Algorithms for Measuring Extremely Rare Events in Statistical Physics

by

David A. Adams

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2011

Doctoral Committee:

Professor Leonard M. Sander, Co-Chair
Professor Robert M. Ziff, Co-Chair
Professor Mark E. Newman
Professor Gregory Tarle
Associate Professor Michal R. Żochowski

To mom and dad for continual inspiration

ACKNOWLEDGEMENTS

This thesis has been possible because of a number of caring people who have helped me along this path. First, I would like to thank my parents Jim and Becky Adams. It was because of their support, both emotional and intellectual, that I started down this path. Next, I would like to thank Beate Schmittmann and Royce Zia. It was my undergraduate work with them which sparked my interest in research, especially statistical physics. They showed the me excitement of discovering something new about the universe. I would also like to thank all of the other professors and graduate students that I worked with as an undergraduate including Alexey Onufriev, Andrew Fenley, Leo Piilonen, Greg Tarle, Dave Gerdes, Tim McKay, and Sharon Glotzer. During my time in graduate school, other graduate students were an invaluable resource and gave helpful guidance both academic and practical. Many people fall into this category including Jack Waddell Sarah Feldt, Casey Schneider-Mizell, Yen Ting Lin, Liz Strahman, Chris Fink, Dan Maruyama, John Schaibley, and Mike Slootsky, and Ibrahima Bah. I would also like to thank my two advisors Len Sander and Bob Ziff whose creativity and intellect made working for them a constant joy. Lastly, I would like to thank my wife Katie Adams. Without her love and support getting through graduate school and thesis writing would have be much more challenging.

Many people contributed to the works in this thesis. These following people deserve acknowledgement for their useful conversations and input: Len Sander, Bob

Ziff, Yen Ting Lin, Ellak Somfai, Charlie Doering, Arthur Voter, Bertrand Duplantier, Chris Fink, Robin Ball, Sorin Tanase-Nicola, Evgeniy Khain, and Mark Dykman. Financial assistance for the work here came from grants from the NSF. The computational resources of the Center for Advanced Computing in the University of Michigan School of Engineering and TeraGrid were used for most simulations presented.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

CHAPTER I

Introduction

1.1 Rare Events

An event that occurs infrequently is called a "rare event." Some rare events can be of significant interest. Monte Carlo sampling can accurately determine expectation values and rates of typical events in systems. In contrast, efficiently measuring rare events requires special techniques because many rare events occur so infrequently, e.g., every 10^{-30} events, that they are effectively impossible to observe during normal simulations. The development and effective use of these techniques is the topic of this thesis.

Interestingly, rare events are seen in many contexts: queue overflow in queuing theory [8], bio-chemical switching [9], extinction of diseases [10], or populations [11], slow chemical reactions [12], first-order phase transitions [13], and journeys deep into labyrinths [14]. In this thesis, many of the systems listed above are studied.

The next three chapters describe techniques for measuring extremely rare excursions into random fractals. These rare events are critical for determining the harmonic measure, see section 1.4. The chapter following these discusses properties of critical Q-state Potts model clusters [15]. The two chapters after that cover the development of techniques to measure the average time between rare transitions.

These techniques are used to measure the disease extinction time of the susceptible-infectious-susceptible (SIS) model [16], the switching time of the bistable Maier-Stein model [4], and the poisoning time of the Ziff-Gulari-Barshad (ZGB) model of heterogeneous catalysis[17].

1.2 Rare Event Techniques

When rare events are studied one or more of several properties are measured: the phase space density of extremely low density states, the probability of ever reaching a rarely visited state in a system with absorbing states, the probability of reaching a rare state before a certain amount of time has lapsed, and the average time it takes to perform a rare transition. The vast majority of rare event techniques are focused on the last property, the average time it takes rare transitions to occur. Most of these rare event techniques fall into one of two categories: either biased sampling or path sampling methods.

Splitting was the first rare event technique developed [18], and uses biased sampling to gain efficiency in measuring rare phase space densities. Splitting consists of running a single non-deterministic trajectory through phase space until it reaches a rare region, bounded by a surface, after which it is split into several independent trajectories whose statistical weights add up to that of the original. This way the region can be efficiently sampled. The phase-space density is approximated by weighted sampling of the space by all of the trajectories. A single splitting is usually insufficient to sample very rare regions, so several splitting surfaces are typically used, which is called multi-level splitting [19]. Two main factors limit the usefulness of multi-level splitting. First, the majority of the low weight trajectories go back to the well sampled regions wasting computational effort. Second, the number of tra-

jectories tend to exponentially grow or shrink depending on how closely the splitting surfaces are spaced [19].

To overcome the first limiting factor of multi-level splitting, RESTART was developed [20]. RESTART works by designating a primary copy when spitting occurs at a barrier. This primary copy is the only copy allowed to backtrack past the splitting barrier without being removed. A splitting-like technique that does not suffer from either of the limitations of multi-level splitting is forward flux sampling (FFS) [21]. FFS uses splitting to measure the small probability of transitioning from the surface of a well-sampled region to another region, using intermediate splitting surfaces. This small probability multiplied by the average rate of leaving the well-sampled region gives the overall transition probability. Lastly, milestoning [22] is a technique that calculates the transition rate between two regions by calculating the flux rates between a number of dividing surfaces. Milestoning differs from FFS in that it equilibrates trajectories along each surface, requiring the system to satisfy detailed balance; FFS does not require detailed balance.

Path sampling, the other major category of rare event techniques, aim to find a set of paths that best represent the transition from one meta-stable state A to another, usually meta-stable, state B. One of the first of these techniques to be developed was transition path sampling (TPS) [23]. TPS performs Monte Carlo sampling over all trajectories of fixed length, i.e., a sequence of neighboring states, that connect A and B. For molecular dynamics, the acceptance rate of a trial move is unity when the first state in the new trajectory (in A) has a higher phase-space density than the first state in the old trajectory. When the first state in the new trajectory has lower phase-space density, the new trajectory is accepted with a probability equal to the ratio of the phase-space density of the first states of the two trajectories. A trial move

is always rejected if the first state is not within A or the last state is not within B. The resulting ensemble of states is then used to calculate the transition rate between A and B with umbrella sampling. Related to TPS and FFS is transition interface sampling (TIS) [24]. TIS samples trajectories like TPS but uses the simple rare event calculation formula of FFS instead of the computationally expensive umbrella sampling used in TPS.

In this thesis we will develop and use biased sampling techniques to measure rare events in non-equilibrium systems.

1.3 Random fractals: Q-state Potts model clusters and Diffusion-limited Aggregates

In this section, we give some background on the systems we study in this thesis.

Undoubtedly, the most well-studied model in statistical mechanics is the Ising model [25], which models magnetic behavior. The hamiltonian of the Ising model can be written as,

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{1.1}$$

where J is the coupling constant and $\langle i, j \rangle$ indicates that the sum is over all nearest neighbor sites i and j. When J is positive, the system is ferromagnetic and anti-ferromagnetic when J is negative. The Ising model can be generalized to Q different spins. One way to accomplishing this is to only allow same spin neighbors to interact, in which case it is called the Q-state Potts model [15],

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta_{s_i,s_j}, \tag{1.2}$$

where δ_{s_i,s_j} is the Kronecker delta, and s_i can take on integer values from 1 to Q. The Potts model is related to a number of different systems [26] including bond percolation [27] ($Q = 1$), the Ising model ($Q = 2$), and mono-layers absorbed on

two-dimensional lattices [28] ($Q = 3$). The first four values of ($Q = 1, 2, 3, 4$) are interesting because they all go through second order, continuous, phrase transitions at their critical temperature, $T_c = J/\ln(1+\sqrt{Q})$. The clusters of like spin that form at the critical temperature are self-similar (fractal) [14]. These fractal clusters can be defined using the Fortuin-Kasteleyn (FK) method of re-writing the partition function in terms of a generalized bond percolation [29]. Directly simulating these systems at their critical temperature is notoriously difficult [30]. To overcome this challenge, Swendsen and Wang [31] developed an extremely efficient dynamic updating scheme for critical Potts model systems. While $Q \geq 2$ can be understood in terms of Ising model behavior, $Q = 1$, percolation, is sufficiently different and warrants further discussion.

Percolation is a simple process that can be performed on any system consisting of sites with bonds connecting them, e.g., regular lattices and graphs. Percolation comes in two varieties: site and bond percolation. The connected clusters generated by keeping only a fraction p of sites from an initially complete square lattice is an example of site percolation. Bond percolation is performed by keeping a fraction of bonds instead of sites. For small values of p, the remaining clusters of connected sites are small. In contrast, the majority of the system remains as a single connected cluster when p is large. A fundamental property of percolating systems is the value of p at which there is a non-zero probability of making a single connected cluster for an extremely large system, p_c. This percolation threshold is known theoretically for some simple lattice systems and known computationally for dozens of other lattices and network models. Like the clusters formed when the other Q-states are prepared at their critical temperature, the percolation clusters formed at p_c are fractal.

Accurately measuring p_c, as well as obtaining large clusters, is difficult using

the simple method of selecting each site (or bond) and choosing to include it with probability p. Methods like the Newman-Ziff algorithm [32] and the Leath algorithm [33] are very efficient ways to obtain the percolation threshold and large percolation clusters, respectively.

In addition to critical Potts model clusters, we studied diffusion-limited aggregate (DLA) clusters [34]. DLA is a simple growth process where a particle starts very far from the cluster and diffuses until it touches a part of it and sticks. This process is performed one particle at a time, starting with an initial stationary particle. The clusters that result from this process are visually striking, complex fractal structures with spindly, spiky branches and deep maze-like fjords. The finger-like growth seen in DLA is surprisingly similar to a wide range of physical phenomena including bacterial growth with limited resources, electrodeposition, crystal growth on surfaces, and viscous fingering, see [35].

The reason why this DLA-like growth gives rise to complex structure is two-fold. First, the deep fjords never fill up because they are screened by the tips of the branches. In fact, we have found that the growth probability deep within the fjords can become as small as 10^{-30} for moderately large clusters [36]. Also, the tips of the DLA fingers are the most likely to grow. The growth on these tips are unstable, due to the Mullins-Sekerka [37] instability, and tend to split with one branch winning out over the other. The mathematical framework and results that are used to understand DLA can be mapped into the other finger-like growth systems due to the similar mathematical forms of the Laplace equation and the source-free diffusion equation. There is another interesting mathematical feature of DLA, namely the distribution of locations at each step where the next particle will adhere, called the harmonic measure. The harmonic measure is especially interesting in the case of

DLA because it is intimately connected to the growth process.

1.4 Harmonic Measure

The solution of the Laplace equation on the surface of a conductor with a fixed charge of unity is called harmonic measure. The harmonic measure, μ, is the normal derivative of the potential at the surface of the object. Another interpretation of the harmonic measure is the distribution of probabilities that a random walker starting far from the object will hit a given section of the perimeter. This quantity is of theoretical [38] as well as practical interest as it determines where diffusing particles will land on rough objects, e.g., catalysts. The harmonic measure of fractals is *multi-fractal* [39], that is, characterized by many generalized dimensions, $D(q)$.

The harmonic measure is commonly analyzed in terms of the spectrum of $D(q)$. By definition, $D(q)$ is the dimension of the surface of an object where the sections of the surface are given weight p_i^q, where p_i is the measure for section i of the surface. When the entire surface is weighted equally ($q = 0$), the fractal dimension of the surface is recovered. $D(q)$ is difficult to calculate for small and negative values of q because the rarely reached regions carry a significant amount of the total weight. We have developed several algorithms to measure these small probabilities and their associated $D(q)$ spectra for on- and off-lattice random fractals in two and three dimensions.

1.5 Non-Equilibrium models with rare transitions: SIS, Maier-Stein, and ZGB models

The majority of rare event techniques aim to determine the frequency of rare transitions in models that satisfy detailed balance, i.e., models that have an underlying energy landscape which determines the steady-state phase space density. If the

expression, $\pi_i P_{ij} = \pi_j P_{ji}$, where π_i is the steady-state phase-space density and P_{ij} is the transition probability from state i to state j, is true for every pair of neighboring states i and j, then the system in question satisfies detailed balance. To make this idea more concrete, consider a two-state system (A and B) with π_A, π_B, P_{AB} and P_{BA}. The system starts with some arbitrary probability of being in A and otherwise being in B. After many steps, the probability of being in state A and state B approach constant values. That is, after a long time there is no net probability flux from A to B or B to A. For this to be true, the probability fluxes from A to B and B to A must exactly cancel. The probability to move from A to B is the probability of being in state A times the probability of moving from A to B, i.e., $\pi_A P_{AB}$; this must equal the flux in the opposite direction, hence, $\pi_A P_{AB} = \pi_B P_{BA}$. Detailed balance is founded in many systems, for example, any system that is in thermodynamic equilibrium satisfies detailed balance.

There are many systems with interesting rare events that lack detailed balance. Part of this thesis focuses on several such models.

1.5.1 SIS model

One of the classic models in epidemiology is the susceptible-infectious-susceptible (SIS) model. The model consists of a well-mixed population of fixed size N. Of these N individuals, I are infected with the disease and S are susceptible. Infected individuals can infect susceptible individuals at a rate $\beta SI/N$ and can recover, becoming susceptible, at a rate γI, where β and γ are infectivity and recovery rates, respectively. Because $S = N - I$, the SIS model is a one-dimensional model in I, or S. The stability of the disease is characterized by $R_0 = \beta/\gamma$. The disease dies out quickly if $R_0 < 1$ and is endemic for $R_0 > 1$. When the disease is endemic, on average

$N(1-\gamma/\beta)$ individuals will be infectious at any give time.

The fraction of individuals in state I fluctuates over time. These fluctuations can lead to the entire population becoming permanently disease-free ($I = 0$); when $I = 0$ there are no infected individuals present to make any susceptible individuals infected. In the case of endemic disease, the average time for the disease to go extinct grows exponentially with population size, making the extinction time nearly impossible to calculate for even modest population sizes, e.g., $N = 50$. Our rare event techniques can help aid this effort by accurately calculating extremely long disease extinction times as well as determining the most likely path to extinction for many epidemiological models.

There are many ways in which the SIS model can be generalized. One such generalization is to allow the population to fluctuate. This can be accomplished by adding three events to the system: allow susceptible individuals to join the population at rate μN, remove susceptible individuals at rate μS, and remove infectious individuals at rate μI [7], μ. These additional transitions change the model from an exactly solvable one-dimensional model to a two-dimensional model without an exact solution. There has been recent theoretical interest in this generalized model with several predictions including the extinction time and the most likely path to extinction. We have verified these predictions using one of our techniques [40].

1.5.2 Maier-Stein model

Maier and Stein introduced a bistable non-equilibrium model, which has received considerable interest [4]. The model consists of two coupled stochastic differential equations:

$$\dot{x} = f_x(\mathbf{x}) + \xi_x(t), \quad \dot{y} = f_y(\mathbf{x}) + \xi_y(t), \tag{1.3}$$

where $\mathbf{x} = (x, y)$ and $\mathbf{f} = (f_x, f_y)$ is a time-independent drift field:

$$f_x = x - x^3 - \alpha x y^2, \quad f_y = -\mu y(1 + x^2). \tag{1.4}$$

The white noise $\boldsymbol{\xi} = (\xi_x, \xi_y)$ is delta function correlated with a variance of ϵ:

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t+\tau)\xi_j(t) \rangle = \epsilon \delta_{ij} \delta(t - \tau). \tag{1.5}$$

For $\alpha = \mu$, the model obeys detailed balance; the force can be written in terms of a potential energy. The model has two stable points at $\mathbf{x} = (\pm 1, 0)$ separated by a separatrix at $x = 0$. This model can be thought of as an over-damped Brownian particle under the influence of some static force. A rare event in this model is a transition from one stable point to the separatrix in the low noise limit ($\epsilon \to 0$). We study this region because the theoretical predictions are only valid as $\epsilon \to 0$.

1.5.3 ZGB model

When a desirable chemical reaction naturally happens slowly, catalysts are frequently used speed-up the rate of reaction. One everyday use of catalysis is to turn carbon monoxide (CO) and molecular oxygen (O_2) into carbon dioxide (CO_2) in the exhaust of automobiles. This reaction takes place on the platinum surfaces in catalytic converters. A simple and effective representation of this reaction is the ZGB model [17]. The platinum surface is modeled as a square lattice where a CO molecule can adsorb on a single site, O_2 disassociates and sticks onto two sites, and if a CO and O are ever on neighboring sites, they desorb instantaneously producing CO_2. The only parameters in the model are the relative probability of attempting to place a CO molecule on the lattice, versus placing an O_2, (p_{CO}) and the size of the system (L by L).

The model displays surprisingly complex behavior including three distinct phases. When the probability of placing a CO and O_2 are roughly equal, the system is

reactive and produces a steady throughput of CO_2. Starting in the reactive state and decreasing p_{CO} causes the system to go through a second-order (continuous) phase transition where the entire lattice becomes covered with oxygen atoms, called the O poisoned state. A similar result occurs when the system is started in the reactive state and p_{CO} is increased. This time, there is a first-order (discontinuous) phase transition where the lattice becomes covered with CO (CO poisoned). Unlike the second-order phase transition, the first-order transition is robust to simple changes in the model, like allowing occasional reactant desorption, and has been seen experimentally.

When p_{CO} is slightly past the transition point, the system is seperated with both reactive and CO poisoned regions present. Co-existance of these phases persists until p_{CO} reaches the spinodal point, at which only the CO poisoned state is present. The same phenomenon can be seen in everyday life. A cup of undisturbed water in the microwave can be heated above its boiling point, superheating it. If the temperature is raised high enough, the cup boils instantaneously. Taking the analogy to the ZGB model, the reactive state corresponds to the liquid water and the CO poisoned state corresponds to the water vapor. Between the spinodal and transition point, "bubbles" of the CO poisoned state are formed. These bubbles have a critical size which depends on p_{CO}, above which they will tend to grow indefinitely and below which they will tend to shrink. These analogies from equilibrium physics all hold, including droplet surface tension, even though the system is far from equilibrium.

1.5.4 Organization of thesis

The organization of the thesis is as follows. Chapter 2 covers the development of our first rare event technique, signposting, and shows its application to the harmonic measure of critical percolation and Ising model clusters. Chapter 3 is a continuation

of the harmonic measure work on critical Potts model clusters and introduces the etching method. Chapter 4 discusses our last published harmonic measure project, which is the harmonic measure of 2D DLA clusters. Chapter 5 covers our measurements of the hull height and length distribution of critical Potts model clusters. Chapter 6 shows the development of our first long transition time algorithm, the barrier method, and its application to the Maier-Stein and SIS models. Chapter 7 introduces our latest algorithm, forward flux sampling in time (FFST), and discusses its application to the ZGB model of heterogenous catalysis. Finally, chapter 8 summarizes the previous results, discusses future work, and gives some guidance towards successful application of rare event techniques.

CHAPTER II

Harmonic Measure for Percolation and Ising Clusters Including Rare Events

The work in this chapter was published in Physical Review Letters in 2008 (Vol. 101, Art. Num. 144102). This chapter represents our first attempt to develop rare event algorithms. We have adapted and used the algorithm developed in this chapter, signposting, to obtain the harmonic measure of 2D and 3D DLA clusters.

2.1 Introduction

The harmonic measure is a fundamental property of geometric objects. It may be defined by considering the object to be a grounded conductor with fixed charge of unity. The harmonic measure, μ, specifies the normal derivative of the potential (a harmonic function) on the surface. That is, it is the distribution of electric field on the hull (surface) of the object. We may also allow many random walkers to start far away from the object and record where they land. The probability density of hitting the hull at a point is μ. This quantity is the focus of much theoretical activity [38, 1, 2, 41] and it is of considerable practical interest because it predicts where particles will diffuse to for adsorption, catalytic reaction, etching, etc. See also [42]. If the shape in question is fractal, μ shows interesting scaling properties.

Here we show how to find μ numerically for two systems that produce fractal clusters in two dimensions, percolation [27] and Fortuin-Kasteleyn (FK) [29] clusters in the Ising model. Our method allows us to sample very small probabilities (of order 10^{-300}) using random walker simulations. In these cases μ is *multifractal*. Our large dynamic range allows us to explore this property fully.

The function μ is non-negative and normalized on the hull: $\int d\mu = 1$. A partition function [43] can be defined by dividing the hull into j boxes of length l,

$$Z_q = \sum_j p_j^q, \tag{2.1}$$

where $p_j = \int d\mu$ over box j. For large fractal clusters Z_q scales as a power-law,

$$Z_q \approx (l/R)^{(q-1)D(q)}, \tag{2.2}$$

where R is the length scale of the cluster. $D(q)$ is called the generalized dimension. In our simulations, we choose the smallest l to be the lattice spacing, and use $2l$, $4l$, etc., until a sufficient range is available to fit Eq. (3.2). We recall some special values of $D(q)$: $D(0)$ is the fractal dimension of the support of the measure, which describes the region the hull covers. Additionally, $D(1) = 1$ is known from Makarov's theorem [44]. A related quantity is the curve $f(\alpha)$, which is the Legendre transform of $\tau(q) \equiv (q-1)D(q)$:

$$f(\alpha) = q\frac{d\tau}{dq} - \tau, \quad \alpha = \frac{d\tau}{dq}. \tag{2.3}$$

We will focus on $D(q)$ in this chapter.

The exact spectrum of $D(q)$ for percolation [1] and the more general Q-state Potts model [2] can be derived from generalized conformal invariance in terms of a central charge c,

$$D(q) = \frac{1}{2} + \left(\sqrt{\frac{24q+1-c}{25-c}} + 1\right)^{-1} \quad q \in \left[-\frac{1-c}{24}, +\infty\right). \tag{2.4}$$

Percolation and FK-Ising clusters correspond to $c = 0$ and $c = 1/2$, respectively.

Eq. (2.4) was derived for the *accessible* or *external* hull of the clusters [45, 46]. For a finite system, the external hull is approximately produced by closing all fjords on the complete hull with a neck size of order unity. This reduces the dimension for percolation and the Ising model from 7/4 and 5/3 for the complete hull to 4/3 and 11/8 for the external hull, respectively [2].

We should note that Eq. (2.4) is based on a computation about a continuum model; in principle, it might not actually apply to the scaling limit of lattice percolation. The prediction was made nearly ten years ago and has never been reliably tested in the significant small-q regime. For percolation, and for large q it agrees with results [38] on relatively small systems, $\approx 10^5$ sites. These simulations did not probe deep into the fractal surface, which is necessary for the small q regime.

The authors of Ref. [38] used the method mentioned above: a large number of walkers were allowed to diffuse until they were absorbed on the hull. This method is able to measure μ to an accuracy of $\approx 10^{-10}$. However, percolation clusters with 10^5 sites can have regions of the hull with probabilities per lattice site smaller than 10^{-100}. Although these regions do not contribute to $D(q)$ for large q, they dominate for small and negative q.

The computation of $D(q)$ is a very difficult numerical problem (as emphasized in [41]). We have solved this once and for all for arbitrary shapes; the algorithm of this chapter can measure probabilities down to 10^{-300}. This accuracy completely samples lattice systems with $\sim 10^4$ hull sites. We have applied the method on systems as large as $\sim 10^6$ hull sites. Here we consider only percolation and Ising clusters on a lattice. Our method is quite general and can also be applied to off-lattice clusters. Ref. [41] is our only real competitor for finding the complete harmonic measure;

however, it relies on a technique that is only applicable for small DLA [34] clusters.

2.2 Models and Methods

To treat rare events, we use *iterative biased sampling* to keep track of the "lucky" walkers that penetrate deep into the fjords of the hull, see Fig.2.1. In the first iteration, N random walkers with weight $1/N$ are released from outside the cluster and allowed to diffuse until they are absorbed on the hull. The weights of the walkers are temporarily added to the probability of the site where they land. This step probes the cluster to find regions of small measure. The hull sites that bound regions below a threshold (say 1/10) are used as the end points of absorbing lines (*signposts*) which mark the depth of our current sampling. Then the probability added in the first step is removed and N more random walkers are released. These walkers can either be absorbed on the hull or on a signpost. The weights of the walkers that touch the hull in this step are permanently added to the probability distribution. After all walkers have been absorbed, the signposts are removed.

Next, the probe step is repeated with N walkers released from the locations along the signposts where walkers absorbed previously. The threshold for small probability is reduced by a constant factor, e.g., 10. These walkers carry a weight given by the fraction of the walkers in the last step that touched a signpost. The method is repeated until small probabilities are sampled. We find that errors build up slowly in the method: even for probabilities of order 10^{-300} the fractional standard deviation over the ensemble is only 10-20%.

In effect, we find the Green's function for the random walkers by summing over intermediate positions. At the intermediate points where the sampling is poor, we enrich it. This is similar to methods used in chemical physics [22]. Fig. 2.2 shows

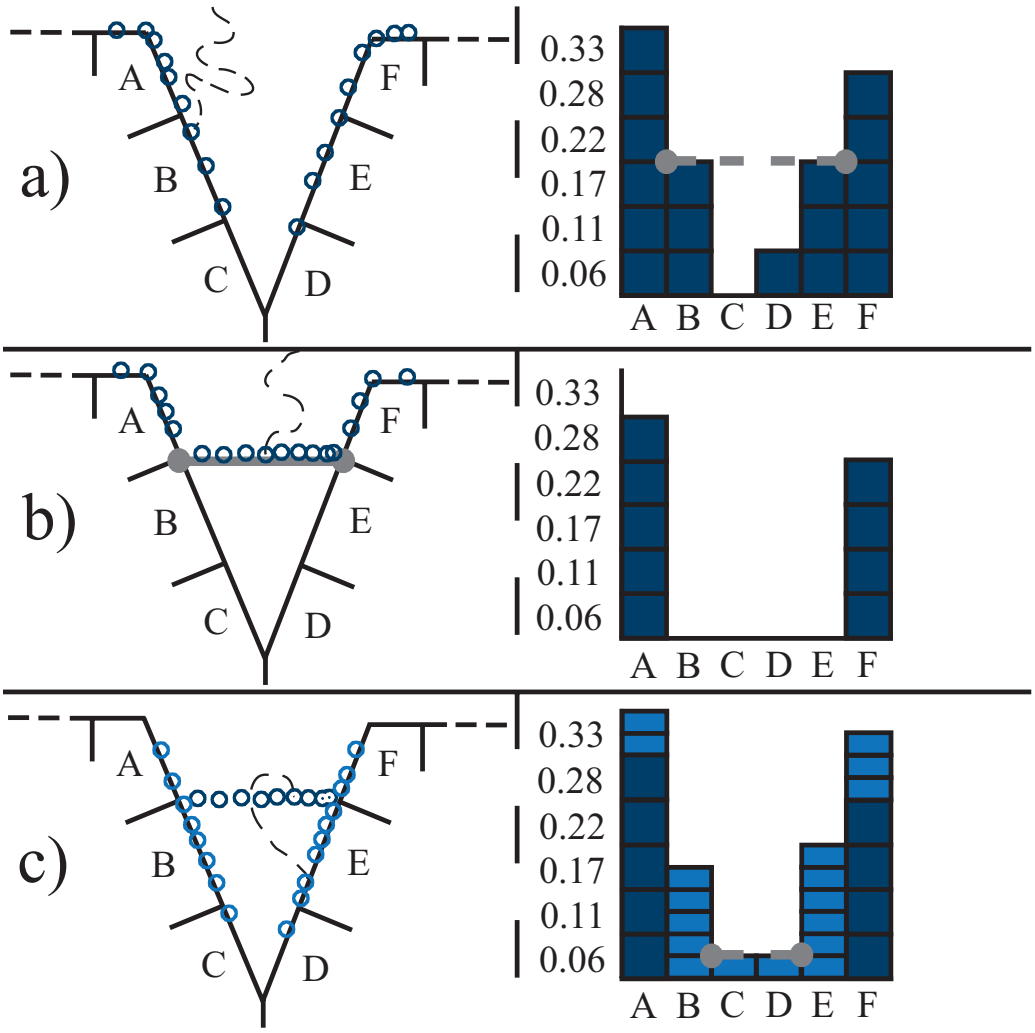

Figure 2.1: The signpost algorithm. Left: hull of the cluster divided into boxes A-F. Absorbed walkers are the $N = 18$ dots. Right: histogram of probability in each box. The initial threshold is 1/6. (a) Probe step: walkers absorb onto the hull. Below the heavy black line on the histogram $p < 1/6$. Thus, boxes B-E should be behind a signpost. (b) N more walkers are released and absorb onto the hull and the signpost (horizontal gray line). In this case, there are $N/2$ walkers on the signpost; histogram shows the probability on the hull. (c) Next probe step: N walkers are released from the signpost where walkers in step (b) landed (open circles). These walkers have half of the weight as the ones released in parts (a) and (b). The heavy line on the histogram shows the new threshold, 1/36. In the next step, boxes C, D must be behind a signpost.

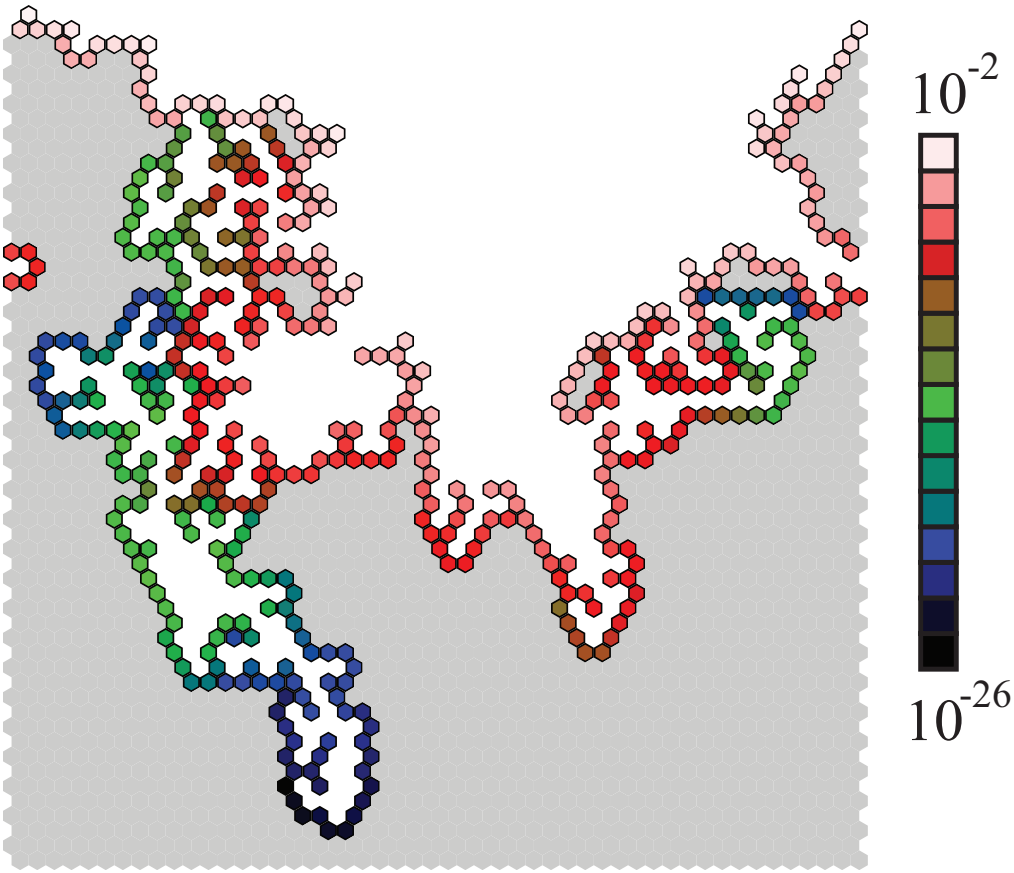

Figure 2.2: Harmonic measure on the complete hull of a percolation cluster. The hull sites are outlined in black and the harmonic measure goes from high to low; light colors are high and dark, low. The scale is given by the color thermometer on the right. Sites outside the cluster are white, and inside, grey.

the harmonic measure of the complete hull of a percolation cluster obtained using this method.

Our simulations are performed on a periodic triangular lattice with height h and width w such that $h = 100w$, so that we obtain clusters that wrap around in width but not in height. One ambiguity which must be resolved is the definition of random walkers touching the hull. Here, we interpret this as the walker hopping *onto* a hull site.

The percolation clusters are grown using the Leath algorithm [33], with p equal

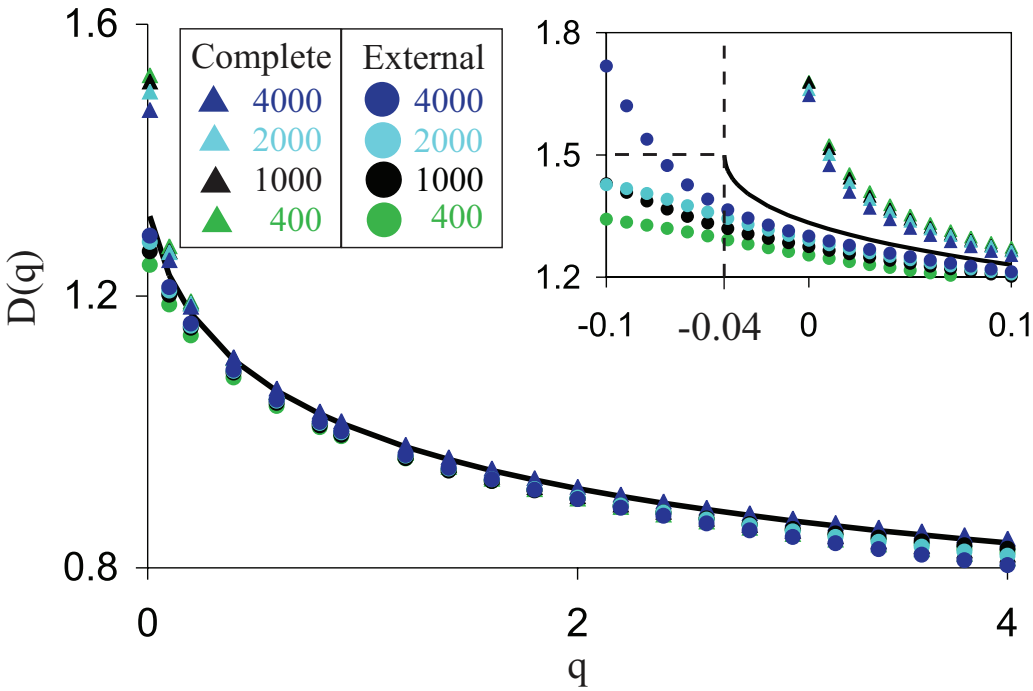

Figure 2.3: $D(q)$ vs. q for the complete and external hulls of percolation clusters for four different widths compared to theory, Eq. (2.4) [1] (black line). Inset: small q behavior of $D(q)$. Dashed lines are the theoretical limit for the external hull; the vertical dashed line is at $-1/24$.

to the site threshold for the triangular lattice, $p_c = 1/2$. If a given cluster spans the width of the system, the top hull of the cluster is found using a simple border walking algorithm related to the method of generating percolation cluster perimeters by random walks [47]. The list of complete hull sites on the cluster is then used in the signpost method to obtain the harmonic measure. If the topmost vacant sites bordering the cluster are used instead of the occupied sites as the adsorbing sites, one obtains the external hull.

For the FK Ising model clusters, bonds are placed between adjacent same-spin sites with probability $p_c = 1 - e^{-\beta J_c}$, where on the triangular lattice $p_c = 1 - 1/\sqrt{3}$. We use the Swendsen-Wang method [31] to equilibrate the system and simultaneously generate the FK clusters. After the system is sufficiently equilibrated, we attempt

to find a spanning cluster. These spanning bond clusters must be converted to site clusters if they are to be used with our algorithm. We do this by making another triangular lattice with half the lattice spacing. Bonds are copied to the new lattice at the even sites on which they are centered. Odd sites are added to the cluster if two adjacent bonds meet at the odd site. Next, the perimeter-walk algorithm is used to record the locations of the hull sites; then, we use the signpost method to find the harmonic measure. As in percolation, an external hull can be obtained. However, for Ising clusters, we need to add artificial vacancies to all sites bordering the cluster.

The signpost method iteratively obtains smaller and smaller probabilities by reducing the weight of the random walkers released in each round, in our case by a factor of 10, on average. We took the number of walkers, N to depend on the system width, w. For example, for $w = 400$ we use $N = 2 \times 10^6$ and for $w = 4000$, $N = 2 \times 10^7$. The signpost method is performed until all probabilities have been measured or until the minimum measurable probability 10^{-300} has been reached. This minimum is close to the smallest value that can be stored in a double precision floating point number. Smaller values, in principle, could be obtained by storing the logarithm of the probability instead of the probability itself.

The locations of the sites and their associated probabilities are then used to obtain $D(q)$ and the histogram of the probability distribution (see below). $D(q)$ is obtained by applying a linear fit to $\log Z_q$ in Eq. (3.1) versus $\log l$, where l is the box length. The fit was performed for a range of l over which the function was linear.

2.3 Results

Simulations of percolation and Ising clusters were performed for a number of system widths. Our results are for $w = 400$, 1000, 2000, and 4000. Small systems,

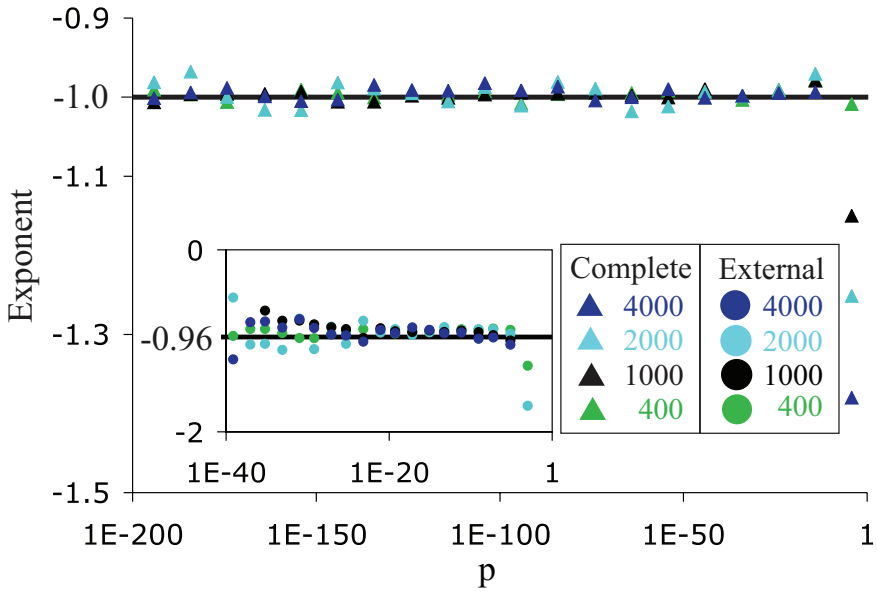

Figure 2.4: Exponent of a power-law fit to the histogram of the probabilities as a function of p for the complete hull of a percolation cluster for several different widths. Inset: a similar plot for the external hull with the associated theoretical prediction (black line) [1].

$w = 400$, have $\approx 5 \times 10^5$ hull sites in the cluster and large systems, $w = 4000$, have $\approx 5 \times 10^6$ hull sites. $D(q)$ and the slope of the power-law fit to the probability distribution were obtained for the complete and external hulls of both percolation and Ising model clusters.

Figure 2.3 shows a comparison between the results of the complete and external hulls of percolation clusters with the theory for the external hulls, Eq. (2.4). There is good agreement among all three for large q, which is not surprising as the complete hull fjords contribute negligibly to $D(q)$ in this case. For small q, there is significant disagreement between the complete hull and the theory as the two must approach different values for $D(0)$. Previous simulations [45] have shown that $D(0)$ increases with increasing width; however we see a peak at a width of 1000; see Fig. 2.3. This is because there is a non-negligible fraction of the hull sites with probabilities less than

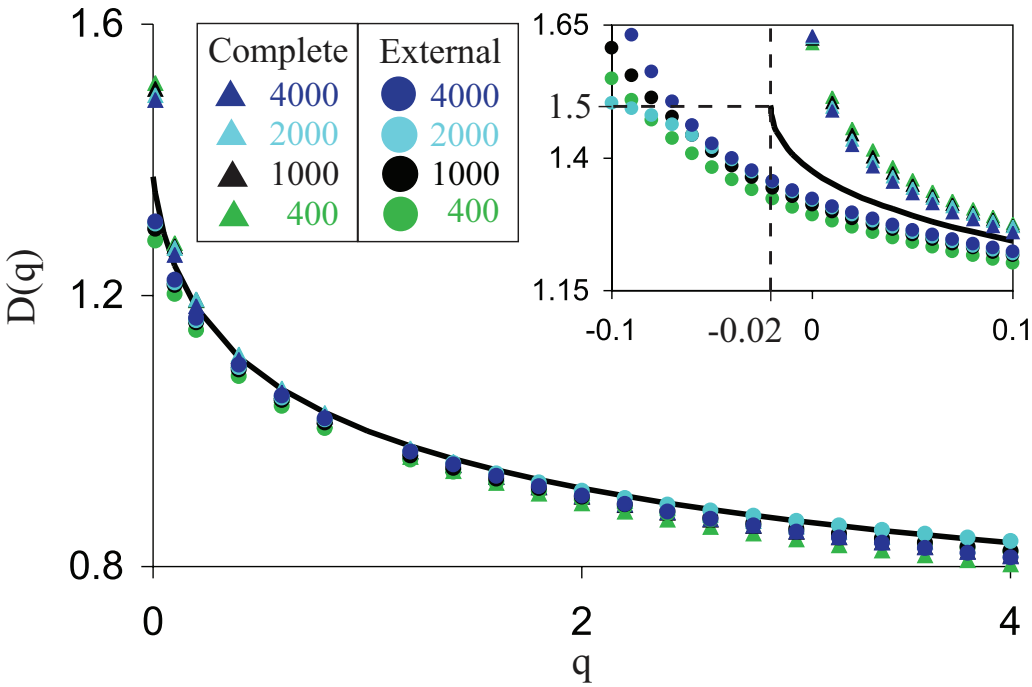

Figure 2.5: $D(q)$ vs. q for the complete and external hulls of Ising clusters for four different widths compared to theory from Eq. (2.4) (black line). Inset: small q behavior of $D(q)$. Dashed lines at $(-1/48, 3/2)$ are the theoretical limit for the external hull.

10^{-300} for large widths. We expect for very large systems, if we are able to record all probabilities, that the complete hull $D(q)$ will be nearly identical to the theory for $q > 0$ because the small probabilities do not contribute. But at $q = 0$ there will be a jump to $D(0) = 7/4$ because we are finding all of the sites. For $q < 0$, $D(q)$ will be ill-defined (unbounded). In comparison, for the external hull, we see good agreement between the data and the theory (2.4) over the entire range of $D(q)$, especially for the largest system sizes.

The histogram of the frequency of occurrence of p_j was tallied using exponentially distributed bin sizes, e.g., the first box has size 1/2, the next 1/4, then 1/8, etc. The histogram is a power law over, (incredibly) more than 150 orders of magnitude. The exponent of the power-law is fit at different probabilities using 5 points which

roughly span an order of magnitude in probability. It is shown for the complete and external hull (inset) in Fig. 2.4. The exponent for the complete hull is -0.996 ± 0.01. We presume that the exact value of the exponent is -1, which implies that $D(q)$ is undefined for $q < 0$. Previous simulations [38] were unable to obtain this result because the smallest probability that could be measured, $\approx 10^{-10}$, is still in the transient regime. The initial overshoot of the power for small systems corresponds to the probability distribution for the *external* hull being picked up by the complete hull. The power-law exponent is also obtained for the external hull, -0.93 ± 0.05, which is consistent with the theoretical prediction of $-23/24 \simeq -0.958$.

Similar results were obtained for the Ising model. Fig. 2.5 shows the comparison between the complete and external hulls of Ising clusters with the theory [2] for $D(q)$. As with percolation, there is good agreement with theory for large q for both the complete and external hulls but significant disagreement at small q for the complete hull, where Eq. (2.4) does not apply. The probability power-law exponents for the complete and external hull are, -0.997±0.012 and -0.920±0.048, respectively for the Ising model. The complete hull exponent again points to $q = 0$ as the discontinuity point for $D(q)$. The external hull exponent agrees roughly with theory which gives $-47/48 = -0.979$.

2.4 Conclusions

In summary, we have described a method to obtain precise values of $D(q)$ by including events of extremely low probability. We probed the internal structure of percolation and Ising model complete cluster hulls. We observe the histogram of occurrences of probability p to be $\sim p^{-1}$. We are not aware of any prediction of this case.

In future work [36], we plan to apply the continuous version of this algorithm to obtain the harmonic measure for Diffusion Limited Aggregation (DLA) [34] for which there are no exact results, though there are several conjectures for the form of $D(q)$ for small and negative q [41]. For DLA, the harmonic measure plays a central role because it represents the growth probability at every point on the cluster at a given time. The best current results for $D(q)$ use iterative conformal maps [48, 49], and are restricted clusters of $\approx 10^4$ sites. Our method can go to much larger sizes, $\approx 10^7$ sites. This is important because the slow crossover of some length scales in DLA [50] suggests large clusters are necessary to approximate the scaling limit. Our method could shed light on the internal structure of DLA about which little is known.

CHAPTER III

The Harmonic Measure for Critical Potts Clusters

The work in this chapter was published in Physical Review E in 2009 (Vol. 80, Art. Num. 031141). This chapter is the followup article to the letter in chapter II, in which only the first two Q states in the Q-state Potts model were studied. Here we obtained the harmonic measure for the first four Q states. We also developed two new harmonic measure algorithms: Etching and the Green's function method. Both of these algorithms influenced our later work: we used etching on 3D percolation clusters, and the Green's function method inspired our work on the barrier method.

3.1 Introduction

3.1.1 Potts model

The Q-state Potts model, a generalization of the Ising model to Q different spins, has been the subject of considerable interest [26]. Two important cases are $Q = 1$ and $Q = 2$, which correspond to percolation [27] and the Ising model, respectively. When a Potts system is prepared at its critical temperature, subsets of the clusters of like spins, the Fortuin-Kastelyn (FK) clusters [51, 29] (to be defined below), are self-similar fractals [14]. For $Q = 1$ the FK clusters are the same as the usual percolation clusters. In this chapter, we will study the harmonic measure of the hulls of these

fractal clusters for $Q = 1, 2, 3, 4$.

The harmonic measure may be thought of as the distribution of the surface electric field on a charged conductor. Since the Laplace equation and the steady-state diffusion equation are identical in form, the harmonic measure is also equal to the distribution of probabilities of random walkers diffusing far from the cluster onto a given section of the hull. In this chapter, we use a biased random-walk sampling technique to obtain the harmonic measure. We also review other methods for measuring small probabilities and give details of our algorithms.

The harmonic measure is of practical interest because of its relation to the anomalous frequency dependence of the impedance of rough electrodes [42] and because of its obvious connection to processes that involve absorption of diffusing particles such as catalysis [52]. It has a deep connection to the structure of diffusion-limited aggregates (DLA) [34], since the harmonic measure determines where each walker will land; that is, for DLA it is the growth probability. In the case of critical Potts clusters and DLA, the harmonic measure is multifractal [39]. Advances in conformal field theory and Schramm-Loewner evolution have brought about renewed interest in the harmonic measure. In particular, certain aspects of the harmonic measure for Potts clusters can be computed in the continuum limit using these methods [1, 2, 53, 54, 55, 56].

Numerical investigation of the harmonic measure of percolation [38] and DLA [38, 57, 58] clusters is difficult because the measure has a huge dynamic range for systems of even moderate size. In refs. [38, 57, 58] one of two methods were used: the first is the straightforward one of releasing a large number of random walkers far from the cluster and determining where they land. The second uses relaxation or equivalent algorithms to solve the Laplace equation. The random walker method can

only measure probabilities down to about 10^{-10} and samples a very small part of the measure for clusters of reasonable size. Relaxation like methods are computationally costly and limited to small clusters and give similar lower limits on the probabilities that can be measured.

For DLA it is possible to go to much smaller probabilities by using the method of iterated conformal maps [48, 59, 60]. However, this technique is only capable of treating moderate size clusters [41]. In an earlier chapter we generalized the random walker method and gave a technique capable of obtaining probabilities down to 10^{-300} for any fractal. We applied it to FK clusters for percolation and the Ising model [61]. This chapter describes a further development of those techniques.

3.1.2 Generalized dimensions

The harmonic measure, the distribution of probabilities that random walkers will hit a given site on the perimeter of a cluster, is very complicated and varies wildly for the cases we are studying; see Fig. 3.1. A popular and useful way to characterize it is in terms of the generalized dimension, $D(q)$, of the measure. We define these objects as follows: we cover the hull with boxes of length L. With each box we associate a probability, p_i, which is the sum of the measure over the sites within the box. We then define a function $Z_L(q)$, sometimes called the partition function:

$$Z_L(q) = \sum_i p_i^q \tag{3.1}$$

where q is some power [43]. If the object in question is fractal, then the partition function will follow a power-law in L:

$$Z_L(q) \sim (R/L)^{-\tau_q} = (R/L)^{-(q-1)D(q)} \tag{3.2}$$

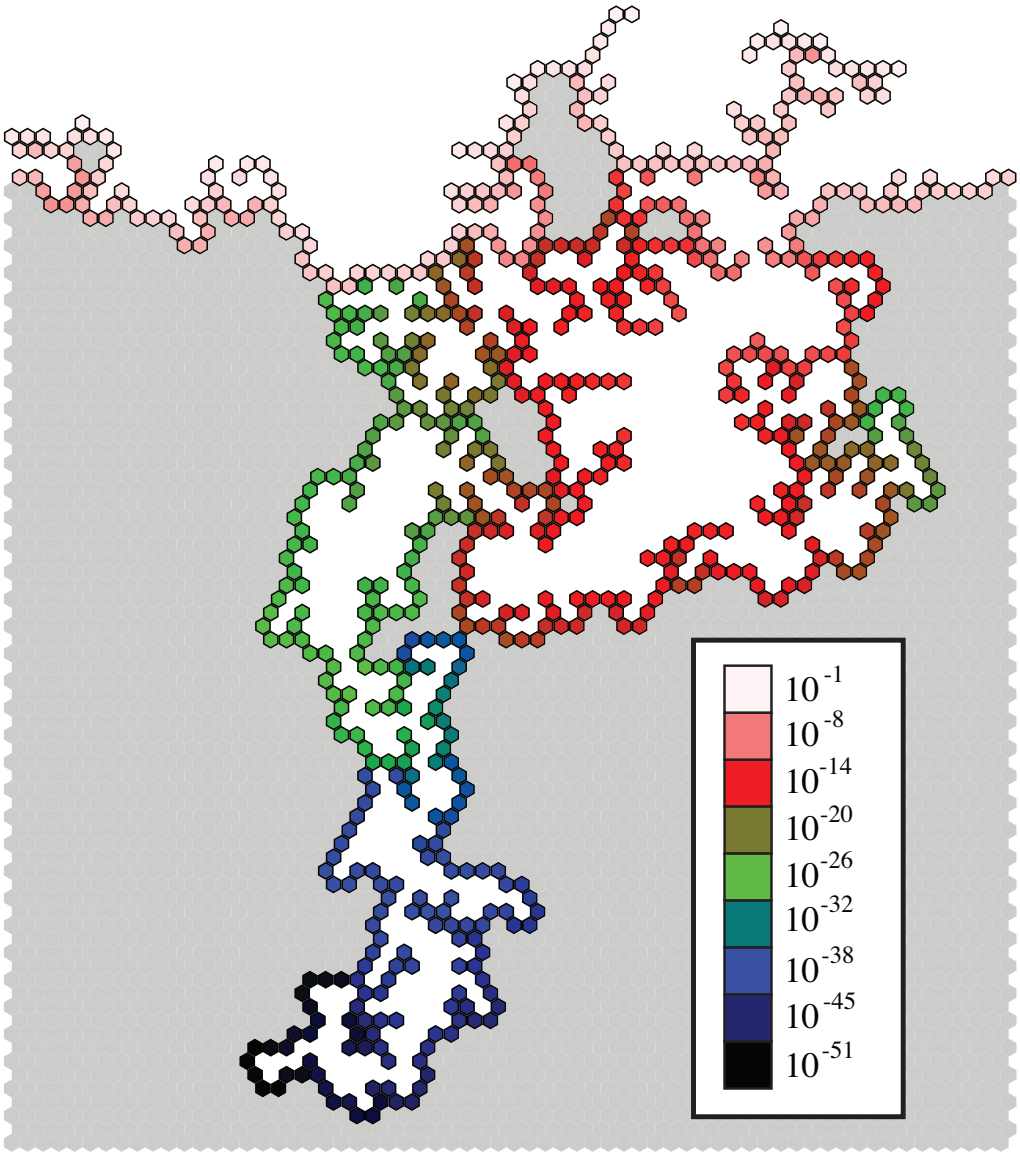

Figure 3.1: The harmonic measure for the complete perimeter of a small, $W = 64$, percolation cluster. The solid grey regions represents the area that is inaccessible to the random walkers diffusing from above the cluster. Every perimeter site is colored according to its measure. The computation was performed using the etching method described below. Note that the measure on this small cluster spans 50 orders of magnitude.

for $(R/L) \to \infty$, where R is the size of the cluster. For integer q, $D(q)$ corresponds to the fractal dimension of the q-point correlation function. There are special values of $D(q)$. $D(0)$ is the fractal, box-counting, dimension of the hull. Also $D(1)$, the information dimension, is always unity by Makarov's theorem [44]. A related function is the singularity spectrum $f(\alpha)$, the Legendre transform of $D(q)$:

$$f(\alpha) = q\frac{d\tau}{dq} - \tau, \quad \alpha = \frac{d\tau}{dq}. \tag{3.3}$$

In this chapter, we will focus exclusively on $D(q)$. The singularity spectrum can be derived from our results using Eq. (4.2).

3.2 Models

3.2.1 Simulations of FK clusters for the Potts model

We produce critical Potts clusters in two ways. For percolation, we use the Leath algorithm [33]. The algorithm starts with a single active site; we attempt to turn its neighbors into active sites with probability p. If a conversion attempt fails, the site is labeled inactive. The process is repeated with neighbors of the active sites which have not been labeled as inactive. The process continues until there are no new active sites. If p equals p_c, the percolation threshold, a critical percolation cluster is produced. The outer layer of active sites is called the complete perimeter. Its fractal dimension is denoted D_H. The cluster of active sites is surrounded by a single layer of inactive sites; this layer is called the accessible (or exterior) perimeter [45] and has a fractal dimension denoted D_{EP}. The accessible perimeter is of interest because, unlike the complete perimeter, it is expected to have a well-behaved limit when clusters are very large and are rescaled. The harmonic measure has been determined in this limit for the accessible perimeter of Potts clusters [1, 2].

To obtain critical Potts clusters for $Q = 2$, 3, and 4, we grow equilibrated FK clusters using the Swendsen-Wang (SW) algorithm [31]. For any configuration of spins, FK clusters are subsets of clusters of like spins formed by a bond percolation process. That is, we consider the clusters formed when adjacent spins are connected with probability $p_c(Q) = 1 - \exp[-K_c(Q)]$ where $K_c(Q)$ is the critical coupling constant. For $Q = 2, 3, 4$, on the triangular lattice, $p_c(Q)$ is known to be $1 - 1/\sqrt{3}$, $1 - 1/[1 + \frac{1}{2}\sqrt{3}\sec(\pi/18)]$, and $1/2$, respectively [62]. To obtain the equilibrium ensemble of FK clusters we iterate two steps until the system settles down (see below). The first step takes every current FK cluster and replaces the spin with one of the Q possible values, at random. In the second step, the bonds connecting the clusters are discarded and bond percolation is performed again, with $p = p_c(Q)$, on all neighboring sites with the same spin. The process is then repeated by updating the spins on the newly formed clusters. These two steps together constitute a spin update.

3.2.2 Parameters and Observables

We grew critical Potts clusters for $Q = 1-4$ on the triangular lattice, as described above. We chose to use a triangular lattice rather than a square lattice because the square lattice does not allow diffusion into fjords bounded by diagonal entrances. We use the width of the system, W, as the characteristic length. The clusters we want span in the width direction but not in the height direction. To make sure the clusters will only span in one direction, we chose very large aspect ratios. The height of the lattices were $100W$ and $8W$ for $Q = 1$ and $Q > 1$, respectively. We looked at six different system widths, $W = 128, 256, 512, 1024, 2048$, and 4096. Because FK clusters are intrinsically bond clusters, we needed to use a trick to turn them into

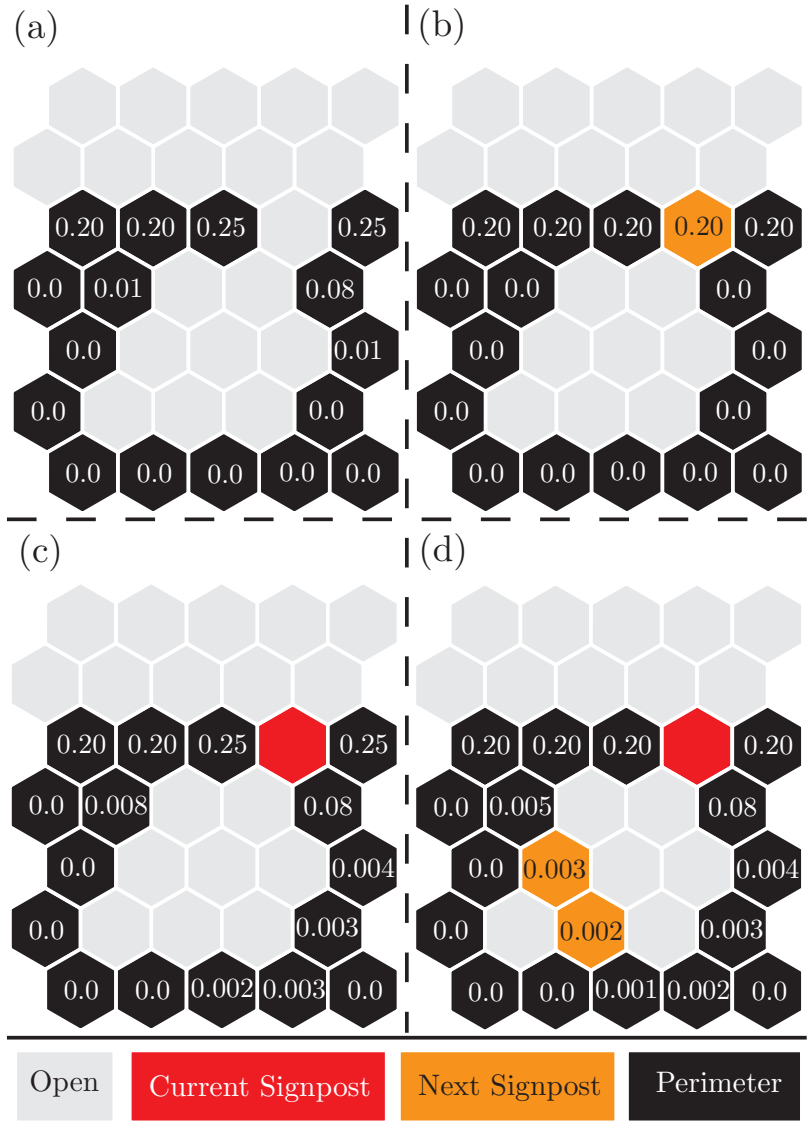

Figure 3.2: The signpost method. The system is periodic in the horizontal direction. (a) First (probe) step: N random walkers are released from the top row and absorb onto the perimeter sites. (b) We choose the first probability threshold as 0.1. Using this threshold, we connect the the bounding sites using signpost sites. In the second (measurement) step we send N more random walkers from above which can absorb onto the signpost or perimeter sites. (c) Second probe step: the random walkers are launched from the signpost sites in the previous measurement step. The walkers released in this step have a weight of p/N where p is the fraction of the random walkers that hit the signpost site in (b). (d) The second threshold, 0.01, is used to determine the location of the new signpost sites.

site clusters. We created a lattice twice as dense as the original and marked every site at the center of a bond and every site where two bonds meet as cluster sites. The FK cluster widths used were $W/2 = 64, 128, 256, 512, 1024, 2048$.

To have proper FK clusters we require equilibration in the SW algorithm. We numerically determined that the equilibration time for $Q = 2$, 3, and 4 is of the order of W spin updates by looking at the relaxation of the average energy per spin and the average largest cluster size. For $Q = 2, 3$ and small W we ran a separate simulation to equilibrium for each spanning cluster which was added to our ensemble. For $Q = 2$ and 3 and $W = 2048$ and 4096 and for all of the $Q = 4$ clusters, the equilibration time was too large to proceed in this way. In these cases we equilibrated the system once and recorded an ensemble of spanning clusters as the simulation proceeded. We conservatively estimate the correlation time as 50 spin updates for all W and Q. This means we recorded a spanning cluster every 50 spin updates.

For each system size we grew a number of clusters. For all Q our ensemble was 2000, 2000, 1000, 1000, 400, and 100 clusters for $W = 128$, 256, 512, 1024, 2048, and 4096, respectively.

3.3 Measuring small probabilities with random walkers

3.3.1 Previous Methods

Small probabilities in the harmonic measure correspond to very unlikely paths. As the simulation proceeds we can think of the event of a random walker landing where the measure is very small as a rare event. Thus, computing small probabilities is a similar task to finding the rate of a rare chemical reaction [12], a rare extinction of a disease [10] or a population [11], or the failure of a queuing system via queue over-flow [8].

Accelerated numerical methods for these problems often involve biased event sampling. The sampling can frequently be cast as a random walk, either through state space or in our case, physical space. For example, one could ask what is the probability that a random walker starting halfway up a hill will successfully climb up to the top before sliding down to the bottom. If the hill is steep, it could be impossible to directly sample the probability to climb the hill. One could place barriers uniformly on the hill, which when crossed by the random walker, will split the random walker into two walkers, each with equal weight which add up to the original weight of the walker. This will aid sampling of the events higher up on the hill. This method is called "splitting" and effectively performs importance sampling [18]. One significant drawback of splitting is that if the barriers are too densely or sparsely spaced, the number of random walkers will tend to diverge or extinguish, respectively.

The methods we detail in this chapter are related to the splitting method, but differ in that our methods do not have the possibility of diverging or extinguishing. Another popular method called "milestoning" [22], does not have a divergence problem, but does require the system studied to be in equilibrium and the location of the barriers to be known *a priori*, whereas our method works for equilibrium and non-equilibrium systems and the barriers are placed "on the fly."

3.3.2 Signposts

We have developed several accelerated methods for the harmonic measure problem. The motivation, as we have stated, is that it is usually impossible to send in enough random walkers to directly obtain the harmonic measure: the clusters will frequently have regions with probabilities of being hit that are smaller than 10^{-100}.

It would require of order of 10^{100} random walkers to sample this region; such a computation is clearly impossible.

We now review the first method we developed, the signpost method [61]. The signpost method consists of two steps which are applied iteratively; see Fig. 3.2. In the first (probe) step we release N diffusing random walkers far from the cluster to determine which regions are rarely visited in straightforward sampling. Next, we block off all poorly sampled regions with signposts (absorbing lines). In the second (measurement) step, N more walkers are released far from the cluster and either absorb on the cluster (or the accessible perimeter) or onto the signposts. The walkers sent in this step have their weight permanently added to the harmonic measure of the perimeter sites where they landed. In the next probe step, the walkers are released from the points on the signposts where the walkers in the previous measurement landed. The new walkers have a weight of p/N, where p is the fraction of random walkers that absorb onto signpost lines in the previous step, to conserve probability. The probe step again helps determine which regions are still poorly sampled, which are subsequently blocked off. Next, another measurement step is performed. This process is repeated until all regions are explored by the random walkers. This algorithm can be applied to on- and off-lattice clusters.

We should note some things about this method. First, one must determine the entire perimeter of the cluster at the beginning of the computation in order to figure out how to block poorly sampled regions. Also, one needs to choose a rate to reduce the threshold for calling areas "poorly sampled" in each iteration. In [61], we moved the threshold down by a power of 10 each iteration, whereas in [36], we reduced it as a function of how many walkers hit the signpost in the previous iteration. When more walkers hit the signposts we moved them even deeper. The second method gave more

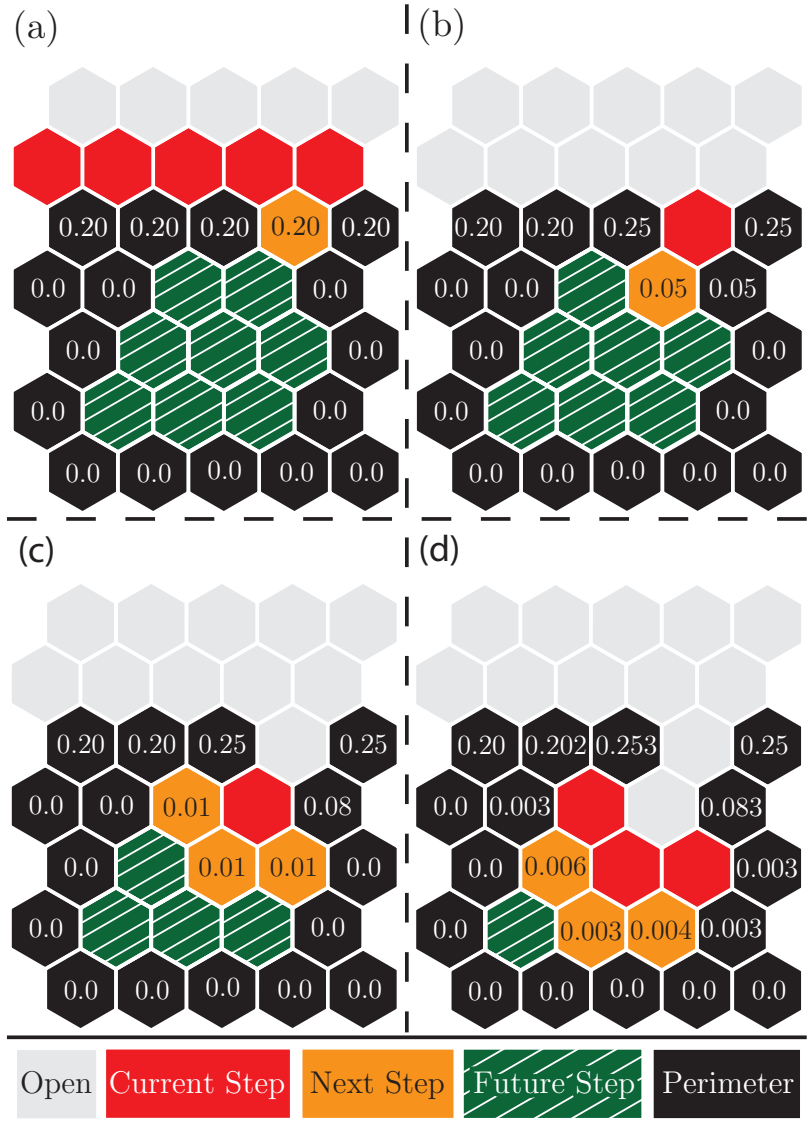

Figure 3.3: The etching method. Walkers are released from the current level sites. The next level of soft sites absorb walkers; they are then relabeled as current level sites. Future sites are all sites which will eventually become current level sites. The first round of random walkers are launched from the row above the cluster, (a). The weight of all of the walkers released is $0.2/N$, where N is the number of walkers released per current level site. 20% of the walker weight is deposited onto the top row of perimeter sites and the next level soft site, which will release N walkers in the next step. (b). One more perimeter site is accessible to the random walkers and 5% of the weight is deposited on the site in the next level. (c) Three sites in the next level each absorb 1% of the walker weight. (d) Due to the reduced weight of the walkers released in the next step, small probabilities are measured on the newly exposed perimeter sites.

consistent walker saturation, which should lead to a slower compounding of error. It is important to note the signpost algorithm is only practical for two-dimensional problems. For higher dimensions, one would need to define signpost *surfaces* to block poorly sampled regions. This is could be very complex for a complicated cluster.

3.3.3 Etching

We now describe the method we use here which we call "etching." Consider the hull of FK clusters grown on a triangular lattice with periodic boundary conditions. We want to find the harmonic measure of the top perimeter from above. To do this, we start by marking all sites that are exterior to the cluster from above as *soft* sites; the soft sites are absorbing like the cluster (or accessible perimeter) sites. The highest row is limited to one level above the highest point on the perimeter.

We next relabel every site on that highest row as a *current level* site; these are not absorbing. We release N random walkers, each with weight $1/(NW)$, from each current level site. The walkers released from these sites are allowed to walk until they deposit their weight onto a soft site or a perimeter site. If they move one level further away from the cluster, they are immediately moved back onto the current level sites using a Green's function which must be determined in advance. However, this is rather simple since it is the Green's function to return to a plane from one site above the plane. This Green's function is used for the entire simulation and limits a walker to backtrack to at most one level above the cluster. After all random walkers are released, the labels on each current level site are removed and every soft site hit in the previous step is labeled as a current level site. From each current level site i we release N random walkers with weight p_i, where p_i is the amount of probability deposited on the site in the previous step divided by N. This process is repeated

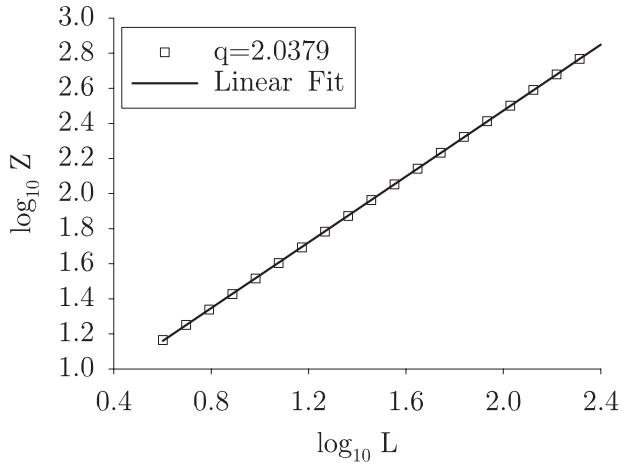

Figure 3.4: An example of the fit of $\log_{10} Z(L,q)$ versus $\log_{10} L$ to a straight line for $q = 2.0379$. The behavior is similar for all q values that we have examined. The slope of the line is $\tau(q) \equiv (q-1)D(q)$.

until there are no more soft sites. See Fig. 3.3.

Etching can be thought of as the limit of the signpost method with the signposts spaced one site apart. However, etching has several benefits over the signpost method. First, the entire perimeter of the cluster does not need to be mapped out before we start. Both algorithms have the same time complexity, $O(W^3)$ for the complete perimeter of Ising clusters, and both methods have similar memory requirements. In contrast to the signpost method, the etching method can be easily generalized to higher-dimensional lattice problems and networks. We have successfully used etching to obtain the harmonic measure of three-dimensional percolation clusters [63].

3.3.4 Green's functions

We have also developed a rare event method which may be significantly more efficient than etching and signposting for some problems. Thus far we have applied this method only to simple test problems. This method manipulates probabilities

directly and does not allow backtracking of probability. To do this, we calculate the Green's function $G(i,j;k,l)$, i.e., the probability to move to any of the sites i,j in the next level from a given site k,l in the current level.

To illustrate our algorithm, consider finding the probability distribution in a channel with absorbing walls on a square lattice. The initial condition is that the probability is uniformly distributed among the sites in the first row of the channel and the zeroth row is a reflecting boundary. All sites that initially have probability are denoted by C. The previous level sites, absorbing sites, and next level sites accessible to the current level sites are denoted by B, A, and N, respectively. (Initially, the previous level is the reflecting boundary.) In each iteration, the goal is to move all of the probability from each current level site to the all the next level and absorbing sites.

We find the Green's function by iteration on an index s. The process begins for some current level site, k,l; $(k,l) \in C$. Initially, probability only resides at k,l so that for $s=0$, $G^s(i,j;k,l) = \delta_{i,k}\delta_{j,l}$. In each iteration, the probability is moved to each of the current level site's neighbors,

$$G^{(s+1)}(i,j;k,l) = \sum_{(m,n)} W(i,j;m,n) G^{(s)}(m,n;k,l), \tag{3.4}$$

using the jump probability,

$$\begin{aligned} W(i,j;m,n) &= \frac{1}{4}(\delta_{i,m+1}\delta_{j,n} + \delta_{i,m-1}\delta_{j,n} + \delta_{i,m}\delta_{j,n+1} \\ &+ \delta_{i,m}\delta_{j,n-1}) \quad (m,n) \in C \\ &= G_B(i,j;m,n) \quad (m,n) \in B \\ &= \delta_{i,m}\delta_{j,n} \quad (m,n) \in A \cup N \end{aligned} \tag{3.5}$$

Here $G_B(i,j;m,n)$ is the Green's function for the previous level, see below, and the last line represents the probability of staying at absorbing and next level sites.

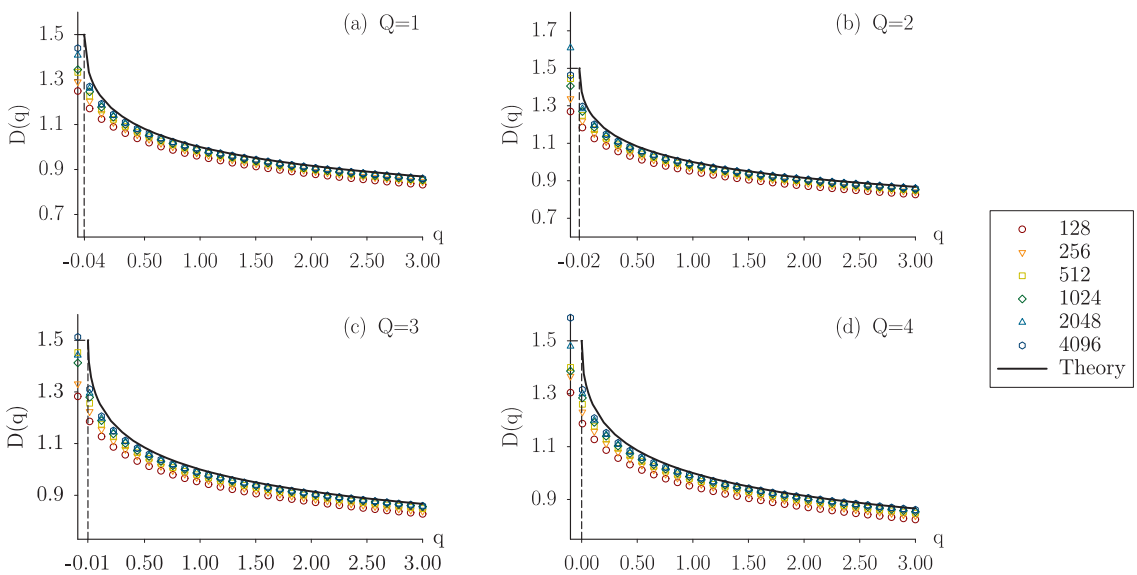

Figure 3.5: The $D(q)$ spectrum for the accessible perimeters of $Q = 1, 2, 3, 4$ clusters, in (a), (b), (c), and (d), respectively. The solid lines are the theory of [2] and the symbols are the results of our simulations for several system widths. The vertical dotted lines marks q_{min} for the theoretical spectra for infinite systems.

$G_B(i, j; m, n)$ takes into account all the processes that would correspond to random walkers backtracking before the previous level. To start the process, the reflecting boundary has $G_B(i, j; 0, n) = \delta_{i,1}\delta_{j,n}$.

For large s, virtually all of the probability will be on absorbing sites and next level sites. In any finite amount of time, some slight probability will remain in the current level, so after some stopping criteria is met, the probabilities recorded on the absorbing and next level sites must be normalized. When this has been achieved, we have the Green's function from a given site in the current level, k, l, to any site in the next level, i, j:

$$G_B(i, j; k, l) = \lim_{s\to\infty} G^{(s)}(i, j; k, l). \tag{3.6}$$

In the next step, this G_B will be used as a jump probability.

This process is repeated for all current level sites so that Green's functions from those sites to the next level sites and absorbing sites are calculated. With these Green's functions, it is easy to determine where the probability from the first level will end up. If the probability in the starting level is $P(k,l)$, then the probability in the next level is,

$$P(i,j) = \sum_{(k,l)\in C} G_B(i,j;k,l)P(k,l) \tag{3.7}$$

Note that (i,j) can be absorbing sites as well as next level sites.

The next step is to relabel all current level sites as previous level sites, relabel all next level sites as current level sites, and mark all sites that are accessible to the new current level sites (which are not previous or absorbing sites) as next level sites. Then the process is repeated.

The end result of this process is that all of the original probability is at absorbing sites, as it would be using signposting or etching. Although this example contained only sites that were completely absorbing or non-absorbing, the Green's function method can easily be generalized to partial absorption problems.

The Green's function method is somewhat more complex to program than the etching method and the simplest implementation involves setting up the Green's function look-ups in sparse arrays. This leads to a memory complexity which grows like W^{2d}, where d is the dimension of the space. The memory complexity would significantly reduce its usefulness, as it would take at least one terabyte to store a two-dimensional cluster with a length scale of 1000 lattice sites. However, it is possible to store the Green's function lookup in an associative array; this reduces the memory complexity to W^{d-1+D}, where D is the fractal dimension of the perimeter. For the external perimeter of two-dimensional percolation clusters the memory complexity grows like $W^{7/3}$, which is quite close to the memory complexity for etching, W^2. For

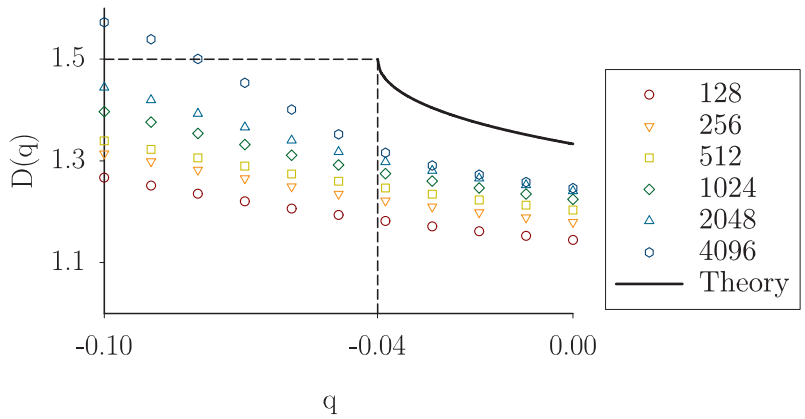

Figure 3.6: The $D(q)$ spectrum for the accessible perimeters of $Q = 1$ clusters for small q. As the system size increases the simulated values increase, presumably to approach infinity for $q < -1/24$.

a cluster with a length scale of 1000 sites, the minimum required memory would be about ten megabytes for the Green's function method.

3.4 Results

We used etching to find the harmonic measure of Q-state Potts model clusters. We analyze the measure by producing $D(q)$ spectra and histograms of the probability distributions. To obtain $D(q)$, we start by sectioning individual clusters into boxes of length L as described above. Because we are using a triangular lattice, it is convenient to use a parallelogram aligned with the lattice as a box. After completely tiling the cluster with boxes, we define the probability within a box $p_{i,L}$ as the sum of the measure of perimeter sites within the box. We then calculate $Z(L, q)$ using Eq. (3.1). $D(q)$ is related to $Z(L, q)$ by $(q - 1)D(q) = m$, where m is the slope of $\log Z(L, q)$ versus $\log L$.

We found that for a given Q and q, all system sizes have similar local slope behavior over a range of L; see Fig. 3.4. In order to average over the ensemble

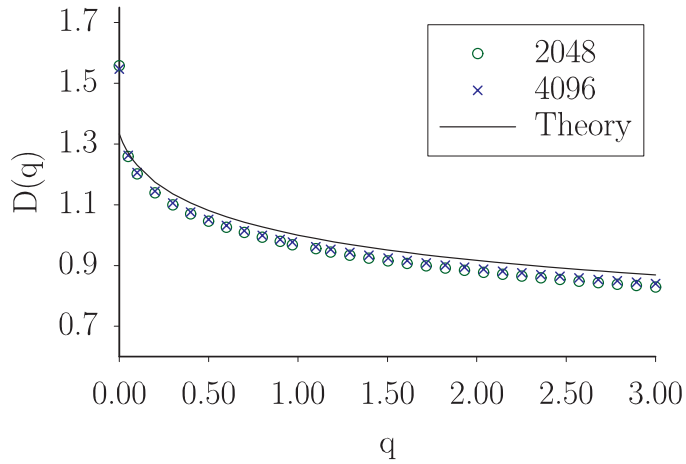

Figure 3.7: The $D(q)$ spectrum for $Q = 1$ for the complete perimeter. There is no theoretical prediction for this quantity. However, for q substantially bigger than 0 we expect this result to be very similar to the result for the accessible perimeter since large probabilities will dominate the sum in Eq. (3.1). The line labeled "theory" is for the accessible perimeter.

we average $\log Z$. However, if we use the slopes for each individual member of the ensemble and average them we get virtually identical results.

The spectra of generalized dimensions for the external hulls of $Q = 1 - 4$ are given in Fig. 3.5. In all cases the results are close to the theoretical predictions [2]. The theoretical predictions include a divergence of $D(q)$ for $q < q_{min}$ for an infinite system, see below. Our simulation results increase rapidly with W for this regime, as expected; see Fig. 3.6.

For completeness, we include the spectrum of generalized dimensions for the complete perimeter for the case $Q = 1$; see Fig. 3.7. There is no theoretical prediction for this quantity. For positive q the results are close to those of the accessible perimeter shown in Fig. 3.5. This is because, for positive q, large probabilities contribute most of the weight in $Z(q)$. Near $q = 0$ the two spectra differ because there are significantly more sites with small measure for the complete hulls.

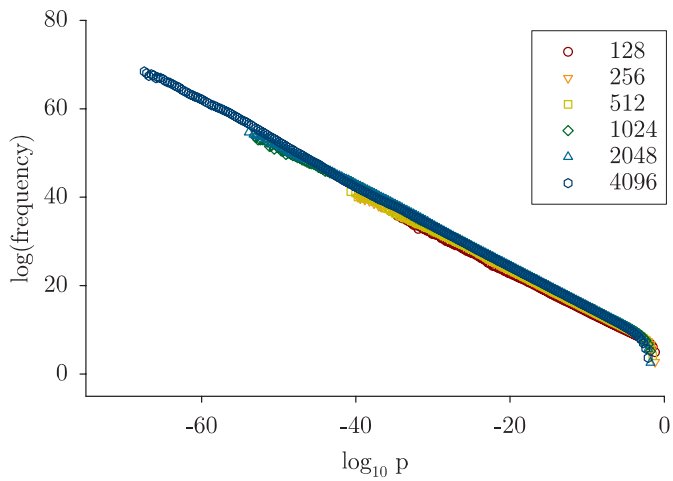

Figure 3.8: The histogram of the frequency of occurrence of the values of p for the accessible perimeter for $Q = 1$. The points for various values of W are superimposed.

We also considered the distribution of the values of p directly, by making histograms of its frequency for all Q and W. The histograms turn out to be power laws with negative powers near -1; for an example see Fig. 3.8. Since the histogram is very accurately a power-law in p, it is useful to plot the local slope of the histogram, which is shown in Fig. 3.9 for the accessible perimeters for $Q = 1 - 4$. We also show the local slope for the complete perimeter of $Q = 1$; see Fig. 3.10. The slope is calculated over about 10 orders of magnitude in p for the accessible perimeter, and more than one order of magnitude for the complete perimeter.

The significance of the slope is that it gives information about the non-scaling aspects of the distribution, and, in particular, the value of q_{min} mentioned above. If we call the slope of the histogram $-\phi$ (so that ϕ is a positive number) we see that the partition function of Eq. (3.1) formally diverges if $q < \phi - 1$, or, said another way, we expect $D(q)$ to be undefined for $q < q_{min} = -1 + \phi$. This means that the partition function is dominated by a few instances of very small probabilities which

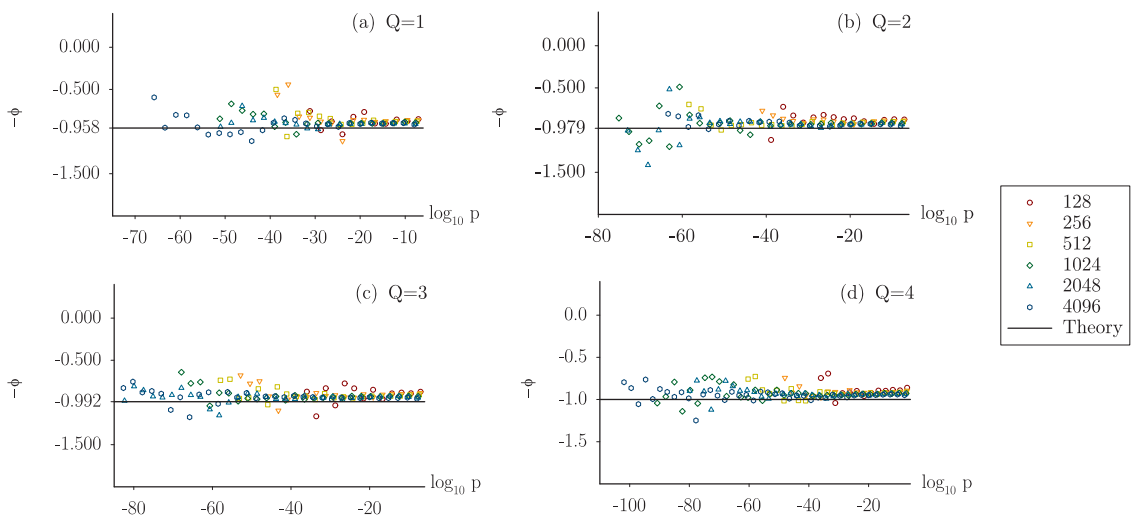

Figure 3.9: The local slope of the histogram of the frequency of occurrence of the values of p for the accessible perimeter for $Q = 1, 2, 3, 4$ in (a), (b), (c), and (d), respectively. Also shown (solid lines) are the theoretical predictions of the local slope from [2]. Note that in (b) the smallest probabilities recorded were not from the largest system size, but were from $W = 1024$. This can be understood by the fact that ten times as many clusters were generated for $W = 1024$. That is, among the many samples at $W = 1024$, a few abnormally deep clusters were recorded which happened to have the smallest probabilities.

does not scale as power law in R/L. The values for the limit of the spectrum agree well with the predictions of Duplantier [1, 2]; see Fig. 3.9. Note that the slopes are very nearly constant over about 40 orders of magnitude in p.

The slopes for the complete perimeter of percolation clusters are also constant over many orders of magnitude; see Fig. 3.10. In this case, we find that ϕ is very close to 1, and the limit of the spectrum is at $q_{min} = 0$. There is no theory for this case and no explanation for this intriguing result.

3.5 Error estimate

Since etching involves sampling the probability, there will be errors due to the finite number of random walkers released at each step. For the results in this chapter, we released 10^3 random-walkers per current level site for all system widths and Q values.

We can estimate the sampling errors as follows: we considered one percolation cluster with $W = 2048$ and made 10 independent computations of the p_i. The variance of the probability over this sample at a given point on the cluster, δp_i, is a measure of the reliability of the measurement. In our case, we found that some points have a rather large percentage error, though always less than a factor of 3, but the average over all the points, $\langle \delta p_i / p_i \rangle$, was 23%. Note that the very small probabilities well inside the cluster have very small errors. There is no build-up of the error as we etch toward the interior, as might have been expected.

If it is necessary to reduce the error further, more random walkers can be used. However, we believe that the ensemble averaging that we did means that the generalized dimensions are much more accurate than the individual probabilities. Our evidence for the last statement is the good quality of the fit in Fig. 3.4, and the closeness of the results in Fig. 3.5 to theory. Note also that $D(0)$ is close to the known fractal dimensions of the exterior perimeters.

3.6 Conclusions

In this chapter, we presented the etching method, a new accelerated technique for computing the harmonic measure. We are able to measure probabilities as small as 10^{-4600}. We showed how this method relates to other methods. We used etching to obtain the harmonic measure for the accessible perimeter of FK clusters for the

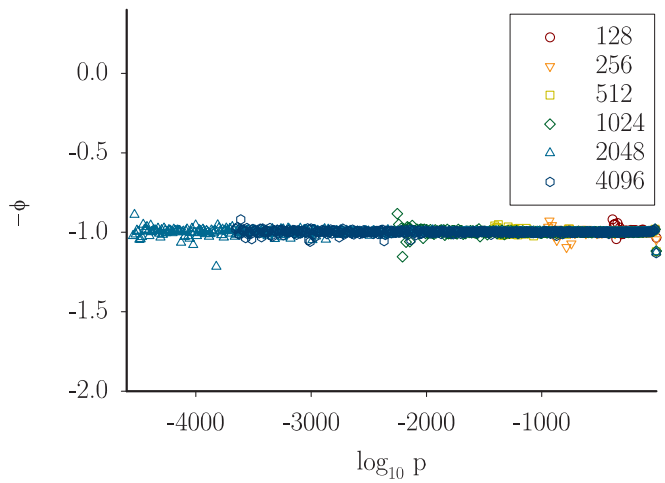

Figure 3.10: The local slope of the histogram of the frequency of occurrence of the values of p for the complete perimeter for $Q = 1$.

Q-state Potts model for $Q = 1 - 4$, for a range of system sizes. We compared this data to theoretical predictions [1, 2]. These theories were produced for a continuum model which, in principle, might not apply to the scaling limit of the Q-state Potts model on a lattice. In fact, we found good agreement between our numerical results and the theoretical predictions for every comparison we made including the $D(q)$ spectra and the slopes of the power-law probability distributions.

For the complete perimeter of percolation clusters, we found the slope to be almost exactly -1 for about 4000 orders of magnitude. This suggests the smallest q for which $D(q)$ is defined is $q = 0$. This means that there are many instances of small probabilities on the complete perimeter of percolation clusters which tend to diverge towards negative infinity faster than any power of R/L.

Etching, signposting, and the Green's function method are three tools which can find very small probabilities. The advantage of signposting is that it is natural to use in off-lattice systems, and, in fact, we have applied it to off-lattice DLA [36]. Etching is simple to program and should be easy to use in higher dimensional on-lattice

systems. Lastly, the Green's function method is likely to be the most efficient of the algorithms for on-lattice and network systems, but it is more difficult to implement and requires more memory than etching. The etching and Green's function methods (but not signposting) can be used in problems which involve absorption probabilities less than unity.

CHAPTER IV

The Harmonic Measure of Diffusion-Limited Aggregates including Rare Events

The work in this chapter was published in Europhysics Letters (EPL) in 2009 (Vol. 87, Art. Num. 20001). This chapter is an expansion of the work in chapter II, using signposting to obtain the complete harmonic measure of DLA clusters. Because of the deep connection between the harmonic measure and the growth process, DLA is perhaps the most interesting system on which to study the harmonic measure. This letter was selected as a highlight by the editors of EPL. It was also selected for the list of "Best of 2009" in EPL; the selection criteria was the number of downloads of the article.

4.1 Introduction

Diffusion-limited aggregation (DLA) is a stochastic model for irreversible growth which gives rise to fractal clusters [34, 35], see figs. 4.1, 4.2. The growth process is defined by releasing a random walker far from the cluster and allowing it to diffuse until it sticks to the surface and becomes part of the cluster. Then another particle is released, and so forth. The probability of sticking at various points on the cluster, *i.e.* the distribution of the growth probability, is a function with very large variations.

It is the subject of this chapter.

Since the Laplace equation is equivalent to the steady-state diffusion equation, this probability distribution is proportional to the perpendicular electric field on the surface of a charged electrode with the shape of the cluster; in this context the probability is called the harmonic measure, and is defined for any surface. For fractal surfaces, including that of DLA, the harmonic measure is usually multifractal [39]. For DLA the harmonic measure is of particular interest because of the connection with the growth probability. For other fractal surfaces this connection is lost. However, the measure is still of substantial practical interest because of its relationship with physical processes such as catalysis [64].

For many interesting equilibrium fractals the harmonic measure can be calculated using conformal field theory [53, 54, 55] or Schramm-Loewner evolution (SLE) [56]. There is no corresponding theory for DLA for which the measure must be found numerically. There are numerous studies in the literature of this quantity, for example [38, 57, 41, 58]. This is a difficult problem because of the very large variation of the growth probability. As we will see, the dynamic range of the function is of the order of 10^{80} even for rather small clusters. This is far out of the range accessible to straightforward random walker sampling.

In this chapter, we use a biased random-walk sampling method. We can obtain extremely small growth probabilities and to accurately obtain the complete harmonic measure for DLA clusters of up to 10^6 particles. The method was previously used on percolation and Ising clusters [61]. For those (equilibrium) systems, we found good agreement with analytic predictions for the harmonic measure [1, 2].

The harmonic measure is usually characterized in terms of the generalized dimensions $D(q)$. For integer q, $D(q)$ corresponds to the fractal dimension of the q

point correlation function. We define $D(q)$ by partitioning the external boundary of a DLA cluster into boxes of length l. The probability that a diffusing particle will hit the section of the perimeter contained in box i is denoted by p_i. These probabilities define a "partition function" $Z_l(q) = \sum_i (p_i)^q$ [43]. If $Z_l(q)$ can be written as a power law in the dimensionless ratio R/l, where R is the overall size of the cluster, then the generalized dimension is given by

$$Z_l(q) = (R/l)^{-\tau_q} = (R/l)^{-(q-1)D(q)}. \tag{4.1}$$

There are special values of $D(q)$ including the box-counting dimension, $D(0) \approx 1.71$ [65]. For two dimensional clusters we always have $D(1) = 1$ [44].

Another quantity of interest is $f(\alpha)$, which is called the singularity spectrum. This function is the Legendre transform of $\tau(q)$:

$$f(\alpha) = q\frac{d\tau}{dq} - \tau, \quad \alpha = \frac{d\tau}{dq}. \tag{4.2}$$

As is the case for $D(q)$, some special values of $f(\alpha)$ are known: $f(1) = 1$ and the largest value of $f(\alpha)$ is equal to $D(0)$. $f(\alpha)$ can have a phase transition, namely a maximum value of α, α_{max}, for which $f(\alpha)$ is defined. There has been significant disagreement as to whether $f(\alpha)$ for DLA has a phase transition. This controversy is summarized in [41].

The main difficulty in resolving this issue is that large α, or small q, corresponds to the smallest probabilities on the cluster. The straightforward method of obtaining the harmonic measure, sending large numbers of random walkers at the cluster, is only capable of measuring probabilities down to $\approx 10^{-10}$; even clusters with only 1000 particles have sections with growth probabilities significantly smaller than that. This issue was partially resolved by Jensen *et al.* [41], who used the method of iterated conformal maps (CM) [48, 59, 60], to obtain significantly smaller probabilities. Their

main result was the determination of the harmonic measure of a single cluster of size $3 \cdot 10^4$, where they found probabilities down to 10^{-35}. This work [41] was a significant advance, though the CM results are not conclusive in giving the asymptotic results for DLA because the CM method is limited to small clusters, and it is known that some features of DLA have slow crossover to asymptotic scaling [50]. There are other, technical, questions about the CM method that we discuss below.

Other groups have obtained the harmonic measure for on-lattice clusters using relaxation methods to solve the Laplace equation. Ball and Spivack [57] grew DLA clusters, corrected for lattice anisotropy, up to 10^5 particles. They then solved Laplace's equation numerically to obtain the measure. Hanan *et al.* [58] measured the complete harmonic measure of DLA clusters using a related relaxation technique. In contrast to [57], these authors first grew the cluster off-lattice, then forced it on-lattice to solve for the measure. The simulations in [58] were also limited to small clusters of $6 \cdot 10^4$ particles.

4.2 Simulation Methods

We grow our DLA clusters by the method that is now standard [35], which includes speeding up the process by allowing the random walker particles to take large jumps. We store the cluster in a data structure which allows the calculation of the size of the jump to be performed in $O(\log(n))$ time, where n is the cluster size. These methods allow us to grow clusters in $O(n \log(n))$ time, a big improvement over the CM method, which is $O(n^2)$. In the CM method the harmonic measure is available at each step. In our case, we use a biased random walker method (the signpost method [61]) to obtain the harmonic measure once the cluster has finished growing.

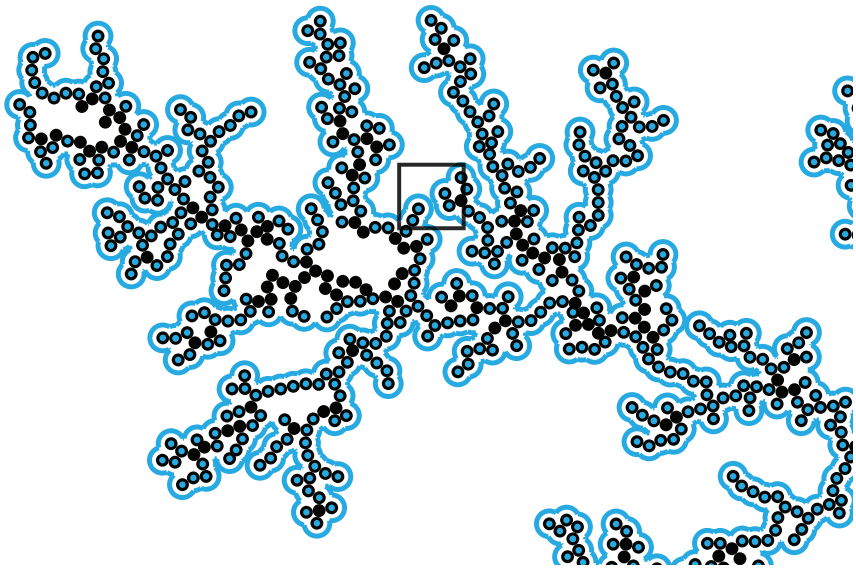

Figure 4.1: A branch of a DLA cluster with an external border representing the Minkowski cover of the cluster. The particles filled blue in the center are accessible to random walkers. The solid black particles can never be reached. Inside the gray box is a narrow neck which marks a low probability pathway for random walkers.

The signpost method consists of two iterated steps: a sampling step and a measurement step. In the first step, a large number of random walker probe particles (N), each with weight $1/N$, are released far from the cluster and diffuse until they hit the cluster. This allows us to determine which areas of the cluster are poorly sampled. Next, we place line segments (signposts) blocking off all regions of the cluster that have sites that are hit by fewer than some percentage of the probe particles, say 10%. In the measurement step, we release N more probe particles far from the cluster and allow them to hit the cluster and signposts. The probe particles that hit the cluster have their weight permanently added to the perimeter site probability distribution. The locations on the signposts where the probe particles hit in the measurement step of the first iteration are used as the initial locations of the N probe particles released in the probe step of the second iteration. To conserve probability, each probe parti-

cle released in the second iteration has weight p/N, where p is the fraction of probe particles that absorbed on the signposts in the first iteration. The probe particles released in the probe step of the second iteration help determine which sections of the cluster are still poorly sampled. More signposts are added to block off the still poorly sampled regions and then the probe particles for the measurement step are released. This process is repeated until the growth probability of all sites has been measured. For a more detailed description of the algorithm, see [61].

This method is similar to a rare event method in chemical physics that uses "milestones"[22]. The main difference between the our method and that of [22] is we choose the locations of the signposts/milestones *dynamically* and that we do not need to know an *a priori* distribution for the random walkers along the milestones. See also [66].

This signpost method allowed us to measure probabilities down to 10^{-300} for percolation and Ising clusters. For DLA we have measured probabilities down to 10^{-80}. Performing the signpost algorithm on DLA clusters is more complex than the percolation and Ising cases. DLA clusters are grown off-lattice, which means that some sections of the exterior of the cluster are almost completely blocked by two branches of the cluster nearly meeting, making a narrow passage slightly larger than the diameter of a probe particle; see fig. 4.1. The probability of a probe particle diffusing through some of these passages without touching the cluster is smaller than 10^{-8}. These passages are treated differently, but in a way consistent with the signpost algorithm. Specifically, we slowly move signposts perpendicular to the passage inward over several iterations until the probe particles can reach other sections of perimeter. In other words, we allow narrow passages to have closely spaced signpost lines so that proper sampling can be achieved.

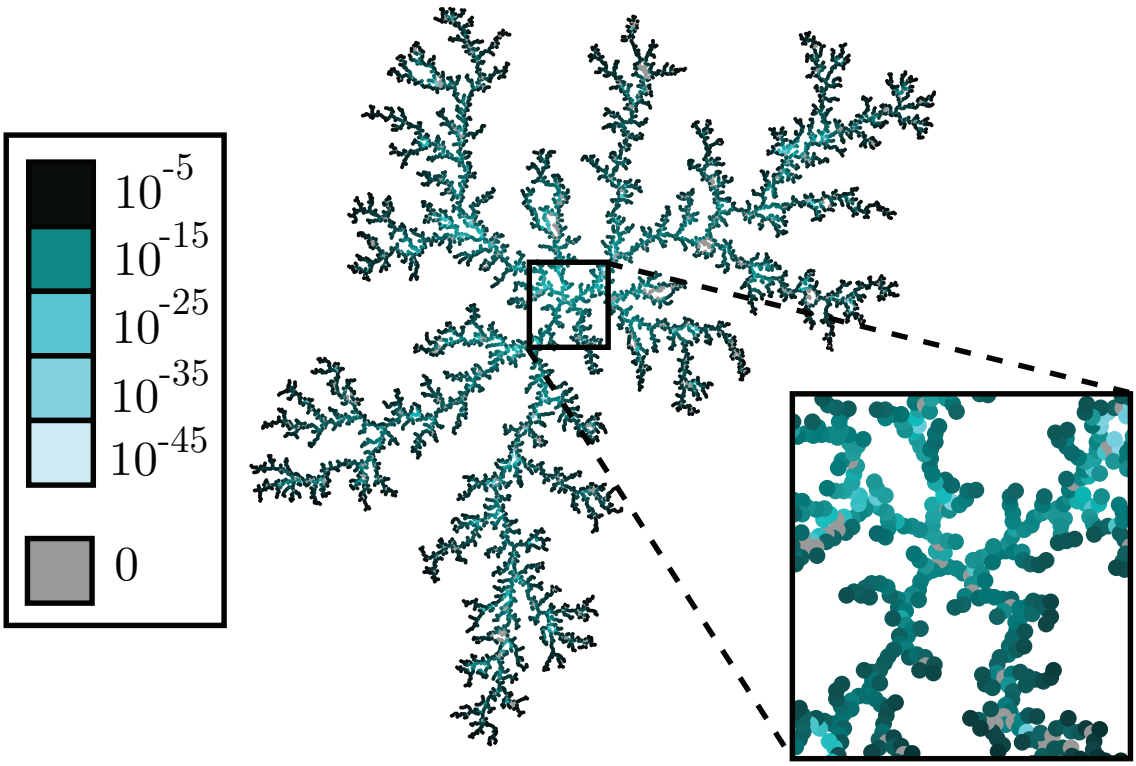

Figure 4.2: The Harmonic Measure for a cluster with 10^4 particles, the lighter the color the smaller the measure. The size of the particles is doubled to represent the cover of the DLA cluster. The smallest measure in the cluster is $\approx 10^{-49}$. Sites that cannot be reached are marked grey.

One minor difference between the signposting we use in this chapter and the one used previously is that we now reduce the threshold for blocking off sections of the cluster as a function of the total number of probe particles that hit signposts in the previous iteration. When more probe particles hit the signposts, we move the signposts much deeper the next time. Previously, we reduced the threshold by a fixed amount each iteration. We found that this dynamic threshold adjustment gave us more consistent particle saturation on the signposts, which in turn decreased the rate at which the error grew from step to step.

Before we can apply the signpost method to a DLA cluster, we first must find the perimeter of the cluster. More precisely, we must find all sites that are accessible to the probe particles. First, we take a ball the size of a probe particle and roll it clockwise around the cluster particle furthest from the center of the cluster until the ball touches a second cluster particle. After that, the ball is rolled clockwise about the second cluster particle until it touches a third particle. This process is repeated until the ball returns to its initial location. Note that a single cluster particle can be visited more than once by the ball. This process finds something akin to the Minkowski cover of the cluster, fig. 4.1. We found that on average only 80% of the particles in a cluster are accessible to random walkers. This means that 20% of the particles have a measure of exactly zero and these regions of the cluster will never grow. We found the 80% accessibility to be constant over a range of large cluster sizes, which shows that the accessible perimeter has the same fractal dimension of the complete perimeter and the cluster itself, in contrast to percolation where the corresponding accessible perimeter has a smaller fractal dimension than the complete perimeter [67, 1].

Our perimeter accessibility results agree qualitatively with other work [68], which

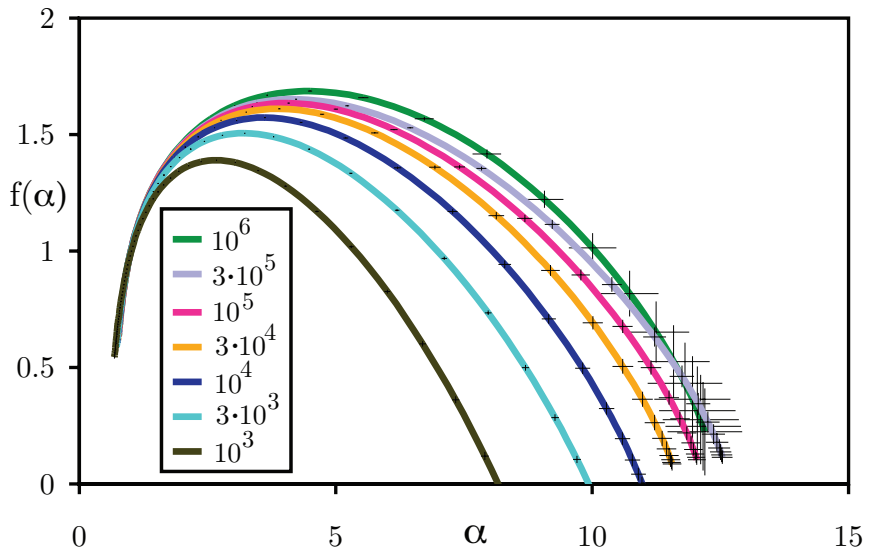

Figure 4.3: $f(\alpha)$ *vs.* α for seven different system sizes with error bars. Note that the spectra appear to be converging to some asymptotic spectrum.

looked at cluster particle accessibility as a function of probe particle size. However, the agreement is not precise because the authors of [68] measured the accessible cluster using 10^5 random walkers, which are extremely unlikely to hit the low measure sections of the perimeter.

We grew DLA clusters of various sizes: 10^3, $3 \cdot 10^3$, 10^4, $3 \cdot 10^4$, 10^5, $3 \cdot 10^5$, and 10^6 particles. For each DLA cluster grown, we obtained the harmonic measure using the signpost algorithm. Fig. 4.2 shows the harmonic measure for a cluster with 10^4 particles. The different cluster sizes required a different number of random walkers per iteration, 10^6, 10^6, $5 \cdot 10^6$, 10^7, $2.5 \cdot 10^7$, 10^8, and $2.5 \cdot 10^8$ for 10^3, $3 \cdot 10^3$, 10^4, $3 \cdot 10^4$, 10^5, $3 \cdot 10^5$, and 10^6 sized clusters, respectively. The number of random walkers needed was estimated by determining the number of walkers required to get at least 10^4 random walkers absorbed on each signpost for every iteration. We believe this is a conservative criterion.

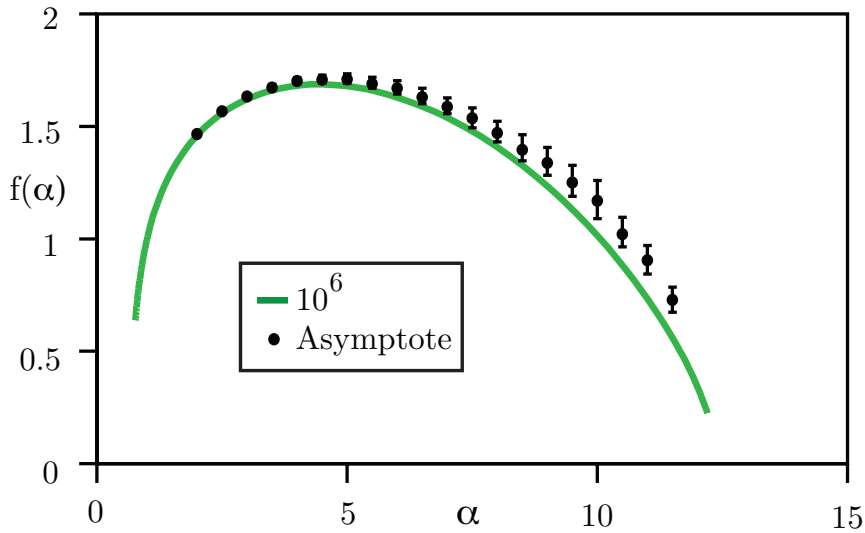

Figure 4.4: $f(\alpha)$ *vs.* α for $n = 10^6$ and the asymptotic estimate, dotted, with error bars estimated from the data in Fig. 4.3. Note that the asymptotic estimate terminates at $\alpha = 11$ only because there were too few system sizes to extrapolate for larger α. We believe the phase-transition in $f(\alpha)$ occurs at $\alpha \approx 14$.

4.3 Results

We use the method described above to obtain $D(q)$. First, we take the space that contains a cluster and section it into boxes of size l and then measure $Z_l(q)$. We do this measurement of $Z_l(q)$ for various values of l for a given q. Next, we calculate the slope of the function $\ln Z_l(q)$ *vs.* $\ln l$; this is $\tau(q)$, which when divided by $(q-1)$, gives $D(q)$. The fit is performed over the range of l for which the log-log plot is linear. This range is about one order of magnitude for the smallest system size and larger than one order of magnitude for larger systems. With this set of $D(q)$'s for individual clusters of various sizes, we can calculate the average values of $D(q)$ for various sizes. We found that our results for large systems are close to the known values for $D(0)$ and $D(1)$, 1.66 and 0.99 respectively.

Using the results of $D(q)$ for individual clusters, we can Legendre-transform the

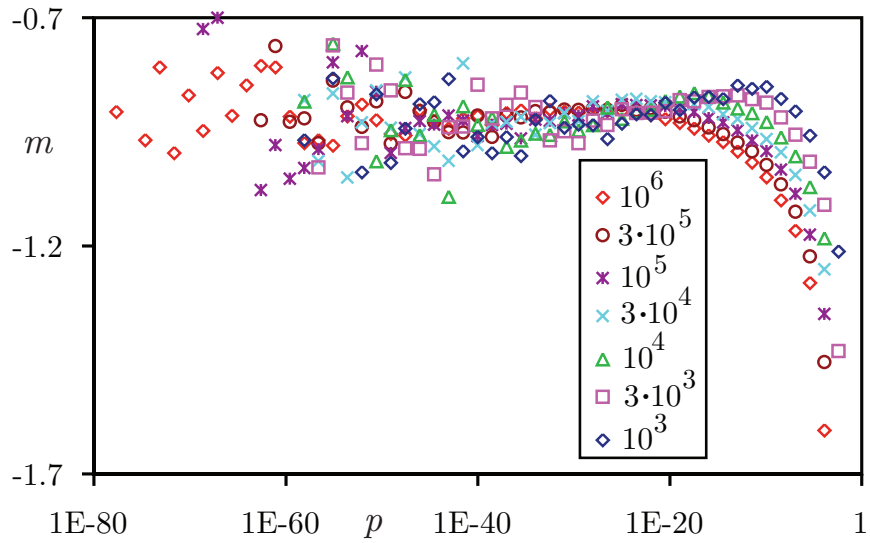

Figure 4.5: The slope of the power-law fit to the probability distribution at various points. The slope at each probability is averaged over about an order of magnitude in probability.

results to obtain $f(\alpha)$ for each cluster; see fig. 4.3. For a range of α's, we estimated the asymptotic value of $f(\alpha)$ using finite-size scaling techniques in n, so that the correction to $f(\alpha)$ is of the form $n^{-\beta}$ where β is a crossover exponent. The asymptotic values were determined by minimizing the residual of the power-law fit, see fig. 4.4. We found the best fit for β was 0.4 for the entire range of α. Note that exponents of 1/3 and 1/2 are also consistent with the data. This means we are consistent with [50] where a crossover exponent of 1/3 was found (for different quantities). The asymptotic $f(\alpha)$ values are consistent with special known values $f(1)$ and the maximum f, measured to be 1.00 and 1.71 respectively. We believe the asymptotic $f(\alpha)$ calculated is the true $f(\alpha)$ for DLA. The last α for which $f(\alpha)$ is defined is more difficult to estimate. From visual inspection, the asymptotic point of the phase transition appears to be about $\alpha \approx 14$. This is significantly smaller than the value found by the authors of [41], $\alpha \approx 18$.

We were also able to obtain a histogram of the growth probabilities for every

system size. The bins were sized logarithmically, to allow for a power law fit to the results. Fig. 4.5 shows the slope of the power law fit to the probability distribution. The slope is fairly consistently -0.9 which corresponds to a smallest q for which $D(q)$ is defined being -0.1. These values agree moderately well with [41].

4.4 Conclusions

In this chapter we applied a rare-event technique to obtain the complete harmonic measure of DLA clusters. We found that the probability distribution is consistent with a power law exponent of -0.9. We also found a slow crossover to infinite-size cluster behavior in $f(\alpha)$, in agreement with previous work [50, 58]. We believe that our extrapolated $f(\alpha)$ is a very good approximation to $f(\alpha)$ for infinite-sized DLA. We found a phase transition in $f(\alpha)$ at $\alpha \approx 14$. This maximum α is related to the opening angle of the branches near the seed point of the cluster. The area around the seed point should have the lowest measure, so the angles in that region, ϕ_{max}, are related to the largest alpha by $\alpha_{max} \sim 1/\phi_{max}$ [41]. By determining the exact relationship between α_{max} and ϕ_{max}, Hanan and Heffernan [58] determined the asymptotic α_{max} as $\alpha_{max} \approx 15$ using results from Mandelbrot et al. [69] for the asymptotic estimate of ϕ_{max}. This is in satisfactory agreement with our results for α_{max}.

Our results differ significantly from those obtained by the CM method [41] in several ways. First, we find a significantly smaller value for α_{max}. Second, we find that finite-size effects are still noticeable on clusters with 10^6 particles. The authors of [41] found no finite-size effects at their largest system size, $3 \cdot 10^4$ particles. This is inconsistent with our findings. Lastly, we find that the smallest probabilities found on clusters of size $3 \cdot 10^4$ are significantly smaller, about 5 to 10 orders of magnitude,

than reported for CM clusters. We do find good agreement for small and moderate values of α, which corresponds to a region of the spectrum which is easily measured by random walker sampling. This explains why the difference between CM and standard DLA clusters was not seen earlier.

Assuming that the signpost method and the CM method are both successful at obtaining the measure for their respective clusters, then the only explanation for the discrepancy is that CM clusters are not the same as DLA clusters grown using particles. Superficially, CM clusters appear to be the cover of DLA clusters. If this were the case, then both methods would obtain the same measure because the measure for a probe particle the same size as a cluster particle hitting a standard cluster is exactly the same as the measure for a point-sized particle hitting the cover of the same cluster. The heart of the issue may be the shape and size of the "bumps" added to the CM clusters during each step. These bumps are designed to have a semicircular shape, and to be of fixed size, but they can distort as noted in [70, 71]. It is important to check that the bumps are, in fact of fixed size, and resize them if necessary. It is not clear that this was done in [41]. Even if this correction were made, the shape of the bumps can be very distorted deep inside the cluster.

CHAPTER V

Fractal Dimensions of the Q-state Potts Model for the Complete and External Hulls

The work in this chapter was published in Journal of Statistical Mechanics: Theory and Experiment in 2010 (Vol. 2010, Art. Num. P03004). This chapter is an outgrowth to our work on the critical Potts model clusters. Given our code base, it was straight forward to accurately measure the hull length and height distributions. In the process, we discovered and resolved an apparent contradiction in the form of the hull and height distributions.

5.1 Introduction

The Q-state Potts model [15] is a well-studied system in condensed matter physics, exhibiting a continuous phase transition for $Q \leq 4$ [26]. It is a generalization of the Ising model [25] with Q different spins, such that like spins interact with a single coupling constant K, and has applications to a range of different physical systems: $Q = 1$ and $Q = 2$ correspond to percolation and the Ising model, respectively, and $Q = 3$ has been used to represent absorbed rare-gas monolayers on graphite surfaces [28, 72]. In this chapter we present results from a numerical study of the fractal dimensions of the hulls of the Fortuin-Kastelyn (FK) clusters [29, 51] for $Q = 1, 2, 3, 4$

using a cylindrical system geometry. In addition we look at the probability distribution of the lengths of these hulls, and find that it has a simple exponential tail. A related quantity, the probability distribution of the height (i.e. the vertical span) of the hulls, also has such a tail.

Fractal clusters enter the Potts model in the following way: for percolation, $Q = 1$, it is well known that the spanning cluster at the critical point is a fractal. The generalization of this for other Q is that at the critical point, K_c, a subset of clusters of like spins, the FK clusters [29, 51], are fractal [73]. FK introduced the clusters by showing that the partition function of the Q-state Potts model can be written as a sum over all bond configurations on a lattice with a bond occupation probability $p = 1-e^{-K}$, multiplied by a weight Q^{N_c} where N_c is the number of clusters in a distinct configuration. To go from a spin configuration to a corresponding bond configuration, bonds are added between like spins with probability p; the connected clusters are the ones we are interested in. The partition function for the Potts model is thus a sum over all possible FK clusters, and at the critical point $p = p_c = 1-e^{-K_c}$, the clusters are fractal. The $p_c(Q)$ for the triangular lattice are given by [62]:

$$\begin{aligned}
Q = 2: &\qquad p_c(2) = 1 - 1/\sqrt{3} \approx 0.42265, \\
Q = 3: &\qquad p_c(3) = 1 - \left[1 + \frac{1}{2}\sqrt{3}\sec\left(\frac{\pi}{18}\right)\right]^{-1} \approx 0.46791, \\
Q = 4: &\qquad p_c(4) = 1/2.
\end{aligned} \tag{5.1}$$

These clusters are characterized by several different fractal dimensions, including D_H, D_{EP}, D_M, D_{SC}, and D_G, corresponding to the complete hull, external hull, mass, singly connected bonds, and narrow-gate fjords. Most previous work on the fractal properties of these clusters has focused on percolation ($Q = 1$) for which

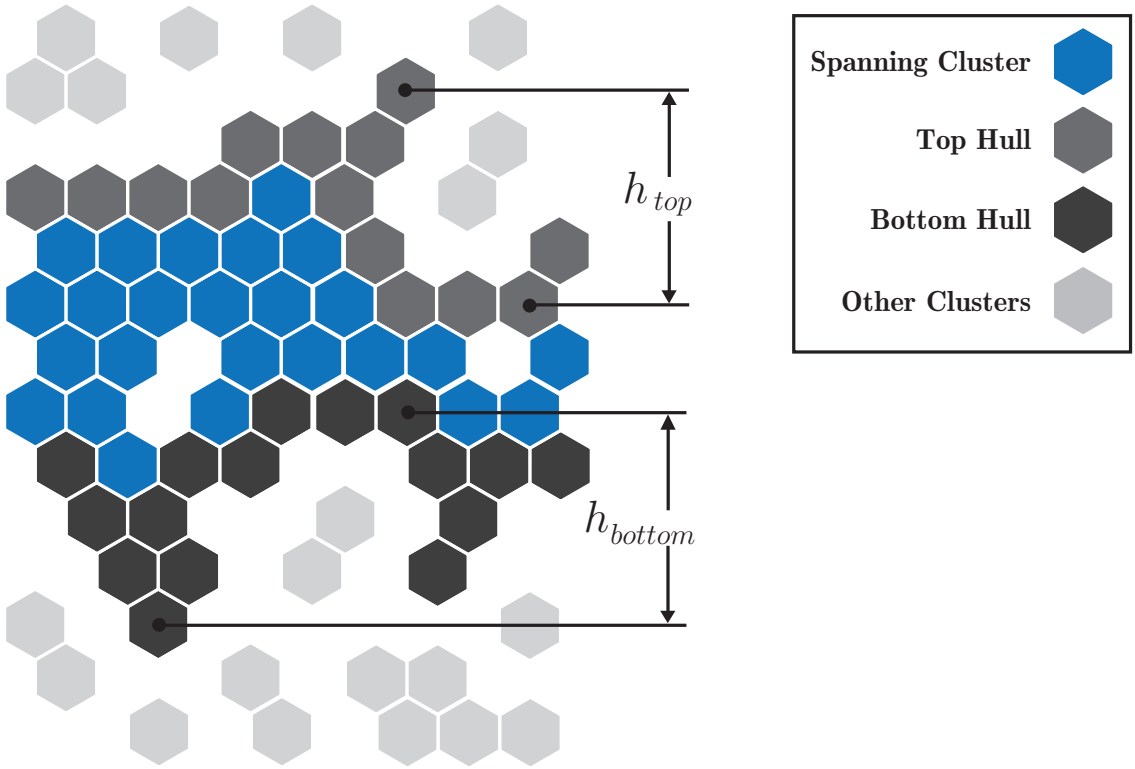

Figure 5.1: Visualization of the spanning cluster, the top and bottom hull, and the height (vertical span) of the hulls.

numerical values have been given for D_H [73, 47, 74, 75], D_{EP} [76], D_{SC} [77], and D_G [78]. Theoretical studies for percolation include studies of D_{SC} [79, 80], D_H [81], D_{EP} [76] and D_G [46]. Theoretical values for D_M, D_H, and D_{SC} have been calculated for all Q values up to the upper critical dimension $Q = 4$ [67, 1, 2]. Hull exponents have also been derived by SLE theory [82].

In this chapter we will present numerical data for D_H and D_{EP} and the height and length distributions for both types of perimeters. In a recent paper, Asikainen *et al* [78] measured D_M, D_H, D_{EP}, D_{SC}, and D_G for $Q = 1$ through $Q = 4$ by studying individual isolated clusters that do not touch the boundary. In the present work, we consider a cylindrical geometry and look at the hulls of the clusters that wrap around it. This method provides an unambiguous measure of the length scale (namely, the

circumference) and leads to accurate numerical results for the fractal dimensions. We also measure the height and length distributions of the hulls themselves.

5.2 Model

Two types of simulations are used. To model $Q = 1$ (percolation), we use the Leath algorithm [33] with site percolation. This method starts with a single "active" site, which attempts to change its undetermined neighbors into active sites with a probability p_c. If a change fails, the site becomes "inactive", and can never be made active. Each new active site attempts to make all of its neighbors active until no more active sites remain. The resulting structure is a percolation cluster. All clusters are grown on the triangular lattice.

For all other Q-states, we generate FK clusters [29, 51] using the Swendsen-Wang (SW) method [31]. The SW algorithm is as follows: after the bonds have been placed for a given configuration, each cluster of sites connected by bonds is labeled with a randomly chosen spin. Then new bonds are put down with probability p between like spins, forming the FK clusters. This process is then repeated. Each cycle constitutes a spin update in that we have flipped entire FK clusters. The SW method allows for fast equilibration of critical clusters.

5.3 Simulation

We grew our clusters on a triangular lattice with periodic boundary conditions along both directions. We made the dimensions of the system elongated so that most large clusters wrapped around in one direction but not the other. The system, though really a torus, was effectively a cylinder. Clusters that wrapped around both directions were rejected. The lattices had an aspect ratio of 100:1 and 8:1 for

Q	theory	measured
1C	$7/4 = 1.750$	1.747
1E	$4/3 = 1.333$	1.330
2C	$5/3 = 1.667$	1.663
2E	$11/8 = 1.375$	1.375
3C	$8/5 = 1.600$	1.602
3E	$17/12 = 1.417$	1.412
4C	$3/2 = 1.500$	1.510
4E	$3/2 = 1.500$	1.534

Table 5.1: The theoretical [67, 1, 2] and measured values for the fractal dimensions for the complete (C) and external (E) hulls of Q-state Potts model

percolation and FK clusters, respectively. We use the width (W), of the lattice as the characteristic length. For $Q = 1$, clusters were grown for $W = 8$, 16, 32, 64, 128, and 256. For $Q = 2$ through 4, we also considered $W = 512$ and 1024 in order to reach the asymptotic fractal dimension.

For each valid cluster grown, the number of sites on the top and bottom of the cluster were recorded as well as the height (vertical span) of the hulls, defined as the highest point on the top hull minus the lowest point on the top hull; see figure 5.1. The top and bottom hulls were considered independently and both were used to calculate the fractal dimension. To record the complete hull, we used the outermost layer of active sites as the hull. For the external hull, we used the layer of inactive sites that pad (are neighbors to) the outermost layer of active sites.

In the case of FK clusters, the system must be equilibrated before recording can start. The equilibration times used are $2W$, $3W$, and $16W$, for $Q = 2$, 3, and 4, respectively. We use longer equilibrium times for $Q = 4$ because of its slower dynamics [83]. To determine the equilibration time, we measure the average energy per spin and the average largest cluster size as a function of the number of spin updates. We find that both observables relax to their steady-state values exponentially quickly

with the same decay time for a given Q and W, though the average energy is always closer to the steady-state value at a given step. For example, we find the average energy was within 1% of the steady-state value within 500 spin updates, whereas it took the average largest cluster size 1500 spin updates to reach 1% of its steady-state value for $Q = 4$ and $W = 256$. Because of its slower relaxation, we believe that the average largest cluster size is a better measure of equilibration.

After the system is equilibrated, we pick a random cluster, and if the cluster wraps around the width of the system, its hull length and height are recorded in a similar fashion as was done with percolation. Note that while the FK clusters are essentially bond percolation clusters, we consider the hulls on the sites of the clusters – that is, we treat the sites in the FK clusters as a site percolation problem. To record the external hulls of these clusters a layer of inactive sites are simply added outside of the top and bottom hulls. (This is the advantage of using the triangular lattice.) Several spin updates are taken between each successive attempt to find a valid cluster. The number of updates between measurements are 8, 12, and 64 for $Q = 2$, 3, and 4, respectively.

The work of Asikainen *et al* [78] differs from ours in several ways. For $Q > 1$, they simulated a single system size, 4096×4096, and picked random clusters after the system was well equilibrated. To equilibrate such a large system, they needed to use a new technique to speed up equilibration. For their system of isolated clusters, the length scale is not as clearly defined as it is in the cylindrical geometry; they used the cluster's radius of gyration for the length scale. Lastly, they used a square lattice for their simulations, for which the complete and external hulls are not as clearly defined as in the triangular case because of the corners in the square system [45]. Our results are consistent with theirs in regards to the fractal dimensions, but

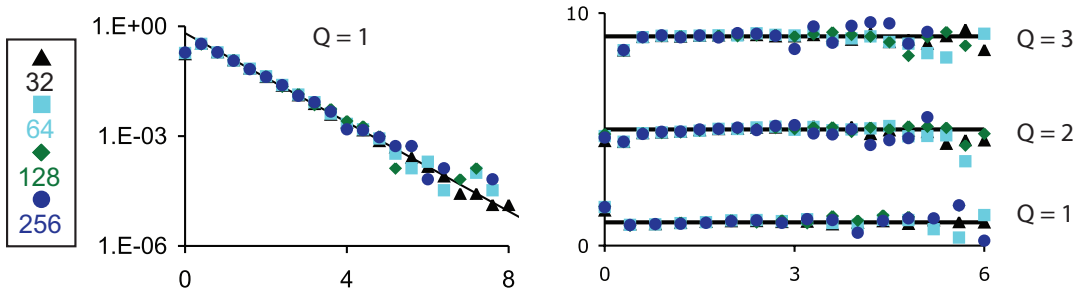

Figure 5.2: The probability distributions for the lengths of the complete hulls, for several system widths: $W = 32$, 64, 128 and 256 for $Q = 1$ as a function of $L/\langle L\rangle$ where $\langle L\rangle = 0.93W^{7/4}$ (left). The data are put into bins of size $C\langle L\rangle$ with $C = 0.4$. The black line is the best fit using an exponential with an inverse decay length of λ^{-1}. The plot on the right is of the residual (difference) of the fit of exponentials to the different Q complete hulls, where different $C = 0.3$ for $Q = 2, 3$. The best fit for the λ^{-1} for $Q = 1$, 2, and 3 are 1.4, 1.65, and 1.7, respectively. The residual data for different Q are offset vertically for clarity.

are significantly more precise.

5.4 Results

For each state Q we obtain a large set of values for the lengths of the complete and external hulls and hull heights for a wide range of system widths. With this data we can compute the fractal dimension of the different hulls. For a fractal the length of the hulls scales as a power $L \propto W^D$, where D is the fractal dimension. We measure D in the conventional way by making a linear fit of $\ln L$ as a function of $\ln W$. The accuracy can be assessed two ways: (1) the value of the square of the Pearson product-moment coefficient, R^2, of the fit and (2) the apparent randomness of the residuals. We focus on the latter as it can be used to determine when finite-size effects are significant, and thus which sizes we can use for the fit. We determined that all measured widths for the complete hull can be used to fit D for $Q < 4$. For

the external hull, there are significant corrections for $W < 256$, so for those systems, we use 256, 512, and 1024. We use the same three sizes for both hulls of $Q = 4$.

Table 5.1 shows a summary of our findings for the fractal dimensions of the complete and external hulls for $Q = 1$, 2, 3, and 4. With the exception of the external hull for $Q = 4$, all the measured values agree with theory within a fraction of a percent. The discrepancy at $Q = 4$ can be interpreted as arising from logarithmic corrections to scaling [84]. In Ref [78] a correction is made for the slow crossover.

We now turn to the probability distribution of the quantities that we measured. We first found that the average height of the hulls scales linearly with system width so that these are isotropic fractals. We then measured the heights and hull lengths of a large number of clusters in order to produce the probability distribution of hull lengths. Figure 5.2 shows the scaled probability distributions of complete hulls for $Q = 1$ and the residuals to the best-fit exponential for the first three values of Q. $Q = 4$ is not shown because the scaling does not work for the size systems we used. Figure 5.3 is a similar plot for the external hulls, with a finer bin size. All scaled hull length distributions have exponential tails. The complete and external hulls appear to have fairly similar decay lengths for the different values of Q.

We also calculate the probability distribution of hull heights for the complete and external hulls, figures 5.4 and 5.5. Again we see that the distributions have exponential tails.

The exponential tails are, in some sense, no surprise in this type of problem as we can see from a simple example. Consider the track of a free random walker in a channel with absorbing walls at $x = \pm W/2$. The walker starts in the middle of the channel, and we seek the probability distribution of the maximum height attained by the walk before it hits the side walls. This is more-or-less what we are doing with

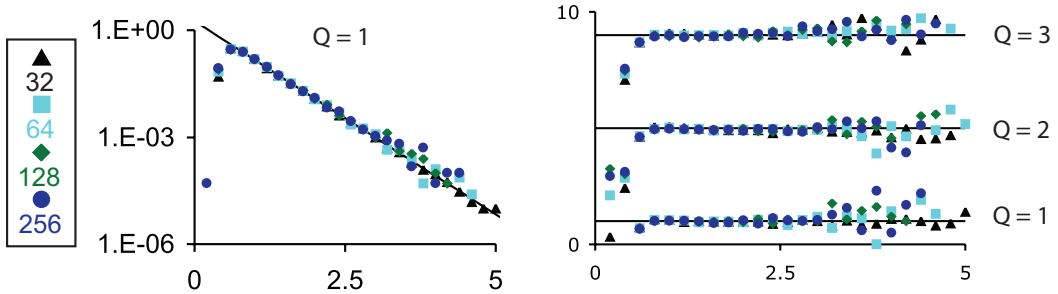

Figure 5.3: The probability distributions for the lengths of the external hulls, for several system widths: $W = 32$, $W = 64$, $W = 128$, $W = 256$ for $Q = 1$ (left) where the different system sizes are scaled in the same way as fig. 5.2. The right shows the residuals for $Q = 1$ through $Q = 3$. The best fit for λ^{-1} for the different Q's are 2.5, 2.4, and 2.35 for $Q = 1$, 2, and 3, respectively. The residual data for different Q are offset vertically for clarity.

our spanning fractals, though they are not free random walks, of course.

It is not hard to see why we get exponential tails in this case, by considering the following steps: first we define an auxiliary problem by putting an absorbing wall at height h above the origin. The number of walks, $N(h)$, that ever hit this wall before being absorbed on the sides is the number of walks with height greater than h. The distribution we seek is proportional to dN/dh. The problem can be solved exactly by going to the continuum limit and using conformal mapping. However, we can guess the solution easily: the probability for a walker to penetrate a channel for height h is clearly exponentially decreasing in h; the conformal map gives $e^{-\pi h/W}$. Then N must have this dependence, along with its derivative.

For percolation or FK clusters, the probability distribution of the maximum height can also be found from the derivative of the probability of crossing, and in this case one also expects exponential behavior for large h [85, 86].

The existence of exponential tails in both the hull length and height distributions

may appear contradictory. We know that the average height, h, is proportional to the system width, W, which is related to hull length, L, through a power-law $W \sim L^{1/D_H}$. One might guess that this should lead to a stretched exponential $\exp(-cL^{1/D_H})$ rather than simple exponential for the distribution of L. In order to clarify this issue, we looked at another well-studied system: the percolation hull walk in an open rectangular region [47] for which it is easy to generate a large ensemble of fractal curves.

We performed simulations of the self-avoiding hull walker on a square lattice [47, 74]. The hull walk algorithm is as follows. At every step the walker either turns right or left. If the current site has not been marked, the walker turns left with probability p, marks the site "active" and takes a step forward. With probability $1-p$ it turns right, marks the site "inactive," and takes a step forward. If the walker steps onto an active or inactive site it always turns left or right, respectively. We used this walk to create closed figures that have the same fractal dimension as the complete hull of percolation clusters, $D = 7/4$. We recorded the average walk length as a function of the maximum height for several values of W.

For every W, the walk length is proportional to the maximum height,i.e., $L \sim m_W h$, as one would expect for large h. However, we find that the *coefficient*, m_W, depends on W as $m_W = aW^{3/4}$. With these expressions, we get back the known power-law relationship between L and W, $\langle L \rangle \sim m_W \langle h \rangle \approx aW^{3/4}W = aW^{7/4}$, so that we recover the fractal dimension of 7/4.

5.5 Conclusion

In this chapter, we measured the fractal dimensions for the complete and external hull lengths, D_H and D_{EP} for $Q = 1$ through $Q = 4$. We used the Leath algorithm

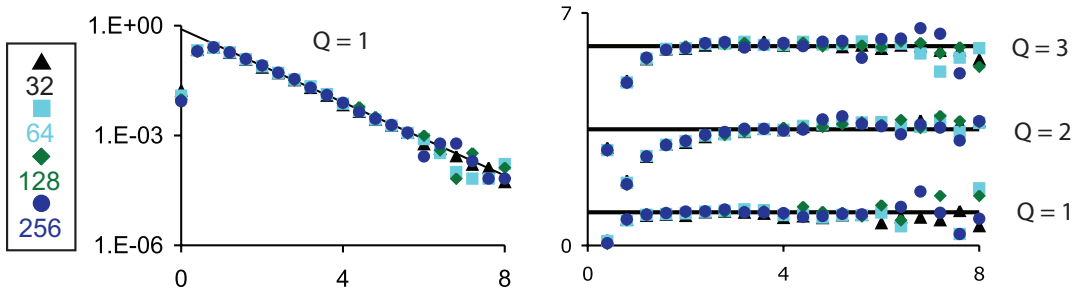

Figure 5.4: The probability distributions for the heights of the complete hulls, for several system widths: $W = 32$, $W = 64$, $W = 128$, $W = 256$ for $Q = 1$ (left). The bin sizes are CW, where $C = 0.4$. The right shows the residuals for $Q = 1$ through $Q = 3$. The best fit for λ^{-1} for the different Q's are 1.15, 1.0, and 1.05 for $Q = 1$, 2, and 3, respectively. The residual data for different Q are offset vertically for clarity.

to grow critical percolation clusters. For $Q = 2$,3, and 4, we used the SW method to generate critical FK clusters. All systems used a triangular lattice with periodic boundary conditions. The aspect ratio of the systems were heavily skewed so that spanning clusters would span in one direction and not the other. The smaller length (W) exactly determined the length scale of the system. We generated a large number of spanning clusters of various system sizes, $W = 8, 16, 32, 64, 128$, and 256 and additionally $W = 512$ and 1024, for percolation and $Q > 1$, respectively. We find excellent agreement between our results and the associated theories [67, 1, 2]. We also measured the distributions of the hull lengths and heights. We found that the distributions for $Q = 1$, 2, and 3 have exponential tails. The values of the inverse decay lengths for the exponentials are given in the figure captions. We discussed the apparent contradiction between the height and length distributions both having exponential tails, as opposed to one being a stretched exponential. We resolved this contradiction by showing that the relationship between average height and width is in fact linear for a fixed W. We found that the power-law scaling formula can be written

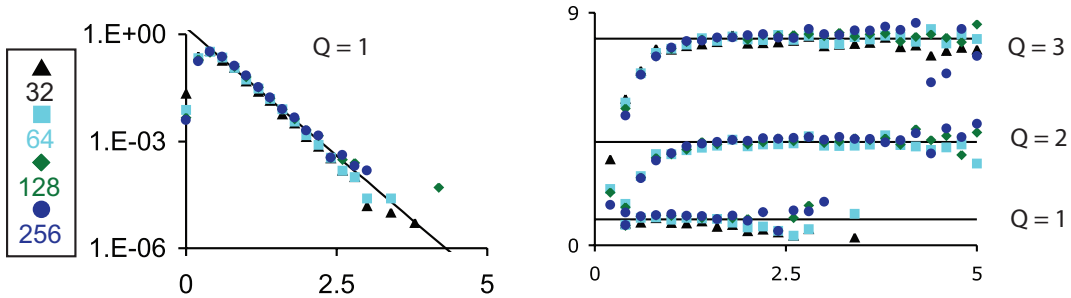

Figure 5.5: The probability distributions for the heights of the external hulls, for several system widths: $W = 32$, $W = 64$, $W = 128$, $W = 256$ for $Q = 1$ (left). The bin sizes are CW, where $C = 0.2$. The right shows the residuals for $Q = 1$ through $Q = 3$. The best fit for λ^{-1} for the different Q's are 3.3, 1.8, and 1.7 for $Q = 1$, 2, and 3, respectively. The residual data for different Q are offset vertically for clarity.

as $\langle L \rangle \sim m_W \langle H \rangle \approx aW^{3/4}W$. This formula clarifies the relationships between L, H, and W. We are not aware of any predictions regarding the exponential tails in the hull length distributions.

CHAPTER VI

The Barrier Method: A Technique for Calculating Very Long Transition Times

The work in this chapter was published in Journal of Chemical Physics in 2010 (Vol. 133, Art. Num. 124103). This chapter is our first foray into long transition times, or infrequent rare events. The algorithm developed in this chapter, the barrier method, was inspired by our work on the Green's function method in Chapter III.

6.1 Introduction

Important events for a transition may have a time scale many orders of magnitude larger than typical events; in this case, they are called "rare events." They have been studied in a number of different contexts, including the extinction of diseases [10] or of populations [11], network queue overflow [8], and slow chemical reactions [12].

To fix our ideas, we consider a very simple problem, that of an endemic disease which fluctuates to extinction. Consider a population of fixed size, N, with S members who are susceptible to an infection, and I who are infected. When S encounters I, the infection is transferred with rate $\beta SI/N$; β measures the infectivity. Infected individuals can spontaneously recover with rate κ, and can be immediately reinfected. This is called the susceptible-infected-susceptible (SIS) model [16]. In this

simple form it can be thought of as a Markov process in the number of infected; note that $S = N - I$. Thus

$$\begin{aligned} W(I \to I+1) &= \beta SI/N \\ W(I \to I-1) &= \kappa I. \end{aligned} \tag{6.1}$$

There is a long-lived state where the disease persists when $R_0 = \beta/\kappa > 1$, namely $I = N(1 - R_0^{-1})$. However, there is an important rare event, namely a fluctuation to $I = 0$, which means that the disease is extinct and cannot return. For this simple form of the model the mean exit time for this transition T can be found exactly [3]. For large N, there is an asymptotic formula

$$T \to \frac{R_0}{(R_0-1)^2}\sqrt{\frac{2\pi}{N}}e^{N(\log R_0 - 1 + 1/R_0)}. \tag{6.2}$$

This formula has features that are generic to the kind of "barrier climbing" problems that we treat here. The exit time is of the form $g(N)\exp(NW)$ where g is a slowly varying prefactor, and W is a generalized barrier height (or quasipotential) scaled by the large parameter, N. A plot of the exact results for the SIS model with $R_0 = 2$ is given in Fig. 6.1 along with results from numerical computations that we will describe below. Note that even for modest-sized systems, the mean time to extinction can be huge; for $N = 300$ we have $T \approx 4 \times 10^{24}$.

As we see from this simple case, these phenomena are frequently out of reach for brute-force simulations. To overcome this problem, many techniques have been developed [87]. In this article, we revisit this problem and present a very efficient technique which we call the *barrier method* and which gives the mean first-passage time for a transition to an unlikely target state. The method does not depend on special features such as knowledge of a steady-state distribution or detailed balance in the process. We need only that the dynamics be stochastic and that the transi-

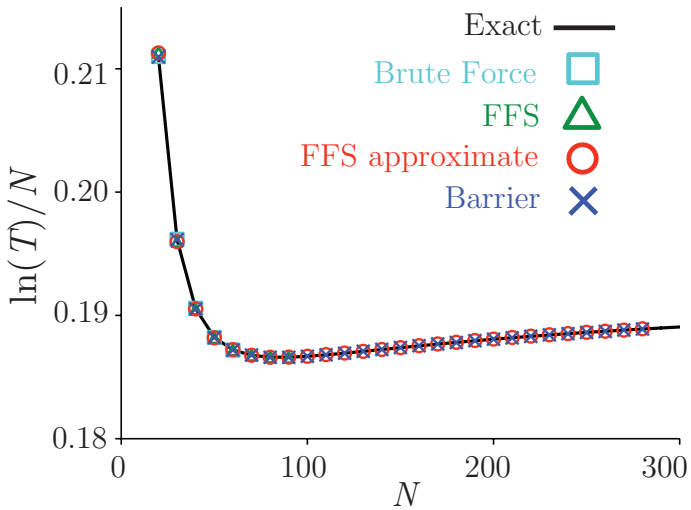

Figure 6.1: The mean time to extinction T as a function of the population size N for the simple SIS model calculated exactly [3] and by two of the numerical methods described in the text. All numerical results agree with the exact result within statistical error.

tion probabilities depend only on the current state. That is, we deal with Markov dynamics, which can be reversible or irreversible.

The essence of the algorithm is to follow the development of an ensemble of systems and oversample the cases that happen to approach the target, and not allow backtracking away from the target. In this respect our method resembles the signposting algorithm [61, 36] that we developed for finding the penetration of a random walker into a fractal.

We believe that the barrier method is the most efficient method available for computations in low-dimensional systems. In this chapter, we explain how it differs from previous techniques, and we apply it to two systems. The first is described by a nonequilibrium model introduced by Maier and Stein.[4] The second is a generalization of the SIS model in which the population is allowed to fluctuate [88]. In this model, we are concerned with the average time for the disease to go extinct. Lastly, we discuss the advantages of our method and future work.

6.2 Background

To study rare events in a Markov process, we must deal with states in state-space that are unlikely to be visited in any simulation of reasonable length [18]. Most sample paths spend the majority of time visiting the most likely states and give good estimates of the corresponding probability density. For rarely visited regions, we need to use special methods.

An example of such a method is biased sampling, in which we arrange our simulation to be biased towards rarely visited regions. Two important subsets of this approach are importance sampling and splitting techniques. Importance sampling, the most commonly used method for equilibrium systems, requires some *a priori* information about the probability distribution. In contrast, if the probability distribution is only accessible via simulation, *splitting* and related techniques are very useful. The latter case is the focus of our work.

6.2.1 Splitting and repetitive simulation trials after reaching threshold

Splitting [18] involves placing a barrier in state-space. When a sample path crosses the barrier, it is split into independent realizations whose statistical weights add up to the original. By placing one of these splitting barriers in a region which would be infrequently visited, that region will subsequently be better sampled. For very difficult-to-reach regions of state-space one barrier is not sufficient. The use of several barriers is called multilevel splitting [19].

Multilevel splitting has two drawbacks. First, if the barriers are too close together or too far apart, the number of simulations will grow or decay exponentially. Further, realizations which have small weight (because they have been split many times) will often *backtrack*, i.e., tend to move back to the well-sampled regions and

waste computational effort. Some of these problems have been solved by REpetitive Simulation Trials After Reaching Threshold (RESTART) [20], which is designed for dealing with queuing problems.

In RESTART [20] one considers nested subsets of phase space, $A \supset B \supset C \supset D$, where D is the target. Barriers are placed between these regions. A sample path is started in A and evolves until it reaches B. Then the sample is split into R "retrials" with equal weight. One of these is designated the primary and all realizations evolve independently. If any of the nonprimary paths backtracks into A, it is terminated. Each barrier is crossed in turn, and the time spent in D by the reweighted paths gives an estimator of the phase space probability in D.

RESTART partially solves the backtracking problem because most samples that exit low probability regions are terminated. However, in the original version [20, 89] RESTART still can lead to a divergence in the number of samples if the barriers are too closely spaced. It has been noted that the barrier placement problem could be partially alleviated by performing fixed effort RESTART instead of fixed splitting RESTART [90], but this does not completely fix the problem. We give another approach to this problem below.

6.2.2 Forward flux sampling

Forward flux sampling (FFS) [21] uses the same principles as splitting in applications to computational chemical physics. Most rare-event techniques in this area [87] require equilibrium ensembles and detailed balance. FFS is unusual because it is applicable to systems without detailed balance. It has been used to study genetic switches [21, 66, 91], nucleation [91, 92, 93] and a model problem due to Maier and Stein [91, 94] which we also treat below.

FFS finds the first-passage time between metastable states A and D as follows. First, we run a single long simulation and count the number of times the sample path exits A through barrier λ_0, which bounds region A. The average flux, k_0, through the barrier λ_0 is calculated by dividing this number by the total simulation time, discounting the time associated with trajectories that reach D and return to A.

If we call λ_M the barrier around D, the transition rate from A to D is

$$k_{AD} = k_0 P(\lambda_M|\lambda_0). \tag{6.3}$$

where $P(\lambda_M|\lambda_0)$ is the probability that a sample path which starts on λ_0 will cross λ_M before going back to A.

We can get $P(\lambda_M|\lambda_0)$ efficiently by introducing intermediate barriers $\lambda_i, i = 1 \ldots (M-1)$ to divide the sample space along level surfaces of some reasonable guess for the reaction coordinate — we call this the order parameter. The probability factors into

$$P(\lambda_M|\lambda_0) = \prod_{i=0}^{M-1} P(\lambda_{i+1}|\lambda_i), \tag{6.4}$$

where $P(\lambda_{i+1}|\lambda_i)$ is the probability of starting at λ_i and reaching λ_{i+1} before going back to λ_0. To measure $P(\lambda_1|\lambda_0)$, R samples are started from the locations along λ_0 where they left A in the first step. The paths are continued until they reach λ_1 or go back inside λ_0. The fraction of that reach λ_1 is the estimator of $P(\lambda_1|\lambda_0)$. Then we proceed to λ_2 and start R samples, etc. The point is to break down a long sample path into a series of short segments.

FFS does not allow the number of samples to diverge, as in splitting. However, it does allow backtracking because samples which start at λ_i must be allowed to return to A. This effect can be somewhat reduced by pruning the backtracking paths [66]. However, if there are metastable states in the region between A and D, backtracking

can take a long time.

Further, the calculation of k_0 requires that the initial long simulation reaches the end state D at least once in order to properly sample the entire region between the start and the target. If one does this, FFS will often be bottlenecked by the calculation of k_0; this defeats the purpose of using a rare-event technique. Fortunately, for systems with featureless barriers, running an initial simulation that crosses λ_0 a fixed number of times, say $10R$, is typically sufficient. We call this version "approximate FFS." For a comparison of the two versions of FFS for the SIS model, see Figs. 6.1 and 6.3. The approximate method of FFS gives good accuracy for this problem and is quite fast. However, for systems with metastable states in the region between A and D, this method of determining k_0 is not sufficient. Our method (see below) overcomes all of these problems.

6.2.3 Equilibrium methods

Milestoning [22] is a technique for equilibrium systems in which one runs simulations of short paths between barriers to find the local first-passage time from one barrier to the next. The equilibrium ensemble on the barriers gives the launching points and there is no backtracking at all. The local first passage times are put into an integral equation to find the global first passage time. As we will see, we avoid using equilibrium considerations on the barriers by keeping track of the landing points of individual paths, but otherwise, our method uses similar ideas.

There are many other useful equilibrium rare-event techniques that have been developed, including transition path sampling (TPS) [23], which collects an ensemble of possible transition paths between metastable regions. Transition interface sampling (TIS) [24] is an improvement on TPS which uses the same rate calculation formula

as FFS. Also, the finite temperature string method [95] finds the most likely path between the start and end regions represented by a smooth curve (string). Because of the equilibrium requirements of these methods, they cannot easily be applied to the models studied in this chapter.

6.3 Barrier method

6.3.1 Algorithm

In the barrier method, we consider the same sort of problem as FFS. For the moment, we assume that the locations of the barriers are known *a priori*, as shown in Fig. 6.2. We will discuss the best way to distribute the barriers below. In the first step, R trial simulations are started in A and run until they reach λ_0. Each trial r ends at $W^r_{\lambda_0}$ at time τ^r_{A,λ_0}. R more trials are started at the locations along λ_0 where each r ended in the previous step and run until they reach λ_1. These trials can backtrack as far as they need to. The locations along λ_1 where the trials stopped are $W^r_{\lambda_1}$ and the transition times from λ_0 to λ_1 are $\tau^r_{\lambda_0,\lambda_1}$. For each r the total time to start from A and reach λ_1 is $\tau^r_{A,\lambda_1} = \tau^r_{A,\lambda_0} + \tau^r_{\lambda_0,\lambda_1}$.

In the next step we eliminate backtracking. We start each trial r at the location on λ_1 where the previous trial r finished, $W^r_{\lambda_1}$. Each sample path continues until it either reaches λ_2 and stops or returns to λ_0. If the trial returns to λ_0, we move it back to λ_1 according to an estimate of where it would have crossed λ_1 starting at where it stopped along λ_0; we also add the time that it would take that estimated path to get from that spot on λ_0 back to λ_1 to the current trial time. In practice, we find the closest $W^s_{\lambda_0}$ to the current location and add $\tau^s_{\lambda_0,\lambda_1}$ to the trial time and continue the trial at $W^s_{\lambda_1}$. That is, we move the sample point "in one step" to the next barrier. Continuing, the sample path can either reach λ_1 and stop or go back

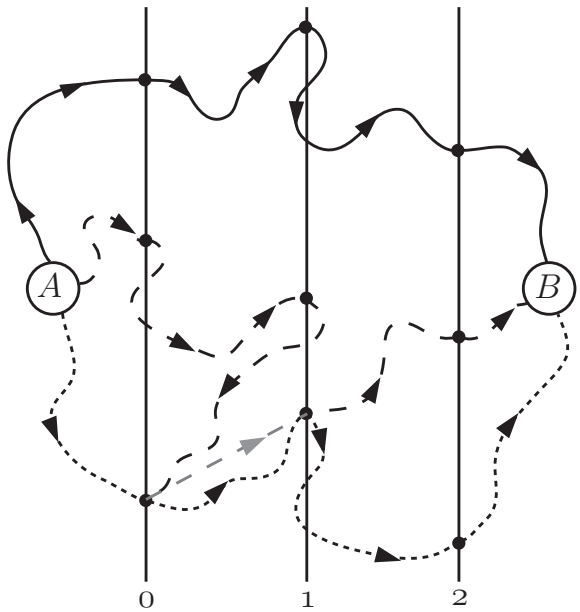

Figure 6.2: Barrier method. Three different paths are started from A. The paths cross barrier 0 then cross barrier 1. The middle path then moves backwards and reaches barrier 0. The end points on barriers 0 and 1 are used to "jump" the path back to the barrier 1, grey path. Paths terminate once they reach B.

to λ_0, in which case we repeat the process of jumping. The locations along λ_2 and the transition time for each r are $W^r_{\lambda_2}$ and $\tau^r_{\lambda_1,\lambda 2}$. The total time to reach λ_2 is $\tau^r_{A,\lambda_2} = \tau^r_{A,\lambda_1} + \tau^r_{\lambda_1,\lambda_2}$. Then we repeat the process for the next barrier and continue until τ^r_{A,λ_M} and W_{λ_M} are calculated. The average transition time from A to λ_M is $\sum_r \tau^r_{A,\lambda_M}/R$.

This method differs from FFS in two ways. First, we work with transition times instead of transition rates and never deal with probabilities directly. Second, the barrier method does not require sample paths to travel from λ_i all the way back to λ_0. This can lead to a dramatic improvement in efficiency over FFS, as we will see in Section 6.3.2. Note that metastable states in the barrier region pose no problem for this method.

6.3.2 Accuracy and Efficiency

The simple SIS model described above is exactly solvable. This allows us to make a direct comparison between FFS and the barrier method. Using $N = 20$ to 200, we found less than 1% difference between the barrier method and the exact results, which was within the variance of the measurement. Both versions of FFS gave similar accuracy; see Fig. 6.1. The exact version of FFS, which samples the whole region between A and D in the estimate of k_0 is impractical for $N > 60$, as shown in Fig. 6.3.

We now compare the efficiency of approximate FFS and the barrier method. Simulation efficiency can be defined by the amount of computation time C needed to obtain the exit time within a relative accuracy of σ; we chose $\sigma = 0.1$. Using this definition, we varied the number of trials R to find C for $R_0 = 2$ for a range of N, from 50 to 290 in steps of ten for FFS and the barrier method. For both, we used $N/10-1$ barriers, placed every five infectious population size steps starting at $I = 5$. No attempt was made to optimize the barrier placement. The comparison is given in Fig. 6.3. We find that both algorithms appear to have a power-law relationship between C and N, with powers of roughly 3 and 2 for FFS and the barrier method, respectively. This shows that for a case where the barrier method is not obviously better, because of the simple landscape, it still significantly outperforms FFS.

6.3.3 Dynamic Barrier Placement

We have also developed a way to determine the locations of all of the λ_i barriers as the simulation progresses. The method outlined in this section can be applied to FFS as well as any other biased sampling methods that use barriers, e.g., RESTART [20]. We have previously used a similar method for signposting [61, 36].

Assuming that the location of the first barrier has been chosen and all of the $W^r_{\lambda_0}$ values have been calculated, p probe trials are started along λ_0 at $W^r_{\lambda_0}$ for some randomly chosen r's. The probe trials are run for some fixed number of simulation steps S. The location of λ_1 is chosen as the average longest excursion along the order parameter of the probe trials. Next, $W^r_{\lambda_1}$ and $\tau^r_{\lambda_0,\lambda_1}$ are recorded. We repeat the process by running more probe trials for S simulation steps, not counting any steps involved with jumping from λ_0 to λ_1, starting at random $W^r_{\lambda_i}$'s. The average furthest excursion is used as the location of λ_2. This process is repeated until λ_M is reached. An alternative approach would be to estimate the location of the next barrier such that a fixed fraction of samples reach λ_{i+1} before reaching λ_{i-1} when starting at λ_i. A rough estimate of the optimal forward fraction is [89] e^{-2}. We did not chose this approach because we found that for small noise the probability of making any forward progress is significantly smaller than e^{-2}. This would cause the barrier method to stop making progress.

There has been some previous work on barrier placement or staging using FFS [96]. The scheme uses two FFS calculations: one with a guess of the best barrier locations and fixed R, and a second with either the barrier locations optimized or the number of trials R for each barrier optimized. The authors found that it is better to optimize the spacing of the barrier to get uniform $p(\lambda_{i+1}|\lambda_i)$ with fixed R than to optimize R for each barrier. When we apply dynamic barrier placement to FFS, we also take this view. The important difference from the approach outlined above is that we determine the location of the barriers *one at a time* instead of all at once. This should lead to comparative computational gains when the trial barrier placement is significantly different from the optimal placement.

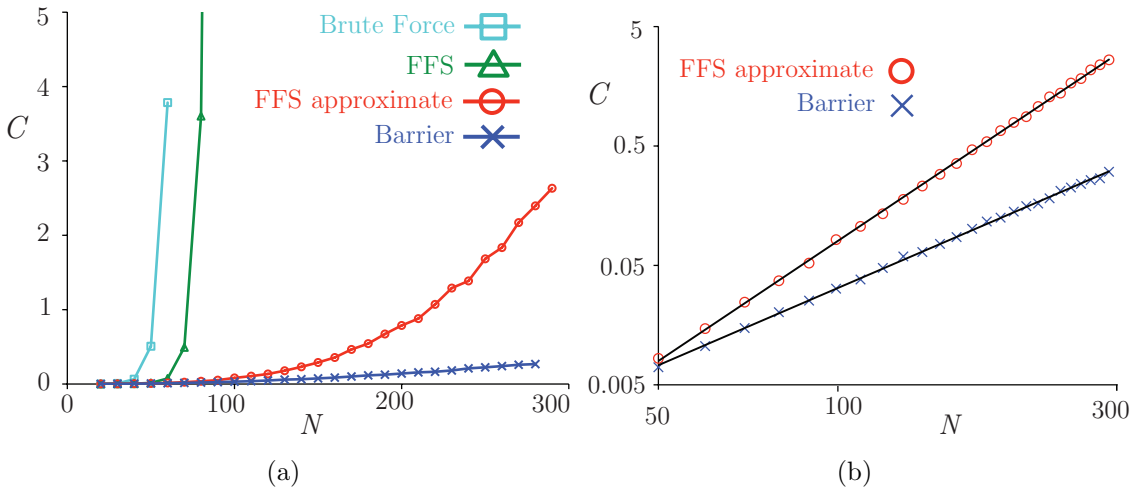

Figure 6.3: (a) CPU time required to calculate T to a fixed precision (arbitrary units) for the SIS model. For FFS and the barrier method, the same number of barriers were used and placed in the same locations (every five steps in number of infected). (b) Efficiency comparison between the (approximate) FFS method and the barrier method. The solid lines represent the best fit to a power-law: with powers 3.3 (FFS) and 2.1 (barrier method). These values should be taken as rough estimates of the powers.

6.3.4 Additional lookups

As with FFS, the barriers should be set up at fixed values of the order parameter. The closer the order parameter is to the true reaction coordinate, the more efficient the method becomes. For the barrier method, an additional issue arises if the most probable exit path (MPEP) crosses a barrier more than once, so that it goes backwards. The original set of lookups moves simulations from a previous barrier to the current barrier *along the paths already discovered*. These lookups cannot move the simulations along the correct MPEP, as seen in Fig. 6.4. This problem can be fixed by allowing additional lookups to be created, if needed. For the barrier method applied to the generalized SIS model, below, we used this feature.

6.3.5 Algorithm assumptions and limitations

To gain efficiency, rare-event techniques typically make assumptions about the problem at hand. These assumptions frequently limit the models for which the technique can be applied efficiently. The performance gain of the barrier method is achieved by using previous trajectories to move simulations "forward" if they travel too far "back." These previous trajectories are an attempt to approximate the transfer function for that location. For discrete systems, the only assumption is that the transfer function for every site on every barrier can be well-approximated by a small number of samples on each barrier. The performance improvement of the barrier method over brute-force simulations is strongly dependent on that assumption. For continuous systems, another assumption is needed: the transfer functions in state-space must vary smoothly over each barrier. This assumption is needed because we pick the nearest lookup on the previous barrier. If the transfer function is significantly different between the location where the trajectory crosses the barrier and the

closest lookup, a systematic bias can result. This assumption is also strongly tied to efficiency: the smoother the transfer functions in state-space, the fewer lookups are required to well-approximate those transfer functions.

Another assumption made in the barrier method is that there is always a reasonable amount of spread between the barriers. If there is no parallel diffusion between barriers, i.e., paths started from a point on λ_i always reach a single point λ_{i+1}, then the lookups become deterministic and trajectories will tend to cycle between lookups, creating artificially long cycles before the next barrier is reached. The effect of this cycling is a dramatic increase in the variance of the transition time. We found this effect in the SIS model only when the birth-death rate μ is four orders of magnitude lower than the other rates in the system. However, we found no apparent problems when more comparable rates were chosen.

The final necessary assumption is that the model being studied is a memory-free (Markov) process. For example, if the dynamics of the Maier–Stein model depended on the particle velocity, the barrier method, as applied in this chapter, would not work. However, this problem can be solved by defining the state-space as the position and the velocity of the particle, making it a four-dimensional state space. In general, any finite-memory process can be made into a Markov process with a larger state-space [12]. Unfortunately, making the state-space larger poses a particular problem for the barrier method.

The efficiency of the barrier method is strongly dependent on the dimensionality of the state-space and smoothness of the transfer functions along every barrier. In the two models treated in this chapter, the state-space is two-dimensional and the barriers are one-dimensional, which made the sampling necessary to generate the lookups fairly simple. In higher dimensions, $\mathcal{D}$, one must well-sample the transfer

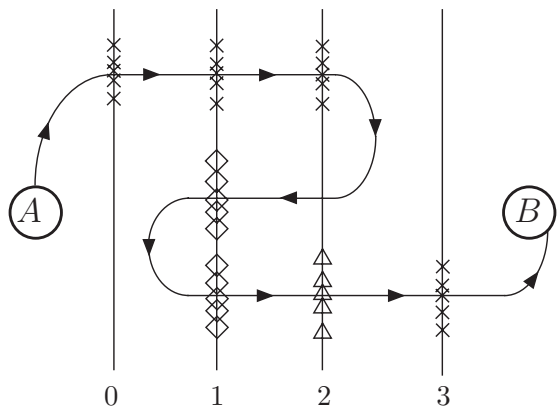

Figure 6.4: An example of the barrier method when the MPEP crosses barriers several times. The line from A to B represents the MPEP. The $\times$'s represent the initial locations where simulations reach a given barrier. These locations are used to generate the original set of lookups. The diamonds represent the new lookups generated by trials originated on the second barrier. Similarly, the triangles are new lookups generated by the trials started on the third barrier. Note that without adding new lookups, the simulation would never finish because the closest lookups on previous barriers would move the simulations backwards along the MPEP when they are moved to the next barrier.

functions on $(\mathcal{D}-1)$-dimensional manifolds. It appears that the number of lookups tend to grow exponentially with $\mathcal{D}$ and this will likely limit the barrier method to models with four or fewer dimensions in the state-space. However, many models exhibit a useful property: the most likely trajectories that connect the start region (A) to the end region (D) are focused along one or several "tubes" in state-space [95]. In these cases, it is only necessary to approximate the transfer functions in small regions on each manifold where the tubes cross. The barrier method will likely have poor performance on high-dimensional systems that lack this property.

In addition to assumptions as limitations, there are also practical considerations when applying the barrier method on a model. As with FFS, an order parameter, which marks progress toward the end region, must be chosen. When applying the barrier method on models with continuous state-spaces, one must define a distance

measure, or metric, on the barriers to determine which lookup is closest. Also, it is, in general, necessary to add more lookups. This requires setting some criterion for when a new lookup is needed, an example of which is given below for the SIS model. The best solutions for these design choices are dependent on the model being studied, and require some knowledge about the model's behavior.

6.4 Maier-Stein model

Maier and Stein [4] introduced an interesting example of a dynamic system which lacks detailed balance. It has received considerable theoretical [97, 5], experimental [98, 99], and computational interest [98, 100, 94, 91]. We study two aspects of this model: the mean exit time T from one of the metastable states and the distribution of exit locations along the separatrix.

The model is specified by two coupled stochastic differential equations

$$\begin{aligned} \dot{x} &= f_x(\mathbf{x}) + \xi_x(t), \\ \dot{y} &= f_y(\mathbf{x}) + \xi_y(t), \end{aligned} \tag{6.5}$$

where $\mathbf{x} = (x, y)$ and $\mathbf{f} = (f_x, f_y)$ is the time-independent drift field

$$\begin{aligned} f_x &= x - x^3 - \alpha x y^2, \\ f_y &= -\mu y(1 + x^2). \end{aligned} \tag{6.6}$$

For $\alpha = \mu$ the model obeys detailed balance. The white noise $\boldsymbol{\xi} = (\xi_x, \xi_y)$ has variance ϵ

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t+\tau)\xi_j(t) \rangle = \epsilon \delta_{ij} \delta(t - \tau). \tag{6.7}$$

We are interested in the small noise case: $\epsilon \to 0$. The model is bistable with the metastable states located at $\mathbf{x} = (\pm 1, 0)$. There is a separatrix at $x = 0$. The exit time, the transition time from one of the metastable states to the separatrix is, in

the small-noise limit

$$T \propto e^{W/\epsilon}, \tag{6.8}$$

where W is the generalized barrier height. In this model, W is a function of α and μ. If $\mu = 1$, for $1 < \alpha < 4$, there is a unique MPEP, and for $\alpha > 4$ there are two MPEP's.

6.4.1 Simulation

To simulate the system, we solve Eq. (6.5) using the Euler method

$$\begin{aligned} x(t+h) &= x(t) + h\left(x(t) - x(t)^3 - \alpha x(t) y(t)^2\right) + \sqrt{\epsilon h}, \\ y(t+h) &= y(t) + h\left(-\mu y(t)(1 + x(t)^2)\right) + \sqrt{\epsilon h}, \end{aligned} \tag{6.9}$$

where h is the time step.

6.4.2 Exit time and distribution of exit points

We start R simulations at the left metastable state $(-1, 0)$ and iterate Eq. (6.9) to obtain the trajectories. We dynamically locate each barrier. The first barrier is located at the average furthest excursion along x, the order parameter, using $S/4$ probe steps. All subsequent barriers are located after S steps. To calculate the exit time, we have three parameters at our control: the number of trials R, the time step h, and the spacing of the barriers controlled by S. To find W, we record the exit time for various values of ϵ and find the slope of $1/\epsilon$ versus $\ln T$.

6.4.3 Results

Using the barrier method, we measured W and the exit distributions along the separatrix $P(y)$. We computed W for $\mu = 1$ and for $\alpha = 1$ to $\alpha = 8$, and compared to the analytic theory[4]. These results are shown in Fig. 6.5.

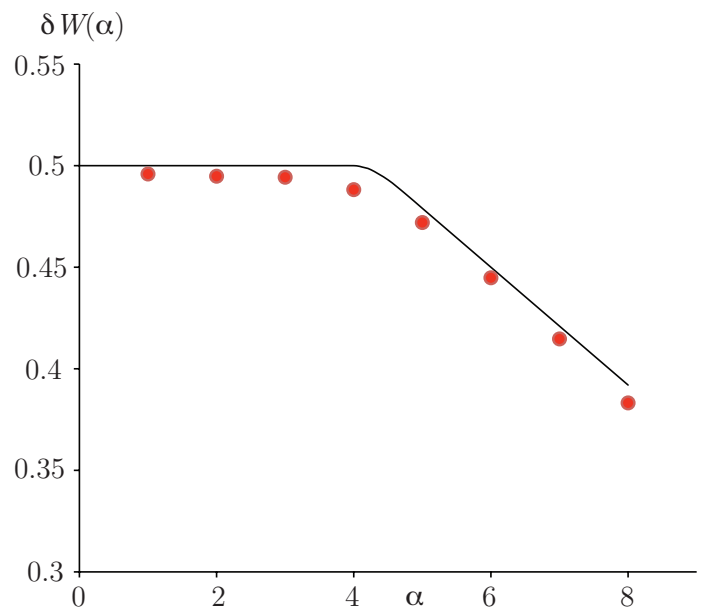

Figure 6.5: Barrier height, W, for the Maier–Stein model as a function of α for $\mu = 1$. Circles, barrier method; line, analytic theory (Ref. [4]) for $\epsilon \to 0$.

In order to use this model, we have to choose h, R, and S. These parameters have different, competing effects. Increasing h increases the efficiency but decreases the accuracy. The opposite is true for S. Also, effects of the values of h and S are connected; small S causes there to be a large number of barriers, which requires a small h to give an accurate result and vice versa. We used $h = 10^{-6}$, $R = 7 \times 10^4$, and $S = 4 \times 10^6$. However, we used a value of S which is four times smaller than the general case to locate the first barrier. The exit time was measured with five independent trials for five values of $1/\epsilon$: 20, 40, 60, 80, and 100, for each α. The calculated value of W and the small-noise theory [4] are shown in Fig. 6.5. The numerical results are consistently smaller than the theory and increase as $1/\epsilon$ increases. The values of W in Fig. 6.5 were consistent over other values of h, R, and S, that we tested.

We ran separate simulations to measure the exit distribution along the separatrix.

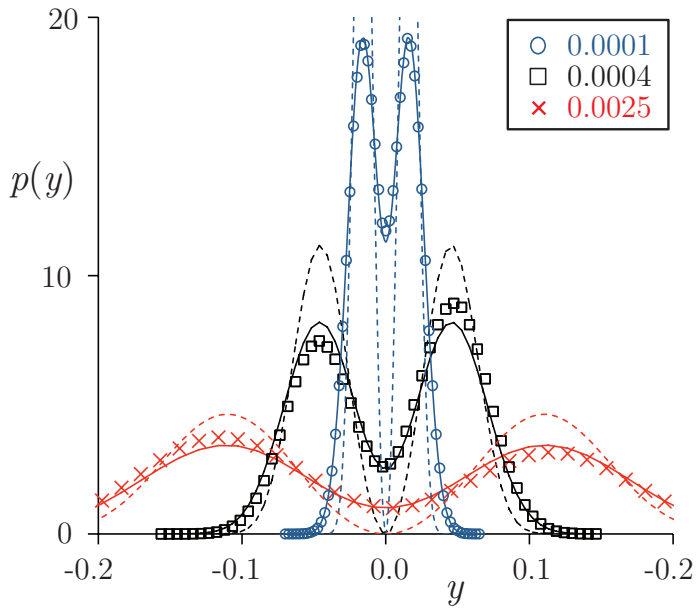

Figure 6.6: The exit distribution for the Maier–Stein model for $\alpha = 10, \mu = 0.67$ and $\epsilon = 0.0025, 0.0004, 0.0001$. The simulation results are denoted by symbols. The dashed line is the symmetrized Weibull distribution $P(y) = N|y|^{2/\mu-1} \exp\left(-|y/A|^{2/\mu}/\epsilon\right)$ from theory (Refs. [5, 6]), where N is the normalization and A is a parameter of order unity. Solid line: Weibull distribution convolved with a Gaussian with $\sigma = B\sqrt{\epsilon}$.

Because we are interested only in the final location along the separatrix and not the time it takes to reach it, we found that we could use significantly larger h. We were able to get reliable results for values of h as large as 0.0001. This allowed us to reach much smaller values of ϵ (see Fig.6.6). The parameter values used to obtain the results in Fig. 6.6 are $h = 0.0001$, $R = 10^5$, and $S = 4 \times 10^4$. The results were averaged over twenty independent simulations for each ϵ. The significant result is that for the smallest noise value, $\epsilon = 0.0001$, the value of $P(0)$ is not close to zero as the theory suggests. Rather the ratio of $P(0)$ to the maximum of $P(y)$ appears instead to be *increasing* as ϵ decreases.

These results are consistent with other simulation [100] and experimental [6] results. Previously, the results were assumed not to match the theory because the values of ϵ were not small enough. The barrier method allows us to reach a value of ϵ which is 50 times smaller than the best previous simulation result [100] and 110 times smaller than the best experimental result [6] without $P(0)/P_{\max}$ getting any closer to zero.

We propose that the reason for this is that the theoretical prediction of the Weibull distribution represents the leading term in ϵ. For finite ϵ, the distribution should be "rounded" over a scale $y = O(\sqrt{\epsilon})$ [6]. Accordingly, we convolved the asymptotic theory, the Weibull distribution given in the caption of Fig. 6.6, with a Gaussian with $\sigma = B\sqrt{\epsilon}$ (see Fig. 6.6). We find good agreement between the simulation results and the convolved theory. The values of B which gave the best fit were 0.8, 0.85, and 0.85 for $\epsilon = 0.0025$, 0.0004, and 0.0001, respectively. The fact that B is roughly constant over a wide range of ϵ gives support to our estimate.

The reason that $P(0)/P_{\max}$ does not tend to zero is that even though the rounding is over a scale that decreases as $\sqrt{\epsilon}$, the location of the maxima of $P(y)$ also moves

toward the origin. These locations are $y_{\max} = \pm 2^{-\mu/2} A \epsilon^{\mu/2} (2-\mu)^{\mu/2}$, so that $|y_{\max}| \propto A\epsilon^{\mu/2} \approx A\epsilon^{1/3}$ for $\mu = 0.67$. If A were constant, $P(0)/P(y_{\max})$ should approach zero rather slowly. We numerically found this rate to be $\epsilon^{0.6}$. However, we find that A is not constant. Our best fit values for A, which is unaffected by the convolution, are 0.94, 0.72, and 0.4 for $\epsilon = 0.0025$, 0.0004, and 0.0001, respectively, so that $P(0)/P(y_{\max})$ does not approach 0 in our computations.

6.5 Generalized SIS model

It is interesting to generalize the SIS model of Eq. (6.1) to allow the fluctuations of the total population by introducing birth and death rates [16, 101, 88, 7]. Now there are two independent stochastic variables; S the number of susceptibles and I the number of infected; and four parameters: μ the birth and death rate assumed equal, β the infectious contact rate, κ the disease-recovery rate, and N the steady-state population size. The transition rates are now

$$\begin{aligned} &W[(S,I) \to (S+1,I)] = \mu N, \quad W[(S,I) \to (S-1,I)] = \mu S, \\ &W[(S,I) \to (S,I-1)] = \mu I, \quad W[(S,I) \to (S+1,I-1)] = \kappa I, \\ &W[(S,I) \to (S-1,I+1)] = \beta S I / N. \end{aligned} \tag{6.10}$$

The model has an endemic state when $R_0 = \beta/(\mu+\kappa) > 1$. There is one stable fixed point, the endemic state $(S,I) = (NR_0^{-1}, N(1-R_0^{-1}))$, and an unstable saddle point $(N,0)$, where the disease is extinct. We seek the transition time from the endemic state to the disease-free state which will be of the form $T \sim \exp(NW)$, as above.

We are interested in the case of small μ so that population fluctuations are slow compared to disease dynamics. It might seem that the situation would be very similar to the case $\mu = 0$ treated above. However, this is not true [7]. Population fluctuations make extinction of the disease much easier: the most likely exit path is

via a population decrease at fixed S followed by extinction along a path of smaller fixed population, and then an increase of population of susceptible individuals to $S = N$.

6.5.1 Simulation

We simulate the SIS system using standard techniques [102, 103]. We define the order parameter as $-I$; the barriers are added at decreasing values of I. Because this system is on a two-dimensional parameter lattice it is much easier to dynamically add new lookups if previously unexplored backward regions are reached or if a given lookup has been used too frequently.

The method we use to add lookups is as follows. Every time a sample moves backwards and reaches the previous barrier, a new lookup for that site is generated with a probability $p_G = L_G/D(S, I)$, where L_G is a constant which controls the growth rate of the lookups and $D(S, I)$ is the number of lookups at site (S, I); if p_G is greater than 1, a new lookup is always added. If a new lookup value is needed, the value is produced by starting a path at (S, I) until it reaches the next barrier and stops or reaches the previous barrier. There a lookup is used to move back to the current barrier to continue. Note that a new lookup can move back to a previous barrier and cause another new lookup to be generated at *that* previous barrier. This cascading effect can continue until the first barrier is reached. This effect makes programming the algorithm more complex, but the cascading is necessary to produce accurate results when the MPEP crosses some of the barriers several times, as illustrated in Fig. 6.4.

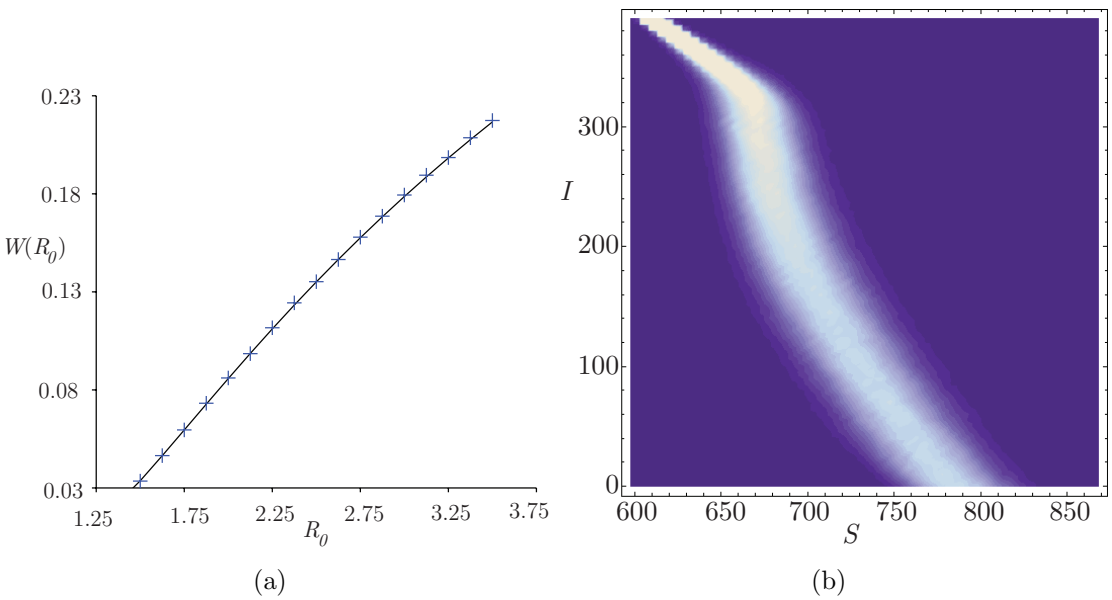

Figure 6.7: (a) Generalized barrier in the SIS model for $\mu = 1, \kappa = 100$ as a function of R_0. The line is the theoretical estimate (Ref. [7]). (b) Distribution of first-passage points on barriers for $N = 1000, R_0 = 1.667, \mu = 0.25$. The exit point on the line $I = 0$ is in good agreement with the prediction (Ref. [7]) $S \approx N/\sqrt{R_0}$.

6.5.2 Results

We found the exit time in the SIS model $T \sim \exp(NW)$ for $\mu = 1$ and $\kappa = 100$, and varied β to obtain different values of R_0. We compare to analytic work [101, 88, 7] in Figs.6.7 and 6.8. In these references the authors chose $\mu = 0.02$. They did this because the very large separation in time scales allowed them to use their analytic techniques.

We calculate W by performing a linear regression of $\ln T$ versus N, as in section 6.4. We ran ten simulations for each of the following N: 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500. For the last three values of N, we did this only for $R_0 < 2.5$. By looking at the residual of the fit, we found for smaller R_0 the smallest sizes were not large enough to reach a constant value of W. The smallest N included in the fits were 350, 350, 300, 300, 250, 250, 200, 200, and 200 for $R_0 = 1.5$, 1.625, 1.75, 1.875,

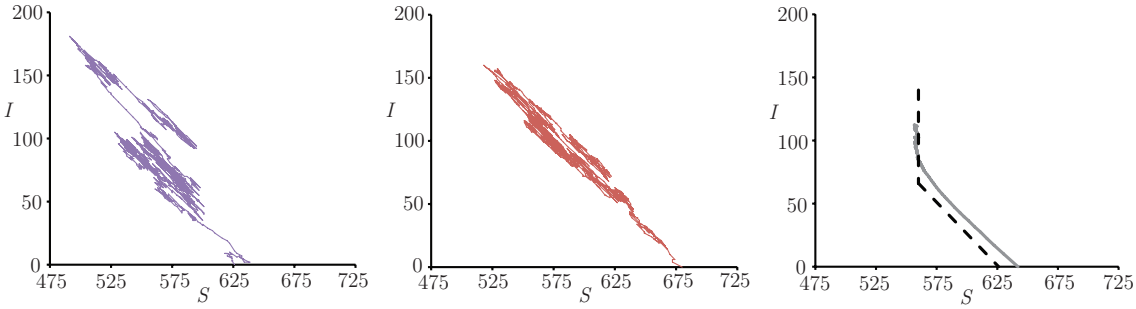

Figure 6.8: Exit paths for the SIS model for $N = 700, R_0 = 1.25$. Left panels, some sample paths; right panel, average over many paths to estimate the MPEP with a dashed line for the corresponding theory (Ref. [7]).

2.0, 2.125, 2.25, 2.375, and 2.5 and greater, respectively. The calculated values of W are shown in Fig. 6.7 and compared to analytic estimates [7].

The barrier method does not give the exit path directly. If we plot the distribution of the W^r's on the barriers, it is the distribution of first passage points. However, the paths that continue are not uniformly distributed on the barriers. Nevertheless, it is interesting to plot the first-passage distribution, Fig.6.7b.

We did a separate, brute-force computation to find the actual MPEP. A few sample paths are shown in Figs. 6.8(a) and 6.8(b), along with the average of many paths, 6.8(c).

6.6 Discussion

In this chapter, we developed a new rare-event technique: the barrier method. We described the relationship between it and other related techniques such as FFS and showed that for a simple model problem, the barrier method is more efficient than FFS.

The barrier method was then used to find exit times and distributions for the Maier–Stein model [4]. We found fairly good agreement with theory on exit times

and good agreement with previous simulation and experimental results on the exit distributions. We determined that convolving the theory with a Gaussian to account for the next correction to the theory gave excellent agreement for the distribution of exit points.

The exit times for a SIS model with births and deaths were then calculated. The results agreed with analytic estimates [101, 88, 7]. The MPEP was also found for this case.

The barrier method is an excellent tool to determine rare events in low dimensional systems. In this chapter, we have treated one and two degrees of freedom and we have preliminary work for three dimensions. We believe the most important aspect of the barrier method is the elimination of practically all "backtracking." This can enable the traversal of landscapes with many metastable states. The method is also general enough to apply to on-lattice and off-lattice problems, equilibrium and nonequilibrium problems, and any system that can be written as a nondeterministic Markov process.

The barrier method is a reliable way to estimate the exit time because it faithfully samples the ensemble of paths that lead to the goal region. For a given point on a barrier λ_i, there is an ensemble of paths that start from that point and eventually reach the next barrier λ_{i+1}, according to the dynamics of the model. These paths can backtrack an arbitrary long distance before reaching λ_{i+1}. The lookups in the barrier method approximate the transfer function that maps the single point on the λ_i^{th} barrier to a distribution of points, and their associated times, on the $\lambda_{i+1}^{\text{th}}$ barrier. Even though we only use the single closest point on λ_i to approximate the function, this approach is valid so long as the lookup points are spaced close enough together such that diffusion of a trajectory close to the barrier could cause the trajectory to

be closest to several different lookups once it touches the barrier. Because we are approximating the transfer function, and not *randomly* picking samples on previous barriers, we are making no assumption of loss of memory between barriers. In fact, in the SIS model discussed above, the lookups from barrier to barrier were strongly correlated, whereas the lookups in the Maier–Stein model were not strongly correlated. The barrier method performs well in both cases and does not depend upon the lookups being uncorrelated. While the barrier method may not work well for high-dimensional systems, we nevertheless believe that it can have wide applicability for many rare-event problems.

CHAPTER VII

Computation of Nucleation at a Nonequilibrium First-order Phase Transition Using a Rare-event Algorithm

The work in this chapter was published in Journal of Chemical Physics in 2010 (Vol. 133, Art. Num. 174107). This chapter is builds on the previous chapter. Here we develop an improvement of FFS, called forward flux sampling in time (FFST) and show the connection between FFS and the barrier method.

7.1 Introduction

In many systems, rare events occur with a very low probability compared to typical events. Sometimes they are of central interest. Examples include the extinction of diseases [10] or of populations [11], network queue overflow [8], and slow chemical reactions [12]. The study of such processes poses a particular challenge to simulation. In the field of chemical physics, many rare-event techniques are commonly used: transition path sampling [23], transition interface sampling [24], milestoning [22], the string method [95], and the weighted-ensemble method [104], to highlight a few methods. A thorough review can be found in Ref. [87]. Most of these methods require that the system being studied has an underlying energy landscape, which precludes their use on nonequilibrium systems, i.e., systems that lack detailed balance.

Forward flux sampling [21, 105, 106] (FFS) is a rare-event technique designed specifically for nonequilibrium systems and has proven useful in studying genetic switches [21, 66, 91], nucleation [91, 92, 93], isomerization of alanine dipeptide [107], and the Maier–Stein model of reaction dynamics [4, 91, 94]. In this chapter, we develop a variant on FFS and use it to study the first-order nonequilibrium phase transition in a catalysis model.

FFS was developed to measure transition rates between two locally stable regions A and B separated by a high, featureless barrier. When the barrier separating these regions contains long-lived metastable states, FFS is generally inaccurate or inefficient, depending on how it is applied. We have overcome this limitation of FFS with a variant, which we call forward flux sampling in time (FFST). In our method, we adjust for long-lived metastable states by measuring the times associated with sampling the region between A and B in the second stage of the FFS algorithm. The method is described in detail in Appendix A.

We apply FFST to the Ziff-Gulari-Barshad (ZGB) catalytic surface-reaction model [17]. This model is of interest because it has a first-order phase transition which acts in many ways like an equilibrium phase transition (it shows critical behavior, nucleation, etc.), but also the model is manifestly nonequilibrium. Many techniques have been applied to the transition in order to tease out its properties. FFST allows us to study the *dynamics* as well as the overall rates. Using FFST, we found transition times for nucleation as large as 10^{40} Monte Carlo steps (MCS), more than 30 orders of magnitude longer than those accessible to direct simulation. The method generates not only the transition times but the ensemble of most likely states as the system progresses from one phase to another. This allows us to measure properties of the ensemble during the transition, which helps determine the pathway.

The outline of this chapter is as follows. In Sec. 7.2 we describe the model and simulation method. In Sec. 7.3 we describe our results, and in Sec. 7.4 we summarize our conclusions. In Appendix A we give the details of the FFST technique, and in Appendix B we test FFST on an exactly solvable one-dimensional system.

7.2 Model and Simulation Method

7.2.1 The ZGB Model

The ZGB model was introduced to study the behavior of the oxidation of carbon monoxide (CO) on platinum surfaces [17]. The catalytic surface is represented by a square lattice on which CO and O_2 can adsorb. The CO takes up one lattice site, whereas the O_2 dissociates into two O atoms that take two adjacent vacant sites. Both species, CO and O, are bound to the surface until the other species adsorbs on a neighboring site. At this point, the CO and O form CO_2 and desorb from the catalyst, leaving two lattice sites empty. The rates of reaction $CO + O \rightarrow CO_2$ and desorption of CO_2 are assumed to be infinite. The state of the catalyst is controlled by the fraction of the time CO is attempted to be placed on the catalyst; this fraction is called p_{CO}.

In this basic model, there exists a region of steady-state reaction, bordered from below at $p_{CO} = p_1$ by a second-order, continuous kinetic phase transition to an O-covered state, and above at $p_{CO} = p_2$ by a discontinuous, first-order kinetic phase transition to a CO-covered state. The first-order transition is robust to small changes in the model (diffusion, small desorption, etc.) and is seen experimentally at lower temperatures as a sharp transition from high to low reactivity [108]. The second-order O-poisoning transition is weak and not seen experimentally. The first-order transition is associated with many complex oscillatory and wave phenomena in the-

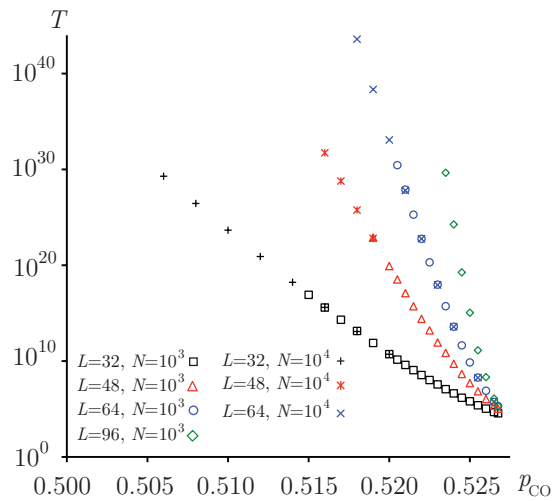

Figure 7.1: The poisoning time T vs p_{CO} for various values of system size L and FFST trials N.

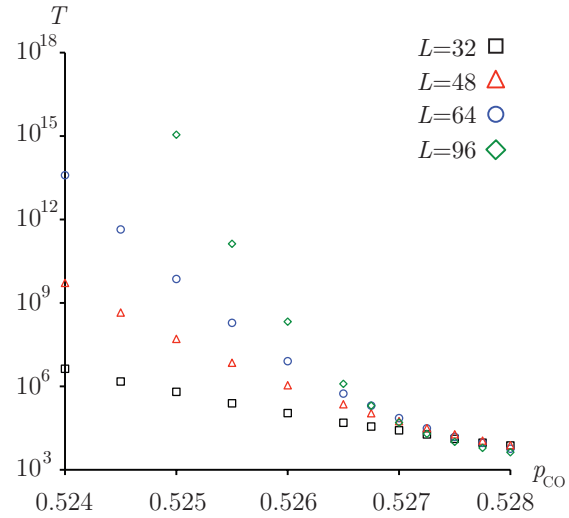

Figure 7.2: The poisoning time T vs p_{CO} for various values of L with $N = 10^3$.

oretical [109, 110] and experimental systems [111, 112] and has thus received much attention. It also serves as a paradigm for general first-order kinetic phase transitions [13].

Studying the first-order transition has proved to be a challenging problem in simulations. Because of the difficulty of nucleating a sufficiently large CO cluster or island, simply increasing p_{CO} from the reactive steady-state misses the transition point, and instead the CO-poisoning is seen to occur at $p_{CO} \approx 0.5277$ [113]. That point is close to the effective spinodal point p^*, where the transition occurs without any kind of barrier [114, 115].

Estimating the value of p_2 and properties of the first-order phase transition have received significant attention. Using a "constant-coverage" technique, the values $p_2 = 0.52560(1)$ [116, 114] and more recently 0.525615(5) [115] have been found. Other methods, including histograms [117], epidemic analysis [118], and epidemics

and trigger waves [119, 120], have also been used to probe the first-order transition.

The results of these studies are that there is a first-order transition at $p_2 \approx 0.5256$, and a spinodal at $p^* \approx 0.527 - 0.528$ [116, 115], although the somewhat lower (and very precise) value $p^* = 0.52675(5)$ has also been proposed [121]. For $p_2 < p_{\mathrm{CO}} < p^*$, there presumably exists a critical CO-cluster size, below which clusters tend to shrink and above which they tend to grow. That critical size changes from ∞ to something of the order of the lattice spacing as p_{CO} goes from p_2 to p^*. For finite systems, the behavior is controlled strongly by the boundaries. Using periodic boundary conditions, as p_{CO} is increased, one finds that the largest CO cluster goes from being isolated, to wrapping around one direction (leading to two interfaces that are flat on the average), to wrapping around in both directions, as illustrated in Fig. 7.8. In the intermediate coverage region, the constant-coverage method gives p_2 accurately with very small finite-size effects.

While the constant-coverage technique maps out the transition, it does not provide any information about the dynamics of the system. For that, it is necessary to study the standard (constant-rate) ensemble. However, in that case, the nucleation barrier makes it virtually impossible to study dynamics except for very close to the spinodal point. To overcome this problem, we use a modified FFS technique to find nucleation dynamics as well as overall rates for a wide range of p_{CO} values.

7.2.2 The simulation method

We study the ZGB model using the constant-rate ensemble on a square $L \times L$ lattice with periodic boundary conditions. The dynamics involve repeated attempts to adsorb the CO or O_2 species. The procedure for an adsorption trial is given below.

- Pick $r \in [0, 1)$, if $r < p_{\mathrm{CO}}$, attempt to place a CO molecule, otherwise attempt

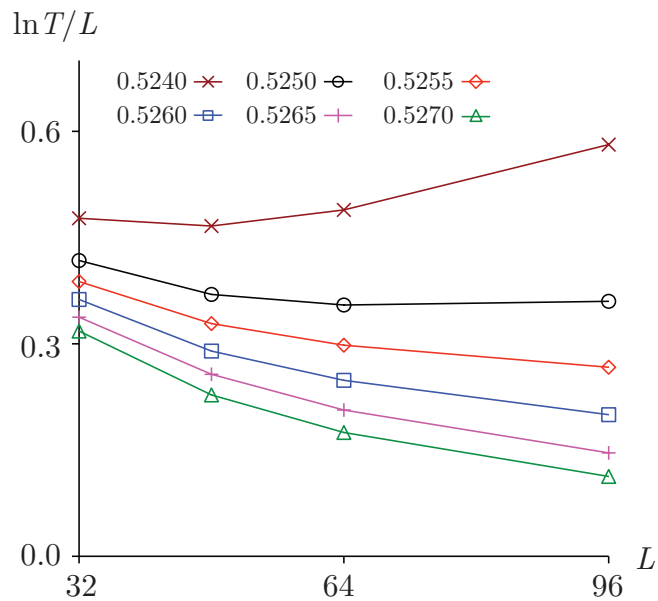

Figure 7.3: The scaled transition time $(\ln T)/L$ vs L for various values of p_{CO}.

to place O_2.

- Pick a random lattice site (x, y) and continue if the site is empty. If placing O_2, also pick a neighboring site $(x \pm 1, 0)$ or $(x, y \pm 1)$ and continue if that site is also empty.

- Place the CO or O_2 (dissociated) onto the empty lattice site(s).

- For each lattice site now occupied, check all of the neighbors of that site and determine if any of those neighbors are of the opposite species. If any are, remove the recently placed species and a randomly chosen neighbor of the opposite species.

Before any adsorption trials can be attempted, the system must be initialized. Because we wish to study the first-order phase transition, we prepare the system in the reactive state. Starting from an initially empty lattice with p_{CO} close to the

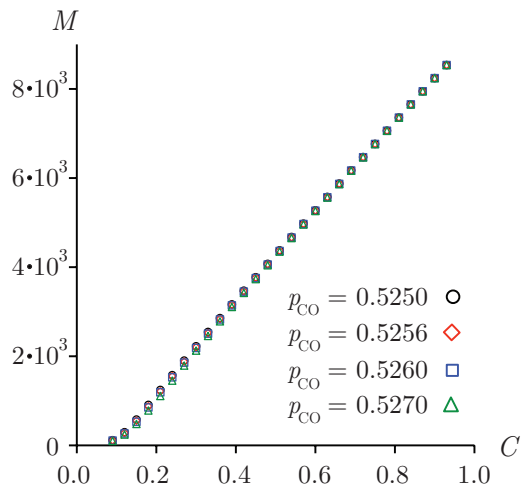

Figure 7.4: The average largest cluster size M vs coverage C for various values of p_{CO} and $L = 96$.

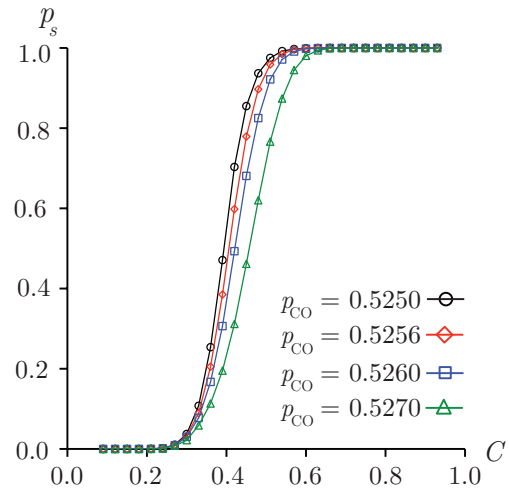

Figure 7.5: The spanning probability vs C for various values of p_{CO} and $L = 96$.

spinodal would frequently lead directly to the poisoned state. Instead, we prepare the system by adsorbing CO with probability $p^i_{\mathrm{CO}} = 0.07$ and O with probability $p^i_{\mathrm{O}} = 0.43$ on every site. This initialization generates invalid states with CO neighboring O. To remedy this problem, we run the simulation for 10 MCS, which drives the system to a valid state. These 10 MCS "burn off" a significant fraction of the original CO and O including those with neighbors of the opposite species.

7.2.3 Forward Flux Sampling in Time

To study the first-order phase transition, we use the FFST algorithm, which is described in detail in Appendix A. A sketch of the algorithm is as follows. Before the simulations begin, we define a starting region in state-space A and an ending region B, bounded by "barriers" λ_0 and λ_M, respectively. We also define λ_1 through λ_{M-1} as dividing surfaces in state-space that effectively mark the distance between A and B. The first step of the algorithm is to run a long simulation starting in A

and recording where the sample path crosses λ_0 going out of A, and the average time spent after crossing back into A before leaving again; we call this the internal return time T_{int}. Next, we start sample paths at different points along λ_0, where the initial simulation crossed, and continue them until they reach λ_1 or go back into A. The fraction of paths that reach λ_1 gives an estimate of the probability of reaching λ_1 without going back into A, $P(\lambda_1|\lambda_0)$. We also keep track of the average time it took to reach λ_1 and go back into A starting from λ_1. In the second step, we continue paths from the locations along λ_1, where the previous paths stopped, and run them until they reach λ_2 or go back inside A, crossing λ_0. The results give estimates for $P(\lambda_2|\lambda_1)$ and the time it takes to reach λ_2 or A from λ_1. This process is repeated, step by step, until λ_M is reached on the Mth step. Finally, we use the results collected to calculate the overall transition time, which is given by the probability of reaching λ_M from λ_0 without going into A ($P(\lambda_M|\lambda_0)$) times the average time it takes to leave A. The overall transition probability is given by the product of the intermediate transition probabilities, $P(\lambda_M|\lambda_0) = \prod_{i=0}^{M-1} P(\lambda_{i+1}|\lambda_i)$. The time to leave A is the average time it takes to return to λ_0 from inside A, T_{int}, plus the average time it takes to return to A from outside λ_0, T_{ext}, which we calculate from the times measured during the second stage of the algorithm. The gains of FFST over FFS are illustrated in Appendix B.

Thus, the technique follows fruitful paths from the reactive state to the CO-poisoned state. To apply the FFST to the ZGB model, we must first define an order parameter, which is used to determine progress towards the poisoned state. This function should smoothly increase as the system transitions from being reactive to poisoned. We chose the fraction of CO on the lattice (C) as our order parameter.

FFST uses the order parameter to make surfaces (barriers) in state space which

are used to mark progress. Simulations are run from locations along a barrier in state space until they reach the next barrier, as defined by the order parameter, or the first barrier.

These barriers can be placed before the sampling begins, which we call *static* barriers. The disadvantage of static barriers is that without *a priori* information about how to best place them, some barriers will have a large effective separation, creating a performance bottleneck. To overcome these performance issues, two methods have recently been developed [96, 40], which determine where to place the barriers automatically. To measure the poisoning time, we use the dynamic barrier placement method [40] to place the barriers during the FFST algorithm. We use static barriers for measurements of the average largest cluster, largest cluster spanning probability, and committor probability [23]; these terms are defined below. We use static barriers for the observables because we want measurements at uniformly spaced intervals of coverage. When placing the barriers dynamically, we space them such that typically 10% of trials make it to the next barrier before returning to the steady state of the reactive region with low CO coverage. We use N trials per step in the FFST algorithm, except for the first barrier, which we use $10N$; we use $N/10$ trials to determine where the next barrier should be placed. We locate the first barrier at the largest value of the order parameter found in the first 50 MCS. For the calculation of poisoning times, we used $N = 10^3$ and 10^4. $N = 10^4$ was used for average largest cluster, largest cluster spanning probability, and committor probability.

7.3 Results

We desire the transition times from the reactive to poisoned state, which we call the *poisoning time* T. We determine T for a range of L: 32, 48, 64, and 96, as well

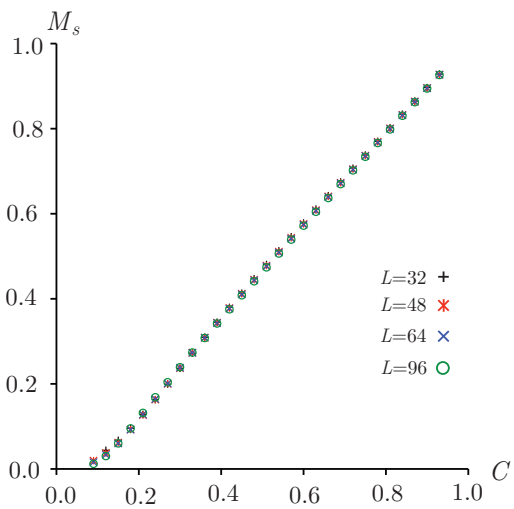

Figure 7.6: The average largest cluster size M_s vs. coverage C, as a function of L for $p_{\mathrm{CO}} = 0.5256$.

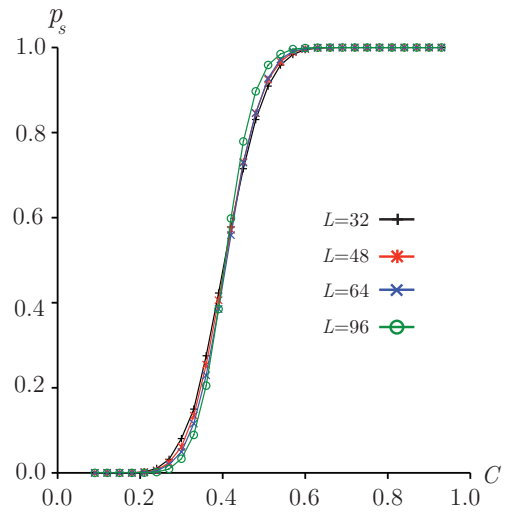

Figure 7.7: The spanning probability p_s vs. coverage C for various L and $p_{\mathrm{CO}} = 0.5256$.

as a number of values of p_{CO} in the critical region, $p_2 < p_{\mathrm{CO}} < p^*$, and in the reactive region, $p_{\mathrm{CO}} < p_2$ down to $p_{\mathrm{CO}} = 0.505$. The results are shown in Figs. 7.1 and 7.2.

Figures 7.1 and 7.2 show that the poisoning time grows smoothly as p_{CO} decreases, with no indication of a transition at p_2. Even if p_{CO} is much smaller than p_2, we found it possible to measure the poisoning time, demonstrating that the reactive state for finite systems is always metastable. We were able to determine maximum poisoning times of $T = 10^{30}$ - 10^{40} MCS. The poisoning times converge to a value that is independent of system size for $p_{\mathrm{CO}} \approx 0.5275$, as shown in Fig. 7.2.

Classical nucleation theory [122], which has been recently used to study the first-order nonequilibrium phase-transitions in the ZGB model [109], the quadratic contact process [123], and Schloegl's second model[124], states that transition (poisoning) times take the form $T \propto v^{-2/3}k^{-1/3}$, where v is the growth velocity of a critical droplet and k is the critical droplet production rate per unit area per unit time. The

above expression is valid when the characteristic length $l = (v/k)^{1/3}$ is much larger than the critical droplet radius and much smaller than the system width, a condition which should hold near p_2 for large enough system sizes. The largest system size we use in this chapter is $L = 96$, which is much smaller than most critical droplet sizes for p_{CO} near p_2. On the other hand, near p^*, the critical droplet size and other length scales become small enough that finite-size effects disappear. This leads us to believe that the spinodal is at $p_{CO} \approx 0.5275$. The poisoning times for all sizes were also checked using brute-force simulations slightly above the spinodal: $p_{CO} = 0.5265$ and 0.52675, and identical results were found.

We attempt to find the form of the transition time as a function of L and p_{CO}. In many nonequilibrium systems, transition times take the form $T \approx e^{WL}$, where W is an effective energy barrier and L characterizes the size of the system. Figure 7.3 shows $(\ln T)/L$ which is an estimate for W. We see that $(\ln T)/L$ does not have a strong system size dependence near p_2, which may indicate that there is an effective Arrhenius energy. If this were the case, then $(\ln T)/L$ would become independent of L for very large system sizes.

Although FFST is primarily a tool for obtaining transition times, it also gives ensembles of states along each barrier as the system progresses towards poisoning. These ensembles at a fixed value of coverage represent essentially what the constant-coverage ensemble attempts to mimic. Expectation values of quantities, like largest cluster size, can be taken over these ensembles, which give insight into the dynamics of the phase transition.

We ran FFST with evenly spaced static barriers and measured the average largest cluster size and the spanning probability for the ensemble captured on every barrier. Spanning occurs when the largest cluster wraps around the periodic boundary and

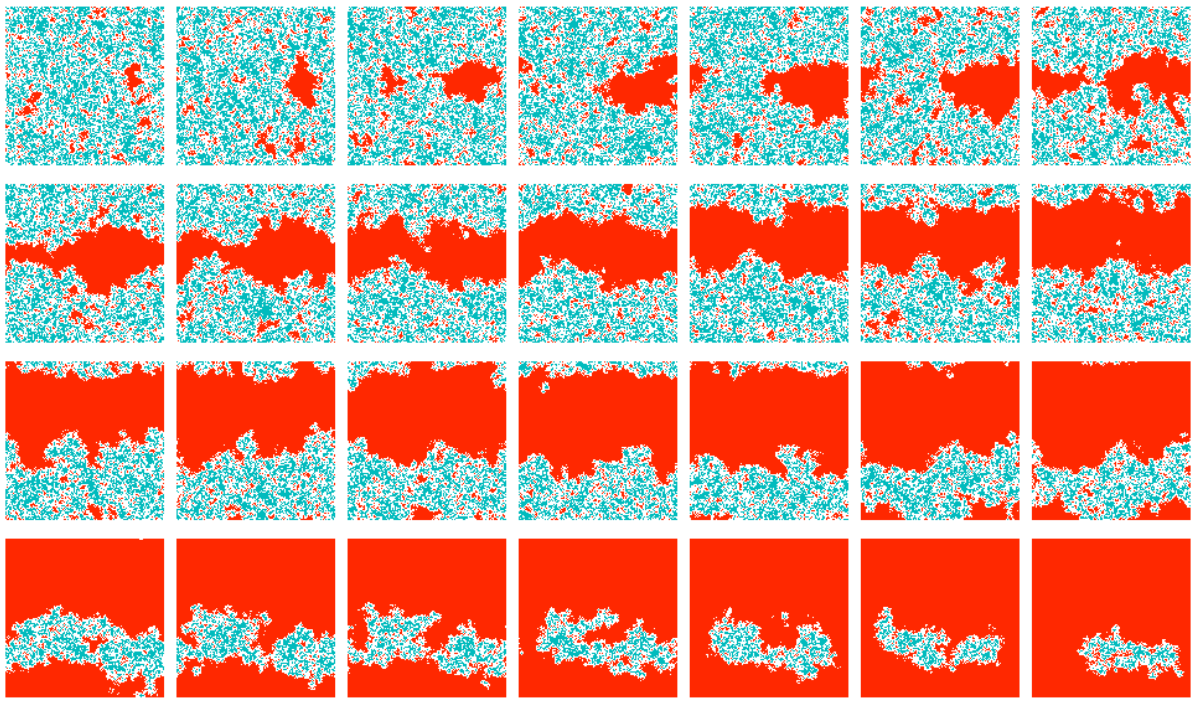

Figure 7.8: The evolution of the most likely path to the poisoned state for $L = 128$ and $p_{\mathrm{CO}} = 0.5256 \approx p_2$. The pictures are in increasing values of coverage fraction from left to right starting at 0.09 and increasing in steps of 0.03.

touches itself. We see a dependence on p_{CO} and L upon the spanning probability, as shown in Figs. 7.5 and 7.7. Figure 7.5 shows that the smaller the value of p_{CO}, the more likely that the largest cluster will wrap earlier in the path to poisoning, at fixed L. Figure 7.7 shows that increasing L appears to narrow the range of coverage for which spanning has a non-negligible probability of occurring or not occurring, at fixed $p_{\mathrm{CO}} \approx p_2$. This suggests that the variation in shape of the largest cluster decreases with increasing L. It also appears that reaching a spanning probability of 50% is achieved at $C \approx 0.42$ independent of L at p_2. This indicates that the average largest cluster is significantly elongated when it begins to span the system, as seen in Fig. 7.8. We also found a linear relationship between the scaled average largest cluster size and coverage, independent of both L and p_{CO}, as shown in Figs. 7.4 and

7.6.

We believe this analysis sheds light on the preferred pathway to poisoning for finite-sized systems with periodic boundaries. For systems with nonperiodic (that is, open) boundaries and $p_{CO} > p_2$, we expect clusters to form along the walls because the open boundary tends to favor CO adsorption over O_2 adsorption. For smaller systems, these outer clusters (which have an net negative curvature) will expand to poison the system, and therefore the poisoning should occur more quickly than in the case of periodic boundary conditions. We have tested this at $p_{\mathrm{CO}} = 0.5265$ on systems with $L = 32$ and 96, and indeed find poisoning for the nonperiodic case is two orders of magnitude faster than for the periodic case. For very large nonperiodic systems, on the other hand, we expect the wall effects to be small because many critical clusters can form in the bulk before being overtaken by a wall cluster. That is, we expect very large periodic and non-periodic systems will both form many critical clusters that merge as the system poisons (the so-called Avrami process), and thus have similar poisoning times [109, 123, 124].

The ensembles of states at different values of the coverage can also be used to directly measure progress towards poisoning. By running every state in the ensemble until it returns to the reactive (A) or poisoned (B) state, one obtains the probability of poisoning from these particular values of the coverage. This probability as a function of the order parameter is called the *committor* probability, p_B. (In the case of the ZGB model, this is only an estimate because the model is nonequilibrium, so the forward-tending ensemble obtained is not necessarily the same as the steady-state nonequilibrium distribution along the barriers.) We measured the committor probability for various L and p_{CO}, as shown in Fig. 7.9. We find that for p_{CO} below p_2, the larger the system, the larger the coverage must be in order to have a

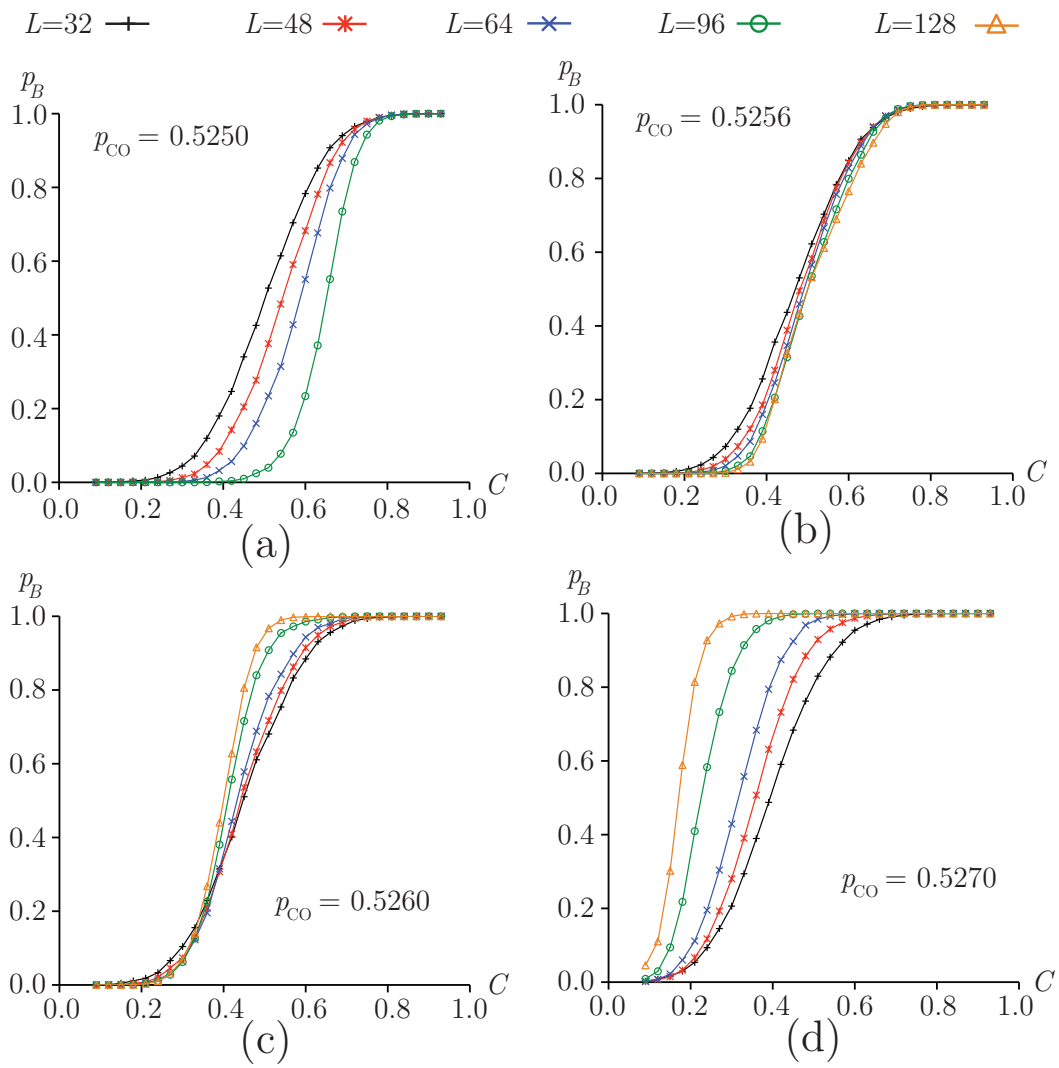

Figure 7.9: The committor probability p_B vs C for four values of p_{CO} and different L. $N = 10^4$ for all data shown.

particular probability of poisoning. The opposite effect is found above p_2. At p_2, the committor probability tends to a single form for the largest system sizes. This shows the significance of p_2, as the point where p_B is independent of L.

By keeping track of which state on a given barrier is responsible for a particular state on the next barrier, we were able to piece together complete trajectories from the reactive state to the poisoned state. Among these reconstructed paths, it is possible to determine a most likely path. Every state j_i on every barrier λ_i has some

probability of continuing to the next barrier before returning to the reactive state denoted by $P^j(\lambda_{i+1}|\lambda_i)$, which is measured during FFST. We define the most likely path as the path connecting λ_0 and λ_M which has the largest value of the product of the intermediate barrier crossing probabilities, $\prod_i P^j(\lambda_{i+1}|\lambda_i)$. The states in Fig. 7.8 show the most likely path for $L = 128$ at the transition point p_2. The preferred path to poisoning involves wrapping around the system and then expanding to complete CO coverage. This path is favored because nucleating droplets have an effective kinetic surface tension which causes droplets to be unfavored. Once a droplet spans the system and the net curvature disappears, the cluster is significantly more favored. This preferred pathway of wrapping and expanding has also been seen in magnetic memory switching [125]. This behavior is to be expected, at least near $C \approx 0.5$, where the cluster usually wraps around in one direction. In that case, at the transition point p_2, the system should be equally likely to poison or return to the reactive state, so $p_B \approx 0.5$ independent of L.

7.4 Conclusions

In this chapter, we introduce forward flux sampling in time (FFST). We use it to analyze the first-order phase transition in the ZGB model. We found a size-independent poisoning time at $p^* \approx 0.5275$, which we associate with the spinodal point. The poisoning time is a continuous function of p_{CO} near the first-order transition for the relatively small system sizes we study. For larger systems, one would expect a more pronounced change in behavior of the poisoning time at p_2, because below p_2, very large clusters will grow at a constant rate and therefore poisoning will take place relatively quickly once a large cluster is formed. However, we expect one would have to go to systems of perhaps thousands of lattice spacing across in order

to see this effect.

By inspecting the ensembles of states measured at each barrier, we found a linear relationship between the scaled average largest cluster size and the total CO coverage, which is practically independent of L and p_{CO}. When the paths from barrier to barrier are connected, they make the ensemble of successful trajectories. We determined the most probable of these trajectories and found that wrapping and then expanding is the preferred path to poisoning.

At $p_{\mathrm{CO}} = 0.5256$, the committor probabilities appears to be independent of system size for large L. We believe that this is a signature of a first-order transition for the following reasons: first, the committor probability is 0.5 for half coverage, which is has been previously used to determine the transition point [17]. Second, the matching of the entire curves for various L stems from the critical cluster size being infinite at the transition point. This implies that the probability of a droplet growing is always less than 1/2. The growth probability cannot be strongly dependent on the cluster size as clusters of arbitrarily large size must all have roughly the same growth probability, slightly less than 1/2. For large lattice sizes, what matters is how close the cluster is to spanning, which is only dependent on the mass of the largest cluster scaled by the total system size. Thus, we have found that the committor probability can be used to locate the first-order transition point of the system. Lastly, we found size-dependent poisoning times for systems well below the transition point, reflecting the difficulty to nucleate a cluster the size of the system in this regime.

For the ZGB model, we found that the efficient implementation of FFS, i.e., a fixed number of crossings in the first stage, gives accurate results. This shows that ZGB does not have extremely long-lived metastable states, which is the primary advantage of FFST. However, we still found that FFST can outperform FFS in

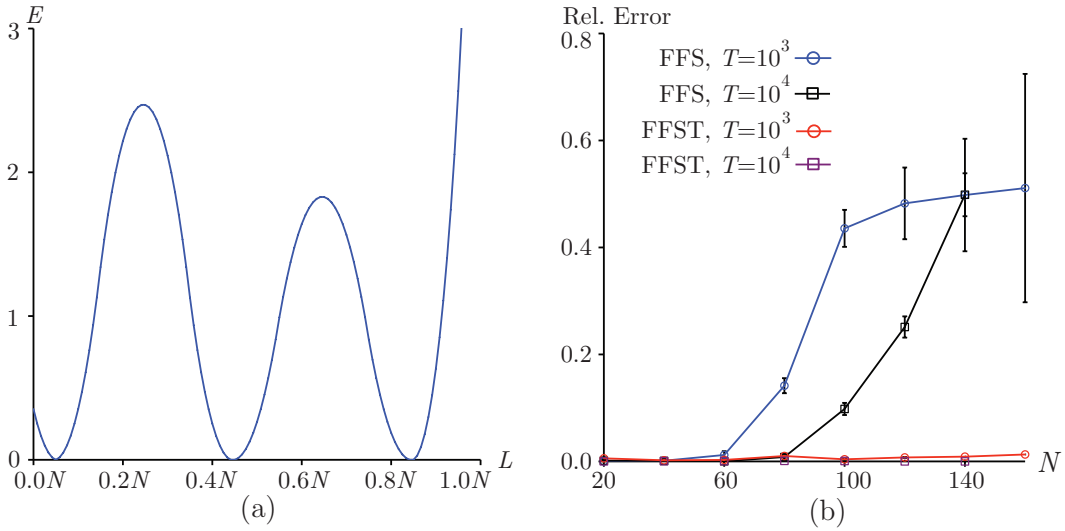

Figure 7.10: (a) A sketch of the effective energy landscape as a function of scaled system length L for a model one-dimensional problem. (b) The relative error of the estimated transition time as a function of system length N.

terms of smaller variance, which translates into better computational efficiency to reach a target variance. Specifically, we found that FFST was effectively 35% more efficient than FFS for a test case.

7.5 Appendix A: The FFST algorithm

The FFS algorithm works well for problems with a featureless barrier, i.e., problems with infrequent but fast transitions. However, in the case of problems with *slow* transitions, typically caused by long-lived metastable states, it has significant shortcomings [106]. In these cases, FFS either grossly underestimates the transition time or becomes nearly as inefficient as direct simulation (see Appendix B for an example). In the "inefficient" approach, where all of the times in the system are sampled in the flux rate calculation, FFS produces a constant gain in simulation efficiency, roughly ten times faster. The "efficient" approach, where a fixed number of crossings is used to calculate the flux rate, can be about as efficient as FFST and

reach gains of 10^{30}, as found in this chapter. In this Appendix, we derive the version of FFS, which we call FFST. We then outline two other variants on FFS which helps us illustrate the connection between FFS and the barrier method, which we recently introduced [40].

As we explained above, FFS separates the problem of finding rare transition events into two steps. The first step is to find the rate of leaving the initial region A. Finding the probability of reaching the final region B from the surface of A without going back into A is the second step. Finding the rate of leaving A typically involves running a single long simulation until it has exited A a fixed number of times, e.g., $10N$, where N is the number of trials per barrier in the second step. The number of crossings divided by the total time of the simulation (discounting paths that reach B) gives an estimate of the rate of leaving A. This calculation of the flux is accurate only if the trajectory captures the important times and features of the entire landscape. For example, if the trajectory does not sample a long-lived, metastable state between A and B, the estimated flux would be higher than the true flux. We show such an example in Appendix B.

A simple and correct way to avoid this problem is to run the initial simulation until it reaches B one or more times. However, this is impractical because it effectively solves the problem using brute force, which is what FFS was designed to avoid, and would lead to miniscule efficiency gains over brute-force simulations. Our version of FFS has the accuracy of the "correct" calculation and the efficiency of the usual flux calculation.

7.5.1 Forward Flux Sampling in Time

FFST, the algorithm we use in this chapter, performs FFS in terms of transition times instead of rates. The problem is decomposed in the same way as FFS [21], utilizing the idea of endpoint regions A and B. In this view, there are three important times: the time to return inside A from the surface of A (T_{ext}), the time to exit A from just inside A (T_{int}), and the time to reach B from the surface of A without going back into A (T_f). The only additional quantity needed to calculate the transition time, T_{tot}, is the probability p of reaching B from the surface of A. Then

$$T_{tot} = \left(\frac{1}{p} - 1\right)(T_{ext} + T_{int}) + T_f. \tag{7.1}$$

If the first term, which represents the total time spent on unsuccessful attempts to reach B, is much larger than the time of a successful attempt to reach B, which is explicitly used in the construction of FFS, $[(1/p) - 1](T_{ext} + T_{int}) \gg T_f$, and p is small, then the above formula simplifies to the inverse of the FFS formula,

$$T_{tot} \approx \frac{1}{p}(T_{ext} + T_{int}) = \frac{1}{P(\lambda_M|\lambda_0)}\left(\frac{h_A}{\Phi_{A,0}}\right) = \frac{1}{k_{AB}}. \tag{7.2}$$

The efficiency gain of FFST comes from measuring T_{int} during the first step and measuring T_{ext} and p in the second step. This is because the features that would require a long flux rate calculation to capture are instead captured in the second step, allowing for less computational effort in the first step. Measuring T_{int} involves counting the time that the initial simulation spent in region A. In the second step, paths are allowed to run from the ith barrier to the $(i+1)^{\text{th}}$ or back to A. We record the times to make these transitions to the next barrier (t_i^{i+1}) and back into A (t_i^0) in addition to the probability of making it to the next barrier $P(\lambda_{i+1}|\lambda_i)$. T_{ext} can be

calculated by properly weighting the time it takes for different paths to return to A,

$$\begin{aligned} T_{ext} &= \sum_{i=0}^{M-1} P(\lambda_i|\lambda_0)P(\lambda_0|\lambda_i)\left(t_0^i + t_i^0\right) / \left(1 - P(\lambda_M|\lambda_0)\right) \\ &= \left[\sum_{i=0}^{M-1}\left(\prod_{j=0}^{i-1} P(\lambda_{j+1}|\lambda_j)\right)\left(1 - P(\lambda_{i+1}|\lambda_i)\right)\left(\sum_{k=0}^{i-1} t_k^{k+1} + t_i^0\right)\right] \\ &\Big/ \left(1 - \prod_{i=0}^{M-1} P(\lambda_{i+1}|\lambda_i)\right). \end{aligned} \tag{7.3}$$

The numerator in the last expression is the sum over return times of all paths sorted by largest excursion. The probability that a path will start at λ_0, make it to λ_i, and then return to λ_0 without making any further progress is the product of the probability of reaching λ_i ($\prod_{j=0}^{i-1} P(\lambda_{j+1}|\lambda_j)$) and the probability of then returning to λ_0 without reaching λ_{i+1} ($1 - P(\lambda_{i+1}|\lambda_i)$). The average time of this excursion is the time it takes to reach λ_i ($\sum_{k=0}^{i-1} t_k^{k+1}$) plus the time it takes to return to λ_0 without making any more progress (t_i^0). The term in the denominator is for normalization and represents the probability of returning to A. The expression for T_f is much simpler: $T_f = \sum_{i=0}^{M-1} t_i^{i+1}$. These values of T_{int}, T_{ext}, and T_f can be combined with $P(\lambda_M|\lambda_0)$ in Eq. (7.1) to obtain T_{tot}. By measuring T_{ext} in the second step of FFS, the value of T_{tot} obtained will be comparatively more accurate and have less variance than FFS. The only extra work done in FFST over FFS is keeping track of the simulation time during the second step, which makes the advantages gained by FFST practically free.

7.5.2 Forward Flux with Quasi-Markov Dynamics

In order to gain more insight into the connection between the FFS and the barrier method, we formulate FFS as quasi-Markov dynamics. That is, we iteratively calculate the time it takes to travel between three barriers. The first step is similar to FFS: run a single long simulation and calculate the time it takes to reach the first barrier (λ_1) from the surface of A (λ_0), and where along λ_0 the trajectory exits A.

N samples are started along λ_1, where the initial trajectory crossed in the first step, and are run until they reach λ_2 or λ_0. The average times to go from λ_0 to λ_1, λ_1 to λ_2, and from λ_1 to λ_0 are t_0^1, t_1^2, and t_1^0, respectively. In general, the time takes to reach λ_j from λ_i is given by t_i^j. The probability of reaching λ_2 from λ_1 without first going back to λ_0 is $P(\lambda_2|\lambda_1)$. These times and probabilities can be used to make a random walk with three states: λ_0, λ_1, and λ_2. The transition time from λ_0 to λ_2 (t_0^2) is given by the weighted times of all possible paths from λ_0 to λ_2. These paths can be organized by the number of times they return to λ_0. Writing out the first few terms in this series exposes the general form,

$$\begin{aligned} t_0^2 &= (t_0^1 + t_1^2)P(\lambda_2|\lambda_1) \\ &\quad + (t_0^1 + t_1^2 + (t_1^0 + t_0^1))P(\lambda_2|\lambda_1)(1 - P(\lambda_2|\lambda_1)) \\ &\quad + (t_0^1 + t_1^2 + 2(t_1^0 + t_0^1))P(\lambda_2|\lambda_1)(1 - P(\lambda_2|\lambda_1))^2 + \dots \\ &= \sum_{k=0}^{\infty}(t_0^1 + t_1^2 + k(t_1^0 + t_0^1))P(\lambda_2|\lambda_1)(1 - P(\lambda_2|\lambda_1))^k. \end{aligned} \tag{7.4}$$

The above expression is the sum over all possible ways to reach λ_2 from λ_0, sorted by the number of times the simulation returned to λ_0. The probability of returning to λ_0 k times before reaching λ_2 is given by the product of the probability of not reaching λ_2 k times ($(1 - P(\lambda_2|\lambda_1))^k$) and then reaching λ_2 on the $(k+1)^{\text{th}}$ try ($P(\lambda_2|\lambda_1)$). The time this takes is given by the sum of the time it takes to go from λ_0 to λ_1 to λ_0 k times ($k(t_1^0 + t_0^1)$) plus the time it takes to make it from λ_0 directly to λ_2 ($t_0^1 + t_1^2$). Note that in practice the sum converges quickly because of the factor, $(1 - P(\lambda_2|\lambda_1))^k$.

The next step is to repeat the same process using λ_0, λ_2, and λ_3 as the three barriers. Brute-force dynamics is used to measure t_2^3, t_2^0, and $P(\lambda_3|\lambda_2)$. Then, the three-barrier calculation from Eq. (7.4) is used with 0, 2, and 3, in place of 0, 1, and

2. The result from the calculation is an estimated value of t_0^3. In general to calculate t_0^{i+1} for the ith step we use

$$t_0^{i+1} = \sum_{k=0}^{\infty} \left[t_0^i + t_i^{i+1} + k(t_i^0 + t_0^i)\right] P(\lambda_{i+1}|\lambda_i)(1 - P(\lambda_{i+1}|\lambda_i))^k. \tag{7.5}$$

This process of performing short brute-force simulations, followed by solving Eq. (7.5), is repeated for every barrier until λ_M is reached. At the end we have $t_0^M = T_{tot}$. This method could be useful for practical simulations. Here we introduce it as a pedagogical device to show that by making small changes to the FFS algorithm, the barrier method can be effectively obtained.

7.5.3 Forward Flux Barriers

Our last FFS variant is a different take on quasi-Markov dynamics. We start by measuring the average time it takes to reach λ_1 starting at λ_0 (t_0^1), while also keeping track of where along λ_1 the trajectory exits A during a single long simulation. The paths are continued from λ_1 until they reach λ_2 or λ_0. If a sample reaches λ_0, then it is restarted at λ_1 at one of the locations where paths ended in the first step, and t_0^1 is added to the time. This process is continued until all samples reach λ_2. We now have an estimate of t_0^2 and the locations along λ_2 where the sample paths ended. From these locations, the paths are continued until they reach λ_3 or λ_0. If they reach λ_0 they are restarted at a location where a previous path stopped at λ_2 and t_0^2 is added to the time. This step is finished once all sample paths reach λ_3. The general step is to start the paths on λ_k and run them until they reach λ_{k+1} or λ_0. If a path reaches λ_0, it is restarted at λ_k with t_0^k added to the time. The step is complete when all paths reach λ_{k+1} and the time gives t_0^{k+1}. This process is repeated until λ_M is reached. The result is a value for t_0^M which is the estimate for the transition time.

This construction shows the relationship between FFS and the barrier method

[40]. This version of FFS measures the average time it takes to reach each barrier during the algorithm, as in the barrier method. Also, both methods avoid characterizing the flux rate from the surface and the transition probability. The main difference between this algorithm and the barrier method is that the barrier method does not need the simulation to go all the way back to λ_0 before jumping back to the current barrier; only the previous barrier need be reached. This is the source of the performance gains of the barrier method. However, this method is currently only tractable for low-dimensional systems as it requires a reasonable sample of previous barriers.

7.6 Appendix B: Testing FFS and FFST on an exactly solvable problem

In this appendix, we use a simple one-dimensional system to show that FFST can give accurate results for transition times when FFS fails. We also briefly discuss the comparative efficiency of the different algorithms applied to the ZGB model.

Consider a discrete hopping process on a line of length L. The probability of jumping from the ith to the $(i-1)^{\text{th}}$ site is p_i. The time of a jump is unity. In the cases where the first site ($i = 0$) is adsorbing and the last site is reflecting ($p_{L-1} = 1$), the system can be solved exactly [3]. The solution can be written in terms of hopping rates instead of hopping probability. λ_i and μ_i are the rates of hopping from i to $(i+1)$ and $(i-1)$, respectively. In terms of p_i these are $\mu_i = p_i$ and $\lambda_i = 1 - p_i$. The average time to reach the adsorbing site $i = 0$ from site $i = n$ is [3]

$$\begin{aligned}\tau_n &= \sum_{m=1}^{n}\left[\frac{1}{\mu_m} + \prod_{i=1}^{m-1}\frac{\mu_i}{\lambda_i}\sum_{j=m+1}^{L-1}\frac{1}{\mu_j}\prod_{k=1}^{j-1}\frac{\lambda_k}{\mu_k}\right], \\ &= \sum_{m=1}^{n}\left[\frac{1}{p_m} + \prod_{i=1}^{m-1}\frac{p_i}{1-p_i}\sum_{j=m+1}^{L-1}\frac{1}{p_j}\prod_{k=1}^{j-1}\left(\frac{1}{p_k} - 1\right)\right].\end{aligned} \tag{7.6}$$

This equation has been used to find the extinction time of a disease within a population in a simple model from epidemiology [3]. Equation (7.6) is general and can be used to construct "energy landscapes." We create a landscape with nonuniform hopping probabilities such that there are three metastable states, regions A, B, and C, all with roughly equal stability, as sketched in Fig. 7.10(a). We measure the time it takes to reach the absorbing state starting near the reflecting boundary. This requires escape from the first metastable region A, then the second metastable region B, to finally reach the absorbing state near the center of the last metastable region C. We measured the time using FFS and FFST for various sizes of systems (well depths) as shown in Fig. 7.10(b). We found that FFS significantly underestimates the transition time by as much as 50%. There is also a significant increase in the variance of the result, roughly an order of magnitude for this model system. This is caused by the flux being strongly influenced by the rare occurrence of a trajectory that makes it to region B, spends a long time there, and then returns to region A.

Even in the absence of long-lived metastable states, FFST can produce transition times with less variance than FFS because it samples the external return time significantly better. We found this to be the case in the ZGB model. The FFS(T) and ZGB parameters used were eleven evenly spaced static barriers starting at $C = 0.06$, $N = 10^3$, $L = 32$, and $p_{\mathrm{CO}} = 0.5268$. Under these circumstances, we found a variance of 5.0% for FFST and 5.8% for FFS; an improvement of about 16%. To equal the variance of the FFST result, FFS which would require roughly 35% more trials, which would translate into a 35% longer run time. In the case of the ZGB model, this effect can be mitigated by choosing λ_0 to be far enough away from the metastable region so that the internal return time T_{int} is much larger than the external return time T_{ext}, bounding the effect on the increase in variance of T_{ext}. In general, moving

λ_0 is not always useful because metastable states can make T_{ext} arbitrarily large.

CHAPTER VIII

Summary and Conclusions

8.1 Summary

The previous chapters showed several applications of rare event techniques.

Our first foray into rare events was the efficient calculation of the harmonic measure. Harmonic measure rare events occur when random walkers make deep excursions into the fjords of fractals. These small probability events make a significant contribution to the $D(q)$ spectrum for small and negative q. Using signposting, we were able to obtain the entire harmonic measure for the complete and external perimeters of percolation and Ising model clusters. Our results for the $D(q)$ spectra for these systems agree with theory [1].

Next, we used signposting to make new measurements of the $f(\alpha)$ spectrum, the Legendre transform of $D(q)$, for DLA, and found significant disagreement with previous iterative conformal mapping results [41]. Our results suggest that these conformal map clusters suffer from small distortions which affect measurement of small probabilities. However, we did find agreement with iterative conformal mapping results in regard to the so called "wedge angle" in DLA. We found evidence that there are many small wedges deep inside of DLA clusters. Next, we developed two more techniques, etching and the Green's function method, and used etching to obtain the

measure of the first four ($Q = 1,2,3,4$) states of the Q-states Potts model. We again found excellent agreement with previous theoretical predictions for the measure [2]. During this work, we found our smallest probability to date: 10^{-4600}.

Along these same lines, we measured other properties of critical Q-state Potts model clusters. Specifically, we measured the height and length distributions of the complete and external hulls of critical clusters for the first four Q states. We found exponential tails for both the hull height and length distributions. This, at first, appears contradictory because of the power-law relationship between the average of system width and hull length, $L \propto W^D$, and the linear relationship between the height and the width $h \propto W$. That is, we expect the hull length and height to have a power-law relationship, which would lead to one of the distributions having a stretched exponential tail. Through additional measurements, we were able to show that there is, in fact, a linear relationship between average hull height and length, *for a fixed system width*, which resolved the apparent contradictions.

Encouraged by our success with harmonic measure rare events, we expanded our work into long transition times in non-equilibrium systems. Our first algorithm developed for these problems was the barrier method, which, like the Green's function method, iteratively develops Green's functions to map a backward moving trajectory forward again. We found that avoiding backtracking leads to superior performance compared to similar rare event techniques when applied to low-dimensional systems. We used the barrier method to study the bi-stable Maier-Stein model and found good agreement between simulation results and theory for the transition time from the meta-stable state to the separatrix. We also measured the exit distribution along the separatrix and found initial disagreement with theory. However, once we convolved the theory with a Gaussian of size $\sqrt{\epsilon}$, the size of the next correction to the theory,

we found excellent agreement. In fact, the necessity of the convolution explains the apparent disagreement between previous computational and experimental results and theory.

Lastly, we exposed a limitation of the popular FFS technique and developed an improved version, called FFST, which overcame this limitation. FFST works by keeping track of the times associated with long excursions from the meta-stable state during the second part of the FFS calculation, where typically only probabilities are recorded. We used FFST to study the first-order phase-transition in the ZGB model of heterogeneous catalysis. We found the transition time from the reactive state to the poisoned state as a function of the two parameters of the model: the system length L and the probability of attempting to place a CO molecule at every step, p_{CO}. We also tracked quantities as the system transitioned from the reactive state to the poisoned state, including the average largest cluster size. Finally, we showed how FFS and FFST are related to the barrier method by introducing two additional algorithms.

8.2 Future Work

In work to be published, we have expanded our harmonic measure work to three dimensions, see Fig. 8.1. 3D harmonic measure is fascinating because practically nothing is known about it. This is because powerful theoretical techniques like conformal field theory and Schramm-Loewner evolution are limited to two dimensions. Our first step into 3D harmonic measure was percolation. We grew critical percolation clusters on a cubic lattice using the Leath algorithm. We then used etching [126] to obtain the measure for all accessible sites. Our $D(q)$ spectra agreed with previous results for the dimension of the hull, $D(0) \approx 2.5$. We also found another

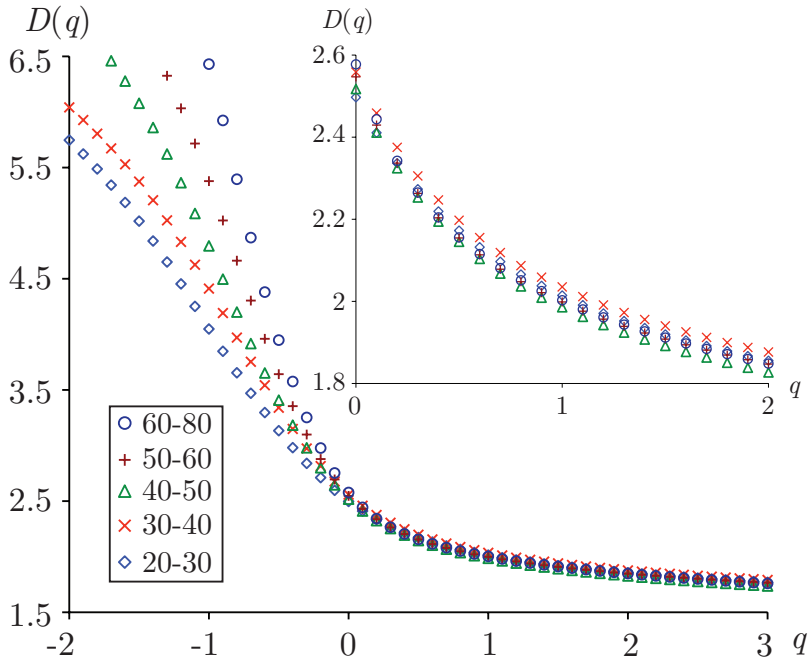

Figure 8.1: The $D(q)$ spectra for 3D percolation clusters. The spectra were computed using the ratios of average values of the partition function $Z(q, R_g)$ for pairs of radii of gyration.

interesting result: the value of $D(1) \approx 2$, which was also found for 3D Koch curves [64]. Although it has been proven that an extension of Makarov's theorem to three dimensions, $D(1) = d - 1$, does not hold exactly [127], it appears to hold closely for these two cases. We also looked at the histogram of the measure and found a surprising result: the histogram was flat. This means that for small probabilities, all probabilities are equally likely to be found. In contrast, in 2D Potts model clusters and DLA clusters the smaller the measure the more likely a site will be found with that measure. We are not aware of any previous observation of this result.

We have also expanded our DLA harmonic measure work to 3D, see Fig. 8.2. Unlike 3D percolation, we could not directly apply any of our previous techniques, since we aimed to study off-lattice clusters (making etching impossible) and bounding low probability regions with signpost planes is exceedingly difficult. Our solution was

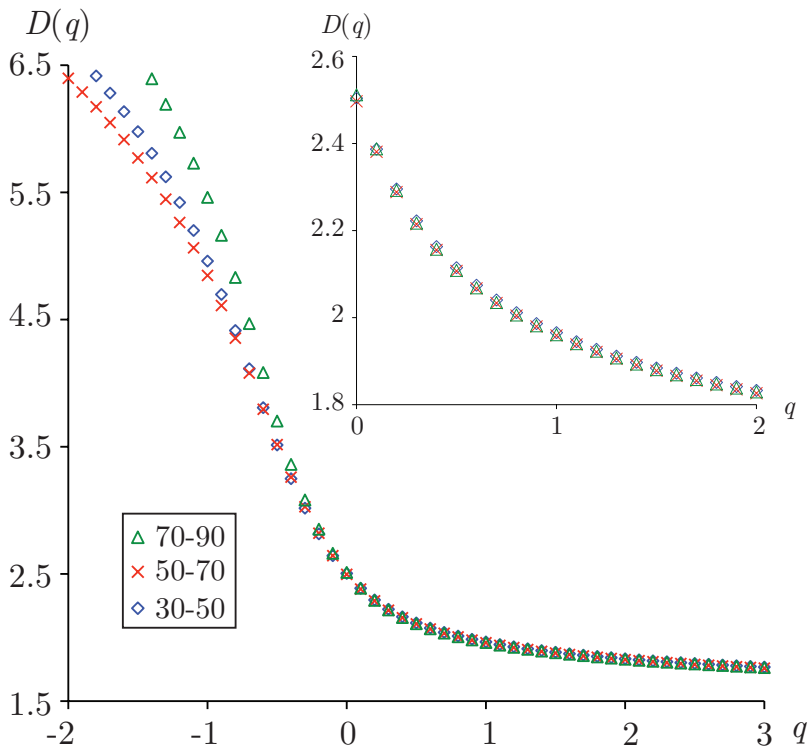

Figure 8.2: The $D(q)$ spectra for 3D DLA clusters. The spectra were computed using the ratios of average values of the partition function $Z(q, R_g)$ for pairs of radii of gyration.

to adapt signposting using spheres instead of planes. The idea is to make a surface out of the intersection of spheres bounding cluster sites that are rarely visited. This guarantees that regions that are poorly sampled will be bounded by a surface. Also, the representation of the surface by spheres makes the simulated diffusion of the random walkers very efficient because of the tree-based data structure used to store the cluster. We call this "bubble signposting," see Fig. 8.3.

Using bubble signposting, we obtained the complete harmonic measure of 3D DLA clusters. Our results for $D(q)$ look similar to that of 3D percolation, which is reasonable considering that both have nearly the same fractal dimension. Also, their $D(q)$ spectra converge for large q suggesting that the tips of percolation and DLA clusters "look similar." One difference between the two is the value of $D(1)$, which is 2% smaller for 3D DLA, i.e., ≈ 1.96; we find this value to be statistically significantly different than 2.0. It is possible that the percolation and DLA spectra are different for negative q, but the variance of our results for both percolation and DLA are too large to be definitive. The similarities continue as we look at the histogram of measure. Again, we find the histogram to be flat indicating all small values of measure are equally likely to be present. This suggests that flat histograms of the measure may be a common feature of 3D fractals.

Another studied aspect of 2D DLA is the terminating point of the $f(\alpha)$ spectrum, e.g., the largest value of α for which $f(\alpha)$ is defined. In 2D DLA, the value is roughly $\alpha_{max} = 14$ [36], and the value of $f(\alpha_{max}) > 0$. Because $f(\alpha_{max}) \neq 0$, it is analogous to a phase transition, and is connected to the so called "wedge angle." The value of $f(\alpha_{max}) \approx 0$ that we measured for 3D DLA hints at the lack of a wedge angle in 3D.

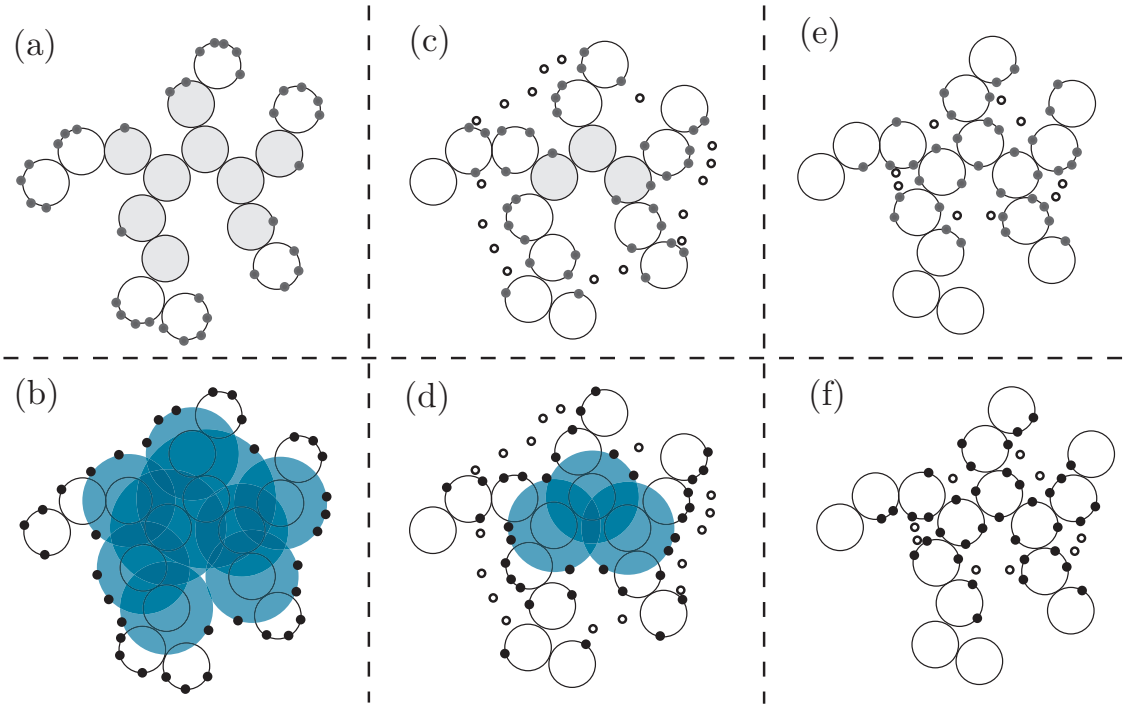

Figure 8.3: Visualization of bubble signposting. The initial probe step is shown in (a). N probe particles are released far from the cluster and allowed to diffuse until they reach the cluster. Every cluster site hit by less than 3 probe particles is under-sampled (grey). In (b), every under-sampled site is covered by a blue bubble. The size of the bubble is determined by the distance between the site and the closest well-sampled site. N more particles are released and those that hit cluster sites add their weight to the sites. The particles that hit the bubbles are temporarily frozen. The locations of the frozen particles are used to launch the reduced weight probe particles in (c). The weight of each new particle is w/N, where w is the total walker weight not yet absorbed onto the cluster sites. Those sites that are still under-sampled are marked grey. In (d), bubbles surround those sites as more particles are released and stick to the cluster sites and bubbles, adding their weight to the cluster sites. The particles that stick to the bubbles are again used to release, even lighter weight probe particles in (e). With all sites well-sampled during the probe step (e), the final particles are released and give the remaining weight to the cluster sites.

8.3 Tips for Successfully using Rare Event Techniques

Rare event techniques, when applied successfully, can make seemingly impossible calculations simple. When used incorrectly, they can be less efficient than direct sampling methods and, more seriously, can give wildly inaccurate results. Rare event techniques gain efficiency by leveraging common system properties. When the properties assumed by a rare event technique are not valid for a system, significant error typically results. Unfortunately, most descriptions of rare event techniques are given without explicit statements about the many assumptions made. Below is an attempt to review the most common assumptions, which algorithms are based on them, and when the assumptions are true.

Assumption 1: Rare events are excursions from a well defined (meta) stable state in phase space. With the exception of the harmonic measure algorithms, all rare event techniques developed in this thesis make this underlying assumption. In fact, *practically all* popular rare event techniques make this assumption. The notable exception is RESTART when applied to events that rarely occur within a fixed time period, like multi-component failure in reliable systems [89]. This assumption is commonly used because it holds for most rare events studied. One model for which this assumption does not hold is that of a population with a steadily increasing carrying capacity; the problem is that the steady-state is not stationary.

Assumption 2: Detailed balance is obeyed and dynamics are reversible.

Many rare event techniques were developed with the aim to speed up calculations of slow chemical reactions and conformational changes in macro-molecules, e.g., proteins. These systems are always near thermodynamic equilibrium, satisfy detailed balance, and are typically described by pair-wise interaction potentials be-

tween atoms and simulated via molecular dynamics or Langevin dynamics. Given a particular configuration c of the system, the total energy $E(c)$ can be calculated. For a given temperature T, the probability that the system will be in that state is given by the Boltzmann weight, $p(c) = e^{-E(c)/kT}/\mathcal{Z}$, where k is the Boltzmann constant and $\mathcal{Z}$ is the partition function. If one is only interested in the ratio of the probability of being in two different states, the partition function does not need to be calculated: $p(c_1)/p(c_2) = e^{(E(c_2)-E(c_1))/kT}$. Therefore, such systems obey detailed balance.

One consequence of detailed balance is that the relative phase-space density estimated by performing Monte Carlo sampling of a subspace of the original system will be the same as the relative phase-space density in that subspace sampled via Monte Carlo sampling of the full system. In other words, when detailed balance exists, Monte Carlo sampling of any area gives an accurate estimate of its relative density. When detailed balance does not exist, little can be definitively said about the phase-space density without sampling *all of* phase-space. Popular algorithms like TPS, TIS, milestoning, and finite temperature string method require detailed balance to be satisfied. Otherwise, full sampling of phase-space, which is typically impractical, is necessary to ensure correct results.

All path sampling methods, e.g., TPS, TIS, and finite temperature string, require that simulations can be faithfully run forwards and backwards in time. This requirement is met for molecular dynamics and Langevin dynamics but does not have to hold in general. In this thesis, all of our algorithms work for systems that lack detailed balance.

Assumption 3: The model is stochastic or chaotic. Biased sampling methods follow fruitful paths to better sample important events. This is done by mak-

ing multiple copies (splitting) of "good" states and occasionally removing (pruning) "bad" states. The multiple copies of a good state then must evolve in time independently from one another and spread out so that splitting and pruning can be applied again. The model being studied must have some non-determinism, stochasticity or chaos, for the multiple copies to fan out in phase-space. This does not generally hold for molecular dynamics simulations; however, several rare event techniques have this requirement, including splitting, RESTART, FFS, FFST, and the barrier method.

Assumption 4: The model does not have long-lived intermediate meta-stable states. One common feature of rare events is the separation in time scales between the short time it takes the desired rare event to complete, e.g., a successful transition between two meta-stable states, and the long time for the rare event to first occur. Such systems are typically described as having a smooth landscape, as compared to a rough landscape or diffusive transitions. Many popular techniques make this assumption including TPS. TIS and FFS can be used without this assumption, but not making this assumption degrades the performance of these algorithms to the point where they are marginally better than direct sampling. Naïvely using these techniques on systems with significant intermediate meta-stable states can, in some cases, cause transition times estimated to be as much as a factor of two off from the real result [128]. In addition to decreased accuracy, long-lived intermediate meta-stable states can degrade algorithm performance. Degradation occurs because sample paths or trajectories tend to spend long periods in the intermediate meta-stable states.

There are only a few algorithms that have good performance and accuracy when applied to problems with meta-stable states, namely milestoning and the barrier method. Both of these methods overcome this issue in the same way: (roughly)

uniformly sampling all sample regions of state-space between the start and end points, i.e., no heavy sampling of meta-stable states. Unfortunately, both milestoning and the barrier method have limited applicably as milestoning only work on systems that obey detailed balance and the barrier method is limited to systems with only a few dimensions.

FFST gives accurate results, but is inefficient because its run-time is limited by the time it takes a trajectory to escape the meta-stable state. Although not ideal, FFST may be the best general purpose solution to systems with long-lived meta-stable states as many algorithms, TPS, TIS, and FFS, are inefficient and give inaccurate transition time estimates.

Assumption 5: The most likely reactive pathways are easy to find.

Underlying practically all rare event techniques is the assumption, or hope, that given some reasonable initial guess for the reaction coordinate (for biased sampling methods) or the most likely path (for path sampling methods) that the algorithm will stumble upon the most likely path(s) and correctly calculate the transition time. Finding the most likely path is similar to computationally minimizing the action functional in classical mechanics, which is a global optimization problem. Global optimization can, in general, be extremely hard and in some cases NP-complete. Fortunately, many physical systems have energy landscapes which are simple and smooth to the degree where stumbling upon the most likely path is difficult *not to do*. Unfortunately, not all landscapes are that simple. The problem of mistaking another path for the most likely path has been encountered and discussed in regard to TPS [23] and FFS [106]. Figure 8.4 shows how this problem can manifest itself in a simple 2D energy landscape.

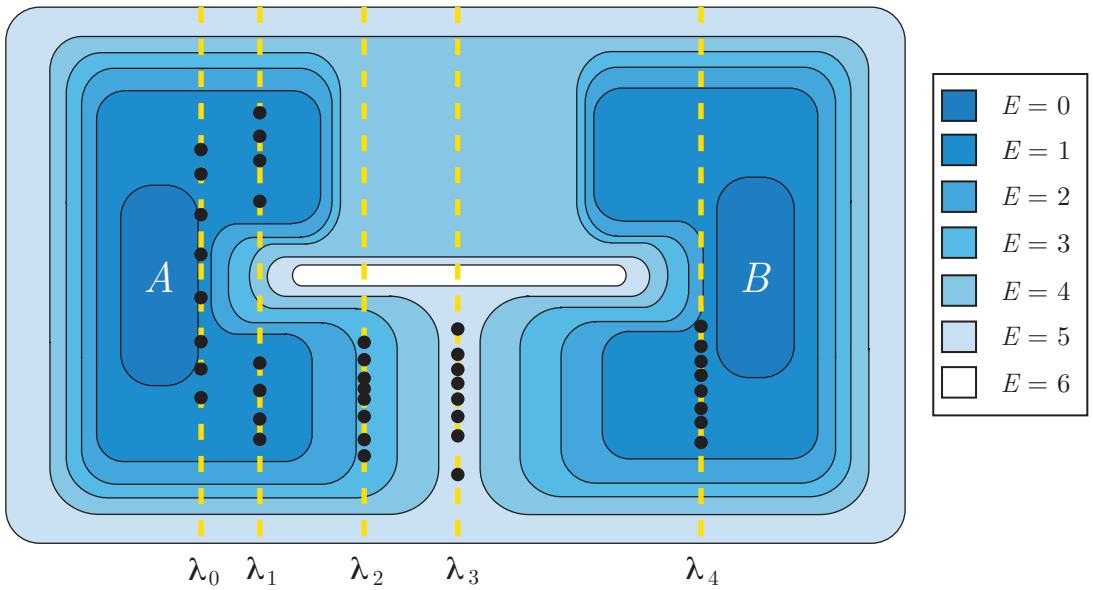

Figure 8.4: Applying FFS to a difficult landscape. FFS is used to measure the transition time between A and B with five barriers. Along the first two barriers both the top and bottom paths are well sampled. Between λ_1 and λ_2 the top path has a change in energy of 3 and the bottom path has a change in energy of 2, i.e., the top path is steeper. Because the bottom path is easier to cross, it ends up being the only path sampled. The resulting estimate of transition time would be $T \sim e^{5/(k_B T)}$. Direct sampling would have resulted in trajectories going primarily over the top path, as it has the lower energy barrier and the resulting transition time would be lower, $T \sim e^{4/(k_B T)}$. The steepness of the top path "fools" FFS into not sampling it, even though it has the smaller barrier. The inaccuracy in T in this example can be arbitrarily large as the relative error grows like $e^{1/(k_B T)}$.

BIBLIOGRAPHY

[1] B. Duplantier, Phys. Rev. Lett. **82**, 3940 (1999)

[2] B. Duplantier, Phys. Rev. Lett. **84**, 1363 (2000)

[3] C.R. Doering, K.V. Sargsyan, L.M. Sander, Multiscale Model. Simul. **3** (2005)

[4] R.S. Maier, D.L. Stein, Phys. Rev. Lett. **71**, 1783 (1993)

[5] R.S. Maier, D.L. Stein, SIAM J. Appl. Math. **57**, 752 (1997)

[6] D.G. Luchinsky, R. Maier, R. Mannella, P.V.E. McClintock, D.L. Stein, Phys. Rev. Lett. **82**, 1806 (1999)

[7] M. Khasin, M.I. Dykman, Phys. Rev. Lett. **103**, 068101 (2009)

[8] J. Medhi, *Stochastic Models in Queueing Theory* (Academic Press, Boston, 2003)

[9] M. Ptashne, *A Genetic Switch* (Cold Spring Harbor, New York, 2004)

[10] H. Andersson, T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis* (Springer, 2000)

[11] M.S. Bartlett, Wiley, New York (1961)

[12] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, 2007)

[13] J. Marro, R. Dickman, *Nonequilibrium Phase Transitions* (Cambridge University Press, 1999)

[14] B.B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, 1982)

[15] R.B. Potts, Camb. Phil. Soc. **48**, 106 (1952)

[16] J.A. Jacquez, C.P. Simon, Math. Biosci. **117**, 77 (1993)

[17] R.M. Ziff, E. Gulari, Y. Barshad, Phys. Rev. Lett. **56**, 2553 (1986)

[18] J.M. Hammersley, D.C. Handscomb, *Monte Carlo Methods* (Methuen, London, 1965)

[19] P. Glasserman, P. Heidelberger, P. Shahabuddin, T. Zajic, Oper. Res. **47**, 585 (1999)

[20] M. Villen-Altamirano, J. Villen-Altamirano, (1991), p. 71

[21] R.J. Allen, P.B. Warren, P.R. ten Wolde, Phys. Rev. Lett. **94**, 18104 (2005)

[22] A. Faradjian, R. Elber, J. Chem. Phys. **120**, 10880 (2004)

[23] P.G. Bolhuis, D. Chandler, C. Dellago, P.L. Geissler, Annu. Rev. Phys, Chem. **53**, 291 (2002)

[24] T.S. van Erp, D. Moroni, P.G. Bolhuis, J. Chem. Phys. **118**, 7762 (2003)

[25] E. Ising, Z. Phys. **21**, 613 (1925)

[26] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982)

[27] D. Stauffer, A. Aharony, *Introduction to Percolation Theory* (CRC press, 1994)

[28] S. Alexander, Phys. Lett. A **54**, 353 (1975)

[29] P.W. Kasteleyn, C.M. Fortuin, J. Phys. Soc. of Japan **26**, 11 (1969)

[30] M.E.J. Newman, G.T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford, New York, 1999)

[31] R.H. Swendsen, J.S. Wang, Phys. Rev. Lett. **58**, 86 (1987)

[32] M.E.J. Newman, R.M. Ziff, Phys. Rev. Lett. **85**, 4104 (2000)

[33] P.L. Leath, Phys. Rev. B **14**, 5046 (1976)

[34] T.A. Witten, L.M. Sander, Phys. Rev. Lett. **47**, 1400 (1981)

[35] L.M. Sander, Contemp. Phys. **41**, 203 (2000)

[36] D.A. Adams, L.M. Sander, E. Somfai, R.M. Ziff, EPL **87**, 20001 (2009)

[37] W.W. Mullins, R.F. Sekerka, J. Appl. Phys. **34** (1963)

[38] P. Meakin, A. Coniglio, H.E. Stanley, Phys. Rev. A **34**, 3325 (1986)

[39] B.B. Mandelbrot, C.J.G. Evertsz, Nature **348**, 143 (1990)

[40] D.A. Adams, L.M. Sander, R.M. Ziff, J. Chem. Phys. **133**, 124103 (2010)

[41] M.H. Jensen, A. Levermann, J. Mathiesen, I. Procaccia, Phys. Rev. E **65**, 046109 (2002)

[42] T.C. Hasley, M. Leibig, Ann. Phys. **219**, 109 (1992)

[43] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, Phys. Rev. A **33**, 1141 (1986)

[44] N.G. Makarov, Proc. London Math. Soc. **51**, 369 (1985)

[45] T. Grossman, A. Aharony, J. Phys. A **20**, 1193 (1987)

[46] M. Aizeman, B. Duplantier, A. Aharony, Phys. Rev. Lett. **83**, 1359 (1999)

[47] R.M. Ziff, P.T. Cummings, G. Stell, J. Phys. A **17**, 3009 (1984)

[48] M.B. Hastings, L.S. Levitov, Physica D **116**, 224 (1998)

[49] B. Davidovitch, H.G.E. Hentschel, Z. Olami, I. Procaccia, L.M. Sander, E. Somfai, Phys. Rev. E **59**, 1368 (1999)

[50] E. Somfai, L.M. Sander, R.C. Ball, Phys. Rev. Lett. **83**, 5523 (1999)

[51] C.M. Fortuin, P.W. Kasteleyn, Physica **57**, 536 (1972)

[52] B. Sapoval, J.S. Andrade, M. Filoche, Chem. Eng. Sci. **56**, 5011 (2001)

[53] A. Belikov, I.A. Gruzberg, I. Rushkin, J. Phys. A **41**, 285006 (2008)

[54] B. Duplantier, I.A. Binder, Nucl. Phys. B **802**, 494 (2008)

[55] E. Bettelheim, I. Rushkin, I.A. Gruzberg, P. Wiegmann, Phys. Rev. Lett. **95**, 170602 (2005)

[56] I.A. Gruzberg, J. Phys. A **39**, 12601 (2006)

[57] R.C. Ball, O.R. Spivack, J. Phys. A **23**, 5295 (1990)

[58] W.G. Hanan, D.M. Heffernan, Phys. Rev. E **77**, 011405 (2008)

[59] B. Davidovitch, A. Levermann, I. Procaccia, Phys. Rev. E **62**, R5919 (2000)

[60] B. Davidovitch, M.H. Jensen, A. Levermann, J. Mathiesen, I. Procaccia, Phys. Rev. Lett. **87**, 164101 (2001)

[61] D.A. Adams, L.M. Sander, R.M. Ziff, Phys. Rev. Lett. **101**, 144102 (2008)

[62] D. Kim, R.I. Joseph, J. Phys. C **7**, L167 (1974)

[63] D.A. Adams, L.M. Sander, R.M. Ziff, (to be published)

[64] D.S. Grebenkov, A.A. Lebedev, M. Filoche, B. Sapoval, Phys. Rev. E **71**, 056121 (2005)

[65] P. Meakin, S. Tolman, Phys. Rev. A **40**, 428 (1989)

[66] R.J. Allen, D. Frenkel, P.R. ten Wolde, J. Chem. Phys. **124**, 024102 (2006)

[67] H. Saleur, B. Duplantier, Phys. Rev. Lett. **58**, 2325 (1987)

[68] A.Y. Menshutin, L.N. Shchur, V.M. Vinokour, Physica A **387**, 6299 (2008)

[69] B.B. Mandelbrot, B. Kol, A. Aharony, Phys. Rev. Lett. **88**, 055501 (2002)

[70] M.G. Stepanov, L.S. Levitov, Phys. Rev. E **63**, 061102 (2001)

[71] E. Somfai, R.C. Ball, J.P. DeVita, L.M. Sander, Phys. Rev. E **68**, 020401 (2003)

[72] M. Bretz, Phys. Rev. Lett. **38**, 501 (1977)

[73] R.F. Voss, J. Phys. A **17**, L373 (1984)

[74] P. Grassberger, J. Phys. A **19**, 2675 (1986)

[75] B. Sapoval, M. Rosso, J.F. Gouyet, J. Physisque Lett. **46**, L149 (1985)

[76] T. Grossman, A. Aharony, J. Phys. A **19**, L745 (1986)

[77] R. Pike, H.E. Stanley, J. Phys. A **14**, L169 (1981)

[78] J. Asikaninen, A. Aharony, B.B. Mandelbrot, E. Rausch, J.P. Hovi, Euro. Phys. J. B **34**, 479 (2003)

[79] A. Coniglio, J. Phys. Lett. **46**, 250 (1981)

[80] A. Coniglio, J. Phys. A **15**, 3829 (1982)

[81] A. Bunde, J.F. Gouyet, J. Phys. A **18**, L285 (1985)

[82] S. Smirnov, W. Werner, Math. Research Lett. **8**, 729 (2001)

[83] X. Li, A.D. Sokal, Phys. Rev. Lett. **63**, 827 (1989)

[84] A. Aharony, J. Asikaninen, Fractals **11**, 3 (2003)

[85] J.L. Cardy, J. Phys. A **25**, L201 (1992)

[86] J. Cardy, J. Stat. Phys. **125**, 1 (2006)

[87] C. Dellago, P.G. Bolhuis, Advanced Computer Simulation Approaches for Soft Matter Sciences III **221**, 167 (2008)

[88] I. Schwartz, L. Billings, M. Dykman, A. Landsman, J. Stat. Mech. **2009**, P01005 (2009)

[89] M. Villen-Altamirano, J. Villen-Altamirano, in *Proceedings of the 26th Conference on Winter Simulation* (Society for Computer Simulation International, 1994), p. 282

[90] M.J.J. Garvels, D.P. Kroese, in *Proceedings of the 30th Conference on Winter Simulation* (Society for Computer Simulation International, 1998), p. 601

[91] C. Valeriani, R.J. Allen, M.J. Morelli, D. Frenkel, P.R. ten Wolde, J. Chem. Phys. **127**, 114109 (2007)

[92] E. Sanz, C. Valeriani, D. Frenkel, M. Dijkstra, Phys. Rev. Lett. **99**, 55501 (2007)

[93] R.J. Allen, C. Valeriani, S. Tănase-Nicola, P.R. ten Wolde, D. Frenkel, J. Chem. Phys. **129**, 134704 (2008)

[94] R.J. Allen, D. Frenkel, P.R. ten Wolde, J. Chem. Phys. **124**, 194111 (2006)

[95] W. E, W.Q. Ren, E. Vanden-Eijnden, J. Phys. Chem. B **109**, 6688 (2005)

[96] E. Borrero, F. Escobedo, J. Chem. Phys. **129**, 024115 (2008)

[97] R.S. Maier, D.L. Stein, J. Stat. Phys. **83**, 291 (1996)

[98] D.G. Luchinsky, R.S. Maier, R. Mannella, P.V.E. McClintock, D.L. Stein, Phys. Rev. Lett. **79**, 3109 (1997)

[99] D.G. Luchinsky, P.V.E. McClintock, Nature **389**, 463 (1997)

[100] G.E. Crooks, D. Chandler, Phys. Rev. E **64**(2), 26109 (2001)

[101] M.I. Dykman, I. Schwartz, A. Landsman, Phys. Rev. Lett. **101**, 78101 (2008)

[102] A. Bortz, M. Kalos, J. Lebowitz, J. Comp. Phys. **17**, 10 (1975)

[103] D. Gillespie, J. Comput. Phys. **22**, 403 (1976)

[104] G.A. Huber, S. Kim, Biophys. J. **70**, 97 (1996)

[105] F.A. Escobedo, E.E. Borrero, C.A. Araque, J. Phys. Condens. Matter **21**, 333101 (2009)

[106] R.J. Allen, C. Valeriani, P.R. ten Wolde, J. Phys: Condens. Matter **21**, 463102 (2009)

[107] C. Velez-Vega, E. Borrero, F. Escobedo, J. Chem. Phys. **130**, 225101 (2009)

[108] M. Ehsasi, M. Matloch, O. Frank, J.H. Block, K. Christmann, F.S. Rys, W. Hirschwald, J. Chem. Phys. **91**, 4949 (1989)

[109] E. Machado, G.M. Buendía, P.A. Rikvold, Physical Review E **71**, 31603 (2005)

[110] E. Machado, G.M. Buendía, P.A. Rikvold, R.M. Ziff, Phys. Rev. E **71**, 16120 (2005)

[111] S. Jakubith, H.H. Rotermund, W. Engel, A. von Oertzen, G. Ertl, Phys. Rev. Lett. **65**, 3013 (1990)

[112] M. Kim, M. Bertram, M. Pollmann, A.v. Oertzen, A.S. Mikhailov, H.H. Rotermund, G. Ertl, Science **292**, 1357 (1998)

[113] P. Meakin, D.J. Scalapino, J. Chem. Phys. **87**, 731 (1987)

[114] B.J. Brosilow, R.M. Ziff, Phys. Rev. A **46**, 4534 (1992)

[115] E.S. Loscar, E.V. Albano, EPL **85**, 30004 (2009)

[116] R.M. Ziff, B.J. Brosilow, Phys. Rev. A **46**, 4630 (1992)

[117] T. Tomé, R. Dickman, Phys. Rev. E **47**, 948 (1993)

[118] J.W. Evans, M.S. Miesch, Phys. Rev. Lett. **66**, 833 (1991)

[119] J.W. Evans, T.R. Ray, Phys. Rev. E **50**, 4302 (1994)

[120] R.H. Goodman, D.S. Graff, L.M. Sander, P. Leroux-Hugon, E. Clément, Phys. Rev. E **52**, 5904 (1995)

[121] E.V. Albano, Physics Letters A **288**, 73 (2001)

[122] M. Avrami, J. Chem. Phys **7**, 1103 (1939)

[123] X. Guo, D. Liu, J.W. Evans, Phys. Rev. E **75**, 061129 (2007)

[124] X. Guo, D. Liu, J.W. Evans, J. Chem. Phys **130**, 074106 (2009)

[125] W. E, W. Ren, E. Vanden-Eijnden, Journal of Applied Physics **93**, 2275 (2003)

[126] D.A. Adams, Y.T. Lin, L.M. Sander, R.M. Ziff, Phys. Rev. E **80**, 031141 (2009)

[127] T.H. Wolff, *Essays on Fourier Analysis in Honor of Elias M. Stein* (Princeton University Press, Princeton, N. J., 1995), pp. 321–384

[128] D.A. Adams, L.M. Sander, R.M. Ziff, J. Chem. Phys. **133**, 174107 (2010)