Asymptotic Analysis and Rare Event Simulation for Failure Probabilities in Discrete Random Media

by

Jeff LaComb

Department of Mathematics
Duke University

Date: _____

Approved:

_____
Jianfeng Lu, Supervisor

_____
Wilkins Aquino

_____
James Nolen

_____
Thomas Witelski

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2019

Abstract

Asymptotic Analysis and Rare Event Simulation for Failure Probabilities in Discrete Random Media

by

Jeff LaComb

Department of Mathematics
Duke University

Date: _____

Approved:

_____
Jianfeng Lu, Supervisor

_____
Wilkins Aquino

_____
James Nolen

_____
Thomas Witelski

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics in the Graduate School of Duke University
2019

Abstract

The problem of material failure is of considerable importance in several applications. We will analyze a discrete atom chain model as a means of studying a material failure problem in a random medium. For different assumptions on the atomistic interaction potential, we determine the conditions necessary for material failure, and conclude failure may only occur in the event of a large deviation in the random model parameters. This observation is then used to derive asymptotic bounds on the probability of failure. Furthermore, we use our theoretical results to motivate the development of an importance sampling algorithm to calculate rare failure probabilities with greater efficiency than standard Monte Carlo methods.

Contents

Abstract **iv**

List of Tables **vii**

List of Figures **viii**

Acknowledgements **x**

1 Introduction **1**

1.1 Material Failure Problems 3

1.1.1 Continuum Descriptions 4

1.1.2 Atom Chain Models 6

1.2 Preliminaries 11

1.2.1 ℓ^p Spaces 11

1.2.2 Properties of Normal Distributions 13

1.2.3 Optimization Algorithms 16

1.2.4 Importance Sampling 17

2 Failure Probabilities for Nearest Neighbor Harmonic Model **20**

2.1 Variational Estimates 21

2.2 Explicit Solution 31

2.3 Failure Probabilities Under a Delta Forcing 34

3 Importance Sampling Algorithm for Nearest Neighbor Harmonic Model **46**

3.1 Algorithm Overview 47
3.2 Numerical Experiments 51
3.3 Efficiency 53

4 Failure Probabilities for Next Nearest Neighbor Harmonic Model 59

4.1 Variational Estimates 60
4.2 Explicit Solution 72

5 Importance Sampling Algorithm for Next Nearest Neighbor Harmonic Model 79

5.1 Algorithm Overview 80
5.2 Numerical Experiments 83
5.3 Efficiency 86

6 Conclusion 92

A Proof of Lemma 1.2.5 94

B List of Notations 101

References 102

List of Tables

3.1 Data for Monte Carlo and Importance Sampling with Nearest Neighbor Model and a Uniform Force 52

3.2 Data for Monte Carlo and Importance Sampling with Nearest Neighbor Model and a Delta Force 53

5.1 Data for Monte Carlo and Importance Sampling with Next Nearest Neighbor Model and a Uniform Force 84

5.2 Data for Monte Carlo and Importance Sampling with Next Nearest Neighbor Model and a Delta Force 85

List of Figures

1.1 Atom Interaction Potentials 8

1.2 Atom Interaction Stencil 8

2.1 Comparison of ξ_k^1 and $\frac{1}{b}$ 30

2.2 Solutions to Nearest Neighbor Model 35

3.1 Number of Break Events for Nearest Neighbor Model with Uniform Force 50

3.2 Number of Break Events for Nearest Neighbor Model with Delta Force 50

3.3 Failure Probabilities for Nearest Neighbor Model with Uniform Force 52

3.4 Failure Probabilities for Nearest Neighbor Model with Delta Force . . 54

3.5 Upper Bound for Nearest Neighbor Failure Probability 54

3.6 Comparison of Computation Times for Nearest Neighbor Model Using Monte Carlo and Importance Sampling (Uniform Force) 55

3.7 Comparison of Computation Times for Nearest Neighbor Model Using Monte Carlo and Importance Sampling (Delta Force) 56

3.8 Importance Sampling for Nearest Neighbor Model Using Truncated Gaussian 58

4.1 Solutions to Next Nearest Neighbor Model with Delta Force 69

4.2 Solutions to Next Nearest Neighbor Model with Uniform Force 70

4.3 Comparison of ξ_k^1, ξ_k^2, and ξ_{k+1}^2 with $\frac{1}{b}$. 71

5.1 Number of Break Events for Next Nearest Neighbor Model with Uniform Force 82

5.2 Number of Break Events for Next Nearest Neighbor Model with Delta Force 82

5.3 Failure Probabilities for Next Nearest Neighbor Model with Uniform Force 84

5.4 Failure Probabilities for Next Nearest Neighbor Model with Delta Force 85

5.5 Upper Bound for Next Nearest Neighbor Failure Probability 86

5.6 Comparison of Computation Times for Next Nearest Neighbor Model Using Monte Carlo and Importance Sampling (Uniform Force) 87

5.7 Comparison of Computation Times for Next Nearest Neighbor Model Using Monte Carlo and Importance Sampling (Delta Force) 88

5.8 Comparison of Different Importance Sampling Shifts 89

5.9 Importance Sampling for Next Nearest Neighbor Model Using Truncated Gaussian 91

Acknowledgements

I would like to take this opportunity to express my appreciation to all of my teachers and professors, particularly the members of my thesis committee, for their time and guidance. A special thanks goes to my adviser, Dr. Jianfeng Lu, for his patience and advice in helping me complete my thesis. Lastly, I'd like to thank my family for their constant support.

1

Introduction

Rare event simulation is an area of particular interest in many disciplines. Several applications can be found in the fields of Materials Science, Biology, Finance and Chemistry, among others (See [2, 16, 17, 26] for a few examples). As their name implies, rare events are those which are observed very infrequently. This may be due to an inherently low probability of the event occurring, or because it occurs on a much longer time scale than what can be achieved during simulation. Despite their rarity, these events are often of considerable significance as they result in major (and sometimes catastrophic) changes to the system. In order to be able to predict these events and prevent any detrimental influences they may have, gaining a thorough understanding of rare events is vital. Unfortunately, this presents a considerable challenge for two main reasons. First, it is often computationally infeasible to simulate rare events through conventional simulation methods. Because the probability of the event is extremely low, a large number of realizations of the system would be necessary to observe even one occurrence of the event in question. Considering that each realization might be expensive to generate, this can quickly become a prohibitive

barrier to direct simulation techniques. Secondly, the problem is further exacerbated by the random nature of the system. It may be extremely time consuming to observe a single occurrence of the rare event, but observation of multiple such occurrences will be necessary to estimate the probability of the event reliably. These challenges necessitate the use of better procedures to calculate probabilities of rare events.

The primary application of these ideas studied in this work is to the failure problem in a random material. Failure is defined as a large deformation in the material's structure, causing it to change shape or break apart. Needless to say, such a failure event could have dire consequences, preventing the material from performing as intended. As such, quantifying the probability of failure based on relevant material properties is important to ensure that failure events are sufficiently rare to lie within acceptable error tolerances. This problem has been studied in the context of a continuum PDE model in [17, 18]. However, it is precisely during the event of interest, when the material breaks, that a continuum model ceases to provide an accurate description. The goal of this thesis is to generalize the results and analyses of the continuum model to an atomistic model of the material.

In this context, there are two primary objectives we aim to address. The first is to provide asymptotic upper bounds on the failure probabilities of the material, in the case of a large failure threshold. The second objective is to develop more efficient algorithms for calculating these failure probabilities, based on importance sampling techniques. These algorithms are informed by a theoretical analysis of the failure problem and designed to yield more occurrences of failure events during simulations.

In the remainder of this chapter, we shall give a brief overview of material failure problems and introduce the model we will study. Next, basic preliminaries necessary for the remainder of the thesis will be laid out. Much of the material covered in this section is well known, and is only reproduced here for the sake of completeness

and to establish notations. Chapter 2 will focus on a theoretical analysis of failure probabilities for an atom chain model with nearest neighbor interactions. In chapter 3, this analysis is utilized to develop an efficient importance sampling algorithm, and its performance is discussed. Chapters 4 and 5 are analogous to 2 and 3, but will instead focus on a more complicated model including next nearest neighbor atomistic interactions. Lastly, we conclude with a summary of our results and a discussion of future work.

1.1 Material Failure Problems

Failure problems have an immense importance in a variety of contexts, and have been extensively studied (see for example [1, 3]). A general framework has been developed in the Engineering literature for determining the reliability of a system in the presence of uncertainty, by determining a quantity known as the failure probability ([9, 13, 22]). The key object in this theory is the limit-state function $g(x, v)$, which represents the state of the system in terms of known design variables x and uncertain quantities v. It is said the system is in a state of failure if, for a given realization of v, $g(x, v) > 0$. One is thus interested in the failure probability,

$$p = P(g(x, v) > 0)$$

There are many possible criteria for material failure which may be incorporated into this framework, but we shall focus on a maximum strain criterion. This states a failure occurs when the largest strain in the material exceeds a specified critical value b, so that if $du(x, v)$ is the maximum strain in the system, then $g = du - b$. We shall discuss how this can be specialized to specific formulations for the model of a material below.

1.1.1 Continuum Descriptions

Material failure problems in random media have previously been studied in the context of an elastic continuum model in [17, 18]. In these works, the displacement field u of a material under external forcing $f(x)$ is assumed to satisfy an elliptic PDE over some domain D:

$$\begin{cases} -\nabla(a(x)\nabla u(x)) = f(x) & \text{for } x \in D \\ u(x) = 0 & \text{for } x \in \partial D \end{cases} \tag{1.1}$$

$a(x)$ encodes the material's stiffness and is assumed to be a log-normal random field,

$$a(x) = e^{-\sigma\nu(x)},$$

where $\nu(x)$ is a Gaussian random field and $\sigma > 0$. With this set-up, the maximum strain criterion becomes:

$$\sup_{x \in D} |\nabla u| > b,$$

for some specified threshold value b. When this condition is satisfied, the material is assumed to have failed. The key quantity of interest is the failure probability,

$$P\left(\sup_{x \in D} |\nabla u| > b\right)$$

There are two interesting regimes in which tail probabilities of this type may be studied. The first is the small noise limit, i.e. $\sigma \to 0$. Even when the noise level is small, the randomness can lead to significant deviations from the deterministic limit. Asymptotic estimates of tail probabilities relating to the solutions of (1.1)

were considered in the recent work [14]. The second regime of interest is when the failure threshold is large, i.e. $b \to \infty$, and we shall focus on this case in what follows.

In [18] the failure probability is examined for large b in the one dimensional case, with a delta function external force. It was shown in Theorem 2.1 of that paper the failure probability satisfies, to leading order,

$$P\left(\sup_{x\in D}|\nabla u| > b\right) \sim \left(\frac{D\sigma}{\ln(b)-\kappa}\exp\left(-\frac{(\ln(b)-\kappa)^2}{2\sigma^2}\right)\right), \tag{1.2}$$

for known constants D, κ. The argument is based on knowledge of the exact solution to the PDE (1.1) in the 1-D case. The numerical computation of the failure probability was addressed in [17]. In that paper, an efficient importance sampling algorithm was developed, where the change of measure is based on the intuition that failure at a location $x \in D$ is associated with a large deviation of the underlying field $a(x)$, specifically that $-\ln(a(x)) \sim \ln(b)$.

Based on the above works, we have a good understanding of the overall behavior of the failure problem in a continuum setting. However, the continuum model assumes we are in the regime of relatively small deformations u in the absence of fractures/defects in the material. Hence, it is precisely when the failure event of interest occurs that the underlying assumptions of the continuum model break down (mathematically speaking, solutions u satisfying the maximum strain criterion will fail to be C^1 in the limit $b \to \infty$). In order to properly capture details of the failure event, it is better to use an atomistic model for the material. Thus, one of our primary objectives is to extend the findings from the continuum models to atomistic ones.

The main advantage of the atomistic approach is the ability to model material behavior in greater detail, including variations in material properties at the micro-

scopic level. The downside, however, is that the atomistic problem generally is more complex both analytically and computationally. Perhaps the ideal approach in practice is to make use of hybrid methods based on atomistic to continuum coupling schemes such as the QNL or QCF (see [5, 15, 20] for more on such methods). In this case, we could use the fully atomistic model in the vicinity of the break location for its greater accuracy and use the continuum model away from this for its efficiency. Even when using such a method, however, a thorough understanding of the fully atomistic model is critical, and that shall be the focus of this work.

1.1.2 Atom Chain Models

A material will be modeled as single chain of N atoms, indexed by the integers $1, 2, \cdots, N$. Furthermore, for convenience in handling boundary conditions, we add four additional "ghost atoms" with indices $-1, 0, N+1$, and $N+2$. We assume that atom i can interact with its nearest neighbors in the chain (atoms $i-1$ and $i+1$) according to a potential V_1, and with its next nearest neighbors (atoms $i-2$ and $i+2$) according to V_2. Typical choices for these potentials are (appropriately scaled versions of) the harmonic potential $V(x) = x^2$ and Lennard-Jones potential $V(x) = \frac{1}{x^{12}} - \frac{2}{x^6}$. These potentials are depicted in Figure 1.1. In this thesis, we will use the harmonic potential as a simplified model, however the Lennard-Jones potential provides a more realistic model of atomistic interactions.

The strength of the interaction between atoms i and $i-1$ is given by a random coefficient ξ_i^1 drawn from a distribution Ξ^1. Similarly, the interaction between atoms i and $i-2$ has a strength given by ξ_i^2, which is obtained from the distribution Ξ^2. The stochasticity in these coefficients takes into account the uncertainty and spatial inhomogeneity in the material's strength at the atomistic scale. We will assume in subsequent analysis Ξ^j are log-normal distributions: this assumption is used in

[17, 18] and justification for using such a model for material properties is given in [24]. Under the log-normal assumption, if $\nu^j = (\nu_1^j, \nu_2^j, \ldots, \nu_n^j)$ is a $\mathcal{N}(0, \Sigma)$ random variable, then $\xi_i^j = e^{-\nu_i^j}$. Σ is a specified covariance matrix with entries σ_{ij}. To generate the ξ^j we set

$$\xi_i^j = \exp((AR^i + \mu)_j), \tag{1.3}$$

where R^i, $i = 1, 2$ are independently generated vectors of i.i.d. $\mathcal{N}(0, I)$ random variables, and $\Sigma = AA^T$ is the Cholesky decomposition of the covariance matrix. We have assumed here that ξ^1, ξ^2 are independent, but this is not necessary for the subsequent analysis. This restriction is mainly made for the ease of presentation of certain results. In practice we will typically choose Σ to be determined by the Gaussian kernel, so

$$\sigma_{ij} = \exp\left(-\gamma(i-j)^2\right), \tag{1.4}$$

for some constant γ. Other choices for the covariance are of course possible, but the entries should decay as we move further away from the main diagonal. This assumption is physically reasonable, as we expect atoms that are very far apart from each other should not have much direct influence on each other. In particular, for $R < 1$ we impose the following decay criterion:

$$|\sigma_{ij}| \leq R^{|i-j|}\sigma_{ii} \tag{1.5}$$

If no external force is applied to the material, then each atom will settle into an equilibrium position x_i. The equilibrium spacing between atoms will be defined by

$$\epsilon = \frac{1}{N} \tag{1.6}$$

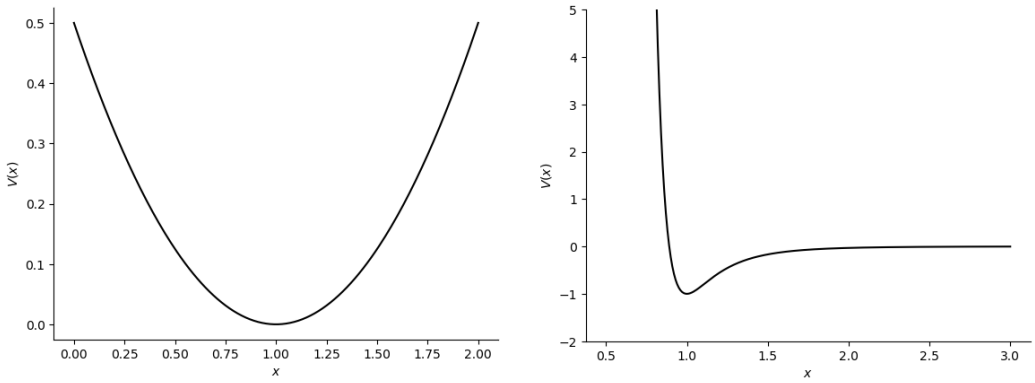

FIGURE 1.1: Examples of atomistic interaction potentials. The left plot shows the harmonic potential, while the right one displays the Lennard-Jones potential.

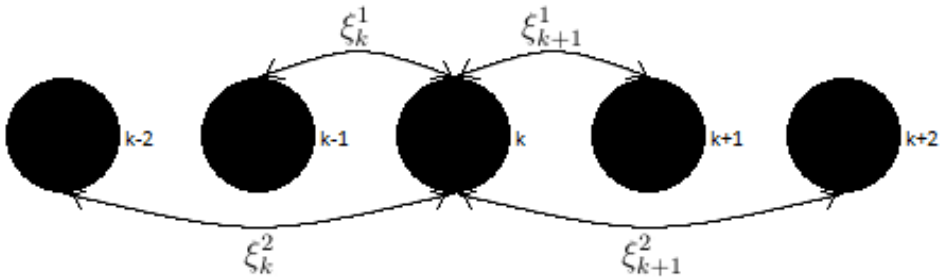

FIGURE 1.2: A stencil for nearest neighbor and next nearest neighbor interactions in an atom chain. "Ghost atoms" are added at the boundaries of the chain, so that each interior atom will exhibit the full interaction with its two closest neighbors on either side.

When performing numerical simulations to validate our results, we choose $N = 10$ and so $\epsilon = .1$. In more practical systems, of course, N shall be quite large, so we can safely assume ϵ is a small parameter. The atom chain model described above is illustrated in Figure 1.2, and has been well studied in the literature (For example, see [15, 20]).

For a given atom configuration $y \in \mathbb{R}^{N+4}$, we can write $y_i = x_i + u_i$, where u_i is the displacement of atom i from its equilibrium position. The energy associated with this configuration can then be expressed as:

$$E^{NNN}(y) = \epsilon \sum_{k=1}^{N+1} \xi_k^1 V_1 \left(\frac{y_k - y_{k-1}}{\epsilon} \right) + \epsilon \sum_{k=1}^{N+2} \xi_k^2 V_2 \left(\frac{y_k - y_{k-2}}{\epsilon} \right) \tag{1.7}$$

Equivalently, by noting that

$$y_i - y_{i-n} = (u_i - u_{i-n}) + (x_i - x_{i-n}) = (u_i - u_{i-n}) + (n\epsilon),$$

for $n = 1, 2$, we can rewrite the energy as a function of the displacement u:

$$E^{NNN}(u) = \epsilon \sum_{k=1}^{N+1} \xi_k^1 V_1 \left(\frac{\delta u_i + \epsilon}{\epsilon} \right) + \epsilon \sum_{k=1}^{N+2} \xi_k^2 V_2 \left(\frac{\delta u_i + \delta u_{i-1} + 2\epsilon}{\epsilon} \right) \tag{1.8}$$

The finite difference operator δu used above is defined component-wise by

$$\delta u_k = u_k - u_{k-1} \tag{1.9}$$

We assume Dirichlet boundary conditions on the displacement field, so $u_{-1} = u_0 = u_{N+1} = u_{N+2} = 0$. This would correspond to holding the ends of the material fixed. If we wish to restrict attention to a model with nearest neighbor interactions only, we can take $V_2 = 0$ to arrive at:

$$E^{NN}(u) = \epsilon \sum_{k=1}^{N+1} \xi_k^1 V_1 \left(\frac{\delta u_i + \epsilon}{\epsilon} \right) \tag{1.10}$$

If we apply a force $f \in \mathbb{R}^N$ to each atom in the chain, it will be perturbed to a new atom configuration $y^* = x + u^*$. It is well known this new configuration will

satisfy the energy minimization problem

$$u^* \in \text{argmin}(\mathcal{M}(u)), \tag{1.11}$$

where the objective function $\mathcal{M}$ is defined by:

$$\mathcal{M}(u) = E(u) - \epsilon \sum_{i=1}^{N} u_i f_i \tag{1.12}$$

As we shall see in what follows, it is often more convenient to calculate only δu^*. However, as the following proposition shows, this is sufficient to characterize the displacement field u^*.

Proposition 1.1.1. *If $u \in \mathbb{R}^n$, then for any $k \in \{1, \cdots, n\}$,*

$$u_k = -\sum_{i=k+1}^{N+1} \delta u_i \tag{1.13}$$

Proof. By the boundary conditions on u, we have

$$\delta u_{N+1} = u_{N+1} - u_N = -u_N \tag{1.14}$$

Using the fact and that the sum in the proposition telescopes, we can then conclude that:

$$-\sum_{i=k+1}^{N+1} \delta u_i = u_N - \sum_{i=k+1}^{N} \delta u_i = u_N - (u_N - u_k) = u_k \tag{1.15}$$

□

For a given threshold value $b > 0$, we will say a failure has occurred, or that the

material has broken, if the solution satisfies the maximum strain criterion:

$$\sup_{i\in\{1,2,\cdots,N+1\}} \left| \frac{\delta u_i^*}{\epsilon} \right| > b \tag{1.16}$$

Our objective is then to understand the failure probability:

$$p(b) = P\left(\sup_i |\delta u_i^*| > \epsilon b \right) \tag{1.17}$$

For the purposes of the analysis, it will also be convenient to define the localized failure probability,

$$p_k(b) = P(|\delta u_k^*| > \epsilon b) \tag{1.18}$$

We will focus on the regime where the failure threshold b is large, so that material failure is indeed a rare event. Our goal is then to gain a theoretical understanding of the solutions to (1.11) under the condition that the material breaks, and develop theoretical bounds as well as efficient importance sampling algorithms to estimate (1.17).

1.2 Preliminaries

1.2.1 ℓ^p Spaces

We think of atom chain configurations as elements of $\mathbb{R}^N$, so it will be helpful to establish an appropriate structure on this space. We define the ℓ^p norm for a vector u in $\mathbb{R}^n$ by:

$$||u||_{\ell^p} = \left(\sum_{i=1}^{n} u_i^p \right)^{\frac{1}{p}} \tag{1.19}$$

For $p = \infty$, we have:

$$||u||_\infty = \sup_{i \in \{1,2,\cdots,n\}} |u_i| \tag{1.20}$$

When working with discrete atom chain models, the weighted norm

$$||u||_{\ell^p_\epsilon} = \left(\epsilon \sum_{i=1}^{n} u_i^p \right)^{\frac{1}{p}} \tag{1.21}$$

is often more convenient to work with. We will typically be most interested in the case $p = 2$, as then we may also define an inner product:

$$\langle u, v \rangle_\epsilon = \epsilon \sum_i u_i v_i \tag{1.22}$$

A useful result in the subsequent analysis is the following discrete analogue of the Poincare inequality.

Proposition 1.2.1.

$$\sum_{k=1}^{n} |u_k|^2 \leq n^2 \sum_{k=1}^{n+1} |\delta u_k| \tag{1.23}$$

Proof. Using the boundary condition $u_0 = 0$ followed by the Cauchy-Schwarz inequality,

$$|u_k| = \left| \sum_{i=1}^{k} \delta u_i \right| \leq \sum_{i=1}^{k} |\delta u_i| \leq \sum_{i=1}^{n} |\delta u_i| \cdot 1 \leq \left(\sum_{i=1}^{n} |\delta u_i|^2 \right)^{\frac{1}{2}} \sqrt{n}$$

Now, making use of the above inequality, we obtain:

$$\sum_{k=1}^{n} |u_k|^2 \leq \sum_{k=1}^{n} \sum_{i=1}^{n} |\delta u_i|^2 n = n^2 \sum_{i=1}^{n} |\delta u_i|^2 \leq n^2 \sum_{i=1}^{n+1} |\delta u_i|^2$$

□

1.2.2 Properties of Normal Distributions

We will denote a normal (or Gaussian) distribution with mean μ and covariance matrix Σ by $\mathcal{N}(\mu, \Sigma)$. The following are standard but useful results about normal distributions which we shall utilize throughout our analysis.

Proposition 1.2.2. *Suppose X is a $\mathcal{N}(\mu, \Sigma)$ random variable on $\mathbb{R}^n$. Decompose X as $X = (x_1, x_2)$, and correspondingly decompose the mean and covariance as $\mu = (\mu_1, \mu_2)$ and $\Sigma = \left[\begin{smallmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{smallmatrix}\right]$. Then for $i = 1, 2$, the marginal distribution of x_i is:*

$$\mathcal{N}(\mu_i, \Sigma_{ii}) \tag{1.24}$$

Furthermore, the conditional distribution of x_i given x_j is

$$\mathcal{N}(\mu_i + \Sigma_{ij}\Sigma_{jj}^{-1}(x_j - \mu_j), \Sigma_{ii} - \Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}) \tag{1.25}$$

Proposition 1.2.3. *Suppose X is a $\mathcal{N}(\mu, \sigma^2)$ random variable. Then for any $x > \mu$:*

$$\frac{1}{\sqrt{2\pi}} \left(\frac{\sigma}{x - \mu} - \frac{\sigma^3}{(x - \mu)^3} \right) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \leq P(X \geq x) \leq \frac{\sigma e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}(x - \mu)} \tag{1.26}$$

We will also have use for the truncated normal distribution, i.e. a normal random variable which is bounded above and/or below. Information on such distributions

can be found, for example, in [12]. For our purposes, we shall only consider the case in which the components of the random variable are bounded from below. If $X \sim \mathcal{N}(\mu, \Sigma)$ satisfies $X_i \geq c_i$ for $i = 1, \ldots, n$, then the the density function for X is given by:

$$f(x) = \begin{cases} \frac{\phi(x,\mu,\Sigma)}{1-\Phi(c,\mu,\Sigma)} & x > c \\ 0 & \text{otherwise} \end{cases} \tag{1.27}$$

Here $\phi(x, \mu, \Sigma)$ and $\Phi(x, \mu, \Sigma)$ denote the pdf and cdf, respectively, of a normal random variable with mean μ and covariance Σ.

Before moving on, we prove a simple result concerning truncated normal distributions which will turn out to be quite useful:

Proposition 1.2.4. *Assume ν is a $\mathcal{N}(\mu, \Sigma)$ random variable on $\mathbb{R}^n$, and $c \in \mathbb{R}^n$ have all positive entries. Let $T > n \sup_j c_j$, Y_j be the event $\{\nu_i > c_i,\ i \neq j\}$, and set $G(\mu, \Sigma, c) = \frac{\sup_j P\left(Y_j \mid \nu_j > \frac{T}{n}\right)}{\int_{y>c} \phi(y, \mu, \Sigma) dy}$. Then:*

$$P(||\nu||_{\ell^2} \geq T \mid \nu > c) \leq G \sum_j P\left(\nu_j > \frac{T}{n}\right) \tag{1.28}$$

Proof. In order for $||\nu||_{\ell^2}$ to exceed T, we must have $\nu_j \geq \frac{T}{n}$ for at least one index j. Then by sub-additivity we have:

$$\begin{aligned} P(||\nu||_{\ell^2} \geq T \mid \nu > c) &\leq P\left(\nu_1 \geq \frac{T}{n} \cup \nu_2 \geq \frac{T}{n} \cup \cdots \cup \nu_n \geq \frac{T}{n} \mid \nu > c\right) \\ &\leq \sum_j P\left(\nu_j \geq \frac{T}{n} \mid \nu > c\right) \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_j \int_{Y_j} \int_{\frac{T}{n}}^{\infty} \frac{\phi(x,\mu,\Sigma)}{\int_{y>c}\phi(y,\mu,\Sigma)dy} dx_j d\hat{x}_j \\
&= \frac{1}{\int_{y>c}\phi(y,\mu,\Sigma)dy} \sum_j \int_{Y_j} \int_{\frac{T}{n}}^{\infty} \phi(x,\mu,\Sigma) dx_j d\hat{x}_j \\
&= \frac{1}{\int_{y>c}\phi(y,\mu,\Sigma)dy} \sum_j P\left(\nu_j > \frac{T}{n}, Y_j\right) \\
&= \frac{1}{\int_{y>c}\phi(y,\mu,\Sigma)dy} \sum_j P\left(\nu_j > \frac{T}{N}\right) P\left(Y_j \mid \nu_j > \frac{T}{N}\right) \\
&\leq G \sum_j P\left(\nu_j > \frac{T}{N}\right)
\end{aligned}
$$

□

Lastly, we state a useful lemma which will be needed for a few results later.

Lemma 1.2.5. *Let $\nu \sim \mathcal{N}(0,\Sigma)$ be a Gaussian random variable on $\mathbb{R}^n$, and suppose $\sigma^M = \sup_i \sigma_{ii}$. Let $j \in \mathbb{N}_{\leq n}$ and $\mathcal{I} \subset \mathbb{N}_{\leq n}$ with $|\mathcal{I}| = d \geq 1$ and $j > \max(\mathcal{I})$. Choose $1 < \alpha < \sqrt{2}$ and pick R small enough that $g(R,d) > 0$ and $\beta^* Q_{\mathcal{I}j}(R,d) > \alpha d$, where*

$$\beta^* = \frac{1 - \sqrt{2d^{-2}Q_{\mathcal{I}j}^2 - 1}}{1 - d^{-2}Q_{\mathcal{I}j}^2}$$

Assume Σ satisfies (1.5) with constant R. Then for any positive constant $c_1 \in \mathbb{R}$, $c_2 \in \mathbb{R}^d$ with all positive entries and $c_{2,i} \leq c_1$, there exists a constant $C > 0$ such that:

$$
\begin{aligned}
&P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) \\
&\leq C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3}\right)^{-1} \frac{\sqrt{\sigma^M}}{c_1} \exp\left(\frac{(1-\alpha^2)c_1^2}{2\sigma^M}\right)
\end{aligned}
\tag{1.29}
$$

The proof is somewhat involved, so it as well as the precise definitions of $Q_{\mathcal{I}j}(R,d)$ and $g(r,d)$ are deferred to the appendix. Here, we merely note this lemma allows us to quantify how unlikely it is to see multiple large entries in ν based on how strong the correlation between its components is. In particular, if we take $\alpha \to \sqrt{2}$ we can recover nearly the same exponential decay as we'd expect if the components of ν were independent, but this is only possible by taking $R \to 0$, placing stronger restrictions on the covariance matrix.

1.2.3 Optimization Algorithms

As we have seen, to determine atom configurations of a chain under a given external force we must solve optimization problems of the form $\min_u \mathcal{M}(u)$. For numerical simulations when an explicit solution to the optimization problem is not available, this will be done using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. This is a quasi-Newton algorithm which can attain superlinear convergence, and has the advantage that the Hessian of the objective function does not need to be calculated exactly. Details of the algorithm can be found, for example, in [21]. For the sake of completeness, the steps are listed here:

Algorithm 1: BFGS

1. Choose a starting point u_0, error tolerance $\epsilon > 0$, and Hessian approximation $H_0 = cI$, for some constant c.

2. Choose a search direction $p_k = -H_k \nabla \mathcal{M}_k$

3. Determine a step size α_k satisfying the Wolfe conditions,

$$\mathcal{M}(u_{k+1}) \leq \mathcal{M}(u_k) + c_1 \alpha_k \nabla \mathcal{M}_k^T p_k$$

$$\nabla \mathcal{M}(u_{k+1})^T p_k \geq c_2 \nabla \mathcal{M}_k^T p_k$$

   Typical values of c_1, c_2 are $c_1 = .0001$, $c_2 = .9$.

4. Update u_k by $u_{k+1} = u_k + \alpha_k p_k$.

5. Set

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$$

   where $y_k = \nabla \mathcal{M}_{k+1} - \nabla \mathcal{M}_k$, $s_k = u_{k+1} - u_k$, and $\rho_k = \frac{1}{y_k^T s_k}$.

6. If $||\nabla \mathcal{M}_{k+1}|| < \epsilon$, terminate the algorithm. Otherwise, set k to $k+1$ and repeat steps 2 through 6.

1.2.4 Importance Sampling

Since failure events are extremely rare in the large b regime, our problem can be interpreted as one of estimating a probability $P(A) \ll 1$ with a high degree of accuracy. The most direct approach to doing this would be through Monte Carlo simulation. In our context, this entails simulating the system T times, and observing the number of times event A occurs, T_A. If we let χ_k denote the characteristic function of A occurring on trial k, then we can then estimate $P(A) \approx p$ as

$$p = \frac{1}{T} \sum_{k=1}^{n} \chi_k = \frac{T_A}{T} \tag{1.30}$$

The variance of this estimator can be calculated as follows:

$$\sigma^2 = \frac{1}{T^2}\sum_{k=1}^{T}(\chi_k - p)^2 = \frac{1}{T^2}(T_A(1-p)^2 + (T-T_A)p^2)$$
$$= \frac{1}{T^2}\left(T_A - \frac{T_A^2}{T}\right) = \frac{1}{T}(p - p^2)$$

In particular, we notice that the variance of the estimator is small simply by virtue of p being small. Since we want our computation to be accurate relative to the magnitude of $P(A)$, a better measure of the quality of our estimate is the relative error,

$$\text{Rel. Err.} = \frac{\sigma^2}{p^2} = \frac{1-p}{Tp} \tag{1.31}$$

From this, we can see that when $p \ll 1$, we must have T at least on the order of $\frac{1}{p}$ to achieve a small relative error. This means direct Monte Carlo simulation will require a huge number of trials, making the method inefficient. If the event A is particularly rare, we may not observe even a single occurrence of A during simulation, making a meaningful estimate of $P(A)$ impossible. As such, it is extremely desirable to have methods for making rare events easier to simulate and reducing variance/relative error when performing Monte Carlo estimates.

The problems with using direct Monte Carlo methods to estimate small probabilities are well known, and many different approaches have been proposed to circumvent these issues ([9],[23],[25]). In general, there is no fixed rule about what approach is best to take, and insight into the specific problem under consideration is often crucial to designing more efficient simulation methods. However, one particularly useful and common method for variance reduction is importance sampling. Instead of trying to draw samples directly from P, we will introduce a new measure Q (with P absolutely

continuous with respect to Q), such that $Q(A) > P(A)$, and hence we can estimate $Q(A)$ using fewer trials. The probability under the original distribution can then be calculated by

$$P(A) = \int_\Omega \chi(A) dP = \int_\Omega \chi(A) \frac{dP}{dQ} dQ, \tag{1.32}$$

where $\frac{dP}{dQ}$ is the Radon-Nikodym derivative. If the proposal measure Q is chosen wisely, the importance sampling estimator will exhibit a reduced variance when compared to direct Monte Carlo methods.

Clearly, the key challenge in developing an effective importance sampling algorithm is in choosing the distribution Q. As is well known, the ideal choice would be

$$Q(B) = \frac{P(B \cap A)}{P(A)}$$

With this choice it is evident $Q(A) = 1$, so that A occurs essentially every trial under Q and the importance sampling estimator will have zero variance. Of course, this is worthless in practice since $P(A)$ is precisely the quantity we are trying to compute: if we had enough information to calculate Q, there would be no need for the importance sampling in the first place. Instead, we will require other methods to choose Q in order to increase $Q(A)$. For our problem, we shall see it is possible to make use of additional knowledge about the behavior of the system to inform our construction of Q.

2

Failure Probabilities for Nearest Neighbor Harmonic Model

In this chapter, we shall focus on a nearest neighbor model for atomistic interactions with the harmonic potential,

$$V(u) = \frac{1}{2}(u-1)^2 \tag{2.1}$$

With these choices, the model (1.10) simplifies to:

$$E_H^{NN}(u) = \frac{\epsilon}{2}\sum_{k=1}^{N+1} \xi_k^1 \left(\frac{\delta u_i}{\epsilon}\right)^2 \tag{2.2}$$

While perhaps the least realistic model we shall consider, this is also the simplest and hence most amenable to mathematical analysis. In section 2.1, we shall argue directly from the minimization of the energy (2.2) to show that failure events are associated with large deviations in the coefficient field ξ^1. In addition, this observation will

be used to obtain preliminary estimates on the failure probability. In section 2.2, we derive Euler-Lagrange equations which allow us to solve (1.11) exactly, and we utilize this result to derive sharper estimates on $p(b)$ for a class of localized forces in section 2.3.

2.1 Variational Estimates

To understand failure probabilities for this model, we shall first show an intuitive result which states the material can only break if there is some location where the bond between atoms is weak.

Lemma 2.1.1. *Suppose u is a solution to (1.11) with $E(u) = E_H^{NN}$ such that $\sup_i |\delta u_i| \geq \epsilon b$. Then*

$$\min_k \xi_k^1 \leq \frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{\epsilon}} \tag{2.3}$$

Proof. We first observe that since $\mathcal{M}(0) = 0$, the minimizer u of (1.11) must satisfy $\mathcal{M}(u) \leq 0$. From (1.12), this implies

$$\frac{\epsilon}{2} \sum_{k=1}^{N+1} \xi_k^1 \left(\frac{\delta u_i}{\epsilon} \right)^2 \leq \epsilon \sum_{i=1}^{N} f_i u_i$$

Now by applying the Cauchy-Schwarz inequality to the right hand side followed by (1.23), we obtain:

$$\frac{\epsilon}{2} \sum_{k=1}^{N+1} \xi_k^1 \left(\frac{\delta u_i}{\epsilon} \right)^2 \leq N||f||_{\ell_\epsilon^2}||\delta u||_{\ell_\epsilon^2} \tag{2.4}$$

Rearranging terms and utilizing (1.6) implies

$$\sum_{k=1}^{N+1} \xi_k^1 \left(\delta u_i\right)^2 \leq 2||f||_{\ell_\epsilon^2}||\delta u||_{\ell_\epsilon^2}$$

The left hand side can be bounded from below by $\frac{1}{\epsilon}\min_k \xi_k^1 ||\delta u||_{\ell_\epsilon^2}^2$, allowing this to be simplified to:

$$\min_k \xi_k^1 \leq \frac{2\epsilon||f||_{\ell_\epsilon^2}}{||\delta u||_{\ell_\epsilon^2}}$$

The failure condition (1.16) implies $||\delta u||_{\ell_\epsilon^2} \geq \epsilon^{\frac{3}{2}} b$. Plugging this into the previous inequality yields the result. □

In fact, we can build upon Lemma 2.1.1 to reach the stronger conclusion the material must break at the same location where the coefficient field ξ^1 is weakest. Before doing so, we shall first need a few technical lemmas to show we can restrict our attention to the case when there is only a single break in the atom chain which does not have $||\delta u||_\infty$ exceed ϵb by too much. To justify neglecting other events, we demonstrate they have a negligible probability when compared to a single break event.

Lemma 2.1.2. *Suppose that for some j, k, the solution u has $|\delta u_k| > \epsilon b$, and $|\delta u_j| > r\epsilon b$, for a given $r < 1$ and $j \neq k$. Further assume $\sup_i |\delta u_i| < R\epsilon b$, for some $R > 1$. Then for at least two distinct indices a, b and $s \in \{a, b\}$:*

$$\xi_s^1 \leq \left(1 + \frac{R^2}{r^2}\right) \frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{1+r^2}\sqrt{\epsilon}} \tag{2.5}$$

Proof. Suppose the smallest ξ^1 value is ξ_a^1, the second smallest is ξ_b^1. The left hand

side of (2.4) can be rewritten as:

$$\frac{\epsilon}{2}\left(\sum_{i\neq a}\xi_i^1\frac{\delta u_i^2}{\epsilon^2}+\xi_a^1\frac{\delta u_a^2}{\epsilon^2}\right)\geq\frac{\epsilon}{2}\sum_{i=1}^{N+1}\xi_b^1\frac{\delta u_i^2}{\epsilon^2}+(\xi_a^1-\xi_b^1)\frac{\delta u_a^2}{2\epsilon}$$

Then manipulating the inequality (2.4) gives:

$$\frac{\epsilon}{2}\sum_{i=1}^{N+1}\xi_b^1\frac{\delta u_i^2}{\epsilon^2}\leq N||f||_{\ell_\epsilon^2}||\delta u||_{\ell_\epsilon^2}+\frac{1}{2\epsilon}(\xi_b^1-\xi_a^1)\delta u_a^2$$

Upon further simplification, this becomes:

$$\frac{1}{2\epsilon^2}\xi_b^1||\delta u||_{\ell_\epsilon^2}^2\leq N||f||_{\ell_\epsilon^2}||\delta u||_{\ell_\epsilon^2}+\frac{1}{2\epsilon}(\xi_b^1-\xi_a^1)\delta u_a^2$$

Dividing by $||\delta u||_{\ell_\epsilon^2}$ and using the lower bounds on $\delta u_j, \delta u_k$ yields:

$$\xi_b^1\leq\frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{1+r^2}\sqrt{\epsilon}}+\frac{\delta u_a^2\xi_b^1}{\sum_{i=1}^{N+1}\delta u_i^2}$$

Solving for ξ_b^1 results in:

$$\begin{aligned}\xi_b^1&\leq\left(1-\frac{\delta u_a^2}{\sum_{i=1}^{N+1}\delta u_i^2}\right)^{-1}\frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{1+r^2}\sqrt{\epsilon}}\\&\leq\left(1-\frac{R^2\epsilon^2b^2}{R^2\epsilon^2b^2+r^2\epsilon^2b^2}\right)^{-1}\frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{1+r^2}\sqrt{\epsilon}}\end{aligned}$$

From here, the proof is completed by simple algebraic calculations, since $\xi_a^1<\xi_b^1$ by assumption. □

Now for a specified r, R, we can describe two events we will want to exclude from future calculations:

$$E_1 = \left\{ \sup_i |\delta u_i| > R\epsilon b \right\}$$

$$E_2 = \left\{ \sup_i |\delta u_i| > \epsilon b \text{ and } \sup_{i \neq k} |\delta u_i| > r\epsilon b \right\}$$

The first set accounts for deviation in δu which is significantly larger than what we expect is required for the material to break, while the second encompasses events in which the material fails at more than one location.

Next, define a "rare" failure set $\mathcal{R}$ which includes these events as well as a "normal" failure set $\mathcal{F}$, which focuses attention on the material breaking in a single location with $\sup_i |\delta u_i|$ only just exceeding b:

$$\begin{aligned} \mathcal{R} &= E_1 \cup E_2 \\ \mathcal{F} &= \left\{ \sup_i |\delta u_i| > \epsilon b \right\} \cap \mathcal{R}^C \end{aligned} \tag{2.6}$$

The next lemma justifies the view of $\mathcal{R}$ as the "rare" set by showing it has negligible probability, even when compared other rare events under consideration.

Lemma 2.1.3. *Choose $1 < \alpha < \sqrt{2}$ and assume Σ satisfies (1.5) with constant $\tilde{R}$ small enough that the conditions of Lemma 1.2.5 are satisfied. Then there exist*

positive constants $\bar{C}$, $k_1 > \frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}$, $k_2 = \frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}$, *and* $k_3 > 0$ *with*

$$P(\mathcal{R}) \leq \bar{C}\Bigg(\frac{1}{\sqrt{2\pi}\ln(k_1 b)} \exp\left(\frac{-(\ln(k_1 b))^2}{2\sigma^M}\right) + \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(k_3 b)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(k_3 b)^3} \right)^{-1} \frac{1}{\ln(k_2 b)^2} \exp\left(\frac{-(\ln(k_2 b))^2}{2\sigma_{ii}}\right) \exp\left(\frac{(1-\alpha^2)\ln(k_2 b)^2}{2\sigma^M}\right) \Bigg) \tag{2.7}$$

Proof. Our first observation is that by additivity of the probability measure,

$$P(\mathcal{R}) = P(E_1) + P(E_2 \cap E_1^C) \tag{2.8}$$

We estimate these two probabilities separately. First, consider $P(E_1)$. By using the result of Lemma 2.1.1, we obtain:

$$P(E_1) \leq P\left(\min_i \xi_i^1 < \frac{2||f||_{\ell^2_\epsilon}}{Rb\sqrt{\epsilon}} \right)$$

This can be rewritten in terms of normal random variables:

$$\begin{aligned} P(E_1) &\leq P\left(\max_i \nu_i^1 \geq \ln\left(\frac{Rb\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}} \right) \right) \\ &= P\left(\bigcup_i \left(\nu_i^1 \geq \ln\left(\frac{Rb\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}} \right) \right) \right) \\ &\leq \sum_{i=1}^{N+1} P\left(\nu_i^1 \geq \ln\left(\frac{Rb\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}} \right) \right) \end{aligned}$$

For convenience, let $k_1 = \frac{R\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}$. To estimate the terms in the sum, we make use of (1.26):

$$\begin{aligned} P(E_1) &\leq \sum_{i=1}^{N+1} \frac{\sqrt{\sigma_{ii}}}{\sqrt{2\pi}\ln(k_1 b)} \exp\left(\frac{-(\ln(k_1 b))^2}{2\sigma_{ii}}\right) \\ &\leq (N+1) \sup_i \frac{\sqrt{\sigma_{ii}}}{\sqrt{2\pi}\ln(k_1 b)} \exp\left(\frac{-(\ln(k_1 b))^2}{2\sigma_{ii}}\right) \end{aligned} \tag{2.9}$$

Now we move on to $P(E_2 \cap E_1^C)$. Observe that on this set, the conditions of Lemma 2.1.2 are satisfied, and consequently we have $\xi_s^1 \leq (1 + \frac{R^2}{r^2})\frac{2||f||_{\ell^2_\epsilon}}{b\sqrt{(1+r^2)\epsilon}}$ for the second smallest ξ^1 coefficient. Let A_i^1 be the event $\left\{\xi_i^1 \leq \frac{2||f||_{\ell^2_\epsilon}}{b\sqrt{\epsilon}}\right\}$ and A_i^2 be the event $\left\{\xi_i^1 \leq (1 + \frac{R^2}{r^2})\frac{2||f||_{\ell^2_\epsilon}}{b\sqrt{(1+r^2)\epsilon}}\right\}$. Then a simple calculation yields:

$$P(E_2 \cap E_1^C) \leq \sum_{i=1}^{N+1} \sum_{j>i} P\left(A_i^1 \cap A_j^2\right) = \sum_{i=1}^{N+1} \sum_{j>i} P\left(A_i^1\right) P(A_j^2 \mid A_i^1) \tag{2.10}$$

Using a similar procedure to that used when estimating $P(E_1)$, we can find

$$P\left(\xi_i \leq \frac{2||f||_{\ell^2_\epsilon}}{\sqrt{\epsilon}b}\right) \leq \frac{\sqrt{\sigma_{ii}}}{\sqrt{2\pi}\ln(k_2 b)} \exp\left(\frac{-(\ln(k_2 b))^2}{2\sigma_{ii}}\right) \tag{2.11}$$

For $P(A_j^2 \mid A_i^1)$, we use Lemma (1.2.5) to obtain the estimate:

$$P(A_j^2 \mid A_i^1) \leq C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(k_3 b)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(k_3 b)^3}\right)^{-1} \frac{\sqrt{\sigma^M}}{\ln(k_2 b)} \exp\left(\frac{(1-\alpha^2)\ln(k_2 b)^2}{2\sigma^M}\right) \tag{2.12}$$

Continuing from (2.10), if we substitute in (2.11) and (2.12) we get the inequality:

$$P(E_2 \cap E_1^C) \leq \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(k_3 b)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(k_3 b)^3} \right)^{-1} \frac{C\sigma^M}{\sqrt{2\pi}\ln(k_2 b)^2} \exp\left(\frac{-(\ln(k_2 b))^2}{2\sigma_{ii}} \right) \exp\left(\frac{(1-\alpha^2)\ln(k_2 b)^2}{2\sigma^M} \right) \tag{2.13}$$

Returning to (2.8), the estimates (2.9) and (2.13) imply:

$$P(\mathcal{R}) \leq \left((N+1)\frac{\sqrt{\sigma^M}}{\sqrt{2\pi}\ln(k_1 b)} \exp\left(\frac{-(\ln(k_1 b))^2}{2\sigma^M} \right) + N(N+1)\inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(k_3 b)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(k_3 b)^3} \right)^{-1} \frac{C\sigma^M}{\sqrt{2\pi}\ln(k_2 b)^2} \exp\left(\frac{-(\ln(k_2 b))^2}{2\sigma_{ii}} \right) \exp\left(\frac{(1-\alpha^2)\ln(k_2 b)^2}{2\sigma^M} \right) \right)$$

The lemma follows immediately from this, by setting

$$\bar{C} = \frac{N(N+1)\max(C\sigma^M, \sqrt{\sigma^M})}{\sqrt{2\pi}}$$

. □

Based on these results, we now understand that "typical" failure events are those coming from the set $\mathcal{F}$. As such, it is sufficient restrict our attention to $\mathcal{F}$ when strengthening the result of Lemma 2.1.1.

Theorem 2.1.4. *Suppose u is a solution to (1.11) with $E(u) = E_H^{NN}$ such that $u \in \mathcal{F}$. Further assume that the break occurs at location k, $\sup_i |\delta u_i| = |\delta u_k|$. Then,*

provided the failure threshold b is sufficiently large,

$$\xi_k^1 \leq \frac{2||f||_{\ell_\epsilon^2}}{b\sqrt{\epsilon}} \tag{2.14}$$

Proof. Suppose the minimizing atom configuration has components $u = (u_1, \ldots, u_N)$. For any $j \neq k$, define a new atom configuration $\hat{u}$ by :

$$\hat{u}_i = \begin{cases} u_i & \text{if } i < \min(j,k) \text{ or } i \geq \max(j,k) \\ u_i + \delta u_j - \delta u_k & \text{if } j > k \text{ and } k \leq i < j \\ u_i + \delta u_k - \delta u_j & \text{if } k > j \text{ and } j \leq i < k \end{cases}$$

It is not difficult to see that the effect of this rearrangement is that $\delta\hat{u}_k = \delta u_j$, and $\delta\hat{u}_j = \delta u_k$, while otherwise $\delta\hat{u}_i = \delta u_i$ if $i \neq j, k$. Since u is the minimizer of $\mathcal{M}$, we must have $\mathcal{M}(u) - \mathcal{M}(\hat{u}) \leq 0$. Consequently, we have the inequality:

$$\begin{aligned} &\frac{\epsilon}{2}\left(\xi_k^1\frac{\delta u_k^2}{\epsilon^2} + \xi_j^1\frac{\delta u_j^2}{\epsilon^2} - \xi_k^1\frac{\delta u_j^2}{\epsilon^2} - \xi_j^1\frac{\delta u_k^2}{\epsilon^2}\right) \leq \epsilon\sum_{i=1}^N f_i(u_i - \hat{u}_i) \\ \implies &\frac{\epsilon}{2}\left(\frac{\delta u_k^2}{\epsilon^2} - \frac{\delta u_j^2}{\epsilon^2}\right)(\xi_k^1 - \xi_j^1) \leq \epsilon\sum_{i=1}^N f_i(u_i - \hat{u}_i) \\ &\leq \epsilon\begin{cases} (\delta u_k - \delta u_j)\sum_{i=k}^{j-1}|f_i| & \text{if } k < j \\ (\delta u_j - \delta u_k)\sum_{i=j}^{k-1}|f_i| & \text{if } j < k \end{cases} \end{aligned} \tag{2.15}$$

Our objective is to show that $\xi_k^1 - \xi_j^1 \leq 0$, provided b is large enough. We consider two possible cases: $j > k$ and $j < k$. If $j > k$, but $\delta u_k - \delta u_j \leq 0$, the right hand side of (2.15) is negative. Because of the assumption $\sup_i |\delta u_i| = |\delta u_k|$, $\delta u_k^2 \geq \delta u_j^2$, and

hence the inequality can only hold if $\xi_k^1 \leq \xi_j^1$. On the other hand, if $\delta u_k - \delta u_j > 0$ we can rearrange (2.15) to obtain:

$$\frac{\epsilon}{2}\left(\frac{\delta u_k}{\epsilon^2} + \frac{\delta u_j}{\epsilon^2}\right)(\xi_k^1 - \xi_j^1) \leq \epsilon \sum_{i=k}^{j-1} |f_i| \tag{2.16}$$

Since $\delta u_k > b$ in this case and $u \in \mathcal{F}$, $(\delta u_k + \delta u_j) \geq (1-r)\epsilon b$, and thus the left hand side of (2.16) tends to ∞ as $b \to \infty$, while the right hand side remains constant. Therefore, the only way (2.16) can hold for b sufficiently large is if $\xi_k^1 \leq \xi_j^1$.
Next, consider the case where $k > j$. In this case, when $\delta u_j - \delta u_k < 0$ the right hand side in (2.15) is negative, so we must have $\xi_k^1 \leq \xi_j^1$. If instead $\delta u_j - \delta u_k > 0$, we can obtain:

$$-\frac{\epsilon}{2}\left(\frac{\delta u_k}{\epsilon^2} + \frac{\delta u_j}{\epsilon^2}\right)(\xi_k^1 - \xi_j^1) \leq \epsilon \sum_{i=k}^{j-1} |f_i| \tag{2.17}$$

In this case δu_k must be negative, and thus $\delta u_k + \delta u_j < -(1-r)\epsilon b$. Hence as $b \to \infty$, $-\frac{\epsilon}{2}\left(\frac{\delta u_k}{\epsilon^2} + \frac{\delta u_j}{\epsilon^2}\right) \to \infty$, and (2.17) can only hold provided $\xi_k \leq \xi_j$ if b is large enough. Since j is arbitrary we have shown $\xi_k^1 \leq \xi_j^1$ in all possible cases, we can conclude $\min_j \xi_j^1 = \xi_k^1$. With this, the proof is completed by referring back to (2.3). □

This result was validated numerically by performing $1,000,000$ Monte Carlo trials and averaging ξ_k^1 at the break location for each solution exhibiting a break. This average was then compared to $\frac{1}{b}$ for various threshold sizes. The results are shown in Figure 2.1. The linear relationship suggests that the estimate provided in Theorem (2.1.4) captures the correct asymptotic behavior. Now by combining the results of this section, we can finally obtain a preliminary estimate on the failure probability $p(b)$.

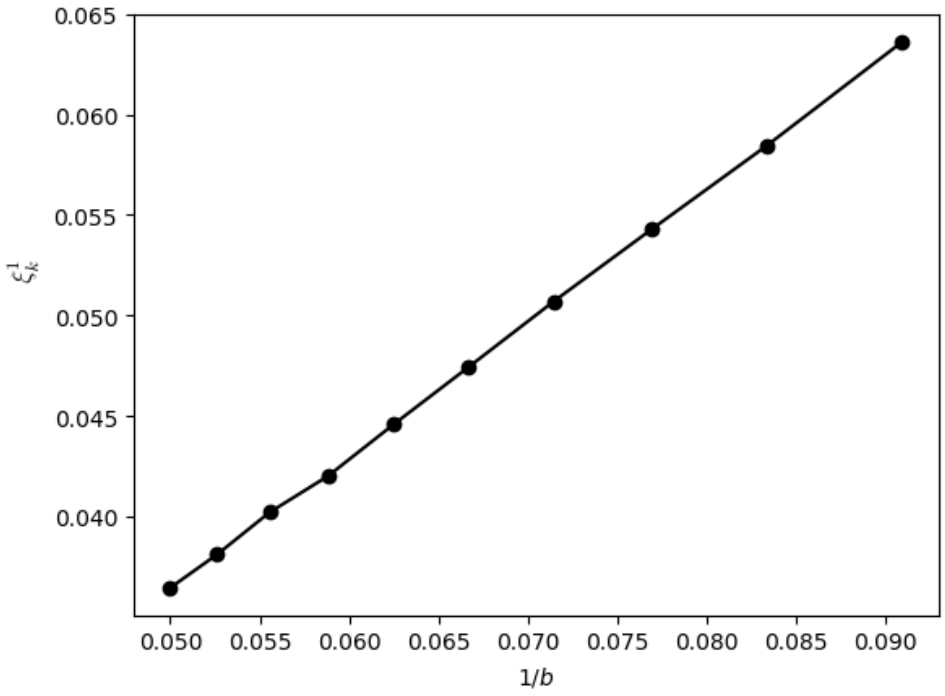

FIGURE 2.1: Comparison of ξ_k at the break location with $\frac{1}{b}$. We take $\epsilon = .1$, $f = 2.5 \cdot \mathbf{1}$, and average the ξ values over all solutions exhibiting breaks in $1,000,000$ trials.

Theorem 2.1.5. *Under the same assumptions as Lemmas 2.1.1 and 2.1.3, if* $\frac{b\sqrt{\epsilon}}{2||f||_{\ell_\epsilon^2}} > 1$*, we have the estimate*

$$p(b) \leq \sum_{i=1}^{N+1} \frac{\sqrt{\sigma_{ii}}}{\sqrt{2\pi}\ln\left(\frac{\sqrt{\epsilon}}{2||f||_{\ell_\epsilon^2}}b\right)} \exp\left(\frac{-\left(\ln\left(\frac{\sqrt{\epsilon}}{2||f||_{\ell_\epsilon^2}}b\right)\right)^2}{2\sigma_{ii}}\right)(1+o(1)) \tag{2.18}$$

as $b \to \infty$*.*

Proof. We begin by observing $p(b) = P(\mathcal{F}) + P(\mathcal{R})$. Based on Lemma 2.1.1,

$$P(\mathcal{F}) \leq P\left(\sup_i |\delta u_i| > \epsilon b)\right) \leq \sum_i P(|\delta u_i| > \epsilon b) \leq \sum_i P\left(\nu_i^1 > \ln\left(\frac{b\sqrt{\epsilon}}{2||f||_{\ell_\epsilon^2}}\right)\right)$$

As long as b is large enough to satisfy the conditions of the theorem, we can apply (1.26) and combine this with Lemma 2.1.3, to arrive at:

$$\begin{aligned} p(b) \leq & \sum_i \frac{\sqrt{\sigma_{ii}}}{\sqrt{2\pi}\ln\left(\frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}b\right)} \exp\left(\frac{-\left(\ln\left(\frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}b\right)\right)^2}{2\sigma_{ii}}\right) \\ & + \bar{C}\left(\frac{1}{\sqrt{2\pi}\ln(k_1 b)}\exp\left(\frac{-(\ln(k_1 b))^2}{2\sigma^M}\right)\right. \\ & + \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(k_3 b)} - \frac{\sqrt{\sigma^3_{ii}}}{\ln(k_3 b)^3}\right)^{-1} \frac{1}{\ln(k_2 b)^2} \\ & \left. \exp\left(\frac{-(\ln(k_2 b))^2}{2\sigma_{ii}}\right)\exp\left(\frac{(1-\alpha^2)\ln(k_2 b)^2}{2\sigma^M}\right)\right) \end{aligned} \tag{2.19}$$

Since $k_1 > \frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}$, $k_2 = \frac{\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}$, and $\alpha > 1$, the $P(\mathcal{R})$ term is negligible compared to the first term as $b \to \infty$, and the result follows. □

2.2 Explicit Solution

To obtain more accurate estimates on the failure probabilities, we require additional knowledge of the solution u to (1.11). In the case of the harmonic potential with nearest neighbor interactions, the model is simple enough that we can in fact calculate the solution to the minimization problem exactly. The first key observation is that the minimization problem can in this case be rewritten as a linear system of equations.

Lemma 2.2.1. *Suppose u is a solution to* (1.11). *Then for $k = 1, \ldots, N$, u_k satisfies the equation:*

$$\xi^1_k \delta u_k - \xi^1_{k+1} \delta u_{k+1} = \epsilon^2 f_k := f_{k,\epsilon} \tag{2.20}$$

Furthermore, we have the boundary conditions $\delta u_1 = u_1$ *and* $\delta u_{N+1} = -\delta u_N$.

Proof. Suppose $u = (u_1, \ldots, u_N)$ is the solution to (1.11). We construct a perturbed atom configuration $u^\alpha = (u_1, \ldots, u_{k-1}, u_k + \alpha\delta u_{k+1}, u_{k+1}, \ldots, u_N)$. Observe $\delta u^\alpha_j = \delta u_j$ for $j \neq k, k+1$, while $\delta u^\alpha_k = \delta u_k + \alpha \delta u_{k+1}$ and $\delta u^\alpha_{k+1} = (1-\alpha)\delta u_{k+1}$. Next, for a given k define a function $g(\alpha)$ by:

$$\begin{aligned} g(\alpha) = \mathcal{M}(u^\alpha) =& \frac{1}{2\epsilon} \left(\sum_{j \neq k, k+1} \xi^1_j \delta u_j^2 + \xi^1_k (\delta u_k + \alpha \delta u_{k+1})^2 + \xi^1_{k+1} (1-\alpha)^2 \delta u^2_{k+1} \right) \\ & - \epsilon \left(\sum_{j \neq k} f_j u_j + f_k (u_k + \alpha \delta u_{k+1}) \right) \end{aligned}$$

We can then calculate:

$$g'(\alpha) = \frac{1}{\epsilon} \left(\xi^1_k (\delta u_k + \alpha \delta u_{k+1}) \delta u_{k+1} - \xi^1_{k+1} (1-\alpha) \delta u^2_{k+1} \right) - \epsilon f_k \delta u_{k+1}$$

By assumption g has a minimum at $\alpha = 0$, so we should have $g'(0) = 0$. From this requirement, we obtain:

$$\frac{1}{\epsilon} \left(\xi^1_k \delta u_k \delta u_{k+1} - \xi^1_{k+1} \delta u^2_{k+1} \right) - \epsilon f_k \delta u_{k+1} = 0$$

If we move the f_k term to the right hand side and divide by δu_{k+1}, we obtain (2.20). The boundary conditions on δu follow immediately from the Dirichlet boundary conditions for u. □

In matrix form, this can be written as the system:

$$\begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix} \begin{bmatrix} \xi_1^1 \delta u_1 \\ \vdots \\ \vdots \\ \xi_N^1 \delta u_N \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \xi_{N+1}^1 \delta u_{N+1} \end{bmatrix} = \begin{bmatrix} f_{1,\epsilon} \\ \vdots \\ \vdots \\ f_{N,\epsilon} \end{bmatrix} \tag{2.21}$$

It is now straightforward to calculate an explicit formula for u. Note that in this expression, we interpret the sums as being 0 if the lower index exceeds the upper one.

Theorem 2.2.2. *The solution to the problem (2.21) is given by*

$$\delta u_k = \frac{1}{\xi_k^1} \left(\sum_{i=k}^{N} f_{i,\epsilon} - \left(\sum_{i=1}^{N+1} \frac{1}{\xi_i^1} \right)^{-1} \sum_{j=1}^{N} \sum_{i=j}^{N} \frac{f_{i,\epsilon}}{\xi_j^1} \right) \tag{2.22}$$

for $k = 1, ..N + 1$.

Proof. We solve the system via back-substitution. The last equation in the system (2.21) implies:

$$\delta u_N = \frac{1}{\xi_N^1} (f_{N,\epsilon} + \xi_{N+1}^1 \delta u_{N+1}) \tag{2.23}$$

Via an elementary induction argument, one can then show the general formula:

$$\delta u_k = \frac{1}{\xi_k^1} \left(\sum_{i=k}^{N} f_{i,\epsilon} + \xi_{N+1}^1 \delta u_{N+1} \right) \tag{2.24}$$

By using the right boundary condition, we have $\delta u_{N+1} = -u_N$. On the other hand, the left boundary condition implies $u_1 = \delta u_1$, so by equating (1.13) and (2.24) with

$k = 1$, we obtain:

$$u_N - \sum_{k=2}^{N} \frac{1}{\xi_k^1} \left(\sum_{i=k}^{N} f_{i,\epsilon} - \xi_{N+1}^1 u_N \right) = \frac{1}{\xi_1^1} \left(\sum_{i=1}^{N} f_{i,\epsilon} - \xi_{N+1}^1 u_N \right) \tag{2.25}$$

Solving for u_N yields the solution

$$u_N = \frac{1}{\xi_{N+1}^1} \left(\sum_{k=1}^{N+1} \frac{\xi_{N+1}^1}{\xi_k^1} \right)^{-1} \sum_{k=1}^{N} \sum_{i=k}^{N} \frac{f_{i,\epsilon}}{\xi_k^1} \tag{2.26}$$

By substituting (2.26) back into (2.24) and using $\delta u_{N+1} = -u_N$, we obtain the desired result. □

Some sample realizations of the solution u under the condition that a break occurs are shown in Figure 2.2. As we expect from the results of the previous section, the break location depends on where the nearest neighbor bond strength ξ^1 is weak. As a corollary of Theorem 2.2.2, we also obtain the following uniqueness result:

Corollary 2.2.3. *The solution to* (1.11) *with Dirichlet boundary conditions for the nearest neighbor harmonic model is unique.*

Proof. Any solution to (1.11) must satisfy (2.21). However, by Theorem. 2.2.2 for a given coefficient field ξ, there is only one atom configuration which solves (2.21), implying uniqueness. □

2.3 Failure Probabilities Under a Delta Forcing

Before continuing our analysis of the failure probability $p(b)$, we will restrict our attention to the case $f = e_\alpha$, where e_α are the standard basis vectors on $\mathbb{R}^N$ and

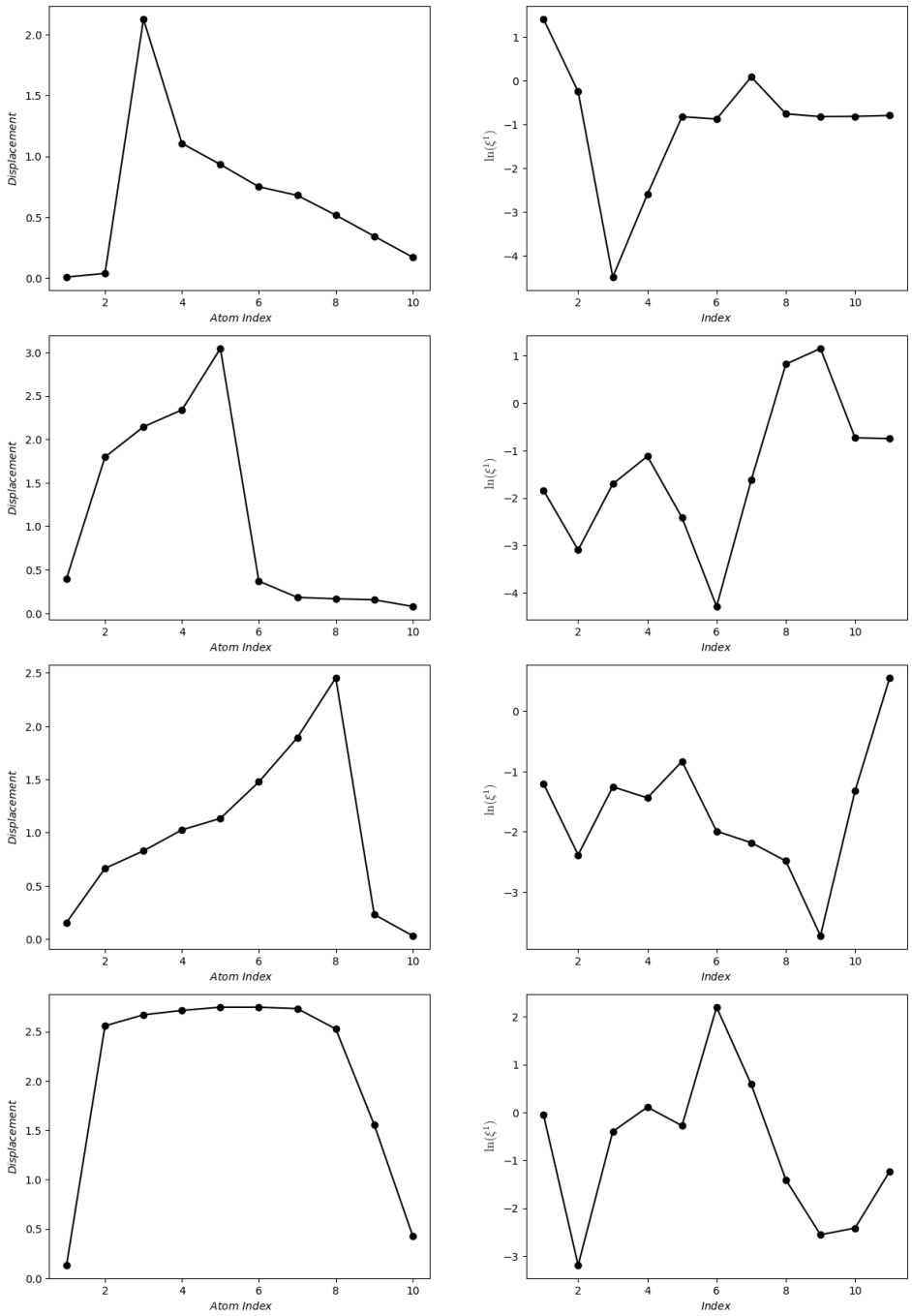

FIGURE 2.2: Solutions to (1.11) exhibiting a break with $b = 20$, $\epsilon = .1$. From top to bottom, we take $f = e_3$, $f = e_5$, $f = e_8$, and $f = 2.5 \cdot \mathbf{1}$. The left column shows the solutions u, while the right column displays the values $\ln(\xi^1)$ of the random coefficient field for the given solution. Notice that in each case, the minimal ξ^1 value occurs at the location where the solution exhibits a break.

$1 \leq \alpha \leq N$. Suppose $u^{(\alpha)}$ solves (2.20) for $f = e_\alpha$. Then for a general force $f = (f_1, \ldots, f_N)$, the configuration $u = \sum_{i=1}^N f_i u^{(i)}$ satisfies:

$$\begin{aligned}\xi^1_k \delta u_k - \xi^1_{k+1}\delta u_{k+1} &= \sum_{i=1}^N f_i(\xi^1_k \delta u^{(i)}_k - \xi^1_{k+1}\delta u^{(i)}_{k+1}) \\ &= f_k(\xi^1_k \delta u^{(k)}_k - \xi^1_{k+1}\delta u^{(k)}_{k+1}) = f_{k,\epsilon}\end{aligned}$$

In other words, by taking a superposition of the $u^{(\alpha)}$, we can obtain the solution for an arbitrary load vector f. As such, understanding the case of $f = e_\alpha$ can be used as a starting point for understanding more general problems (although we will not pursue that direction here). In the following we will scale the load vector as $f = \frac{1}{\epsilon} e_\alpha$, so that it forms a discrete analogue of a delta function.

With this simplification in mind, we shall wish to study the case of a localized external force in more detail. To this end, the following corollary to Theorem 2.2.2 is useful:

Corollary 2.3.1. *The solution to* (1.11) *with* $f = \frac{1}{\epsilon} e_\alpha$ *is given by*

$$\begin{aligned}\delta u_k &= \begin{cases} \frac{\epsilon}{\xi^1_k}\left(\sum_{j=1}^{N+1} \frac{1}{\xi^1_j}\right)^{-1} \sum_{j=\alpha+1}^{N+1} \frac{1}{\xi^1_j} & k \leq \alpha \\ -\frac{\epsilon}{\xi^1_k}\left(\left(\sum_{j=1}^{N+1} \frac{1}{\xi^1_j}\right)^{-1} \sum_{j=1}^{\alpha} \frac{1}{\xi^1_j}\right) & k > \alpha \end{cases} \\ &:= \frac{\epsilon}{\xi^1_k} F_k\end{aligned} \tag{2.27}$$

Proof. Let χ be the characteristic function, $\chi(k \leq \alpha) = \begin{cases} 1 & k \leq \alpha \\ 0 & k > \alpha \end{cases}$. Then substi-

tuting $f = \frac{1}{\epsilon} e_\alpha$ into (2.22), we get:

$$\begin{aligned}
\delta u_k =& \frac{\epsilon}{\xi_k^1} \left(\chi(k \leq \alpha) - \left(\sum_{j=1}^{N+1} \frac{1}{\xi_j^1} \right)^{-1} \sum_{j=1}^{\alpha} \frac{1}{\xi_j^1} \right) \\
=& \frac{\epsilon}{\xi_k^1} \begin{cases} \frac{\sum_{j=1}^{N+1} \frac{1}{\xi_j^1}}{\sum_{j=1}^{N+1} \frac{1}{\xi_j^1}} - \left(\sum_{j=1}^{N+1} \frac{1}{\xi_j^1} \right)^{-1} \sum_{j=1}^{\alpha} \frac{1}{\xi_j^1} & k \leq \alpha \\ - \left(\sum_{j=1}^{N+1} \frac{1}{\xi_j^1} \right)^{-1} \sum_{j=1}^{\alpha} \frac{1}{\xi_j^1} & k > \alpha \end{cases}
\end{aligned}$$

Simplification yields (2.27) □

Recall that ξ^1 can be written as $\xi^1 = [e^{-\nu_1^1}, \cdots, e^{-\nu_{N+1}^1}]^T$, where ν^1 is an $\mathcal{N}(0, \Sigma)$ random variable. With the explicit solution of (1.11) at our disposal, we can now easily observe that in order for $\left| \frac{\delta u_k}{\epsilon} \right|$ to exceed b, we must have the corresponding ν_k^1 be sufficiently large.

Proposition 2.3.2. *If u is the solution to (1.11) with $f = \frac{1}{\epsilon} e_\alpha$, and $|\delta u_k| > \epsilon b$ for some $k \in \{1, 2, \ldots, N+1\}$, then $\nu_k^1 > \ln(b)$.*

Proof. The key observation from (2.27) is that $|F_k| \leq 1$, so:

$$\epsilon e^{\nu_k^1} \geq |\delta u_k| > \epsilon b \implies \nu_k^1 > \ln(b)$$

□

This is a similar type of result to Theorem 2.1.4, so it is worth comparing the two. In the case of $f = \frac{1}{\epsilon} e_\alpha$ for some α, Theorem 2.1.4 will give

$$\nu_k^1 > \ln\left(\frac{\epsilon b}{2} \right)$$

In other words, we obtain the same asymptotic dependence on b, but the estimate obtained in Proposition 2.3.2 is sharper for small ϵ. However, Proposition 2.3.2 is more restricted in its use than Theorem 2.1.4 since we have made additional assumptions on f. In either case, we note that this observation of the relationship between large deviations of δu and large deviations of ν is similar to known results in the context of the continuum model (1.1), and is critical to our understanding of the system.

The converse of Proposition 2.3.2 is not true, $|\delta u_k| > \epsilon b$ is not sufficient to guarantee a break will occur in the material at index k due to the F_k factor. Hence, in order to understand $p_k(b)$, we must first better understand the behavior of F_k. The next lemma is a technical result which will be useful in our subsequent analysis.

Lemma 2.3.3. *Suppose* $\frac{1}{\xi_k^1} = M > b$. *Let* $w = \sum_{j\neq k} \frac{e^{\nu_j^1}}{M}$ *and define* $S_b = \{w < 1\}$. *Then provided* $\left|\frac{\sigma_{jk}}{\sigma_{kk}}\right| < 1$ *for all* $j, k \in \{1, 2, \ldots, N+1\}$ *with* $j \neq k$ *and* $\ln(b) > \left(1 - \frac{\sigma_{jk}}{\sigma_{kk}}\right)\ln(N)$, *we have the estimate:*

$$P(S_b^C \mid \nu_k^1 = \ln(M)) \leq \sum_{j\neq k} \frac{1}{t_{j,k}\sqrt{2\pi}} e^{-\frac{t_{j,k}^2}{2}}, \tag{2.28}$$

where $t_{j,k} = \dfrac{\left(1 - \frac{\sigma_{jk}}{\sigma_{kk}}\right)\ln(M) - \ln(N)}{\sqrt{\tilde{\sigma}_{j,k}}}$ *and* $\tilde{\sigma}_{j,k} = \sigma_{jj} - \frac{\sigma_{jk}^2}{\sigma_{kk}}$

Proof. We first observe that in order for w to exceed 1, we must have $e^{\nu_j^1} > \frac{M}{N}$ for at least one j, since there are a total of N terms in the sum. Consequently, using the

inclusion-exclusion principle,

$$P(S_b^C \mid \nu_k^1 = \ln(M)) \leq P(\nu_j^1 > \ln(M) - \ln(N) \text{ for some } j \neq k \mid \nu_k^1 = \ln(M))$$

$$\leq \sum_{j \neq k} P(\nu_j^1 > \ln(M) - \ln(N) \mid \nu_k^1 = \ln(M))$$

For a given index $j \neq k$, $(\nu_j^1 \mid \nu_k^1 = \ln(M))$ is normal with mean $\frac{\sigma_{jk}}{\sigma_{kk}} \ln(M) := \tilde{\mu}_{j,k} \ln(M)$ and variance $\tilde{\sigma}_{j,k}$ given by (1.25). Because $\ln(M) - \ln(N) > \tilde{\mu}_{j,k} \ln(M)$ by assumption, the probabilities in the sum can each be calculated using the tail distribution for a normal random variable, given in (1.26) and we obtain:

$$P(S_b^C \big| \nu_k^1 = \ln(M)) \leq \sum_{j \neq k} \frac{1}{t_{j,k}\sqrt{2\pi}} e^{-\frac{t_{j,k}^2}{2}} \tag{2.29}$$

□

Note this lemma immediately implies that as long as $\frac{\sigma_{jk}}{\sigma_{kk}} < 1$, we may conclude $P(S_b^C \mid \nu_k^1 = \ln(M)) \to 0$ as $b \to \infty$, so it will be sufficient to focus on the "nice" set S_b, provided the threshold for failure is sufficiently large. Let $\mathbb{E}_A(X)$ denote the expectation of X over the set A. Our next result will help us to better understand the behavior of F_k on the set S_b.

Lemma 2.3.4. *With the same notations/assumptions as in Lemma 2.3.3, further define* $A_k = \left\{ \sup_j \frac{1}{\xi_j^1} = \frac{1}{\xi_k} = M \right\}$, *for* $M > b$. *Then there exists a* B *such that for*

all $b > B$, we have:

$$\mathbb{E}_{S_b}\left(|F_k| \mid A_k\right) \leq (1+O(w)) \begin{cases} \frac{1}{M}\sum_{j=1}^{\alpha} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & \text{for } k > \alpha \\ \frac{1}{M}\sum_{j=\alpha+1}^{N+1} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & \text{for } k \leq \alpha \end{cases}, \tag{2.30}$$

where we have set $\bar{\mu}_{j,k} = \frac{\sigma_{jk}}{\sigma_{kk}} \ln(M)$.

Proof. We start with the case $k > \alpha$. In this case, we have the expression

$$F_k = \left(\sum_{j=1}^{N+1} \frac{1}{\xi_j^1}\right)^{-1} \sum_{j=1}^{\alpha} \frac{1}{\xi_j^1} = \sum_{j=1}^{\alpha} \frac{e^{\nu_j^1}}{M + \sum_{j \neq k} e^{\nu_j^1}} = \sum_{j=1}^{\alpha} \frac{1}{M} e^{\nu_j^1} (1+w)^{-1}$$

On the set S_b, we can expand $(1+w)^{-1}$ via a Taylor series to obtain:

$$F_k = \sum_{j=1}^{\alpha} \frac{1}{M} e^{\nu_j^1} (1 + O(w))$$

Consequently, we have:

$$\mathbb{E}_{S_b}(F_k \mid A_k) = \frac{1}{M} \sum_{j=1}^{\alpha} \mathbb{E}_{S_b}(e^{\nu_j^1} \mid \nu_k^1 = \ln(M))(1 + O(w)) \tag{2.31}$$

By a conditional Gaussian calculation, we observe for each j, $(\nu_j^1 \mid \nu_k^1 = \ln(M))$ is a $\mathcal{N}(\bar{\mu}_{j,k}, \tilde{\sigma}_{j,k})$ random variable. Hence by the properties of the log-normal distribution,

$$\mathbb{E}_{S_b}(F_k \mid A_k) = \frac{1}{M} \sum_{j=1}^{\alpha} (\mathbb{E}(e^{\nu_j^1} | \nu_k^1 = \ln(M))) - (\mathbb{E}_{S_b^C}(e^{\nu_j^1} | \nu_k^1 = \ln(M)))(1 + O(w)) \tag{2.32}$$

$$\leq \frac{1}{M}\sum_{j=1}^{\alpha}\exp\left(\bar{\mu}_{j,k}+\frac{\tilde{\sigma}_{j,k}}{2}\right)(1+O(w))$$

We now consider the case $k \leq \alpha$, and proceed in a similar manner as above. Provided we restrict ourselves to the set S_b, we can once again expand F_k in terms of w:

$$F_k = -\left(\sum_{j=1}^{N+1}\frac{1}{\xi_j^1}\right)^{-1}\sum_{j=\alpha+1}^{N+1}\frac{1}{\xi_j^1} = -\sum_{j=\alpha+1}^{N+1}\frac{e^{\nu_j^1}}{M+\sum_{j\neq k}e^{\nu_j^1}}$$
$$= -\sum_{j=\alpha+1}^{N+1}\frac{1}{M}e^{\nu_j^1}(1+w)^{-1} = -\sum_{j=\alpha+1}^{N+1}\frac{1}{M}e^{\nu_j^1}(1+O(w))$$

Notice we have arrived at almost the same expression as before, but the sum is over different indices. Thus by an essentially identical argument to the one given above, we can obtain:

$$\mathbb{E}_{S_b}(|F_k| \mid A_k) \leq \frac{1}{M}\sum_{j=\alpha+1}^{N+1}\exp\left(\bar{\mu}_{j,k}+\frac{\tilde{\sigma}_{j,k}}{2}\right)(1+O(w)) \tag{2.33}$$

Combining our results for $k > \alpha$ and $k \leq \alpha$ yields the result stated above. □

We next consider $p_k(b)$ for a fixed k. From the exact solution (2.27), we know $\delta u_k = \frac{1}{\xi_k^1}F_K$, and so for $|\delta u_k|$ to exceed b, we must have $|F_k| > \frac{\xi_k^1 b}{\epsilon}$. Using the previous results, we can estimate the probability this occurs.

Lemma 2.3.5. *Suppose for all $k, j \in \{1, 2, \ldots, N+1\}$, we have $\left|\frac{\sigma_{jk}}{\sigma_{kk}}\right| < 1$. Then for*

any $h > 0$, $R > 1$, there exists a B such that for all $b > B$,

$$p_k(b) \leq h + \frac{(1+O(w))\sqrt{\sigma_{kk}}}{\epsilon b \ln(b)\sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{\sigma_{kk}}\right) \begin{cases} \sum_{j=1}^{\alpha} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_j}{2}\right) & k > \alpha \\ \sum_{j=\alpha+1}^{N+1} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_j}{2}\right) & k \leq \alpha \end{cases}, \tag{2.34}$$

with $\tilde{\mu}_{j,k} = \frac{\sigma_{jk}}{\sigma_{kk}} \ln(Rb)$.

Proof. By Proposition 2.3.2, we know that a break occurring at location k implies $\nu_k^1 > \ln(b)$, so if D_k is the event $\{\nu_k^1 > \ln(b)\}$, we can conclude

$$\begin{aligned} P(|\delta u_k| > \epsilon b) &= P(|\delta u_k| > \epsilon b \cap D_k) = P(|\delta u_k| > \epsilon b \mid D_k) P(D_k) \\ &= P(|F_k| > \epsilon b e^{-\nu_k^1} \mid D_k) P(D_k) \\ &= \int_{D_k} P(|F_k| > \epsilon b e^{-\nu_k^1} \mid \nu_k^1 = \ln(M)) \phi_c(M) dM P(D_k), \end{aligned} \tag{2.35}$$

where $\phi_c(M)$ is the conditional distribution of ν_k^1 given D_k. Now by applying the conditional form of Markov's inequality, we obtain:

$$\begin{aligned} P(|\delta u_k| > \epsilon b) &\leq \int_{D_k} \frac{M}{\epsilon b} \mathbb{E}\Big(|F_k| \mid \nu_k^1 = \ln(M)\Big) \phi_c(M) dM P(D_k) \\ &\leq \int_{D_k} \frac{M}{\epsilon b} \Bigg(\mathbb{E}_{S_b}\Big(|F_k| \mid \nu_k^1 = \ln(M)\Big) \phi_c(M) P(D_k) \\ &\qquad + \mathbb{E}_{S_b^C}\Big(|F_k| \mid \nu_k^1 = \ln(M)\Big) \phi_c(M) P(D_k) \Bigg) dM \end{aligned} \tag{2.36}$$

Note $P(D_k)$ can be easily estimated using the tail distribution for a Gaussian random

variable:

$$P(D_k) \leq \frac{\sqrt{\sigma_{kk}}}{\ln(b)\sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{kk}}\right)$$

The expectation over S_b^C can be bounded as follows. We recall from the proof of the previous lemma that $|F_k| \leq \sum_{j=1}^{N+1} \frac{1}{M} e^{\nu_j^1}(1 + O(w))$. Also, we know by assumption that $e^{\nu_j^1} \leq e^{\nu_k^1} = M$ so using (2.28) we have:

$$\mathbb{E}_{S_b^C}(e^{\nu_j^1} \mid \nu_k^1 = \ln(M)) \leq MP(S_b^C \mid \nu_k^1 = \ln(M)) \leq M \sum_{j \neq k} \frac{1}{t_{j,k}\sqrt{2\pi}} e^{-\frac{t_{j,k}^2}{2}} \tag{2.37}$$

Because we've stipulated $\left|\frac{\sigma_{jk}}{\sigma_{kk}}\right| < 1$, we can conclude $M\frac{1}{t_{j,k}\sqrt{2\pi}}e^{-\frac{t_{j,k}^2}{2}} \to 0$ as $b \to \infty$. Consequently, for b sufficiently large, we can guarantee for any $h > 0$,

$$\frac{M}{\epsilon b} P(D_k)\phi_c(M)\mathbb{E}_{S_b^C}(|F_k| \mid \nu_k^1 = \ln(M)) \leq \frac{h}{2}$$

So, provided b is large enough, we can return to (2.36) and ensure that

$$\frac{M}{\epsilon b} P(D_k)\mathbb{E}\Big(|F_k| \mid \nu_k^1 = \ln(M)\Big) \leq \frac{M}{\epsilon b} P(D_k)\mathbb{E}_{S_b}\Big(|F_k| \mid \nu_k^1 = \ln(M)\Big) + \frac{h}{2} \tag{2.38}$$

Next, applying (2.30) gives us:

$$p_k(b) \leq \int_{D_k} \left(\frac{(1 + O(w))\sqrt{\sigma_{kk}}}{\epsilon b \ln(b)\sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{k,k}}\right) \begin{cases} \sum_{j=1}^{\alpha} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k > \alpha \\ \sum_{j=\alpha+1}^{N+1} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k \leq \alpha \end{cases} \right.$$

$$\left. + \frac{h}{2} \right) \phi_c(M) dM \tag{2.39}$$

Now let $S_M = \{M < Rb\}$, $R > 1$, and divide the integral into two pieces: one over S_M and one over S_M^C. Note the integrand in (2.39) is bounded thanks to the assumption $\left|\frac{\sigma_{jk}}{\sigma_{kk}}\right|$ and the $\phi_C(M)$ factor. Let us call this bound Y. It is clear that $P(S_M^C) \to 0$ as $b \to \infty$, so by taking b larger if necessary, we can ensure $P(S_M^C) \leq \frac{h}{2Y}$. Hence by continuing from (2.39),

$$
\begin{aligned}
p_k(b) \leq & \int_{S_M} \left(\frac{(1+O(w))\sqrt{\sigma_{kk}}}{\epsilon b \ln(b)\sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{k,k}}\right) \begin{cases} \sum_{j=1}^{\alpha} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k > \alpha \\ \sum_{j=\alpha+1}^{N+1} \exp\left(\bar{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k \leq \alpha \end{cases} \right. \\
& \left. + \frac{h}{2} \right) \phi_c(M) dM + \frac{h}{2} \\
\leq & \left(\frac{(1+O(w))\sqrt{\sigma_{kk}}}{\epsilon b \ln(b)\sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{k,k}}\right) \begin{cases} \sum_{j=1}^{\alpha} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k > \alpha \\ \sum_{j=\alpha+1}^{N+1} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2}\right) & k \leq \alpha \end{cases} \right. \\
& \left. + \frac{h}{2} \right) \int_{S_M} \phi_c(M) dM + \frac{h}{2}
\end{aligned}
$$

Since the integral of a pdf cannot exceed 1, this concludes the proof. □

With the above preparations, we are now able to easily prove the main result regarding the failure probability $p(b)$ for the nearest neighbor model with a harmonic potential.

Theorem 2.3.6. *With the same setup as in Lemma 2.3.5, for the problem* (2.20)

with $f = \frac{1}{\epsilon} e_\alpha$,

$$
\begin{aligned}
p(b) \leq & \left(\sum_{k=1}^{\alpha} \frac{\sqrt{\sigma_{kk}}}{\epsilon b \ln(b) \sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{kk}} \right) \sum_{j=\alpha+1}^{N+1} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2} \right) \right. \\
& \left. + \sum_{k=\alpha+1}^{N+1} \frac{\sqrt{\sigma_{kk}}}{\epsilon b \ln(b) \sqrt{2\pi}} \exp\left(-\frac{\ln(b)^2}{2\sigma_{kk}} \right) \sum_{j=1}^{\alpha} \exp\left(\tilde{\mu}_{j,k} + \frac{\tilde{\sigma}_{j,k}}{2} \right) \right) (1 + O(w)) \\
& + h(N+1)
\end{aligned}
\tag{2.40}
$$

Proof. We first observe that, using sub-additivity,

$$
p\left(\sup_k |\delta u_k| > \epsilon b \right) = p(\cup_k \{ |\delta u_k| > \epsilon b \}) \leq \sum_{k=1}^{N+1} p(|\delta u_k| > \epsilon b)
$$

By applying Lemma 2.3.5 to bound the terms in this sum, we arrive at the result through straightforward algebraic manipulations. □

Very roughly speaking, the theorem suggests an asymptotic upper bound for the break probability behaves like $\frac{1}{b \ln(b)} \sup_{j,k} \exp(\tilde{\mu}_{j,k} - \ln(b)^2) \approx \frac{1}{\ln(b)} \exp(-\ln(b)^2)$. It is worth pointing out the similarity of this result to the estimate (1.2) obtained in [18] for a delta external force. Furthermore, this will in general give a sharper estimate of the failure probability than Theorem 2.1.5. However, we do note the usefulness of this result is somewhat limited by the restriction on the load vector f.

3

Importance Sampling Algorithm for Nearest Neighbor Harmonic Model

As seen in Theorem 2.1.5, the failure probabilities $p(b)$ decay with b at a rapid rate, and hence the probability of observing a failure event is quite small for large b values. This means that if we use a direct Monte Carlo method, i.e. simulate the system for a large number of randomly generated ξ^1 values and see how many exhibit a break, we will require millions of trials to obtain highly accurate results. For especially large values of the threshold b, it may not be possible to observe any failure events in practice in a feasible amount of time. This suggests we need better methods for simulating these rare events. In this chapter, we will use the intuition from our previous analysis, that a large value of $|\delta u|$ can only occur if we observe a correspondingly large value in the coefficient field ν^1, to develop a more efficient importance sampling algorithm for computing failure probabilities. Not only will this method yield estimates of the failure probability for larger values of b than Monte Carlo simulation, using the importance sampling estimator will also result in signif-

icantly reduced variance in our calculation of $p(b)$.

In section 3.1 we provide an overview of our algorithm and its theoretical foundation, with supporting numerical experiments provided in section 3.2. Lastly, we shall briefly discuss the performance of our algorithm, as well as potential methods of improving it, in section 3.3.

3.1 Algorithm Overview

To begin we choose a threshold value τ which we expect $\sup_i \nu_i^1$ should exceed in order for a break to occur. The choice of τ is informed by our analysis in chapter 2: for example, $\tau_u = \ln\left(\frac{b\sqrt{\epsilon}}{2||f||_{\ell^2_\epsilon}}\right)$ is a good choice for a general forcing, while in the special case of a localized f, we might try simply $\tau_\delta = \ln(b)$. The idea of our importance sampling algorithm is that we can switch to a probability measure under which we are more likely to see a failure event occur by making large values in ν^1 more likely to occur.

Let P be the original probability distribution for ν^1, and Q the modified distribution we draw our samples from. The rare event $|\delta u_i| > \epsilon b$ will far more likely under Q than under P, so $p_Q(b)$ can be estimated using far fewer realizations than are required by a direct Monte Carlo method. Once we obtain an estimate for $Q\left(\sup_i |\delta u_i| > \epsilon b\right)$, we can obtain $p(b)$ from

$$P\left(\sup_i |\delta u_i| > \epsilon b\right) = \int_\Omega \chi\left(\sup_i |\delta u_i| > \epsilon b\right) dP = \int_\Omega \chi\left(\sup_i |\delta u_i| > \epsilon b\right) \frac{dP}{dQ} dQ, \tag{3.1}$$

where $\frac{dP}{dQ}$ is the Radon-Nikodym derivative. If P and Q are probability distributions with densities f_P and f_Q, respectively, then a straight-forward calculation shows:

$$\begin{aligned} P(A) &= \int_\Omega \chi_A(\omega) dP = \int_{\mathbb{R}^{N+1}} \chi_A(x) f_p(x) dx \\ &= \int_{\mathbb{R}^{N+1}} \chi_A(x) \frac{f_P(x)}{f_Q(x)} f_Q(x) dx = \int_\Omega \chi_A \frac{f_P}{f_Q} dQ \end{aligned}$$

From this we can see directly from the definition of the Radon-Nikodym derivative that $\frac{dP}{dQ} = \frac{f_P}{f_Q}$.

We shall now address how to choose the proposal distribution Q. To generate a sample from Q, we begin by choosing a random index j uniformly from the set $\{1, 2, 3, \ldots, N+1\}$. j shall be the location where we expect the chain to break, so we shift the distribution of ν_j^1 to be $\mathcal{N}(\tau, \sigma_{jj}^P)$. Let f_j^Q be the probability density function for this distribution. Next, we generate the remaining ν_i^1, $i \neq j$, according to their original distribution conditioned on the realized value of ν_j^1 (using (1.25)). The resulting density for distribution of the ν_i^1 shall be denoted as $\tilde{f}_j^Q$, and its mean and covariance by $\tilde{\mu}^j, \tilde{\Sigma}^j$.

From this procedure, for the chosen value of j we obtain a normal random vector ν^1, with mean $\mu_Q^j = (\tilde{\mu}_1^j, \ldots, \tilde{\mu}_{j-1}^j, \tau, \tilde{\mu}_{j+1}^j, \ldots \tilde{\mu}_{N+1}^j)$ and covariance matrix:

$$\Sigma_Q^j = \begin{bmatrix} \tilde{\Sigma}_{11}^j & 0 & \tilde{\Sigma}_{12}^j \\ 0 & \sigma_P^{jj} & 0 \\ \tilde{\Sigma}_{21}^j & 0 & \tilde{\Sigma}_{22}^j \end{bmatrix} \tag{3.2}$$

Note the 0's are row/column vectors as appropriate. $\tilde{\Sigma}_{11}^j$ is the collection of entries of $\tilde{\Sigma}^j$ with both row and column indices less than j, $\tilde{\Sigma}_{12}^j$ is the collection of entries

with row index less than j, column index greater than j ($\tilde{\Sigma}^j_{21}$ is the reverse of this), and $\tilde{\Sigma}^j_{22}$ the collection of entries with both indices greater than j.

Now that we have described how to generate samples of Q, we must calculate the density f_Q so we can determine the Radon-Nikodym derivative. For a set $A \subseteq \mathbb{R}^{N+1}$,

$$\begin{aligned} Q(\nu \in A) &= \sum_{i=1}^{N+1} Q(\nu \in A | j = i) P(j = i) \\ &= \frac{1}{N+1} \sum_{i=1}^{N+1} \int_{\mathbb{R}} Q(\nu \in A | j = i, \nu_i = x_i) f_i^Q dx_i \\ &= \frac{1}{N+1} \sum_{i=1}^{N+1} \int_{A_j} \int_{\hat{A}_j} f_i^Q \tilde{f}_i^Q d\hat{x}_i dx_i = \frac{1}{N+1} \int_A \sum_{i=1}^{N+1} f_i^Q \tilde{f}_i^Q dx \end{aligned}$$

From the calculation, we conclude $f_Q = \frac{1}{N+1} \sum_{i=1}^{N+1} f_i^Q \tilde{f}_i^Q$. Since the product $f_i^Q \tilde{f}_i^Q$ can be seen to be just the pdf for the $\mathcal{N}(\mu_Q^i, \Sigma_Q^i)$ distribution, this sum can easily be calculated for each sample we generate from Q.

For the new probability measure Q to give more efficient estimates, it should produce more failure events than sampling from the original distribution in the same number of trials. This is verified in Figures 3.1 and 3.2. This disparity in the number of failure events between the two methods is particularly pronounced in Figure 3.2, since in that case we have a sharper value of τ due to the analysis of the exact solution.

Compiling our calculations thus far leads us to the following importance sampling algorithm:

Algorithm 2: NN Importance Sampling

1. Choose an index $j \in \{1, 2, 3, \ldots, N+1\}$ with uniform probability.

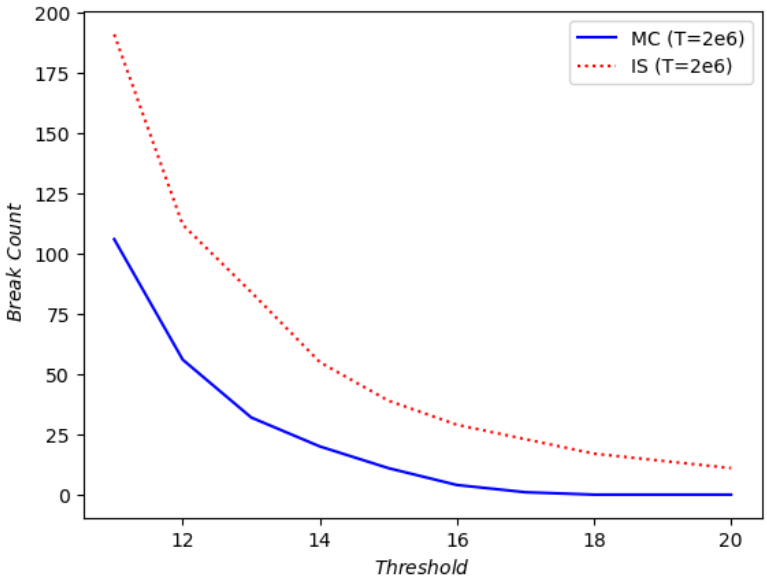

FIGURE 3.1: Comparison of the number of failure events observed from Monte Carlo and our importance sampling algorithm for a single run of 2 million trials and a uniform force $f = \mathbf{1}$.

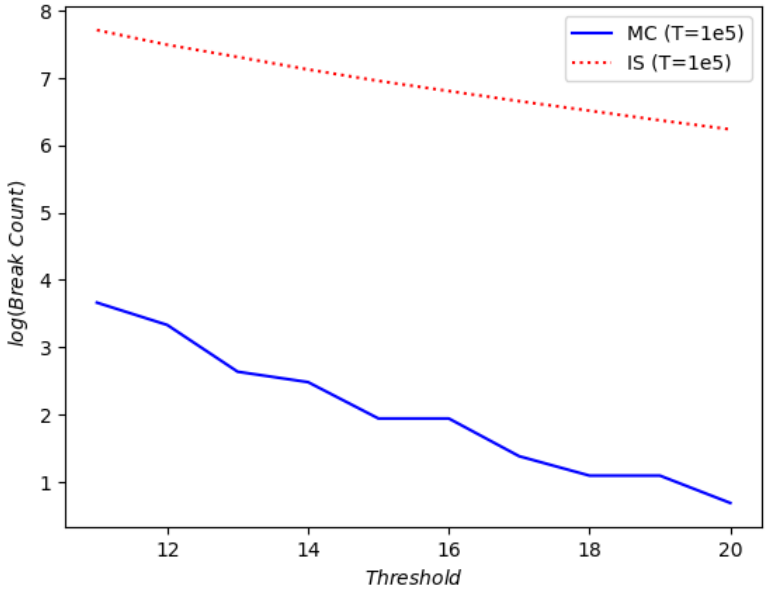

FIGURE 3.2: Comparison of the number of failure events observed from Monte Carlo and our importance sampling algorithm for 100,000 trials and a localized force $f = \frac{1}{\epsilon} e_6$.

2. Generate ν_j^1 according to the distribution $\mathcal{N}(\tau, \sigma_{jj}^P)$

3. Conditional on the realized ν_j^1, generate the remaining ν_i.

4. Solve for δu given the generated ν^1 and check if $\left|\sup_i \delta u_i\right| > \epsilon b$.

5. Repeat steps 1-4 T_{IS} times.

6. Our estimate for $P\left(\sup_i |\delta u_i| > \epsilon b\right)$ is $\frac{1}{T_{IS}} \sum_{k=1}^{T_{IS}} \frac{dP}{dQ}((\nu^1)^{(k)}) \chi\left(\sup_i |\delta u_k^{(i)}| > \epsilon b\right)$.

3.2 Numerical Experiments

To illustrate the effectiveness of our algorithm, we shall consider two test problems: one with a uniform force $f = \mathbf{1}$, and one with $f = \frac{1}{\epsilon} e_6$. For numerical simulations, we take $N = 10$ and generate a baseline estimate of $p(b)$ by performing $5 \cdot 10^7$ Monte Carlo trials. For the uniform forcing, we use the general estimate for the importance sampling shift τ_u and perform $2 \cdot 10^6$ trials using both direct Monte Carlo and importance sampling, while for the localized force we shall use the more aggressive shift τ_δ and perform 10^5 trials for both methods.

The results of the simulations for a uniform force are listed in Table 3.1. We observe from the data that the IS method provides better estimates of the failure probability than direct Monte Carlo, and also reduces the relative error of the estimator. The comparison between the three methods of calculating $p(b)$ is displayed graphically on a log scale in Figure 3.3.

Similar data for $f = \frac{1}{\epsilon} e_6$ is presented in Table 3.2 and Figure 3.4. One thing we observe is that the advantage of the IS algorithm is even more pronounced in this case. Even for a fairly small number of trials, it produces an estimate of $p(b)$

Table 3.1: Numerical results for a uniform force. For MC(2e6) and IS, we use 2e6 trials. The MC(5e7) gives a baseline estimate for comparison, and is generated using a MC simulation with 5e7 trials. A - indicates no estimate was obtained because no failure events were observed during the simulation.

b	-	12	14	16	18	20
$p(b)$	MC(2e6)	2.8e-5	1.0e-5	2.0e-6	-	-
	IS	2.09e-5	7.47e-6	2.58e-6	1.05e-6	4.58e-7
	MC(5e7)	2.49e-5	8.5e-6	3.16e-6	1.2e-6	4.2e-7
Std. Dev.	MC(2e6)	3.74e-6	2.24e-6	1.00e-6	-	-
	IS	2.10e-6	1.03e-6	4.90e-7	2.62e-7	1.42e-7
	MC(5e7)	7.05e-7	4.12e-7	2.51e-7	1.54e-7	9.12e-8
Rel. Err.	MC(2e6)	.134	.224	.500	-	-
	IS	.100	.137	.190	.248	.309
	MC(5e7)	.028	.049	.080	.129	.218

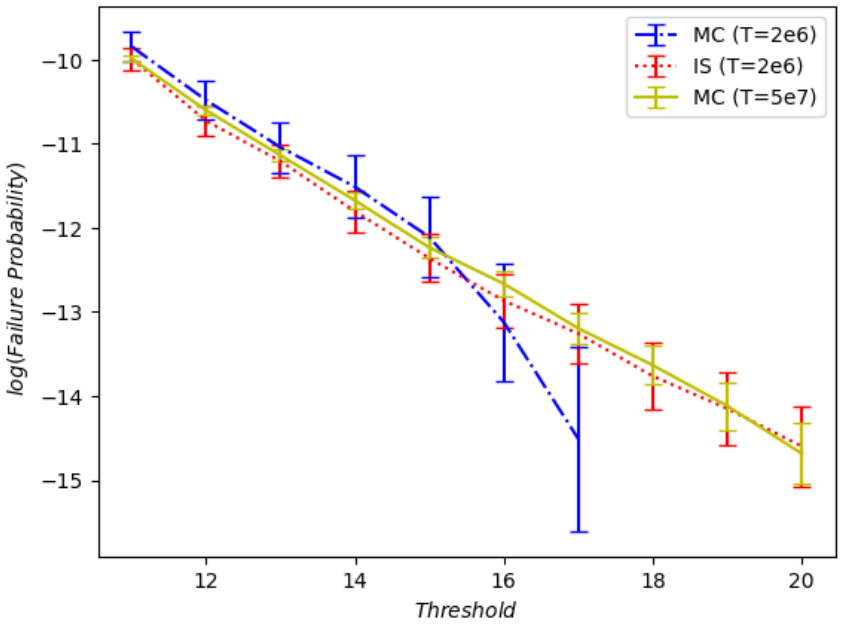

FIGURE 3.3: A comparison of the two MC estimates, and the IS estimate of $p(b)$ for various b values in the case of a uniform f. The IS algorithm is more accurate than MC across all threshold values and is able to generate estimates for a larger range of threshold values as well.

which has low relative error and agrees well with the longer Monte Carlo simulation across all b values. The reason for this improvement is due to the fact we have a

more precise estimate of the threshold ν_j^1 must exceed for a failure event to occur, based on our analysis of the solution to (1.11) in chapter 2. This allowed us to select a better value for τ in constructing the probability measure Q. If a similar analysis could be performed in the uniform f case, it is likely we could improve the efficiency of the algorithm in that scenario as well.

While it is difficult to generate accurate estimates of $p(b)$ in a reasonable amount

Table 3.2: Numerical results for a force $f = \frac{1}{\epsilon}e_6$. For MC(1e5) and IS, we use 1e5 trials. The MC(5e7) gives a baseline estimate for comparison, and is generated using a MC simulation with 5e7 trials.

b	-	12	14	16	18	20
$p(b)$	MC(1e5)	2.8e-4	1.2e-4	7.0e-5	3.0e-5	2.0e-5
	IS	2.72e-4	1.04e-4	4.23e-5	1.94e-5	8.92e-6
	MC(5e7)	2.83e-4	1.09e-4	4.73e-5	2.16e-5	1.06e-5
Std. Dev.	MC(1e5)	5.29e-5	3.46e-5	2.65e-5	1.73e-5	1.41e-5
	IS	8.97e-6	4.35e-6	2.16e-6	1.19e-6	6.50e-7
	MC(5e7)	2.38e-6	1.48e-6	9.72e-7	6.58e-7	4.61e-7
Rel. Err.	MC(1e5)	.189	.289	.378	.577	.707
	IS	.033	.042	.051	.061	.073
	MC(5e7)	.008	.014	.021	.030	.043

of time using MC when b is large, this is less of a restriction for the IS algorithm. Therefore, we may now use the IS algorithm to verify the asymptotic behavior of $p(b)$ derived in Theorem 2.3.6. The results are shown in Figure 3.5. The constant gap between the two curves suggests the theorem provides the correct b dependence for $p(b)$ but does not have the best possible prefactor.

3.3 Efficiency

While the results of the previous section support the superior performance of the IS algorithm compared to MC simulations, we must also consider that each trial of the IS simulation will be more costly than a trial of MC: generating samples from Q

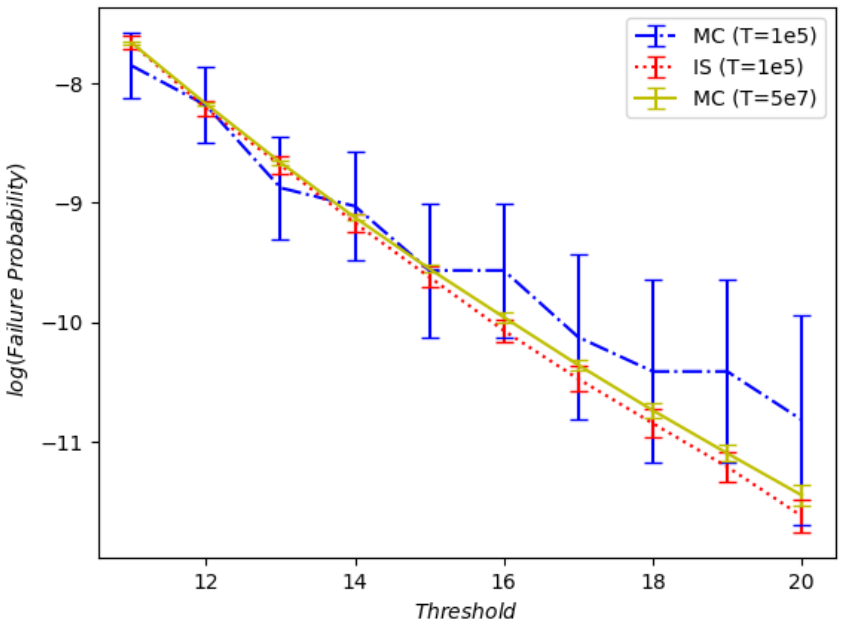

FIGURE 3.4: A comparison of the two MC estimates, and the IS estimate of $p(b)$ for various b values in the case of $f = \frac{1}{\epsilon}e_6$.

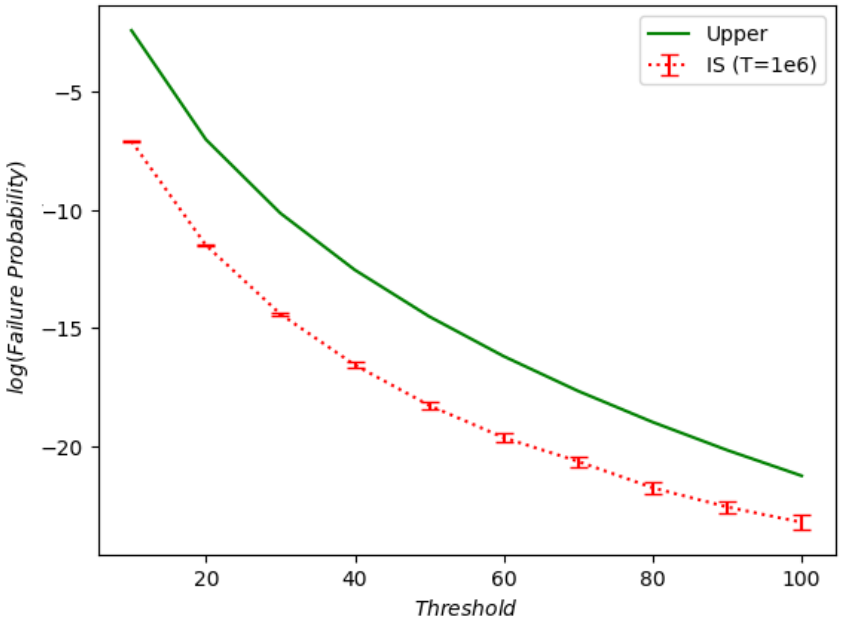

FIGURE 3.5: Comparison of the upper bound in (2.3.6) and numerical calculation of $p(b)$ for b values ranging from 10 to 100.

is more expensive than generating samples of P, and the IS method requires additional calculations to recover $p(b)$ using the Radon-Nikodym derivative. This creates a trade-off between the reduced number of trials to achieve a low relative error and the increased computation time per trial. To ensure IS is indeed saving time over MC, we compared the relative error of the two estimators based on the length of the simulation for several numbers of trials. Figures 3.6 and 3.7 illustrate that for a given threshold value, the IS estimator provides a much smaller error than MC in the same amount of time dedicated to the simulation.

These results, however, can be somewhat misleading because they only consider

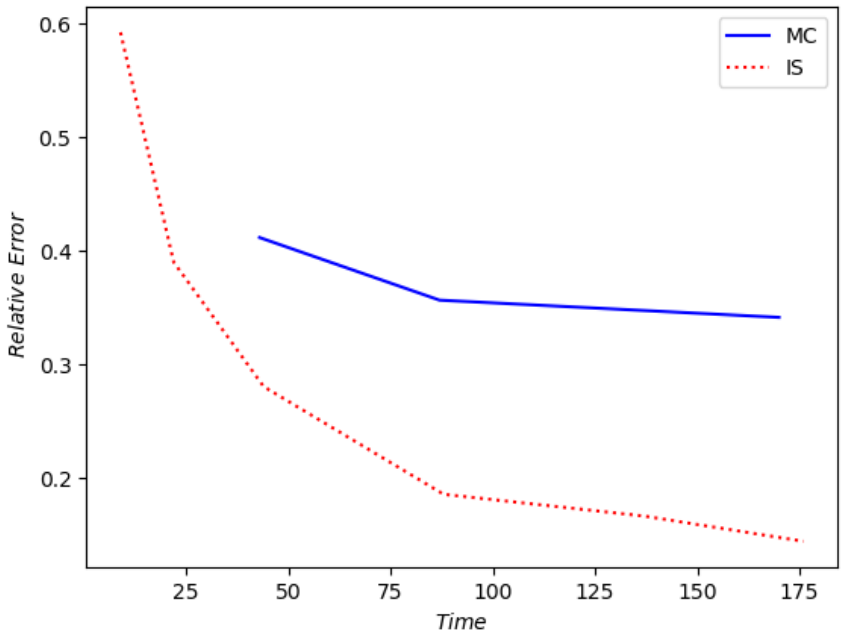

FIGURE 3.6: Comparison of computation time (in seconds) and the resulting error for MC and IS, in the case of $f = \mathbf{1}$ and $b = 15$.

the calculation for a single value of the failure threshold. A major computational disadvantage of the importance sampling method is that the choice of τ depends on b, so if we are interested in considering a wide range of b values (as in Figures 3.3 and 3.4) then we must recompute the coefficients ν^1 and the solution u for each thresh-

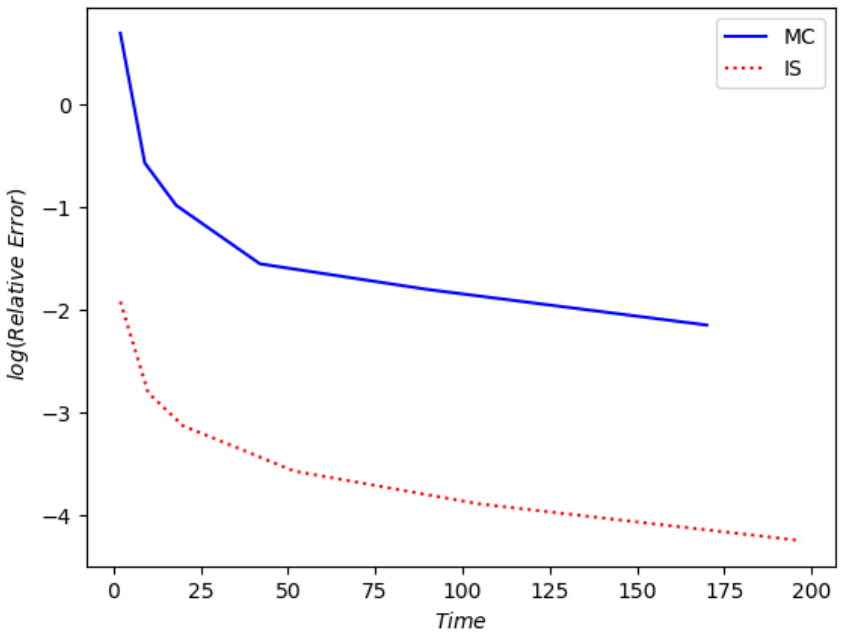

FIGURE 3.7: Comparison of computation time (in seconds) and the resulting error for MC and IS, in the case of $f = \frac{1}{\epsilon}e_6$ and $b = 15$.

old, effectively multiplying the number of trials being performed by the number of threshold values under consideration. On the other hand, in the MC method we can easily test for failure events at all threshold levels after generating only a single realization of u. As a result, MC with a large number of trials may be more useful if we wish to calculate $p(b)$ for a large number of relatively small b values. On the other hand, the advantages of the IS method are more pronounced if we are interested in a smaller number of threshold values or large values of b.

One possible remedy for this problem would be to fix τ for all threshold values under consideration (say, choose the τ given by the smallest b value), and use that as the shift in the distribution Q. We shall not pursue this direction here, but one would expect this would significantly reduce computation time for IS at the cost of some accuracy for higher threshold values due to the choice of a sub-optimal τ. Another approach to further increase efficiency of the importance sampling method

would be to try and improve the proposal distribution Q so that samples drawn from Q are more tightly focused on the most significant events contributing to the failure probability. For example, in step 2 of Algorithm 2, we might consider drawing ν_j^1 from a truncated normal distribution with lower bound $c = \tau$. This will ensure all samples drawn from Q will at least satisfy the minimal criteria for a break event, namely $\sup_j \nu_j^1 > \tau$. Using our previous calculations, it is not difficult in this case to see that the new density for Q still has the form:

$$\frac{1}{N+1} \sum_{i=1}^{N+1} f_i^Q \tilde{f}_i^Q$$

However, now f_i^Q will be the density for a truncated normal distribution, given by (1.27).

If the truncation value τ is well chosen, this can further improve the accuracy of the IS method, as is illustrated in Figure 3.8 for a localized external force. However, if τ is chosen too small or too large, this method may actually perform worse than using a simple shifted Gaussian, as it focuses the samples of Q on the wrong area. Thus we will not pursue this approach for more general problems where our choice of τ is more conservative.

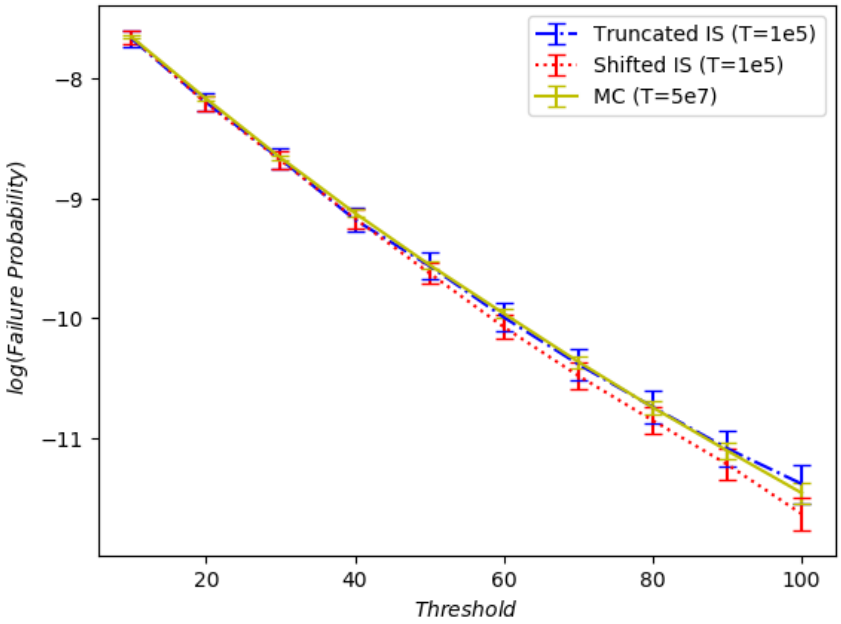

FIGURE 3.8: Comparison of different choices for the importance sampling measure Q for the case $f = \frac{1}{\epsilon}e_6$ and 100,000 trials for each IS method. We can see that using a truncated normal distribution provides greater agreement with the "true" probability value obtained from a long MC simulation.

4

Failure Probabilities for Next Nearest Neighbor Harmonic Model

In this chapter, we shall consider a more complicated model which allows for atoms to interact with their second nearest neighbors as well. We again focus on a harmonic potential (2.1) for V_1, and the potential V_2 is chosen to be:

$$V_2(u) = \frac{1}{2}(u-2)^2 \tag{4.1}$$

Under these conditions, (1.8) becomes:

$$E_H^{NNN}(u) = \frac{\epsilon}{2}\sum_{k=1}^{N+1}\xi_k^1\left(\frac{\delta u_k}{\epsilon}\right)^2 + \frac{\epsilon}{2}\sum_{k=1}^{N+2}\xi_k^2\left(\frac{\delta u_k + \delta u_{k-1}}{\epsilon}\right)^2 \tag{4.2}$$

We shall proceed using ideas analogous to those developed in chapter 2. First, in section 4.1, we will obtain probability estimates by using variational arguments similar to those in chapter 2.1. We will also calculate the exact solution to the

energy minimization problem (1.11) using Euler-Lagrange equations in section 4.2: however, unlike in the nearest neighbor case, the solution is quite complicated and not amenable to further analysis.

4.1 Variational Estimates

Based on the intuition gained from our earlier study of the nearest neighbor model, we expect that a failure event will only occur if the ξ coefficients near the location where $|\delta u|$ exceeds ϵb are particularly small. However, unlike in the nearest neighbor case, terms involving δu_k are directly influenced by three ξ coefficients: ξ_k^1, ξ_k^2, and ξ_{k+1}^2. Our first goal is to make precise the notion that, with high probability, these particular ξ coefficients must be small for a break in the atom chain to occur at location k. To this end, it shall be useful to introduce a couple new notations.

First, we let $\xi = (\xi^1, \xi^2) \in \mathbb{R}^{2N+3}$, and define a set $B_W(\xi)$ to keep track of the number of small coefficients in ξ:

$$B_W(\xi) = \left\{ k \in \mathbb{Z}_{2N+3} \mid \xi_k < \frac{W}{\sqrt{\epsilon} b} \right\} \tag{4.3}$$

Although it is a trivial observation, it is useful to note that $B_W(\xi) \subseteq B_V(\xi)$ if $V > W$. Since we will want to show that the most probable failure events involve only three ξ coefficients being small, it is useful to establish $P(|B_W(\xi)| \geq 4)$ is extremely small, so that failure events requiring more than three large deviations of the ξ field are negligible. That is the goal of the next lemma.

Lemma 4.1.1. *Choose $1 < \alpha < \sqrt{2}$ and assume Σ satisfies the decay criteria* (1.5) *with R small enough to satisfy the conditions of Lemma 1.2.5 for all $d \leq 3$. Then*

for some positive constant $\bar{C}$*,*

$$P(|B_W(\xi)| \geq 4) \leq \bar{C} \inf_i \left(\frac{\ln(b_W)^5}{(\sqrt{\sigma_{ii}}(\ln(b_W)^2 - \sigma_{ii}))^3} \right) \exp\left(\frac{(3(1-\alpha^2)-1)\ln(b_W)^2}{2\sigma^M} \right) \tag{4.4}$$

where:

$$b_W = \frac{\sqrt{\epsilon} b}{W}$$

Proof. Before beginning the proof, we make a definition for ease of notation.

$$A_{\mathcal{I}} = \{\nu_i > \ln(b_W), \forall i \in \mathcal{I}\}$$

Our first observation is that for a coefficient field ξ to satisfy $|B_W(\xi)| \geq 4$, it is necessary that ξ has at least four small coordinates, so that:

$$\begin{aligned} P(|B_W(\xi)| \geq 4) &\leq \sum_{i \leq j \leq k \leq \ell} P\left(\xi_a < \frac{1}{b_W}\right) \\ &= \sum_{i \leq j \leq k \leq \ell} P\left(\nu_a > \ln(b_W) \text{ for } a \in \{i,j,k,\ell\}\right) \\ &= \sum_{i \leq j \leq k \leq \ell} P\left(\nu_i > \ln\left(b_W\right)\right) P\left(\nu_j > \ln\left(b_W\right) \mid A_i\right) \\ &\qquad P\left(\nu_k > \ln\left(b_W\right) \mid A_{ij}\right) P\left(\nu_\ell > \ln\left(b_W\right) \mid A_{ijk}\right) \end{aligned} \tag{4.5}$$

We must estimate each of the four probabilities in the product. The first term is

straightforward to bound using (1.26):

$$P\left(\nu_i > \ln(b_W)\right) \leq \frac{\sqrt{\sigma_{ii}}\exp\left(\frac{-\ln(b_W)^2}{2\sigma_{ii}}\right)}{\sqrt{2\pi}\ln(b_W)}$$

The remaining three terms all have a similar structure and are handled via Lemma 1.2.5. Provided R is small enough the conditions of the lemma are satisfied for each of the three factors, we can obtain the estimates:

$$\begin{aligned}
&P\left(\nu_s > \ln\left(b_W\right) \mid A_{\mathcal{I}_s}\right) \\
&\leq C_s \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(b_W)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(b_W)^3}\right)^{-1} \frac{\sqrt{\sigma^M}}{\ln(b_W)} \exp\left(\frac{(1-\alpha^2)\ln(b_W)^2}{2\sigma^M}\right)
\end{aligned}$$

for $s \in \{j,k,\ell\}$, $\mathcal{I}_j = \{i\}$, $\mathcal{I}_k = \{i,j\}$, $\mathcal{I}_\ell = \{i,j,k\}$, and constants C_s as specified by the lemma. Substituting our results back into (4.5) implies:

$$\begin{aligned}
P(|B_W(\xi)| \geq 4) \leq N^4 \inf_i &\left(\frac{\sqrt{\sigma_{ii}}}{\ln(b_W)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(b_W)^3}\right)^{-3} \frac{C_j C_k C_\ell (\sigma^M)^2}{\sqrt{2\pi}\ln(b_W)^4} \\
&\exp\left(\frac{(3(1-\alpha^2)-1)\ln(b_W)^2}{2\sigma^M}\right),
\end{aligned}$$

This completes the proof once we set $\bar{C} = \frac{N^4 C_j C_k C_\ell (\sigma^M)^2}{\sqrt{2\pi}}$ and simplify. □

We now define a set of "rare" and "normal" failure events, denoted by $\mathcal{R}_W$ and

$\mathcal{F}_W$, respectively, by:

$$\begin{aligned} \mathcal{R}_W &= \{\xi \mid |B_W(\xi)| \geq 4 \text{ and } \sup_i |\delta u_i| > \epsilon b\} \\ \mathcal{F}_W &= \{\xi \mid |B_W(\xi)| \leq 3 \text{ and } \sup_i |\delta u_i| > \epsilon b\} \end{aligned} \tag{4.6}$$

Our next result allows to obtain an upper bound on how large $||\delta u||_{\ell^2_\epsilon}$ can be on the set $\mathcal{F}_1$.

Lemma 4.1.2. *Suppose $\xi \in \mathcal{F}_1$. Then for any $K > \max(\sqrt{6}, 4||f||_{\ell^2_\epsilon})$ we have the estimate*

$$||\delta u||_{\ell^2_\epsilon} \leq K\epsilon^{\frac{3}{2}} M, \tag{4.7}$$

where $M = \sup_i |\delta u_i|$.

Proof. Assume to the contrary $||\delta u||_{\ell^2_\epsilon} > K\epsilon^{\frac{3}{2}} M$. Because $\xi \in \mathcal{F}_1$, we observe $\xi^1_i \geq \frac{1}{\sqrt{\epsilon}b} \geq \frac{1}{\sqrt{\epsilon}M}$ for all but at most three indices i. We note by considering $u = 0$ that the solution of 1.11 will satisfy $E(u) \leq 0$, and thus for the minimizing u we have:

$$\frac{\epsilon}{2} \sum_{k=1}^{N+1} \xi^1_k \left(\frac{\delta u_k}{\epsilon} \right)^2 + \frac{\epsilon}{2} \sum_{k=1}^{N+2} \xi^2_k \left(\frac{\delta u_k + \delta u_{k-1}}{\epsilon} \right)^2 \leq \epsilon \sum_{i=1}^{N} u_i f_i$$

Applying the Cauchy-Schwarz and Poincare inequalities on the right hand side, we arrive at:

$$\frac{\epsilon}{2} \sum_{k=1}^{N+1} \xi^1_k \left(\frac{\delta u_k}{\epsilon} \right)^2 + \frac{\epsilon}{2} \sum_{k=1}^{N+2} \xi^2_k \left(\frac{\delta u_k + \delta u_{k-1}}{\epsilon} \right)^2 \leq N||f||_{\ell^2_\epsilon}||\delta u||_{\ell^2_\epsilon} \tag{4.8}$$

By dropping the next nearest neighbor terms and performing some rearrangements,

this reduces to:

$$\sum_{B_1} \xi_i^1 \delta u_i^2 + \sum_{B_1^C} \xi_i^1 \delta u_i^2 \leq 2||f||_{\ell_\epsilon^2} \sqrt{\epsilon} \sqrt{\sum_{B_1} \delta u_i^2 + \sum_{B_1^C} \delta u_i^2} \tag{4.9}$$

Since $||\delta u||_{\ell_\epsilon^2} > K\epsilon^{\frac{3}{2}}M$, we know that either $\sum_{B_1} \delta u_i^2 > \frac{K^2}{2}\epsilon^2 M^2$ or $\sum_{B_1^C} \delta u_i^2 > \frac{K^2}{2}\epsilon^2 M^2$. Suppose for the sake of contradiction the latter is true. Then since $|B_1| < 4$ on $\mathcal{F}_1$ and ϵM is the supremum of $|\delta u|$,

$$\sum_{B_1} \delta u_i^2 \leq 3\epsilon^2 M^2 \leq \sum_{B_1^C} \delta u_i^2$$

In the second inequality, we've made use of the fact that $K^2 > 6$.

Next, by dropping the sum over B_1 on the left hand side (it is a strictly positive term) in (4.9), we obtain :

$$\sum_{B_1^C} \xi_i^1 \delta u_i^2 \leq 2\sqrt{2}||f||_{\ell_\epsilon^2} \sqrt{\epsilon} \sqrt{\sum_{B_1^C} \delta u_i^2} \tag{4.10}$$

Recall that for $i \notin B_1$, we can bound ξ_i^1 below:

$$\frac{1}{\sqrt{\epsilon}M} \sum_{B_1^C} \delta u_i^2 \leq 2\sqrt{2}||f||_{\ell_\epsilon^2} \sqrt{\epsilon} \sqrt{\sum_{B_1^C} \delta u_i^2} \tag{4.11}$$

Rearranging the inequality, we finally arrive at:

$$1 \leq 2\sqrt{2}||f||_{\ell_\epsilon^2} \sqrt{\epsilon} \frac{\sqrt{\epsilon}M}{\sqrt{\sum_{B_1^C} \delta u_i^2}} \leq \frac{4}{K}||f||_{\ell_\epsilon^2} \tag{4.12}$$

However, by the definition of K, the right hand side of this inequality is smaller than 1, which provides a contradiction.

Therefore, we conclude $\sum_{B_1} \delta u_i^2 > \frac{K^2}{2}\epsilon^2 M^2 > 3\epsilon^2 M^2$. Since the left hand side contains at most three terms each of which is bounded by $\epsilon^2 M^2$, this is a contradiction as well. Consequently, we must have $||\delta u||_{\ell^2} \leq K\epsilon^{\frac{3}{2}} M$. □

Having completed the above preparations, we are now prepared to prove the analogue of Theorem 2.1.4 for the next nearest neighbor interaction model.

Theorem 4.1.3. *Choose positive constants* $\lambda_1, \lambda_2, \rho_1, \rho_2 < 1, K > \max(\sqrt{6}, 4||f||_{\ell^2_\epsilon})$ *and suppose* $\xi \in \mathcal{F}_W$*, where we define* W *as:*

$$W = \max\left(1, \frac{2K||f||_{\ell^2_\epsilon}}{\rho_1^2(1-\rho_2)^2}, \frac{2K||f||_{\ell^2_\epsilon}}{\rho_1^2\rho_2^2}, \frac{2K||f||_{\ell^2_\epsilon}}{\lambda_1^2(1-\lambda_2)^2}, \frac{2K||f||_{\ell^2_\epsilon}}{\lambda_1^2\lambda_2^2}\right)$$

Further assume $\sup_i |\delta u_i| = |\delta u_k| = \epsilon M > \epsilon b$*. Then we have the following bounds on the* ξ *coefficients:*

$$\xi_k^1 \leq \frac{2K||f||_{\ell^2_\epsilon}}{\sqrt{\epsilon}b} \tag{4.13}$$

$$\xi_k^2 \leq \frac{2K||f||_{\ell^2_\epsilon}}{(1-\lambda_1)^2\sqrt{\epsilon}b} \tag{4.14}$$

$$\xi_{k+1}^2 \leq \frac{2K||f||_{\ell^2_\epsilon}}{(1-\rho_1)^2\sqrt{\epsilon}b} \tag{4.15}$$

Proof. From (4.8) we have the inequality:

$$\sum_{k=1}^{N+1} \xi_k^1 \left(\delta u_k\right)^2 + \sum_{k=1}^{N+2} \xi_k^2 \left(\delta u_k + \delta u_{k-1}\right)^2 \leq 2||f||_{\ell^2_\epsilon}||\delta u||_{\ell^2_\epsilon}$$

Since $\xi \in \mathcal{F}_1$, we can use (4.7) to obtain:

$$\sum_{i=1}^{N+1} \xi_i^1 \left(\delta u_i\right)^2 + \sum_{i=1}^{N+2} \xi_i^2 \left(\delta u_i + \delta u_{i-1}\right)^2 \leq 2K\epsilon^{\frac{3}{2}} ||f||_{\ell_\epsilon^2} M \tag{4.16}$$

By dropping all terms on the LHS of (4.16) except the nearest neighbor term at the break location, we can conclude:

$$\xi_k^1 \delta u_k^2 \leq 2K\epsilon^{\frac{3}{2}} ||f||_{\ell_\epsilon^2} M$$

By rearranging and using the fact $|\delta u_k|^2 > \epsilon^2 M^2$, we reach (4.13):

$$\xi_k^1 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\sqrt{\epsilon} M} \leq \frac{2K||f||_{\ell_\epsilon^2}}{\sqrt{\epsilon} b}$$

Now that we have the desired bound on the nearest neighbor coefficient, we turn our attention to the next nearest neighbor terms. By dropping all of these terms except the one with $i = k$ in (4.16), we arrive at:

$$\xi_k^2 (\delta u_k + \delta u_{k-1})^2 \leq 2K\epsilon^{\frac{3}{2}} ||f||_{\ell_\epsilon^2} M$$

We consider two possible cases, which we shall denote as case L1 and case L2. In the L1 case, we assume $sgn(\delta u_k)\delta u_{k-1} \geq -\lambda_1 \epsilon M$, for some $\lambda_1 < 1$ of our choosing. Then we can bound the $(\delta u_k + \delta u_{k-1})$ term below, yielding:

$$\xi_k^2 (1 - \lambda_1)^2 \epsilon^2 M^2 \leq 2K\epsilon^{\frac{3}{2}} ||f||_{\ell_\epsilon^2} M$$

Therefore, by a simple rearrangement, we get:

$$\xi_k^2 \leq \frac{2K||f||_{\ell_\epsilon^2}}{(1-\lambda_1)^2\sqrt{\epsilon}M} \leq \frac{2K||f||_{\ell_\epsilon^2}}{(1-\lambda_1)^2\sqrt{\epsilon}b}$$

In case L2, we observe $sgn(\delta u_k)\delta u_{k-1} < -\lambda_1\epsilon M$ so the solution u nearly exhibits a break in a second location, $|\delta u_{k-1}| > \lambda_1\epsilon M$. By repeating the analysis of the proof at index $k-1$, we thus obtain:

$$\xi_{k-1}^1 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\lambda_1^2\sqrt{\epsilon}b}$$

In addition, we can further divide case $L2$ into cases $L2a$ and $L2b$ depending on whether $sgn(\delta u_{k-1})\delta u_{k-2} \geq \lambda_2\lambda_1\epsilon M$ or $sgn(\delta u_{k-1})\delta u_{k-2} < \lambda_2\lambda_1\epsilon M$. Analogously to the above arguments, we can then conclude that at least one of the following must hold:

$$\xi_{k-1}^2 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\lambda_1^2(1-\lambda_2)^2\sqrt{\epsilon}b}$$

$$\xi_{k-2}^1 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\lambda_1^2\lambda_2^2\sqrt{\epsilon}b}$$

The important thing to notice here is that case L1 requires only one additional small ξ coefficient, while case $L2$ requires at least 2. We can then return to (4.16) and carry out a similar analysis for the $k+1$ next nearest neighbor coefficient (The only changes required will be to replace indices $k-1$, $k-2$ with indices $k+1, k+2$, λ with ρ, and cases L1, L2 with R1, R2). From this, we obtain in case R1 the bound:

$$\xi_{k+1}^2 \leq \frac{2K||f||_{\ell_\epsilon^2}}{(1-\rho_1)^2\sqrt{\epsilon}b}$$

In case R2, we get both

$$\xi_{k+1}^1 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\rho_1^2\sqrt{\epsilon}b}$$

and at least one of the two inequalities:

$$\xi_{k+2}^2 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\rho_1^2(1-\rho_2)^2\sqrt{\epsilon}b}$$

$$\xi_{k+2}^1 \leq \frac{2K||f||_{\ell_\epsilon^2}}{\rho_1^2\rho_2^2\sqrt{\epsilon}b}$$

Combining results, there are four possible case combinations to consider. With W defined as in the statement of the theorem, we observe that if we have L2+R1, L1+R2, or L2+R2, then $|B_W(\xi)| \geq 4$, which contradicts our assumption that $\xi \in \mathcal{F}_W$. Consequently, we must be in the cases L1 and R1, and so the proof is complete.

□

Some solutions to the energy minimization problem which exhibit a break in the atom chain, along with the corresponding ξ fields, are shown in Figures 4.1 and 4.2. Figure 4.1 is for an external force concentrated on atom 6, while 4.2 uses a uniform external force. Notice that in each case, the ξ^1 and ξ^2 coefficients achieve their smallest values near the break location, in accordance with Theorem 4.1.3.

To see the bounds in Theorem 4.1.3 capture the correct dependence on b, we plotted the relationship between $\frac{1}{b}$ and $\xi_k^1, \xi_k^2, \xi_{k+1}^2$ at the break location across several threshold values. The results are shown in Figure 4.3. The linear relationships suggest the asymptotic dependence on b is correct, but as we shall discuss later when we pursue importance sampling, the prefactors are not optimal for obtaining sharp bounds.

The results proven above allow us to achieve our goal of deriving an asymptotic

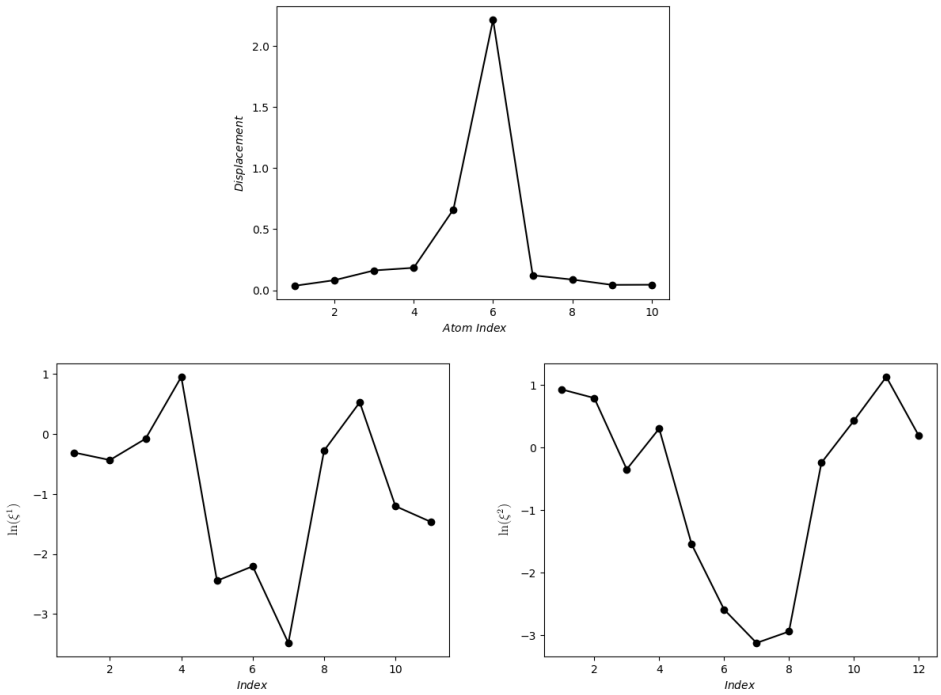

FIGURE 4.1: Solutions to (1.11) exhibiting a break with $b = 20$, $\epsilon = .1$, and $f = \frac{5}{\epsilon}e_6$. The top row depicts the displacement field, while the bottom row shows the log of the ξ coefficients.

estimate for the failure probability when the threshold b is large.

Theorem 4.1.4. *Choose λ_1, λ_2, ρ_1, ρ_2, K, W as in Theorem 4.1.3, and choose $\sqrt{\frac{3}{2}} < \alpha < \sqrt{2}$. Assume Σ satisfies the decay condition with R small enough to satisfy the conditions of Lemma 1.2.5. Then:*

$$p(b) \leq \sum_{k=1}^{N+1} C \inf_i \frac{(\sigma^M)^{\frac{3}{2}}}{2\pi\sqrt{\sigma_{ii}}(\ln(b_W)^2 - \sigma_{ii})} \exp\left(\left(\frac{-1-\alpha^2}{2\sigma^M}\right)\ln(b_W)^2\right)(1+o(1)) \quad (4.17)$$

as $b \to \infty$.

Proof. By the law of total probability,

$$p(b) = P(\mathcal{F}_W) + P(\mathcal{R}_W) \quad (4.18)$$

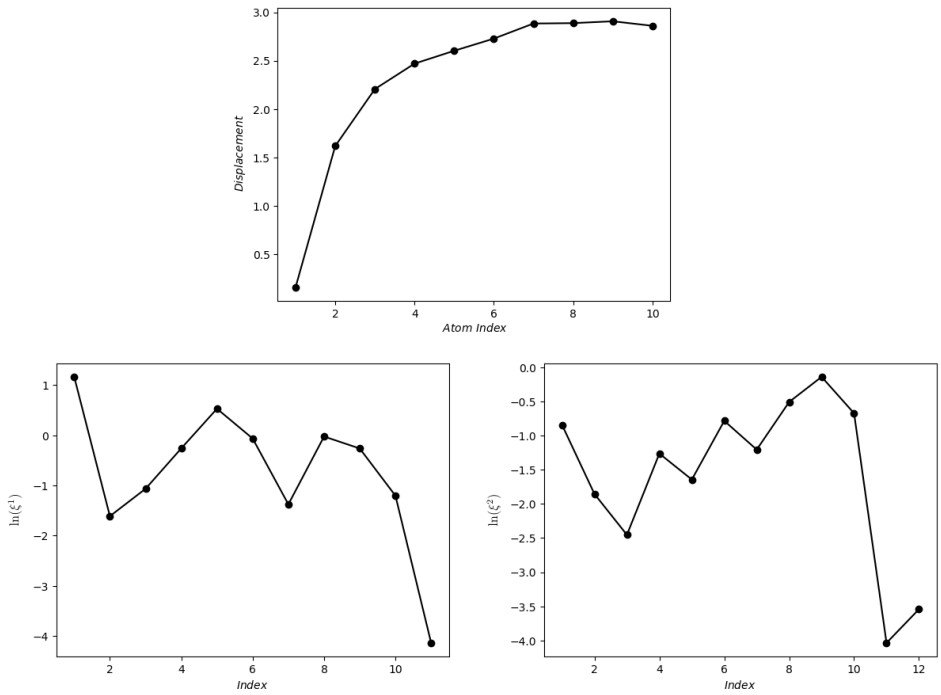

FIGURE 4.2: Solutions to (1.11) exhibiting a break with $b = 20$, $\epsilon = .1$, and $f = 10 \cdot \mathbf{1}$. The top row depicts the displacement field, while the bottom row shows the log of the ξ coefficients. In this example, the break occurs at the right edge of the atom chain, between atom 10 and the "ghost atom" at index 11.

On the set $\mathcal{F}_W$, Theorem 4.1.3 applies, and hence by using subadditivity to consider all possible break locations,

$$\begin{aligned} P(\mathcal{F}_W) &\leq \sum_{k=1}^{N+1} P\left(\xi_k^1 \leq \frac{W}{\sqrt{\epsilon}b} \cap \xi_k^2 \leq \frac{W}{\sqrt{\epsilon}b} \cap \xi_{k+1}^2 \leq \frac{W}{\sqrt{\epsilon}b}\right) \\ &\leq \sum_k P\left(\nu_k^1 \geq \ln(b_W) \cap \nu_k^2 \geq \ln(b_W) \cap \nu_{k+1}^2 \geq \ln(b_W)\right) \\ &= \sum_k P(\nu_k^1 \geq \ln(b_W)) P(\nu_k^2 \geq \ln(b_W)) P(\nu_{k+1}^2 \geq \ln(b_W) \mid \nu_k^2 \geq \ln(b_W)) \end{aligned}$$

In the last line, we used the independence of ν^1 and ν^2 (If one wished to remove the independence assumption, it would simply leave us with more conditional prob-

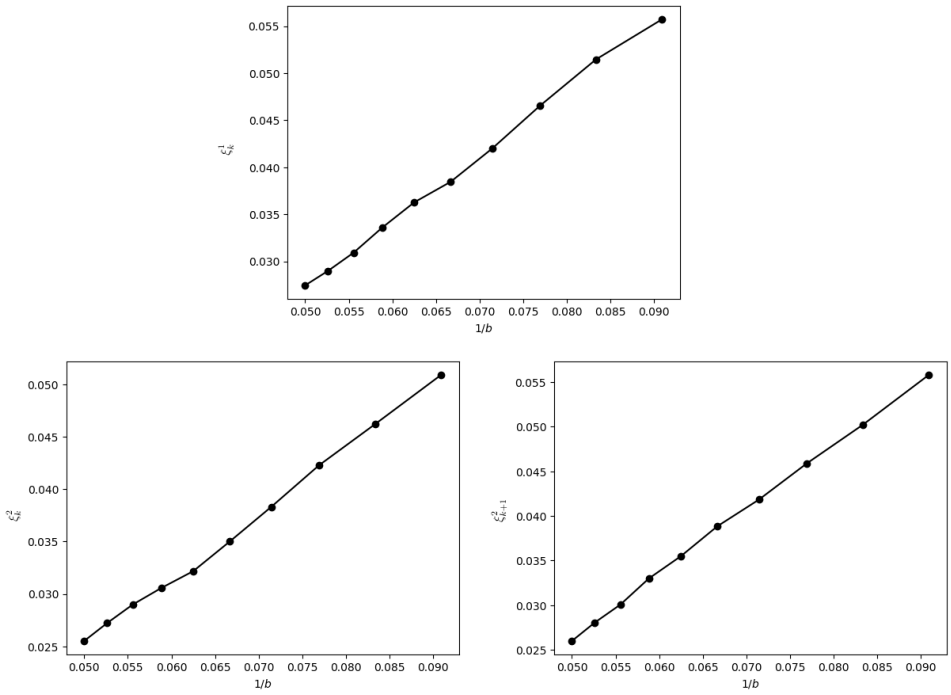

FIGURE 4.3: Comparison of specified ξ coefficients at the break location with $\frac{1}{b}$ for $f = 10 \cdot \mathbf{1}$.

abilities to compute, but would not significantly alter the argument). We have dealt with probabilities of this form several times previously, and by again using (1.26) and (1.29) we obtain:

$$\begin{aligned} P(\mathcal{F}_W) &\leq \sum_{k=1}^{N+1} C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{\ln(b_W)} - \frac{\sqrt{\sigma_{ii}^3}}{\ln(b_W)^3} \right)^{-1} \frac{(\sigma^M)^{\frac{3}{2}}}{2\pi \ln(b_W)^3} \exp\left(\left(\frac{-1-\alpha^2}{2\sigma^M} \right) \ln(b_W)^2 \right) \\ &\leq \sum_{k=1}^{N+1} C \inf_i \frac{(\sigma^M)^{\frac{3}{2}}}{2\pi\sqrt{\sigma_{ii}}(\ln(b_W)^2 - \sigma_{ii})} \exp\left(\left(\frac{-1-\alpha^2}{2\sigma^M} \right) \ln(b_W)^2 \right) \end{aligned} \tag{4.19}$$

On the other hand, by Lemma 4.1.1 and the assumption on α,

$$P(\mathcal{R}_W) \leq P(|B_w| \geq 4) = o\left(\inf_i \frac{1}{(\ln(b_W)^2 - \sigma_{ii})} \exp\left(\left(\frac{-1-\alpha^2}{2\sigma^M}\right)\ln(b_W)^2\right)\right) \tag{4.20}$$

Substitution into (4.18) completes the proof. □

It is worth noting we need the assumption on R since in $\mathcal{R}_{\mathcal{W}}$, the four elements of B_W could all correspond to ξ^1 coefficients, so there is no guarantee we can exploit the independence with ξ^2 the way we do in the calculation over $\mathcal{F}$. Hence the strength of the estimate in Lemma 4.1.1 is dependent on the strength of the correlation of the ξ coefficients, which necessitates the additional assumption that $\alpha > \sqrt{\frac{3}{2}}$. The Gaussian correlation matrix we use, (1.4), will still be able to satisfy (1.5) with a small enough R provided γ is large enough.

An important point to draw from this analysis is that when compared to the nearest neighbor model, failure events are significantly more rare when next nearest neighbor interactions are included. The extra atomistic interactions make the chain more cohesive, requiring more large deviations in the ξ field before a failure event can occur. This makes the need for efficient methods of sampling failure events even more pronounced, and this shall be addressed in the next chapter.

4.2 Explicit Solution

As was the case with the nearest neighbor model, the first step towards solving the energy minimization problem in the next nearest neighbor case is to convert the optimization problem (1.11) into a system of equations. This is accomplished through the following lemma, which is analogous to Lemma 2.2.1.

Lemma 4.2.1. *Suppose u is a solution to* (1.11). *Then for $k = 1, \ldots, N$, u_k satisfies*

the equation:

$$\xi_k^1 \delta u_k - \xi_{k+1}^1 \delta u_{k+1} + \xi_k^2(\delta u_k + \delta u_{k-1}) - \xi_{k+2}^2(\delta u_{k+1} + \delta u_{k+2}) = f_{k,\epsilon} \tag{4.21}$$

Furthermore, we have the boundary conditions $\delta u_1 = u_1$ *and* $\delta u_{N+1} = -u_N$.

Proof. We define a perturbed atom configuration

$$u^\alpha = (u_1, \ldots, u_{k-1}, u_k + \alpha\delta u_{k+1}, u_{k+1}, \ldots, u_N)$$

Further define a function $g(\alpha)$ by:

$$\begin{aligned} g(\alpha) = \mathcal{M}(u^\alpha) =& \frac{1}{2\epsilon}\Bigg(\sum_{j\neq k,k+1} \xi_j \delta u_j^2 + \xi_k(\delta u_k + \alpha\delta u_{k+1})^2 + \xi_{k+1}(1-\alpha)^2 \delta u_{k+1}^2 \\ &+ \sum_{j\neq k,k+1,k+2} \xi_j^2(\delta u_j + \delta u_{j-1})^2 + \xi_k^2(\delta u_k + \delta u_{k-1} + \alpha\delta u_{k+1})^2 \\ &+ \xi_{k+1}^2(\delta u_{k+1} + \delta u_k)^2 + \xi_{k+2}^2(\delta u_{k+2} + (1-\alpha)\delta u_{k+1})^2 \Bigg) \\ &- \epsilon\left(\sum_{j\neq k} f_j u_j + f_k(u_k + \alpha\delta u_{k+1}) \right) \end{aligned}$$

From this, we calculate:

$$\begin{aligned} g'(\alpha) =& \frac{1}{\epsilon}\big(\xi_k^1(\delta u_k + \alpha\delta u_{k+1})\delta u_{k+1} - \xi_{k+1}^1(1-\alpha)\delta u_{k+1}^2 \\ &+ \xi_k^2 \delta u_{k+1}(\delta u_{k-1} + \delta u_k + \alpha\delta u_{k+1}) - \xi_{k+2}^2 \delta u_{k+1}(\delta u_{k+2} + (1-\alpha)\delta u_{k+1}) \big) \\ &- \epsilon f_k \delta u_{k+1} \end{aligned}$$

By assumption g has a minimum at $\alpha = 0$, so we should have $g'(0) = 0$. From this requirement, we obtain:

$$\frac{1}{\epsilon}\left(\xi^1_k \delta u_k \delta u_{k+1} - \xi^1_{k+1}\delta u^2_{k+1} + \xi^2_k \delta u_{k+1}(\delta u_k + \delta u_{k-1}) - \xi^2_{k+2}\delta u_{k+1}(\delta u_{k+1} + \delta u_{k+2})\right)$$

$$= \epsilon f_k \delta u_{k+1}$$

If we divide both sides by δu_{k+1}, we obtain (4.21). The boundary conditions on δu follow immediately from the Dirichlet boundary conditions for u. □

Written in matrix form, the system we must solve is:

$$A\begin{bmatrix}\delta u_1 \\ \vdots \\ \vdots \\ \vdots \\ \delta u_N\end{bmatrix} - \begin{bmatrix}0 \\ \vdots \\ 0 \\ \xi^2_{N+1}\delta u_{N+1} \\ (\xi^1_{N+1} + \xi^2_{N+2})\delta u_{N+1}\end{bmatrix} = \begin{bmatrix}f_{1,\epsilon} \\ \vdots \\ \vdots \\ \vdots \\ f_{N,\epsilon}\end{bmatrix}, \tag{4.22}$$

where the matrix A is defined by:

$$A = \begin{bmatrix}
(\xi^1_1 + \xi^2_1) & -(\xi^1_2 + \xi^2_3) & -\xi^2_3 & 0 & \cdots & \cdots & 0 \\
\xi^2_2 & (\xi^1_2 + \xi^2_2) & -(\xi^1_3 + \xi^2_4) & -\xi^2_4 & \ddots & & \vdots \\
0 & \xi^2_3 & (\xi^1_3 + \xi^2_3) & -(\xi^1_4 + \xi^2_5) & -\xi^2_5 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \xi^2_N \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & -(\xi^1_N + \xi^2_{N+1}) \\
0 & \cdots & \cdots & \cdots & \cdots & \xi^2_N & (\xi^1_N + \xi^2_N)
\end{bmatrix}$$

Due the band structure of the matrix, we can write down the solution to this system of equations.

Theorem 4.2.2. *The solution to the problem (4.22) is given by*

$$\begin{aligned} \delta u_k &= \alpha_k + \beta_k \delta u_{N+1} \text{ for } k = 1, ..N \\ \delta u_{N+1} &= \frac{-\sum_{j=1}^{N+1} \alpha_j}{\sum_{j=1}^{N+1} \beta_j}, \end{aligned} \tag{4.23}$$

where the coefficients α_k, β_k, and γ are defined by

$$\begin{aligned} \alpha_k &= a_k + \frac{b_k}{\gamma}\left(f_{1,\epsilon} - (\xi_1^1 + \xi_1^2)a_1 + (\xi_2^1 + \xi_3^2)a_2 + \xi_3^2 a_3\right) \\ \beta_k &= \frac{b_k}{\gamma}\left(-(\xi_1^1 + \xi_1^2)c_1 + (\xi_2^1 + \xi_3^2)c_2 + \xi_3^2 c_3 + \frac{\gamma}{b_k}c_k\right) \\ \gamma &= b_1(\xi_1^1 + \xi_1^2) - b_2(\xi_2^1 + \xi_3^2) - b_3\xi_3^2 \end{aligned} \tag{4.24}$$

with $\alpha_{N+1} = 0$, $\beta_{N+1} = 1$. The constants a_k, b_k, c_k are chosen to satisfy the recurrence relations

$$\begin{aligned} a_{k-1} &= \frac{1}{\xi_k^2}(f_{k,\epsilon} - (\xi_k^1 + \xi_k^2)a_k + (\xi_{k+1}^1 + \xi_{k+2}^2)a_{k+1} + \xi_{k+2}^2 a_{k+2}) \\ b_{k-1} &= \frac{1}{\xi_k^2}(-(\xi_k^1 + \xi_k^2)b_k + (\xi_{k+1}^1 + \xi_{k+2}^2)b_{k+1} + \xi_{k+2}^2 b_{k+2}) \\ c_{k-1} &= \frac{1}{\xi_k^2}(-(\xi_k^1 + \xi_k^2)c_k + (\xi_{k+1}^1 + \xi_{k+2}^2)c_{k+1} + \xi_{k+2}^2 c_{k+2}), \end{aligned} \tag{4.25}$$

with initial conditions $a_{N+1} = b_{N+1} = 0$, $c_{N+1} = 1$, $a_N = c_N = 0$, $b_N = 1$, $a_{N-1} = \frac{1}{\xi_N^2} f_{N,\epsilon}$, $b_{N-1} = -\frac{1}{\xi_N^2}(\xi_N^1 + \xi_N^2)$, and $c_{N-1} = \frac{1}{\xi_N^2}(\xi_{N+1}^1 + \xi_{N+2}^2)$.

Proof. Suppose for some k and all $j \geq k$, we can express δu_j in the form

$$\delta u_j = a_j + b_j \delta u_N + c_j \delta u_{N+1} \tag{4.26}$$

It is clear that $a_{N+1} = b_{N+1} = 0$, $c_{N+1} = 1$, and $a_N = c_N = 0$, $b_N = 1$. Now consider the $k = N$ equation in (4.21). Rearranging this equation yields:

$$\delta u_{N-1} = \frac{1}{\xi_N^2}(f_{N,\epsilon} - (\xi_N^1 + \xi_N^2)\delta u_N + (\xi_{N+1}^1 + \xi_{N+2}^2)\delta u_{N+1}) \tag{4.27}$$

Hence we can determine $a_{N-1} = \frac{1}{\xi_N^2} f_{N,\epsilon}$, $b_{N-1} = -\frac{1}{\xi_N^2}(\xi_N^1 + \xi_N^2)$, and $c_{N-1} = \frac{1}{\xi_N^2}(\xi_{N+1}^1 + \xi_{N+2}^2)$. Now that we have this information, we can set up a set of recurrence relations for the coefficients a_k, b_k, and c_k. By substituting (4.26) into (4.21) and solving for δu_{k-1}, we obtain the equation:

$$\begin{aligned}\delta u_{k-1} = \frac{1}{\xi_k^2}\Big(& (f_{k,\epsilon} - (\xi_k^1 + \xi_k^2)a_k + (\xi_{k+1}^1 + \xi_{k+2}^2)a_{k+1} + \xi_{k+2}^2 a_{k+2}) \\ & (-(\xi_k^1 + \xi_k^2)b_k + (\xi_{k+1}^1 + \xi_{k+2}^2)b_{k+1} + \xi_{k+2}^2 b_{k+2}) \\ & (-(\xi_k^1 + \xi_k^2)c_k + (\xi_{k+1}^1 + \xi_{k+2}^2)c_{k+1} + \xi_{k+2}^2 c_{k+2})\Big)\end{aligned}$$

This implies the recurrence relations (4.25) for the coefficients a_k, b_k, and c_k.

Next, we must make use of (4.21) in the case of $k = 1$. Substituting (4.26) into this equation allows us to express δu_N in terms of δu_{N+1}:

$$\begin{aligned}\delta u_N = \frac{1}{\gamma}\Big(& f_{1,\epsilon} - (\xi_1^1 + \xi_1^2)a_1 + (\xi_2^1 + \xi_3^2)a_2 \\ & + \xi_3^2 a_3 - ((\xi_1^1 + \xi_1^2)c_1 - (\xi_2^1 + \xi_3^2)c_2 - \xi_3^2 c_3)\delta u_{N+1}\Big)\end{aligned} \tag{4.28}$$

By plugging (4.28) into (4.26), we can then solve for all δu_k in terms of only δu_{N+1},

$$\begin{aligned}\delta u_k =& \left(a_k + \frac{b_k}{\gamma}(f_{1,\epsilon} - (\xi_1^1 + \xi_1^2)a_1 + (\xi_2^1 + \xi_3^2)a_2 + \xi_3^2 a_3)\right) \\ &+ \frac{b_k}{\gamma}\left(-(\xi_1^1 + \xi_1^2)c_1 + (\xi_2^1 + \xi_3^2)c_2 + \xi_3^2 c_3 + \frac{\gamma}{b_k}c_k\right)\delta u_{N+1} \\ :=& \alpha_k + \beta_k \delta u_{N+1}\end{aligned} \tag{4.29}$$

With the convention $\alpha_{N+1} = 0$, $\beta_{N+1} = 1$, this equation trivially holds for δu_{N+1} as well. Also note α_k and β_k are all known provided we have solved (4.25). Next by (1.13), we see that $\delta u_{N+1} = -\sum_{i=1}^{N} \delta u_i$. Substituting (4.29) into this equation, we obtain:

$$\delta u_{N+1} = \frac{-\sum_{i=j}^{N+1} \alpha_j}{\sum_{j=1}^{N+1} \beta_j} \tag{4.30}$$

Once δu_{N+1} is known explicitly, we obtain all δu_k from (4.29). □

While it is not easy to solve the relations (4.25) explicitly, the coefficients can be computed algorithmically on a computer with minimal effort, since the initial conditions for each of these recurrence relations are all known. As before, we immediately obtain a uniqueness result as a corollary.

Corollary 4.2.3. *The solution to* (1.11) *with Dirichlet boundary conditions for the next nearest neighbor harmonic model is unique.*

Proof. Any solution to (1.11) must satisfy (4.22). As the coefficients in (4.25) are all well defined given their initial conditions, we see that for a given pair of coefficient fields ξ^1, ξ^2, there is exactly one atom configuration which solves (4.22). □

In principle, one might hope to extract useful information regarding failure probabilities from the solution provided by Theorem 4.2.2. However, we find this path is not especially fruitful due to the complexity of the result, and in particular the difficulty in finding a closed form for the a, b, c coefficients. Hence, we use the theorem primarily as a computational tool for solving (1.11) during Monte Carlo or importance sampling simulations in the next chapter.

5

Importance Sampling Algorithm for Next Nearest Neighbor Harmonic Model

As we have previously observed, the rapid rate of decay of $p(b)$ as $b \to \infty$ means that failure events for the next nearest neighbor model are exceedingly rare, even more so than for the model with nearest neighbor interactions. As such, the difficulties associated with simulating failure events are amplified in this model. In particular, while we were able to generate an essentially "exact" estimate of $p(b)$ (i.e. an estimate with very small relative error) by running a long Monte Carlo simulation with many trials in the nearest neighbor case, that shall no longer be possible here. While we still attempt to use such a long simulation as a baseline for comparison of our importance sampling method, we shall see that for larger threshold values this is simply not a feasible approach. In fact, the importance sampling procedure can yield better results than Monte Carlo even with a few orders of magnitude fewer trials. This makes the practical importance of the algorithm presented here even more considerable than in the context of chapter 3. However, we shall also see there

are some additional difficulties associated with the algorithm that were not present in the simpler case.

In section 5.1 we will explain the details of the algorithm, with its performance on a few sample problems presented in section 5.2. In section 5.3 we shall very briefly discuss the performance of the importance sampling algorithm and possible improvements.

5.1 Algorithm Overview

As was the case in the nearest neighbor model, the foundation of our importance sampling method is the intuition that $\sup_i |\delta u_i|$ exceeding b is associated with large deviations in the coefficient fields ξ^1 and ξ^2. More specifically, Theorem 4.1.3 tells us that we must have $\nu_k^1 > \ln(c_1 b) := \tau_1$, $\nu_k^2, \nu_{k+1}^2 > \ln(c_2 b) := \tau_2$ for some constants $c_1, c_2 > 0$. Choosing c_1 and c_2 for good algorithm performance is more difficult than it was in the nearest neighbor case, and we shall discuss this in more detail later. For now, we merely note that in practice it is often easiest to set $c_1 = c_2$, as this means we only have one parameter to choose and experimentally the algorithm still performs quite well even with this reduction.

The form of the algorithm is quite similar to what we have previously presented in chapter 3. If P is our original probability distribution and we wish to sample from a proposal distribution Q, we may still use (3.1) to relate the two probability measures. Thus, the only new question we must address is how to choose Q and what is the corresponding density f_Q.

Under P, we assume $\xi = (\xi^1, \xi^2) \in \mathbb{R}^{2N+3}$ has distribution $\mathcal{N}(0, \Sigma_f)$, where $\Sigma_f = \left[\begin{smallmatrix} \Sigma^1 & 0 \\ 0 & \Sigma^2 \end{smallmatrix}\right]$. Σ^1, Σ^2 are matrices of size $N+1$ and $N+2$, respectively, with entries σ_{ij}^1 and σ_{ij}^2 given by (1.4). To generate a sample from Q, we begin by choosing an

index $j \in \{1, 2, \ldots, N+1\}$ uniformly at random. We first generate ν^1 by shifting ν_j by τ and generating the remaining ν_i conditional on the realized ν_j, identically to what was done in chapter 3. We suppose the resulting distribution for ν has distribution $f^Q_{1,j}$. Once this is done, we generate (ν^2_j, ν^2_{j+1}) according to a $\mathcal{N}((\tau, \tau), \Sigma^2_j)$ distribution, where Σ^2_j is the principal submatrix of Σ^2 obtained from deleting all rows and columns except the j and $j+1^{st}$. We then generate ν^2_i for $i \neq j, j+1$ by following their original distribution conditioned on the realized ν^2_j, ν^2_{j+1}. If $f^Q_{2,j}$ is the distribution for the resulting ν_2, then by a calculation analogous to the one done for nearest neighbor interactions we may conclude that:

$$f_Q = \frac{1}{N+1} \sum_{i=1}^{N+1} f^Q_{1,i} f^Q_{2,i} \tag{5.1}$$

Our goal in choosing Q is to ensure that break events are more frequent under this probability measure than under P. To test this, we ran $5 \cdot 10^5$ simulations, generating ξ according to both P and Q and comparing the number of breaks which occurred. The results are displayed in Figures 5.1 and 5.2. From the figures, we can verify that sampling from Q allows us to observe a greater number of failure events.

In summary, with the choices above we have the following importance sampling algorithm:

Algorithm 3: NNN Importance Sampling

1. Choose an index $j \in \{1, 2, 3, \ldots, N+1\}$ with uniform probability.
2. Generate ν^1_j according to the distribution $\mathcal{N}(\tau, \sigma^1_{jj})$
3. Conditional on the realized ν_j, generate the remaining ν^1_i.

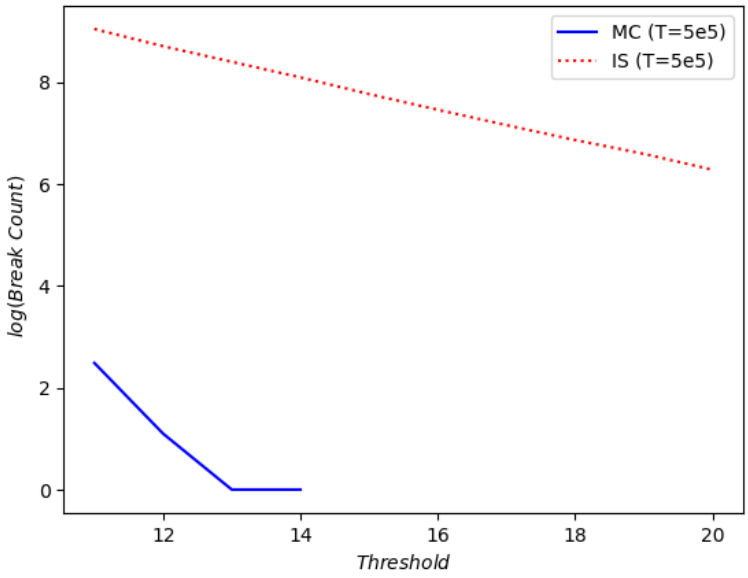

FIGURE 5.1: Comparison of the number of break events for Monte Carlo and importance sampling methods in 500,000 trials, for $f = 10 \cdot \mathbf{1}$. In the IS algorithm, we use $c = 1$.

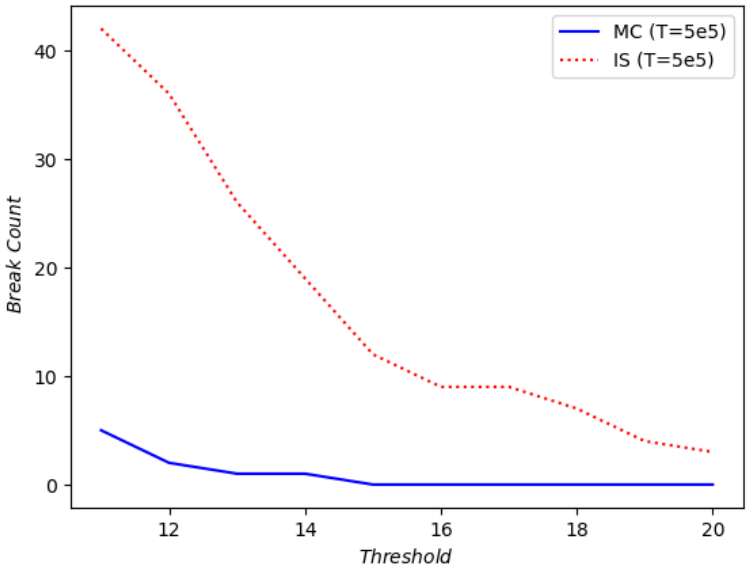

FIGURE 5.2: Comparison of the number of break events for Monte Carlo and importance sampling methods in 500,000 trials, for $f = \frac{5}{\epsilon} e_6$. In the IS algorithm, we use $c = \frac{1}{5}$.

4. Generate ν_j^2, ν_{j+1}^2 according to the distribution $\mathcal{N}((\tau,\tau),\Sigma_j^2)$.

5. Conditional on the realized ν_j^2, ν_{j+1}^2, generate the remaining ν_i^2.

6. Solve for δu given the generated ν and check if $\sup_i |\delta u_i| > \epsilon b$.

7. Repeat steps 1-4 T_{IS} times.

8. Our estimate for $P\left(\sup_i |\delta u_i| > \epsilon b\right)$ is $\frac{1}{T_{IS}} \sum_{k=1}^{T_{IS}} \frac{dP}{dQ}(\nu^{(k)}) \chi\left(\sup_i |\delta u_k^{(i)}| > \epsilon b\right)$.

5.2 Numerical Experiments

We once again consider two test problems for the method, one with $f = \frac{5}{\epsilon} \cdot e_6$ and one with $f = 10 \cdot \mathbf{1}$. In each case, we choose $N = 10$ and consider b values ranging from 10 to 20. The constants 5 and 10 serve to strengthen the force and consequently raise the break probability. This is done to allow us to make comparisons to direct Monte Carlo techniques; MC simulations for weaker forces generally do not produce any break events even when a large number of trials is used. As a baseline estimate of the probability, we generate a value for $p(b)$ using $5 \cdot 10^7$ MC trials. However, it is important to notice that the relative error in this estimate grows quite large for the higher b values. We compare this estimate to $5 \cdot 10^5$ trials of both the MC and IS methods. For the uniform force IS method, we choose the constant $c = 1$, while for the example with concentrated force we choose $c = \frac{1}{5}$.

The results of our simulations for a uniform force are shown in Table 5.1. The IS method outperforms even MC with 100 times as many trials, yielding estimates with lower variance and reaching higher values of b. Figure 5.3 displays the values of $p(b)$ for the three methods. Similarly, the results for the example $f = \frac{5}{\epsilon} e_6$ are displayed

in Table 5.2 and Figure 5.4.

Lastly, using our importance sampling algorithm we can test the decay rate of

Table 5.1: Numerical results for a force $f = 10 \cdot \mathbf{1}$. For MC(5e5) and IS, we use 5e5 trials. The MC(5e7) gives a baseline estimate for comparison, and is generated using a MC simulation with 5e7 trials. A - indicates no estimate was obtained because no failure events were observed during the simulation.

b	-	12	14	16	18	20
$p(b)$	MC(5e5)	6.0e-6	2.0e-6	-	-	-
	IS	7.02e-6	1.20e-6	2.57e-7	5.82e-8	3.76e-9
	MC(5e7)	7.18e-6	1.28e-6	2.4e-7	4.0e-8	-
Std. Dev.	MC(5e5)	3.46e-6	2.0e-6	-	-	-
	IS	3.48e-7	9.00e-8	3.48e-8	1.33e-8	6.50e-7
	MC(5e7)	3.79e-7	1.6e-7	6.93e-8	2.82e-8	-
Rel. Err.	MC(5e5)	.58	1.0	-	-	-
	IS	.05	.07	.14	.23	.27
	MC(5e7)	.05	.125	.29	.70	-

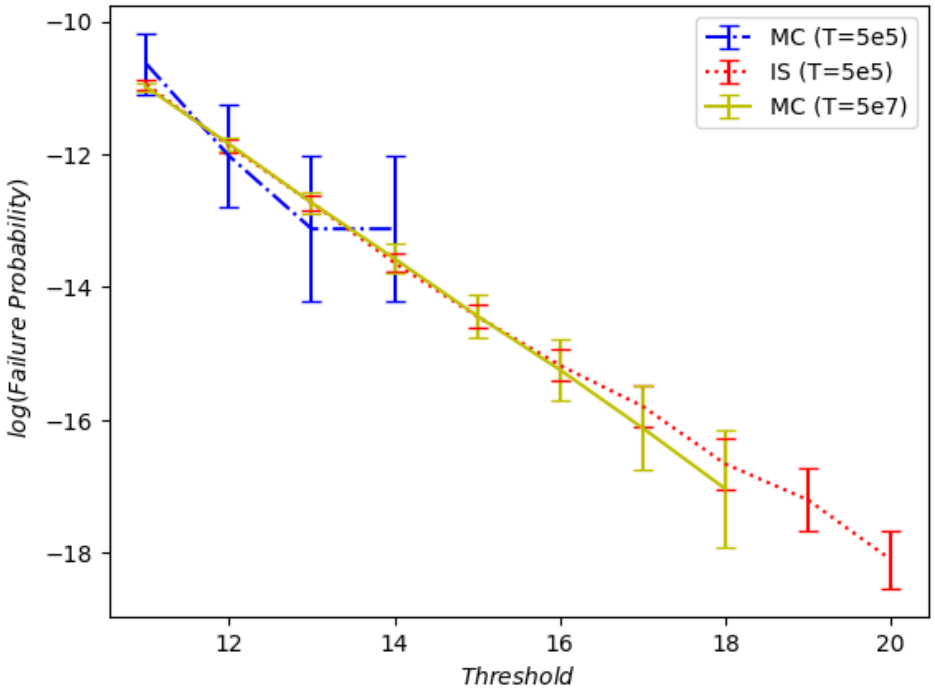

FIGURE 5.3: A comparison of the two MC estimates, and the IS estimate of $p(b)$ for various b values, in the case $f = 10 \cdot \mathbf{1}$.

Table 5.2: Numerical results for a force $f = \frac{5}{\epsilon}e_6$. For MC(5e5) and IS, we use 5e5 trials. The MC(5e7) gives a baseline estimate for comparison, and is generated using a MC simulation with 5e7 trials. A - indicates no estimate was obtained because no failure events were observed during the simulation.

b	-	12	14	16	18	20
$p(b)$	MC(5e5)	4.0e-6	2.0e-6	-	-	-
	IS	4.33e-6	1.07e-6	2.35e-7	7.09e-8	1.86e-8
	MC(5e7)	4.8e-6	1.14e-6	4.0e-7	8.0e-8	4.0e-8
Std. Dev.	MC(5e5)	2.83e-6	2.0e-6	-	-	-
	IS	7.98e-7	2.82e-7	9.02e-8	3.18e-8	1.47e-8
	MC(5e7)	3.10e-7	1.51e-7	8.94e-8	4.0e-8	2.83e-8
Rel. Err.	MC(5e5)	.71	1.0	-	-	-
	IS	.18	.26	.38	.45	.79
	MC(5e7)	.09	.13	.22	.50	.71

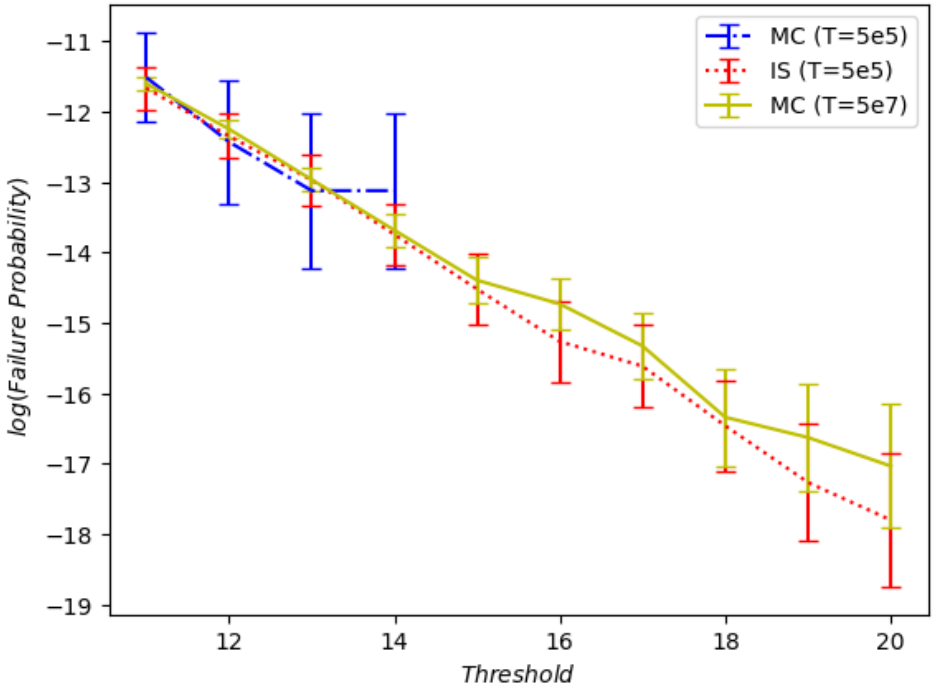

FIGURE 5.4: A comparison of the two MC estimates, and the IS estimate of $p(b)$ for various b values, in the case $f = \frac{5}{\epsilon}e_6$.

$p(b)$ for large b values. While we have an estimate of this rate from Theorem 4.1.4, the constant prefactor in b_W obtained from the variational argument is not sharp.

A similar phenomenon occurred in the nearest neighbor model as well, but there we were able to strengthen the bound using the exact solution, at least in certain cases. While it is not so easy to do so here, we conjecture that Theorem 4.1.4 could be improved to replace b_W by simply b. Under this assumption, the upper bound in Theorem 4.1.4 is compared to the numerical estimation of $p(b)$ for a uniform force in Figure 5.5. We can see that this conjecture appears to capture the correct decay rate.

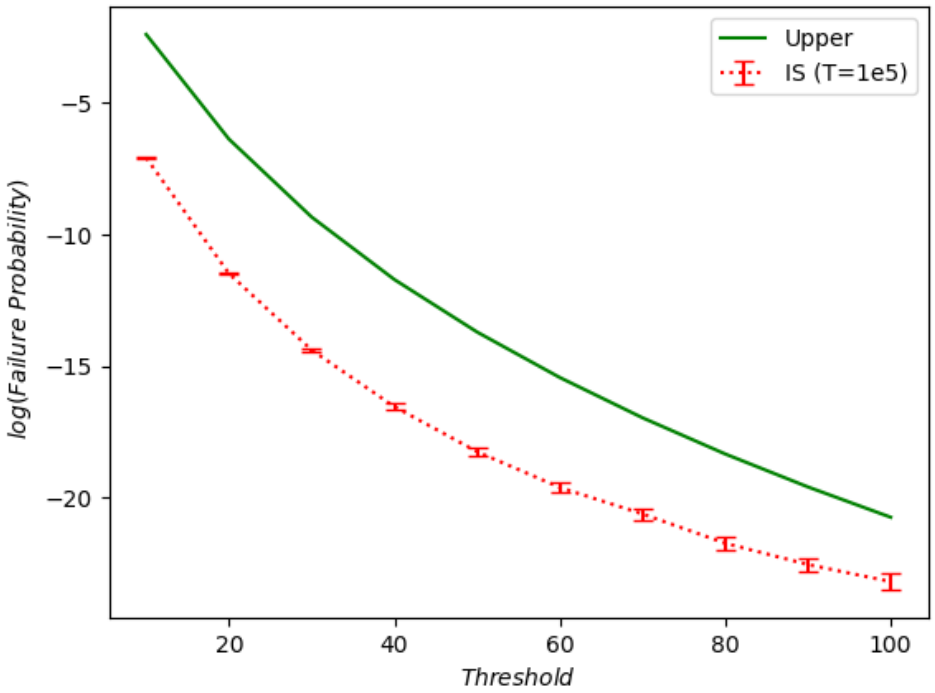

FIGURE 5.5: A comparison of the IS estimate of $p(b)$ to the upper bound of Theorem 4.1.4 with $b_W = b$.

5.3 Efficiency

As we've seen in the previous section, using importance sampling techniques can improve our estimation of $p(b)$ when compared to direct Monte Carlo. However, several of the considerations which were concerns for the importance sampling algorithm

in the case of nearest neighbor interactions, namely that each trial is more expensive than MC and must be run independently for each threshold value, are still relevant in the next nearest neighbor case. To provide another measure of how effective the proposed algorithm is when compared to more standard methods, we compare the relative error achieved by each estimator to the computation time for the simulation in Figures 5.6 and 5.7. The results support the improved efficiency of our algorithm when compared with direct Monte Carlo.

Another important consideration which was not as much of a concern in chapter

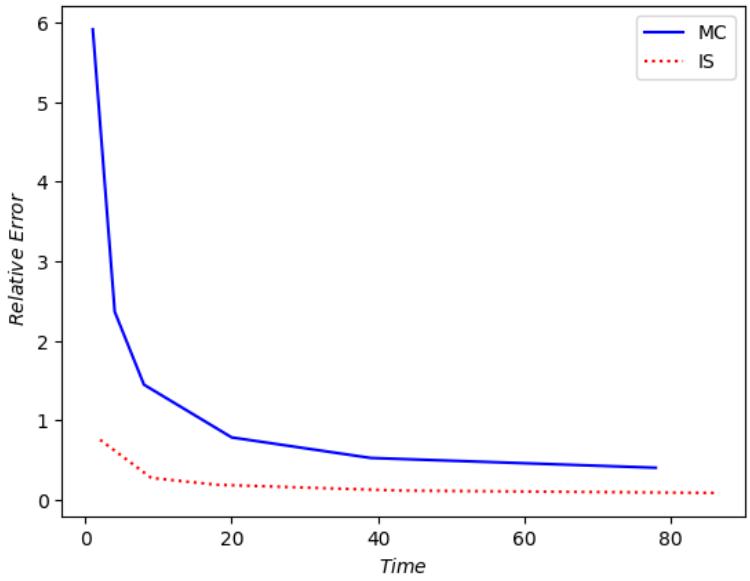

FIGURE 5.6: Amount of accuracy achieved for various computation times (in seconds) using both MC and IS algorithms. Here we take $f = 10 \cdot \mathbf{1}$, $b = 10$.

3 is the need to wisely choose the value of $\tau = \ln(cb)$. While we found the theoretical estimates provided by our analysis of the nearest neighbor model were sufficient for the importance sampling shift in chapter 3, here the constants suggested by Theorem 4.1.3 are too conservative. For example, we compare the results of using the constants in the theorem to the more aggressive choice $\tau = \ln(b)$ in Figure 5.8. Clearly, the more aggressive choice of τ yields considerably better results, even with fewer

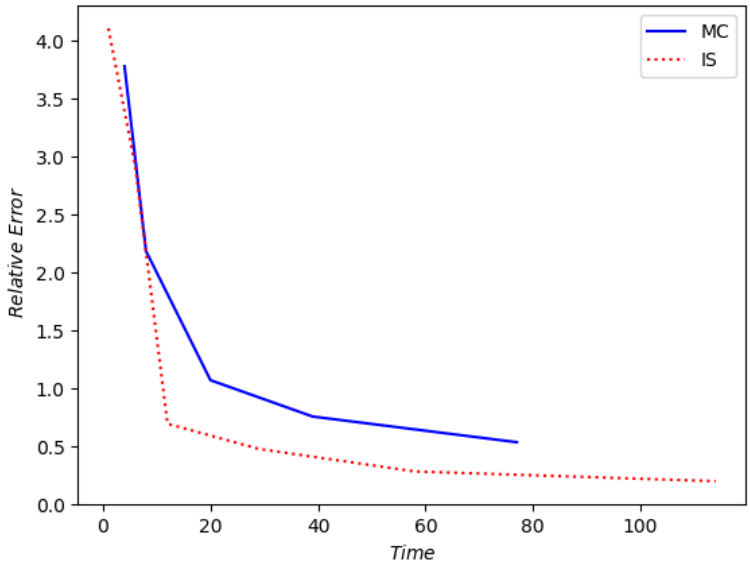

FIGURE 5.7: Amount of accuracy achieved for various computation times (in seconds) using both MC and IS algorithms. Here we take $f = \frac{5}{\epsilon}e_6$, $b = 10$.

trials.

It is evident that choosing an ideal value of τ is critical. Given that failure events are associated with large deviations of ν, choosing an extremely large τ guarantees that we will see a large number of breaks occurring. However, if τ is chosen too large, we may only sample extremely rare failure events while the more common ones lie in the left tail of Q. This can affect the accuracy of the algorithm, as the most important failure events are simply not observed during the simulation due to a finite number of trials. For example, in the case of the localized force, qualitative experiments suggest $\tau = \ln(b)$ produces larger errors than $\tau = \ln\left(\frac{b}{5}\right)$.

Unfortunately, there is no fixed rule for how best to choose τ for a given f. One possible approach is to run several simulations with comparatively few trials (say 10^4) while gradually increasing τ. Once τ is large enough that it is possible to observe failure events, this value could be used in a longer simulation to obtain higher accuracy results. One could also try raising τ further to improve efficiency, but in

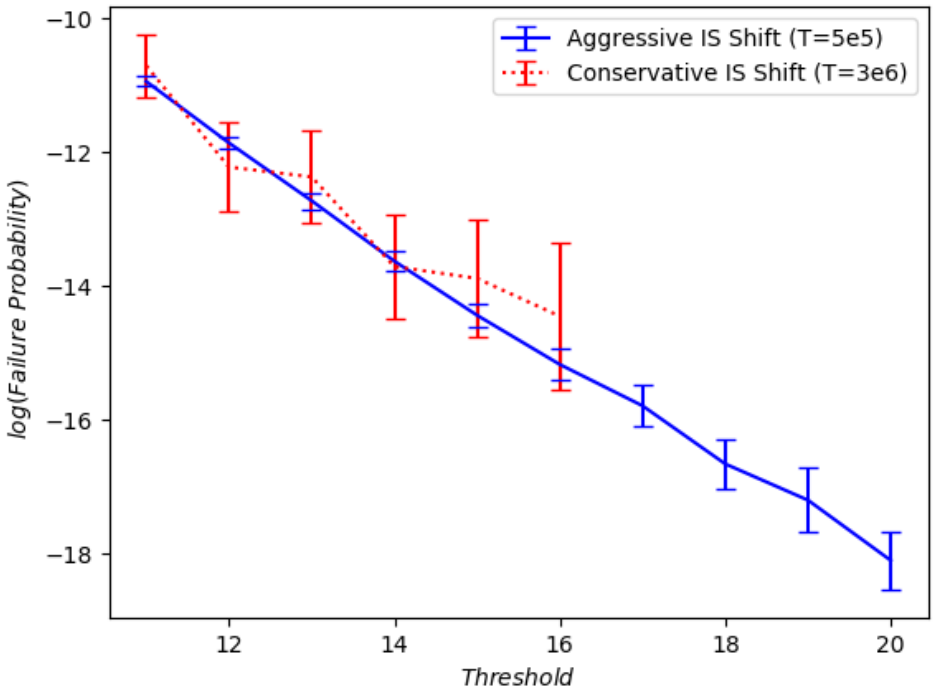

FIGURE 5.8: Comparison of IS methods for a uniform force. For the "aggressive" method, we choose a shift $\tau = \ln(b)$ and perform $5 \cdot 10^5$, while the "conservative" method uses a smaller shift in accordance with the constants in Theorem 4.1.3 and $3 \cdot 10^6$ trials.

general it is better to make a somewhat conservative choice to ensure accuracy of the method.

Lastly, we shall briefly discuss the possibility of improving efficiency by modifying the proposal distribution, Q. Specifically, in step 2 we choose ν_j^1 according to a $\mathcal{N}(0, \sigma_{jj}^1)$ distribution truncated to (τ, ∞) and in step 4 we generate (ν_j^2, ν_{j+1}^2) using a $\mathcal{N}(0, \Sigma_j^2)$ distribution truncated to $[\tau_2, \infty] \times [\tau_2, \infty]$. The only adjustment required to the algorithm is in the computation of the Radon-Nikodym derivative; we shall now find the density under Q is

$$\frac{dP}{dQ} = \frac{1}{N+1} \sum_{i=1}^{N+1} f_{1,i}^Q \tilde{f}_{1,i}^Q f_{2,i}^Q \tilde{f}_{2,i}^Q$$

where $f_{1,i}^{Q}$ is the truncated normal pdf for ν_i^1, $\tilde{f}_{1,i}^{Q}$ is the distribution for the remaining ν_j^1 conditioned on ν_i^1, and $f_{2,i}^{Q}$, $\tilde{f}_{2,i}^{Q}$ play analogous roles for the NNN coefficients.

Figure 5.9 compares the results of the two choices for Q. We can see that using the truncated Gaussian does as well as the shifted one for the smaller b values, but fails for larger values of b. This is again due to difficulty in choosing the cut-off τ used in the truncation. With the shifted Gaussian, we can still potentially observe events in the left tail of Q, so that even if we choose τ slightly too large the method may still perform well. By contrast, the truncated distribution cannot observe events where the coefficients at the break location are less than τ. As a result, these events will be overlooked and accuracy lost if we choose τ too aggressively. This necessitates a more cautious choice of τ than was needed for the shifted Gaussian, leading to the decreased effectiveness of the method for larger b. However, if we had a sharp estimate on the ideal τ to choose, we expect the use of the truncated Gaussian distribution would outperform the shifted one.

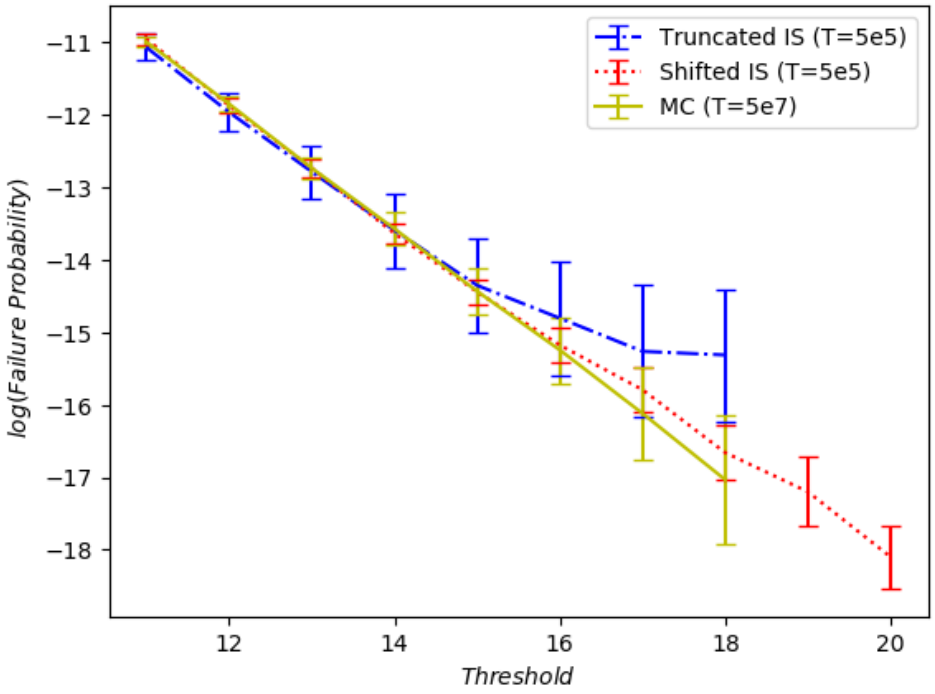

FIGURE 5.9: Comparison of IS methods using shifted and truncated Gaussian distributions for Q.

6

Conclusion

In this thesis we have studied a material failure problem using a discrete, one dimensional atom chain model. We have considered two models of this form, using a harmonic interaction potential and either nearest neighbor or next nearest neighbor interaction ranges. In both cases, we have established a connection between failure events and large deviations in the coefficient field ξ near the break location. Using this connection, we have derived asymptotic upper bounds on the probability of failure and developed more efficient numerical algorithms for calculating $p(b)$, based on importance sampling techniques.

There are many possible extensions and interesting directions for future work on this problem. Firstly, we may wish to improve the bounds on $p(b)$ in the next nearest neighbor case. For example, as per the discussion of Figure 5.5, we expect it is possible to improve the prefactors in b_W. Next, while we have only considered harmonic potentials for atomistic interactions, this is not particularly realistic. A better model would use the Lennard-Jones potential. As can be seen from Figure 1.1, one distinctive feature of this model which contrasts with the harmonic one is

the existence of a concave down region in which the strength of the interaction levels off. This means that the distance between atoms can be increased significantly while paying only a minimal energy penalty. As a result of this feature, we can no longer expect as simple a correlation between large deviations in ξ and large δu values. This presents new challenges in determining how to efficiently sample failure events for such a model.

Another possibility for future work is the extension of our techniques to higher dimensions. In this case, a failure event could be defined as a large gap between atoms along a $d-1$ dimensional separating subspace of the atom lattice. It is expected that importance sampling techniques could again provide a framework for more efficient computation of failure probabilities in this context.

Appendix A

Proof of Lemma 1.2.5

We shall now present the proof of Lemma 1.2.5. First, we introduce some notations. For $0 < R < 1$, $d \in \mathbb{N}$, let

$$g(R,d) = \left(1 - \frac{2R\left(1 - R^{\lfloor \frac{d}{2} \rfloor}\right)}{1-R}\right)$$

Next, for a subset $\mathcal{I}$ of $\mathbb{N}$ and an index $j \in \mathbb{N}$, define

$$Q_{\mathcal{I}j}(R,d) = \frac{g(R,d)\inf_{i\in\mathcal{I}}\sigma_{ii}}{R\sigma_{jj}\sqrt{1+R^2+R^4+\cdots+R^{2(d-1)}}}$$

Lemma 1.2.5. *Let $\nu \sim \mathcal{N}(0,\Sigma)$ be a Gaussian random variable on $\mathbb{R}^n$, and suppose $\sigma^M = \sup_i \sigma_{ii}$. Let $j \in \mathbb{N}_{\leq n}$ and $\mathcal{I} \subset \mathbb{N}_{\leq n}$ with $|\mathcal{I}| = d \geq 1$ and $j > \max(\mathcal{I})$. Choose*

$1 < \alpha < \sqrt{2}$ *and pick* R *small enough that* $g(R, d) > 0$ *and* $\beta^* Q_{\mathcal{I}j} > \alpha d$*, where*

$$\beta^* = \frac{1 - \sqrt{2d^{-2}Q_{\mathcal{I}j}^2 - 1}}{1 - d^{-2}Q_{\mathcal{I}j}^2}$$

Assume Σ *satisfies* (1.5) *with constant* R*. Then for any positive constant* $c_1 \in \mathbb{R}$*,* $c_2 \in \mathbb{R}^d$ *with all positive entries and* $c_{2,i} \leq c_1$*, there exists a constant* $C > 0$ *such that:*

$$\begin{aligned} &P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) \\ &\leq C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3} \right)^{-1} \frac{\sqrt{\sigma^M}}{c_1} \exp\left(\frac{(1-\alpha^2)c_1^2}{2\sigma^M} \right) \end{aligned}$$

Proof. Let $\nu_{\mathcal{I}} = (\nu_{i_1}\nu_{i_2}, \ldots, \nu_{i_d})$ with $i_1 < i_2 < \cdots < i_d$. By the law of total probability,

$$P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) = \int_{y > c_2} P(\nu_j > c_1 \mid \nu_{\mathcal{I}} = y)\phi_c(y)dy, \tag{A.1}$$

where $\phi_c(y)$ denotes the pdf of $\nu_{\mathcal{I}}$ conditioned on $\nu_{\mathcal{I}} > c_2$ (i.e., a truncated normal pdf). Let $\tilde{\mu} = \sigma_{j\mathcal{I}}^T S_{\mathcal{I}}^{-1} y$ and $\tilde{\sigma} = \sigma_{jj} - \sigma_{j\mathcal{I}}^T S_{\mathcal{I}}^{-1} \sigma_{j\mathcal{I}}$ be the conditional mean and variance of ν_j given $\nu_{\mathcal{I}} = y$, where

$$\sigma_{j\mathcal{I}} = \begin{bmatrix} \sigma_{ji_1} \\ \vdots \\ \sigma_{ji_d} \end{bmatrix}$$

and S is the principle submatrix of Σ obtained by retaining only the row and columns

corresponding to indices in $\mathcal{I}$,

$$S_{\mathcal{I}} = \begin{bmatrix} \sigma_{i_1 i_1} & \sigma_{i_1 i_2} & \cdots & \sigma_{i_1 i_d} \\ \sigma_{i_2 i_1} & \sigma_{i_2 i_2} & \cdots & \sigma_{i_2 i_d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i_d i_1} & \sigma_{i_d i_2} & \cdots & \sigma_{i_d i_d} \end{bmatrix}$$

Choose $\beta^* \in (0,1)$ as specified and define $c_3 = ||\sigma_{j\mathcal{I}}||_{\ell^2}^{-1} ||S_{\mathcal{I}}^{-1}||_{\ell^2}^{-1} \beta^* c_1$. We must obtain an estimate on the size of c_3 before proceeding. If λ_m is the smallest eigenvalue of $S_{\mathcal{I}}$, then by the Gershgorin Circle theorem and the decay hypothesis on Σ,

$$|\lambda_m - \sigma_{ii}| \leq \sum_{k \neq i} |\sigma_{ik}| \leq 2\sigma_{ii} \sum_{k=1}^{\lfloor \frac{d}{2} \rfloor} R^k \leq 2\sigma_{ii} R \frac{1 - R^{\lfloor \frac{d}{2} \rfloor}}{1 - R},$$

for some $i \in \mathcal{I}$. Consequently, we arrive at the inequality

$$\lambda_m \geq \inf_{i \in \mathcal{I}} \sigma_{ii} \left(1 - \frac{2R\left(1 - R^{\lfloor \frac{d}{2} \rfloor}\right)}{1 - R} \right) = \inf_{i \in \mathcal{I}} \sigma_{ii} g(R, d)$$

We note that as long as R is sufficiently small, this gives a positive lower bound for λ_m $\big($For the purposes of this thesis, we'll have $d \leq 3$ and so $R < \frac{1}{2}$ suffices here$\big)$. Using this bound, we estimate the norm of $S_{\mathcal{I}}$ by:

$$||S_{\mathcal{I}}^{-1}||_{\ell^2} = \frac{1}{\lambda_m} \leq \frac{1}{g(R,d) \inf_{i \in \mathcal{I}} \sigma_{ii}}$$

In addition, since $i_1 < i_2 < \cdots < j$, the decay hypothesis implies:

$$||\sigma_{j\mathcal{I}}||_{\ell^2} = \sqrt{\sigma_{ji_1}^2 + \cdots + \sigma_{ji_d}} \leq R\sigma_{jj}\sqrt{1 + R^2 + R^4 + \cdots + R^{2(d-1)}}$$

Combining our results, we may conclude:

$$c_3 \geq \frac{g(R,d)\inf_{i\in\mathcal{I}}\sigma_{ii}}{R\sigma_{jj}\sqrt{1 + R^2 + R^4 + \cdots + R^{2(d-1)}}}\beta^* c_1 = Q_{\mathcal{I}j}(R,d)\beta^* c_1 \tag{A.2}$$

Take R small enough so that $\beta^* Q_{\mathcal{I}j} > \alpha d$ (It is easy to check this is possible as $\beta^* Q_{\mathcal{I}j} \to \sqrt{2}d$ as $R \to 0$). This ensures $\frac{c_3}{d} > \alpha c_1 > \alpha \sup_i c_{2,i}$.

Now we split the integral (A.1) into two pieces, based on whether $\tilde{\mu}$ exceeds $\beta^* c_1$ or not.

$$\begin{aligned} P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) \leq & \int_{||y||_{\ell^2} > c_3} P(\nu_j > c_1, \nu_{\mathcal{I}} = y)\phi_c(y)dy \\ & + \int_{||y||_{\ell^2} \leq c_3, y > c_2} P(\nu_j > c_1, \nu_{\mathcal{I}} = y)\phi_c(y)dy \end{aligned} \tag{A.3}$$

We focus our attention on the first integral in (A.3). Clearly,

$$\begin{aligned} \int_{||y||_{\ell^2} > c_3} P(\nu_j > c_1, \nu_{\mathcal{I}} = y)\phi_c(y)dy &\leq \int_{||y||_{\ell^2} > c_3} \phi_c(y)dy \\ &= P(||\nu_{\mathcal{I}}||_{\ell^2} > c_3 \mid \nu_{\mathcal{I}} > c_2) \end{aligned}$$

By making use of (1.28) followed by (1.26), this probability can be estimated as:

$$\int_{||y||_{\ell^2} > c_3} P(\nu_j > c_1, \nu_{\mathcal{I}} = y) \leq G(0, \Sigma, c_2) \sum_{u\in\mathcal{I}} P\left(\nu_i > \frac{c_3}{d}\right) \tag{A.4}$$

$$\leq G \sum_{i \in \mathcal{I}} \frac{d\sqrt{\sigma_{ii}}}{\sqrt{2\pi} c_3} \exp\left(\frac{-c_3^2}{2d^2\sigma_{ii}}\right)$$

In the definition of G, for any index k we can write:

$$\int_{y > c_2} \phi(y, 0, \Sigma) dy = P(\nu_k > c_{2,k}) P(\nu_\ell > c_\ell, \ell \neq k \mid \nu_k > c_{2,k})$$

Therefore, by choosing k as the index which maximizes the numerator of G and recalling $\frac{c_3}{d} > \sup_i c_{2,i}$, G can be bounded by using the lower bound in (1.26):

$$G \leq \frac{1}{P(\nu_k > c_{2,k})} \leq \sqrt{2\pi} \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3}\right)^{-1} \exp\left(\frac{c_{2,i}^2}{2\sigma_{ii}}\right)$$

Therefore, we can obtain:

$$\begin{aligned}
\int_{||y||_{\ell^2} > c_3} P(\nu_j > c_1, \nu_{\mathcal{I}} = y) \leq & \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3}\right)^{-1} \exp\left(\frac{c_{2,i}^2}{2\sigma_{ii}}\right) \\
& \sum_{i \in \mathcal{I}} \frac{d\sqrt{\sigma_{ii}}}{c_3} \exp\left(\frac{-c_3^2}{2d^2\sigma_{ii}}\right) \\
\leq & \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3}\right)^{-1} \\
& \sum_{i \in \mathcal{I}} \frac{d\sqrt{\sigma_{ii}}}{c_3} \exp\left(\frac{(d^2 - Q_{\mathcal{I}j}^2(\beta^*)^2) c_1^2}{2d^2\sigma_{ii}}\right)
\end{aligned} \tag{A.5}$$

We now turn our attention to the second integral in (A.3). Focusing on the first part

of the integrand, by (1.26) we conclude:

$$P(\nu_j > c_1 \mid \nu_i = y) \leq \frac{\sqrt{\tilde{\sigma}}}{\sqrt{2\pi}(c_1 - \tilde{\mu})} \exp\left(\frac{-(c_1 - \tilde{\mu})^2}{2\tilde{\sigma}}\right)$$

However, recall $\tilde{\mu} < \beta^* c_1$ in this integral, so we can get the inequality:

$$P(\nu_j > c_1 \mid \nu_i = y) \leq \frac{\sqrt{\tilde{\sigma}}}{\sqrt{2\pi}(1 - \beta^*)c_1} \exp\left(\frac{-((1 - \beta^*)c_1)^2}{2\tilde{\sigma}}\right)$$

Note this is a uniform bound independent of the value of y, so the full integral can be calculated to be:

$$\begin{aligned}
&\int_{||y||_{\ell^2} \leq c_3, y > c_2} P(\nu_j > c_1, \nu_{\mathcal{I}} = y)\phi_c(y)dy \\
&\leq \frac{\sqrt{\tilde{\sigma}}}{\sqrt{2\pi}(1 - \beta^*)c_1} \exp\left(\frac{-((1 - \beta^*)c_1)^2}{2\tilde{\sigma}}\right) \int_{||y||_{\ell^2} \leq c_3, y > c_2} \phi_c(y)dy \qquad \text{(A.6)} \\
&\leq \frac{\sqrt{\tilde{\sigma}}}{\sqrt{2\pi}(1 - \beta^*)c_1} \exp\left(\frac{-((1 - \beta^*)c_1)^2}{2\tilde{\sigma}}\right),
\end{aligned}$$

where we have used that the integral of the pdf ϕ_c cannot exceed 1.

Substituting (A.5) and (A.6) into (A.3) now yields

$$\begin{aligned}
&P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) \\
&\leq \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3}\right)^{-1} \sum_{i \in \mathcal{I}} \frac{d\sqrt{\sigma_{ii}}}{c_3} \exp\left(\frac{(d^2 - Q_{\mathcal{I}j}^2(\beta^*)^2)c_1^2}{2d^2\sigma^M}\right) \\
&+ \frac{\sqrt{\tilde{\sigma}}}{\sqrt{2\pi}(1 - \beta^*)c_1} \exp\left(\frac{-((1 - \beta^*)c_1)^2}{2\sigma^M}\right),
\end{aligned}$$

where $\sigma^M = \sup_{i \leq n} \sigma_{ii}$. β^* was chosen so that the two exponential factors are now equal, and we obtain:

$$
\begin{aligned}
& P(\nu_j > c_1 \mid \nu_{\mathcal{I}} > c_2) \\
& \leq C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3} \right)^{-1} \frac{\sqrt{\sigma^M}}{c_1} \exp\left(\frac{(d^2 - Q_{\mathcal{I}j}^2(\beta^*)^2)c_1^2}{2d^2\sigma^M} \right) \\
& \leq C \inf_i \left(\frac{\sqrt{\sigma_{ii}}}{c_{2,i}} - \frac{\sqrt{\sigma_{ii}^3}}{c_{2,i}^3} \right)^{-1} \frac{\sqrt{\sigma^M}}{c_1} \exp\left(\frac{(1-\alpha^2)c_1^2}{2\sigma^M} \right),
\end{aligned}
$$

with $C = \max\left(\frac{1}{\alpha}, \frac{1}{\sqrt{2\pi}(1-\beta^*)}\right)$ □

Appendix B

List of Notations

δ – Finite difference operator, defined page 10.

$\mathcal{N}(\mu, \Sigma)$ – Normal distribution with mean μ, covariance Σ.

ϕ – Probability density function for the normal distribution.

Φ – Cumulative density function for the normal distribution.

G – Defined page 14.

P – Probability measure.

Q – Importance sampling measure.

$d\hat{x}_i$ – Denotes integration over all components of x except x_i.

N – Number of atoms.

ϵ – Default spacing between atoms.

V – Atomisitc interaction potential.

ξ^1 – Vector of nearest neighbor interaction coefficients, defined page 7.

ξ^2 – Vector of next nearest neighbor interaction coefficients, defined page 7.

ν^i – When ξ^i is log-normal, this is the corresponding normal random variable.

σ_{ij} – $(i, j)^{th}$ entry of the covariance matrix Σ.

u – Vector of atom displacements.

E – Potential energy from atom interactions, defined page 9.

f – External force applied to atom chain.

$\mathcal{M}$ – Total energy in atom chain, defined page 10.

b – Threshold for material failure.

M – Maximum value attained by $|\delta u_k|$.

$p(b)$ – Failure probability for threshold level b, defined page 11.

$p_k(b)$ – Localized failure probability for threshold level b, defined page 11.

$\mathcal{F}$ – Set of common failure events, defined pages 24 and 63.

$\mathcal{R}$ – Set of rare failure events, defined pages 24 and 63.

$\tilde{}$ – Used to denote conditional quantities (mean, variance, etc.).

$f_{k,\epsilon}$ – Rescaled force, defined page 32.

$\mathbf{1}$ – A vector of ones.

e_i – The i^{th} standard basis vector on $\mathbb{R}^n$.

$\chi(A)$ – Characteristic function of the set A.

F_k – Defined page 36.

S_b – Defined page 38.

B_W – Defined page 60.

b_W – Rescaled threshold value. Defined page 61.

T – Number of trials for Monte Carlo or importance sampling algorithms.

References

[1] Zdenek Bazant and Luigi Cedolin, *Stability of Structures*. World Scientific, 2010.

[2] Jose Blanchet, Jingchen Liu, and Peter Glynn, *State-Dependent Importance Sampling and Large Deviations*. valuetools '06 Proc. 1^{st} Int. Conf. on Perf. Eval. Meth. and Tools, New York, ACM, 20, 2006.

[3] Abhijit Dasgupta and Michael Pecht, *Material Failure Mechanisms and Damage Models*. IEEE Trans. Rel. **40**(1991), No. 5, 531-536

[4] O. Ditlevson and H.O. Madsen, *Structural Reliability Methods*. Wiley, 1996.

[5] Weinan E, Jianfeng Lu, and Jerry Zhang, *Uniform Accuracy of the Quasicontinuum Method*. Phys. Rev. B, **74**(2006) Iss. 21

[6] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden, *Minimum Action Method for the Study of Rare Events*. Comm. Pure and App. Math, **57**, 2004, Iss. 5, 637-656

[7] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden, *Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier Crossing Events*. Journ. Chem. Phys., **126**, 2007, Iss. 16

[8] Weinan E, Eric Vanden-Eijnden, *Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events*. Ann. Rev. Phys. Chem., **61**, 2010, 391-420

[9] S. Engelund, and R. Rackwitz, *A Benchmark Study on Importance Sampling Techniques in Structural Reliability*. Structural Safety, **12**(1993) Iss. 4, 255-276

[10] Mark Freidlin and Alexander Wentzell, *Random Perturbations of Dynamical Systems*. Springer-Verlag, Berlin Heidelberg, 2012.

[11] Paul Glasserman, *Monte Carlo methods in Financial Engineering*. Springer-Verlag, New York, 2003.

[12] William C. Horrace, *Some Results on the Multivariate Truncated Normal Distribution*. Journ. Multivar. Anal. **94**(2015), Iss. 1, 209-221

[13] Marwan Harajli, Tyrrell Rockafellar, and Johannes Royset, *Importance Sampling in the Evaluation and Optimization of Buffered Failure Probability*. 12^{th} Int. Conf. on App. of Stat. and Prob. in Civil Eng., Vancouver, 2015.

[14] Xiaoou Li, Jingchen Liu, Jianfeng Lu, and Xiang Zhou. *Moderate Deviation for Random Elliptic PDE with Small Noise*. Ann. App. Prob. **28**(2018), No. 5, 2781-2813

[15] Xiantao Li and Pingbing Ming, *A Study on the Quasicontinuum Approximations of a One-Dimensional Fracture Model*. Multi-Scale Model. and Sim. **12**, 1379-1400

[16] Florian Lipsmeier and Ellen Baake, *Rare Event Simulation for T-cell Activation*. Journ. Stat. Phys., **134**(2009) Iss. 3, 537-566

[17] Jingchen Liu, Jianfeng Lu, and Xiang Zhou, *Efficient Rare Event Simulation for Failure Problems in Random Media*. SIAM Journ. Sci. Comput. **37**(2015), Iss. 2, A609-A624

[18] Jingchen Liu and Xiang Zhou, *On the Failure Probability of One Dimensional Random Material Under Delta External Force*. Commun. Math Sci. **11**(2013), No. 2, 499-521

[19] Jingchen Liu, Xiang Zhou, Rohit Patra, and Weinan E, *Failure of Random Materials: A Large Deviation and Computational Study*. Proc. 2011 Wint. Sim. Conf. (2011)

[20] Mitchell Luskin and Christoph Ortner, *Atomistic to Continuum Coupling*. Acta Numerica, **22**(2013), 397-508

[21] Jorge Nocedal and Stephen Wright, *Numerical Optimization*. Springer, New York, 1999.

[22] R.T. Rockafellar and J. Royset, *On Buffered Failure Probability in Design and Optimization of Structures*. Rel. Eng. and Sys. Safe. **95** (2010), Iss. 5, 499-510

[23] John Shortle and Pierre L'Ecuyer, *Introduction to Rare Event Simulation*. Wiley Encyclopedia of Operations Research and Management Science, 2011.

[24] R.J. Torrent, *The Log-Normal Distribution: A Better Fitness for the Results of Mechanical Testing of Materials*. Mat. Constr. **11** (1978), Iss. 4, 235-245

[25] Elisabeth Ullmann and Iason Papaioannou, *Multilevel Estimation of Rare Events*. SIAM/ASA Journ. Uncert. Quant. **3** (2015), Iss. 1, 922-953

[26] Maddalena Venturoli, Eric Vanden-Eijnden, and Giovanni Ciccotti, *Kinetics of Phase Transitions in Two Dimensional Ising Models Studied with the String Method*. Journ. Math. Chem. **45** (2009), Iss. 1, 188-222

[27] Xiaoliang Wan and Xiang Zhou, *Asymptotically Efficient Simulation of Elliptic Problems with Small Random Forcing*. SIAM Journ. Sci. Comput. **40**(2018), No. 1, A548-A572