^2 \rangle \simeq \frac{(n-1)T}{D + H/2}. \tag{2.29}$$

We take $\sigma = \sqrt{\langle \theta^2 \rangle}/2$. Fig. 2.1 shows the effect of making such a choice on autocorrelation times for the energy for a critical 3D XY (O(2)) model. At small fields both methods perform the same as zero field Wolff. Intermediate field values see efficiency gains for both methods. At large field the uniform sampling method sees correlation times grow rapidly without bound, while for the sampling method described here the correlation time crosses over to a constant. A similar behavior holds for the critical O(3) model, though in that case the constant value the correlation time approaches at large field is larger than its minimum value (see Fig. 2.2). This behavior isn't particularly worrisome, since the very large field regime corresponds to correlation lengths comparable to the lattice spacing and is efficiently simulated by other algorithms. More detailed discussion on correlation times and these numeric

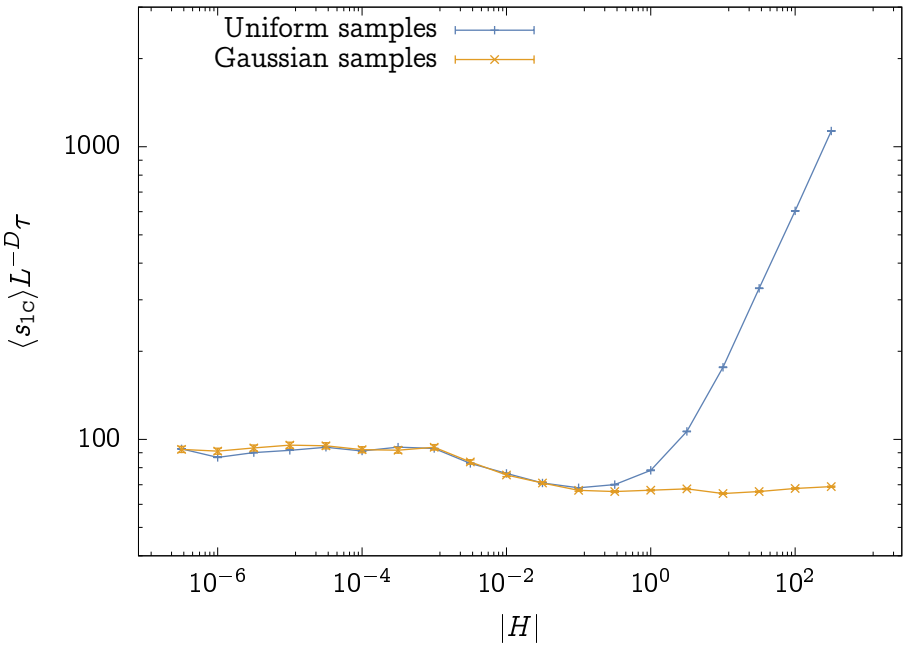

Figure 2.1: The scaled autocorrelation time of the energy $\mathcal{H}$ for the Wolff algorithm on a $32 \times 32 \times 32$ XY model at its critical temperature as a function of applied vector field magnitude $|H|$. Blue points correspond to reflections sampled uniformly, while the yellow points represent reflections sampled as described in section 16.

experiments can be found in section 17.

The Potts model

In the q-state Potts model spins are described by elements of $\{1, \ldots, q\}$. Its symmetry group is the symmetric group S_n of permutations of its elements. The element $(i_1, \ldots, i_q)$ takes the spin s to i_s. There are potentially many elements of order two, but the two-element swaps alone are sufficient to both generate the group and act transitively on $\{1, \ldots, q\}$, providing ergodicity. Equations of state have been predicted from ϵ-expansions for the three-state

Potts model with symmetry-breaking fields—it may be interesting to compare this with efficient numeric results.[161,162]

Clock models

In the q-state clock model spins are described by elements of $\mathbb{Z}/q\mathbb{Z}$, the set of integers modulo q. Its symmetry group is the dihedral group $D_q = \{r_0, \dots, r_{q-1}, s_0, \dots, s_{q-1}\}$, the group of symmetries of a regular q-gon. The element r_n represents a rotation by $2\pi n/q$, and the element s_n represents a reflection composed with the rotation r_n. The group acts on spins by permutation: $r_n \cdot m = n + m \pmod q$ and $s_n \cdot m = -(n + m) \pmod q$. This is the natural action of the group on the vertices of a regular polygon that have been numbered 0 through $q - 1$. The elements of D_q of order 2 are all reflections and $r_{q/2}$ if q is even, though the former can generate the latter. While reflections do not necessarily generate the entire group, their action on $\mathbb{Z}/q\mathbb{Z}$ is transitive and therefore the algorithm is ergodic.

Roughening models

Though not often thought of as a spin model, roughening of surfaces can be described in this framework. These models are used to describe what happens to the interface between magnetic or crystalline orders, which can experience a phase transition in which the correlation length of the height of the interface diverges. The "spins," which describe the height of the interface at each point in space, are described by integers $\mathbb{Z}$ and their symmetry group is the infinite dihedral group $D_\infty = \{r_i, s_i \mid i \in \mathbb{Z}\}$, whose action on the spin $j \in \mathbb{Z}$ is given by $r_i \cdot j = i + j$ and $s_i \cdot j = -i - j$. The elements of order two are reflections

s_i, whose action on $\mathbb{Z}$ is transitive. The coupling can be any function of the absolute difference $|i-j|$. Because uniform choice of reflection will almost always result in energy changes so large that the whole system is flipped, it is better to select random reflections about integers or half-integers close to the state of the seed. Besides randomly selecting such a reflection plane, choosing it to coincide with another random spin, different from the seed, has proven quite successful, though it isn't enough to give ergodicity on its own.[163–165]. Detailed balance is still satisfied if the bond probabilities (2.21) are modified by adding a constant $0 < x \leq 1$ with

$$p_r(s_m, s_j \mid x) = \min\{0, 1 - xe^{\beta(Z(r\cdot s_m, s_j) - Z(s_m, s_j))}\}. \tag{2.30}$$

When $x < 1$ transformations that do not change the energy of a bond can still activate it in the cluster, which allows nontrivial clusters to be seeded when the height of the starting site is also the plane of reflection. This modification is likely useful in general for systems with large yet discrete state spaces.

§17. Performance

No algorithm is worthwhile if it doesn't run efficiently. This algorithm, being an extension of the Wolff algorithm into a new domain, should be considered successful if it likewise extends the efficiency of the Wolff algorithm into that domain. Some systems are not efficient under Wolff, and we don't expect them to fare better when extended in a field. For instance, Ising models with random fields or bonds technically can be treated with Wolff,[166] but it is not efficient because the clusters formed do not scale naturally with the

correlation length.[140,152] Other approaches, like replica methods, should be relied on instead.[140–142]

At a critical point, correlation time τ scales with system size $L = N^{-D}$ as $\tau \sim L^z$. Cluster algorithms are celebrated for their small dynamic exponents z. In the vicinity of an ordinary critical point, the renormalization group predicts scaling behavior for the correlation time as a function of temperature t and field h of the form

$$\tau = h^{-z\nu/\beta\delta}\mathcal{T}(ht^{-\beta\delta}, hL^{\beta\delta/\nu}). \tag{2.31}$$

If a given dynamics for a system at zero field results in scaling like L^z, one should expect its natural extension in the presence of a field to scale roughly like $h^{-z\nu/\beta\delta}$ and collapse appropriately as a function of $hL^{\beta\delta/\nu}$.

We measured the autocorrelation time τ of the energy $\mathcal{H}$ for a variety of models at critical temperature with many system sizes and canonical fields (see Table 2.1 with $h = \beta H$) using standard methods for obtaining the value and uncertainty from time series.[167] Since the computational effort expended in each step of the algorithm depends linearly on the size of the associated cluster, these values are then scaled by the average cluster size per site $\langle s_{1c}\rangle/L^D$ to produce something proportional to machine time per site. The resulting scaling behavior, plotted in Figure 2.2, is indeed consistent with an extension to finite field of the behavior at zero field, with an eventual finite-size crossover to constant autocorrelation time at large field. This crossover isn't always kind to the efficiency, e.g., in the O(3) model, but in the large-field regime where the crossover happens the correlation length is on the scale of the lattice spacing

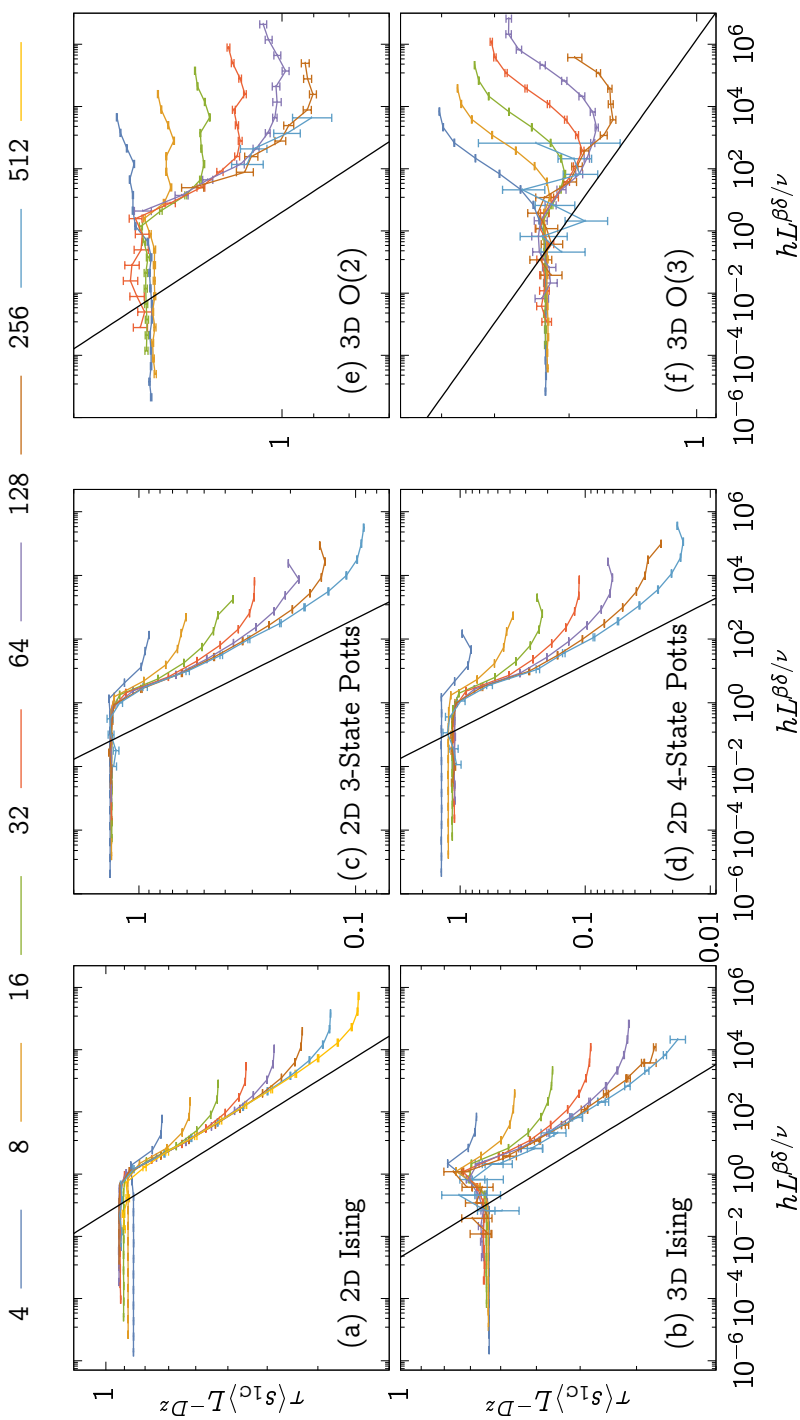

Figure 2.2: Scaling collapse of autocorrelation times τ for the energy $\mathcal{H}$ scaled by the average cluster size as a function of external field for various models of Table 2.1. Critical exponents are model-dependent. Colored lines and points depict values as measured by the extended algorithm. Solid black lines show a plot proportional to $h^{-z\nu/\beta\delta}$ for each model. The dynamic exponents z are roughly measured as 2D Ising: 0.23(5), 3D Ising: 0.28(5), 2D 3-State Potts: 0.55(5), 2D 4-State Potts: 0.94(5), 3D O(2): 0.17(5), 3D O(3): 0.13(5). O(n) models use the distribution of transformations described in §16. The curves stop collapsing at high fields when the correlation length falls to near the lattice spacing; here non-cluster algorithms can be efficiency used.

and better algorithms exist, like Bortz–Kalos–Lebowitz for the Ising model.[168] Also plotted are lines proportional to $h^{-z\nu/\beta\delta}$, which match the behavior of the correlation times in the intermediate scaling region as expected. Values of the critical exponents for the models were taken from the literature[169–171] with the exception of z for the energy in the Wolff algorithm, which was determined for each model by making a power law fit to the constant low field behavior. These exponents are imprecise and are provided in the figure with only qualitative uncertainty.

Since the formation and flipping of clusters is the hallmark of Wolff dynamics, another way to ensure that the dynamics with field scale like those without is to analyze the distribution of cluster sizes. The success of the algorithm at zero field is related to the fact that the clusters formed undergo a percolation transition at models' critical point. According to the scaling theory of percolation,[41] the distribution of cluster sizes in a full Swendsen–Wang decomposition—where the whole system is decomposed into clusters with every bond activated with probability (2.21)—of the system scales consistently near the critical point if it has the form

$$P_{\mathrm{SW}}(s) = s^{-\tau} f(ts^{\sigma}, th^{-1/\beta\delta}, tL^{1/\nu}). \tag{2.32}$$

The distribution of cluster sizes in the Wolff algorithm can be computed from this using the fact that the algorithm selects clusters with probability propor-

tional to their size, or

$$\begin{aligned}\langle s_{1c}\rangle &= \sum_s sP_{1c}(s) = \sum_s s\frac{s}{N}P_{SW}(s)\\ &= L^{\gamma/\nu}g(ht^{-\beta\delta}, hL^{\beta\delta/\nu}).\end{aligned} \tag{2.33}$$

For the Ising model, an additional scaling relation can be written. Since the average cluster size is the average squared magnetization, it can be related to the scaling functions of the magnetization and susceptibility per site by (with $ht^{-\beta\delta}$ dependence dropped)

$$\begin{aligned}\langle s_{1c}\rangle &= L^D\langle M^2\rangle = \beta\langle\chi\rangle + L^D\langle M\rangle^2\\ &= L^{\gamma/\nu}\left[(hL^{\beta\delta/\nu})^{-\gamma/\beta\delta}\beta\mathcal{Y}(hL^{\beta\delta/\nu}, ht^{-\beta\delta})\right.\\ &\quad \left. + (hL^{\beta\delta/\nu})^{2/\delta}\mathcal{M}(hL^{\beta\delta/\nu}, ht^{-\beta\delta})\right].\end{aligned} \tag{2.34}$$

We therefore expect that, for the Ising model, $\langle s_{1c}\rangle L^{-\gamma/\nu}$ should go as $(hL^{\beta\delta/\nu})^{2/\delta}$ for large argument. We further conjecture that this scaling behavior should hold for other models whose critical points correspond with the percolation transition of Wolff clusters. This behavior is supported by our numeric work along the critical isotherm for various Ising, Potts, and O(n) models, shown in Fig. 2.3. Fields are the canonical ones referenced in Table 2.1. As can be seen, the average cluster size collapses for each model according to the scaling hypothesis, and the large-field behavior likewise scales as we expect from the naïve Ising conjecture.

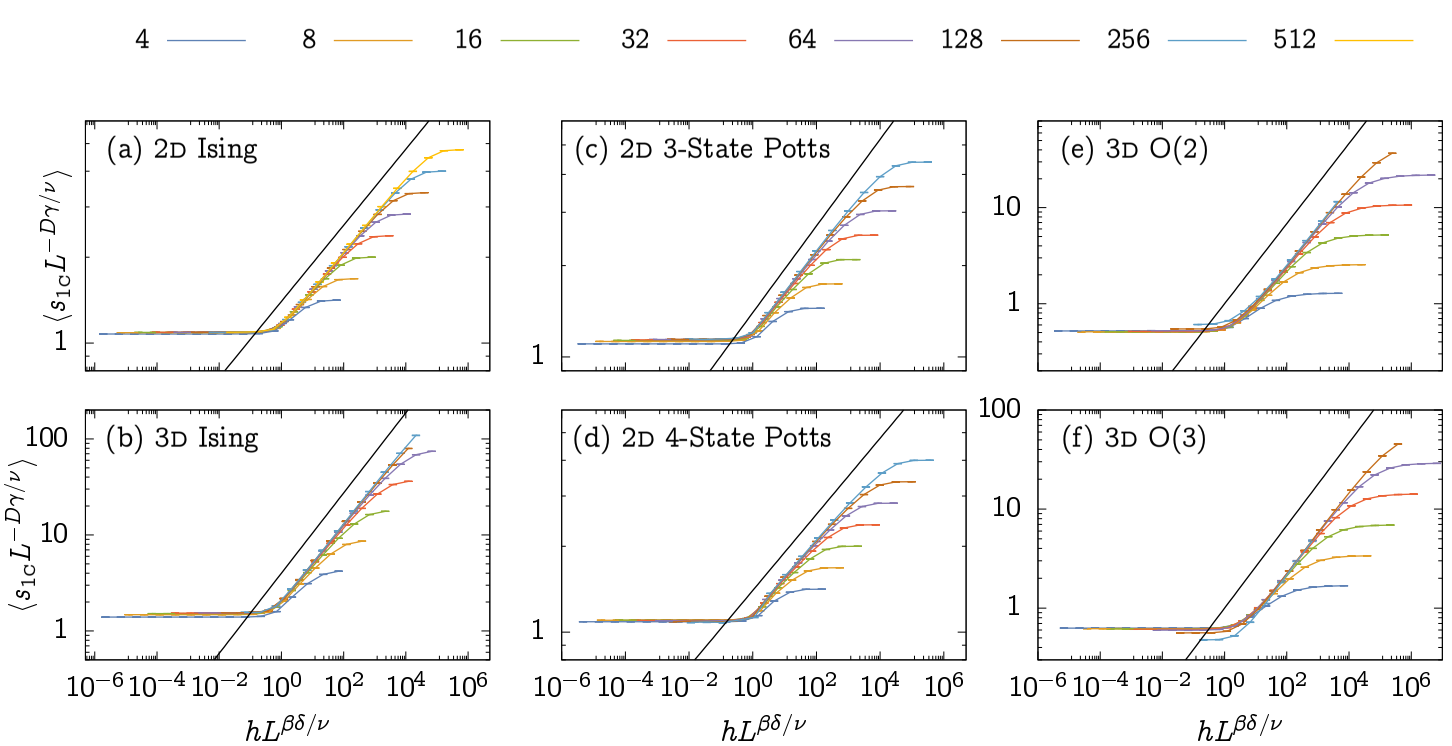

Figure 2.3: Collapses of rescaled average Wolff cluster size $\langle s \rangle_{1C} L^{-\gamma/\nu}$ as a function of field scaling variable $hL^{\beta\delta/\nu}$ for a variety of models. Critical exponents γ, ν, β, and δ are model-dependent. Colored lines and points depict values as measured by the extended algorithm. Solid black lines show a plot of $g(0, x) \propto x^{2/\delta}$ for each model.

§18. Applying nonlinear fields to the XY model

Thus far our numeric work has quantified the performance of existing techniques. Briefly, we demonstrate our general framework in a new way: harmonic perturbations to the low-temperature XY, or 2D O(2), model. We consider fields of the form $B_n(s) = h_n \cos(n\theta(s))$, where θ is the angle made between s and the x-axis. Corrections of these types are expected to appear due to the presence of the lattice or substrate in realistic models of systems naïvely expected to exhibit Berezinskiĭ–Kosterlitz–Thouless critical behavior. Whether these fields are relevant or irrelevant in the renormalization group sense determines whether those systems spoil or admit that critical behaviour. Among many fascinating[139,154,172–174] results that emerge from systems with one or more of these fields applied, it is predicted that h_4 is relevant while h_6 is not at some sufficiently high temperatures below the Berezinskiĭ–Kosterlitz–Thouless point.[154] The sixfold fields are expected to be present, for instance, in the otherwise Berezinskiĭ–Kosterlitz–Thouless-type two-dimensional melting of argon on a graphite substrate.[175]

We made a basic investigation of this result using our algorithm. Since we ran the algorithm at fairly high fields we did not choose reflections though the origin uniformly. Instead, we choose the planes of reflection first by rotating our starting spin by $\pi m/n$ for m uniformly taken from $1, \ldots, n$ and generating a normal to the plane from that direction as described in §16. The resulting susceptibilities as a function of system size are shown for various field strengths in Figure 2.4. In the fourfold case, for each field strength there is a system size at which the divergence in the susceptibility is cut off, while

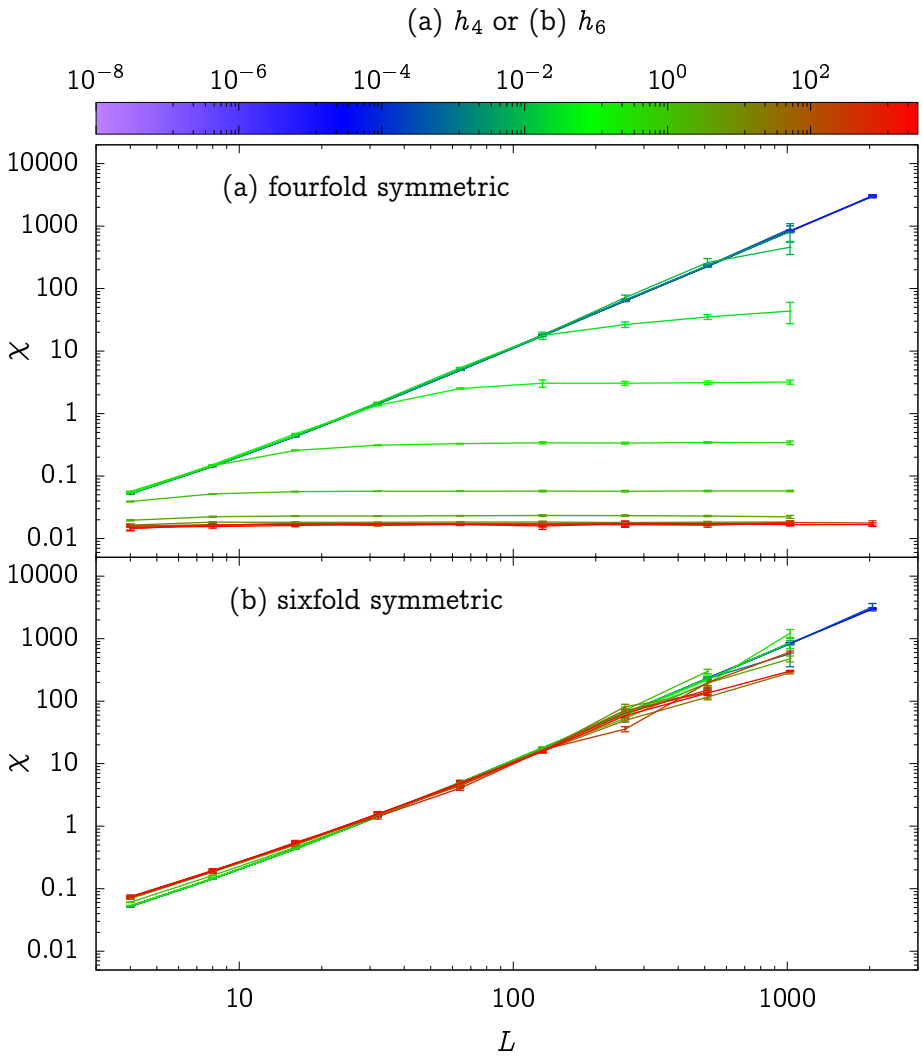

Figure 2.4: Susceptibilities as a function of system size for a 2D O(2) model at $T = 0.7$ and with (top) fourfold symmetric and (bottom) sixfold symmetric perturbing fields. Different field strengths are shown in different colors.

for the sixfold case we measured no such cutoff, even up to strong fields. This conforms to the expected result, that even in a strong field the sixfold perturbations preserve the critical behavior. Previous work has used Monte Carlo to investigate similar symmetry-breaking fields and used a hybrid cluster–metropolis method.[139] To our knowledge, no application of a direct cluster method has been applied to this problem before now.

§19. Directly measuring metastable scaling functions

Anyone who has even casually tried to measure some magnetization-related property of the standard Ising model has had to reckon with an exceedingly subtle aspect of equilibrium statistical mechanics. Record the magnetization and susceptibility at a variety of temperatures and plot. At first glance, a catastrophe: the susceptibility indeed increases as the critical point is approached from above, but *continues* to increase below it—and the magnetization looks like random, near-zero gibberish! What has happened?

The answer lies in the fact that finite-size simulation of statistical mechanics models with broken symmetry have behavior that is qualitatively different from the same models in the thermodynamic limit. At any finite size, equilibrium follows typical ergodic arguments and the magnetization is zero, while at infinite size a superselection rule kicks in and, at low temperatures, dynamics will never bring a system from positive to negative magnetization.[176]

Though in practice we always simulate at finite size (unless very clever[177]), we are usually interested in properties at infinite size. We therefore must invent schemes for taking data found from simulations at finite size and re-

liably and quickly extrapolate them to infinite size. For the Ising model's magnetic properties, a tradition formula is to take the absolute value of the magnetization as the basis for a measurement. If we're looking to measure low-temperature properties at zero external field, this decision makes a lot of sense: after all, the state with net up and that with net down should have identical properties (once one is flipped to match the other), and therefore we might as well sum twice as much data to help our averages converge well.

When the external field isn't zero, this reasoning makes somewhat less sense. Typical spin configurations that have a net magnetization against the field should not resemble those with a net magnetization towards the field. Therefore rather than take the average of values or absolute values of properties of microstates we shouldn't expect to resemble those in the thermodynamic limit, we should take averages only over microstates that we suspect belong to the equilibrium state we wish to study. One could envision, in systems with subtle differences between states, an involved procedure for taking a microstate at finite size and determining what phase it most likely belongs to (machine learning?!), but for the Ising model and other systems where the order parameter space is clearly partitioned into disjoint sectors by a symmetry, a crude rule emerges with very little effort: associate a microstate with whatever phase its magnetization aligns toward. This means, for the equilibrium Ising model in a positive field: only average positive magnetization samples!

What about the negative magnetization samples? Should we throw them away? No! If we expect that Monte Carlo averages over samples with positive magnetization at finite size should converge well to expectation values of a

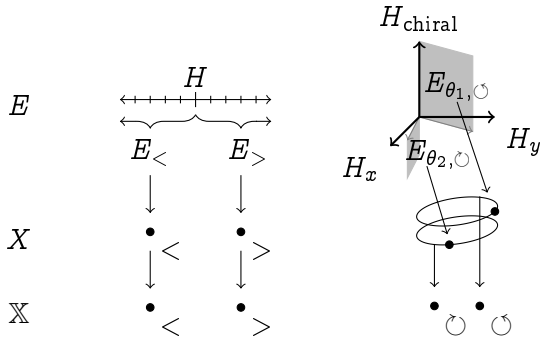

Figure 2.5: Two examples of parameter space E, order parameter space X, and the connected components of order parameter space $\mathbb{X}$. Left: The Ising model has a scalar field. Negative (or zero) values correspond to negative magnetization equilibria, while positive (or zero) values correspond to positive magnetization equilibria. The order parameter is the direction of the magnetization, and the order parameter space is simply two points. The connected components of this space is simply the space again. Right: A chiral planar spin model might have a vector field in the plane, and a field that favors different chiralities, like some vector spin glasses.[178] The order parameter is the direction of the planar spins and their chirality, making order parameter space $S_1 \times S_1$ or $\mathrm{O}(2)$. There are two connected components in this space, so the set of connected components is again two points.

phase with positive magnetization in the thermodynamic limit, then we should also expect that averages over samples with *negative* magnetization should converge well to expectation values of a phase with *negative* magnetization in the thermodynamic limit. But what significance should be afforded a *negative* magnetization phase in a *positive* magnetic field? Traditionally, this is the notion of the metastable state.

This reasoning lies behind most direct measurements of metastable properties of standard statistical models in numerics. With rare exception,[179,180] studies making those measurements set up a numeric apparatus in the expected metastable state, then average quantities until the system has fallen

out of this state into the true equilibrium.[181–184] The alternative, which is to sample with abandon and collect the samples into bins by magnetization afterward, is inaccessible to Monte Carlo methods on all but the smallest systems.[179] That is, without Wolff in a field!

Wolff is marked for the unphysical nature of its updates. Unlike traditional local Monte Carlo methods, whose updates make a good stand-in for true time evolution and become virtually trapped in a single-magnetization state at zero field when the system size becomes even moderate, Wolff rapidly samples back and forth between opposing magnetizations, as in Figure 2.6 (top). In small fields, this property persists (Figure 2.6 (middle)), and frequent samples from the set of microstates whose properties we expect to converge on those of the metastable state occur. However, this well is not infinite, and eventually the sorry fact that the metastable configurations have vanishingly small probability compared to the stable ones independent of dynamics comes into play, and no switching is seen (Figure 2.6 (bottom)).

Data taken within a regime where the switching happens offers four different possible estimates of the finite-size magnetization based on different partitions of the empirical distribution, depicted in Figure 2.7. The first, $\langle M \rangle$, is the standard average over all states, which is so quickly abandoned at zero field because it must be zero. The second, $\langle |M| \rangle$, is the average over all states of the absolute value, a usual fix at zero temperature. Finally, $\langle M \rangle_>$ and $\langle M \rangle_<$ are the average over only microstates for the $M > 0$ and $M < 0$, respectively. These measurements are plotted for a sample system as a function of the applied field in Figure 2.8, with a detail of the region very close to zero field in Figure 2.9. Notice in that second figure how, while the positive aver-

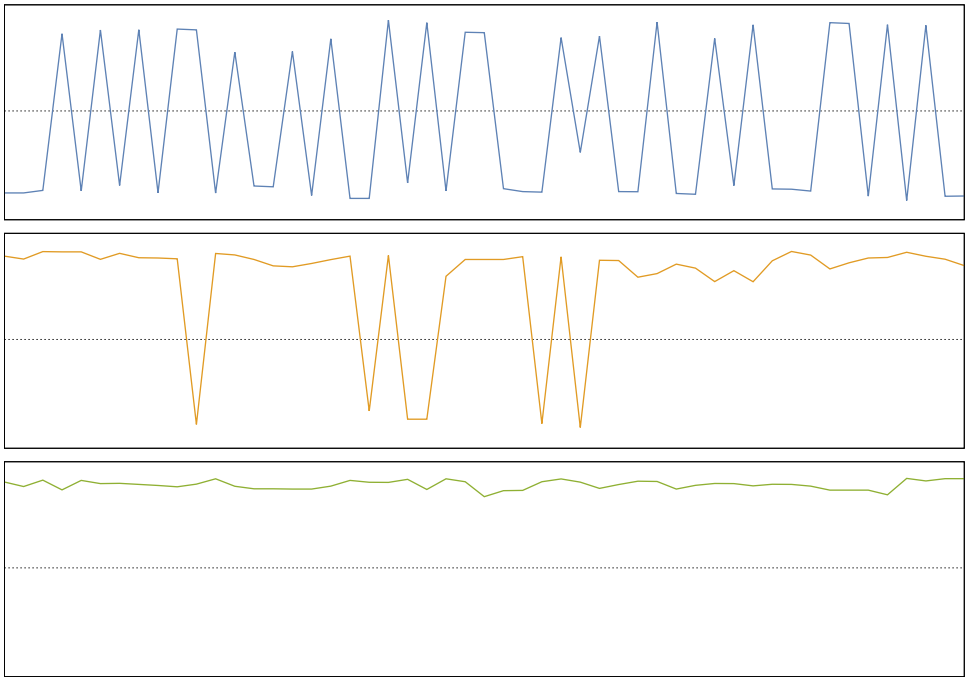

Figure 2.6: Magnetization as a function of step for 50 steps of a Wolff simulation of a 2D Ising model with 64×64 spins at $T = 2.23794$ and (Top) $H = 0$ (Middle) $H = 0.001$ (Bottom) $H = 0.01$.

age and absolute value average coincide at $H = 0$, the approach is markedly different, and in particular the absolute value estimator does a shoddy job of estimating the slope and curvature of the magnetization on its approach.

The differences become more striking when the estimated susceptibilities are examined. Recall that the susceptibility—which is defined as the slope of the magnetization with respect to H—can be written as the variance in the

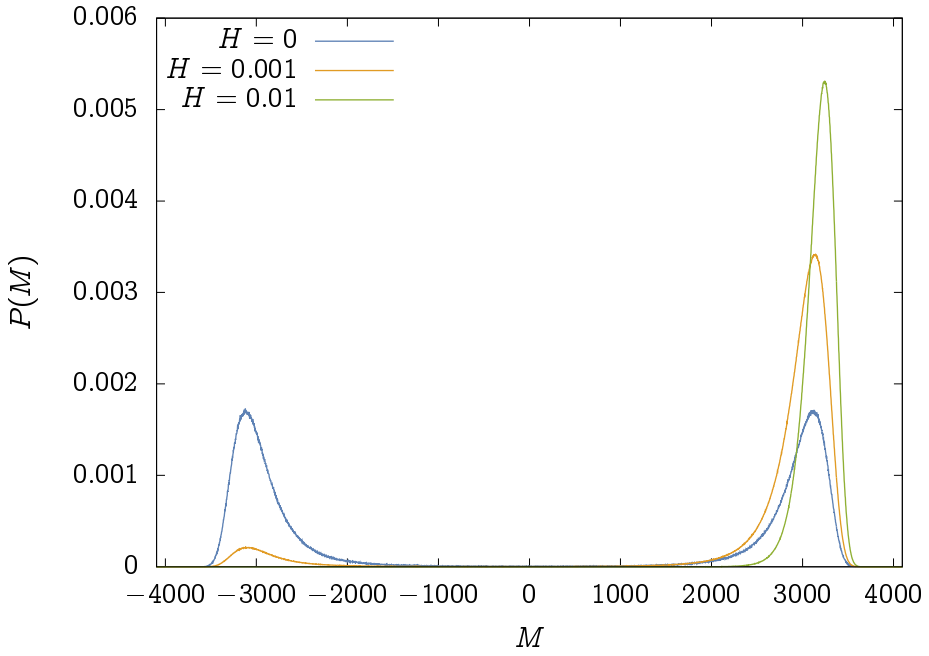

Figure 2.7: Probability distribution for the magnetization of a 64×64 Ising model at $T = 2.23794$ at three different external fields.

magnetization via a fluctuation–response relationship:

$$\begin{aligned}\chi &= \frac{\partial \langle M \rangle}{\partial H} = \frac{\partial}{\partial H} \frac{\operatorname{Tr} M e^{-\beta \mathcal{H}_0 + \beta H M}}{\operatorname{Tr} e^{-\beta \mathcal{H}_0 + \beta H M}} \\ &= \beta \left[\frac{\operatorname{Tr} M^2 e^{-\beta \mathcal{H}_0 + \beta H M}}{\operatorname{Tr} e^{-\beta \mathcal{H}_0 + \beta H M}} + \left(\frac{\operatorname{Tr} M e^{-\beta \mathcal{H}_0 + \beta H M}}{\operatorname{Tr} e^{-\beta \mathcal{H}_0 + \beta H M}} \right)^2 \right] \\ &= \beta (\langle M^2 \rangle - \langle M \rangle^2)\end{aligned} \tag{2.35}$$

This relationship is unchanged when the averages involved are restricted to subsets of the configuration space, but it *is* changed when M is replaced by $|M|$. Figure 2.10 shows the susceptibilities measured for the same system as a function of field, defined by the fluctuations in the magnetization. Notice that

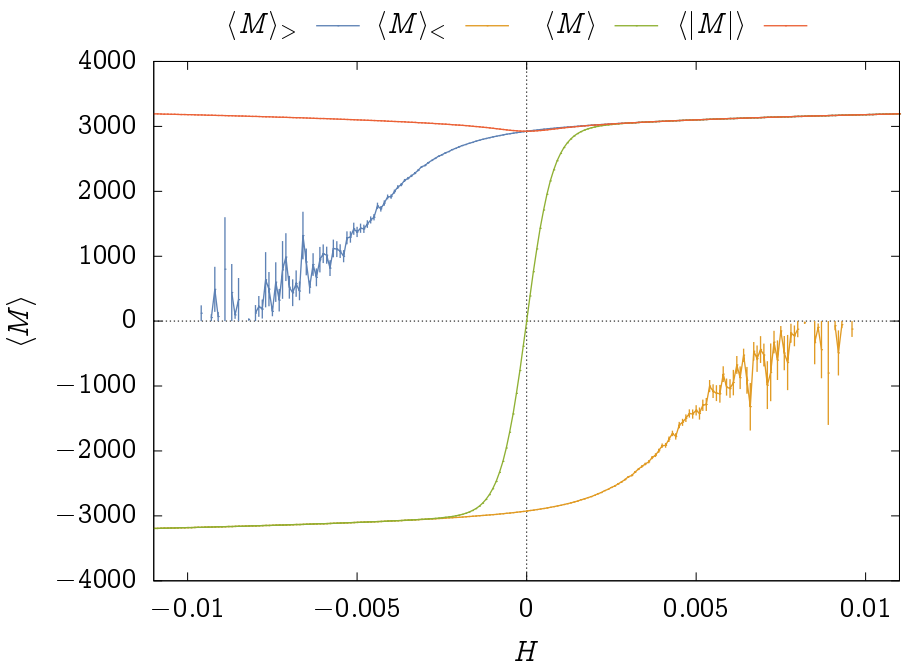

Figure 2.8: Various finite-size measurements of the magnetization of a 64×64 Ising model at $T = 2.23794$ as a function of external field H.

while the true finite-size susceptibility and restricted-ensemble susceptibilities correspond with the slopes of the magnetizations in Figure 2.8, the absolute value susceptibility does not—the corresponding magnetization has zero slope at the origin, but the susceptibility is not zero.

In all of these examples, some very interesting nontrivial behavior is happening in the negative field, positive magnetization ensembles. The magnetization approaches zero around the region where the number of flipped samples goes away, and the susceptibility appears nonmonotonic, eventually turning over! While an investigation very far into the metastable region is unlikely to be very fruitful, since the frequency of the appearance of the metastable

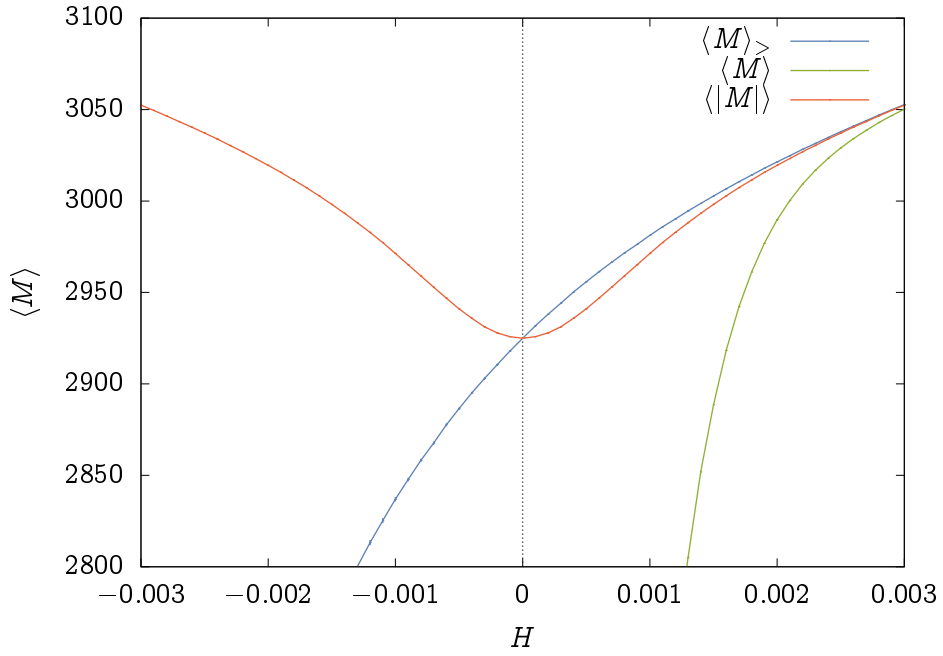

Figure 2.9: Various finite-size measurements of the magnetization of a 64×64 Ising model at $T = 2.23794$ as a function of external field H, zoomed into the region around $H = 0$.

samples vanishes, a shrewd use of finite size scaling to dance around this barrier may provide insight about critical scaling functions much farther into the metastable region than is currently known.

A first foray into a measurement like this can be seen in Figure 2.11, which depicts the susceptibility of a 128×128 Ising model as a function of the scaling invariants $h|t|^{-\beta\delta}$ and $L|t|^{\nu}$. The typical universal scaling function is found when the latter goes to infinity. Also shown are the first seven terms of a divergent series for the susceptibility, which quickly and unphysically diverges from the measured value in the, e.g., large field region because an essential

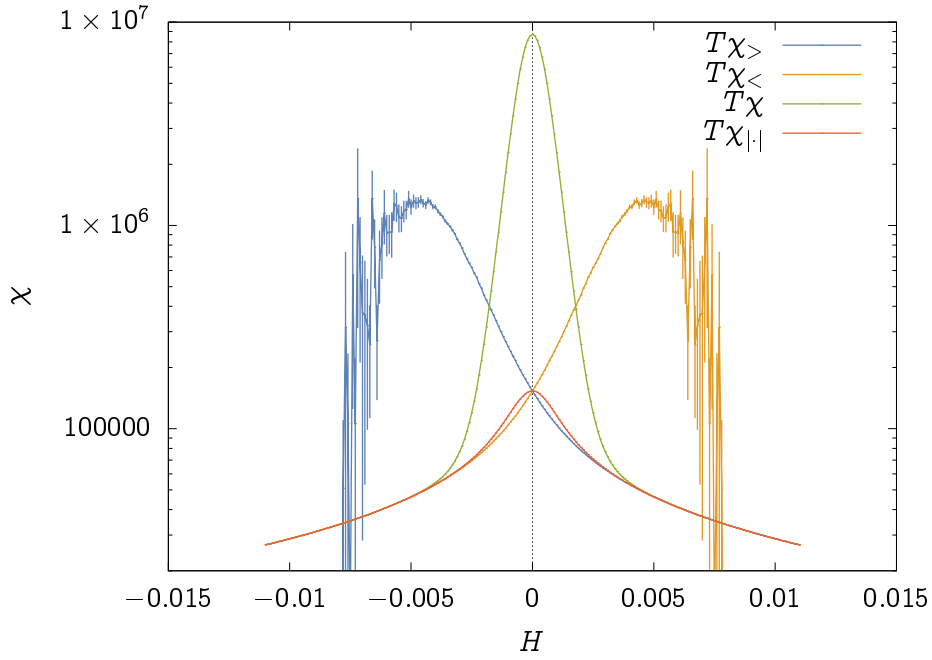

Figure 2.10: Various finite-size measurements of the susceptibility of a 64×64 Ising model at $T = 2.23794$ as a function of external field H.

singularity exists at $h = 0$.* Even at this modest system size, the equilibrium susceptibility converges to its infinite-size value well. Whether the same can be said for the metastable susceptibility will require more careful analysis, including an informed ansatz for the form of the scaling function in that phase. Notice that at $L|t|^\nu = 4$ we are pushing up against the point where very few metastable samples are being taken. Analyzing data like this, spread across many system sizes, could begin to put bounds on the form of metastable properties. Other models with disconnected order parameter spaces could also be studied this way, e.g., the chiral Potts model using the geometric

*An essential what?! See Chapter 4!

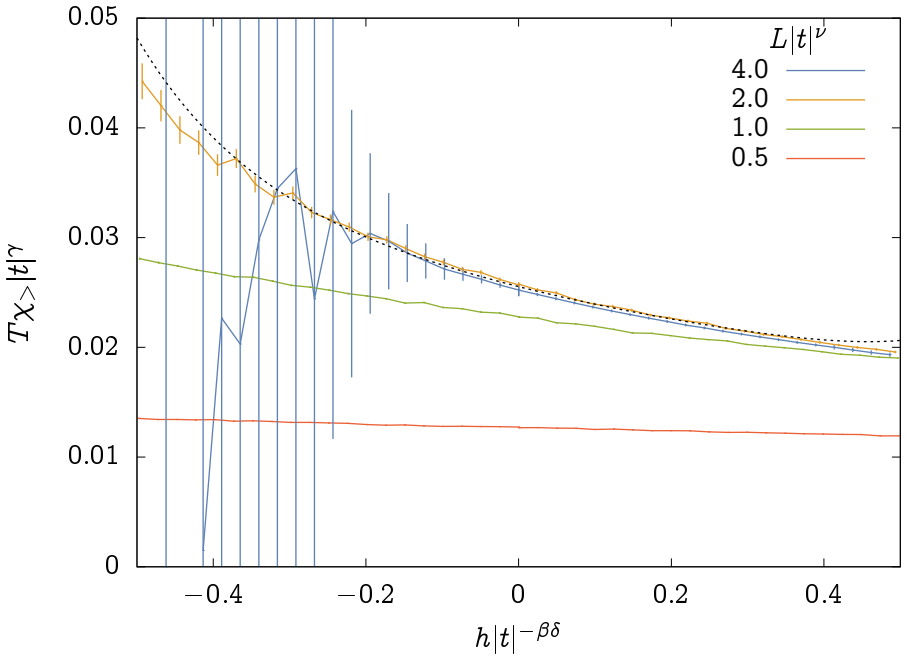

Figure 2.11: The susceptibility $\chi_>$ of a 128×128 Ising model as a function of the scaling invariant $h|t|^{-\beta\delta}$ for values of the scaling invariant $L|t|^\nu$. The dashed line shows the sum of the first seven terms of a divergent expansion for the susceptibility scaling function about zero field.[187]

cluster methods described in the next section, which might give insight into metastable chirality in liquid crystals and other similar systems.[185,186]

§20. Conclusions

We have taken several disparate extensions of cluster methods to spin models in an external field and generalized them to work for any model of a broad class. The resulting representation involves the introduction of not a ghost spin, but a ghost transformation. We provide a C++ library with example implementations of all models described here. We provided evidence that al-

gorithmic extensions deriving from this method are the natural way to extend cluster methods in the presence of a field, in the sense that they appear to reproduce the scaling of dynamic properties in a field that would be expected from renormalization group predictions.

In addition to uniting several extensions of cluster methods under a single description, our approach allows the application of fields not possible under prior methods. Instead of simply applying a spin-like field, this method allows for the application of *arbitrary functions* of the spins. For instance, theoretical predictions for the effect of symmetry-breaking perturbations on spin models can be tested numerically.[154,159,160,188]

CHAPTER III

CLUSTER MONTE CARLO WITH BACKGROUND LANDSCAPES

In the previous chapter I discussed how fast cluster Monte Carlo algorithms for general lattice models can be used in the presence of on-site potentials by adding a new degree of freedom to the original Hamiltonian and then taking modified expectation values on the new model. Throughout, the *lattice* aspect of those models was at the forefront of our physics imaginary: the models consist of degrees of freedom ("spins") that occupy the sites of a graph (the "lattice") and interact via its edges (between "neighboring" sites). Though this was intended, we need not restrict ourselves to that application. In this chapter, I will describe how the same method can be used to construct cluster algorithms for atomistic, colloidal, and other models defined by *spatial* degrees of freedom interacting via arbitrary pair potentials.

In one sense this chapter will be a corollary to the previous one, since once we make the formal correspondence between its language and spatial models the technique follows immediately. However, in another sense this chapter will discuss many details of implementation and simulation strategy that were less relevant before, and so stands alone in its own right. After introducing the correspondence, I will discuss several examples, strengths of the method for

general atomistic simulation, subtleties in making an efficient implementation, and speculative hope for insight into the glass problem. Though the validity of the algorithms described here is shown and schematic techniques for their efficient use are detailed, much of the work to compare their efficiency against other methods is still ongoing.

§21. The formal correspondence

Consider a model of particles that exist in space, have some internal structure, interact with each other, and interact with an energetic background. The ith particle has a coordinate in $r_i \in \mathcal{S}$, which describes its position in space (center of mass, say). It also potentially has static properties $\rho_i \in \mathcal{R}$ (like mass or radius) and dynamic properties $\eta_i \in \mathcal{E}$ (like orientation). The pair interaction $V : (\mathcal{S} \times \mathcal{R} \times \mathcal{E}) \times (\mathcal{S} \times \mathcal{R} \times \mathcal{E}) \to \mathbb{R}$ between particles depends on both their positions and their properties. Finally, the energetic background $U : \mathcal{S} \times \mathcal{R} \times \mathcal{E} \to \mathbb{R}$ also depends on all these properties. The Hamiltonian of a model like this is given by

$$\mathcal{H} = \sum_{ij} V\left[(r_i, \rho_i, \eta_i), (r_j, \rho_j, \eta_j)\right] + \sum_i U(r_i, \rho_i, \eta_i) \tag{3.1}$$

Here's the trick: call $X = \mathcal{S} \times \mathcal{E}$, $G = (V, E)$ the complete graph on N elements, i.e, $V = \{1, \dots, N\}$ and $E = \{\{i, j\} \mid 1 \le i \le N, i < j \le N\}$, and

$$\begin{aligned} Z_{ij}[(r_i, \eta_i), (r_j, \eta_j)] &= -V\left[(r_i, \rho_i, \eta_i), (r_j, \rho_j, \eta_j)\right] \\ B_i[(r_i, \eta_i)] &= -U(r_i, \rho_i, \eta_i) \end{aligned} \tag{3.2}$$

It follows that for $s \in X^N$,

$$\mathcal{H}(s) = -\sum_{\{i,j\}\in E} Z_{ij}(s_i, s_j) - \sum_{i\in V} B_i(s_i) \tag{3.3}$$

which is precisely the form of (2.18) with bond and site dependence, which don't modify any of the conclusions of the previous chapter. While the graph describing "neighboring" pairs is complete and any object therefore a "neighbor" of any other, this is a reference to interaction neighbors, not positional neighbors. Any two, e.g., spheres must be a neighbor because they have a pair interaction, even if they are so distant in a particular configuration that their pair interaction is zero. In practice, using neighbor-finding algorithms to identify positional neighbors saves time by preventing the useless evaluation of many zero-energy pairs.

We're not quite done, though: besides identifying a 'spin' space and a 'lattice' graph, we need a symmetry group with sufficient structure to provide the basis for constructing clusters. Our 'space,' recall, is the product of actual space and mutable particle properties: what are its symmetries? In the absence of a background potential U, and if the dependence of the pair potential on the mutable properties are sensible, then the relevant symmetries are the symmetries of space itself. If $\mathcal{S} = \mathbb{R}^D$, then the symmetry group is the Euclidean group. Since global rotations, translations, inversions, and reflections are all symmetries of Euclidean space, they are also symmetries of our Hamiltonian. Table 3.1 details several common examples.

Space ($\mathcal{S}$)	Symmetry group	Order-two elements
Real space ($\mathbb{R}^D$)	Euclidean group	Reflections, inversions, π rotations
Discrete space ($\mathbb{Z}^D$)	Coxeter group R_D	Half-integer reflections, inversions, π rotations
Torus ($\mathbb{T}^D$)	Torus group	Translations & reflections along axes, inversions, π rotations
Sphere (S^n)	Orthogonal group	Reflections through origin, π rotations

Table 3.1: Common single-particle coordinate spaces, their symmetry groups, and member elements of order two.

§22. Lattice models at constant magnetization

When the state of the spins are fixed, traditional lattice models can be viewed as varieties of lattice gas. The correspondence between the Ising model and a hard-center lattice gas with neighbor attraction is a classic exercise.[189] We can use the relationship between these to make cluster moves with potentials that preserve the net state of these models. Versions of this method without a field were developed by Heringa and Blöte and have been used in other studies under the name "geometric cluster algorithms."[190–193]

Lattice gases live on a lattice, either infinite, e.g., $\mathbb{Z}^D$, or finite, e.g. $(\mathbb{Z}/L\mathbb{Z})^D$. In either case, order-two symmetries of the space include reflections through half-integer planes orthogonal to any principal axis, π rotations about half-integer sites, and inversions about half-integer sites. The model's spins have an internal state s, e.g., up or down, and interact via a pair inter-

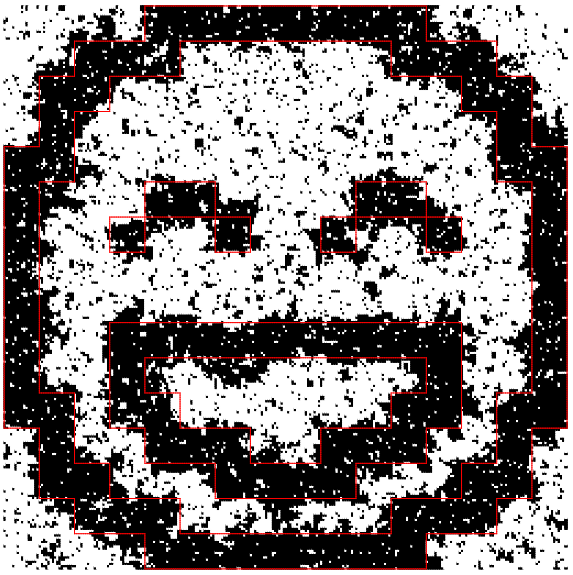

Figure 3.1: A 256×256 Ising model at its critical temperature and held at constant magnetization. The area enclosed by red lines favors black spins over white ones.

action like

$$V[(r_i, s_i), (r_j, s_j)] = \begin{cases} \infty & r_i = r_j \\ -s_i s_j & \|r_i - r_j\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

The external field can be any function of the lattice site, e.g., a fun pattern as in Figure 3.1. The symmetry group element on the ghost site can be represented by an integer-valued orthogonal matrix along with an integer-valued translation.

A basic cluster flip is performed by picking a reflection at random and a spin at random to seed the cluster. The spin residing on the site our seed spin is destined to land will certainly also join the cluster, because the relative change in their bond energy is infinite, and things progress from there. Notice that in this language, simulating a depleted lattice, e.g., the Blume–Capel

model, is natural. This formulation can be used to simulate models we could not use clusters with in the previous section, because their pair potential was asymmetric: the chiral Potts model has spins with a $\mathbb{Z}/p\mathbb{Z}$ symmetry, which does not have enough order-two elements to give ergodicity with ordinary clusters, but *simultaneous* inversion of space and spin flips are symmetries of the chiral Potts model with order two![194]

However, a severe inefficiency arises quickly in this scheme. Consider the seemingly trivial case of a trial move that lands a spin directly on top of an identical spin. If we were so foolish as to waste our time making these moves, what *should* happen is nothing: the net change in bond energy due to simply exchanging the spins is zero, so the bond activation probability for each neighbor should be zero. But that's not what happens when we move one spin at a time! When the seed spin leaves its original site, *all* bonds whose energies were changed in the move have their probabilities evaluated, both the very unhappy spin the seed landed on *and* the spins neighboring both sites! A spin that was energetically happy with its neighbor now sees a void, and the resulting change in bond energy may result in an activated bond and a cluster formation that *doesn't change any relative configuration of the spins*—what a calamity!

The solution is to construct *joint moves*. If at any time during the course of cluster formation a spin is asked to move to a place that frustrates an infinite energetic bond, rather than adding that bond to the queue for later processing, the degree of freedom at the other end of the bond is included in the elementary step. For a lattice gas, this process is fairly simple, and in fact was just the default procedure when the algorithm without a field was first

introduced: reflections that take spins on top one another just swap the spins, and all the bonds of both spins are considered for activation with an energy change equal to the net change for the swap.

This procedure preserves detailed balance because the construction of the joint moves is equally likely in the forward and backward processes; if a spin at site a impinged on one at site b in the forward process, the spin at site b would do so for site a in the reverse. And once the joint move is constructed, it's as if we are working with a strange degree of freedom in the usual course of the algorithm; an Ising model embedded into our original model in a particularly strange but nonetheless valid way.

Assessment of the efficiency of geometric cluster algorithms on lattice models without a field indicate a similar speedup that ordinary cluster algorithms offer, measuring $z \simeq 0.2$ for the 2D Ising model.[192] A careful study of the efficiency effect of adding a field using the method described here has not yet been made, but preliminary results like those in Figure 3.2 indicate that, at least for modestly strong fields, the algorithm is more efficient in a field near the critical point than at the critical point. Eyeballing the data in that one-system-size plot, it looks as though the efficiency falls off with a logarithm instead of a power law as in Figure 2.2, though this is a conclusion drawn from extremely limited data.

§23. Hard spheres, soft spheres, & other particles

Hard spheres are one of the most useful simple models in modern statistical physics, exhibiting all kinds of interesting structural and nonequilibrium tran-

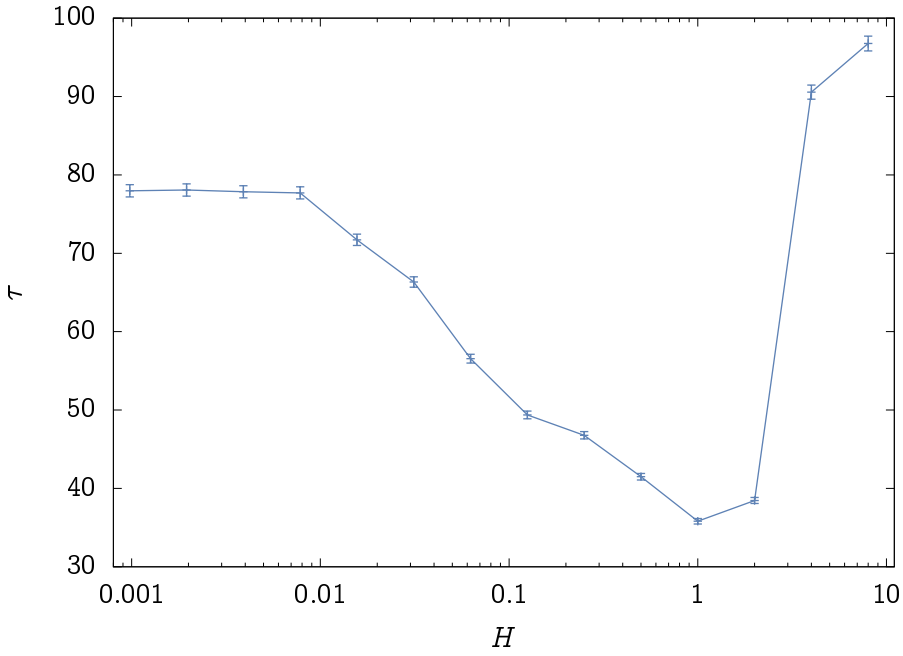

Figure 3.2: Autocorrelation times measured for a 32×32 2D fixed-magnetization Ising model at its critical temperature and in a field $U(x) = H\cos(2\pi x/L)$ as a function of H.

sitions. Spheres live in continuous space, either infinite, e.g., $\mathbb{R}^D$, or finite, e.g., $\mathbb{T}^D$. The order-two symmetries of space are again reflections, inversions, and π rotations, now about any point. Hard spheres have a radius s, and pair interaction of the form

$$V[(r_i, s_i), (r_j, s_j)] = \begin{cases} \infty & \|r_i - r_j\|/(s_i + s_j) < 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

The symmetry group element on the ghost site can be represented by a real orthogonal matrix and a real translation.

Because the only energies at play are infinity or zero in the absence of an

external potential, all microstates are either permitted with equal probability or not regardless of temperature. In fact, because all probabilities are zero and one, all the fancy tuning we did to preserve detailed balance is unimportant for a purely hard system, and one can dispense even with the need to use global symmetry transformations of order-two; *any* global transformation, e.g., small rotations or translations, are valid: for any transformation and any seed, the forward and reverse probabilities are always one, the energies of the initial and final states are always the same, and detailed balance is guaranteed. The formulation of a cluster algorithm for hard spheres in the absence of potentials was first done by Dress and Krauth and for soft spheres by Liu and Luijten, and both have been used in other studies.[195–199]

Hard spheres with hard external potential (also infinite or zero) lend one of the clearest ways to intuit what is going on with the cluster algorithm. Pick a seed, and try to move it using a global symmetry transformation. If it hits anyone, move them with the same transformation. Keep going until no one else gets hit. In the worst case, everyone moves but the state is still valid because the transformation was a symmetry of space. Otherwise, some subset of spheres is now oriented differently with respect to the rest. The addition of a hard external potential, like a box, doesn't change this picture at all: what is a hard box if not a big, weirdly shaped hard sphere? If one of the spheres hits the box, transform the box as well—the same considerations apply. A demonstration of these steps is shown in Figure 3.3. Studies have predicted the container-shape dependent free energy of fluids, and this method could provide a direct way to measure these.[200]

A more useful demonstration of this method is to simulate a common lab-

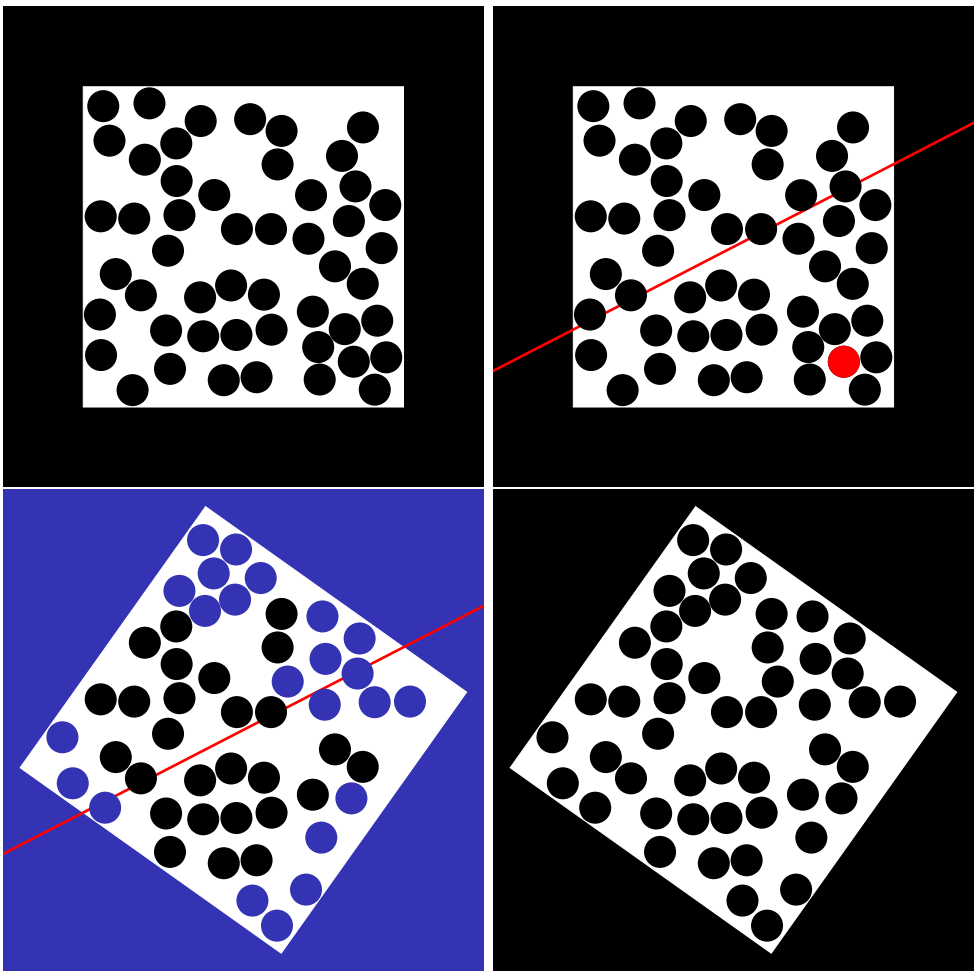

Figure 3.3: Hard spheres in a hard box. On the top left, they start in a legal configuration. On the top right, a cluster flip is started using the red seed and the red reflection plane. On the bottom left, the final positions of all transformed objects ore shown in blue, while those that were not transformed remain black. Having no further overlaps, the bottom right leaves the system in a new state.

oratory situation, which is spheres in a gravitational potential. Recent work has studied the phase boundary between solid, liquid, and *hexatic* states[201] of approximately hard spheres on slanted planes, and the cluster method with a field, perhaps along with other Monte Carlo steps, is a good candidate for speeding up simulations of such a set-up. Similar studies have been done for spheres in the absence of a field using event-chain Monte Carlo.[202–204] A snapshot of a simulation of hard spheres in a gravitational potential can be seen in Figure 3.4.

Soft spheres may seem like a corollary at this point, but in practice there are new complexities that make them worth addressing all on their own. A soft sphere model is defined by a pair potential that is a function of the distance between the spheres relative to their radius, or

$$V[(r_i, s_i), (r_j, s_j)] = v(\|r_i - r_j\|/(s_i + s_j)) \tag{3.6}$$

All the same considerations apply as for hard spheres. However, because soft spheres are often somewhat *attractive*, the same consideration that led to complexity in the choice of elementary steps for lattice models arises here: if we very nearly swap two spheres with each other in a move that would, in isolation, leave their close neighbors energetically happy, how do we ensure that their neighbors' bonds are not activated by the difference between their presence and the void left in the intermediate step?

Unlike in lattice gases, we don't necessarily have a deterministic procedure for this, since set of cases where one sphere lands *exactly* on top another should be vanishingly small if we hope to mix the system quickly. How can

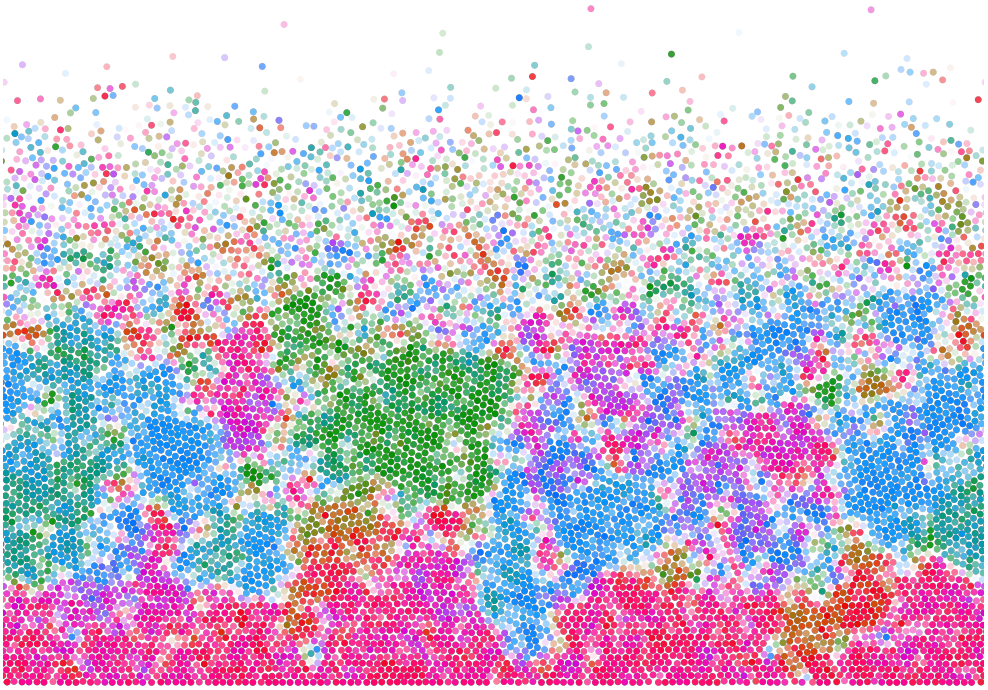

Figure 3.4: 10 000 hard spheres equilibrated(?) in a periodic box in a gravitational field using only cluster flips. Their hue and opacity is given by the argument and magnitude, respectively, of the hexatic order parameter $\Psi_i = \frac{1}{N_i} \sum_{\langle ij \rangle} e^{i6\theta_{ij}}$, where $N_i = \sum_{\langle ij \rangle} 1$ is the number of nearest neighbors of particle i and θ_{ij} is the angle the vector $r_i - r_j$ makes with the horizontal. The argument corresponds with the orientation of the local hexagonal crystal, and the magnitude corresponds roughly with how closely the particles' neighbors are to being close-packed.

we dynamically decide, when a sphere is about to be moved, whether to associate one near its destination with the move inherently or not? There turn out to be many schemes for doing this, but we will describe a simple one here, with an eye on expanding this in the future.

Here is the scheme: say you have popped a sphere from the stack and are now about to transform it to a new place. Before you do, evaluate the energy change that each of the affected bonds will experience. If the largest of these changes is ΔE, make instead a joint move incorporating the sphere on the other end of the bond with probably $\max\{0, 1 - e^{-\beta\Delta E}\}$. One can see that this preserves detailed balance the same way that normal cluster construction does. Any bond in a cluster that is activated as part of one of these joint moves remains the maximum energy bond for the same symmetry transformation in the ensuing configuration, and the probability of it being activated in the reverse direction is equal. Any bond in a cluster that is tried but *not* activated as part of a joint move will result in the new system in a bond with unfavorable energy ΔE and will be activated for a joint move with probability zero in the reverse process. It therefore follows that for such a system,

$$\frac{P_{\text{forward}}}{P_{\text{reverse}}} = \frac{1 - (1 - e^{-\beta\Delta E})}{1} = e^{-\beta\Delta E}, \tag{3.7}$$

which is precisely the relative Boltzmann factor associated with the increased bond energy as a result of the forward process. It is important that, if a bond is chosen for use in the formation of a joint move and the move fails to be constructed, that bond *not* be evaluated in the normal Wolff process,

since the probability of its second failure adds a second factor of $e^{-\beta\Delta E}$ to the *total* ratio of forward to reverse probabilities for the cluster flip; no double jeopardy!

Our choice to only built joint moves in pairs, and for those pairs to be constructed only from the bond with the largest energy change, were somewhat arbitrary, and in fact more complicated schemes may result in more efficient joint moves.[195] This sort of consideration is likely very important if we hope to use techniques like cluster flips at the densities encountered in solid and glassy state problems.

Dimers & more

Dimers are made from pairs of spheres joined together with some overlap. They're aren't particularly special for the purposes of this example, except that attempting to compute the overlap of two dimers is easy (it's the same as computing the overlap of four spheres) while computing the overlap of two ellipses is difficult.

The treatment of dimers is only somewhat more complicated than that of spheres. The position variable of the dimer is the same, and it has internal properties like a radius or offset between the component spheres, but it also has a dynamic property: its orientation. Symmetry transformations on space like a reflection transform both the positions of the dimers in a system, and their orientation.

The internal degrees of freedom of the dimer present a challenge for ergodicity when simulated with periodic boundary conditions. Since the only symmetries of the torus are π rotations and reflections along the principal

axes, *no global symmetry transformation exists than can tilt a dimer any way but at right angles.* Therefore, cluster dynamics in periodic boundary conditions cannot by itself explore a phase space that includes internal orientational degrees of freedom. Previous studies using cluster algorithms on dimers and other asymmetric particles, like cubes, have been restricted to lattice sites, or to keeping the faces parallel.[205–207]

If the torus is out of the picture for an ergodic cluster algorithm for dimers, free space doesn't necessarily look much better: the dimers can just scatter to the wind! Luckily, we have the ability to apply confining fields, which allow control over, e.g., the density of a simulation, with all the global symmetries of free space to work with. Examples of dimers relaxed in a linear potential both on periodic boundary conditions and in free space are shown in Figure 3.5. One could also simulate these systems on the surface of a sphere, where reflections across planes through the center of the sphere do provide ergodicity, though for dimers and other crystals the curved space would lead to topological defects, and has previously been seen to affect glassy properties in colloidal simulations.[208,209] However, anisotropic particles confined to a cell with boundaries also have interesting phases in their own right, and this algorithm is uniquely suitable for studying them.[210]

§24. The choice of reflections & other tweaks

Much more so than in the previous chapter, the freedom to choose the Ising embeddings with which to define clusters is extremely important in the efficient simulation of atomistic models. For a dense system of stiff particles,

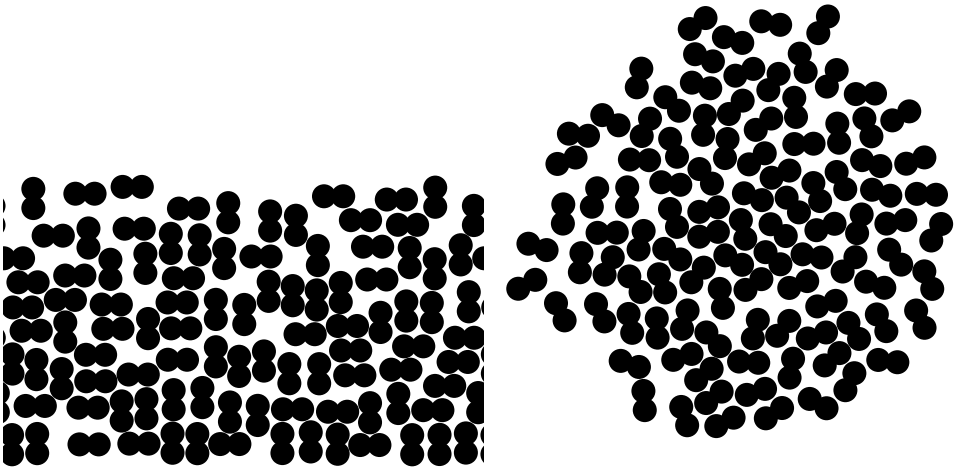

Figure 3.5: Dimers! (Left) An attempt to relax dimers in a gravitational field with periodic boundary conditions. (Right) Relaxing dimers in a radially symmetric 'gravitational' field in free space.

haphazard choice of reflections will lead to clumsy moves in which the whole system is pointlessly flipped this way and that without actually decorrelating it from its starting point.

When envisioning moves that will lead to fast mixing, it helps to think of the types of particle configurations we expect to see, and then push up to the limits allowed by the constraints of detailed balance. Recall, as per the previous chapter, that detailed balance requires that the distribution by which a transformation is chosen be invariant under the transformation, so that it is equally probable to choose the same transformation in the reverse process. Previously, we used a distribution of reflection planes in the O(n) models that was only a function of the absolute angle between the plane and the seed spin. Once the seed spin is flipped, it is equally likely we will choose the same plane to go in the reverse direction, as the absolute angle between

the flipped seed and the plane has not changed.

In dense simulations of solids, liquids, hexatics, or glasses, particles are locally close-packed or nearly so. Moves which insert a seed directly between two or more nestled particles are unlikely to lead to new configurations. A good scheme for seeds, then, is to pick a pair of particles and *nearly* swap them. This can be done either by picking an inversion point almost directly between them, or (if in real space) by picking a reflection plane that nearly brings one to the other. When the amount by which the inversion or reflections are perturbed from direct swaps is properly tuned to neither be too large nor trivially small, these pair seeds dramatically improve the situation compared to that of random transformations of random seeds. A systematic procedure for coming to an optimal or even nearly optimal tuning is ongoing work.

One might be tempted to continue this trend in a direction that favors *local* pairs, since nearly swapping distant pairs is unlikely to align with, e.g., a lattice direction. Unfortunately, making certain pairs more probable based on dynamic properties like their relative position is almost always illegal. The problem is subtle—after all, the distance between any pair of particles that are both transformed by a global symmetry is invariant under that transformation. But unless every other particle is also transformed (which we don't want anyway), the distances between *other* pairs of particles is changed, and therefore also is the normalization for the distribution of possible reflection planes! Consider a system of three particles, two that begin nearby but both farther from a third. If the probability of choosing a pair seed is proportional to the distance between that pair, then the two that begin nearby are more likely to be chosen as a seed than the pairs consisting of either with the dis-

tant third. Once that nearby pair is transformed, the move could result in one member of that pair ending very close to the formerly distant third particle. Our procedure would then dictate that the pair consisting of that particle and the newly close third one be more probable to seed a cluster in the reverse process than the original transformed pair, spoiling detailed balance.

Some freedom does remain, though, and in a sense you can effectively favor local moves. Once a pair of particles is selected at random to seed the cluster, the amount by which their near exchange is perturbed can be distributed by a function of the distance between them, or their relative orientation, or anything else invariant under the transformation. Seed pairs within a certain proximity can be assigned larger perturbations, for instance, or those outside it can be left alone.

These considerations also apply to other efficient seed transformations. Another, simpler, example is to simply nudge a particle to a position near its current one by picking a reflection or inversion nearby its center. If the reflection plane or inversion point is made to preserve the value of the external field, even if only for that seed particle, then the magnitude of the nudge can depend on the height in that field. Considerations like this may be important for large-scale simulation of sedimented particles, where denser regions at lower points merit smaller perturbations from their current position than rarefied regions higher up.

§25. Potential relevance to swap Monte Carlo

Cluster algorithms for particles resemble in many ways the hugely successful swap algorithms that have recently revolutionized the computer modelling of structural glasses.[211–217] These computer glasses are formed by equilibrating soft spheres with quenched polydispersity at extremely low temperatures and high densities, where the timescales of any traditional dynamics diverge. The dynamics of glasses in this regime involves particles trapped in extremely deep wells, and equilibration becomes impossible on computer timescales.

Swap is a fairly simple idea. Every so often, pick a pair of differently sized particles in the system, and try to swap them, with a probability of accepting the swap that corresponds to the usual Metropolis step for the energy change the swap would create. If the rate of these swap steps compared to regular evolution is carefully tuned (sound familiar?), a fantastic speedup occurs and equilibration timescales reduce by many orders of magnitude.

With an appropriately chosen reflection plane or inversion point, cluster moves are alike to swap: two particles may have their positions exchanged exactly. Metropolis swap would accept the change with probability

$$\exp\left(-\beta\sum_b \Delta E_b\right) \tag{3.8}$$

for bond energy changes ΔE_b, while a cluster swap would accept the change *as being the end of the cluster* (and therefore resulting in the same end state

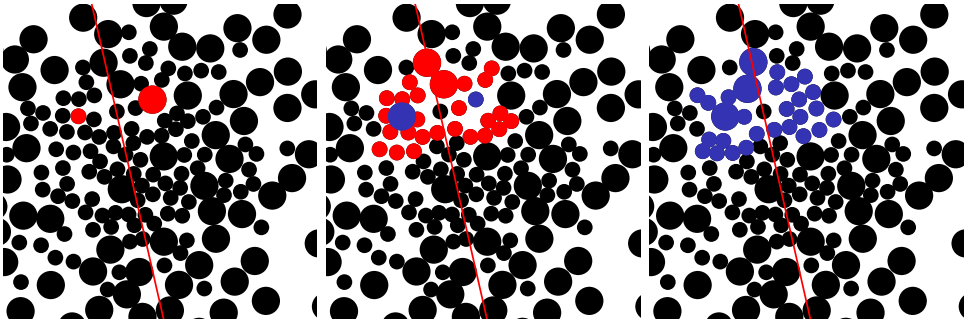

Figure 3.6: An example of nontrivial differing behavior between a cluster 'swap' and a true swap step with some soft, polydisperse spheres. (Left) The red spheres about to be swapped by reflection over the red line. (Center) Their swap (now blue) seeds a cluster containing all of the red spheres. (Right) A new equilibrium configuration of spheres is reached.

as the ordinary swap) with probability

$$\exp\left(-\beta \sum_{b,\Delta E_b>0} \Delta E_b\right) \tag{3.9}$$

These probabilities coincide if all bonds are happy with the change or none of them are, but in intermediate cases the cluster swap is more likely to abandon that particular microstate than the metropolis swap. However, the *way* that the cluster method abandons the microstate is completely different: it keeps the change but activates unhappy bonds adjacent to the swapped particles, potentially swapping more than just the original two. In the worst case scenario, the swap really is 'rejected' because *all* particles are added to the cluster, and the system is left where it started at a much higher cost than a metropolis rejection step.

It's possible that—at the high densities and with the polydispersity that are used in these glass simulations—cluster flips will never produce nontriv-

ial changes without enormous cost. But some models of the glass transition indicate that there is a relevant length scale over which cooperative rearrangements occur,[218] and inducing dynamics at those length scales seems like something a cluster-like method could accomplish. It is also established that the geometric clusters moves describe here outperform ordinary Monte Carlo for high polydispersity, just like swap does.[196,199,207,219,220] Would it require very careful tuning? Definitely. Perhaps even elaborate cluster-shaping steps like those discussed briefly at the end of the previous section are necessary. But the extension of these methods to contribute to the equilibration of computer glasses is something we are looking into.

§26. Pressure and chemical potential

So far we have embedded Ising models into a lot of strange places, but this section may well take the cake. Here, we will describe how a similar trick to that used previously—introducing a new degree of free that restores a global symmetry to a symmetry-broken model—can be used to describe ensembles of polydisperse particles at constant pressure. If the method implied by the formal extension described was efficient, it would be a boon for cluster simulations, since systems of constant pressure are more often of interest than those at constant volume. The method is not a panacea, though: the polydispersity is a *necessary* feature, and it is an annealed polydispersity at that—not likely to exactly coincide with the common soft-sphere models of glassy physics.

Consider a model of soft spheres with pair potential

$$V[(r_i, s_i), (r_j, s_j)] = v(\|r_i - r_j\|/(s_i + s_j)) \tag{3.10}$$

where $r \in \mathbb{R}^D$ is the position and s is the radius. Suppose further that the spheres are confined to a potential $U_L(r)$ of characteristic size L, e.g., the side lengths of the box. This model, in isolation, has a global symmetry of a kind we haven't discussed yet: scale invariance. That is, a configuration of particles $\{(r, s)\}$ and box size L has the same energy as one with $\{(\Lambda r, \Lambda s)\}$ and box size ΛL for any constant Λ. If we allow the particle and box size to be dynamic degrees of freedom, this model isn't very interesting: entropy dictates the box blows up to infinity and the particles spread out into a rarefied gas with no characteristic size.

We therefore want to constrain both the box size and the particle sizes. The former is done by the pressure, which contributes a term pL^D to the Hamiltonian. The latter is done by a term of the form

$$\sum_i f(s_i) \tag{3.11}$$

with $f(s) = -\delta(s - \bar{s})$ for monodisperse spheres but could be taken to be Gaussian, e.g., $f(s) \propto e^{-(s-\bar{s}^2)/\delta s^2}$. Both of these terms break the scale invariance described previously. We will restore it, as before, by introducing a new degree of freedom into the system: a positive real number ζ that enters the Hamiltonian like

$$\mathcal{H} = \sum_{ij} V[(r_i, s_i), (r_j, s_j)] + \sum_i U_L(r_i) + p(\zeta L)^D + \sum_i f(\zeta s_i) \tag{3.12}$$

and where the action of a dilation by Λ of space on ζ is defined by $\zeta \mapsto \zeta/\Lambda$.

This ghost variable couples to every sphere with pair potential

$$\tilde{V}[(r_i, s_i), \zeta] = f(\zeta s_i) \tag{3.13}$$

and to the *other* ghost variable associated with the confining potential like $p(\zeta L)^D$.

We can imagine what a cluster 'flip' using this ghost site would be like. Starting with a seed particle, we pick a dilation of space about a point near the particle. The particle grows a bit and translates slightly, causing it to potentially bump other particles in its vicinity, that also dilate. If the process goes too far and too many particles have ventured from their characteristic size, the ghost site is triggered and the scale at which all particles and the box itself are measured against is increased.

This is where the crank we've been turning to produce algorithms hits a snag: the symmetry group whose transformations would yield changes in the particle and system sizes—the multiplicative group of positive real numbers—does not have *any* elements of order two.* This means that rejection-free clusters of the type we have been constructing cannot be formed. It helps at this point to review why: when the bond probabilities are tuned appropriately, the probability of the forward and reverse process have a ratio

$$\frac{P(\mathbf{s} \to \mathbf{s}')}{P(\mathbf{s}' \to \mathbf{s})} = \frac{e^{-\mathcal{H}(\mathbf{s}')}}{e^{-\mathcal{H}(\mathbf{s})}} \prod_{\{i,j\} \in C} \frac{p_r(s_i, s_j)}{p_{r^{-1}}(s_i', s_j')} \tag{3.14}$$

where C is the set of activated bonds that make up the interior of the activated

*If only particle sizes were complex...

cluster. Every term in the product usually vanishes, since for a group element of order two $r = r^{-1}$, and since every pair s_i' and s_j' that are both members of the cluster have not changed their bond energy the ratio is one. If $r \neq r^{-1}$ we can no longer assume this to be true, and detailed balance would not hold if we attempted to build clusters and 'flip' them without rejection. We can still build clusters, however, if we only accept them with a relative probability that cancels out this ratio. So what is that acceptance rate?

For bonds between regular particles, the forward acceptance probability is

$$\begin{aligned} p_\Lambda[(r_i, s_i), (r_j, s_j)] &= 1 - \exp\Big\{ -\beta(V[\Lambda \cdot (r_i, s_i), (r_j, s_j)] - V[(r_i, s_i), (r_j, s_j)])\Big\} \\ &= 1 - \exp\left\{ -\beta\left[v\left(\frac{\|\Lambda r_i - r_j\|}{\Lambda s_i + s_j}\right) - v\left(\frac{\|r_i - r_j\|}{s_i + s_j}\right)\right]\right\} \end{aligned} \tag{3.15}$$

while the reverse is

$$\begin{aligned} p_{\Lambda^{-1}}[(r_i', s_i'), (r_j', s_j')] &= 1 - \exp\Big\{ -\beta(V[\Lambda^{-1} \cdot (r_i', s_i'), (r_j', s_j')] - V[(r_i', s_i'), (r_j', s_j')])\Big\} \\ &= 1 - \exp\left\{ -\beta\left[v\left(\frac{\|r_i - \Lambda r_j\|}{s_i + \Lambda s_j}\right) - v\left(\frac{\|r_i - r_j\|}{s_i + s_j}\right)\right]\right\} \\ &= p_\Lambda[(r_j, s_j), (r_i, s_i)] \end{aligned} \tag{3.16}$$

Here we see another interpretation of why transformations of order two are so important: neither the pair potential nor the activation probability are symmetric under exchanging the group action in their arguments. However, this offers us our rejection step. If, every time a bond as activated because

site i has been transformed and modified the energy of its bond with site j, we stop the cluster process and throw it away with probability

$$1-p_\Lambda[(r_j,s_j),(r_i,s_i)] = \exp\left\{-\beta\left[v\left(\frac{\|r_i-\Lambda r_j\|}{s_i+\Lambda s_j}\right)-v\left(\frac{\|r_i-r_j\|}{s_i+s_j}\right)\right]\right\} \quad (3.17)$$

then the offending factor is cancelled in the ratio of forward to reverse probabilities. This is the probability that the bond *wouldn't* have been activated because site j had been transformed and modified its energy with site i. This probability is one if we have activated a bond because of site i's transformation that would not have had its energy changed because of site j's transformation, which can certainly happen, as in Figure 3.7. That our cluster growth is cancelled so frequency is bad news for this idea, but there is a degree of freedom remaining: the exponential factor in all of the activation probabilities can be freely scaled by some constant $0 < a \leq 1$. If a is taken to be relatively small, we will accept more bonds and need to evaluate this probability more often, but it will no longer be rejection-definite.

Whether this strange formal use of the ghost site idea performs in actual simulation comparably to existing methods for annealed polydispersity at constant pressure is unknown. This exploration for now is a demonstration on how important the existence of order-two transformations is for the cluster method—rejection steps are much easier to construct when the symmetry is broken instead!

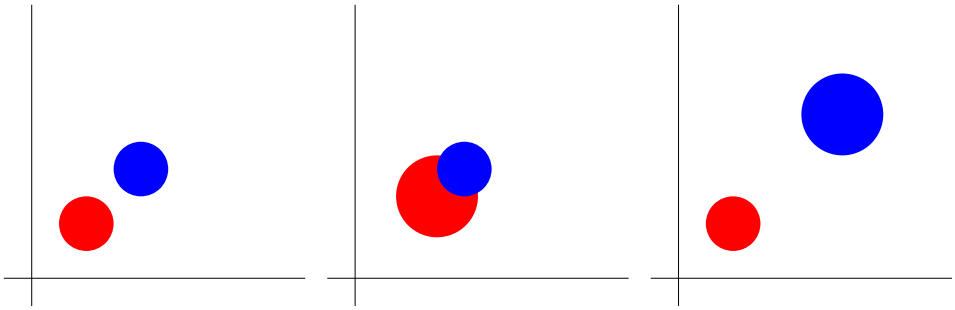

Figure 3.7: A demonstration of the asymmetry of dilation. (Left) The starting configuration of two spheres. (Center) If the red sphere is dilated with respect to the origin by a factor 1.5, it intersects the blue sphere and their relative energies change. (Right) If the blue sphere is so dilated, the spheres do not intersect and their relative energy remains the same.

CHAPTER IV

CRITICAL DROPLETS & SCALING

Scaling and the renormalization group usually focus their attention on continuous phase transitions. This makes sense, because the emergent scale invariance at continuous phase transitions requires their use to make quantitative predictions. Proportionally less attention has been paid to scaling implications in the vicinity of abrupt transitions, which do not need scaling machinery to describe. However, abrupt transitions do not always occur nicely separated from scale-invariance. In the most canonical model of both an abrupt transition and a critical phenomenon, the two are nestled up against each other where the abrupt separatrix merges into the critical point.

The Ising model is the canonical example of a system with a continuous phase transition, and the study of its singular properties marked the first success of the renormalization group (RG) method in statistical physics.[222,223] Its status makes sense: it's a simple model whose continuous phase transition contains all the essential features of more complex ones, but admits RG methods in a straightforward way and has exact solutions in certain dimensions and for certain parameter restrictions. The Ising critical point is not simply

*Portions of the material of this chapter were posted online in J Kent-Dobias and JP Sethna, "Essential singularities in universal scaling functions at the Ising coexistence line", (2017), arXiv:1707.03791v2 [cond-mat.stat-mech].

a continuous transition: it also ends a line of abrupt phase transitions extending from it at zero field below the critical temperature. Though typically neglected in RG scaling analyses of the critical point, we demonstrate that the strange and subtle analytic properties of the abrupt transition well-describe the universal scaling functions of the Ising model in a field, and develop a perturbative approximate form for those scaling functions that should converge better than any analytic approximation near the abrupt transition line.

§27. Metastability and essential singularities

RG analysis predicts that the singular part of the free energy per site F as a function of reduced temperature $t = 1 - T_c/T$ and field $h = H/T$ in the vicinity of the critical point takes the scaling form $F(t,h) = |t|^{2-\alpha}\mathcal{F}(h|t|^{-\beta\delta})$ for the low temperature phase $t < 0$.[189] When studying the properties of the Ising critical point, it is nearly always assumed that the universal scaling function $\mathcal{F}$ is analytic, i.e., has a convergent Taylor series. However, it has long been known that there exists an essential singularity in $\mathcal{F}$ at zero argument, though its effects have long been believed to be undetectable by direct thermodynamic measurements,[224] or simply just neglected.[225–232] With careful analysis, we have found that assuming the presence of the essential singularity is predictive of the scaling form of, for instance, the susceptibility and magnetization.

The provenance of the essential singularity can be understood using the ideas of critical droplet theory for the decay of an Ising system in a metastable state, i.e., an equilibrium Ising state for $T < T_c$, $H > 0$ subjected to a small negative external field $H < 0$. The existence of an essential singularity has

also been suggested by transfer matrix studies,[187,233–236] RG methods,[237,238] perturbative conformal field theory,[239,240] and direct study of the partition function.[241] A different kind of essential singularity is known to exist in the zero-temperature susceptibility.[242–247] It has long been known that the decay rate Γ of metastable states in statistical mechanics is often related to the metastable free energy F by $\Gamma \propto \operatorname{Im} F$.[248–252]

'Metastable free energy' can be thought of as either an analytic continuation of the free energy through the abrupt phase transition, or restriction of the partition function trace to states in the vicinity of the local free energy minimum that characterizes the metastable state. We've already seen an empirical definition of metastable properties in §19 of this very thesis! In any case, the free energy develops a nonzero imaginary part in the metastable region. Heuristically, this can be thought of as similar to what happens in quantum mechanics with a non-unitary Hamiltonian: the imaginary part describes loss rate of probability that the system occupies any 'accessible' state, which corresponds to decay.

Critical droplet theory posits that the metastable state decays when a domain of the equilibrium state forms whose surface-energy cost for growth is outweighed by bulk-energy benefits. There is numerical evidence that, near the critical point, these droplets are spherical,[253,254] and we will assume that we are in a regime where lattice effects that would make the droplets, e.g., square-like are not relevant. One can see the following argument in the context of a course-grained theory.

The free energy cost of the surface of a spherical droplet of radius R is $\Sigma S_d R^{d-1}$ and that of its bulk is $-\Delta M|H|V_d R^d$, where S_d and V_d are the

surface area and volume of a $(d-1)$-sphere, respectively, ΔM is the change in magnetization between the equilibrium and metastable states, and Σ is the surface tension of the equilibrium–metastable interface. The critical droplet size is found by minimizing the change with respect to R, with

$$0 = \left.\frac{\partial \Delta E}{\partial R}\right|_{R=R_c} = (d-1)\Sigma S_d R_c^{d-2} - d\Delta M|H|V_d R_c^{d-1} \tag{4.1}$$

giving the critical droplet radius

$$R_c = \frac{d-1}{d}\frac{S_d}{V_d}\frac{\Sigma}{\Delta M|H|} \tag{4.2}$$

The free energy of the critical droplet is then given by substituting R_c for R in the expressions above, or

$$\Delta E_c = \left(\frac{S_d \Sigma}{d}\right)^d \left(\frac{d-1}{V_d \Delta M|H|}\right)^{d-1} \tag{4.3}$$

As the metastable field is brought to zero, the size of the droplet required to collapse the state diverges, as does the free energy cost of forming it.

The previous analysis amounts to solving a mean-field and coarse-grained picture, where droplets are little uniform spheres in a uniform background with infinitely thin boundaries. Introducing an interface thickness complicates the picture by making the energy contribution of both the bulk and surface have more detailed R dependence,[255] but in the limit of small field these introduce analytic corrections to R_c and ΔE_c and don't affect the singular behavior in $|H|$.

The addition of fluctuations has the possibility of changing the picture dra-

matically. First, it is reasonable to suspect that, near the critical point, the shape of critical droplets would deviate from circular and instead have some fractal structure. The introduction of a fractal bulk and surface dimension to the droplet idea above for $d_b - d_s \neq 1$ *would* change the singularity of ΔE_c in $|H|$. However, numeric evidence both directly measuring the imaginary free energy using transfer matrix methods[253] and indirectly by measuring the decay of the metastable state[256,257] find close agreement between the predictions of the spherical theory and data.

A second effect of fluctuations is to give the critical droplet a characteristic entropy independent of the entropy associated with the equilibrium state it contains in its bulk. The principal contributions to this entropy are from zero modes of the droplet. First, there is an entropy associated with the position of the droplet, which is present event without fluctuations but is independent of the droplet size, and so doesn't affect the height of the free energy barrier as a function of H. Second, there is an entropy associated with long-range fluctuations of the droplet surface. These do depend on R_c and therefore H, and their influence can be found by computing the partition function for a field theory of the surface degrees of freedom in the vicinity of the spherical droplet solution, following instanton methods.[258] The result gives $\Delta S_c \sim b \log R_c$, where $b = (d-3)d/2$ for $d = 2, 4$ and $b = 7/3$ for $d = 3$.[258,259] This gives a total free energy cost for the droplet of the form

$$\Delta F_c = \Delta E_c - T\Delta S_c \tag{4.4}$$

Droplet theory then predicts an asymptotic decay of the metastable state as

following an Arrhenius form

$$\Gamma = \Gamma_0 e^{-\beta \Delta F_c} = \Gamma_0 e^{\Delta S_c} e^{-\beta \Delta E_c} \tag{4.5}$$

In the decay rate, the singularity in E_c as a function of $|H|$ becomes an essential singularity. Making the association between the decay rate and an imaginary component of the free energy, we see the emergence of an essential singularity in the free energy as well.

§28. Scaling considerations

Though we assumed in the previous section that we were working close enough to the critical point for droplets to be spherical, it wasn't strictly necessary that we be near the critical point. The evidence for nearly spherical droplets is maintained in the square lattice 2D Ising model down to nearly half of T_c.[253] However, we will now fix our attention to the critical region, and invoke many of the scaling properties of the Ising fixed point to work the imaginary free energy above into an ansatz for the asymptotic form of the free energy's universal scaling function, ignoring corrections to scaling.

As the critical point is approached from below, both the magnetization and the surface free energy vanish with power laws in the reduced temperature t, multiplied by universal scaling functions of the invariant $h|t|^{-\beta\delta}$.* They

*For now, we will suppose we are sufficiently close to the critical point to take $u_t \simeq t$ and $u_h \simeq h$.

therefore have the form

$$\begin{aligned} \Sigma/T &= |t|^{\mu}\mathcal{S}(h|t|^{-\beta\delta}) \\ M &= |t|^{\beta}\mathcal{M}(h|t|^{-\beta\delta}) \end{aligned} \tag{4.6}$$

Asymptotically close to the abrupt transition, the metastable magnetization is equal to minus the equilibrium magnetization and $\Delta M \simeq 2M$. Combined with the exponent relations $\mu = -\nu + \gamma + 2\beta$,[260] $2 = \alpha + 2\beta + \gamma$, and $\alpha + \beta(1+\delta) = 2$,[189] the critical droplet radius can be brought to the form

$$\begin{aligned} R_{\mathrm{c}} &= \frac{d-1}{2d}\frac{S_d}{V_d}\frac{T|t|^{\mu}\mathcal{S}(h|t|^{-\beta\delta})}{|t|^{\beta}\mathcal{M}(h|t|^{-\beta\delta})|H|} \\ &= |t|^{-\nu}\frac{d-1}{2d}\frac{S_d}{V_d}\frac{\mathcal{S}(h|t|^{-\beta\delta})}{\mathcal{M}(h|t|^{-\beta\delta})}(-h|t|^{-\beta\delta})^{-1} \\ &= |t|^{-\nu}\bigl[\mathcal{R}(h|t|^{-\beta\delta})\bigr]^{-1} \end{aligned} \tag{4.7}$$

which suggests a critical scaling form for the droplet size. Notice that R_{c} is proportional to the correlation length $\xi = \xi_0^- |t|^{-\nu}$ for fixed $h|t|^{-\beta\delta}$. Upon substitution into the critical energy, we have

$$\begin{aligned} \Delta E_{\mathrm{c}} &= \left(\frac{S_d T|t|^{\mu}\mathcal{S}(h|t|^{-\beta\delta})}{d}\right)^{d}\left(\frac{d-1}{2V_d|t|^{\beta}\mathcal{M}(h|t|^{-\beta\delta})|H|}\right)^{d-1} \\ &= T\frac{S_d^d(d-1)^{d-1}}{(2V_d)^{d-1}d^d}\frac{\mathcal{S}^d(h|t|^{-\beta\delta})}{\mathcal{M}^{d-1}(h|t|^{-\beta\delta})}(-h|t|^{-\beta\delta})^{-(d-1)} \\ &= T\bigl[\mathcal{G}(h|t|^{-\beta\delta})\bigr]^{-(d-1)} \end{aligned} \tag{4.8}$$

for the free energy change due to the critical droplet. Based on the way we have defined the scaling function $\mathcal{R}$ and $\mathcal{G}$, and since $\mathcal{M}$ and $\mathcal{S}$ go to nonzero values as the field goes to zero, these new scaling functions have expansions

of the form

$$\begin{aligned} \mathcal{R}(X) &= -R_0 X + \mathcal{O}(X^2) \\ \mathcal{G}(X) &= -BX + \mathcal{O}(X^2) \end{aligned} \tag{4.9}$$

where the linear coefficients are defined by

$$\begin{aligned} R_0 &= \frac{2d}{d-1}\frac{V_d}{S_d}\frac{\mathcal{M}(0)}{\mathcal{S}(0)} \\ B &= \frac{2V_d\mathcal{M}(0)}{d-1}\left(\frac{d}{S_d\mathcal{S}(0)}\right)^{d/(d-1)} \end{aligned} \tag{4.10}$$

Neither of these are universal, since the scaling functions are only universal up to a constant factor, but the combination

$$\frac{B}{R_0\xi_0^-} = \left(\frac{d}{S_d(\xi_0^-)^{d-1}\mathcal{S}(0)}\right)^{1/(d-1)} \tag{4.11}$$

is universal, since $\mathcal{S}(0)(\xi_0^-)^{d-1}$ is a universal amplitude ratio.[261] The critical droplet size written in units of the correlation length and taken to the $(d-1)$th power over the critical droplet energy in units of temperature

$$\frac{(R_c/\xi)^{d-1}}{\Delta E_c/T} \simeq \left(\frac{B}{R_0\xi_0^-}\right)^{d-1} \tag{4.12}$$

is likewise universal.

Based on the considerations in the last section, the imaginary part of the free energy takes the form

$$\operatorname{Im} F \propto \Gamma = \Gamma_0 e^{-\Delta F_c/T} = \tilde{\Gamma}_0 R_c^b e^{-\mathcal{G}(h|t|^{-\beta\delta})^{-(d-1)}} \tag{4.13}$$

If the imaginary part of the free energy is part of the universal scaling around the Ising critical point, it must scale like $|t|^{2-\alpha}$, and we therefore suspect that it ultimately can be written

$$\begin{aligned}\operatorname{Im} F &\propto |t|^{2-\alpha+b\nu} R_c^b e^{-\mathcal{G}(h|t|^{-\beta\delta})^{-(d-1)}} \\ &\propto |t|^{2-\alpha} \mathcal{R}(h|t|^{-\beta\delta})^{-b} e^{-\mathcal{G}(h|t|^{-\beta\delta})^{-(d-1)}} \\ &= |t|^{2-\alpha} \operatorname{Im} \mathcal{F}(h|t|^{-\beta\delta})\end{aligned} \tag{4.14}$$

where we have defined the scaling function

$$\operatorname{Im} \mathcal{F}(x) = A\Theta(-x)\mathcal{R}(x)^{-b} e^{-\mathcal{G}(x)^{-(d-1)}} \tag{4.15}$$

where Θ is the Heaviside theta function, required for the extension to all values of x. Expanding in x, we arrive at the familiar form of the essential singularity[262,263]

$$\operatorname{Im} \mathcal{F}(x) = A\Theta(-x)\Big[R_0 x + \mathcal{O}(x^2)\Big]^{-b} \exp\Big\{ -\Big[- Bx + \mathcal{O}(x^2)\Big]^{d-1}\Big\} \tag{4.16}$$

In 2D, the relative coefficient for the $\mathcal{O}(x^2)$ term is known from perturbative conformal field theory.[239]

§29. On 'simplest' scaling forms

In the following, we are going to take (4.15) and, assuming that it correctly captures the singularity at $x = 0$, use it to write down approximations of the real free energy that likewise capture that singularity. In working with this form, however, we will want to make truncations and simplifications of the

arbitrary functions it is built from. It is therefore worth investigating what effect such truncations and simplifications have on the singularity.

We will make this explicit in the following sense: we will say that a truncation $\operatorname{Im}\tilde{\mathcal{F}}$ of the free energy preserves its singular structure if there exists a near-identity analytic transformation $\tilde{x}(x) = x + \mathcal{O}(x^2)$ such that

$$\operatorname{Im}\tilde{\mathcal{F}}(\tilde{x}(x)) = \operatorname{Im}\mathcal{F}(x) \tag{4.17}$$

With such a transformation, one could consistently take the imaginary part of the free energy, simplify it by transforming the coordinates, bring the simplified version through, e.g., a Kramers–Kronig transformation and then invert the coordinate transformation on the other side to recover the untruncated version.

One could choose such a function to either be $\tilde{x}(x) = \mathcal{G}(x)/B$ or $\tilde{x}(x) = \mathcal{R}(x)/R_0$ to greatly simplify the form of the free energy, but both cannot be true at once. Taking the former, which ensures only *one* singular exponential factor is present in the scaling form and not $d-1$ of them, provides

$$\operatorname{Im}\tilde{\mathcal{F}}(\tilde{x}) = A\mathcal{R}(B\mathcal{G}^{-1}(\tilde{x}))^{-b}e^{-1/(-B\tilde{x})^{d-1}} = A\tilde{\mathcal{R}}(B\tilde{x})^{-b}e^{-1/(-B\tilde{x})^{d-1}} \tag{4.18}$$

for a function $\tilde{\mathcal{R}}$. Excepting the singular exponential factor, this has a series expansion of the form

$$\tilde{\mathcal{R}}(X)^{-b} = (-X)^{-b}\sum_{n=0}^{\infty} r_n X^n \tag{4.19}$$

We will start by truncating this series to its smallest order, but we should ex-

pect corrections that come from terms like $X^{n-b}e^{-1/(-B\tilde{x})^{d-1}}$. More informed corrections might come from considering more of the free energy's analytic properties in the vicinity of the abrupt transition line.[264]

§30. Recovering the real free energy

Griffith's analyticity guarantees that the real part of $\mathcal{F}$ is analytic in the upper complex plane of its argument,[265] and therefore the real part of F in the equilibrium state can be extracted from this imaginary metastable free energy using the Kramers–Kronig relation

$$\operatorname{Re}\mathcal{F}(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\operatorname{Im}\mathcal{F}(x')}{x'-x}\,dx' = \frac{1}{\pi}\int_{-\infty}^{0} \frac{\operatorname{Im}\mathcal{F}(x')}{x'-x}\,dx' \tag{4.20}$$

This relationship has been used to compute high-order moments of the free energy with H in good agreement with transfer matrix expansions away from the critical point.[266] Here, we will use it to get *functional forms* to iteratively approximate the whole real free energy.

Before we begin, let's look at some of the weaknesses of this technique. We have an asymptotic form for the imaginary free energy scaling function at zero argument, and we are looking to use it for approximating the real free energy via an integral over all values of the imaginary free energy. Since we have a poor understanding of the imaginary free energy at large x, we shouldn't expect results that rely on its value at large x to be reliable. This is why the Kramers–Kronig relation is most reliable for high derivatives of the

free energy, evaluated at zero: those derivatives in general are

$$\frac{d^n}{dx^n}\operatorname{Re}\mathcal{F}(x) = \frac{n!}{\pi}\int_{-\infty}^{0}\frac{\operatorname{Im}\mathcal{F}(x')}{(x'-x)^{n+1}}\,dx' \tag{4.21}$$

Evaluated at $x = 0$, the integrand becomes more and more strongly peaked around $x' = 0$, evaluating the imaginary part exactly where we know it best.[267,268] Lower derivatives rely more heavily on parts of the function we know worst. Therefore, we should expect low-order products of these transformations to be inaccurate, and need corrections.

In order to better facilitate the addition of these corrections, we will start by using the integral relation to match the scaling function for the susceptibility, which then will have the zeroth and first order corrections for the free energy emerge from constants of integration. Its scaling function is given by

$$\begin{aligned}\mathcal{Y}(x) = T\chi|t|^{-\gamma} &= T|t|^{-\gamma}\frac{\partial M}{\partial H} = -|t|^{-\gamma}\frac{\partial^2 F}{\partial h^2}\\ &= -|t|^{2-\alpha-\gamma}\frac{\partial^2\mathcal{F}}{\partial h^2} = -\frac{\partial^2\mathcal{F}}{\partial x^2} = -\frac{2}{\pi}\int_{-\infty}^{0}\frac{\operatorname{Im}\mathcal{F}(x')}{(x'-x)^3}\,dx'\end{aligned} \tag{4.22}$$

The other benefit of progressing using the susceptibility scaling function rather than the free energy is that, for the 2D truncated form of the imaginary free energy, the free energy integral does not converge, while this one does.

For a first approximation, let's take the full truncation of the imaginary free energy and run it through this machine. These functions can be integrated

explicitly in terms of named functions, giving

$$\begin{aligned}
\mathcal{Y}_{2\text{D}}(x/B) &= \frac{AB^2}{\pi x^3}\Big[x(x-1) - e^{1/x}\,\mathrm{Ei}(-x^{-1})\Big] \\
\mathcal{Y}_{3\text{D}}(x/B) &= \frac{AB^2}{54\pi x^8}\Big[e^{-1/x^2}(18-69x^2+35x^4)(2x\Gamma(2/3)\,\mathrm{E}_{5/3}(-x^{-2}) - 3\Gamma(7/6)\,\mathrm{E}_{7/6}(-x^{-2})) \\
&\qquad\qquad 6x^2(x(13x^2-6)\Gamma(2/3) + 3(3-8x^2)\Gamma(7/6))\Big] \\
\mathcal{Y}_{4\text{D}}(x/B) &= \frac{9AB^2}{8\pi^3 x} G^{4,3}_{0,0}\left(\begin{matrix} 0 & 1/3 & 2/3 & \\ 2/3 & 1 & 1 & 4/3 \end{matrix}\,\middle|\, x^{-3}\right)
\end{aligned} \tag{4.23}$$

where Ei and E are exponential integrals and G is the Meijer G function. In what follows, we will restrict our discussion to only the two and three dimensional Ising models, for which we have data to compare with. All remarks apply equally well to the four-dimensional model.

§31. Comparison with data

How predictive are these scaling forms in the proximity of the critical point and the abrupt transition line? We simulated the 2D and 3D Ising models on a square or cubic lattice using the Wolff algorithm modified for use in an external field, described in Chapter 2. Data was then taken for susceptibility and magnetization for $T_\mathrm{c} - T, H \leq 0.1$. This data, rescaled as appropriate to collapse onto a single curve, is plotted in Fig. 4.1.

Without any sort of fitting, we can fully restrain our truncated scaling forms. First, the constant B can be set. For the 2D Ising model on a square lattice, exact results at zero temperature have $\mathcal{S}(0) = 4/T_\mathrm{c}$ and $\mathcal{M}(0) =$

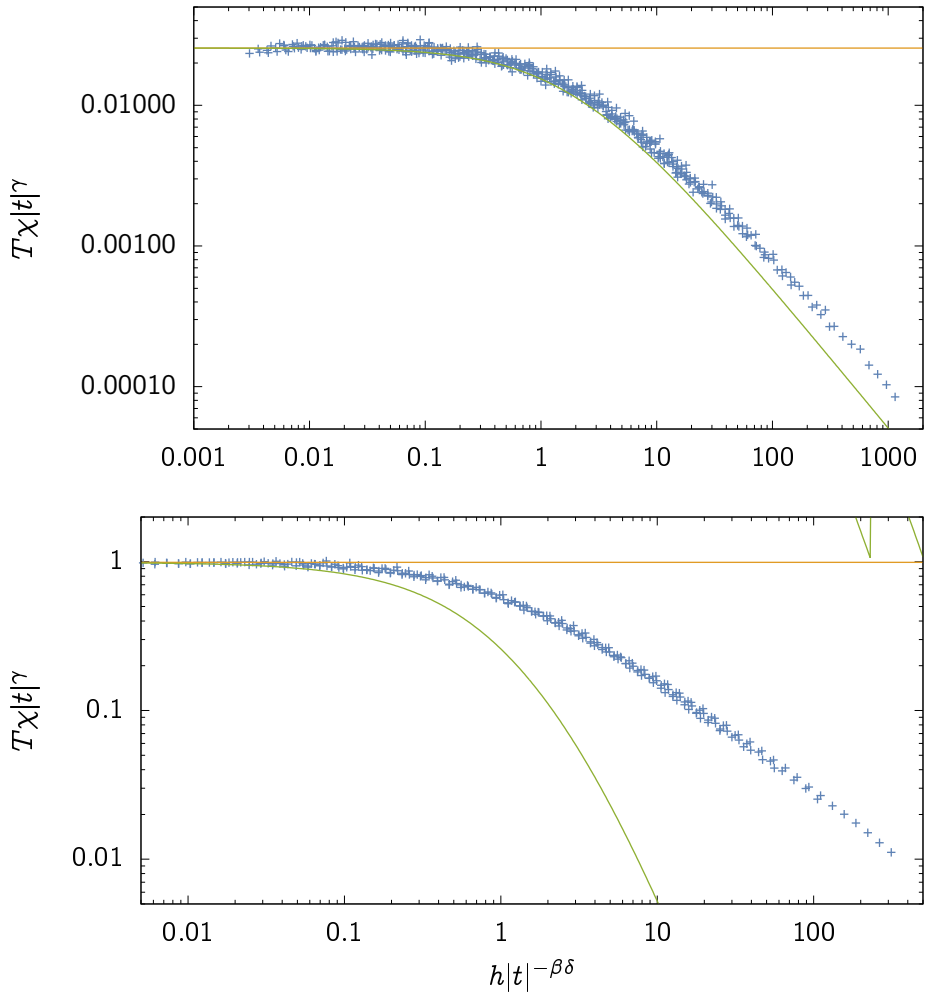

Figure 4.1: Measured susceptibilities along with asymptotic approximations for a (Top) 4096×4096 2D Ising model and (Bottom) $128 \times 128 \times 128$ 3D Ising model. Data shown as blue points, critical amplitude for the susceptibility shown in yellow, and lowest-order truncated scaling function in green.

$(2^{5/2}\operatorname{arcsinh}1)^{\beta}$ for $\beta = 1/8$.[269] This gives

$$B_{2D} = \frac{8V_2\mathcal{M}(0)}{S_d^2\mathcal{S}(0)^2} = (2^{11/16}\pi(\operatorname{arcsinh}1)^{15/8})^{-1} \simeq 0.250448 \tag{4.24}$$

Our value of B_{2D} matches that obtained by perturbative conformal field theory.[239] For the 3D Ising model on a cubic lattice, $\mathcal{S}(0) \simeq 10.083/T_c^{\mu}$ for $\mu = 2\nu \simeq 1.25994$ [270], $T_c \simeq 4.51152480$ [271] and $\mathcal{M}(0) = 1.6919045$ [272]. This gives

$$B_{3D} \simeq 0.0464597 \tag{4.25}$$

Having fixed B, we only have A, which will be set by the critical amplitude of the susceptibility. For the square lattice 2D Ising model $\mathcal{Y}(0) = 0.0255369719$ [273], while for the cubic lattice 3D Ising model $\mathcal{Y}(0) = 0.99051$ [274]. These critical amplitudes are shown as yellow lines in Figure 4.1. These set the constants A in the limit of zero field.

The truncated scaling functions with constants set as above are shown as green lines in Figure 4.1. The 2D function describes the initial crossover from the zero-field behavior well, but loses that description in the tail. The 3D case appears a disaster. These become even more dramatic in the magnetization, since integrating out over a tail which does not match adds up the discrepancies. This lack of matching in the tails is due to a fundamental fault in our scaling forms, which we will at least partially address in the next section.

These considerations are largely to do with the low-order matching of the function, while the high-order agreement should be better. This is seen in the

behavior of the series predicted for scaling function about zero field, compared with direct numerical results for the same series coefficients.[236,275] These data are shown in Figure 4.2. The 2D function performs extremely well against 12 known coefficients, while the 3D function does very poorly for the many fewer known coefficients. In the 2D case, is worth emphasizing that to get this agreement, we have only set two numbers: B, which depends on quantities not directly related to these coefficients, and A, which was used to set the 0th coefficient and therefore matches its exact value. The seemingly poor performance of the 3D scaling functions may be due to a misestimation of the constant B, or of the true asymptotic form of the singularity on the imaginary side.

§32. Fixing the tails

Like mopey Eeyores, we have lost our tails. What happened? It first helps to reflect what the tails should look like in the first place. As one approaches the critical isotherm, or approaches $t = 0$ at finite h, the free energy has a scaling form

$$F(t,h) = h^{(2-\alpha)/\beta\delta}\Phi(th^{-1/\beta\delta}) \tag{4.26}$$

There is no transition or singularity on the critical isotherm, and so Φ is analytic in its argument. It follows that

$$\mathcal{F}(x) = x^{(2-\alpha)/\beta\delta}\Phi(x^{-1/\beta\delta}) = x^{(2-\alpha)/\beta\delta}\Big(\Phi(0) + \mathcal{O}(x^{-1/\beta\delta})\Big) \tag{4.27}$$

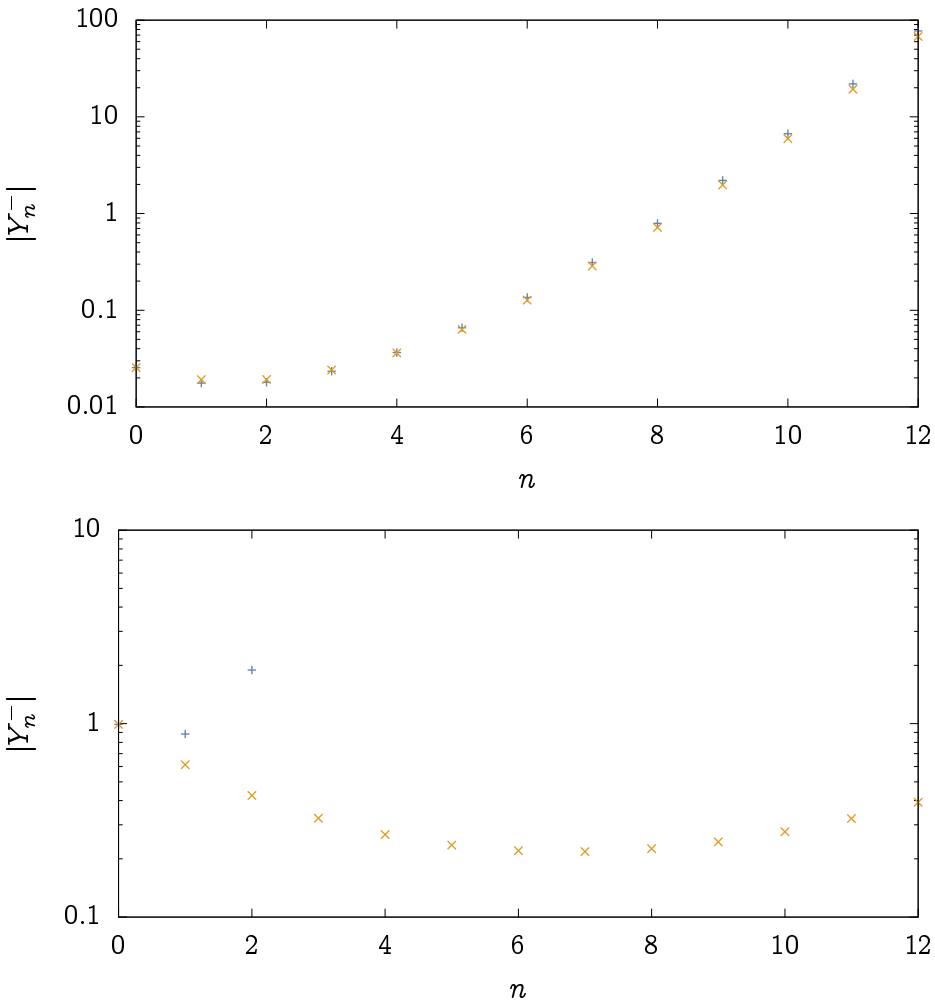

Figure 4.2: Series coefficients for the (top) 2D or (bottom) 3D Ising susceptibility scaling function near the abrupt transition. (Blue) Measured using variational transfer matrix methods for 2D [236] and high temperature expansions for 3D [275]. (Yellow) Predicted by our zeroth-order scaling function.

for large x, and moreover that

$$\begin{aligned}\mathcal{Y}(x) &= -\frac{\partial^2 \mathcal{F}}{\partial x^2} \\ &= -\left[\left(\frac{2-\alpha}{\beta\delta}\right)^2 - \frac{2-\alpha}{\beta\delta}\right]\Phi(0)x^{(2-\alpha)/\beta\delta-2} + \mathcal{O}(x^{(1-\alpha)/\beta\delta-2})\end{aligned} \tag{4.28}$$

For the 2D Ising model, the leading behavior goes as $x^{-14/15}$, while for the 3D Ising model it goes as $x^{-0.791221}$. The amplitude of these corrections is again fixed by known amplitudes: for the 2D Ising model, $\Phi(0) = -0.992798$ [187], while for the 3D Ising model $\Phi(0) = -4.20274$ [225]. Neither of our approximate scaling functions has the correct behavior at large x, as both are analytic in x^{-1} and have in particular

$$\begin{aligned}\mathcal{Y}_{2\text{D}}(x) &= \frac{AB}{\pi x} + O(x^{-2}) \\ \mathcal{Y}_{3\text{D}}(x) &= \frac{2A\Gamma(7/3)}{2\pi B x^3} + O(x^{-4})\end{aligned} \tag{4.29}$$

In the 2D case we simply lucked out: $1 \simeq 14/15$. We can also see one reason why the discrepancy in the 3D case is so severe: $3 \gg 0.79$.

That these limits do not match is not surprising, since the large-field limit depends heavily on the large negative field behavior of the imaginary free energy, which we do not know much about. Suppose that for some argument

x_*, $\operatorname{Im} F(x) \simeq A_\infty(-x)^\zeta$ for all $|x| > |x_*|$. For $x \gg |x_*|$,

$$\begin{aligned}
\mathcal{Y}(x) &= -\frac{2}{\pi}\int_{-\infty}^{0} \frac{\operatorname{Im}\mathcal{F}(x')}{(x'-x)^3}\,dx' \\
&= -\frac{2}{\pi}\int_{x_*}^{0} \frac{\operatorname{Im}\mathcal{F}(x')}{(x'-x)^3}\,dx' - \frac{2}{\pi}\int_{-\infty}^{x_*} \frac{\operatorname{Im}\mathcal{F}(x')}{(x'-x)^3}\,dx' \\
&\simeq -\frac{2}{\pi}\int_{x_*}^{0} \frac{\operatorname{Im}\mathcal{F}(x')}{x^3}\,dx' - \frac{2}{\pi}\int_{-\infty}^{x_*} \frac{A_\infty(-x)^\zeta}{(x'-x)^3}\,dx' \\
&\simeq -\frac{2}{\pi}\frac{1}{x^3}\int_{x_*}^{0} \operatorname{Im}\mathcal{F}(x')\,dx' - \frac{A_\infty}{\pi}\zeta(\zeta-1)\Gamma(2-\zeta)\Gamma(\zeta-1)x^{\zeta-2}
\end{aligned} \tag{4.30}$$

In the 2D case, $\zeta = 1$, and the asymptotic behavior is x^{-1}, as we find when we expand the integrated function about infinity.* In the 3D case, $\zeta = -7/3$, and the asymptotic behavior is *not* $x^{-13/3} = x^{-4.33}$ because the x^{-3} term dominates.

This analysis presents us with an possibility to fix the large x behavior of our functions, by ensuring they match the expected behavior in that limit. We might add a crossover term to our imaginary free energies so that they take the form

$$\operatorname{Im}\mathcal{F}(x) = A\Theta(-x)(-Bx)^{-b}e^{-1/(-Bx)^{d-1}}\left[1+c(-Bx^n)\right]^{(b+(2-\alpha)/\beta\delta)/n} \tag{4.31}$$

This seems like a good choice because it does not modify the singularity of $\mathcal{F}$ near $x = 0$, as a crossover like $1 + cBx^\zeta$ would. The large x behavior is

$$\operatorname{Im}\mathcal{F}(x) \simeq Ac^{b+(2-\alpha)/\beta\delta}(-Bx)^{(2-\alpha)/\beta\delta} \tag{4.32}$$

*Careful readers will notice that the *amplitude* of the term cannot be correct, since $\zeta - 1 = 0$. There is a special amplitude for $\zeta = 1$ equal to $\frac{1}{2}$, which precisely reproduces what we see when we expand the integrated result. This can also be found by taking the limit $\zeta \to 1$.

which matches the asymptotic behavior for large x is as desired. We can now try to set the constant c by requiring that the amplitude of the tails match. Note that setting nonzero c changes the value of $\mathcal{Y}(0)$ as well, so A also need to be rematched.

Unfortunately, we can no longer write the resulting integral as a jumble of officially named functions, but numerically integrating still works fine. Doubly unfortunately, this approach is very poor. For the 2D Ising model, both the zero and infinite argument limits cannot be simultaneously fit by a function of this form, while for the 3D model the addition fits both limits but leaves an unsightly and unphysical camel hump in the middle. The choice appears unnatural, in a sense, and while it might be suitable for a project to *fit* the scaling function to intermediate data, it is not a controlled way to match it to its asymptotic series expansions and expect the intermediate regions to converge.

§33. Patching things together with Schofield coordinates

We've had limited success modifying the imaginary free energy scaling function to get the desired high-field behavior in the real free energy. However, analytic corrections can still be made on the real side, and substantial machinery exists to write down parametrically defined scaling forms that have the correct analytic behavior in each limit by definition.

Let us review this behavior, for the susceptibility, very briefly. Near the

abrupt transition line the susceptibility has a scaling form

$$T\chi = |t|^\gamma \mathcal{Y}^-(h|t|^{-\beta\delta}) = |t|^\gamma \sum_{n=0}^{\infty} Y_n^-(h|t|^{-\beta\delta})^n \tag{4.33}$$

where though the function is nonanalytic and therefore the series nonconvergent it is nonetheless a series in integer powers. Near the critical isotherm, it has

$$T\chi = |h|^{\gamma/\beta\delta} \mathcal{Y}^*(th^{-1/\beta\delta}) = |h|^{\gamma/\beta\delta} \sum_{n=0}^{\infty} Y_n^*(th^{-1/\beta\delta})^n \tag{4.34}$$

which is perfectly analytic, and near $h = 0$ in the high temperature phase it has the form

$$T\chi = |t|^\gamma \mathcal{Y}^+(h|t|^{-\beta\delta}) = |t|^\gamma \sum_{n=0}^{\infty} Y_n^+(h|t|^{-\beta\delta})^n \tag{4.35}$$

which is again analytic, and where $\mathcal{Y}^+$ is an even function, since the magnetization is odd.

The susceptibility and other scaling functions functions can be put into a form that also incorporates known properties of the scaling functions in the rest of the configuration space using a Schofield-like parameterization.[227,229] Those parameters are R, which is sort of like a radius from the critical point, and θ, which is sort of like an angle from the high-temperature zero-field axis. They are defined by

$$t = R(1 - \theta^2) \qquad\qquad h = h_0 R^{\beta\delta} g(\theta) \tag{4.36}$$

where h_0 is a constant and g is an arbitrary odd function whose first finite zero $\theta_c > 1$ corresponds to the abrupt transition. Normally g is chosen such that the coordinates obey a third relationship, $M = m_0 R^\beta \theta$, which give special significance to the lines of constant R. However, this is impossible to do while preserving the analyticity of g if we consider the abrupt transition line as part of our domain, and not infinitesimally outside of it: there is an essential singularity in M as a function of h, and so there must be in θ and R as functions of H at fixed T! We will therefore abandon this last constraint, since the variables prove useful enough on their own.

In these coordinates the invariant combination $h|t|^{-\beta\delta}$ near the abrupt transition line is given by

$$h|t|^{-\beta\delta} = \frac{h_0 g(\theta)}{|1-\theta^2|^{\beta\delta}} = \frac{h_0(-g'(\theta_c))}{(\theta_c^2-1)^{\beta\delta}}(\theta_c - \theta) + O\big((\theta_c-\theta)^2\big), \tag{4.37}$$

which is an analytic function of θ about θ_c, asymptotically linear in θ. The invariant $th^{-1/\beta\delta}$ near the critical isotherm is

$$th^{-1/\beta\delta} = \frac{(1-\theta^2)}{[h_0 g(\theta)]^{1/\beta\delta}} = \frac{2}{(h_0 g(1))^{1/\beta\delta}(1-\theta)} + O\big((1-\theta)^2\big), \tag{4.38}$$

also an analytic function of θ, asymptotically linear in θ. Near $h = 0$ in the high-temperature phase we get a similar result to (4.37).

Because of these nice properties, dependence we would normally write as polynomial expansions of the scaling invariants that have non-polynomial expansions in different limits can instead be written as polynomial functions

of θ and have all scaling dependence naturally captured. If we define

$$T\chi = R^{\gamma}\mathcal{Z}(\theta) \tag{4.39}$$

it has all the right limits built in! On the low or high temperature zero-field line, we have

$$|t|^{-\gamma}\mathcal{Y}^{\pm}(h|t|^{-\beta\delta}) = \mathcal{R}^{-\gamma}|1-\theta^2|^{-\gamma}\mathcal{Y}^{\pm}(h|t|^{-\beta\delta}) = R^{-\gamma}\mathcal{Z}(\theta) \tag{4.40}$$

while near the critical isotherm we have

$$h^{-\gamma/\beta\delta}\mathcal{Y}^{*}(th^{-1/\beta\delta}) = R^{-\gamma}(h_0 g(\theta))^{-\gamma/\beta\delta}\mathcal{Y}^{*}(th^{-1/\beta\delta}) = R^{-\gamma}\mathcal{Z}(\theta) \tag{4.41}$$

This sets the correspondence between familiar functions of our scaling variables and functions of the new Schofield coordinate. Note that the Schofield coordinates are not analytically invertible, so to make this correspondence formal takes quite it a bit of work.

As a first order attempt, let's take our zeroth order functions and construct the simplest scaling form consistent with the value of $T\chi$ along the critical isotherm and at both zero field limits, making sure it contains our singularity. This function is

$$\mathcal{Z}(\theta) = Z(\theta) + \mathcal{Y}\left(\frac{h_0(-g'(\theta_c))}{(\theta_c^2-1)^{\beta\delta}}(\theta_c-\theta)\right) + \mathcal{Y}\left(\frac{h_0(-g'(\theta_c))}{(\theta_c^2-1)^{\beta\delta}}(\theta_c+\theta)\right) \tag{4.42}$$

where $Z(\theta)$ is an analytic, even function of θ. For now, we will take it to lowest order, with $Z(\theta) \simeq Z_0$. As θ goes to $\pm\theta_c$, this function has an essential

singularity of exactly our form. It is analytic in the range $-\theta_c < \theta < \theta_c$.

The function g needs to be odd, have its first zero at θ_c, and near-identity. This is satisfied generally by

$$g(\theta) = \left(1 - \frac{\theta^2}{\theta_c^2}\right) \sum_{n=0}^{\infty} g_n \theta^{2n+1} \tag{4.43}$$

but we will also take this to lowest order, or

$$g(\theta) \simeq \theta\left(1 - \frac{\theta^2}{\theta_c^2}\right) \tag{4.44}$$

We are looking to match our function in three places, but have *four* unknowns: A (embedded in the definition of $\mathcal{Y}$), Z_0, h_0, and θ_c. We will therefore also pick one derivative to match, and make sure that the first derivative of $\mathcal{Y}^-$ is also matched by our scaling function.

Fixing χ at three values gives the consistency equations

$$\begin{aligned} \mathcal{Z}(0) &= Y_0^+ \\ \mathcal{Z}(1) &= (h_0 g(1))^{-\gamma/\beta\delta} Y_0^* \\ \mathcal{Z}(\theta_c) &= |1 - \theta_c^2|^{-\gamma} Y_0^- \end{aligned} \tag{4.45}$$

where for the 2D Ising model, $Y_0^- = 0.0255369719$ [273], $Y_0^* = -\Big[(2-\alpha)^2/(\beta\delta)^2 - (2-\alpha)\beta\delta)\Big]\Phi(0) = 0.070599$ [187], and $Y_0^+ = 0.962582$ while for the 3D Ising model, $Y_0^- = 0.99051$ [274], $Y_0^* = 1.07716$ [225], and $Y_0^+ = 5.060$ [275].

Matching a derivative is more difficult because the coordinate transformations are not analytically invertible. First, let us try to match the first

	2D	3D
A	15.9885	1.86074
h_0	0.599909	1.1843
Z_0	0.93902	2.24137
θ_0	1.04621	1.22238

Table 4.1: Parameters used to fit the parametric version of the asymptotic scaling function to its known values.

derivative by the abrupt transition. If we call the invariant $x = h|t|^{-\beta\delta}$, then

$$\begin{aligned}\frac{\mathcal{Z}'(\theta_c)}{|1-\theta_c^2|^{-\gamma}} - \frac{2\gamma\theta_c\mathcal{Z}(\theta_c)}{|1-\theta_c^2|^{-\gamma-1}} &= \frac{\partial}{\partial\theta}\left(\frac{\mathcal{Z}(\theta)}{|1-\theta^2|^{-\gamma}}\right)_{\theta=\theta_c} \\ &= \left.\frac{\partial\mathcal{Y}^-}{\partial\theta}\right|_{\theta=\theta_c} = \left.\frac{\partial\mathcal{Y}^-}{\partial x}\frac{\partial x}{\partial\theta}\right|_{\theta=\theta_c} = -Y_1^- \frac{h_0(-g'(\theta_c))}{(\theta_c^2-1)^{\beta\delta}}\end{aligned} \tag{4.46}$$

which connects the derivative of $\mathcal{Z}$ to the coefficient Y_1^-. For the 2D Ising model, $Y_1^- = -0.0176358$ [236] and for the 3D Ising model $Y_1^- = -0.883$ [275]. Other expressions like this exist for the other expansions, and they get quite tedious for higher derivatives. The process can be made automatic with computer algebra software, but for our purposes we will not need it: we only need four constraints to match the four variables of our second-attempt functions.

The values of the free parameters that result from making the matching described in this section can be found in Table 4.1. There is little reason to believe that these parameters would be stable to these values upon the further expansion of the approximate scaling function to higher orders in $\mathcal{Z}$ and g. The resulting forms for the susceptibility are plotted in Figure 4.3 in brown, alongside the asymptotic forms that the scaling functions were matched to (yellow and red) and the lowest-order truncated scaling form (green).

The use of these parametric coordinates has done wonders for the ability

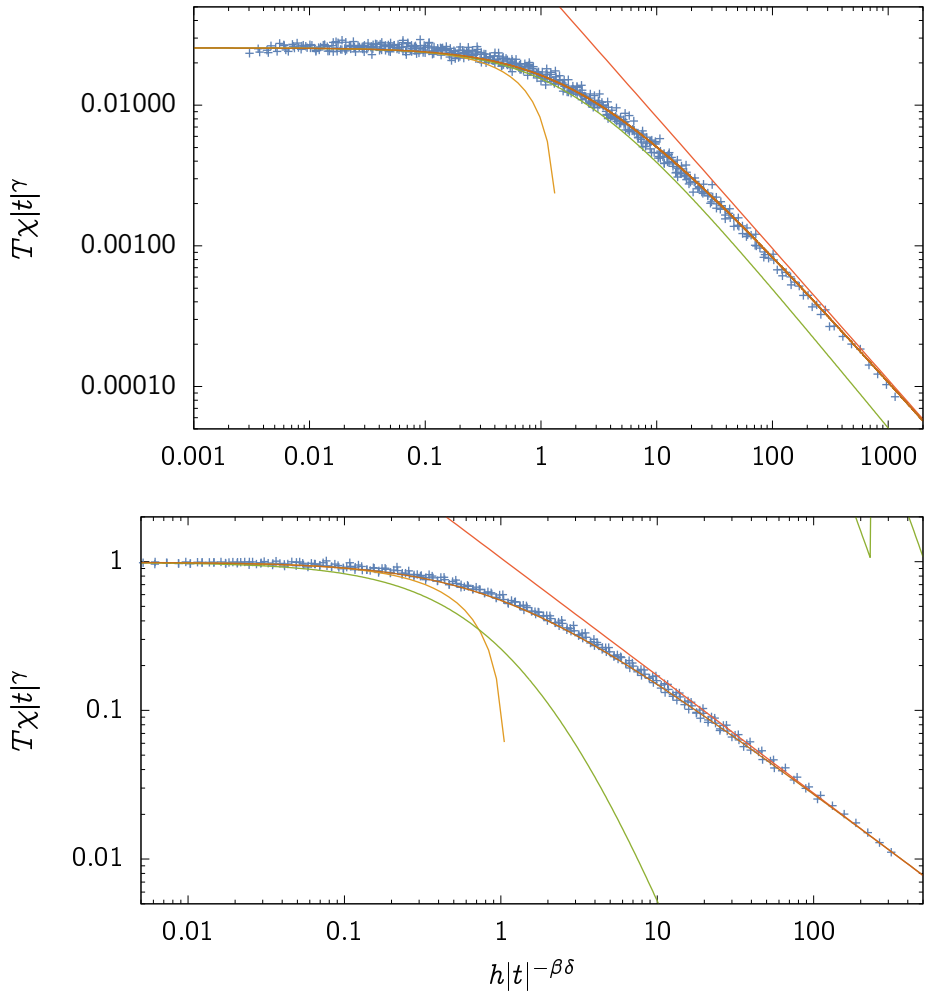

Figure 4.3: Measured susceptibilities along with asymptotic approximations for a (Top) 4096×4096 2D Ising model and (Bottom) $128 \times 128 \times 128$ 3D Ising model. Data shown as blue points, first two small-argument terms in the susceptibility shown in yellow, and lowest-order truncated scaling function in green, the first large-argument term in the susceptibility shown in red, and the next-order scaling function in brown.

of the scaling functions to describe the tails, having only matched the value of the scaling function in three places and its derivative in one! Even the 3D Ising model, whose lowest-order truncated scaling function did a very poor job of describing the actual scaling function, has been brought much closer to its true form. Iterative fitting of this type to higher orders in g should have good convergence, though this is work that still needs to be done.

Having modified our singular scaling functions with analytic corrections on the real side, we should reassess how the high-order coefficients behave. These are pictured, alongside empirical coefficients and those of our lowest-order truncation, in Figure 4.4. The extraordinary initial closeness between the 2D coefficients and those of the truncated function mean that, by adding a correction that fixes another of those coefficients, we have make the agreement slightly worse. For the 3D case, the growth of the coefficient values picks up dramatically, though still falls short of predicting well the third coefficient. It's difficult to draw conclusions from the discrepancies in the 3D case here, though, because so few of its low-temperature scaling function coefficients are known.

Finally, having produced a basically competent approximation of the susceptibility that includes its essential singularity, we are posed to finally address the problem of the real *free energy*, which we started discussing exclusively but lost ground to the susceptibility at some point along the way. This can be achieved by taking the forms above and integrating with respect to h. Once again, the inability to analytically invert the Schofield coordinates raises its ugly head, and makes it very difficult to perform such an integral analytically. However, once the form is set, a numeric integral can be carried out,

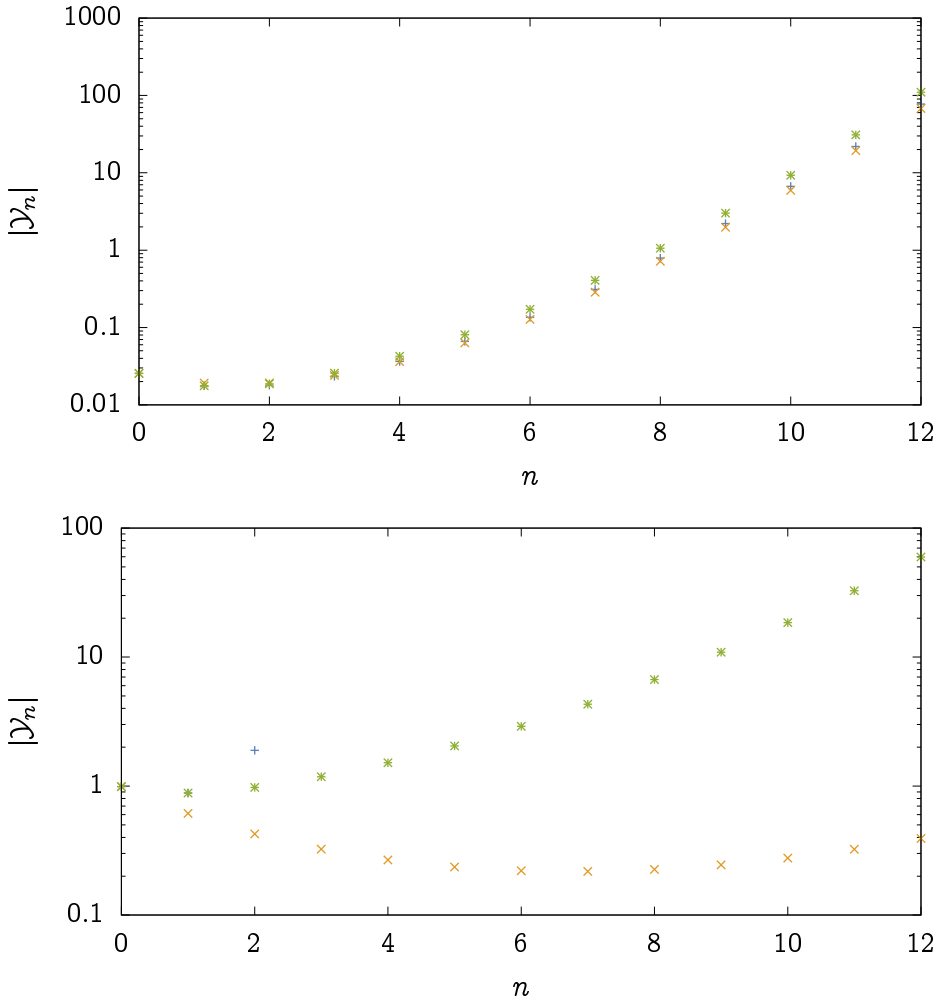

Figure 4.4: Series coefficients for the (top) 2D or (bottom) 3D Ising susceptibility scaling function near the abrupt transition. (Blue) Measured using variational transfer matrix methods for 2D [236] and high temperature expansions for 3D [275]. (Yellow) Predicted by our zeroth-order scaling function from §31. (Green) Predicted by our parametric scaling function from §33.

with the result for the magnetization seen in Figure 4.5. There is reasonable agreement, though the deviation in the tails shows the accumulated effects of integrating small deviations though the course of the susceptibility.

§34. Discussion, conclusions, & next steps

We have seen that, with very little information about its precise form, asymptotic scaling forms derived from analytic continuation from the metastable state can well-describe the behavior of the Ising critical scaling functions. In two dimensions, the description appears natural from the outset, while in three the description requires corrections to even qualitatively match the behavior. As mentioned above, part of the gross discrepancy in the three dimensional case is a result of bad luck in the truncated form's tail, which isn't a reflection of the singular behavior at zero field, but the equally gross discrepancy with the (admittedly few) higher-order coefficients in the function should lead to pause. Why is the 2D version so much more successful than the 3D version? One thought is related to the fact that the droplet entropy has special scaling: you might have noticed that the exponent b multiplying the exponential had an exceptional form for three dimensions. This is a result of unique logarithmic divergences in the renormalization group analysis for the long-wavelength fluctuations on the surfaces of the 3D droplets; perhaps 3D is special for more than just this reason, or perhaps this exceptional form is incorrect. Following this analysis with the 4D model should shed some light on this question, since markedly better performance in 4D would suggest 3D is uniquely bad, while similar performance would suggest 2D is uniquely good. We would also like to

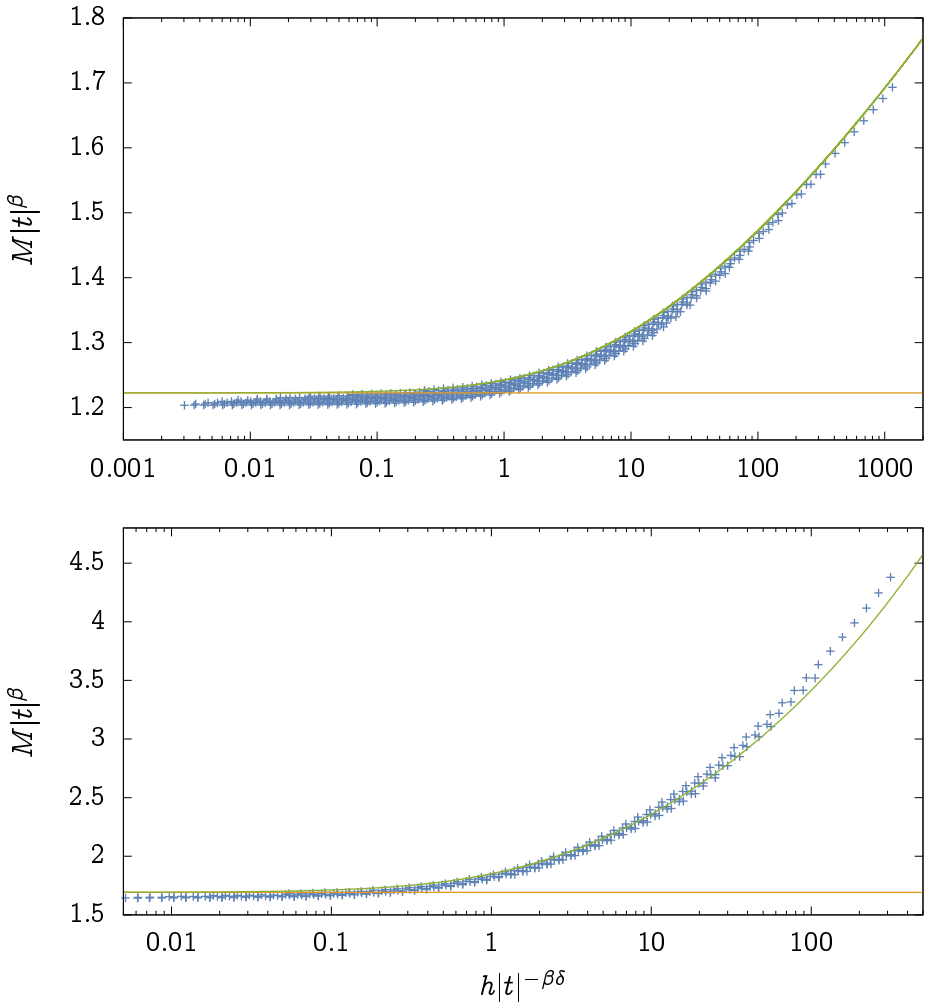

Figure 4.5: Measured magnetizations along with numerically integrated asymptotic approximations for the susceptibility from (Top) a 4096×4096 2DIsing model and (Bottom) a $128 \times 128 \times 128$ 3D Ising model. Data shown as blue points, critical amplitude for the magnetization shown in yellow, and integrated approximation in green.

attempt to directly measure the exponent of this prefactor, perhaps by simulating metastable decay directly. Some work using the renormalization group suggests that the essential singularity predicted by a naïve droplet theory is not the only singularity at the transition line, and our predictions may be doing so badly in 3D without further matching because we have missed an aspect of the singularity.[238,276,277] Improving existing direct numerical studies of the 3D Ising decay rate in the vicinity of the critical point may help resolve some of these questions.[278]

In the case of the 2D model, where the method converges exceptionally well, our results should allow improved high-precision functional forms for the free energy,[229] and should have implications for the scaling of correlation functions.[279,280] A more careful self-consistent analysis would start with the singular form for the *free energy*—not the susceptibility—and put it in the parametric Schofield form to undergo the same limit-matching we saw previewed here. Developing a generalization of Schofield coordinates that naturally captures the metastable region may dramatically improve things, taking considerations of global renormalization flow into account and not just those local to the critical point.[281–283] One might be able to work with such coordinates directly in the imaginary free energy, removing the need to incorporate them in the real free energy in an *ad hoc* way, as we did here. Our methods might be generalized to predict similar singularities in systems where nucleation and metastability are proximate to continuous phase transitions, such as 2D superfluid transitions,[284,285] the melting of 2D crystals,[286] and freezing transitions in glasses, spin glasses, and other disordered systems.

CHAPTER V

NORMAL FORMS FOR INFINITE-ORDER PHASE TRANSITIONS

The renormalization group made firm how the universality enters statistical physics. Different physical systems with different microscopic laws and different general phenomenology exhibit the same critical behavior if their critical points lie in the basin of attraction of the same fixed point under RG. All systems with critical points in the same basin of attraction are said to belong to the same universality class.

This provides a heuristic path to a taxonomy of critical phenomena: label each critical point by the fixed point it flows into. This leaves an important step in any classification project implicit: how does one label a fixed point? The question is more serious than choosing a convention. If I set up one RG procedure for a model and another RG procedure for another model, and after painstaking work can tell you facts about the fixed point of each, what exactly is needed to establish that the fixed points of both are the same? How many exponents or critical amplitude ratios need to match? What about

*Portions of the material of this chapter were published in A Raju, CB Clement, LX Hayden, J Kent-Dobias, DB Liarte, DZ Rocklin, and JP Sethna, "Normal form for renormalization groups", Physical Review X **9**, 10.1103/physrevx.9.021014 (2019). Some of the work was done in collaboration with the listed authors.

fixed points whose critical behavior doesn't have a power law form? Do the corrections to scaling need to match? Which ones?

The failure of simply using exponents as the tool for classification is especially clear in the case of *infinite order* transitions, so-called because every derivative of the free energy is continuous at the transition. In these situations, a wide array of critical phenomena usually described by slapping an exponent on it now has a different form than a power law. Researchers often fall back on the familiar anyway, citing the values of parameters in stretched exponentials as "anomalous exponents" and using the appearance of even a single stretched exponential that matches one in a known theory to associate the two.

Luckily, we need not guess in the dark. A classification developed by my colleagues provides a natural and rigorous way to differentiate fixed points of arbitrary strangeness, and even tell which corrections to scaling are inherent traits of a given fixed point and which are not.[287] In this chapter I will briefly review that classification, then reflect on its treatment of two canonical models with infinite-order transitions: the 2D XY model and the hexatic–solid transition. We will then look at progress towards classifying a model with a phenomenological infinite order transition that lacks a corresponding renormalization group theory. Finally, we will review a few other oddities in the collection of transitions of this type.

§35. Normal form theory & the classification of fixed points

The renormalization group is a continuous mapping from the space of possible models onto itself. A physical model with parameters $x_i(0)$ is mapped by a flow over ℓ to a model with parameters $x_i(\ell)$ whose properties are identical to properties of the original model coarse-grained by a factor e^ℓ. Sinks, or stable fixed points, are places in model space that are mapped onto themselves, and all models that flow into those sinks are described by the same phase, since by definition their long-range behavior is identical. Phase transitions occur when a model, having smoothly varied its physical parameters, moves from the basin of attraction of one sink to that of another. These basins are separated by critical surfaces that flow on the unstable border between the basins to unstable fixed points, which act as separatrices between basins.[222,223,288]

The typical picture connecting critical phenomena in physical systems to the behavior of renormalization group flow near unstable fixed points is to write down a set of differential equations governing the evolution of model space under renormalization. For our parameters x_i, they would take the form

$$\frac{dx_i}{d\ell} = f_i(x_1, \ldots, x_N), \tag{5.1}$$

where f_i is an arbitrary function, but must be analytic because the coarse-graining operation that gives rise to the renormalization group flow is itself

analytic.[189] At a fixed point $x_i = x_i^*$, we expand these functions,

$$\frac{dx_i}{d\ell} = f_i(x_1^*, \ldots, x_N^*) + \sum_j \frac{\partial f_i}{\partial x_j}(x_j - x_j^*) + \cdots \tag{5.2}$$

Since a fixed point must have no flow, $f_i(x^*) = 0$ for all i and the principal behavior of the flow is the linear term in the expansion. Usually, the flow is then taken along eigendirections y_i of the Jacobian matrix $\partial f_i/\partial x_j$,

$$\frac{dy_i}{d\ell} = \frac{\partial f_i}{\partial y_i} y_i + \cdots = \lambda_i y_i + \cdots \tag{5.3}$$

which introduces the scaling dimensions λ_i which give rise to the famous power law exponents in the solutions to these flow equations. Also, implicitly, one transforms from a picture with the original model variables x_i (or linear combinations thereof) and uses *scaling fields* u_i defined by the property that $u_i = y_i + h_i(y_1, \ldots, y_n)$ is near identity, and that

$$\frac{du_i}{d\ell} = \lambda_i u_i \tag{5.4}$$

with no trailing dots. So simplified, the flow equations can be solved analytically to find, e.g., that combinations of the form $u_i u_j^{-\lambda_i/\lambda_j}$ are invariant under the flow and can be used for collapses and the like. Analytic corrections arise from the nonlinear dependence of the scaling fields on the original parameters.[289] A well-behaved fixed point whose flow equations have this simple form is known as hyperbolic.

If we wish to use the great utility that comes with these simplifications of the renormalization group picture of a fixed point, we need to take care that

we do note introduce anything in our analysis that jeopardises the connection between our physical parameters and the behavior of our convenient scaling fields. If one were, for instance, to write $u_i = y_i^\alpha$, we would change our precious exponent! In order for our model to have the same singular dependence on the scaling fields that it does on our critical parameters, they must be connected by an *analytic* change of coordinates. The functions h_i that make these corrections must therefore be analytic.

Usually this can be done in a generic way. Consider a higher-order term in the flow equations for the variable y_i of the form $cy_1^{\alpha_1} \cdots y_n^{\alpha_n}$ for some set of integers $\alpha_1, \ldots, \alpha_n$. It can be shown, as it is in [287], that this term can be removed, leaving only changes to the flow equations at polynomial order higher than $\sum_j \alpha_j$, by introducing a term of the form $a\theta_1^{\alpha_1} \cdots \theta_n^{\alpha_n}$ in the function h_i with

$$a = \frac{c}{\lambda_i - \sum_j \lambda_j \alpha_j} \tag{5.5}$$

The procedure for forming the scaling fields from flow equations follows. At each polynomial order, solve the equation (5.5), using the results to define the polynomial transformations h to the same order. These partially defined functions define partially linearized scaling fields with no nonlinear terms up to and at that polynomial order. The process is then repeated on these new variables at the next highest order.

Looking at (5.5), one can immediately see that issues arise in this procedure. If $\lambda_i = \sum_j \lambda_j a_j$ for a nonlinear term in the flow equations for some variable, the solution is singular and the term cannot be removed. A reso-

nance like this is responsible for the anomalous logarithm in the scaling for the 2D Ising model. These often appear when an exponent is tuned through zero, which happens in the upper and lower critical dimensions of many models.

If we cannot linearize every flow equation to then compare exponents, we will have to do the next best thing. Having agreed on an order in which to try to remove nonlinear terms from the flow equations, going order by order and trying to do so (perhaps with more general equations than (5.5)) will always leave the same terms behind for the same model, even if it is analytically reparameterized. The result, known as its *simplest normal form*, is the identifying fingerprint of the flow equation, and the coefficients of the remaining terms, linear and otherwise, are all universal. Constants of the flow are constructed from integrating these equations, and can differ from the familiar form of power law products if unremovable nonlinear terms are present. These invariant combinations, while more complicated in form, are actually analytically related to the physical parameters.

At an infinite order fixed point there are no relevant variables, but there are marginal ones whose nonlinear terms cause them to flow away from the fixed point. This means there is no Jacobian to diagonalize, no exponents, and the formula (5.5) is useless for removing higher terms. The continuity of the free energy arises from the continuity of the scaling invariants that the free energy is a function of. If u_i and u_j are two marginal scaling fields for a fixed point with flows

$$\begin{aligned} \frac{du_i}{d\ell} &= c_i u_i^{\alpha_{ii}} u_j^{\alpha_{ij}} \\ \frac{du_j}{d\ell} &= c_j u_j^{\alpha_{ji}} u_j^{\alpha_{jj}} \end{aligned} \tag{5.6}$$

for constants c and integers α, then any function of

$$\frac{c_j u_i^{1-\alpha_{ii}+\alpha_{ji}}}{1-\alpha_{ii}+\alpha_{ji}} - \frac{c_i u_j^{1+\alpha_{ij}-\alpha_{jj}}}{1+\alpha_{ij}-\alpha_{jj}} \tag{5.7}$$

is an invariant of the flow. The universal scaling functions for quantities like the free energy are functions of this invariant, which for integer as such that the denominators are nonzero is infinitely continuous and in fact *analytic* in both variables.

§36. The XY model & the BKT transition

The 2D XY model is a remarkable system for several reasons. It was the site of recently celebrated insight into the connection between ground-state topology and phase transitions.[290] Thermodynamic quantities have essential singularities at its phase transition, not ordinary power laws, and their derivatives remain continuous to arbitrary order, making its phase transition infinite order.[291–294] This is related the fact that its RG flow equations are inherently nonlinear: they have no relevant and two marginal state variables and the procedure laid out by (5.5) for removing higher order terms from the flow equations contributes nothing to their simplification.

The XY model is usually posed as ferromagnetically interacting planar spins. Its partition function is exactly equivalent to the product of a trivial Gaussian model—corresponding to spin wave degrees of freedom—with a neutral Coulomb gas—corresponding to the interaction of spin vortices,[295–297] or to roughening models of 2D interfaces.[298] The latter component contains the interesting critical behavior, which is characterized by these vortices going

through an unbinding transition. The flow equations for a Coulomb gas in dimension d can be computed by renormalizing the vortex–vortex interaction screened by bound vortices, and are given by

$$dK/dl = -K(\tfrac{1}{4}Ky^2 + d - 2) + \cdots \tag{5.8}$$

$$dy/dl = -y(K - d) + \cdots \tag{5.9}$$

where $K \sim T^{-1}$ and y is the fugacity* of the vortices,[299] which for an XY model is a function of temperature and cannot be tuned independently but is a free parameter in other equivalent models, e.g., the Coulomb gas itself. For $d > 2$ there is no phase transition in this system, and for $d < 2$ a nontrivial unstable fixed point appears and there is a phase transition in the hyperbolic universality family. It is worth noting that these flow equations do not describe the XY model for any dimension besides $d = 2$; 2 is the *upper* critical dimension of the Coulomb gas and these flow equations, while it is the *lower* critical dimension for the XY model. At $d = 2$ the flow equations undergo a novel bifurcation: there appears a line of stable fixed points at $y = 0$ for all $K > 2$, terminating at $K = 2$. This termination is the Berezinskiĭ–Kosterlitz–Thouless (BKT) critical point.

The flow equation near this point with $x = K - 2$ is, written to linear order,

$$\begin{aligned} dx/dl &= -y^2 + \cdots \\ dy/dl &= -xy + \cdots \end{aligned} \tag{5.10}$$

*Recall that the *fugacity* of a system in the grand canonical ensemble is proportional to the Boltzmann weight per particle, i.e., $y \propto e^{\mu}$ for chemical potential μ.

Note that, at this order, there is nothing resembling an exponent here, and that isn't an accident: because of the nonlinear nature of the flow equations, any physical parameters that result in flow equations with the same simplest normal form as (5.10) but with constant coefficients on the quadratic terms can be brought into the form above, with no quadratic coefficients, by constant rescaling of the parameters. These flow equations are zero to linear order and have zero Jacobian at the fixed point.

In principle arbitrary higher-order terms in these equations exist, but there are several constraints on their form. There is a symmetry $y \to -y$ in the partition function arising from a neutrality condition. In the context of vortices, this is related to the fact that a lone vortex has an energy that falls off logarithmically, and so therefore is infinitely costly in the thermodynamic limit. All vortices must come in pairs, or at least the vorticity of all extant vortices must cancel out, and each microstate in the partition function gets one factor of the fugacity y per unit *vorticity squared*, not per vortex. The fugacity therefore enters the partition function in factors of $y^{-\sum_r n_r^2}$ for $\sum_r n_r = 0$. This implies that dx/dl be even in y and dy/dl be odd. In addition, when the fugacity is zero the model is trivial and x cannot flow, meaning that for $y = 0$ we have $dx/dl = 0$ identically, and we must only have terms proportional to y. Note that, with these constraints and a single tuning of dimension, we uniquely arrive at the truncation (5.10) to quadratic order.

Having applied these constraints, the simplest normal form has been proven

by induction in polynomial order (Appendix A of [300]) to take the form

$$\begin{aligned} du_x/dl &= -u_y^2 - b_0 u_x u_y^2 - b_1 u_x^3 u_y^2 + \cdots \\ &= -u_y^2(1 + u_x f(u_x^2)) \\ du_y/dl &= -u_x u_y \end{aligned} \tag{5.11}$$

For the BKT point in the sine–Gordon model, which is thought to display to the same universality as the XY model, it is known that $b_0 = 3/2$.[300–302] An infinite number of coefficients remain, represented here in the form of the Taylor coefficients of an analytic function f. These numbers are universal in the sense that there is no redefinition of u_x and u_y such that the flow equations take on the form above and contain different coefficient values. This bifurcation does not have a named classification as far as we know.

A true constant of the RG flow can be found by integrating these forms. First, dividing the equations the two equations in (5.11) we find

$$\frac{du_y}{dl} \Big/ \frac{du_x}{dl} = \frac{u_x}{u_y(1 + u_x f(u_x^2))}, \tag{5.12}$$

which separates into

$$u_y \frac{du_y}{dl} = \frac{u_x}{1 + u_x f(u_x^2)} \frac{du_x}{dl}. \tag{5.13}$$

Integrating both sides and choosing l_0 such that $u_x(l_0) = 0$, we find

$$\frac{1}{2}(u_y(l)^2 - u_y(l_0)^2) = \int_{u_y(l_0)}^{u_y(l)} y\, du_y = \int_{l_0}^{l} u_y \frac{du_y}{dl} dl \tag{5.14}$$

$$= \int_{l_0}^{l} \frac{u_x}{1 + u_x f(u_x^2)} \frac{du_x}{dl} dl = \int_0^{u_x(l)} \frac{u_x}{1 + u_x f(u_x^2)} dx. \tag{5.15}$$

It follows that

$$Q(u_x, u_y) = u(l_0)^2 = u_y(l)^2 - 2\int_0^{u_x(l)} \frac{u_x}{1 + u_x f(u_x^2)} du_x \tag{5.16}$$

$$= u_y(l)^2 - u_x(l)^2 + \frac{2}{3} b_0 u_x(l)^3 - \frac{1}{2} b_0 u_x(l)^4 \tag{5.17}$$

$$+ \frac{2}{5}(b_0^3 + b_1) u_x(l)^5 + O(u_x(l)^6) \tag{5.18}$$

is a constant of the flow. The expansion of the integral can be taken to arbitrary order with ordinary computer algebra software. Its truncation to second order yields the usually cited form of the BKT invariant. The finite-size behavior of the flow is rather complicated and doesn't yield closed form results, details can be found in [300]. The line that separates the unbound and bound phases near the BKT fixed point is parameterized by $Q = 0$.

The correlation length in the approach to the BKT transition can be found by requiring that $L^{-1}\xi$ be an invariant of the flow. This yields a differential equation, which can be integrated to yield

$$\xi(u_x, u_y) = \mathcal{J}(Q) \exp\left\{ \int_1^{u_x} \left[(1 + f(v^2)v)\left(Q + 2\int_1^s \frac{s}{1 + s f(s^2)} ds \right) \right]^{-1} dv \right\} \tag{5.19}$$

for an undetermined function $\mathcal{J}$. If the higher-order terms are truncated, or $f = 0$, this can be integrated to yield

$$\xi(u_x, u_y) \simeq \mathcal{J}(Q) \exp\left\{ Q^{-1/2} \tan^{-1} \frac{u_x}{\sqrt{Q}} \right\} \tag{5.20}$$

which in the limit of small Q has the form $\xi(u_x, u_y) \simeq \mathcal{J}(0) e^{\pi/2Q^{1/2}}$. Since Q is an analytic function of u_x and u_y, which are in turn analytic functions of the physical control variables, e.g., the temperature, this establishes the stretched exponential form of the correlation length divergence as the critical line is approached from high temperature.[292] This gives the "anomalous exponent" of $\sigma = \frac{1}{2}$. Systematic corrections to scaling exist due to both the higher order terms and finite-size effects—the first correction at infinite size takes the form

$$\xi_\infty = \xi_0 e^{\pi/\sqrt{Q}} \left[1 + I_1 \sqrt{Q} + O(Q) \right] \tag{5.21}$$

where $I_1 = -\pi b_0^2/12$.[300,303]

Conformal field theory predicts the presence of infinitely many models with this anomalous exponent.[304] The value of σ been shown to be fixed by the quadratic-order truncation of the system's flow equation, independent of any higher-order terms.[305] There are six possible quadratic-order terms in flow equations with two variables. Of these, two can be removed by linear transformations of the two variables. Two more can be set to 1 by rescaling the variables. Hence, there are two parameters at quadratic order which determine the universality family that the system belongs to, and infinite number

of subsequent terms which determine the universality class. Giving a full classification of possibilities is beyond the scope of this paper but we give some examples below.

§37. The hexatic fixed point

When the requirement of symmetry under $y \to -y$ is lifted, the flow equations can no longer be brought to the form (5.11). This occurs naturally in the case of the *hexatic–solid* transition, which occurs during the melting of certain two-dimensional crystals. Starting from a crystalline state, there are two relevant topological defects of the lattice: dislocations and disclinations. Dislocations can move more easily, and like the vortices in the XY model have a critical temperature at which they unbind and move freely within the crystal, destroying long-range translational order but preserving the orientational order. At a higher temperature, the disclinations go through a similar unbinding transition, melting the orientational order and reverting to a liquid. Both dislocations and disclinations are topological defects within the lattice and have an interaction energy that falls off with a logarithm in two dimensions.

Though this phenomenology is very similar in both circumstances to the XY model, only one turns out to share the XY universality class. Disclinations occur at points and have no relative orientation, and only have a scalar topological charge. They therefore obey the same neutrality rules that XY vortices do, and the hexatic–liquid transition is in the BKT universality class. Dislocations, on the other hand, have a charge given by their Burger's vector, which is a vector quantity. As a result, an *odd* number of topological

charges can exist in the system simultaneously and still have net zero charge: just point three length-one arrows away from a central point. As a result, the symmetry restriction of the BKT critical point we discussed in the last section is lifted, and these transitions belong to a different class.

The effect on the flow equations at quadratic order are the addition of two more terms, giving

$$\begin{aligned} \frac{dx}{d\ell} &= a_1 y^2 + a_2 xy \cdots \\ \frac{dy}{d\ell} &= b_1 xy + b_2 y^2 + \cdots \end{aligned} \tag{5.22}$$

However, this is not the simplest quadratic form. We can make a coordinate change, redefining x and y at *linear* order in a way that preserves the fact that the flow in x is identically zero at zero y. If $x' = c_1 x + c_2 y$, we have (to quadratic order)

$$\begin{aligned} \frac{dx'}{d\ell} &= (b_1 c_2 - a_2) x' y + \frac{b_1 c_2^2 + b_2 c_2 - a_1 a_2 c_2}{c_1} y^2 + \cdots \\ \frac{dy}{d\ell} &= b_1 c_2 x' y + (b_1 c_2 + b_2) y^2 + \cdots \end{aligned} \tag{5.23}$$

If we choose $c_2 = a_2/b_1$ and then rescale x and y so that the terms that have negative coefficients in the BKT flow have the same ones here, we have

$$\begin{aligned} \frac{d\tilde{x}}{d\ell} &= -\tilde{y}^2 + \cdots \\ \frac{d\tilde{y}}{d\ell} &= -\tilde{x}\tilde{y} + 2A\tilde{y}^2 + \cdots \end{aligned} \tag{5.24}$$

Truncated to this order, these equations lead to a critical correlation length

identical to that of the BKT transition, but with

$$\sigma = 1 - \frac{1}{2}\left[1 + A^2 + A(1+A^2)^{1/2}\right]^{-1} \tag{5.25}$$

an anomalous exponent that varies continuously with A.[306] There are physical reasons to believe that $A^2 \simeq 0.072939$ for the actual hexatic transition—with equal magnitudes of the angular and radial stiffnesses—but for more general physical scenarios A, and therefore the anomalous exponent, can take other values.

No proof exists right now for what simplest normal form the hexatic flow equations should take. We have implemented a method for matching the terms in differential equations under analytic transformations term by term in order to see how the hexatic flow equations differ from those of the BKT, at least at lower orders.[307,308] The term y^3 in the flow equations for u_x cannot be removed, and it seems like a new infinite set of terms is likewise created. Up to seventh order, the flow equations with quadratic form (5.24) can be brought to the form

$$\begin{aligned} \frac{du_x}{d\ell} &= -u_y^2 + Bu_y^3 + c_0 u_x u_y^2 + c_1 u_x^2 u_y^2 + c_2 u_x^3 u_y^2 + c_3 u_x^4 u_y^2 + \cdots \\ &= -u_y^2(1 + u_x f(u_x)) + Bu_y^3 \\ \frac{du_y}{d\ell} &= -u_x u_y - 2Au_y^2 + \cdots \end{aligned} \tag{5.26}$$

At seventh order and above, the order-by-order matching becomes strenuous enough to take a desktop computer too long to finish within a day. A rigorous mathematical attempt should be made to formally bring these equations into

normal form.

§38. Percolation in grown networks

Grown networks are graphs whose edges are connected probabilistically as vertices are added, meant to model the structure of real-world networks that develop incrementally. A simple model—first introduced by Callaway, Hopcroft, Kleinberg, Newman, and Strogatz—is as follows.[309] Start with a graph on one vertex. Every timestep, add one vertex, and with probability p choose two vertices at random and add an edge between them. The connection to real-world networks emerges from the heuristic structure that develops: early vertices are more likely to have higher degrees and more connectivity, while later ones are connected more sparsely.

In the long-time limit, this model has a novel phase transition at $p_c = 1/8$. For $p < p_c$ the largest connected cluster of vertices scales more slowly than the total elapsed time T, and all clusters are finite. For $p > p_c$ the largest connected cluster has an extensive weight that scales like the distance $\Delta p = p - p_c$ from the critical point like

$$S = \lim_{T\to\infty} \frac{s_{\text{largest}}}{T} \simeq S_0 \exp\left[-\frac{\pi}{2\sqrt{2}} \Delta p^{-1/2} \right], \tag{5.27}$$

with $S_0 \simeq 0.590$.[310] At the critical point, the cluster size distribution has the novel asymptotic form

$$n_s \sim s^{-3} \log^{-2} s \tag{5.28}$$

*Much of the thought and work done in this section was in collaboration with David Hathcock.

Far below p_c, the cluster size distribution has the form $n_s \sim s^{-4\sqrt{2\Delta p}/(1-2\sqrt{2\Delta p})-1}$, i.e. as a power law with constantly varying exponent, and as the critical point is approached from below it has the form $n_s \sim s^{-3-4\sqrt{2\Delta p}}$ for $\log s \gg \Delta p^{-1/2}$. Above the critical point, the distribution has a cutoff for large s with $s \gg (S/T)^{-1}$ of the form

$$n_s \sim s^{-3/2} \exp\left[-\frac{Ss}{2e}\right] \tag{5.29}$$

giving the cutoff the form $s_{\max} \simeq 2e/S$.

The form of (5.27) has an essential singularity that closely resembles that present in the correlation length at a BKT transition. Because the BKT correlation length grows steadily as the transition is approached from above, diverges at the transition, and remains infinite below it, while the weight of the infinite cluster is zero below and at the transition but grows steadily above it, many have take to referring to this transition and similar ones that appear in other network models an "inverted BKT" transition. However, insofar as it can be considered a BKT transition, there need not be anything "inverted" about it. Though the weight of the infinite cluster here resembles the form of a correlation length, the weight of the infinite cluster in standard percolation resembles the form of a critical correlation length in the sense that it has a power law form, but the weight of the infinite cluster *is not* the correlation length. At the percolation critical point, the weight of the infinite cluster is *zero*, since even as the correlation length ξ diverges it grows with a power law ξ^{d_f} for fractal dimension $d_f < d$ less than the dimension of the percolation lattice. As the weight of the infinite cluster becomes finite with a

power law above the transition, the correlation length *decreases*, corresponding roughly to the maximum linear extent of finite clusters. Since we see here that $\xi^{d_f} \sim s_{\max} \simeq S^{-1}$, it would be more appropriate to take the correlation length as diverging just as it does approaching the BKT transition—nothing inverted about it.[311] Since the clusters remain power-law distributed below the transition, the correlation length remains infinite there, even while the weight of the infinite cluster is zero.

Besides the phenomenological similarities to the BKT transition, no formal correspondence has been established between these transitions in grown networks and BKT. As we have seen in the case of the hexatic, the BKT transition is somewhat specially tuned, and why some network model should check all its requirements is strange. For instance, what symmetry of the network model preserves the requirement that the flow equations be even in the fugacity-like variable? It's hard to say, because no good renormalization-group technique has been developed to treat the network models directly and even write down flow equations. Therefore, in order to investigate the BKTness of the grown network model, we will have to make use of the same phenomenological comparisons that led to their hazy identification in the first place. We will focus on the tangible question: is the scaling of the grown network percolation consistent with the value of the first subleading flow coefficient for BKT, $b_0 = 3/2$?

Working with the exact solution

Looking at the facts about this model that I rattled off above—with πs, es, and other order one constants strewn about—one might wonder what sort

of insight can be gained from existing analytic information about the model. After all, we know from Pelissetto and Vicari exactly what the corrections to scaling for the correlation length are in terms of b_0; with all that information, why not just write down the subleading term and compare coefficients? Alas, the method by which the exact leading behavior (5.27) is come by does not generalize easily to higher orders. Here is the approach up to the point that we have stalled, largely following the solution by Dorogovtsev, Mendes, and Samukhin.[310]

The weight of the infinite component can be found from evaluating a generating function g, with $S = 1 - g(1)$. The generating function obeys the differential equation

$$x g'(x) = \frac{1}{2p} \frac{x - g(x)}{1 - g(x)} \tag{5.30}$$

with boundary condition $g(0) = 0$. Note that even numerically solving this equation is challenging, since the solution $g(1)$ approaches 1 as p goes to p_c, making the equation singular. Note further that the relevance of $p_c = 1/8$ is completely opaque. We make it more clear with a change of variables, defining $u(y) = 1 - g(1 - x)$, which yields

$$(1 - y) u'(y) = \frac{1}{2p} \frac{u(y) - y}{u(y)} \tag{5.31}$$

This equation is no more integrable than the previous one, but in the limit of small y we can approximate $1 - y \simeq 1$. This makes the equation separable, as

we can now write $u(y) = v(y)y$, and

$$v'(y)y = -\frac{1}{2p}\frac{2pv^2(y) - v(y) + 1}{v(y)} \tag{5.32}$$

Finally, a hint of the significance of p_c emerges: the roots of the polynomial in the numerator are

$$v(y) = \frac{1 \pm \sqrt{1 - 8p}}{4p} \tag{5.33}$$

and therefore at $p_c = 1/8$ they become degenerate. At exactly that value, the equation is integrated to give

$$u(y)|_{p=1/8} = 2y(1 - f(y)) \tag{5.34}$$

where the function f is the solution to the equation

$$\log(yf(y)) + \frac{1}{f(y)} = \log c \tag{5.35}$$

for a numeric constant c. This "threshold" equation will become useful later. Having separated (5.32), we can integrate it, yielding

$$\begin{aligned} C + \log y = \int \frac{dy}{y} &= -2p \int \frac{v}{2pv^2 - v + 1}\, dv \\ &= \frac{1}{\sqrt{8p-1}} \tan^{-1}\left(\frac{4pv - 1}{\sqrt{8p-1}}\right) - \frac{1}{2}\log(1 - v + 2pv^2) \end{aligned} \tag{5.36}$$

or with the function u restored,

$$C = \frac{1}{\sqrt{8p-1}} \tan^{-1}\left(\frac{4pu(y)/y - 1}{\sqrt{8p-1}}\right) - \frac{1}{2}\log(y^2 - u(y)y + 2pu^2(y)) \quad (5.37)$$

The constant C is set by requiring $u(0) = S$, which gives

$$C = -\frac{\pi}{2\sqrt{8p-1}} - \frac{1}{2}\log(2pS^2) \quad (5.38)$$

Expanding right hand side of (5.37) in powers of $\Delta p = p - p_c$, we find

$$-\frac{\pi}{4\sqrt{2\Delta p}} - \frac{1}{2}\log(2pS^2) = \frac{\pi}{4\sqrt{2\Delta p}} - \frac{1}{1 - u(y)/2y} - \frac{1}{2}\log\left(\frac{1}{4}(1 - u(y)/2y)^2\right) + O(\Delta p^{1/2}) \quad (5.39)$$

The remaining dependence on y can be eliminated using the threshold equation, which should be approximately satisfied for small Δp. Substituting,

$$\begin{aligned} -\frac{\pi}{2\sqrt{2\Delta p}} - \log(S/2) &= -\frac{1}{f(y)} - \frac{1}{2}\log\left(\frac{1}{4}f^2(y)\right) \\ &= -\log c + \log 2 \end{aligned} \quad (5.40)$$

which allows us to finally solve for the equation (5.27).

Notice the places that an ultimately uncontrolled approximation enters this calculation if we wish to make a consistent expansion in Δp. First, we approximate the differential equation near the point we would eventually like to evaluate it. Second, we use the solution to our approximate equation at the critical point as approximately true in the vicinity of the point. The one

controlled approximation, expanding the inverse tangent in powers of Δp, is useless for our purposes, since—after the first constant term of $-\pi/2$—that expansion only yields terms *odd* in $\sqrt{\Delta p}$, which in turn only yield even powers of $\sqrt{\Delta p}$ as corrections, *and therefore miss the primary correction we are hoping to match.* The less-controlled approximations related to the differential equation are necessary in order to get the $\sqrt{\Delta p}$ correction we're looking for.

There are a couple possible routes here, though I am not certain of the tools that exist to confront them. One involves focus on the approximation that renders the differential equation separable: given an approximate solution to a differential equation that is *nearly* separable, can a self-consistent expansion be made in the magnitude of the non-separable term? The second is related to connecting the critical solution into the equation to get the finial result—can that critical solution be usefully made perturbative in the distance from criticality? More work needs to be done.

Numeric studies

With an analytic route stymied by a messy set of uncontrolled approximations, we now turn to numeric options. A direct approach, attempting to measure the infinite cluster size either by solving the differential equation for the generating function or by simulating the growth process and then fitting the result as a function of p, requires unreasonable precision, since we are looking to fit the singular correction to an essential singularity, which vanishes extremely quickly and dominates the behavior.

We therefore will resort to finite size scaling.[312] We have written a routine

that generates grown networks of size N and computes their cluster size distributions, and the distribution of their largest component, using the Newman–Ziff algorithm.[55] We make the following scaling theory. First, we suppose the infinite-size correlation length is of the form (5.21), following Pelissetto and Vicari.[300] We expect that $Q \propto \Delta p$ for sufficiently small Δp. We expect that the infinite-size weight of the infinite cluster above p_c should be related to the correlation length by $S_\infty = \xi_\infty^{-d_f}$. This, the exact asymptotic solution, and the relationship between ξ_∞ and Q from (5.21) uniquely set $Q = 8d_f^2 \Delta p$.

Finite size effects enter in the measure of any observable X compared to its infinite-size value X_∞ via the relationship

$$\frac{X(Q,L)}{X_\infty(Q)} = A_X(\xi_\infty/L) + QB_X(\xi_\infty/L) + O(Q^2) \tag{5.41}$$

so long as we are not at the critical point.[300] At the critical point, observables scale like

$$\frac{X(L)}{L^{-\alpha_X}} = A\left(\mu + \frac{b_0}{3}\log\mu\right)^{\beta_X}[1 + \mathcal{O}(\mu^{-1})] \tag{5.42}$$

for $\mu = \log(Le^{-K})$ for some nonuniversal constant K.[300] For the weight of the infinite cluster, we therefore expect finite-size scaling of the form

$$S(p_c, N) = aN^{-y_S}\left(\mu + \frac{b_0}{3}\log\mu\right)^{\beta_S} \tag{5.43}$$

where y_S is the codimension of the infinite cluster. Since we are working with a model in infinite dimensions, it's not obvious how L should be replaced in the logarithmic corrections—two simple choices are $L = N$, and

$L = \lim^{d\to\infty} N^{1/d} = 1$ (for which we would see no logarithmic corrections and the project to use this form is in vain). Approaching the critical point from above, the weight of the infinite cluster should have a scaling form

$$S(p,N) = S_\infty(p)\Big[A_S(\xi_\infty/L) + QB_S(\xi_\infty/L) + O(Q^2)\Big] \tag{5.44}$$

where the same considerations about L apply.

A fit of (5.43) to simulated network data is shown in Figure 5.1 as an orange line, and yields $y_S \simeq 0.54$. A green line is also shown with a pure power law with the same exponent for comparison. The corrections to scaling for the infinite cluster weight at the critical point appear very small, and the fit does not fix b_0 at all, allow it to vary over several orders of magnitude.

Slightly more discerning data comes from above p_c. That data is pictured collapsed in Figure 5.2. A fit gives $\xi_\infty = S_\infty^{-d_f}$ with $d_f \simeq 1/y_S$, which lends support to the notion that the natural length scale in this problem is $L \sim \xi \sim S^{-1/y_S} \sim N$. The curve traced by the collapsed data is a measurement of the function A_S for this system. An inset shows detail of the collapse at small ξ_∞/N, a region accessible when the system size is much bigger than the correlation length. The failure of the collapse in this region is expected given the finite-size scaling form (5.44), which predicts corrections to scaling that grow linearly with $Q \simeq \Delta p$; as the color indicates, the deviation from the collapse appears to grow linearly with Δp, and may be the sign of a nonzero function B_S. The fit at present does nothing to restrict I_1 in (5.21), which like b_0 varies freely over several orders of magnitude. Perhaps a more careful fit of data in the region of the inset will yield more information about these

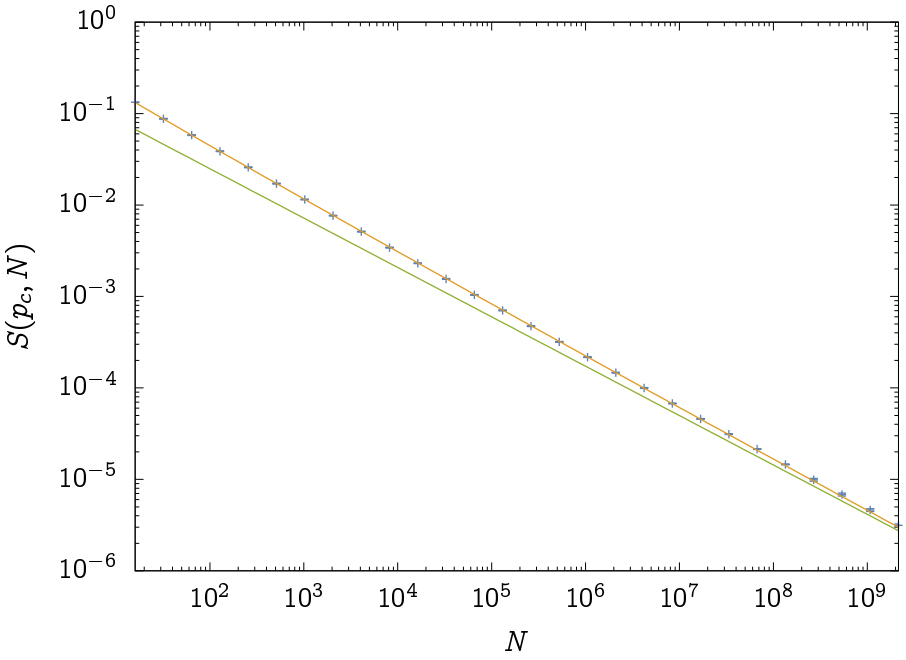

Figure 5.1: The average weight of the largest cluster in a grown network as a function of size N. A fit to the scaling form is shown (yellow), along with a fit to a pure power law (green).

corrections to scaling.

The finite size scaling of the full cluster size distribution might offer better access to corrections to scaling, though the principal scaling of that distribution is already challenging. Recall that at the critical point, the cluster size distribution scales like $s^{-3}\log^{-2} s$. We empirically see something of this form, as seen in Figure 5.3. Assuming that the distribution is cutoff like $S^{-1} \sim \xi^{df}$, we would write

$$n_s = s^{-3}\log^{-2} s\tilde{\mathcal{N}}(sS, N\xi^{-1}) = s^{-3}\log^{-2} s\mathcal{N}(sN^{-d_f}, \xi N^{-1}) \qquad (5.45)$$

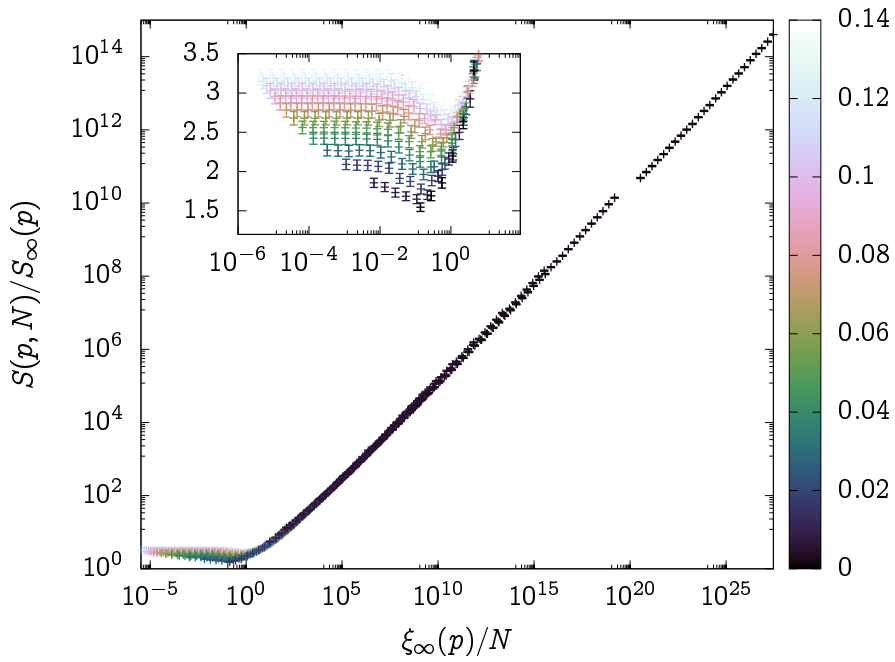

Figure 5.2: The ratio of the average weight of the largest cluster in a grown network to the weight of an infinite cluster, as a function of an estimate for the infinite-size correlation length ξ_∞ over the network size N. The color shows the value of Δp for which the data was taken.

This gives moments of the form

$$\begin{aligned}\langle s^n\rangle &= \int s^n n_s\,ds = \int s^{n-3}\log^{-2} s\mathcal{N}(sN^{-d_f},\xi N^{-1})\,ds \\ &= \int (yN^{d_f})^{n-3}\log^{-2}(yN^{d_f})\mathcal{N}(y,\xi N^{-1})N^{d_f}\,dy \\ &= N^{d_f(n-2)}\int y^{n-3}\log^{-2}(yN^{d_f})\mathcal{N}(y,\xi N^{-1})\,dy\end{aligned} \tag{5.46}$$

The logarithm inside the integral makes it difficult to draw conclusions about the finite-size scaling of the moments from ordinary scaling arguments.

*I made the mistake of having my program take N as an `unsigned int`, which on my architecture is 32 bits and therefore has a maximum value of $2^{32}-1$. Luckily I was forward

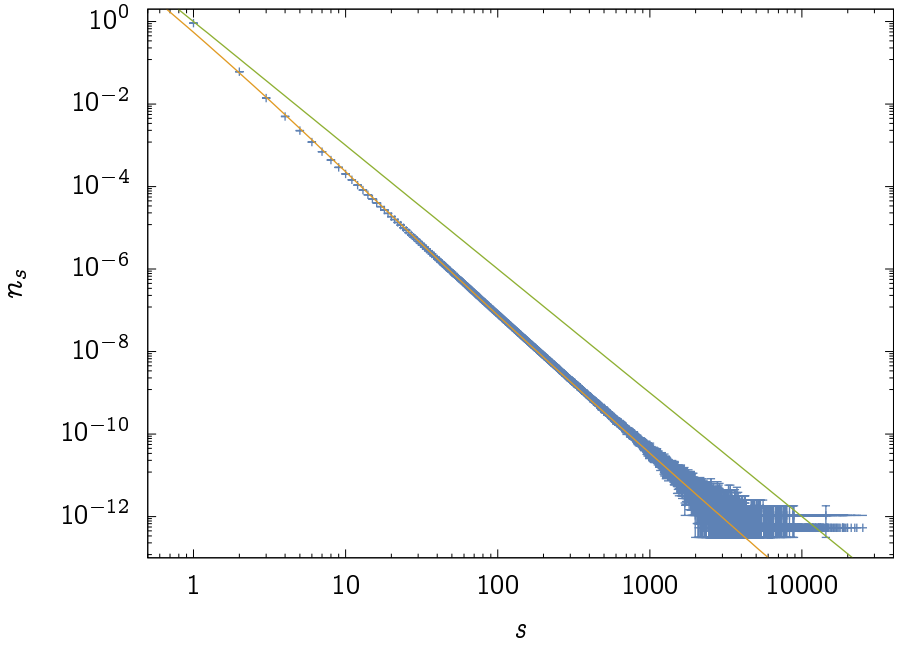

Figure 5.3: The cluster size distribution at criticality measured for networks of size $N = 2^{31}$.* A fit to the asymptotic form is shown in yellow, while a pure s^{-3} power law is shown in green for reference.

§39. Conclusions

Other universality classes and families definitely do exist, characterized by novel values for σ. The level-1 SU(N) Wess–Zumino–Witten model has been found to be characterized by $\sigma = N/(N+2)$.[313] Dislocated-mediated melting alone has produced a melange of anomalous exponents, with $\sigma = 1/2$, $\sigma = 2/5$, and $\sigma = 0.369\,63\ldots$ depending on precise specification of the model and the lattice geometry.[306,314] Topological transitions in systems whose vortices are non-Abelian produce several series of σ values dependent on particular

thinking enough to make the bins `uint64`s or even this data (which counts a few hundred times more than 2^{32} size-one clusters) would be spoiled.

symmetry.[315] Each value of σ indicates either a different universality family or merely a different class within the same family depending on how it affects the terms at quadratic order.

There are also many other network models that produce infinite order transitions. Percolation on branching simplicial complexes has recently been shown to have an infinite cluster that emerges with a stretched exponential with anomalous exponents $\sigma = (s-3)/(s-1)$ that vary with integer $s > 3$.[316] The authors claim the case for $s = 5$ that results in $\sigma = 1/2$ reduces to the BKT singularity, but it is unknown whether the full symmetry of the BKT theory is reproduced by this choice or if choice of s has simply tuned them there in only the quadratic order. Other complex networks see similarly singular percolation transitions.[317,318] Besides percolation transitions, taking certain scale-free networks and slapping spins on their vertices produce lattice models that have novel BKT-like phenomena.[318–320]

Other traditional statistical mechanics models, including ones that have exact solutions or detailed perturbative theories, are thought to exhibit BKT universality on the basis of their anomalous exponent alone. The Kondo problem, which is the model of a spin interacting with a bath of fermions, is one such example, and we hope to examine its exact solution to see if it is amenable to the extraction of corrections to scaling.[321–323] The 1D Ising model with long-ranged inverse square interactions also has a BKT-like transition, and may also be amenable to a systematic study of corrections to scaling.[322,324–328]

A classification of possible bifurcations and corresponding simplest normal forms is in order for flow equations whose leading order is quadratic, and whose

expansions are constrained or not by various symmetries. This would be the first step in developing techniques for distinguishing between universality classes and families of this type using experimental or simulation data.

CHAPTER VI

MODULATED NEMATIC AS HIDDEN ORDER IN URu_2Si_2

We have spent the last five chapters with the closest connection to reality being, at best, numeric experiments and the expectation of universality. Now, reality comes crashing down, and we shall see an example of the application of mean-field scaling ideas to explain actual experimental data. In this chapter, we'll examine a set of simple explanations for a strange feature in ultrasound data near a mysterious phase transition in a uranium compound, and discover than only one of those explanations is consistent.

§40. Introduction

URu_2Si_2 is a paradigmatic example of a material with an ordered state whose broken symmetry remains unknown. This state, known as *hidden order* (HO), sets the stage for unconventional superconductivity that emerges at even lower temperatures. At sufficiently large hydrostatic pressures, both superconductivity and HO give way to local moment antiferromagnetism (AFM).[330] Modern

*The material of this chapter was posted online in J Kent-Dobias, M Matty, and B Ramshaw, "Elastic properties of hidden order in URu_2Si_2 are reproduced by a staggered nematic", (2019), arXiv:1910.01669v2 [cond-mat.str-el]. The work and writing was done in collaboration with the listed authors.

theories[331–348] propose associating any of a variety of broken symmetries with HO. This work analyzes a family of phenomenological models with order parameters of general symmetry that couple linearly to strain. Of these, only one is compatible with two experimental observations: first, the B_{1g} "nematic" elastic susceptibility $(C_{11} - C_{12})/2$ softens anomalously from room temperature down to $T_{HO} = 17.5$ K;[349] and second, a B_{1g} nematic distortion is observed by x-ray scattering under sufficient pressure to destroy the HO state.[350]

Recent resonant ultrasound spectroscopy (RUS) measurements were used to examine the thermodynamic discontinuities in the elastic moduli at T_{HO}.[351] The observation of discontinues only in compressional, or A_{1g}, elastic moduli requires that the point-group representation of HO be one-dimensional. This rules out many order parameter candidates[340–344,348,352] in a model-independent way, but doesn't differentiate between those that remain.

Recent x-ray experiments discovered rotational symmetry breaking in URu_2Si_2 under pressure.[350] Above 0.13–0.5 GPa (depending on temperature), URu_2Si_2 undergoes a B_{1g} nematic distortion, which might be related to the anomalous softening of the B_{1g} elastic modulus $(C_{11} - C_{12})/2$ that occurs over a broad temperature range at zero pressure.[353,354] Motivated by these results—which hint at a B_{1g} strain susceptibility associated with the HO state—we construct a phenomenological mean field theory for an arbitrary OP coupled to strain, and then determine the effect of its phase transitions on the elastic response in different symmetry channels.

We find that only one OP representation reproduces the anomalous B_{1g} elastic modulus, which softens in a Curie–Weiss-like manner from room temperature and then cusps at T_{HO}. That theory associates HO with a B_{1g} OP

modulated along the *c*-axis, the high pressure state with uniform B_{1g} order, and the triple point between them with a Lifshitz point. In addition to the agreement with the ultrasound data across a broad temperature range, the theory predicts uniform B_{1g} strain at high pressure—the same distortion that was recently seen in x-ray scattering experiments.[350] This theory strongly motivates future ultrasound experiments under pressure approaching the Lifshitz point, which should find that the $(C_{11} - C_{12})/2$ modulus diverges as the uniform B_{1g} strain of the high pressure phase is approached.

§41. Model & phase diagram

The point group of URu_2Si_2 is D_{4h}, and any theory must locally respect this symmetry in the high-temperature phase. Our phenomenological free energy density contains three parts: the elastic free energy, the OP, and the interaction between strain and OP. The most general quadratic free energy of the strain ϵ is $f_{\text{ELASTIC}} = C^0_{ijkl}\epsilon_{ij}\epsilon_{kl}$.* The form of the bare moduli tensor C^0 is further restricted by symmetry. Linear combinations of the six independent components of strain form five irreducible components of strain in D_{4h} as

$$\begin{aligned} \epsilon_{A_{1g},1} &= \epsilon_{11} + \epsilon_{22} & \epsilon_{B_{1g}} &= \epsilon_{11} - \epsilon_{22} \\ \epsilon_{A_{1g},2} &= \epsilon_{33} & \epsilon_{B_{2g}} &= 2\epsilon_{12} \\ \epsilon_{E_g} &= 2\{\epsilon_{11}, \epsilon_{22}\}. \end{aligned} \tag{6.1}$$

*Components of the elastic modulus tensor C were given in the popular Voigt notation in the introduction. Here and henceforth the notation used is that natural for a rank-four tensor.

All quadratic combinations of these irreducible strains that transform like $\mathrm{A_{1g}}$ are included in the free energy,[355]

$$f_{\text{ELASTIC}} = \frac{1}{2}\sum_{\mathrm{X}} C^0_{\mathrm{X},ij}\epsilon_{\mathrm{X},i}\epsilon_{\mathrm{X},j}, \tag{6.2}$$

where the sum is over irreducible representations of the point group and the bare elastic moduli C^0_{X} are

$$\begin{aligned} C^0_{\mathrm{A_{1g}},11} &= \tfrac{1}{2}(C^0_{1111}+C^0_{1122}) & C^0_{\mathrm{B_{1g}}} &= \tfrac{1}{2}(C^0_{1111}-C^0_{1122}) \\ C^0_{\mathrm{A_{1g}},22} &= C^0_{3333} & C^0_{\mathrm{B_{2g}}} &= C^0_{1212} \\ C^0_{\mathrm{A_{1g}},12} &= C^0_{1133} & C^0_{\mathrm{E_g}} &= C^0_{1313}. \end{aligned} \tag{6.3}$$

The interaction between strain and an OP η depends on the point group representation of η. If this representation is X, the most general coupling to linear order is

$$f_{\text{INT}} = -b^{(i)}\epsilon^{(i)}_{\mathrm{X}}\eta. \tag{6.4}$$

Many high-order interactions are permitted, and in §44 another of the form $\epsilon^2\eta^2$ is added to the following analysis. If there exists no component of strain that transforms like the representation X then there can be no linear coupling. The next-order coupling is linear in strain, quadratic in order parameter, and the effect of this coupling at a continuous phase transition is to produce a jump in the $\mathrm{A_{1g}}$ elastic moduli if η is single-component,[356–358] and jumps in other elastic moduli if multicomponent.[351] Because we are interested in physics that anticipates the phase transition—for instance, that the growing

OP susceptibility is reflected directly in the elastic susceptibility—we will focus our attention on OPs that can produce linear couplings to strain. Looking at the components present in (6.1), this rules out all of the u-reps (which are odd under inversion), the A_{2g} irrep, and all half-integer (spinor) representations.

If the OP transforms like A_{1g} (e.g. a fluctuation in valence number), odd terms are allowed in its free energy and without fine-tuning any transition will be first order and not continuous. Since the HO phase transition is second-order,[349] we will henceforth rule out A_{1g} OPs as well. For the OP representation X as any of those remaining—B_{1g}, B_{2g}, or E_g—the most general quadratic free energy density is

$$f_{\text{OP}} = \frac{1}{2}\left[r\eta^2 + c_{\parallel}(\nabla_{\parallel}\eta)^2 + c_{\perp}(\nabla_{\perp}\eta)^2 + D_{\perp}(\nabla_{\perp}^2\eta)^2\right] + u\eta^4, \tag{6.5}$$

where $\nabla_{\parallel} = \{\partial_1, \partial_2\}$ transforms like E_u, and $\nabla_{\perp} = \partial_3$ transforms like A_{2u}. Other quartic terms are allowed—especially many for an E_g OP—but we have included only those terms necessary for stability when either r or $c_{\perp}$ become negative. The full free energy functional of η and ϵ is

$$\begin{aligned} F[\eta, \epsilon] &= F_{\text{OP}}[\eta] + F_{\text{ELASTIC}}[\epsilon] + F_{\text{INT}}[\eta, \epsilon] \\ &= \int dx\, (f_{\text{OP}} + f_{\text{ELASTIC}} + f_{\text{INT}}). \end{aligned} \tag{6.6}$$

Rather than analyze this two-argument functional directly, we begin by tracing out the strain and studying the behavior of the OP alone. Later we will invert this procedure and trace out the OP when we compute the effective elastic moduli. The only strain relevant to an OP of representation X at linear coupling is ϵ_X, which can be traced out of the problem exactly in mean field

theory. Extremizing the functional (6.6) with respect to ϵ_{X} gives

$$0 = \left.\frac{\delta F[\eta,\epsilon]}{\delta \epsilon_{\mathrm{X}}(x)}\right|_{\epsilon=\epsilon_\star} = C^0_{\mathrm{X}}\epsilon^\star_{\mathrm{X}}(x) - b\eta(x), \tag{6.7}$$

which in turn gives the strain field conditioned on the state of the OP field as $\epsilon^\star_{\mathrm{X}}[\eta](x) = (b/C^0_{\mathrm{X}})\eta(x)$ at all spatial coordinates x, and $\epsilon^\star_{\mathrm{Y}}[\eta] = 0$ for all other irreps $\mathrm{Y} \neq \mathrm{X}$. Upon substitution into (6.6), the resulting single-argument free energy functional $F[\eta, \epsilon_\star[\eta]]$ has a density identical to f_{OP} with the identification $r \to \tilde{r} = r - b^2/2C^0_{\mathrm{X}}$.

With the strain traced out, (6.5) describes the theory of a Lifshitz point at $\tilde{r} = c_\perp = 0$.[359,360] The properties discussed in the remainder of this section can all be found in a standard text, e.g., in Chapter 4 §6.5 of Chaikin & Lubensky.[361] For a one-component OP ($\mathrm{B_{1g}}$ or $\mathrm{B_{2g}}$) and positive $c_\parallel$, it is traditional to make the field ansatz $\langle\eta(x)\rangle = \eta_* \cos(q_* x_3)$. For $\tilde{r} > 0$ and $c_\perp > 0$, or $\tilde{r} > c_\perp^2/4D_\perp$ and $c_\perp < 0$, the only stable solution is $\eta_* = q_* = 0$ and the system is unordered. For $\tilde{r} < 0$ there are free energy minima for $q_* = 0$ and $\eta_*^2 = -\tilde{r}/4u$ and this system has uniform order of the OP representation, e.g., $\mathrm{B_{1g}}$ or $\mathrm{B_{2g}}$. For $c_\perp < 0$ and $\tilde{r} < c_\perp^2/4D_\perp$ there are free energy minima for $q_*^2 = -c_\perp/2D_\perp$ and

$$\eta_*^2 = \frac{c_\perp^2 - 4D_\perp\tilde{r}}{12D_\perp u} = \frac{\tilde{r}_c - \tilde{r}}{3u} = \frac{|\Delta\tilde{r}|}{3u}, \tag{6.8}$$

with $\tilde{r}_c = c_\perp^2/4D_\perp$ and the system has modulated order. The transition between the uniform and modulated orderings is first order for a one-component OP and occurs along the line $c_\perp = -2\sqrt{-D_\perp\tilde{r}/5}$.

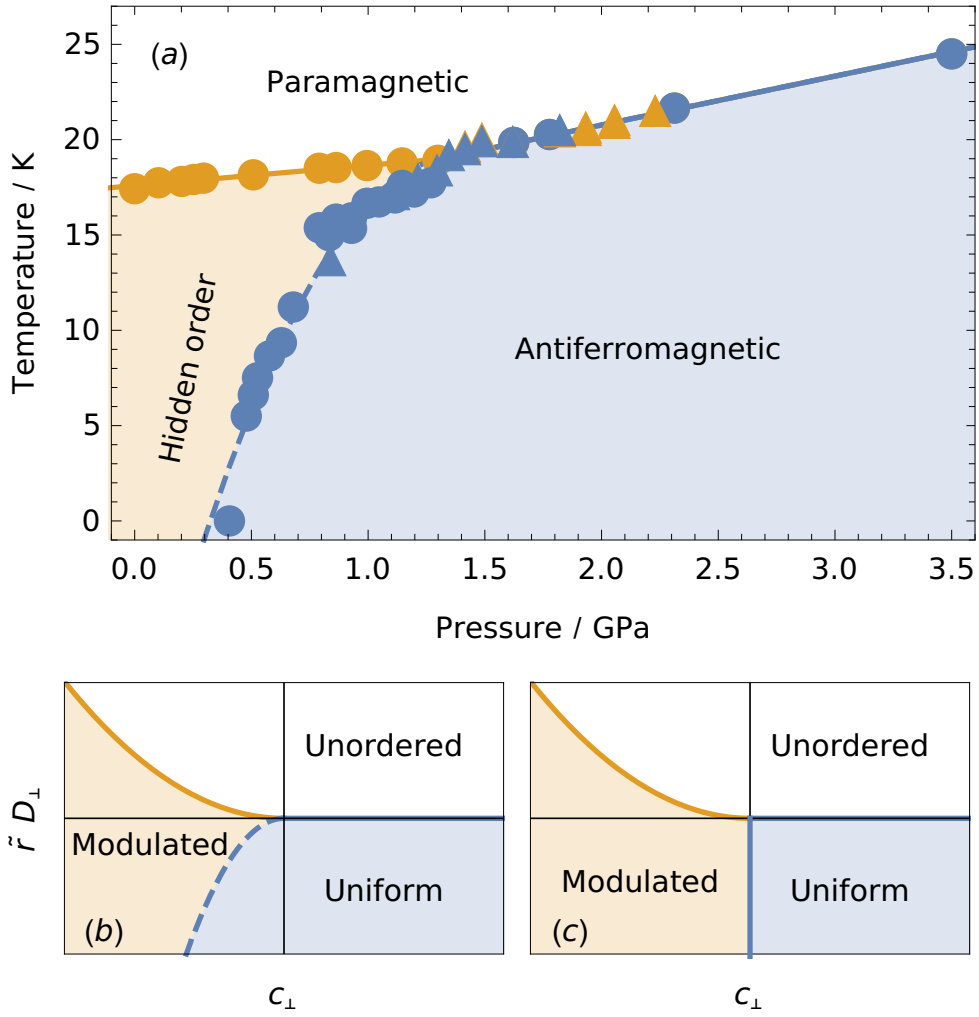

Figure 6.1: Phase diagrams for (a) URu_2Si_2 from experiments (neglecting the superconducting phase)[330] (b) mean field theory of a one-component (B_{1g} or B_{2g}) Lifshitz point (c) mean field theory of a two-component (E_g) Lifshitz point. Solid lines denote continuous transitions, while dashed lines denote first order transitions. Later, when we fit the elastic moduli predictions for a B_{1g} OP to data along the ambient pressure line, we will take $\Delta\tilde{r} = \tilde{r} - \tilde{r}_c = a(T - T_c)$.

For a two-component OP (E_g) we must also allow a relative phase between the two components of the OP. In this case the uniform ordered phase is only stable for $c_\perp > 0$, and the modulated phase is now characterized by helical order with $\langle \eta(x) \rangle = \eta_* \{\cos(q_* x_3), \sin(q_* x_3)\}$. The uniform to modulated transition is now continuous. This does not reproduce the physics of URu_2Si_2, whose HO phase is bounded by a line of first order transitions at high pressure, and so we will henceforth neglect the possibility of a multicomponent order parameter. Schematic phase diagrams for both the one- and two-component models are shown in Figure 6.1.

§42. Susceptibility & elastic moduli

We will now derive the effective elastic tensor C that results from the coupling of strain to the OP. The ultimate result, found in (6.17), is that C_X differs from its bare value C_X^0 only for the representation X of the OP. Moreover, this modulus does not vanish at the unordered to modulated transition—as it would if the transition were a $q = 0$ phase transition—but instead ends in a cusp. In this section we start by computing the susceptibility of the OP at the unordered to modulated transition, and then compute the elastic modulus for the same.

The susceptibility of a single-component (B_{1g} or B_{2g}) OP is

$$\begin{aligned} \chi^{\{-1\}}(x, x') &= \left. \frac{\delta^2 F[\eta, \epsilon_\star[\eta]]}{\delta\eta(x)\delta\eta(x')} \right|_{\eta=\langle\eta\rangle} \\ &= \left[\tilde{r} - c_\parallel \nabla_\parallel^2 - c_\perp \nabla_\perp^2 + D_\perp \nabla_\perp^4 + 12u\langle\eta(x)\rangle^2\right]\delta(x - x'), \end{aligned} \tag{6.9}$$

where $\{-1\}$ indicates a functional reciprocal defined as

$$\int dx''\, \chi^{\{-1\}}(x,x'')\chi(x'',x') = \delta(x-x'). \tag{6.10}$$

Taking the Fourier transform and integrating out q' gives

$$\chi(q) = \Big(\tilde{r} + c_{\|}q_{\|}^2 + c_{\perp}q_{\perp}^2 + D_{\perp}q_{\perp}^4 + 12u\sum_{q'}\langle\tilde{\eta}_{q'}\rangle\langle\tilde{\eta}_{-q'}\rangle\Big)^{-1}. \tag{6.11}$$

Near the unordered to modulated transition this yields

$$\begin{aligned}\chi(q) &= \left[c_{\|}q_{\|}^2 + D_{\perp}(q_*^2 - q_{\perp}^2)^2 + |\Delta\tilde{r}|\right]^{-1} \\ &= \frac{1}{D_{\perp}}\frac{\xi_{\perp}^4}{1+\xi_{\|}^2q_{\|}^2 + \xi_{\perp}^4(q_*^2-q_{\perp}^2)^2},\end{aligned} \tag{6.12}$$

with $\xi_{\perp} = (|\Delta\tilde{r}|/D_{\perp})^{-1/4} = \xi_{\perp 0}|t|^{-1/4}$ and $\xi_{\|} = (|\Delta\tilde{r}|/c_{\|})^{-1/2} = \xi_{\|0}|t|^{-1/2}$, where $t = (T-T_c)/T_c$ is the reduced temperature and $\xi_{\perp 0} = (D_{\perp}/aT_c)^{1/4}$ and $\xi_{\|0} = (c_{\|}/aT_c)^{1/2}$ are the bare correlation lengths perpendicular and parallel to the plane, respectively. The static susceptibility $\chi(0) = (D_{\perp}q_*^4 + |\Delta\tilde{r}|)^{-1}$ does not diverge at the unordered to modulated transition. Though it anticipates a transition with Curie–Weiss-like divergence at the lower point $a(T-T_c) = \Delta\tilde{r} = -D_{\perp}q_*^4 < 0$, this is cut off with a cusp at $\Delta\tilde{r} = 0$.

The elastic susceptibility, which is the reciprocal of the effective elastic modulus, is found in a similar way to the OP susceptibility: we must trace over η and take the second variation of the resulting effective free energy

functional of ϵ alone. Extremizing over η yields

$$0 = \left.\frac{\delta F[\eta,\epsilon]}{\delta\eta(x)}\right|_{\eta=\eta_\star} = \left.\frac{\delta F_{\text{OP}}[\eta]}{\delta\eta(x)}\right|_{\eta=\eta_\star} - b\epsilon_{\text{X}}(x), \tag{6.13}$$

which implicitly gives $\eta_\star[\epsilon]$, the OP conditioned on the configuration of the strain. Since $\eta_\star$ is a functional of ϵ_{X} alone, only the modulus C_{X} will be modified from its bare value C^0_{X}.

Though the differential equation for $\eta_\star$ cannot be solved explicitly, we can use the inverse function theorem to make use of (6.13) anyway. First, denote by $\eta_\star^{-1}[\eta]$ the inverse functional of $\eta_\star$ implied by (6.13), which gives the function ϵ_{X} corresponding to each solution of (6.13) it receives. This we can immediately identify from (6.13) as $\eta_\star^{-1}[\eta](x) = b^{-1}(\delta F_{\text{OP}}[\eta]/\delta\eta(x))$. Now, we use the inverse function theorem to relate the functional reciprocal of the derivative of $\eta_\star[\epsilon]$ with respect to ϵ_{X} to the derivative of $\eta_\star^{-1}[\eta]$ with respect to η, yielding

$$\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\text{X}}(x')}\right)^{\{-1\}} = \left.\frac{\delta\eta_\star^{-1}[\eta](x)}{\delta\eta(x')}\right|_{\eta=\eta_\star[\epsilon]} = b^{-1}\left.\frac{\delta^2 F_{\text{OP}}[\eta]}{\delta\eta(x)\delta\eta(x')}\right|_{\eta=\eta_\star[\epsilon]}. \tag{6.14}$$

Next, (6.13) and (6.14) can be used in concert with the ordinary rules of

functional calculus to yield the second variation

$$\begin{aligned}
\frac{\delta^2 F[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon_{\rm X}(x)\delta\epsilon_{\rm X}(x')} &= C^0_{\rm X}\delta(x-x') - 2b\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')} - b\int dx''\,\frac{\delta^2\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')\delta\epsilon_{\rm X}(x'')}\epsilon_{\rm X}(x'') \\
&\quad + \int dx''\,\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon_{\rm X}(x)\delta\epsilon_{\rm X}(x')}\frac{\delta F_{\rm OP}[\eta]}{\delta\eta(x'')}\bigg|_{\eta=\eta_\star[\epsilon]} + \int dx''\,dx'''\,\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon_{\rm X}(x)}\frac{\delta\eta_\star[\epsilon](x''')}{\delta\epsilon_{\rm X}(x')}\frac{\delta^2 F_{\rm OP}[\eta]}{\delta\eta(x'')\delta\eta(x''')}\bigg|_{\eta=\eta_\star[\epsilon]} \\
&= C^0_{\rm X}\delta(x-x') - 2b\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')} - b\int dx''\,\frac{\delta^2\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')\delta\epsilon_{\rm X}(x'')}\epsilon_{\rm X}(x'') \\
&\quad + \int dx''\,\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon_{\rm X}(x)\delta\epsilon_{\rm X}(x')}(b\epsilon_{\rm X}(x'')) + b\int dx''\,dx'''\,\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon_{\rm X}(x)}\frac{\delta\eta_\star[\epsilon](x''')}{\delta\epsilon_{\rm X}(x')}\left(\frac{\partial\eta_\star[\epsilon](x'')}{\partial\epsilon_{\rm X}(x''')}\right)^{\{-1\}} \\
&= C^0_{\rm X}\delta(x-x') - 2b\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')} + b\int dx''\,\delta(x-x'')\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon_{\rm X}(x')} \\
&= C^0_{\rm X}\delta(x-x') - b\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')}.
\end{aligned} \tag{6.15}$$

The elastic modulus is given by the second variation (6.15) evaluated at the extremized strain $\langle\epsilon\rangle$. To calculate it, note that evaluating the second variation of $F_{\rm OP}$ in (6.14) at $\langle\epsilon\rangle$ (or $\eta_\star(\langle\epsilon\rangle) = \langle\eta\rangle$) yields

$$\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon_{\rm X}(x')}\right)^{\{-1\}}\bigg|_{\epsilon=\langle\epsilon\rangle} = b^{-1}\chi^{\{-1\}}(x,x') + \frac{b}{C^0_{\rm X}}\delta(x-x'), \tag{6.16}$$

where $\chi^{\{-1\}}$ is the OP susceptibility given by (6.9). Upon substitution into (6.15) and taking the Fourier transform of the result, we finally arrive at

$$C_{\rm X}(q) = C^0_{\rm X} - b\Big(\frac{1}{b\chi(q)} + \frac{b}{C^0_{\rm X}}\Big)^{-1} = C^0_{\rm X}\Big(1 + \frac{b^2}{C^0_{\rm X}}\chi(q)\Big)^{-1}. \tag{6.17}$$

Though not relevant here, this result generalizes to multicomponent OPs.

What does (6.17) predict in the vicinity of the HO transition? Near the

disordered to modulated transition—the zero-pressure transition to the HO state—the static modulus is given by

$$C_{\mathrm{X}}(0) = C^0_{\mathrm{X}}\left[1 + \frac{b^2}{C^0_{\mathrm{X}}}\left(D_\perp q_*^4 + |\Delta\tilde{r}|\right)^{-1}\right]^{-1}. \tag{6.18}$$

This corresponds to a softening in the X-modulus approaching the transition that is cut off with a cusp of the form $|\Delta\tilde{r}|^\gamma \propto |T - T_c|^\gamma$ with $\gamma = 1$. This is our main result. The only OP irreps that couple linearly with strain and reproduce the topology of the URu_2Si_2 phase diagram are B_{1g} and B_{2g}. For either of these irreps, the transition into a modulated rather than uniform phase masks traditional signatures of a continuous transition by locating thermodynamic singularities at nonzero $q = q_*$. The remaining clue at $q = 0$ is a particular kink in the corresponding modulus.

§43. Comparison to experiment

RUS experiments[351] yield the individual elastic moduli broken into irreps; data for the B_{1g} and B_{2g} components defined in (6.1) are shown in Figures 6.2(a–b). The B_{2g} modulus in Fig. 6.2(a) doesn't appear to have any response to the presence of the transition, exhibiting the expected linear stiffening upon cooling from room temperature, with a low-temperature cutoff at some fraction of the Debye temperature.[362] The B_{1g} modulus Fig. 6.2(b) has a dramatic response, softening over the course of roughly 100 K and then cusping at the HO transition. The data in the high-temperature phase can be fit to the theory (6.18), with a linear background modulus $C^0_{\mathrm{B}_{1g}}$ and $\tilde{r} - \tilde{r}_c = a(T - T_c)$, and the result is shown in Figure 6.2(b).

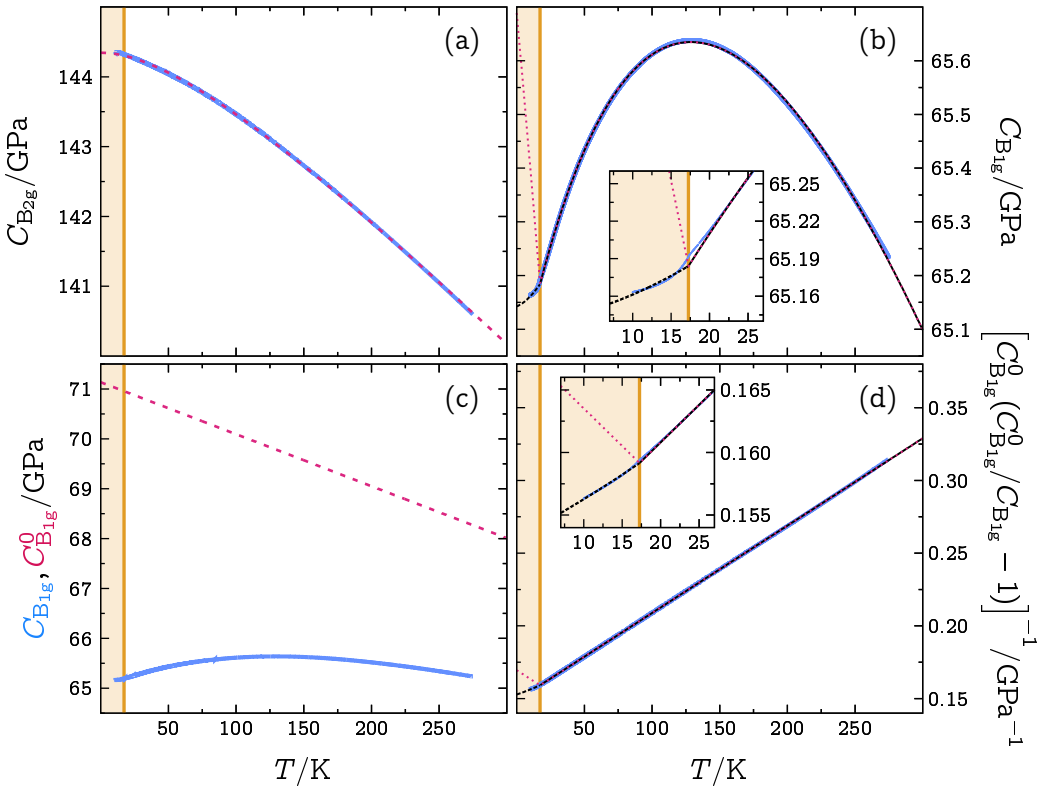

Figure 6.2: RUS measurements of the elastic moduli of URu_2Si_2 at ambient pressure as a function of temperature from recent experiments[351] (blue, solid) alongside fits to theory (magenta, dashed and black, solid). The solid yellow region shows the location of the HO phase. (a) B_{2g} modulus data and a fit to the standard form.[362] (b) B_{1g} modulus data and a fit to (6.18) (magenta, dashed) and a fit to (6.41) (black, solid). The fit gives $C^0_{B_{1g}} \simeq \left[73 - (0.012\,\mathrm{K}^{-1})T\right]$ GPa, $D_\perp q_*^4/b^2 \simeq 0.12\,\mathrm{GPa}^{-1}$, and $a/b^2 \simeq 3.7 \times 10^{-4}\,\mathrm{GPa}^{-1}\,\mathrm{K}^{-1}$. Addition of a quadratic term in $C^0_{B_{1g}}$ was here not needed for the fit.[362] (c) B_{1g} modulus data and the fit of the *bare* B_{1g} modulus. (d) B_{1g} modulus data and the fits transformed by $[C^0_{B_{1g}}(C^0_{B_{1g}}/C_{B_{1g}} - 1)]]^{-1}$, which is predicted from (6.18) to equal $D_\perp q_*^4/b^2 + a/b^2|T - T_c|$, e.g., an absolute value function.

The behavior of the modulus below the transition does not match (6.18) well, but this is because of the truncation of the free energy expansion used above. Higher order terms like $\eta^2\epsilon^2$ contribute to the modulus starting at order η_*^2, and therefore while they do not affect the behavior above the transition, they change the behavior below it. To demonstrate this, in §44 we compute the modulus in a theory where the interaction free energy is truncated after fourth order with new term $\frac{1}{2}g\eta^2\epsilon^2$. The thin solid black line in Fig. 6.2 shows the fit of the RUS data to (6.41) and shows that high-order corrections can account for the low-temperature behavior.

The data and theory appear quantitatively consistent, suggesting that HO can be described as a B_{1g}-nematic phase that is modulated at finite q along the $c-$axis. The predicted softening appears over hundreds of Kelvin; Figures 6.2(c–d) show the background modulus $C^0_{B_{1g}}$ and the OP–induced response isolated from each other.

We have seen that the mean-field theory of a B_{1g} OP recreates the topology of the HO phase diagram and the temperature dependence of the B_{1g} elastic modulus at zero pressure. This theory has several other physical implications. First, the association of a modulated B_{1g} order with the HO phase implies a *uniform* B_{1g} order associated with the high pressure phase, and moreover a uniform B_{1g} strain of magnitude $\langle\epsilon_{B_{1g}}\rangle^2 = b^2\tilde{r}/4u(C^0_{B_{1g}})^2$, which corresponds to an orthorhombic structural phase. The onset of orthorhombic symmetry breaking was recently detected at high pressure in URu_2Si_2 using x-ray diffraction, a further consistency of this theory with the phenomenology of URu_2Si_2.[350]

Second, as the Lifshitz point is approached from low pressure, this theory

predicts that the modulation wavevector q_* should vanish continuously. Far from the Lifshitz point we expect the wavevector to lock into values commensurate with the space group of the lattice, and moreover that at zero pressure, where the RUS data here was collected, the half-wavelength of the modulation should be commensurate with the lattice spacing $a_3 \simeq 9.68\,\text{Å}$, or $q_* = \pi/a_3 \simeq 0.328\,\text{Å}^{-1}$.[363–367] In between these two regimes, mean field theory predicts that the ordering wavevector shrinks by jumping between ever-closer commensurate values in the style of the devil's staircase.[368] In reality the presence of fluctuations may wash out these transitions.

This motivates future ultrasound experiments done under pressure, where the depth of the cusp in the B_{1g} modulus should deepen (perhaps with these commensurability jumps) at low pressure and approach zero as $q_*^4 \sim (c_\perp/2D_\perp)^2$ near the Lifshitz point. Alternatively, RUS done at ambient pressure might examine the heavy Fermi liquid to AFM transition by doping. Though previous RUS studies have doped URu_2Si_2 with rhodium,[369] the magnetic rhodium dopants likely promote magnetic phases. A non-magnetic dopant such as phosphorous may more faithfully explore the transition out of the HO phase. Our work also motivates experiments that can probe the entire correlation function—like x-ray and neutron scattering—and directly resolve its finite-q divergence. The presence of spatial commensurability is known to be irrelevant to critical behavior at a one-component disordered to modulated transition, and therefore is not expected to modify the thermodynamic behavior otherwise.[370]

There are two apparent discrepancies between the orthorhombic strain in the phase diagram presented by recent x-ray data[350], and that predicted by

our mean field theory if its uniform B_{1g} phase is taken to be coincident with URu_2Si_2's AFM. The first is the apparent onset of the orthorhombic phase in the HO state at slightly lower pressures than the onset of AFM. As the recent x-ray research[350] notes, this misalignment of the two transitions as function of doping could be due to the lack of an ambient pressure calibration for the lattice constant. The second discrepancy is the onset of orthorhombicity at higher temperatures than the onset of AFM. We note that magnetic susceptibility data sees no trace of another phase transition at these higher temperatures.[371] It is therefore possible that the high-temperature orthorhombic signature in x-ray scattering is not the result of a bulk thermodynamic phase, but instead marks the onset of short-range correlations, as it does in the high-T_c cuprates[372] (where the onset of CDW correlations also lacks a thermodynamic phase transition).

Three dimensions is below the upper critical dimension $4\frac{1}{2}$ of a one-component disordered-to-modulated transition, and so mean field theory should break down sufficiently close to the critical point due to fluctuations, at the Ginzburg temperature.[373,374] Magnetic phase transitions tend to have a Ginzburg temperature of order one. Our fit above gives $\xi_{\perp 0} q_* = (D_\perp q_*^4 / a T_c)^{1/4} \simeq 2$, which combined with the speculation of $q_* \simeq \pi / a_3$ puts the bare correlation length $\xi_{\perp 0}$ on the order of lattice constant, which is about what one would expect for a generic magnetic transition. The agreement of this data in the $t \sim 0.1$–10 range with the mean field exponent suggests that this region is outside the Ginzburg region, but an experiment may begin to see deviations from mean field behavior within approximately several Kelvin of the critical point. An ultrasound experiment with more precise temperature resolution near the

critical point may be able to resolve a modified cusp exponent $\gamma \simeq 1.31$,[171] since according to one analysis the universality class of a uniaxial modulated one-component OP is that of the O(2), 3D XY transition.[370]

§44. Adding a higher-order interaction

In this section, we compute the B_{1g} modulus for a theory with a higher-order interaction truncation to better match the low-temperature behavior. Consider the free energy density $f = f_{\text{ELASTIC}} + f_{\text{INT}} + f_{\text{OP}}$ with

$$\begin{aligned} f_{\text{ELASTIC}} &= \frac{1}{2} C_0 \epsilon^2 \\ f_{\text{INT}} &= -b\epsilon\eta + \frac{1}{2} g \epsilon^2 \eta^2 \\ f_{\text{OP}} &= \frac{1}{2}\left[r\eta^2 + c_{\parallel}(\nabla_{\parallel}\eta)^2 + c_{\perp}(\nabla_{\perp}\eta)^2 + D(\nabla_{\perp}^2\eta)^2\right] + u\eta^4. \end{aligned} \tag{6.19}$$

The mean-field stain conditioned on the order parameter is found from

$$\begin{aligned} 0 &= \left.\frac{\delta F[\eta,\epsilon]}{\delta \epsilon(x)}\right|_{\epsilon=\epsilon_\star[\eta]} \\ &= C_0 \epsilon_\star[\eta](x) - b\eta(x) + g\epsilon_\star[\eta](x)\eta(x)^2, \end{aligned} \tag{6.20}$$

which yields

$$\epsilon_\star[\eta](x) = \frac{b\eta(x)}{C_0 + g\eta(x)^2}. \tag{6.21}$$

Upon substitution into (6.19) and expanded to fourth order in η, $F[\eta, \epsilon_\star[\eta]]$ can be written in the form $F_{\text{OP}}[\eta]$ alone with $r \to \tilde{r} = r - b^2/C_0$ and $u \to \tilde{u} = u + b^2 g / 2C_0^2$. The phase diagram in η follows as before with the shifted

coefficients, and namely $\langle \eta(x) \rangle = \eta_* \cos(q_* x_3)$ for $\tilde{r} < c_\perp^2/4D = \tilde{r}_c$ with $q_*^2 = -c_\perp/2D$ and

$$\eta_*^2 = \frac{c_\perp^2 - 4D\tilde{r}}{12D\tilde{u}} = \frac{|\Delta\tilde{r}|}{3\tilde{u}}. \tag{6.22}$$

We would like to calculate the q-dependent modulus

$$C(q) = \frac{1}{V}\int dx\, dx'\, C(x,x') e^{-iq(x-x')}, \tag{6.23}$$

where

$$\begin{aligned} C(x,x') &= \left.\frac{\delta^2 F[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')}\right|_{\epsilon=\langle\epsilon\rangle} \\ &= \frac{\delta^2 F_{\text{ELASTIC}}[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')} + \frac{\delta^2 F_{\text{INT}}[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')} + \left.\frac{\delta^2 F_{\text{OP}}[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')}\right|_{\epsilon=\langle\epsilon\rangle} \end{aligned} \tag{6.24}$$

and $\eta_\star$ is the mean-field order parameter conditioned on the strain defined implicitly by

$$0 = \left.\frac{\delta F[\eta,\epsilon]}{\delta\eta(x)}\right|_{\eta=\eta_\star[\epsilon]} = -b\epsilon(x) + g\epsilon(x)^2\eta_\star[\epsilon](x) + \left.\frac{\delta F_{\text{OP}}[\eta]}{\delta\eta(x)}\right|_{\eta=\eta_\star[\epsilon]}. \tag{6.25}$$

We will work this out term by term. The elastic term is the most straightforward, giving

$$\frac{\delta^2 F_{\text{ELASTIC}}[\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')} = \frac{1}{2}C_0 \frac{\delta^2}{\delta\epsilon(x)\delta\epsilon(x')}\int dx''\,\epsilon(x'')^2 = C_0\delta(x-x'). \tag{6.26}$$

The interaction term gives

$$\begin{aligned}
&\frac{\delta^2 F_{\text{INT}}[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')} \\
&\quad= -b\frac{\delta^2}{\delta\epsilon(x)\delta\epsilon(x')}\int dx''\,\epsilon(x'')\eta_\star[\epsilon](x'') + \frac{1}{2}g\frac{\delta^2}{\delta\epsilon(x)\delta\epsilon(x')}\int dx''\,\epsilon(x'')^2\eta_\star[\epsilon](x'')^2 \\
&\quad= -b\frac{\delta\eta_\star[\epsilon](x')}{\delta\epsilon(x)} - b\frac{\delta}{\delta\epsilon(x)}\int dx''\,\epsilon(x'')\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x')} + g\frac{\delta}{\delta\epsilon(x)}\left[\epsilon(x')\eta_\star[\epsilon](x')^2\right] \\
&\qquad\quad + g\frac{\delta}{\delta\epsilon(x)}\int dx''\,\epsilon(x'')^2\eta_\star[\epsilon](x'')\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x')} \\
&\quad= -2(b - 2g\epsilon(x)\eta_\star[\epsilon](x))\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')} - b\int dx''\,\epsilon(x'')\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon(x)\delta\epsilon(x')} + g\eta_\star[\epsilon](x)^2\delta(x-x') \\
&\qquad\quad + g\int dx''\,\epsilon(x'')^2\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x)}\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x')} + g\int dx''\,\epsilon(x'')^2\eta_\star[\epsilon](x'')\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon(x)\delta\epsilon(x')}.
\end{aligned} \tag{6.27}$$

The order parameter term relies on some other identities. First, (6.25) implies

$$\left.\frac{\delta F_{\text{OP}}[\eta]}{\delta\eta(x)}\right|_{\eta=\eta_\star[\epsilon]} = b\epsilon(x) - g\epsilon(x)^2\eta_\star[\epsilon](x), \tag{6.28}$$

and therefore that the functional inverse $\eta_\star^{-1}[\eta]$ is

$$\eta_\star^{-1}[\eta](x) = \frac{b}{2g\eta(x)}\left(1 - \sqrt{1 - \frac{4g\eta(x)}{b^2}\frac{\delta F_{\text{OP}}[\eta]}{\delta\eta(x)}}\right). \tag{6.29}$$

The inverse function theorem further implies (with substitution of (6.28) after the derivative is evaluated) that

$$\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\right)^{\{-1\}} = \left.\frac{\delta\eta_\star^{-1}[\eta](x)}{\delta\eta(x')}\right|_{\eta=\eta_\star[\epsilon]} = \frac{g\epsilon(x)^2\delta(x-x') + \left.\frac{\delta^2 F_{\text{OP}}[\eta]}{\delta\eta(x)\delta\eta(x')}\right|_{\eta=\eta_\star[\epsilon]}}{b - 2g\epsilon(x)\eta_\star[\epsilon](x)}$$

(6.30)

and therefore that

$$\left.\frac{\delta^2 F_{\rm OP}[\eta]}{\delta\eta(x)\delta\eta(x')}\right|_{\eta=\eta_\star[\epsilon]} = (b-2g\epsilon(x)\eta_\star[\epsilon](x))\Big(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\Big)^{\{-1\}} - g\epsilon(x)^2\delta(x-x'). \tag{6.31}$$

Finally, we evaluate the order parameter term, using (6.28) and (6.31) which give

$$\begin{aligned}\frac{\delta^2 F_{\rm OP}[\eta_\star[\epsilon]]}{\delta\epsilon(x)\delta\epsilon(x')} &= \frac{\delta}{\delta\epsilon(x)}\int dx''\,\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x')}\frac{\delta F_{\rm OP}[\eta]}{\delta\eta(x'')}\Big|_{\eta=\eta_\star[\epsilon]} \\ &= \int dx''\,\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon(x)\delta\epsilon(x')}\frac{\delta F_{\rm OP}[\eta]}{\delta\eta(x'')}\Big|_{\eta=\eta_\star[\epsilon]} + \int dx''dx'''\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x)}\frac{\delta\eta_\star[\epsilon](x''')}{\delta\epsilon(x')}\frac{\delta^2 F_{\rm OP}[\eta]}{\delta\eta(x'')\delta\eta(x''')}\Big|_{\eta=\eta_\star[\epsilon]} \\ &= \int dx''\,\frac{\delta^2\eta_\star[\epsilon](x'')}{\delta\epsilon(x)\delta\epsilon(x')}(b\epsilon(x)-g\epsilon(x)^2\eta_\star[\epsilon](x)) + (b-2g\epsilon(x)\eta_\star[\epsilon](x))\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')} \\ &\quad - g\int dx''\,\epsilon(x'')^2\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x)}\frac{\delta\eta_\star[\epsilon](x'')}{\delta\epsilon(x')}.\end{aligned} \tag{6.32}$$

Summing all three terms, we see a great deal of cancellation, with

$$\frac{\delta^2 F[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')} = C_0\delta(x-x')+g\eta_\star[\epsilon](x)^2\delta(x-x')-(b-2g\epsilon(x)\eta_\star[\epsilon](x))\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}. \tag{6.33}$$

We new need to evaluate this at $\langle\epsilon\rangle$. First, $\eta_\star[\langle\epsilon\rangle] = \langle\eta\rangle$, and

$$\left.\frac{\delta^2 F[\eta_\star[\epsilon],\epsilon]}{\delta\epsilon(x)\delta\epsilon(x')}\right|_{\epsilon=\langle\epsilon\rangle} = C_0\delta(x-x')+g\langle\eta(x)\rangle^2\delta(x-x')-(b-2g\langle\epsilon(x)\rangle\langle\eta(x)\rangle)\left.\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\right|_{\epsilon=\langle\epsilon\rangle}.$$

(6.34)

Computing the final functional derivative is the most challenging part. We will first compute its functional inverse, take the Fourier transform of that, and then use the basic relationship between Fourier functional inverses to find the form of the non-inverse. First, we note

$$\left.\frac{\delta^2 F_{\text{OP}}[\eta]}{\delta\eta(x)\delta\eta(x')}\right|_{\eta=\langle\eta\rangle} = \left[r - c_\perp\nabla_\perp^2 - c_\parallel\nabla_\parallel^2 + D\nabla_\perp^4 + 12u\langle\eta(x)\rangle^2\right]\delta(x-x'), \quad (6.35)$$

which gives

$$\begin{aligned}\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\right)^{\{-1\}}\Bigg|_{\epsilon=\langle\epsilon\rangle} &= \frac{1}{b - 2g\langle\epsilon(x)\rangle\langle\eta(x)\rangle}\left[g\langle\epsilon(x)\rangle^2\delta(x-x') + \frac{\delta^2 F_{\text{OP}}[\eta]}{\delta\eta(x)\delta\eta(x')}\right]_{\eta=\langle\eta\rangle} \\ &= \frac{1}{b - 2g\langle\epsilon(x)\rangle\langle\eta(x)\rangle}\left[g\langle\epsilon(x)\rangle^2 + r - c_\perp\nabla_\perp^2 - c_\parallel\nabla_\parallel^2 + D\nabla_\perp^4 + 12u\langle\eta(x)\rangle^2\right]\delta(x-x').\end{aligned} \quad (6.36)$$

Upon substitution of (6.21) and expansion to quadratic order it $\langle\eta(x)\rangle$, we find

$$\begin{aligned}\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\right)^{\{-1\}}\Bigg|_{\epsilon=\langle\epsilon\rangle} = \frac{1}{b}\delta(x-x')\Bigg\{&r - c_\perp\nabla_\perp^2 - c_\parallel\nabla_\parallel^2 + D\nabla_\perp^4 \\ &+ \langle\eta(x)\rangle^2\left[12u + \frac{b^2 g}{C_0^2} + \frac{2g}{C_0}(r - c_\perp\nabla_\perp^2 - c_\parallel\nabla_\parallel^2 + D\nabla_\perp^4)\right] + O(\langle\eta\rangle^4)\Bigg\}\end{aligned} \quad (6.37)$$

Defining $\widehat{\langle\eta\rangle^2} = \int dq'\, \langle\hat{\eta}(q')\rangle\langle\hat{\eta}(-q')\rangle$, its Fourier transform is then

$$\begin{aligned} G(q) &= \frac{1}{V}\int dx\, dx'\, e^{-iq(x-x')}\left(\frac{\delta\eta_\star[\epsilon](x)}{\delta\epsilon(x')}\right)^{\{-1\}}\Big|_{\epsilon=\langle\epsilon\rangle} \\ &= \frac{1}{b}\Bigg\{ r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + D q_\perp^4 + \widehat{\langle\eta\rangle^2}\Big[12u + \frac{b^2 g}{C_0^2} \\ &\qquad + \frac{2g}{C_0}(r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + D q_\perp^4)\Big] + O(\langle\hat{\eta}\rangle^4)\Bigg\}. \end{aligned} \tag{6.38}$$

We can now compute $C(q)$ by taking its Fourier transform, using the convolution theorem for the second term:

$$\begin{aligned} C(q) &= C_0 + g\widehat{\langle\eta\rangle^2} - \int dq''\left(b\delta(q'') - \frac{gb}{C_0}\int dq'\langle\hat{\eta}_{q'}\rangle\langle\hat{\eta}_{q''-q'}\rangle\right)/G(q-q'') \\ &= C_0 + g\widehat{\langle\eta\rangle^2} - b^2\Bigg(\frac{1}{r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + D q_\perp^4} \\ &\qquad - \widehat{\langle\eta\rangle^2}\frac{12u + b^2 g/C_0^2 + \frac{2g}{C_0}(r + c_\perp q^2 + c_\parallel q_\parallel^2 + D q_\perp^4)}{(r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + D q_\perp^4)^2}\Bigg) \\ &\quad + \frac{gb^2}{C_0}\int dq'\, dq''\frac{\langle\hat{\eta}_{q'}\rangle\langle\hat{\eta}_{q''-q'}\rangle}{r + c_\perp(q_\perp - q_\perp'')^2 + c_\parallel(q_\parallel - q_\parallel'')^2 + D(q_\perp - q_\perp'')^4} + O(\langle\hat{\eta}\rangle^4). \end{aligned} \tag{6.39}$$

Upon substitution of $\langle\hat{\eta}_q\rangle = \frac{1}{2}\eta_*[\delta(q_\perp - q_*) + \delta(q_\perp + q_*)]\delta(q_\parallel)$, we have

$$\begin{aligned} C(q) = C_0 + \frac{1}{4}g\eta_*^2 - b^2\Bigg(&\frac{1}{r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + Dq_\perp^4} \\ &- \frac{\eta_*^2}{4}\frac{12u + b^2 g/C_0^2 + \frac{2g}{C_0}(r + c_\perp q^2 + c_\parallel q_\parallel^2 + Dq_\perp^4)}{(r + c_\perp q_\perp^2 + c_\parallel q_\parallel^2 + Dq_\perp^4)^2}\Bigg) \\ + \frac{gb^2\eta_*^2}{4C_0}\Bigg(&\frac{2}{r + c_\parallel q_\parallel^2 + c_\perp q_\perp^2 + Dq_\perp^4} + \frac{1}{r + c_\parallel q_\parallel^2 + c_\perp(q_\perp - 2q_*)^2 + D(q_\perp - 2q_*)^4} \\ &+ \frac{1}{r + c_\parallel q_\parallel^2 + c_\perp(q_\perp + 2q_*)^2 + D(q_\perp + 2q_*)^4}\Bigg) + O(\eta_*^4). \end{aligned} \tag{6.40}$$

Evaluating at $q = 0$, we have

$$C(0) = C_0 - \frac{b^2}{r} + \frac{\eta_*^2}{4}\left(g + \frac{b^2}{r^2}(12u + b^2 g/C_0^2) + \frac{2gb^2}{C_0 r}\frac{16Dq_*^4 + 3r}{8Dq_*^4 + r}\right) \tag{6.41}$$

Above the transition this has exactly the form of (6.18) for any g; below the transition it has the same form at $g = 0$ to order η_*^2. With $r = a\Delta T + c^2/4D + b^2/C_0$, $u = \tilde{u} - b^2 g/2C_0^2$, and

$$\eta_*^2 = \begin{cases} 0 & \Delta T > 0 \\ -a\Delta T/3\tilde{u} & \Delta T \le 0, \end{cases} \tag{6.42}$$

we can fit the ratios $b^2/a = 1665\,\mathrm{GPa\,K}$, $b^2/Dq_*^4 = 6.28\,\mathrm{GPa}$, and $b\sqrt{-g/\tilde{u}} = 14.58\,\mathrm{GPa}$ with $C_0 = (71.14 - (0.010426 \times T)/\mathrm{K})\,\mathrm{GPa}$. The resulting fit the dashed black line in Fig. 6.2.

§45. Conclusion & outlook

We have developed a general phenomenological treatment of HO OPs that have the potential for linear coupling to strain. The two representations with mean field phase diagrams that are consistent with the phase diagram of URu_2Si_2 are B_{1g} and B_{2g}. Of these, only a staggered B_{1g} OP is consistent with zero-pressure RUS data, with a cusp appearing in the associated elastic modulus. In this picture, the HO phase is characterized by uniaxial modulated B_{1g} order, while the high pressure phase is characterized by uniform B_{1g} order. The staggered nematic of HO is similar to the striped superconducting phase found in LBCO and other cuperates.[375]

The coincidence of our theory's orthorhombic high-pressure phase and URu_2Si_2's AFM is compelling, but our mean field theory does not make any explicit connection with the physics of AFM. Neglecting this physics could be reasonable since correlations often lead to AFM as a secondary effect, like what occurs in many Mott insulators. An electronic theory of this phase diagram may find that the AFM observed in URu_2Si_2 indeed follows along with an independent high-pressure orthorhombic phase associated with uniform B_{1g} electronic order.

The corresponding prediction of uniform B_{1g} symmetry breaking in the high pressure phase is consistent with recent diffraction experiments,[350] except for the apparent earlier onset in temperature of the B_{1g} symmetry breaking, which we believe may be due to fluctuating order at temperatures above the actual transition temperature. This work motivates both further theoretical work regarding a microscopic theory with modulated B_{1g} order, and

preforming symmetry-sensitive thermodynamic experiments at pressure, such as ultrasound, that could further support or falsify this idea.

BIBLIOGRAPHY

[1]O Tange, "GNU parallel: the command-line power tool", ;login: The USENIX Magazine **36**, 42–47 (2011).

Chapter 1: Quasibrittle fracture

[2]JL Chaboche, "Continuous damage mechanics — a tool to describe phenomena before crack initiation", Nuclear Engineering and Design **64**, 233–247 (1981).

[3]JL Chaboche, "Continuum damage mechanics: present state and future trends", Nuclear Engineering and Design **105**, 19–33 (1987).

[4]GC Sih, *Fracture mechanics of concrete: Structural application and numerical calculation* (Springer Netherlands, Dordrecht, 1984).

[5]P Acker, P Rossi, and JM Torrenti, "Cracking and fracture of concrete: size effect or scaling law?", in *Disorder and fracture* (Springer US, 1990), pp. 269–278.

[6]HJ Herrmann and L de Arcangelis, "Scaling in fracture", in *Disorder and fracture* (Springer US, 1990), pp. 149–163.

[7]DS Fisher, "Collective transport in random media: From superconductors to earthquakes", Physics Reports **301**, 113–150 (1998).

[8]MJ Alava, PKVV Nukala, and S Zapperi, "Fracture size effects from disordered lattice models", International Journal of Fracture **154**, 51–59 (2008).

[9]AB Harris and TC Lubensky, "Randomly diluted xy and resistor networks near the percolation threshold", Physical Review B **35**, 6964–6986 (1987).

[10]MJ Alava, PKVV Nukala, and S Zapperi, "Statistical models of fracture", Advances in Physics **55**, 349–476 (2006).

[11]A Buchel and JP Sethna, "Statistical mechanics of cracks: Fluctuations, breakdown, and asymptotics of elastic theory", Physical Review E **55**, 7669–7690 (1997).

[12]A Hansen, EL Hinrichsen, and S Roux, "Scale-invariant disorder in fracture and related breakdown phenomena", Physical Review B **43**, 665–678 (1991).

[13]MJ Alava, PKVV Nukala, and S Zapperi, "Role of disorder in the size scaling of material strength", Physical Review Letters **100**, 10.1103/physrevlett.100.055502 (2008).

[14]S Lherminier, R Planet, V Levy dit Vehel, G Simon, L Vanel, KJ Måloy, and O Ramos, "Continuously sheared granular matter reproduces in detail seismicity laws", Physical Review Letters **122**, 10.1103/physrevlett.122.218501 (2019).

[15]C Manzato, A Shekhawat, PKVV Nukala, MJ Alava, JP Sethna, and S Zapperi, "Fracture strength of disordered media: universality, interactions, and tail asymptotics", Physical Review Letters **108**, 10.1103/physrevlett.108.065504 (2012).

[16]Z Bertalan, A Shekhawat, JP Sethna, and S Zapperi, "Fracture strength: stress concentration, extreme value statistics, and the fate of the Weibull

distribution", Physical Review Applied **2**, 10.1103/physrevapplied.2.034008 (2014).

[17]A Shekhawat, S Zapperi, and JP Sethna, "From damage percolation to crack nucleation through finite size criticality", Physical Review Letters **110**, 10.1103/physrevlett.110.185505 (2013).

[18]B Kahng, GG Batrouni, S Redner, L de Arcangelis, and HJ Herrmann, "Electrical breakdown in a fuse network with random, continuously distributed breaking strengths", Physical Review B **37**, 7625–7637 (1988).

[19]PKVV Nukala, S Šimunović, and S Zapperi, "Percolation and localization in the random fuse model", Journal of Statistical Mechanics: Theory and Experiment **2004**, P08001 (2004).

[20]M Sahimi and S Arbabi, "Percolation and fracture in disordered solids and granular media: approach to a fixed point", Physical Review Letters **68**, 608–611 (1992).

[21]A Hansen and J Schmittbuhl, "Origin of the universal roughness exponent of brittle fracture surfaces: stress-weighted percolation in the damage zone", Physical Review Letters **90**, 10.1103/physrevlett.90.045504 (2003).

[22]PKVV Nukala, S Zapperi, and S Šimunović, "Statistical properties of fracture in a random spring model", Physical Review E **71**, 10.1103/physreve.71.066106 (2005).

[23]MEJ Newman, *Networks: An introduction*, 1st ed. (Oxford University Press, Oxford, 2010).

[24]F Chung, *Spectral graph theory* (Published for the Conference Board of the mathematical sciences by the American Mathematical Society, Providence, R.I, 1997).

[25]L de Arcangelis, S Redner, and HJ Herrmann, "A random fuse model for breaking processes", Journal de Physique Lettres **46**, 585–590 (1985).

[26]PM Duxbury, PL Leath, and PD Beale, "Breakdown properties of quenched random systems: The random-fuse network", Physical Review B **36**, 367–380 (1987).

[27]A Gilabert, C Vanneste, D Sornette, and E Guyon, "The random fuse network as a model of rupture in a disordered medium", Journal de Physique **48**, 763–770 (1987).

[28]M Seeger, "Low rank updates for the Cholesky decomposition", (2004).

[29]Y Chen, TA Davis, WW Hager, and S Rajamanickam, "Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate", ACM Transactions on Mathematical Software **35**, 1–14 (2008).

[30]HJ Herrmann, DC Hong, and HE Stanley, "Backbone and elastic backbone of percolation clusters obtained by the new method of 'burning'", Journal of Physics A: Mathematical and General **17**, L261–L266 (1984).

[31]S Roux and A Hansen, "A new algorithm to extract the backbone in a random resistor network", Journal of Physics A: Mathematical and General **20**, L1281–L1285 (1987).

[32]D Eppstein, "Subgraph isomorphism in planar graphs and related problems", Journal of Graph Algorithms and Applications **3**, 1–27 (1999).

[33]JT Chayes, L Chayes, JR Franz, JP Sethna, and SA Trugman, "On the density of state for the quantum percolation problem", Journal of Physics A: Mathematical and General **19**, L1173–L1177 (1986).

[34]AA Middleton, "Asymptotic uniqueness of the sliding state for charge-density waves", Physical Review Letters **68**, 670–673 (1992).

[35]AA Middleton and DS Fisher, "Critical behavior of charge-density waves below threshold: numerical and scaling analysis", Physical Review B **47**, 3530–3552 (1993).

[36]JP Sethna, K Dahmen, S Kartha, JA Krumhansl, BW Roberts, and JD Shore, "Hysteresis and hierarchies: Dynamics of disorder-driven first-order phase transformations", Physical Review Letters **70**, 3347–3350 (1993).

[37]S Zapperi, P Ray, HE Stanley, and A Vespignani, "First-order transition in the breakdown of disordered media", Physical Review Letters **78**, 1408–1411 (1997).

[38]S Zapperi, P Ray, HE Stanley, and A Vespignani, "Analysis of damage clusters in fracture processes", Physica A: Statistical Mechanics and its Applications **270**, 57–62 (1999).

[39]S Roux, A Hansen, H Herrmann, and E Guyon, "Rupture of heterogeneous media in the limit of infinite disorder", Journal of Statistical Physics **52**, 237–244 (1988).

[40]D Stauffer and A Aharony, *Introduction to percolation theory*, 2nd (Taylor & Francis Group, Boca Raton, 1991).

[41]D Stauffer, "Scaling theory of percolation clusters", Physics Reports **54**, 1–74 (1979).

[42]RM Ziff, "Correction-to-scaling exponent for two-dimensional percolation", Physical Review E **83**, 10.1103/physreve.83.020107 (2011).

[43]X Xu, J Wang, Z Zhou, TM Garoni, and Y Deng, "Geometric structure of percolation clusters", Physical Review E **89**, 10.1103/physreve.89.012120 (2014).

[44]Z Zhou, J Yang, Y Deng, and RM Ziff, "Shortest-path fractal dimension for percolation in two and three dimensions", Physical Review E **86**, 10.1103/physreve.86.061101 (2012).

[45]AA Moreira, CLN Oliveira, A Hansen, NAM Araújo, HJ Herrmann, and JS Andrade, "Fracturing highly disordered materials", Physical Review Letters **109**, 10.1103/physrevlett.109.255701 (2012).

[46]R Toussaint and SR Pride, "Interacting damage models mapped onto ising and percolation models", Physical Review E **71**, 10.1103/physreve.71.046127 (2005).

[47]A Weinrib, "Long-range correlated percolation", Physical Review B **29**, 387–395 (1984).

[48]S Prakash, S Havlin, M Schwartz, and HE Stanley, "Structural and dynamical properties of long-range correlated percolation", Physical Review A **46**, R1724–R1727 (1992).

[49]M Sahimi and S Mukhopadhyay, "Scaling properties of a percolation model with long-range correlations", Physical Review E **54**, 3870–3880 (1996).

[50]B Nienhuis and M Nauenberg, "First-order phase transitions in renormalization-group theory", Physical Review Letters **35**, 477–479 (1975).

[51]ME Fisher and AN Berker, "Scaling for first-order phase transitions in thermodynamic and finite systems", Physical Review B **26**, 2507–2513 (1982).

[52]GF Mazenko, OT Valls, and FC Zhang, "Kinetics of first-order phase transitions: Monte Carlo simulations, renormalization-group methods, and scaling for critical quenches", Physical Review B **31**, 4453–4464 (1985).

[53]AM Becker and RM Ziff, "Percolation thresholds on two-dimensional voronoi networks and delaunay triangulations", Physical Review E **80**, 10.1103/physreve.80.041101 (2009).

[54]L Girard, J Weiss, and D Amitrano, "Damage-cluster distributions and size effect on strength in compressive failure", Physical Review Letters **108**, 10.1103/physrevlett.108.225502 (2012).

[55]MEJ Newman and RM Ziff, "Fast Monte Carlo algorithm for site or bond percolation", Physical Review E **64**, 10.1103/physreve.64.016706 (2001).

[56]S Zapperi, P Ray, HE Stanley, and A Vespignani, "Avalanches in breakdown and fracture processes", Physical Review E **59**, 5049–5057 (1999).

[57]MM Driscoll, BGg Chen, TH Beuman, S Ulrich, SR Nagel, and V Vitelli, "The role of rigidity in controlling material failure", Proceedings of the National Academy of Sciences **113**, 10813–10817 (2016).

[58]L de Arcangelis, S Redner, and A Coniglio, "Anomalous voltage distribution of random resistor networks and a new model for the backbone at the percolation threshold", Physical Review B **31**, 4725–4727 (1985).

[59]B Fourcade and AMS Tremblay, "Universal multifractal properties of circle maps from the point of view of critical phenomena I: Phenomenology", Journal of Statistical Physics **61**, 607–637 (1990).

[60]B Fourcade and AMS Tremblay, "Universal multifractal properties of circle maps from the point of view of critical phenomena II: Analytical results", Journal of Statistical Physics **61**, 639–665 (1990).

[61]S Roux and A Hansen, "Introduction to multifractality", in *Disorder and fracture* (Springer US, 1990), pp. 17–30.

[62]R Rammal, MA Lemieux, and AMS Tremblay, "Comment on "ϵ-expansion for the conductivity of a random resistor network"", Physical Review Letters **54**, 1087–1087 (1985).

[63]B Fourcade and AMS Tremblay, "Anomalies in the multifractal analysis of self-similar resistor networks", Physical Review A **36**, 2352–2358 (1987).

[64]B Fourcade and AMS Tremblay, "Amplitudes of multifractal moments at the onset of chaos: Universal ratios and crossover functions", Physical Review Letters **64**, 2659–2662 (1990).

[65]U Frisch and M Vergassola, "A prediction of the multifractal model: the intermediate dissipation range", Europhysics Letters (EPL) **14**, 439–444 (1991).

[66]R Rammal, C Tannous, and AMS Tremblay, "$1/f$ noise in random resistor networks: fractals and percolating systems", Physical Review A **31**, 2662–2671 (1985).

[67]AMS Tremblay, RR Tremblay, G Albinet, and B Fourcade, "How many correlation lengths for multifractals?", Physica A: Statistical Mechanics and its Applications **183**, 398–410 (1992).

[68]LJ Pyrak-Nolte, LR Myer, and DD Nolte, "Fractures: finite-size scaling and multifractals", in *Fractals and chaos in the earth sciences* (Birkhäuser Basel, 1992), pp. 679–706.

[69]HJ Herrmann, A Hansen, and S Roux, "Fracture of disordered, elastic lattices in two dimensions", Physical Review B **39**, 637–648 (1989).

[70]S Redner, "Fractal and multifractal scaling of electrical conduction in random resistor networks", (2007), arXiv:0710.1105v1 [cond-mat.stat-mech].

[71]R Rammal, C Tannous, P Breton, and AMS Tremblay, "Flicker ($1/f$) noise in percolation networks: a new hierarchy of exponents", Physical Review Letters **54**, 1718–1721 (1985).

[72]AB Harris, "Field-theoretic formulation of the randomly diluted nonlinear resistor network", Physical Review B **35**, 5056–5065 (1987).

[73]B Fourcade and AMS Tremblay, "Field theory and second renormalization group for multifractals in percolation", Physical Review E **51**, 4095–4104 (1995).

[74]L de Arcangelis, S Redner, and A Coniglio, "Multiscaling approach in random resistor and random superconducting networks", Physical Review B **34**, 4656–4673 (1986).

[75]MH Jensen, G Paladin, and A Vulpiani, "Multiscaling in multifractals", Physical Review Letters **67**, 208–211 (1991).

[76]B Fourcade, P Breton, and AMS Tremblay, "Multifractals and critical phenomena in percolating networks: Fixed point, gap scaling, and universality", Physical Review B **36**, 8925–8928 (1987).

[77]A Carpinteri and B Chiaia, "Multifractal scaling laws in the breaking behaviour of disordered materials", Chaos, Solitons & Fractals **8**, 135–150 (1997).

[78]L de Arcangelis, A Hansen, HJ Herrmann, and S Roux, "Scaling laws in fracture", Physical Review B **40**, 877–880 (1989).

[79]G Albinet, RR Tremblay, and AMS Tremblay, "Scaling behavior of multifractal-moment distributions near criticality", Journal de Physique I **3**, 323–330 (1993).

[80]L de Arcangelis and HJ Herrmann, "Scaling and multiscaling laws in random fuse networks", Physical Review B **39**, 2678–2684 (1989).

[81]AMS Tremblay and B Fourcade, "Observable infinite sets of exponents in multifractals and in critical phenomena: the role of symmetry", in Xviith international colloquium on group theoretical methods in physics, Vol. World Scientific, Singapore, edited by Y Saint-Aubin and L Vinet (1989), pp. 137–153.

[82]A Ebrahimkhanlou, A Farhidzadeh, and S Salamone, "Multifractal analysis of crack patterns in reinforced concrete shear walls", Structural Health Monitoring: An International Journal **15**, 81–92 (2016).

[83]SV Buldyrev, S Havlin, and HE Stanley, "Anisotropic percolation and the d-dimensional surface roughening problem", Physica A: Statistical Mechanics and its Applications **200**, 200–211 (1993).

[84]D Ertaş and M Kardar, "Critical dynamics of contact line depinning", Physical Review E **49**, R2532–R2535 (1994).

[85]D Ertaş and M Kardar, "Anisotropic scaling in depinning of a flux line", Physical Review Letters **73**, 1703–1706 (1994).

[86]E Bouchaud, G Lapasset, and J Planès, "Fractal dimension of fractured surfaces: a universal value?", Europhysics Letters (EPL) **13**, 73–79 (1990).

[87]JP Bouchaud, E Bouchaud, G Lapasset, and J Planès, "Models of fractal cracks", Physical Review Letters **71**, 2240–2243 (1993).

[88]L Ponson, D Bonamy, H Auradou, G Mourot, S Morel, E Bouchaud, C Guillot, and JP Hulin, "Anisotropic self-affine properties of experimental fracture surfaces", International Journal of Fracture **140**, 27–37 (2006).

[89]YJ Chen, S Zapperi, and JP Sethna, "Crossover behavior in interface depinning", Physical Review E **92**, 10.1103/physreve.92.022146 (2015).

[90]K Sneppen, "Self-organized pinning and interface growth in a random medium", Physical Review Letters **69**, 3539–3542 (1992).

[91]C Poirier, M Ammi, D Bideau, and JP Troadec, "Experimental study of the geometrical effects in the localization of deformation", Physical Review Letters **68**, 216–219 (1992).

[92]S Ramanathan, D Ertaş, and DS Fisher, "Quasistatic crack propagation in heterogeneous media", Physical Review Letters **79**, 873–876 (1997).

[93]E Bouchaud, "Scaling properties of cracks", Journal of Physics: Condensed Matter **9**, 4319–4344 (1997).

[94]B Skjetne, T Helle, and A Hansen, "Roughness of crack interfaces in two-dimensional beam lattices", Physical Review Letters **87**, 10.1103/physrevlett.87.125503 (2001).

[95]A Hansen, EL Hinrichsen, and S Roux, "Roughness of crack interfaces", Physical Review Letters **66**, 2476–2479 (1991).

[96]J Kertész, VK Horváth, and F Weber, "Self-affine rupture lines in paper sheets", Fractals **01**, 67–74 (1993).

[97]MA Muñoz, R Dickman, A Vespignani, and S Zapperi, "Avalanche and spreading exponents in systems with absorbing states", Physical Review E **59**, 6175–6179 (1999).

[98]LI Salminen, AI Tolvanen, and MJ Alava, "Acoustic emission from paper fracture", Physical Review Letters **89**, 10.1103/physrevlett.89.185503 (2002).

[99]D Sornette, "Elasticity and failure of a set of elements loaded in parallel", Journal of Physics A: Mathematical and General **22**, L243–L250 (1989).

[100]D Sornette, "Mean-field solution of a block-spring model of earthquakes", Journal de Physique I **2**, 2089–2096 (1992).

[101]HE Daniels, "The statistical theory of the strength of bundles of threads I", Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **183**, 405–435 (1945).

[102]PC Hemmer and A Hansen, "The distribution of simultaneous fiber failures in fiber bundles", Journal of Applied Mechanics **59**, 909–914 (1992).

[103]SL Phoenix and HM Taylor, "The asymptotic strength distribution of a general fiber bundle", Advances in Applied Probability **5**, 200–216 (1973).

[104]JP Sethna, KA Dahmen, and CR Myers, "Crackling noise", Nature **410**, 242–250 (2001).

[105]S Papanikolaou, F Bohn, RL Sommer, G Durin, S Zapperi, and JP Sethna, "Universality beyond power laws and the average avalanche shape", Nature Physics **7**, 316–320 (2011).

[106]YJ Chen, S Papanikolaou, JP Sethna, S Zapperi, and G Durin, "Avalanche spatial structure and multivariable scaling functions: Sizes, heights, widths, and views through windows", Physical Review E **84**, 10.1103/physreve.84.061103 (2011).

[107]DR Nelson, "Crossover scaling functions and renormalization-group trajectory integrals", Physical Review B **11**, 3504–3519 (1975).

[108]D O'Connor and CR Stephens, "Crossover scaling: a renormalization group approach", Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **444**, 287–296 (1994).

[109]E Frey, UC Täuber, and F Schwabl, "Crossover from self-similar to self-affine structures in percolation", Europhysics Letters (EPL) **26**, 413–418 (1994).

[110]E Frey, UC Täuber, and F Schwabl, "Crossover from isotropic to directed percolation", Physical Review E **49**, 5058–5072 (1994).

[111]M Porto, N Schwartz, S Havlin, and A Bunde, "Optimal paths in disordered media: scaling of the crossover from self-similar to self-affine behavior", Physical Review E **60**, R2448–R2451 (1999).

[112]LAN Amaral, AL Barabási, SV Buldyrev, ST Harrington, S Havlin, R Sadr-Lahijany, and HE Stanley, "Avalanches and the directed percolation depinning model: experiments, simulations, and theory", Physical Review E **51**, 4655–4673 (1995).

[113]JS Wang, "Anisotropic finite-size scaling analysis of a two-dimensional driven diffusive system", Journal of Statistical Physics **82**, 1409–1427 (1996).

[114]Kt Leung, "Finite-size scaling of driven diffusive systems: theory and Monte Carlo studies", Physical Review Letters **66**, 453–456 (1991).

[115]S Ramanathan and DS Fisher, "Dynamics and instabilities of planar tensile cracks in heterogeneous media", Physical Review Letters **79**, 877–880 (1997).

Chapter 2: Lattice Monte Carlo

[41] D Stauffer, "Scaling theory of percolation clusters", Physics Reports **54**, 1–74 (1979).

[116] J Kent-Dobias and JP Sethna, "Cluster representations and the Wolff algorithm in arbitrary external fields", Physical Review E **98**, 10.1103/physreve.98.063306 (2018).

[117] U Wolff, "Critical slowing down", Nuclear Physics B: Proceedings Supplements **17**, 93–102 (1990).

[118] PC Hohenberg and BI Halperin, "Theory of dynamic critical phenomena", Reviews of Modern Physics **49**, 435–479 (1977).

[119] CW Liu, A Polkovnikov, and AW Sandvik, "Dynamic scaling at classical phase transitions approached through nonequilibrium quenching", Physical Review B **89**, 10.1103/physrevb.89.054307 (2014).

[120] RH Swendsen and JS Wang, "Nonuniversal critical dynamics in Monte Carlo simulations", Physical Review Letters **58**, 86–88 (1987).

[121] U Wolff, "Collective Monte Carlo updating for spin systems", Physical Review Letters **62**, 361–364 (1989).

[122] W Janke, "Nonlocal Monte Carlo algorithms for statistical physics applications", Mathematics and Computers in Simulation **47**, 329–346 (1998).

[123] A Coniglio and W Klein, "Clusters and Ising critical droplets: A renormalisation group approach", Journal of Physics A: Mathematical and General **13**, 2775–2780 (1980).

[124]U Wolff, "Comparison between cluster Monte Carlo algorithms in the Ising model", Physics Letters B **228**, 379–382 (1989).

[125]J Du, B Zheng, and JS Wang, "Dynamic critical exponents for Swendsen–Wang and Wolff algorithms obtained by a nonequilibrium relaxation method", Journal of Statistical Mechanics: Theory and Experiment **2006**, P05004–P05004 (2006).

[126]JS Wang and RH Swendsen, "Cluster monte carlo algorithms", Physica A: Statistical Mechanics and its Applications **167**, 565–579 (1990).

[127]CF Baillie and PD Coddington, "Comparison of cluster algorithms for two-dimensional Potts models", Physical Review B **43**, 10617–10621 (1991).

[128]U Wolff, "High precision simulation techniques for lattice field theory", International Journal of Modern Physics C **04**, 451–458 (1993).

[129]BB Beard, M Pepe, S Riederer, and UJ Wiese, "Efficient cluster algorithm for $CP(N-1)$ models", Computer Physics Communications **175**, 629–634 (2006).

[130]RG Edwards and AD Sokal, "Dynamic critical behavior of Wolff's collective-mode Monte Carlo algorithm for the two-dimensional $O(n)$ nonlinear σ model", Physical Review D **40**, 1374–1377 (1989).

[131]S Caracciolo, RG Edwards, A Pelissetto, and AD Sokal, "Dynamic critical behaviour of Wolff's algorithm for RP^N σ-models", Nuclear Physics B - Proceedings Supplements **26**, 595–597 (1992).

[132]A Sokal, "Monte Carlo methods in statistical mechanics: foundations and new algorithms", in *Functional integration* (Springer US, 1997), pp. 131–192.

[133]A Coniglio, F de Liberto, G Monroy, and F Peruggi, "Exact relations between droplets and thermal fluctuations in external field", Journal of Physics A: Mathematical and General **22**, L837–L842 (1989).

[134]Z Alexandrowicz, "Swendsen-Wang simulation of Ising spins and a precise definition of critical clusters", Physica A: Statistical Mechanics and its Applications **160**, 310–320 (1989).

[135]JS Wang, "Clusters in the three-dimensional Ising model with a magnetic field", Physica A: Statistical Mechanics and its Applications **161**, 249–268 (1989).

[136]TS Ray and JS Wang, "Metastability and nucleation in Ising models with Swendsen-Wang dynamics", Physica A: Statistical Mechanics and its Applications **167**, 580–588 (1990).

[137]C Destri, FD Renzo, E Onofri, P Rossi, and GP Tecchiolli, "Swendsen-Wang Monte Carlo study of the Ising model with external field", Physics Letters B **278**, 311–316 (1992).

[138]PG Lauwers and V Rittenberg, "The critical 2d Ising model in a magnetic field: A Monte Carlo study using a Swendesen-Wang algorithm", Physics Letters B **233**, 197–200 (1989).

[139]T Ala-Nissila, E Granato, K Kankaala, JM Kosterlitz, and SC Ying, "Numerical studies of the two-dimensional XY model with symmetry-breaking fields", Physical Review B **50**, 12692–12701 (1994).

[140]O Redner, J Machta, and LF Chayes, "Graphical representations and cluster algorithms for critical points with fields", Physical Review E **58**, 2749–2752 (1998).

[141]L Chayes, J Machta, and O Redner, "Graphical representations for Ising systems in external fields", Journal of Statistical Physics **93**, 17–32 (1998).

[142]J Machta, MEJ Newman, and LB Chayes, "Replica-exchange algorithm and results for the three-dimensional random field Ising model", Physical Review E **62**, 8782–8789 (2000).

[143]V Martin-Mayor and D Yllanes, "Cluster Monte Carlo algorithm with a conserved order parameter", Physical Review E **80**, 10.1103/physreve.80.015701 (2009).

[144]V Martin-Mayor, B Seoane, and D Yllanes, "Tethered Monte Carlo: managing rugged free-energy landscapes with a Helmholtz-potential formalism", Journal of Statistical Physics **144**, 554–596 (2011).

[145]HG Evertz, HM Erkinger, and W von der Linden, "New cluster method for the Ising model", in *Springer proceedings in physics* (Springer Berlin Heidelberg, 2002), pp. 123–133.

[146]CM Fortuin and PW Kasteleyn, "On the random-cluster model I: Introduction and relation to other models", Physica **57**, 536–564 (1972).

[147]CM Fortuin, "On the random-cluster model II: The percolation model", Physica **58**, 393–418 (1972).

[148]CM Fortuin, "On the random-cluster model III: The simple random-cluster model", Physica **59**, 545–570 (1972).

[149]RG Edwards and AD Sokal, "Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm", Physical Review D **38**, 2009–2012 (1988).

[150]S Caracciolo, RG Edwards, A Pelissetto, and AD Sokal, "Wolff-type embedding algorithms for general nonlinear σ-models", Nuclear Physics B **403**, 475–541 (1993).

[151]O Loos, *Symmetric spaces I: General theory* (1969).

[152]H Rieger, "Monte Carlo studies of Ising spin glasses and random field systems", in *Annual reviews of computational physics II* (World Scientific, Mar. 1995), pp. 295–341.

[153]D Kandel, R Ben-Av, and E Domany, "Cluster monte carlo dynamics for the fully frustrated ising model", Physical Review B **45**, 4700–4709 (1992).

[154]JV José, LP Kadanoff, S Kirkpatrick, and DR Nelson, "Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model", Physical Review B **16**, 1217–1241 (1977).

[155]S Caracciolo, RG Edwards, A Pelissetto, and AD Sokal, "Generalized Wolff-type embedding algorithms for nonlinear σ-models", Nuclear Physics B - Proceedings Supplements **20**, 72–75 (1991).

[156]U Wolff, "Lattice field theory as a percolation process", Physical Review Letters **60**, 1461–1463 (1988).

[157]M Hasenbusch, "Improved estimators for a cluster updating of O(n) spin models", Nuclear Physics B **333**, 581–592 (1990).

[158]I Dimitrović, P Hasenfratz, J Nager, and F Niedermayer, "Finite-size effects, Goldstone bosons and critical exponents in the $d = 3$ Heisenberg model", Nuclear Physics B **350**, 893–905 (1991).

[159]AD Bruce and A Aharony, "Coupled order parameters, symmetry-breaking irrelevant scaling fields, and tetracritical points", Physical Review B **11**, 478–499 (1975).

[160]D Blankschtein and D Mukamel, "Fluctuation-induced first-order transitions and symmetry-breaking fields: The $n = 3$-component cubic model", Physical Review B **25**, 6939–6951 (1982).

[161]JP Straley and ME Fisher, "Three-state potts model and anomalous tricritical points", Journal of Physics A: Mathematical, Nuclear and General **6**, 1310–1326 (1973).

[162]MC Barbosa and WK Theumann, "Equations of state for the three-state potts model with symmetry-breaking perturbations", Journal of Physics: Condensed Matter **1**, 6059–6070 (1989).

[163]HG Evertz, M Hasenbusch, M Marcu, K Pinn, and S Solomon, "Stochastic cluster algorithms for discrete gaussian (SOS) models", Physics Letters B **254**, 185–191 (1991).

[164]HG Evertz, M Hasenbusch, M Marcu, and K Pinn, "The solid-on-solid surface width around the roughening transition", Physica A: Statistical Mechanics and its Applications **199**, 31–39 (1993).

[165]M Hasenbusch, M Marcu, and K Pinn, "High precision renormalization group study of the roughening transition", Physica A: Statistical Mechanics and its Applications **208**, 124–161 (1994).

[166]VS Dotsenko, W Selke, and AL Talapov, "Cluster Monte Carlo algorithms for random Ising models", Physica A: Statistical Mechanics and its Applications **170**, 278–281 (1991).

[167]G Ossola and AD Sokal, "Dynamic critical behavior of the Swendsen–Wang algorithm for the three-dimensional Ising model", Nuclear Physics B **691**, 259–291 (2004).

[168]AB Bortz, MH Kalos, and JL Lebowitz, "A new algorithm for Monte Carlo simulation of Ising spin systems", Journal of Computational Physics **17**, 10–18 (1975).

[169]FY Wu, "The Potts model", Reviews of Modern Physics **54**, 235–268 (1982).

[170]S El-Showk, MF Paulos, D Poland, S Rychkov, D Simmons-Duffin, and A Vichi, "Solving the 3d Ising model with the conformal bootstrap II: *c*-minimization and precise critical exponents", Journal of Statistical Physics **157**, 869–914 (2014).

[171]R Guida and J Zinn-Justin, "Critical exponents of the N-vector model", Journal of Physics A: Mathematical and General **31**, 8103–8121 (1998).

[172]K Kankaala, T Ala-Nissila, and SC Ying, "Theory of adsorbate-induced surface reconstruction on W(100)", Physical Review B **47**, 2333–2343 (1993).

[173]SB Dierker, R Pindak, and RB Meyer, "Consequences of bond-orientational order on the macroscopic orientation patterns of thin tilted hexatic liquid-crystal films", Physical Review Letters **56**, 1819–1822 (1986).

[174]JV Selinger and DR Nelson, "Theory of hexatic-to-hexatic transitions", Physical Review Letters **61**, 416–419 (1988).

[175]QM Zhang and JZ Larese, "Melting of monolayer argon adsorbed on a graphite substrate", Physical Review B **43**, 938–946 (1991).

[176]R Peierls, "On Ising's model of ferromagnetism", Mathematical Proceedings of the Cambridge Philosophical Society **32**, 477–481 (1936).

[177]HG Evertz and W von der Linden, "Simulations on infinite-size lattices", Physical Review Letters **86**, 5164–5167 (2001).

[178]CL Henley, "Defect concepts for vector spin glasses", Annals of Physics **156**, 368–411 (1984).

[179]K Binder and DP Landau, "Finite-size scaling at first-order phase transitions", Physical Review B **30**, 1477–1485 (1984).

[180]K Binder, "Theory of first-order phase transitions", Reports on Progress in Physics **50**, 783–859 (1987).

[181]E Stoll and T Schneider, "Computer simulation of critical properties and metastable states in a finite square ising system", Physical Review A **6**, 429–432 (1972).

[182]K Binder and H Müller-Krumbhaar, "Investigation of metastable states and nucleation in the kinetic ising model", Physical Review B **9**, 2328–2353 (1974).

[183]K Binder, "Monte carlo computer experiments on critical phenomena and metastable states", Advances in Physics **23**, 917–939 (1974).

[184]I Shteto, J Linares, and F Varret, "Monte carlo entropic sampling for the study of metastable states and relaxation paths", Physical Review E **56**, 5128–5137 (1997).

[185]JV Selinger, ZG Wang, RF Bruinsma, and CM Knobler, "Chiral symmetry breaking in langmuir monolayers and smectic films", Physical Review Letters **70**, 1139–1142 (1993).

[186] J Pang and NA Clark, "Observation of a chiral-symmetry-breaking twist-bend instability in achiral freely suspended liquid-crystal films", Physical Review Letters **73**, 2332–2335 (1994).

[187] VV Mangazeev, MT Batchelor, VV Bazhanov, and MY Dudalev, "Variational approach to the scaling function of the 2d Ising model in a magnetic field", Journal of Physics A: Mathematical and Theoretical **42**, 042005 (2008).

[188] JM Carmona, A Pelissetto, and E Vicari, "N-component Ginzburg-Landau Hamiltonian with cubic anisotropy: A six-loop study", Physical Review B **61**, 15136–15151 (2000).

Chapter 3: Geometric Monte Carlo

[189] J Cardy, *Scaling and renormalization in statistical physics* (Cambridge University Press, Cambridge New York, 1996).

[190] JR Heringa and HWJ Blöte, "The simple-cubic lattice gas with nearest-neighbour exclusion: ising universality", Physica A: Statistical Mechanics and its Applications **232**, 369–374 (1996).

[191] JR Heringa and HWJ Blöte, "Cluster dynamics and universality of Ising lattice gases", Physica A: Statistical Mechanics and its Applications **251**, 224–234 (1998).

[192] JR Heringa and HWJ Blöte, "Geometric cluster monte carlo simulation", Physical Review E **57**, 4976–4978 (1998).

[193] HWJ Blöte, JR Heringa, and MM Tsypin, "Three-dimensional Ising model in the fixed-magnetization ensemble: a Monte Carlo study", Physical Review E **62**, 77–82 (2000).

[194]HWJ Blöte, JR Heringa, and E Luijten, "Cluster Monte Carlo: extending the range", Computer Physics Communications **147**, 58–63 (2002).

[195]C Dress and W Krauth, "Cluster algorithm for hard spheres and related systems", Journal of Physics A: Mathematical and General **28**, L597–L601 (1995).

[196]J Liu and E Luijten, "Rejection-free geometric cluster algorithm for complex fluids", Physical Review Letters **92**, 10.1103/physrevlett.92.035504 (2004).

[197]J Liu and E Luijten, "Generalized geometric cluster algorithm for fluid simulation", Physical Review E **71**, 10.1103/physreve.71.066701 (2005).

[198]E Luijten, "Introduction to cluster Monte Carlo algorithms", in *Computer simulations in condensed matter systems: from materials to chemical biology volume 1* (Springer Berlin Heidelberg), pp. 13–38.

[199]C Day, "New algorithm speeds up computer simulations of complex fluids", Physics Today **57**, 25–27 (2004).

[200]PM König, R Roth, and KR Mecke, "Morphological thermodynamics of fluids: shape dependence of free energies", Physical Review Letters **93**, 10.1103/physrevlett.93.160601 (2004).

[201]AL Thorneywork, JL Abbott, DGL Aarts, and RPA Dullens, "Two-dimensional melting of colloidal hard spheres", Physical Review Letters **118**, 10.1103/physrevlett.118.158001 (2017).

[202]EP Bernard, W Krauth, and DB Wilson, "Event-chain monte carlo algorithms for hard-sphere systems", Physical Review E **80**, 10.1103/physreve.80.056704 (2009).

[203]EP Bernard and W Krauth, "Two-step melting in two dimensions: first-order liquid-hexatic transition", Physical Review Letters **107**, 10.1103/physrevlett.107.155704 (2011).

[204]M Michel, SC Kapfer, and W Krauth, "Generalized event-chain monte carlo: constructing rejection-free global-balance algorithms from infinitesimal steps", The Journal of Chemical Physics **140**, 054116 (2014).

[205]W Krauth and R Moessner, "Pocket Monte Carlo algorithm for classical doped dimer models", Physical Review B **67**, 10.1103/physrevb.67.064503 (2003).

[206]DA Huse, W Krauth, R Moessner, and SL Sondhi, "Coulomb and liquid dimer models in three dimensions", Physical Review Letters **91**, 10.1103/physrevlett.91.167004 (2003).

[207]A Buhot and W Krauth, "Numerical solution of hard-core mixtures", Physical Review Letters **80**, 3787–3790 (1998).

[208]F Sausset, G Tarjus, and P Viot, "Tuning the fragility of a glass-forming liquid by curving space", Physical Review Letters **101**, 10.1103/physrevlett.101.155701 (2008).

[209]P Fonda, M Rinaldin, DJ Kraft, and L Giomi, "Thermodynamic equilibrium of binary mixtures on curved surfaces", Physical Review E **100**, 10.1103/physreve.100.032604 (2019).

[210]LBG Cortes, Y Gao, RPA Dullens, and DGAL Aarts, "Colloidal liquid crystals in square confinement: isotropic, nematic and smectic phases", Journal of Physics: Condensed Matter **29**, 064003 (2016).

[211]D Gazzillo and G Pastore, "Equation of state for symmetric non-additive hard-sphere fluids: an approximate analytic expression and new Monte Carlo results", Chemical Physics Letters **159**, 388–392 (1989).

[212]TS Grigera and G Parisi, "Fast Monte Carlo algorithm for supercooled soft spheres", Physical Review E **63**, 10.1103/physreve.63.045102 (2001).

[213]A Cavagna, TS Grigera, and P Verrocchio, "Dynamic relaxation of a liquid cavity under amorphous boundary conditions", The Journal of Chemical Physics **136**, 204502 (2012).

[214]R Gutiérrez, S Karmakar, YG Pollack, and I Procaccia, "The static length-scale characterizing the glass transition at lower temperatures", EPL (Europhysics Letters) **111**, 56009 (2015).

[215]L Berthier, D Coslovich, A Ninarello, and M Ozawa, "Equilibrium sampling of hard spheres up to the jamming density and beyond", Physical Review Letters **116**, 10.1103/physrevlett.116.238002 (2016).

[216]A Ninarello, L Berthier, and D Coslovich, "Models and algorithms for the next generation of glass transition studies", Physical Review X **7**, 10.1103/physrevx.7.021039 (2017).

[217]L Berthier, P Charbonneau, D Coslovich, A Ninarello, M Ozawa, and S Yaida, "Configurational entropy measurements in extremely supercooled liquids that break the glass ceiling", Proceedings of the National Academy of Sciences **114**, 11356–11361 (2017).

[218]L Berthier, G Biroli, JP Bouchaud, and G Tarjus, "Can the glass transition be explained without a growing static length scale?", The Journal of Chemical Physics **150**, 094501 (2019).

[219]L Santen and W Krauth, "Absence of thermodynamic phase transition in a model glass former", Nature **405**, 550–551 (2000).

[220]JG Malherbe and S Amokrane, "True mixture versus effective one component fluid models of asymmetric binary hard sphere mixtures: a comparison by simulation", Molecular Physics **99**, 355–361 (2001).

Chapter 4: Droplet singularities

[187]VV Mangazeev, MT Batchelor, VV Bazhanov, and MY Dudalev, "Variational approach to the scaling function of the 2d Ising model in a magnetic field", Journal of Physics A: Mathematical and Theoretical **42**, 042005 (2008).

[189]J Cardy, *Scaling and renormalization in statistical physics* (Cambridge University Press, Cambridge New York, 1996).

[221]J Kent-Dobias and JP Sethna, "Essential singularities in universal scaling functions at the Ising coexistence line", (2017), arXiv:1707.03791v2 [cond-mat.stat-mech].

[222]KG Wilson, "Renormalization group and critical phenomena I: Renormalization group and the Kadanoff scaling picture", Physical Review B **4**, 3174–3183 (1971).

[223]KG Wilson, "Renormalization group and critical phenomena II: Phase-space cell analysis of critical behavior", Physical Review B **4**, 3184–3205 (1971).

[224]ME Fisher, "The theory of condensation and the critical point", Physics Physique Fizika **3**, 255–283 (1967).

[225]R Guida and J Zinn-Justin, "3d Ising model: The scaling equation of state", Nuclear Physics B **489**, 626–652 (1997).

[226]S Seide and C Wetterich, "Equation of state near the endpoint of the critical line", Nuclear Physics B **562**, 524–546 (1999).

[227]P Schofield, "Parametric representation of the equation of state near a critical point", Physical Review Letters **22**, 606–608 (1969).

[228]P Schofield, JD Litster, and JT Ho, "Correlation between critical coefficients and critical exponents", Physical Review Letters **23**, 1098–1102 (1969).

[229]M Caselle, M Hasenbusch, A Pelissetto, and E Vicari, "The critical equation of state of the two-dimensional Ising model", Journal of Physics A: Mathematical and General **34**, 2923–2948 (2001).

[230]BD Josephson, "Equation of state near the critical point", Journal of Physics C: Solid State Physics **2**, 1113–1115 (1969).

[231]ME Fisher, S Zinn, and PJ Upton, "Trigonometric models for scaling behavior near criticality", Physical Review B **59**, 14533–14545 (1999).

[232]J Engels, L Fromme, and M Seniuch, "Numerical equation of state and other scaling functions from an improved three-dimensional Ising model", Nuclear Physics B **655**, 277–299 (2003).

[233]RJ McCraw and LS Schulman, "Metastability in the two-dimensional Ising model", Journal of Statistical Physics **18**, 293–308 (1978).

[234]IG Enting and RJ Baxter, "An investigation of the high-field series expansions for the square lattice Ising model", Journal of Physics A: Mathematical and General **13**, 3723–3734 (1980).

[235]GA Baker and D Kim, "Ising model phase boundary", Journal of Physics A: Mathematical and General **13**, L103–L106 (1980).

[236]VV Mangazeev, MY Dudalev, VV Bazhanov, and MT Batchelor, "Scaling and universality in the two-dimensional Ising model with a magnetic field", Physical Review E **81**, 10.1103/physreve.81.060103 (2010).

[237]W Klein, DJ Wallace, and RKP Zia, "Essential singularities at first-order phase transitions", Physical Review Letters **37**, 639–642 (1976).

[238]W Klein, "Droplet models, renormalization group, and essential singularities at first-order phase transitions", Physical Review B **21**, 5254–5261 (1980).

[239]P Fonseca and A Zamolodchikov, "Ising field theory in a magnetic field: analytic properties of the free energy", Journal of Statistical Physics **110**, 527–590 (2003).

[240]A Zamolodchikov and I Ziyatdinov, "Inelastic scattering and elastic amplitude in ising field theory in a weak magnetic field at . perturbative analysis", Nuclear Physics B **849**, 654–674 (2011).

[241]SN Isakov, "Nonanalytic features of the first order phase transition in the ising model", Communications in Mathematical Physics **95**, 427–443 (1984).

[242]WP Orrick, B Nickel, AJ Guttmann, and JHH Perk, "The susceptibility of the square lattice Ising model: New developments", Journal of Statistical Physics **102**, 795–841 (2001).

[243]Y Chan, AJ Guttmann, BG Nickel, and JHH Perk, "The Ising susceptibility scaling function", Journal of Statistical Physics **145**, 549–590 (2011).

[244]AJ Guttmann and IG Enting, "Solvability of some statistical mechanical systems", Physical Review Letters **76**, 344–347 (1996).

[245]B Nickel, "On the singularity structure of the 2d Ising model susceptibility", Journal of Physics A: Mathematical and General **32**, 3889–3906 (1999).

[246]B Nickel, "Addendum to 'on the singularity structure of the 2d Ising model susceptibility'", Journal of Physics A: Mathematical and General **33**, 1693–1711 (2000).

[247]M Assis, JL Jacobsen, I Jensen, JM Maillard, and BM McCoy, "Analyticity of the Ising susceptibility: An interpretation", Journal of Physics A: Mathematical and Theoretical **50**, 365203 (2017).

[248]JS Langer, "Statistical theory of the decay of metastable states", Annals of Physics **54**, 258–275 (1969).

[249]O Penrose and JL Lebowitz, "Towards a rigorous molecular theory of metastability", in *Fluctuation phenomena*, edited by EW Montroll and JL Lebowitz (Elsevier, 1987), pp. 7–293.

[250]B Gaveau and LS Schulman, "Metastable decay rates and analytic continuation", Letters in Mathematical Physics **18**, 201–208 (1989).

[251]V Privman and LS Schulman, "Analytic properties of thermodynamic functions at first-order phase transitions", Journal of Physics A: Mathematical and General **15**, L231–L238 (1982).

[252]V Privman and LS Schulman, "Analytic continuation at first-order phase transitions", Journal of Statistical Physics **29**, 205–229 (1982).

[253]CCA Günther, PA Rikvold, and MA Novotny, "Numerical transfer-matrix study of metastability in the $d = 2$ Ising model", Physical Review Letters **71**, 3898–3901 (1993).

[254]S Ryu and W Cai, "Validity of classical nucleation theory for Ising models", Physical Review E **81**, 10.1103/physreve.81.030601 (2010).

[255]A Nußbaumer, E Bittner, T Neuhaus, and W Janke, "Monte Carlo study of the evaporation/condensation transition of Ising droplets", Europhysics Letters (EPL) **75**, 716–722 (2006).

[256]H Tomita and S Miyashita, "Statistical properties of the relaxation processes of metastable states in the kinetic Ising model", Physical Review B **46**, 8886–8893 (1992).

[257]PA Rikvold, H Tomita, S Miyashita, and SW Sides, "Metastable lifetimes in a kinetic Ising model: Dependence on field and system size", Physical Review E **49**, 5080–5090 (1994).

[258]NJ Gunther, DJ Wallace, and DA Nicole, "Goldstone modes in vacuum decay and first-order phase transitions", Journal of Physics A: Mathematical and General **13**, 1755–1767 (1980).

[259]A Houghton and TC Lubensky, "The metastable Ising magnet in a negative field", Physics Letters A **77**, 479–480 (1980).

[260]F Ramos-Gómez and B Widom, "Noncritical interface near a critical end point II", Physica A: Statistical Mechanics and its Applications **104**, 595–620 (1980).

[261]Sy Zinn and ME Fisher, "Universal surface-tension and critical-isotherm amplitude ratios in three dimensions", Physica A: Statistical Mechanics and its Applications **226**, 168–180 (1996).

[262]JS Langer, "Theory of the condensation point", Annals of Physics **41**, 108–157 (1967).

[263]CK Harris, "Metastability in the $(1+1)$d Ising model: a primitive droplet model calculation", Journal of Physics A: Mathematical and General **17**, 1767–1770 (1984).

[264]BM McCoy and TT Wu, "Two-dimensional ising field theory in a magnetic field: breakup of the cut in the two-point function", Physical Review D **18**, 1259–1267 (1978).

[265]RB Griffiths, "Thermodynamic functions for fluids and ferromagnets near the critical point", Physical Review **158**, 176–187 (1967).

[266]MJ Lowe and DJ Wallace, "Instantons and the Ising model below T_c", Journal of Physics A: Mathematical and General **13**, L381–L385 (1980).

[267]G Parisi, "Asymptotic estimates in perturbation theory", Physics Letters B **66**, 167–169 (1977).

[268]EB Bogomolny, "Calculation of the Green functions by the coupling constant dispersion relations", Physics Letters B **67**, 193–194 (1977).

[269]L Onsager, "Crystal statistics I: A two-dimensional model with an order-disorder transition", Physical Review **65**, 117–149 (1944).

[270]M Caselle, M Hasenbusch, and M Panero, "The interface free energy: Comparison of accurate Monte Carlo results for the 3d Ising model with effective interface models", Journal of High Energy Physics **2007**, 117–117 (2007).

[271]M Campostrini, A Pelissetto, P Rossi, and E Vicari, "25th-order high-temperature expansion results for three-dimensional Ising-like systems on the simple-cubic lattice", Physical Review E **65**, 10.1103/physreve.65.066127 (2002).

[272]AL Talapov and HWJ Blöte, "The magnetization of the 3d Ising model", Journal of Physics A: Mathematical and General **29**, 5727–5733 (1996).

[273]E Barouch, BM McCoy, and TT Wu, "Zero-field susceptibility of the two-dimensional Ising model near T_c", Physical Review Letters **31**, 1409–1411 (1973).

[274]HWJ Blöte, E Luijten, and JR Heringa, "Ising universality in three dimensions: a Monte Carlo study", Journal of Physics A: Mathematical and General **28**, 6289–6313 (1995).

[275]P Butera and M Pernici, "Free energy in a magnetic field and the universal scaling equation of state for the three-dimensional ising model", Phys. Rev. B **83**, 054433 (2011).

[276]RKP Zia, "The variety of singularities in models of first order phase transitions", Zeitschrift für Physik B Condensed Matter **41**, 129–138 (1981).

[277]W Klein, "Percolation, droplet models, and spinodal points", Physical Review Letters **47**, 1569–1572 (1981).

[278]D Stauffer, A Coniglio, and DW Heermann, "Monte carlo experiment for nucleation rate in the three-dimensional ising model", Physical Review Letters **49**, 1299–1302 (1982).

[279]YJ Chen, NM Paquette, BB Machta, and JP Sethna, *Universal scaling function for the two-dimensional Ising model in an external field: A pragmatic approach*, arXiv:1307.6899 [cond-mat.stat-mech].

[280]TT Wu, BM McCoy, CA Tracy, and E Barouch, "Spin-spin correlation functions for the two-dimensional Ising model: Exact theory in the scaling region", Physical Review B **13**, 316–374 (1976).

[281] A Parola, D Pini, and L Reatto, "First-order phase transitions, the maxwell construction, and the momentum-space renormalization group", Physical Review E **48**, 3321–3332 (1993).

[282] J Alexandre, V Branchina, and J Polonyi, "Global renormalization group", Physical Review D **58**, 10.1103/physrevd.58.016002 (1998).

[283] JD Gunton and MC Yalabik, "Renormalization-group analysis of the mean-field theory of metastability: a spinodal fixed point", Physical Review B **18**, 6199–6205 (1978).

[284] V Ambegaokar, BI Halperin, DR Nelson, and ED Siggia, "Dissipation in two-dimensional superfluids", Physical Review Letters **40**, 783–786 (1978).

[285] V Ambegaokar, BI Halperin, DR Nelson, and ED Siggia, "Dynamics of superfluid films", Physical Review B **21**, 1806–1826 (1980).

[286] AJ Dahm, MA Stan, and RG Petschek, "Dynamics of dislocation-mediated melting in a two-dimensional lattice in the presence of an oscillatory applied strain", Physical Review B **40**, 9006–9016 (1989).

Chapter 5: Normal form of nonlinear RGs

[55] MEJ Newman and RM Ziff, "Fast Monte Carlo algorithm for site or bond percolation", Physical Review E **64**, 10.1103/physreve.64.016706 (2001).

[189] J Cardy, *Scaling and renormalization in statistical physics* (Cambridge University Press, Cambridge New York, 1996).

[222] KG Wilson, "Renormalization group and critical phenomena I: Renormalization group and the Kadanoff scaling picture", Physical Review B **4**, 3174–3183 (1971).

[223] KG Wilson, "Renormalization group and critical phenomena II: Phase-space cell analysis of critical behavior", Physical Review B **4**, 3184–3205 (1971).

[287] A Raju, CB Clement, LX Hayden, J Kent-Dobias, DB Liarte, DZ Rocklin, and JP Sethna, "Normal form for renormalization groups", Physical Review X **9**, 10.1103/physrevx.9.021014 (2019).

[288] KG Wilson and J Kogut, "The renormalization group and the ϵ expansion", Physics Reports **12**, 75–199 (1974).

[289] A Aharony and ME Fisher, "Nonlinear scaling fields and corrections to scaling near criticality", Physical Review B **27**, 4394–4400 (1983).

[290] JM Kosterlitz, "Nobel lecture: Topological defects and phase transitions", Reviews of Modern Physics **89**, 10.1103/revmodphys.89.040501 (2017).

[291] VL Berezinskiĭ, "Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group I: Classical systems", Soviet Journal of Experimental and Theoretical Physics **32**, 493 (1971).

[292] JM Kosterlitz and DJ Thouless, "Ordering, metastability and phase transitions in two-dimensional systems", Journal of Physics C: Solid State Physics **6**, 1181–1203 (1973).

[293] JM Kosterlitz, "The critical properties of the two-dimensional xy model", Journal of Physics C: Solid State Physics **7**, 1046–1060 (1974).

[294] JM Kosterlitz, "Kosterlitz–Thouless physics: a review of key issues", Reports on Progress in Physics **79**, 026001 (2016).

[295]HJF Knops, "Exact relation between the solid-on-solid model and the XY model", Physical Review Letters **39**, 766–769 (1977).

[296]D Boyanovsky, "Field-theoretical renormalisation and fixed-point structure of a generalised Coulomb gas", Journal of Physics A: Mathematical and General **22**, 2601–2614 (1989).

[297]DJ Amit, YY Goldschmidt, and S Grinstein, "Renormalisation group analysis of the phase transition in the 2d Coulomb gas, sine-Gordon theory and XY-model", Journal of Physics A: Mathematical and General **13**, 585–620 (1980).

[298]ST Chui and JD Weeks, "Phase transition in the two-dimensional coulomb gas, and the interfacial roughening transition", Physical Review B **14**, 4978–4982 (1976).

[299]JM Kosterlitz, "The d-dimensional Coulomb gas and the roughening transition", Journal of Physics C: Solid State Physics **10**, 3753–3760 (1977).

[300]A Pelissetto and E Vicari, "Renormalization-group flow and asymptotic behaviors at the Berezinskii-Kosterlitz-Thouless transitions", Physical Review E **87**, 10.1103/physreve.87.032105 (2013).

[301]J Balog, M Niedermaier, F Niedermayer, A Patrascioiu, E Seiler, and P Weisz, "The intrinsic coupling in integrable quantum field theories", Nuclear Physics B **583**, 614–670 (2000).

[302]J Balog and Á Hegedus, "Two-loop beta-functions of the sine-Gordon model", Journal of Physics A: Mathematical and General **33**, 6543–6547 (2000).

[303]LP Kadanoff and AB Zisook, "Correlation functions on the critical line of the two-dimensional planar model: logarithms and corrections to scaling", Nuclear Physics B **180**, 61–76 (1981).

[304]P Ginsparg, *Applied conformal field theory*, arXiv:hep-th/9108028v1.

[305]C Itoi and H Mukaida, "Renormalization group for renormalization-group equations toward the universality classification of infinite-order phase transitions", Physical Review E **60**, 3688–3700 (1999).

[306]AP Young, "Melting and the vector Coulomb gas in two dimensions", Physical Review B **19**, 1855–1866 (1979).

[307]P Yu, "Computation of the simplest normal forms with perturbation parameters based on Lie transform and rescaling", Journal of Computational and Applied Mathematics **144**, 359–373 (2002).

[308]P Yu and AYT Leung, "The simplest normal form and its application to bifurcation control", Chaos, Solitons & Fractals **33**, 845–863 (2007).

[309]DS Callaway, JE Hopcroft, JM Kleinberg, MEJ Newman, and SH Strogatz, "Are randomly grown graphs really random?", Physical Review E **64**, 10.1103/physreve.64.041902 (2001).

[310]SN Dorogovtsev, JFF Mendes, and AN Samukhin, "Anomalous percolation properties of growing networks", Physical Review E **64**, 10.1103/physreve.64.066110 (2001).

[311]C Borgs, JT Chayes, H Kesten, and J Spencer, "The birth of the infinite cluster: Finite-size scaling in percolation", Communications in Mathematical Physics **224**, 153–204 (2001).

[312]T Nogawa, T Hasegawa, and K Nemoto, "Generalized scaling theory for critical phenomena including essential singularities and infinite dimensionality", Physical Review Letters **108**, 10.1103/physrevlett.108.255703 (2012).

[313]C Itoi and MH Kato, "Extended massless phase and the Haldane phase in a spin-1 isotropic antiferromagnetic chain", Physical Review B **55**, 8295–8303 (1997).

[314]DR Nelson and BI Halperin, "Dislocation-mediated melting in two dimensions", Physical Review B **19**, 2457–2484 (1979).

[315]SA Bulgadaev, "Berezinskii-Kosterlitz-Thouless phase transitions in two-dimensional systems with internal symmetries", Journal of Experimental and Theoretical Physics **89**, 1107–1113 (1999).

[316]G Bianconi, I Kryven, and RM Ziff, "Percolation on branching simplicial and cell complexes and its relation to interdependent percolation", Physical Review E **100**, 10.1103/physreve.100.062311 (2019).

[317]T Hasegawa, T Nogawa, and K Nemoto, "Profile and scaling of the fractal exponent of percolations in complex networks", EPL (Europhysics Letters) **104**, 16006 (2013).

[318]SN Dorogovtsev, AV Goltsev, and JFF Mendes, "Critical phenomena in complex networks", Reviews of Modern Physics **80**, 1275–1335 (2008).

[319]M Hinczewski and AN Berker, "Inverted Berezinskii-Kosterlitz-Thouless singularity and high-temperature algebraic order in an Ising model on a scale-free hierarchical-lattice small-world network", Physical Review E **73**, 10.1103/physreve.73.066126 (2006).

[320]AN Berker, M Hinczewski, and RR Netz, "Critical percolation phase and thermal Berezinskii-Kosterlitz-Thouless transition in a scale-free network with short-range and long-range random bonds", Physical Review E **80**, 10.1103/physreve.80.041118 (2009).

[321]KG Wilson, "The renormalization group: critical phenomena and the Kondo problem", Reviews of Modern Physics **47**, 773–840 (1975).

[322]PW Anderson and G Yuval, "Some numerical results on the Kondo problem and the inverse square one-dimensional Ising model", Journal of Physics C: Solid State Physics **4**, 607–620 (1971).

[323]PB Wiegmann, "Exact solution of the s-d exchange model (Kondo problem)", Journal of Physics C: Solid State Physics **14**, 1463–1478 (1981).

[324]E Luijten and HWJ Blöte, "Classical critical behavior of spin models with long-range interactions", Physical Review B **56**, 8945–8958 (1997).

[325]MJ Wragg and GA Gehring, "The Ising model with long-range ferromagnetic interactions", Journal of Physics A: Mathematical and General **23**, 2157–2164 (1990).

[326]J Bhattacharjee, S Chakravarty, JL Richardson, and DJ Scalapino, "Some properties of a one-dimensional Ising chain with an inverse-square interaction", Physical Review B **24**, 3862–3865 (1981).

[327]E Luijten and H Meßingfeld, "Criticality in one dimension with inverse square-law potentials", Physical Review Letters **86**, 5305–5308 (2001).

[328]PW Anderson, G Yuval, and DR Hamann, "Scaling theory for the Kondo and one-dimensional Ising models", Solid State Communications **8**, 1033–1037 (1970).

Chapter 6: Hidden order in URu_2Si_2

[171]R Guida and J Zinn-Justin, "Critical exponents of the N-vector model", Journal of Physics A: Mathematical and General **31**, 8103–8121 (1998).

[329]J Kent-Dobias, M Matty, and B Ramshaw, "Elastic properties of hidden order in URu_2Si_2 are reproduced by a staggered nematic", (2019), arXiv:1910.01669v2 [cond-mat.str-el].

[330]E Hassinger, G Knebel, K Izawa, P Lejay, B Salce, and J Flouquet, "Temperature-pressure phase diagram of URu_2Si_2 from resistivity measurements and ac calorimetry: hidden order and Fermi-surface nesting", Physical Review B **77**, 115117 (2008).

[331]S Kambe, Y Tokunaga, H Sakai, T Hattori, N Higa, TD Matsuda, Y Haga, RE Walstedt, and H Harima, "Odd-parity electronic multipolar ordering in URu_2Si_2: conclusions from Si and Ru NMR measurements", Physical Review B **97**, 235142 (2018).

[332]K Haule and G Kotliar, "Arrested Kondo effect and hidden order in URu_2Si_2", Nature Physics **5**, 796–799 (2009).

[333]H Kusunose and H Harima, "On the hidden order in URu_2Si_2—antiferro hexadecapole order and its consequences", Journal of the Physical Society of Japan **80**, 084702 (2011).

[334]HH Kung, RE Baumbach, ED Bauer, VK Thorsmolle, WL Zhang, K Haule, JA Mydosh, and G Blumberg, "Chirality density wave of the "hidden order" phase in URu_2Si_2", Science **347**, 1339–1342 (2015).

[335]F Cricchio, F Bultmark, O Grånäs, and L Nordström, "Itinerant magnetic multipole moments of rank five as the hidden order in URu_2Si_2", Physical Review Letters **103**, 107202 (2009).

[336]FJ Ohkawa and H Shimizu, "Quadrupole and dipole orders in URu_2Si_2", Journal of Physics: Condensed Matter **11**, L519–L524 (1999).

[337]P Santini and G Amoretti, "Crystal field model of the magnetic properties of URu_2Si_2", Physical Review Letters **73**, 1027–1030 (1994).

[338]A Kiss and P Fazekas, "Group theory and octupolar order in URu_2Si_2", Physical Review B **71**, 054415 (2005).

[339]H Harima, K Miyake, and J Flouquet, "Why the hidden order in URu_2Si_2 is still hidden—one simple answer", Journal of the Physical Society of Japan **79**, 033705 (2010).

[340]P Thalmeier and T Takimoto, "Signatures of hidden-order symmetry in torque oscillations, elastic constant anomalies, and field-induced moments in URu_2Si_2", Physical Review B **83**, 165110 (2011).

[341]S Tonegawa, K Hashimoto, K Ikada, YH Lin, H Shishido, Y Haga, TD Matsuda, E Yamamoto, Y Onuki, H Ikeda, Y Matsuda, and T Shibauchi, "Cyclotron resonance in the hidden-order phase of URu_2Si_2", Physical Review Letters **109**, 036401 (2012).

[342]JG Rau and HY Kee, "Hidden and antiferromagnetic order as a rank-5 superspin in URu_2Si_2", Physical Review B **85**, 245112 (2012).

[343]SC Riggs, MC Shapiro, AV Maharaj, S Raghu, ED Bauer, RE Baumbach, P Giraldo-Gallo, M Wartenbe, and IR Fisher, "Evidence for a nematic component to the hidden-order parameter in URu_2Si_2 from differential elastoresistance measurements", Nature Communications **6**, 6425 (2015).

[344]S Hoshino, J Otsuki, and Y Kuramoto, "Resolution of entropy $\ln\sqrt{2}$ by ordering in two-channel Kondo lattice", Journal of the Physical Society of Japan **82**, 044707 (2013).

[345]H Ikeda and Y Ohashi, "Theory of unconventional spin density wave: a possible mechanism of the micromagnetism in U-based heavy fermion compounds", Physical Review Letters **81**, 3723–3726 (1998).

[346]P Chandra, P Coleman, and R Flint, "Hastatic order in the heavy-fermion compound URu_2Si_2", Nature **493**, 621–626 (2013).

[347]N Harrison and M Jaime, *Hidden valence transition in URu_2Si_2?*, arXiv:1902.06588v2 [cond-mat.str-el].

[348]H Ikeda, MT Suzuki, R Arita, T Takimoto, T Shibauchi, and Y Matsuda, "Emergent rank-5 nematic order in URu_2Si_2", Nature Physics **8**, 528–533 (2012).

[349]A de Visser, FE Kayzel, AA Menovsky, JJM Franse, J van den Berg, and GJ Nieuwenhuys, "Thermal expansion and specific heat of monocrystalline URu_2Si_2", Physical Review B **34**, 8168–8171 (1986).

[350]J Choi, O Ivashko, N Dennler, D Aoki, Kv Arx, S Gerber, O Gutowski, MH Fischer, J Strempfer, Mv Zimmermann, and J Chang, "Pressure-induced rotational symmetry breaking in URu_2Si_2", Physical Review B **98**, 241113 (2018).

[351]S Ghosh, M Matty, R Baumbach, ED Bauer, KA Modic, A Shekhter, JA Mydosh, EA Kim, and BJ Ramshaw, "One-component order parameter in URu_2Si_2 uncovered by resonant ultrasound spectroscopy and machine learning", Science Advances **6**, eaaz4074 (2020).

[352]P Chandra, P Coleman, and R Flint, "Origin of the large anisotropy in the χ_3 anomaly in URu_2Si_2", Journal of Physics: Conference Series **449**, 012026 (2013).

[353]B Wolf, W Sixl, R Graf, D Finsterbusch, G Bruls, B Lüthi, EA Knetsch, AA Menovsky, and JA Mydosh, "Elastic properties of the heavy fermion superconductor URu_2Si_2", Journal of Low Temperature Physics **94**, 307–324 (1994).

[354]K Kuwahara, H Amitsuka, T Sakakibara, O Suzuki, S Nakamura, T Goto, M Mihalik, AA Menovsky, A de Visser, and JJM Franse, "Lattice instability and elastic response in the heavy electron system URu_2Si_2", Journal of the Physical Society of Japan **66**, 3251–3258 (1997).

[355]LD Landau, EM Lifshitz, AM Kosevich, and LP Pitaevskiĭ, *Theory of elasticity*, 3rd, Vol. 7, Landau and Lifshitz Course of Theoretical Physics (Pergamon Press, 1986).

[356]B Lüthi and TJ Moran, "Sound propagation near the structural phase transition in strontium titanate", Physical Review B **2**, 1211–1214 (1970).

[357]BJ Ramshaw, A Shekhter, RD McDonald, JB Betts, JN Mitchell, PH Tobash, CH Mielke, ED Bauer, and A Migliori, "Avoided valence transition in a plutonium superconductor", Proceedings of the National Academy of Sciences **112**, 3285–3289 (2015).

[358]A Shekhter, BJ Ramshaw, R Liang, WN Hardy, DA Bonn, FF Balakirev, RD McDonald, JB Betts, SC Riggs, and A Migliori, "Bounding the pseudogap with a line of phase transitions in $YBa_2Cu_3O_{6+\delta}$", Nature **498**, 75–77 (2013).

[359]EM Lifshitz, "On the theory of phase transitions of the second order I: Changes of the elementary cell of a crystal in phase transitions of the second order", Proceedings of the USSR Academy of Sciences Journal of Physics **6**, 61 (1942).

[360]EM Lifshitz, "On the theory of phase transitions of the second order II: Phase transitions of the second order in alloys", Proceedings of the USSR Academy of Sciences Journal of Physics **6**, 251 (1942).

[361]PM Chaikin and TC Lubensky, *Principles of condensed matter physics*, 1st (Cambridge University Press, 1995).

[362]YP Varshni, "Temperature dependence of the elastic constants", Physical Review B **2**, 3952–3958 (1970).

[363]JQ Meng, PM Oppeneer, JA Mydosh, PS Riseborough, K Gofryk, JJ Joyce, ED Bauer, Y Li, and T Durakiewicz, "Imaging the three-dimensional Fermi-surface pairing near the hidden-order transition in URu_2Si_2 using angle-resolved photoemission spectroscopy", Physical Review Letters **111**, 127002 (2013).

[364]C Broholm, H Lin, PT Matthews, TE Mason, WJL Buyers, MF Collins, AA Menovsky, JA Mydosh, and JK Kjems, "Magnetic excitations in the heavy-fermion superconductor URu_2Si_2", Physical Review B **43**, 12809–12822 (1991).

[365]CR Wiebe, JA Janik, GJ MacDougall, GM Luke, JD Garrett, HD Zhou, YJ Jo, L Balicas, Y Qiu, JRD Copley, Z Yamani, and WJL Buyers, "Gapped itinerant spin excitations account for missing entropy in the hidden-order state of URu_2Si_2", Nature Physics **3**, 96–99 (2007).

[366]F Bourdarot, E Hassinger, S Raymond, D Aoki, V Taufour, LP Regnault, and J Flouquet, "Precise study of the resonance at $Q_0 = (1, 0, 0)$ in URu_2Si_2", Journal of the Physical Society of Japan **79**, 064719 (2010).

[367]E Hassinger, G Knebel, TD Matsuda, D Aoki, V Taufour, and J Flouquet, "Similarity of the Fermi surface in the hidden order state and in the antiferromagnetic state of URu_2Si_2", Physical Review Letters **105**, 216409 (2010).

[368]P Bak, "Commensurate phases, incommensurate phases and the devil's staircase", Reports on Progress in Physics **45**, 587–629 (1982).

[369]T Yanagisawa, "Ultrasonic study of the hidden order and heavy-fermion state in URu_2Si_2 with hydrostatic pressure, Rh-doping, and high magnetic fields", Philosophical Magazine **94**, 3775–3788 (2014).

[370]T Garel and P Pfeuty, "Commensurability effects on the critical behaviour of systems with helical ordering", Journal of Physics C: Solid State Physics **9**, L245–L249 (1976).

[371]T Inoue, K Kindo, H Okuni, K Sugiyama, Y Haga, E Yamamoto, TC Kobayashi, Y Uwatoko, and Y Onuki, "High-field magnetization of URu_2Si_2 under high pressure", Physica B: Condensed Matter **294-295**, 271–275 (2001).

[372]G Ghiringhelli, M Le Tacon, M Minola, S Blanco-Canosa, C Mazzoli, NB Brookes, GM De Luca, A Frano, DG Hawthorn, F He, T Loew, MM Sala, DC Peets, M Salluzzo, E Schierle, R Sutarto, GA Sawatzky, E Weschke, B Keimer, and L Braicovich, "Long-range incommensurate charge fluctuations in (Y, Nd)$Ba_2Cu_3O_{6+x}$", Science **337**, 821–825 (2012).

[373]RM Hornreich, "The Lifshitz point: phase diagrams and critical behavior", Journal of Magnetism and Magnetic Materials **15-18**, 387–392 (1980).

[374]VL Ginzburg, "Some remarks on phase transitions of the second kind and the microscopic theory of ferroelectric materials", Soviet Physics, Solid State **2**, 1824–1834 (1961).

[375]E Berg, E Fradkin, SA Kivelson, and JM Tranquada, "Striped superconductors: how spin, charge and superconducting orders intertwine in the cuprates", New Journal of Physics **11**, 115004 (2009).