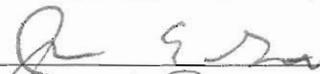


OPTIMAL SAMPLING OF RANDOM FIELDS
FOR TOPOLOGICAL ANALYSIS

by

Gregory Scott Cochran
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Mathematics

Committee:

	Dr. Thomas Wanner, Dissertation Director
	Dr. David Singman, Committee Member
	Dr. Daniel Anderson, Committee Member
	Dr. James Gentle, Committee Member
	Dr. Stephen Saperstone, Department Chair
	Dr. Timothy Born, Associate Dean, Student and Academic Affairs
	Dr. Vikas Chandhoke, Dean, College of Science

Date: 07/26/2011

Summer Semester 2011
George Mason University
Fairfax, VA

Optimal Sampling of Random Fields for Topological Analysis

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Gregory Scott Cochran
Bachelor of Science
Virginia Polytechnic Institute and State University, 2006

Director: Thomas Wanner, Professor
Department of Mathematical Sciences

Summer Semester 2011
George Mason University
Fairfax, VA

Copyright © 2011 by Gregory Scott Cochran
All Rights Reserved

Dedication

I dedicate this dissertation to my dad.

Acknowledgments

I would first like to thank my advisor Dr. Thomas Wanner for his patience. I also want to give special thanks for my committee members. I would like to thank my Texas Family - Mom, Dusty, Rebecca, Ma, Papa, Tyler, Callie and my Virginia Family - Janey, Rob, Kate, and Robert.

In addition, I want to thank Robert, Javed, Keith, Lars, Andy, Ester, and the new people in 242. I would also like to give very special thanks to Christine.

Lastly I want to thank Catherine, without whom I could not have written this thesis. I love you.

Table of Contents

	Page
List of Tables	vii
List of Figures	viii
Abstract	x
1 Introduction	1
1.1 Motivation	1
1.2 Cubical Homology	3
1.3 Computational Homology	17
1.4 Random Fields and Analysis	18
1.4.1 Orthogonal Expansions of Random Fields	26
1.5 Cahn-Hilliard Simulations	29
1.6 Previous Work	33
2 Guaranteed Homology Computations	36
2.1 Overview	36
2.2 Previous Algorithm	36
2.3 A Randomized Subdivision Process	41
2.3.1 Range Enclosure	41
2.3.2 Randomized Subdivisions	46
2.4 Numerical Experiments	49
2.4.1 Double Well Function	49
3 Previous Probabilistic Bounds	53
3.1 Introduction	53
3.2 Homogenous Bounds in One Dimension	57
3.3 Homogenous Random Fields in Two-Dimensions	63
3.4 Non-Homogenous Bounds	71
4 Non-Homogenous Bounds	80
4.1 Introduction	80
4.2 Three Point Sign Configurations	87
4.3 Four Point Square	91
4.4 Four Point Skewed Square	95

4.5	Five Point Sign Configuration	99
4.6	Admissibility of Rectangles	103
5	Minimization Algorithm	113
5.1	Introduction	113
5.2	Algorithm in One Dimension	113
5.3	Two Dimensional Algorithm	117
6	Computational Results	124
6.1	Introduction	124
6.2	Random Chebyshev Series	124
6.3	Random Cosine Series	128
7	Conclusion	131
7.1	Conclusion	131
7.2	Further Directions	133
A	Appendix	134
B	Appendix	140
	Bibliography	155

List of Tables

Table	Page
2.1 Eccentricities of boxes for validated homology using the Golden Ratio . . .	49

List of Figures

Figure	Page
1.1 Effect of applying the boundary operator to the elementary cube Q	9
1.2 Geometries With Only One Connected Component and No Voids	12
1.3 Geometries With One Connected Component and One Void	12
1.4 Geometry With One Connected Component and Two Voids	13
1.5 Geometry of Nodal Domains for Cahn-Hilliard Simulation	14
1.6 Series of Elementary Collapses on the Space $[0, 1] \times [1, 2]$	16
1.7 Five Realizations of $g_2(\omega)x^2 + g_1(\omega)$ over $[-1, 1]$, Mean, and Variance . . .	24
1.8 Betti Numbers for a Simulation of the Cahn-Hilliard-Cook model.	31
1.9 Averaged Zero Betti Number for a Simulation of the Cahn-Hilliard-Cook model.	32
1.10 Betti Numbers for a Simulation of the Viscous Cahn-Hilliard-Cook model.	33
1.11 Effect of noise intensity of the Betti numbers for Cahn-Hilliard-Cook. . . .	33
2.1 Possible Sign Structures for Vertices.	37
2.2 Possible Nodal Lines with Given Sign Structures.	38
2.3 Problem with Nodal Lines for Last Sign Configuration	39
2.4 Example boxes of validated homology routine using the Golden Ratio. . . .	48
2.5 Sample Grids From Validation Algorithm	51
2.6 Key Parameters for Double Well Function Using New Algorithm.	52
2.7 [30] Key Parameters for Double Well Function Using Original Algorithm. .	52
3.1 Correct and Incorrect Cubical Approximation of the Nodal Domain	54
3.2 Nodal Domains for Sample Random Field	56
3.3 Example of a random field with no crossover on $[\alpha, \beta]$	58
3.4 Example of a random field that has a single crossover over $[\alpha, \beta]$	58
3.5 Example of Double Crossover	59
3.6 Forbidden Sign Configurations for B admissibility.	64
3.7 Allowed sign configurations at next dyadic subdivision for B -admissibility. .	65
3.8 Forbidden Sign Configurations for I_4 admissibility.	66
3.9 Forbidden Sign Configurations for I_5 admissibility.	66

3.10	Sign Configurations Required for I -Admissibility.	67
3.11	Possible Sign Structures for Squares Using I_4 Admissibility	68
3.12	Necessity of I -admissibility.	69
3.13	Example of Homogenous Random Field and Non-Homogenous Random Field	72
4.1	Sample Newton Polygon	85
4.2	Newton polygon for equation $F(x, y) = y^5 - 14xy^2 + 5x^2y + x^3$	86
4.3	Newton polygon for three point sign configuration	89
4.4	Four Point Sign Configuration	91
4.5	Newton polygon for four point sign configuration	93
4.6	Skewed Four Point Sign Configuration	96
4.7	Newton polygon for five point sign configuration	102
4.8	Forbidden sign configurations for B -admissibility.	104
4.9	Forbidden Sign Configurations for I_4 admissibility.	104
4.10	Forbidden Sign Configuration for I_5 admissibility.	104
4.11	Sign Configurations Required for I -Admissibility.	105
6.1	Correctness Probability of Random Chebyshev Series for Different Orders. .	125
6.2	Exponent for Probability Threshold for Chebyshev Polynomials	126
6.3	Density of zeros function (<i>top</i>) and the function $\mathcal{C}_0(x)$ (<i>bottom</i>).	127
6.4	Betti Numbers for Chebyshev Series	127
6.5	Correctness Probability of Random Cosine Series for Different Dimensions.	128
6.6	Exponent for Probability Threshold for Cosine Series	129
6.7	Density of zeros function (<i>top</i>) and the function $\mathcal{C}_0(x)$ (<i>bottom</i>).	130
6.8	Betti Numbers for Cosine Series	130

Abstract

OPTIMAL SAMPLING OF RANDOM FIELDS FOR TOPOLOGICAL ANALYSIS

Gregory Scott Cochran, PhD

George Mason University, 2011

Dissertation Director: Dr. Thomas Wanner

Algebraic topology is becoming an increasingly important tool in applied mathematics. In particular, homology theory allows one to distinguish different topologies while being tractable to compute. An important application is the study of nodal domains for solutions to stochastic partial differential equations. These are the sets where the function value is greater than zero and less than zero.

In order to compute the homology of the nodal sets computationally, we must discretize the domains. However, in this discretization process, we can make mistakes in the topology. Can we develop a method that will allow us to determine a proper discretization size a priori?

One approach is to use an algorithm that is guaranteed to return the correct homology. The original algorithm has a few shortcomings. We will present these shortcomings and develop methods to overcome these issues.

The other approach is to establish explicit probability bounds for the making the correct homology. This is an a priori approach that will return the probability for a fixed discretization size and also determines the optimal location of the sampled points.

Chapter 1: Introduction

1.1 Motivation

Physical systems have a natural tendency to vary in time. Many can be described by evolution equations. These systems can also form interesting patterns that evolve with time. How can we analyze the patterns formed? On one hand, there are natural questions about existence and uniqueness of solutions to partial differential equations and any stochastic extensions. On the other hand, we can analyze the solutions using numerics. Although both are useful, can we somehow use techniques that look only at the patterns formed and make a quantitative assessment of the patterns?

Algebraic topology gives a useful tool for studying these complex patterns. It gives us the ability to extract simple topological information from the data set. It can give us the number of connected components of our solution set, the number of holes, and the number of cavities as well as other information. In order to use algebraic topology, we associate a group to the object in question. If we restrict ourselves to abelian groups, then we can use homology. Homology allows us to associate coarse information about the geometry of the object with the ability to easily compute this information. This ability comes at the price of losing more detailed information about the geometry.

Even though computing the homology of the sets is possible, to study time varying systems we cannot expect to be able to compute these objects for all time. We instead make use of numerics and use computational homology. This will allow us to get a time series of the homology of the patterns produced from each model.

The natural question is if we discretize our domain into small subsections, is the topology of our discretization the same as the true topology? The more general question asked is if we want to sample a random manifold with n points, can we create a mesh of the manifold

such that the homology of the mesh and the manifold agree? The first attempt to answer this question is from [62]. Although they provide an answer, it is of limited applicability for the topics we have in mind. The approach they have in mind is a stochastic algorithm which samples the manifold with N points and then finds bounds on the probability of the convergence of the algorithm. This probability depends on M and τ which is a condition number related to the manifold itself.

As previously stated, we are interested in the patterns produced by nonlinear evolution equations. If we perturb our equation with a small random disturbance, what happens to the patterns? To solve such equations, we will discretize our domain and solve the problem numerically. If we discretize our domain into M sub-regions, how can we be guaranteed the homology of the discretized domain matches that of the true domain? As the patterns evolve in time, what happens to the homology of the patterns?

While the answers to the above questions are yes, there is another equally important question to ask. The algorithms for computing the homology of the sets in question are contained in [46]. Since we must replace the true geometries with something that is computationally feasible, will these algorithms produce the desired results?

Suppose we want to do multiple simulations of a physical model and collect statistics about the homologies produced. If we then add noise to the model, how will the homologies change? Can we still ensure the homologies are correct? Since we are adding random noise to the models, we cannot give a simple answer. However, we can give probabilities of correctly computing the homologies.

To begin, we need to give specific details about the collection of functions we will work with. In particular, we will work with Gaussian Random fields. We then give previously known bounds on computing the correct homology for particular examples of Gaussian Random fields. Since we want to somehow validate the bounds, we then give a recursive algorithm to compute the homology.

Most of the remainder of this chapter can be found in various references. The next sections on cubical homology and computational homology can be found in [46], and the homology

software can be obtained through [25]. Since one of our ultimate goals is to compute the homology associated with random fields, we then present relevant topics in probability and random field theory. There are numerous references to these topics but a few relevant to our study are [3], [12], and [13].

Chapter two discusses an algorithm that creates a grid on which we can correctly compute the homology of nodal domains. The original form of this algorithm can be found in [30]. Once the original algorithm has been presented, we will discuss a few shortcomings of this algorithm and discuss improvements we have made to this algorithm.

Chapter three presents the previous probability bounds for making the correct homology computation of random fields. These results can be found in [53], [54], [29], and the references therein. This chapter contains no new material but is essential for the results that we will establish in Chapter four. As such, Chapter four contains new results on the probability for making the correct homology computation of non-homogenous random fields. Chapter five also contains new results on how to maximize this probability. This method is based on the barrier method in optimization. The barrier method can be found in [41] and [7]. Lastly, in Chapter six, we discuss numerical experiments to confirm our new results.

1.2 Cubical Homology

Since we are interested in using homology to analyze patterns, we must first understand what homology really is. In this section, we present an overview of homology alongside computational considerations. Much of the following is from [46]. To begin, we need the following definitions.

Definition 1.1. *An elementary cube is a set*

$$Q = I_1 \times I_2 \times \dots \times I_d \subset \mathbb{R}^d,$$

where $I_j = [l, l + 1]$ or $I_j = [l, l]$, and $l, k \in \mathbb{Z}$. Note, we allow degenerate intervals. The

set of all elementary cubes in \mathbb{R}^d is denoted \mathcal{K}^d . The set of all elementary cubes is denoted

$$\mathcal{K} = \bigcup_d \mathcal{K}^d.$$

Definition 1.2. For an elementary cube Q , the embedding number of $Q \subset \mathbb{R}^d$ is defined to be d . It is denoted $\text{emb } Q$. The dimension of Q is defined to be the number of non-degenerate intervals in Q . Let

$$\mathcal{K}_k := \{Q \in \mathcal{K} \mid \dim Q = k\}$$

and

$$\mathcal{K}_k^d = \mathcal{K}^d \cap \mathcal{K}_k.$$

With these definitions, we can now build more complex sets from elementary cubes.

Definition 1.3. A set $X \subset \mathbb{R}^d$ is a cubical set if $X = \bigcup_{i=1}^n Q_i$, where Q_i is an elementary cube. A cubical set is a finite union of elementary cubes. For a cubical set X , we denote

$$\mathcal{K}(X) := \{Q \in \mathcal{K} \mid Q \subset X\}$$

and

$$\mathcal{K}_k(X) := \{Q \in \mathcal{K}(X) \mid \dim Q = k\}.$$

Since the goal of algebraic topology is to assign algebraic objects to geometric objects, we now turn to the algebraic side of homology. With the previous definitions, we now assign algebraic objects and operations to these sets.

Definition 1.4. For each elementary k -cube $Q \in \mathcal{K}_k^d$, we associate an abstract algebraic object \widehat{Q} called an elementary k -chain. The set of all elementary k -chains of \mathbb{R}^n is denoted

$$\widehat{\mathcal{K}}_k^d := \{\widehat{Q} \mid Q \in \mathcal{K}_k^d\}.$$

This is analogous to a basis in \mathbb{R}^n . For a finite collection of k -chains, we define finite sums

by

$$c = \alpha_1 \widehat{Q}_1 + \dots + \alpha_m \widehat{Q}_m,$$

where $\alpha_i \in \mathbb{Z}$. The set of all k -chains is denoted C_k^d . For chains $c_1 = \sum_{j=1}^m \alpha_j \widehat{Q}_j$ and $c_2 = \sum_{j=1}^m \beta_j \widehat{Q}_j$, define the addition of $c_1 + c_2$ as

$$c_1 + c_2 = \sum_{j=1}^m \alpha_j \widehat{Q}_j + \sum_{j=1}^m \beta_j \widehat{Q}_j := \sum_{j=1}^m (\alpha_j + \beta_j) \widehat{Q}_j.$$

For any k -chain c , define the inverse chain $-c$ by $c + (-c) = 0$. We thus have C_k^d is a free abelian group with basis $\widehat{\mathcal{K}}_k^d$.

While this definition is straightforward, there is an equivalent definition that is more algebraic.

Definition 1.5. For each $Q \in \mathcal{K}_k^d$, define $\widehat{Q} : \mathcal{K}_k^d \rightarrow \mathbb{Z}$ by

$$\widehat{Q}(P) := \begin{cases} 1 & \text{if } P = Q \\ 0 & \text{if } P \neq Q \end{cases}$$

The group C_k^d is the free abelian group generated by the elementary chains. The elements of C_k^d are functions $c : \mathcal{K}_k^d \rightarrow \mathbb{Z}$ such that $c(Q) = 0$ for all but a finite number of $Q \in \mathcal{K}_k^d$. We thus have the basis $\widehat{\mathcal{K}}_k^d$ for C_k^d .

We have a bijection between \mathcal{K}_k^d and $\widehat{\mathcal{K}}_k^d$. To continue with associating algebraic objects to cubical sets we need to move from elementary cubes to cubical sets.

Definition 1.6. For $c_1, c_2 \in C_k^d$, for $c_1 = \sum_{j=1}^m \alpha_j \widehat{Q}_j$ and $c_2 = \sum_{j=1}^m \beta_j \widehat{Q}_j$, the scalar product

is then

$$\langle c_1, c_2 \rangle := \sum_{j=1}^m \alpha_j \beta_j.$$

Note that the scalar product is bilinear. Given elementary cubes $P \in \mathcal{K}_k^d$ and $Q \in \mathcal{K}_{k'}^{d'}$, set

$$\widehat{P} \diamond \widehat{Q} := \widehat{P \times Q}.$$

For $c_1 \in C_k^d$ and $c_2 \in C_{k'}^{d'}$ define the cubical product of c_1 and c_2 by

$$c_1 \diamond c_2 := \sum_{P \in \mathcal{K}_k, Q \in \mathcal{K}_{k'}} \langle c_1, \widehat{P} \rangle \langle c_2, \widehat{Q} \rangle \widehat{P \times Q}.$$

One of the most important consequences of all the above is that we can decompose cubical sets in way that is computationally feasible. In particular, we have the following theorem.

Theorem 1.1. [46] *Let \widehat{Q} be an elementary cubical chain in \mathbb{R}^d with $d > 1$. Then there exist unique elementary cubical chains \widehat{I} and \widehat{P} with $\text{emb } \widehat{I} = 1$ and $\text{emb } \widehat{P} = d - 1$ such that*

$$\widehat{Q} = \widehat{I} \diamond \widehat{P}.$$

As an example, let $Q = [0, 1] \times [1, 2]$. Then Q is an elementary cube with elementary chain \widehat{Q} . We have $\widehat{Q} = [0, 1] \times [1, 2] = [0, 1] \diamond [1, 2]$.

We now move from elementary cubes to sets built from many elementary cubes.

Definition 1.7. *For $X \subset \mathbb{R}^d$ a cubical set, define*

$$\widehat{\mathcal{K}}_k(X) := \{\widehat{Q} \mid Q \in \mathcal{K}(X)\}.$$

Then we have $C_k(X)$ is the subgroup of C_k^d generated by the elements of $\widehat{\mathcal{K}}_k(X)$ and is the

set of k -chains of X .

Proposition 1.1. [46] For any chain c , we have

$$c = \sum_{Q \in \mathcal{K}_k(X)} \langle c, \widehat{Q} \rangle \widehat{Q}.$$

One of the most important tool in homology is the concept of the boundary operator. This is a mapping from the cubical set X to the boundary of X and will be a homomorphism between C_k^d and C_{k-1}^d . Because this map will be a homomorphism, we can investigate the kernel and images of such maps. This will be crucial to our study of homology.

Definition 1.8. For $k \in \mathbb{Z}$, the cubical boundary operator

$$\partial_k : C_k^d \rightarrow C_{k-1}^d$$

is the homomorphism of free abelian groups which is defined recursively by induction on the embedding number d as follows. First, consider $d = 1$. Then Q is an elementary interval so we have $Q = [l, l] \in \mathcal{K}_0^1$ or $Q = [l, l + 1] \in \mathcal{K}_1^1$. Then define

$$\partial_k \widehat{Q} := \begin{cases} 0 & \text{if } Q = [l, l] \\ \widehat{[l+1]} - \widehat{[l]} & \text{if } Q = [l, l + 1]. \end{cases}$$

We now extend this definition for $d > 1$. Suppose $I = I_1(Q)$ and $P = I_2(Q) \times \dots \times I_d(Q)$. Then by the above proposition, we have $\widehat{Q} = \widehat{I} \diamond \widehat{P}$. Then define

$$\partial_k \widehat{Q} := (\partial_{k_1} \widehat{I}) \diamond \widehat{P} + (-1)^{\dim I} \widehat{I} \diamond \partial_{k_2} \widehat{P},$$

where $k_1 = \dim I$ and $k_2 = \dim P$. Lastly, extend this definition to all chains linearly via

$$\partial_k c = \alpha_1 \partial_k \widehat{Q}_1 + \alpha_2 \partial_k \widehat{Q}_2 + \dots + \alpha_m \partial_k \widehat{Q}_m.$$

As an example, let $Q = [0, 1] \times [1, 2]$. Then from the previous example we have $\widehat{Q} = \widehat{[0, 1]} \diamond \widehat{[1, 2]}$. This gives

$$\begin{aligned} \partial_2 \widehat{Q} &= \partial_1 \widehat{[0, 1]} \diamond \widehat{[1, 2]} + (-1) \widehat{[0, 1]} \diamond \partial_1 \widehat{[1, 2]} \\ &= (\widehat{[1]} - \widehat{[0]}) \diamond \widehat{[1, 2]} - \widehat{[0, 1]} \diamond (\widehat{[2]} - \widehat{[1]}) \\ &= \widehat{[1]} \diamond \widehat{[1, 2]} - \widehat{[0]} \diamond \widehat{[1, 2]} - \widehat{[0, 1]} \diamond \widehat{[2]} + \widehat{[0, 1]} \diamond \widehat{[1]} \\ &= \widehat{[1] \times [1, 2]} - \widehat{[0] \times [1, 2]} + \widehat{[0, 1] \times [1]} - \widehat{[0, 1] \times [2]} \\ &= \widehat{B}_1 - \widehat{A}_1 + \widehat{A}_2 - \widehat{B}_2, \end{aligned}$$

where

$$A_1 = [0] \times [1, 2]$$

$$B_1 = [1] \times [1, 2]$$

$$A_2 = [0, 1] \times [1]$$

$$B_2 = [0, 1] \times [2].$$

Notice that this is a sum of the chains for the boundary of Q . This is shown in Figure 1.1. The first image in the figure is Q and the second image is the boundary of Q . Notice that we have specified a direction for A_1 , A_2 , B_1 , and B_2 . This direction is important for 1-chains and thus the boundary operator.

We now simplify the notation of the boundary operator ∂_k to ∂ when it is clear from the

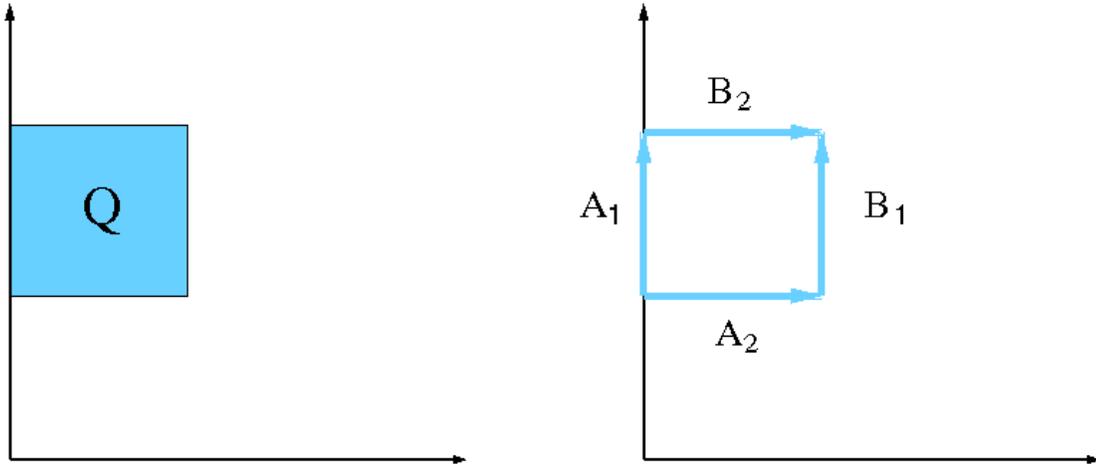


Figure 1.1: Effect of applying the boundary operator to the elementary cube Q

context what the dimension is. While this definition can be applied to arbitrary k -chains, we will often need to take the boundary operator of cubical products. The next proposition gives a method for doing this operation.

Proposition 1.2. [46] *Let c_1 and c_2 be cubical chains. Then we have*

$$\partial(c_1 \diamond c_2) = \partial c_1 \diamond c_2 + (-1)^{\dim c_1} c_1 \diamond \partial c_2.$$

Since the boundary operator maps the cubical set X onto its boundary, what happens if we then take the boundary operator again. Since we have mapped X onto the boundary, we need to remain on the boundary.

Proposition 1.3. [46] *For the boundary operator, we have*

$$\partial \circ \partial = 0.$$

Proposition 1.4. [46] *Let $X \subset \mathbb{R}^d$ be a cubical set. Then we have*

$$\partial_k(C_k(X)) \subset C_{k-1}(X).$$

Definition 1.9. *The boundary operator for the cubical set X is defined to be*

$$\partial_k^X : C_k(X) \rightarrow C_{k-1}(X),$$

by restricting $\partial_k : C_k^d \rightarrow C_{k-1}^d$ to $C_k(X)$. The cubical chain complex for a cubical set $X \subset \mathbb{R}^d$ is

$$\mathcal{C}(X) := \{C_k(X), \partial_k^X\}_{k \in \mathbb{Z}},$$

where $C_k(X)$ are the groups of cubical k -chains generated by $\mathcal{K}_k(X)$ and ∂_k^X is the cubical boundary operator restricted to X .

We are finally ready to give a formal definition of the homology of a cubical set.

Definition 1.10. *For a cubical set $X \subset \mathbb{R}^d$, a k -chain $z \in C_k(X)$ is called a cycle in X if $\partial z = 0$. Thus the set of all k -cycles in X , denoted $Z_k(X)$ is $\ker \partial_k^X$ and forms a subgroup of $C_k(X)$. Thus we have*

$$Z_k(X) := \ker \partial_k^X = C_k(X) \cap \ker \partial_k \subset C_k(X).$$

A k -chain $z \in C_k(X)$ is called a boundary in X if there exists $c \in C_{k+1}(X)$ such that $\partial c = z$. The set of all boundaries in $C_k(X)$, denoted $B_k(X)$ is the image of ∂_{k+1}^X . Note $B_k(X)$ is a subgroup of $C_k(X)$. We thus have

$$B_k(X) := \text{im } \partial_{k+1}^X = \partial_{k+1}(C_{k+1}(X)) \subset C_k(X).$$

Since we know that for a boundary z , $\partial c = z$, we have $\partial z = \partial^2 c = \partial \partial c = 0$. We will be interested in cycles that are not boundaries. We thus treat cycles that are boundaries as trivial. We say that two cycles $z_1, z_2 \in Z_k(X)$ are homologous if $z_1 - z_2$ is a boundary in X , so $z_1 - z_2 \in B_k(X)$. We write this as $z_1 \sim z_2$. This is an equivalence relation and we can form the quotient group. Finally, we define the cubical homology group $H_k(X)$ as the

quotient group

$$H_k(X) := Z_k(X)/B_k(X).$$

The collection of all homology groups of X is the homology of X and is denoted

$$H_*(X) := \{H_k(X)\}.$$

For a cubical set $X \subset \mathbb{R}^d$, it can be shown that all the homology groups $H_i(X)$ must be of the form

$$H_i(X) = \mathbb{Z}^{\beta_i} \times \mathbb{Z}_{b_{i,1}} \times \mathbb{Z}_{b_{i,2}} \times \cdots \times \mathbb{Z}_{b_{i,p_i}},$$

for $i = 1, \dots, d$. The number β_i is the i -th Betti number for X and b_k are the torsion coefficients. This follows from the Fundamental Theorem for Abelian Groups. This particular form of $H_i(X)$ is also given in such a way that \mathbb{Z}_{b_k} is the integers modulo b_k with the constraint b_k divides b_{k+1} .

It can be shown that for $d < 3$, there do not exist any torsion coefficients. Also it can be shown that $H_i(X) = 0$ for $i > d$. The Betti numbers can be interpreted as follows:

- a.) $\beta_0(X)$ is the number of connected components of X
- b.) $\beta_1(X)$ is the number of loops or holes in X
- c.) $\beta_k, k > 1$ is the number of k dimensional holes in X .

Formalizing this gives us the following.

Theorem 1.2. *For a cubical set X , $H_0(X)$ is a free abelian group. Also, if*

$$\{P_i | i = 1, \dots, n\}$$

is a collection of vertices in X with one vertex for each connected component of X , then

$$\{|\widehat{P}_i| \in H_0(X) | i = 1, \dots, n\}$$



Figure 1.2: Geometries With Only One Connected Component and No Voids

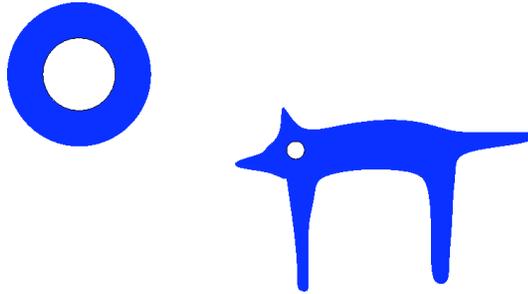


Figure 1.3: Geometries With One Connected Component and One Void

forms a basis of $H_0(X)$.

The important consequence of this theorem is that if the rank of $H_0(X)$ is r , then there are r connected components in X . For example, let $Q = [0, 1] \times [1, 2]$. Then there are four vertices in this cube. Take any of these vertices, in particular, take $P_1 = [0] \times [1]$. Then the $\widehat{[0] \times [1]}$ forms a basis for $H_0(Q)$. Notice that this basis only has one element and that the number of connected components is also one for Q .

If we look at Figure 1.2, each geometry has one connected component. By the preceding theorem, a basis for $H_0(X)$ has only one element for each geometry. If on the other hand, we regard the image as a single geometry, then there are three separate components that are not connected. Then a basis for $H_0(X)$ must have three elements. If we examine Figure 1.3, then each geometry has one connected component but now each also has a hole for a void in the interior of the geometry. This suggests that the basis for $H_0(X)$ has one element and the basis for $H_1(X)$ has one element. However, if we regard the entire image as a single geometry, then there are two connected components and two holes in the interior. Lastly examine Figure 1.4. In this case, there is one connected component but there are two holes in the interior. It should be noted that the previous examples are not cubical sets. However, we can represent each geometry by cubical sets. The process of moving



Figure 1.4: Geometry With One Connected Component and Two Voids

from these geometries to cubical sets must be done with care. We need to enforce that the homologies of the cubical sets to be homotopic to the images shown in the figures. This will ensure that the homology of the cubical sets is the same as the homology of the images shown in the figures. This process will be made clear in later chapters.

For the previous examples, it was clear how many components and holes there were for each geometry. In many of the applications we have in mind, this is not the case. As an example, we now examine the nodal domains for a realization of the Cahn-Hilliard equation. If we look at figure Figure 1.5, we can see there are two different regions of interest. One region is the light blue region B_1 and the other region is the dark blue region B_2 . It should be clear that $B_1 \cap B_2 = \emptyset$. If we count the number of regions of each we see there are nine dark blue regions and five light blue regions so we have $H_0(B_1) = \mathbb{Z}^5$ and $H_0(B_2) = \mathbb{Z}^9$. For the dark blue region, there are no holes in any of the components. However, for the light blue region, there are two holes in the components. Note that in this case, the two holes in the light blue region correspond to connected components in the dark blue region. Thus we have $H_1(B_1) = \mathbb{Z}^2$ and $H_1(B_2) = 0$. In terms of Betti numbers, we have $\beta_0(B_1) = 5$, $\beta_0(B_2) = 9$, $\beta_1(B_1) = 2$, and $\beta_1(B_2) = 0$. This example is one of our primary motivations and will be formalized in later sections.

For a more complex geometry, it should be apparent that computing the homology can become quite tedious. Luckily, there is a method that reduces the overall complexity.

Definition 1.11. *For a cubical set X , let $P, Q \in \mathcal{K}(X)$. If $Q \subset P$, then Q is a face of P*

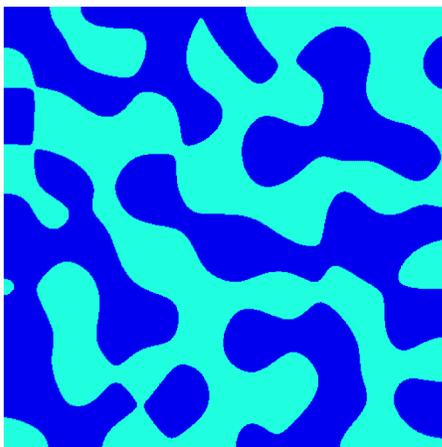


Figure 1.5: Geometry of Nodal Domains for Cahn-Hilliard Simulation

and is denoted $Q \preceq P$. If $Q \preceq P$ and $Q \neq P$, then Q is a proper face of P , denoted $Q \prec P$. Q is a primary face of P , if Q is a face of P and $\dim Q = \dim P - 1$. If Q is not a proper face for some $P \in \mathcal{K}(X)$, then Q is a maximal face. Denote the set of all maximal faces by $\mathcal{K}_{\max}(X)$. A face that is a proper face of only one elementary cube in X is a free face of X .

Let Q be a free face of X and let P be the unique cube in $\mathcal{K}(X)$ such that Q is a proper face of P . Let

$$\mathcal{K}'(X) := \mathcal{K}(X) - \{Q, P\}.$$

Define

$$X' := \bigcup_{R \in \mathcal{K}'(X)} R.$$

Then X' is a cubical space from X by means of an elementary collapse of P by Q .

We can use elementary collapses to simplify the geometry of our cubical set X . This must be done in such a way that the resulting homology is the same as the homology of X .

Proposition 1.5. [46] *If X' is a cubical space obtained from X by an elementary collapse of P by Q , then*

$$\mathcal{K}(X') = \mathcal{K}'(X).$$

Theorem 1.3. *Let X be a cubical set and assume X' is obtained from X by an elementary collapse of $P_0 \in \mathcal{K}_k(X)$ by $Q_0 \in \mathcal{K}_{k-1}(X)$. Then*

$$H_*(X') \cong H_*(X).$$

We can repeat elementary collapses to greatly reduce the complexity of the geometry to something easier to compute. For example, in \mathbb{R}^2 , a cube has the homology as a single point. Again, take the cube $Q = [0, 1] \times [1, 2]$. For this cube we have

$$\mathcal{K}_2 = \{[0, 1] \times [1, 2]\}$$

$$\mathcal{K}_1 = \{[0] \times [1, 2], [1] \times [1, 2], [0, 1] \times [1], [0, 1] \times [2]\}$$

$$\mathcal{K}_0 = \{[0] \times [1], [0] \times [2], [1] \times [1], [1] \times [2]\}.$$

Using these, there are four free faces for Q , namely the elements in \mathcal{K}_1 . If we let $R_1 = [0, 1] \times [2]$, then $R_1 \prec Q$. We can then collapse Q into the space $Q_1 = [0] \times [1, 2] \cup [1] \times [1, 2] \cup [0, 1] \times [1]$. This was done by the elementary collapse by using R_1 and Q .

We now have

$$\mathcal{K}_2 = \emptyset$$

$$\mathcal{K}_1 = \{[0] \times [1, 2], [1] \times [1, 2], [0, 1] \times [1]\}$$

$$\mathcal{K}_0 = \{[0] \times [1], [0] \times [2], [1] \times [1], [1] \times [2]\}.$$

We can now do another elementary collapse. This time take the free face $[0] \times [1, 2]$. We can keep repeating this process until we are left with a space that has no free faces. This is shown in Figure 1.6.

This shows that we are able to collapse our original space $Q = [0, 1] \times [1, 2]$ down to the point $[0] \times [1]$. We have already seen that this point was a basis for Q . Spaces that are able to collapse to a single point are known as acyclic spaces.

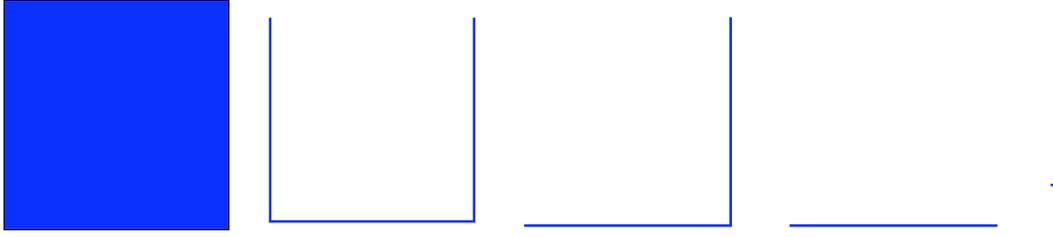


Figure 1.6: Series of Elementary Collapses on the Space $[0, 1] \times [1, 2]$

Definition 1.12. *A cubical set X is acyclic if*

$$H_*(X) \cong \begin{cases} \mathbb{Z} & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$$

Theorem 1.4. *All elementary cubes are acyclic.*

This theorem shows that the example $Q = [0, 1] \times [1, 2]$ is acyclic. Even though it is a two dimensional set that is embedded in \mathbb{R}^2 it has the same homology as a single point. In higher dimensions, the same result still holds: if Q is an n -dimensional connected set, then Q has the same homology as a single point in \mathbb{R}^n .

If we are given more than one acyclic set, what happens to the intersection of the sets? Does it remain acyclic? This is the content of the next proposition.

Proposition 1.6. *[46] Let $K, L \in \mathbb{R}^d$ be cubical sets. Then*

$$C_k(K \cup L) = C_k(K) + C_k(L).$$

Theorem 1.5. *Let X and Y be cubical sets in \mathbb{R}^d . If X, Y , and $X \cap Y$ are acyclic, then $X \cup Y$ is acyclic.*

1.3 Computational Homology

We have given an overview of cubical homology. For large data sets, it is infeasible to compute the homology by hand. We now turn our attention to algorithms which are well suited for computational purposes. The key to such computations is linear algebra. The computation of homology groups is equivalent to matrix algebra over the integers. We require basic algorithms for computing addition and multiplication of integer matrices and also a row reduction algorithm. However, the key to computing such groups is computing the Smith Normal Form.

Since we know that the homology groups will be free abelian groups of the form

$$H_i(X) = \mathbb{Z}^{\beta_i} \times \mathbb{Z}_{b_1} \times \mathbb{Z}_{b_2} \times \cdots \times \mathbb{Z}_{b_p},$$

we have a basis for the group. Since we have a basis, we can find a matrix representation for the boundary operator, call this matrix A . This represents a homomorphism of abelian groups $A : \mathbb{Z}^{d_2} \rightarrow \mathbb{Z}^{d_2}$. Note this matrix will be defined over the integers \mathbb{Z} . We are interested in computing both the kernel of this matrix and the image of this matrix transformation. This means we need a method for calculating $x \in \mathbb{Z}^{d_2}$ such that

$$Ax = 0$$

and finding $z \in \mathbb{Z}^{d_2}$ such that

$$Ax = z,$$

for some $x \in \mathbb{Z}^{d_1}$.

Using normal techniques such as row reduction and Gaussian Elimination are insufficient.

We must then resort to computing the Smith Normal Form.

Suppose we have a matrix $A \in \mathbb{Z}^{d \times d}$. Then there are integer matrices P and Q such that $\det P = \pm 1$, $\det Q = \pm 1$ and

$$D = PAQ$$

is a diagonal matrix with entries p_1, p_2, \dots, p_s such that p_1 divides p_2 , p_2 divides p_3 and so forth. In particular, we have $p_1 | p_2 | \dots | p_s$. The matrix D is the Smith Normal Form of A . Once we have computed the Smith Normal Form D of the matrix representation, we can then use D to compute the image and kernel of A . This will then give us the generators for the homology groups.

Although it has now been demonstrated how to compute the homology for a cubical set, it should be noted that these algorithms are computationally expensive in their current form. More sophisticated algorithms exist which can greatly reduce the number of overall computations required thus greatly improving running time.

1.4 Random Fields and Analysis

Much of our interest lies in the study of the homology of random fields. In particular, we are interested in the topology of solutions to either stochastic partial differential equations or partial differential equations with a random ensemble of initial conditions. The main interest is in obtaining averaged topological information. Before we can accomplish our goal, we need a thorough foundation of probability and random fields.

In order to discuss probability, we require an underlying probability space which consists of the triple $(\Omega, \mathcal{F}, \mathbb{P})$. The first quantity Ω is the sample space that consists of all possible outcomes. The second quantity \mathcal{F} is a family of subsets of possible events. Formally, \mathcal{F} is a σ -algebra. In particular, this definition requires the following

1. We have $\emptyset \in \mathcal{F}$,
2. If $F \in \mathcal{F}$, then $F^c \in \mathcal{F}$ (the complement of F),
3. If $F_1, F_2, \dots \in \mathcal{F}$, then $F = \bigcup_{j=1}^{\infty} F_j \in \mathcal{F}$.

Lastly, we have the measure \mathbb{P} prescribing the probability for each event. The probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ is then a measure space with measure \mathbb{P} .

The family \mathcal{F} and the probability measure \mathbb{P} satisfy the following properties:

1. The events Ω and \emptyset are both in \mathcal{F} .
2. If E and F are events, then E^c , $E \cap F$, and $E \cup F$ are all events in \mathcal{F} .
3. For all events E , we have $0 \leq \mathbb{P}(E) \leq 1$.
4. If E is an event, then $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.
5. $\mathbb{P}(\Omega) = 1$.
6. $\mathbb{P}(\emptyset) = 0$.
7. If E_1, \dots , are events, then $\bigcup_{j=1}^{\infty} E_j$ and $\bigcap_{j=1}^{\infty} E_j$ are also events.
8. If E_1, \dots , are all pairwise disjoint events, then $\mathbb{P}(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mathbb{P}(E_j)$.

Given a family of subsets \mathcal{U} of Ω , we can always find a smallest σ -algebra containing \mathcal{U} . One can show that this smallest σ -algebra has the form

$$\mathcal{F}_{\mathcal{U}} = \bigcap \{ \mathcal{F} : \mathcal{U} \subset \mathcal{F}, \mathcal{F} \text{ is a } \sigma\text{-algebra} \}.$$

If $X : \Omega \rightarrow \mathbb{R}$ is a function, then the σ -algebra \mathcal{F}_X generated by X is the smallest σ -algebra that contains all the sets $X^{-1}(U)$, where $U \subset \mathbb{R}$ open. This generated σ -algebra has the form

$$\mathcal{F}_X = \{ X^{-1}(B) : B \in \mathcal{B} \},$$

where \mathcal{B} is the Borel σ -algebra on \mathbb{R} and the Borel σ -algebra is the smallest σ -algebra containing all open sets.

Given two measure spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ a function $f : \Omega_1 \rightarrow \Omega_2$ is measurable if

$$f^{-1}(F_2) = \{ \omega_1 \in \Omega_1 : f(\omega_1) \in F_2 \} \in \mathcal{F}_1$$

for all $F_2 \in \mathcal{F}_2$. A random variable $X(\omega)$ or $g(\omega)$ is a measurable function from our probability space to the the real numbers. All random variables $X(\omega)$ induce a probability

measure μ_X on \mathbb{R} defined by

$$\mu_X(E) = \mathbb{P}(X^{-1}(E)).$$

The expected value of a random variable is

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P},$$

provided $\int_{\Omega} |X(\omega)| d\mathbb{P} < \infty$. The variance of a random variable $X(\omega)$ is defined as

$$\text{var}(X) = \mathbb{E} \left[(X - \mathbb{E}[X])^2 \right].$$

An equivalent form is

$$\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

An important concept that we will use often is independence of random variables. Two subsets $E, F \in \mathcal{F}$ are independent if

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

This states that the probability of one event occurring does not affect the probability of the other event to occur. We can extend this definition to arbitrary collections of events. A collection of random variables $\{X_i\}_{i \in I}$ is independent if the collection of the σ -algebras generated are independent. If this is the case, then we have

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y].$$

A random variable $X(\omega)$ has the probability density function ϕ if

$$\mathbb{P}[a \leq X \leq b] = \int_a^b \phi(x) dx,$$

for all $a, b \in \mathbb{R} \cup \{\pm\infty\}$. The probability density function we will be working with is the Gaussian distribution. If $X(\omega)$ is a Gaussian random variable, then the probability density function has the form

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Using this density function, the mean of X is μ and the variance is σ^2 . We write this as $X \sim N(\mu, \sigma^2)$. If the mean is zero and the variance is one, we call this distribution the Standard Normal distribution.

With the above properties of probability and expected value, we can now start defining random fields. Much of the following is from [3].

Definition 1.13. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X a topological space. Then a mapping $f : \Omega \rightarrow \mathbb{R}^X$ is called a real-valued random field. If f is measurable, we say f is a measurable random field.*

The simplest way to interpret a random field is as an object which randomly chooses a function according to some probability distribution. So we have that $f(\cdot, \omega)$ is a function $f(\cdot, \omega) : X \rightarrow \mathbb{R}$ and $f(x, \omega)$ is the realization at $x \in X$.

Given a random field $f(x, \omega)$, the mean or expected value of f is

$$m(x) = m_x := \mathbb{E}[f(x, \omega)] = \int_{\Omega} f(x, \omega) d\mathbb{P}.$$

This is an integration with respect to the probability measure over Ω . We can interpret the variable x as stationary. The covariance function between two points $x_1, x_2 \in X$ is

$$R(x_1, x_2) = \mathbb{E}[(f(x_1, \omega) - m_{x_1})(f(x_2, \omega) - m_{x_2})].$$

This is a measure of how much the function values of x_1 and x_2 change together. With these definitions, the variance at a point $x \in X$ is the covariance of a point with itself.

Namely, we have

$$\text{var}(x) = \sigma_x^2 = R(x, x) = \mathbb{E} \left[(f(x, \omega) - m_x)(f(x, \omega) - m_x) \right].$$

The variance is a measure of how much x can deviate from the mean value.

It should be noted that the covariance function is positive definite. A matrix $C_n \in \mathbb{R}^{n \times n}$ is positive definite if $z^T C_n z \geq 0$ for $z \in \mathbb{R}^n$. A function $R : X \times X \rightarrow \mathbb{R}$ is positive definite if the matrices formed from $(R(x_i, x_j))$ are positive definite for $(x_1, \dots, x_n) \in \mathbb{R}^n$ and $1 \leq n < \infty$. To see this, let $z \in \mathbb{R}^n$ and $x_j \in D$, $j = 1, \dots, n$, where D is the domain of the random field $u(\cdot, \omega)$. Then we have

$$\begin{aligned} z^t R(x, x) z &= \sum_{j=1}^n \sum_{k=1}^n z_j R(x_j, x_k) z_k \\ &= \sum_{j=1}^n \sum_{k=1}^n z_j \mathbb{E} [(u(x_j, \cdot) - \mathbb{E}[u(x_j, \cdot)]) (u(x_k, \cdot) - \mathbb{E}[u(x_k, \cdot)])] z_k \\ &= \mathbb{E} \left[\sum_{j=1}^n \sum_{k=1}^n (u(x_j, \cdot) - \mathbb{E}[u(x_j, \cdot)]) (u(x_k, \cdot) - \mathbb{E}[u(x_k, \cdot)]) z_j z_k \right] \\ &= \mathbb{E} \left[\sum_{j=1}^n \sum_{k=1}^n y_j y_k z_j z_k \right] \\ &= \mathbb{E} [(y^t z)^2] \geq 0, \end{aligned}$$

where we have defined $y_j = u(x_j, \cdot) - \mathbb{E}[u(x_j, \cdot)]$. This shows that the covariance matrix formed from all finite dimensional distributions is positive definite.

As an example, let $X = [-1, 1]$ and let $g_1(\omega), g_2(\omega)$ be two independent, random variables from a Standard Normal distribution. Define

$$f(x, \omega) = g_2(\omega)x^2 + g_1(\omega).$$

Then the expected value is

$$\begin{aligned}
 m(x) &= \mathbb{E} \left[f(x, \omega) \right] \\
 &= \int_{\Omega} (g_2(\omega)x^2 + g_1(\omega)) d\mathbb{P} \\
 &= x^2 \int_{\Omega} g_2(\omega) d\mathbb{P} + \int_{\Omega} g_1(\omega) d\mathbb{P} \\
 &= 0,
 \end{aligned}$$

since the expected values of $g_1(\omega)$ and $g_2(\omega)$ are zero. Since we know the expected value of this random field is zero, we can calculate the covariance as

$$\begin{aligned}
 R(x_1, x_2) &= \mathbb{E} \left[f(x_1, \omega) f(x_2, \omega) \right] \\
 &= \mathbb{E} \left[(g_2(\omega)x_1^2 + g_1(\omega)) (g_2(\omega)x_2^2 + g_1(\omega)) \right] \\
 &= \mathbb{E} \left[x_1^2 x_2^2 g_2^2(\omega) + g_1(\omega) g_2(\omega) x_1^2 + g_1(\omega) g_2(\omega) x_2^2 + g_1(\omega)^2 \right] \\
 &= x_1^2 x_2^2 + 1,
 \end{aligned}$$

where we have used $\mathbb{E}[g_i^2(\omega)] = 1$ since $g_1(\omega)$ and $g_2(\omega)$ are drawn from a Standard Normal distribution and the $g_1(\omega)$ and $g_2(\omega)$ are independent. Lastly, we know that the variance at each point is simply the covariance between the point and itself we have

$$\sigma_x^2 = R(x, x) = x^4 + 1.$$

We plot a few realizations of this random field, the mean and the variance in Figure 1.7.

A random field is called homogenous if for any two points $x_1, x_2 \in X$, the covariance

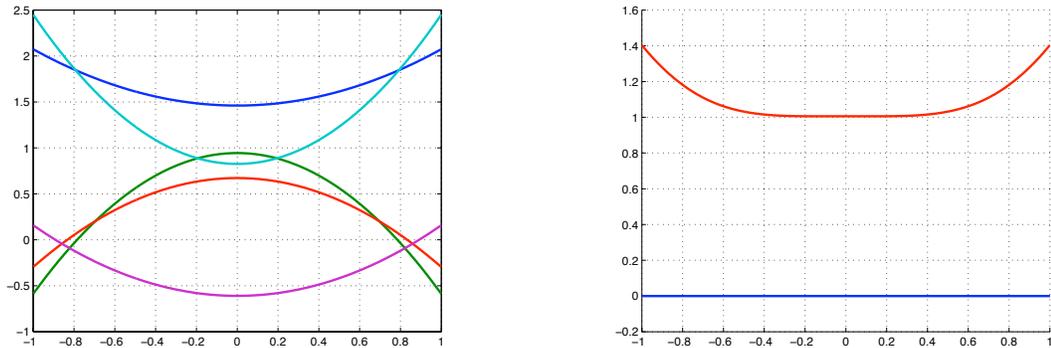


Figure 1.7: Five Realizations of $g_2(\omega)x^2 + g_1(\omega)$ over $[-1, 1]$, Mean, and Variance

depends only on the distance between them. We can then write

$$R(x, y) = r(x - y),$$

or

$$R(x, y) = r(y - x),$$

for an appropriate function r . As a quintessential example, let $g_1(\omega)$ and $g_2(\omega)$ be independent random variables from a Standard Normal distribution. Let $X = [0, 1]$ and $x \in X$.

Define

$$f(x, \omega) = g_1(\omega) \cos 2\pi x + g_2(\omega) \sin 2\pi x,$$

where $g_1(\omega)$ and $g_2(\omega)$ are independent Gaussian random variables with mean zero and variance one. Then we have the mean is zero since

$$\begin{aligned} \mathbb{E}[f(x, \omega)] &= \mathbb{E}[g_1(\omega) \cos(2\pi x) + g_2(\omega) \sin(2\pi x)] \\ &= \cos(2\pi x) \mathbb{E}[g_1(\omega)] + \sin(2\pi x) \mathbb{E}[g_2(\omega)] \\ &= 0. \end{aligned}$$

We can write the covariance between $x_1, x_2 \in X$ as

$$\begin{aligned}
R(x_1, x_2) &= \mathbb{E} \left[f(x_1, \omega) f(x_2, \omega) \right] \\
&= \mathbb{E} \left[(g_1(\omega) \cos 2\pi x_1 + g_2(\omega) \sin 2\pi x_1) (g_1(\omega) \cos 2\pi x_2 + g_2(\omega) \sin 2\pi x_2) \right] \\
&= \mathbb{E} \left[g_1^2(\omega) \cos 2\pi x_1 \cos 2\pi x_2 + g_1(\omega) g_2(\omega) \cos 2\pi x_1 \sin 2\pi x_2 \right. \\
&\quad \left. + g_1(\omega) g_2(\omega) \sin 2\pi x_1 \cos 2\pi x_2 + g_2^2(\omega) \sin 2\pi x_1 \sin 2\pi x_2 \right] \\
&= \cos 2\pi x_1 \cos 2\pi x_2 + \sin 2\pi x_1 \sin 2\pi x_2 \\
&= \cos(2\pi(x_2 - x_1)) \\
&= r(x_2 - x_1).
\end{aligned}$$

This shows that the covariance does not depend on where we sample our points, but on the distance between the points. If we calculate the covariance between two points separated by a distance $\delta > 0$, then we can shift the sampled points within the domain and the covariance remains the same as long as the points are still separated by a distance $\delta > 0$. We can extend this same idea to the random field

$$f(x, \omega) = \sum_{j=1}^N a_j (g_{2j}(\omega) \cos(2\pi j x) + g_{2j-1}(\omega) \sin(2\pi j x)),$$

for coefficients a_j if the $g_j(\omega)$ are pairwise independent. Then the expected value is also zero, and the covariance is then

$$R(x_1, x_2) = \sum_{j=1}^N a_j^2 \cos(2\pi j(x_2 - x_1)),$$

for $g_k(\omega)$ independent, identically normally distributed random variables. This covariance also only depends on the distance between the points. For a homogenous random field f , we want to be able to write f as the series

$$f(x, \omega) = \sum_{j=1}^{\infty} a_j (g_{2j}(\omega) \cos 2\pi jx + g_{2j-1}(\omega) \sin 2\pi jx).$$

Also, we hope to write a general random field as

$$f(x, \omega) = \sum_{j=0}^{\infty} g_j(\omega) \phi_j(x),$$

for $g_j(\omega)$ a random variable, and ϕ_j an appropriate basis function defined over the domain. In order to justify these random series, we must determine appropriate convergence properties of such series.

1.4.1 Orthogonal Expansions of Random Fields

Up to now, we have been given a class of random fields and calculated the covariance function. However, in order to justify the convergence of these infinite series we need a proper space of functions. The main theorem from this section will state that every Gaussian random field with a continuous covariance function can be represented as

$$f(x, \omega) = \sum_{j=1}^{\infty} \xi_j \phi_j(x),$$

where ξ_j are independent, Gaussian random variables and $\phi_j(x)$ are functions defined on the domain X determined by the covariance function R . First, we require an appropriate space of functions. To start, suppose we have a positive definite function $R : X \times X \rightarrow \mathbb{R}$.

Define the space S to be

$$S = \left\{ u : X \rightarrow \mathbb{R} : u(\cdot) = \sum_{i=1}^n a_i R(x_i, \cdot), a_i \in \mathbb{R}, x_i \in X \right\}$$

Equip S with the inner product

$$\begin{aligned} \langle u, v \rangle_H &= \left\langle \sum_{i=1}^n a_i R(x_i, \cdot), \sum_{j=1}^m b_j R(y_j, \cdot) \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^m a_i b_j R(x_i, y_j). \end{aligned}$$

This space has the following interesting property. Let $u \in S$, then

$$\begin{aligned} \langle u, R(y, \cdot) \rangle_H &= \left\langle \sum_{i=1}^n a_i R(x_i, \cdot), R(y, \cdot) \right\rangle_H \\ &= \sum_{i=1}^n a_i R(x_i, y) \\ &= u(y) \end{aligned}$$

Define $H(R)$ to be closure of S under the norm $\|u\|_H^2 = \langle u, u \rangle$. Then $H(R)$ is called the *Reproducing Kernel Hilbert Space of R* . Using the above inner product, we can then give theorems about convergence of sequences within the space.

Now define $\mathcal{H} = \text{span} \{f(x, \omega) : x \in X\} \subset L^2(\mathbb{P})$, where \mathcal{H} inherits the inner product from $L^2(\mathbb{P})$. It is important to note that \mathcal{H} is the span of our random field associated to the covariance function R . Next define the mapping $\Theta : S \rightarrow \mathcal{H}$ by

$$\Theta(u) = \Theta \left(\sum_{i=1}^n a_i R(x_i, \cdot) \right) = \sum_{i=1}^n a_i f(x_i, \omega).$$

Note that

$$\left\| \Theta \left(\sum_{i=1}^n a_i R(x_i, \cdot) \right) \right\|_H^2 = \sum_{i,j=1}^n a_i R(x_i, x_j) a_j = \left\| \sum_{i=1}^n a_i f(x_i, \omega) \right\|_2^2,$$

and also $\Theta(U)$ is Gaussian for all $u \in S$. This shows that Θ extends to $H(R)$ with the range being the same as \mathcal{H} . We now build an orthonormal basis for \mathcal{H} . Suppose $\{\phi_n\}$ is an orthonormal basis for $H(R)$. Then set

$$\eta_n = \Theta(\phi_n).$$

Then $\{\eta_n\}$ is an orthonormal basis for \mathcal{H} . Also $\eta_n \sim N(0, 1)$ and

$$f(x, \omega) = \sum_{j=1}^{\infty} \eta_j \mathbb{E} \{f(x, \omega) \eta_j\},$$

where this series converges in $L^2(\mathbb{P})$. The next few theorems from [3] give convergence results for this space.

Theorem 1.6. [3] *If $\{\phi_n\}$ is an orthonormal basis for $H(R)$, then $f \in H(R)$ has the representation*

$$f(x, \omega) = \sum_{k=1}^{\infty} \xi_k \phi_k(x),$$

where $\{\xi_n\}$ is an orthonormal sequence of centered Gaussian variables and convergence is in L^2 .

Theorem 1.7. [3] *If f is almost-surely continuous, then the above sum converges uniformly on X with probability 1.*

Theorem 1.8. [3] *Let $\{\phi_n\}_n$ be an orthonormal basis for $H(R)$. If the covariance function*

R is continuous, each ϕ_n is continuous and

$$\sum_{n=1}^{\infty} \phi_n^2(x)$$

converges uniformly to $R(x, x)$ for $x \in X$.

In order to state the next theorem, define the operator $\mathcal{C} : L^2(X) \rightarrow L^2(X)$ by

$$\mathcal{R}(\phi)(x) = \int_X R(x, y)\phi(y)dy.$$

Next, let ϕ_n and λ_n be eigenfunctions and eigenvalues respectively to the problem

$$\int_X R(x, y)\phi(x)dx = \lambda\phi(y).$$

Theorem 1.9 (Mercer's Theorem). [3] *Let R be a covariance function. Let ϕ_n and λ_n be the eigenfunctions and eigenvalues for the above eigenvalue problem. Then*

$$R(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x)\phi_i(y),$$

and the series converges uniformly.

1.5 Cahn-Hilliard Simulations

We have now given sections on computational homology and random fields. In this section, we apply this topics to a specific application. The application we have in mind is phase separation of binary alloys. The Cahn-Hilliard-Cook equation is

$$u_t = -\Delta (\epsilon^2 \Delta u + f(u)) + \sigma \cdot \eta_t.$$

The parameter ϵ can be interpreted as the interaction length and η_t is a white noise process. In this equation, values of u close to 1 correspond to a high concentrate of one alloy and values of u close to -1 correspond to the other alloy. This model is mass preserving. Another model of phase separation is the Allen-Cahn equation

$$u_t = \epsilon^2 \Delta u + f(u) + \frac{1}{m(D)} \int_D u dx + \sigma \cdot \eta_t.$$

Lastly, the Viscous Cahn-Hilliard equation is

$$\alpha u_t - (1 - \alpha) \epsilon^2 \Delta u_t = -\Delta(\epsilon^2 \Delta u + f(u)) + \sigma \cdot \eta_t.$$

This is an interpolation between the Cahn-Hilliard-Cook equation and the Allen-Cahn equation. For $\alpha = 0$, we recover the Allen-Cahn equation and for $\alpha = 1$ we recover the Cahn-Hilliard equation. All of these models are defined on a domain D with suitable boundary conditions.

For the simulations we present, we take the domain $D = [0, 1] \times [0, 1]$ with the boundary conditions

$$\frac{\partial u}{\partial \nu} = \frac{\partial \Delta u}{\partial \nu} = 0,$$

when $\alpha \neq 0$, and the boundary conditions

$$\frac{\partial u}{\partial \nu} = 0,$$

when $\alpha = 0$. Also, we take the nonlinear term to be $f(u) = u^3 - u$. This is the derivative of the double well potential $F(u) = 1/4(u^2 - 1)^2$. Qualitatively, the models are similar. Our main goal is to use homology to investigate the dynamics of the patterns produced. For our simulations, we will assume equal mass in each component.

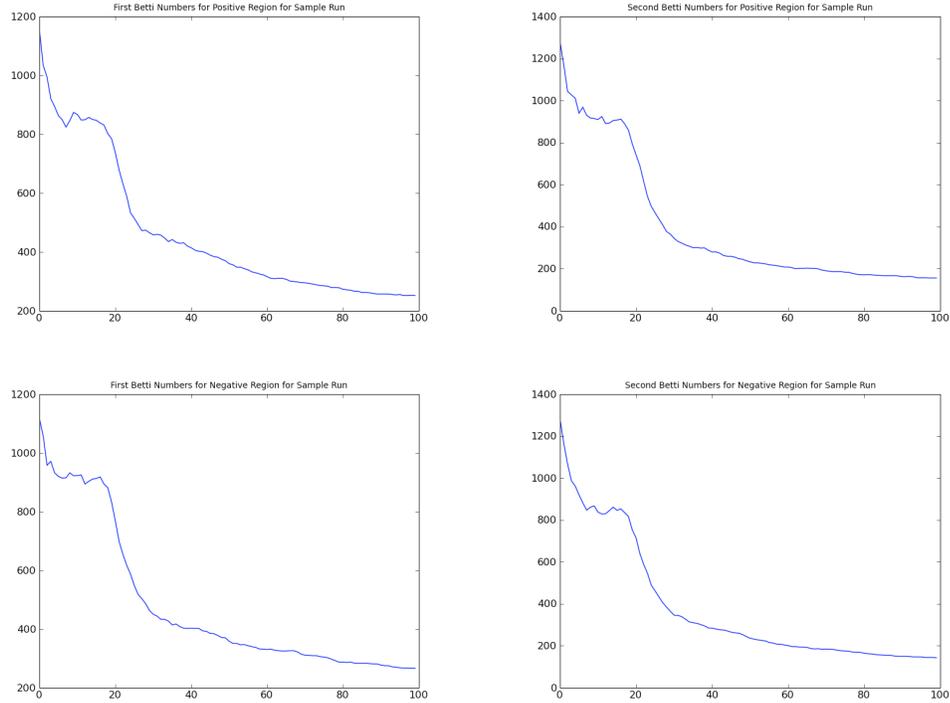


Figure 1.8: Betti Numbers for a Simulation of the Cahn-Hilliard-Cook model.

We are interested in the topology of the nodal domains

$$N^\pm(t) = \{(x, y) \in [0, 1] \times [0, 1] : \pm u(x, y, t) \geq 0\}.$$

Computationally, we will instead work with the sets

$$Q_M^\pm(t_k) = \{Q_{l,n} : \pm u(x_l, y_n, t_k) \geq 0\},$$

where t_k is the time step of the numerical scheme. For the simulations we present, we will use a spectral method to find the solution u . We do 10,000 time steps and every 100 time steps, we compute the homology of the nodal domains.

In Figure 1.8, we see a sample of the Betti numbers produced for a simulation. The parameters were a small interaction length, equal mass, and a small noise intensity. Notice that initially the Betti numbers are fairly large but quickly decay. Also notice that there is

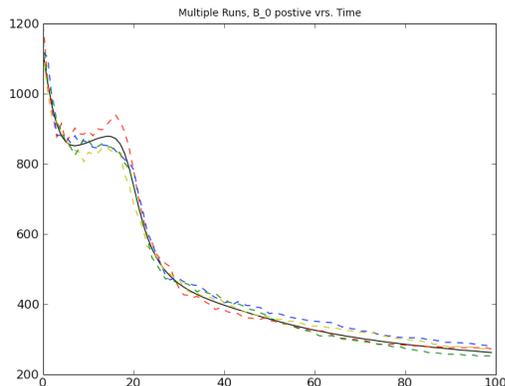


Figure 1.9: Averaged Zero Betti Number for a Simulation of the Cahn-Hilliard-Cook model.

a period where the Betti numbers level off before they decay again. This was for a single simulation with $\sigma = .0001$.

Is this behavior typical? If we repeat the simulations with the same parameters and only vary the initial conditions, we can see that this behavior should be observed on each simulation. The results of 1,000 simulations are shown in Figure 1.9. This image shows the zero'th Betti number for the positive nodal domain. The solid black line is the averaged Betti number at each time step, and the dashed lines are the zero'th Betti number for different simulations. Notice that this period where the Betti numbers level off is observed on average.

The previous simulations were only for the Cahn-Hilliard-Cook model. We also performed simulations for the Viscous Cahn-Hilliard-Cook equation for the parameters $\alpha = 0.0, .25, .5, .75, \text{ and } 1.0$. The results of these simulations are shown in Figure 1.10. This figure shows the averaged zero'th and first Betti number for the positive nodal domains. Observe that the period where the Betti numbers level off is present in almost all the parameter values.

The period where the Betti numbers level off is present in all the above simulations. However, we note that the noise intensity was the same small parameter for all of these simulations. The effect of increasing the noise intensity is shown in Figure 1.11. This figure shows the zero'th Betti numbers of the Positive nodal domain for the Cahn-Hilliard-Cook

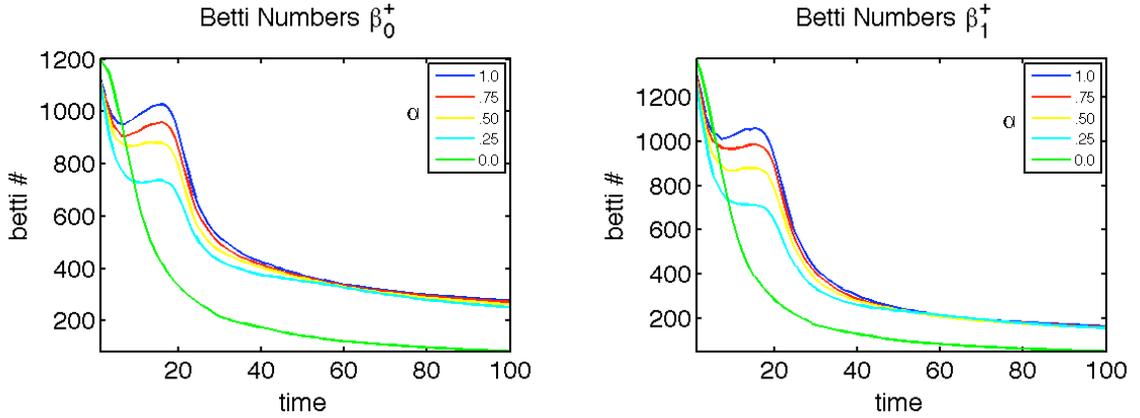


Figure 1.10: Betti Numbers for a Simulation of the Viscous Cahn-Hilliard-Cook model.

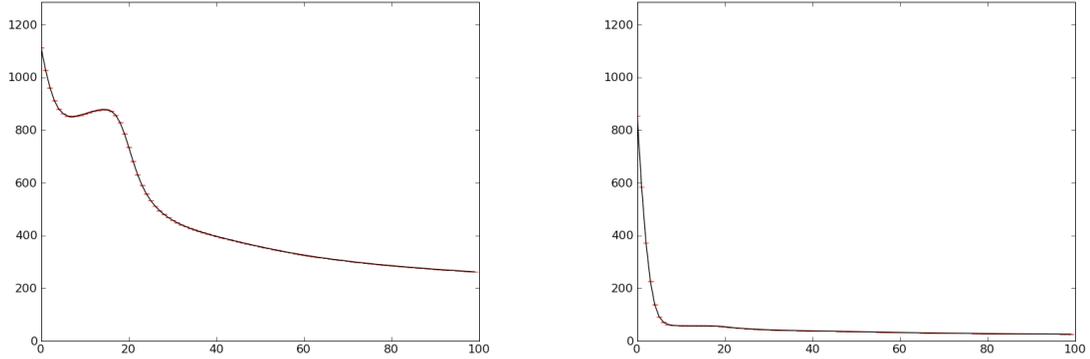


Figure 1.11: Effect of noise intensity of the Betti numbers for Cahn-Hilliard-Cook.

equation. The left image is for a small noise intensity, $\sigma = .0005$, and the right image is for a large noise intensity, $\sigma = .05$. This figure shows that noise has a pronounced influence on pattern formation.

1.6 Previous Work

In this chapter, we have presented an overview of homology and computation of homology, probability and random fields, and given examples of used homology to study the nodal domains of the viscous Cahn-Hilliard equation. The main goal of this dissertation is to provide explicit measures for the correctness of homology computations for nodal domains

of random fields. This final section presents a recap of previous work for this goal.

The first major work in this area was in [62]. The authors propose a stochastic algorithm to randomly sample a topological manifold $X \subset \mathbb{R}^n$ to compute the homology. A random sample of M points is taken and the authors derive explicit probability bounds for the correct homology. The probability bounds depend on M and a parameter called the condition number $1/\tau$. The parameter τ depends on the local curvature of X and other properties related to the separation of X . The quantity $1/\tau$ is the largest number such that the open bundle of radius r is embedded in \mathbb{R}^n .

The quantity $1/\tau$ is particularly difficult to compute for the applications we have in mind. Also, since we are interested in the time evolution of patterns, it is conceivable that the topology of the nodal domains will change. As the topology changes, $1/\tau$ becomes unbounded.

In [53], the authors present a different approach to this problem. Instead of randomly sampling a manifold, the authors deterministically sample random manifolds. The random manifolds are the sets we are interested in, nodal domains. In this work, the authors consider one and two dimensional nodal domains. An equi-distance partition of M or M^2 intervals or squares is formed. The authors then derive explicit probability bounds on the correctness of the homology of the nodal domains.

In this work, there are a few restrictions. The nodal domains must be the zero level sets for homogenous random fields. Further, the random fields must be Gaussian.

A generalization of this work is found in [54]. In this work, the authors have the same goal in mind: to derive explicit probability bounds for the correctness of homology computations of random fields. The authors loosen the assumption that the nodal domains must be zero level sets. In particular, the authors are able to derive bounds for generalized nodal domains of the form $N_\mu^\pm = \{\pm(u - \mu) \geq 0\}$, where μ is a threshold function.

The authors also weaken the assumption that the random fields must be homogenous. As a consequence, the sampled points must be chosen carefully. Explicit characterizations for the sampling of these points is given. This work was restricted to generalized nodal domains

in one dimension.

While the above references give probability bounds for the correctness, another approach is possible. In [30], the authors devise an algorithm that gives the correct homology for nodal domains. This algorithm uses interval arithmetic and provides a computer assisted proof of the correctness. However, the original algorithm presented has a number of shortcomings. Issues related to the use of interval arithmetic restricted the applications that could be considered. A few issues are long computational times and sampling of points that are too close to a zero.

In this dissertation, I will make extensions to the above. The next chapter gives modifications to the validated homology algorithm. This results in short computational times, smaller complexes for computing the homology, and alleviates grid alignment issues for patterns that evolve as a function of time. Chapter Three rigorously presents the existing probability bounds for the correctness. This recaps work done in [53], [54], and the references therein. Chapter Four then extends these results to non-homogenous random fields in two spatial dimensions. We derive explicit probability bounds for making an incorrect homology computation on each box in a non-uniform decomposition of the domain. Chapter Five then presents an algorithm to minimize the probability for making an incorrect homology computation by determining the optimal location of where to sample. Chapter Six then presents results of all this work.

Chapter 2: Guaranteed Homology Computations

2.1 Overview

We now present an algorithm that correctly computes the homology of the nodal domains N^\pm . This will provide a computer assisted proof for the correctness of $H_*(N^\pm)$ which uses interval arithmetic. It tests what happens on the vertices and grid lines to determine the topology. If we cannot determine what happens in the interior and the grid lines, we subdivide the grid and determine what happens on each new grid line. The precise algorithm will be presented in the next few sections. The material from the first section can be found in [30]. The authors devise the original algorithm. However, this algorithm has a few shortcomings that can lead to frequent failures. Once the original algorithm is formulated, we will present the shortcomings and address methods to improve the algorithm in subsequent sections in this chapter.

2.2 Previous Algorithm

In this section, we present an overview of the previous algorithm. This was given in [30]. In the next section, we will present new modifications to this algorithm which provide decreased computational time and smaller grids on which we can compute the homology. In order to design a computer assisted proof for the correct homology, we rely on interval arithmetic. This is a generalization of ordinary arithmetic of numbers that allows us to perform arithmetic on intervals and Cartesian products of intervals. For the intervals $[a, b]$ and $[c, d]$, define

$$[a, b] \oplus [c, d] := [a + c, b + d].$$

$$[a, b] \ominus [c, d] := [a - c, b - d],$$

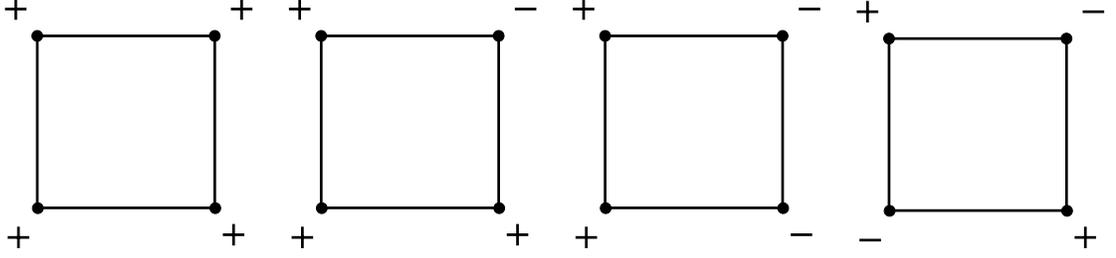


Figure 2.1: Possible Sign Structures for Vertices.

$$[a, b] \otimes [c, d] := [\min\{ac, ad, bc, db\}, \max\{ac, ad, bc, bd\}],$$

$$[a, b] \circ [c, d] := [\min\{a/c, a/d, b/c, d/b\}, \max\{a/c, a/d, b/c, b/d\}],$$

where $0 \notin [c, d]$ for \circ . Using these definitions, we now define interval evaluations for functions by

$$f([a, b]) := \{f(x) : x \in [a, b]\}.$$

Computationally, we will use outward rounding to obtain an interval $[\bar{c}, \bar{d}]$, from which we know $f([a, b]) = [c, d] \subset [\bar{c}, \bar{d}]$. This gives us an interval which contains the range of the function.

We will use interval arithmetic for the computations in our validation scheme. Given a domain $D = [a_1, b_1] \times [a_2, b_2]$, and a function $f : D \rightarrow \mathbb{R}$, we want to determine the topology of $N^\pm = \{x \in D : \pm f(x) > 0\}$. Without loss of generality, we use the algorithm to determine N^+ . If we can obtain a grid that accurately represents N^+ , then we are guaranteed the grid also can resolve the topology of N^- . The first step is to make an initial subdivision of D into a uniform grid of any size. For each box in the grid, compute the sign of the function at each vertex of the box. Thus at each vertex v , a small interval $\widetilde{f(v)}$ is computed in which the actual function value lies. If $\widetilde{f(v)}$ does not contain zero, we have correctly computed the sign of the vertex. The algorithm automatically fails if we cannot determine the sign of a vertex.

For each box B in the grid, a *verification step* is performed to determine the topology of

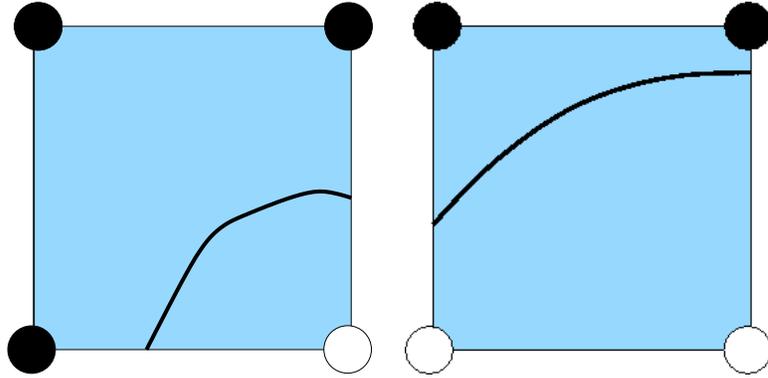


Figure 2.2: Possible Nodal Lines with Given Sign Structures.

$N^+ \cap B$. This step depends on the sign configuration of the box vertices. Up to rotation and negation, we have the 4 possible sign configurations in Figure 2.1. For each sign configuration, we want to ensure the nodal lines are simple. In particular, we want to ensure that the nodal lines stretch to the boundary of the box. If the verification step fails, we then subdivide B into smaller boxes and recursively perform the verification step on each smaller box. We continue this process until all boxes in this grid pass the verification step or the grid is refined below a certain specified threshold.

We handle each of the four sign configurations separately.

For the first case a.), we have all the same sign on the vertices on a box B . We must verify that $N^+ \cap B = B$. This step uses the first lemma below. Given this sign configuration, we test whether or not B has any region of opposite sign. Thus we know the topology is represented by the vertices alone.

For the next two cases, the sign structure indicates that a sign change must occur in the box B . We need to ensure this sign change is simple and does not occur more than once. If this happened, we could not guarantee the topology. This indicates we want to ensure the topology of $\partial N^+ \cap B$ looks like one of the pictures in Figure 2.2. In Case b.), we try to ensure that f is monotone in the x-direction and y-direction on B . For Case c.), we try to ensure that f is bounded away from 0 on the top and bottom edges of B and then test to see if f is monotone in the y-direction. These tests are designed to exclude regions of opposite sign. If the function is monotone in the appropriate directions, there cannot be

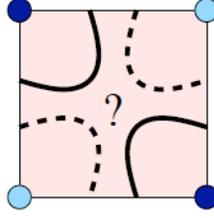


Figure 2.3: Problem with Nodal Lines for Last Sign Configuration

any regions of opposite sign that are strictly enclosed in B . While testing for monotonicity can be too much, it is the only way to ensure a region of opposite sign is strictly enclosed in the box B .

For Case d.), the sign structure cannot resolve the topology. If we examine Figure 2.3, it should be clear that this sign structure cannot give us information about which region is connected in the box. The positive nodal domain may be connected or the negative nodal domain may be connected. This sign structure alone cannot tell this information. For this case, the verification step fails and the box must be subdivided.

To perform the verification step, we need to check conditions that ensure the topology is captured by the sign structure of the vertices. We check that the function is either bounded away from zero or is monotonic along appropriate rays. To perform these checks, we employ the following lemmas and definitions from [30].

Lemma 2.1. [30] *Suppose the function $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) . Then*

$$f([a, b]) \subset f\left(\frac{a+b}{2}\right) \oplus \frac{b-a}{2} \otimes f'([a, b]) \otimes [-1, 1]. \quad (2.1)$$

Corollary 2.1. [30] *Let $D = [a_1, b_1] \times [a_2, b_2]$ and $f : D \rightarrow \mathbb{R}$ be a C^1 function. If*

$$0 \notin f\left(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}\right) \oplus \frac{b_1-a_1}{2} \otimes f_x(D) \otimes [-1, 1] \oplus \frac{b_2-a_2}{2} \otimes f_y(D) \otimes [-1, 1] \quad (2.2)$$

then f is bounded away from 0 on D .

Definition 2.1. Let $D = [a_1, b_1] \times [a_2, b_2]$. We say $f : D \rightarrow \mathbb{R}$ is monotone in the x -direction if $0 \notin f_x([a_1, b_1], y)$ for each $y \in [a_2, b_2]$. Also, $f : D \rightarrow \mathbb{R}$ is monotone in the y -direction if $0 \notin f_y(x, [a_2, b_2])$ for each $x \in [a_1, b_1]$.

Corollary 2.2. [30] Let $D = [a_1, b_1] \times [a_2, b_2]$ and $f : D \rightarrow \mathbb{R}$ be a C^2 function. If

$$0 \notin f_x\left(\frac{a_1 + b_1}{2}, [a_2, b_2]\right) \oplus \frac{b_1 - a_1}{2} \otimes f_{xx}(D) \otimes [-1, 1], \quad (2.3)$$

then f is monotone in the x -direction. Also, if

$$0 \notin f_y\left([a_1, b_1], \frac{a_2 + b_2}{2}\right) \oplus \frac{b_2 - a_2}{2} \otimes f_{yy}(D) \otimes [-1, 1], \quad (2.4)$$

then f is monotone in the y -direction.

We use these lemmas and corollaries in the algorithm in order to obtain bounds on the function and derivatives. The first corollary is used for the sign structure that has all the same sign. The next corollary is used for the remaining two cases. This tests whether the function is monotone in either the x or y direction and that the nodal lines are simple enough to resolve the topology.

If this process terminates successfully, we obtain a nonuniform grid decomposition of the topology. We choose the smallest size box in the grid, namely we choose a box of size $\frac{1}{M}$. We then decompose D into a uniform grid of boxes of size $\frac{1}{M}$, where M becomes the discretization number. We can then use standard computational homology techniques to determine the topology.

2.3 A Randomized Subdivision Process

If the original algorithm terminates successfully, we have a grid on which we can compute the homology and are guaranteed it is the correct homology. However, the original algorithm has a few shortcomings.

The first shortcoming comes from the use of interval arithmetic. Since we use interval arithmetic to find interval bounds on the function and its derivatives, we want the actual range of these bounds. However, since interval arithmetic uses outward rounding for each computation, the ranges are often over extended. This results in both lengthy computation times and also frequent failure of the algorithm. The second shortcoming is more subtle. For time varying patterns, the nodal lines will of course evolve with time. As these lines evolve, they can come arbitrarily close to the dyadic subdivision lines in the grid decomposition which unfortunately can lead to failure of the algorithm. As a result, more subdivisions must be performed to resolve the topology. The other alternative is to use different initial decompositions of the domain.

2.3.1 Range Enclosure

The first problem we address is enclosing the range. In many of the applications we have in mind, we are interested in the homology of the nodal domains for random sums of the form

$$f(x, y, \omega) = \sum_{j=1}^N \sum_{k=1}^N a_{j,k} g_{j,k}(\omega) \phi_j(x) \psi_k(y),$$

for an integer N , specified coefficients $a_{j,k}$, $g_{j,k}(\omega)$ independent, identically distributed normal random variables, and ϕ_j and ψ_k basis functions. Since we are dealing with sums, interval arithmetic uses outward rounding arithmetic for each addition. The repeated outward rounding produces ranges that are often extended far beyond their true range. This results in using far more subdivisions than are actually needed and thus a much finer grid.

In this section, we present an algorithm that will test for upper and lower bounds by enclosing the range.

The original algorithm is known as the Skelboe-Moore algorithm. It was first presented in [67]. This algorithm uses a range tolerance with branch and bound techniques to eliminate regions of the domain where the lower bound cannot occur. Before proceeding, we need to understand interval arithmetic better. For an interval $X = [\underline{x}, \bar{x}]$, the width of X is $w(X) = \bar{x} - \underline{x}$, the center of X is $c(X) = (\underline{x} + \bar{x})/2$. If $f(x_1, \dots, x_n)$ is a function with real valued variables x_1, \dots, x_n , then the interval function $F(X_1, \dots, X_n)$ associated to f is formed by replacing the real valued variables x_1, \dots, x_n by the intervals X_1, \dots, X_n and replacing the ordinary arithmetic operations with their corresponding interval valued arithmetic operations. The interval valued function F is known as the interval extension of f . However, the quantity we are interested in is not F , but the function

$$\bar{F}(X_1, \dots, X_n) := \{f(x_1, \dots, x_n) : x_i \in X_i, i = 1, \dots, n\}.$$

The result of this computation will give us the actual range of our function f . It should be noted that

$$\bar{F}(X_1, \dots, X_n) \subset F(X_1, \dots, X_n).$$

With these definitions, Skelboe in [67] showed the following theorems are true.

Theorem 2.1. [67] *Let $F(X_1, \dots, X_n)$ be an interval extension of $f(x_1, \dots, x_n)$. Suppose*

$$X'_1 \subset X_1, X'_2 \subset X_2, \dots, X'_n \subset X_n$$

Then

$$f(X'_1, X'_2, \dots, X'_n) \subset f(X_1, X_2, \dots, X_n).$$

Theorem 2.2. [67] Let $F(X_1, \dots, X_n)$ be an interval extension of $f(x_1, \dots, x_n)$. If $X_j = \bigcup_{i=1}^N X_j^{(i)}$, then

$$f(X_1, X_2, \dots, X_j, \dots, X_n) = \bigcup_{i=1}^N f(X_1, X_2, \dots, X_j^{(i)}, \dots, X_n).$$

Theorem 2.3. [67] Let $F(X_1, \dots, X_n)$ be an interval extension of $f(x_1, \dots, x_n)$. Subdivide the intervals X_1, \dots, X_n such that

$$X_j = \bigcup_{i=1}^N X_j^{(i)} \text{ with } w(X_j^{(i)}) = \frac{1}{N} w(X_j).$$

Then there is a positive number k such that

$$\bigcup_{i_1=1}^N \dots \bigcup_{i_n=1}^N f(X_1^{(i_1)}, \dots, X_n^{(i_n)}) = \bar{f}(X_1, \dots, X_n) + E_n,$$

where

$$w(E_n) \leq \frac{k}{N^2} \max_{j=1, \dots, n} (w(X_j)^2).$$

Theorem 2.3 will be important for the algorithm. Without loss of generality, suppose we are interested in finding a lower bound for \bar{F} . Perform an initial subdivision $X \rightarrow X_2^{(1)}, X_2^{(2)}$ by dividing X in half. Thus we have $w(X_2^{(1)}) = w(X_2^{(2)}) = 1/2w(X)$. Because of Theorem 2.2, we have

$$\bigcup_{i=1}^2 F(X_2^{(i)}) \subset F(X).$$

More importantly, we know that

$$\overline{F}(X) \subset \bigcup_{i=1}^2 F(X_2^{(i)}).$$

We can subdivide each of $X_2^{(1)}$ and $X_2^{(2)}$ to get four intervals of equal length, then form the union of the interval functions and we know that the true range is contained in this new set. However, instead of taking repeated subdivisions, we can expedite this process significantly. Since we are interested in finding the true range $[z_1, z_2]$ of our function f , we will find a lower bound \bar{z}_1 and an upper bound \bar{z}_2 . We will then drive the lower bound upward and drive the upper bound down. This will enclose the true range and will allow us to get arbitrarily close to the actual range. We will accomplish this goal by isolating out possible regions where the lower bounds and upper bounds can occur. We will first work on finding a tight lower bound and then a tight upper bound.

On the subdivision of $X \rightarrow X_1^{(1)}, X_2^{(2)}$, we take the lower bounds of each new interval. Next form a list \mathcal{L} of the new intervals and arrange them in increasing order according to the lowest bound.

We now subdivide again. However, instead of subdividing both intervals, we simply remove the first element in \mathcal{L} and subdivide the interval in half, thus obtaining two new intervals. From this subdivision, we obtain two new intervals. We evaluate the interval function on each interval and once again add these intervals to the list \mathcal{L} according to the lowest bound. Now take the first element in the list \mathcal{L} , remove it from \mathcal{L} , subdivide it in half and evaluate the interval function on each new interval. Put both intervals in \mathcal{L} according to the lowest bound. We keep doing this process until some tolerance is met.

Why does this method work? By Theorem 2.2, we know the true range will be contained in the union of these interval functions. By removing the first element in the list \mathcal{L} , subdividing it in half, and arranging the elements in \mathcal{L} , we are effectively discarding regions of the domain that are unlikely to contain the lower bound of our function \overline{F} . To find an upper bound,

we repeat the same procedure but arrange the list with the largest upper bound.

This was for an interval function of only one variable. We can do almost the same process with one modification. The first initial subdivision is done along the edge with longer side length. We then arrange the boxes in the list according to the lowest bound. Remove the first element from the list. To subdivide this new box, we subdivide along the longest side length and then add these boxes to the list. We keep doing this procedure until a lower bound is found within some specified tolerance.

While the above algorithm finds a lower bound for the interval function, it is too exhaustive for our purposes. This form of the algorithm has no increase in performance or decrease in computational times. In fact, it often has a longer computational time than the original validated homology algorithm. This can be attributed to establishing both an upper and a lower bound to a pre-specified tolerance. However, in our validated homology algorithm, we do not need precise upper and lower bounds. For example, suppose we are working with a box that is to be validated. Also, suppose the vertices all have a positive sign. Then we do not need an upper bound; we only need a lower bound for this case. In addition, we do not even need a sharp lower bound. We only need to ensure that the lower bound is itself positive. We can handle the other sign configurations with similar considerations. If we have a box with only one opposite sign, then we only need to ensure it is monotone in the appropriate directions. However, this is the same as testing for an upper or lower bound on the partial derivatives.

We now present the formal algorithm we have used. Again, without loss of generality assume we are trying to find a lower bound on the box B that is positive. If we need to find an upper bound that is negative, we can simply replace f with $-f$.

Set f_l equal to an upper bound of $\min f(c(B))$. This should be a rigorous upper bound on the lower bounds. Our goal will be to drive this upper bound on lower bound downward. Create 2 empty lists \mathcal{L} and \mathcal{C} and assume $\epsilon > 0$ is a specified tolerance. The list \mathcal{L} will contain an ordered list of the possible lower bounds of f . The list is ordered so that the first element is the smallest lower bound. The list \mathcal{C} will contain boxes that contain the lower

bound which are within a specified tolerance. We will think of this list as the confirmed list of lower bounds.

Bisect B along the direction of longer length to get 2 new boxes B_1 and B_2 . Set $f_l = \min \{f_l, \overline{f(B_1)}, \overline{f(B_2)}\}$, where we define $\overline{f(B)}$ to be the right endpoint of $f(B)$ using outward rounding interval arithmetic and also $\underline{f(B)}$ to be the left endpoint of $f(B)$ using outward rounding arithmetic. The value f_l will be an upper bound on the true lower bound that we seek. If

$$\max \{ \overline{f(B_1)}, \overline{f(B_2)} \} - \min \{ \underline{f(B_1)}, \underline{f(B_2)} \} < \epsilon,$$

then put B_1 and B_2 into the list \mathcal{C} in order. This condition specifies that the width of the lower bounds for boxes in \mathcal{C} are within ϵ . If $\mathcal{L} \neq \emptyset$, then remove the first item from \mathcal{L} and set this as the new B . If $\underline{f(B)} > f_l$, then return the lower bound to be f_l . If $\mathcal{L} = \emptyset$, then return with the lower bound from the first box in \mathcal{C} .

If on the other hand, we have $\max \{ \overline{f(B_1)}, \overline{f(B_2)} \} - \min \{ \underline{f(B_1)}, \underline{f(B_2)} \} \geq \epsilon$, then put $(B_1, f(B_1))$ and $(B_2, f(B_2))$ in order into the list \mathcal{L} . Set B equal to the box from the first element in \mathcal{L} and remove from the list. If \mathcal{L} is not empty, then go back to the bisection step and repeat.

For each bisection, we are driving the upper bound f_l for the lower bound downwards and then comparing this with the computed lower bound. If at any step in the bisection, we compute the lower bound and it is positive then we have correctly computed the the sign of the box. This check results in a tremendous computational speedup in comparison to the original Skelboe-Moore algorithm. It also results in a speedup with regards to the original validation algorithm.

2.3.2 Randomized Subdivisions

For time varying patterns, the nodal lines will evolve. As these lines move, they can come arbitrarily close to the grid lines in the subdivision grid. As a result, the algorithm must

then subdivide again in order to resolve the topology. This creates a tremendous bottle-neck in the computation. In order to alleviate this problem, the authors in [30], took different initial subdivisions. If the nodal line comes close to a grid line with one initial subdivision, then for a different initial subdivision the nodal line may be far enough away to resolve the topology. The authors took initial subdivisions of $M_0 = 2, 3, 5, 7,$ and 9 . This repeated process creates a tremendous computational effort in order to capture the topology.

We have proposed an alternative to taking different initial subdivisions. Without loss of generality, assume our domain is D . For the new algorithm, we do not take an initial decomposition of our domain. We then proceed as before and run the verification steps on the original domain and keep subdividing as needed. For each box in the grid, we run the verification steps with our new range enclosure algorithm. If the box cannot be verified, we must subdivide it into new boxes that must be verified. However, in the subdivision step, we now only subdivide in the coordinate direction which has longer edge length. If the side lengths are equal, we randomly choose which direction to subdivide.

Let $\alpha_1 < \alpha_2$ be two fixed ratios with $\alpha_1, \alpha_2 \in (0, 1)$. If a box $B = [a, b] \times [c, d]$ cannot be verified, we then randomly choose a ratio and subdivide the box along the edge with longer side length into two boxes according to which ratio was chosen. For a rectangle, define the eccentricity to be the ratio of the longer edge length to the smaller edge length. Denote this as χ . One of the benefits of choosing between fixed ratios is that the eccentricities of all the boxes in the grid will remain bounded. For our simulations, we choose to use the Golden ratio. That is, we may randomly subdivide with the ratios $\alpha = \frac{\sqrt{5}-1}{2}$ and also $1 - \alpha = \frac{3-\sqrt{5}}{2}$. Looking at Figure 2.4, shows that the boxes indeed remain bounded but it also appears that the boxes only can have a few possible eccentricities. Using the Golden Ratio, we may state the following lemma.

Lemma 2.2. *Suppose the algorithm is started on a box of equal side lengths. If the two ratios possible are the Golden Ratio α and $1 - \alpha$, then the only possible eccentricities are $1, 1/\alpha$ and $1/(1 - \alpha)$.*

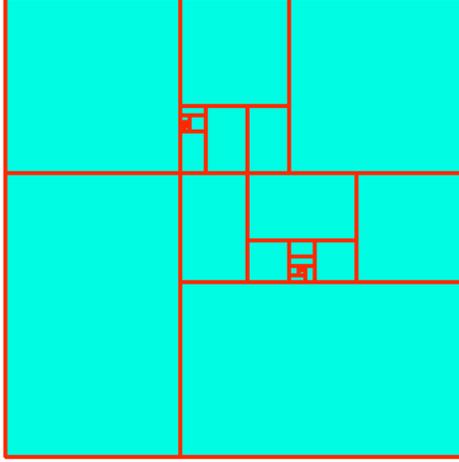


Figure 2.4: Example boxes of validated homology routine using the Golden Ratio.

Proof. To begin, we use $\alpha > 0$, such that α solves $\alpha^2 = 1 - \alpha$, which is $\alpha = \frac{\sqrt{5}-1}{2}$. From this, the two ratios we have are $0 < 1 - \alpha < \alpha < 1$. Without loss of generality, assume we have the domain is $D = [0, 1] \times [0, 1]$. The initial eccentricity is $\chi_0 = 1$. After the first subdivision we have two eccentricities:

$$\chi_1 = \frac{1}{\alpha}$$

$$\chi_1 = \frac{1}{1 - \alpha}.$$

For the next subdivision, we must work with each eccentricity separately.

For $\chi_1 = \frac{1}{\alpha}$, the eccentricities for the next subdivision are

$$\chi_2 = 1$$

$$\chi_2 = \frac{\alpha}{1 - \alpha} = \frac{1}{\alpha},$$

where we have used $\alpha^2 = 1 - \alpha$. For $\chi_1 = \frac{1}{1 - \alpha}$, the eccentricities for the next subdivision

Table 2.1: Eccentricities of boxes for validated homology using the Golden Ratio

Eccentricity	Numeric Value
1	1
$\frac{1}{\alpha}$	$\frac{2}{\sqrt{5}-1}$
$\frac{1}{1-\alpha}$	$\frac{2}{3-\sqrt{5}}$

are

$$\chi_2 = 1$$

$$\chi_2 = \frac{\alpha}{1-\alpha} = \frac{1}{\alpha}.$$

This finishes the proof since we have found what happens to all possible eccentricities. In particular, for the eccentricity $\chi_k = 1$, the only eccentricities for the next subdivision are $\chi_{k+1} = \frac{1}{\alpha}$ or $\chi_{k+1} = \frac{1}{1-\alpha}$. If $\chi_k = \frac{1}{\alpha}$, the only possible eccentricities for the next subdivision are $\chi_{k+1} = 1$ or $\chi_{k+1} = \frac{1}{\alpha}$. For $\chi_k = \frac{1}{1-\alpha}$, the eccentricities for the next subdivision must be $\chi_{k+1} = 1$ or $\chi_{k+1} = \frac{1}{\alpha}$. This furnishes a bounded sequence of eccentricities. \square

Using the Golden Ratio on a domain with eccentricity one, the only possible eccentricities of the subdivided boxes are shown in Table 2.1.

2.4 Numerical Experiments

In this section, we will test the performance of the new algorithm by applying it to specific situations. These situations were considered in [30].

2.4.1 Double Well Function

The first test case we consider is a double well function. We will concentrate on rotations of the function

$$H_C(x, y) = \frac{1}{2}x^2 - \frac{1}{4}x^4 - \frac{1}{2}y^2 + C. \quad (2.5)$$

For $C \leq -1/4$, the positive nodal domain is empty. For $-1/4 \leq C \leq 0$, the positive nodal domain consists of two disks which merge as $C \rightarrow 0^-$ and contract to two points as $C \rightarrow -1/4$. When $C \geq 0$, the positive nodal domain is connected. As mentioned in [30], at the values $C = -1/4$ and $C = 0$, the algorithm cannot resolve the topology and thus fails.

In order to determine the performance of the algorithm in this setting, we will not work with H_C directly. Instead, we consider the nodal domains of H_C rotated through a random angle $\theta \in [0, 2\pi)$ about the point

$$r_c = \left(\frac{3\sqrt{3}}{10}, \frac{2\sqrt{2}}{5} \right).$$

Namely, we consider the function

$$H_{C,\theta}(x, y) = H_C(5R_\theta^{-1}((x, y) - r_c)^t), \quad R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

A few images of the nodal domains are shown in Figure 2.5. The top images are for $C \approx -1/4 + 10^{-8}$ and the bottom images are for $C \approx -1/4 + 10^{-3}$. The left images show the nodal domains, the middle images show the grids resulting from the original validation algorithm, and the right images show the grids result from our new modified algorithm. Since the algorithm will fail for $C = 0, -1/4$, we test our algorithm for perturbations around these critical values, i.e., we consider

$$C = c_0 + c_s \gamma, \quad \gamma = 2^{-k}, \quad k = 1, \dots, 49, \quad c_0 = 0, -1/4, \quad c_s = \pm 1$$

For each of these values, we choose 5000 random angles θ from a uniform distribution and apply the algorithm to the unit square $[0, 1]^2$. We record the number of boxes, the number of calls to the verification function, the number of interval evaluations, and the logarithm

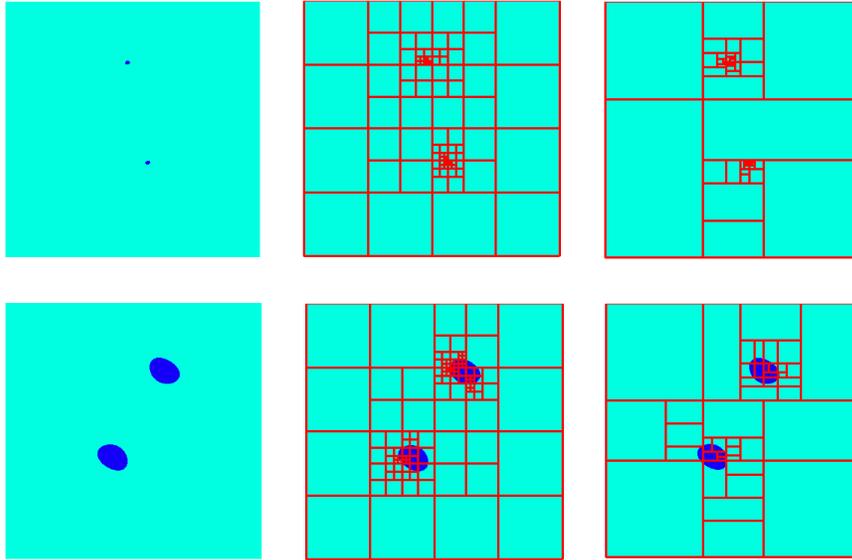


Figure 2.5: Sample Grids From Validation Algorithm

of the discretization size M . The discretization size is now defined to be the smallest edge length of the boxes in the final decomposition. After the runs are complete, we then average each value. We have plotted the key parameters in Figure 2.6, where the dependence is on the absolute value $\gamma = |C - c_0|$. The solid blue, dashed green, and solid red curves correspond to the values $C = -1/4$, $C = 0^-$, and $C = 0^+$, respectively.

Comparing these key parameters to the original algorithm in [30], we see the most dramatic effect in the number of boxes in the final grid. The key parameters for the original algorithm are shown in Figure 2.7. The original algorithm required around 350 boxes in the final grid for $C \approx 0$, while the new algorithm only requires 200 boxes. By enclosing the range on each box, we are able to capture the topology with far fewer boxes. The remaining two key parameters are the number of calls to the validated rectangle function in our code and also the logarithm of the discretization size M . We note here that the discretization size M in our new algorithm is taken to be the smallest side length of all the boxes in our algorithm.

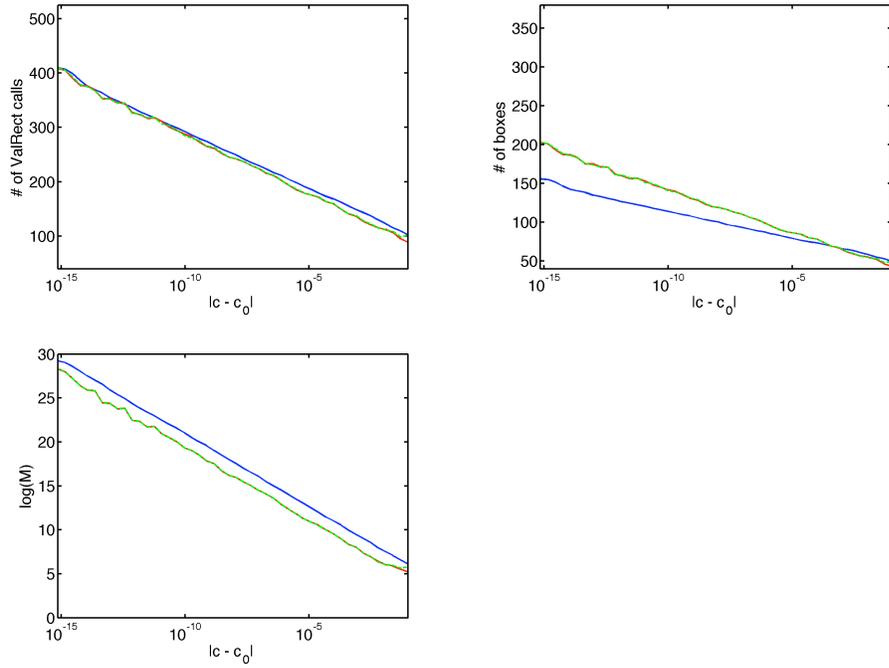


Figure 2.6: Key Parameters for Double Well Function Using New Algorithm.

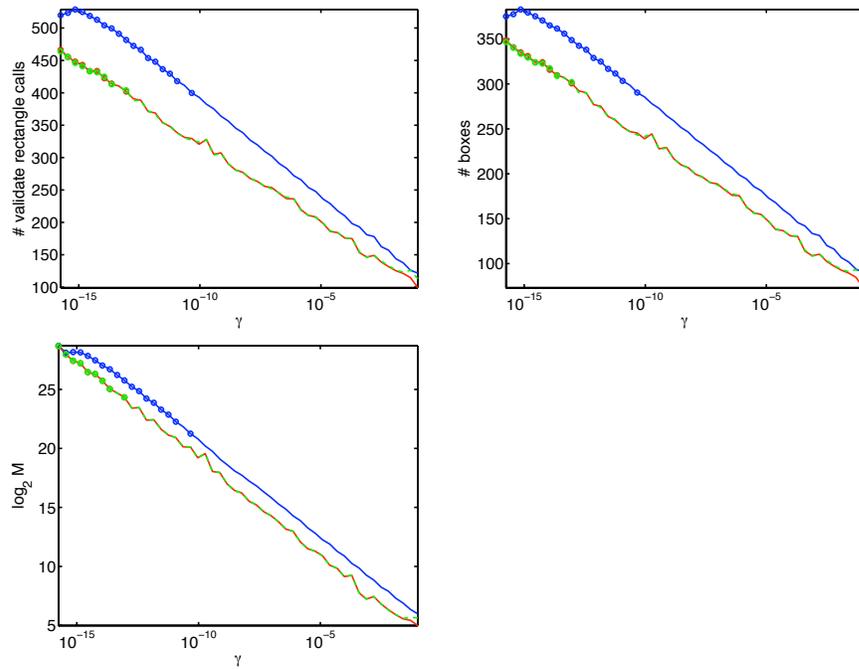


Figure 2.7: [30] Key Parameters for Double Well Function Using Original Algorithm.

Chapter 3: Previous Probabilistic Bounds

3.1 Introduction

The previous chapter gave an algorithm that correctly computes the homology of N^\pm . While the new modifications decrease the computational time, alleviate grid alignment issues, and reduce the number of boxes to be validated, it still suffers one bottleneck. Since our interest lies in computing the nodal domains for finite sums, the interval ranges expand far beyond their true range. With the range enclosure we can get precise ranges or determine bounds for the range, small computational errors in floating point arithmetic quickly propagate. These small errors are magnified by interval arithmetic. For random sums of the form

$$f(x, y, \omega) = \sum_{j=0}^N \sum_{k=0}^N a_{j,k} g_{j,k}(\omega) \phi_j(x) \psi_k(y)$$

the algorithm is infeasible for large N . In fact, for multiple simulations it is still not practical to do values larger than $N = 20$.

In this chapter and the next, we present a different approach. Instead of guaranteeing the homology is correct, we try and give tight probability bounds for finding the correct homology with a fixed discretization size M . In the process, we will also devise an algorithm that determines the optimal placement of the points to be sampled. The remainder of this chapter is devoted to giving previous bounds. Most of the following can be found in [53], [54], [29], and [30]. In [53], the authors present probabilistic bounds for the correctness of such homology computations in one and two spatial dimensions. However, their work is limited to homogenous random fields. In [54], the authors extend their results to non-homogenous random fields in one dimension. In this chapter, we present results from both

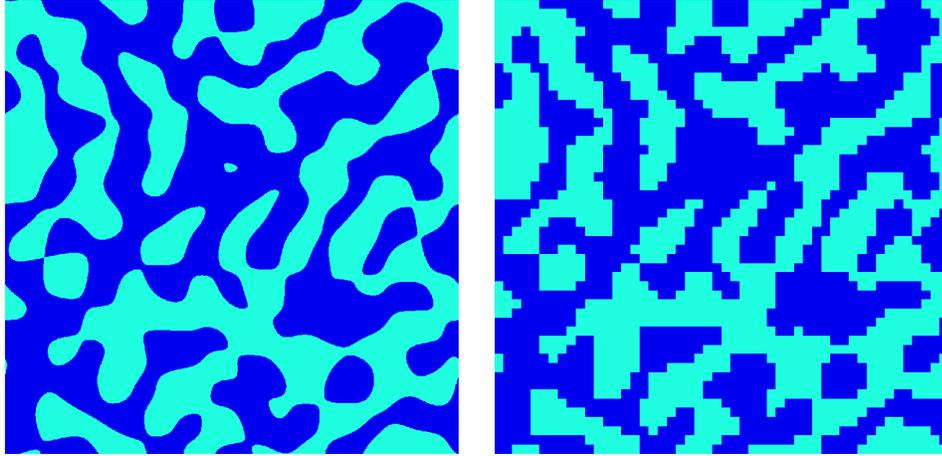


Figure 3.1: Correct and Incorrect Cubical Approximation of the Nodal Domain

of these and the next chapter will extend these results to two-dimensional, non-homogenous random fields.

Suppose we have a realization of a random field $f(\cdot, \omega) : D \rightarrow \mathbb{R}$. We are interested in the nodal domains of $f(\cdot, \omega)$. We are particularly interested in computing the homology of these sets. Since we must discretize our domains, what can go wrong? In particular, if we discretize our domain into M pieces in one dimension or M^2 pieces in two dimensions, are we computing the correct homology? Does this approximation accurately represent the homology of N^\pm ? For example, suppose we have a square domain and use equi-spaced points to decompose the domain. Using this decomposition, we test the sign of the lower left vertex in each box and depending on the sign, add the box to either N^+ or N^- . If we look at Figure 3.1, we can see that we have not sampled enough points to determine the homology accurately. If we examine the images, we see that there are regions in the true nodal regions that are connected. When we discretize the nodal domains, some of these regions become disconnected. Also upon inspection, there is an island of one component of N^\pm that is entirely lost with this approximation. This will of course give a different topology. First, we formalize the sets we shall be working with. Given a domain $D = [a, b]^d$, let $u : D \rightarrow \mathbb{R}$.

Definition 3.1. *The generalized nodal domains of u are given by the two sets*

$$N_\mu^\pm := \{x \in D \mid \pm (u(x) - \mu(x)) \geq 0\},$$

for a continuous function $\mu : D \rightarrow \mathbb{R}$.

In any computational scheme, these nodal domains must be discretized in some appropriate manner. For all the following probability bounds, we use cubical approximations and cubical homology.

Definition 3.2. *Let $u : [a, b]^d \subset \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous and M a positive integer. Define the equidistance partition of $[a, b]$ to be the collection of points*

$$x_k = a + k \frac{b - a}{M}$$

for $k = 0, \dots, M$. Then we define the cubical approximations

$$Q_M^\pm := \bigcup \left\{ \prod_{l=1}^d [k_l, k_{l+1}] \mid \pm u(x_{1,k_1}, \dots, x_{d,k_d}) \geq 0 \right\},$$

where $x_{j,0}, \dots, x_{j,M}$ are the equidistant M -discretizations of the j 'th component of D .

Note that while the original domain is $[a, b]^d$, the homology domain will be a subset of \mathbb{Z}^d . This is a requirement of the code for using cubical homology. Chapter 2 and [46] provide a thorough overview of this homology theory and the computational considerations. If we look at Figure 3.2, we see possible differences between the nodal domains and a cubical approximation. In the first figure, the nodal domain for the positive region is shown in the first graph and the negative region just below this plot. In the second image, we randomly select 8 points in our domain to sample, and form the cubical approximations from these points. It is clear that the homologies of these particular Q^\pm do not coincide.

Our ultimate goal is to either ensure the homologies of N^\pm and Q_M^\pm agree or establish a

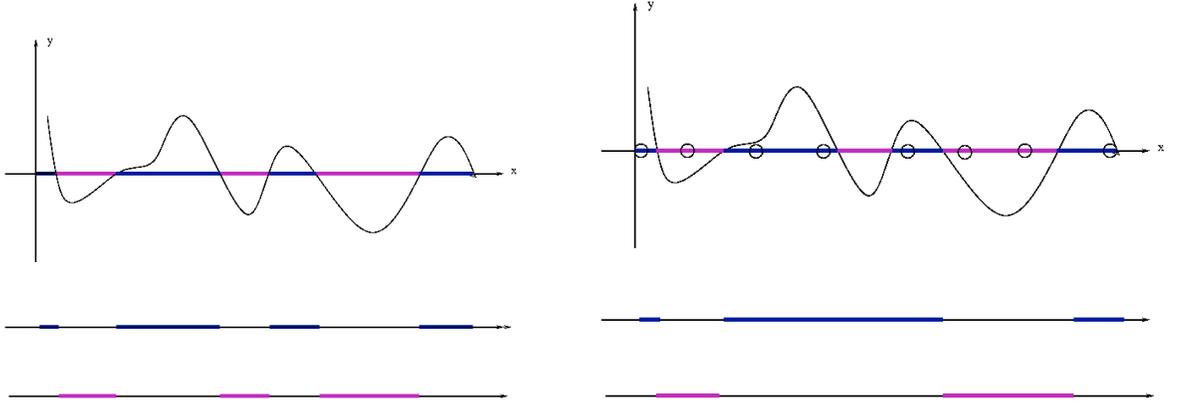


Figure 3.2: Nodal Domains for Sample Random Field

measure of how much they disagree. It is clear that for small M , the homologies can be quite different as can be seen in Figure 3.2. Throughout the remainder, we will make the following assumptions:

(A1) for every $x \in D$, we have $\mathbb{P}\{(u(x, \omega) - \mu(x)) = 0\} = 0$

(A2) we have that $\mathbb{P}\{0 \text{ is a critical value of } u - \mu\} = 0$.

In addition to these assumptions, we also assume the threshold function μ is at least of class C^1 . These assumptions exclude situations with probability one in which the sampled point is equal to zero. If this happens, then we cannot add the box to either of Q_M^\pm . However, with these assumptions, we can make the following statement from [53].

Theorem 3.1. [53] Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a domain $D := [a, b]^d \subset \mathbb{R}^d$ and a random field $u : D \times \Omega \rightarrow \mathbb{R}$ over $(\Omega, \mathcal{F}, \mathbb{P})$ such that for \mathbb{P} -almost all $\omega \in \Omega$, the function $u(\cdot, \omega) : D \rightarrow \mathbb{R}$ is twice continuously differentiable. For each $\omega \in \Omega$, denote the nodal domains by $N^\pm(\omega)$ and the cubical approximations by $Q_M^\pm(\omega)$.

If (A1) and (A2) are satisfied, then for \mathbb{P} -almost all $\omega \in \Omega$, the following holds. For all sufficiently large M , the homology of the cubical approximations Q_M^\pm matches that of N^\pm . In other words, there exists a random variable $\bar{M} : \Omega \rightarrow \mathbb{N}$ such that

$$\mathbb{P}\{\text{for all } M \geq \bar{M}, \text{ one has } H_*(N^\pm) \cong H_*(Q_M^\pm)\} = 1.$$

Note that the random variable \bar{M} is neither constant nor bounded.

This shows that if we take M large enough, then we can be assured that the homologies agree. However, what happens when M is not large enough? What is a large enough value of M ? The value of M is not known a priori, it is a random variable that is associated to a specific realization of the nodal domains. This suggests the following problem. For a given M , find sharp lower bounds on

$$\mathbb{P}\{H_*(N^\pm) \cong H_*(Q_M^\pm)\},$$

and where to sample our points to maximize this probability.

3.2 Homogenous Bounds in One Dimension

To answer the above question, we first restrict ourselves to specific cases. We first investigate one-dimensional, homogenous random fields with zero threshold $\mu(x) := 0$. In this case, it is straightforward to determine when the homologies of N^\pm and Q_M^\pm agree. The sets we wish to investigate are

$$N^\pm = \left\{ x \in [a, b] \mid \pm u(x) \geq 0 \right\}.$$

However computationally, we work with the sets

$$Q_M^\pm = \bigcup \left\{ [k, k+1] \mid \pm u(x_k) \geq 0 \right\}.$$

In this setting, we can easily characterize when the homologies agree. They will coincide if for each $[x_k, x_{k+1}]$ either of the following are true:

- a.) the function u does not have a zero on $[x_k, x_{k+1}]$
- b.) it has one zero and different signs on the endpoints on $[x_k, x_{k+1}]$.

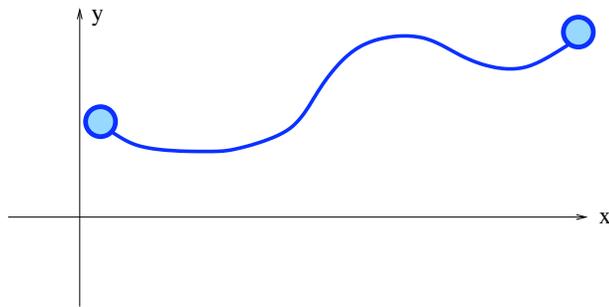


Figure 3.3: Example of a random field with no crossover on $[\alpha, \beta]$.

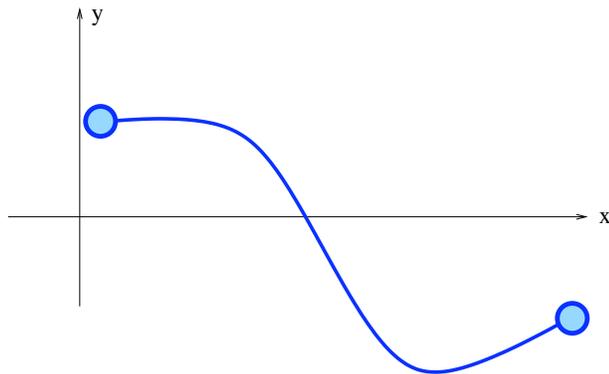


Figure 3.4: Example of a random field that has a single crossover over $[\alpha, \beta]$.

If we look at Figures 3.3 and 3.4, we see intervals which will computationally give us the correct homologies.

The above conditions must hold for every sampled subinterval of D . Suppose a sampled interval is positive on the endpoints and negative at some point in between. Since the function is continuous, we know there is a region that is negative. However, our sampled points are simply the endpoints, so the enclosed negative region is ignored and thus gives the incorrect homology.

Since we are working with random fields, we cannot state whether the function will have a zero in the interior of $[x_k, x_{k+1}]$. To overcome this, we work within the context of a double crossover.

Definition 3.3. *Let $u : [a, b] \rightarrow \mathbb{R}$ be a continuous function and $[\alpha, \beta] \subset [a, b]$. Then u has*

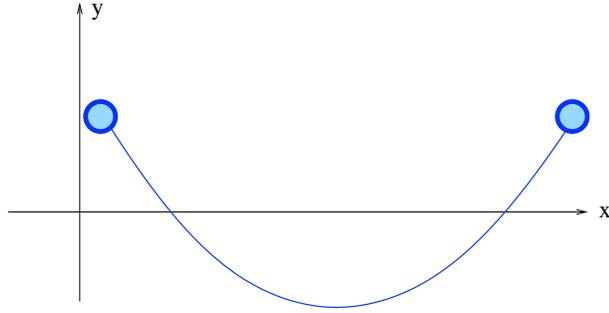


Figure 3.5: Example of Double Crossover

a double crossover on $[\alpha, \beta]$ if

$$\sigma \cdot u(\alpha) \geq 0, \quad \sigma \cdot u\left(\frac{\alpha + \beta}{2}\right) \leq 0, \quad \sigma \cdot u(\beta) \geq 0$$

for $\sigma \in \{\pm 1\}$.

If we have a double crossover on any $[x_k, x_{k+1}]$, then the cubical approximation will ignore a section of the nodal domains. We wish to minimize the probability of having a double crossover on any of the subintervals. To do this, we introduce the following.

Definition 3.4. Let $u : [a, b] \rightarrow \mathbb{R}$ be continuous and $J = [\alpha, \beta] \subset [a, b]$.

- The dyadic points of J are

$$d_{n,k} = \alpha + (\beta - \alpha) \frac{k}{2^n}$$

- The dyadic subintervals of J are then $[d_{n,k}, d_{n,k+1}]$
- We say that the interval J is admissible for the function u if it does not have a double crossover on any of the dyadic subintervals of J .

The existence of a double crossover implies there exist at least two zeros. Thus we can write

$$\mathbb{P}[\text{double crossover on } [\alpha, \beta]] = 1 - \mathbb{P}[\text{less than two zeros on } [\alpha, \beta]].$$

The term on the right may also be written

$$\mathbb{P}[\text{less than two zeros on } [\alpha, \beta]] = \mathbb{P}[\text{one zero on } [\alpha, \beta]] + \mathbb{P}[\text{no zero on } [\alpha, \beta]].$$

By the Intermediate Value Theorem, we can find an upper bound on the probability that each of these terms occur. Thus, we have

$$\mathbb{P}[\text{no double crossover on } [\alpha, \beta]] \leq C,$$

for some number C . However, using this will thus furnish

$$\mathbb{P}[\text{double crossover on } [\alpha, \beta]] \geq 1 - C.$$

However, this is the incorrect bound. Since our ultimate goal is to rule out double crossovers, we want to be able to find an upper bound on the probability for the existence of a double crossover. This method only gives us a lower bound. By working with dyadic subdivisions, we will be able to find an upper bound on the probability of a double crossover on an interval.

With these two notions, we can then formulate the following validation criterion from [53].

Proposition 3.1. *Let $u : [a, b] \rightarrow \mathbb{R}$ be a continuous function and let $M \in \mathbb{N}$ be arbitrary. Let N^\pm be the nodal domains and Q_M^\pm be the cubical approximation. Assume the following hold:*

- a.) *the function u is nonzero at all x_k*
- b.) *the function u has no double zero in (a, b)*
- c.) *each interval $[x_k, x_{k+1}]$ between consecutive discretization points is admissible.*

Then we have

$$H_*(N^\pm) \cong H_*(Q_M^\pm).$$

In practice, the first two conditions are automatically true for Gaussian random fields. Since we are dealing with random functions, we cannot state whether the last condition will always be met. We can however, give upper bounds to the probability of the interval not being admissible. This gives rise to the following theorem.

Theorem 3.2. [53] *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a compact interval $G = [a, b]$ and a random field $u : G \times \Omega \rightarrow \mathbb{R}$ over $(\Omega, \mathcal{F}, \mathbb{P})$ for which the original assumptions (A1) and (A2) hold. Assume u is twice continuously differentiable \mathbb{P} -almost always. In addition, for $x \in G$ and $\delta > 0$ with $x + \delta \in G$, define*

$$p_\sigma(x, \delta) = \mathbb{P} \left\{ \sigma \cdot u(x, \omega) \geq 0, \sigma \cdot u\left(x + \frac{\delta}{2}\right) \leq 0, \sigma \cdot u(x + \delta) \geq 0 \right\},$$

and assume that there is a constant $\mathcal{C}_0 > 0$ such that

$$p_\sigma(x, \delta) \leq \mathcal{C}_0 \delta^3 \quad \text{for all } \sigma \in \{\pm 1\} \text{ and } x \in G \text{ with } x + \delta \in G.$$

Then for every discretization size M , the probability that the homologies of $N^\pm(\omega)$ and $Q_M^\pm(\omega)$ coincide satisfies

$$\mathbb{P}\{H_*(N^\pm) \cong H_*(Q_M^\pm)\} \geq 1 - \frac{8\mathcal{C}_0(b-a)^3}{3M^2}.$$

This is the first result that gives an explicit lower bound on the probability that the homologies of the discretized nodal domains coincide with the true nodal domains. This bound is also supported by numerical evidence which shows the same $\frac{1}{M^2}$ asymptotic behavior. In practice, we must be able to find the constant \mathcal{C}_0 or approximate it.

Random Periodic Functions

As a particular application of the above bounds, we will work with random periodic functions. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $G = [0, L]$ and consider the following random Fourier series $u : G \times \Omega \rightarrow \mathbb{R}$ of the form

$$u(x, \omega) = \sum_{k=0}^{\infty} a_k \left(g_{2k}(\omega) \cos \frac{2\pi kx}{L} + g_{2k-1}(\omega) \sin \frac{2\pi kx}{L} \right).$$

Assume the following are true

- the Gaussian random variables $g_l(\omega)$ are independent and normally distributed with mean zero and variance 1
- the constants a_k are arbitrary such that at least 2 of them do not vanish and

$$\sum_{k=0}^{\infty} a_k^2 k^6 < \infty$$

The above constants a_k are directly related to the smoothness of the random field. Define

$$A_l = \sum_{k=0}^{\infty} a_k^2 k^{2l}.$$

Then we have that

$$\mathbb{E} \|D_x^l u\|_{L^2(0,L)}^2 = (2\pi)^{2l} L^{1-2l} A_l.$$

This amounts to A_l containing the averaged information in the L^2 norm about the l -th derivative over $(0, L)$ of u .

Random Fourier series of this type are convenient to study since they are homogenous random fields. This can be seen through

$$\begin{aligned}
R(x, y) &= \mathbb{E}(u(x)u(y)) \\
&= \mathbb{E} \left\{ \left(\sum_{k=0}^{\infty} a_k \left(g_{2k}(\omega) \cos \frac{2\pi kx}{L} + g_{2k-1}(\omega) \sin \frac{2\pi kx}{L} \right) \right) \right. \\
&\quad \cdot \left. \left(\sum_{l=0}^{\infty} a_l \left(g_{2l}(\omega) \cos \frac{2\pi lx}{L} + g_{2l-1}(\omega) \sin \frac{2\pi lx}{L} \right) \right) \right\} \\
&= \sum_{k=0}^{\infty} a_k^2 \cos \frac{2\pi k(x-y)}{L}
\end{aligned}$$

We now wish to find an explicit lower bound on the probability that the homologies coincide for this particular class of random fields. The following theorem from [53] gives us a lower bound.

Theorem 3.3. [53] *Consider the random Fourier series defined over $G = [0, L]$. Let M be an arbitrary natural number. Then the probability that the homologies coincide satisfies*

$$\mathbb{P}\{H_*(Q_M^\pm) \cong H_*(N^\pm)\} \geq 1 - \frac{\pi^2}{6M^2} \frac{A_0 A_2 - A_1^2}{A_0^{3/2} A_1^{1/2}} + O\left(\frac{1}{M^3}\right),$$

where A_l was defined above.

This theorem transforms the abstract estimate into an estimate that is directly related to smoothness properties of the underlying random field.

3.3 Homogenous Random Fields in Two-Dimensions

For functions of a single variable, the above bounds are fairly straightforward to characterize. The homologies of N^\pm and Q_M^\pm will differ when we have a double crossover on an interval.

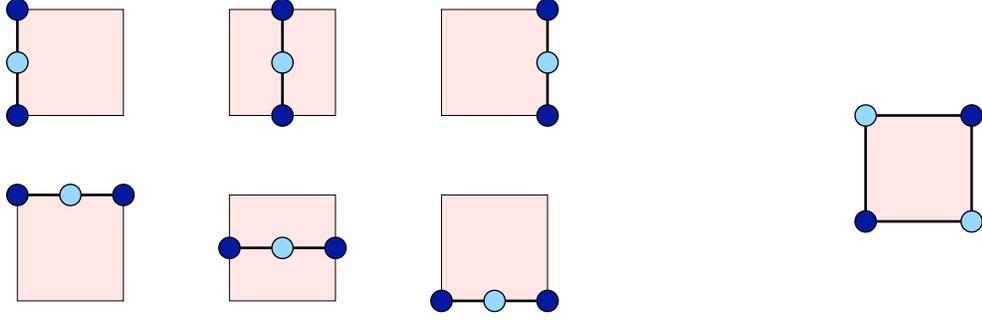


Figure 3.6: Forbidden Sign Configurations for B admissibility.

We then said that an interval was admissible if it does not contain a double crossover on any of its dyadic subintervals. How does this translate to higher dimensions? How can we characterize what an admissible square $[x_k, x_{k+1}] \times [y_l, y_{l+1}]$ will be? To begin, we start by defining a dyadic subsquare.

Definition 3.5. Let $J = [\alpha, \alpha + \delta] \times [\beta, \beta + \delta] \subset \mathbb{R}^2$ and let $d_{n,k}$ denote the dyadic points in the interval $[0, \delta]$. Then the dyadic points in the square J are the points

$$d_{n,k,l} = (\alpha + d_{n,k}, \beta + d_{n,l}) \in \mathbb{R}^2,$$

for all $k, l = 0, \dots, 2^n$, and $n \in \mathbb{N}$. The dyadic subsquares of J are the squares $d_{n,k,l} + [0, \delta/2^n]^2$ for all $k, l = 0, \dots, 2^n - 1$ and $n \in \mathbb{N} \cup \{0\}$.

We can now begin talking about admissibility for squares.

Definition 3.6. Let $u : G \rightarrow \mathbb{R}$ be an arbitrary function and $J \subset G$ be a square. Then J is B -admissible for u if none of the dyadic subsquares contain any horizontal or vertical translations of the first sign configuration shown in Figure 3.6 and the dyadic subsquares do not contain the given four point configuration.

This definition is enough to validate the homology computations. Upon examining the sign configurations for the next dyadic subdivision, shown in Figure 3.7, we group the allowed sign configurations by examining what happens on the outer vertices. Two observations can be made by examining these:

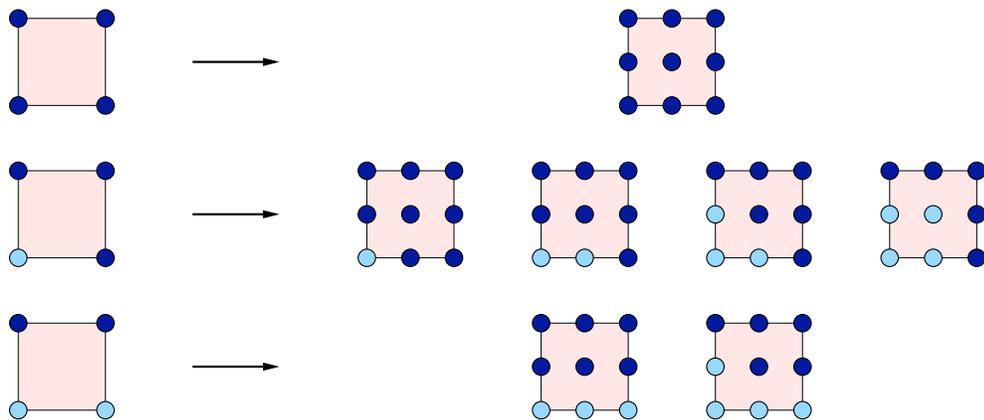


Figure 3.7: Allowed sign configurations at next dyadic subdivision for B -admissibility.

- a.) if all function values of u at the corners of the box have the same sign, then all function values at the nine dyadic points must have the same sign,
- b.) if both signs can be observed at the corners of the box, then there are two sides of the box with both positive and negative function values.

This demonstrates that if a box J is B -admissible and the function u is positive at the corners, then u cannot take on negative function values in J . This suggests that it is possible to pin down the nodal lines with these two observations. In fact, one can show that the following proposition is true.

Proposition 3.2. [53] *Let $G \subset \mathbb{R}^2$ denote a square, let $u : G \rightarrow \mathbb{R}$ be C^2 such that 0 is not a critical value of u and let $J \subset G$ be a B -admissible square. Then the following hold:*

- a.) *If u is strictly positive at the corners of J , then u is strictly positive on J . Also, If u is strictly negative at the corners of J , then u is strictly negative on J .*
- b.) *If u takes both positive and negative function values at the corners of J , then the nodal line of u inside J is a simple smooth curve which connects one side of J with another side of J .*

While B -admissibility will generate the correct homology, it will not give the correct asymptotics for the convergence rate. It can be shown that these configurations will generate

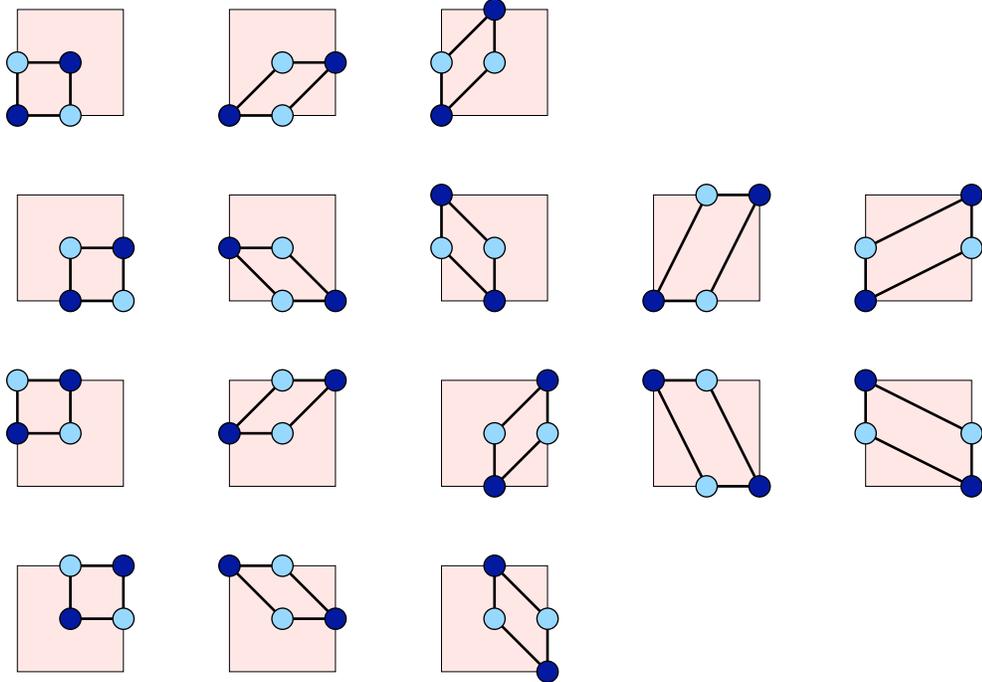


Figure 3.8: Forbidden Sign Configurations for I_4 admissibility.

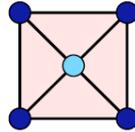


Figure 3.9: Forbidden Sign Configurations for I_5 admissibility.

a bound of the form $\mathbb{P}[J] = C_B \delta^3$, for the square J that has side lengths δ . However, in the two dimensional case, we have M^2 squares, and summing the probability for each square being B -admissible will give us a $O(1/M)$ probability bound. This is sub-optimal and we want to be able to find tighter bounds. In particular, we want to obtain a $O(1/M^2)$ bound. We must also use another notion of admissibility of squares to achieve this bound.

Definition 3.7. Let $u : G \rightarrow \mathbb{R}$ and $J = [\alpha, \alpha + \delta] \times [\beta, \beta + \delta] \subset G$. Then J is I_4 -admissible if it does not contain any of the sign configurations in Figure 3.8. The square J is I_5 admissible if it does not contain the sign configuration in Figure 3.9.

Definition 3.8. Let $u : G \rightarrow \mathbb{R}$ and $J = [\alpha, \alpha + \delta] \times [\beta, \beta + \delta] \subset \mathbb{R}^2$. Assume u and J

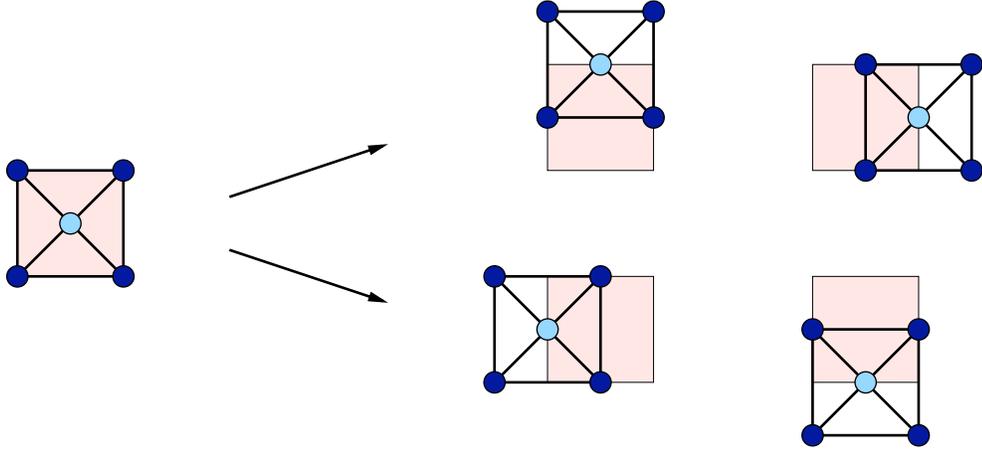


Figure 3.10: Sign Configurations Required for I -Admissibility.

are such that the $\delta/2$ translations up and down and left and right are contained in G . Then J is I -admissible if every dyadic subsquare J^* the following 5 squares are both I_4 and I_5 admissible:

- the dyadic subsquare J^*
- the four shifted squares obtained by translating J^* horizontally and vertically $\delta/2^n$ in either direction.

These are shown in Figure 3.10. We can now give a criterion for when we have made the correct global homology computations.

Proposition 3.3. [53] Let $G \subset \mathbb{R}^2$ be a square domain and $u : G \rightarrow \mathbb{R}$ be twice continuously differentiable. Let N^\pm denote the nodal domains and Q_M^\pm be the cubical approximations for fixed M . Assume u is nonzero at each sampled point and that either of the following hold:

- if the square J lies on the boundary of G , then J is B -admissible
- if the square lies in the interior of G , then J is I -admissible.

Then we have

$$H_*(N^\pm) \cong H_*(Q_M^\pm),$$

so the homologies of the nodal domains match that of the cubical approximations.

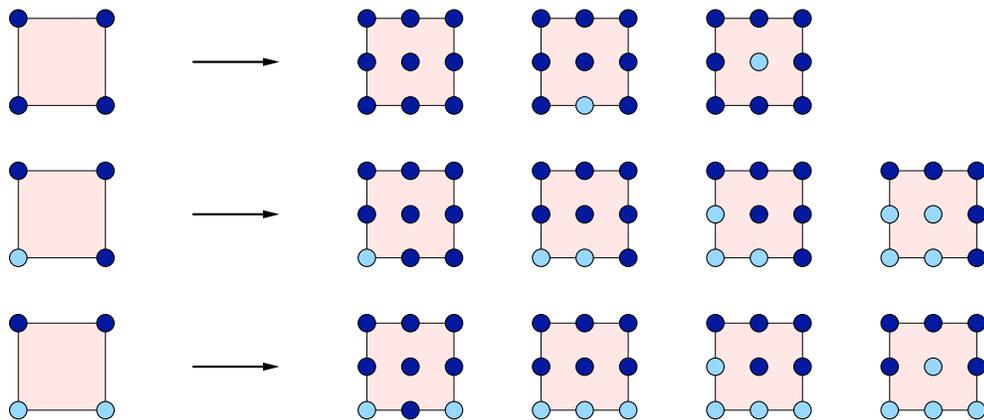


Figure 3.11: Possible Sign Structures for Squares Using I_4 Admissibility

Why is this true? The ultimate goal is simply to be able to test the function value sign at the vertices and compute the correct homology. We need to ensure that the signs give the correct geometry of the nodal lines within the region. Suppose a square contains all the same sign on the vertices. Then using I_4 admissibility, the only sign configurations that are still allowed at the next refinement are shown in Figure 3.11. Note that the last configuration in the first row is problematic. It suggests that there may be region of opposite sign enclosed in the box. It is this possible configuration that makes I_4 -admissibility insufficient and thus we require I_5 admissibility as well.

Suppose we have the sign configuration in Figure 3.11. Then we can see the only possible sign configurations for the remaining vertices that are still allowed using I_4 -admissibility. However, I_4 and I_5 admissibility are still not sufficient to capture the topology. If we look at Figure 3.12, we can see problems. The middle image has the same sign on all the vertices of two adjacent boxes. However, on the edge connecting the boxes, we have an opposite sign that is not ruled out by I_4 or I_5 admissibility. The bottom image shows a similar problem where the edge connecting two adjacent boxes is of the opposite sign at the vertices, but at the next refinement we can still have a problem. For these reasons, we need to also rule out the I_4 and I_5 sign configurations on the shifted squares as well.

With this proposition, we can finally furnish precise asymptotic bounds for the probability. This is the subject of the following theorem.

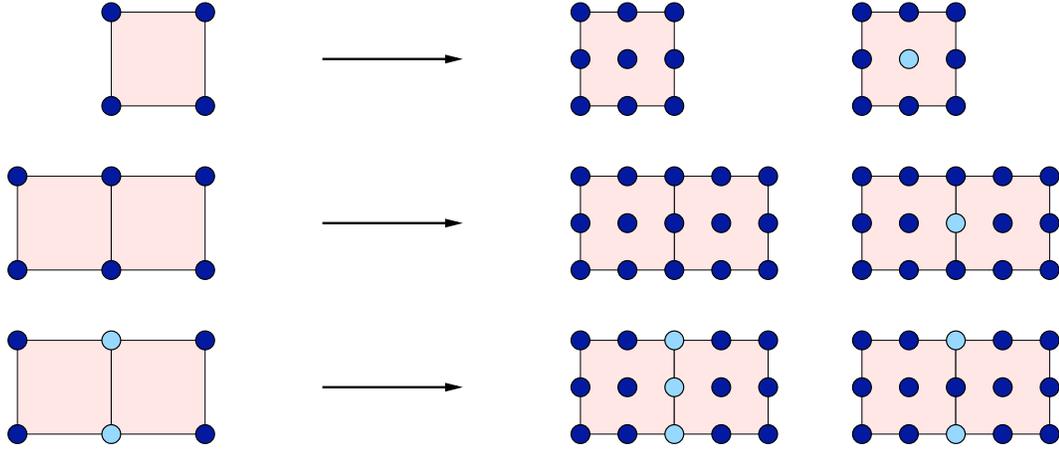


Figure 3.12: Necessity of I -admissibility.

Theorem 3.4. [53] Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a square domain $G = [0, L] \times [0, L]$. Let $u : G \rightarrow \mathbb{R}$ denote a random field satisfying assumptions A1 and A2:

(A1) for every $x \in D$, we have $\mathbb{P}\{u(x, \omega) - \mu(x) = 0\} = 0$

(A2) we have that $\mathbb{P}\{0 \text{ is a critical value of } u - \mu\} = 0$.

Assume $u(\cdot, \omega)$ is twice continuously differentiable for \mathbb{P} -almost all $\omega \in \Omega$. For each $\omega \in \Omega$ denote by $N^\pm(\omega)$ the nodal domains and let $Q_M^\pm(\omega)$ be the corresponding cubical approximations for a fixed M . For $x = (x_1, x_2)$, $\delta > 0$, and $J = [x_1, x_1 + \delta] \times [x_2, x_2 + \delta]$ consider the following:

- i.) let $E_B(x, \delta)$ denote the set of all $\omega \in \Omega$ for which $u(\cdot, \omega)$ exhibits at least one of the seven sign patterns in Figure 3.6 that are not B -admissible.
- ii.) let $E_I(x, \delta)$ denote the set of all $\omega \in \Omega$ for which $u(\cdot, \omega)$ exhibits at least one of the seventeen sign patterns in Figures 3.8 and 3.9 that are not I -admissible.

Assume there are positive constants \mathcal{C}_1 and \mathcal{C}_2 such that

$$\mathbb{P}(E_B(x, \delta)) \leq \mathcal{C}_1 \delta^3$$

and

$$\mathbb{P}(E_I(x, \delta)) \leq \mathcal{C}_2 \delta^4,$$

for all $x \in G$ and $\delta > 0$ for which J lies in G . Then for a fixed discretization size M , the probability that the homologies of $N^\pm(\omega)$ and $Q_M^\pm(\omega)$ coincide satisfies

$$\mathbb{P}\{H_*(N^\pm) \cong H_*(Q_M^\pm)\} \geq 1 - \frac{24\mathcal{C}_1 L^3 + 20\mathcal{C}_2 L^4}{3M^2}.$$

Random Periodic Functions

We now apply the previous theorem to obtain bounds for random periodic functions. Assume we have a function over the domain $[0, L] \times [0, L]$ of the form:

$$\begin{aligned} u(x_1, x_2, \omega) = \sum_{k,l=0}^{\infty} a_{k,l} \left(g_{k,l,1}(\omega) \cos \frac{2\pi k x_1}{L} \cos \frac{2\pi l x_2}{L} \right. \\ + g_{k,l,2}(\omega) \cos \frac{2\pi k x_1}{L} \sin \frac{2\pi l x_2}{L} \\ + g_{k,l,3}(\omega) \sin \frac{2\pi k x_1}{L} \cos \frac{2\pi l x_2}{L} \\ \left. + g_{k,l,4}(\omega) \sin \frac{2\pi k x_1}{L} \sin \frac{2\pi l x_2}{L} \right), \end{aligned}$$

where we assume each $g_{k,l,i}(\omega)$ are independent, normally distributed random variables defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and at least two $a_{k,l}$ are nonzero. Also assume that the random field u is such that

$$\sum_{k,l=0}^{\infty} (k^6 + l^6) a_{k,l}^2 < \infty,$$

which is directly related to the smoothness of the random field u . Define

$$A_{p,q} = \sum_{k,l=0}^{\infty} k^{2p} l^{2q} a_{k,l}^2.$$

Then it can be shown that

$$\mathbb{E} \|D_{x_1}^p D_{x_2}^q u\|_{L^2(G)}^2 = (2\pi)^{2p+2q} L^{2-2p-2q} A_{p,q}.$$

Again, these particular functions are convenient to study since we can write the covariance as

$$R(x_1, x_2, y_1, y_2) = \sum_{k,l=0}^{\infty} a_{k,l}^2 \cos \frac{2\pi k(y_1 - x_1)}{L} \cos \frac{2\pi l(y_2 - x_2)}{L}.$$

Then from [53], the probability that the homologies $N^\pm(\omega)$ and $Q_M^\pm(\omega)$ agree satisfies

$$\mathbb{P} \left[H_*(Q^\pm) \cong H_*(N^\pm) \right] \geq 1 - \frac{1067\pi^2}{18M^2} \frac{(A_{2,0} + A_{1,1} + A_{0,2})^2}{A_{0,0}^{1/2} A_{0,1}^{1/2} A_{1,0}^{1/2} A_{1,1}^{1/2}} + O\left(\frac{1}{M^3}\right).$$

3.4 Non-Homogenous Bounds

The above sections provide explicit lower bounds on the correctness of homology computations in one and two spatial dimensions. In each case, an abstract bound was first found. This bound was then applied to a concrete class of random fields, namely random periodic functions. These random fields are nice to study because of their homogeneity property. This permitted us to sample the points at equispaced distances. However, this is a severe restriction since we expect to encounter non-homogenous random fields. If we do encounter them, we cannot expect to obtain optimal bounds on the correctness probability for our

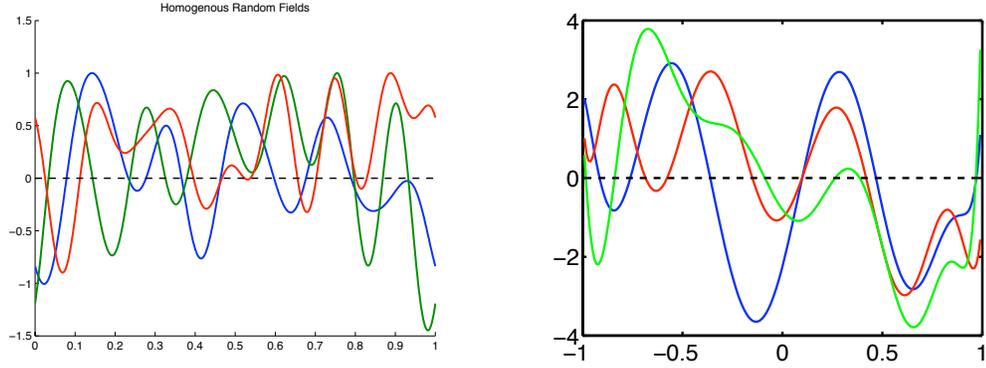


Figure 3.13: Example of Homogenous Random Field and Non-Homogenous Random Field

homology computations with equispaced points. Intuitively this results from a nonuniform clustering of the zeros of the random field.

As an illustration, we show three realizations of the random periodic function

$$u(x, \omega) = \sum_{i=0}^N (g_{2i}(\omega) \cos 2\pi i x + g_{2i-1}(\omega) \sin 2\pi i x),$$

with domain $D = [0, 1]$ and three realizations of the non-homogenous random field

$$u(x, \omega) = \sum_{i=0}^N g_i(\omega) \cos(i \arccos x),$$

with domain $D = [-1, 1]$ in Figure 3.13, where $g_i(\omega)$ is independent identically distributed normal distributions. If we compare the plots for the random periodic functions, we observe that the zeros are fairly equispaced. If we then compare the non-homogenous case, we see that the zeros tend to be spaced near the boundary. This is evidence that equispaced sampling of our random field will give suboptimal correctness bounds for non-homogenous random fields.

We now update the definitions to account for non-homogenous random fields.

Definition 3.9. Let $G = [a, b] \subset \mathbb{R}$ be a compact interval, $\mu : G \rightarrow \mathbb{R}$ be a threshold

function, and $u : G \rightarrow \mathbb{R}$ be continuous function. An M -discretization of $[a, b]$ is a collection of $M + 1$ points

$$a = x_0 < x_1 < \cdots < x_M = b.$$

The generalized nodal domains about the threshold function μ are the sets

$$N_\mu^\pm := \{x \in [a, b] : \pm(u(x) - \mu(x)) \geq 0\}.$$

These are again the sub and super-level sets about μ . The cubical approximations $Q_{\mu, M}^\pm$ of the generalized nodal domains are the sets

$$Q_{\mu, M}^\pm := \bigcup \{[x_k, x_{k+1}] : \pm(u(x) - \mu(x)) \geq 0\}.$$

Again our goal is to find lower bounds for which

$$\mathbb{P}\{H_*(N_\mu^\pm) \cong H_*(Q_{\mu, M}^\pm)\}.$$

In order to proceed, assume μ is a continuous threshold function. We make the following assumptions:

(A1) For every $x \in D$, we have $\mathbb{P}\{u(x, \omega) - \mu = 0\} = 0$

(A2) We have that $\mathbb{P}\{u - \mu \text{ has double zero in } [a, b]\} = 0$.

(A3) For $\sigma \in \{\pm 1\}$, $x \in G$ and $\delta > 0$ with $x + \delta \in G$ define

$$p_\sigma(x, \delta) = \mathbb{P}\left\{\sigma \cdot u(x) \geq \mu(x), \sigma \cdot u\left(x + \frac{\delta}{2}\right) \leq \sigma \cdot \mu\left(x + \frac{\delta}{2}\right), \sigma \cdot u(x + \delta) \geq \sigma \cdot \mu(x + \delta)\right\},$$

Then there is a continuously differentiable function $\mathcal{C}_0 : G \rightarrow \mathbb{R}$ and a positive constant

\mathcal{C}_1 such that

$$p_{+1}(x, \delta) + p_{-1}(x, \delta) \leq \mathcal{C}_0(x)\delta^3 + \mathcal{C}_1\delta^4.$$

These assumptions are quite similar to the previous ones except now they depend on the sampled points. They also depend on a threshold function μ . It is also clear that the function \mathcal{C}_0 will vary over $[a, b]$ and give different probabilities for different x but constant δ . The next theorem furnishes where we need to sample our points to get optimal bounds for our homology computations.

Theorem 3.5. [54] *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a continuous threshold function $\mu : [a, b] \rightarrow \mathbb{R}$. Let $u : [a, b] \times \Omega \rightarrow \mathbb{R}$ be a random field over $(\Omega, \mathcal{F}, \mathbb{P})$ such that for \mathbb{P} -almost all $\omega \in \Omega$ the function $u(\cdot, \omega) : [a, b] \rightarrow \mathbb{R}$ is continuous. Choose the sample points $a = x_0 < x_1 < \dots < x_M = b$ such that*

$$\int_{x_{k-1}}^{x_k} \sqrt[3]{\mathcal{C}_0(x)} dx = \frac{1}{M} \int_a^b \sqrt[3]{\mathcal{C}_0(x)} dx.$$

Consider the generalized nodal domains $N_\mu^\pm(\omega)$ and the cubical approximations $Q_{\mu, M}^\pm(\omega)$. If the assumptions (A1), (A2), and (A3) hold, then

$$\mathbb{P}\{H^*(N_\mu^\pm) \cong H^*(Q_{\mu, M}^\pm)\} \geq 1 - \frac{4}{3M^2} \left(\int_a^b \sqrt[3]{\mathcal{C}_0(x)} dx \right)^3.$$

How do we calculate $\mathcal{C}_0(x)$ for an arbitrary random field and threshold function μ ?
Recalling that the covariance function is

$$R(x, y) = \mathbb{E}((u(x) - \mathbb{E}u(x))(u(y) - \mathbb{E}u(y))),$$

we can then write

$$\begin{aligned} R_{k,l}(x) &= \frac{\partial^{k+l} R}{\partial x^k \partial y^l}(x, x) \\ &= \mathbb{E}((u^{(k)}(x) - \mathbb{E}u^{(k)}(x))(u^{(l)}(y) - \mathbb{E}u^{(l)}(y))). \end{aligned}$$

To write an explicit formula for \mathcal{C}_0 , assume the following matrix is positive definite:

$$\mathcal{R}(x) := \begin{pmatrix} R_{0,0}(x) & R_{1,0}(x) & R_{2,0}(x) \\ R_{1,0}(x) & R_{1,1}(x) & R_{2,1}(x) \\ R_{2,0}(x) & R_{2,1}(x) & R_{2,2}(x) \end{pmatrix}.$$

Now, define the following relations.

$$\mathcal{R}_{33}^m := R_{0,0}R_{1,1} - R_{1,0}^2$$

$$\mathcal{R}_{32}^m := R_{0,0}R_{2,1} - R_{1,0}R_{2,0}$$

$$\mathcal{R}_{31}^m := R_{1,0}R_{2,1} - R_{1,1}R_{2,0}$$

These relations are simply the determinants of the minors of \mathcal{R} . We can now state the following theorem which gives the optimal probability bound in terms of an explicit representation of \mathcal{C}_0

Theorem 3.6. [54] *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a continuous threshold function $\mu : [a, b] \rightarrow \mathbb{R}$ and a random field $u : [a, b] \times \Omega \rightarrow \mathbb{R}$ over $(\Omega, \mathcal{F}, \mathbb{P})$ such that for \mathbb{P} -almost all $\omega \in \Omega$ the function $u(\cdot, \omega) : [a, b] \rightarrow \mathbb{R}$ is of class C^3 . Also assume μ is of class C^3 . Choose the sample points $a = x_0 < x_1 < \dots < x_M = b$ such that*

$$\int_{x_{k-1}}^{x_k} \sqrt[3]{\mathcal{C}_0(x)} dx = \frac{1}{M} \int_a^b \sqrt[3]{\mathcal{C}_0(x)} dx.$$

We then have

$$\mathbb{P}\{H^*(N^\pm) \cong H^*(Q_M^\pm)\} \geq 1 - \frac{1}{M^2} \left(\int_a^b \sqrt[3]{\mathcal{C}_0(x)} dx \right)^3,$$

where

$$\mathcal{C}_0 = \frac{\det \mathcal{R}(x)}{48\pi \mathcal{R}_{3,3}^m(x)^{3/2}} (1 + \mathcal{A}(x)) e^{-\mathcal{B}(x)},$$

for

$$\mathcal{A}(x) = \frac{(\mathcal{R}_{3,1}^m(x)\mu(x) - \mathcal{R}_{3,2}^m(x)\mu'(x) + \mathcal{R}_{3,3}^m(x)\mu''(x))^2}{\mathcal{R}_{3,3}^m(x) \det \mathcal{R}(x)}$$

and

$$\mathcal{B}(x) = \frac{(R_{1,0}(x)\mu(x) - R_{0,0}(x)\mu'(x))^2 + \mathcal{R}_{3,3}^m(x)\mu(x)^2}{2R_{0,0}(x)\mathcal{R}_{3,3}^m(x)}.$$

Asymptotic Sign-Change Probabilities

All of the above probabilities rely on being able to identify problematic sign change configurations. In the one dimensional case, the problematic configuration occurs when we have a double crossover. In the two dimensional situation, we had to identify many more problematic configurations. Since we are dealing with random functions, we cannot state whether a problematic sign configuration will occur. We can however give precise probabilities as to when we can expect a particular sign configuration.

To begin, let $T(\delta) = (T_1(\delta), T_2(\delta), \dots, T_n(\delta))$ denote a one parameter family of \mathbb{R}^n -valued random Gaussian variables indexed by $\delta > 0$. Choose a particular sign configuration from $(s_1, s_2, \dots, s_n) \in \{\pm 1\}^n$. Let $\tau(\delta)$ be an arbitrary threshold vector. Using this, we

are interested in the asymptotic behavior as $\delta \rightarrow 0$ of

$$P(\delta) = \mathbb{P}\{s_i(T_i(\delta) - \tau_i(\delta)) \geq 0, \text{ for all } i = 1, \dots, n\}$$

Theorem 3.7. [54] *Let $(s_1, \dots, s_n) \in \{\pm 1\}^n$ denote a fixed sign sequence and consider the one parameter families of a threshold vector $\tau(\delta) \in \mathbb{R}^n$ and an \mathbb{R}^n -values random Gaussian variable $T(\delta)$ over $(\Omega, \mathcal{F}, \mathbb{P})$ for $\delta > 0$. Assume the holding conditions hold:*

- a.) *For each $\delta > 0$, assume that the Gaussian random variable $T(\delta)$ has mean $0 \in \mathbb{R}^n$ and a positive definite covariance matrix $C(\delta) \in \mathbb{R}^{n \times n}$, whose positive eigenvalues are given by $0 < \lambda_1(\delta) < \dots < \lambda_n(\delta)$ with corresponding normalized eigenvectors $\nu_1(\delta), \dots, \nu_n(\delta)$.*
- b.) *There exists a vector $\bar{\nu}_1 = (\bar{\nu}_{1,1}, \dots, \bar{\nu}_{1,n}) \in \mathbb{R}^n$ such that $\nu_1(\delta) \rightarrow \bar{\nu}_1$ as $\delta \rightarrow 0$ and $s_i \bar{\nu}_{1,i} > 0$ for all $i = 1, \dots, n$.*
- c.) *The quotient $\lambda_1(\delta)/\lambda_k(\delta)$ converges to 0 as $\delta \rightarrow 0$, for all $k = 2, \dots, n$.*
- d.) *There exists a vector $\alpha = (\alpha_1, \dots, \alpha_n)^t \in \mathbb{R}^n$ such that*

$$\lim_{\delta \rightarrow 0} \frac{\tau(\delta) \cdot \nu_k(\delta)}{\sqrt{\lambda_k(\delta)}} = \alpha_k, \text{ for all } k = 1, \dots, n.$$

Furthermore, for α define

$$S_\alpha = \frac{2}{2^{n/2} \cdot \Gamma(n/2)} e^{-\sum_{k=2}^n \alpha_k^2/2} \cdot \int_{\alpha_1}^{\infty} (s - \alpha_1)^{n-1} e^{-s^2/2} ds.$$

Then the probability $P(\delta)$ satisfies

$$\lim_{\delta \rightarrow 0} P(\delta) \sqrt{\frac{\det C(\delta)}{\lambda_1(\delta)^n}} = \frac{\Gamma(n/2) \cdot S_\alpha}{2 \cdot \pi^{n/2} \cdot (n-1)!} \cdot \left| \prod_{j=1}^n \bar{\nu}_{1,j} \right|^{-1}.$$

This theorem gives us a method to determine the probability of the sign configurations shown. The above theorem takes the threshold into consideration. For a zero threshold $\mu = 0$, we can greatly simplify the above theorem as follows.

Theorem 3.8. [53] *Let $(s_1, \dots, s_n) \in \{\pm 1\}^n$ be a fixed sign sequence and consider the one parameter family of \mathbb{R}^n -values random Gaussian variable $T(\delta)$ over $(\Omega, \mathcal{F}, \mathbb{P})$ for $\delta > 0$. Assume the holding conditions hold:*

- a.) *For each $\delta > 0$, assume that the Gaussian random variable $T(\delta)$ has mean $0 \in \mathbb{R}^n$ and a positive definite covariance matrix $C(\delta) \in \mathbb{R}^{n \times n}$, whose positive eigenvalues are given by $0 < \lambda_1(\delta) < \dots < \lambda_n(\delta)$ with corresponding normalized eigenvectors $\nu_1(\delta), \dots, \nu_n(\delta)$.*
- b.) *There exists a vector $\bar{\nu}_1 = (\bar{\nu}_{1,1}, \dots, \bar{\nu}_{1,n}) \in \mathbb{R}^n$ such that $\nu_1(\delta) \rightarrow \bar{\nu}_1$ as $\delta \rightarrow 0$ and $s_i \bar{\nu}_{1,i} > 0$ for all $i = 1, \dots, n$.*
- c.) *The quotient $\lambda_1(\delta)/\lambda_k(\delta)$ converges to 0 as $\delta \rightarrow 0$, for all $k = 2, \dots, n$.*

Then the probability $P(\delta)$ satisfies

$$\lim_{\delta \rightarrow 0} P(\delta) \sqrt{\frac{\det C(\delta)}{\lambda_1(\delta)^n}} = \frac{\Gamma(n/2)}{2 \cdot \pi^{n/2} \cdot (n-1)!} \cdot \left| \prod_{j=1}^n \bar{\nu}_{1,j} \right|^{-1}.$$

This theorem gives the precise asymptotic behavior of most the sign configurations as the side lengths decrease to zero. However, it does not hold for the five point configuration. One can provide another theorem that has weaker hypothesis that holds for the five point configuration.

Theorem 3.9. [53] *Let $(s_1, \dots, s_n) \in \{\pm 1\}^n$ be a fixed sign sequence and consider the one parameter family of \mathbb{R}^n -valued random Gaussian variable $T(\delta)$ over $(\Omega, \mathcal{F}, \mathbb{P})$ for $\delta > 0$. Assume the holding conditions hold:*

a.) For each $\delta > 0$, assume that the Gaussian random variable $T(\delta)$ has mean $0 \in \mathbb{R}^n$ and a positive definite covariance matrix $C(\delta) \in \mathbb{R}^{n \times n}$, whose positive eigenvalues are given by $0 < \lambda_1(\delta) < \dots < \lambda_n(\delta)$ with corresponding normalized eigenvectors $\nu_1(\delta), \dots, \nu_n(\delta)$.

b.) There exists a vector $\bar{\nu}_1 = (\bar{\nu}_{1,1}, \dots, \bar{\nu}_{1,n}) \in \mathbb{R}^n$ such that $\nu_1(\delta) \rightarrow \bar{\nu}_1$ as $\delta \rightarrow 0$ and $s_i \bar{\nu}_{1,i} > 0$ for all $i = 1, \dots, n$.

Then the probability $P(\delta)$ satisfies

$$\lim_{\delta \rightarrow 0} P(\delta) \sqrt{\frac{\det C(\delta)}{\lambda_1(\delta)^n}} \leq \frac{\Gamma(n/2)}{2 \cdot \pi^{n/2} \cdot (n-1)!} \cdot \left| \prod_{j=1}^n \bar{\nu}_{1,j} \right|^{-1}.$$

Chapter 4: Non-Homogenous Bounds

4.1 Introduction

Assume we are given a random field $u(x, y, \omega)$ defined over $[a, b] \times [c, d]$ such that we can write

$$u(x, y, \omega) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{j,k} g_{j,k}(\omega) \phi_j(x) \psi_k(y),$$

where $g_{j,k}(\omega)$ are independent, identically distributed Gaussian random variables with mean zero and variance one, and $\phi_j(x)$ and $\psi_k(y)$ are complete orthogonal basis functions.

Our main objective is to correctly identify the homology of the nodal domains

$$N^{\pm}(\omega) := \{(x, y) : \pm u(x, y, \omega) \geq 0\}$$

We again have two main obstacles to overcome in order to address the correctness of the homology. First, we will be working with discretized versions of the nodal domains. In particular, we will work with the sets

$$Q_M^{\pm}(\omega) := \bigcup_{i,j=0}^{M-1} \{[i, i+1] \times [j, j+1] : \pm u(x_i, y_j, \omega) \geq 0\}$$

The second difficulty is the probabilistic nature of our random fields. Since we are dealing with random variables, we will have different nodal domains for each realization. Thus we do not know a priori what the nodal domains will look like and where to sample. Our main

goal is thus to find sharp lower bounds for

$$\mathbb{P} \{H_* (N^\pm) \cong H_* (Q_M^\pm)\}.$$

As such our approach will be as follows. We will make heavy use of Theorems 3.7, 3.8, and 3.9 to find the local probability for each forbidden sign configuration in Q_M^\pm . Once the local probabilities have been calculated, we can then sum over all elements in Q_M^\pm to obtain the total probability of making mistakes. Lastly, we minimize the total failure probability by changing our sampled points (x_i, y_j) .

Due to the structure of the random fields under consideration, we can write the covariance between the coordinate points (x_1, y_1) and (x_2, y_2) as

$$\begin{aligned} R(x_1, x_2, y_1, y_2) &= \mathbb{E}[(u(x_1, y_1) - \mathbb{E}(u(x_1, y_1))) (u(x_2, y_2) - \mathbb{E}(u(x_2, y_2)))] \\ &= \mathbb{E}[u(x_1, y_1)u(x_2, y_2)] \\ &= \mathbb{E} \left[\left(\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{j,k} g_{j,k}(\omega) \phi_j(x_1) \psi_k(y_1) \right) \left(\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} a_{l,m} g_{l,m}(\omega) \phi_l(x_2) \psi_m(y_2) \right) \right] \\ &= \mathbb{E} \left[\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} a_{j,k} a_{l,m} g_{j,k} g_{l,m} \phi_j(x_1) \phi_l(x_2) \psi_j(y_1) \psi_m(y_2) \right] \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} a_{j,k} a_{l,m} \mathbb{E}[g_{j,k} g_{l,m}] \phi_j(x_1) \phi_l(x_2) \psi_j(y_1) \psi_m(y_2) \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{j,k}^2 \phi_j(x_1) \phi_j(x_2) \psi_k(y_1) \psi_k(y_2), \end{aligned}$$

where we have used the fact that $g_{j,k}(\omega)$ are independent Gaussian random variables with mean zero and variance one. The above makes sense by Mercer's theorem assuming the basis functions are continuous and complete over the domain.

Since we are mainly interested in the asymptotic behavior of the forbidden sign configurations, we begin by performing a Taylor expansion of R about the point (x_1, x_2, y_1, y_2) to the point $(x_1 + a, x_2 + b, y_1 + c, y_2 + d)$. This yields

$$\begin{aligned}
R = r_0 &+ \left(r_1 a + r_2 b + r_3 c + r_4 d \right) + \frac{1}{2!} \left(r_{11} a^2 + 2r_{12} ab + 2r_{13} ac + 2r_{14} ad + r_{22} b^2 \right. \\
&+ 2r_{23} bc + 2r_{24} bd + r_{33} c^2 + 2r_{34} cd + r_{44} d^2 \left. \right) \\
&+ \frac{1}{3!} \left(r_{111} a^3 + 3r_{112} aab + 3r_{113} aac + 3r_{114} aad + 3r_{122} abb + 6r_{123} abc + 6r_{124} abd \right. \\
&+ 3r_{133} acc + 6r_{134} acd + 3r_{144} add + r_{222} b^3 + 3r_{223} bcb + 3r_{233} bcc \\
&+ 6r_{234} bcd + r_{333} c^3 + 3r_{334} ccd + 3r_{344} cdd + r_{444} d^3 + 3r_{244} bdd + 3r_{224} bbd \left. \right) \\
&+ \frac{1}{4!} \left(r_{1111} aaaa + 4r_{1112} aaab + 4r_{1113} aaac + 4r_{1114} aaad + 6r_{1122} aabb + 12r_{1123} aabc \right. \\
&+ 12r_{1124} aabd + 6r_{1133} aacc + 12r_{1134} aacd + 6r_{1144} aadd + 24r_{1234} abcd \\
&+ 12r_{1233} abcc + 12r_{1244} abdd + 4r_{1222} abbb + 12r_{1223} abbc + 12r_{1224} abbd + 4r_{1333} accc \\
&+ 12r_{1334} accd + 12r_{1443} addc + 4r_{1444} addd + r_{2222} bbbb + 4r_{2223} bbbc + 4r_{2224} bbbd \\
&+ 6r_{2233} bbcc + 12r_{2234} bbcd + 6r_{2244} bbdd + 4r_{2333} bccc + 12r_{2334} bccd + 12r_{2344} bcdd \\
&+ 4r_{2444} bddd + r_{3333} cccc + 4r_{3334} cccd + 6r_{3344} ccdd + 4r_{3444} cddd + r_{4444} dddd \left. \right) \\
&+ O(|(a, b, c, d)|^5),
\end{aligned}$$

where $r_D := r_D(x_1, x_2, y_1, y_2)$ is the $D = (d_1, d_2, d_3, d_4)$ derivative of R evaluated at the sampled points (x_1, x_2, y_1, y_2) and also $r_0 = r(x_1, x_2, y_1, y_2)$.

We now derive all the local probabilities for the forbidden sign configurations. We will show

explicitly how to do this for four situations. The remaining sign configurations are similar in procedure to the ones explicitly derived in the next sections. In particular, we show the derivations for the three point configuration, the four point square, the four point skewed square, and the five point configuration.

We employ one technique in our calculations that is non-standard. Known as the Newton polygon method, this tool determines small solutions to polynomial equations. It has practical applications in bifurcation theory, differential equations, and dynamical systems. A few good references are [47],[66], and [69].

Assuming we wish to find small solutions to the equation

$$F(\nu, w) = 0 \quad \text{where } F(0, 0) = 0,$$

around $(0, 0)$. Assume the following:

- Let \mathcal{O} be open with $(0, 0) \in \mathcal{O}$ and $F : \mathcal{O} \rightarrow \mathbb{R}$ analytic. Then we can write

$$F(\nu, w) = \sum_{j,k=0}^{\infty} c_{j,k} w^k \nu^j.$$

- The functions $F(\cdot, w)$ and $F(\nu, \cdot)$ are not identically zero. This implies there exists $\bar{j}, \bar{k} \in \mathbb{N}$ such that

$$c_{\bar{j},0} \neq 0 \quad \text{and} \quad c_{0,\bar{k}} \neq 0.$$

Our desire is to find small solutions of the form

$$w = a\nu^\alpha + o(\nu^\alpha)$$

as $\nu \rightarrow 0$. Notice that when we put this expression into the term $c_{j,k} w^k \nu^j$ we get

$$c_{j,k} w^k \nu^j = a^k c_{j,k} \nu^{k\alpha+j}.$$

This gives that $k\alpha + j = \hat{k}\alpha + \hat{j}$ if and only if the points (j, k) and (\hat{j}, \hat{k}) lie on a line with slope $-\alpha$ so that

$$-\alpha = \frac{\hat{j} - j}{\hat{k} - k}.$$

Also notice that $k\alpha + j < \hat{k}\alpha + \hat{j}$ if and only if (\hat{j}, \hat{k}) lies above the line through (j, k) with slope $-\alpha$. These two observations are the crux of the Newton Polygon. The following provides a rigorous definition of the Newton polygon.

Definition 4.1. *Let \mathbb{K} be the convex hull of all points (j, k) such that $c_{j,k} \neq 0$. The Newton Polygon is the polygon along the lower left boundary of \mathbb{K} between the points $(0, k)$ and $(j, 0)$. This provides a sequence of segments whose slopes are $-\alpha_1 < -\alpha_2 < \dots < -\alpha_r$.*

This definition gives us a simple method to determine small solutions for the equation $F(\nu, w)$. We first form the Newton polygon given by F . Then the slopes $-\alpha_s$ determine the exponents for the equation $w = a\nu^\alpha$. Then we plug this expression into F , collection the terms which contribute to that slope and solve for a . An example is shown in Figure 4.1. The outer lines form the Newton Polygon and the lower line determines the exponents for the small solutions. The exact formulation is given in Theorem 4.1.

Theorem 4.1. [47] *Let F be an analytic function satisfying:*

a.) *Let \mathcal{O} be open with $(0, 0) \in \mathcal{O}$ and $F : \mathcal{O} \rightarrow \mathbb{R}$ analytic. Then we can write*

$$F(\nu, w) = \sum_{j,k=0}^{\infty} c_{j,k} w^k \nu^j.$$

b.) *The functions $F(\cdot, w)$ and $F(\nu, \cdot)$ are not identically zero. This implies there exists $\bar{j}, \bar{k} \in \mathbb{N}$ such that*

$$c_{\bar{j},0} \neq 0 \quad \text{and} \quad c_{0,\bar{k}} \neq 0.$$

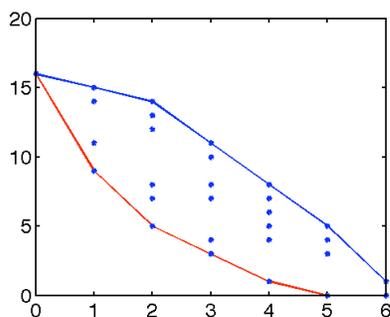


Figure 4.1: Sample Newton Polygon

Let L_1, \dots, L_r denote the line segments of the Newton Polygon with slopes $-\alpha_1 < \dots < -\alpha_r$.

Then

$$F(\nu, w) = 0$$

has a solution

$$\nu = aw^\alpha + o(w^\alpha)$$

close to $(0, 0)$ if and only if

1. $\alpha = \alpha_i$, for $i = 1, \dots, r$,
2. $a \neq 0$ is a solution of

$$\sum_{(k,l) \in L_\gamma, c_{k,l} \neq 0} c_{k,l} a^k = 0.$$

For example, suppose we wish to find small solutions to the equation

$$F(x, y) = y^5 - 14xy^2 + 5x^2y + x^3,$$

in terms of $x = ay^s + o(y^s)$. The Newton Polygon is shown in Figure 4.2. The first image shows the entire Newton Polygon, the second image shows the line with greatest slope in red, and the last image shows the remaining line in red. Examining the diagram gives us

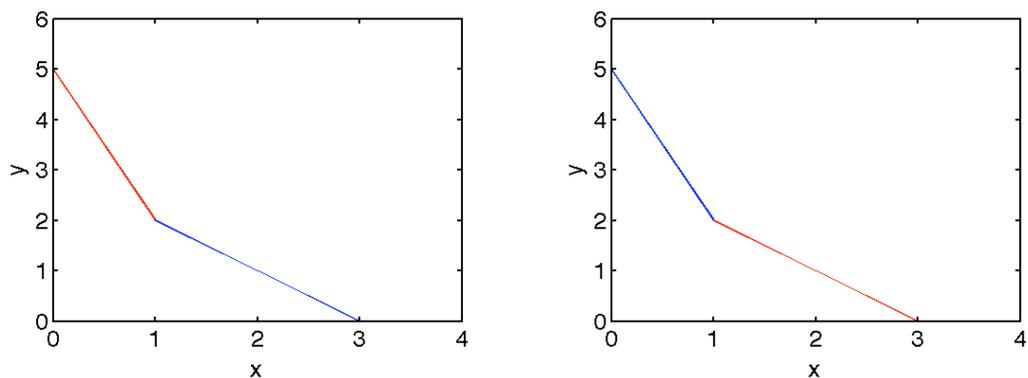


Figure 4.2: Newton polygon for equation $F(x, y) = y^5 - 14xy^2 + 5x^2y + x^3$.

the slopes $-\alpha_1 = 3$ and $-\alpha_2 = 1$. The small solutions are then of the form

$$x = a_1y^3 + o(y^3)$$

$$x = a_2y + o(y).$$

Now we find the coefficients a_1 and a_2 . To find a_1 , we substitute the expression $x = a_1y^3$ into the terms that contributed to the line with slope $-\alpha_1 = 3$. This gives

$$\begin{aligned} 0 &= y^5 - 14xy^2 \\ &= y^5 - 14(a_1y^3)y^2 \\ &= y^5(1 - 14a_1). \end{aligned}$$

This gives $a_1 = 1/14$. We do a similar procedure to find a_2 . We substitute $x = a_2y$ into the terms of F that contributed the line with slope $-\alpha_2 = 1$. This gives

$$\begin{aligned} 0 &= -14xy^2 + 5x^2y + x^3 \\ &= -14(a_2y)y^2 + 5(a_2y)^2y + (a_2y)^3 \\ &= y^3(a_2^2 + 5a_2 - 14), \end{aligned}$$

which solving for a_2 gives $a_2 = 2$ and $a_2 = -7$. Then for the equation $F(x, y) = y^5 - 14xy^2 + 5x^2y + x^3$, the small solutions are

$$x = \frac{1}{14}y^3 + o(y^3)$$

$$x = 2y + o(y)$$

$$x = -7y + o(y).$$

4.2 Three Point Sign Configurations

We start with the three point sign configuration in the x direction. In this case, we obviously have $n = 3$ and the sign vector $(s_1, s_2, s_3) = (+1, -1, +1)$. We are interested in the asymptotics of the random vector $T(\delta_1, \delta_2) = (T_1(\delta_1, \delta_2), T_2(\delta_1, \delta_2), T_3(\delta_1, \delta_2)) : \Omega \rightarrow \mathbb{R}^3$, defined by

$$T_1(\delta_1, \delta_2) = u(x_1, y_1), \quad T_2(\delta_1, \delta_2) = u(x_1 + \delta_1/2, y_1), \quad T_3(\delta_1, \delta_2) = u(x_1 + \delta_1, y_1).$$

Then $T(\delta_1, \delta_2)$ is a Gaussian random variable with mean $(0, 0, 0) \in \mathbb{R}^3$. The covariance matrix of $T(\delta_1, \delta_2)$ about the point (x_1, y_1) is then $\mathcal{C}_1(\delta_1, \delta_2) =$

$$\begin{pmatrix} R(x_1, x_1, y_1, y_1) & R(x_1, x_1 + \delta_1/2, y_1, y_1) & R(x_1, x_1 + \delta_1, y_1, y_1) \\ R(x_1 + \delta_1/2, x_1, y_1, y_1) & R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1, y_1) & R(x_1 + \delta_1/2, x_1 + \delta_1, y_1, y_1) \\ R(x_1 + \delta_1, x_1, y_1, y_1) & R(x_1 + \delta_1, x_1 + \delta_1/2, y_1, y_1) & R(x_1 + \delta_1, x_1 + \delta_1, y_1, y_1) \end{pmatrix}.$$

Using the Taylor expansion for R , the determinant of \mathcal{C}_1 is then

$$\det \mathcal{C}_1 = \frac{1}{64} \mathcal{A}_1(x_1, y_1) \delta_1^6 + O(\delta_1^7),$$

where we have defined

$$\mathcal{A}_1(x_1, y_1) := (r_0 r_{1122} r_{12} - r_0 r_{112} r_{122} - r_1 r_{1122} r_2 + r_{11} r_{122} r_2 + r_1 r_{112} r_{22} - r_{11} r_{12} r_{22}).$$

We note here that we used Mathematica for the symbolic calculations. We first expanded the covariance function in a symbolic Taylor series. Then based upon the sign configuration we are using, we formed the covariance matrix. The determinant was then calculated using Mathematica's determinant command. This gives the full determinant, so to find the leading order terms, we then collected all the leading order terms.

Since the covariance matrix is positive definite for δ_1, δ_2 sufficiently small, this immediately implies that the determinant is positive. The next step is to obtain the leading order terms for the eigenvalues of \mathcal{C}_1 . To do this, we will use the Newton Polygon. In particular, we will use this method for the characteristic polynomial $\det(\mathcal{C}_1(\delta_1, \delta_2) - \lambda I)$. To leading order in each term, we obtain the characteristic polynomial

$$\frac{1}{64} \mathcal{A}_1 \cdot \delta_1^6 \cdot \lambda^0 + \mathcal{B}_1 \cdot \delta_1^2 \cdot \lambda^1 + \mathcal{D}_1 \cdot \delta_1^0 \cdot \lambda^2 - \lambda^3,$$

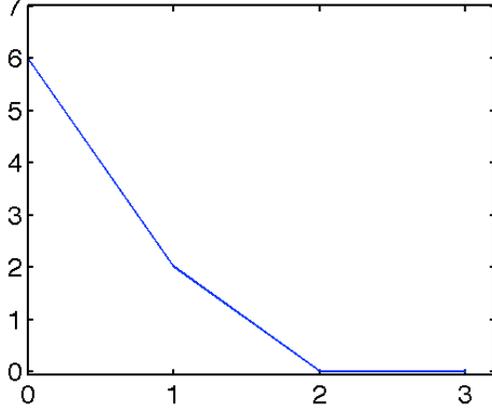


Figure 4.3: Newton polygon for three point sign configuration

where we have

$$\mathcal{B}_1 := \mathcal{B}_1(x_1, y_1) = -\frac{3}{2}(r_0 r_{12} - r_1 r_2)$$

$$\mathcal{D}_1 := \mathcal{D}_1(x_1, y_1) = 3r_0.$$

Using the Newton Polygon shown in Figure 4.3, we find the slopes $l_1 = -4$, $l_2 = -2$, and $l_3 = 0$. To leading order, the eigenvalues are

$$\lambda_1 = -\frac{1}{96} \frac{\mathcal{A}_1}{\mathcal{B}_1} \delta_1^4 + O(\delta_1^5)$$

$$\lambda_2 = -\frac{1}{2} \frac{\mathcal{B}_1}{\mathcal{D}_1} \delta_1^2 + O(\delta_1^3)$$

$$\lambda_3 = 3r_{00} + O(\delta_1^1),$$

which is

$$\lambda_1 = -\frac{1}{96} \frac{r_0 r_{1122} r_{12} - r_0 r_{112} r_{122} - r_1 r_{1122} r_2 + r_{11} r_{122} r_2 + r_1 r_{112} r_{22} - r_{11} r_{12} r_{22}}{r_0 r_{12} - r_1 r_2} \delta_1^4 + O(\delta_1^5)$$

$$\lambda_2 = \frac{1}{2} \frac{r_0 r_{12} - r_1 r_2}{r_0} \delta_1^2 + O(\delta_1^3)$$

$$\lambda_3 = 3r_0 + O(\delta_1^1).$$

Lastly, we must find the asymptotic behavior of the eigenvectors. To do this, we consider the adjoint matrix, which can be calculated as

$$\text{adj } \mathcal{C}_1(\delta_1, \delta_2) = \frac{3}{2} (r_0 r_{12} - r_1 r_2) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} \delta_1^2 + O(\delta_1^3).$$

This was again found symbolically using Mathematica. This matrix has eigenvalues 0, 0, and 6. The eigenvalue 6 has the associated eigenvector $(1, -2, 1)^t / \sqrt{6}$. The eigenspace of the largest eigenvalue of the adjoint corresponds to the eigenspace of the smallest eigenvalue $\lambda_1(\delta_1, \delta_2)$ of the covariance matrix \mathcal{C}_1 . Hence, we can conclude that in the limit $(\delta_1, \delta_2) \rightarrow (0, 0)$ we have $\nu_1(\delta_1, \delta_2) \rightarrow (1, -2, 1)^t / \sqrt{6}$.

We have now established the validity of all the assumptions for Theorem 3.8. We can therefore conclude that

$$\begin{aligned} \mathbb{P}_1(x_1, y_1, \delta_1, \delta_2) &= \frac{1}{128\pi^2} \sqrt{\frac{\lambda_1^3}{\det \mathcal{C}_1}} \delta_1^3 + O(\delta_1^3) \\ &= \frac{r_0 r_{1122} r_{12} - r_0 r_{112} r_{122} - r_1 r_{1122} r_2 + r_{11} r_{122} r_2 + r_1 r_{112} r_{22} - r_{11} r_{12} r_{22}}{128\pi^2 (r_0 r_{12} - r_1 r_2)^{3/2}} \delta_1^3 + O(\delta_1^4), \end{aligned}$$

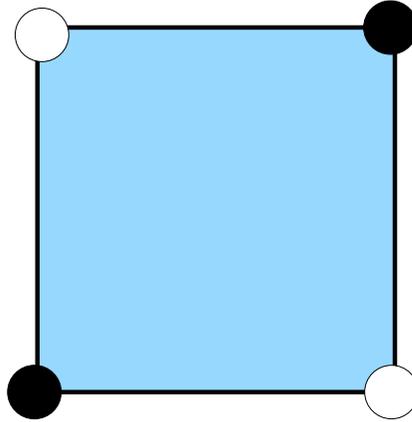


Figure 4.4: Four Point Sign Configuration

where each of the coefficients r_d depends on the sampled coordinates (x_1, y_1) .

4.3 Four Point Square

We turn our attention to the four point square shown in Figure 4.4. The same process will be followed as in the three point configuration. In this case, we are investigating the behavior of the sign vector $s = (+1, -1, +1, -1)$ for the random vector $T_3(\delta_1, \delta_2) = (u(x_1, y_1), u(x_1 + \delta_1, y_1), u(x_1 + \delta_1, y_1 + \delta_2), u(x_1, y_1 + \delta_2))$. The covariance matrix of T_3 is then

$$c_3 = \begin{pmatrix} S_{1,1} & S_{1,2} & S_{1,3} & S_{1,4} \\ S_{2,1} & S_{2,2} & S_{2,3} & S_{2,4} \\ S_{3,1} & S_{3,2} & S_{3,3} & S_{3,4} \\ S_{4,1} & S_{4,2} & S_{4,3} & S_{4,4} \end{pmatrix},$$

where

$$S_{1,1} = R(x_1, x_1, y_1, y_1)$$

$$S_{1,2} = R(x_1, x_1 + \delta_1, y_1, y_1)$$

$$S_{1,3} = R(x_1, x_1 + \delta_1, y_1, y_1 + \delta_2)$$

$$S_{1,4} = R(x_1, x_1, y_1, y_1 + \delta_2)$$

$$S_{2,1} = R(x_1 + \delta_1, x_1, y_1, y_1)$$

$$S_{2,2} = R(x_1 + \delta_1, x_1 + \delta_1, y_1, y_1)$$

$$S_{2,3} = R(x_1 + \delta_1, x_1 + \delta_1, y_1, y_1 + \delta_2)$$

$$S_{2,4} = R(x_1 + \delta_1, x_1, y_1, y_1 + \delta_2)$$

$$S_{3,1} = R(x_1 + \delta_1, x_1, y_1 + \delta_2, y_1)$$

$$S_{3,2} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2, y_1)$$

$$S_{3,3} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{3,4} = R(x_1 + \delta_1, x_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{4,1} = R(x_1, x_1, y_1 + \delta_2, y_1)$$

$$S_{4,2} = R(x_1, x_1 + \delta_1, y_1 + \delta_2, y_1)$$

$$S_{4,3} = R(x_1, x_1 + \delta_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{4,4} = R(x_1, x_1, y_1 + \delta_2, y_1 + \delta_2)$$

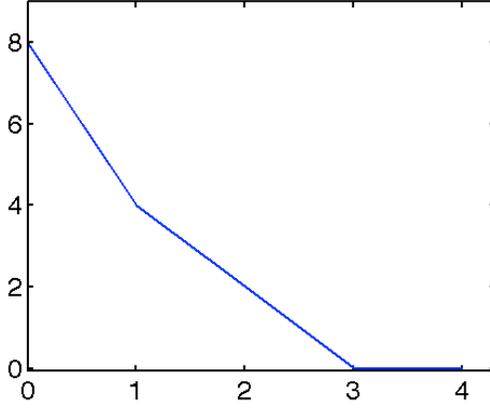


Figure 4.5: Newton polygon for four point sign configuration

Using the Taylor series expansion, we can calculate the determinant of \mathcal{C}_3 to obtain

$$\det \mathcal{C}_3 = \mathcal{A}_3(x_1, y_1) \delta_1^4 \delta_2^4 + O(|\delta_1, \delta_2|^9)$$

where the term $\mathcal{A}_3(x_1, y_1)$ is given by

$$\begin{aligned} \mathcal{A}_3(x_1, y_1) = & r_{00}r_{124}r_{134}r_{23} - r_{00}r_{1234}r_{14}r_{23} - r_{00}r_{12}r_{134}r_{234} + r_{00}r_{123}r_{14}r_{234} \\ & + r_{11}r_{134}r_2r_{234} - r_{13}r_{14}r_2r_{234} - r_{11}r_{134}r_{23}r_{24} + r_{13}r_{14}r_{23}r_{24} \\ & - r_{124}r_{134}r_2r_3 + r_{1234}r_{14}r_2r_3 + r_{12}r_{134}r_{24}r_3 - r_{123}r_{14}r_{24}r_3 \\ & + r_{00}r_{12}r_{1234}r_{34} - r_{00}r_{123}r_{124}r_{34} - r_{11}r_{1234}r_2r_{34} + r_{124}r_{13}r_2r_{34} \\ & + r_{11}r_{123}r_{24}r_{34} - r_{12}r_{13}r_{24}r_{34} + r_{11}r_{1234}r_{23}r_4 - r_{124}r_{13}r_{23}r_4 \\ & - r_{11}r_{123}r_{234}r_4 + r_{12}r_{13}r_{234}r_4 - r_{12}r_{1234}r_3r_4 + r_{123}r_{124}r_3r_4. \end{aligned}$$

We now turn our attention to the eigenvalues of \mathcal{C}_3 . The Newton Polygon is shown in Figure 4.5. To leading order terms of (δ_1, δ_2) , the characteristic polynomial $\det(\mathcal{C}_3 - \lambda I) =$

0 may be written as

$$\mathcal{A}_3 \cdot \delta_1^4 \delta_2^4 \cdot \lambda^0 + \mathcal{B}_3 \cdot \delta_1^2 \delta_2^2 \cdot \lambda^1 + (\mathcal{D}_{3,1} \delta_1^2 + \mathcal{D}_{3,2} \delta_2^2) \cdot \lambda^2 + \mathcal{E}_3 \delta_1^0 \delta_2^0 \cdot \lambda^3 + \lambda^4 = 0.$$

We want to find small solutions such that $\lambda = \lambda(\delta_1, \delta_2)$. However, we cannot use the Newton Polygon for the characteristic polynomial in the above form. Here we make the assumption that as $(\delta_1, \delta_2) \rightarrow (0, 0)$, this limit occurs linearly. In particular, we assume we can write $\delta_1 = \kappa_1 \cdot t$ and $\delta_2 = \kappa_2 \cdot t$, where t is a small parameter. This is justified since we cannot decrease one side length of the box at a much faster rate than the other side length. Using this assumption, we may then write the characteristic polynomial as

$$\mathcal{A}_3 \kappa_1^4 \kappa_2^4 t^8 \lambda^0 + \mathcal{B}_3 \kappa_1^2 \kappa_2^2 t^4 \lambda^1 + (\mathcal{D}_{3,1} \kappa_1^2 + \mathcal{D}_{3,2} \kappa_2^2) t^2 \lambda^2 + \mathcal{E}_3 \kappa_1^0 \kappa_2^0 t^0 \lambda^3 + \lambda^4 = 0.$$

We can now write the small solutions of λ in terms of t using the Newton Polygon. Doing this gives the smallest eigenvalue

$$\lambda_1 = -\frac{\kappa_1^2 \kappa_2^2 \mathcal{A}_3}{4 \mathcal{B}_3} t^4 + O(t^5),$$

which upon substituting κ_1 and κ_2 , we obtain

$$\lambda_1 = -\frac{\delta_1^2 \delta_2^2 \mathcal{A}_3}{4 \mathcal{B}_3} + O(|(\delta_1, \delta_2)|^5).$$

We note here that the coefficient \mathcal{B}_3 is given by the formula

$$\mathcal{B}_3 = r_{00} r_{14} r_{23} - r_{14} r_2 r_3 - r_{00} r_{12} r_{34} + r_{11} r_2 r_{34} - r_{11} r_{23} r_4 + r_{12} r_3 r_4.$$

In order to determine the behavior of the eigenvectors of \mathcal{C}_3 , we again use the classical

adjoint matrix. Computing this yields

$$adj\mathcal{C}_3 = -4\mathcal{B}_3 \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix} \delta_1^2 \delta_2^2 + O(|\delta_1, \delta_2|^5)$$

The constant matrix has the eigenvalue 0 with multiplicity three and the eigenvalue 4 with multiplicity 1. The nonzero eigenvalue has the corresponding normalized eigenvector $(-1/2, 1/2, -1/2, 1/2)^t$. By the same reasoning as before we conclude

$$\nu_3(\delta_1, \delta_2) \rightarrow \left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)^t.$$

Using all this information, we can conclude from Theorem 3.8 that the probability of having this forbidden sign configuration satisfies the equation

$$\mathbb{P}_3 = \frac{1}{12\pi^2} \frac{\mathcal{A}_3^{3/2}}{\mathcal{B}_3^2} \delta_1^2 \delta_2^2 + O(|(\delta_1, \delta_2)|^5).$$

4.4 Four Point Skewed Square

In this section, we investigate the behavior of the skewed four point sign configuration in Figure 4.6. We are again using the sign vector $s = (+1, -1, +1, -1)$ with the random vector $T_5(\delta_1, \delta_2) = (u(x_1, y_1), u(x_1 + \delta_1/2, y_1), u(x_1 + \delta_1, y_1 + \delta_2/2), u(x_1 + \delta_1/2, y_1 + \delta_2/2))$. The

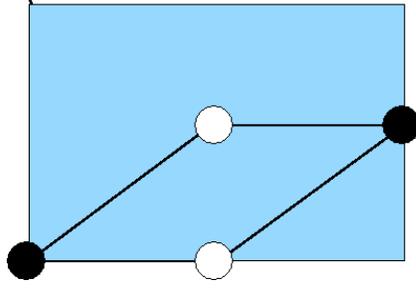


Figure 4.6: Skewed Four Point Sign Configuration

covariance matrix of T_5 is then

$$\mathcal{C}_5(\delta_1, \delta_2) = \begin{pmatrix} S_{1,1} & S_{1,2} & S_{1,3} & S_{1,4} \\ S_{2,1} & S_{2,2} & S_{2,3} & S_{2,4} \\ S_{3,1} & S_{3,2} & S_{3,3} & S_{3,4} \\ S_{4,1} & S_{4,2} & S_{4,3} & S_{4,4} \end{pmatrix},$$

where

$$S_{1,1} = R(x_1, x_1, y_1, y_1)$$

$$S_{1,2} = R(x_1, x_1 + \delta_1/2, y_1, y_1)$$

$$S_{1,3} = R(x_1, x_1 + \delta_1, y_1, y_1 + \delta_2/2)$$

$$S_{1,4} = R(x_1, x_1 + \delta_1/2, y_1, y_1 + \delta_2/2)$$

$$S_{2,1} = R(x_1 + \delta_1/2, x_1, y_1, y_1)$$

$$S_{2,2} = R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1, y_1)$$

$$S_{2,3} = R(x_1 + \delta_1/2, x_1 + \delta_1, y_1, y_1 + \delta_2/2)$$

$$S_{2,4} = R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1, y_1 + \delta_2/2)$$

$$S_{3,1} = R(x_1 + \delta_1, x_1, y_1 + \delta_2/2, y_1)$$

$$S_{3,2} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2/2, y_1)$$

$$S_{3,3} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2/2, y_1 + \delta_2/2)$$

$$S_{3,4} = R(x_1 + \delta_1, x_1, y_1 + \delta_2/2, y_1 + \delta_2/2)$$

$$S_{4,1} = R(x_1 + \delta_1/2, x_1, y_1 + \delta_2/2, y_1)$$

$$S_{4,2} = R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1 + \delta_2/2, y_1)$$

$$S_{4,3} = R(x_1 + \delta_1/2, x_1 + \delta_1, y_1 + \delta_2/2, y_1 + \delta_2/2)$$

$$S_{4,4} = R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1 + \delta_2/2, y_1 + \delta_2/2)$$

The determinant may be computed to leading order as

$$\det \mathcal{C}_5 = \frac{1}{256} (\mathcal{A}_{5,1} \delta_1^6 \delta_2^2 + \mathcal{A}_{5,2} \delta_1^5 \delta_2^3 + \mathcal{A}_{5,3} \delta_1^4 \delta_2^4) + O(|(\delta_1, \delta_2)|^9).$$

The eigenvalues of \mathcal{C}_5 are again found symbolically by computing $\det(\mathcal{C}_5 - \lambda I) = 0$. Again, we make the assumption that $\delta_1 = \kappa \cdot t$ and $\delta_2 = \kappa_2 \cdot t$, where t is the small parameter. Using the Newton Polygon which is the same as the four point square case, we compute the smallest eigenvalue as

$$\begin{aligned} \lambda_1 &= \frac{1}{64} \frac{(\mathcal{A}_{5,1} \kappa_1^4 + \mathcal{A}_{5,2} \kappa_1^3 \kappa_2 + \mathcal{A}_{5,3} \kappa_1^2 \kappa_2^2)}{\mathcal{B}_5} \cdot t^4 + O(t^5) \\ &= \frac{1}{64} \frac{(\mathcal{A}_{5,1} \delta_1^4 + \mathcal{A}_{5,2} \delta_1^3 \delta_2 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2)}{\mathcal{B}_5} + O(|(\delta_1, \delta_2)|^5) \end{aligned}$$

For this situation, the constant matrix in adjoint matrix is the same as the four point square. This gives the same eigenvalues and eigenvectors for the adjoint. So we again have

$$\nu_5(\delta_1, \delta_2) \rightarrow \left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)^t.$$

Combining all this information with Theorem 3.8, we have the probability of this sign configuration is

$$\mathbb{P}_5 = \frac{1}{192\pi^2} \frac{(\mathcal{A}_{5,1} \delta_1^4 + \mathcal{A}_{5,2} \delta_1^3 \delta_2 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_5^2 \delta_1 \delta_2} + O(|(\delta_1, \delta_2)|^5)$$

4.5 Five Point Sign Configuration

In this section, we show the calculations for the five point configuration. We are interested in the quantities

$$s = \begin{pmatrix} +1 \\ +1 \\ -1 \\ +1 \\ +1 \end{pmatrix}, \quad T_{10}(\delta_1, \delta_2) = \begin{pmatrix} u(x_1, y_1) \\ u(x_1 + \delta_1, y_1) \\ u(x_1 + \delta_1/2, y_1 + \delta_2/2) \\ u(x_1, y_1 + \delta_2) \\ u(x_1 + \delta_1, y_1 + \delta_2) \end{pmatrix}$$

and,

$$\mathcal{C}_{10} = \begin{pmatrix} S_{1,1} & S_{1,2} & S_{1,3} & S_{1,4} & S_{1,5} \\ S_{2,1} & S_{2,2} & S_{2,3} & S_{2,4} & S_{2,5} \\ S_{3,1} & S_{3,2} & S_{3,3} & S_{3,4} & S_{3,5} \\ S_{4,1} & S_{4,2} & S_{4,3} & S_{4,4} & S_{4,5} \\ S_{5,1} & S_{5,2} & S_{5,3} & S_{5,4} & S_{5,5} \end{pmatrix},$$

where

$$S_{1,1} = R(x_1, x_1, y_1, y_1)$$

$$S_{1,2} = R(x_1, x_1 + \delta_1, y_1, y_1)$$

$$S_{1,3} = R(x_1, x_1 + \delta_1/2, y_1, y_1 + \delta_2/2)$$

$$S_{1,4} = R(x_1, x_1, y_1, y_1 + \delta_2)$$

$$S_{1,5} = R(x_1, x_1 + \delta_1, y_1, y_1 + \delta_2)$$

$$S_{2,1} = R(x_1 + \delta_1, x_1, y_1, y_1)$$

$$S_{2,2} = R(x_1 + \delta_1, x_1 + \delta_1, y_1, y_1)$$

$$S_{2,3} = R(x_1 + \delta_1, x_1 + \delta_1/2, y_1, y_1 + \delta_2/2)$$

$$S_{2,4} = R(x_1 + \delta_1, x_1, y_1, y_1 + \delta_2)$$

$$S_{2,5} = R(x_1 + \delta_1, x_1 + \delta_1, y_1, y_1 + \delta_2)$$

$$S_{3,1} = R(x_1 + \delta_1/2, x_1, y_1 + \delta_2/2, y_1)$$

$$S_{3,2} = R(x_1 + \delta_1/2, x_1 + \delta_1, y_1 + \delta_2/2, y_1)$$

$$S_{3,3} = R(x_1 + \delta_1/2, x_1 + \delta_1/2, y_1 + \delta_2/2, y_1 + \delta_2/2)$$

$$S_{3,4} = R(x_1 + \delta_1/2, x_1, y_1 + \delta_2/2, y_1 + \delta_2)$$

$$S_{3,5} = R(x_1 + \delta_1/2, x_1 + \delta_1, y_1 + \delta_2/2, y_1 + \delta_2)$$

$$S_{4,1} = R(x_1, x_1, y_1 + \delta_2, y_1)$$

$$S_{4,2} = R(x_1, x_1 + \delta_1, y_1 + \delta_2, y_1)$$

$$S_{4,3} = R(x_1, x_1 + \delta_1/2, y_1 + \delta_2, y_1 + \delta_2/2)$$

$$S_{4,4} = R(x_1, x_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{4,5} = R(x_1, x_1 + \delta_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{5,1} = R(x_1 + \delta_1, x_1, y_1 + \delta_2, y_1)$$

$$S_{5,2} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2, y_1)$$

$$S_{5,3} = R(x_1 + \delta_1, x_1 + \delta_1/2, y_1 + \delta_2, y_1 + \delta_2/2)$$

$$S_{5,4} = R(x_1 + \delta_1, x_1, y_1 + \delta_2, y_1 + \delta_2)$$

$$S_{5,5} = R(x_1 + \delta_1, x_1 + \delta_1, y_1 + \delta_2, y_1 + \delta_2).$$

To leading order, we can calculate the determinant as

$$\det \mathcal{C}_{10} = \frac{1}{64} (\mathcal{A}_{10,1} \delta_1^8 \delta_2^4 + \mathcal{A}_{10,2} \delta_1^6 \delta_2^6 + \mathcal{A}_{10,3} \delta_1^4 \delta_2^8) + O(|(\delta_1, \delta_2)|^{13}),$$

where $\mathcal{A}_{10,i}$ are coefficients specified in the appendix. Computing the characteristic equation $\det \mathcal{C}_{10} - \lambda I$ gives equation

$$\begin{aligned} & \frac{1}{64} \left(\mathcal{A}_{10,1} \delta_1^8 \delta_2^4 + \mathcal{A}_{10,2} \delta_1^6 \delta_2^6 + \mathcal{A}_{10,3} \delta_1^4 \delta_2^8 \right) \cdot \lambda^0 + \left(\mathcal{B}_{10,1} \delta_1^6 \delta_2^2 + \mathcal{B}_{10,2} \delta_1^4 \delta_2^4 \right. \\ & \quad \left. + \mathcal{B}_{10,3} \delta_1^2 \delta_2^6 \right) \cdot \lambda^1 + \left(\mathcal{C}_{10,1} \delta_1^4 + \mathcal{C}_{10,2} \delta_1^2 \delta_2^2 + \mathcal{C}_{10,3} \delta_2^4 \right) \cdot \lambda^2 \\ & \quad + \left(\mathcal{D}_{10,1} \delta_1^2 + \mathcal{D}_{10,2} \delta_2^2 \right) \cdot \lambda^3 + \mathcal{E}_{10} \cdot \lambda^4 + \mathcal{F}_{10} \cdot \lambda^5 = 0, \end{aligned}$$

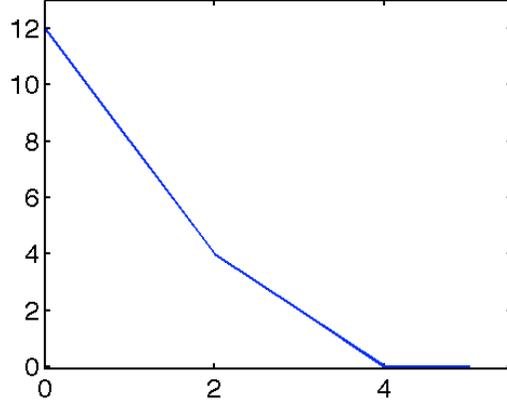


Figure 4.7: Newton polygon for five point sign configuration

where we have only shown the leading order terms for each $\lambda^r, r = 0, \dots, 5$. Using the assumption $\delta_1 = \kappa_1 t$ and $\delta_2 = \kappa_2 t$, we get the equation

$$\begin{aligned} & \frac{1}{64} \left(\mathcal{A}_{10,1} \kappa_1^8 \kappa_2^4 + \mathcal{A}_{10,2} \kappa_1^6 \kappa_2^6 + \mathcal{A}_{10,3} \kappa_1^4 \kappa_2^8 \right) \cdot t^{12} \cdot \lambda^0 + \left(\mathcal{B}_{10,1} \kappa_1^6 \kappa_2^2 + \mathcal{B}_{10,2} \kappa_1^4 \kappa_2^4 \right. \\ & \quad \left. + \mathcal{B}_{10,3} \kappa_1^2 \kappa_2^6 \right) \cdot t^8 \cdot \lambda^1 + \left(\mathcal{C}_{10,1} \kappa_1^4 + \mathcal{C}_{10,2} \kappa_1^2 \kappa_2^2 + \mathcal{C}_{10,3} \kappa_2^4 \right) \cdot t^4 \cdot \lambda^2 \\ & \quad + \left(\mathcal{D}_{10,1} \kappa_1^2 + \mathcal{D}_{10,2} \kappa_2^2 \right) \cdot t^2 \cdot \lambda^3 + \mathcal{E}_{10} \cdot \lambda^4 + \mathcal{F}_{10} \cdot \lambda^5 = 0, \end{aligned}$$

which we simplify as

$$\mathcal{A}_{10} \cdot t^{12} \cdot \lambda^0 + \mathcal{B}_{10} \cdot t^8 \cdot \lambda^1 + \mathcal{C}_{10} \cdot t^4 \cdot \lambda^2 + \mathcal{D}_{10} \cdot t^2 \cdot \lambda^3 + \mathcal{E}_{10} \cdot \lambda^4 + \mathcal{F}_{10} \cdot \lambda^5 = 0.$$

The Newton Polygon in terms of t and λ is shown in Figure 4.7. When we solve for the small solutions, we arrive at two solutions that have the same asymptotic behavior as $t \rightarrow 0$.

These are

$$\lambda_{10,-} = \frac{-\mathcal{B}_{10} - \sqrt{\mathcal{B}_{10}^2 - 4\mathcal{A}_{10}\mathcal{C}_{10}}}{2\mathcal{C}} t^4$$

$$\lambda_{10,+} = \frac{-\mathcal{B}_{10} + \sqrt{\mathcal{B}_{10}^2 - 4\mathcal{A}_{10}\mathcal{C}_{10}}}{2\mathcal{C}} t^4.$$

We use $\lambda_{10} = \lambda_{10,-}$ as the smallest eigenvalue.

The next item we need to calculate is the asymptotic behavior of the eigenvector corresponding to λ_{10} . Using Mathematica, we find

$$\nu_{10}(\delta_1, \delta_2) \rightarrow \frac{1}{2\sqrt{5}}(1, 1, -4, 1, 1)^t.$$

We can then use Theorem 3.9 to combine all this information to calculate the probability this sign configuration occurs for this domain.

4.6 Admissibility of Rectangles

The last sections gave derivations for the probability that each sign configuration occurs on a rectangle. These local sign probabilities will be used to discuss the admissibility of such a rectangle. Let $D_{m,n} = [x_m, x_{m+1}] \times [y_n, y_{n+1}]$ with $x_{m+1} - x_m = \delta_1$ and $y_{n+1} - y_n = \delta_2$. Recall:

- a.) the rectangle $D_{m,n}$ is B -admissible if it does not contain any of the forbidden sign configurations in Figure 4.8 on any of the dyadic subdivisions on $D_{m,n}$.
- b.) the rectangle $D_{m,n}$ is I_4 -admissible if it does not contain any of the forbidden sign configurations in Figure 4.9 on any of the dyadic subdivisions on $D_{m,n}$.
- c.) the rectangle $D_{m,n}$ is I_5 -admissible if it does not contain the five point sign configuration in Figure 4.10 on any of the dyadic subdivisions.

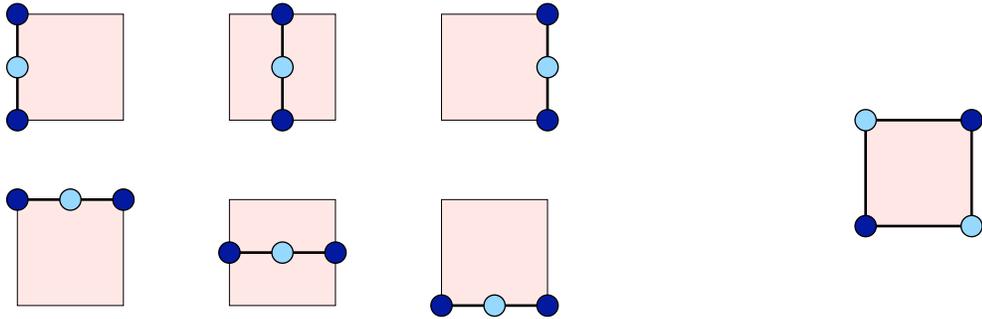


Figure 4.8: Forbidden sign configurations for B -admissibility.

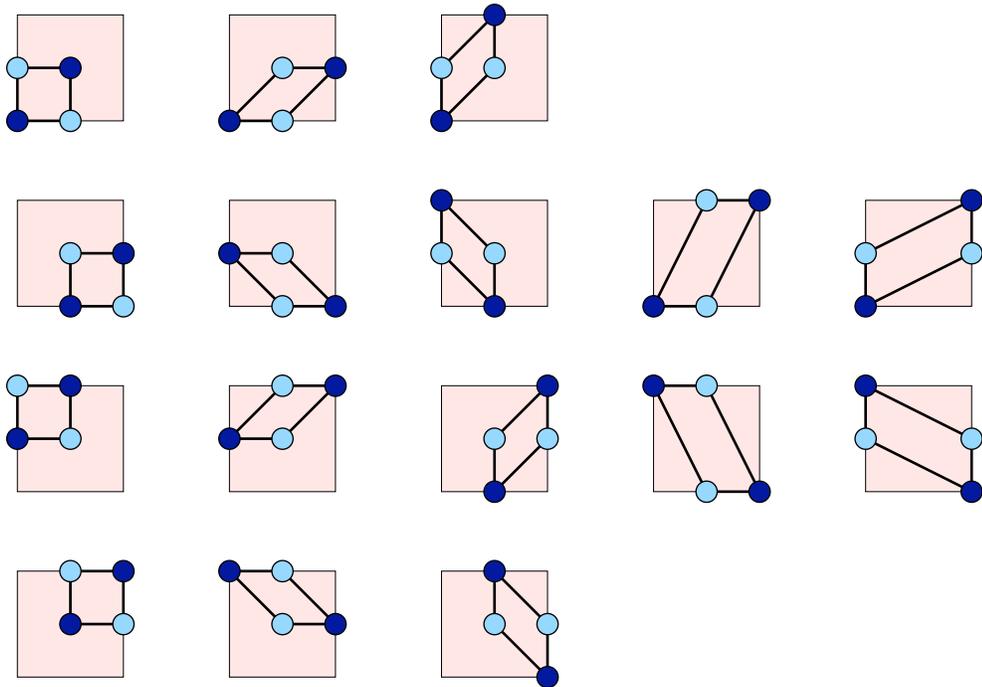


Figure 4.9: Forbidden Sign Configurations for I_4 admissibility.

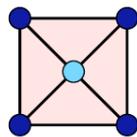


Figure 4.10: Forbidden Sign Configuration for I_5 admissibility.

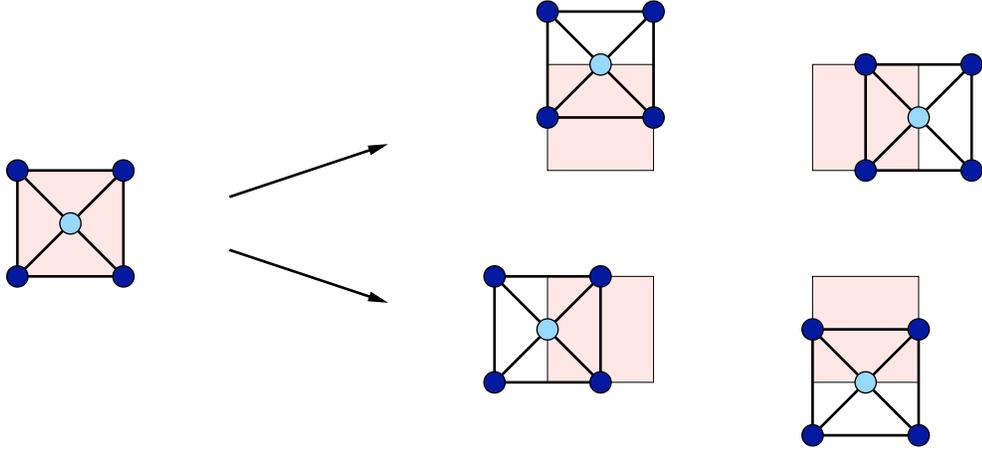


Figure 4.11: Sign Configurations Required for I -Admissibility.

d.) It is I -admissible if it is I_4 and I_5 admissible on $D_{m,n}$ and the shifted squares shown in Figure 4.11.

The last few sections derived explicit probabilities for each sign configuration on the actual rectangle itself. However, for a square to be admissible we need exclude the forbidden sign configurations on all the dyadic subdivisions of the rectangle. The following lemma gives bounds on the probability of a rectangle being either B or I admissible.

Lemma 4.1. *Given the rectangle D , for each of the forbidden sign configurations, let E_{sign} be the event that a particular sign configuration occurs on any of the dyadic subdivisions of D and the shifts of D . Then we have*

$$\mathbb{P}\{E_{sign}\} \leq C \cdot \mathbb{P}\{\text{sign configuration on } D\},$$

where C is a constant depending on whether the rectangle D touches the boundary or is an interior rectangle. In particular, if D touches the boundary, $C = 2$ and if D is an interior rectangle we have $C = 20/3$.

Proof. We will only show the results for a boundary term and a skewed four point configuration. The remaining sign configurations can be handled similarly.

To begin, we will work with A_5 , the skewed four point configuration. Let $D = [x_m, x_{m+1}] \times [y_n, y_{n+1}]$ denote the rectangle with $x_{m+1} - x_m = \delta_1$ and $y_{n+1} - y_n = \delta_2$. We first note that we can bound the probability of this sign configuration evaluated at any dyadic point by the probability at the corner plus higher order terms of δ_1 and δ_2 . To see this, denote the dyadic points as $x_{j,k}$ and $y_{j,k}$. We can then write the probability of the A_5 sign configuration occurring at $(x_{j,k}, y_{j,k})$ with side lengths $\delta_{1,j,k}$ and $\delta_{2,j,k}$ as

$$\begin{aligned} \mathbb{P}_5 &= \frac{(\mathcal{A}_{5,1}(x_{j,k}, y_{j,k})\delta_1^4 + \mathcal{A}_{5,2}(x_{j,k}, y_{j,k})\delta_1^3\delta_2 + \mathcal{A}_{5,3}(x_{j,k}, y_{j,k})\delta_1^2\delta_2^2)^{3/2}}{192\pi^2\mathcal{B}_5^2(x_{j,k}, y_{j,k})\delta_1\delta_2} + O(|(\delta_1, \delta_2)|^5) \\ &= \frac{(\mathcal{G}_{5,1}(x_{j,k}, y_{j,k})\delta_1^4 + \mathcal{G}_{5,2}(x_{j,k}, y_{j,k})\delta_1^3\delta_2 + \mathcal{G}_{5,3}(x_{j,k}, y_{j,k})\delta_1^2\delta_2^2)^{3/2}}{192\pi^2\delta_1\delta_2} + O(|(\delta_1, \delta_2)|^5), \end{aligned}$$

where we have defined $\mathcal{G}_{5,i}(x, y) = \frac{\mathcal{A}_{5,i}(x, y)}{\mathcal{B}_5(x, y)^{4/3}}$. Then by the mean value theorem, there is a constant c such that

$$\mathcal{G}_{5,1}(x_{j,k}, y_{j,k}) - \mathcal{G}_{5,1}(x, y) = \nabla\mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \cdot ((x_{j,k}, y_{j,k}) - (x, y)),$$

and similarly for $\mathcal{G}_{5,2}$ and $\mathcal{G}_{5,3}$. This result immediately implies

$$\mathcal{G}_{5,1}(x_{j,k}, y_{j,k}) = \mathcal{G}_{5,1}(x, y) + \nabla\mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \cdot ((x_{j,k}, y_{j,k}) - (x, y)).$$

The last term on the right is the sum of the following two expressions

$$\begin{aligned} &\left(\frac{\partial}{\partial x} \mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \right) (x_{j,k} - x) \\ &\left(\frac{\partial}{\partial y} \mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \right) (y_{j,k} - y). \end{aligned}$$

Denote $M_{5,1,x} = \max \left| \frac{\partial}{\partial x} \mathcal{G}_{5,1} \right|$ and $M_{5,1,y} = \max \left| \frac{\partial}{\partial y} \mathcal{G}_{5,1} \right|$, where the maximum is taken over the entire rectangle $[x, x + \delta_1] \times [y, y + \delta_2]$. Using the fact that $|x_{j,k} - x| \leq \delta_1$ and

$|y_{j,k} - y| \leq \delta_2$, we can bound the above expressions by

$$\begin{aligned} \left(\frac{\partial}{\partial x} \mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \right) (x_{j,k} - x) &\leq M_{5,1,x} \delta_1 \\ \left(\frac{\partial}{\partial y} \mathcal{G}_{5,1}((1-c) \cdot (x_{j,k}, y_{j,k}) + c \cdot (x, y)) \right) (y_{j,k} - y) &\leq M_{5,1,y} \delta_2. \end{aligned}$$

This furnishes the bound

$$\mathcal{G}_{5,1}(x_{j,k}, y_{j,k}) \leq \mathcal{G}_{5,1}(x, y) + M_{5,1,x} \delta_1 + M_{5,1,y} \delta_2.$$

By similar reasoning, we can conclude

$$\mathcal{G}_{5,2}(x_{j,k}, y_{j,k}) \leq \mathcal{G}_{5,2}(x, y) + M_{5,2,x} \delta_1 + M_{5,2,y} \delta_2$$

$$\mathcal{G}_{5,3}(x_{j,k}, y_{j,k}) \leq \mathcal{G}_{5,3}(x, y) + M_{5,3,x} \delta_1 + M_{5,3,y} \delta_2.$$

When we substitute these expressions into the first expression, we get

$$\begin{aligned} &\frac{(\mathcal{G}_{5,1}(x_{j,k}, y_{j,k}) \delta_1^4 + \mathcal{G}_{5,2}(x_{j,k}, y_{j,k}) \delta_1^3 \delta_2 + \mathcal{G}_{5,3}(x_{j,k}, y_{j,k}) \delta_1^2 \delta_2^2)^{3/2}}{192 \pi^2 \delta_1 \delta_2} \\ &\leq \frac{1}{192 \pi^2 \delta_1 \delta_2} \left((\mathcal{G}_{5,1}(x, y) + M_{5,1,x} \delta_1 + M_{5,1,y} \delta_2) \delta_1^4 + (\mathcal{G}_{5,2}(x, y) + M_{5,2,x} \delta_1 + M_{5,2,y} \delta_2) \delta_1^3 \delta_2 \right. \\ &\quad \left. + (\mathcal{G}_{5,3}(x, y) + M_{5,3,x} \delta_1 + M_{5,3,y} \delta_2) \delta_1^2 \delta_2^2 \right)^{3/2}. \end{aligned}$$

We collect in terms of powers of δ_1 and δ_2 . Keeping only leading order terms we arrive at

$$\begin{aligned} &\frac{(\mathcal{G}_{5,1}(x_{j,k}, y_{j,k}) \delta_1^4 + \mathcal{G}_{5,2}(x_{j,k}, y_{j,k}) \delta_1^3 \delta_2 + \mathcal{G}_{5,3}(x_{j,k}, y_{j,k}) \delta_1^2 \delta_2^2)^{3/2}}{192 \pi^2 \delta_1 \delta_2} \\ &\leq \frac{(\mathcal{G}_{5,1}(x, y) \delta_1^4 + \mathcal{G}_{5,2}(x, y) \delta_1^3 \delta_2 + \mathcal{G}_{5,3}(x, y) \delta_1^2 \delta_2^2)^{3/2}}{192 \pi^2 \delta_1 \delta_2} + O(|(\delta_1, \delta_2)|^5). \end{aligned}$$

This shows that

$$\mathbb{P}(x_{j,k}, y_{j,k}) \leq \mathbb{P}(x, y) + O(|(\delta_1, \delta_2)|^5).$$

Thus we are able to bound the probability at all the dyadic points by the probability of the sign configuration on the original box. In order to calculate the probability of the event E_{sign} , we need to sum the probability at each of the dyadic points. This may be written

$$\begin{aligned} & \sum_{n=0}^{\infty} \sum_{j=0}^{2^n-1} \sum_{k=0}^{2^n-1} \frac{1}{192\pi^2} \frac{\left(\mathcal{A}_{5,1} \left(\frac{\delta_1}{2^n}\right)^4 + \mathcal{A}_{5,2} \left(\frac{\delta_1}{2^n}\right)^3 \left(\frac{\delta_2}{2^n}\right)^1 + \mathcal{A}_{5,3} \left(\frac{\delta_1}{2^n}\right)^2 \left(\frac{\delta_2}{2^n}\right)^2 \right)^{3/2}}{\mathcal{B}_5^2 \left(\frac{\delta_1}{2^n}\right) \left(\frac{\delta_2}{2^n}\right)} \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^{2^n-1} \sum_{k=0}^{2^n-1} \frac{1}{192\pi^2} \frac{\left(\mathcal{A}_{5,1} \delta_1^4 + \mathcal{A}_{5,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2 \right)^{3/2}}{\mathcal{B}_5^2 \delta_1 \delta_2} \cdot \left(\frac{1}{2^n} \right)^4 \\ &= \frac{4}{3} \frac{1}{192\pi^2} \frac{\left(\mathcal{A}_{5,1} \delta_1^4 + \mathcal{A}_{5,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2 \right)^{3/2}}{\mathcal{B}_5^2 \delta_1 \delta_2}, \end{aligned}$$

where we have used the above bound on the probability evaluated at the dyadic points. This establishes the probability this sign configuration occurs on any of the dyadic subdivisions of the original box. For I -admissibility, we also need the probability that this sign configuration occurs on the four shifted boxes as well. By a similar analysis, we can find the same bound for all shifted squares. This furnishes the constant $C = 5\frac{4}{3} = \frac{20}{3}$. The remaining sign configurations can be handled in similar ways. \square

The above result furnishes a bound on the probability that a particular sign configuration occurs on any of the dyadic subdivisions of a rectangle. The probability that a rectangle is I -admissible is obtained by summing over all possible excluded sign configurations. Once we have excluded all problematic sign configurations on the dyadic subdivisions for a box, we then sum the probability of each box being admissible.

Definition 4.2. Let $u(\cdot, \omega) : D \rightarrow \mathbb{R}$ be a Gaussian random field. Let J be a rectangle

$[x_1, y_1] \times [x_2, y_2] \subset D$. Then define

$$\mathbb{P}_{lo}(x_1, x_2, y_1, y_2)$$

to be the local probability that the rectangle J is admissible. If J touches the boundary, then we are referring to B -admissibility, otherwise we are referring to I -admissibility.

The next result establishes the desired result.

Theorem 4.2. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the domain $G = [a, b] \times [c, d]$. Let $u : G \times \Omega \rightarrow \mathbb{R}$ denote a random field satisfying the two assumptions A1 and A2. Additionally, assume $u(\cdot, \omega)$ is sufficiently smooth. Denote $N^\pm(\omega)$ as the nodal domains of u . Let $\vec{x}_M = (a, x_1, x_2, \dots, x_{m-1}, b)^t$ and $\vec{y}_M = (c, y_1, y_2, \dots, y_{m-1}, d)^t$ denote a collection of x and y sampling points, and denote $Q_M^\pm(\omega)$ denote the decomposition of G into M^2 rectangles. Let $z = (x, y)$, $\delta_1 > 0$, $\delta_2 > 0$ such that $J = [x, x + \delta_1] \times [y, y + \delta_2] \subset G$. Denote $\delta = (\delta_1, \delta_2)$. Assume the following:

- i.) Let $E_B(z, \delta)$ denote the set of $\omega \in \Omega$ such that for which u has one of the sign configurations for the boundary on any of the dyadic subdivisions of J
- ii.) Let $E_I(z, \delta)$ denote the set of $\omega \in \Omega$ such that for which u has one of the sign configurations for the interior on any of the dyadic subdivisions of J .

Assume we can bound $\mathbb{P}[E_B(z, \delta)]$ by a term in the appendix and also we can bound $\mathbb{P}[E_I(z, \delta)]$ by an appropriate term in the appendix as well.

Then the probability that the homologies $N^\pm(\omega)$ and $Q_M^\pm(\omega)$ agree satisfies

$$\mathbb{P}\{N^\pm(\omega) \cong Q_M^\pm(\omega)\} \geq 1 - \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \mathbb{P}_{lo}(x_j, y_k, x_{j+1}, y_{k+1}),$$

where $\mathbb{P}_{lo}(x_j, y_k, x_{j+1}, y_{k+1})$ is the sum of the local probabilities for the rectangle $[x_j, x_{j+1}] \times [y_k, y_{k+1}]$.

Proof. To begin, fix a discretization of G into M^2 rectangles. Then the homologies of $N^\pm(\omega)$ and $Q_M^\pm(\omega)$ can be different if and only if at least one of the following occur:

- The function u is zero at one of the dyadic points $x_{j,k}$ and $y_{j,k}$,
- The function u has a double zero,
- One of the rectangles on the boundary is not B admissible,
- One of the rectangles in the interior is not I admissible.

By assumption, the probability that the first two events occur is zero. For the remaining two, we have already established explicit probabilities that they occur. Thus if we can exclude these events from occurring on all the appropriate rectangles, we have the correct homology of $Q_M^\pm(\omega)$. This furnishes

$$1 - \mathbb{P} \{N^\pm(\omega) \cong Q_M^\pm(\omega)\} \leq \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \mathbb{P}_{I_0}(x_j, y_k, x_{j+1}, y_{k+1}),$$

which is the desired result. □

The above theorem establishes an explicit probability for the correct homology. However, due to the complex nature of the local probabilities, it is difficult to establish a convergence rate in this form. The next theorem provides a means of determining the convergence rate.

Theorem 4.3. *Let the above theorem hold. Additionally assume $G = [a, a+L] \times [c, c+L]$, for a fixed length L . Then for an equi-partitioned discretization of G , we have the convergence*

$$\mathbb{P} \{H_*(N^\pm(\omega)) \cong H_*(Q_M^\pm(\omega))\} \geq 1 - \frac{24\mathcal{C}_1 L^3 + 20\mathcal{C}_2 L^4}{3M^2},$$

where $\mathcal{C}_1 = \max \mathbb{P}_B$ and $\mathcal{C}_2 = \max \mathbb{P}_I$ and \mathbb{P}_B is the local probability for B admissibility of a unit square and \mathbb{P}_I is the local probability for I admissibility with of a unit square.

Proof. The main ingredients are to establish explicit probabilities for B and I admissibility for all squares. If we are able to do this, then the result follows from Theorem 3.8, [53]. To begin, we start with B admissibility. Assume we have partitioned G into an equi-partition. Then we only have sub-squares, each with side length $\delta = \frac{L}{M}$. Then the B admissible probability may be written

$$\begin{aligned}
\mathbb{P} &= 2(\mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3) \\
&= 2\left(\frac{1}{128\pi^2} \frac{\mathcal{A}_1}{\mathcal{B}_2^{3/2}} \delta_1^3 + \frac{1}{128\pi^2} \frac{\mathcal{A}_2}{\mathcal{B}_2^{3/2}} \delta_2^3 + \frac{1}{12\pi^2} \frac{(\mathcal{A}_3)^{3/2}}{\mathcal{B}_3^3} \delta_1^2 \delta_2^2\right) \\
&= 2\left(\frac{1}{128\pi^2} \frac{\mathcal{A}_1}{\mathcal{B}_2^{3/2}} \delta^3 + \frac{1}{128\pi^2} \frac{\mathcal{A}_2}{\mathcal{B}_2^{3/2}} \delta^3 + \frac{1}{12\pi^2} \frac{(\mathcal{A}_3)^{3/2}}{\mathcal{B}_3^3} \delta^4\right) \\
&\leq C_1 \delta^3.
\end{aligned}$$

We take $\mathcal{C}_1 = \max C_1$, where the maximum is taken over all G . This implies that for all squares where B admissibility is required, we have

$$\mathbb{P}[E_B] \leq \mathcal{C}_1 \delta^3.$$

By a similar argument, we can find

$$\mathbb{P}[E_I] \leq \mathcal{C}_2 \delta^4,$$

for all squares that require I -admissibility. The result follows from Theorem 3.8, [53]. \square

The above theorem establishes a $1/M^2$ convergence rate. However, in establishing this result we took the constants \mathcal{C}_1 and \mathcal{C}_2 to be the maximum admissible probabilities. While this does give the correct asymptotic convergence rate, the constants \mathcal{C}_1 and \mathcal{C}_2 are often very large and for a specified random field cannot be decreased. However, this theorem only holds for equi-spaced points. This suggests to increase the likelihood of making the correct

homology computations, we cannot sample with equi-distant points. In the next chapter, we will propose an algorithm to find the optimal location of our discretization points to maximize the correctness probability.

Chapter 5: Minimization Algorithm

5.1 Introduction

Due to the complex and implicit nature of the local probabilities, it seems unlikely to find an explicit formula that can show us where to sample our points. Instead, we now present an algorithmic approach that seeks to minimize the probability of making an incorrect homology computation. Instead of maximizing the probability of correctness, we instead minimize the probability of failure. This method uses a barrier method to transform a constrained minimization problem into an unconstrained problem. We then use a Newton type iteration scheme to update the sampling points. To illustrate the ideas, we separate the one dimensional and two dimensional cases.

5.2 Algorithm in One Dimension

To begin, we present the algorithm in one spatial dimension. This will make the main points of the algorithm clear. It is based on a barrier method in constrained optimization with a quasi-Newton method.

We can consider the problem as follows. Suppose we are given the domain $D = [a, b]$ and for each sub-interval $[x_k, x_{k+1}]$ we have the local probability of failure $\mathbb{P}_{lo}(x_k, x_{k+1})$. Then for the decomposition of D into M subintervals with $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$, the total probability of failure can be written

$$\mathbb{P}_T(\vec{x}) = \sum_{j=0}^{n-1} \mathbb{P}_{lo}(x_j, x_{j+1}),$$

where $\vec{x} = (x_1, x_2, \dots, x_{n-1})^t$. In particular, we view the total probability of failure as a function of the interior points of D since for the endpoints it must hold that $x_0 = a$ and $x_n = b$. The goal is to find $\min \mathbb{P}_T$. The constraints are then

$$c_1(\vec{x}) = x_1 - a > 0$$

$$c_2(\vec{x}) = x_2 - x_1 > 0$$

⋮

$$c_{n-1}(\vec{x}) = x_{n-1} - x_{n-2} > 0$$

$$c_n(\vec{x}) = b - x_{n-1} > 0,$$

which we write compactly as $\vec{C}(\vec{x}) > 0$. Using the Barrier method, we transform this problem to the unconstrained problem

$$\begin{aligned} & \min \left[\sum_{j=0}^{n-1} \mathbb{P}_{lo}(x_j, x_{j+1}) - \mu \sum_{k=1}^n \log(c_k(\vec{x})) \right] \\ &= \min [f(\vec{x}) - \mu g(\vec{x})] \\ &= \min \beta_\mu(\vec{x}). \end{aligned}$$

We can now use Newton's method to iteratively find the minimum. This suggests we use the scheme

$$\vec{x}_{k+1} = \vec{x}_k - H_k^{-1}(\vec{x}_k) \cdot \nabla \beta_\mu(\vec{x}_k),$$

where $\nabla \beta_\mu$ is the gradient of β_μ with respect to \vec{x} and H_k is an approximation of the Hessian of β_μ with respect to \vec{x} . We can write $\nabla \beta_\mu$ as $\nabla \beta_\mu = \nabla f - \mu \nabla g$. We will work with each

quantity separately. First we have

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(\vec{x}) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} f(\vec{x}) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} (\mathbb{P}_{lo}(x_0, x_1) + \mathbb{P}_{lo}(x_1, x_2)) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} (\mathbb{P}_{lo}(x_{n-2}, x_{n-1}) + \mathbb{P}_{lo}(x_{n-1}, x_n)) \end{bmatrix}.$$

We can break this into two terms $(\nabla f)_i = (\nabla (\mathbb{P}_{lo}(x_{i-1}, x_i) + \mathbb{P}_{lo}(x_i, x_{i+1})))_i$. Now we use a centered finite difference scheme with h sufficiently small. For this one-dimensional algorithm, we choose $h \approx 10^{-8}$. For each term, we have

$$\frac{\partial}{\partial x_i} \mathbb{P}_{lo}(x_{i-1}, x_i) \approx \frac{\mathbb{P}_{lo}(x_{i-1}, x_i + h) - \mathbb{P}_{lo}(x_{i-1}, x_i - h)}{2h},$$

and

$$\frac{\partial}{\partial x_i} \mathbb{P}_{lo}(x_i, x_{i+1}) \approx \frac{\mathbb{P}_{lo}(x_i + h, x_{i+1}) - \mathbb{P}_{lo}(x_i - h, x_{i+1})}{2h}.$$

Note that we must use $x_0 = a$ and $x_n = b$ in these expressions. For $\nabla \beta_\mu$, we also need to calculate ∇g . This can be written

$$\begin{aligned} \nabla g(\vec{x}) &= \begin{bmatrix} \frac{\partial}{\partial x_1} g(\vec{x}) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} g(\vec{x}) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} (\log(x_1 - a) + \log(x_2 - x_1)) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} (\log(x_{n-1} - x_{n-2}) + \log(b - x_{n-1})) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{x_1 - a} - \frac{1}{x_2 - x_1} \\ \vdots \\ \frac{1}{x_{n-1} - x_{n-2}} - \frac{1}{x_n - x_{n-1}} \end{bmatrix}. \end{aligned}$$

We can now combine these to write $\nabla\beta_\mu$.

The next item we need to calculate is H_k , the Hessian approximation of β_μ . We will not compute the Hessian at each step. We will use a Quasi-Newton update scheme instead. We choose a Broyden update method. For a fixed μ value, the process is as follows. Beginning with an initial approximation to the Hessian H , for our case we choose $H = I$, the identity matrix, we will update H as H_k on each iteration of the Newton scheme.

For the k -th iterate of \vec{x}_k , solve the linear system

$$H_k \vec{p}_k = -\nabla\beta_\mu(\vec{x}_k)$$

for the direction of descent \vec{p}_k . Once we have \vec{p}_k , we perform a simple line search and find α_k such that $\beta_\mu(\vec{x}_k + \alpha_k \vec{p}_k)$ is minimized and set

$$\vec{x}_{k+1} = \vec{x}_k + \alpha_k \vec{p}_k.$$

In particular, by finding α_k , we find the minimum of f in the direction of \vec{p}_k starting at \vec{x}_k . Next, we set $\vec{s}_k = \alpha_k \vec{p}_k$ and $\vec{y}_k = \nabla\beta_\mu(\vec{x}_{k+1}) - \nabla\beta_\mu(\vec{x}_k)$. Lastly, we update the Hessian approximation as

$$H_{k+1} = H_k + \frac{\vec{y}_k \vec{y}_k^t}{\vec{y}_k^t \vec{s}_k} - \frac{(H_k \vec{s}_k)(H_k \vec{s}_k)^t}{\vec{s}_k^t H_k \vec{s}_k}.$$

We perform this on each iteration in the Newton scheme. The primary reason for the above updating scheme is that it retains the positive definiteness of the Hessian approximation.

The above scheme was for a fixed μ value. The next part is to steadily decrease μ as this will allow the points to approach the boundary if needed. It should be noted that in our case, the minimized solution cannot lie on the boundary since we must have $x_i < x_{i+1}$ with strict inequality.

5.3 Two Dimensional Algorithm

In this section, we extend the numerical method to minimize the total probability of failure to two dimensions. While many of the ideas are the same, there are a few differences that must be addressed.

To begin, suppose we are working with the fixed domain $D = [a, b] \times [c, d]$ and decompose D into M^2 sub-rectangles. In particular, we have $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ and $c = y_0 < y_1 < \dots < y_{n-1} < y_n = d$. Denote $\vec{x} = (x_1, x_2, \dots, x_{n-1})^t$ and $\vec{y} = (y_1, y_2, \dots, y_{n-1})^t$. We can then write the total probability of failure as

$$P_T(\vec{x}, \vec{y}) = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{lo}(x_j, x_{j+1}, y_k, y_{k+1}),$$

where $\mathbb{P}_{lo}(x_j, x_{j+1}, y_k, y_{k+1})$ is the local probability of failure for the rectangle with bottom left point (x_j, y_k) and side lengths $\delta_{1,j} = x_{j+1} - x_j$ and $\delta_{2,k} = y_{k+1} - y_k$.

The minimization problem can then be stated as

$$\min P_T(\vec{x}, \vec{y}),$$

subject to the constraints

$$c_1(\vec{x}) = x_1 - a > 0$$

$$c_2(\vec{x}) = x_2 - x_1 > 0$$

$$\vdots$$

$$c_{n-1}(\vec{x}) = x_{n-1} - x_{n-2} > 0$$

$$c_n(\vec{x}) = b - x_{n-1} > 0$$

$$d_1(\vec{y}) = y_1 - c > 0$$

$$d_2(\vec{y}) = y_2 - y_1 > 0$$

$$\vdots$$

$$d_{n-1}(\vec{y}) = y_{n-1} - y_{n-2} > 0$$

$$d_n(\vec{y}) = d - y_{n-1} > 0.$$

Denote $\vec{z} = \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix}$ and $G(\vec{z}) = \begin{bmatrix} \vec{c}(\vec{x}) \\ \vec{d}(\vec{y}) \end{bmatrix} > 0$, component wise. Using the Barrier Method

again, we can transform the unconstrained problem into the unconstrained problem via

$$\min \left(\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{l_0}(x_j, x_{j+1}, y_k, y_{k+1}) - \mu \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \right)$$

where μ is the barrier parameter that will steadily be decreased. Compactly, this will be written $\min (f(\vec{z}) - \mu g(\vec{z})) = \min \beta_\mu(\vec{z})$, where $\beta_\mu : \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$.

We now provide details for the minimization scheme which is again accomplished using a Quasi-Newton method. To begin, fix the parameter μ . Initially specify an initial guess \vec{z}_0 and an approximation for the Hessian matrix H_0 , usually taken to be the identity matrix. Iterate over all i and do the following for each i solve $H_i \vec{p}_i = -\nabla_{\vec{x}, \vec{y}} \beta_\mu(\vec{z}_i)$. Then perform a line search to determine α_i such that $\beta_\mu(\vec{z}_i + \alpha_i \vec{p}_i)$ is a local minimum. Update $\vec{z}_{i+1} = \vec{z}_i + \alpha_i \vec{p}_i$. Next compute $s_i = \vec{z}_{i+1} - \vec{z}_i$ and $w_i = \nabla_{\vec{x}, \vec{y}} \beta_\mu(\vec{z}_{i+1}) - \nabla_{\vec{x}, \vec{y}} \beta_\mu(\vec{z}_i)$. Lastly, update B_{i+1} as

$$H_{i+1} = H_i - \frac{(H_i s_i)(H_i s_i)^t}{s_i^t H_i s_i} + \frac{w_i s_i^t}{w_i^t s_i}.$$

There are two items to address. The first is the Hessian approximation H_i . It should be noted that as the iterations continue H_i becomes a better approximation for the Hessian. By construction, H_i retains the crucial aspect of being positive definite since we are performing a search for a minimum.

The second issue to address is the gradient $\nabla_{\vec{x}, \vec{y}} \beta_\mu(\vec{z})$. Since we have two separate notions of admissibility, namely B and I -admissibility. Hence, we have different notions of local probability, which we must take into consideration in this construction. It should be clear that

$$\begin{aligned} \nabla_{\vec{x}, \vec{y}}(\beta_\mu(\vec{z})) &= \nabla_{\vec{x}, \vec{y}}(P_T(\vec{z}) - \mu g(\vec{z})) \\ &= \nabla_{\vec{x}, \vec{y}}(P_T(\vec{z})) - \mu \nabla_{\vec{x}, \vec{y}}(g(\vec{z})), \end{aligned}$$

so we can split the gradient into two terms which we will handle separately. The first term is

$$\nabla_{\vec{x}, \vec{y}} P_T(\vec{z}) = \begin{bmatrix} \frac{\partial}{\partial x_1} P_T(\vec{z}) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} P_T(\vec{z}) \\ \frac{\partial}{\partial y_1} P_T(\vec{z}) \\ \vdots \\ \frac{\partial}{\partial y_{n-1}} P_T(\vec{z}) \end{bmatrix},$$

which is the same as

$$\nabla_{\vec{x}, \vec{y}} P_T(\vec{z}) = \begin{bmatrix} \frac{\partial}{\partial x_1} \left(\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{l_0}(x_j, x_{j+1}, y_k, y_{k+1}) \right) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} \left(\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{l_0}(x_j, x_{j+1}, y_k, y_{k+1}) \right) \\ \frac{\partial}{\partial y_1} \left(\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{l_0}(x_j, x_{j+1}, y_k, y_{k+1}) \right) \\ \vdots \\ \frac{\partial}{\partial y_{n-1}} \left(\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{P}_{l_0}(x_j, x_{j+1}, y_k, y_{k+1}) \right) \end{bmatrix}.$$

We can now do a few simplifications. Since each local probability depends on at most two points, we can simplify this to

$$\nabla_{\vec{x}, \vec{y}} P_T(\vec{z}) = \begin{bmatrix} \frac{\partial}{\partial x_1} \left(\sum_{k=0}^{n-1} \mathbb{P}_{lo}(x_0, x_1, y_k, y_{k+1}) + \mathbb{P}_{lo}(x_1, x_2, y_k, y_{k+1}) \right) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} \left(\sum_{k=0}^{n-1} \mathbb{P}_{lo}(x_{n-2}, x_{n-1}, y_k, y_{k+1}) + \mathbb{P}_{lo}(x_{n-1}, x_n, y_k, y_{k+1}) \right) \\ \frac{\partial}{\partial y_1} \left(\sum_{j=0}^{n-1} \mathbb{P}_{lo}(x_j, x_{j+1}, y_0, y_1) + \mathbb{P}_{lo}(x_j, x_{j+1}, y_1, y_2) \right) \\ \vdots \\ \frac{\partial}{\partial y_{n-1}} \left(\sum_{j=0}^{n-1} \mathbb{P}_{lo}(x_j, x_{j+1}, y_{n-2}, y_{n-1}) + \mathbb{P}_{lo}(x_j, x_{j+1}, y_{n-1}, y_n) \right) \end{bmatrix}.$$

The crucial difference here now comes from the boundary terms. Recall $x_0 = a, x_n = b, y_0 = c$, and $y_n = d$. So these points are fixed and cannot move. Also, in each of the above partial derivatives, we must deal with local probability for the boundary.

Lastly, due to the complex nature of the terms involved, we perform a centered finite difference approximation for each partial derivative. It is this step that requires a careful analysis. For simplicity, we will only calculate $\frac{\partial}{\partial x_i}$ for all $i = 1, \dots, n-1$, as the same analysis can be performed on the y_i 's. To begin, assume $i = 2, \dots, n-2$. We will handle the cases $i = 1$ and $i = n-1$ separately. We may explicitly write a term in the gradient as

$$\begin{aligned} \frac{\partial}{\partial x_i} P_T &= \frac{\partial}{\partial x_i} \left(\sum_{k=0}^{n-1} \mathbb{P}_{lo}(x_{i-1}, x_i, y_k, y_{k+1}) + \sum_{k=0}^{n-1} \mathbb{P}_{lo}(x_i, x_{i+1}, y_k, y_{k+1}) \right) \\ &= \frac{\partial}{\partial x_i} \left(\mathbb{P}_{lo}^B(x_{i-1}, x_i, y_0, y_1) + \sum_{k=1}^{n-2} \mathbb{P}_{lo}^I(x_{i-1}, x_i, y_k, y_{k+1}) + \mathbb{P}_{lo}^B(x_{i-1}, x_i, y_{n-1}, y_n) \right. \\ &\quad \left. + \mathbb{P}_{lo}^B(x_i, x_{i+1}, y_0, y_1) + \sum_{k=1}^{n-2} \mathbb{P}_{lo}(x_i, x_{i+1}, y_k, y_{k+1}) + \mathbb{P}_{lo}^B(x_i, x_{i+1}, y_{n-1}, y_n) \right), \end{aligned}$$

where we have separated terms that touch the boundary out as \mathbb{P}_{lo}^B from the terms that are contained in the interior \mathbb{P}_{lo}^I . On each of these terms we can do a standard centered finite difference scheme as

$$\frac{\partial}{\partial x_i} \mathbb{P}_{lo}^{B/I}(x_{i-1}, x_i, y_k, y_{k+1}) \approx \frac{\mathbb{P}_{lo}^{B/I}(x_{i-1}, x_i + h, y_k, y_{k+1}) - \mathbb{P}_{lo}^{B/I}(x_{i-1}, x_i - h, y_k, y_{k+1})}{2h},$$

and

$$\frac{\partial}{\partial x_i} \mathbb{P}_{lo}^{B/I}(x_i, x_{i+1}, y_k, y_{k+1}) \approx \frac{\mathbb{P}_{lo}^{B/I}(x_i + h, x_{i+1}, y_k, y_{k+1}) - \mathbb{P}_{lo}^{B/I}(x_i + h, x_{i+1} - h, y_k, y_{k+1})}{2h},$$

for h sufficiently small. We chose to use values of h ranging from $h = 10^{-6}$ to $h = 10^{-8}$. For values smaller than $h = 10^{-8}$, roundoff errors begin to cause the points to violate the constraints c_i and d_i .

Note that for the cases $i = 1$ and $i = n - 1$, the situation will be similar. However, we should note that one side will consist entirely of boundary terms.

The next item we must compute is the gradient of the log-constraints. In particular, we must calculate

$$\nabla_{\vec{x}, \vec{y}} g(\vec{z}) = \begin{bmatrix} \frac{\partial}{\partial x_1} g(\vec{z}) \\ \frac{\partial}{\partial x_2} g(\vec{z}) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} g(\vec{z}) \\ \frac{\partial}{\partial y_1} g(\vec{z}) \\ \frac{\partial}{\partial y_2} g(\vec{z}) \\ \vdots \\ \frac{\partial}{\partial y_{n-1}} g(\vec{z}) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \\ \frac{\partial}{\partial x_2} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \\ \vdots \\ \frac{\partial}{\partial x_{n-1}} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \\ \frac{\partial}{\partial y_1} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \\ \frac{\partial}{\partial y_2} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \\ \vdots \\ \frac{\partial}{\partial y_{n-1}} \left(\sum_{l=1}^n \log c_l(\vec{x}) + \sum_{m=1}^n \log d_m(\vec{y}) \right) \end{bmatrix}.$$

However, this is straightforward since c_l depends only on the x_i values and d_m only depends on the y_i values. In this case, no finite difference approximations are needed and we can calculate the gradient explicitly as

$$\nabla_{\vec{x}, \vec{y}} g(\vec{z}) = \begin{bmatrix} \frac{1}{x_1 - a} - \frac{1}{x_2 - x_1} \\ \frac{1}{x_2 - x_1} - \frac{1}{x_3 - x_2} \\ \vdots \\ \frac{1}{x_{n-1} - x_{n-2}} - \frac{1}{b - x_{n-1}} \\ \frac{1}{y_1 - c} - \frac{1}{y_2 - y_1} \\ \frac{1}{y_2 - y_1} - \frac{1}{y_3 - y_2} \\ \vdots \\ \frac{1}{y_{n-1} - y_{n-2}} - \frac{1}{d - y_{n-1}} \end{bmatrix}.$$

Chapter 6: Computational Results

6.1 Introduction

In this chapter, we discuss computational results of our methods. We will discuss results of three particular random fields. First, we will work with random Chebyshev series of the form

$$f_1(x, y, \omega) = \sum_{j=0}^N \sum_{k=0}^N a_{j,k} g_{j,k}(\omega) \cos(j \arccos x) \cos(k \arccos y),$$

where $a_{j,k}$ are coefficients to be specified and $g_{j,k}(\omega)$ are independent Gaussian random variables with mean zero and variance one. The domain for f_1 will be $[-1, 1] \times [-1, 1]$. We will also be discussing random Cosine series of the form

$$f_2(x, y, \omega) = \sum_{j=0}^N \sum_{k=0}^N a_{j,k} g_{j,k}(\omega) \cos j\pi x \cos k\pi y,$$

with $a_{j,k}$ specified coefficients, $g_{j,k}(\omega)$ independent Gaussian random variables with mean zero and variance one, and domain $[0, 1] \times [0, 1]$.

6.2 Random Chebyshev Series

We begin our discussion of the computational results with the random Chebyshev Series of the form

$$f_1(x, y, \omega) = \sum_{j=0}^N \sum_{k=0}^N a_{j,k} g_{j,k}(\omega) \cos(j \arccos x) \cos(k \arccos y),$$

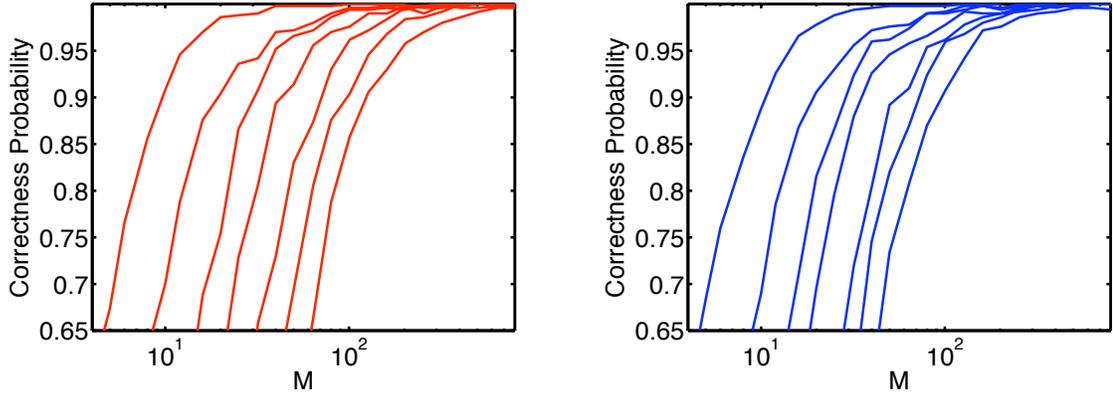


Figure 6.1: Correctness Probability of Random Chebyshev Series for Different Orders.

where $a_{j,k}$ are coefficients to be specified, $g_{j,k}(\omega)$ are independent Gaussian random variables with mean zero and variance one, with domain $D = [-1, 1] \times [-1, 1]$. For simplicity, we will assume the coefficients $a_{j,k}$ will be 1 for all j, k .

The simulations we performed were for orders $N = 3, \dots, 10$. For each order, we performed 500 simulations. For each simulation, we used the new validation routine to compute the homology of N^+ . Next we performed homology calculations for a variety of M values. For each M value, we computed the homology using equi-spaced points and also the points from our minimization routine.

The images in Figure 6.1 shows the results of these simulations. Each curve shows the probability of the correct homology as a function of the discretization size M for increasing orders N . The red curves are for equi-spaced points and the blue curves are the non-uniform sampling. Notice for small orders N , the uniform sampling gives a better probability but the minimization routine quickly becomes much better. For a specified probability threshold, we are interested in the asymptotic behavior of the discretization size M . We fit a curve of the form $M = C_N N^{\alpha_p}$ for different probability thresholds. Figure 6.2 shows the exponents α_p for $p = .7, \dots, .95$. The uniform discretization is shown in red and the optimal sampling is shown in blue. Notice that the non-uniform sampling decreases the exponent by .5. Also in Figure 6.2, we show the curves $M = C_N N^{\alpha_p}$ for the threshold probabilities $p = .9$ and $p = .95$. The red curves is the uniform sampling, the blue curves are the optimal sampling,

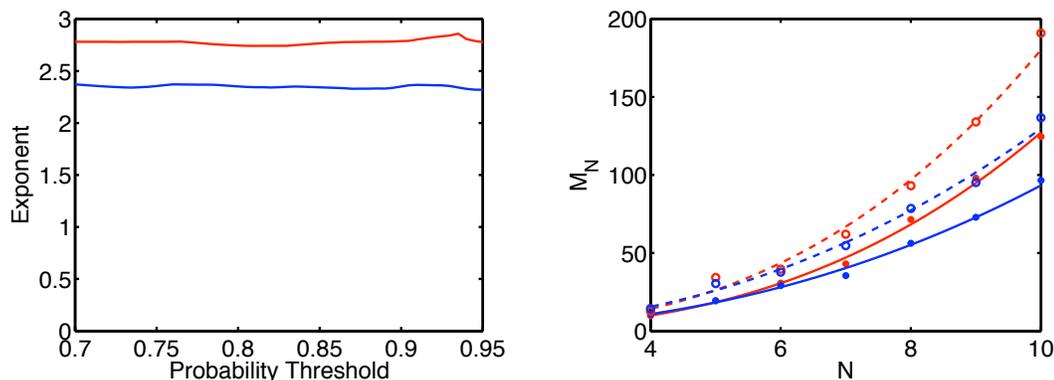


Figure 6.2: Exponent for Probability Threshold for Chebyshev Polynomials

the solid lines are for $p = .9$, and the dashed lines are for the probability threshold $p = .95$.

The minimization routine for these series moves points closer to the boundary. This results from the Chebyshev basis functions having more sign changes closer to the boundary. Notice that the validation refines closer to the boundary, while the interior has relatively large squares that were validated. We give two realizations for dimensions 3, 5, 9, and 14. In order to understand why the sampling needs to be closer to the boundary, we will examine what happens in one dimension. In one dimension, the random field is

$$u(x, \omega) = \sum_{j=1}^N g_j(\omega) \cos(j(\arccos(x))),$$

where $g_j(\omega)$ is an independent Gaussian random variable with mean zero and variance one defined over the domain $D = [-1, 1]$. The density of zeros for this random field is shown in Figure 6.3 along with the function $\mathcal{C}_0(x)$ in one dimension. Notice that the density of zeros is much higher close to the boundary than the interior. This suggests that in one dimension we need to sample more points closer to the boundary. While there is not an analog to density of zeros in two dimensions, it does suggest that the nodal lines will be closer to the boundary. This suggests that we must again sample closer to the boundary in

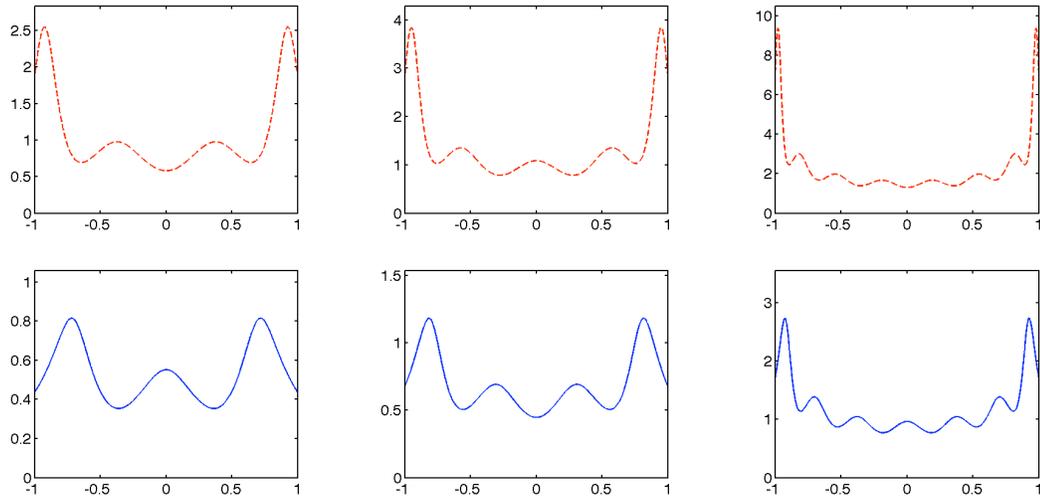


Figure 6.3: Density of zeros function (*top*) and the function $C_0(x)$ (*bottom*).

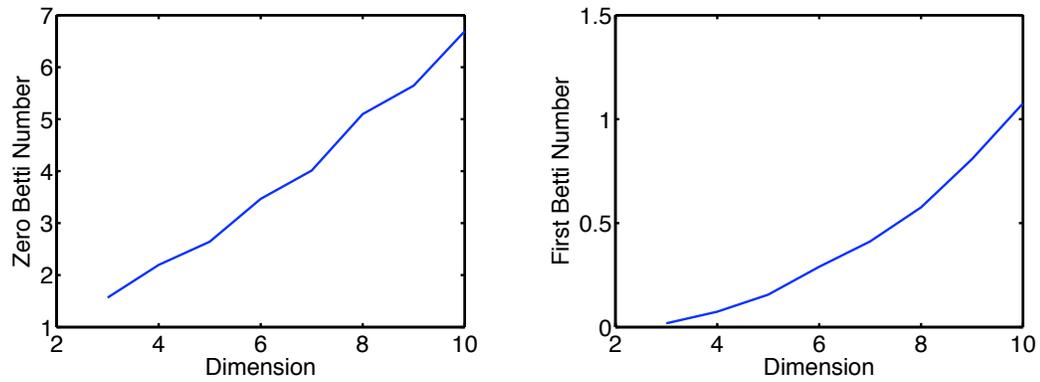


Figure 6.4: Betti Numbers for Chebyshev Series

two dimensions.

Lastly, we show figures for the Betti numbers as a function of the dimension. These are shown in Figure 6.4. The first image is the average zero'th Betti number and the second image is the average first Betti number.

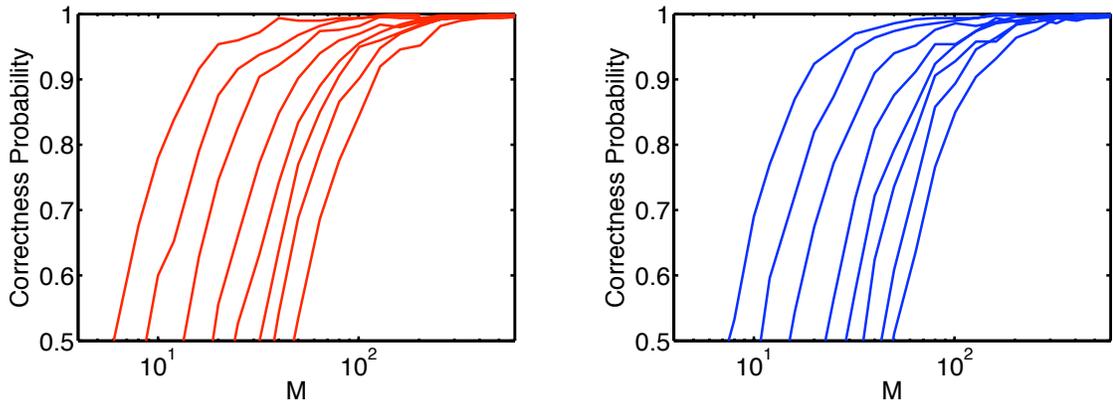


Figure 6.5: Correctness Probability of Random Cosine Series for Different Dimensions.

6.3 Random Cosine Series

In this section, we give results pertaining to the cosine random field of the form

$$f_2(x, y, \omega) = \sum_{j=0}^N \sum_{k=0}^N a_{j,k} g_{j,k}(\omega) \cos j\pi x \cos k\pi y,$$

where $a_{j,k}$ are coefficients and $g_{j,k}(\omega)$ are independent Gaussian random variables with mean zero and variance one. The domain for this random field will be $[0, 1] \times [0, 1]$. For simplicity, we will assume $a_{j,k} = 1$ for all $1 \leq j, k \leq N$. This process is similar to that of the random Chebyshev series. We do 500 simulations for $N = 3, \dots, 10$. For each simulation, we use the validation routine to correctly identify the homology of the positive nodal domain. Then we partition the domain into M^2 rectangles and compute the homology of each discretization. We use equi-partitioned squares and the points obtained from the minimization algorithm.

The images in Figure 6.5 shows the results of these simulations. Each curve shows the probability of the correct homology as a function of the discretization size M for increasing dimensions N . The red curves are for equi-spaced points and the blue curves are the non-uniform sampling. Notice that the uniform sampling appears to give better probabilities. However, this is only true for small dimensions N . As N increases, the minimization routine

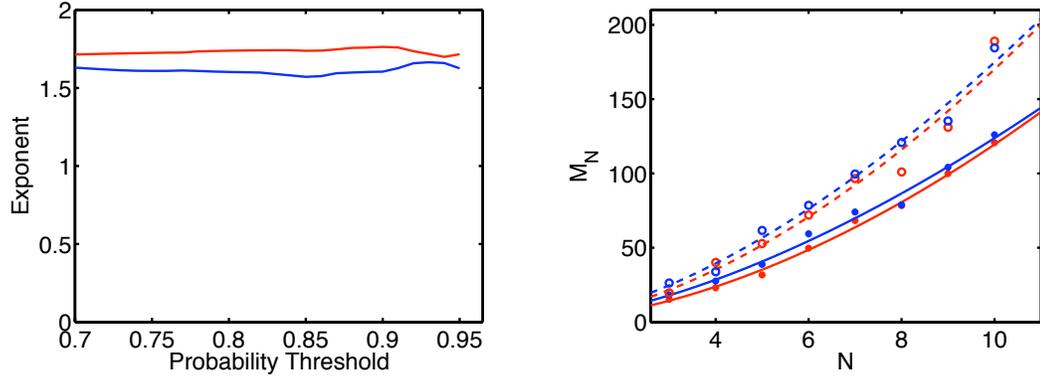


Figure 6.6: Exponent for Probability Threshold for Cosine Series

begins to give better results. To understand this point better, we investigate the best fit curve for specified probabilities. Figure 6.6 shows the exponents α_p for $p = .7, \dots, .95$. The uniform discretization is shown in red and the optimal sampling is shown in blue. Notice that the non-uniform sampling decreases the exponent by .1. Also in Figure 6.6, we show the curves $M = C_N N^{\alpha_p}$ for the threshold probabilities $p = .9$ and $p = .95$. The red curves are the uniform sampling, the blue curves are the optimal sampling, the solid lines are for $p = .9$, and the dashed lines are for the probability threshold $p = .95$.

The exponent for the uniform sampling is higher than the non-uniform counterpart. This shows that for higher dimensions, the minimization routine gives one of two points depending on the viewpoint. We can interpret this as giving a smaller discretization value M for a specified probability tolerance, or we can interpret this as giving a higher probability for a fixed discretization size. Another interesting fact can be observed from Figures 6.6 and 6.2. The decrease in the exponent for the cosine series is much smaller than the decrease for the Chebyshev series.

This difference can be explained by again examining the density of zeros for the random Cosine fields. In one dimension, the random field is

$$u(x, \omega) = \sum_{j=1}^N g_j(\omega) \cos(2\pi jx)$$

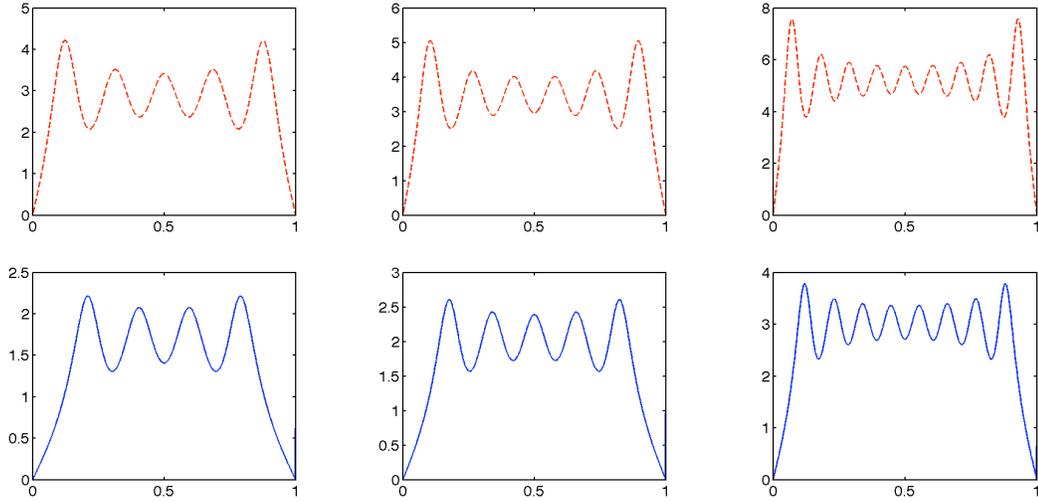


Figure 6.7: Density of zeros function (*top*) and the function $\mathcal{C}_0(x)$ (*bottom*).

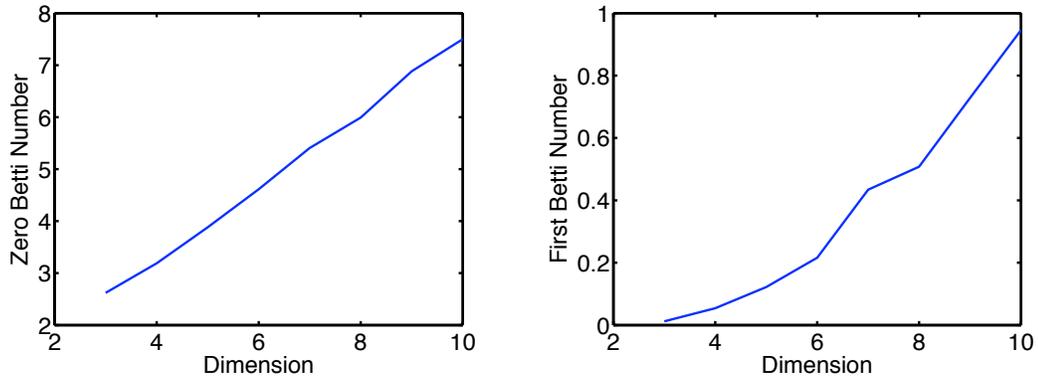


Figure 6.8: Betti Numbers for Cosine Series

where $g_j(\omega)$ is an independent Gaussian random variable with mean zero and variance one defined over the domain $D = [0, 1]$. The density of zeros for this random field is shown in Figure 6.7 along with the function $\mathcal{C}_0(x)$ in one dimension. Notice that the density of zeros is fairly constant in the interior. This shows that a uniform sampling process can almost capture the topology. We should expect the same type of behavior in two dimensions.

Lastly, we show the asymptotic behavior of the expected Betti numbers for the random cosine series. These are shown in Figure 6.8.

Chapter 7: Conclusion

7.1 Conclusion

In this final chapter, we recap the major contributions of this dissertation. Our main goal is to compute the homology of nodal domains N^\pm for a function $f : D \rightarrow \mathbb{R}$, for an appropriate function f and domain D . The homology groups are denoted $H_*(N^\pm)$. However, from a computational point of view, we must instead work with cubical approximations to the nodal domains, denoted Q_M^\pm , for a discretization size M . This discretization can of course lead to incorrect homology computations.

One approach to computing the homology of the nodal domains is to use validated homology. This algorithm was first presented in [30]. First an initial decomposition of the domain is made. The authors explicitly laid out verification steps that are performed on each box in the domain. These verification steps use interval arithmetic to obtain appropriate bounds on the function and its derivatives. If the verification step is passed, the box is then added to either Q_M^+ or Q_M^- as indicated by the verification step.

The authors proved that if the algorithm terminates successfully, the homology computation on the grid produced will be correct. However, there are a few shortcomings to this original algorithm. The primary shortcoming came from the use of interval arithmetic and outward rounding. Interval arithmetic was used to obtain bounds on the function and its derivatives. However, these bounds are often overextended far beyond the true range of the function. This results in frequent failure of the verification steps and as a result, more subdivisions are required.

In this dissertation, we proposed a new method to test for interval bounds. Since we only need a lower bound or upper bound as required from the verification step, we found a method to push the lower bounds upward. This is done by subdividing the current box and

putting the boxes into a list order by the smallest lower bound. We then drove this smallest lower bound upward until we reached a specified tolerance or the lower bound was strictly positive. This process results in a grid that is usually as coarse as possible.

The other main problem with the original algorithm was the subdivision lines can line up with the nodal lines. This phenomena thus results in the verification steps failing close to the nodal lines. As a result, many subdivisions must be performed to resolve the topology. We have proposed a modification to the original algorithm by randomly subdividing. We fix two ratios, in our case we chose the Golden Ratio, and if the verification step fails, we then randomly choose amongst these two ratios and subdivide the box into two boxes by the ratio chosen. This procedure seems to eliminate any grid alignment issues.

The next major contribution of this dissertation is an extension of probabilistic bounds for homology computations of random fields. In [53] and [54], the authors present a probabilistic formula for the correctness of random field homology computations. The authors showed bounds for homogenous random fields in [53] in one and two spatial dimensions. In [54], the authors gave bounds for non-homogenous random fields. In this work, the authors also state the optimal location of the sampled points.

In this dissertation, we have extended the results of these two previous works to include non-homogenous random fields in two spatial dimensions. These bounds first find the local probability of certain sign configurations occurring. We then found bounds that these sign configurations occur on any of the dyadic subdivisions of the box in question. The total probability of failure is then the sum of these local failure probabilities.

This total probability gives little indication as to the optimal location of the points to be sampled. We then proposed an algorithm to determine the best sampling for random fields. This is a quasi-Newton method that transforms the constrained optimization problem into an unconstrained optimization problem.

7.2 Further Directions

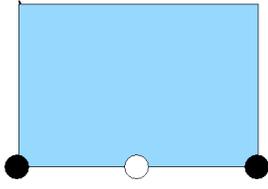
This line of research has many interesting directions one can take.

Since we have probability bounds for both homogenous and non-homogenous random fields in one and two spatial dimensions, the first direction would be to extend these to higher dimensions. This will require establishing the notion of admissibility in higher dimensions and then finding the local probability that each forbidden sign configuration occurs.

Another possible direction to take this line of research is a three dimensional validation routine. The first step in this direction is to determine new verification steps based on what happens on the vertices.

Appendix A : Local Probabilities

In this first section, we present tables showing the probabilities that each sign configuration occurs. For each figure and table, we assume the random field is $u : D \times \Omega \rightarrow \mathbb{R}$ and u is a Gaussian random field with mean zero. In addition, we assume that each box is sampled at the point (x, y) and the edge length along the x -axis is δ_1 and the edge length along the y -axis is δ_2 .



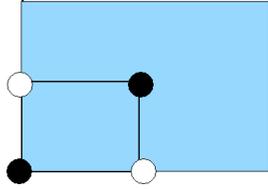
$\det \mathcal{C}_1 =$	$\mathcal{A}_1 \delta_1^6$
$\lambda_1 =$	$\frac{\mathcal{A}_2}{\mathcal{B}_2} \delta_1^4$
$\nu_1 =$	$(\sqrt{\frac{1}{6}}, -\sqrt{\frac{2}{3}}, \sqrt{\frac{1}{6}})$
$\mathbb{P}_1 =$	$\frac{1}{128\pi^2} \frac{\mathcal{A}_1}{\mathcal{B}_1^{3/2}} \delta_1^3$



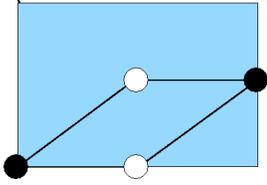
$\det \mathcal{C}_2 =$	$\mathcal{A}_2 \delta_2^6$
$\lambda_1 =$	$\frac{\mathcal{A}_2}{\mathcal{B}_2} \delta_2^4$
$\nu_2 =$	$(\sqrt{\frac{1}{6}}, -\sqrt{\frac{2}{3}}, \sqrt{\frac{1}{6}})$
$\mathbb{P}_2 =$	$\frac{1}{128\pi^2} \frac{\mathcal{A}_2}{\mathcal{B}_2^{3/2}} \delta_2^3$



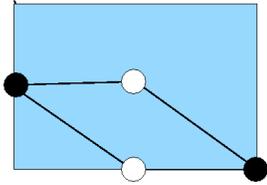
$\det \mathcal{C}_3 =$	$\mathcal{A}_3 \delta_1^4 \delta_2^4$
$\lambda_1 =$	$-\frac{1}{4} \frac{\mathcal{A}_3}{\mathcal{B}_3} \delta_1^2 \delta_2^2$
$\nu_3 =$	$(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$
$\mathbb{P}_3 =$	$\frac{1}{12\pi^2} \frac{(\mathcal{A}_3)^{3/2}}{\mathcal{B}_3^2} \delta_1^2 \delta_2^2$



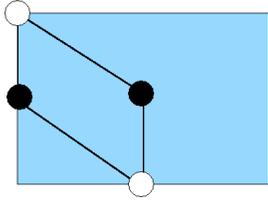
$\det \mathcal{C}_4 =$	$\frac{1}{256} \mathcal{A}_3 \delta_1^4 \delta_2^4$
$\lambda_1 =$	$-\frac{1}{64} \frac{\mathcal{A}_3}{\mathcal{B}_3} \delta_1^2 \delta_2^2$
$\nu_4 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$
$\mathbb{P}_4 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_3)^{3/2}}{\mathcal{B}_3^3} \delta_1^2 \delta_2^2$



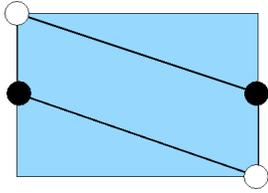
$\det \mathcal{C}_5 =$	$\frac{1}{256} (\mathcal{A}_{4,1} \delta_1^6 \delta_2^2 + \mathcal{A}_{4,2} \delta_1^5 \delta_2^3 + \mathcal{A}_{4,3} \delta_1^4 \delta_2^4)$
$\lambda_1 =$	$-\frac{1}{64} \frac{(\mathcal{A}_{4,1} \delta_1^4 + \mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{4,3} \delta_1^2 \delta_2^2)}{\mathcal{B}_4}$
$\nu_5 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$
$\mathbb{P}_5 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{4,1} \delta_1^4 + \mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{4,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_4^2 \delta_1 \delta_2}$



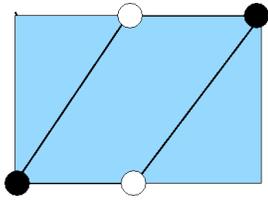
$\det \mathcal{C}_6 =$	$\frac{1}{256} (\mathcal{A}_{4,1} \delta_1^6 \delta_2^2 + \mathcal{A}_{4,2} \delta_1^5 \delta_2^3 + \mathcal{A}_{4,3} \delta_1^4 \delta_2^4)$
$\lambda_1 =$	$-\frac{1}{64} \frac{(\mathcal{A}_{4,1} \delta_1^4 + \mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{4,3} \delta_1^2 \delta_2^2)}{\mathcal{B}_4}$
$\nu_6 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$
$\mathbb{P}_6 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{4,1} \delta_1^4 + \mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + \mathcal{A}_{4,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_4^2 \delta_1 \delta_2}$



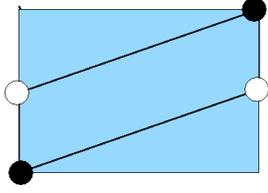
$\det \mathcal{C}_7 =$	$\frac{1}{256} (\mathcal{A}_{5,1} \delta_1^2 \delta_2^6 + \mathcal{A}_{5,2} \delta_1^3 \delta_2^5 + \mathcal{A}_{5,3} \delta_1^4 \delta_2^4)$
$\lambda_1 =$	$-\frac{1}{64} \frac{\mathcal{A}_{5,1} \delta_2^4 + \mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2}{\mathcal{B}_3}$
$\nu_7 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_7 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{5,1} \delta_2^3 + \mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



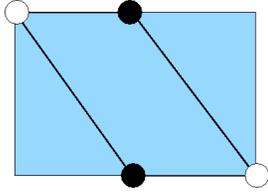
$\det \mathcal{C}_8 =$	$\left(\frac{1}{64} \mathcal{A}_{5,1} \delta_1^2 \delta_2^6 + \frac{1}{32} \mathcal{A}_{5,2} \delta_1^3 \delta_2^5 + \frac{1}{16} \mathcal{A}_{5,3} \delta_1^4 \delta_2^4 \right)$
$\lambda_1 =$	$-\frac{1}{64} \frac{\mathcal{A}_{5,1} \delta_2^4 + 2\mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + 4\mathcal{A}_{5,3} \delta_1^2 \delta_2^2}{\mathcal{B}_3}$
$\nu_8 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_8 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{5,1} \delta_2^4 + 2\mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + 4\mathcal{A}_{5,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



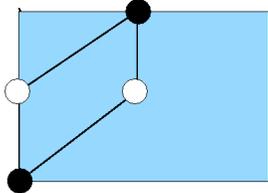
$\det \mathcal{C}_9 =$	$\left(\frac{1}{64} \mathcal{A}_{4,1} \delta_1^6 \delta_2^2 + \frac{1}{32} \mathcal{A}_{4,2} \delta_1^5 \delta_2^3 + \frac{1}{16} \mathcal{A}_{4,3} \delta_1^4 \delta_2^4 \right)$
$\lambda_1 =$	$-\frac{1}{64} \frac{(\mathcal{A}_{4,1} \delta_1^4 + 2\mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + 4\mathcal{A}_{4,3} \delta_1^2 \delta_2^2)}{\mathcal{B}_3}$
$\nu_9 =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_9 =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{4,1} \delta_1^4 + 2\mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + 4\mathcal{A}_{4,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



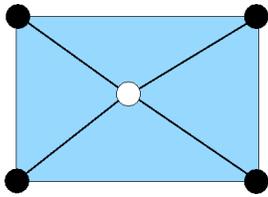
$\det \mathcal{C}_{11} =$	$\left(\frac{1}{64} \mathcal{A}_{5,1} \delta_1^2 \delta_2^6 + \frac{1}{32} \mathcal{A}_{5,2} \delta_1^3 \delta_2^5 + \frac{1}{16} \mathcal{A}_{5,3} \delta_1^4 \delta_2^4 \right)$
$\lambda_1 =$	$-\frac{1}{64} \frac{\mathcal{A}_{5,1} \delta_2^4 + 2\mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + 4\mathcal{A}_{5,3} \delta_1^2 \delta_2^2}{\mathcal{B}_3}$
$\nu_{11} =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_{11} =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{5,1} \delta_2^3 + 2\mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + 4\mathcal{A}_{5,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



$\det \mathcal{C}_{12} =$	$\left(\frac{1}{64} \mathcal{A}_{4,1} \delta_1^6 \delta_2^2 + \frac{1}{32} \mathcal{A}_{4,2} \delta_1^5 \delta_2^3 + \frac{1}{16} \mathcal{A}_{4,3} \delta_1^4 \delta_2^4 \right)$
$\lambda_1 =$	$-\frac{1}{64} \frac{(\mathcal{A}_{4,1} \delta_1^4 + 2\mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + 4\mathcal{A}_{4,3} \delta_1^2 \delta_2^2)}{\mathcal{B}_3}$
$\nu_{12} =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_{12} =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{4,1} \delta_1^4 + 2\mathcal{A}_{4,2} \delta_1^3 \delta_2^1 + 4\mathcal{A}_{4,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



$\det \mathcal{C}_{13} =$	$\frac{1}{256} (\mathcal{A}_{5,1} \delta_1^2 \delta_2^6 + \mathcal{A}_{5,2} \delta_1^3 \delta_2^5 + \mathcal{A}_{5,3} \delta_1^4 \delta_2^4)$
$\lambda_1 =$	$-\frac{1}{64} \frac{\mathcal{A}_{5,1} \delta_2^4 + \mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2}{\mathcal{B}_3}$
$\nu_{13} =$	$\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$
$\mathbb{P}_{13} =$	$\frac{1}{192\pi^2} \frac{(\mathcal{A}_{5,1} \delta_2^4 + \mathcal{A}_{5,2} \delta_1^1 \delta_2^3 + \mathcal{A}_{5,3} \delta_1^2 \delta_2^2)^{3/2}}{\mathcal{B}_3^2 \delta_1 \delta_2}$



$\det \mathcal{C}_{10} =$	$\frac{1}{64} \mathcal{A}_{10}$
$\lambda_1 =$	$\frac{-\mathcal{B}_{10} - \sqrt{\mathcal{B}_{10}^2 - 4\mathcal{A}_{10}\mathcal{C}_{10}}}{2\mathcal{C}_{10}}$
$\nu_{10} =$	$\frac{1}{2\sqrt{5}} (1, 1, -4, 1, 1)$
$\mathbb{P}_{10} =$	$\frac{1}{200\sqrt{5}\pi^2} \sqrt{\frac{\lambda_{10}^5}{\mathcal{A}_{10}}}$

Appendix B : Coefficients

$$\mathcal{A}_1 = r_{00}r_{1122}r_{12} - r_{00}r_{112}r_{122} - r_1r_{1122}r_2 + r_{11}r_{122}r_2 + r_1r_{112}r_{22} - r_{11}r_{12}r_{22}$$

$$\mathcal{A}_2 = r_{00}r_{3344}r_{34} - r_{00}r_{334}r_{344} - r_3r_{3344}r_4 + r_{33}r_{344}r_4 + r_3r_{334}r_{44} - r_{33}r_{34}r_{44}$$

$$\mathcal{A}_3 = -r_{13}r_{14}r_2r_{234} + r_{13}r_{14}r_{23}r_{24} - r_{124}r_{134}r_2r_3 + r_{1234}r_{14}r_2r_3 + r_{12}r_{134}r_{24}r_3$$

$$- r_{123}r_{14}r_{24}r_3 + r_{124}r_{13}r_2r_{34} - r_{12}r_{13}r_{24}r_{34} + r_{00}(r_{124}r_{134}r_{23} - r_{1234}r_{14}r_{23})$$

$$- r_{12}r_{134}r_{234} + r_{123}r_{14}r_{234} + r_{12}r_{1234}r_{34} - r_{123}r_{124}r_{34}) - r_{124}r_{13}r_{23}r_4$$

$$+ r_{12}r_{13}r_{234}r_4 - r_{12}r_{1234}r_3r_4 + r_{123}r_{124}r_3r_4 + r_1(r_{134}r_2r_{234} - r_{134}r_{23}r_{24})$$

$$- r_{1234}r_2r_{34} + r_{123}r_{24}r_{34} + r_{1234}r_{23}r_4 - r_{123}r_{234}r_4).$$

$$\mathcal{A}_{4,1} = (r_{11}r_{14}r_{22}r_{23} - r_{11}r_{14}r_2r_{232} - r_{114}r_{122}r_2r_3 + r_{1122}r_{14}r_2r_3 + r_{114}r_{12}r_{22}r_3$$

$$- r_{112}r_{14}r_{22}r_3 + r_{11}r_{122}r_2r_{34} - r_{11}r_{12}r_{22}r_{34} + r_{00}(r_{114}r_{122}r_{23} - r_{1122}r_{14}r_{23})$$

$$- r_{114}r_{12}r_{232} + r_{112}r_{14}r_{232} + r_{1122}r_{12}r_{34} - r_{112}r_{122}r_{34}) - r_{11}r_{122}r_{23}r_4$$

$$+ r_{11}r_{12}r_{232}r_4 - r_{1122}r_{12}r_3r_4 + r_{112}r_{122}r_3r_4 + r_1(-r_{114}r_{22}r_{23} + r_{114}r_2r_{232},$$

$$- r_{1122}r_2r_{34} + r_{112}r_{22}r_{34} + r_{1122}r_{23}r_4 - r_{112}r_{232}r_4))$$

$$\begin{aligned}
\mathcal{A}_{4,2} = & (r_{13}r_{14}r_{22}r_{23} - r_{13}r_{14}r_2r_{232} - r_{11}r_{14}r_2r_{234} + r_{11}r_{14}r_{23}r_{24} - r_{114}r_{124}r_2r_3 \\
& - r_{122}r_{134}r_2r_3 + r_{1124}r_{14}r_2r_3 + r_{1223}r_{14}r_2r_3 + r_{12}r_{134}r_{22}r_3 - r_{123}r_{14}r_{22}r_3 \\
& + r_{114}r_{12}r_{24}r_3 - r_{112}r_{14}r_{24}r_3 + r_{111}r_{124}r_2r_{34} + r_{122}r_{13}r_2r_{34} - r_{12}r_{13}r_{22}r_{34} \\
& - r_{11}r_{12}r_{24}r_{34} + r_{00}(r_{114}r_{124}r_{23} + r_{122}r_{134}r_{23} - r_{1124}r_{14}r_{23} - r_{1223}r_{14}r_{23} \\
& - r_{12}r_{134}r_{232} + r_{123}r_{14}r_{232} - r_{114}r_{12}r_{234} + r_{112}r_{14}r_{234} + r_{1124}r_{12}r_{34} \\
& + r_{12}r_{1223}r_{34} - r_{122}r_{123}r_{34} - r_{112}r_{124}r_{34}) - r_{11}r_{124}r_{23}r_4 - r_{122}r_{13}r_{23}r_4 \\
& + r_{12}r_{13}r_{232}r_4 + r_{11}r_{12}r_{234}r_4 - r_{1124}r_{12}r_3r_4 - r_{12}r_{1223}r_3r_4 + r_{122}r_{123}r_3r_4 \\
& + r_{112}r_{124}r_3r_4 + r_1(-r_{134}r_{22}r_{23} + r_{134}r_2r_{232} + r_{114}r_2r_{234} - r_{114}r_{23}r_{24} \\
& - r_{1124}r_2r_{34} - r_{1223}r_2r_{34} + r_{123}r_{22}r_{34} + r_{112}r_{24}r_{34} + r_{1124}r_{23}r_4 \\
& + r_{1223}r_{23}r_4 - r_{123}r_{232}r_4 - r_{112}r_{234}r_4)),
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{4,3} = & (-r_{13}r_{14}r_2r_{234} + r_{13}r_{14}r_{23}r_{24} - r_{124}r_{134}r_2r_3 + r_{1234}r_{14}r_2r_3 + r_{12}r_{134}r_{24}r_3 \\
& - r_{123}r_{14}r_{24}r_3 + r_{124}r_{13}r_2r_{34} - r_{12}r_{13}r_{24}r_{34} + r_{00}(r_{124}r_{134}r_{23} - r_{1234}r_{14}r_{23} \\
& - r_{12}r_{134}r_{234} + r_{123}r_{14}r_{234} + r_{12}r_{1234}r_{34} - r_{123}r_{124}r_{34}) - r_{124}r_{13}r_{23}r_4 \\
& + r_{12}r_{13}r_{234}r_4 - r_{12}r_{1234}r_3r_4 + r_{123}r_{124}r_3r_4 + r_1(r_{134}r_2r_{234} - r_{134}r_{23}r_{24} \\
& - r_{1234}r_2r_{34} + r_{123}r_{24}r_{34} + r_{1234}r_{23}r_4 - r_{123}r_{234}r_4))
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{5,1} = & (r_{14}r_2r_3r_{3344} - r_{144}r_{233}r_{34} - r_{14}r_2r_{33}r_{344} + r_{14}r_2r_{334}r_{344} + r_{00}(r_{144}r_{23}r_{334} \\
& - r_{14}r_{23}r_{3344} - r_{144}r_{233}r_{34} + r_{12}r_{3344}r_{34} + r_{14}r_{233}r_{344} - r_{12}r_{334}r_{344}) \\
& + r_{14}r_{23}r_{3344}r_4 - r_{12}r_{33}r_{3344}r_4 - r_{14}r_{233}r_{344}r_4 + r_{12}r_{33}r_{344}r_4 + r_{144}(-r_{23}r_{334} \\
& + r_{23}r_{33}r_{34} + r_{233}r_{34} - r_{23}r_{33}r_4) - r_{14}r_{233}r_{344} + r_{14}r_{23}r_{33}r_{44} \\
& - r_{14}r_{23}r_{334}r_{44} + r_{12}r_{33}r_{334}r_{44} + r_{14}r_{233}r_{34}r_{44} - r_{12}r_{33}r_{34}r_{44}),
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{5,2} = & (-r_{14}r_{1443}r_2r_3 - r_{14}r_2r_{2334}r_3 + r_{14}r_{233}r_{24}r_3 + r_{14}r_2r_{234}r_{33} - r_{14}r_{23}r_{24}r_{33} \\
& - r_{14}r_2r_{234}r_{334} + r_{14}r_{23}r_{24}r_{334} + r_{124}r_2r_3r_{334} - r_{12}r_{24}r_3r_{334} - r_{13}r_{144}r_2r_{34} \\
& + r_{14}r_{1443}r_2r_{34} + r_{14}r_2r_{2334}r_{34} - r_{14}r_{233}r_{24}r_{34} - r_{124}r_2r_{33}r_{34} + r_{12}r_{24}r_{33}r_{34} \\
& + r_{13}r_{14}r_2r_{344} + r_{00}(-r_{134}r_{144}r_{23} + r_{14}r_{1443}r_{23} + r_{14}r_{23}r_{2334} - r_{14}r_{233}r_{234} \\
& - r_{124}r_{23}r_{334} + r_{12}r_{234}r_{334} + r_{123}r_{144}r_{34} - r_{12}r_{1443}r_{34} + r_{124}r_{233}r_{34} \\
& - r_{12}r_{2334}r_{34} + r_{12}r_{134}r_{344} - r_{123}r_{14}r_{344}) + r_{13}r_{144}r_{23}r_4 - r_{14}r_{1443}r_{23}r_4 \\
& - r_{14}r_{23}r_{2334}r_4 + r_{14}r_{233}r_{234}r_4 - r_{123}r_{144}r_3r_4 + r_{12}r_{1443}r_3r_4 - r_{124}r_{233}r_3r_4 \\
& + r_{12}r_{2334}r_3r_4 + r_{124}r_{23}r_{33}r_4 - r_{12}r_{234}r_{33}r_4 + r_{14}r_{123}r_{344}r_4 - r_{12}r_{13}r_{344}r_4 \\
& - r_{13}r_{14}r_{23}r_{44} + r_{123}r_{14}r_3r_{44} - r_{14}r_{123}r_{34}r_{44} + r_{12}r_{13}r_{34}r_{44} + r_{134}(r_{144}r_2r_3 \\
& - r_{14}r_2r_{344} + r_{14}r_{23}r_{44} - r_{12}r_3r_{44})),
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{5,3} = & (-r_{13}r_{14}r_2r_{234} + r_{13}r_{14}r_{23}r_{24} - r_{124}r_{134}r_2r_3 + r_{1234}r_{14}r_2r_3 + r_{12}r_{134}r_{24}r_3 \\
& - r_{123}r_{14}r_{24}r_3 + r_{124}r_{13}r_2r_{34} - r_{12}r_{13}r_{24}r_{34} + r_{00}(r_{124}r_{134}r_{23} - r_{1234}r_{14}r_{23} \\
& - r_{12}r_{134}r_{234} + r_{123}r_{14}r_{234} + r_{12}r_{1234}r_{34} - r_{123}r_{124}r_{34}) - r_{124}r_{13}r_{23}r_4 \\
& + r_{12}r_{13}r_{234}r_4 - r_{12}r_{1234}r_3r_4 + r_{123}r_{124}r_3r_4 + r_1(r_{134}r_2r_{234} - r_{134}r_{23}r_{24} \\
& - r_{1234}r_2r_{34} + r_{123}r_{24}r_{34} + r_{1234}r_{23}r_4 - r_{123}r_{234}r_4))
\end{aligned}$$

$$\mathcal{B}_1 = r_{00}r_{12} - r_1r_2$$

$$\mathcal{B}_2 = r_{00}r_{34} - r_3r_4$$

$$\mathcal{B}_3 = r_{00}r_{14}r_{23} - r_{14}r_2r_3 - r_{00}r_{12}r_{34} + r_1r_2r_{34} - r_1r_{23}r_4 + r_{12}r_3r_4$$

$$\begin{aligned}
\mathcal{A}_{10,1} = & r_{00}r_{114}r_{122}r_{1234}r_{23} - r_{00}r_{114}r_{1223}r_{124}r_{23} - r_{00}r_{1124}r_{122}r_{134}r_{23} + r_{00}r_{1122}r_{124}r_{134}r_{23} \\
& + r_{00}r_{1124}r_{1223}r_{14}r_{23} - r_{00}r_{1122}r_{1234}r_{14}r_{23} - r_1r_{114}r_{1234}r_{22}r_{23} + r_{114}r_{124}r_{13}r_{22}r_{23} \\
& + r_1r_{1124}r_{134}r_{22}r_{23} - r_{11}r_{124}r_{134}r_{22}r_{23} + r_{11}r_{1234}r_{14}r_{22}r_{23} - r_{1124}r_{13}r_{14}r_{22}r_{23} \\
& - r_{00}r_{114}r_{12}r_{1234}r_{232} + r_{00}r_{114}r_{123}r_{124}r_{232} + r_{00}r_{1124}r_{12}r_{134}r_{232} - r_{00}r_{112}r_{124}r_{134}r_{232} \\
& - r_{00}r_{1124}r_{123}r_{14}r_{232} + r_{00}r_{112}r_{1234}r_{14}r_{232} + r_1r_{114}r_{1234}r_{2}r_{232} - r_{114}r_{124}r_{13}r_{2}r_{232} \\
& - r_1r_{1124}r_{134}r_{2}r_{232} + r_{11}r_{124}r_{134}r_{2}r_{232} - r_{11}r_{1234}r_{14}r_{2}r_{232} + r_{1124}r_{13}r_{14}r_{2}r_{232} \\
& + r_{00}r_{114}r_{12}r_{1223}r_{234} - r_{00}r_{114}r_{122}r_{123}r_{234} - r_{00}r_{1122}r_{12}r_{134}r_{234} + r_{00}r_{112}r_{122}r_{134}r_{234} \\
& - r_{00}r_{112}r_{1223}r_{14}r_{234} + r_{00}r_{1122}r_{123}r_{14}r_{234} - r_1r_{114}r_{1223}r_{2}r_{234} + r_{114}r_{122}r_{13}r_{2}r_{234} \\
& + r_1r_{1122}r_{134}r_{2}r_{234} - r_{11}r_{122}r_{134}r_{2}r_{234} + r_{11}r_{1223}r_{14}r_{2}r_{234} - r_{1122}r_{13}r_{14}r_{2}r_{234} \\
& + r_1r_{114}r_{123}r_{22}r_{234} - r_{114}r_{12}r_{13}r_{22}r_{234} - r_1r_{112}r_{134}r_{22}r_{234} + r_{11}r_{12}r_{134}r_{22}r_{234} \\
& - r_{11}r_{123}r_{14}r_{22}r_{234} + r_{112}r_{13}r_{14}r_{22}r_{234} + r_1r_{114}r_{1223}r_{23}r_{24} - r_{114}r_{122}r_{13}r_{23}r_{24} \\
& - r_1r_{1122}r_{134}r_{23}r_{24} + r_{11}r_{122}r_{134}r_{23}r_{24} - r_{11}r_{1223}r_{14}r_{23}r_{24} + r_{1122}r_{13}r_{14}r_{23}r_{24} \\
& - r_1r_{114}r_{123}r_{232}r_{24} + r_{114}r_{12}r_{13}r_{232}r_{24} + r_1r_{112}r_{134}r_{232}r_{24} - r_{11}r_{12}r_{134}r_{232}r_{24} \\
& + r_{11}r_{123}r_{14}r_{232}r_{24} - r_{112}r_{13}r_{14}r_{232}r_{24} - r_{114}r_{122}r_{1234}r_{2}r_{3} + r_{114}r_{1223}r_{124}r_{2}r_{3} \\
& + r_{1124}r_{122}r_{134}r_{2}r_{3} - r_{1122}r_{124}r_{134}r_{2}r_{3} - r_{1124}r_{1223}r_{14}r_{2}r_{3} + r_{1122}r_{1234}r_{14}r_{2}r_{3} \\
& + r_{114}r_{12}r_{1234}r_{22}r_{3} - r_{114}r_{123}r_{124}r_{22}r_{3} - r_{1124}r_{12}r_{134}r_{22}r_{3} + r_{112}r_{124}r_{134}r_{22}r_{3}
\end{aligned}$$

$$\begin{aligned}
& + r_{1124}r_{123}r_{14}r_{22}r_3 - r_{112}r_{1234}r_{14}r_{22}r_3 - r_{114}r_{12}r_{1223}r_{24}r_3 + r_{114}r_{122}r_{123}r_{24}r_3 \\
& + r_{1122}r_{12}r_{134}r_{24}r_3 - r_{112}r_{122}r_{134}r_{24}r_3 + r_{112}r_{1223}r_{14}r_{24}r_3 - r_{1122}r_{123}r_{14}r_{24}r_3 \\
& - r_{00}r_{1124}r_{12}r_{1223}r_{34} + r_{00}r_{1124}r_{122}r_{123}r_{34} + r_{00}r_{1122}r_{12}r_{1234}r_{34} - r_{00}r_{112}r_{122}r_{1234}r_{34} \\
& + r_{00}r_{112}r_{1223}r_{124}r_{34} - r_{00}r_{1122}r_{123}r_{124}r_{34} + r_{1}r_{1124}r_{1223}r_{2}r_{34} - r_{1}r_{1122}r_{1234}r_{2}r_{34} \\
& + r_{11}r_{122}r_{1234}r_{2}r_{34} - r_{11}r_{1223}r_{124}r_{2}r_{34} - r_{1124}r_{122}r_{13}r_{2}r_{34} + r_{1122}r_{124}r_{13}r_{2}r_{34} \\
& - r_{1}r_{1124}r_{123}r_{22}r_{34} + r_{1}r_{112}r_{1234}r_{22}r_{34} - r_{11}r_{12}r_{1234}r_{22}r_{34} + r_{11}r_{123}r_{124}r_{22}r_{34} \\
& + r_{1124}r_{12}r_{13}r_{22}r_{34} - r_{112}r_{124}r_{13}r_{22}r_{34} - r_{1}r_{112}r_{1223}r_{24}r_{34} + r_{11}r_{12}r_{1223}r_{24}r_{34} \\
& + r_{1}r_{1122}r_{123}r_{24}r_{34} - r_{11}r_{122}r_{123}r_{24}r_{34} - r_{1122}r_{12}r_{13}r_{24}r_{34} + r_{112}r_{122}r_{13}r_{24}r_{34} \\
& - r_{1}r_{1124}r_{1223}r_{23}r_4 + r_{1}r_{1122}r_{1234}r_{23}r_4 - r_{11}r_{122}r_{1234}r_{23}r_4 + r_{11}r_{1223}r_{124}r_{23}r_4 \\
& + r_{1124}r_{122}r_{13}r_{23}r_4 - r_{1122}r_{124}r_{13}r_{23}r_4 + r_{1}r_{1124}r_{123}r_{232}r_4 - r_{1}r_{112}r_{1234}r_{232}r_4 \\
& + r_{11}r_{12}r_{1234}r_{232}r_4 - r_{11}r_{123}r_{124}r_{232}r_4 - r_{1124}r_{12}r_{13}r_{232}r_4 + r_{112}r_{124}r_{13}r_{232}r_4 \\
& + r_{1}r_{112}r_{1223}r_{234}r_4 - r_{11}r_{12}r_{1223}r_{234}r_4 - r_{1}r_{1122}r_{123}r_{234}r_4 + r_{11}r_{122}r_{123}r_{234}r_4 \\
& + r_{1122}r_{12}r_{13}r_{234}r_4 - r_{112}r_{122}r_{13}r_{234}r_4 + r_{1124}r_{12}r_{1223}r_{3}r_4 - r_{1124}r_{122}r_{123}r_{3}r_4 \\
& - r_{1122}r_{12}r_{1234}r_{3}r_4 + r_{112}r_{122}r_{1234}r_{3}r_4 - r_{112}r_{1223}r_{124}r_{3}r_4 + r_{1122}r_{123}r_{124}r_{3}r_4
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{10,2} = & r_{00}r_{1144}r_{124}r_{134}r_{23} - r_{00}r_{1144}r_{1234}r_{14}r_{23} + r_{00}r_{114}r_{1234}r_{144}r_{23} - r_{00}r_{1124}r_{134}r_{144}r_{23} \\
& - r_{00}r_{114}r_{124}r_{1443}r_{23} + r_{00}r_{1124}r_{14}r_{1443}r_{23} + r_{00}r_{124}r_{134}r_{2233}r_{23} - r_{00}r_{1234}r_{14}r_{2233}r_{23} \\
& - r_{00}r_{124}r_{134}r_{232}r_{233} + r_{00}r_{1234}r_{14}r_{232}r_{233} - r_{00}r_{122}r_{134}r_{23}r_{2334} + r_{00}r_{1223}r_{14}r_{23}r_{2334} \\
& + r_{1}r_{134}r_{22}r_{23}r_{2334} - r_{13}r_{14}r_{22}r_{23}r_{2334} + r_{00}r_{12}r_{134}r_{232}r_{2334} - r_{00}r_{123}r_{14}r_{232}r_{2334} \\
& - r_{1}r_{134}r_{2}r_{232}r_{2334} + r_{13}r_{14}r_{2}r_{232}r_{2334} - r_{00}r_{1144}r_{12}r_{134}r_{234} + r_{00}r_{1144}r_{123}r_{14}r_{234} \\
& - r_{00}r_{114}r_{123}r_{144}r_{234} + r_{00}r_{112}r_{134}r_{144}r_{234} + r_{00}r_{114}r_{12}r_{1443}r_{234} - r_{00}r_{112}r_{14}r_{1443}r_{234} \\
& + r_{1}r_{1144}r_{134}r_{2}r_{234} - r_{1144}r_{13}r_{14}r_{2}r_{234} + r_{114}r_{13}r_{144}r_{2}r_{234} - r_{11}r_{134}r_{144}r_{2}r_{234} \\
& - r_{1}r_{114}r_{1443}r_{2}r_{234} + r_{11}r_{14}r_{1443}r_{2}r_{234} - r_{00}r_{12}r_{134}r_{2233}r_{234} + r_{00}r_{123}r_{14}r_{2233}r_{234} \\
& + r_{1}r_{134}r_{2}r_{2233}r_{234} - r_{13}r_{14}r_{2}r_{2233}r_{234} + r_{00}r_{122}r_{134}r_{233}r_{234} - r_{00}r_{1223}r_{14}r_{233}r_{234} \\
& - r_{1}r_{134}r_{22}r_{233}r_{234} + r_{13}r_{14}r_{22}r_{233}r_{234} - r_{1}r_{1144}r_{134}r_{23}r_{24} + r_{1144}r_{13}r_{14}r_{23}r_{24} \\
& - r_{114}r_{13}r_{144}r_{23}r_{24} + r_{11}r_{134}r_{144}r_{23}r_{24} + r_{1}r_{114}r_{1443}r_{23}r_{24} - r_{11}r_{14}r_{1443}r_{23}r_{24} \\
& - r_{1}r_{134}r_{2233}r_{23}r_{24} + r_{13}r_{14}r_{2233}r_{23}r_{24} + r_{1}r_{134}r_{232}r_{233}r_{24} - r_{13}r_{14}r_{232}r_{233}r_{24} \\
& - r_{1144}r_{124}r_{134}r_{2}r_{3} + r_{1144}r_{1234}r_{14}r_{2}r_{3} - r_{114}r_{1234}r_{144}r_{2}r_{3} + r_{1124}r_{134}r_{144}r_{2}r_{3} \\
& + r_{114}r_{124}r_{1443}r_{2}r_{3} - r_{1124}r_{14}r_{1443}r_{2}r_{3} - r_{124}r_{134}r_{2}r_{2233}r_{3} + r_{1234}r_{14}r_{2}r_{2233}r_{3} \\
& + r_{124}r_{134}r_{22}r_{233}r_{3} - r_{1234}r_{14}r_{22}r_{233}r_{3} + r_{122}r_{134}r_{2}r_{2334}r_{3} - r_{1223}r_{14}r_{2}r_{2334}r_{3}
\end{aligned}$$

$$\begin{aligned}
& - r_{12}r_{134}r_{22}r_{2334}r_3 + r_{123}r_{14}r_{22}r_{2334}r_3 + r_{1144}r_{12}r_{134}r_{24}r_3 - r_{1144}r_{123}r_{14}r_{24}r_3 \\
& + r_{114}r_{123}r_{144}r_{24}r_3 - r_{112}r_{134}r_{144}r_{24}r_3 - r_{114}r_{12}r_{1443}r_{24}r_3 + r_{112}r_{14}r_{1443}r_{24}r_3 \\
& + r_{12}r_{134}r_{2233}r_{24}r_3 - r_{123}r_{14}r_{2233}r_{24}r_3 - r_{122}r_{134}r_{233}r_{24}r_3 + r_{1223}r_{14}r_{233}r_{24}r_3 \\
& - r_{124}r_{134}r_{22}r_{23}r_{33} + r_{1234}r_{14}r_{22}r_{23}r_{33} + r_{124}r_{134}r_{2}r_{232}r_{33} - r_{1234}r_{14}r_{2}r_{232}r_{33} \\
& - r_{122}r_{134}r_{2}r_{234}r_{33} + r_{1223}r_{14}r_{2}r_{234}r_{33} + r_{12}r_{134}r_{22}r_{234}r_{33} - r_{123}r_{14}r_{22}r_{234}r_{33} \\
& + r_{122}r_{134}r_{23}r_{24}r_{33} - r_{1223}r_{14}r_{23}r_{24}r_{33} - r_{12}r_{134}r_{232}r_{24}r_{33} + r_{123}r_{14}r_{232}r_{24}r_{33} \\
& + r_{00}r_{122}r_{1234}r_{23}r_{334} - r_{00}r_{1223}r_{124}r_{23}r_{334} - r_{1}r_{1234}r_{22}r_{23}r_{334} + r_{124}r_{13}r_{22}r_{23}r_{334} \\
& - r_{00}r_{12}r_{1234}r_{232}r_{334} + r_{00}r_{123}r_{124}r_{232}r_{334} + r_{1}r_{1234}r_{2}r_{232}r_{334} - r_{124}r_{13}r_{2}r_{232}r_{334} \\
& + r_{00}r_{12}r_{1223}r_{234}r_{334} - r_{00}r_{122}r_{123}r_{234}r_{334} - r_{1}r_{1223}r_{2}r_{234}r_{334} + r_{122}r_{13}r_{2}r_{234}r_{334} \\
& + r_{1}r_{123}r_{22}r_{234}r_{334} - r_{12}r_{13}r_{22}r_{234}r_{334} + r_{1}r_{1223}r_{23}r_{24}r_{334} - r_{122}r_{13}r_{23}r_{24}r_{334} \\
& - r_{1}r_{123}r_{232}r_{24}r_{334} + r_{12}r_{13}r_{232}r_{24}r_{334} - r_{122}r_{1234}r_{2}r_{3}r_{334} + r_{1223}r_{124}r_{2}r_{3}r_{334} \\
& + r_{12}r_{1234}r_{22}r_{3}r_{334} - r_{123}r_{124}r_{22}r_{3}r_{334} - r_{12}r_{1223}r_{24}r_{3}r_{334} + r_{122}r_{123}r_{24}r_{3}r_{334} \\
& + r_{00}r_{1144}r_{12}r_{1234}r_{34} - r_{00}r_{1144}r_{123}r_{124}r_{34} + r_{00}r_{1124}r_{123}r_{144}r_{34} - r_{00}r_{112}r_{1234}r_{144}r_{34} \\
& - r_{00}r_{1124}r_{12}r_{1443}r_{34} + r_{00}r_{112}r_{124}r_{1443}r_{34} - r_{1}r_{1144}r_{1234}r_{2}r_{34} + r_{1144}r_{124}r_{13}r_{2}r_{34} \\
& + r_{11}r_{1234}r_{144}r_{2}r_{34} - r_{1124}r_{13}r_{144}r_{2}r_{34} + r_{1}r_{1124}r_{1443}r_{2}r_{34} - r_{11}r_{124}r_{1443}r_{2}r_{34} \\
& + r_{00}r_{12}r_{1234}r_{2233}r_{34} - r_{00}r_{123}r_{124}r_{2233}r_{34} - r_{1}r_{1234}r_{2}r_{2233}r_{34} + r_{124}r_{13}r_{2}r_{2233}r_{34} \\
& - r_{00}r_{122}r_{1234}r_{233}r_{34} + r_{00}r_{1223}r_{124}r_{233}r_{34} + r_{1}r_{1234}r_{22}r_{233}r_{34} - r_{124}r_{13}r_{22}r_{233}r_{34}
\end{aligned}$$

$$\begin{aligned}
& - r_{00}r_{12}r_{1223}r_{2334}r_{34} + r_{00}r_{122}r_{123}r_{2334}r_{34} + r_1r_{1223}r_2r_{2334}r_{34} - r_{122}r_{13}r_2r_{2334}r_{34} \\
& - r_1r_{123}r_{22}r_{2334}r_{34} + r_{12}r_{13}r_{22}r_{2334}r_{34} + r_1r_{1144}r_{123}r_{24}r_{34} - r_{1144}r_{12}r_{13}r_{24}r_{34} \\
& - r_{11}r_{123}r_{144}r_{24}r_{34} + r_{112}r_{13}r_{144}r_{24}r_{34} - r_1r_{112}r_{1443}r_{24}r_{34} + r_{11}r_{12}r_{1443}r_{24}r_{34} \\
& + r_1r_{123}r_{2233}r_{24}r_{34} - r_{12}r_{13}r_{2233}r_{24}r_{34} - r_1r_{1223}r_{233}r_{24}r_{34} + r_{122}r_{13}r_{233}r_{24}r_{34} \\
& + r_{122}r_{1234}r_2r_{33}r_{34} - r_{1223}r_{124}r_2r_{33}r_{34} - r_{12}r_{1234}r_{22}r_{33}r_{34} + r_{123}r_{124}r_{22}r_{33}r_{34} \\
& + r_{12}r_{1223}r_{24}r_{33}r_{34} - r_{122}r_{123}r_{24}r_{33}r_{34} - r_{00}r_{114}r_{12}r_{1234}r_{344} + r_{00}r_{114}r_{123}r_{124}r_{344} \\
& + r_{00}r_{1124}r_{12}r_{134}r_{344} - r_{00}r_{112}r_{124}r_{134}r_{344} - r_{00}r_{1124}r_{123}r_{14}r_{344} + r_{00}r_{112}r_{1234}r_{14}r_{344} \\
& + r_1r_{114}r_{1234}r_2r_{344} - r_{114}r_{124}r_{13}r_2r_{344} - r_1r_{1124}r_{134}r_2r_{344} + r_{11}r_{124}r_{134}r_2r_{344} \\
& - r_{11}r_{1234}r_{14}r_2r_{344} + r_{1124}r_{13}r_{14}r_2r_{344} - r_1r_{114}r_{123}r_{24}r_{344} + r_{114}r_{12}r_{13}r_{24}r_{344} \\
& + r_1r_{112}r_{134}r_{24}r_{344} - r_{11}r_{12}r_{134}r_{24}r_{344} + r_{11}r_{123}r_{14}r_{24}r_{344} - r_{112}r_{13}r_{14}r_{24}r_{344} \\
& + r_1r_{1144}r_{1234}r_{23}r_4 - r_{1144}r_{124}r_{13}r_{23}r_4 - r_{11}r_{1234}r_{144}r_{23}r_4 + r_{1124}r_{13}r_{144}r_{23}r_4 \\
& - r_1r_{1124}r_{1443}r_{23}r_4 + r_{11}r_{124}r_{1443}r_{23}r_4 + r_1r_{1234}r_{2233}r_{23}r_4 - r_{124}r_{13}r_{2233}r_{23}r_4 \\
& - r_1r_{1234}r_{232}r_{233}r_4 + r_{124}r_{13}r_{232}r_{233}r_4 - r_1r_{1223}r_{23}r_{2334}r_4 + r_{122}r_{13}r_{23}r_{2334}r_4 \\
& + r_1r_{123}r_{232}r_{2334}r_4 - r_{12}r_{13}r_{232}r_{2334}r_4 - r_1r_{1144}r_{123}r_{234}r_4 + r_{1144}r_{12}r_{13}r_{234}r_4 \\
& + r_{11}r_{123}r_{144}r_{234}r_4 - r_{112}r_{13}r_{144}r_{234}r_4 + r_1r_{112}r_{1443}r_{234}r_4 - r_{11}r_{12}r_{1443}r_{234}r_4 \\
& - r_1r_{123}r_{2233}r_{234}r_4 + r_{12}r_{13}r_{2233}r_{234}r_4 + r_1r_{1223}r_{233}r_{234}r_4 - r_{122}r_{13}r_{233}r_{234}r_4
\end{aligned}$$

$$\begin{aligned}
& - r_{1144}r_{12}r_{1234}r_3r_4 + r_{1144}r_{123}r_{124}r_3r_4 - r_{1124}r_{123}r_{144}r_3r_4 + r_{112}r_{1234}r_{144}r_3r_4 \\
& + r_{1124}r_{12}r_{1443}r_3r_4 - r_{112}r_{124}r_{1443}r_3r_4 - r_{12}r_{1234}r_{2233}r_3r_4 + r_{123}r_{124}r_{2233}r_3r_4 \\
& + r_{122}r_{1234}r_{233}r_3r_4 - r_{1223}r_{124}r_{233}r_3r_4 + r_{12}r_{1223}r_{2334}r_3r_4 - r_{122}r_{123}r_{2334}r_3r_4 \\
& - r_{122}r_{1234}r_{23}r_{33}r_4 + r_{1223}r_{124}r_{23}r_{33}r_4 + r_{12}r_{1234}r_{232}r_{33}r_4 - r_{123}r_{124}r_{232}r_{33}r_4 \\
& - r_{12}r_{1223}r_{234}r_{33}r_4 + r_{122}r_{123}r_{234}r_{33}r_4 + r_{1}r_{1124}r_{123}r_{344}r_4 - r_{1}r_{112}r_{1234}r_{344}r_4 \\
& + r_{11}r_{12}r_{1234}r_{344}r_4 - r_{11}r_{123}r_{124}r_{344}r_4 - r_{1124}r_{12}r_{13}r_{344}r_4 + r_{112}r_{124}r_{13}r_{344}r_4 \\
& - r_{1}r_{114}r_{1234}r_{23}r_{44} + r_{114}r_{124}r_{13}r_{23}r_{44} + r_{1}r_{1124}r_{134}r_{23}r_{44} - r_{11}r_{124}r_{134}r_{23}r_{44} \\
& + r_{11}r_{1234}r_{14}r_{23}r_{44} - r_{1124}r_{13}r_{14}r_{23}r_{44} + r_{1}r_{114}r_{123}r_{234}r_{44} - r_{114}r_{12}r_{13}r_{234}r_{44} \\
& - r_{1}r_{112}r_{134}r_{234}r_{44} + r_{11}r_{12}r_{134}r_{234}r_{44} - r_{11}r_{123}r_{14}r_{234}r_{44} + r_{112}r_{13}r_{14}r_{234}r_{44} \\
& + r_{114}r_{12}r_{1234}r_3r_{44} - r_{114}r_{123}r_{124}r_3r_{44} - r_{1124}r_{12}r_{134}r_3r_{44} + r_{112}r_{124}r_{134}r_3r_{44} \\
& + r_{1124}r_{123}r_{14}r_3r_{44} - r_{112}r_{1234}r_{14}r_3r_{44} - r_{1}r_{1124}r_{123}r_{34}r_{44} + r_{1}r_{112}r_{1234}r_{34}r_{44} \\
& - r_{11}r_{12}r_{1234}r_{34}r_{44} + r_{11}r_{123}r_{124}r_{34}r_{44} + r_{1124}r_{12}r_{13}r_{34}r_{44} - r_{112}r_{124}r_{13}r_{34}r_{44}
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{10,3} = & -r_{00}r_{134}r_{144}r_{23}r_{2334} + r_{00}r_{14}r_{1443}r_{23}r_{2334} + r_{00}r_{134}r_{144}r_{233}r_{234} - r_{00}r_{14}r_{1443}r_{233}r_{234} \\
& + r_{134}r_{144}r_{2}r_{2334}r_3 - r_{14}r_{1443}r_2r_{2334}r_3 - r_{134}r_{144}r_{233}r_{24}r_3 + r_{14}r_{1443}r_{233}r_{24}r_3 \\
& - r_{134}r_{144}r_2r_{234}r_{33} + r_{14}r_{1443}r_2r_{234}r_{33} + r_{134}r_{144}r_{23}r_{24}r_{33} - r_{14}r_{1443}r_{23}r_{24}r_{33} \\
& + r_{00}r_{1234}r_{144}r_{23}r_{334} - r_{00}r_{124}r_{1443}r_{23}r_{334} - r_{00}r_{123}r_{144}r_{234}r_{334} + r_{00}r_{12}r_{1443}r_{234}r_{334} \\
& + r_{13}r_{144}r_2r_{234}r_{334} - r_{1}r_{1443}r_2r_{234}r_{334} - r_{13}r_{144}r_{23}r_{24}r_{334} + r_{1}r_{1443}r_{23}r_{24}r_{334} \\
& - r_{1234}r_{144}r_2r_3r_{334} + r_{124}r_{1443}r_2r_3r_{334} + r_{123}r_{144}r_{24}r_3r_{334} - r_{12}r_{1443}r_{24}r_3r_{334} \\
& + r_{00}r_{124}r_{134}r_{23}r_{3344} - r_{00}r_{1234}r_{14}r_{23}r_{3344} - r_{00}r_{12}r_{134}r_{234}r_{3344} + r_{00}r_{123}r_{14}r_{234}r_{3344} \\
& + r_{1}r_{134}r_2r_{234}r_{3344} - r_{13}r_{14}r_2r_{234}r_{3344} - r_{1}r_{134}r_{23}r_{24}r_{3344} + r_{13}r_{14}r_{23}r_{24}r_{3344} \\
& - r_{124}r_{134}r_2r_3r_{3344} + r_{1234}r_{14}r_2r_3r_{3344} + r_{12}r_{134}r_{24}r_3r_{3344} - r_{123}r_{14}r_{24}r_3r_{3344} \\
& - r_{00}r_{1234}r_{144}r_{233}r_{34} + r_{00}r_{124}r_{1443}r_{233}r_{34} + r_{00}r_{123}r_{144}r_{2334}r_{34} - r_{00}r_{12}r_{1443}r_{2334}r_{34} \\
& - r_{13}r_{144}r_2r_{2334}r_{34} + r_{1}r_{1443}r_2r_{2334}r_{34} + r_{13}r_{144}r_{233}r_{24}r_{34} - r_{1}r_{1443}r_{233}r_{24}r_{34} \\
& + r_{1234}r_{144}r_2r_{33}r_{34} - r_{124}r_{1443}r_2r_{33}r_{34} - r_{123}r_{144}r_{24}r_{33}r_{34} + r_{12}r_{1443}r_{24}r_{33}r_{34} \\
& + r_{00}r_{12}r_{1234}r_{3344}r_{34} - r_{00}r_{123}r_{124}r_{3344}r_{34} - r_{1}r_{1234}r_2r_{3344}r_{34} + r_{124}r_{13}r_2r_{3344}r_{34} \\
& + r_{1}r_{123}r_{24}r_{3344}r_{34} - r_{12}r_{13}r_{24}r_{3344}r_{34} - r_{00}r_{124}r_{134}r_{233}r_{344} + r_{00}r_{1234}r_{14}r_{233}r_{344} \\
& + r_{00}r_{12}r_{134}r_{2334}r_{344} - r_{00}r_{123}r_{14}r_{2334}r_{344} - r_{1}r_{134}r_2r_{2334}r_{344} + r_{13}r_{14}r_2r_{2334}r_{344}
\end{aligned}$$

$$\begin{aligned}
& + r_1 r_{134} r_{233} r_{24} r_{344} - r_{13} r_{14} r_{233} r_{24} r_{344} + r_{124} r_{134} r_2 r_{33} r_{344} - r_{1234} r_{14} r_2 r_{33} r_{344} \\
& - r_{12} r_{134} r_{24} r_{33} r_{344} + r_{123} r_{14} r_{24} r_{33} r_{344} - r_{00} r_{12} r_{1234} r_{334} r_{344} + r_{00} r_{123} r_{124} r_{334} r_{344} \\
& + r_1 r_{1234} r_2 r_{334} r_{344} - r_{124} r_{13} r_2 r_{334} r_{344} - r_1 r_{123} r_{24} r_{334} r_{344} + r_{12} r_{13} r_{24} r_{334} r_{344} \\
& + r_{13} r_{144} r_{23} r_{2334} r_4 - r_1 r_{1443} r_{23} r_{2334} r_4 - r_{13} r_{144} r_{233} r_{234} r_4 + r_1 r_{1443} r_{233} r_{234} r_4 \\
& + r_{1234} r_{144} r_{233} r_3 r_4 - r_{124} r_{1443} r_{233} r_3 r_4 - r_{123} r_{144} r_{2334} r_3 r_4 + r_{12} r_{1443} r_{2334} r_3 r_4 \\
& - r_{1234} r_{144} r_{23} r_{33} r_4 + r_{124} r_{1443} r_{23} r_{33} r_4 + r_{123} r_{144} r_{234} r_{33} r_4 - r_{12} r_{1443} r_{234} r_{33} r_4 \\
& + r_1 r_{1234} r_{23} r_{3344} r_4 - r_{124} r_{13} r_{23} r_{3344} r_4 - r_1 r_{123} r_{234} r_{3344} r_4 + r_{12} r_{13} r_{234} r_{3344} r_4 \\
& - r_{12} r_{1234} r_3 r_{3344} r_4 + r_{123} r_{124} r_3 r_{3344} r_4 - r_1 r_{1234} r_{233} r_{344} r_4 + r_{124} r_{13} r_{233} r_{344} r_4 \\
& + r_1 r_{123} r_{2334} r_{344} r_4 - r_{12} r_{13} r_{2334} r_{344} r_4 + r_{12} r_{1234} r_{33} r_{344} r_4 - r_{123} r_{124} r_{33} r_{344} r_4 \\
& + r_1 r_{134} r_{23} r_{2334} r_{44} - r_{13} r_{14} r_{23} r_{2334} r_{44} - r_1 r_{134} r_{233} r_{234} r_{44} + r_{13} r_{14} r_{233} r_{234} r_{44} \\
& + r_{124} r_{134} r_{233} r_3 r_{44} - r_{1234} r_{14} r_{233} r_3 r_{44} - r_{12} r_{134} r_{2334} r_3 r_{44} + r_{123} r_{14} r_{2334} r_3 r_{44} \\
& - r_{124} r_{134} r_{23} r_{33} r_{44} + r_{1234} r_{14} r_{23} r_{33} r_{44} + r_{12} r_{134} r_{234} r_{33} r_{44} - r_{123} r_{14} r_{234} r_{33} r_{44} \\
& - r_1 r_{1234} r_{23} r_{334} r_{44} + r_{124} r_{13} r_{23} r_{334} r_{44} + r_1 r_{123} r_{234} r_{334} r_{44} - r_{12} r_{13} r_{234} r_{334} r_{44} \\
& + r_{12} r_{1234} r_3 r_{334} r_{44} - r_{123} r_{124} r_3 r_{334} r_{44} + r_1 r_{1234} r_{233} r_{34} r_{44} - r_{124} r_{13} r_{233} r_{34} r_{44} \\
& - r_1 r_{123} r_{2334} r_{34} r_{44} + r_{12} r_{13} r_{2334} r_{34} r_{44} - r_{12} r_{1234} r_{33} r_{34} r_{44} + r_{123} r_{124} r_{33} r_{34} r_{44}
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}_{10,1} = & -r_{00}r_{114}r_{122}r_{23} + r_{00}r_{112}r_{14}r_{23} + r_1r_{114}r_{22}r_{23} - r_{11}r_{14}r_{22}r_{23} \\
& + r_{00}r_{114}r_{12}r_{232} - r_{00}r_{112}r_{14}r_{232} - r_1r_{114}r_{2}r_{232} + r_{11}r_{14}r_{2}r_{232} \\
& + r_{114}r_{122}r_{2}r_3 - r_{1122}r_{14}r_{2}r_3 - r_{114}r_{12}r_{22}r_3 + r_{112}r_{14}r_{22}r_3 \\
& - r_{00}r_{1122}r_{12}r_{34} + r_{00}r_{112}r_{122}r_{34} + r_1r_{1122}r_{2}r_{34} - r_{11}r_{122}r_{2}r_{34} \\
& - r_1r_{112}r_{22}r_{34} + r_{11}r_{12}r_{22}r_{34} - r_1r_{1122}r_{23}r_4 + r_{11}r_{122}r_{23}r_4 \\
& + r_1r_{112}r_{232}r_4 - r_{11}r_{12}r_{232}r_4 + r_{1122}r_{12}r_3r_4 - r_{112}r_{122}r_3r_4
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}_{10,2} = & -r_{00}r_{114}r_{133}r_{23} - 20r_{00}r_{124}r_{134}r_{23} + r_{00}r_{1144}r_{14}r_{23} + 20r_{00}r_{1234}r_{14}r_{23} \\
& - r_{00}r_{114}r_{144}r_{23} + r_{00}r_{14}r_{2233}r_{23} - r_{00}r_{14}r_{232}r_{233} + 20r_{00}r_{12}r_{134}r_{234} \\
& - 20r_{00}r_{123}r_{14}r_{234} - 20r_1r_{134}r_2r_{234} + 20r_{13}r_{14}r_2r_{234} + 20r_1r_{134}r_{23}r_{24} \\
& - 20r_{13}r_{14}r_{23}r_{24} + r_{114}r_{133}r_2r_3 + 20r_{124}r_{134}r_2r_3 - r_{1144}r_{14}r_2r_3 \\
& - 20r_{1234}r_{14}r_2r_3 + r_{114}r_{144}r_2r_3 - r_{14}r_2r_{2233}r_3 + r_{14}r_{22}r_{233}r_3 \\
& - 20r_{12}r_{134}r_{24}r_3 + 20r_{123}r_{14}r_{24}r_3 - r_{14}r_{22}r_{23}r_{33} + r_{14}r_2r_{232}r_{33} \\
& - r_{00}r_{122}r_{23}r_{334} + r_1r_{22}r_{23}r_{334} + r_{00}r_{12}r_{232}r_{334} - r_1r_2r_{232}r_{334} \\
& + r_{122}r_2r_3r_{334} - r_{12}r_{22}r_3r_{334} - r_{00}r_{1144}r_{12}r_{34} - 20r_{00}r_{12}r_{1234}r_{34} \\
& + 20r_{00}r_{123}r_{124}r_{34} + r_{00}r_{112}r_{133}r_{34} + r_{00}r_{112}r_{144}r_{34} + r_1r_{1144}r_2r_{34} \\
& + 20r_1r_{1234}r_2r_{34} - 20r_{124}r_{13}r_2r_{34} - r_{11}r_{133}r_2r_{34} - r_{11}r_{144}r_2r_{34} \\
& - r_{00}r_{12}r_{2233}r_{34} + r_1r_2r_{2233}r_{34} + r_{00}r_{122}r_{233}r_{34} - r_1r_{22}r_{233}r_{34} \\
& - 20r_1r_{123}r_{24}r_{34} + 20r_{12}r_{13}r_{24}r_{34} - r_{122}r_2r_{33}r_{34} + r_{12}r_{22}r_{33}r_{34} \\
& + r_{00}r_{114}r_{12}r_{344} - r_{00}r_{112}r_{14}r_{344} - r_1r_{114}r_2r_{344} + r_{11}r_{14}r_2r_{344} \\
& - r_1r_{1144}r_{23}r_4 - 20r_1r_{1234}r_{23}r_4 + 20r_{124}r_{13}r_{23}r_4 + r_{11}r_{133}r_{23}r_4 \\
& + r_{11}r_{144}r_{23}r_4 - r_1r_{2233}r_{23}r_4 + r_1r_{232}r_{233}r_4 + 20r_1r_{123}r_{234}r_4 \\
& - 20r_{12}r_{13}r_{234}r_4 + r_{1144}r_{12}r_3r_4 + 20r_{12}r_{1234}r_3r_4 - 20r_{123}r_{124}r_3r_4
\end{aligned}$$

$$\begin{aligned}
& -r_{112}r_{133}r_3r_4 - r_{112}r_{144}r_3r_4 + r_{12}r_{2233}r_3r_4 - r_{122}r_{233}r_3r_4 \\
& + r_{122}r_{23}r_{33}r_4 - r_{12}r_{232}r_{33}r_4 + r_1r_{112}r_{344}r_4 - r_{11}r_{12}r_{344}r_4 \\
& + r_1r_{114}r_{23}r_{44} - r_{11}r_{14}r_{23}r_{44} - r_{114}r_{12}r_3r_{44} + r_{112}r_{14}r_3r_{44} \\
& - r_1r_{112}r_{34}r_{44} + r_{11}r_{12}r_{34}r_{44}
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}_{10,3} = & -r_{00}r_{133}r_{23}r_{334} - r_{00}r_{144}r_{23}r_{334} + r_{133}r_2r_3r_{334} + r_{144}r_2r_3r_{334} \\
& + r_{00}r_{14}r_{23}r_{3344} - r_{14}r_2r_3r_{3344} + r_{00}r_{133}r_{233}r_{34} + r_{00}r_{144}r_{233}r_{34} \\
& - r_{133}r_2r_{33}r_{34} - r_{144}r_2r_{33}r_{34} - r_{00}r_{12}r_{3344}r_{34} + r_1r_2r_{3344}r_{34} \\
& - r_{00}r_{14}r_{233}r_{344} + r_{14}r_2r_{33}r_{344} + r_{00}r_{12}r_{334}r_{344} - r_1r_2r_{334}r_{344} \\
& - r_{133}r_{233}r_3r_4 - r_{144}r_{233}r_3r_4 + r_{133}r_2r_3r_{33}r_4 + r_{144}r_2r_3r_{33}r_4 \\
& - r_1r_{23}r_{3344}r_4 + r_{12}r_3r_{3344}r_4 + r_1r_{233}r_{344}r_4 - r_{12}r_{33}r_{344}r_4 \\
& + r_{14}r_{233}r_3r_{44} - r_{14}r_2r_{33}r_{44} + r_1r_{23}r_{334}r_{44} - r_{12}r_3r_{334}r_{44} \\
& - r_1r_{233}r_{34}r_{44} + r_{12}r_{33}r_{34}r_{44}
\end{aligned}$$

Lastly, we set

$$\mathcal{A}_{10} = \delta_1^8 \delta_2^4 \mathcal{A}_{10,1} + \delta_1^6 \delta_2^6 \mathcal{A}_{10,2} + \delta_1^4 \delta_2^8 \mathcal{A}_{10,3}$$

$$\mathcal{B}_{10} = \delta_1^6 \delta_2^2 \mathcal{B}_{10,1} + \delta_1^4 \delta_2^4 \mathcal{B}_{10,2} + \delta_1^2 \delta_2^6 \mathcal{B}_{10,3}$$

$$\mathcal{C}_{10} = r_{00}r_{14}r_{23} - r_{14}r_2r_3 - r_{00}r_{12}r_{34} + r_1r_2r_{34} - r_1r_{23}r_4 + r_{12}r_3r_4$$

Bibliography

Bibliography

- [1] Robert A. Adams, *Sobolev spaces*, Academic Press, San Diego – London, 1978.
- [2] Robert A. Adams and John J. F. Fournier, *Sobolev spaces*, second ed., Elsevier/Academic Press, Amsterdam, 2003.
- [3] Robert Adler and Jonathon Taylor, *Random fields and geometry*, Springer, New York, 2007.
- [4] Robert J. Adler, *The geometry of random fields*, John Wiley & Sons Ltd., Chichester, 1981.
- [5] Ludwig Arnold, *Random dynamical systems*, Springer-Verlag, Berlin – Heidelberg – New York, 1998.
- [6] Jean-Marc Azaïs and Mario Wschebor, *Level sets and extrema of random processes and fields*, Wiley, Hoboken, 2009.
- [7] R. Baker Kearfott Azmy S. Ackleh, Edward James Allen and Padmanabhan Seshaiyer, *Classical and modern numerical analysis: Theory, methods and practice*, Chapman and Hall/CRC Numerical Analysis and Scientific Computation Series, Chapman and Hall, New York, 2009.
- [8] F. Bai, C. M. Elliott, A. Gardiner, A. Spence, and A. M. Stuart, *The viscous Cahn-Hilliard equation. Part I: Computations*, **8** (1995), 131–160.
- [9] Fengshan Bai, Alastair Spence, and Andrew M. Stuart, *Numerical computations of coarsening in the one-dimensional Cahn-Hilliard model of phase separation*, **78** (1994), 155–165.
- [10] Frank S. Bates, *Block copolymer thermodynamics: Theory and experiment*, Annual Review of Physical Chemistry **41** (1990), 525–557.
- [11] Peter W. Bates and Paul C. Fife, *The dynamics of nucleation for the Cahn-Hilliard equation*, **53** (1993), no. 4, 990–1008.
- [12] Heinz Bauer, *Probability theory*, de Gruyter, Berlin, 1996.
- [13] Patrick Billingsley, *Probability and measure*, third ed., Wiley, New York, 1995.
- [14] G. Birkhoff and S. MacLane, *A survey of modern algebra*, 4th ed., Macmillan Publishing Co., New York, 1977.

- [15] John P. Boyd, *Chebyshev and Fourier Spectral Methods*, second ed., Dover Publications Inc., Mineola, NY, 2001.
- [16] Glen E. Bredon, *Topology and geometry*, Graduate Texts in Mathematics, vol. 139, Springer-Verlag, New York, 1997.
- [17] Robert F. Brown, *A topological introduction to nonlinear analysis*, Birkhäuser, Boston, 1993.
- [18] John W. Cahn, *Free energy of a nonuniform system. II. Thermodynamic basis*, **30** (1959), 1121–1124.
- [19] ———, *Phase separation by spinodal decomposition in isotropic systems*, **42** (1965), 93–99.
- [20] ———, *The later stages of spinodal decomposition and the beginnings of particle coarsening*, **14** (1966), 1685–1692.
- [21] ———, *Spinodal decomposition*, **242** (1968), 166–180.
- [22] John W. Cahn and J. E. Hilliard, *Free energy of a nonuniform system I. Interfacial free energy*, **28** (1958), 258–267.
- [23] ———, *Free energy of a nonuniform system III. Nucleation in a two-component incompressible fluid*, **31** (1959), 688–699.
- [24] Gunnar Carlsson and Vin de Silva, *Topological approximation by small simplicial complexes*, Preprint (2003).
- [25] CHomP, *Computational Homology Project*, Software available at <http://chomp.rutgers.edu> (2002–2011).
- [26] H. Cook, *Brownian motion in spinodal decomposition*, **18** (1970), 297–306.
- [27] Richard Courant and David Hilbert, *Methods of mathematical physics*, Intersciences, New York, 1953.
- [28] A. C. Davison and D. V. Hinkley, *Bootstrap methods and their application*, Cambridge University Press, Cambridge, 1997.
- [29] Sarah Day, William Kalies, Konstantin Mischaikow, and Thomas Wanner, *Probabilistic and numerical validation of homology computations for nodal domains*, Electronic Research Announcements of the American Mathematical Society **13** (2007), 60–73.
- [30] Sarah Day, William Kalies, and Thomas Wanner, *Verified homology computations for nodal domains*, SIAM Journal on Multiscale Modeling and Simulation **7** (2009), no. 4, 1695–1726.
- [31] Sarah Day, Jean-Philippe Lessard, and Konstantin Mischaikow, *Validated continuation for equilibria of PDEs*, SIAM Journal on Numerical Analysis **45** (2007), no. 4.

- [32] Paweł Dłotko, Tomasz Kaczynski, Marian Mrozek, and Thomas Wanner, *Coreduction homology algorithm for regular CW-complexes*, Discrete & Computational Geometry (to appear).
- [33] Herbert Edelsbrunner and John L. Harer, *Computational topology*, American Mathematical Society, Providence, 2010.
- [34] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian, *Topological persistence and simplification*, Discrete Computational Geometry **28** (2002), no. 1-3, 511–533.
- [35] David J. Eyre, *Coarsening dynamics for solutions of the Cahn-Hilliard equation in one dimension*, (1995), In press.
- [36] Kambiz Farahmand, *Topics in random polynomials*, Pitman Research Notes in Mathematics, vol. 393, Longman, Harlow, 1998.
- [37] Paul C. Fife, *Pattern dynamics for parabolic PDE's*, Preprint (1989).
- [38] Marcio Gameiro, William Kalies, and Konstantin Mischaikow, *Topological characterization of spatial-temporal chaos*, Physical Review E **70** (2004), no. 3, 1–4.
- [39] Marcio Gameiro, Konstantin Mischaikow, and Thomas Wanner, *Evolution of pattern complexity in the cahn-hilliard theory of phase separation*, Acta Materialia **53** (2005), no. 3, 693–704.
- [40] Marcio Gameiro and Pawel Pilarczyk, *Automatic homology computation with application to pattern classification*, RIMS Kkyroku Bessatsu B3 (2007), 1–10.
- [41] Igor Griva, Stephen Nash, and Ariela Sofer, *Linear and nonlinear optimization*, SIAM, Philadelphia, 2009.
- [42] E. Hairer and G. Wanner, *Solving ordinary differential equations II*, Springer-Verlag, Berlin, 1991.
- [43] Tina Hartley and Thomas Wanner, *A semi-implicit spectral method for stochastic non-local phase-field models*, Discrete and Continuous Dynamical Systems, Series A **25** (2009), no. 2, 399–429.
- [44] Dan Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Mathematics, vol. 840, Springer-Verlag, Berlin – Heidelberg – New York, 1981.
- [45] Luc Jaulin, Michel Kieffer, Olivier Didrit, and Éric Walter, *Applied interval analysis*, Springer-Verlag London Ltd., London, 2001.
- [46] Tomasz Kaczynski, Konstantin Mischaikow, and Marian Mrozek, *Computational homology*, Applied Mathematical Sciences, Springer, 2004.
- [47] Hansjörg Kielhöfer, *Bifurcation theory: An introduction with applications to pdes*, Applied Mathematical Sciences, Springer, New York, 2003.
- [48] Daniel A. Klain and Gian-Carlo Rota, *Introduction to geometric probability*, Cambridge University Press, Cambridge, 1997.

- [49] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Springer-Verlag, Berlin, 1992.
- [50] Michael B. Marcus and Gilles Pisier, *Random fourier series with applications to harmonic analysis*, Annals of Mathematics Studies, vol. 101, Princeton University Press, Princeton, 1981.
- [51] William S. Massey, *A basic course in algebraic topology*, Graduate Texts in Mathematics, vol. 127, Springer-Verlag, New York, 1991.
- [52] Konstantin Mischaikow and Marian Mrozek, *Conley index theory*, Handbook of Dynamical Systems III: Towards Applications (Bernold Fiedler, Gerard Iooss, and Nancy Kopell, eds.), Elsevier, to appear.
- [53] Konstantin Mischaikow and Thomas Wanner, *Probabilistic validation of homology computations for nodal domains*, Annals of Applied Probability **17** (2007), no. 3, 980–1018.
- [54] ———, *Topology-guided sampling of nonhomogenous random processes*, Annals of Applied Probability **20** (2010), no. 3, 1068–1097.
- [55] Ramon E. Moore, *Methods and applications of interval analysis*, SIAM Studies in Applied Mathematics, vol. 2, Society for Industrial and Applied Mathematics, Philadelphia, 1979.
- [56] Ramon E. Moore, R. Baker Kearfott, and Michael J. Cloud, *Introduction to interval analysis*, SIAM, Philadelphia, 2009.
- [57] J. E. Morral and John W. Cahn, *Spinodal decomposition in ternary systems*, **19** (1971), 1037–1045.
- [58] Marian Mrozek, *Cech type approach to computing homology of maps*, Discrete and Computational Geometry (2010), In print.
- [59] Marian Mrozek and Thomas Wanner, *Coreduction homology algorithm for inclusions and persistent homology*, Computers & Mathematics with Applications **60** (2010), no. 10, 2812–2833.
- [60] James R. Munkres, *Elements of algebraic topology*, Addison-Wesley, Menlo Park, 1984.
- [61] Arnold Neumaier, *Interval methods for systems of equations*, Encyclopedia of Mathematics and its Applications, vol. 37, Cambridge University Press, Cambridge, 1990.
- [62] Partha Niyogi, Stephen Smale, and Shmuel Weinberger, *Finding the homology of submanifolds with high confidence from random samples*, Discrete and Continuous Geometry **39** (2008), no. 4, 419–441.
- [63] Michael Renardy and Robert C. Rogers, *An introduction to partial differential equations*, Springer-Verlag, New York, 1993.
- [64] Evelyn Sander and Thomas Wanner, *Monte Carlo simulations for spinodal decomposition*, **95** (1999), no. 5–6, 925–948.

- [65] _____, *Unexpectedly linear behavior for the Cahn-Hilliard equation*, **60** (2000), no. 6, 2182–2202.
- [66] Nikolay Sidorov, Boris Loginov, Aleksandr Sinityn, and Michail Falaleev, *Lyapunov-Schmidt methods in nonlinear analysis and applications*, Mathematics and its Applications, vol. 550, Kluwer Academic Publishers, Dordrecht, 2002.
- [67] Stig Skelboe, *Computation of rational interval functions*, Nordisk Tidskrift for Informationsbehandling (BIT) **14** (1974), 87–95.
- [68] Lloyd N. Trefethen, *Spectral methods in MATLAB*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.
- [69] M. M. Vaĭnberg and V. A. Trenogin, *Theory of branching of solutions of non-Linear equations*, Noordhoff International Publishing, Leyden, 1974.
- [70] Thomas Wanner, Edwin R. Fuller Jr., and David M. Saylor, *Homological characterization of microstructure response fields in polycrystals*, Acta Materialia **58** (2010), no. 1, 102–110.
- [71] Eberhard Zeidler, *Nonlinear functional analysis and its applications. i: Fixed-point theorems*, Springer-Verlag, New York – Berlin – Heidelberg, 1986.

Curriculum Vitae

Gregory Cochran was born in Houston, Texas in July 1982. He went to Virginia Tech for undergraduate studies where he majored in Mathematics with a minor in Physics. He started his graduate studies in Mathematics in August 2006. He has accepted a position as a Data Analyst beginning August 2011.