

Simulation and Analysis of Random Fields

by Gordon A. Fenton

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Civil Engineering and Operations Research

January, 1990

©Copyright by Gordon Ashley Fenton 1990
All Rights Reserved

Abstract

Increased interest in realistic modeling of natural phenomena leads naturally to their description as random processes in space and time. Methods of generating realizations of multi-dimensional random processes, specifically the Fast Fourier Transform, Turning Bands, and Moving Average algorithms, are critically evaluated and means of improving accuracy and computational efficiency are suggested. As an alternative and complement to these methods, a technique called Local Average Subdivision (LAS) is introduced which produces realizations of locally averaged random processes in one, two, or three dimensions. The main advantage of the LAS method is that it can be easily conditioned on known data and that changes in resolution of the field are properly represented statistically.

The LAS method is employed in a simulation-based study of the statistics of excursions and extrema of two-dimensional Gauss-Markov processes. Empirical relationships for the average number of isolated excursions and their areas are presented and compared with existing theories. A measure related to the degree of clustering of the excursions is also proposed. Some common extreme value cumulative distribution functions are compared to the simulation-based distributions.

Best linear estimation techniques in the frequency domain are incorporated in a new approach to the simulation of optionally conditioned stationary or non-stationary space-time processes and applied to earthquake ground motion simulation. This method is used along with the Local Average Subdivision algorithm in a liquefaction risk case study where the soil is modeled as a three-dimensional stochastic medium and input ground motions come from a space-time random field. The liquefaction analysis is performed by a non-linear multi-phase finite element model for which the LAS realizations are ideally suited as they give random properties representing the average over each element.

Acknowledgements

The author would like to express his sincere thanks to his advisor, Dr. E. H. Vanmarcke for his kindness, guidance and support and to the other members of his committee, Dr. G. Watson, Dr. J. Prevost, Dr. E. Cinlar, and Dr. M. Shinozuka for their valuable suggestions and insightful comments. Financial support for the study came from the National Science Foundation under Grant Number ECE-8611521 and from the National Center for Earthquake Engineering Research under Grant NCEER 87-6003. Use of the supercomputer was made possible by the generous support of the John von Neumann Supercomputer Center under grants LAC 21023 and NAC 21046 and by help from the JVNCC staff in implementing the codes. Any opinions, findings, and conclusions or recommendations are those of the author and do not necessarily reflect the views of the above-mentioned organizations. Finally the author would like to thank his wife, Terry, and son, Paul, for their unending patience, care, and encouragement during the preparation of this dissertation.

Table of Contents

	Page
Abstract	iii
Acknowledgements	iv
List of Illustrations	vii
List of Tables	xvi
1. Introduction	1
1.1 General	1
1.2 Basic Properties of Random Fields	3
1.3 Spectral Representation of Homogeneous Processes	6
1.4 Local Average Processes	8
1.5 Parameter Estimation and Ergodicity	11
2. Fourier Transform Techniques	13
2.1 One-Dimensional Processes	13
2.1.1 <i>Spectral Formulation</i>	19
2.1.2 <i>Covariance Formulation</i>	21
2.2 Multi-Dimensional Extensions	23
2.3 Summary	29
3. Moving Average and Turning Bands Methods	41
3.1 Introduction	41
3.2 Moving Average Method	42
3.3 Turning Bands Method	46
4. Local Average Subdivision	54
4.1 Introduction	54
4.2 One-Dimensional Local Average Subdivision	55
4.2.1 <i>Accuracy</i>	61

	Page
4.2.2 <i>Boundary Conditions and Neighborhood Size</i>	64
4.3 Multi-Dimensional Extensions	68
5. Space-Time Processes	78
5.1 Introduction	78
5.2 Two Point Case	79
5.3 Multiple Point Case	81
5.4 Conditioning the Field	82
5.5 Case Studies	85
5.6 Non-Stationary Extensions	91
6. Statistics of Level Excursions and Extrema	100
6.1 Introduction	100
6.2 Local Average Processes	102
6.3 Analysis of Realizations	104
6.4 Total Area of Excursion Regions	107
6.5 Expected Number of Isolated Excursions	107
6.6 Expected Area of Isolated Excursions	113
6.7 Expected Number of Holes Appearing in Excursion Regions	116
6.8 Integral Geometric Characteristic of 2-D Random Fields	119
6.9 Clustering of Excursion Regions	123
6.10 Distribution of the Global Maxima	125
7. Soil Liquefaction Study	128
7.1 Introduction	128
7.2 Wildlife Liquefaction Site	129
7.3 Stochastic Soil Model	134
7.3.1 <i>Permeability</i>	135
7.3.2 <i>Porosity</i>	137
7.3.3 <i>Elastic Modulus and Poisson's Ratio</i>	139
7.3.4 <i>Dilation Angle</i>	141

	Page
7.4 Finite Element Model	143
7.5 Definition of Liquefaction	146
7.6 Monte-Carlo Analysis	149
8. Conclusions and Recommendations	154
References	157
Appendices	164
A Some Additional Results for Multi-Dimensional FFT Simulation	164
A.1 2-D Covariance Formulation	164
A.2 3-D Spectral Formulation	166
A.3 3-D Covariance Formulation	167
B Spectral Analysis of One-Dimensional LAS Processes	170
B.1 Spectral Analysis	170
B.2 Ornstein-Uhlenbeck Process	171
B.3 Damped Oscillatory Noise	172
B.4 Fractional Gaussian Noise	173

List of Illustrations

Figure		Page
1.1	Distances characterizing the relative location of the volumes D_α and D_β in three-dimensional space (courtesy of Vanmarcke [67]).	10
2.1	Comparison of estimated and exact covariance function for a process generated using the Fast Fourier Transform (averaged over 200 realizations).	17
2.2	Example of overly coarse frequency discretization resulting in an over-estimation of point variance ($T = 5$ and $b = 2$).	18
2.3	Fourier coefficient symmetry over the plane of their indices. Dashed arrows denote symmetric half-lines and double arrows denote symmetric planes.	25
2.4	Sample function of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = 4$. For such a large scale of fluctuation, the process varies relatively smoothly and one gets larger dark and light regions.	30
2.5	Sample function of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = 4$	31
2.6	Estimated statistics of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = 4$. Statistics are averaged over 100 fields.	32
2.7	Estimated statistics of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = 4$. Statistics are averaged over 100 fields.	32
2.8	Sample function of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Notice the increased ‘roughness’ corresponding to a shorter scale of fluctuation.	33
2.9	Sample function of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$	34
2.10	Estimated statistics of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Statistics are averaged over 100 fields.	35

Figure	Page	
2.11	Estimated statistics of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Statistics are averaged over 100 fields.	35
2.12	Estimated statistics of a 3-D field (spectral formulation) with covariance function given by (2.55) for $\theta = 4$. Statistics are averaged over 50 fields.	36
2.13	Estimated statistics of a 3-D field (covariance formulation) with covariance function given by (2.55) for $\theta = 4$. Statistics are averaged over 50 fields.	36
2.14	Estimated statistics of a 3-D field (spectral formulation) with covariance function given by (2.55) for $\theta = \frac{1}{2}$. Statistics are averaged over 50 fields.	37
2.15	Estimated statistics of a 3-D field (covariance formulation) with covariance function given by (2.55) for $\theta = \frac{1}{2}$. Statistics are averaged over 50 fields.	37
2.16	Sample function of an anisotropic 2-D field (spectral formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$	38
2.17	Sample function of an anisotropic 2-D field (covariance formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$	39
2.18	Estimated statistics of an anisotropic 2-D field (spectral formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$	40
2.19	Estimated statistics of an anisotropic 2-D field (covariance formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$	40
3.1	Schematic representation of the moving average process in one dimension.	44
3.2	Two-dimensional moving average process. Z_{ij} is formed by summing the contributions from the underlying white noise process in the shaded region.	45
3.3	The Turning Bands Method: contributions from the line process $Z_i(\xi_i)$ at $\mathbf{x}_k \cdot \mathbf{u}_i$ are summed into the field process $Z(\mathbf{x})$ at \mathbf{x}_k	47
3.4	Sample function of a 2-D field via TBM using 16 lines with covariance function given by (3.17) for $\theta = 4$. Notice the banded or streaked appearance.	50

Figure	Page	
3.5	Sample function of a 2-D field via TBM using 64 lines with covariance function given by (3.17) for $\theta = 4$. The streaked appearance has largely disappeared.	51
3.6	Comparison of estimated statistics of 2-D fields generated via TBM using scale $\theta = 4$ for randomly versus evenly oriented lines and ‘ergodic’ versus ‘non-ergodic’ generation of the line processes. Solid lines denote the exact covariance and dashed lines represent estimated horizontal, vertical, and diagonal covariances averaged over 10 realizations.	52
3.7	Comparison of estimated statistics of 2-D fields generated via TBM using scale $\theta = 1$ for randomly versus evenly oriented lines and ‘ergodic’ versus ‘non-ergodic’ generation of the line processes. Solid lines denote the exact covariance and dashed lines represent estimated horizontal, vertical, and diagonal covariances averaged over 10 realizations.	53
4.1	Top-down approach to the LAS construction of a random process	57
4.2	Sample function of an Ornstein-Uhlenbeck process with $B(\tau) = e^{-\frac{ \tau }{2}}$	57
4.3	1-D LAS indexing scheme for stage i (top) and stage $i + 1$ (bottom).	59
4.4	Comparison of algorithmic and exact covariance between adjacent cells across a parent cell boundary for varying effective averaging lengths $2D/\theta$	63
4.5	Comparison of exact and estimated covariance function for $\theta = 4$ averaged over 200 realizations.	63
4.6	Effect of neighborhood size for a) $n = 1$ and b) $n = 2$ on the modeling of damped oscillatory noise (4.23) with $\theta = 4$ and $\omega = 8$	65
4.7	a) Sample function of $1/f$ type noise ($H = 0.95$) and (b) corresponding estimated versus exact covariance structure (averaged over 200 realizations).	67
4.8	First three stages and indexing scheme of the 2-D LAS algorithm (stage 0 = \boxtimes , stage 1 = +, and stage 2 = \circ).	69
4.9	Sample function of the 2-D LAS generated process given by (4.33) with $\theta = 4$	72

Figure	Page
4.10	Sample function of the 2-D LAS generated process given by (4.33) with $\theta = \frac{1}{2}$ 73
4.11	Comparison of estimated and exact covariance structure of the LAS generated 2-D process (4.33), with $\theta = 4$, averaged over 10 realizations. . . . 74
4.12	Comparison of estimated and exact covariance structure of the LAS generated 2-D process (4.33), with $\theta = \frac{1}{2}$, averaged over 10 realizations. . . . 74
4.13	Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = 4$, averaged over 10 realizations. Estimates, shown dashed, are made in various directions through the volume. 75
4.14	Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = 4$, averaged over 50 realizations. Estimates, shown dashed, are made in various directions through the volume. 76
4.15	Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = \frac{1}{2}$, averaged over 10 realizations. Estimates, shown dashed, are made in various directions through the volume. 76
5.1	Spectral density function used in examples. 86
5.2	Four point unconditional time history simulation for various scales. Points are arranged on the vertices of a 100 <i>m</i> square. 88
5.3	Four point conditional time history simulation for various scales with three points known to be moving sinusoidally. 89
5.4	Three point conditional time history simulation at two different scales. The known motion of the two points is established in the uppermost plot. . . 90
5.5	North-South component of down-hole (7.5 <i>m</i>) earthquake acceleration motion recorded Nov. 24, 1987 at the Wildlife Liquefaction Array site, Imperial County, California (Superstition Hills event, magnitude 6.6) showing the four subwindow stationary approximation. 91
5.6	Estimates of the power spectral density function within each time window obtained using FFT and Maximum Entropy methods. 92
5.7	Four point unconditioned non-stationary simulation with points arranged on the vertices of a 100 <i>m</i> square for scale $s = 100$ 96

Figure	Page
5.8	Four point unconditioned non-stationary simulation with points arranged on the vertices of a $100 m$ square for scale $s = 10$ 97
5.9	Two point conditioned non-stationary simulation for scale $s = 100$ 98
5.10	Two point conditioned non-stationary simulation for scale $s = 10$ 99
6.1	Examples of ‘weakly’ surrounded holes - (a) and (b) are found to be holes while (c) and (d) are not. 105
6.2	Sample function of the binary field Y (6.19). Regions shown in gray represent regions of \mathcal{Z} which exceed the threshold $b = 1\sigma_D$. \mathcal{Z} is generated via the 2-D LAS algorithm according to (6.18) with $\theta = \frac{1}{2}$ 106
6.3	Average total area of excursion ratio, $\bar{A}_{b,D}/A_T$, as a function of threshold b 107
6.4	Average number of isolated excursions, $\bar{N}_{b,D}$, estimated from 400 realizations of the locally averaged two-dimensional Gauss-Markov process (6.18). 108
6.5	Comparison of theoretical fit by (6.23) with the observed average number of isolated excursions obtained by simulation. 111
6.6	Comparison of empirical fit by (6.24) with the observed average number of isolated excursions obtained by simulation. 112
6.7	Average area of isolated excursion regions estimated from 400 realizations of the locally averaged two-dimensional Gauss-Markov process. 113
6.8	Comparison of semi-empirical fit by (6.25) with the observed average area of isolated excursions obtained by simulation. 114
6.9	Comparison of fit by (6.26), using the empirically derived variances $\sigma_{\mathcal{Z}_D}^2$, with the observed average area of isolated excursions obtained by simulation. 115
6.10	Average number of holes appearing in excursion regions. 116
6.11	Comparison of empirical fit by (6.27) with observed average number of holes obtained by simulation. 118
6.12	Average values of Adler’s IG characteristic, $\bar{\Gamma}$, obtained from 400 realizations of the locally averaged Gauss-Markov process. 119

Figure	Page
6.13	Comparison of theoretically predicted IG characteristic (6.30) with observed average values obtained by simulation. 121
6.14	Comparison of empirically predicted IG characteristic (6.31) with observed average values obtained by simulation. 122
6.15	Average values of the cluster measure $\bar{\Psi}$ estimated from 200 realizations of the locally averaged Gauss-Markov process. 124
6.16	Observed cumulative distribution of the global maximum of each realization compared to the one-parameter extreme value distribution given by (6.37). 126
6.17	Observed cumulative distribution of the global maximum of each realization compared to the Type I distribution given by (6.38). 127
7.1	Location of Wildlife Management Area and epicenters of the Elmore Ranch ($M_s = 6.2$) and Superstition Hills ($M_s = 6.6$) events. 130
7.2	Cross-section of flood plain at the Wildlife site. Lower figure shows a closeup of the strata and CPT measurements at a number of locations. . . 132
7.3	Location of instrumentation at the Wildlife site. 133
7.4	Typical realization of soil permeabilities over depth using vertical and horizontal scales of fluctuation of $\theta_v = 1$ and $\theta_h = 42.7$ m respectively. Four adjacent columns of the soil model are shown. 137
7.5	Probability density function of the bounded function (7.7) for four different variances. 138
7.6	Typical realization of soil porosities over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown. 139
7.7	Typical realization of soil elastic moduli over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown. 140
7.8	Typical realization of soil Poisson ratios over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown. 141
7.9	Typical realization of soil dilation angles over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown. 142

Figure	Page
7.10	Realization of base acceleration time histories at the four corner columns of the $80 \times 80 m$ soil model. 145
7.11	Surface response computed by DYNA1D for a particular soil column realization. 146
7.12	Contour maps of the planes having the highest average stress ratio q drawn from two realizations of the soil mass and analyzed by DYNA1D. Contours are drawn at q values of 0.3, 0.5, 0.7, and 0.9. Notice the significant areas with $q > 0.9$ 149
7.13	Average area of the maximal plane, \bar{A}_q , having liquefaction indices in excess of the indicated q thresholds. Averaging is performed over 9 realizations for each line. 152
7.14	Estimated standard deviations, s , of the excursion areas A_q 152
7.15	Estimated number of isolated excursion areas, \bar{N}_q , above the liquefaction index thresholds q . Each line represents the average over 9 realizations and the coefficient of variation of the estimate averages about 1.0. 153
7.16	Estimated cluster measure, Ψ , of the excursion areas above the liquefaction index thresholds q . Each line represents the average over 9 realizations and the coefficient of variation of the estimate averages about 0.5. 153
B.1	Spectral estimates of the locally averaged Ornstein-Uhlenbeck process for $\theta = \frac{1}{2}$ averaged over 500 realizations. 171
B.2	Spectral estimates of the damped oscillatory noise for $\lambda = 8$ and $\theta = 4$ using a neighborhood size of 3. 172
B.3	Spectral estimates of the damped oscillatory noise for $\lambda = 8$ and $\theta = 4$ using a neighborhood size of 5. 173
B.4	Spectral estimates of fractional Gaussian noise with self-similarity parameter $H = 0.95$ and a physical process length of 5 divided into 256 intervals. 174
B.5	Spectral estimates of fractional Gaussian noise with self-similarity parameter $H = 0.95$ and a physical process length of 500 divided into 256 intervals. 175

List of Tables

Table		Page
4.1	Comparison of execution times on a Cyber 205 super-computer.	77
6.1	Computed variances of the local average derivative process.	109
6.2	Empirically determined parameters of (6.24) and variances of the derivative process.	110
6.3	Empirically determined parameters of (6.27) based on the observed average number of holes obtained by simulation.	117
6.4	Empirically determined parameters of (6.31) based on the observed average IG characteristic $\bar{\Gamma}$ obtained by simulation.	120
6.5	Empirically determined effective number of independent samples n_{eff} and parameters of the Type I extreme distribution (6.38).	125
7.1	Material parameters for the layered soil deposit at Wildlife site as estimated by Keane and Prevost [33].	131
7.2	Parameters used in the modeling of soil permeability, k , at the Wildlife site.	136
7.3	Parameters used in the lognormal distribution model of the soil elastic moduli, E , at the Wildlife site.	140
7.4	Parameters used in the modeling of soil dilation angle, Φ , at the Wildlife site.	142
7.5	Summary of random soil properties used in the soil model.	142
7.6	Ground motion scaling factors and vertical scales of fluctuation considered in the probabilistic liquefaction analysis.	150
7.7	Average liquefaction indices $Q_{i^*\ell^*}$ of the maximal planes are averaged over the nine realizations to obtain \bar{Q} . The standard deviation of $Q_{i^*\ell^*}$, σ_Q , is estimated for each case and shown along with the average depth, \bar{D} , to the maximal plane.	150

Chapter 1

Introduction

1.1 General

One of the most magnificent and perhaps disturbing aspect of the universe around us is its “randomness”. Whether or not the universe is truly random is still an open question (which may be largely philosophical in nature). However on a practical level its staggering complexity alone renders it unknowable in exact detail and one must settle for approximations. It is this uncertainty that leads naturally to the representation of nature through analytical or numerical models of random phenomena. Such models have great importance in applications ranging from reliability estimation, optimization, and even human thought. The use of these models is not yet commonplace, however there are a number of reasons for this: First of all they depend heavily on extensive databases to ensure that they are reasonably realistic. This data is often not available for a specific application and can be expensive to obtain. Secondly the solution of complex problems using analytical models has often proved to be difficult if not impossible. The advent of high-speed computers is beginning to allow the solution of many of these but the techniques are as yet not general knowledge. Finally the theory governing all but relatively simple random processes has not been well developed. For example, although it is known that many ‘failures’ are initiated at extrema, little is known about the statistics of extrema in more than one dimension nor how to incorporate such statistics into the analysis or design of a system.

The paucity of data is particularly evident for multi-dimensional processes. A geotechnical engineer attempting to evaluate the reliability of a footing will find very little information dealing with the spatial variability of soil properties. Similarly the spatial distribution of wind speeds in many natural settings is poorly understood as is the distribution of defects or micro-cracks in solids. Although no attempt is made herein to address this lack of data, it may be that the availability of useful multi-dimensional random models will encourage the necessary data collection.

It is the formulation and solution of complex stochastic problems to which this dissertation is primarily devoted. In general if the analytical solution to a problem cannot be found,

the very powerful Monte Carlo approach can in principle always be used. In many ways Monte Carlo simulations are analogous to real life – for example much of our knowledge about the safety of structures is based on years of full-scale ‘simulations’, i.e. the construction and observation of thousands of buildings. Of course the use of computer simulation is much less expensive, not to mention the risk to human life that such full-scale ‘testing’ entails. Optimization problems also often entail the generation of many realizations of the phenomena being modeled. For both of these issues reliable and accurate simulation techniques are essential.

With this in mind, the first goal of this thesis is to evaluate and implement a variety of random field generators. To accomplish this the general properties of the types of random fields to be simulated are examined in this chapter following the work of Yaglom [71], Adler [2], Pugachev [57], and Vanmarcke [25]. In Chapter 2 the simulation of random fields in one, two, or three dimensions using Fourier Transform techniques is explored. Emphasis is placed on implementation issues such as means of improving accuracy. While the computational speed of Fast Fourier transform simulations is impressive, considerable care must be taken to ensure reasonable accuracy. In Chapter 3, the Turning Bands and Moving Average techniques are investigated. The Turning Bands technique is attributed originally to Mathéron [45]. Practical application and implementation issues follow the work of Mantoglou and Wilson [43] and Tompson and Ababou [64]. Although the Turning Bands method can be exceptionally efficient, bands or streaks are apparent in the realizations. Also an equivalent one-dimensional spectral or covariance structure must be determined through an integral equation. This reduces the method’s value as a general purpose easy-to-use simulator. The Moving Average approach shares this drawback with the Turning Bands Method – a weighting function must be determined through an integral equation. Once this has been found, generation of one-dimensional process via the Moving Average method is relatively efficient and accurate. However in higher dimensions, the method becomes extremely cumbersome and was abandoned due to its high cost. Chapter 4 introduces a technique dubbed ‘Local Average Subdivision’ (LAS) of producing one, two, or three-dimensional realizations of ‘local average’ random processes. The LAS method depends on the knowledge of a so-called variance function which was developed and investigated in considerable detail by Vanmarcke [25] and is an alternative to the specification of second-order properties of random processes through their covariance functions. The Local Average Subdivision method is found to offer considerable accuracy, efficiency, along with the ability to condition the field easily. In Chapter 5, the FFT techniques are combined with linear estimation theory to enable the production of a set of spatially correlated non-stationary time histories (which

could represent ground motion at a series of points) and optional conditioning of the field (if some ground motions are known).

The final two chapters of the dissertation apply the simulation methods proposed in Chapters 4 and 5 to illustrate that the methods work and to develop methodology for future studies. In particular, Chapter 6 looks at some of the statistics of excursions and extrema of the two-dimensional locally averaged Gaussian process having a first-order Markov correlation function, herein referred to as the Gauss-Markov process. Existing theories concerning statistics of excursions are presented and compared to data obtained through the analysis of a sequence of realizations. Empirical relationships are developed which more closely fit the data over a range of thresholds and scales of fluctuation. Chapter 7 presents a liquefaction case study in which a soil mass is modeled as a three-dimensional random field subjected to earthquake excitation. A Monte-Carlo analysis is performed employing a non-linear finite element code written by Prevost [55] for two different earthquake intensities and two different soil models.

1.2 Basic Properties of Random Fields

In the following Z is defined to be a random variable taking values from the set Ω which is the sample space. Now let \mathcal{E} be an arbitrary set of elements $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$. Then the scalar random process Z is defined over the set \mathcal{E} to be the family of random variables $Z(\mathbf{x}_1), Z(\mathbf{x}_2), \dots$ corresponding to all elements $\mathbf{x}_1, \mathbf{x}_2, \dots$ in the set \mathcal{E} . In general, the set \mathcal{E} will be considered to be infinite, made up of points in some space R^n . Thus $Z(\mathbf{x})$ can be thought of as a random function on \mathcal{E} and is said to be fully specified if for any k elements $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ in the set \mathcal{E} the cumulative distribution function

$$F_{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k}(b_1, b_2, \dots, b_k) = \mathbf{P} [Z(\mathbf{x}_1) < b_1, Z(\mathbf{x}_2) < b_2, \dots, Z(\mathbf{x}_k) < b_k] \quad (1.1)$$

is given.

The random function $Z(\mathbf{x})$ is called strictly homogeneous if the distribution functions (1.1) are invariant under the translation τ , i.e., if

$$F_{\mathbf{x}_1+\tau, \mathbf{x}_2+\tau, \dots, \mathbf{x}_k+\tau}(b_1, b_2, \dots, b_k) = F_{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k}(b_1, b_2, \dots, b_k). \quad (1.2)$$

In particular, homogeneity implies that all the one-dimensional distribution functions ($k = 1$) are the same and independent of position

$$F_{\mathbf{x}_i}(b) = F(b), \quad \forall \mathbf{x}_i \in \mathcal{E}, \quad (1.3)$$

and that all the two-dimensional distribution functions are only dependent on lags and direction

$$F_{\mathbf{x}_i, \mathbf{x}_j}(b_i, b_j) = F_{\mathbf{x}_i - \mathbf{x}_j, 0}(b_i, b_j). \quad (1.4)$$

The complete description of a random function given by (1.1) is usually far too complex to deal with easily and so we shall restrict ourselves herein to a description based on just the first two moments of the probability distribution. Using Stieltjes integrals these moments are defined as follows

$$m(\mathbf{x}) = \mathbb{E} [Z(\mathbf{x})] = \int_{\Omega} b dF_{\mathbf{x}}(b) \quad (1.5)$$

$$B(\mathbf{x}_i, \mathbf{x}_j) + m(\mathbf{x}_i)m(\mathbf{x}_j) = \mathbb{E} [Z(\mathbf{x}_i)Z(\mathbf{x}_j)] = \int_{\Omega} \int_{\Omega} b_1 b_2 dF_{\mathbf{x}_i, \mathbf{x}_j}(b_1, b_2), \quad (1.6)$$

where $m(\mathbf{x})$ and $B(\mathbf{x}_i, \mathbf{x}_j)$ are usually referred to as the mean and covariance functions respectively. If the mean defined by (1.5) is independent of \mathbf{x} (thus constant) and if the covariance function defined by (1.6) is dependent on just the lag and direction, i.e., if

$$\mathbb{E} [Z(\mathbf{x})] = m, \quad (1.7)$$

$$\mathbb{E} [Z(\mathbf{x}_i)Z(\mathbf{x}_j)] - m^2 = B(\mathbf{x}_i - \mathbf{x}_j),$$

then Z is said to be weakly homogeneous or homogeneous in the wide sense. In this work, the word homogeneous will always be meant in this sense. Furthermore if the covariance function is also independent of direction,

$$\mathbb{E} [Z(\mathbf{x}_i)Z(\mathbf{x}_j)] - m^2 = B(|\mathbf{x}_i - \mathbf{x}_j|), \quad (1.8)$$

then Z is isotropic.

In general the mean and covariance functions will not uniquely describe the random function $Z(\mathbf{x})$, as all the higher-order moments are needed. This is not true, however, of a Gaussian or normal process which is completely described by the first and second moments and for which there is no difference between strict homogeneity and wide-sense homogeneity. Because of this and since a large number of random functions encountered in practice are at least approximately Gaussian, we will restrict ourselves to the study of homogeneous Gaussian processes. Gaussian processes can often be transformed into non-Gaussian processes with little loss in accuracy [72] allowing full (albeit approximate) stochastic representation of a large class of phenomena. In addition we shall always consider the Gaussian

process to have mean zero – a mean function can always be easily added. Thus the only information needed about the homogeneous process $Z(\mathbf{x})$ is carried in the covariance function $B(\boldsymbol{\tau})$.

Examining now some of the properties of the covariance function $B(\boldsymbol{\tau})$ for homogeneous zero-mean processes, one sees from the definition (1.7) that

$$B(0) > 0, \quad (1.9a)$$

$$B(-\boldsymbol{\tau}) = B(\boldsymbol{\tau}), \quad (1.9b)$$

$$|B(\boldsymbol{\tau})| \leq B(0), \quad (1.9c)$$

where property (c) follows from the inequality

$$\mathbb{E} [Z(\mathbf{x} + \boldsymbol{\tau}) \pm Z(\mathbf{x})]^2 = \mathbb{E} [Z^2(\mathbf{x} + \boldsymbol{\tau}) \pm 2Z(\mathbf{x} + \boldsymbol{\tau})Z(\mathbf{x}) + Z^2(\mathbf{x})] \geq 0. \quad (1.10)$$

If, in addition to property (1.9b), the covariance function $B(\boldsymbol{\tau})$ is an even function with respect to each component of the lag vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)$, i.e.,

$$B(\tau_1, \dots, \tau_k, \dots, \tau_n) = B(\tau_1, \dots, -\tau_k, \dots, \tau_n), \quad (1.11)$$

for any $k \in \{1, 2, \dots, n\}$, then $B(\boldsymbol{\tau})$ is said to be quadrant symmetric (q.s.) [25]. Functions possessing this property are fully defined if the function is known only over the lag-space quadrant of positive lags.

Equation (1.7) also implies that the inequality

$$\sum_{j,k=1}^m B(\boldsymbol{\tau}_j - \boldsymbol{\tau}_k) a_j a_k \geq 0 \quad (1.12)$$

holds for any m real numbers a_1, a_2, \dots, a_m and any $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_m$. This follows from the relation

$$\begin{aligned} \sum_{j,k=1}^m B(\boldsymbol{\tau}_j - \boldsymbol{\tau}_k) a_j a_k &= \sum_{j,k=1}^m \mathbb{E} [Z(\boldsymbol{\tau}_j) Z(\boldsymbol{\tau}_k)] a_j a_k \\ &= \mathbb{E} \left[\sum_{j=1}^m Z(\boldsymbol{\tau}_j) a_j \right]^2 \geq 0 \end{aligned}$$

and so $B(\boldsymbol{\tau})$ is a positive definite function. Khinchin [35] also shows that the converse result is true, i.e., every positive definite function of a real (or integral) argument is the covariance function of a homogeneous random process (or sequence).

1.3 Spectral Representation of Homogeneous Processes

In the following it is useful to allow the random function $Z(\mathbf{x})$ to take values from the set of complex numbers. In this case $Z(\mathbf{x})$ can be expressed as

$$Z(\mathbf{x}) = R(\mathbf{x}) + iQ(\mathbf{x}), \quad (1.13)$$

where $R(\mathbf{x})$ and $Q(\mathbf{x})$ are real zero-mean random functions. The associated covariance function also becomes complex in general and is given by

$$B(\boldsymbol{\tau}) = \text{E} [Z(\mathbf{x} + \boldsymbol{\tau}) \overline{Z(\mathbf{x})}], \quad (1.14)$$

where the overbar denotes the complex conjugate. Properties (1.9a) and (1.9c) remain valid but (1.9b) becomes

$$B(-\boldsymbol{\tau}) = \overline{B(\boldsymbol{\tau})} \quad (1.15)$$

and the inequality (1.12) showing positive definiteness becomes

$$\sum_{j,k=1}^m B(\boldsymbol{\tau}_j - \boldsymbol{\tau}_k) a_j \overline{a_k} \geq 0, \quad (1.16)$$

where a_1, a_2, \dots, a_m are arbitrary complex numbers.

An important result by Khinchin [35] allows the covariance function of any mean-square continuous homogeneous zero-mean random process to be expressed in the form of an integral

$$B(\boldsymbol{\tau}) = \int_{-\infty}^{\infty} e^{i\boldsymbol{\omega} \cdot \boldsymbol{\tau}} d\Sigma(\boldsymbol{\omega}), \quad (1.17)$$

where $\Sigma(\boldsymbol{\omega})$ is a real non-decreasing bounded function. A sufficient condition for mean-square continuity of $Z(\mathbf{x})$ is that the covariance function $B(\boldsymbol{\tau})$ be continuous at the point $\boldsymbol{\tau} = 0$. In the following, we will assume that this condition is satisfied. If the absolute value of the correlation function decreases sufficiently rapidly as $|\boldsymbol{\tau}| \rightarrow \infty$ such that

$$\int_{-\infty}^{\infty} |B(\boldsymbol{\tau})| d\boldsymbol{\tau} < \infty \quad (1.18)$$

then the spectral distribution function $\Sigma(\boldsymbol{\omega})$ can be written

$$\Sigma(\boldsymbol{\omega}) = \int_{-\infty}^{\omega_1} \int_{-\infty}^{\omega_2} \cdots \int_{-\infty}^{\omega_n} S(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (1.19)$$

where $\omega_1, \omega_2, \dots, \omega_n$ are the components of $\boldsymbol{\omega}$. The function $S(\boldsymbol{\omega})$ is called the spectral density function of the process $Z(\mathbf{x})$ and the condition that $\Sigma(\boldsymbol{\omega})$ be non-decreasing implies that $S(\boldsymbol{\omega}) \geq 0$ for all $\boldsymbol{\omega}$. The use of (1.19) in (1.17) leads to the Wiener-Khinchin relationships

$$B(\boldsymbol{\tau}) = \int_{-\infty}^{\infty} e^{i\boldsymbol{\omega} \cdot \boldsymbol{\tau}} S(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (1.20)$$

$$S(\boldsymbol{\omega}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} e^{-i\boldsymbol{\omega} \cdot \boldsymbol{\tau}} B(\boldsymbol{\tau}) d\boldsymbol{\tau}. \quad (1.21)$$

The restriction (1.18) is satisfied by most covariance functions of practical interest. One notable exception is that of the fractional noise process developed by Mandelbrot and van Ness [41] and discussed in Chapter 4. Introducing now a random field $W(\boldsymbol{\omega})$ with orthogonal increments satisfying

$$E [W(\boldsymbol{\omega})] = 0 \quad (1.22a)$$

$$E [W(\boldsymbol{\omega})\overline{W(\boldsymbol{\omega})}] = \Sigma(\boldsymbol{\omega}) \quad (1.22b)$$

$$E [W(\Delta\boldsymbol{\omega})\overline{W(\Delta\boldsymbol{\omega})}] = \Sigma(\Delta\boldsymbol{\omega}) \quad (1.22c)$$

$$E [W(\Delta\boldsymbol{\omega}_i)\overline{W(\Delta\boldsymbol{\omega}_j)}] = 0 \quad \text{if } \Delta\boldsymbol{\omega}_i \cap \Delta\boldsymbol{\omega}_j = \phi \quad (1.22d)$$

where ϕ denotes the empty set, $\Delta\boldsymbol{\omega} = (\omega_1, \omega_1 + \Delta\omega_1] \times \cdots \times (\omega_n, \omega_n + \Delta\omega_n]$ is an n -dimensional interval starting at the point $(\omega_1, \omega_2, \dots, \omega_n)$, and

$$\Sigma(\Delta\boldsymbol{\omega}) = \Sigma((\omega_1, \omega_1 + \Delta\omega_1] \times \cdots \times (\omega_n, \omega_n + \Delta\omega_n]) \quad (1.23)$$

is the amount of spectral power contained in the interval $\Delta\boldsymbol{\omega}$. For a mean square continuous homogeneous zero-mean random process, the spectral representation theorem tells us that $Z(\boldsymbol{x})$ has the following mean square integral representation

$$Z(\boldsymbol{x}) = \int_{-\infty}^{\infty} e^{i\boldsymbol{x} \cdot \boldsymbol{\omega}} W(d\boldsymbol{\omega}). \quad (1.24)$$

Notice that if (1.19) holds then $E [W(d\boldsymbol{\omega})\overline{W(d\boldsymbol{\omega})}] = S(\boldsymbol{\omega})d\boldsymbol{\omega}$.

The importance of (1.24) is that it allows a realization of the field $Z(\boldsymbol{x})$ to be created simply by generating a sequence of realizations of the random variable $W(d\boldsymbol{\omega})$ according to the distribution given by (1.22a–1.22d) and summing them up. Notice that $W(d\boldsymbol{\omega}_i)$ is independent of $W(d\boldsymbol{\omega}_j)$ if $i \neq j$. In fact if $Z(\boldsymbol{x})$ is defined on a regular lattice then (1.24) becomes

$$\begin{aligned} Z(\boldsymbol{x}_m) &= \int_{-\pi}^{\pi} e^{i\boldsymbol{x}_m \cdot \boldsymbol{\omega}} W(d\boldsymbol{\omega}) \\ &= \lim_{K \rightarrow \infty} \sum_{k=-K}^K e^{i\boldsymbol{x}_m \cdot \boldsymbol{\omega}_k} W(\Delta\boldsymbol{\omega}_k), \end{aligned} \quad (1.25)$$

where it is recognized that the sum is actually an n -dimensional sum and each component ω_{jk} of the vector ω_k has value

$$\omega_{jk} = \frac{k\pi}{K}. \quad (1.26)$$

In practice K is taken as a finite number large enough to adequately subdivide the frequency interval $[-\pi, \pi]$ so that little loss in accuracy occurs.

If $Z(x)$ is mean square continuous, homogeneous, and takes values from the real numbers only, then (1.6) implies that $B(\tau)$ must also be real. This in turn implies that $S(-\omega) = S(\omega)$ since the imaginary part of (1.20) must drop out. For such a real process (1.24) can be written as

$$Z(x) = \lim_{K \rightarrow \infty} \sum_{k=-K}^K \left\{ \mathcal{A}(\Delta\omega_k) \cos(x \cdot \omega_k) + \mathcal{B}(\Delta\omega_k) \sin(x \cdot \omega_k) \right\}, \quad (1.27)$$

where $\mathcal{A}(\Delta\omega_k)$ and $\mathcal{B}(\Delta\omega_k)$ are independent identically distributed random interval functions with

$$\begin{aligned} \text{E} [\mathcal{A}(\Delta\omega_k)] &= \text{E} [\mathcal{B}(\Delta\omega_k)] = 0 \\ \text{E} [\mathcal{A}(\Delta\omega_k)]^2 &= \text{E} [\mathcal{B}(\Delta\omega_k)]^2 = S(\omega_k)\Delta\omega \\ \text{E} [\mathcal{A}(\Delta\omega_k)\mathcal{A}(\Delta\omega_j)] &= \text{E} [\mathcal{B}(\Delta\omega_k)\mathcal{B}(\Delta\omega_j)] = 0 \quad \text{if } k \neq j \end{aligned} \quad (1.28)$$

in the limit as $\Delta\omega \rightarrow 0$.

1.4 Local Average Processes

Virtually all the information gathered concerning the world around us comes to us in the form of an average. The human eye cannot resolve infinite detail and so delivers averages at some scale of resolution - more averaging being performed for some of us than others! Similarly both laboratory and field measurements usually represent averages over some region. It is thus appropriate to investigate the properties of random processes that have been averaged over some domain. Let $Z_D(x)$ be defined as the average of $Z(x)$ over some domain D of size $|D| = T_1 T_2 \cdots T_n$ centered at x ,

$$Z_D(x) = \frac{1}{|D|} \int_{x_1-T_1/2}^{x_1+T_1/2} \cdots \int_{x_n-T_n/2}^{x_n+T_n/2} Z(\xi_1, \dots, \xi_n) d\xi_1 \cdots d\xi_n.$$

Obviously, if $Z(\mathbf{x})$ is a zero mean function then so is $Z_D(\mathbf{x})$. The covariance between local averages can be derived by considering two averaging domains of size $|D_\alpha|$ and $|D_\beta|$ centered at the points \mathbf{x}_α and \mathbf{x}_β and taking expectations,

$$\mathbb{E} [Z_{D_\alpha}(\mathbf{x}_\alpha) Z_{D_\beta}(\mathbf{x}_\beta)] = \frac{1}{|D_\alpha| |D_\beta|} \int_{D_\alpha} \int_{D_\beta} \mathbb{E} [Z(\boldsymbol{\xi}) Z(\boldsymbol{\eta})] d\boldsymbol{\xi} d\boldsymbol{\eta} \quad (1.29)$$

which for a homogeneous process is equivalent to

$$B_{D_\alpha D_\beta}(\mathbf{x}_\alpha - \mathbf{x}_\beta) = \frac{1}{|D_\alpha| |D_\beta|} \int_{D_\alpha} \int_{D_\beta} B(\boldsymbol{\xi} - \boldsymbol{\eta}) d\boldsymbol{\xi} d\boldsymbol{\eta}. \quad (1.30)$$

Letting $|D_\alpha| = |D_\beta| = |D|$ and taking $\mathbf{x}_\alpha = \mathbf{x}_\beta = (\frac{T_1}{2}, \frac{T_2}{2}, \dots, \frac{T_n}{2})$, the point variance of the homogeneous process $Z_D(\mathbf{x})$, $B_D(0)$, can be shown to change with the averaging region D according to

$$B_D(0) = \frac{1}{|D|^2} \int_0^{T_1} \int_0^{T_1} \cdots \int_0^{T_n} \int_0^{T_n} B(\xi_1 - \eta_1, \dots, \xi_n - \eta_n) d\xi_1 d\eta_1 \cdots d\xi_n d\eta_n, \quad (1.31)$$

a $2n$ -fold integral. Equation (1.31) can also be written in the form

$$B_D(0) = \sigma^2 \gamma(T_1, T_2, \dots, T_n), \quad (1.32)$$

where $\sigma^2 = B(0)$ is the point variance of $Z(\mathbf{x})$. The function $\gamma(D)$ is called the variance function of the process Z and is essentially defined by (1.32). It relates the point variance of the process Z to the variance of the averaged process Z_D . The variance function can thus be viewed as a measure of the reduction of the point variance σ^2 under local averaging. Formally,

$$\begin{aligned} \gamma(T_1, \dots, T_n) &= \frac{1}{|D|^2} \int_0^{T_1} \int_0^{T_1} \cdots \int_0^{T_n} \int_0^{T_n} \rho(\xi_1 - \eta_1, \dots, \xi_n - \eta_n) d\xi_1 d\eta_1 \cdots d\xi_n d\eta_n, \\ &= \frac{1}{|D|^2} \int_{-T_1}^{T_1} \cdots \int_{-T_n}^{T_n} \left(|T_1| - |\xi_1| \right) \cdots \left(|T_n| - |\xi_n| \right) \rho(\xi_1, \dots, \xi_n) d\xi_1 \cdots d\xi_n, \end{aligned} \quad (1.33)$$

where $\rho(\boldsymbol{\tau}) = B(\boldsymbol{\tau})/\sigma^2$. The fact that $\rho(-\boldsymbol{\tau}) = \rho(\boldsymbol{\tau})$ was used to reduce (1.33) from a $2n$ -fold integration to an n -fold integration. It can be easily seen from (1.32) and (1.33) that if $|D| = 0$ then $\gamma(0) = 1$. Also since $0 \leq \rho(\boldsymbol{\tau}) \leq 1$ then $0 \leq \gamma(T_1, \dots, T_n) \leq 1$.

The covariance between two local averages (1.30) can also be expressed in terms of the variance function; if D_α is a cube centered at \mathbf{x}_α with volume $|D_\alpha| = T_1 T_2 T_3$ and D_β a cube centered at \mathbf{x}_β with volume $|D_\beta| = T'_1 T'_2 T'_3$ then

$$B_{D_\alpha D_\beta}(\mathbf{x}_\alpha - \mathbf{x}_\beta) = \frac{\sigma^2}{2^3 |D_\alpha| |D_\beta|} \sum_{j=0}^3 \sum_{k=0}^3 \sum_{\ell=0}^3 (-1)^j (-1)^k (-1)^\ell |D_{jk\ell}|^2 \gamma(T_{1j}, T_{2k}, T_{3\ell}). \quad (1.34)$$

The dimensions T_{ij} are shown in Figure 1.1 and $|D_{jkl}| = T_{1j}T_{2k}T_{3l}$. The general relationship for the n -dimensional case is

$$B_{D_\alpha D_\beta}(\mathbf{x}_\alpha - \mathbf{x}_\beta) = \frac{\sigma^2}{2^n |D_\alpha| |D_\beta|} \sum_{j_1=0}^3 \cdots \sum_{j_n=0}^3 (-1)^{j_1} \cdots (-1)^{j_n} |D_{j_1 j_2 \dots j_n}|^2 \gamma(T_{1j_1}, \dots, T_{nj_n}). \quad (1.35)$$

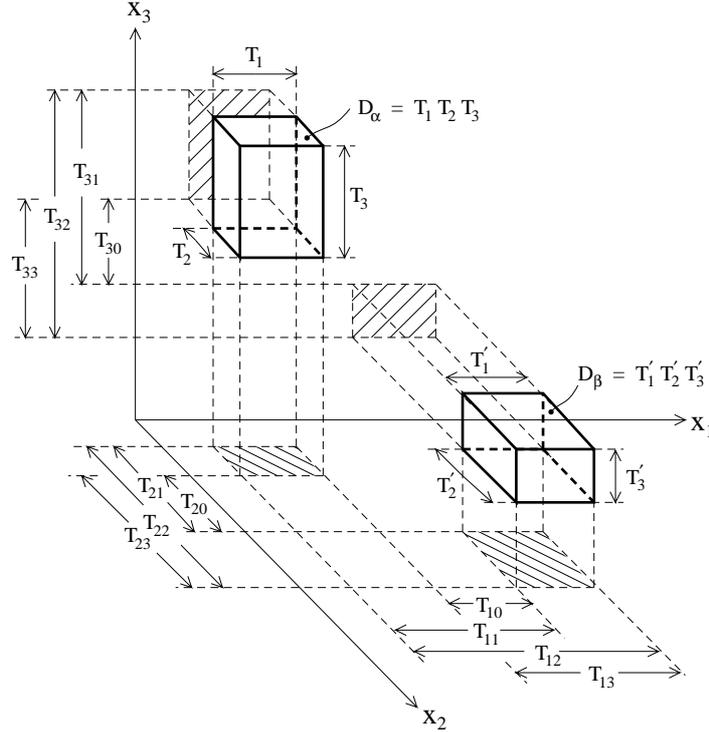


Figure 1.1 Distances characterizing the relative location of the volumes D_α and D_β in three-dimensional space (courtesy of Vanmarcke [25]).

Another useful measure of the statistics of random fields is the scale of fluctuation θ defined by Vanmarcke for one-dimensional processes to be

$$\theta = \lim_{t \rightarrow \infty} T \gamma(T). \quad (1.36)$$

This limit will exist if

$$\lim_{|\tau| \rightarrow \infty} \tau B(\tau) = 0. \quad (1.37)$$

If (1.37) is satisfied, an alternative relationship giving the scale of fluctuation is

$$\theta = \frac{2}{\sigma^2} \int_0^\infty B(\tau) d\tau, \quad (1.38)$$

which can be seen to be a measure of how rapidly the correlation function falls off to zero.

1.5 Parameter Estimation and Ergodicity

In the following chapters where several methods of simulating random fields are presented, a means of evaluating the output of an algorithm is essential. This will be accomplished by comparing the estimated covariance function with the known (or assumed) covariance function. A one-dimensional estimator will be used exclusively since it can always be applied in various directions to estimate the multi-dimensional covariance structure. Since all the simulation techniques used will employ a known zero mean over the ensemble, the following unbiased estimator along a line will be used

$$\hat{B}(j\Delta x) = \frac{1}{N_f} \sum_{n=1}^{N_f} \frac{1}{N-j} \sum_{k=1}^{N-j} Z_n(k\Delta x) Z_n((k+j)\Delta x), \quad j = 0, 1, \dots, N-1 \quad (1.39)$$

where N_f is the number of realizations of the process over which to average, N is the number of discrete field points along the line in question and the subscript on Z signifies that it is a realization. It is assumed that the field points are equispaced and are at locations $(\Delta x, 2\Delta x, \dots, N\Delta x)$ to simplify notation. It is easy to see that for a known mean zero process, $E[\hat{B}(\tau)] = B(\tau)$.

Another aspect of the evaluation of a simulation algorithm is its rate of convergence to the desired mean and covariance structure. If a homogeneous random process is ergodic then its mean and correlation function can be found from a single realization of *infinite* extent

$$m = E[Z(\mathbf{x})] = \lim_{|D| \rightarrow \infty} \frac{1}{|D|} \int_D Z(\mathbf{x}) d\mathbf{x}, \quad (1.40)$$

$$B(\tau) + m^2 = E[Z(\mathbf{x} + \tau)Z(\mathbf{x})] = \lim_{|D| \rightarrow \infty} \frac{1}{|D|} \int_D Z(\mathbf{x} + \tau)Z(\mathbf{x}) d\mathbf{x}. \quad (1.41)$$

In order to guarantee the validity of the above relationships, two conditions must be imposed on the homogeneous random function $Z(\mathbf{x})$. For Gaussian processes these conditions are

$$\lim_{|D| \rightarrow \infty} \frac{1}{|D|} \int_D B(\tau) d\tau = 0, \quad (1.42a)$$

$$\lim_{|D| \rightarrow \infty} \frac{1}{|D|} \int_D |B(\tau)|^2 d\tau = 0, \quad (1.42b)$$

which are clearly met if

$$\lim_{\tau \rightarrow \infty} B(\tau) = 0. \quad (1.43)$$

Thus ergodicity implies that the correlation between points separated by a large distance is negligible. A realization obtained from a particular algorithm is said to be ergodic if the

desired mean and correlation structure can be obtained using equations (1.40) and (1.41) respectively. Of course realizations of infinite extent are never produced and so one cannot expect a finite realization to be ‘ergodic’ (the word loses meaning in this context). In fact for finite-domain realizations averaging must be performed over the whole space Ω of outcomes in order to exactly calculate m and $B(\tau)$. Although some algorithms may produce realizations which more closely approximate the desired statistics when averaged over a fixed (small) number of realizations than others, this becomes a matter of judgement. There is also the argument that since most natural processes are far from ergodic, why should a simulation of the process be ergodic? In general this issue will only be mentioned in passing. All comparisons between the statistics estimated from the realizations of a given algorithm and the exact will be computed using (1.39) and will note the number of realizations over which averaging was performed. It is usually easy to tell if the estimated statistics are correctly approaching the desired statistics.

Chapter 2

Fourier Transform Techniques

2.1 One-Dimensional Processes

Much of the theory involving Fourier transform techniques was developed in the context of time series analysis and can be found in Brillinger *et al.* [8] and Priestley [56]. Although these works are directed primarily towards the spectral analysis of time series, they contain the basic concepts pertinent to simulation using Fourier transforms. Through the spectral representation theorem, it was shown in Chapter 1 that a mean square continuous, homogeneous real process can be expressed as a sum of sinusoids. For the one-dimensional case, the process $\mathcal{Z}(x)$ can be defined as

$$\mathcal{Z}(x) = \sum_{k=-K}^K \{ \mathcal{A}_k \cos(x\omega_k) + \mathcal{B}_k \sin(x\omega_k) \}, \quad (2.1)$$

which, for finite K , is an approximation to the exact process, Z , defined by (1.27). For numerical work, this approximation can be made as close as desired by increasing K and the frequency range. Throughout this and following chapters the symbol \mathcal{Z} will be used to denote the algorithmic process. The coefficients \mathcal{A}_k and \mathcal{B}_k are independent zero-mean random variables which are taken to be Gaussian. Their variances are determined by the spectral density function of the process,

$$E [\mathcal{A}_k]^2 = E [\mathcal{B}_k]^2 = S(\Delta\omega_k), \quad (2.2)$$

and $\mathcal{A}_k \perp \mathcal{A}_j$ if $k \neq j$ (similarly for \mathcal{B}_k). To make the following explanations clearer, the interval function notation used in Chapter 1 and in (2.2) will now be abandoned and by $\Delta\omega$ we will henceforth mean the length of the interval so that (2.2) will be written

$$E [\mathcal{A}_k]^2 = E [\mathcal{B}_k]^2 = S(\omega_k)\Delta\omega_k, \quad (2.3)$$

and for $\omega_1 < \omega_2 < \dots < \omega_K$, the interval length $\Delta\omega_k$ is defined by

$$\Delta\omega_k = \frac{1}{2}(\omega_{k+1} - \omega_{k-1}). \quad (2.4)$$

This interval is usually taken to be constant, however in some cases, such as when the spectral density function changes rapidly, the ability to change spectral resolution is valuable. On the other hand, doing so eliminates (in general) the use of the Fast Fourier Transform techniques to be discussed here and so we will consider the frequency interval length to be constant.

Since $S(-\omega) = S(\omega)$, the one-sided spectral density function $G(\omega)$ can be defined as

$$G(\omega) = 2S(\omega), \quad \omega \geq 0 \quad (2.5)$$

such that the point variance, σ^2 , is preserved when integrating over the positive frequencies only,

$$\sigma^2 = \int_{-\infty}^{\infty} S(\omega) d\omega = \int_0^{\infty} G(\omega) d\omega. \quad (2.6)$$

Using the one-sided power spectral density function, (2.1) can be written in terms of the non-negative frequencies

$$\mathcal{Z}(x) = \sum_{k=1}^K \{ \mathcal{A}_k \cos(x\omega_k) + \mathcal{B}_k \sin(x\omega_k) \}. \quad (2.7)$$

The variance of \mathcal{A}_k and \mathcal{B}_k become

$$\text{E} [\mathcal{A}_1]^2 = \text{E} [\mathcal{B}_1]^2 = \frac{1}{2} G(\omega_1) \Delta\omega, \quad (2.8)$$

$$\text{E} [\mathcal{A}_k]^2 = \text{E} [\mathcal{B}_k]^2 = G(\omega_k) \Delta\omega, \quad k = 2, 3, \dots, K$$

where it was assumed that $\omega_1 = 0$. The simulation thus involves generating realizations for \mathcal{A}_k and \mathcal{B}_k in (2.7).

An equivalent way of writing (2.7) is

$$\mathcal{Z}(x) = \sum_{k=1}^K \mathcal{C}_k \cos(x\omega_k + \Phi_k), \quad (2.9)$$

in which Φ_k is a random phase angle uniformly distributed on the interval $[0, 2\pi]$ and \mathcal{C}_k is Rayleigh distributed with

$$\mathbf{E} [C_k] = \sqrt{\frac{\pi}{2}G(\omega_k)\Delta\omega}, \quad (2.10)$$

$$\mathbf{E} [C_k]^2 = 2G(\omega_k)\Delta\omega. \quad (2.11)$$

Taking C_k to be deterministic, as proposed by Shinozuka and Jan [62],

$$C_k = \sqrt{2G(\omega_k)\Delta\omega}, \quad (2.12)$$

yields a process with the exact (in the limit) spectral power at each frequency reflected in every realization. This formulation has the advantage of converging more rapidly to the desired statistics but the upper bound of $Z(x)$ over the space of outcomes is

$$Z \leq \sum_{k=1}^K \sqrt{2G(\omega_k)\Delta\omega} \quad (2.13)$$

which may be an unrealistic restriction. As it is often the extremal properties of $Z(x)$ which are of interest, (2.7) will be used with random coefficients in the following.

Consider now the special case of the discrete process $Z_j \simeq Z(x_j)$, $j = 1, 2, \dots, K$ where the coordinates x_j and the frequencies ω_k are selected according to

$$x_j = (j - 1)\Delta x = (j - 1) \frac{L}{K - 1}, \quad j = 1, 2, \dots, K \quad (2.14)$$

$$\omega_k = (k - 1)\Delta\omega = (k - 1) \frac{2\pi(K - 1)}{KL}, \quad k = 1, 2, \dots, K, \quad (2.15)$$

where L is the physical length of the process ($L = x_K$). The coefficients can be combined to form a single complex coefficient

$$\mathcal{X}_k = \mathcal{A}_k - i\mathcal{B}_k, \quad (2.16)$$

which leads to the discrete Fourier transform (DFT) pair

$$Z_j = \sum_{k=1}^K \mathcal{X}_k e^{i2\pi(k-1)(j-1)/K}, \quad (2.17)$$

$$\mathcal{X}_k = \frac{1}{K} \sum_{j=1}^K Z_j e^{-i2\pi(k-1)(j-1)/K}. \quad (2.18)$$

It can be seen that $\mathcal{Z}_{K+j} = \mathcal{Z}_j$ and so the process \mathcal{Z} is periodic with period T where

$$T = \frac{LK}{K-1}. \quad (2.19)$$

Furthermore if \mathcal{Z} is real then

$$\begin{aligned} \mathcal{A}_k &= \frac{1}{K} \sum_{j=1}^K \mathcal{Z}_j \cos\left(\frac{2\pi(k-1)(j-1)}{K}\right), & \mathcal{B}_k &= \frac{1}{K} \sum_{j=1}^K \mathcal{Z}_j \sin\left(\frac{2\pi(k-1)(j-1)}{K}\right), \\ &= \mathcal{A}_{K-k+2} & &= -\mathcal{B}_{K-k+2}. \end{aligned} \quad (2.20)$$

From a practical point of view, the generation of a process in the form of (2.17) can be accomplished using the very efficient Fast Fourier Transform (FFT) algorithm developed by Cooley and Tukey [16]. FFT algorithms are commonly available [30, 58, 52, 73] and efficient implementation issues are discussed in some detail by Nobile and Roberto [51], Harris *et al.* [28], Skinner [63], McClellan and Rader [47], and Polge *et al.* [54]. For one-dimensional processes on scalar machines, the FFT program given by Newland [50] was employed which used a standard radix-2 butterfly (restricted to processes of length 2^p). On the CYBER 205 supercomputer, a vectorized FFT routine from the MAGEV library based on a prime factor algorithm ($N = 2^p 3^q 5^r$) was used.

The considerable savings in computer time that the FFT algorithm affords does not come without a price. Care must be taken to ensure that the statistics of the process are preserved reasonably accurately. First of all it is easy to show that the covariance function of a real FFT process is symmetric about a lag of $T/2$. Writing $\hat{B}_k = \hat{B}(k\Delta x)$ to denote the covariance function of the FFT process (for this proof, $k = 0, 1, \dots, K-1$),

$$\begin{aligned} \hat{B}_k &= \mathbf{E} \left[\mathcal{Z}_{\ell+k} \overline{\mathcal{Z}_\ell} \right] \\ &= \mathbf{E} \left[\sum_{j=1}^K \mathcal{X}_j \exp \left\{ i \left(\frac{2\pi(\ell+k-1)(j-1)}{K} \right) \right\} \sum_{m=1}^K \overline{\mathcal{X}_m} \exp \left\{ -i \left(\frac{2\pi(\ell-1)(m-1)}{K} \right) \right\} \right] \\ &= \sum_{j=1}^K \mathbf{E} \left[\mathcal{X}_j \overline{\mathcal{X}_j} \right] \exp \left\{ i \left(\frac{2\pi(j-1)k}{K} \right) \right\}, \end{aligned} \quad (2.21)$$

where use was made of the fact that $\mathbf{E} \left[\mathcal{X}_j \overline{\mathcal{X}_m} \right] = 0$ for $j \neq m$ (overbar denotes the complex conjugate). Similarly one can derive

$$\hat{B}_{K-k} = \sum_{j=1}^K \mathbf{E} \left[\mathcal{X}_j \overline{\mathcal{X}_j} \right] \exp \left\{ -i \left(\frac{2\pi(j-1)k}{K} \right) \right\}$$

$$= \overline{\hat{B}_k} \quad (2.22)$$

since $E[\mathcal{X}_j \overline{\mathcal{X}_j}]$ is real. The covariance function of a real process is also real in which case (2.22) becomes simply

$$\hat{B}_{K-k} = \hat{B}_k. \quad (2.23)$$

This property is illustrated in Figure 2.1 for a process whose desired covariance function is the simple exponential (Ornstein-Uhlenbeck process)

$$B(\tau) = \sigma^2 e^{-\frac{|\tau|}{2}}. \quad (2.24)$$

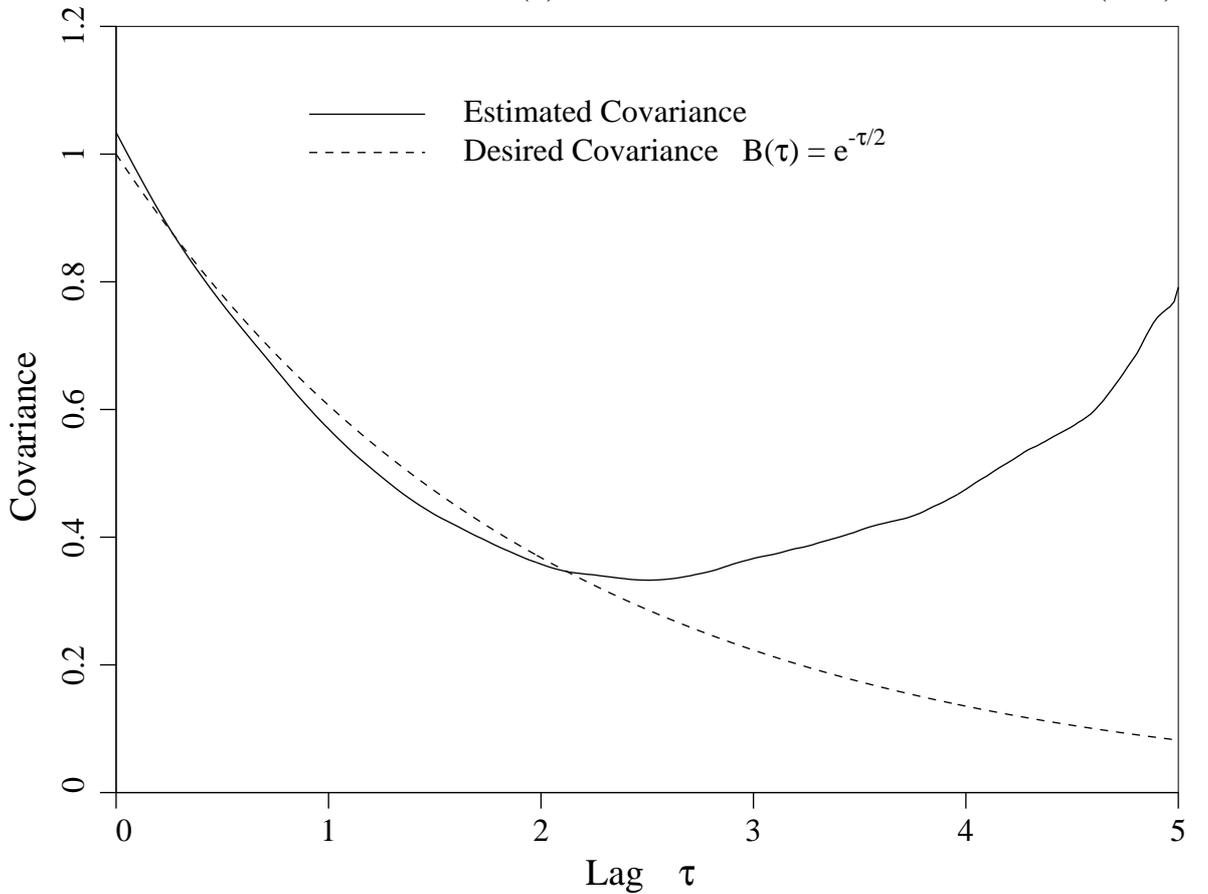


Figure 2.1 Comparison of estimated and exact covariance function for a process generated using the Fast Fourier Transform (averaged over 200 realizations).

It can be seen that for relatively long scales of fluctuation, θ , (where the covariance function decreases slowly compared to the process length) a physical field size greater than that required must be selected and the excess ignored. In general if the distance τ^* is defined to be the lag at which the covariance function decreases in magnitude to a sufficiently small value (say 0.1) then the minimum field size, L_{min} , can be determined approximately from the desired field size L as

$$L_{min} \simeq L + \tau^*. \quad (2.25)$$

For the Ornstein-Uhlenbeck process with covariance function

$$B(\tau) = \sigma^2 e^{-\frac{2|\tau|}{\theta}}, \quad (2.26)$$

the value of τ^* can be taken equal to the scale of fluctuation,

$$\tau^* = \theta. \quad (2.27)$$

so that for such a process $L_{min} = L + \theta$ implying that a field size of at least 9 should have been used for the process shown in Figure (2.1) and the last four units ignored.

The second problem associated with the use of the FFT method has to do with the discretization of the frequency axis. Figure (2.2) illustrates an example in which an overly coarse frequency increment results in a poor estimation of the point variance. The spectral density function shown is that corresponding to the Ornstein-Uhlenbeck process (2.26) with $\theta = 4$.

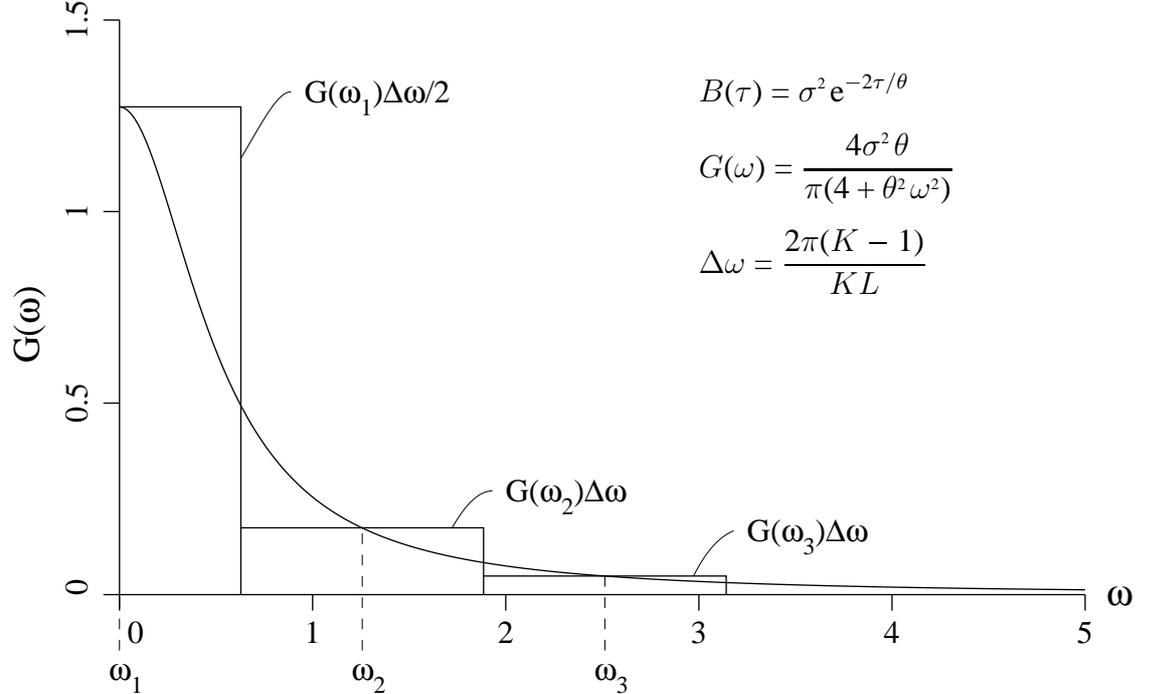


Figure 2.2 Example of overly coarse frequency discretization resulting in a poor estimation of point variance ($L = 5$ and $\theta = 4$).

The frequency increment $\Delta\omega = 2\pi(K - 1)/KL = 2\pi/T$ must be small enough so that the sequence $\frac{1}{2}G(\omega_1)\Delta\omega, G(\omega_2)\Delta\omega, \dots$ adequately approximates the exact spectral density function, particularly in regions where $G(\omega)$ changes rapidly. This represents a further limitation on the minimum size of the field, L_{min} and its discretization.

In the following sub-sections two methods of improving the accuracy of the simulation using the FFT approach are suggested. Both aim at approximating more closely the true point variance of the process, the most important of its second-order statistics.

2.1.1 Spectral Formulation

If the process $\mathcal{Z}(x)$ is to be real, a review of Equation (2.20) shows that the random coefficients \mathcal{A}_{K-k+2} and \mathcal{B}_{K-k+2} must be determined by their counterparts \mathcal{A}_k and \mathcal{B}_k for $k = 2, 3, \dots, \frac{K}{2}$. This in turn implies that \mathcal{A}_k and \mathcal{B}_k need only have realizations generated for them over $k = 1, 2, \dots, 1 + \frac{K}{2}$ according to

$$\begin{aligned} \mathbb{E} [\mathcal{A}_1]^2 &= \frac{1}{2}G(\omega_1)\Delta\omega, \\ \mathbb{E} [\mathcal{A}_k]^2 &= \mathbb{E} [\mathcal{B}_k]^2 = \frac{1}{4}G(\omega_k)\Delta\omega, \quad k = 2, 3, \dots, \frac{K}{2} \end{aligned} \quad (2.28)$$

$$\mathbb{E} [\mathcal{A}_{1+K/2}]^2 = G(\omega_k)\Delta\omega$$

with the added requirement that $\mathcal{B}_1 = \mathcal{B}_{1+K/2} = 0$.

The variances of the coefficients \mathcal{A}_k and \mathcal{B}_k can be obtained in a more consistent fashion using the inverse relationships (2.20)

$$\mathbb{E} [\mathcal{A}_k]^2 = \frac{1}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K \mathbb{E} [Z_j Z_\ell] \cos\left(\frac{2\pi(k-1)(j-1)}{K}\right) \cos\left(\frac{2\pi(k-1)(\ell-1)}{K}\right) \quad (2.29)$$

using the exact covariance $\mathbb{E} [Z_j Z_\ell]$ which is given by (1.20) for a zero-mean process. A discrete approximation to (1.20) is

$$\mathbb{E} [Z_j Z_\ell] \simeq \Delta\omega \sum_{m=1}^K G(\omega_m) \cos\left(\frac{2\pi(m-1)(j-\ell)}{K}\right). \quad (2.30)$$

Substituting (2.30) into (2.29) leads to

$$\begin{aligned} \mathbb{E} [\mathcal{A}_k]^2 &= \frac{\Delta\omega}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K \sum_{m=1}^K G(\omega_m) \cos\left(\frac{2\pi(m-1)(j-\ell)}{K}\right) C_{kj} C_{k\ell} \\ &= \frac{\Delta\omega}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K \sum_{m=1}^K G(\omega_m) \{C_{mj} C_{kj} C_{m\ell} C_{k\ell} + S_{mj} C_{kj} S_{m\ell} C_{k\ell}\} \\ &= \frac{\Delta\omega}{K^2} \sum_{m=1}^K G(\omega_m) \sum_{j=1}^K C_{mj} C_{kj} \sum_{\ell=1}^K C_{m\ell} C_{k\ell} \\ &\quad + \frac{\Delta\omega}{K^2} \sum_{m=1}^K G(\omega_m) \sum_{j=1}^K S_{mj} C_{kj} \sum_{\ell=1}^K S_{m\ell} C_{k\ell}, \end{aligned} \quad (2.31)$$

where

$$C_{kj} = \cos\left(\frac{2\pi(k-1)(j-1)}{K}\right),$$

$$S_{kj} = \sin\left(\frac{2\pi(k-1)(j-1)}{K}\right).$$

To reduce (2.31) further, use is made of the following two identities

$$1) \quad \sum_{k=1}^K \sin\left(\frac{2\pi(m-1)(k-1)}{K}\right) \cos\left(\frac{2\pi(j-1)(k-1)}{K}\right) = 0$$

$$2) \quad \sum_{k=1}^K \cos\left(\frac{2\pi(m-1)(k-1)}{K}\right) \cos\left(\frac{2\pi(j-1)(k-1)}{K}\right) = \begin{cases} 0, & \text{if } m \neq j \\ \frac{K}{2}, & \text{if } m = j \text{ or } K - j + 2 \\ K, & \text{if } m = j = 1 \text{ or } 1 + \frac{K}{2} \end{cases}$$

By identity (1), the second term of (2.31) is zero. The first term is also zero, except when $m = k$ or $m = K - k + 2$, leading to the results

$$E [A_k]^2 = \begin{cases} \frac{1}{2}G(\omega_k)\Delta\omega, & \text{if } k = 1 \\ \frac{1}{4}\{G(\omega_k) + G(\omega_{K-k+2})\}\Delta\omega, & \text{if } k = 2, \dots, \frac{K}{2} \\ G(\omega_k)\Delta\omega, & \text{if } k = 1 + \frac{K}{2} \end{cases} \quad (2.32)$$

remembering that for $k = 1$ the frequency interval is $\frac{1}{2}\Delta\omega$. An entirely similar calculation leads to

$$E [B_k]^2 = \begin{cases} 0, & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{1}{4}\{G(\omega_k) + G(\omega_{K-k+2})\}\Delta\omega, & \text{if } k = 2, \dots, \frac{K}{2} \end{cases} \quad (2.33)$$

The generation of coefficients using (2.32) and (2.33) and the symmetries

$$A_{K-k+2} = A_k, \quad k = 2, 3, \dots, \frac{K}{2} \quad (2.34)$$

$$B_{K-k+2} = -B_k, \quad k = 2, 3, \dots, \frac{K}{2}$$

gives considerably better point variance estimates when used in the FFT simulation of real processes than obtained using (2.28) and represent a new contribution by the author to the art of FFT simulation. In fact the relationships (2.28) are not valid when the symmetry relationships (2.34) are used in (2.17) to produce real processes.

If an analytical form of $G(\omega)$ is known then an even better approximation can be obtained by using $G^*(\omega_k)$ in (2.32) and (2.33) defined as

$$G^*(\omega_k) = \int_{\omega_k - \frac{1}{2}\Delta\omega}^{\omega_k + \frac{1}{2}\Delta\omega} G(\omega) d\omega, \quad (2.35)$$

which is just the area under the power spectral density function within the interval of length $\Delta\omega$ centered at ω_k . Although this improvement should yield better point variance estimates, it was not implemented in the code. The philosophy held by the author is that these algorithms should be as simple as possible for the user, requiring just the specification of the functional form of $G(\omega)$ or supplying discrete estimates of $G(\omega)$. The use of (2.35) would require either the user to evaluate the functional form of G^* for the desired $G(\omega)$ or the numerical integration of $G(\omega)$ at each frequency step. In higher dimensions, the latter could become ponderous. The following technique leads to an improved point variance estimate at a fraction of the overhead.

2.1.2 Covariance Formulation

Equation (2.29) suggests an entirely new approach to the determination of the statistics of the coefficients \mathcal{A}_k and \mathcal{B}_k . Instead of introducing the discrete approximation (2.30), the known covariance function can be used,

$$\begin{aligned}
\mathbb{E} [\mathcal{A}_k]^2 &= \frac{1}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K \mathbb{E} [Z_j Z_\ell] C_{kj} C_{k\ell} \\
&= \frac{1}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K B_{|j-\ell|} C_{kj} C_{k\ell} \\
&= \frac{\sigma^2}{K^2} \sum_{\ell=1}^K [C_{k\ell}]^2 + \frac{2}{K^2} \sum_{j=1}^{K-1} B_j \sum_{\ell=1}^{K-j} C_{k\ell} C_{k,\ell+j},
\end{aligned} \tag{2.36}$$

similarly,

$$\begin{aligned}
\mathbb{E} [\mathcal{B}_k]^2 &= \frac{1}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K \mathbb{E} [Z_j Z_\ell] S_{kj} S_{k\ell} \\
&= \frac{1}{K^2} \sum_{j=1}^K \sum_{\ell=1}^K B_{|j-\ell|} S_{kj} S_{k\ell} \\
&= \frac{\sigma^2}{K^2} \sum_{\ell=1}^K [S_{k\ell}]^2 + \frac{2}{K^2} \sum_{j=1}^{K-1} B_j \sum_{\ell=1}^{K-j} S_{k\ell} S_{k,\ell+j},
\end{aligned} \tag{2.37}$$

where the notation $B_j = B(j\Delta x)$ was employed. The additional trigonometric identities

$$3) \quad \sum_{\ell=1}^K C_{k\ell} C_{k\ell} = \begin{cases} K & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{K}{2} & \text{if } k = 2, 3, \dots, \frac{K}{2} \end{cases}$$

$$\begin{aligned}
4) \quad \sum_{\ell=1}^K S_{k\ell} S_{k\ell} &= \begin{cases} 0 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{K}{2} & \text{if } k = 2, 3, \dots, \frac{K}{2} \end{cases} \\
5) \quad \sum_{\ell=1}^{K-j} C_{k\ell} C_{k,\ell+j} &= \begin{cases} K-j & \text{if } k = 1 \\ \frac{1}{2}\{(K-j-1)C_{k,j+1} - S_{kj}/S_{k1}\} & \text{if } k = 2, 3, \dots, \frac{K}{2} \\ (-1)^j(K-j) & \text{if } k = 1 + \frac{K}{2} \end{cases} \\
6) \quad \sum_{\ell=1}^{K-j} S_{k\ell} S_{k,\ell+j} &= \begin{cases} 0 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{1}{2}\{(K-j-1)C_{k,j+1} - S_{k,j+2}/S_{k1}\} & \text{if } k = 2, 3, \dots, \frac{K}{2} \end{cases}
\end{aligned}$$

allows the reduction of (2.36) and (2.37) to

$$\mathbb{E} [A_k]^2 = \begin{cases} \frac{\sigma^2}{K} + \frac{2}{K^2} \sum_{j=1}^{K-1} (K-j) B_j & \text{if } k = 1 \\ \frac{\sigma^2}{2K} + \frac{1}{K^2} \sum_{j=1}^{K-1} B_j [(K-j-1)C_{k,j+1} - S_{kj}/S_{k1}] & \text{if } k = 2, 3, \dots, \frac{K}{2} \\ \frac{\sigma^2}{K} + \frac{2}{K^2} \sum_{j=1}^{K-1} (-1)^j (K-j) B_j & \text{if } k = 1 + \frac{K}{2} \end{cases} \quad (2.38)$$

$$\mathbb{E} [B_k]^2 = \begin{cases} 0 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{\sigma^2}{2K} + \frac{1}{K^2} \sum_{j=1}^{K-1} B_j [(K-j-1)C_{k,j+1} - S_{k,j+2}/S_{k1}] & \text{if } k = 2, 3, \dots, \frac{K}{2} \end{cases} \quad (2.39)$$

Equations (2.38) and (2.39) can be evaluated using a pair of inverse FFT's making the calculation of the variances very efficient. Letting

$$\begin{aligned}
X_k &= \frac{1}{K^2} \sum_{j=1}^K \delta_j B_{j-1} (K-j+1) \cos\left(\frac{2\pi(k-1)(j-1)}{K}\right), \\
Y_k &= \frac{1}{K^2} \sum_{j=1}^K \delta_j B_{j-1} \sin\left(\frac{2\pi(k-1)(j-1)}{K}\right),
\end{aligned} \quad (2.40)$$

where

$$\delta_j = \begin{cases} \frac{1}{2} & \text{if } j = 1 \\ 1 & \text{otherwise} \end{cases}$$

then the variances of the Fourier coefficients can be expressed as

$$\mathbb{E} [A_k]^2 = E_k X_k - F_k Y_k, \quad (2.41)$$

$$\mathbb{E} [B_k]^2 = G_k X_k + F_k Y_k,$$

where

$$E_k = \begin{cases} 2 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$G_k = \begin{cases} 0 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$E_k = \begin{cases} 0 & \text{if } k = 1 \text{ or } 1 + \frac{K}{2} \\ \frac{\cos(2\pi(k-1)/K)}{\sin(2\pi(k-1)/K)} & \text{otherwise} \end{cases}$$

The process shown previously in Figure 2.1 was generated using this procedure which guarantees that the estimated point variance will tend to the desired point variance no matter what frequency interval is used.

It should be noted at this point that in order to eliminate the symmetry in the estimated covariance function of the simulated process, all the joint moments $E[\mathcal{A}_k \mathcal{A}_m]$, $E[\mathcal{A}_k \mathcal{B}_m]$ and $E[\mathcal{B}_k \mathcal{B}_m]$ which are *not* zero when using (2.17) (as can be easily verified using (2.20) in a manner similar to the above) must be included. This, however, is a formidable task involving the prior simulation of the correlated (and cross-correlated) processes \mathcal{A}_k and \mathcal{B}_k . It will be assumed that these joint moments are zero and a correction for the symmetry in the estimated covariance function will be made by adjusting the size of the process using (2.25) and ignoring the excess.

2.2 Multi-Dimensional Extensions

Following the results of Section 2.1, the discrete Fourier transform of the 2-D discrete real process $\mathcal{Z}_{ij} \simeq Z((i-1)\Delta x_1, (j-1)\Delta x_2)$ is

$$\mathcal{Z}_{ij} = \sum_{m=1}^{K_1} \sum_{n=1}^{K_2} \left\{ \mathcal{A}_{mn} \cos \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right) + \mathcal{B}_{mn} \sin \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right) \right\} \quad (2.42)$$

in which the Fourier coefficients \mathcal{A}_{mn} and \mathcal{B}_{mn} are assumed to be independent zero mean normally distributed random variates. Equation (2.42) will be evaluated using a two-dimensional FFT algorithm. On scalar machines, the 2-D FFT was coded using a radix-2x2 butterfly algorithm following the concepts discussed by Dudgeon and Mersereau [18]. This algorithm was found to run in about 80% of the time required by the usual row-column decomposition algorithm. In three dimensions, a radix-2x2x2 algorithm was coded that runs in about 50% of the time required for the 3-D row-column code. Both of these algorithms require that the number of field points in each direction be an integer power of 2. On the CYBER 205 supercomputer, the vectorized multi-dimensional FFT algorithms provided by

the MAGEVFFT library were used. Other implementation issues of the multi-dimensional FFT are discussed in [9, 31, 59].

Higher dimensional FFT's involve additional summations and associated terms in the trigonometric functions of (2.42). For illustrative purposes, the two-dimensional process will be concentrated on here and the corresponding results for the three-dimensional case just stated. The inverse transforms corresponding to (2.42) are

$$\mathcal{A}_{mn} = \frac{1}{K_1 K_2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_1} \mathcal{Z}_{ij} \cos \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right), \quad (2.43)$$

$$\mathcal{B}_{mn} = \frac{1}{K_1 K_2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_1} \mathcal{Z}_{ij} \sin \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right), \quad (2.44)$$

which imply that the following planar symmetries must apply for real processes

$$\mathcal{A}_{K_1-m+2, K_2-n+2} = \mathcal{A}_{mn}, \quad (2.45)$$

$$\mathcal{A}_{K_1-m+2, n} = \mathcal{A}_{m, K_2-n+2},$$

for $m, n = 2, 3, \dots, \frac{K_\alpha}{2}$ where K_α is either K_1 or K_2 appropriately. In addition the line symmetries

$$\mathcal{A}_{\ell, K_2-n+2} = \mathcal{A}_{\ell, n}, \quad (2.46)$$

$$\mathcal{A}_{K_1-m+2, \ell} = \mathcal{A}_{m, \ell},$$

apply over the same range in m, n , and for $\ell = 1$ or $\ell = 1 + \frac{K_\alpha}{2}$. Similar relationships exist for the coefficients \mathcal{B}_{mn} except that they are anti-symmetries, i.e. $\mathcal{B}_{K_1-m+2, K_2-n+2} = -\mathcal{B}_{mn}$ and so on. The planar and line symmetries are shown graphically in Figure 2.3. Thus in two dimensions, the Fourier coefficients must be generated over the two darker shaded quadrants of Figure 2.3 and the 4 half-lines along $\ell = 1$ and $\ell = 1 + \frac{K_\alpha}{2}$, the rest of the coefficients obtained using the symmetry relations. If the process is quadrant symmetric then the additional symmetry relation

$$\mathcal{A}_{K_1-m+2, n} = \mathcal{A}_{mn} \quad (2.47)$$

applies and so only one quadrant of the \mathcal{A}_{mn} and \mathcal{B}_{mn} coefficients need be specified.

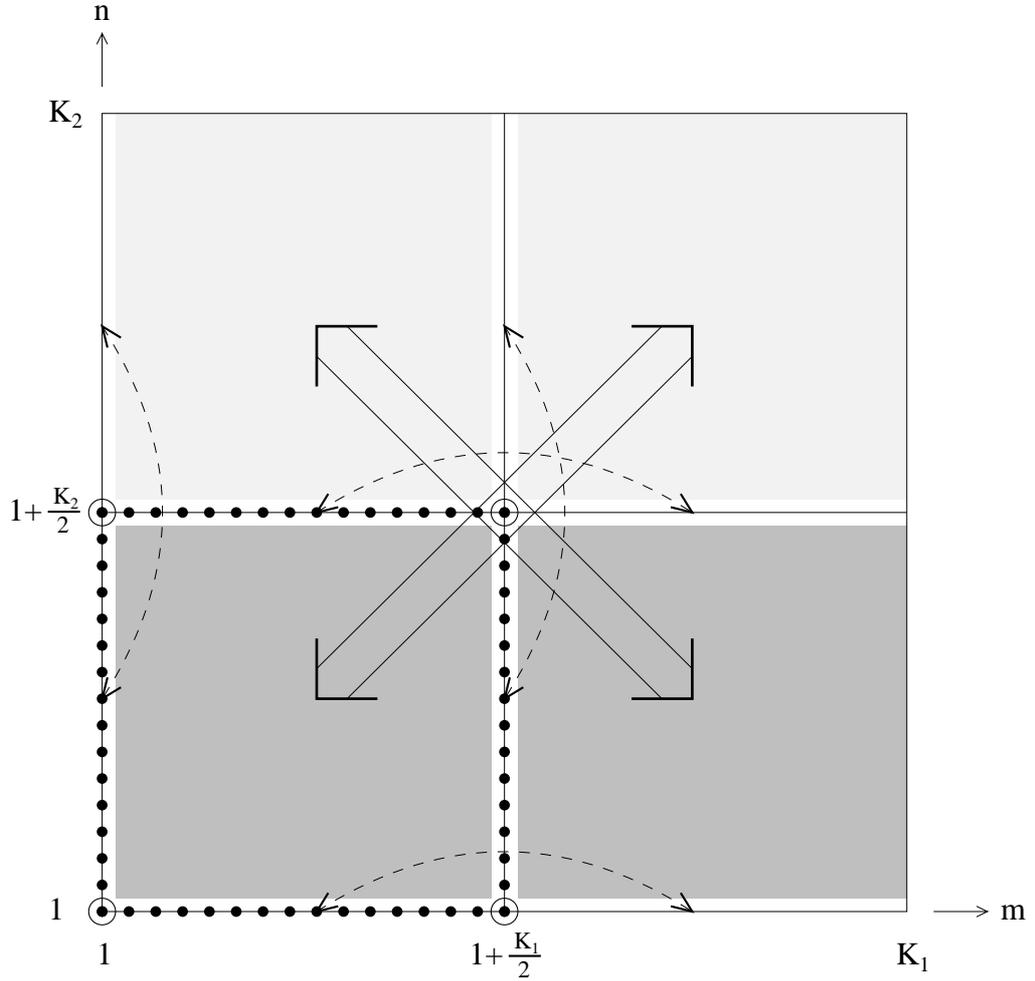


Figure 2.3 Fourier coefficient symmetry over the plane of their indices. Dashed arrows denote symmetric half-lines and double arrows denote symmetric planes.

The three-dimensional symmetries are similar except that the coefficients must be specified over 4 quadrant volumes, 6 half-planes corresponding to the indexes $\ell = 1$ and $\ell = 1 + \frac{K_\alpha}{2}$ ($\alpha = 1, 2, 3$), and the 12 half-lines corresponding to the edges of the $[1, 1 + \frac{K_1}{2}] \times [1, 1 + \frac{K_2}{2}] \times [1, 1 + \frac{K_3}{2}]$ cube. Again, if the process is quadrant symmetric additional symmetries in the coefficients apply and only one quadrant volume need be specified.

The variances of the Fourier coefficients can be obtained from the inverse transforms (2.43) and (2.44) as

$$E [A_{mn}]^2 = \frac{1}{(K_1 K_2)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} E [Z_{ij} Z_{k\ell}] C_{im,jn} C_{km,\ell n}, \quad (2.48)$$

$$\mathbb{E} [\mathcal{B}_{mn}]^2 = \frac{1}{(K_1 K_2)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \mathbb{E} [Z_{ij} Z_{k\ell}] S_{im,jn} S_{km,\ell n}. \quad (2.49)$$

where,

$$C_{im,jn} = \cos \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right),$$

$$S_{im,jn} = \sin \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right).$$

Assuming the covariance of the process, $B_{jk} = B(j\Delta x_1, k\Delta x_2)$, to be quadrant symmetric which in turn implies that the spectral density function, $S(\omega_1, \omega_2)$, is also quadrant symmetric, the discrete approximation to (1.20) becomes

$$\mathbb{E} [Z_{ij} Z_{k\ell}] \simeq \Delta\omega_1 \Delta\omega_2 \sum_{m=1}^{K_1} \sum_{n=1}^{K_2} G(\omega_{1m}, \omega_{2n}) \cos \left(\frac{2\pi(m-1)(i-k)}{K_1} \right) \cos \left(\frac{2\pi(n-1)(j-\ell)}{K_2} \right), \quad (2.50)$$

where $G(\omega) = 2^n S(\omega)$, for a process in R^n , is the uni-quadrant spectral density function. When (2.50) is substituted into (2.48) and (2.49) the following (spectral formulation) variances of the Fourier coefficients are obtained

$$\mathbb{E} [\mathcal{A}_{mn}]^2 = \begin{cases} \Delta\omega_1 \Delta\omega_2 G^d(\omega_m, \omega_n), & m = 1, 1 + \frac{K_1}{2} \text{ and } n = 1, 1 + \frac{K_2}{2} \\ \frac{1}{4} \Delta\omega_1 \Delta\omega_2 (G^d(\omega_m, \omega_n) + G^d(\omega_m, \omega_{K_2-n+2})), & m = 1, 1 + \frac{K_1}{2} \\ \frac{1}{4} \Delta\omega_1 \Delta\omega_2 (G^d(\omega_m, \omega_n) + G^d(\omega_{K_2-m+2}, \omega_n)), & n = 1, 1 + \frac{K_2}{2} \\ \frac{1}{8} \Delta\omega_1 \Delta\omega_2 (G^d(\omega_m, \omega_n) + G^d(\omega_m, \omega_{K_2-n+2}) \\ \quad + G^d(\omega_{K_1-m+2}, \omega_n) + G^d(\omega_{K_1-m+2}, \omega_{K_2-n+2})), & \text{otherwise.} \end{cases} \quad (2.51)$$

The variances of B_{mn} are identical except at the four corner points where $B_{mn} = 0$, $m, n = 1, 1 + \frac{K_\alpha}{2}$. The superscript d denotes the correction made on $G(\omega)$ to account for the discrete nature of the spectral density function: In n -dimensional space, set

$$G^d(\omega) = \frac{G(\omega)}{2^d} \quad (2.52)$$

where d is the number of components of $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ which are equal to zero. (In the one-dimensional case, this was accomplished by using half the frequency interval at $\omega_1 = 0$.)

Using the actual covariance function in (2.48) and (2.49) leads to the variances of the Fourier coefficients in the so-called covariance formulation. These results, along with the spectral formulation results for three-dimensional processes, are shown in Appendix A.

As in the one-dimensional case, the problems associated with the multi-dimensional FFT simulation methods have to do with discretization of the frequency space and symmetry of the estimated covariance function. The former usually leads to poor estimation of the

point variance of the process and can be solved either by increasing the size of the field or by employing the covariance formulation. The latter problem can only be solved by increasing the field size in each component direction using (2.25) and the directional scales of fluctuation (see Vanmarcke [25] for definitions of directional scales of fluctuation). Figures 2.4 and 2.5 represent sample functions of the 2-D process generated using the spectral formulation and covariance formulation respectively for a process having the target covariance function of the continuous Markov process

$$B(\tau_1, \tau_2) = \sigma^2 \exp \left\{ -\frac{2}{\theta} \sqrt{\tau_1^2 + \tau_2^2} \right\}, \quad (2.53)$$

with associated spectral density function

$$G(\omega_1, \omega_2) = \frac{4\theta^2 \sigma^2}{\pi \left(4 + \theta^2(\omega_1^2 + \omega_2^2) \right)^{\frac{3}{2}}}. \quad (2.54)$$

In these processes, the directional scales of fluctuation $\theta_1 = \theta_2 = \theta = 4$ are quite large compared to the 5×5 physical size of the field shown in Figures 2.4 and 2.5. Figures 2.6 and 2.7 show the estimated and exact covariance functions for the spectral and covariance formulations respectively. Notice that although the physical length of each side of the field is only 5, the horizontal scales on Figures 2.6 and 2.7 are long enough to cover the length of the diagonal since the covariance is also estimated in that direction. It can be seen that the spectral formulation (without the correction given by 2.35) significantly over-estimates the point variance of the process. Although the covariance formulation correctly captures the point variance, its sample function has a somewhat striated appearance similar to what might be obtained if aliasing were present. The use of symmetries in the Fourier coefficients eliminates direct aliasing but it is possible that the Fourier transform of the covariance function (see 2.40) is introducing some aliasing of the covariance structure.

When shorter scales of fluctuation are used, both the striated appearance of the covariance formulation and the poorly estimated variance of the spectral formulation tend to disappear and so, under these conditions, the spectral formulation is preferred for its simplicity. Figures 2.8 and 2.9 show sample functions for the spectral and covariance formulations respectively using scales of fluctuation $\theta_1 = \theta_2 = \theta = \frac{1}{2}$. In this case Figures 2.10 and 2.11 show the correct point variance in the estimated statistics and Figure 2.9 has lost much of its striated appearance.

A comment should be made at this point about the rendering of these realizations which consist of 256×256 discrete points. In all cases, a grey scale mapping is used in which the minimum value of the realization is mapped to white and the maximum to black. No

direct smoothing is performed and the printer is actually displaying a 256×256 grid of squares each having an associated grey level. Color renderings were experimented with, but it was found that extreme care had to be exercised in the selection of palettes and brightness levels (for each realization) in order to pick out details such as striations and streaks. Grey scale renderings seem to give these details with little trouble. Printing and reproduction of these realizations is also sometimes troublesome. Variations in toner and drum quality can give apparent artifacts which may be misleading – particularly to the reader of a third-hand copy! Nevertheless, the visual appraisal of these grey-scale renderings was considered by the author as being an essential means of establishing the quality of an algorithm, in terms of individual realizations, that was superior to color renderings or contour plots.

Figure 2.12 show the estimated statistics for a three-dimensional process of physical size $5 \times 5 \times 5$ using the target covariance function

$$B(\tau_1, \tau_2, \tau_3) = \sigma^2 \exp \left\{ -\frac{2}{\theta} \sqrt{\tau_1^2 + \tau_2^2 + \tau_3^2} \right\}, \quad (2.55)$$

with $\theta = 4$ using the spectral formulation. The dashed lines are the estimated covariances along various directions within the $5 \times 5 \times 5$ cube (i.e. vertical, horizontal, depth, plane diagonals and body diagonal). It can be seen that the point variance is significantly over-estimated by the spectral formulation due to the rather large frequency increment. Figure 2.13 shows the estimated statistics of the same process (2.55) using the covariance formulation. A much better estimate of the point variance is obtained using this method. When shorter scales of fluctuation are used in the 3-D case, both algorithms yield good estimates of the covariance structure as illustrated in Figures 2.14 and 2.15, using the spectral and covariance formulations respectively, for $\theta = \frac{1}{2}$. This suggests that, in general, it is best to use a sufficiently large field (compared to the directional scales of fluctuation) when simulating via the FFT algorithm.

Although the examples discussed above are isotropic processes, the programs written to perform the simulations are only restricted to quadrant symmetric processes. As an illustration of an anisotropic process, Figure 2.16 shows a sample function of a 2-D process with the target covariance function

$$B(\tau_1, \tau_2) = \sigma^2 \exp \left\{ -\sqrt{\left(\frac{2\tau_1}{\theta_1}\right)^2 + \left(\frac{2\tau_2}{\theta_2}\right)^2} \right\}, \quad (2.56)$$

with $\theta_1 = 0.8$, $\theta_2 = 0.2$. Figure 2.17 shows the equivalent process simulated using the covariance formulation and Figures 2.18 and 2.19 show their estimated statistics averaged over 100 realizations. It can be seen that anisotropy is handled adequately by the FFT algorithms.

2.3 Summary

To obtain reasonably good realizations using an FFT algorithm, it is suggested that the following simple guidelines be followed;

- 1) ensure that the size of the field to be simulated is larger than that required and that the excess is ignored. How much larger depends on the scale of fluctuation of the process but could be as much as twice the size in any direction.
- 2) ensure that the spectral density above the Nyquist limit $\pi/\Delta x$ is negligible. This represents a limitation on the resolution of the field.
- 3) ensure that $\Delta\omega = 2\pi/T$ is small enough to adequately capture rapid changes in the spectral density function. Using the covariance formulation discussed above or explicitly determining the area under the spectral density function at each frequency step are possible alternatives to this criteria.

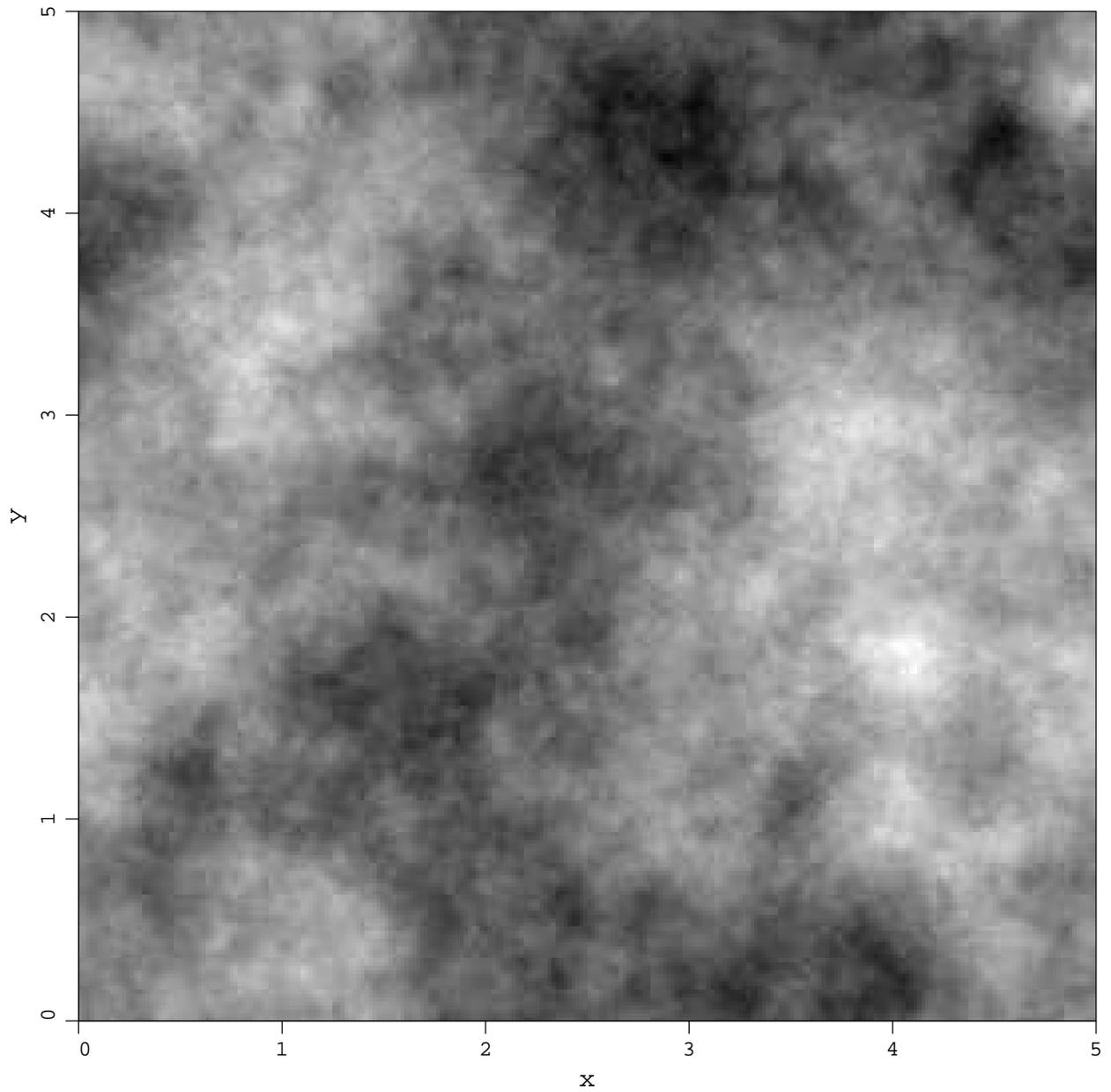


Figure 2.4 Sample function of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = 4$. For such a large scale of fluctuation, the process varies relatively smoothly and one gets larger dark and light regions.

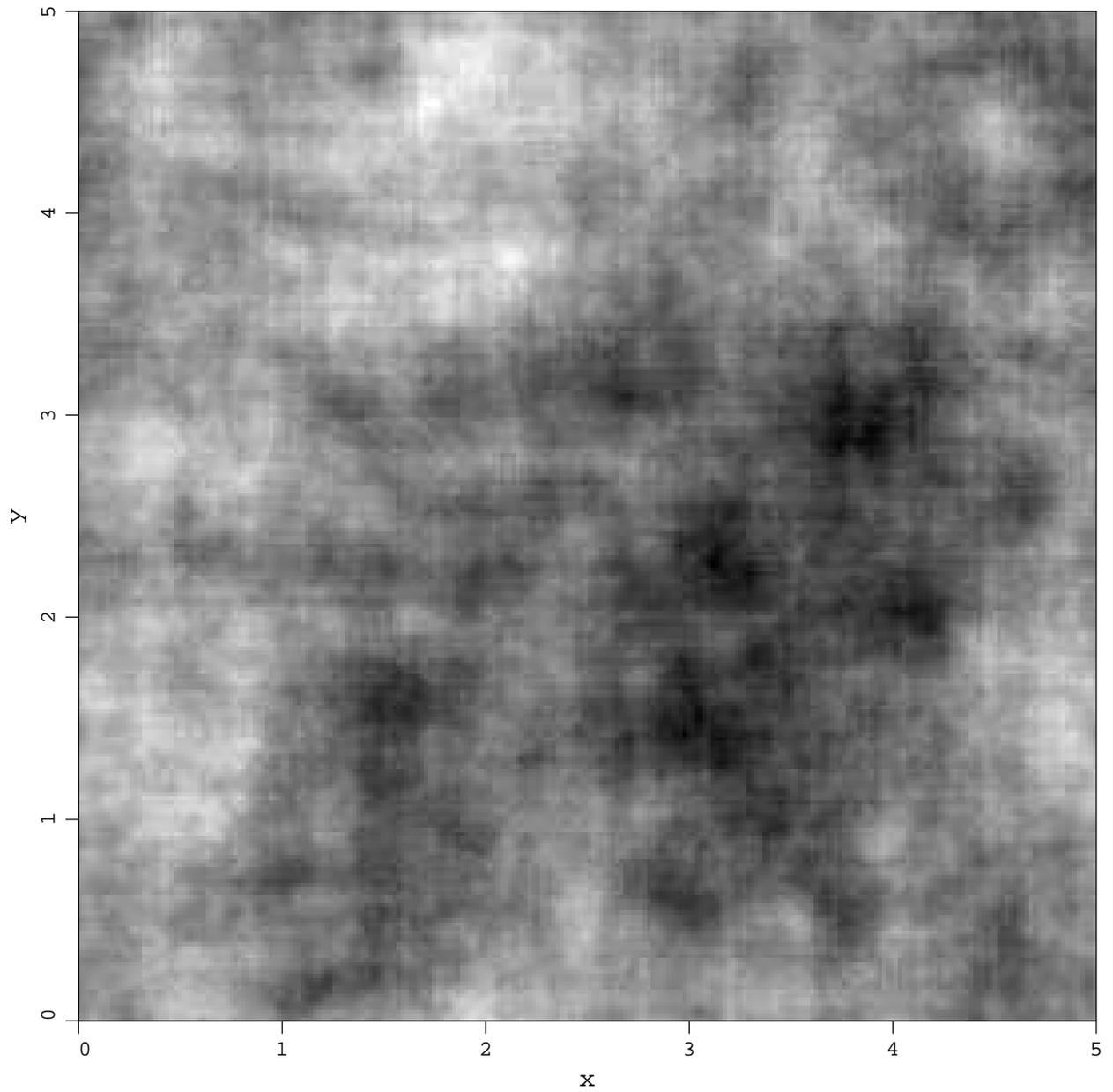


Figure 2.5 Sample function of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = 4$.

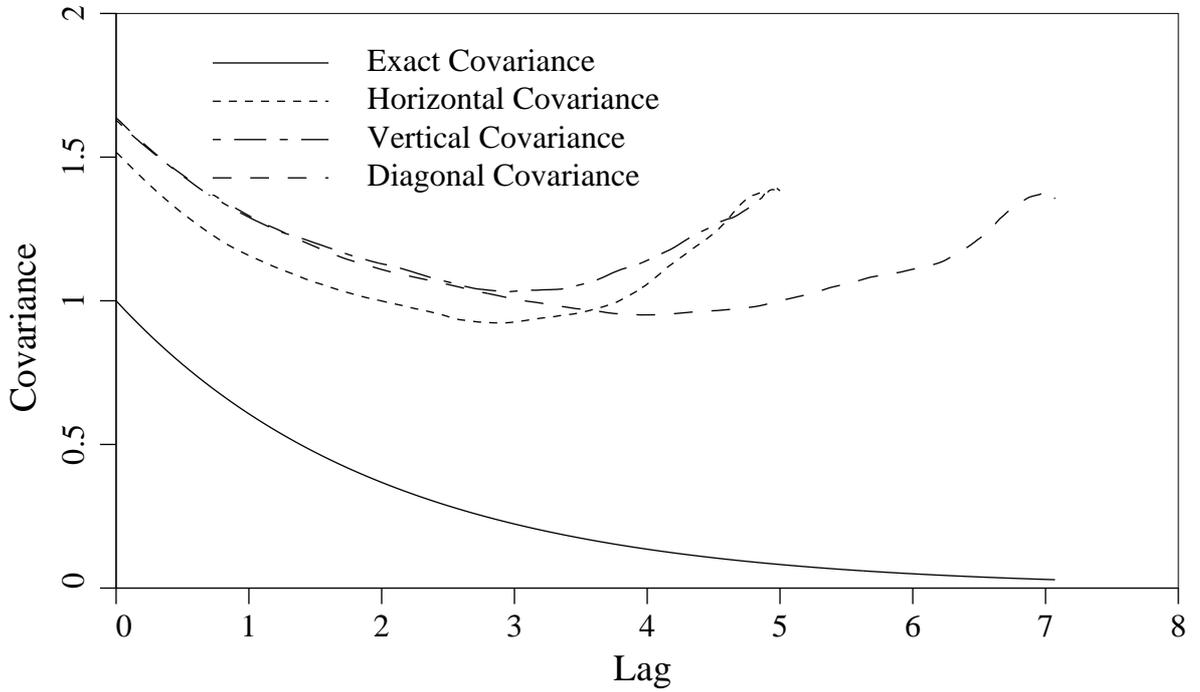


Figure 2.6 Estimated statistics of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = 4$. Statistics are averaged over 100 fields.

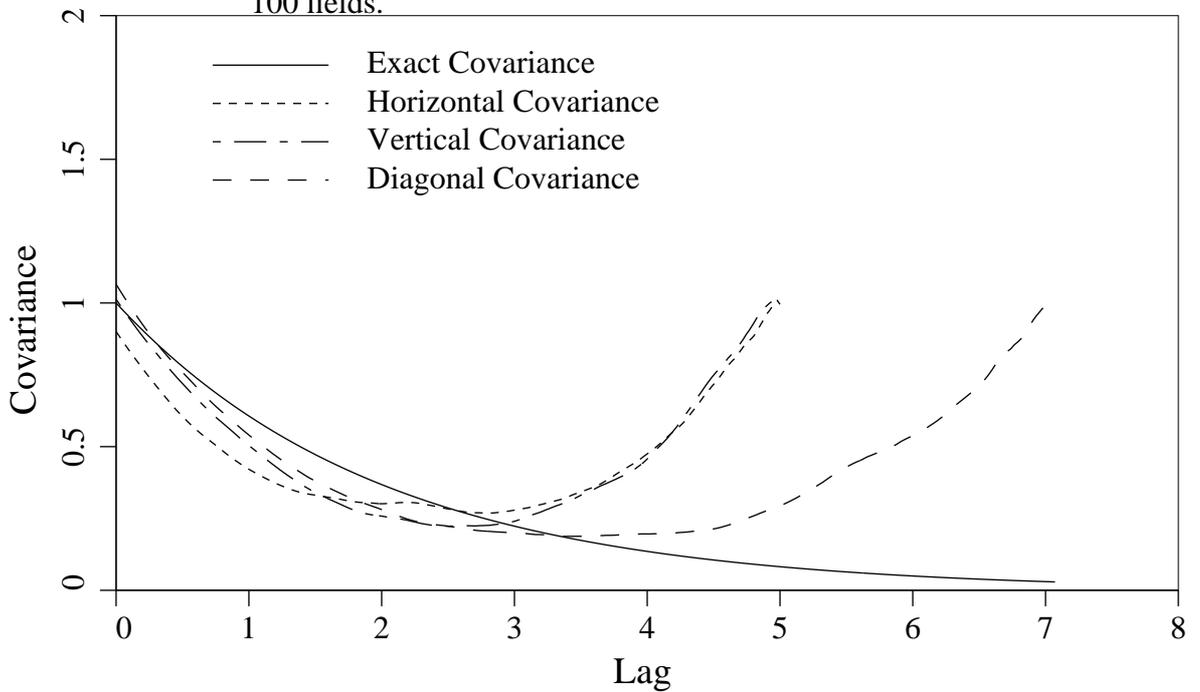


Figure 2.7 Estimated statistics of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = 4$. Statistics are averaged over 100 fields.

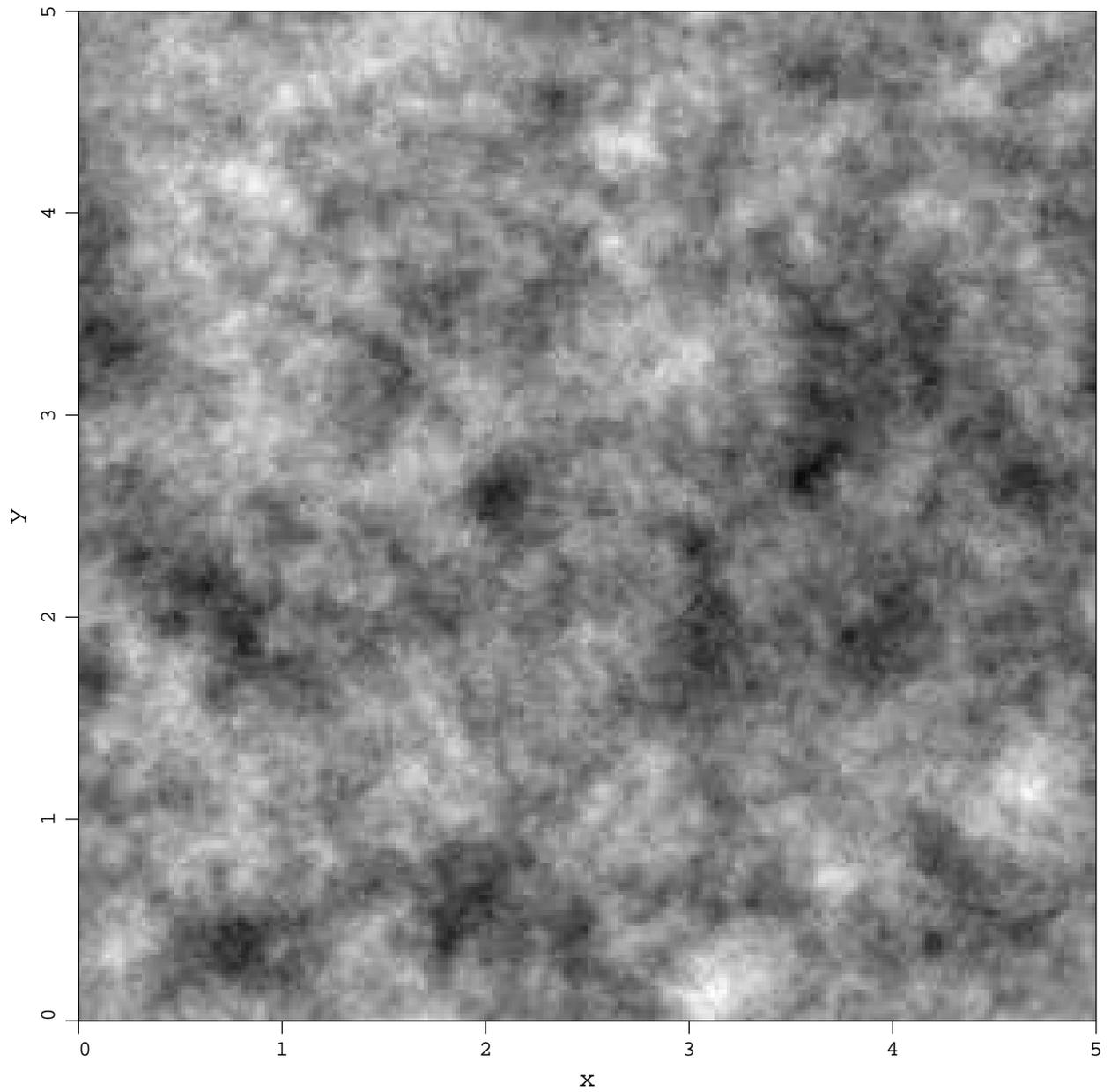


Figure 2.8 Sample function of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Notice the increased ‘roughness’ corresponding to a shorter scale of fluctuation.

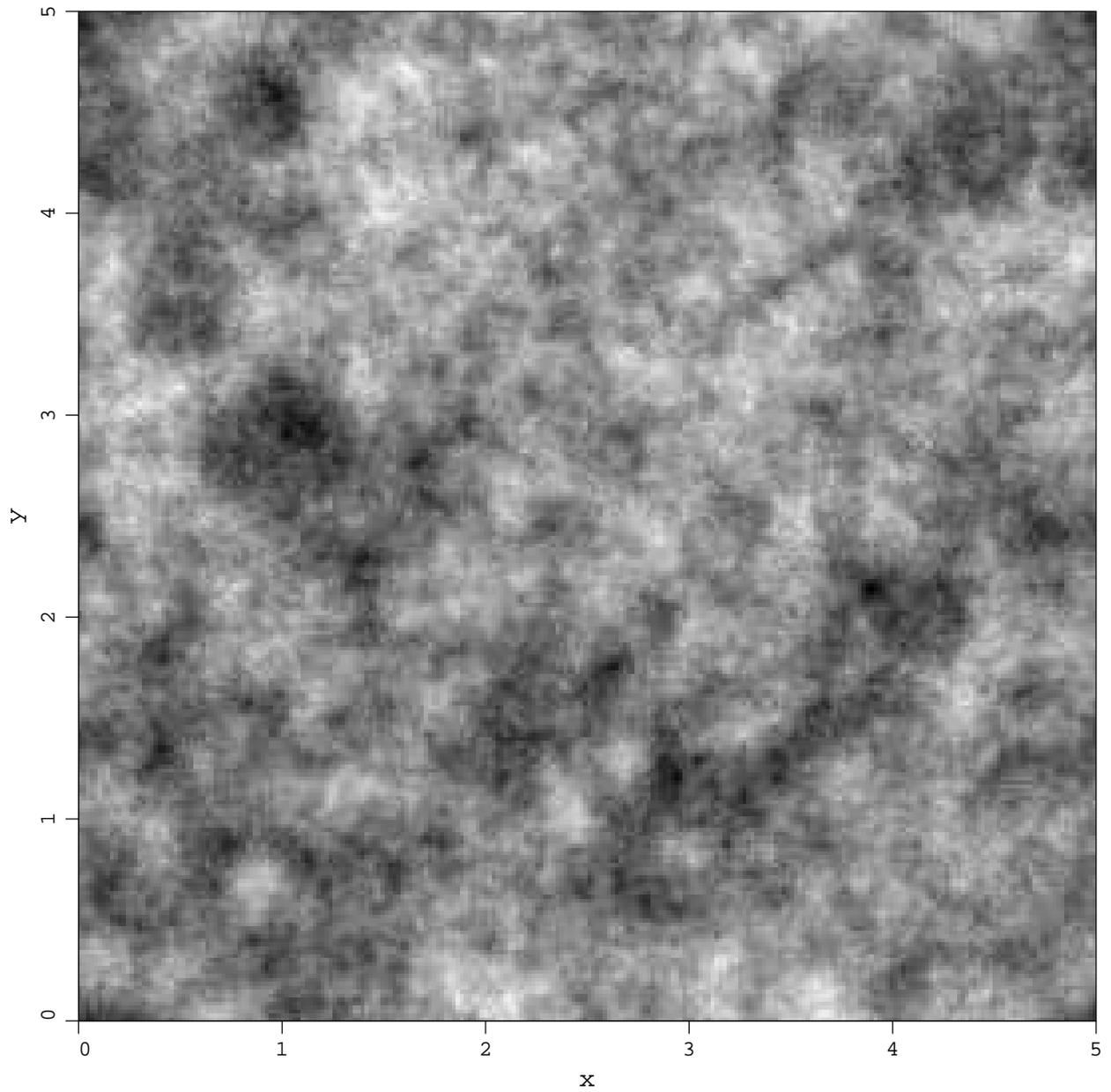


Figure 2.9 Sample function of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$.

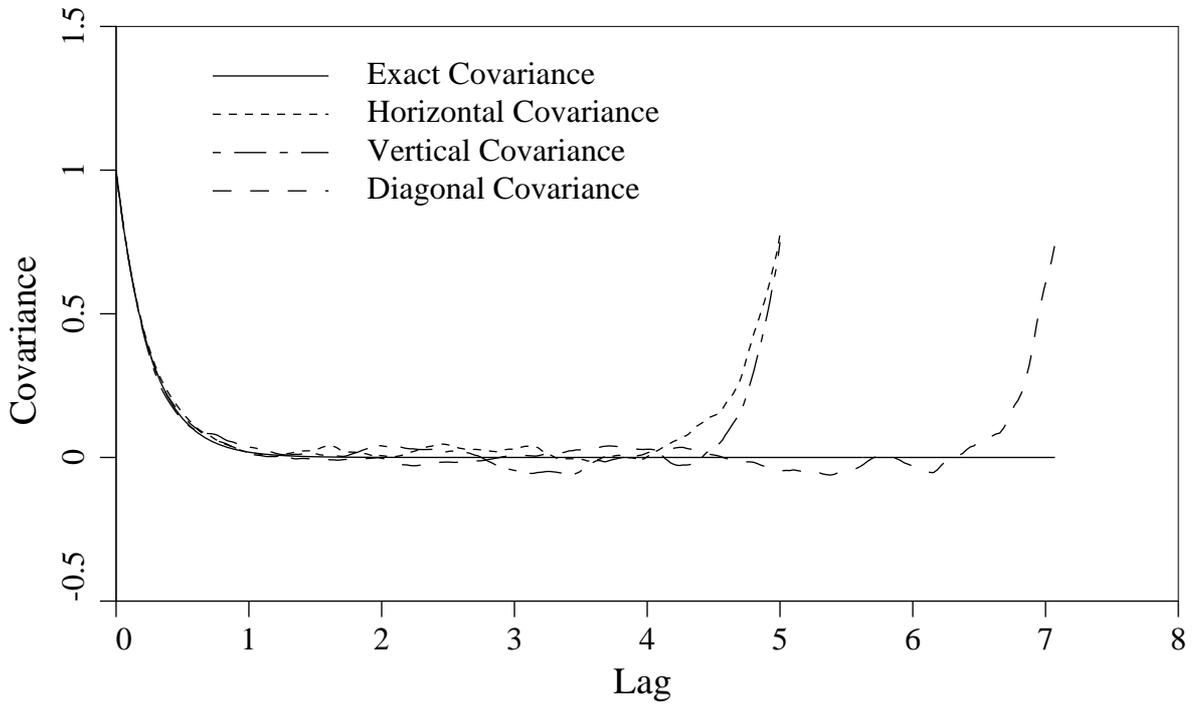


Figure 2.10 Estimated statistics of a 2-D field (spectral formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Statistics are averaged over 100 fields.

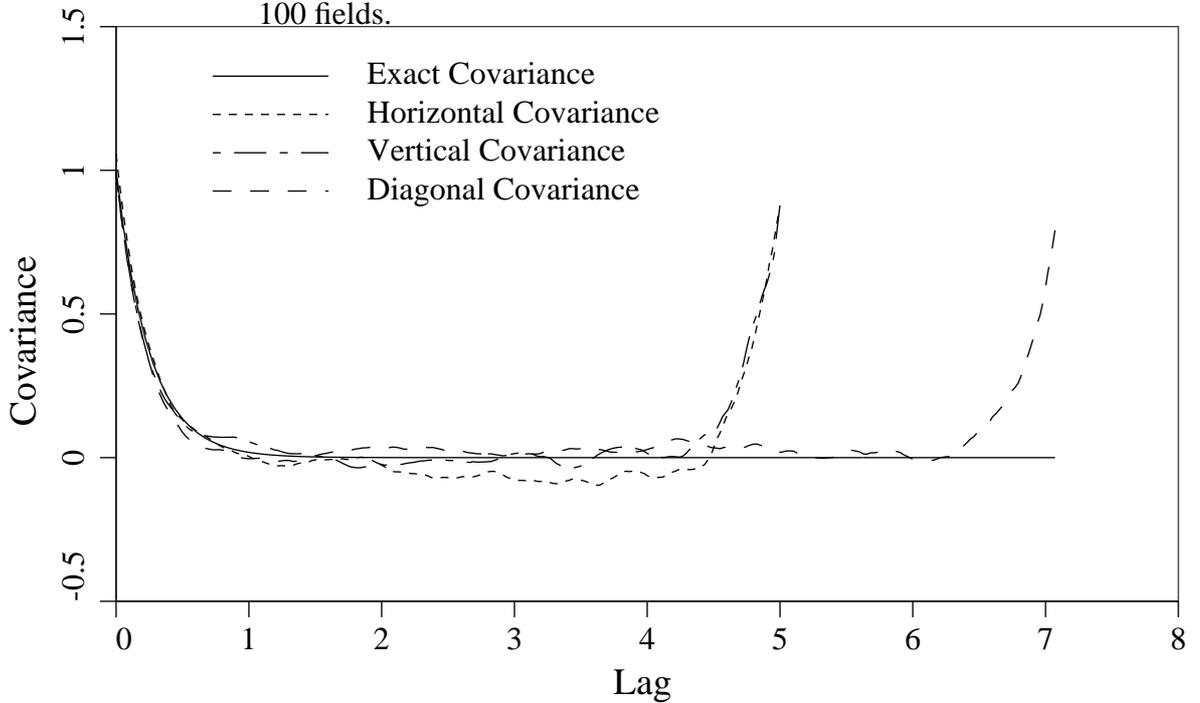


Figure 2.11 Estimated statistics of a 2-D field (covariance formulation) with covariance function given by (2.53) for $\theta = \frac{1}{2}$. Statistics are averaged over 100 fields.

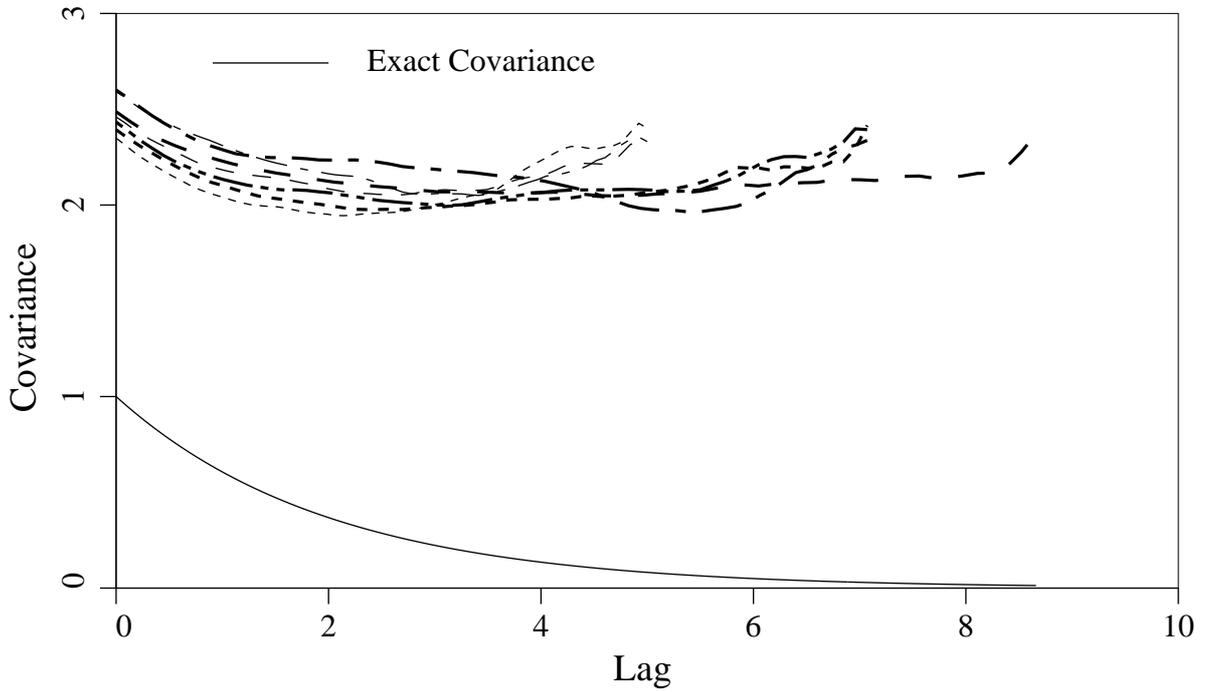


Figure 2.12 Estimated statistics of a 3-D field (spectral formulation) with covariance function given by (2.55) for $\theta = 4$. Statistics are averaged over 50 fields.

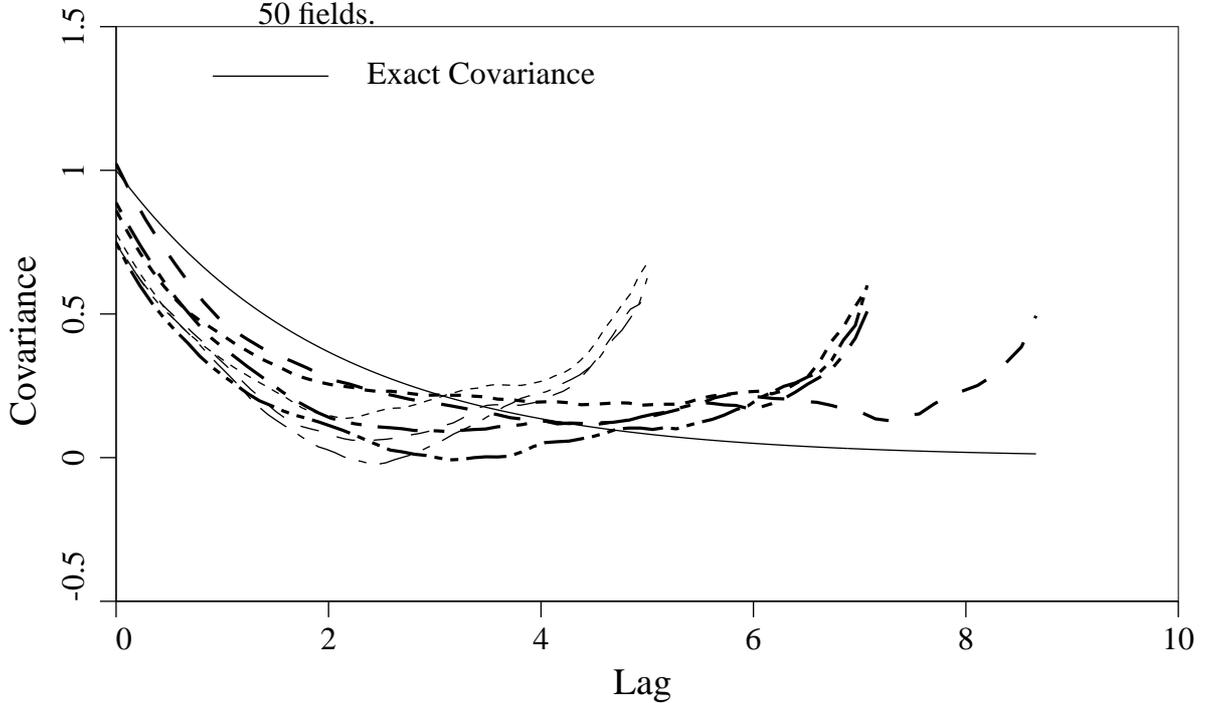


Figure 2.13 Estimated statistics of a 3-D field (covariance formulation) with covariance function given by (2.55) for $\theta = 4$. Statistics are averaged over 50 fields.

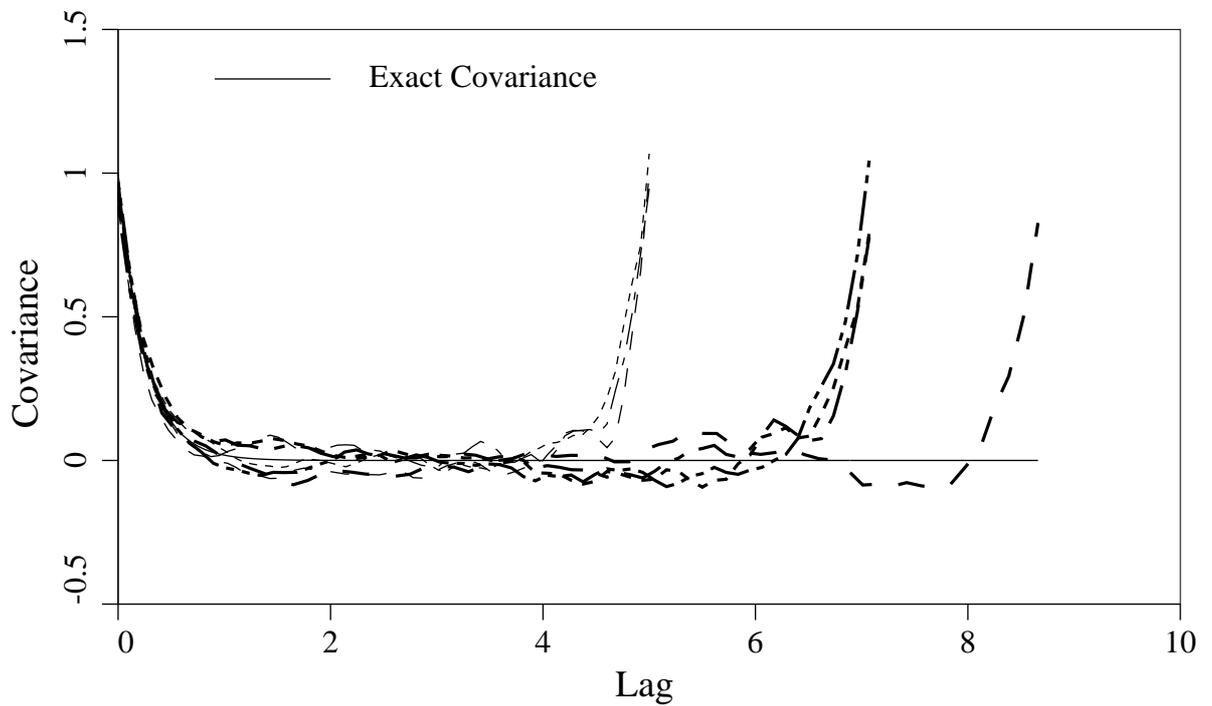


Figure 2.14 Estimated statistics of a 3-D field (spectral formulation) with covariance function given by (2.55) for $\theta = \frac{1}{2}$. Statistics are averaged over 50 fields.

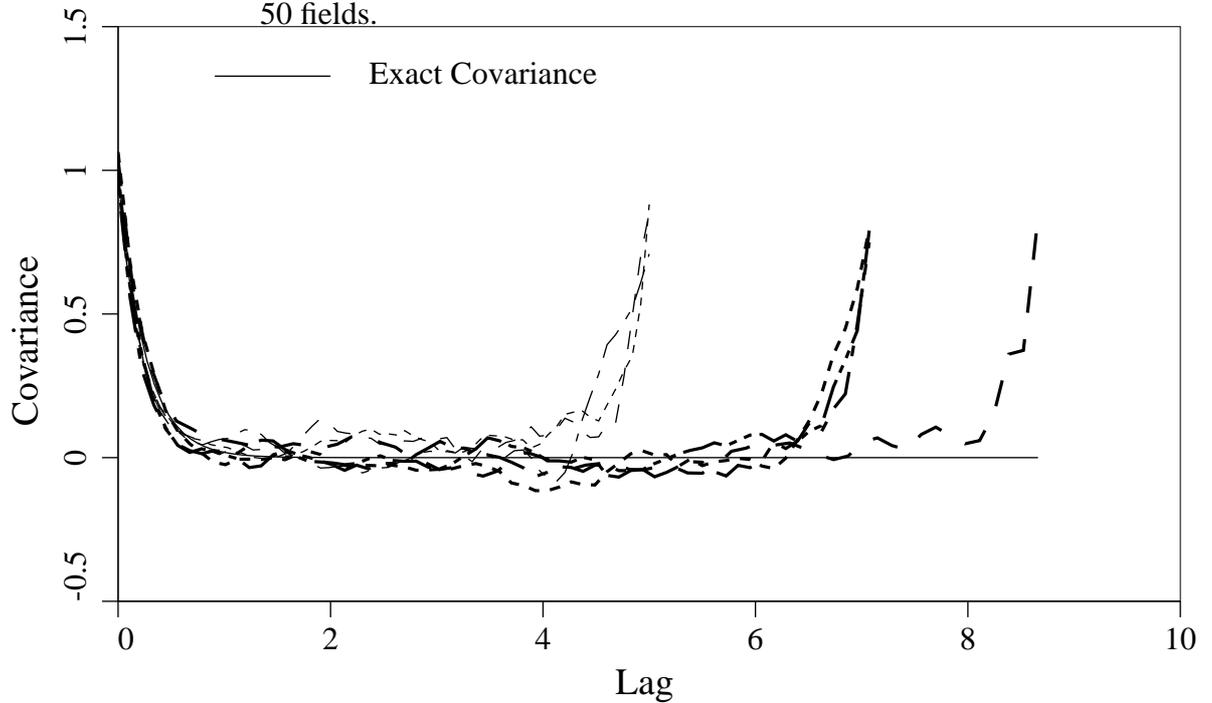


Figure 2.15 Estimated statistics of a 3-D field (covariance formulation) with covariance function given by (2.55) for $\theta = \frac{1}{2}$. Statistics are averaged over 50 fields.

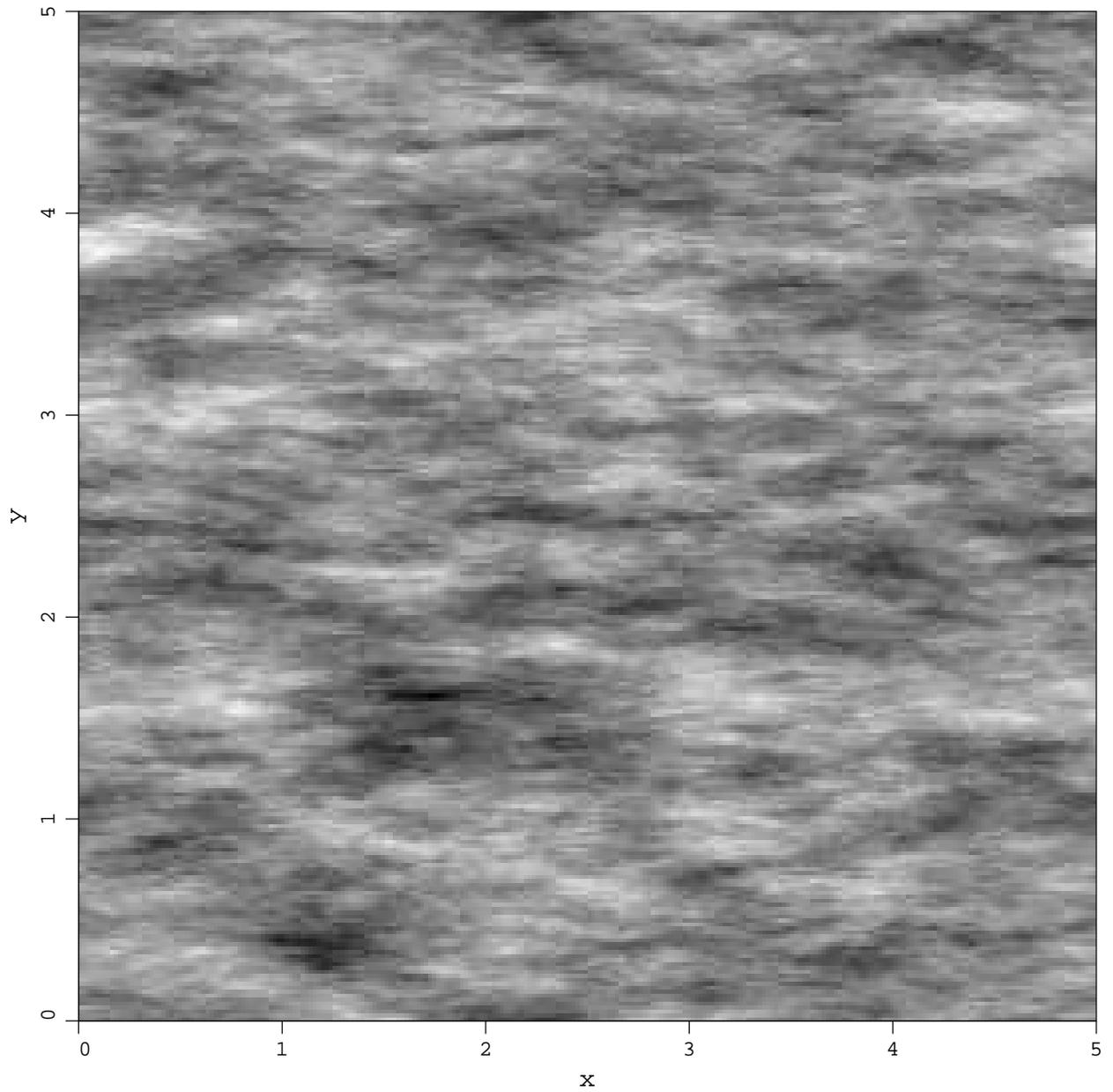


Figure 2.16 Sample function of an anisotropic 2-D field (spectral formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$.

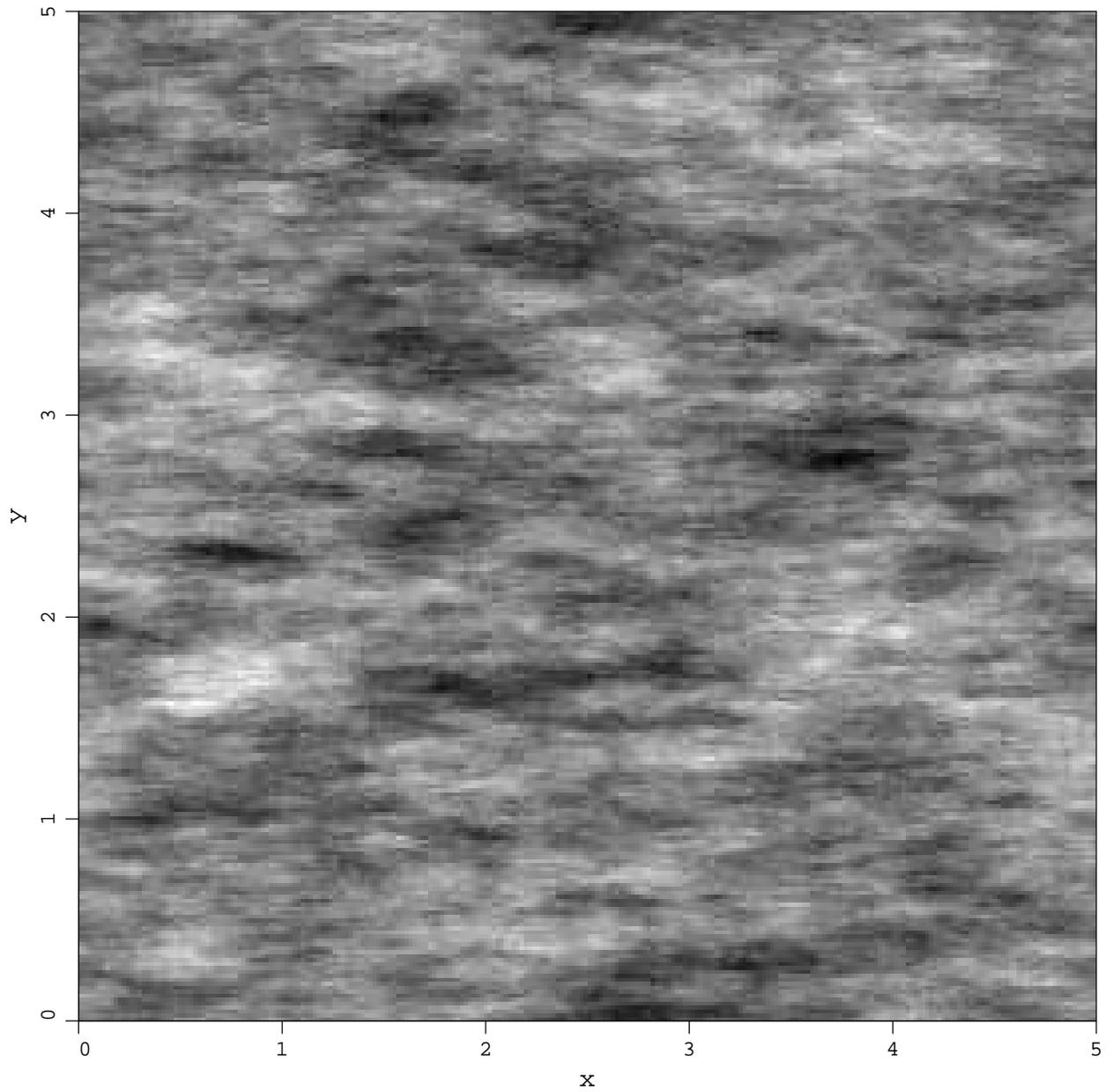


Figure 2.17 Sample function of an anisotropic 2-D field (covariance formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$.

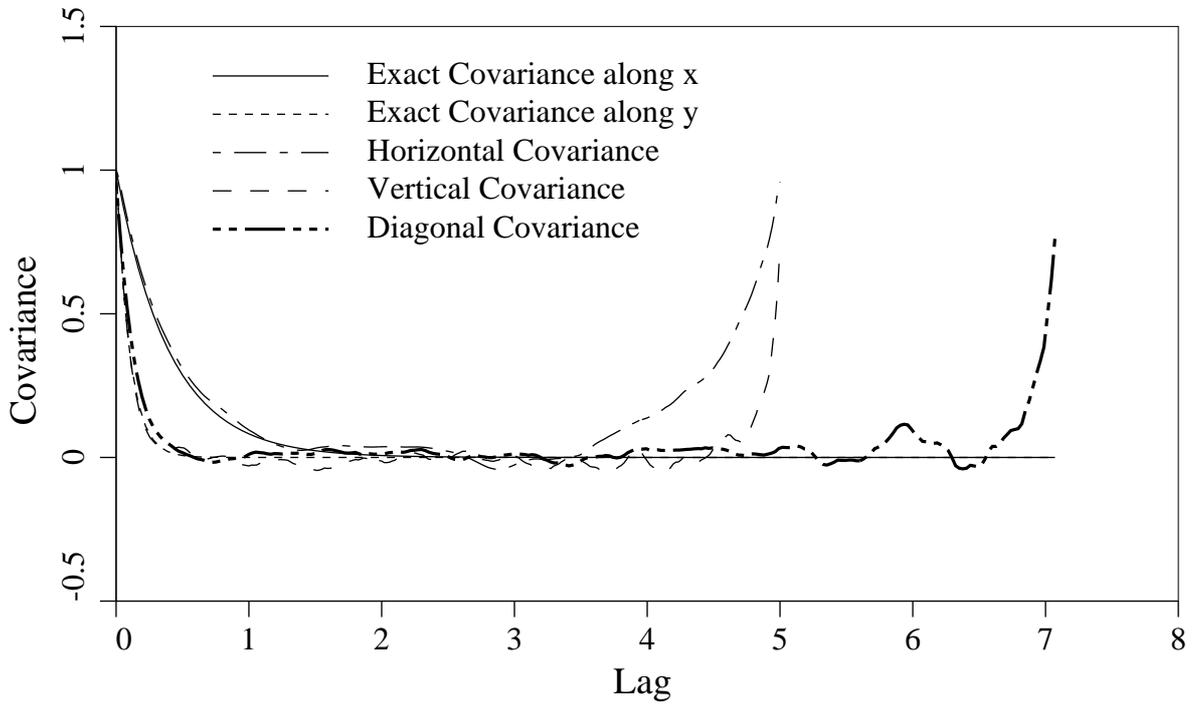


Figure 2.18 Estimated statistics of an anisotropic 2-D field (spectral formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$.

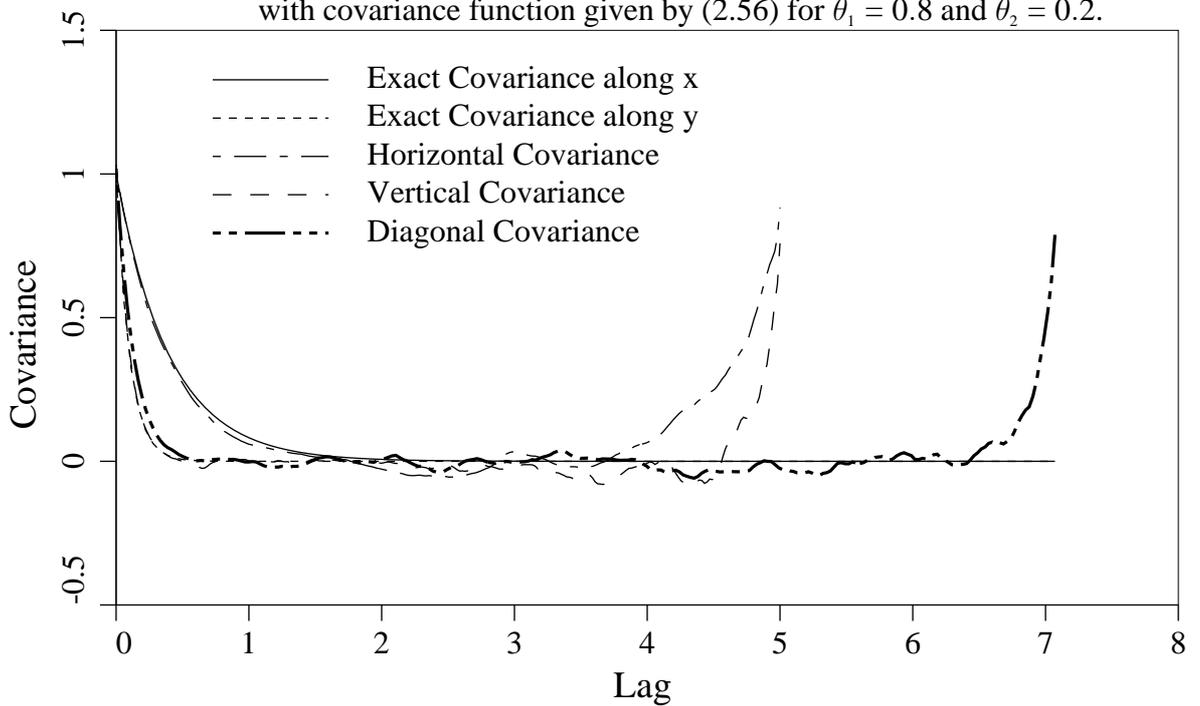


Figure 2.19 Estimated statistics of an anisotropic 2-D field (covariance formulation) with covariance function given by (2.56) for $\theta_1 = 0.8$ and $\theta_2 = 0.2$.

Chapter 3

Moving Average and Turning Bands Methods

3.1 Introduction

In this chapter, two simulation methods will be investigated that both depend on the solution of certain integral equations for their implementation. The first is the Moving Average technique which is usually combined with the Auto-Regressive method to form what is called an ARMA model. The auto-regressive technique involves generating values of a process along a line based on past values of the process,

$$\mathcal{Z}(x_i) = aW_i + \sum_{j=1}^p b_j \mathcal{Z}(x_{i-j}), \quad (3.1)$$

in which a and b_j are coefficients determined by the desired first and second-order statistics of the process and W_i is a discrete white noise process having zero mean and unit variance. Although auto-regression is a very efficient means of generating one-dimensional processes, it does not have a simple interpretation in higher dimensions since the ‘past’ of a point $\mathcal{Z}(x_1, x_2)$ has little meaning. For this reason, the auto-regressive techniques have not been pursued in this work and only the moving average techniques, which are well defined in higher dimensions, have been considered. It should be noted that Naganuma *et al.* [49] have developed an ARMA model for two-dimensional processes that defines the ‘past’ as lying in the lower left quadrant below the point in question. Although there is some question as to the validity of such an assumption, they achieve very good results in matching the estimated covariance structure with the desired which is the more important criteria.

3.2 Moving Average Method

The Moving Average (MA) technique of simulating random processes is a well known approach involving the expression of the process as an average of an underlying white noise process. Formally, if $Z(\mathbf{x})$ is the desired process then

$$Z(\mathbf{x}) = \int_{-\infty}^{\infty} f(\boldsymbol{\xi}) dW(\mathbf{x} + \boldsymbol{\xi}), \quad (3.2a)$$

or equivalently,

$$Z(\mathbf{x}) = \int_{-\infty}^{\infty} f(\boldsymbol{\xi} - \mathbf{x}) dW(\boldsymbol{\xi}), \quad (3.2b)$$

in which $dW(\boldsymbol{\xi})$ is the incremental white noise process at the location $\boldsymbol{\xi}$ with statistical properties

$$\mathbb{E} [dW(\boldsymbol{\xi})] = 0,$$

$$\mathbb{E} [dW(\boldsymbol{\xi})]^2 = d\boldsymbol{\xi}, \quad (3.3)$$

$$\mathbb{E} [dW(\boldsymbol{\xi})dW(\boldsymbol{\xi}')] = 0, \quad \text{if } \boldsymbol{\xi} \neq \boldsymbol{\xi}',$$

and $f(\boldsymbol{\xi})$ is a weighting function determined from the second order statistics of $Z(\mathbf{x})$

$$\begin{aligned} \mathbb{E} [Z(\mathbf{x}) Z(\mathbf{x} + \boldsymbol{\tau})] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\boldsymbol{\xi} - \mathbf{x}) f(\boldsymbol{\xi}' - \mathbf{x} - \boldsymbol{\tau}) \mathbb{E} [dW(\boldsymbol{\xi}) dW(\boldsymbol{\xi}')], \\ &= \int_{-\infty}^{\infty} f(\boldsymbol{\xi} - \mathbf{x}) f(\boldsymbol{\xi} - \mathbf{x} - \boldsymbol{\tau}) d\boldsymbol{\xi}. \end{aligned} \quad (3.4)$$

If $Z(\mathbf{x})$ is homogeneous, then the dependence on \mathbf{x} disappears, and (3.4) can be written in terms of the covariance function (note by 3.3 that $\mathbb{E} [Z(\mathbf{x})] = 0$),

$$B(\boldsymbol{\tau}) = \int_{-\infty}^{\infty} f(\boldsymbol{\xi}) f(\boldsymbol{\xi} - \boldsymbol{\tau}) d\boldsymbol{\xi}. \quad (3.5)$$

Defining the Fourier transform pair corresponding to $f(\boldsymbol{\xi})$ in \mathbb{R}^k to be,

$$F(\boldsymbol{\omega}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} f(\boldsymbol{\xi}) e^{-i\boldsymbol{\omega} \cdot \boldsymbol{\xi}} d\boldsymbol{\xi}, \quad (3.6a)$$

$$f(\boldsymbol{\xi}) = \int_{-\infty}^{\infty} F(\boldsymbol{\omega}) e^{i\boldsymbol{\omega} \cdot \boldsymbol{\xi}} d\boldsymbol{\omega}, \quad (3.6b)$$

then by the convolution theorem (3.5) can be expressed as

$$B(\boldsymbol{\tau}) = (2\pi)^n \int_{-\infty}^{\infty} F(\boldsymbol{\omega}) F(-\boldsymbol{\omega}) e^{-i\boldsymbol{\omega} \cdot \boldsymbol{\tau}} d\boldsymbol{\omega}, \quad (3.7)$$

from which a solution can be obtained from the Fourier transform of $B(\tau)$,

$$F(\omega)F(-\omega) = \frac{1}{(2\pi)^{2n}} \int_{-\infty}^{\infty} B(\tau) e^{-i\omega \cdot \tau} d\tau. \quad (3.8)$$

Note that the symmetry in the left hand side of (3.8) comes about due to the symmetry $B(\tau) = B(-\tau)$. It is still necessary to assume something about the relationship between $F(\omega)$ and $F(-\omega)$ in order to arrive at a final solution through the inverse transform. Usually the function $F(\omega)$ is assumed to be either even or odd.

Weighting functions corresponding to several common one-dimensional covariance functions have been determined by a number of authors, notably Journel and Huijbregts [32] and Mantoglou and Wilson [43]. In higher dimensions, the calculation of weighting functions becomes quite complex and is often done numerically through FFT's. The non-uniqueness of the weighting function and the difficulty in finding it, particularly in higher dimensions, renders this method of questionable value to the user who wishes to be able to handle arbitrary covariance functions.

Leaving this issue for the moment, the implementation of the MA method is itself a rather delicate problem. For a discrete process in one dimension, (3.2a) can be written

$$Z_i = \sum_{j=-\infty}^{\infty} f_j W_{i,j}, \quad (3.9)$$

where $W_{i,j}$ is a discrete white noise process taken to have zero mean and unit variance. To implement this, the sum must be restricted to some range p , usually chosen such that $f_{\pm p}$ is negligible,

$$Z_i = \sum_{j=-p}^p f_j W_{i,j}. \quad (3.10)$$

The next concern is how to discretize the underlying white noise process. If Δx is the increment of the physical process such that $Z_i = Z((i-1)\Delta x)$ and Δu is the incremental distance between points of the underlying white noise process, such that

$$W_{i,j} = W((i-1)\Delta x + j\Delta u), \quad (3.11)$$

then $f_j = f(j\Delta u)$ and Δu should be chosen such that the quotient $r = \Delta x/\Delta u$ is an integer for simplicity. Figure 3.1 illustrates the relationship between Z_i and the discrete white noise process. For finite Δu , the discrete approximation (3.10) will introduce some error into the estimated covariance of the realization. This error can often be removed through a multiplicative correction factor as shown by Journel and Huijbregts [32] but in general is reduced by taking Δu as small as practically possible (and thus p as large as possible).

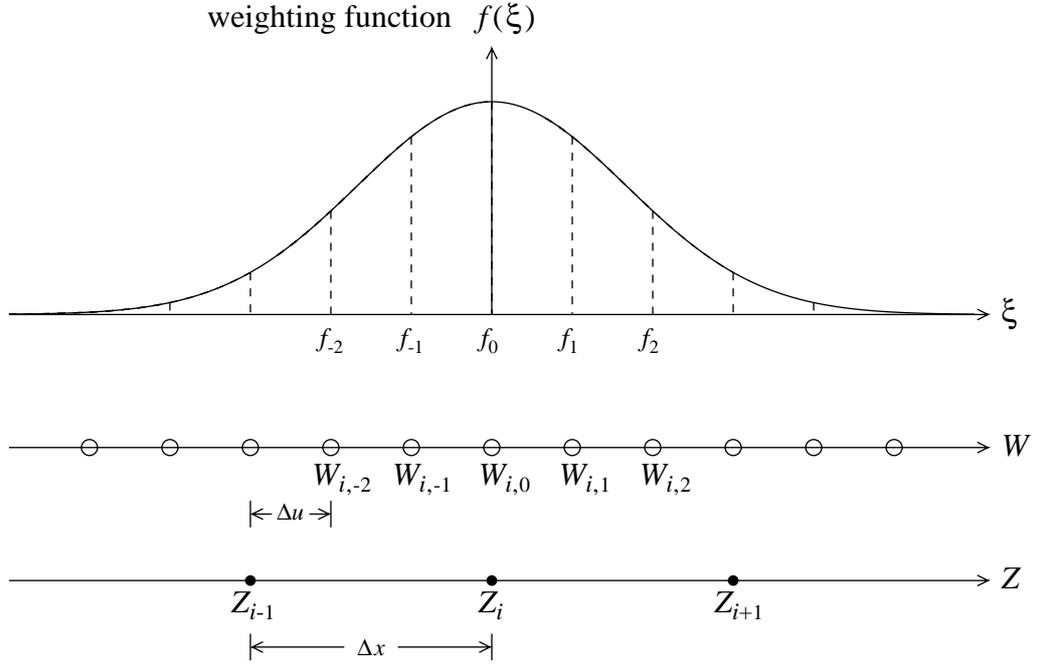


Figure 3.1 Schematic representation of the moving average process in one dimension.

Once the discretization of the underlying white noise process and the range p has been determined, the implementation of (3.10) in one dimension is quite straightforward and usually quite efficient for reasonable values of p . In higher dimensions, the method rapidly becomes cumbersome. Figure 3.2 shows a typical portion of a 2-D discrete process Z_{ij} , marked by X's, and the underlying white noise field, marked by dots. The entire figure represents the upper right corner of a 2-D field. The process Z_{ij} is now formed by the double summation

$$Z_{ij} = \sum_{k=-p_1}^{p_1} \sum_{\ell=-p_2}^{p_2} f_{k\ell} W_{i,j,k,\ell}, \quad (3.12)$$

where $f_{k\ell}$ is the 2-D weighting function and $W_{i,j,k,\ell}$ is the discrete white noise process centered at the same position as Z_{ij} . The i and j subscripts on W are for bookkeeping purposes so that the sum is performed over a centered neighborhood of discrete white noise values.

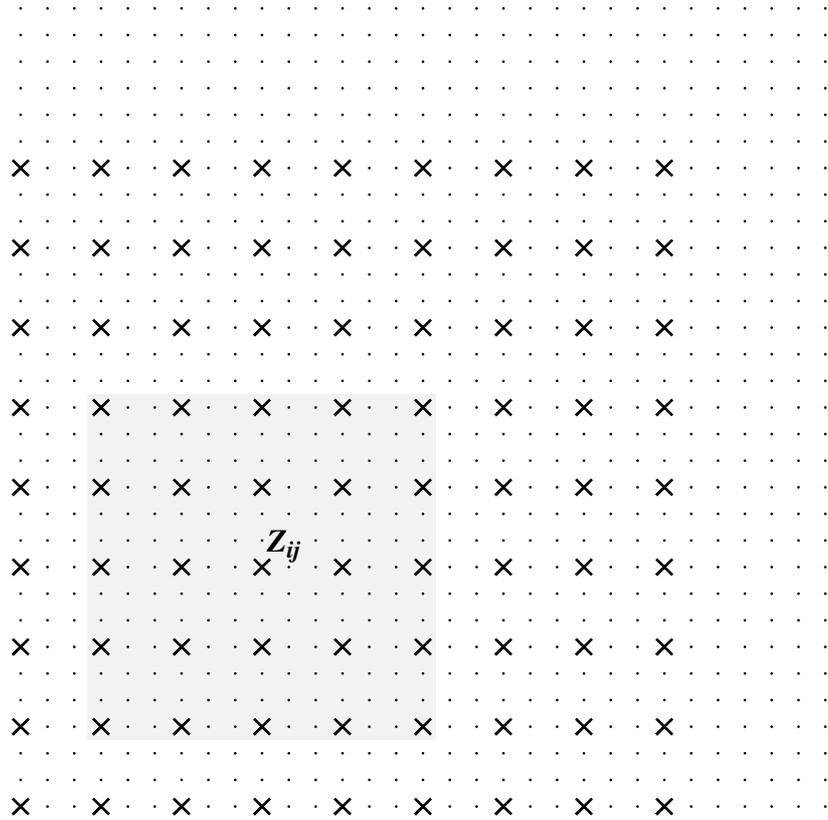


Figure 3.2 Two-dimensional moving average process. Z_{ij} is formed by summing the contributions from the underlying white noise process in the shaded region.

In the typical example illustrated in Figure 3.2, the discretization of the white noise process is such that $r = \Delta u / \Delta x = 3$ and a relatively short scale of fluctuation was used so that $p = 6$. This means that if a $K_1 \times K_2$ field is to be simulated, the total number of white noise realizations to be generated must be,

$$N_w = \left(1 + 2p_1 + r_1(K_1 - 1)\right) \left(1 + 2p_2 + r_2(K_2 - 1)\right), \quad (3.13)$$

or in the neighborhood of $(rK)^2$ for a square field. This can be contrasted immediately with the FFT approach which requires the generation of about $\frac{1}{2}K^2$ random values for a quadrant symmetric process (note that the factor of one-half is a consequence of the periodicity of the generated field). When $r = 3$, some 18 times as many white noise realizations must be generated for the moving average algorithm as for the FFT method. Also the construction of each field point requires a total of $(2p + 1)^2$ additions and multiplications which, for the not unreasonable example given above, is $13^2 = 169$. This means that the entire field will be generated using $K^2(2p + 1)^2$ or about 11 million additions and multiplications for a $200 \times$

200 field. Again this can be contrasted to the two-dimensional FFT method (radix-2, row-column algorithm) which requires some $4K^2 \log_2 K$ or about 2 million multiply-adds. In most cases, the moving average approach in two dimensions was found to run at least 10 times slower than the FFT approach. In three dimensions, the moving average method used to generate a $64 \times 64 \times 64$ field with $p = 6$ was estimated to run over 100 times slower than the corresponding FFT approach. For this reason, and since the weighting function is generally difficult to find, the moving average method as a general method of producing realizations of multi-dimensional random fields was abandoned by the author.

It can be noted in passing that the two-dimensional ARMA model suggested by Naganum *et al.* [49] requires about 50 to 150 multiply-adds (depending on the type of covariance structure modeled) for each field point. This is about 2 to 6 times slower than the FFT approach. While this is quite competitive for certain covariance functions, the corresponding run speeds for three-dimensional processes are estimated to be 15 to 80 times slower than the FFT approach depending on the choice of parameters p and r .

3.3 Turning Bands Method

The Turning Bands Method (TBM), as originally suggested by Matheron [45], involves the simulation of isotropic random fields in two- or higher-dimensional space by using a sequence of one-dimensional processes along lines crossing the space. With reference to Figure 3.3, the algorithm can be described as follows,

- 1) choose an arbitrary origin within or near the domain of the field to be generated,
- 2) select a line i crossing the domain having a direction given by the unit vector \mathbf{u}_i which may be chosen either randomly or from some fixed set,
- 3) generate a realization of a one-dimensional process, $\mathcal{Z}_i(\xi_i)$, along the line i having zero mean and covariance function $B_1(\tau_i)$ where ξ_i and τ_i are measured along line i ,
- 4) orthogonally project each field point \mathbf{x}_k onto the line i to define the coordinate $\xi_{ki} = \mathbf{x}_k \cdot \mathbf{u}_i$ of the one-dimensional process value $\mathcal{Z}_i(\xi_{ki})$,
- 5) add the component $\mathcal{Z}_i(\xi_{ki})$ to the field value $\mathcal{Z}(\mathbf{x}_k)$ for each \mathbf{x}_k ,
- 6) return to step (2) and generate a new one-dimensional process along a subsequent line until L lines have been produced,
- 7) normalize the field $\mathcal{Z}(\mathbf{x}_k)$ by dividing through by the factor \sqrt{L} .

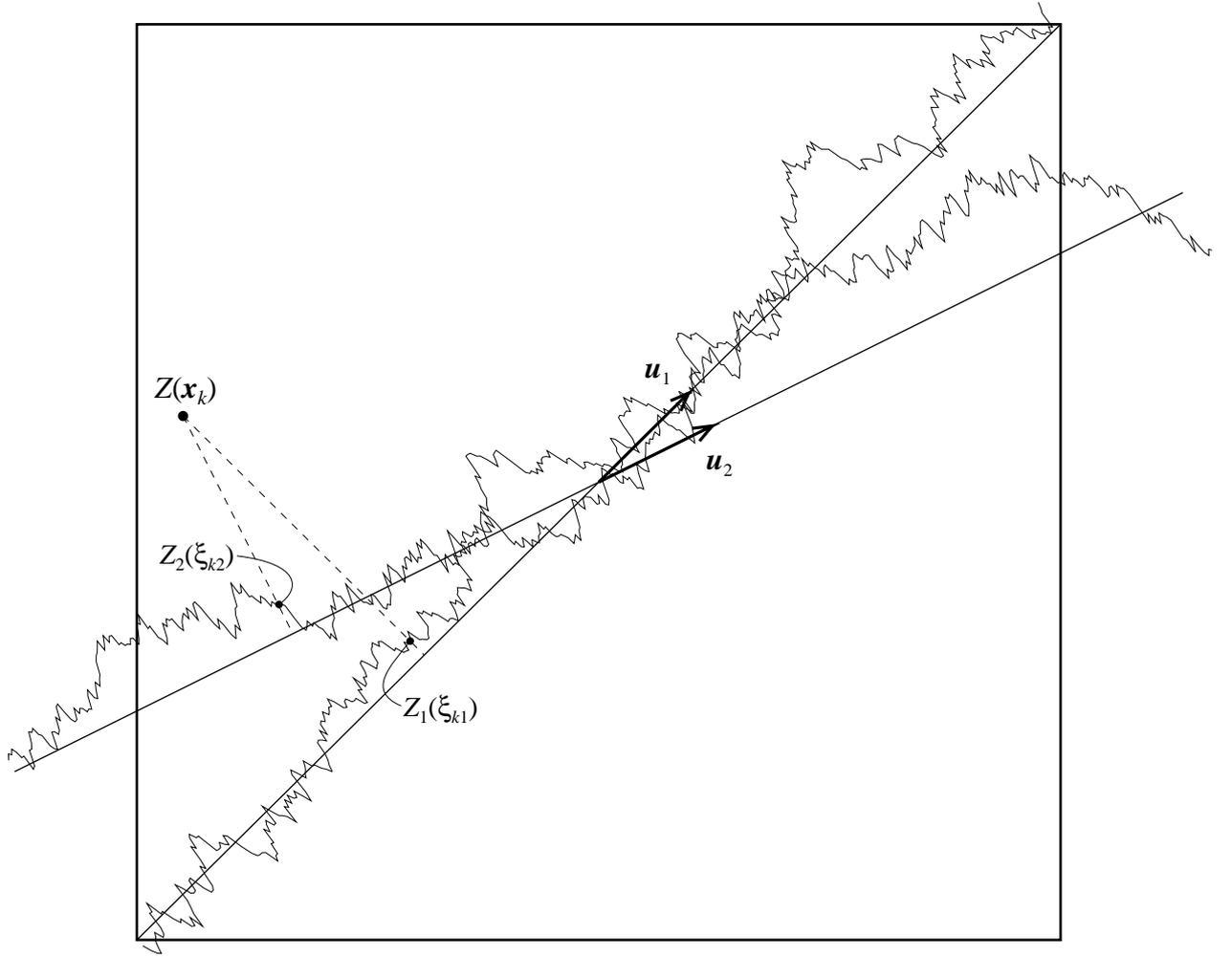


Figure 3.3 The Turning Bands Method: contributions from the line process $Z_i(\xi_i)$ at $x_k \cdot u_i$ are summed into the field process $Z(x)$ at x_k .

Essentially, the generating equation for the zero-mean discrete process $\mathcal{Z}(x)$ is given by

$$\mathcal{Z}(x_k) = \frac{1}{\sqrt{L}} \sum_{i=1}^L \mathcal{Z}_i(x_k \cdot u_i), \quad (3.14)$$

which can be an exceptionally fast algorithm, particularly as the number of dimensions of the process increases. It depends on knowledge of the one-dimensional covariance function, $B_1(\tau)$. Once this is known, the line processes can be produced using some efficient 1-D algorithm such as auto-regressive, moving average, or FFT techniques.

The covariance function $B_1(\tau)$ is chosen such that the multi-dimensional covariance structure $B_n(\tau)$ in \mathbb{R}^n is reflected in each realization or over the ensemble. For two-dimensional isotropic processes, Mantoglou and Wilson [43] give the following relationship between $B_2(\tau)$ and $B_1(\xi)$ for $r = |\tau|$,

$$B_2(r) = \frac{2}{\pi} \int_0^r \frac{B_1(\xi)}{\sqrt{r^2 - \xi^2}} d\xi, \quad (3.15)$$

which is an integral equation to be solved for $B_1(\xi)$. In three dimensions, the relationship between the isotropic $B_3(r)$ and $B_1(\xi)$ is particularly simple,

$$B_1(\xi) = \frac{d}{d\xi} \left(\xi B_3(\xi) \right). \quad (3.16)$$

Mantoglou and Wilson supply explicit solutions for either the equivalent one-dimensional covariance function or the equivalent one-dimensional spectral density function for a variety of common multi-dimensional covariance structures. In particular for the exponential type covariance function,

$$B_2(\tau_1, \tau_2) = \sigma^2 \exp \left\{ -\frac{2}{\theta} \sqrt{\tau_1^2 + \tau_2^2} \right\}, \quad (3.17)$$

the corresponding one-dimensional one-sided spectral density function is shown to be

$$G_1(\omega) = \frac{2\omega\sigma^2\theta^2}{[4 + \theta^2\omega^2]^{3/2}}. \quad (3.18)$$

In this implementation of the TBM, the line processes were constructed using a 1-D FFT algorithm as discussed in Chapter 2. Line lengths were chosen to be twice that of the field diagonal to avoid the symmetric covariance problem inherent with the FFT method. To reduce errors arising due to overly coarse discretization of the lines, the ratio between the incremental distance along the lines, Δu , and the minimum incremental distance in the field along any coordinate, Δx , was selected to be $\Delta u/\Delta x = \frac{1}{2}$. Figure 3.4 represents a realization of a 2-D process with covariance function given by (3.17) in which the mid-point of each line was located at the center of the domain as was illustrated in Figure 3.3. The finite number of lines used, in this case 16, results in a streaked appearance of the realization which is more pronounced if fewer lines are used and less pronounced as the number of lines increases. These artifacts are still evident using 32 lines, but, as shown in Figure 3.5, are almost invisible when using 64 lines (the use of number of lines which are powers of 2 is arbitrary). Since the 16 line case runs at about the same speed as the 2-D FFT approach, the elimination of the streaks in the realization comes at a price of running about 3 to 5 times as slow as the FFT method. Other origin locations were tried including the use of all four corners (the particular corner selected as an origin dependent on which quadrant the unit vector u_i points into) with no particular success in improving the appearance of realizations for a small number of lines.

The orientation of the lines can either be chosen randomly or selected from a set of prescribed directions which evenly divide the unit circle or sphere. In three dimensions, the maximum number of lines which will subdivide the unit sphere into equal solid angles is 15 [32]. If more lines are desired (as is likely to eliminate the streaking phenomena) then the sphere can only be approximately subdivided. Both of the realizations shown in Figures 3.4

and 3.5 were created using evenly spaced lines on the unit circle. By using an ‘ergodic’ uni-dimensional generator (see Equations 2.9 and 2.12), Mantoglou and Wilson state that evenly oriented lines will produce an ergodic realization. This assertion was tested to a limited extent by the author, as shown in Figures 3.6 and 3.7, and it was found that while neither approach yielded a strictly ergodic process over the limited domain, the use of evenly spaced lines did render estimated statistics which converged somewhat more quickly to the desired. This is more evident when the scale of fluctuation is relatively large, as in Figure 3.6. No particular difference was evident between those realization produced with an ‘ergodic’ 1-D FFT line generating algorithm (2.9) and those produced with a ‘non-ergodic’ generator (2.7). Perhaps more importantly, the use of evenly oriented lines reduced the streaks apparent in the realizations.

Note that the Turning Bands Method does not suffer from the symmetric covariance structure that is inherent in the FFT approach. However, the necessity of finding an equivalent 1-D covariance or spectral density function through an integral equation along with the streaked appearance of the realization when an insufficient number of lines are used makes the method less attractive. In two dimensions, the efficiency of the method was found to be reasonably competitive with the 2-D FFT algorithm, running at about the same speed if streaks are acceptable or 3 to 5 times slower if the streaks are to be eliminated. Since it is difficult to visualize 3-D realizations, no tests were run to evaluate their appearance and so no estimate of the number of lines required to eliminate streaking was done. One could presume that a similar ‘density’ of lines is required in the 3-D case as in the 2-D case and so the speed comparisons between the 3-D TBM and 3-D FFT approach may be roughly the same as in the 2-D case.

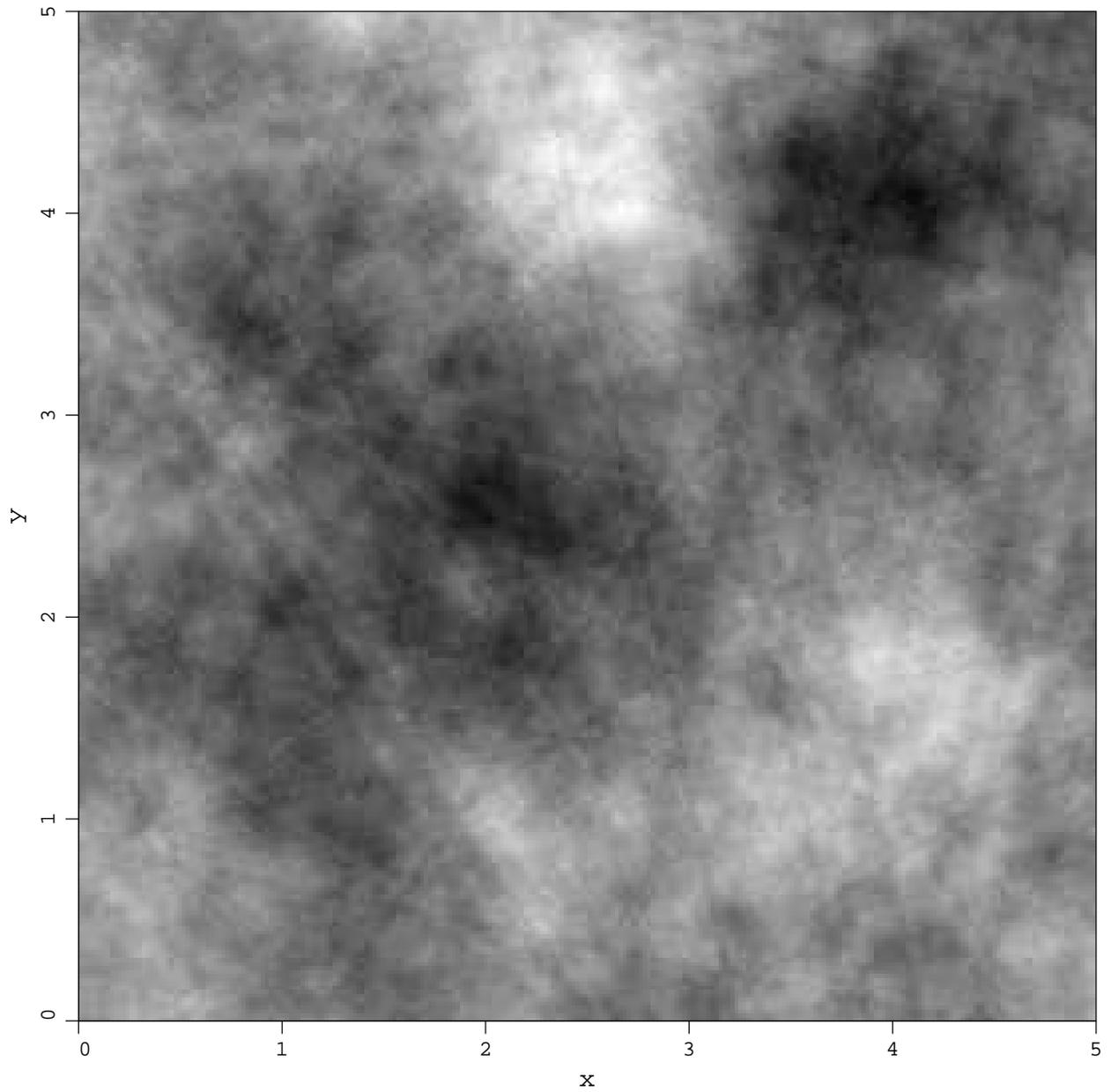


Figure 3.4 Sample function of a 2-D field via TBM using 16 lines with covariance function given by (3.17) for $\theta = 4$. Notice the banded or streaked appearance.

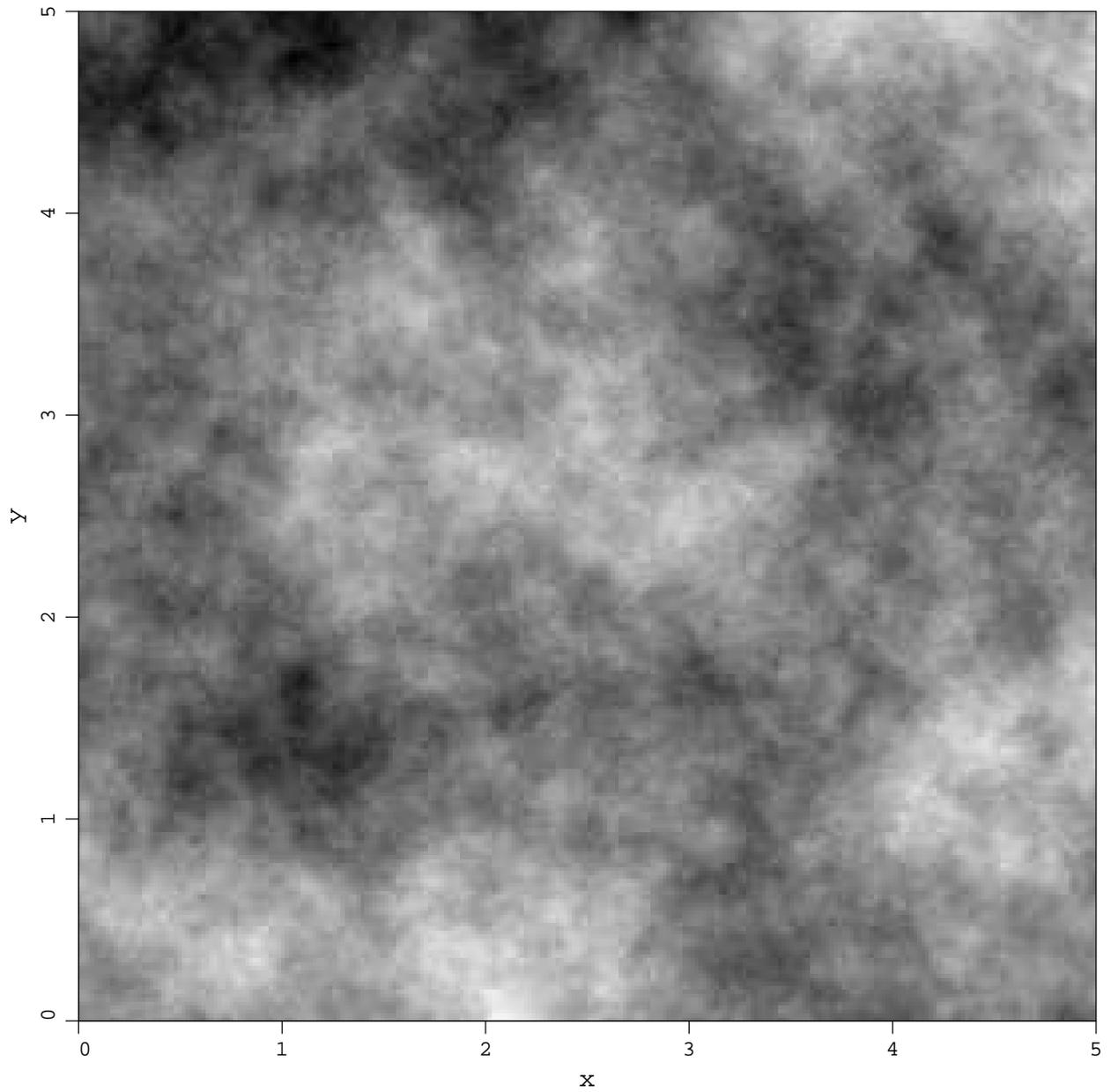


Figure 3.5 Sample function of a 2-D field via TBM using 64 lines with covariance function given by (3.17) for $\theta = 4$. The streaked appearance has largely disappeared.

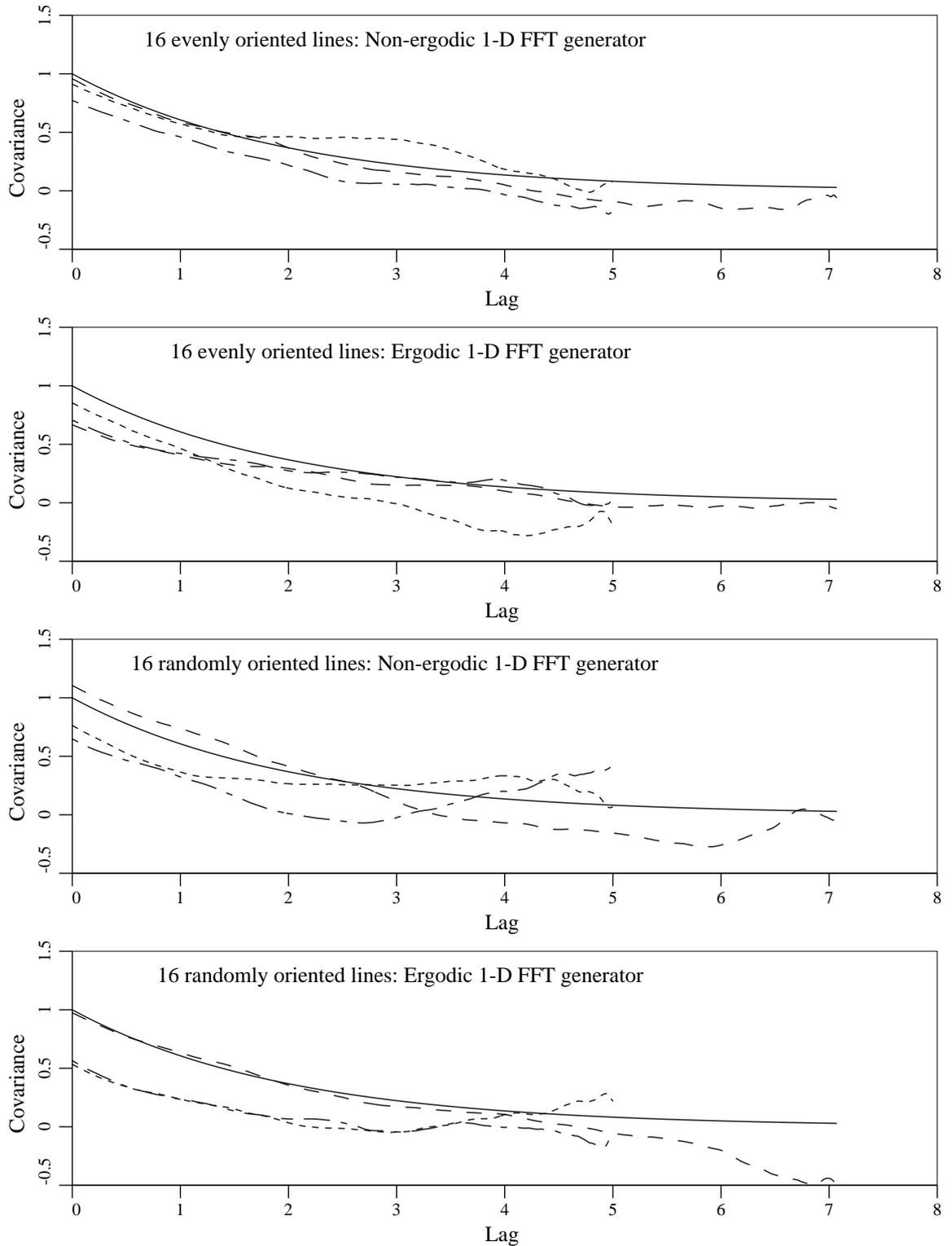


Figure 3.6 Comparison of estimated statistics of 2-D fields generated via TBM using scale $\theta = 4$ for randomly versus evenly oriented lines and ‘ergodic’ versus ‘non-ergodic’ generation of the line processes. Solid lines denote the exact covariance and dashed lines represent estimated horizontal, vertical, and diagonal covariances averaged over 10 realizations.

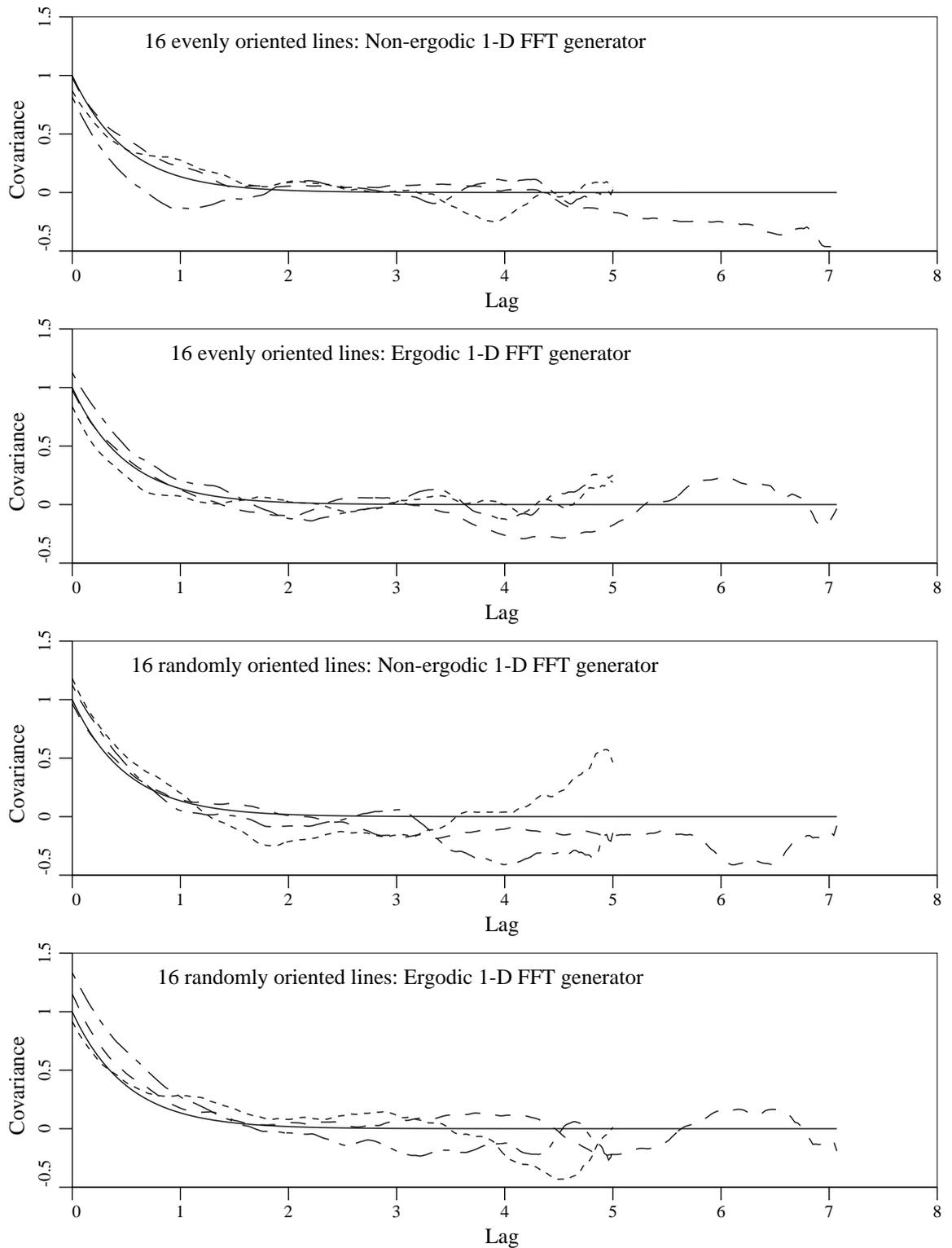


Figure 3.7 Comparison of estimated statistics of 2-D fields generated via TBM using scale $\theta = 1$ for randomly versus evenly oriented lines and ‘ergodic’ versus ‘non-ergodic’ generation of the line processes. Solid lines denote the exact covariance and dashed lines represent estimated horizontal, vertical, and diagonal covariances averaged over 10 realizations.

Chapter 4

Local Average Subdivision

4.1 Introduction

In this chapter a fast and accurate method of producing realizations of a discrete local average random process is presented. The motivation for such an approach arises out of a need to account for the fact that most engineering measurements are only defined over some finite domain and thus represent a local average of the property. For example, soil porosity is ill-defined at the micro-scale – it is measured using samples of finite volume and the variability of the values obtained is often significantly effected by the volume tested. The same is true of strength measurements, say, of concrete cylinders, or radar measurements of cloud densities (see also Rodriguez-Iturbe [60]). A properly defined local average process is therefore more easily related to actual measurements made at any scale and those measures more easily incorporated.

A further advantage of the method proposed herein is that it is ideally suited to stochastic finite element modeling using efficient, low order, interpolation functions. Each discrete local average given by a realization becomes the average property within each discrete element. In this context, the ability to easily change the resolution of a region of the domain while maintaining internal consistency gives finite element modelers the freedom of changing mesh resolution in regions of interest.

The concept behind the Local Average Subdivision (LAS) approach arose out of the stochastic subdivision algorithm described by Carpenter [12] and Fournier *et al.* [21]. Their method is limited to modeling power spectra having a ω^{-d} form and suffered from problems with aliasing and ‘creasing’. Lewis [39] generalized the approach to allow the modeling of arbitrary power spectra without eliminating the aliasing. Such midpoint displacement algorithms involve recursively subdividing the domain by generating new midpoint values randomly selected according to some distribution. Once chosen, the value at a point remains fixed and at each stage in the subdivision only half the points in the process are determined (the others created in previous iterations). Aliasing arises because the power spectral density is not modified at each stage to reflect the increasing Nyquist frequency associated with each increase in resolution. Voss [53, Chap. 1] attempted to eliminate this problem

with considerable success by adding randomness to all points at each stage in the subdivision in a method called ‘successive random additions’. However the internal consistency easily achieved by the midpoint displacement methods (their ability to return to previous states while decreasing resolution through decimation) is largely lost with the successive random additions technique. The property of internal consistency in the midpoint displacement approach implies that certain points retain their value throughout the subdivision and other points are created to remain consistent with them with respect to correlation. In the LAS approach, internal consistency implies that certain regions maintain a constant average throughout the subdivision. The property of internal consistency is important because it allows the process to be easily conditioned.

The method proposed here solves the problems associated with the stochastic subdivision methods and incorporates into it concepts of local averaging theory. The general procedure is presented first for a one-dimensional stationary process characterized by its second-order statistics. The algorithm is illustrated by a Ornstein-Uhlenbeck process, having a simple exponential correlation function, as well as by a fractional Gaussian noise process as defined by Mandelbrot [41]. The simulation procedure in two and three dimensions is then described. Finally some comments concerning the relative efficiency of the method are made.

4.2 One-Dimensional Local Average Subdivision

The construction of a local average process via LAS essentially proceeds in a top-down recursive fashion as illustrated in Figure 4.1. In Stage 0, a global average is generated for the process. At Stage 1, the domain is subdivided into two regions whose ‘local’ averages must in turn average to the global (or parent) value. Subsequent stages are obtained by subdividing each ‘parent’ cell and generating values for the resulting two regions while preserving upwards averaging. Note that the global average remains constant throughout the subdivision – a property that is ensured merely by requiring that the average of each pair generated is equivalent to the parent cell average. This ‘constant average’ is also a property of any cell being subdivided – such internal consistency allows for simple conditioning of the process. Specifically, the algorithm proceeds as follows;

- 1) generate a normally distributed global average (labeled Z_1^0 in Figure 4.1) with mean zero and variance obtained from local averaging theory,
- 2) subdivide the field into two equal parts,
- 3) generate two normally distributed values, Z_1^1 and Z_2^1 , whose means and variances are selected so as to satisfy three criteria:
 - a) that they show the correct variance according to local averaging theory,

- b) that they are properly correlated with one another,
- c) that they average to the parent value, $\frac{1}{2}(Z_1^1 + Z_2^1) = Z_1^0$.

That is, the distributions of Z_1^1 and Z_2^1 are conditioned on the value of Z_1^0 ,

- 4) subdivide each cell in stage 1 into two equal parts,
- 5) generate two normally distributed values, Z_1^2 and Z_2^2 , whose means and variances are selected so as to satisfy four criteria:
 - a) that they show the correct variance according to local averaging theory,
 - b) that they are properly correlated with one another,
 - c) that they average to the parent value, $\frac{1}{2}(Z_1^2 + Z_2^2) = Z_1^1$,
 - d) that they are properly correlated with Z_3^2 and Z_4^2 .

The third criteria implies conditioning of the distributions of Z_1^2 and Z_2^2 on the value of Z_1^1 . The fourth criteria will only be satisfied approximately by conditioning their distributions also on Z_2^1 .

and so on in this fashion. The approximations in the algorithm come about in two ways: first, as was already mentioned, the correlation with adjacent cells across parent boundaries is accomplished through use of the parent values (which are already known having been previously generated). Second the range of parent cells on which to condition the distributions will be limited to some neighborhood. The remainder of this chapter is devoted largely to the determination of these conditional Gaussian distributions at each stage in the subdivision and to an estimation of the algorithmic errors. In the following, the term ‘parent cell’ refers to the previous stage cell being subdivided and ‘within cell’ means within the domain defined by the boundary of the parent cell. The symbol Z is used to denote the algorithmic process and Z to denote the exact process throughout.

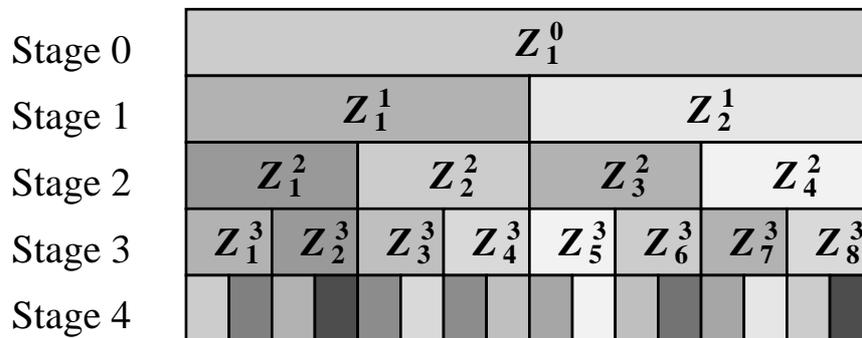


Figure 4.1 Top-down approach to the LAS construction of a random process

Consider first a continuous stationary scalar random function $Z(x)$ in one dimension, a sample of which may appear as shown in Figure 4.2, and define a domain of interest $[0, T]$ within which a realization is to be produced. Two comments should be made at this point: First, as it is currently implemented the method is restricted to stationary processes fully

described by their second-order statistics (mean, variance and autocorrelation function or, equivalently, spectral density function). Second, the subdivision procedure depends on the physical size of the domain being defined since the dimension over which local averaging is to be performed must be known.

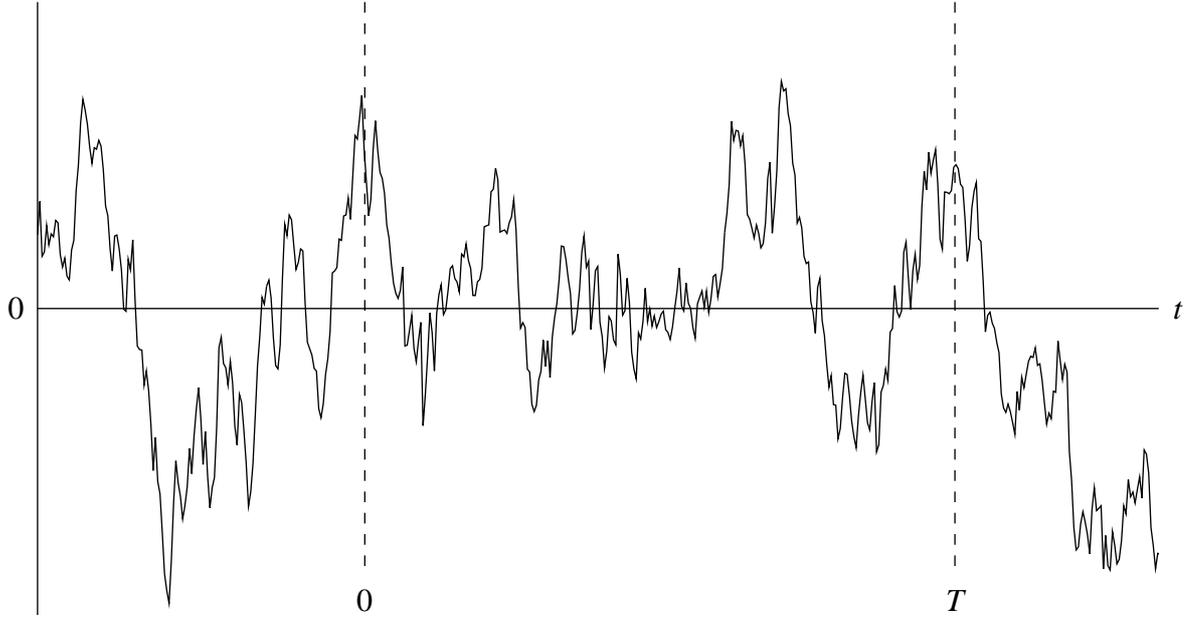


Figure 4.2 Sample function of an Ornstein-Uhlenbeck process with $B(\tau) = e^{-\frac{|\tau|}{2}}$.

The average of $Z(x)$ over the domain $[0, T]$ is given by

$$Z_1^0 = \frac{1}{T} \int_0^T Z(x) dx, \quad (4.1)$$

where Z_1^0 is a random variable whose statistics

$$\mathbf{E} [Z_1^0] = \mathbf{E} [Z], \quad (4.2)$$

$$\mathbf{E} [(Z_1^0)^2] = \frac{1}{T^2} \int_0^T \int_0^T \mathbf{E} [Z(\xi) Z(\xi')] d\xi d\xi' = \mathbf{E}^2[Z] + \frac{2}{T^2} \int_0^T (T - \tau) B(\tau) d\tau, \quad (4.3)$$

can be found by making use of stationarity and the fact that $B(\tau)$, the covariance function of $Z(t)$, is an even function of lag τ according to (1.9b). Assuming $Z(t)$ to be a mean-zero Gaussian random function then (4.2) and (4.3) give sufficient information to generate a realization of Z_1^0 which becomes Z_1^0 (Stage 0) in the LAS method.

Consider now the general case where stage i is known and stage $i + 1$ is to be generated. In the following the superscript i denotes the stage under consideration. Define

$$T^i = \frac{T}{2^i}, \quad i = 0, 1, 2, \dots, L, \quad (4.4)$$

where the desired number of discrete local averages in the realization is $N = 2^L$, and define Z_k^i to be the average of $Z(x)$ over the interval $(k-1)T^i < x \leq kT^i$ centered at $x_k = (k-\frac{1}{2})T^i$, i.e.

$$Z_k^i = \frac{1}{T^i} \int_{(k-1)T^i}^{kT^i} Z(x) dx \quad (4.5)$$

where $E[Z_k^i] = E[Z] = 0$. The target covariance between local averages separated by lag nT^i , $n > 0$, between centers is

$$\begin{aligned} E[Z_k^i Z_{k+n}^i] &= E \left[\left(\frac{1}{T^i} \right)^2 \int_{(k-1)T^i}^{kT^i} \int_{(k+n-1)T^i}^{(k+n)T^i} Z(\xi) Z(\xi') d\xi d\xi' \right] \\ &= \left(\frac{1}{T^i} \right)^2 \int_0^{T^i} \int_{nT^i}^{(n+1)T^i} B(\xi - \xi') d\xi d\xi' \\ &= \left(\frac{1}{T^i} \right)^2 \int_{(n-1)T^i}^{nT^i} (\xi - (n-1)T^i) B(\xi) d\xi + \left(\frac{1}{T^i} \right)^2 \int_{nT^i}^{(n+1)T^i} ((n+1)T^i - \xi) B(\xi) d\xi. \end{aligned} \quad (4.6)$$

A much simpler formulation is possible by introducing the concept of a variance function defined by (1.33) for the 1-D process as follows

$$\gamma(T^i) = \left(\frac{1}{\sigma T^i} \right)^2 \int_0^{T^i} \int_0^{T^i} B(\xi - \xi') d\xi d\xi' = 2 \left(\frac{1}{\sigma T^i} \right)^2 \int_0^{T^i} (|T^i| - |\tau|) B(\tau) d\tau, \quad (4.7)$$

where $\sigma^2 = B(0)$. Vanmarcke [25] has determined this function for a variety of processes. In terms of the variance function, (4.6) becomes for $n \geq 0$

$$E[Z_k^i Z_{k+n}^i] = \frac{\sigma^2}{2} \left[(n-1)^2 \gamma((n-1)T^i) - 2n^2 \gamma(nT^i) + (n+1)^2 \gamma((n+1)T^i) \right]. \quad (4.8)$$

	j		$j+1$		
	$2j-1$	$2j$	$2j+1$	$2j+2$	

Figure 4.3 1-D LAS indexing scheme for stage i (top) and stage $i+1$ (bottom).

With reference to Figure 4.3, the construction of stage $i+1$ values given stage i is obtained by estimating a mean for Z_{2j}^{i+1} and adding a zero mean discrete white noise $c^{i+1} W_j^{i+1}$ having variance $(c^{i+1})^2$,

$$Z_{2j}^{i+1} = M_{2j}^{i+1} + c^{i+1} W_j^{i+1}, \quad (4.9)$$

where the symbol \mathcal{Z} is used to denote the algorithmic process which will be an approximation to Z . The best linear estimate for the mean M_{2j}^{i+1} can be determined by a linear combination of stage i values in some neighborhood $j - n, \dots, j + n$,

$$M_{2j}^{i+1} = \sum_{k=j-n}^{j+n} a_{k-j}^i \mathcal{Z}_k^i. \quad (4.10)$$

Multiplying (4.9) through by \mathcal{Z}_m^i , taking expectations and using the fact that W_j^{i+1} is uncorrelated with the stage i values allows the determination of the coefficients a in terms of the desired covariances,

$$\mathbf{E} [Z_{2j}^{i+1} Z_m^i] = \sum_{k=j-n}^{j+n} a_{k-j}^i \mathbf{E} [Z_k^i Z_m^i], \quad (4.11)$$

i.e., a system of equations ($m = j - n, \dots, j + n$) from which the coefficients a_ℓ^i , $\ell = -n, \dots, n$, can be solved. Notice that the exact process Z is used when evaluating expectations. The covariance matrix multiplying the vector $\{a_\ell^i\}$ is both symmetric and Toeplitz (elements along each diagonal are equal). For $W_j^{i+1} \sim N(0, 1)$, the variance of the noise term is $(c^{i+1})^2$ which can be verified by squaring (4.9), taking expectations, and employing the results of (4.11)

$$(c^{i+1})^2 = \mathbf{E} [(Z_{2j}^{i+1})^2] - \sum_{k=j-n}^{j+n} a_{k-j}^i \mathbf{E} [Z_{2j}^{i+1} Z_k^i]. \quad (4.12)$$

The adjacent cell value, \mathcal{Z}_{2j-1}^{i+1} , is determined by ensuring that upwards averaging is preserved – that the average of each stage $i + 1$ pair equals the value of the stage i parent,

$$\mathcal{Z}_{2j-1}^{i+1} = 2 \mathcal{Z}_j^i - \mathcal{Z}_{2j}^{i+1}, \quad (4.13)$$

which incidentally gives a means of evaluating the cross-stage covariances

$$\mathbf{E} [Z_{2j}^{i+1} Z_m^i] = \frac{1}{2} \mathbf{E} [Z_{2j}^{i+1} Z_{2m-1}^{i+1}] + \frac{1}{2} \mathbf{E} [Z_{2j}^{i+1} Z_{2m}^{i+1}]. \quad (4.14)$$

All the expectations in Equations (4.11) to (4.14) are evaluated using (4.6) or (4.8) at the appropriate stage.

For stationary processes, the set of coefficients $\{a_\ell^i\}$ and c^i are independent of position since the expectations in (4.11) and (4.12) are just dependent on lags. The generation procedure can be restated as follows;

- 1) for $i = 0, 1, 2, \dots, L$ compute the coefficients $\{a_\ell^i\}$, $\ell = -n, \dots, n$ using (4.11) and the coefficients c^{i+1} using (4.12),
- 2) starting with $i = 0$, generate a realization for the global mean using (4.2) and (4.3),
- 3) subdivide the domain,

- 4) for each $j = 1, 2, \dots, 2^i$, generate realizations for Z_{2j}^{i+1} and Z_{2j-1}^{i+1} using (4.9) and (4.13),
- 5) increment i and, if not greater than L , return to step 3.

Because the LAS procedure is recursive, obtaining stage $i + 1$ values using the previous stage, it is relatively easy to condition the field simply by specifying the values of the local averages at a particular stage. So, for example, if the global mean of a process is known *a priori*, then the stage 0 value Z_1^0 can be set to this mean and the LAS procedure started at stage 1. Similarly if the resolution is to be refined in a certain region, then the values in that region become the starting values and the subdivision resumed at the next stage.

Although the LAS method yields a local average process, when the discretization size becomes small enough it is virtually indistinguishable from the limiting continuous process. Thus the method can be used to approximate continuous functions as well.

Finally it should be noted that the calculation of the coefficients a and c need only be done once for a particular process. Subsequent realizations can then be produced extremely efficiently by starting at step 2 in the procedure listed above.

4.2.1 Accuracy

It is instructive to investigate how closely the algorithm approximates the target statistics of the process from one subdivision to the next. Assuming the stage i values come from the exact process, the LAS scheme can be written

$$Z_{2j}^{i+1} = c^{i+1} W_j^{i+1} + \sum_{k=j-n}^{j+n} a_{k-j}^i Z_k^i, \quad (4.15)$$

$$Z_{2j-1}^{i+1} = 2 Z_j^i - Z_{2j}^{i+1}. \quad (4.16)$$

It is easy to see that the expectation of Z is zero since $E [Z_k^i] = 0$, as desired, while the variance is

$$\begin{aligned} E [(Z_{2j}^{i+1})^2] &= E \left[\left(c^{i+1} W_j^{i+1} + \sum_{k=j-n}^{j+n} a_{k-j}^i Z_k^i \right)^2 \right] \\ &= (c^{i+1})^2 + \sum_{k=j-n}^{j+n} a_{k-j}^i \sum_{\ell=j-n}^{j+n} a_{\ell-j}^i E [Z_k^i Z_\ell^i] \\ &= E [(Z_{2j}^{i+1})^2] - \sum_{k=j-n}^{j+n} a_{k-j}^i E [Z_{2j}^{i+1} Z_k^i] + \sum_{k=j-n}^{j+n} a_{k-j}^i E [Z_{2j}^{i+1} Z_k^i] \\ &= E [(Z_{2j}^{i+1})^2], \end{aligned} \quad (4.17)$$

in which the coefficients c^{i+1} and a_ℓ^i were calculated using (4.11) and (4.12) as before. Similarly, the ‘within cell’ covariance at lag T^{i+1} is

$$\begin{aligned}
\mathbf{E} [Z_{2j-1}^{i+1} Z_{2j}^{i+1}] &= \mathbf{E} \left[\left(2Z_j^i - c^{i+1} W_j^{i+1} - \sum_{k=j-n}^{j+n} a_{k-j}^i Z_k^i \right) \left(c^{i+1} W_j^{i+1} + \sum_{\ell=j-n}^{j+n} a_{\ell-j}^i Z_\ell^i \right) \right] \\
&= 2 \sum_{\ell=j-n}^{j+n} a_{\ell-j}^i \mathbf{E} [Z_\ell^i Z_j^i] - \mathbf{E} [(Z_{2j}^{i+1})^2] \\
&= 2 \mathbf{E} [Z_{2j}^{i+1} Z_j^i] - \mathbf{E} [(Z_{2j}^{i+1})^2] \\
&= \mathbf{E} [Z_{2j-1}^{i+1} Z_{2j}^{i+1}], \tag{4.18}
\end{aligned}$$

using the results of (4.17) along with (4.14). Thus the covariance structure within a cell is preserved *exactly* by the subdivision algorithm. Some approximation does occur across cell boundaries as can be seen by considering

$$\begin{aligned}
\mathbf{E} [Z_{2j}^{i+1} Z_{2j+1}^{i+1}] &= \mathbf{E} \left[\left(c^{i+1} W_j^{i+1} + \sum_{k=j-n}^{j+n} a_{k-j}^i Z_k^i \right) \left(2Z_{j+1}^i - c^{i+1} W_{j+1}^{i+1} - \sum_{\ell=j-n+1}^{j+n+1} a_{\ell-j-1}^i Z_\ell^i \right) \right] \\
&= 2 \sum_{k=j-n}^{j+n} a_{k-j}^i \mathbf{E} [Z_k^i Z_{j+1}^i] - \sum_{\ell=j-n+1}^{j+n+1} a_{\ell-j-1}^i \sum_{k=j-n}^{j+n} a_{k-j}^i \mathbf{E} [Z_k^i Z_\ell^i] \\
&= \mathbf{E} [Z_{2j}^{i+1} Z_{2j+1}^{i+1}] + \mathbf{E} [Z_{2j}^{i+1} Z_{2j+2}^{i+1}] - \sum_{\ell=j-n+1}^{j+n+1} a_{\ell-j-1}^i \mathbf{E} [Z_{2j}^{i+1} Z_\ell^i]. \tag{4.19}
\end{aligned}$$

The algorithmic error in this covariance comes from the last two terms. Using a neighborhood size of $n = 1$, the discrepancy between (4.19) and the exact covariance is illustrated numerically in Figure 4.4 for a zero mean Ornstein-Uhlenbeck process having covariance and variance functions

$$B(\tau) = \sigma^2 \exp \left\{ -\frac{2|\tau|}{\theta} \right\}, \tag{4.20}$$

$$\gamma(D) = \frac{\theta^2}{2D^2} \left[\frac{2D}{\theta} + \exp \left\{ \frac{-2D}{\theta} \right\} - 1 \right], \tag{4.21}$$

where D is the averaging dimension (in Figure 4.4, $D = T^{i+1}$). Although Figure 4.4 shows a wide range in the product $2D/\theta$ (and thus a wide range in effective cell sizes), the error is typically very small.

To address the issue of errors at larger lags and the possibility of errors accumulating from stage to stage, it is useful to look at the exact versus estimated ensemble statistics of the process. Figure 4.5 illustrates this comparison for the Ornstein-Uhlenbeck process. It can be seen from this example and from the fractional Gaussian noise example to come, that the errors seem to be self-correcting and the algorithmic correlation structure tends to the exact correlation function when averaged over several realizations. The within-cell rate of convergence of the estimated statistics to the exact is $\frac{1}{N_f}$, where N_f is the number of realizations. The overall rate of convergence of the LAS realizations to the exact statistics is about the same.

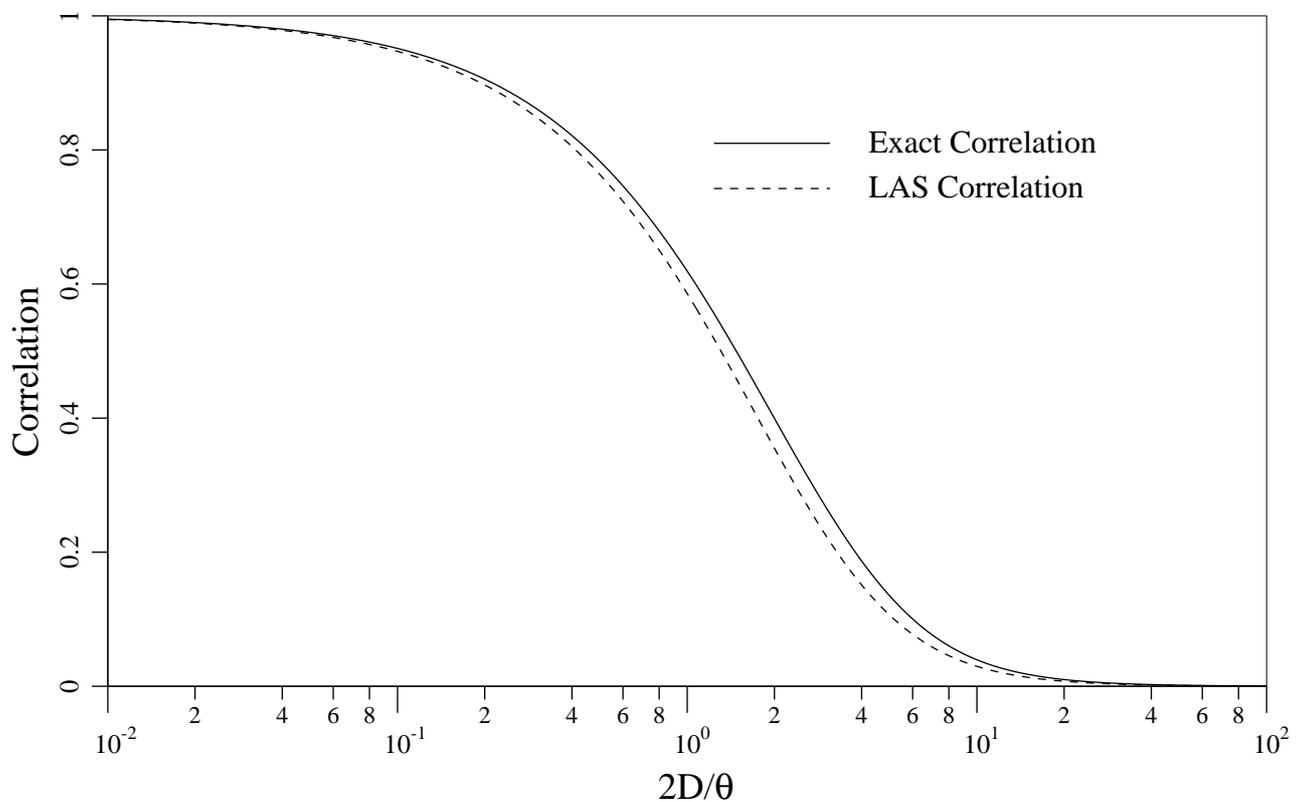


Figure 4.4 Comparison of algorithmic and exact covariance between adjacent cells across a parent cell boundary for varying effective averaging lengths $2D/\theta$.

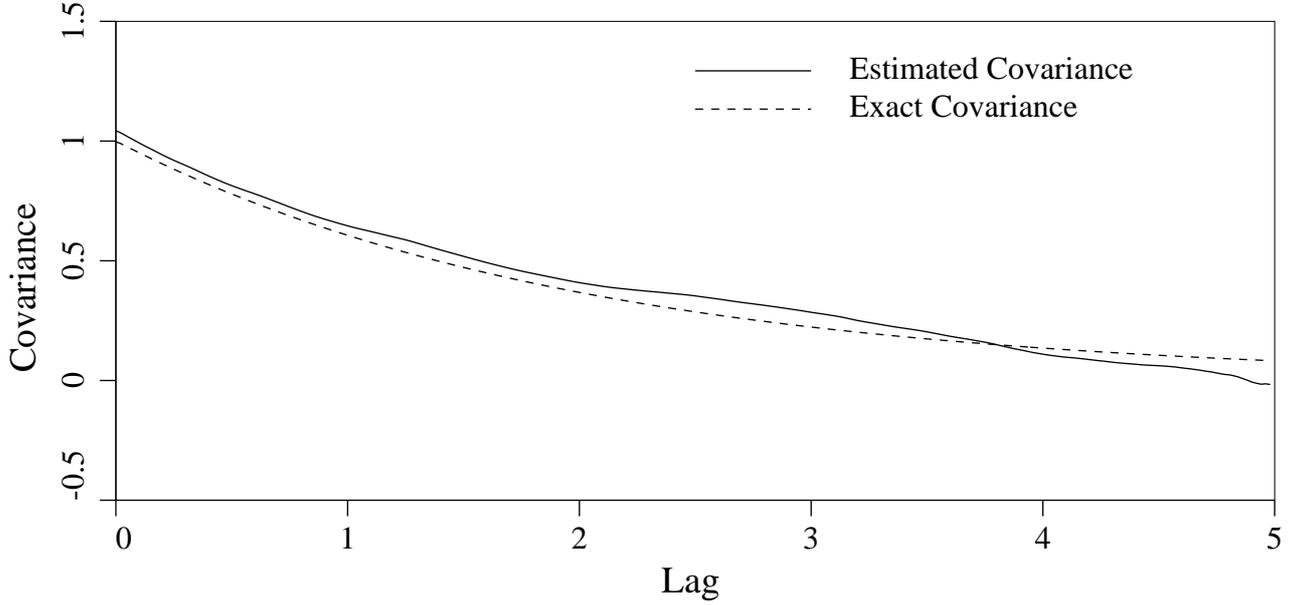


Figure 4.5 Comparison of exact and estimated covariance function for $\theta = 4$ averaged over 200 realizations.

4.2.2 Boundary Conditions and Neighborhood Size

When the neighborhood size $(2n + 1)$ is greater than 1 ($n > 0$), the construction of values near the boundary may require values from the previous stage which lie outside the boundary. This problem is handled by assuming that what happens outside the domain $[0, T]$ is of no interest and uncorrelated with what happens within the domain. The generating relationship (4.9) near either boundary becomes

$$\mathcal{Z}_{2j}^{i+1} = c^{i+1} W_j^{i+1} + \sum_{k=j-p}^{j+q} a_{k-j}^i \mathcal{Z}_k^i, \quad (4.22)$$

where $p = \min(n, j - 1)$, $q = \min(n, 2^i - j)$ and the coefficients a_ℓ^i need only be determined for $\ell = -p, \dots, q$. The periodic boundary conditions mentioned by Lewis [39] are not appropriate if the target covariance structure is to be preserved since they lead to a covariance which is symmetric about lag $T/2$ (unless the desired covariance is also symmetric about this lag).

In the implementation of the 1-D LAS method, a neighborhood size of 3 was used ($n = 1$), the parent cell plus its two adjacent cells. Because of the top-down approach, there seems to be little justification to using a larger neighborhood for processes with covariance functions which decrease monotonically or which are relatively smooth. When the covariance function is oscillatory, a larger neighborhood is required in order to successfully

approximate the function. In Figure 4.6 the observed and theoretical covariances are compared for a process with

$$B(\tau) = \sigma^2 \cos(\omega\tau) e^{-2\tau/\theta}, \quad (4.23)$$

which has the corresponding variance function

$$\gamma(D) = \frac{2\theta^2}{D^2(4 + \theta^2\omega^2)^2} \left\{ e^{-2D/\theta} [(4 - \theta^2\omega^2) \cos \omega D - 4\theta\omega \sin \omega D] + \frac{2D}{\theta} (4 + \theta^2\omega^2) - 4 + \theta^2\omega^2 \right\}. \quad (4.24)$$

Considerable improvement in the model is obtained when a neighborhood size of 5 is used ($n = 2$). This improvement comes at the expense of taking about twice as long to generate the realizations. Many practical models of natural phenomena employ monotonically decreasing covariance functions, often for simplicity, and so the $n = 1$ implementation is usually preferable.

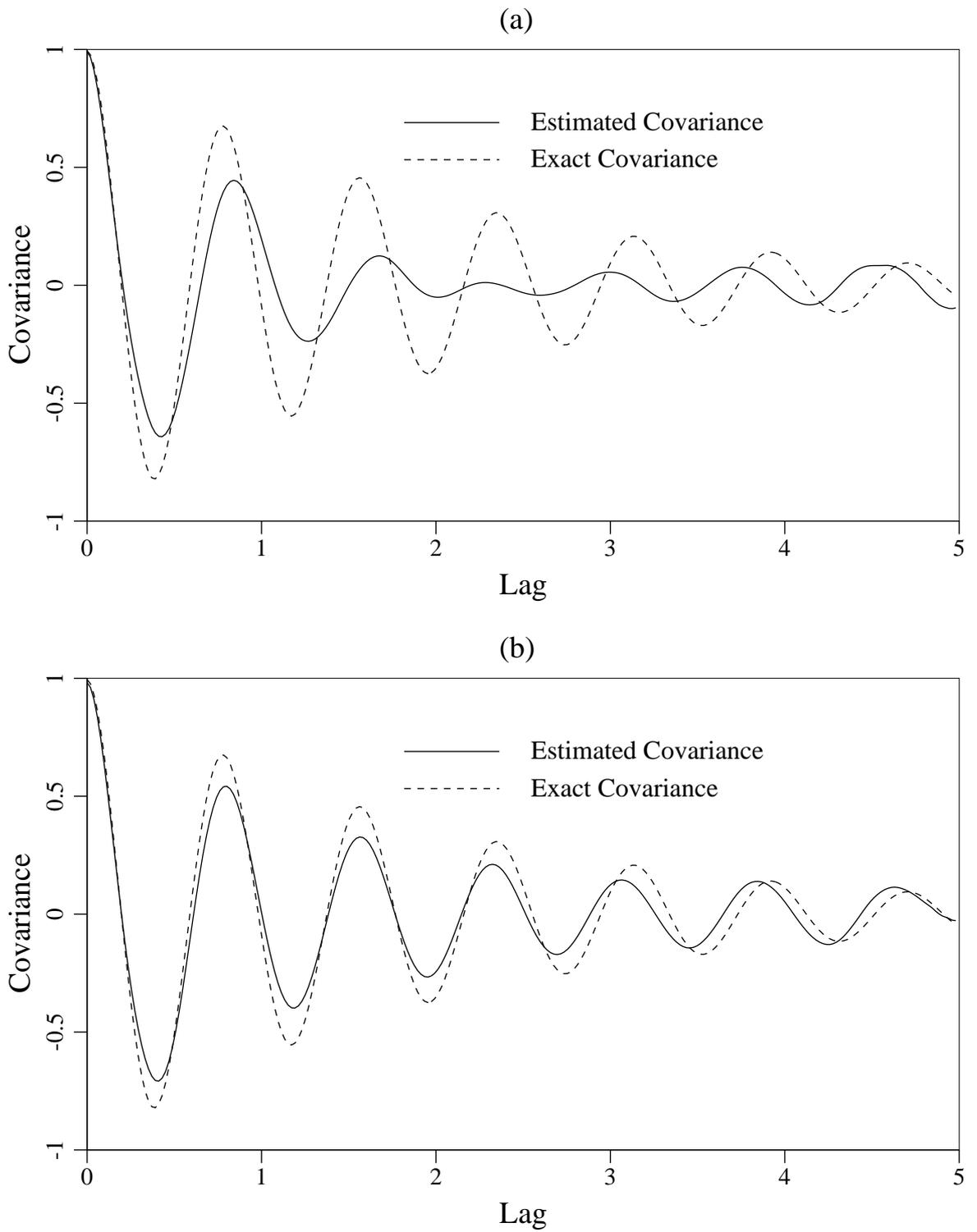


Figure 4.6 Effect of neighborhood size for a) $n = 1$ and b) $n = 2$ on the modeling of damped oscillatory noise (4.23) with $\theta = 4$ and $\omega = 8$.

Fractional Gaussian Noise

As a further demonstration of the LAS method, a self-similar process called fractional Gaussian noise was simulated. Fractional Gaussian noise (fGn) is defined by Mandelbrot and Ness [41] to be the derivative of fractional Brownian motion (fBm), obtained by averaging the fBm over a small interval δ . The resulting process has covariance and variance functions

$$B(\tau) = \frac{\sigma^2}{2\delta^{2H}} \left[|\tau + \delta|^{2H} - 2|\tau|^{2H} + |\tau - \delta|^{2H} \right], \quad (4.25)$$

$$\gamma(D) = \frac{|D + \delta|^{2H+2} - 2|D|^{2H+2} + |D - \delta|^{2H+2} - 2\delta^{2H+2}}{D^2(2H + 1)(2H + 2)\delta^{2H}}, \quad (4.26)$$

defined for $0 < H < 1$. The case $H = 0.5$ corresponds to white noise and $H \rightarrow 1$ gives $1/f$ type noise. In practice δ is taken to be equal to the smallest lag between field points ($\delta = T/2^L$) to ensure that when $H = 0.5$ (white noise), $B(\tau)$ becomes zero for all $\tau \geq T/2^L$. A sample function and its corresponding ensemble statistics are shown in Figure 4.7 for $1/f$ type noise ($H = 0.95$). The self-similar type processes have been demonstrated by Mandelbrot [42], Voss [68], and many others [48, 53, 69] to be representative of a large variety of natural forms and patterns, for example music, terrains, crop yields, and chaotic systems.

Estimates of the power spectral density function of the fractional Gaussian noise, along with the oscillatory and exponential noises discussed earlier, obtained from the LAS realizations can be found in Appendix B. As with the covariance estimates, they show excellent agreement with the exact.

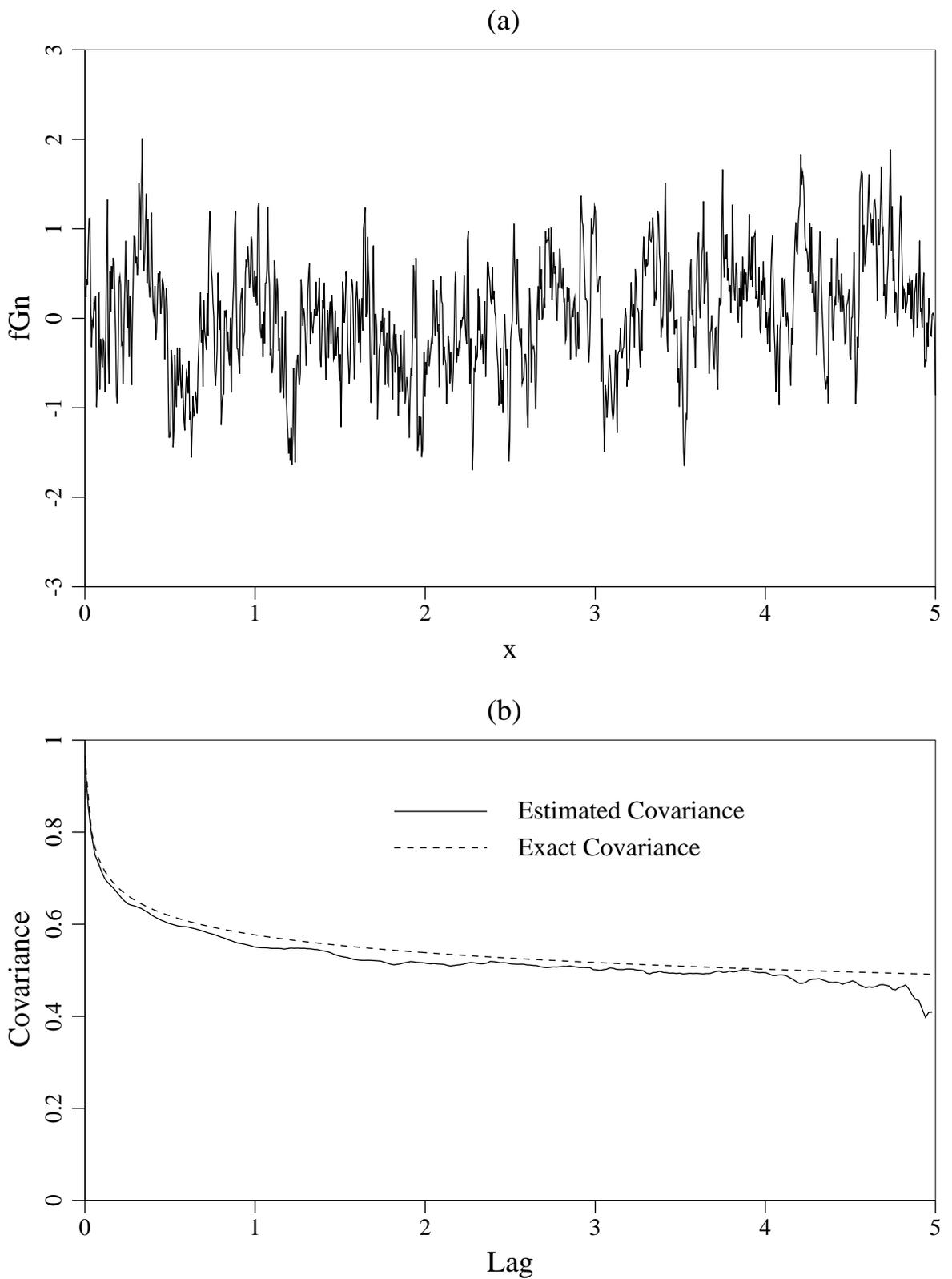


Figure 4.7 a) Sample function of $1/f$ type noise ($H = 0.95$) and (b) corresponding estimated versus exact covariance structure (averaged over 200 realizations).

4.3 Multi-Dimensional Extensions

In two dimensions, a rectangular domain is defined and the subdivision proceeds by dividing rectangles into 4 equal areas at each stage. In order to preserve the exact ‘within cell’ covariance structure, three random noises are added to three of the cell quadrants and the fourth quadrant is determined such that upwards averaging is preserved. Figure 4.8 presents the 2-D LAS scheme for the first 3 stages in which the center of each local average is marked with a different symbol for each stage. The generating relationships are,

$$\begin{aligned}
Z_1^{i+1} = Z_{2j,2k}^{i+1} &= c_{11}^{i+1} W_{1jk}^{i+1} + \sum_{\ell=1}^{n_{xy}} a_{\ell 1}^i Z_{m(\ell),n(\ell)}^i, \\
Z_2^{i+1} = Z_{2j,2k-1}^{i+1} &= c_{21}^{i+1} W_{1jk}^{i+1} + c_{22}^{i+1} W_{2jk}^{i+1} + \sum_{\ell=1}^{n_{xy}} a_{\ell 2}^i Z_{m(\ell),n(\ell)}^i, \\
Z_3^{i+1} = Z_{2j-1,2k}^{i+1} &= c_{31}^{i+1} W_{1jk}^{i+1} + c_{32}^{i+1} W_{2jk}^{i+1} + c_{33}^{i+1} W_{3jk}^{i+1} + \sum_{\ell=1}^{n_{xy}} a_{\ell 3}^i Z_{m(\ell),n(\ell)}^i, \\
Z_4^{i+1} = Z_{2j-1,2k-1}^{i+1} &= 4 Z_{jk}^i - Z_{2j,2k}^{i+1} - Z_{2j,2k-1}^{i+1} - Z_{2j-1,2k}^{i+1},
\end{aligned} \tag{4.27}$$

where W is a discrete zero-mean, unit variance Gaussian white noise and $m(\ell)$, $n(\ell)$ are indexing functions traversing in a fixed pattern the $n_{xy} = (2n_x + 1) \times (2n_y + 1)$ neighborhood of Z_{jk}^i . In this implementation, $n_x = n_y = 1$ and the boundary conditions are handled in the same fashion as for the 1-D case. The coefficients $\{a_{\ell r}^i\}$ can be calculated from the linear equations

$$\begin{aligned}
\mathbf{E} [Z_{2j,2k}^{i+1} Z_{m(p),n(p)}^i] &= \sum_{\ell=1}^{n_{xy}} a_{\ell 1}^i \mathbf{E} [Z_{m(\ell),n(\ell)}^i Z_{m(p),n(p)}^i], & p = 1, 2, \dots, n_{xy} \\
\mathbf{E} [Z_{2j,2k-1}^{i+1} Z_{m(p),n(p)}^i] &= \sum_{\ell=1}^{n_{xy}} a_{\ell 2}^i \mathbf{E} [Z_{m(\ell),n(\ell)}^i Z_{m(p),n(p)}^i], & p = 1, 2, \dots, n_{xy} \\
\mathbf{E} [Z_{2j-1,2k}^{i+1} Z_{m(p),n(p)}^i] &= \sum_{\ell=1}^{n_{xy}} a_{\ell 3}^i \mathbf{E} [Z_{m(\ell),n(\ell)}^i Z_{m(p),n(p)}^i], & p = 1, 2, \dots, n_{xy}
\end{aligned} \tag{4.28}$$

in which the matrices on the right hand sides are symmetric but no longer Toeplitz in general. The coefficient matrix \mathbf{c}^{i+1} is assumed to be lower triangular satisfying

$$\mathbf{c}^{i+1} \cdot (\mathbf{c}^{i+1})^T = \mathbf{R}, \tag{4.29}$$

where \mathbf{R} is symmetric and given by

$$R_{rs} = \text{E} [Z_r^{i+1} Z_s^{i+1}] - \sum_{\ell=1}^{n_{xy}} a_{\ell r}^i \text{E} [Z_{m(\ell),n(\ell)}^i Z_s^{i+1}] \quad r, s = 1, 2, 3 \quad (4.30)$$

using the indexing notation defined at the extreme left of (4.27). Notice that the assumption of homogeneity vastly decreases the number of coefficients that need to be calculated and stored since $\{a_{\ell r}^i\}$ and c^{i+1} become independent of position. As in the 1-D case, the coefficients need only be calculated prior to the first realization – they can be re-used in subsequent realizations reducing the effective cost of their calculation.

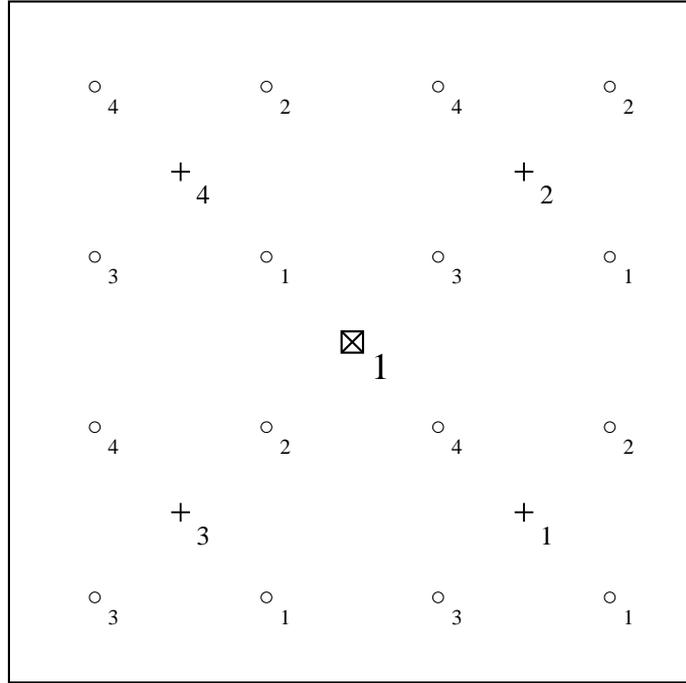


Figure 4.8 First three stages and indexing scheme of the 2-D LAS algorithm (stage 0 = \boxtimes , stage 1 = +, and stage 2 = \circ).

The expectations used in equations (4.28) to (4.30) can be determined from the two dimensional variance function of the process

$$\begin{aligned} \text{E} [Z_{jk}^i Z_{j+m,k+n}^i] &= \frac{1}{4} \sigma^2 \left[(m-1)^2 (n-1)^2 \gamma^i(m-1, n-1) - 2(m-1)^2 n^2 \gamma^i(m-1, n) + (m-1)^2 (n+1)^2 \gamma^i(m-1, n+1) \right. \\ &\quad - 2m^2 (n-1)^2 \gamma^i(m, n-1) + 4(mn)^2 \gamma^i(m, n) - 2m^2 (n+1)^2 \gamma^i(m, n+1) \\ &\quad \left. + (m+1)^2 (n-1)^2 \gamma^i(m+1, n-1) - 2(m+1)^2 n^2 \gamma^i(m+1, n) + (m+1)^2 (n+1)^2 \gamma^i(m+1, n+1) \right] \quad (4.31) \end{aligned}$$

where $\gamma^i(m, n)$ was used to denote $\gamma(mT_1^i, nT_2^i)$, T_1^i and T_2^i being the dimensions of the individual averaging rectangles at stage i . For a quadrant symmetric covariance structure, $\gamma(\cdot)$ is defined by Vanmarcke [25] to be

$$\gamma(D_1, D_2) = \left(\frac{1}{\sigma_{D_1 D_2}} \right)^2 \int_{-D_1}^{D_1} \int_{-D_2}^{D_2} (|D_1| - |\tau_1|)(|D_2| - |\tau_2|) B(\tau_1, \tau_2) d\tau_1 d\tau_2 \quad (4.32)$$

Figure 4.9 shows a sample function of a process with covariance function

$$B(\tau_1, \tau_2) = \sigma^2 \exp\left\{-\frac{2}{\theta} \sqrt{\tau_1^2 + \tau_2^2}\right\}, \quad (4.33)$$

for $\theta = 4$, which was generated using the approximate variance function

$$\gamma(D_1, D_2) = \frac{1}{2} \left[\gamma(D_2) \gamma(D_1|D_2) + \gamma(D_1) \gamma(D_2|D_1) \right], \quad (4.34)$$

where,

$$\gamma(D_i) = \left[1 + \left(\frac{D_i}{\theta} \right)^{\frac{3}{2}} \right]^{-\frac{2}{3}}, \quad (4.35)$$

$$\gamma(D_i|D_j) = \left[1 + \left(\frac{D_i}{\theta_j^i} \right)^{\frac{3}{2}} \right]^{-\frac{2}{3}}, \quad (4.36)$$

$$\theta_j^i = s_i \left[c_\alpha + (1 - c_\alpha) \exp \left\{ - \left(\frac{D_j}{\theta c_\alpha} \right)^2 \right\} \right]. \quad (4.37)$$

For the exponential covariance function (4.33), the value of c_α should be taken as $\frac{\pi}{2}$. Other forms of this approximate 2-D variance function are given by Vanmarcke [25, Chap. 6]. The estimated covariances along three different directions are seen in Figure 4.11 to show very good agreement with the exact (4.33). Figure 4.10 shows a sample function of the same process for $\theta = \frac{1}{2}$ and its corresponding estimated statistics are shown in Figures 4.12.

Although the within-cell covariance structure is reflected exactly by the LAS method, the overall statistics of anisotropic processes are rather poorly preserved. The generated field tends to become isotropic with a scale of fluctuation equal to the minimum scale specified. At this time, it is better to create the anisotropy through post-processing of an isotropic field: generate an isotropic field using a single scale of fluctuation and stretch the resulting field in the direction of the other directional scale of fluctuation to obtain an ellipsoidal correlation structure.

In three dimensions, the LAS method involves recursively subdividing rectangular parallelepipeds into 8 equal volumes at each stage. The generating relationships are essentially the same as in the 2-D case except now 7 random noises are used in the subdivision of each parent volume at each stage

$$\mathcal{Z}_s^{i+1} = \sum_{r=1}^s c_{rs}^{i+1} W_{sjkl}^{i+1} + \sum_{\ell=1}^{n_{xyz}} a_{\ell s}^i \mathcal{Z}_{m(\ell),n(\ell),p(\ell)}^i \quad s = 1, 2, \dots, 7 \quad (4.38)$$

$$Z_8^{i+1} = 8Z_{jkl}^i - \sum_{s=1}^7 Z_s^{i+1}, \quad (4.39)$$

in which Z_s^{i+1} denotes a particular octant of the subdivided cell centered at Z_{jkl}^i . For a neighborhood size, n_{xyz} , of $3 \times 3 \times 3$, Figures 4.13, 4.14, and 4.15 show exact versus estimated statistics for a process having isotropic covariance

$$B(\tau_1, \tau_2, \tau_3) = \sigma^2 \exp \left\{ -\frac{2}{\theta} \left(|\tau_1| + \sqrt{\tau_2^2 + \tau_3^2} \right) \right\}. \quad (4.40)$$

Notice that (4.40) has a partially separable form. This is not necessary but is a form used in Chapter 7 for modeling of soil. The approximate variance function corresponding to (4.40) is given by

$$\gamma(D_1, D_2, D_3) = \gamma(D_1)\gamma(D_2, D_3), \quad (4.41)$$

where $\gamma(D_2, D_3)$ is given by (4.34) with appropriate changes in subscripts and $\gamma(D_1)$ by (4.35).

When compared to Figure 4.13, Figure 4.14 illustrates the rate of convergence of the estimated statistics to the exact as the number of realizations is increased.

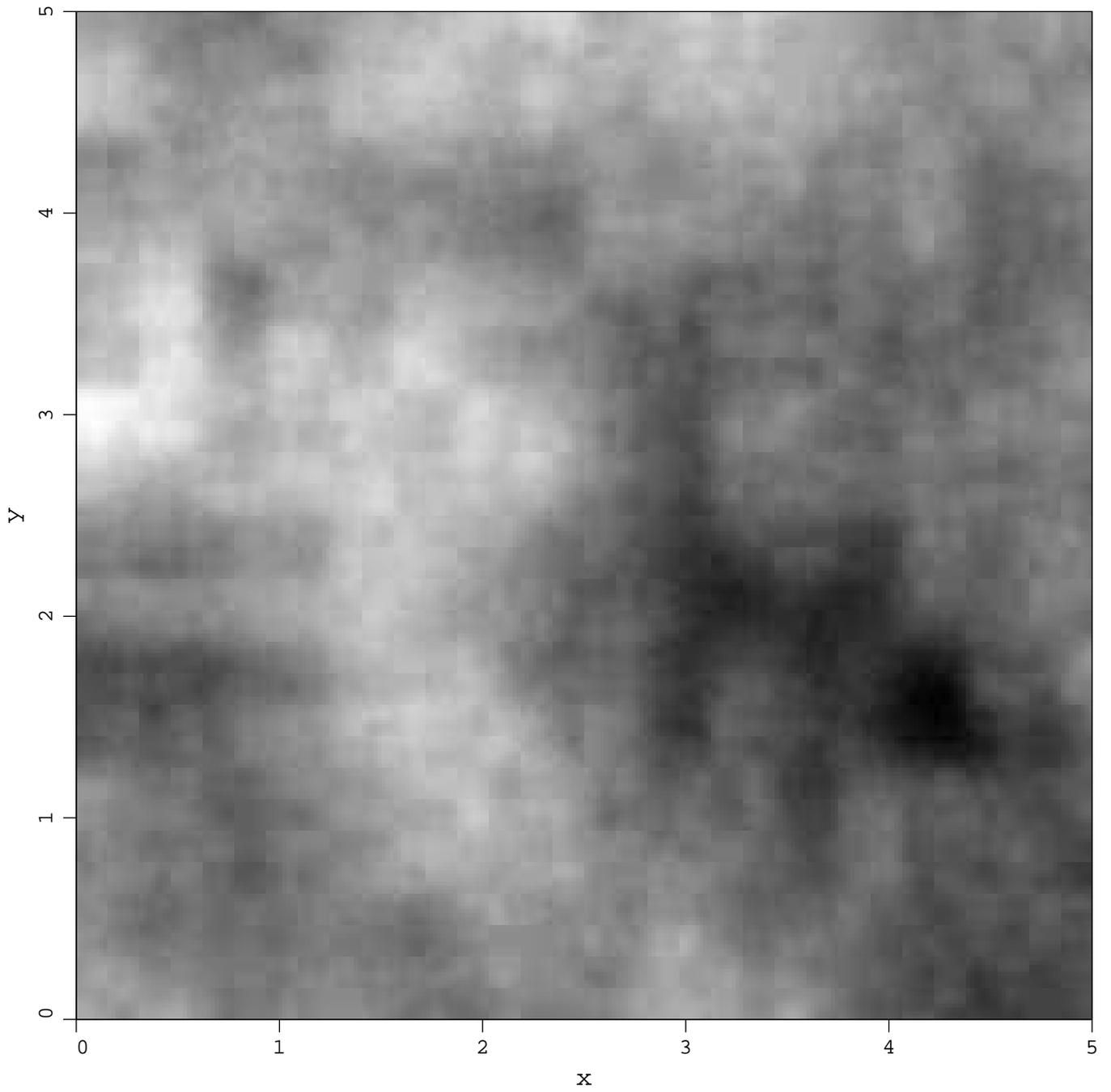


Figure 4.9 Sample function of the 2-D LAS generated process given by (4.33) with $\theta = 4$.

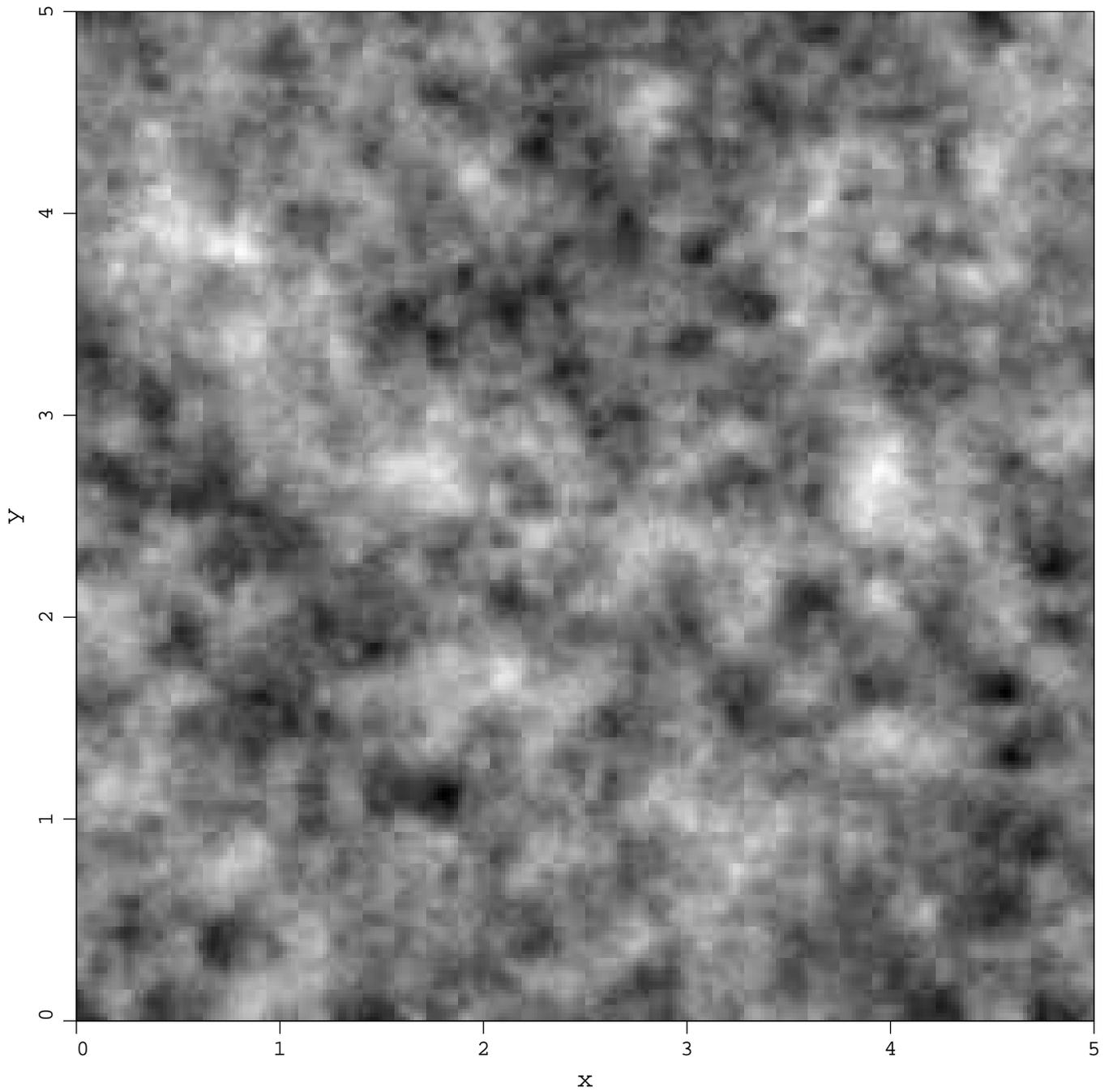


Figure 4.10 Sample function of the 2-D LAS generated process given by (4.33) with $\theta = \frac{1}{2}$.

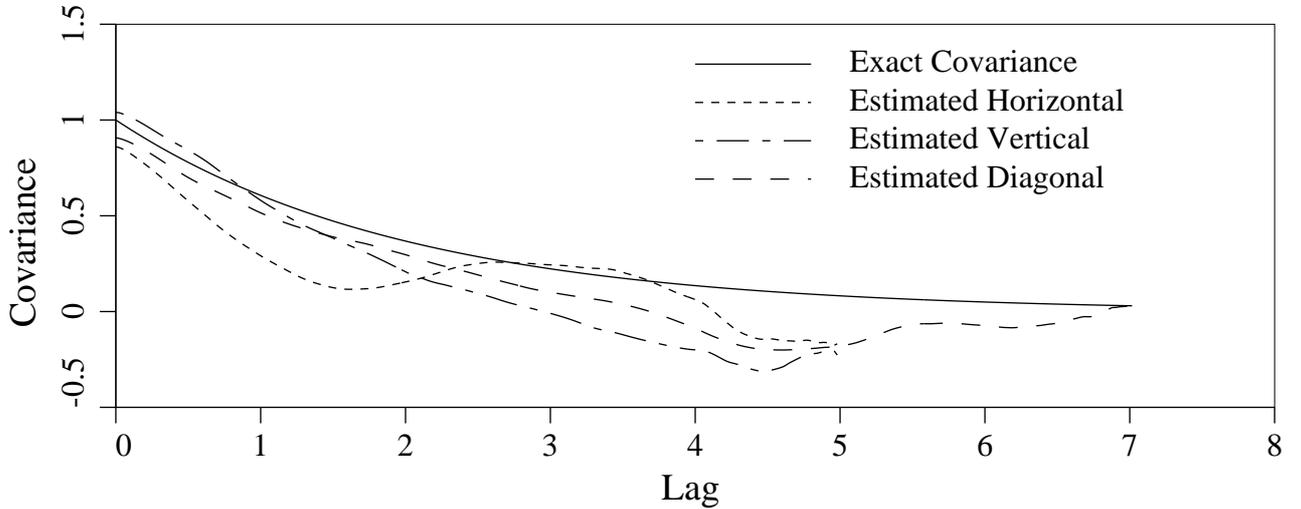


Figure 4.11 Comparison of estimated and exact covariance structure of the LAS generated 2-D process (4.33), with $\theta = 4$, averaged over 10 realizations.

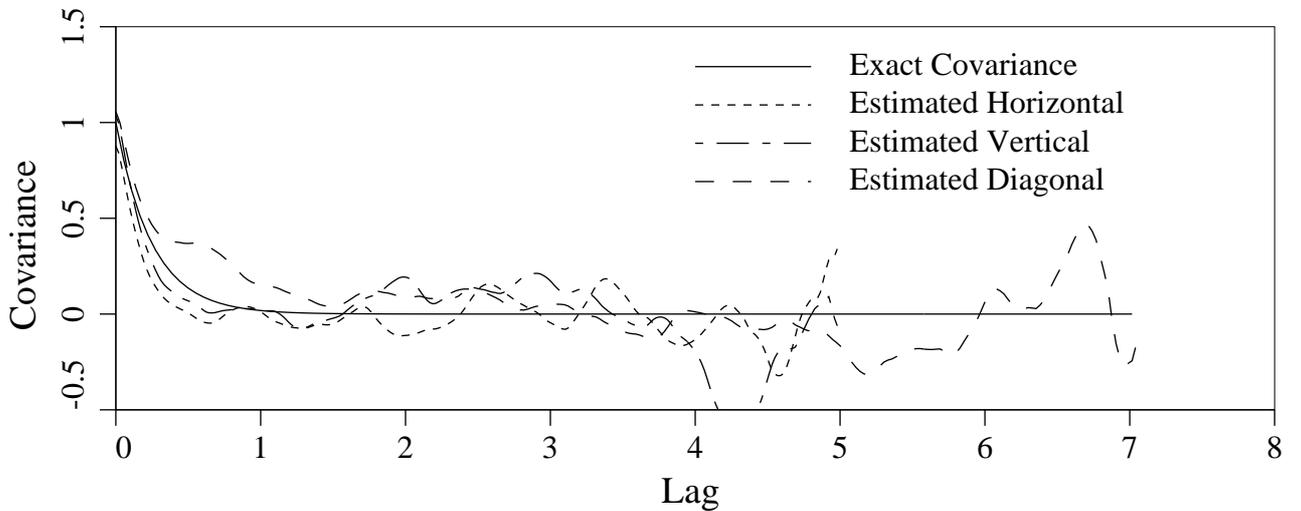


Figure 4.12 Comparison of estimated and exact covariance structure of the LAS generated 2-D process (4.33), with $\theta = \frac{1}{2}$, averaged over 10 realizations.

Implementation and Efficiency

In order to calculate stage $i+1$ values, the values at stage i must be known. This implies that storage must be provided for at least $1.5N$ values where $N = 2^L$ is the desired resolution of the process in one dimension. The implementation described herein stores all the previous stages, a storage requirement of $(2N - 1)$ in 1-D, $\frac{4}{3}(N \times N)$ in 2-D, and $\frac{8}{7}(N \times N \times N)$ in 3-D. This allows rapid ‘zooming out’ of the field. The coefficients $\{a^i\}$ and c^i , which must also be stored, can be efficiently calculated using LU factorization (see Equation (4.29)) and

successive backsubstitutions (see Equation (4.28)). The Toeplitz property of the matrix in Equation (4.11) was not taken advantage of for neighborhood sizes greater than 3.

The LAS method is also fairly competitive with the FFT approach discussed in Chapter 2. Table 4.1 compares times of the two methods running on a Cyber 205 (CDC) supercomputer for one, two, and three-dimensional realizations. In one dimension, using a neighborhood size of 3, LAS runs slightly faster than the FFT approach. Both methods have negligible setup times for the coefficient calculations. In two dimensions, the LAS approach took about 1.5 times longer than the FFT method and the three-dimensional case took about twice as long. It should be pointed out that these comparisons are made to the uncorrected FFT simulation. If the symmetric covariance structure is corrected in the FFT simulation by increasing the field size, then the performance of the multi-dimensional LAS method is seen to approach and even surpass that of the FFT method.

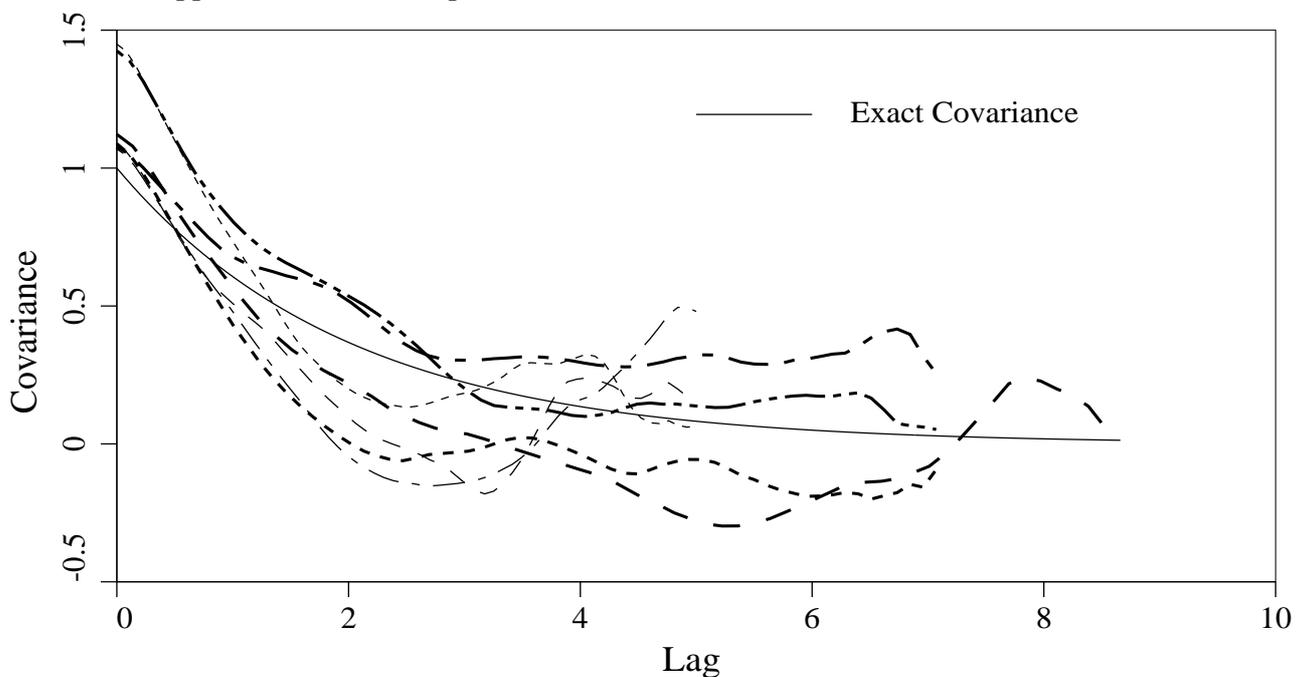


Figure 4.13 Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = 4$, averaged over 10 realizations. Estimates, shown dashed, are made in various directions through the volume.

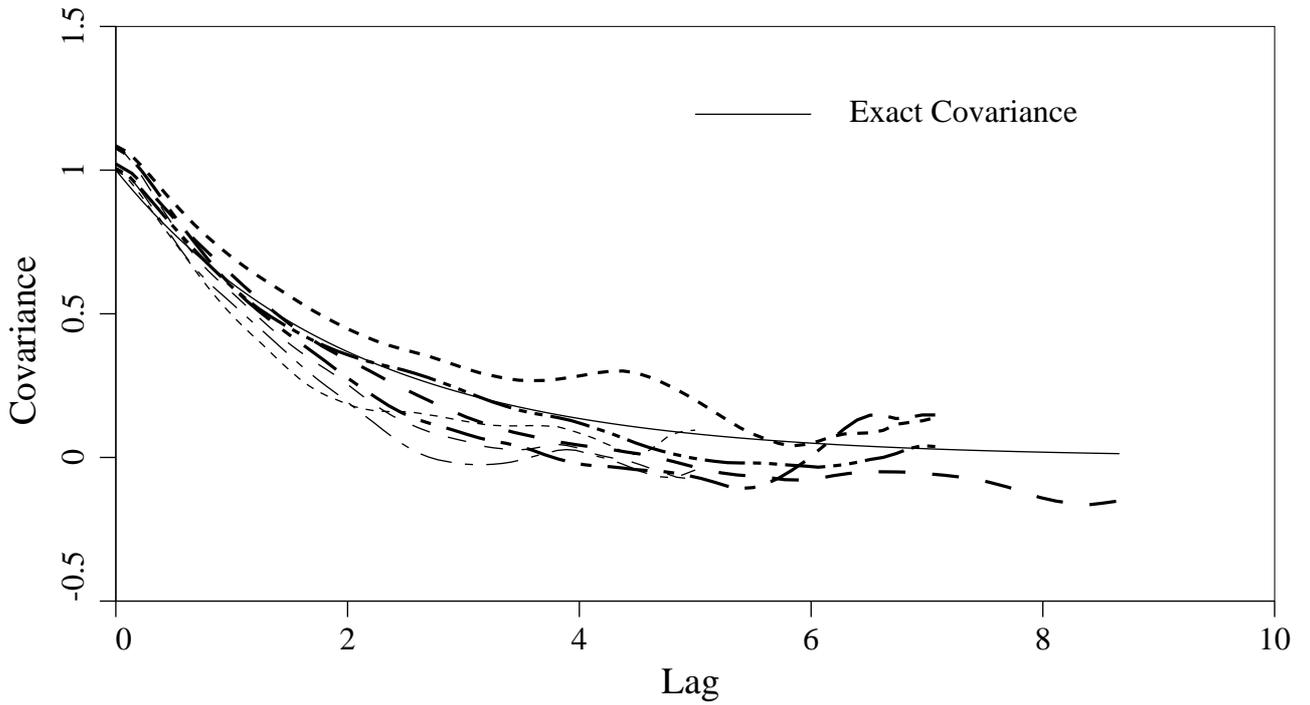


Figure 4.14 Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = 4$, averaged over 50 realizations. Estimates, shown dashed, are made in various directions through the volume.

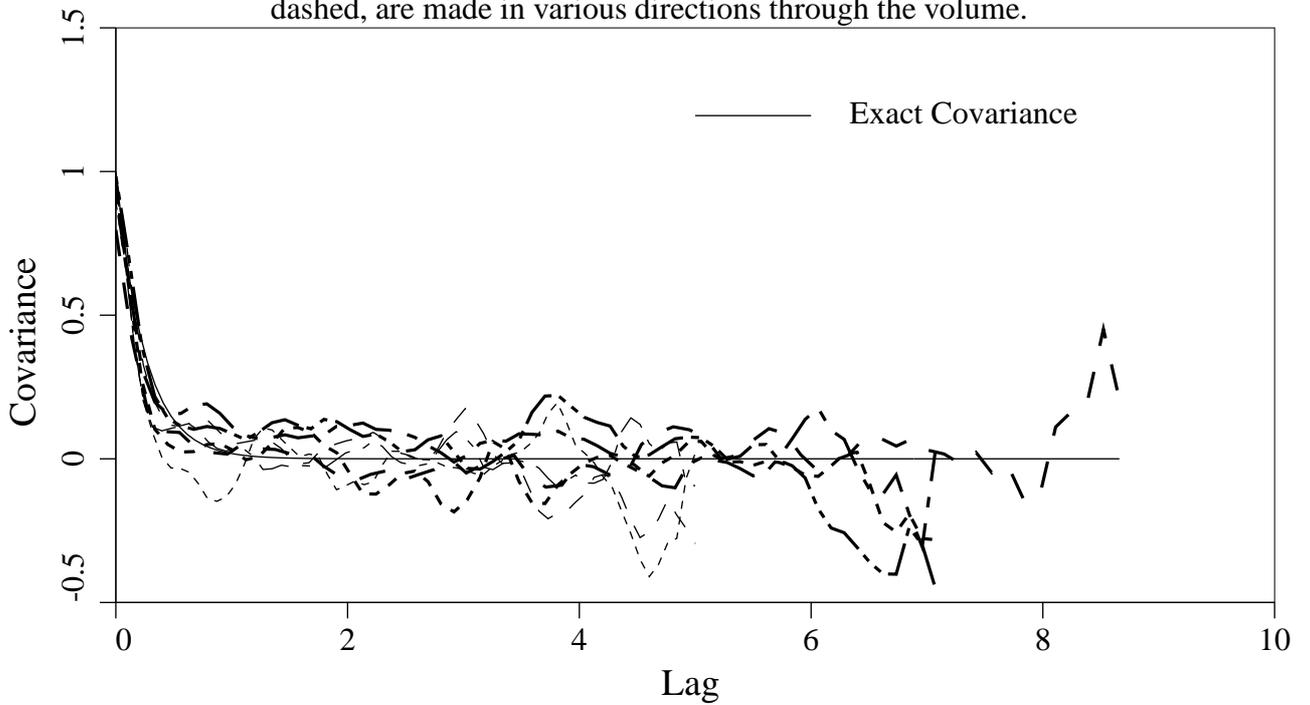


Figure 4.15 Comparison of estimated and exact covariance structure of the LAS generated 3-D process, $\theta = \frac{1}{2}$, averaged over 10 realizations. Estimates, shown dashed, are made in various directions through the volume.

Table 4.1 Comparison of execution times on a Cyber 205 super-computer.

Type	Size	Number of Simulations	Time (seconds)	
			Setup	Generation
1-D FFT	256	200	0.0013	0.1803
1-D LAS	256	200	0.0017	0.1486
2-D FFT	256×256	100	0.1265	15.2002
2-D LAS	256×256	100	0.1156	23.0100
3-D FFT	64×64×64	50	0.1517	48.7742
3-D LAS	64×64×64	50	6.1740	100.5700

Chapter 5

Space-Time Processes

5.1 Introduction

In this chapter attention will focus on spatial processes which evolve in time and thus belong to a class of space-time homogeneous Gaussian random functions. The function may represent, for example, one component of earthquake ground motion spatially distributed over some area. Improved dynamic analyses of multi-support structures require more than a single ground motion time history as input and an appropriate space-time process may be used to this end. For instance, the motion experienced by one pier of a bridge is likely to be somewhat different than that at other piers; the further the piers are separated, the greater the likely difference in their support motions. Also high-frequency components of earthquake ground motion will tend to be weakly correlated in space and low-frequency components highly correlated. It is assumed that the spatial correlation structure as well as the spectral density function at each spatial point is known *a priori*. These functions may be obtained from random field theory combined with empirical data from dense strong-motion accelerograph arrays [1, 7, 26, 27].

Another situation arises when accelerograms are recorded at some points on the free field in the vicinity of a site of interest and one wishes to simulate the ground motion at the site. Linear estimation theory is used to simulate conditioned strong ground motions, compatible with the known time histories, at arbitrary but prescribed locations on the free-field surface.

The problem of generating two time histories with known cross-correlation is first examined. The concepts are then generalized to allow the simulation of spatially correlated time series at any number of locations. Various implementation aspects are investigated; first, the use of the FFT algorithm as a means of generating realizations of random time series, and second, the use of best linear estimation techniques to condition the field. Finally, non-stationary extensions to the model are described and implemented.

5.2 Two Point Case

Consider two points \mathbf{x}_1 and \mathbf{x}_2 on the ground separated by $\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2$. The correlation of the ground motion at these two points is assumed to be described in terms of the frequency-dependent spatial correlation function $\rho_\omega(\mathbf{r})$ that depends on the separation distance vector \mathbf{r} , and frequency ω . This function is real for quadrant-symmetric random fields [25]. As was shown in Chapter 1, the zero-mean Gaussian random motion at \mathbf{x}_1 can be expressed as a sum of independent random amplitude sinusoids at constant frequency intervals, $\Delta\omega$,

$$\mathcal{Z}_1(t) = \sum_{k=1}^K [\mathcal{A}_{1k} \cos(\omega_k t) + \mathcal{B}_{1k} \sin(\omega_k t)], \quad (5.1)$$

where \mathcal{A}_{1k} and \mathcal{B}_{1k} are zero-mean Gaussian random variables and the symbol \mathcal{Z} is used to denote the algorithmic process which is an approximation to the true process Z . The true process will be used in the evaluation of expectations. Consider the k^{th} component of $\mathcal{Z}_1(t)$,

$$\mathcal{Z}_{1k}(t) = \mathcal{A}_{1k} \cos \omega_k t + \mathcal{B}_{1k} \sin \omega_k t, \quad (5.2)$$

whose variance is

$$\begin{aligned} \sigma_{1k}^2 &= \text{Var} [Z_{1k}] = G_1(\omega_k) \Delta\omega = \text{E} [(\mathcal{A}_{1k} \cos \omega_k t + \mathcal{B}_{1k} \sin \omega_k t)^2] \\ &= \frac{1}{2} \text{E} [\mathcal{A}_{1k}^2] + \frac{1}{2} \text{E} [\mathcal{B}_{1k}^2], \end{aligned} \quad (5.3)$$

where $G_1(\omega)$ is the one-sided power spectral density function associated with \mathbf{x}_1 , and

$$\text{E} [\mathcal{A}_{1k}^2] = \text{E} [\mathcal{B}_{1k}^2] = G_1(\omega_k) \Delta\omega. \quad (5.4)$$

Sample functions of $\mathcal{Z}_1(t)$ can be obtained by simulating \mathcal{A}_{1k} and \mathcal{B}_{1k} as independent normally distributed variables with mean zero and variances given by (5.4).

Consider now the “two point problem”, i.e., the generation of two correlated time series. At the frequency ω_k , the components of the motion at \mathbf{x}_1 and \mathbf{x}_2 are

$$\mathcal{Z}_{1k}(t) = \mathcal{A}_{1k} \cos \omega_k t + \mathcal{B}_{1k} \sin \omega_k t \quad (5.5)$$

and

$$\mathcal{Z}_{2k}(t) = \mathcal{A}_{2k} \cos \omega_k t + \mathcal{B}_{2k} \sin \omega_k t. \quad (5.6)$$

The component processes associated with disjoint frequency intervals are always uncorrelated, $\text{E} [Z_{1k} Z_{2j}] = 0$ if $j \neq k$, by virtue of the spectral representation theorem for mean square continuous, zero-mean, homogeneous random fields. The coefficient of correlation between Z_{1k} and Z_{2k} is $\rho_{\omega_k}(\mathbf{r})$ which can be formally expressed as

$$\rho_{\omega_k}(\mathbf{r}) = \frac{\text{E} [Z_{1k} Z_{2k}]}{\sigma_{1k} \sigma_{2k}} = \frac{\frac{1}{2} \text{E} [\mathcal{A}_{1k} \mathcal{A}_{2k}] + \frac{1}{2} \text{E} [\mathcal{B}_{1k} \mathcal{B}_{2k}]}{\sqrt{G_1(\omega_k) G_2(\omega_k)} \Delta\omega}, \quad (5.7)$$

and thus,

$$\mathbb{E} [\mathcal{A}_{1k}\mathcal{A}_{2k}] = \mathbb{E} [\mathcal{B}_{1k}\mathcal{B}_{2k}] = \rho_{\omega_k}(\mathbf{r})\sqrt{G_1(\omega_k)G_2(\omega_k)}\Delta\omega \equiv C_{12}(\omega_k), \quad (5.8)$$

where, for a strictly homogeneous Gaussian random field, $G_1(\omega) = G_2(\omega)$. The distinction between spectral density functions will be retained throughout the chapter to accommodate “nearly homogeneous” approximations and certain limiting cases such as spatial independence. The joint probability density function corresponding to (5.8) is the bivariate normal pdf,

$$f_{\mathcal{A}_1\mathcal{A}_2}(a_1, a_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{a_1^2}{\sigma_1^2} - 2\rho\frac{a_1a_2}{\sigma_1\sigma_2} + \frac{a_2^2}{\sigma_2^2} \right] \right\}, \quad (5.9)$$

where we used \mathcal{A}_1 , σ_1 , and ρ in place of \mathcal{A}_{1k} , σ_{1k} , and $\rho_{\omega_k}(\mathbf{r})$ for simplicity. Associated with equation (5.9) is the conditional density function

$$f_{\mathcal{A}_2|\mathcal{A}_1}(a_2|a_1) = \frac{1}{\sigma_2\sqrt{2\pi(1-\rho^2)}} \exp \left\{ -\frac{1}{2} \left[\frac{a_2 - (\rho\frac{\sigma_2}{\sigma_1})a_1}{\sigma_2\sqrt{1-\rho^2}} \right]^2 \right\}, \quad (5.10)$$

which is equivalent to $f_{\mathcal{B}_2|\mathcal{B}_1}$. Notice that when $\rho_{\omega_k}(\mathbf{r}) = 0$, (5.10) becomes the marginal distribution of \mathcal{A}_{2k} , as expected. When $\rho_{\omega_k}(\mathbf{r}) = 1$, (5.10) becomes a dirac delta function implying that $\mathcal{A}_{2k} = \mathcal{A}_{1k}$ in the perfect-correlation case. The two correlated time series can now be obtained by first generating realizations for \mathcal{A}_{1k} and \mathcal{B}_{1k} (with mean zero and variance σ_{1k}^2) and then generating \mathcal{A}_{2k} and \mathcal{B}_{2k} using the conditional distribution given by (5.10).

5.3 Multiple Point Case

Equation (5.1) can be generalized to describe the time histories at a number of locations \mathbf{x}_i ,

$$\mathcal{Z}_i(t) = \sum_{k=1}^K [\mathcal{A}_{ik} \cos(\omega_k t) + \mathcal{B}_{ik} \sin(\omega_k t)]. \quad (5.11)$$

Again focusing on the k^{th} component, the random coefficients \mathcal{A}_{ik} and \mathcal{B}_{ik} have statistics

$$\mathbb{E} [\mathcal{A}_{ik}] = \mathbb{E} [\mathcal{B}_{ik}] = 0, \quad (5.12)$$

$$\mathbb{E} [\mathcal{A}_{ik}^2] = \mathbb{E} [\mathcal{B}_{ik}^2] = G_i(\omega_k)\Delta\omega, \quad (5.13)$$

$$\mathbb{E} [\mathcal{A}_{ik}\mathcal{A}_{jk}] = \mathbb{E} [\mathcal{B}_{ik}\mathcal{B}_{jk}] = \rho_{\omega_k}(\mathbf{r}_{ij})\Delta\omega\sqrt{G_i(\omega_k)G_j(\omega_k)} \equiv C_{ij}(\omega_k), \quad (5.14)$$

where $\mathbf{r}_{ij} = \mathbf{x}_i - \mathbf{x}_j$. The covariance matrix $C_{ij}(\omega_k)$ is positive definite and can be diagonalized to the set of positive eigenvalues ψ_{mk} through the use of an orthogonal transformation matrix \mathbf{Q} . Using the summation convention, the defining equations are

$$Q_{il}C_{ij}(\omega_k)Q_{jm} = \psi_{mk}\delta_{lm}, \quad (5.15)$$

$$Q_{il}Q_{ij} = \delta_{lj}, \quad (5.16)$$

where δ_{lj} is the kronecker delta, and summation is suspended on underscored indices. The matrix \mathbf{Q} also relates the variables \mathcal{A}_{ik} and \mathcal{B}_{ik} to their uncorrelated counterparts U_{mk} and V_{mk} ,

$$\mathcal{A}_{ik} = Q_{im}U_{mk}, \quad (5.17)$$

$$\mathcal{B}_{ik} = Q_{im}V_{mk}. \quad (5.18)$$

Substituting (5.17) and (5.18) into (5.14) and using (5.15) it is easily seen that

$$\text{Cov} [U_{ik}, U_{jk}] = \begin{cases} 0 & \text{if } i \neq j \\ \psi_{ik} & \text{if } i = j \end{cases} \quad (5.19)$$

and similarly for V_{ik} . Thus $\psi_{ik} = \text{Var} [U_{ik}]$, and $[U_{ik}, U_{jk}]$, $[V_{ik}, V_{jk}]$ represent independent sets of uncorrelated normally distributed random variables with mean zero and variance ψ_{ik} . Using this information, realizations of U_{ik} and V_{ik} can be generated and the resulting values used in (5.17) and (5.18) to obtain \mathcal{A}_{ik} and \mathcal{B}_{ik} .

Equation (5.11) can be efficiently evaluated using the Fast Fourier Transform method. If $\mathcal{Z}_i(t)$ is to be real, certain symmetry conditions must be applied to the coefficients as shown in Chapter 2,

$$\mathcal{A}_{i,K-k+2} = \mathcal{A}_{ik}, \quad \mathcal{B}_{i,K-k+2} = -\mathcal{B}_{ik}, \quad \text{for } k = 2, 3, \dots, \frac{K}{2} \quad (5.20)$$

in which case the ensemble point variance is better preserved if one takes,

$$C_{ij}(\omega_k) = \begin{cases} \rho_{\omega_k}(\mathbf{r}_{ij})\Delta\omega\sqrt{G_i(\omega_k)G_j(\omega_k)} & \text{for } k = 1, 1 + \frac{K}{2} \\ \frac{1}{4}\rho_{\omega_k}(\mathbf{r}_{ij})\Delta\omega\sqrt{\{G_i(\omega_k) + G_i(\omega_k^*)\}\{G_j(\omega_k) + G_j(\omega_k^*)\}} & \text{for } k = 2, 3, \dots, \frac{K}{2} \end{cases} \quad (5.21)$$

where

$$\omega_k = \frac{2\pi(k-1)}{K\Delta t}, \quad \omega_k^* = \frac{2\pi(K-k+1)}{K\Delta t}, \quad \Delta\omega = \frac{2\pi}{K\Delta t}.$$

The covariance matrix is only evaluated for frequencies up to $\omega_{1+K/2} = \pi/\Delta t$. After diagonalizing and generating realizations for their uncorrelated counterparts, the coefficients \mathcal{A}_{ik} and \mathcal{B}_{ik} are determined using (5.17) and (5.18) for $k = 1, 2, \dots, 1 + K/2$, and the remainder found using the symmetries (5.20). Applying an inverse FFT will yield a set of stationary time histories representing a realization of the unconditioned field of ground motions.

5.4 Conditioning the Field

The procedure required to produce a random field in which some of the time histories are known and the others are conditioned on this knowledge is investigated in this section. It can be accomplished through the theory of best linear estimation [36, 37, 46, 70], referred to as ‘kriging’ in the context of geostatistics [32]. The quantity Z herein represents any scalar field value of interest, not necessarily a function of time (as will be shown later, it actually represents the Fourier coefficients derived in the previous sections).

Consider a discrete set of field points \mathbf{x}_ν composed of two disjoint subsets \mathbf{x}_α and \mathbf{x}_β . The subset \mathbf{x}_α , $\alpha = \{1, 2, \dots, n_\alpha\}$ is made up of points at which the field process $Z_\alpha(\mathbf{x}_\alpha)$ is known, whereas the subset \mathbf{x}_β , $\beta = \{1 + n_\alpha, 2 + n_\alpha, \dots, n_\beta + n_\alpha\}$ contains points at which the field process $Z_\beta(\mathbf{x}_\beta)$ is unknown, to be estimated using the following procedure. The subscript on Z indicates that it is not a continuous function in space, but is defined at the discrete points \mathbf{x}_ν . The best linear unbiased estimate of $Z_\beta(\mathbf{x}_\beta)$, $Z_\beta^*(\mathbf{x}_\beta)$, can be obtained by kriging

$$Z_\beta^*(\mathbf{x}_\beta) = \sum_{\alpha=1}^{n_\alpha} \eta_{\alpha\beta} Z_\alpha(\mathbf{x}_\alpha), \quad (5.22)$$

where $\eta_{\alpha\beta}$ is the set of kriging weights to be determined.

The mean of the field process $E [Z(\mathbf{x})] = m(\mathbf{x})$ can be estimated from the known data and expressed as a polynomial,

$$m(\mathbf{x}) = \sum_{i=0}^d a_i f_i(\mathbf{x}). \quad (5.23)$$

This allows drift in the mean to be accounted for if its functional form is known. In the following, $m(\mathbf{x})$ is taken to be either constant, linear, or quadratic in \mathbf{x} , corresponding to 0th, 1st, or 2nd order kriging, respectively. The conditions imposed to ensure that the estimate is unbiased, $E [Z - Z^*] = 0$, are

$$\sum_{\alpha=1}^{n_\alpha} \eta_{\alpha\beta} f_i(\mathbf{x}_\alpha) = f_i(\mathbf{x}_\beta), \quad i = 0, 1, \dots, d. \quad (5.24)$$

If $C(\mathbf{x}_i, \mathbf{x}_j) = E [Z_i(\mathbf{x}_i) Z_j(\mathbf{x}_j)]$ is defined as the covariance between the field processes at \mathbf{x}_i and \mathbf{x}_j , then the variance of the estimate becomes,

$$E \left[\left\{ Z_\beta(\mathbf{x}_\beta) - Z_\beta^*(\mathbf{x}_\beta) \right\}^2 \right] = C(\mathbf{x}_\beta, \mathbf{x}_\beta) - 2 \sum_{\alpha=1}^{n_\alpha} \eta_{\alpha\beta} C(\mathbf{x}_\beta, \mathbf{x}_\alpha) + \sum_{\alpha=1}^{n_\alpha} \sum_{\gamma=1}^{n_\alpha} \eta_{\alpha\beta} \eta_{\gamma\beta} C(\mathbf{x}_\alpha, \mathbf{x}_\gamma) \quad (5.25)$$

which must be minimized subject to the $(d + 1)$ non-bias constraints (5.24). Using the Lagrangian technique, a system of $(n_\alpha + d + 1)$ linear equations in the same number of unknowns

is obtained. The unknowns are the n_α weights $\eta_{\alpha\beta}$ for the given locations \mathbf{x}_β and the $(d+1)$

Lagrange parameters μ_i ,

$$\begin{aligned} \sum_{\alpha=1}^{n_\alpha} \eta_{\alpha\beta} C(\mathbf{x}_\gamma, \mathbf{x}_\alpha) - \sum_{i=0}^d \mu_i f_i(\mathbf{x}_\gamma) &= C(\mathbf{x}_\gamma, \mathbf{x}_\beta), \quad \gamma = 1, 2, \dots, n_\alpha \\ \sum_{\alpha=1}^{n_\alpha} \eta_{\alpha\beta} f_i(\mathbf{x}_\alpha) &= f_i(\mathbf{x}_\beta), \quad i = 0, 1, \dots, d. \end{aligned} \quad (5.26)$$

This system has a unique solution if and only if;

- 1) the covariance matrix C is positive definite,
- 2) the $(d+1)$ functions $f_i(\mathbf{x})$ are linearly independent on the set of n_α data (this is satisfied by the choice of constant, linear, or quadratic functions).

The system of equations (5.26) may be written in matrix form,

$$\left[K_\alpha \right] \left\{ \eta_{\alpha\beta} \right\} = \left\{ M_\beta \right\}, \quad (5.27)$$

where

$$\left[K_\alpha \right] = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \cdot & \cdot & \cdot & C_{1n_\alpha} & 1 & f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \cdot & \cdot & \cdot & f_d(\mathbf{x}_1) \\ C_{21} & C_{22} & C_{23} & \cdot & \cdot & \cdot & C_{2n_\alpha} & 1 & f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \cdot & \cdot & \cdot & f_d(\mathbf{x}_2) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ C_{n_\alpha 1} & C_{n_\alpha 2} & C_{n_\alpha 3} & \cdot & \cdot & \cdot & C_{n_\alpha n_\alpha} & 1 & f_1(\mathbf{x}_{n_\alpha}) & f_2(\mathbf{x}_{n_\alpha}) & \cdot & \cdot & \cdot & f_d(\mathbf{x}_{n_\alpha}) \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ f_1(\mathbf{x}_1) & f_1(\mathbf{x}_2) & f_1(\mathbf{x}_3) & \cdot & \cdot & \cdot & f_1(\mathbf{x}_{n_\alpha}) & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ f_2(\mathbf{x}_1) & f_2(\mathbf{x}_2) & f_2(\mathbf{x}_3) & \cdot & \cdot & \cdot & f_2(\mathbf{x}_{n_\alpha}) & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ f_d(\mathbf{x}_1) & f_d(\mathbf{x}_2) & f_d(\mathbf{x}_3) & \cdot & \cdot & \cdot & f_d(\mathbf{x}_{n_\alpha}) & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}$$

in which $C_{\alpha\gamma} = C(\mathbf{x}_\alpha, \mathbf{x}_\gamma)$ was used for convenience, and

$$\left\{ \eta_{\alpha\beta} \right\} = \begin{pmatrix} \eta_{1\beta} \\ \eta_{2\beta} \\ \eta_{3\beta} \\ \vdots \\ \eta_{n_\alpha\beta} \\ -\mu_0 \\ -\mu_1 \\ -\mu_2 \\ \vdots \\ -\mu_d \end{pmatrix} \quad \left\{ M_\beta \right\} = \begin{pmatrix} C(\mathbf{x}_1, \mathbf{x}_\beta) \\ C(\mathbf{x}_2, \mathbf{x}_\beta) \\ C(\mathbf{x}_3, \mathbf{x}_\beta) \\ \vdots \\ C(\mathbf{x}_{n_\alpha}, \mathbf{x}_\beta) \\ 1 \\ f_1(\mathbf{x}_\beta) \\ f_2(\mathbf{x}_\beta) \\ \vdots \\ f_d(\mathbf{x}_\beta) \end{pmatrix}.$$

As $[K_\alpha]$ depends solely on the covariance structure and geometry of the known points, \mathbf{x}_α , it can be inverted once and used repeatedly to find the weights $\eta_{\alpha\beta}$ for each point \mathbf{x}_β . The actual covariance matrix to be used, $C(\mathbf{x}_\alpha, \mathbf{x}_\gamma) = C_{\alpha\gamma}(\omega_k)$, is dependent on both spatial distance as well as frequency. The conditioning thus involves kriging the Fourier coefficients at each frequency ω_k as follows;

- 1) Fourier decompose the known time histories to obtain the set of known Fourier coefficients $\bar{\mathcal{A}}_{\alpha k}$ and $\bar{\mathcal{B}}_{\alpha k}$ at the points \mathbf{x}_α ,
- 2) kriging the known Fourier coefficients to obtain estimates $\bar{\mathcal{A}}_{\beta k}^*$ and $\bar{\mathcal{B}}_{\beta k}^*$ at the unknown points \mathbf{x}_β ,
- 3) generate the unconditional set of Fourier coefficients $\mathcal{A}_{\nu k}$ and $\mathcal{B}_{\nu k}$ using the procedure outlined in the previous sections for all the field points \mathbf{x}_ν ,
- 4) kriging the simulated Fourier coefficients over the set \mathbf{x}_α to obtain the set of simulated estimates $\mathcal{A}_{\beta k}^*$ and $\mathcal{B}_{\beta k}^*$ at the unknown points \mathbf{x}_β ,
- 5) compute the conditioned simulated Fourier coefficients over the set \mathbf{x}_β to be

$$\tilde{\mathcal{A}}_{\beta k} = \mathcal{A}_{\beta k} + \bar{\mathcal{A}}_{\beta k}^* - \mathcal{A}_{\beta k}^*,$$

$$\tilde{\mathcal{B}}_{\beta k} = \mathcal{B}_{\beta k} + \bar{\mathcal{B}}_{\beta k}^* - \mathcal{B}_{\beta k}^*.$$

Notice that the conditioned values are only generated over the ‘unknown’ points since the Fourier coefficients over the ‘known’ points are already determined in step 1.

In this way the set of conditioned Fourier coefficients are built up for each $\omega_k, k = 1, 2, \dots, 1 + \frac{K}{2}$ (the other half obtained using symmetry as discussed in the previous section) and then an inverse Fourier transform is applied to obtain the final conditioned sample field.

5.5 Case Studies

A computer program was designed and implemented to perform the conditional simulation of correlated earthquake ground motions. Several example runs were performed to verify:

- 1) the ability of the program to produce properly correlated stationary time histories with no field values known *a priori*,
- 2) the ability of the program to produce properly conditioned stationary time histories given one or more prescribed time histories.

The spectral density function shown in Figure 5.1 was used in all the examples. A time increment of 0.005 seconds and length of 1024 values gives time history durations of 5.115 seconds. The maximum Fourier frequency is thus $\omega_{max} = \pi/\Delta t = 628 \text{ rad/sec}$ and it can be seen that $G(\omega)$ is negligible above this value. However, $G(\omega)$ was derived independent of this criteria – it was predicted given an earthquake of magnitude 5.2 at a distance of 15 km according to a procedure outlined by Vanmarcke [65].

The correlation function used in the examples has the following simple isotropic and exponential form

$$\rho_{\omega_k}(\mathbf{r}_{ij}) = \exp \left\{ \frac{-\omega_k |\mathbf{r}_{ij}|}{2\pi cs} \right\}, \quad (5.28)$$

where c is the shear wave velocity in the medium (1500 m/sec) and s is the scale of the process (not to be confused with the scale of fluctuation θ which in this case varies with ω). Equation (5.28) is only intended to illustrate the procedures, but does show some of the characteristics of empirical correlation functions obtained from dense accelerograph array data.

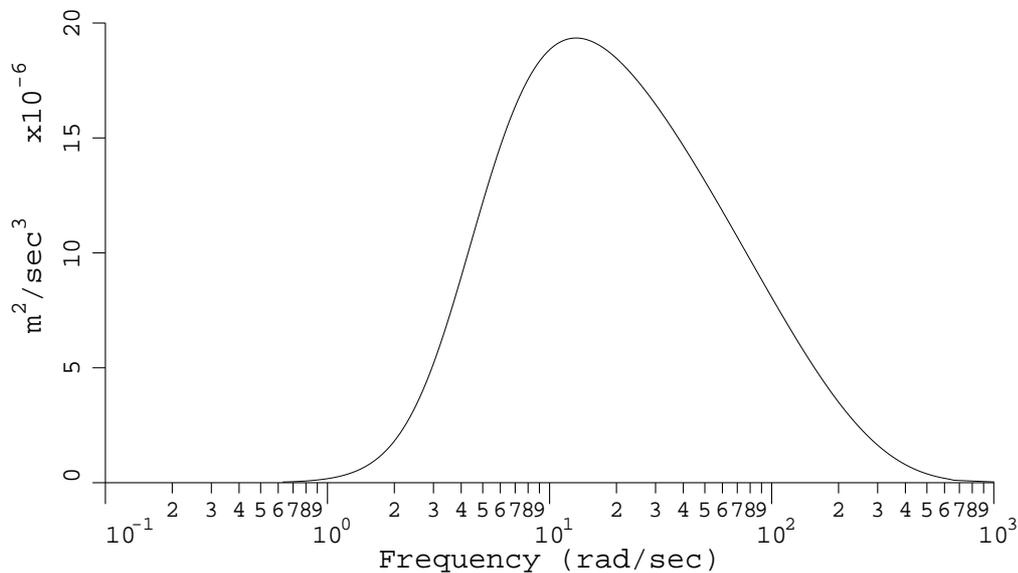


Figure 5.1 Spectral density function used in examples.

Figure 5.2 shows the first 0.6 seconds of a four-point correlated motion example, in which no sample functions are known, for gradually decreasing scales. The points are located on the vertices of a 100 metre square. For a scale about the same as the distance between points, the relatively high correlation shown is to be expected. At the other extreme, for scales negligible compared to the inter-point distances, the motions appear uncorrelated. The estimated point variance over an ensemble of realizations is within 5% of that predicted theoretically by integrating $G(\omega)$.

To verify the kriging algorithm, an extreme example was considered in which three field points were assumed to be moving sinusoidally (at a single frequency) and the fourth point's motion was to be simulated. The known points were located at (x, y) coordinates of $(0,0)$, $(0,100)$, and $(100,0)$ and the fourth unknown point at $(50,50)$. For a large scale, Figure 5.3 shows the three superimposed sinusoids followed fairly closely by the motion at the fourth point, as expected. Notice how the motion at the fourth point becomes increasingly random at smaller scales.

Now consider a more realistic three-point example, where an unknown point is located midway between two "known points" which are separated by 100 metres. The motions at the two "known points" were obtained by running the program once assuming that no points were known and using a scale $s = 1.0$. The resulting time histories shown at the top of Figure 5.4 become the known time histories when the third intermediate point is introduced. Figure 5.4 shows how the motion at the unknown point lies essentially between the two known motions when the scale is relatively large. As the scale decreases the motion at the unknown point becomes increasingly random. It should be noted, however, that because the motions at the two "known points" are so similar, and because there are two of them, the

motion at the “unknown point” appears similar even for very short scales. The relatively high shear wave velocity of 1500 m/sec used also results in higher correlation over most frequencies of interest. In soils, the shear wave velocity is much lower and so ground motions simulated in such a medium would show greater independence for the same scales and inter-point distances. These same comments could have been made about Figure 5.3.

In summary, it appears that the technique proposed here is an effective method of producing stationary conditioned correlated time histories, the first and key step in the process of producing compatible earthquake time histories at closely spaced points. The four-point kriged problem illustrated in Figure 5.3 with 1024 time steps (and thus 513 frequency steps) takes only about 10 seconds to run on a VaXstation. Of course as either of these dimensions is increased, an increase in required computer time is expected particularly as the number of field points increases. However for most practical problems the efficiency and accuracy of the algorithm is excellent.

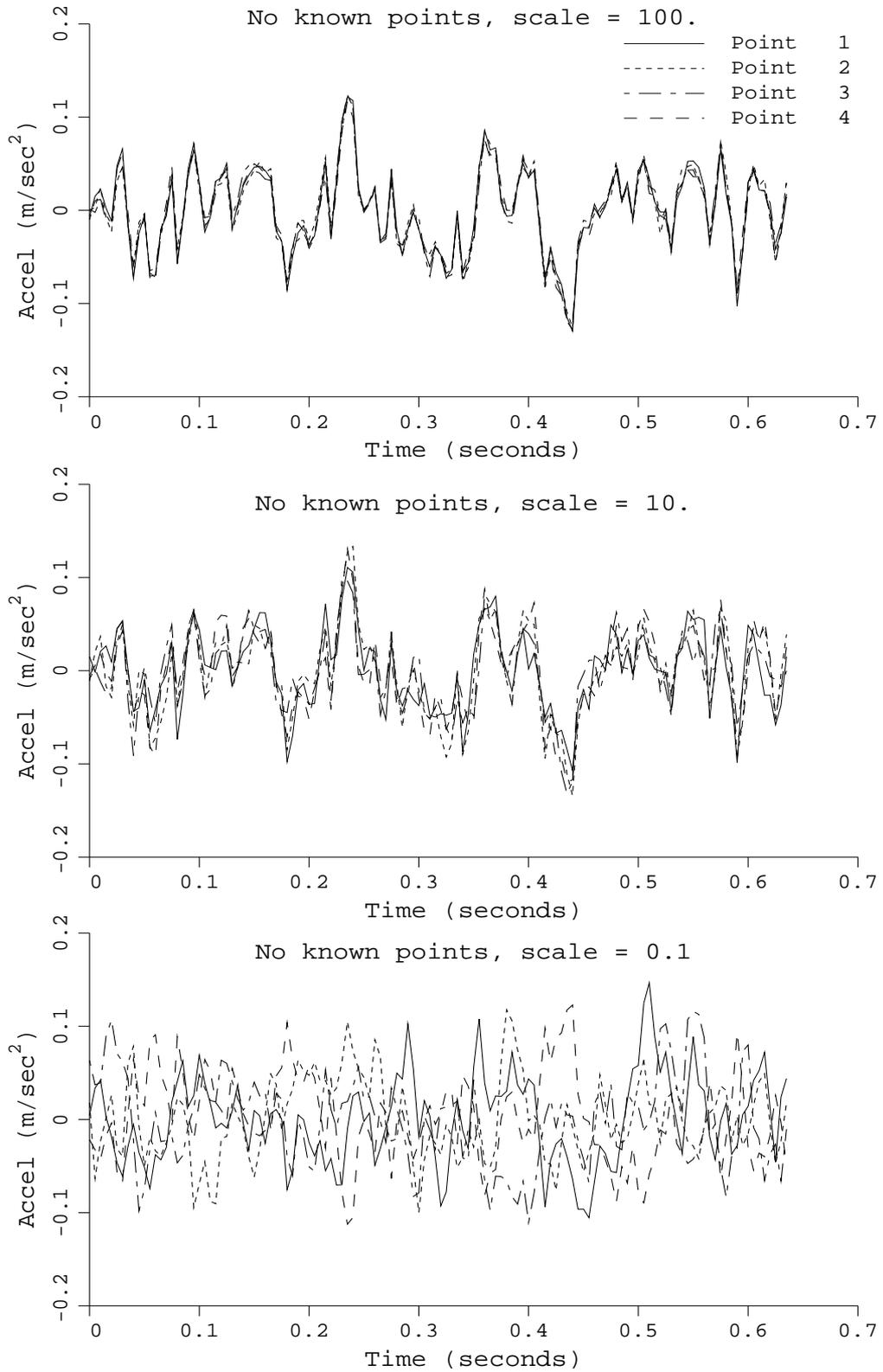


Figure 5.2 Four point unconditional time history simulation for various scales. Points are arranged on the vertices of a 100 m square.

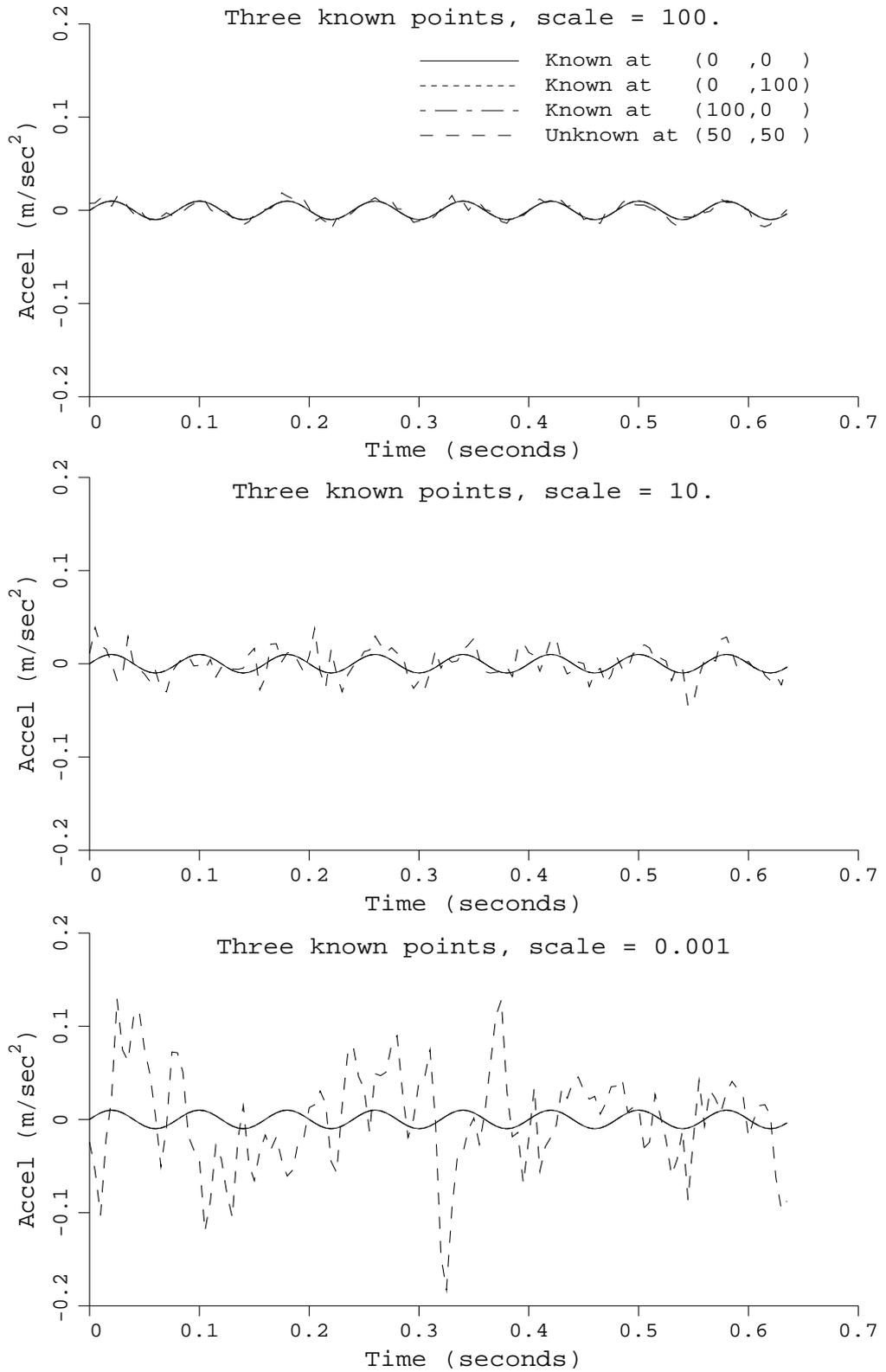


Figure 5.3 Four point conditional time history simulation for various scales with three points known to be moving sinusoidally.

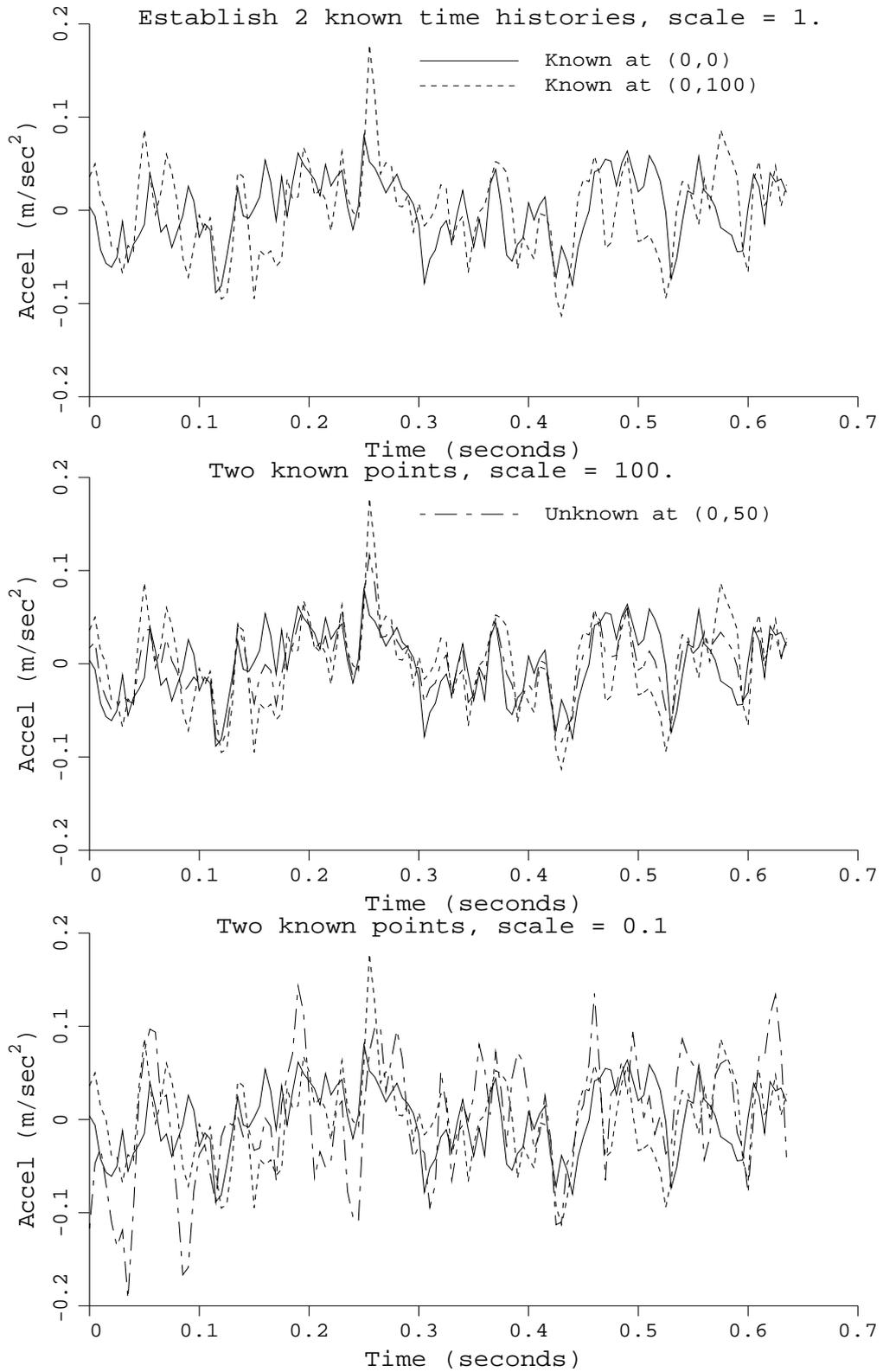


Figure 5.4 Three point conditional time history simulation at two different scales. The known motion of the two points is established in the uppermost plot.

5.6 Non-Stationary Extensions

To become a useful design tool the procedures proposed in this chapter should be able to generate more realistic earthquake ground motions. This implies that non-stationary motions in which both the frequency content and amplitude evolve with time would be desirable. Such an extension can be easily accomplished simply by simulating the time histories within a series of time windows each of which can be considered stationary. Figure 5.5 shows an earthquake motion accelerogram recorded at the Wildlife Liquefaction Array site, Imperial County, California.

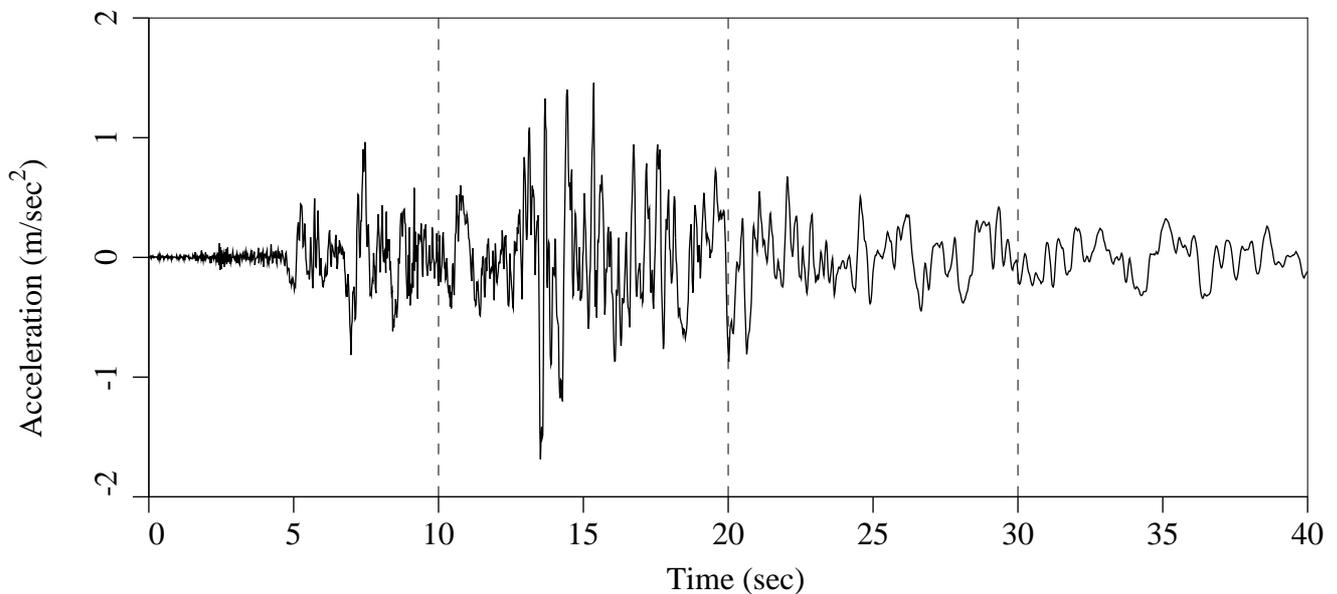


Figure 5.5 North-South component of down-hole (7.5 m) earthquake acceleration motion recorded Nov. 24, 1987 at the Wildlife Liquefaction Array site, Imperial County, California (Superstition Hills event, magnitude 6.6) showing the four subwindow stationary approximation.

The accelerogram was triggered by the Superstition Hills event (magnitude 6.6 on Nov. 24, 1987) and was measured at a depth of 7.5 m below the ground surface. Only the North-South component is shown here. This time history was arbitrarily divided into four equal segments and power spectral density functions were estimated, as shown in Figure 5.6, for the segments within each window. The power spectral density functions were estimated in two ways. The first, shown in dashed lines in Figure 5.6, involved digital spectral analysis using the FFT and a rectangular smoothing window as discussed by Blackman and Tukey [6] and Newland [50]. An 11 data-point rectangular smoothing window (± 5 data points) was used to reduce the variance of the estimate. The second method was the maximum

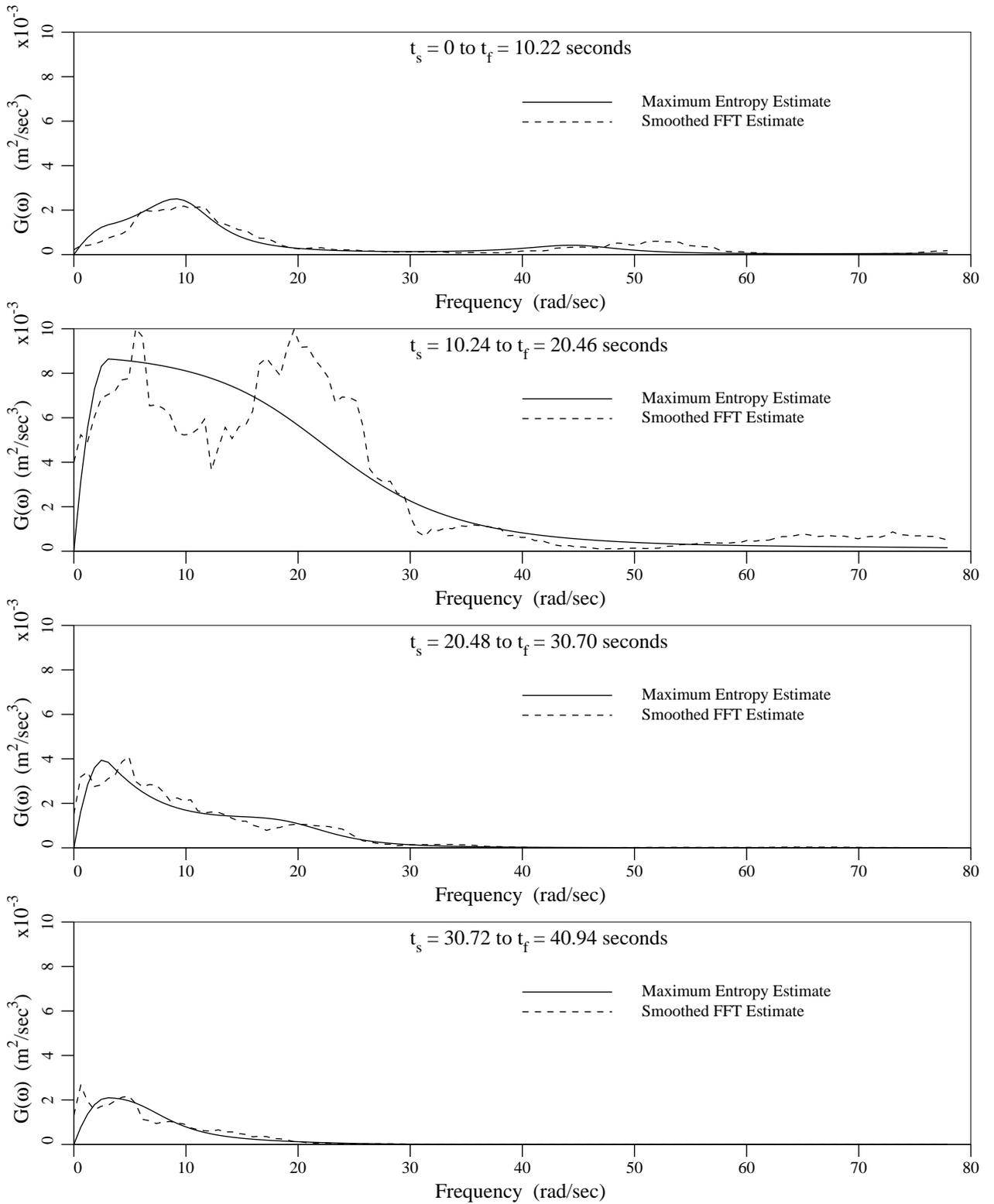


Figure 5.6 Estimates of the power spectral density function within each time window obtained using FFT and Maximum Entropy methods. entropy approach outlined by Anderson [4] and Burg [11]. The maximum entropy estimates

were found to remain relatively high at low frequencies and so were arbitrarily reduced by ignoring the first $m - 1$ frequency steps and fitting a parabola to the m^{th} frequency step that passed through the origin. Using a value of $m = 6$ (3.1 rad/sec) in Figure 5.6 gave maximum entropy estimates in good agreement with those obtained via the FFT approach. For simulation purposes, power spectral estimates which are representative over the ensemble are believed to be superior. Therefore the corrected maximum entropy estimates were used in the following simulations since they are related to maximum likelihood estimates [10].

The next step in the creation of a non-stationary time history from stationary segments is the piecing together of the segments. This must be done such that the correlation between time histories at various points in space is not effected. It turns out that this can be accomplished through a simple linear combination of the time history before the join with the time history after the join. Consider two adjacent segments of the time history, the first segment starting at time $(n_f - K_e + 1)\Delta t$ and ending at $n_f\Delta t$ and the second segment starting at time $(n_f + 1)\Delta t$ and ending at $(n_f + K_e)\Delta t$, where K_e denotes the length of each stationary segment. Calling the first segment $\mathcal{Z}^i(t_j)$ and the second segment $\mathcal{Z}^{i+1}(t_j)$, the final non-stationary time series $\mathcal{Z}(t_j)$ can be obtained as follows

$$\mathcal{Z}(t_j) = \begin{cases} \mathcal{Z}^i(t_j), & \text{if } j < n_f - n_\nu \\ w(n_f - j)\mathcal{Z}^i(t_j) + (1 - w(n_f - j))\mathcal{Z}^{i+1}(t_j), & \text{if } n_f - n_\nu \leq j \leq n_f + n_\nu \\ \mathcal{Z}^{i+1}(t_j), & \text{if } j > n_f + n_\nu \end{cases} \quad (5.29)$$

where $w(j)$ is a discrete weighting function defined over the transition region $[-n_\nu, n_\nu]$. It varies linearly such that $w(n_\nu) = 1$ and $w(-n_\nu) = 0$. Thus $\mathcal{Z}(t_{n_f})$ is just equal to the average of $\mathcal{Z}^i(t_{n_f})$ and $\mathcal{Z}^{i+1}(t_{n_f})$ and has increasing contributions from the appropriate segments as the transition boundaries are approached. One may notice that (5.29) requires the knowledge of $\mathcal{Z}^i(t_j)$ beyond the time t_{n_f} and $\mathcal{Z}^{i+1}(t_j)$ before the time t_{n_f} , both of which are outside their domains. To accommodate this, use can be made of the fact that the time segments will be formed via the FFT and so the periodicity relationships

$$\begin{aligned} \mathcal{Z}^i(t_{n_f+j}) &= \mathcal{Z}^i(t_j), & j > 0 \\ \mathcal{Z}^{i+1}(t_{n_f-j}) &= \mathcal{Z}^{i+1}(t_{n_f+K_e-j}), & j \geq 0 \end{aligned} \quad (5.30)$$

apply.

In the implementation of this non-stationary extension, the size of the transition region n_ν is chosen so that it covers a few oscillations of the motion. Although (5.29) guarantees continuity of $\mathcal{Z}(t_j)$ even when $n_\nu = 0$, such a small value will result in a discontinuous derivative. The use of larger values of n_ν will lead to a continuous derivative. A further

advantage to the transition averaging technique given above becomes apparent when one recalls that the FFT simulation method gives a covariance structure which is symmetric about a lag of $\frac{1}{2}K_e\Delta t$. Within a given segment of the simulated motion $Z^i(t_{n_f-K_{e+1}})$ is almost perfectly correlated with $Z^i(t_{n_f})$. The averaging with the independently simulated $Z^{i+1}(t_{n_f})$ will reduce this correlation. However, this also implies that the size of the windows used to create the segments should be large compared to the scale of fluctuation of the process.

The technique of generating segments longer than required and ignoring the excess could also be used to reduce this symmetric covariance effect when generating unconditioned simulations. However, when generating conditioned simulations, this correction cannot be used since the known time histories cannot be arbitrarily lengthened within each segment. Thus, such a correcting procedure was not employed in this implementation – it was assumed that the reduction in covariance at lags $K_e\Delta t$ due to the averaging of adjacent independent segments of the time series was sufficient to reduce the error to tolerable levels.

Figure 5.7 shows an unconditioned four-point non-stationary simulation in which the points are arranged on the vertices of a 100 m square and a scale $s = 100$ used. Similarly, Figure 5.8 shows an unconditioned four-point non-stationary simulation using a scale $s = 10$. In both cases, the shear wave velocity used was $c = 130$ m/sec in (5.28) which was the average in-situ velocity measured at the Wildlife site at a depth of 7.5 m. The non-stationary nature of the simulated records is clearly evident, as is the increasing independence of the four motions as the scale decreases.

To illustrate the conditional simulation of the non-stationary motions, the original recorded earthquake motion was taken as known and a second conditional motion was simulated as shown in Figures 5.9 and 5.10 for two different scales. The tendency of the two motions to become more similar as the scale increases is clearly evident. However, it can also be seen that the amplitude of the simulated motion is largely governed by its spectral density function. If this differs substantially from that of the known motion (which it does in the first 5 seconds of the record), then there will be amplitude scaling of the simulated motion that does not disappear even at very high scales s (for which the patterns may be identical but the amplitudes are different). This phenomena is reduced as the number of ‘known’ records increases relative to the number of ‘unknown’ points (see for example Figure 5.3) and can also be cured by increasing the number of time window subdivisions and/or aligning the window boundaries with natural changes in spectral power.

A potentially important conclusion that can be reached on the basis of Figures 5.9 and 5.10 is that the method can be extended to non-homogeneous fields in which the (evolutionary) spectral density function is changed from point to point in space. This may allow the simulation of motion on a soil layer conditioned by ground motions recorded on bedrock or below the surface. These observations have yet to be verified and are beyond the scope of the present work.

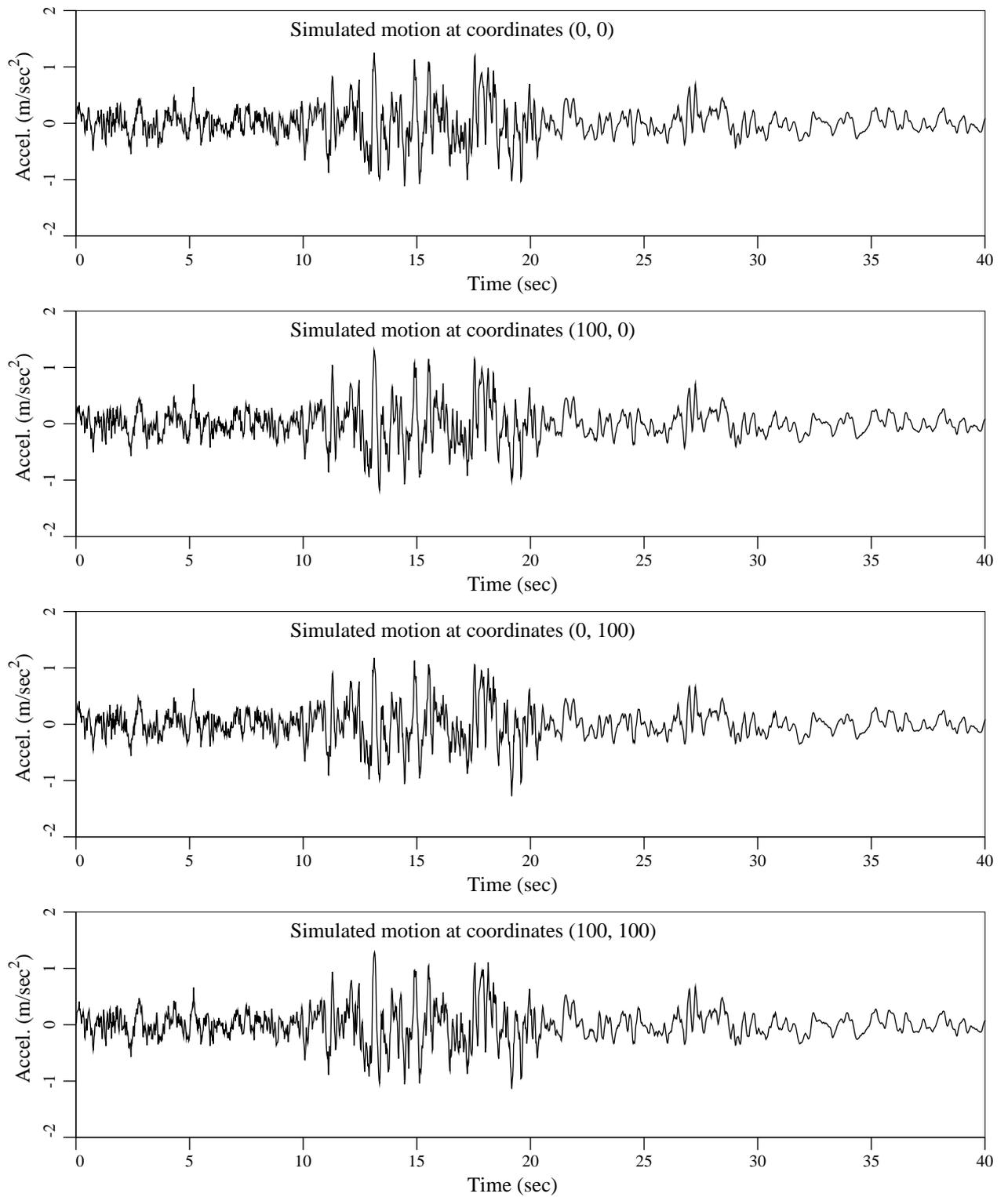


Figure 5.7 Four point unconditioned non-stationary simulation with points arranged on the vertices of a 100 m square for scale $s = 100$.

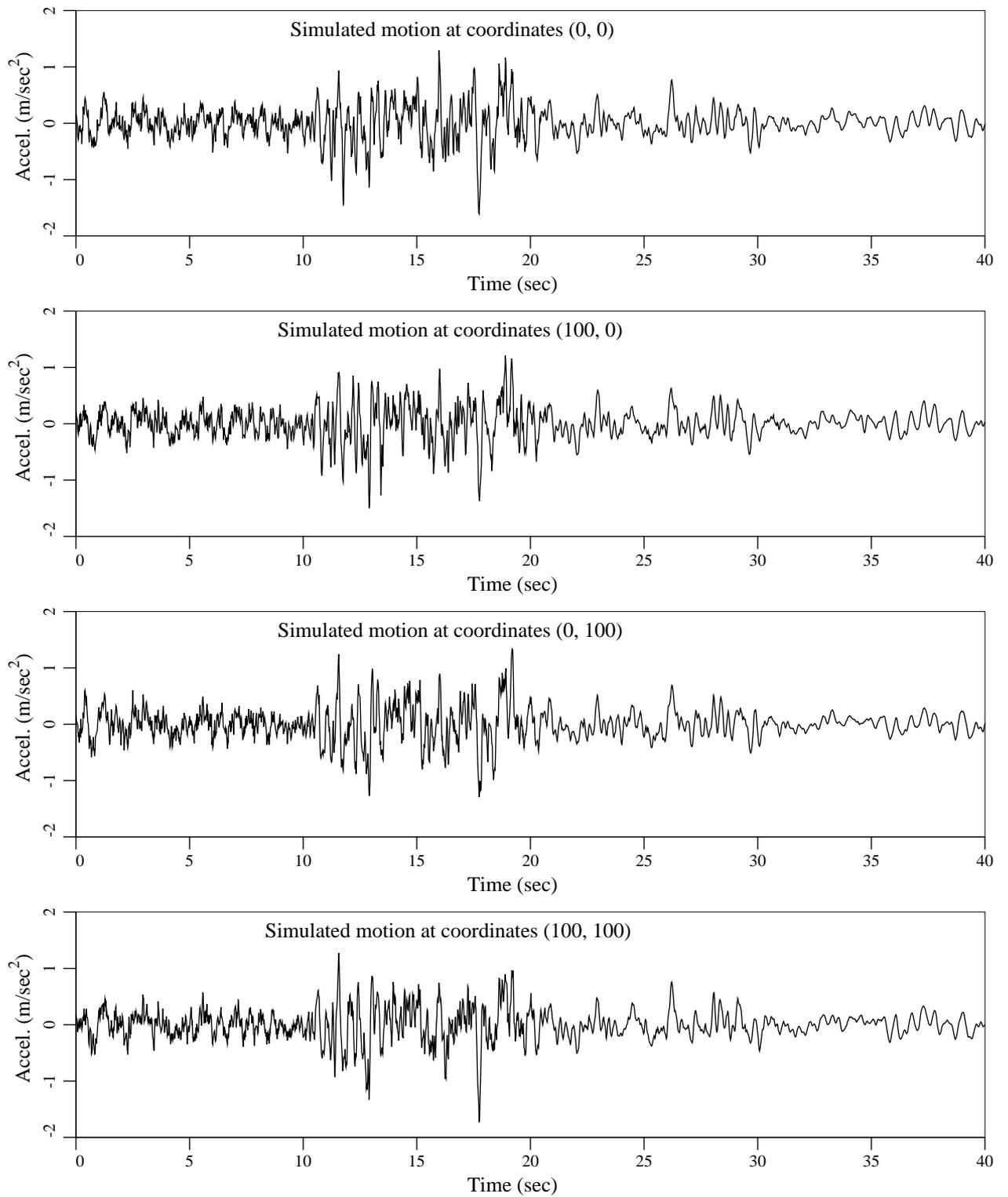


Figure 5.8 Four point unconditioned non-stationary simulation with points arranged on the vertices of a 100 m square for scale $s = 10$.

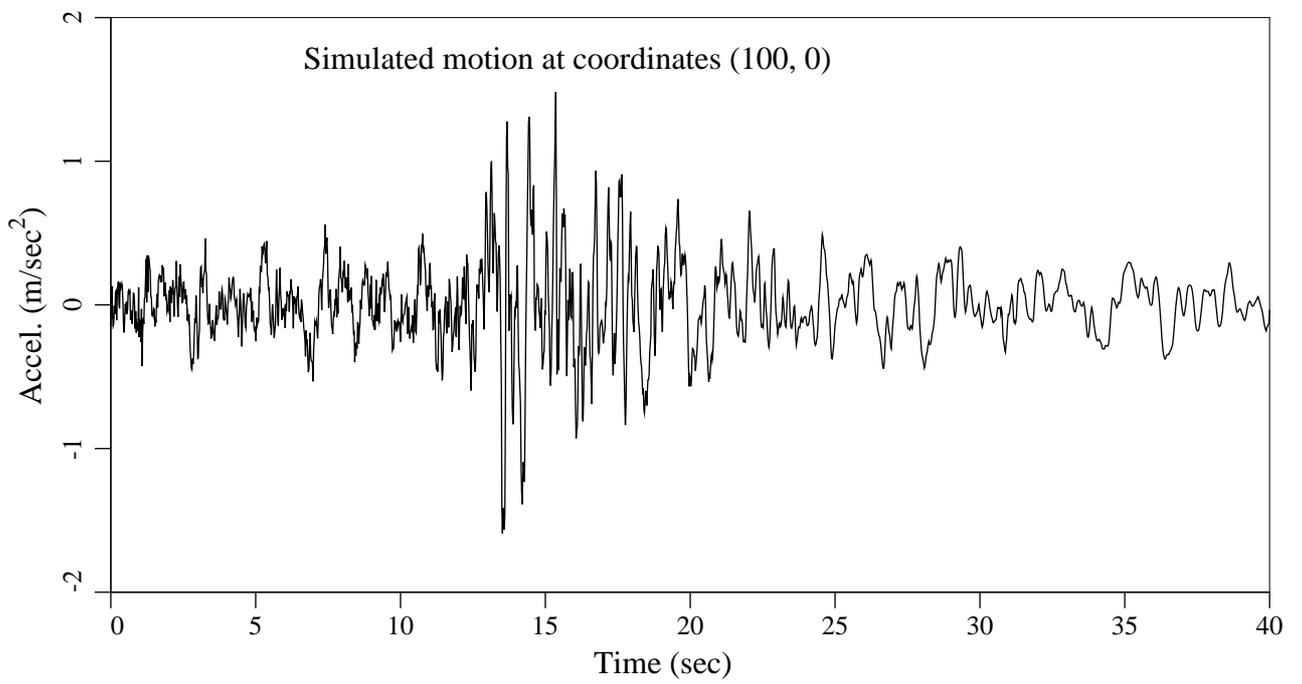
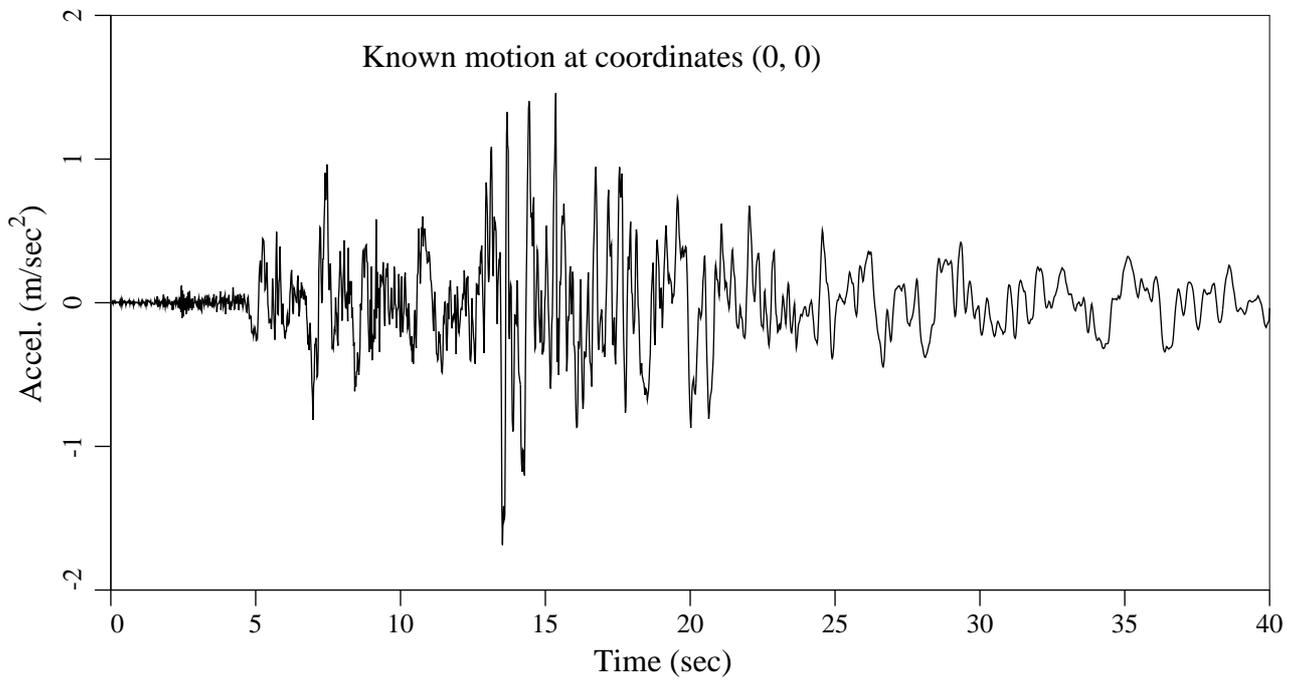


Figure 5.9 Two point conditioned non-stationary simulation for scale $s = 100$.

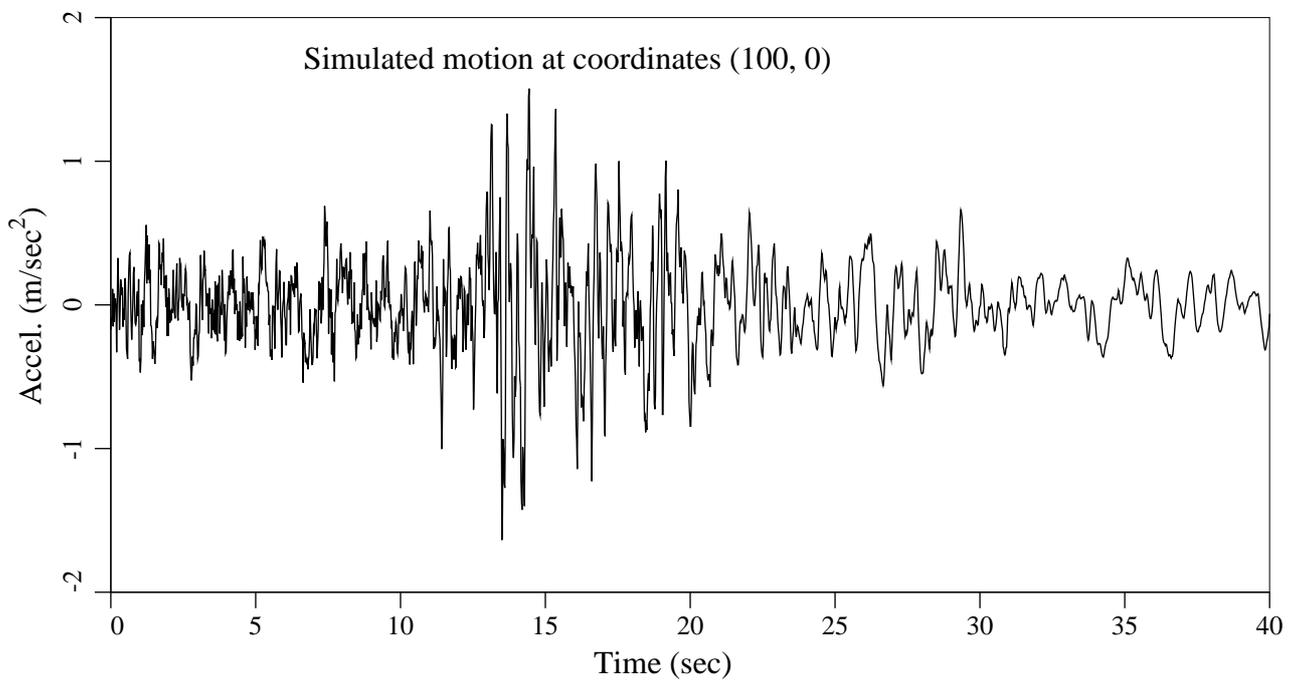
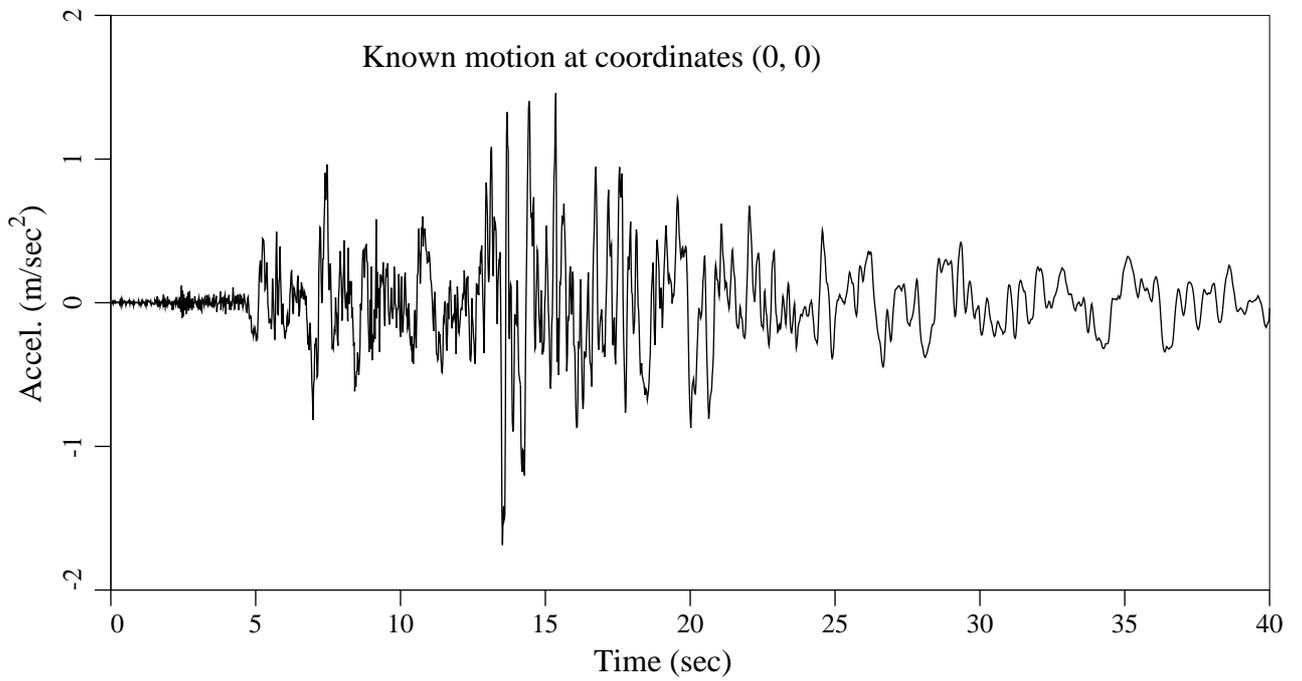


Figure 5.10 Two point conditioned non-stationary simulation for scale $s = 10$.

Chapter 6

Statistics of Level Excursions and Extrema

6.1 Introduction

In both design and analysis contexts, the extremes of random processes are typically of considerable interest. Many reliability problems are defined in terms of threshold excursions – when load exceeds strength for example. Most theories governing extremal statistics of random fields deal with excursion regions, regions in which the process Z exceeds some threshold and the few exact results that exist usually only apply asymptotically when the threshold level approaches infinity. A large class of random functions are not amenable to existing extrema theory at all and for such processes the analysis of a sequence of realizations is currently the only way to obtain their extrema statistics. In this chapter, a methodology for simulation-based estimation of the statistics of level excursions and extrema will be developed. The treatment herein is limited to the two-dimensional case although the procedure is easily extended to higher dimensions. Seven quantities having to do with level excursions and extrema of two-dimensional random fields are examined,

- 1) the total area of excursion regions within a given domain (A_b),
- 2) the number of isolated excursion regions (N_b),
- 3) the area of isolated excursion regions (A_e),
- 4) the number of holes appearing in excursion regions (N_h),
- 5) an integral geometric characteristic defined by Adler [2] (Γ),
- 6) a measure of ‘clustering’ defined herein (Ψ),
- 7) the distribution of the global maxima.

These quantities will be estimated for a single class of random functions, namely Gaussian processes with Markovian covariance structure (Gauss-Markov processes), over a limited range of scales of fluctuation and threshold levels and so the study is by no means complete and should be viewed primarily as a new approach to the determination of these statistics. Hopefully the appearance of empirical relationships will lead eventually to exact analytical results.

Within a given domain $\mathcal{V} = [0, T_1] \times [0, T_2]$ of area A_T , the total excursion area, A_b , can be defined by

$$A_b = \int_{\mathcal{V}} I_{\mathcal{V}}(Z(\mathbf{x}) - b\sigma) d\mathbf{x}, \quad (6.1)$$

where $b\sigma$ is the threshold of interest, σ^2 being the variance of the process, and $I_{\mathcal{V}}(\cdot)$ is the indicator function defined on \mathcal{V}

$$I_{\mathcal{V}}(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}. \quad (6.2)$$

For a homogeneous process, the expected value of A_b is simply

$$\mathbb{E}[A_b] = A_T \mathbb{P}[Z(\mathbf{0}) \geq b\sigma], \quad (6.3)$$

which, for a zero-mean Gaussian process yields

$$\mathbb{E}[A_b] = A_T [1 - \Phi(b)], \quad (6.4)$$

where Φ is the standard normal distribution function. The total excursion area A_b is made up of the areas of isolated (disjoint) excursions A_e as follows

$$A_b = \sum_{i=1}^{N_b} A_{ei}, \quad (6.5)$$

for which the isolated excursion regions can be defined using a point set representation

$$\begin{aligned} \mathcal{A}_{ei} &= \{\mathbf{x} \in \mathcal{V} : Z(\mathbf{x}) \geq b\sigma, \mathbf{x} \notin \mathcal{A}_{ej} \forall j \neq i\}, \\ A_{ei} &= \mathcal{L}(\mathcal{A}_{ei}), \end{aligned} \quad (6.6)$$

where $\mathcal{L}(\mathcal{A}_{ei})$ denotes the Lebesgue measure (or area) of the point set \mathcal{A}_{ei} . Given this definition, Vanmarcke [25] expresses the expected area of isolated excursions as a function of the second-order spectral moments

$$\mathbb{E}[A_{ei}] = 2\pi \left(\frac{F^c(b\sigma)}{f(b\sigma)} \right)^2 |\mathbf{A}_{11}|^{-1/2}, \quad (6.7)$$

in which F^c is the complementary distribution function (for a Gaussian process, $F^c(b\sigma) = 1 - \Phi(b)$), f is the corresponding probability density function, and \mathbf{A}_{11} is the matrix of second-order spectral moments

$$\mathbf{A}_{11} = \begin{bmatrix} \lambda_{20} & \lambda_{11} \\ \lambda_{11} & \lambda_{02} \end{bmatrix}. \quad (6.8)$$

Equation (6.7) assumes that the threshold level is sufficiently high so that the pattern of occurrence of excursions tends toward a two-dimensional Poisson point process. The joint spectral moments $\lambda_{k\ell}$ can be obtained either by integrating the spectral density function,

$$\lambda_{k\ell} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_1^k \omega_2^\ell S(\omega_1, \omega_2) d\omega_1 d\omega_2, \quad (6.9)$$

or through the partial derivatives of the covariance function evaluated at the origin,

$$\lambda_{k\ell} = - \left[\frac{\partial^{k+\ell} B(\boldsymbol{\tau})}{\partial \tau_1^k \partial \tau_2^\ell} \right]_{\boldsymbol{\tau}=\mathbf{0}} \quad (6.10)$$

The above relations presume the existence of the second-order spectral moments of $Z(\boldsymbol{x})$ which is a feature of a mean-square differentiable process. A necessary and sufficient condition for mean square differentiability is

$$\left[\frac{\partial B(\boldsymbol{\tau})}{\partial \tau_1} \right]_{\boldsymbol{\tau}=\mathbf{0}} = \left[\frac{\partial B(\boldsymbol{\tau})}{\partial \tau_2} \right]_{\boldsymbol{\tau}=\mathbf{0}} = 0. \quad (6.11)$$

A quick check of the Gauss-Markov process whose covariance function is given by

$$B(\boldsymbol{\tau}) = \sigma^2 \exp\left\{-\frac{2}{\theta}|\boldsymbol{\tau}|\right\} \quad (6.12)$$

verifies that it is not mean square differentiable. Most of the existing theories governing extrema or excursion regions of random fields depend on this property. Other popular models which are not mean square differentiable and so remain intractable in this respect are:

- 1) the ideal white noise process,
- 2) the moving average of ideal white noise (uniformly weighted window),
- 3) fractal processes.

6.2 Local Average Processes

One of the major motivations for the development of local average theory for random processes is to convert random functions which are not mean square differentiable into processes which are. Vanmarcke shows that even a very small amount of local averaging will produce finite covariances of the derivative process. For a two-dimensional local average process, $Z_D(\boldsymbol{x})$, formed by averaging $Z(\boldsymbol{x})$ over $D = T_1 T_2$, Vanmarcke [25] presents the following relationships for the variance of the derivative process \dot{Z}_D in the two coordinate directions,

$$\text{Var} \left[\dot{Z}_D^{(1)} \right] = \frac{2}{T_1^2} \sigma^2 \gamma(T_2) [1 - \rho(T_1|T_2)], \quad (6.13)$$

$$\text{Var} \left[\dot{Z}_D^{(2)} \right] = \frac{2}{T_2^2} \sigma^2 \gamma(T_1) [1 - \rho(T_2|T_1)], \quad (6.14)$$

where,

$$\dot{Z}_D^{(i)} = \frac{\partial}{\partial x_i} Z_D(\mathbf{x}),$$

$$\gamma(T_1) = \gamma(T_1, 0),$$

$$\gamma(T_2) = \gamma(0, T_2),$$

$$\rho(T_i|T_j) = \frac{1}{T_j^2 \sigma^2 \gamma(T_j)} \int_{-T_j}^{T_j} (T_j - |\tau_j|) B(T_i, \tau_j) d\tau_j. \quad (6.15)$$

Furthermore, he shows that the joint second-order spectral moment of the local average process is always zero for $D > 0$, i.e.,

$$\text{Cov} \left[\dot{Z}_D^{(1)}, \dot{Z}_D^{(2)} \right] = 0, \quad \forall D > 0. \quad (6.16)$$

This result implies that the determinant of the second-order spectral moment matrix for the local average process can be expressed as the product of the two directional derivative process variances,

$$|\mathbf{A}_{11,D}|^{1/2} = \sigma_{\dot{Z}_D}^2 = \left(\text{Var} \left[\dot{Z}_D^{(1)} \right] \text{Var} \left[\dot{Z}_D^{(2)} \right] \right)^{1/2}. \quad (6.17)$$

Since the theory governing statistics of level excursions and extrema for mean square differentiable random functions is reasonably well established for high thresholds (see for example Cramer and Leadbetter [17], Adler [2], and Vanmarcke [25]) attention will now be focused on an empirical and theoretical determination of similar measures for processes which are not mean square differentiable. This will be accomplished through the use of a small amount of local averaging employing the results just stated. In particular the seven quantities specified in the second paragraph of this chapter will be evaluated for the two-dimensional Gauss-Markov process

$$B(\tau_1, \tau_2) = \sigma^2 \exp\left\{-\frac{2}{\theta} \sqrt{\tau_1^2 + \tau_2^2}\right\}, \quad (6.18)$$

realizations of which will be generated using the 2-D LAS method described in Chapter 4. Since the LAS approach automatically involves local averaging of the non-mean square differentiable point process (6.18), the realizations will in fact be drawn from a mean square differentiable process. The subscript D will be used to stress the fact that the results will be for the local average process.

6.3 Analysis of Realizations

Two-dimensional LAS generated realizations of homogeneous, zero-mean, isotropic, Gaussian processes are to be analyzed individually to determine various properties of the discrete binary field, Y , defined by

$$Y_{jk,D} = I_{\mathcal{V}}\left(\mathcal{Z}_{jk,D} - b\sigma_D\right), \quad (6.19)$$

where σ_D^2 is the variance of the local average process. The indicator function $I_{\mathcal{V}}$ is given by (6.2) and so $Y_D(\mathbf{x})$ has value 1 where the function \mathcal{Z}_D exceeds the threshold and 0 elsewhere. In the following, each discrete value of $Y_{jk,D}$ will be referred to as a pixel which is ‘on’ if $Y_{jk,D} = 1$ and ‘off’ if $Y_{jk,D} = 0$. A space filling algorithm was devised and implemented to both determine the area of each simply connected isolated excursion region, $\mathcal{A}_{ei,D}$, according to (6.6), as well as to find the number of ‘holes’ in these regions. In this case, the Lebesgue measure is simply

$$A_{ei,D} = \mathcal{L}(\mathcal{A}_{ei,D}) = \sum \Delta A_{ei,D}, \quad (6.20)$$

where

$$\Delta A_{ei,D} = I_{\mathcal{A}_{ei,D}}\left(\mathcal{Z}_D(\mathbf{x}) - b\sigma_D\right) \Delta A \quad (6.21)$$

is just the incremental area of each pixel which is ‘on’ within the discrete set of points $\mathcal{A}_{ei,D}$ constituting the i^{th} simply connected region. In practice, the sum is performed only over those pixels which are elements of the set $\mathcal{A}_{ei,D}$. Note that the area determined in this fashion is typically slightly less than that obtained by computing the area within a smooth contour obtained by linear interpolation. The difference, however, is expected to be minor at a suitably fine level of resolution.

A ‘hole’ is defined as a set of one or more contiguous ‘off’ pixels which are surrounded by ‘on’ pixels. With reference to Figure 6.1, it can be seen that situations arise in which the hole is only ‘weakly’ surrounded by ‘on’ pixels. The algorithm was devised in such a way that only about half of these weakly surrounded regions are determined to be holes. In addition, if an ‘off’ region intersect with the boundary of the domain, then it is not classified as a hole even if it is surrounded on all other sides by ‘on’ regions.

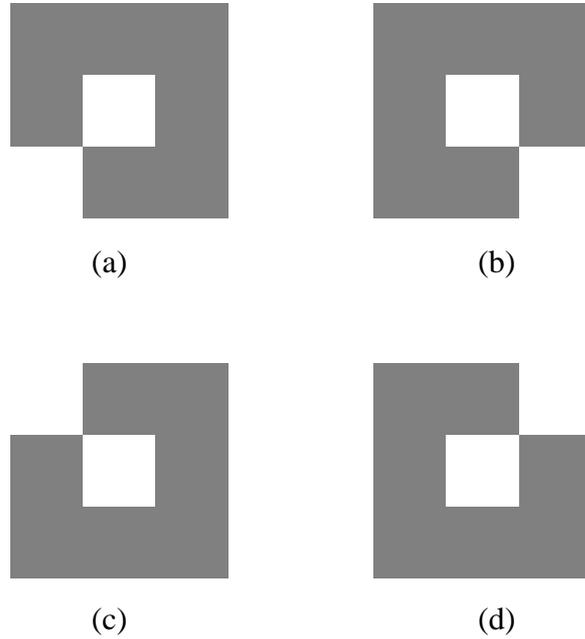


Figure 6.1 Examples of ‘weakly’ surrounded holes - (a) and (b) are found to be holes while (c) and (d) are not.

A comment deserves to be made about the extrema statistics obtained from realizations of the LAS algorithm. As explained in Chapter 4, the LAS method produces a local average process. Thus the statistics obtained are, strictly speaking, those of a local average process and will be affected by the size of the averaging domain. Noting that as the resolution of the field is increased the local average process approaches that of the point process, we will restrict ourselves herein to an analysis of a high resolution field. Our concentration will be primarily on the variation of extrema statistics with scale of fluctuation and threshold level and the dependence on the size of the averaging domain left for later work.

The fields to be generated will have resolution 128×128 and physical size 5×5 . This gives a fairly small averaging domain having edge sizes of $T_1 = T_2 = 5/128$ for which the variance function defined by (4.34) to (4.37) corresponding to (6.18) ranges in value from 0.971 to 0.999 for $\theta = \frac{1}{2}$ to $\theta = 4$. In all cases, the variance of the governing equation (6.18) will be taken as unity and so σ_D^2 equals the variance function.

Figure 6.2 shows a typical realization of the binary field Y obtained by determining the $b = 1$ excursion regions of Z for a scale of fluctuation $\theta = \frac{1}{2}$. Also shown in Figure 6.2 are the $b = 1$ contours which follow very closely the ‘on’ regions. The centroid of each excursion is marked with a darker pixel.

In the sections to follow, trial functions are matched to the observed data and their parameters estimated. All curve fitting was performed by visual matching since it was found that existing least squares techniques for fitting complex non-linear functions were in general unsatisfactory. In most cases the statistics were obtained as averages from 400 realizations.

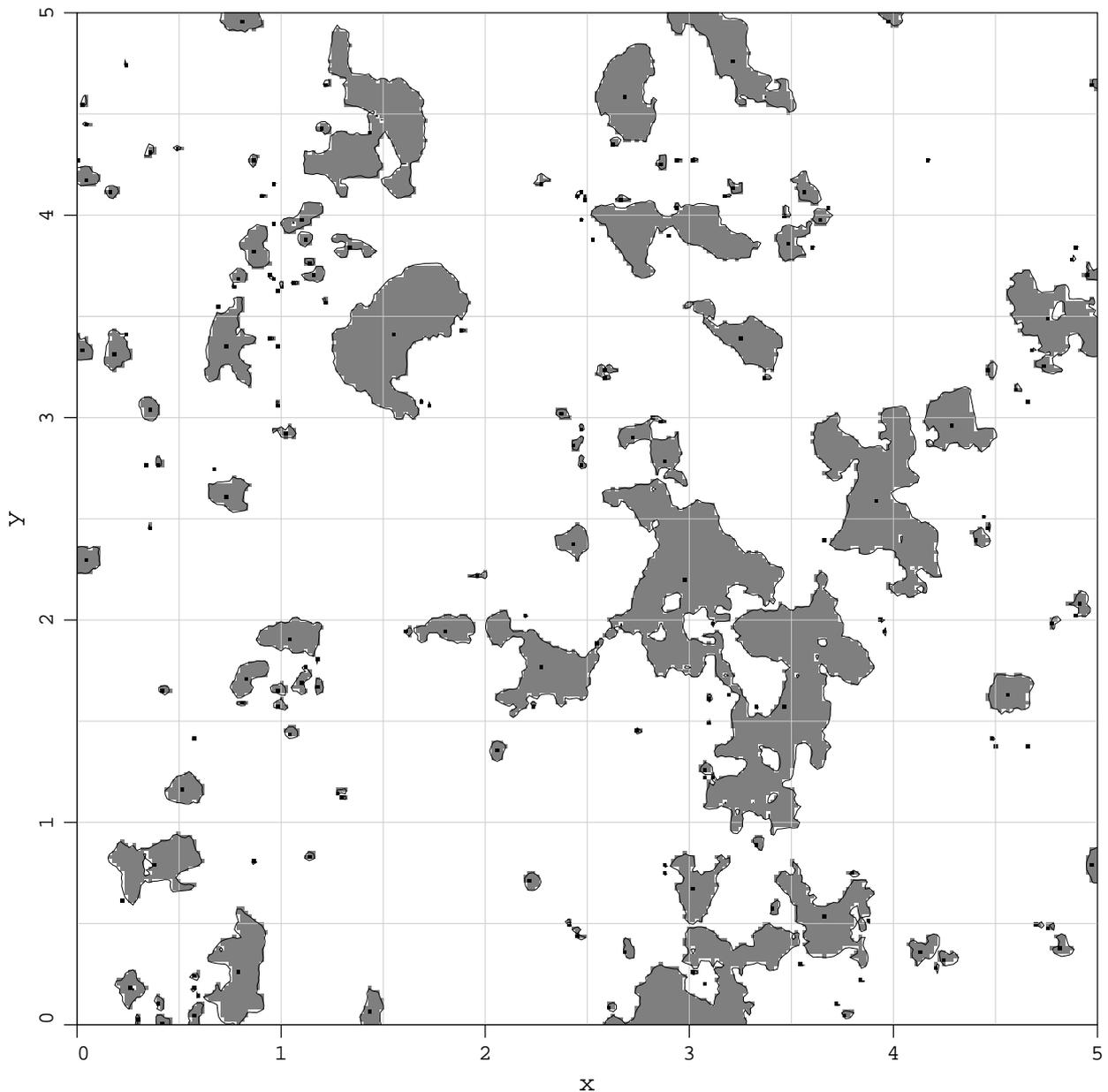


Figure 6.2 Sample function of the binary field Y (6.19). Regions shown in gray represent regions of Z which exceed the threshold $b = 1\sigma_D$. Z is generated via the 2-D LAS algorithm according to (6.18) with $\theta = \frac{1}{2}$.

6.4 Total Area of Excursion Regions

Since an exact relationship for the expected total area of excursion regions within a given domain, (6.4), is known for a Gaussian process, an estimation of this quantity from a series of realizations represents a further check on the accuracy of the simulation method. Figure 6.3 shows the normalized average total area of excursions, $\bar{A}_{b,D}/A_T$, for $A_T = 25$. Here and to follow, the overbar denotes the quantity obtained by averaging over the realizations. The estimated area ratios show excellent agreement with the exact.

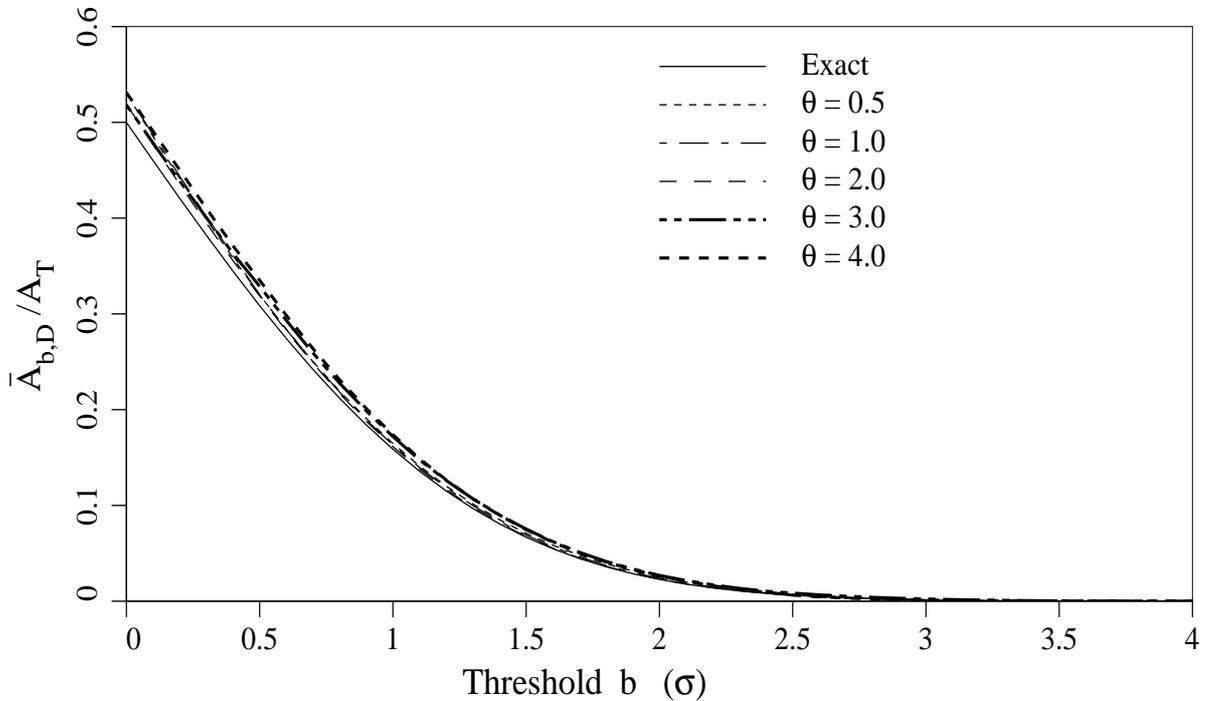


Figure 6.3 Average total area of excursion ratio, $\bar{A}_{b,D}/A_T$, as a function of threshold b .

6.5 Expected Number of Isolated Excursions

Figure 6.4 shows the average number of isolated excursion regions observed within the domain, $\bar{N}_{b,D}$, as a function of scale and threshold. Here the word ‘observed’ will be used to denote the average number of excursion regions seen in the individual realizations. A similar definition will apply to other quantities of interest in the remainder of the chapter. The observed $\bar{N}_{b,D}$ is seen in Figure 6.4 to be a relatively smooth function defined all the way out to thresholds in excess of $3\sigma_D$.

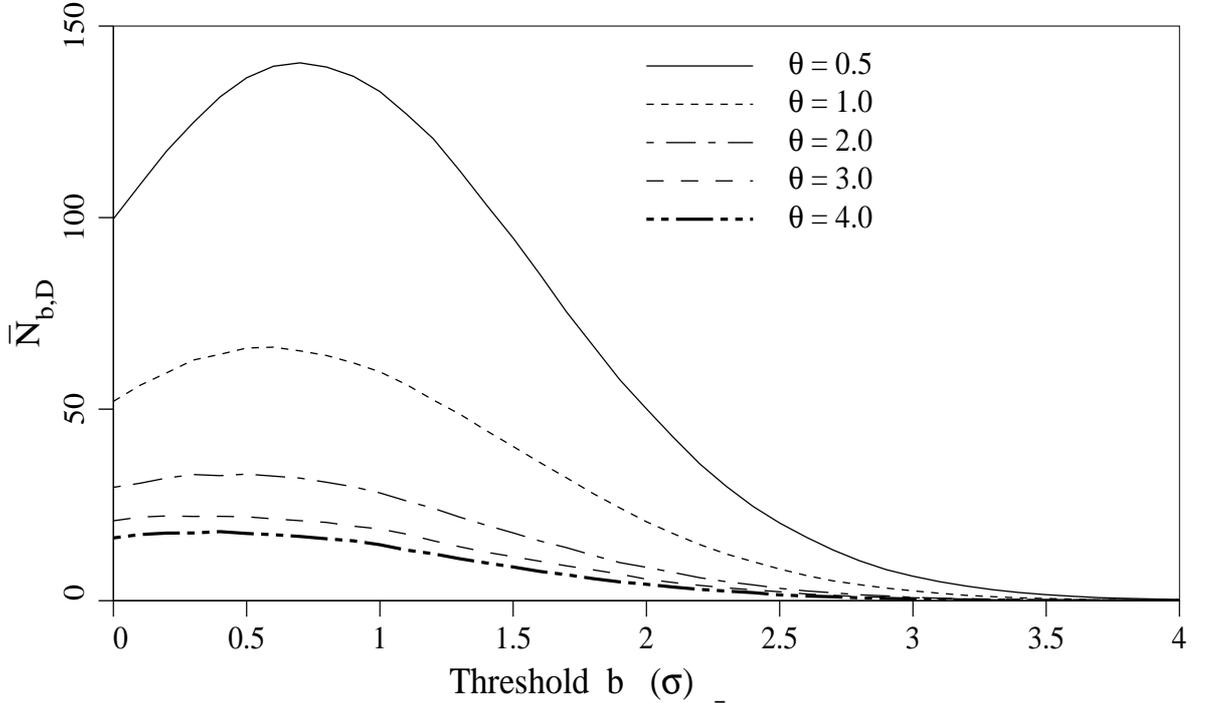


Figure 6.4 Average number of isolated excursions, $\bar{N}_{b,D}$, estimated from 400 realizations of the locally averaged two-dimensional Gauss-Markov process (6.18).

An attempt will be made to fit the theoretical results which describe the mean number of excursions of a local average process above a relatively high threshold to the data shown in Figure 6.4. This expectation is [25]

$$\mathbb{E} [N_{b,D}] = \frac{A_T f_D^2(b\sigma_D)}{2\pi F_D^c(b\sigma_D)} \sigma_{Z_D}^2, \quad (6.22)$$

in which f_D and F_D^c are the pdf and complementary cdf of the local average process respectively. $\sigma_{Z_D}^2$ is the geometric mean of the directional variances of the derivative process as defined by (6.17). For the Gaussian process, (6.22) becomes

$$\mathbb{E} [N_{b,D}] = \frac{A_T e^{-b^2}}{4\pi^2 \sigma_D^2 [1 - \Phi(b)]} \sigma_{Z_D}^2. \quad (6.23)$$

To determine $\sigma_{Z_D}^2$ the functions $\rho(T_1|T_2)$ and $\rho(T_2|T_1)$ must first be calculated using (6.15). Consider $\rho(T_1|T_2)$ for the quadrant symmetric Gauss-Markov process

$$\begin{aligned} \rho(T_1|T_2) &= \frac{2}{T_2^2 \sigma^2 \gamma(T_2)} \int_0^{T_2} (T_2 - \tau_2) B(T_1, \tau_2) d\tau_2 \\ &= \frac{2}{T_2^2 \gamma(T_2)} \int_0^{T_2} (T_2 - \tau_2) \exp\left\{-\frac{2}{\theta} \sqrt{T_1^2 + \tau_2^2}\right\} d\tau_2. \end{aligned}$$

Making the substitution $r^2 = T_1^2 + \tau_2^2$ gives

$$\rho(T_1|T_2) = \frac{2}{T_2^2 \sigma^2 \gamma(T_2)} \int_{T_1}^{\sqrt{T_1^2 + T_2^2}} \left[\frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} - r e^{-2r/\theta} \right] dr.$$

To avoid trying to numerically integrate a function with a singularity at its lower bound, the first term in the integrand can be evaluated as follows

$$\begin{aligned} \int_{T_1}^{\sqrt{T_1^2+T_2^2}} \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} dr &= \int_{T_1}^{\infty} \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} dr - \int_{\sqrt{T_1^2+T_2^2}}^{\infty} \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} dr \\ &= T_2 T_1 K_1\left(\frac{2T_1}{\theta}\right) - \int_{\sqrt{T_1^2+T_2^2}}^a \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} dr - \int_a^{\infty} \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} dr. \end{aligned}$$

The second integral on the right hand side can now be evaluated numerically and for a chosen sufficiently large, the last integral has the simple approximation $\frac{1}{2}\theta T_2 \exp\{-2a/\theta\}$. The function K_1 is the modified Bessel function of order 1. Unfortunately, for small T_1 , the evaluation of this integral is extremely delicate as it involves the small differences of very large numbers. An error of only 0.1% in the estimation of either K_1 or the integrals on the right hand side can result in a drastic change in the value of $\sigma_{\dot{Z}_D}^2$ particularly at larger scales of fluctuation. The following results were obtained using $T_1 = T_2 = \frac{5}{128}$, for which $\rho(T_1|T_2) = \rho(T_2|T_1)$, and a 20 point Gaussian quadrature integration scheme.

Table 6.1 Computed variances of the local average derivative process.

Scale	$\rho(T_1 T_2)$	$\sigma_{\dot{Z}_D}^2$
0.5	0.8482	196.18
1.0	0.9193	105.18
2.0	0.9592	53.32
3.0	0.9741	33.95
4.0	0.9822	23.30

Using these variances, (6.22) was plotted against the observed $\bar{N}_{b,D}$ in Figure 6.5. The relatively poor agreement achieved may be as a result of the combination of the difficulty in accurately determining $\sigma_{\dot{Z}_D}^2$ for small averaging dimensions and the fact that (6.22) is an asymptotic relationship, valid only for $b \rightarrow \infty$. A much better fit in the tails ($b > 1.5$) was obtained using the empirically determined values of $\sigma_{\dot{Z}_D}^2$ shown in Table 6.2 which are typically about one-half to one-third those shown in Table 6.1. Using these values, the fit is still relatively poor at lower threshold levels.

An alternative approach to the description of $\bar{N}_{b,D}$ involves selecting a trial function and determining its parameters. A trial function of the form

$$\bar{N}_{b,D} \simeq A_T (a_1 + a_2 b) \exp\{-\frac{1}{2} b^2\}, \quad (6.24)$$

where the symbol \simeq is used to denote an empirical relationship, was chosen and a much closer fit to the observed data, as shown in Figure 6.6, was obtained using the coefficients shown in Table 6.2. The functional form of (6.24) was chosen so that it exhibits the correct trends beyond the range of thresholds for which its coefficients were derived.

Table 6.2 Empirically determined parameters of (6.24) and variances of the derivative process.

Scale	a_1	a_2	$\sigma_{\dot{Z}_D}^2$
0.5	3.70	5.20	90.0
1.0	2.05	1.90	40.0
2.0	1.18	0.65	17.5
3.0	0.81	0.41	11.3
4.0	0.66	0.29	8.5

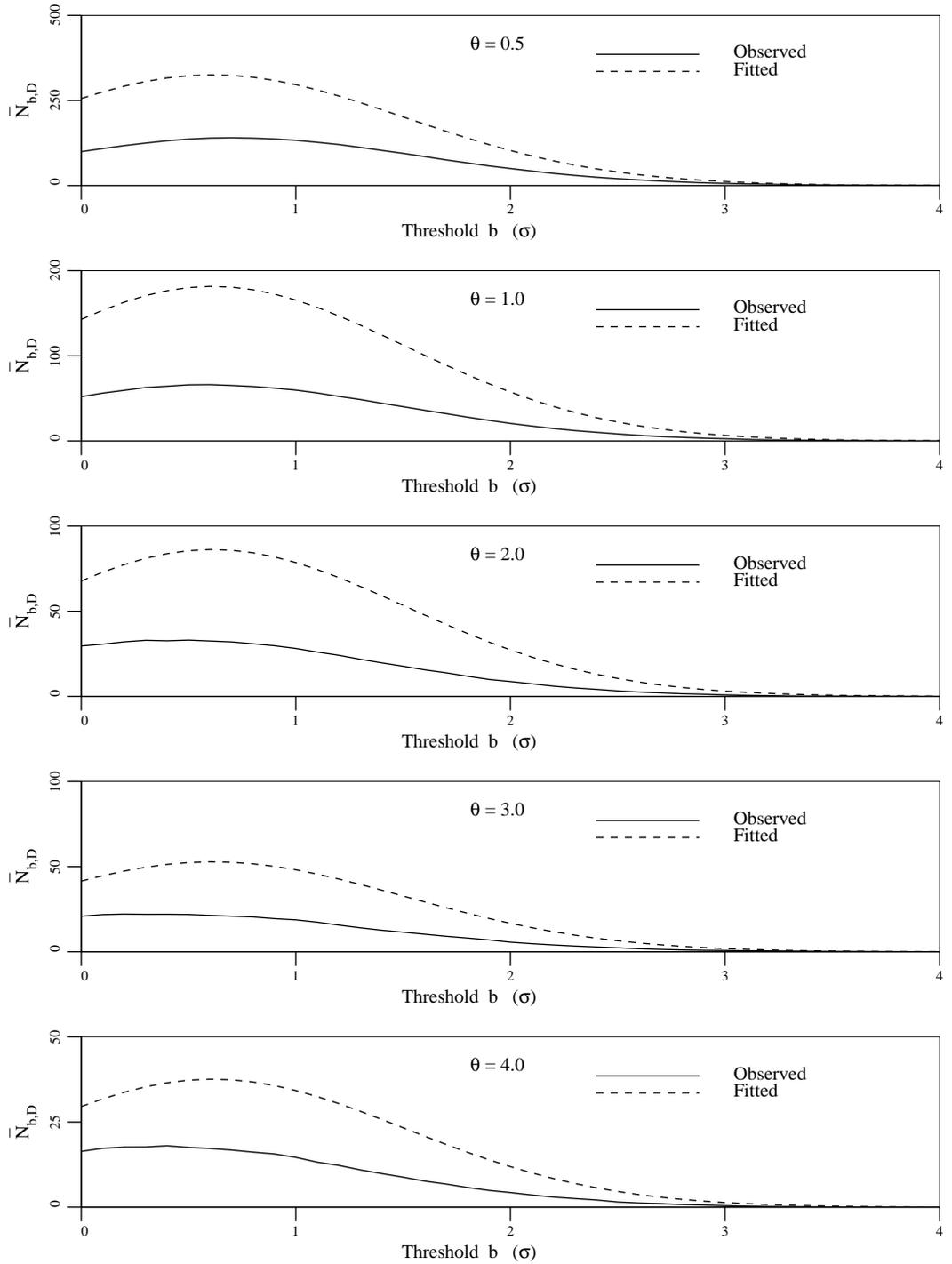


Figure 6.5 Comparison of theoretical fit by (6.23) with the observed average number of isolated excursions obtained by simulation.

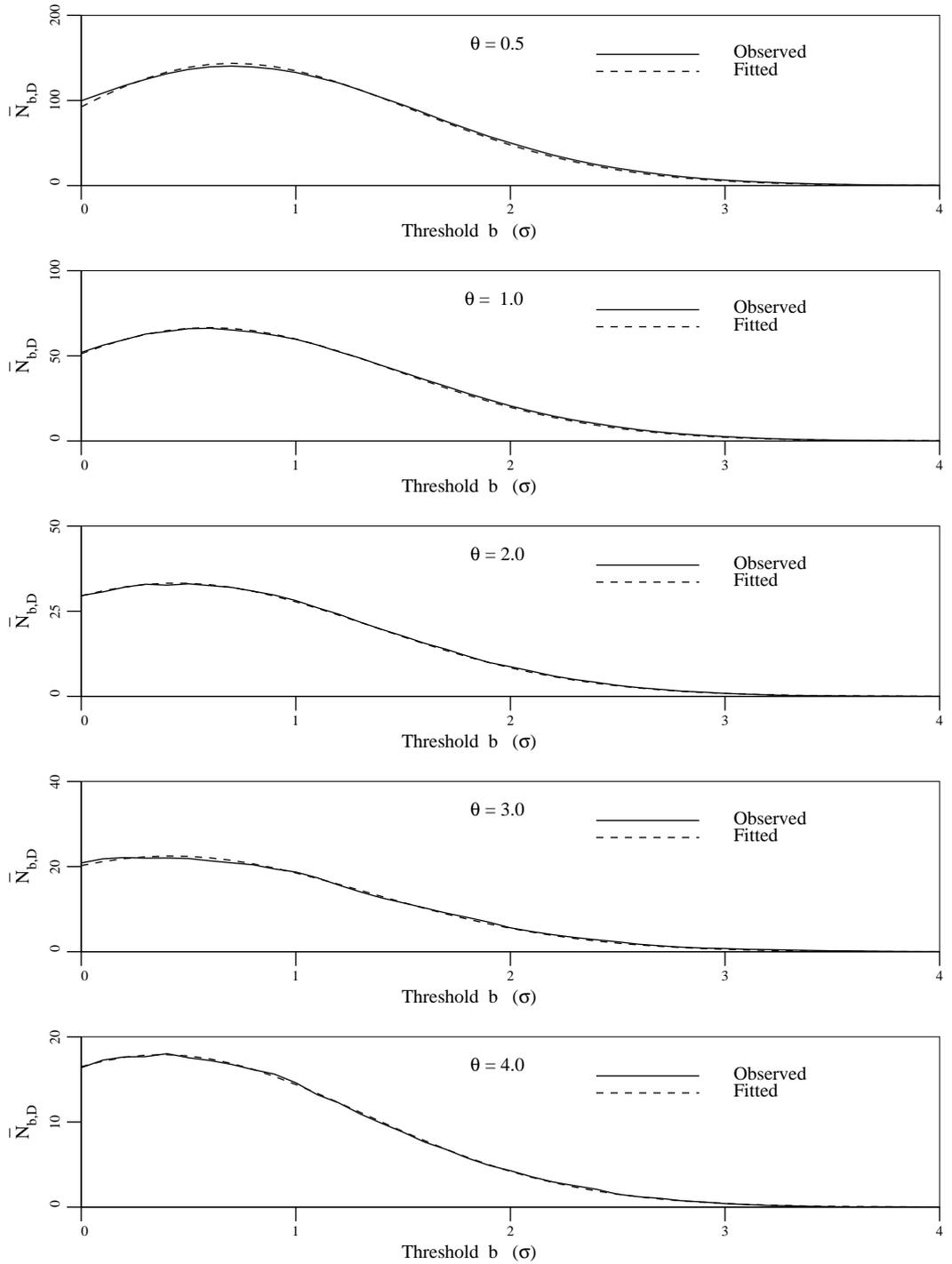


Figure 6.6 Comparison of empirical fit by (6.24) with the observed average number of isolated excursions obtained by simulation.

6.6 Expected Area of Isolated Excursions

Within each realization, the average area of isolated excursions, $\bar{A}_{e,D}$, is obtained by dividing the total excursion area by the number of isolated areas. Further averaging over the 400 realizations leads to the mean excursion areas shown in Figure 6.7 which are again referred to as the ‘observed’ results. The empirical relationship of the previous section, (6.24), can be used along with the theoretically expected total excursion area (6.4) to obtain the semi-empirical relationship

$$\bar{A}_{e,D} \simeq \frac{[1 - \Phi(b)]e^{\frac{1}{2}b^2}}{a_1 + a_2 b}, \quad (6.25)$$

which is compared to the observed in Figure 6.8 and is seen to show very good agreement.

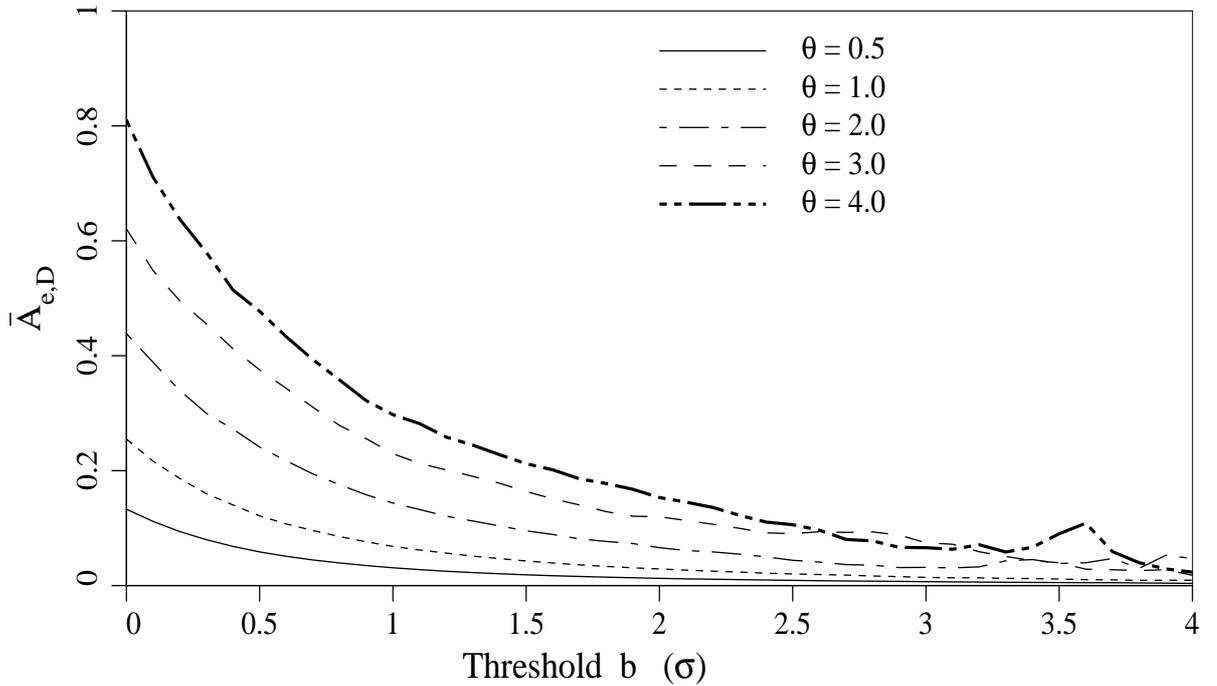


Figure 6.7 Average area of isolated excursion regions estimated from 400 realizations of the locally averaged two-dimensional Gauss-Markov process.

For relatively high thresholds, dividing (6.4) by (6.23) and assuming independence between the number of regions and their total size, yields the expected area to be

$$E [A_{e,D}] = 4\pi^2 [1 - \Phi(b)]^2 e^{b^2} \left(\frac{\sigma_D^2}{\sigma_{Z_D}^2} \right). \quad (6.26)$$

Again the use of $\sigma_{Z_D}^2$ as calculated from (6.17) gives a rather poor fit. Using the empirically derived variances shown in Table 6.2 improves the fit in the tails, as shown in Figure 6.9, but loses accuracy at lower thresholds at most scales.

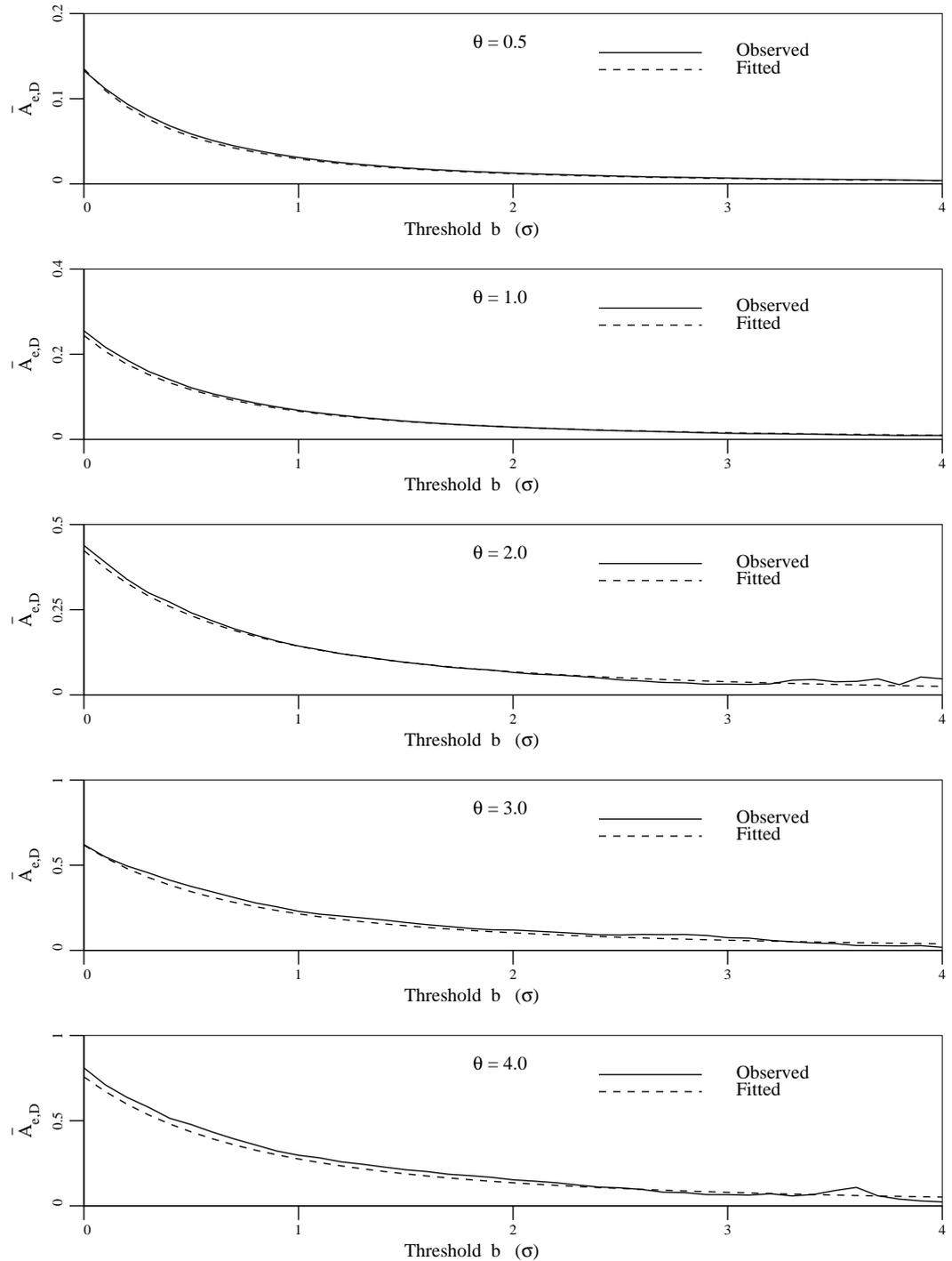


Figure 6.8 Comparison of semi-empirical fit by (6.25) with the observed average area of isolated excursions obtained by simulation.

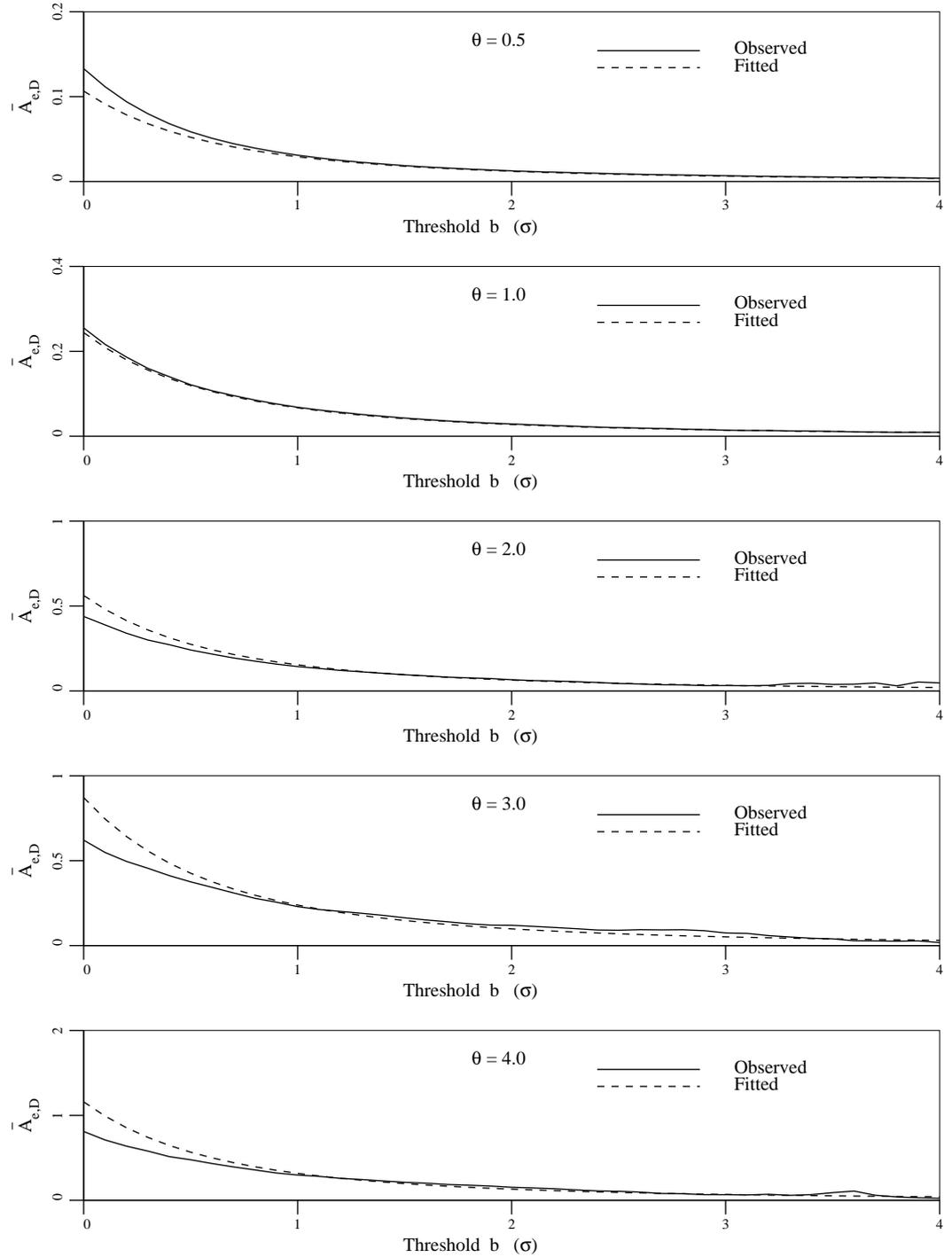


Figure 6.9 Comparison of fit by (6.26), using the empirically derived variances $\sigma_{Z_D}^2$, with the observed average area of isolated excursions obtained by simulation.

6.7 Expected Number of Holes Appearing in Excursion Regions

As will be shown later, the number of holes ('off' regions surrounded by 'on' regions) appearing in excursion regions is to be used in the determination of Adler's [2] integral geometric characteristic of two-dimensional random fields. Since the data is being gathered, an empirical measure relating the average number of holes, $\bar{N}_{h,D}$, with the threshold level and the scale of fluctuation will be derived here. The estimated $\bar{N}_{h,D}$ curves, obtained by finding the number of holes in each realization and averaging over 400 realizations, are shown in Figure 6.10.

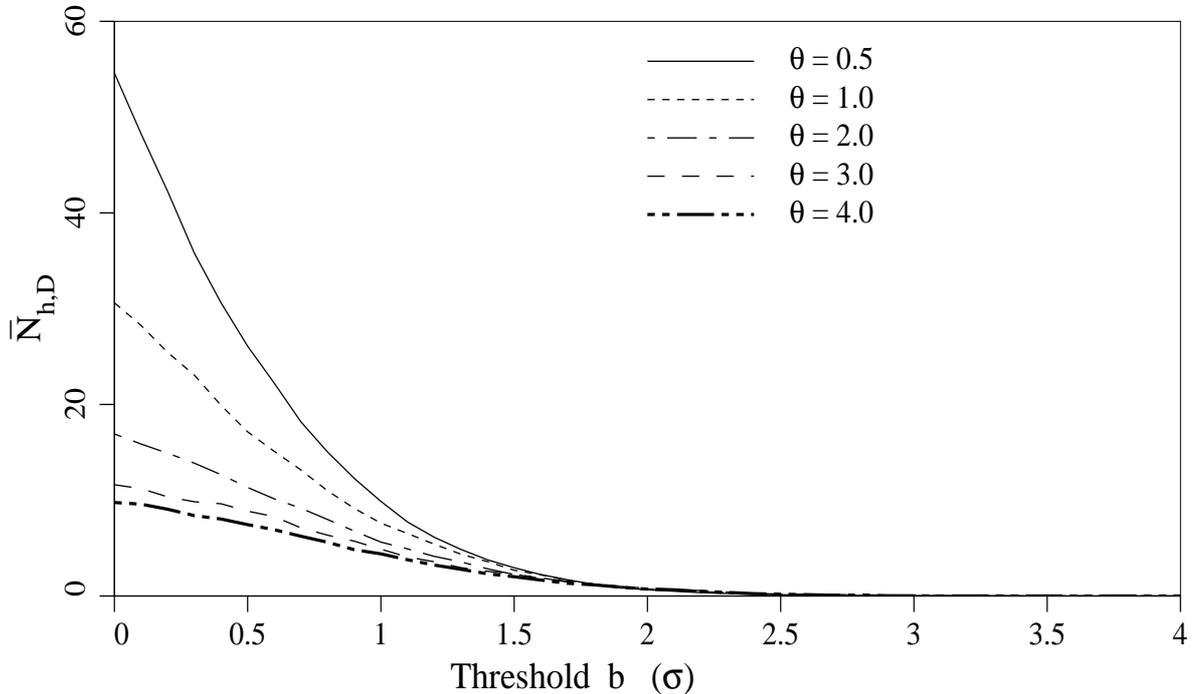


Figure 6.10 Average number of holes appearing in excursion regions.

The empirical model used to fit these curves was

$$\bar{N}_{h,D} \simeq A_T(h_1 + h_2 b)[1 - \Phi(b)], \quad (6.27)$$

where the parameters giving the best fit are shown in Table 6.3 and the comparison is made in Figure 6.11.

Table 6.3 Empirically determined parameters of (6.27) based on the observed average number of holes obtained by simulation.

Scale	h_1	h_2
0.5	4.45	-2.00
1.0	2.49	-0.55
2.0	1.39	0.06
3.0	0.97	0.25
4.0	0.80	0.28

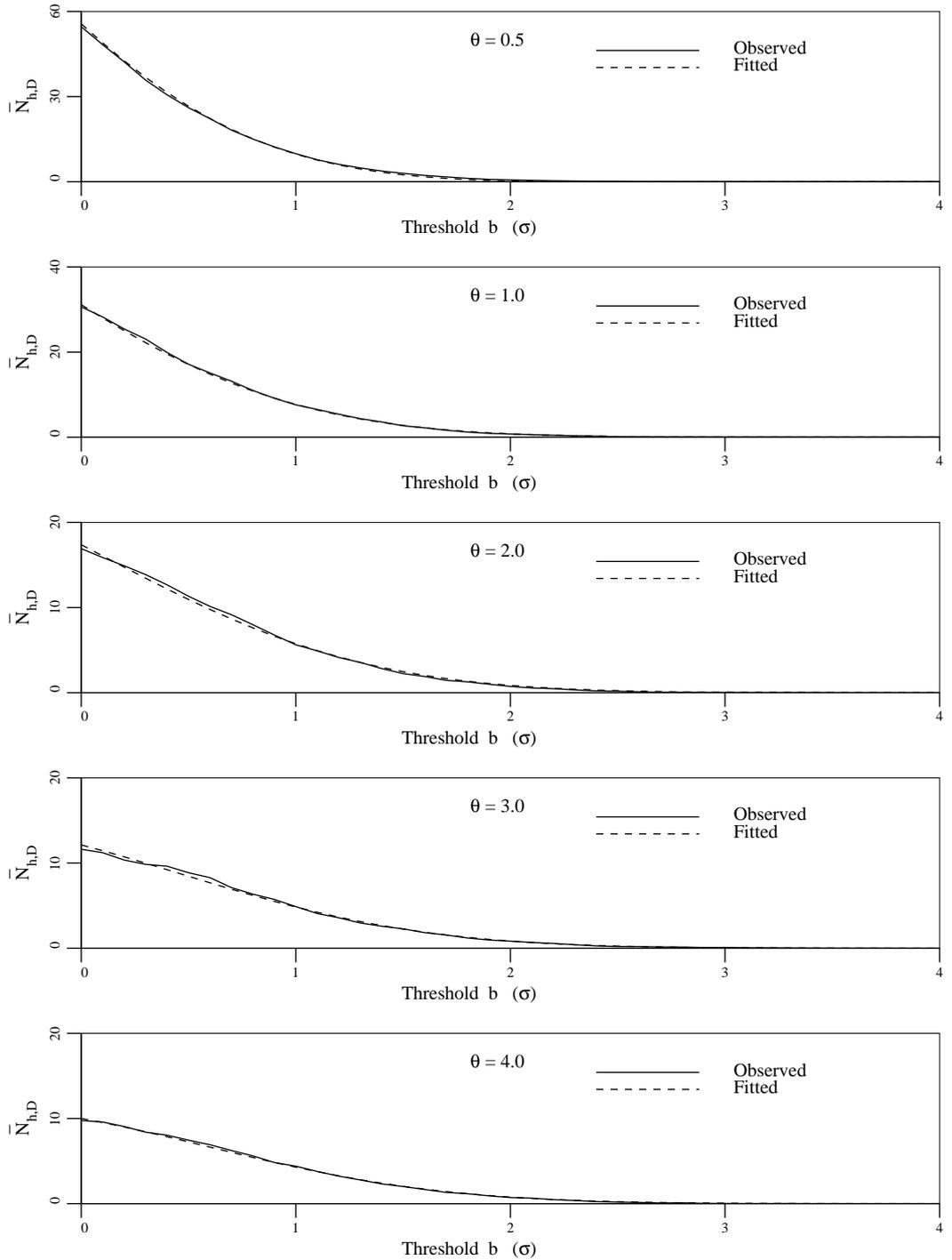


Figure 6.11 Comparison of empirical fit by (6.27) with observed average number of holes obtained by simulation.

6.8 Integral Geometric Characteristic of 2-D Random Fields

In his thorough treatment of the geometrical properties of random fields, Adler [2] develops a so-called integral geometric (IG) characteristic, $\Gamma(\mathcal{A}_{b,D})$, as a statistical measure of two-dimensional random fields. The definition of $\Gamma(\mathcal{A}_{b,D})$ will be shown here specifically for the two-dimensional case although a much more general definition is given by Adler. First, using a point set representation, the excursion set $\mathcal{A}_{b,D}$ can be defined as the set of points in $\mathcal{V} = [0, T_1] \times [0, T_2]$ for which $Z_D(\mathbf{x}) \geq b\sigma_D$,

$$\mathcal{A}_{b,D} = \{\mathbf{x} \in \mathcal{V} : Z_D(\mathbf{x}) \geq b\sigma_D\}. \quad (6.28)$$

The Hadwiger characteristic of $\mathcal{A}_{b,D}$, $\varphi(\mathcal{A}_{b,D})$, is equal to the number of connected components of $\mathcal{A}_{b,D}$ (the number of isolated excursion regions) minus the number of holes in $\mathcal{A}_{b,D}$. Finally, if $\hat{\mathcal{V}}$ is defined as the edges of \mathcal{V} which pass through the origin (the coordinate axes), then the IG characteristic is formally defined as

$$\Gamma(\mathcal{A}_{b,D}) = \varphi(\mathcal{A}_{b,D}) - \varphi(\mathcal{A}_{b,D} \cap \hat{\mathcal{V}}). \quad (6.29)$$

Essentially, $\Gamma(\mathcal{A}_{b,D})$ is equal to the number of isolated excursion areas which do not intersect the coordinate axes minus the number of holes in them. Figure 6.12 shows the average value of the IG characteristic, $\bar{\Gamma}(\mathcal{A}_{b,D})$, obtained from the locally averaged Gauss-Markov process realizations.

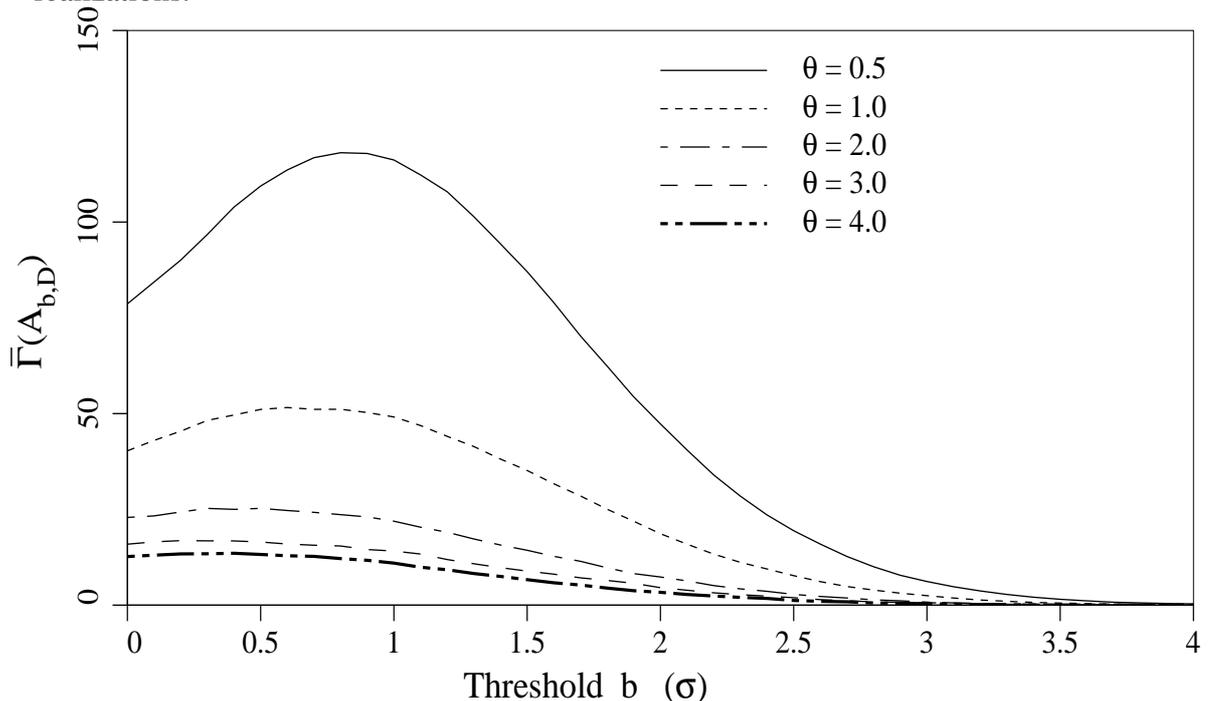


Figure 6.12 Average values of Adler's IG characteristic, $\bar{\Gamma}$, obtained from 400 realizations of the locally averaged Gauss-Markov process.

Adler presents an analytic result for the expected value of $\Gamma(\mathcal{A}_{b,D})$ which has been modified here to account for local averaging of a Gaussian process

$$E [\Gamma(\mathcal{A}_{b,D})] = \frac{b A_T}{(2\pi)^{3/2} \sigma_D^2} \exp\{-\frac{1}{2}b^2\} \sigma_{Z_D}^2. \quad (6.30)$$

Figure 6.13 shows the comparison between (6.30) and the observed data using the empirically estimated variances $\sigma_{Z_D}^2$ shown in Table 6.2. The fit at higher thresholds appears to be quite reasonable.

Using a function of the same form as (6.24),

$$\bar{\Gamma}(\mathcal{A}_{b,D}) \simeq A_T(g_1 + g_2 b) \exp\{-\frac{1}{2}b^2\}, \quad (6.31)$$

yields a much closer fit over the entire range of thresholds by using the empirically determined parameters shown in Table 6.4. Figure 6.14 illustrates the comparison.

Table 6.4 Empirically determined parameters of (6.31) based on the observed average IG characteristic $\bar{\Gamma}$ obtained by simulation.

Scale	g_1	g_2
0.5	2.70	5.10
1.0	1.50	1.80
2.0	0.87	0.58
3.0	0.61	0.32
4.0	0.50	0.22

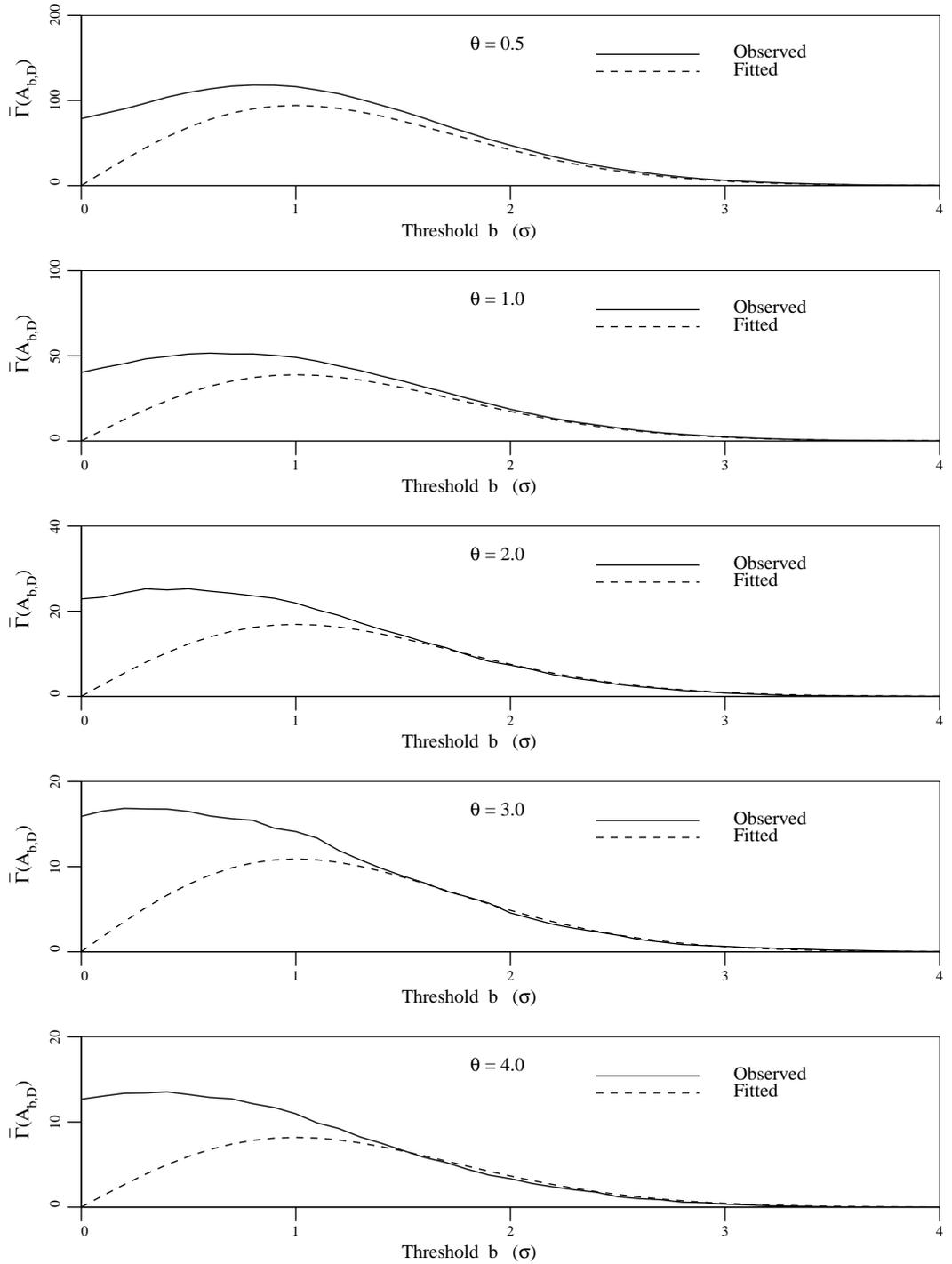


Figure 6.13 Comparison of theoretically predicted IG characteristic (6.30) with observed average values obtained by simulation.

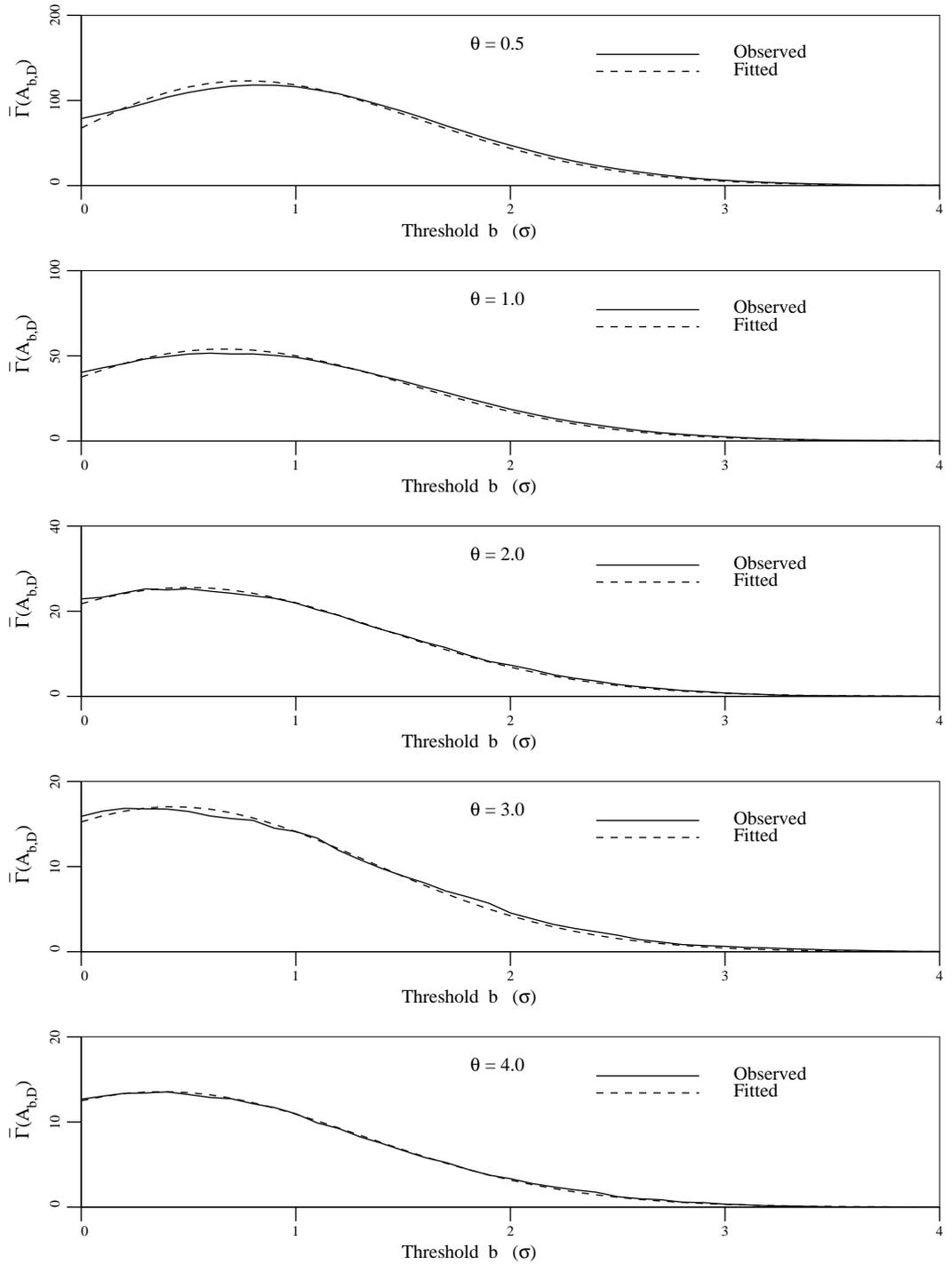


Figure 6.14 Comparison of empirically predicted IG characteristic (6.31) with observed average values obtained by simulation.

6.9 Clustering of Excursion Regions

Once the total area of an excursion and the number of components which make it up have been determined, a natural question to ask is how the components are distributed – do they tend to be clustered together or are they more uniformly distributed throughout the domain? It would be useful to define a measure, herein called Ψ , which varies from 0 to 1 and denotes the degree of clustering, 0 corresponding to a uniform distribution and larger values corresponding to denser clustering. The determination of such a measure involves first defining a reference domain within which the measure will be calculated. This is necessary since a homogeneous process over infinite space always has excursion regions throughout the space. On such a scale, the regions will always appear uniformly distributed (unless the scale of fluctuation also approaches infinity). For example, at scales approaching the boundaries of the known universe, the distribution of galaxies appears very uniform. It is only when attention is restricted to smaller volumes of space that one begins to see the local clustering of stars. Thus an examination of the tendency of excursions to occur in groups must involve a comparison within the reference domain of the existing pattern of excursions against the two extremes of uniform distribution and perfect clustering.

A definition for Ψ which satisfies these criteria can be stated as follows

$$\Psi = \frac{J_u - J_b}{J_u - J_c}, \quad (6.32)$$

where J_b is the polar moment of inertia of the excursion areas about their combined centroid, J_c is the polar moment of inertia of all the excursion areas concentrated within a circle, and J_u is the polar moment of inertia about the same centroid if the excursion area were distributed uniformly throughout the domain. Specifically

$$J_b = \sum_i^{N_{b,D}} J_{ei} + A_{ei,D} |\bar{\mathbf{x}}_b - \bar{\mathbf{x}}_i|^2, \quad (6.33)$$

$$J_{ei} = \sum_j \Delta A_{ei,D} |\bar{\mathbf{x}}_i - \mathbf{x}_j|^2, \quad (6.34)$$

$$J_u = \frac{A_{b,D}}{A_T} \int_{\mathcal{V}} |\bar{\mathbf{x}}_b - \mathbf{x}|^2 d\mathbf{x}, \quad (6.35)$$

$$J_c = \frac{A_{b,D}^2}{2\pi}, \quad (6.36)$$

where J_{ei} is the polar moment of inertia of the i^{th} excursion region of area A_{ei} about its own centroid, $\bar{\mathbf{x}}_i$. $\Delta A_{ei,D}$ is as defined by (6.21) and $\bar{\mathbf{x}}_b$ is the centroid of all the excursion regions. The second moment of area was used in the definition since it is invariant under rotations.

It can be easily seen that this definition will result in $\Psi = 0$ when the excursion regions are uniformly distributed over the space ($J_b \rightarrow J_u$) and $\Psi \rightarrow 1$ when the excursion regions are clustered within a small region ($J_b \rightarrow J_c$). It is also possible for Ψ to take negative values, indicating the occurrence of two local clusters at opposite sides of the domain. This information is just as valuable as positive values for Ψ , but in practice has not been observed to occur on average.

All that remains is to define Ψ in the limiting cases. Equation (6.32) ensures that Ψ will be quite close to 1 in the case of only a single excursion region. It seems natural then to take $\Psi = 1$ if no excursions occur. At the other extreme, as $A_{b,D} \rightarrow A_T$, both the denominator and numerator of (6.32) become very small. Although the limit for non-circular domains is zero, it appears that the measure becomes somewhat unstable as $A_{b,D} \rightarrow A_T$. This situation is of limited interest since the cluster measure of a domain which entirely exceeds a threshold has little meaning. It is primarily a measure of the scatter of isolated excursions.

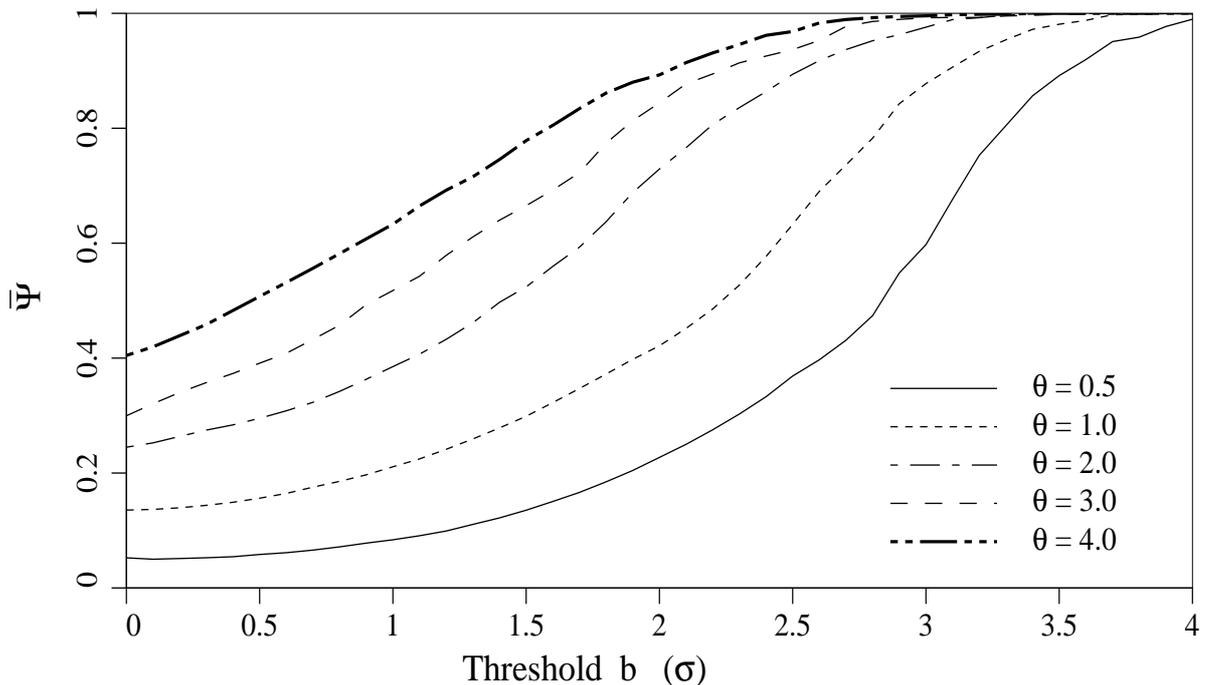


Figure 6.15 Average values of the cluster measure $\bar{\Psi}$ estimated from 200 realizations of the locally averaged Gauss-Markov process.

Individual realizations were analyzed to determine the cluster measure Ψ and then averaged over 200 realizations to obtain the results shown in Figure 6.15. Definite, relatively smooth trends both with scale of fluctuation and threshold level are evident indicating that the measure might be a useful one to categorize the degree of clustering.

6.10 Distribution of the Global Maxima

Extracting the maximum value of Z_D from each realization allows the estimation of its corresponding probability density function (or equivalently the cumulative distribution) with reasonable accuracy given a sufficient number of realizations. A total of 2200 realizations of the locally averaged Gauss-Markov process were generated for each scale of fluctuation considered. Conceptually it is not unreasonable to expect the cumulative distribution of the global maximum $F_{max}(b)$ to have the form of an extreme value distribution for a Gaussian process

$$F_{max}(b) = [\Phi(b)]^{n_{eff}}, \quad (6.37)$$

where n_{eff} is the *effective* number of independent samples in each realization. As the scale of fluctuation approaches zero, n_{eff} should approach the total number of field points (128×128) and as the scale becomes much larger than the field size, n_{eff} is expected to approach 1 (the field becomes totally correlated). Except at the shortest scale of fluctuation considered, $\theta = 0.5$, the function defined by (6.37) was disappointing in its match with the CDF obtained from the realizations. Figure 6.16 illustrates the comparison for the empirically determined values of n_{eff} shown in Table 6.5. The better fit at the smallest scale of fluctuation is to be expected since at very small scales the field consists of a set of (almost) independent random variables and thus satisfies the conditions under which (6.37) applies. Not surprisingly, an improved match is obtained using a two-parameter Type I extreme value distribution having the double exponential form

$$F_{max}(b) = \exp\{-e^{-\alpha(b-\mu)}\}, \quad (6.38)$$

where the parameters α and μ , estimated by an order statistics method developed by Leiblein [38] using the simulation data, are presented in Table 6.5 for each scale of fluctuation. The comparison between the simulation-based cumulative distribution and that predicted by the Type I extreme value distribution is shown in Figure 6.17.

Table 6.5 Empirically determined effective number of independent samples n_{eff} and parameters of the Type I extreme distribution (6.38).

Scale	n_{eff}	α	μ
0.5	2900	3.14	3.41
1.0	900	2.49	3.05
2.0	180	2.05	2.52
3.0	70	1.78	2.15
4.0	35	1.62	1.86

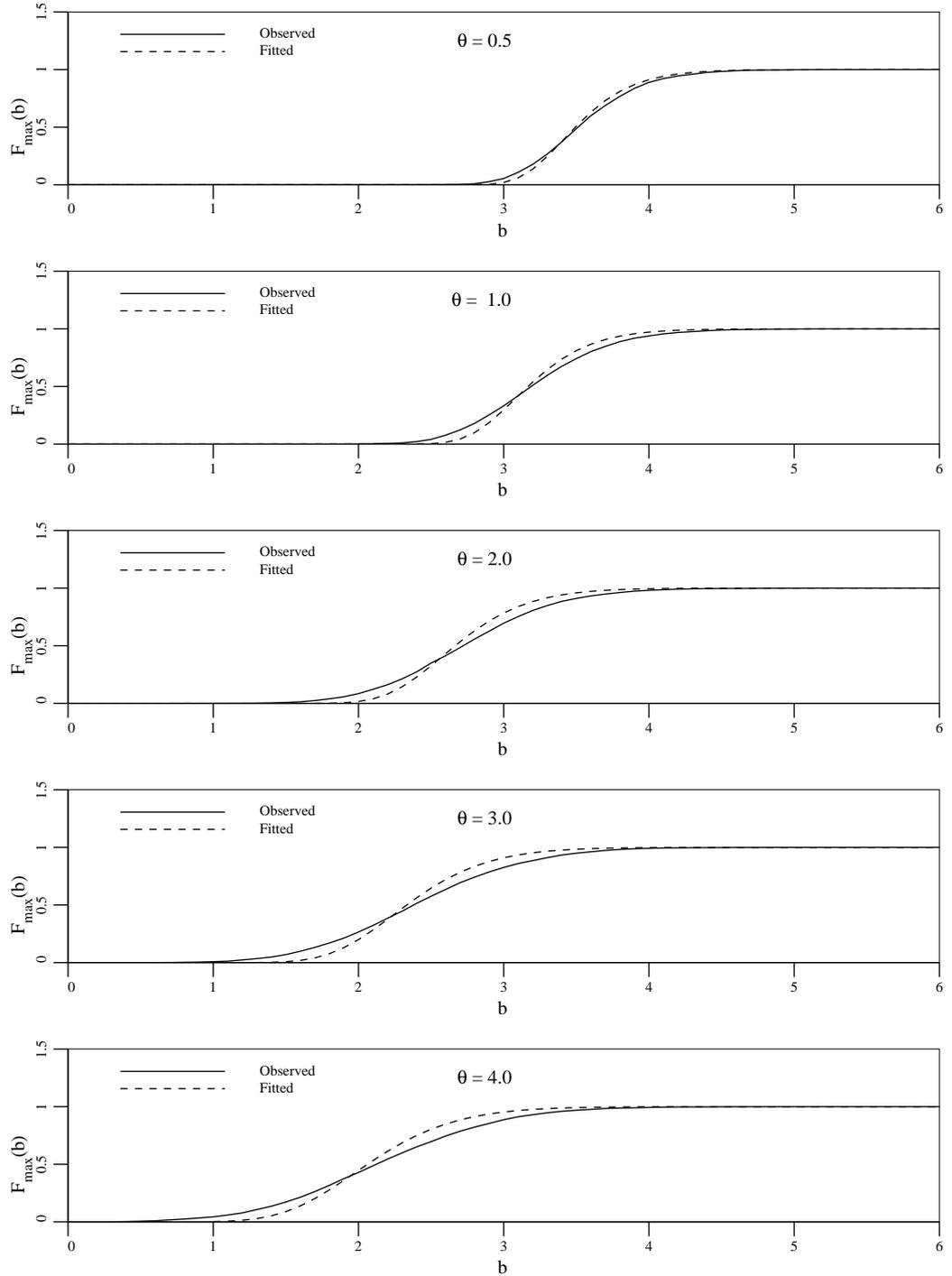


Figure 6.16 Observed cumulative distribution of the global maximum of each realization compared to the one-parameter extreme value distribution given by (6.37).

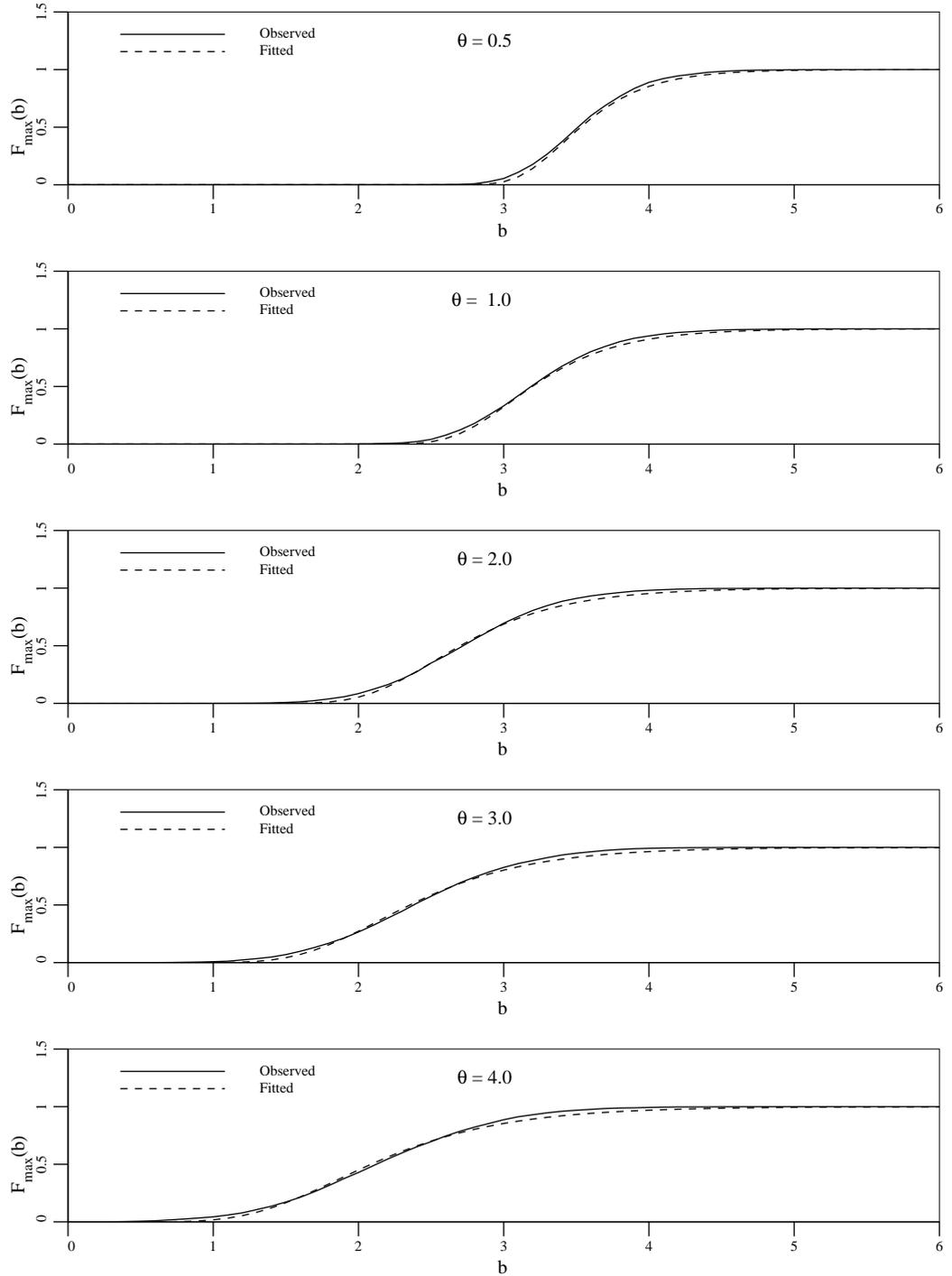


Figure 6.17 Observed cumulative distribution of the global maximum of each realization compared to the Type I distribution given by (6.38).

Chapter 7

Soil Liquefaction Study

7.1 Introduction

This chapter is devoted to a probabilistic liquefaction case study in which a soil deposit, located at the Wildlife Management Area in California, is represented by a three-dimensional stochastic model and subjected to earthquake shaking. The earthquake motions used as input to the model are realizations of a space-time field generated according to the methods presented in Chapter 5. Four cases will be investigated by varying the following two model parameters:

- 1) The evolutionary spectral density functions of the input motion are estimated using an actual time history recorded at the site and a set of unconditioned correlated ground motions are produced (see Section 5.6). In recognition of the major effect earthquake intensity has on the likelihood of liquefaction, a second set of input motions is obtained simply by scaling the first.
- 2) The scale of fluctuation governing the stochastic soil properties is varied from a 'best estimate' to a significantly higher value to investigate its effect on the spatial distribution of liquefaction.

All other parameters are held constant within each case, including the starting pseudo-random number generator seed, so that direct comparison of the results can in principle be carried out. This is desirable since only nine realizations are analyzed for each case and the variance of estimates is in general quite high. The number of realizations considered is held to a minimum since the analysis of each realization is quite time consuming, owing to the non-linear nature of the computations. The study was performed on a Cyber 205, a serial vector supercomputer. It should be pointed out that a Monte-Carlo analysis such as this is ideally suited to parallel architecture computers which could accomplish the same result in a small fraction of the time with little modification of the code.

The major contribution of this chapter is its development of stochastic spatial analysis of liquefaction potential. The trends discovered in the spatial distribution of liquefaction indices and their statistics are believed to be general results but the actual values obtained are site or model specific. Many more analyses such as this are needed at other sites and for other earthquakes before statements about liquefaction likelihood can be confidently made.

7.2 Wildlife Liquefaction Site

An earthquake of magnitude $M_s = 6.0$, on April 26, 1981 in the Imperial Valley near Westmorland, California, caused a significant amount of damage – in many cases by liquefaction. This prompted a detailed geological survey of the valley and the selection of a site for the installation of accelerometers and piezometers to record ground motions and pore-water pressure changes during future earthquakes. The site chosen was the Wildlife Management Area located 3 km south of Calipatria in the Imperial Wildfowl Management Area and lying on the west side of the incised flood plain of the Alamo River, as shown in Figure 7.1. Penetration test and samples were taken by the U.S. Geological Survey [5] identifying seven geological units in the upper 26.5 m of which the three topmost units were considered to be the most significant as they lie within a zone of high liquefaction probability. A section across the flood plain showing these three units is seen in Figure 7.2 and the arrangement of the instrumentation is depicted in Figure 7.3. The instrumentation was installed in 1982 and consisted of surface and down-hole (7.5 m depth) accelerometers and 6 pore-water pressure transducers. Also shown in Figures 7.2 and 7.3 are the cone penetration test (CPT) results at a few lateral positions as a function of depth. Although there appears to be significant vertical variation, the lateral variation is relatively small, consistent with the sedimentary (layered) nature of the deposits.

Figure 7.1 Location of Wildlife Management Area and epicenters of the Elmore Ranch ($M_s = 6.2$) and Superstition Hills ($M_s = 6.6$) events.

Within the upper three units, a closer examination by Holzer *et al.* [29] revealed five geological soil strata to the level of the downhole accelerometer identified as follows:

- 1) Layer 1 (0.0 to 1.2 *m*): very loose silt
- 2) Layer 2 (1.2 to 2.5 *m*): very loose silt
- 3) Layer 3 (2.5 to 3.5 *m*): very loose to loose sandy silt
- 4) Layer 4 (3.5 to 6.8 *m*): loose to medium dense silty sand
- 5) Layer 5 (6.8 to 7.5 *m*): medium to stiff clayey silt

Table 7.1 contains the estimated and measured soil properties for each layer as compiled by Keane and Prevost [34]. The water table level was at a depth of 1.2 *m* and forms the boundary between Layers 1 and 2.

Table 7.1 Material parameters for the layered soil deposit at Wildlife site as estimated by Keane and Prevost [34].

Layer	1	2	3	4	5
Depth (<i>m</i>)	0.0 to 1.2	1.2 to 2.5	2.5 to 3.5	3.5 to 6.8	6.8 to 7.5
Shear Wave Velocity (<i>m/sec</i>)	99	99	116	116	130
Total Density (<i>kg/m³</i>)	1600	1940	1970	1970	2000
Shear Modulus (solid) (<i>N/m²</i>)	1.57×10^7	1.47×10^7	2.08×10^7	2.08×10^7	2.70×10^7
Bulk Modulus (solid) (<i>N/m²</i>)	2.61×10^7	2.44×10^7	4.50×10^7	4.50×10^7	5.83×10^7
Coefficient of Permeability (<i>m/sec</i>)	—	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-4}	1.0×10^{-6}
Poisson's Ratio	0.25	0.25	0.30	0.30	0.30
Void Ratio	0.6799	0.7955	0.7400	0.7400	0.6878
Porosity	0.4047	0.4431	0.4253	0.4253	0.4075
Friction Angle	21.3°	20.0°	22.0°	22.0°	35.0°
Dilation Angle	21.3°	20.0°	19.0°	18.0°	5.0°
Reference Mean Normal Stress (<i>N/m²</i>)	1.15×10^4	2.95×10^4	4.10×10^4	6.10×10^4	8.00×10^4

Figure 7.2 Cross-section of flood plain at the Wildlife site. Lower figure shows a closeup of the strata and CPT measurements at a number of locations.

Figure 7.3 Location of instrumentation at the Wildlife site.

The random medium representation of the soil properties and the deterministic finite element program used to analyze the statistical nature of liquefaction at the Wildlife site are described in the following sections. Recognizing that little information concerning spatial variation of the soil properties and the earthquake ground motion is available, the model is necessarily idealized requiring many of its parameters assumed. For the same reason, a detailed sensitivity analysis was judged to be both too expensive and unwarranted. Only the intensity of the earthquake excitation and the scales of fluctuation of the soil properties were selected as parameters to be varied for the purpose of sensitivity analysis, as discussed in further detail in the next section.

7.3 Stochastic Soil Model

Soil properties are well known to exhibit high spatial variabilities in most natural deposits. For this study, five soil properties expected to have the greatest impact on liquefaction likelihood and site response were selected to be modeled as three-dimensional random fields. They are,

- 1) permeability, k ,
- 2) porosity, n ,
- 3) modulus of elasticity (solid phase), E ,
- 4) Poisson's ratio (solid phase), ν ,
- 5) dilation angle, Φ .

Beyond CPT tests performed at a small number of spatial locations, the available published site information [5, 29] contains barely enough data to establish the mean properties listed in Table 7.1. Estimates of the second moment statistical nature of the above properties must therefore come from a combination of engineering judgement and a review of the literature. Before discussing the specific models to be used, the generation method deserves some attention. In all cases the random material properties will be obtained through appropriate transformations of 3-D zero-mean unit-variance homogeneous Gaussian fields, realizations of which will be obtained using the three-dimensional LAS method. Letting $Q_i(\mathbf{x})$ represent the i^{th} property listed above and z the depth below the surface,

$$Q_i(\mathbf{x}) = \mathcal{T}_i \left(u_i(z) + s_i(z) Z_i(\mathbf{x}) \right), \quad (7.1)$$

where $u_i(z)$ is the mean, $s_i(z)$ is the standard deviation, and \mathcal{T}_i is a transformation taking the Gaussian process into the distribution appropriate for property i . Notice that such a formulation allows trends in the mean and variance as a function of depth to be incorporated. Furthermore, if the field Z_i is generated completely independently of Z_j for $j \neq i$ (as they will be), then Q_i will be independent of Q_j . A point-by-point correlation between the properties can be achieved by writing

$$Q_i(\mathbf{x}) = \mathcal{T}_i \left(u_i(z) + s_i(z) \sum_{j=1}^{n_p} L_{ij} Z_j(\mathbf{x}) \right), \quad (7.2)$$

where n_p is the number of random properties and \mathbf{L} is a lower triangular matrix obtained by Cholesky decomposition of the inter-property correlation matrix, Σ ,

$$\mathbf{L} \cdot \mathbf{L}^T = \Sigma. \quad (7.3)$$

If the transformation \mathcal{T}_i is linear then the components of Σ are

$$\Sigma_{ij} = \frac{E [Q_i Q_j] - E [Q_i] E [Q_j]}{\sigma_i \sigma_j}, \quad (7.4)$$

where σ_i^2 is the variance of the i^{th} property. If the transformation \mathcal{T}_i is not linear, as is usually the case, then (7.4) does not apply and the correlations Σ_{ij} must be chosen approximately (for example, by linearization of \mathcal{T}_i at the mean). This formulation has been implemented in the computer model and represents a tractable means of introducing correlation between soil properties.

7.3.1 Permeability

A soil's permeability is perhaps the single most important property influencing liquefaction. Water trapped within the soil structure carries an increasing fraction of the stress as the soil densifies during shaking. Eventually the inter-granular stresses may become so low that relative movement between particles becomes possible and the medium effectively liquefies.

The mean, depth dependent permeabilities for the Wildlife site were estimated by Keane *et al.* [34] and are listed in Table 7.1. The soil permeability is assumed to be lognormally distributed and so the transformation \mathcal{T}_1 is the exponential

$$k(\mathbf{x}) = \exp \left\{ \mu_{\ln k} + \sigma_{\ln k} \sum_{j=1}^{n_p} L_{1j} Z_j(\mathbf{x}) \right\}, \quad (7.5)$$

where $\mu_{\ln k}$ and $\sigma_{\ln k}^2$ are the mean and variances of the logarithm of permeability, respectively. Both are functions of depth.

A review of the literature concerning the spatial variability of soil properties [3, 13, 14, 15, 19, 23, 22, 24, 33, 40, 44] reveals that little is known about the variability and scales of fluctuation of soil permeability. Denoting by μ_k the mean of the permeability and by σ_k^2 its variance, Gomez-Hernandez *et al.* [22] use a coefficient of variation (σ_k/μ_k) of 1.0 while de Marsily [44] quotes typical values ranging from 0.5 to 1.5. Gelhar [23] estimates that the standard deviation of $\ln k$ varies from 1.2 to 1.7 for unsaturated soils. Table 7.2 shows the variabilities chosen for this study which are roughly in the range indicated by these three researchers. The variance of permeability in the 4th layer (from 3.5 to 6.8 *m* depth) is assumed to be somewhat higher than that of the other layers for demonstrative purposes.

Table 7.2 Parameters used in the modeling of soil permeability, k , at the Wildlife site.

Depth (m)	μ_k	$\mu_{\ln k}$	$\sigma_{\ln k}^2$
0.0 - 1.2	1×10^{-5}	-11.7	0.6
1.2 - 2.5	1×10^{-5}	-11.7	0.6
2.5 - 3.5	1×10^{-5}	-11.9	0.8
3.5 - 6.8	1×10^{-4}	-9.7	1.0
6.8 - 7.5	1×10^{-6}	-14.1	0.5

The covariance function used to model the spatial variability of all the random soil properties is of a simple exponential form parameterized by θ , the scale of fluctuation,

$$B(\tau_1, \tau_2, \tau_3) = \sigma^2 \exp \left\{ -\frac{2}{\theta} \left(|\tau_1| + \sqrt{\tau_2^2 + \tau_3^2} \right) \right\}, \quad (7.6)$$

with associated variance function given by (4.41). In the vertical direction, de Marsily expects the scale of fluctuation of soil permeability to be in the order of 1 m and so, initially, a vertical scale of fluctuation, θ_v , of 1 m is used. Scales of fluctuation in the horizontal direction, θ_h , are highly dependent on the uniformity of the stratification. It appears from Figures 7.2 and 7.3 that the layers are fairly uniform and so a ratio of horizontal to vertical scales of $\theta_h/\theta_v \simeq 40$ was selected implying a horizontal scale of fluctuation $\theta_h \simeq 40$ m . This is in the same range as Vanmarcke's [66] estimate of 55 m for the compressibility index of a sand layer. Although compressibility and permeability are different engineering properties, it is felt that the scale of fluctuation is largely dependent on the geotechnical process of layer deposition rather than the actual property studied. Based on this reasoning, all the random soil properties are modeled using the same scale of fluctuation as well as the same form of the covariance function.

The case studies will be repeated using a larger scale of fluctuation of $\theta_v = 5$ and the ratio of horizontal to vertical scales will be held constant at about 40. In the following, we will refer only to the vertical scale of fluctuation when defining the cases.

Figure 7.4 shows a typical realization of the permeabilities obtained for the Wildlife site using the means and variances given in Table 7.2 and vertical scale of fluctuation $\theta_v = 1.0$ m . This realization was produced using the three-dimensional LAS algorithm discussed in Chapter 4. Changes in mean with depth are clearly evident. The figure shows 4 adjacent columns of the $32 \times 32 \times 32$ realization. In order to obtain the lateral scale of fluctuation of about 40 times the vertical scale, the field was first generated as a 7.5 m cube then stretched in the lateral directions to a size of 320×320 m giving each cell dimensions 0.23 m in height

by 10×10 m laterally. Using this transformation, the final horizontal scale of fluctuation is $\theta_h = 42.7 \theta_v$.

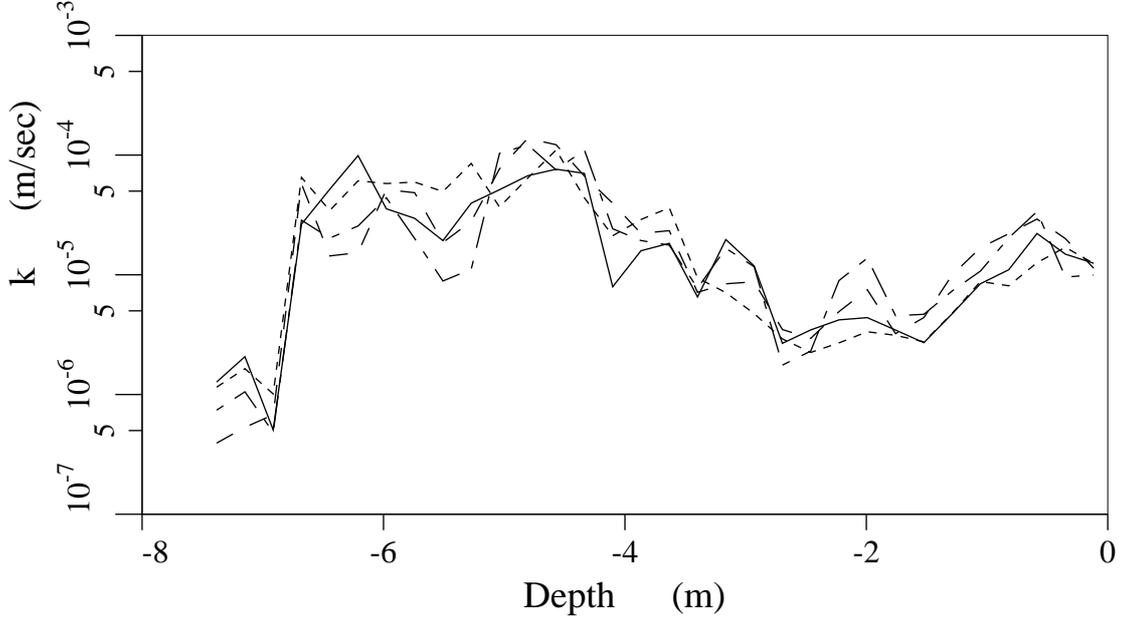


Figure 7.4 Typical realization of soil permeabilities over depth using vertical and horizontal scales of fluctuation of $\theta_v = 1$ and $\theta_h = 42.7$ m respectively. Four adjacent columns of the soil model are shown.

7.3.2 Porosity

Porosity has little direct influence on the occurrence of liquefaction but is related to soil density which in turn affects the initial vertical stresses in the medium as well as the shear wave velocities. The average porosity, n , over depth at the Wildlife site is assumed constant at 0.42. Recognizing that n must be bounded, a transformation \mathcal{T}_2 (see 7.2) must be found to take a normally distributed variate into a bounded distribution. One such transform is

$$X = \mathcal{T}_2(Y) = \frac{1}{2} \left\{ 1 + \tanh \left(\frac{Y}{2\pi} \right) \right\}, \quad (7.7)$$

which is a one-to-one mapping of $Y \in (-\infty, \infty)$ into $X \in (0, 1)$ where Y is obtained from the random fields Z according to (7.2)

$$Y(\mathbf{x}) = u_2(z) + s_2(z) \sum_{j=1}^{n_p} L_{2j} Z_j(\mathbf{x}). \quad (7.8)$$

The probability function corresponding to (7.7) for $Y \sim N(0, s_2)$ is

$$f_X(x) = \frac{\sqrt{\frac{\pi}{2}}}{x(1-x)s_2} \exp \left\{ -\frac{1}{2} \left(\frac{\pi \ln \left(\frac{x}{1-x} \right)}{s_2} \right)^2 \right\}, \quad x \in (0, 1), \quad (7.9)$$

which is symmetric about $x = \frac{1}{2}$. Figure 7.5 illustrates this distribution for a variety of variances.

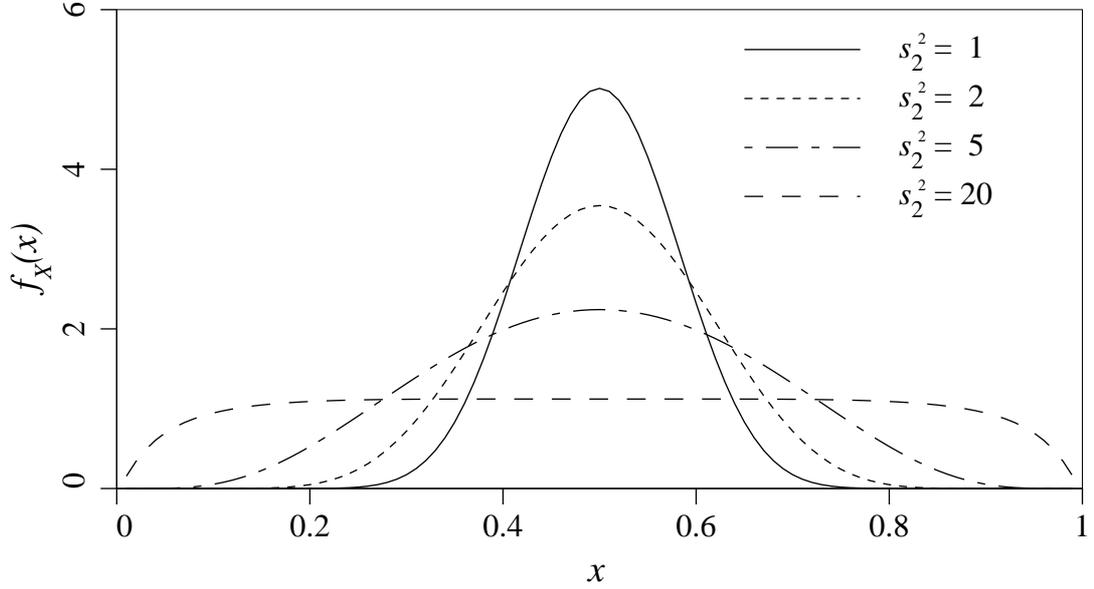


Figure 7.5 Probability density function of the bounded function (7.7) for four different variances.

A further improvement can be made to (7.7) by introducing variable bounds. In this case if $X \in (a, b)$ then (7.7) becomes

$$X = a + (b - a)\mathcal{T}_2(Y). \quad (7.10)$$

For this site, it is assumed that $n \in (0.22, 0.62)$ and that $u_2 = 0$ and $s_2 = 1.0$ so that

$$n = 0.22 + 0.4\mathcal{T}_2(Y). \quad (7.11)$$

The upper bound for n is chosen so that the soil does not initially ‘float’ in the water by requiring

$$(1 - n)\rho_s - \rho_w > 0, \quad (7.12)$$

where ρ_s and ρ_w are the solid and fluid phase mass densities, respectively. The mass density ρ_s was taken to be 2687 kg/m^3 [34] below the water table giving an upper bound on n of 0.627. Because it is well known that soil porosity is related to permeability, a value of $\Sigma_{12} = 0.5$ was assumed in the model to reflect their mutual correlation. Figure 7.6 shows a realization of the porosities (corresponding to Figure 7.4) in four adjacent columns of the field using the above relationships and parameters.

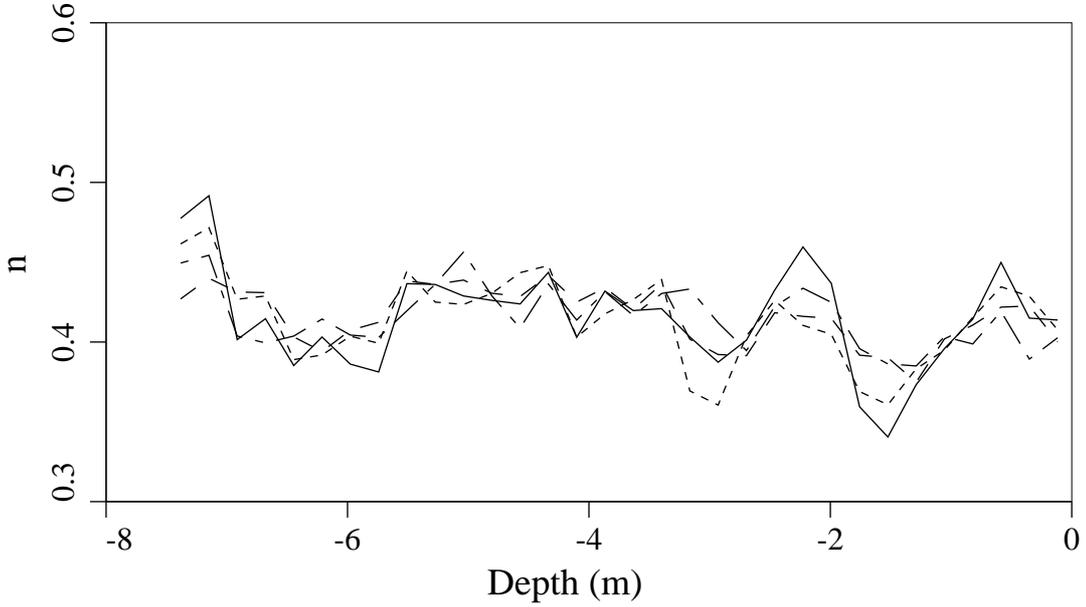


Figure 7.6 Typical realization of soil porosities over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown.

7.3.3 Elastic Modulus and Poisson's Ratio

In terms of the shear moduli G , bulk moduli B , and Poisson's ratio ν , appearing in Table 7.1, the mean elastic moduli can be calculated as a function of depth. Assuming the elastic moduli to follow a lognormal distribution, the calculated means, μ_E , and parameters of the distribution, $\mu_{\ln E}$ and $\sigma_{\ln E}$, are shown in Table 7.3 where the variances, $\sigma_{\ln E}$, of the log-moduli are assumed. The elastic moduli are then obtained using the exponential transformation

$$E(\mathbf{x}) = \exp \left\{ \mu_{\ln E} + \sigma_{\ln E} \sum_{j=1}^{n_p} L_{3j} Z_j(\mathbf{x}) \right\}. \quad (7.13)$$

Figure 7.7 illustrates a typical realization of the elastic moduli over depth for four adjacent columns of the soil model.

Poisson's ratio is chosen to be a bounded variate, $\nu \in (0.075, 0.475)$, according to (7.7) with constant mean 0.275 and $s_4 = 1.0$ where now Y is given by

$$Y(\mathbf{x}) = s_4 \sum_{j=1}^{n_p} L_{4j} Z_j(\mathbf{x}), \quad (7.14)$$

so that

$$\nu = 0.075 + 0.4 \mathcal{T}_4(Y), \quad (7.15)$$

and the transformation \mathcal{T}_4 is the same as \mathcal{T}_2 in (7.7). A typical realization of Poisson's ratio in four adjacent columns of the soil model can be seen in Figure 7.8.

Table 7.3 Parameters used in the lognormal distribution model of the soil elastic moduli, E , at the Wildlife site.

Depth (m)	μ_E (N/m^2)	$\mu_{\ln E}$	$\sigma_{\ln E}^2$
0.0 - 1.2	3.9×10^7	17.1	0.8
1.2 - 2.5	3.7×10^7	17.1	0.6
2.5 - 3.5	5.4×10^7	17.4	0.8
3.5 - 6.8	5.4×10^7	17.2	1.2
6.8 - 7.5	7.0×10^7	17.7	0.8

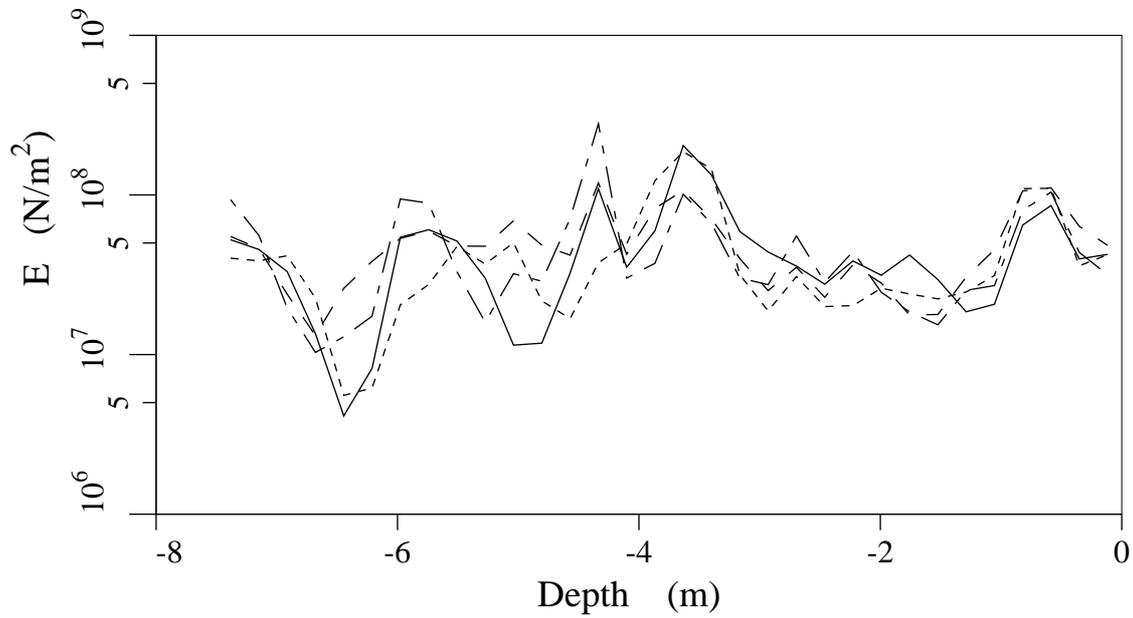


Figure 7.7 Typical realization of soil elastic moduli over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown.

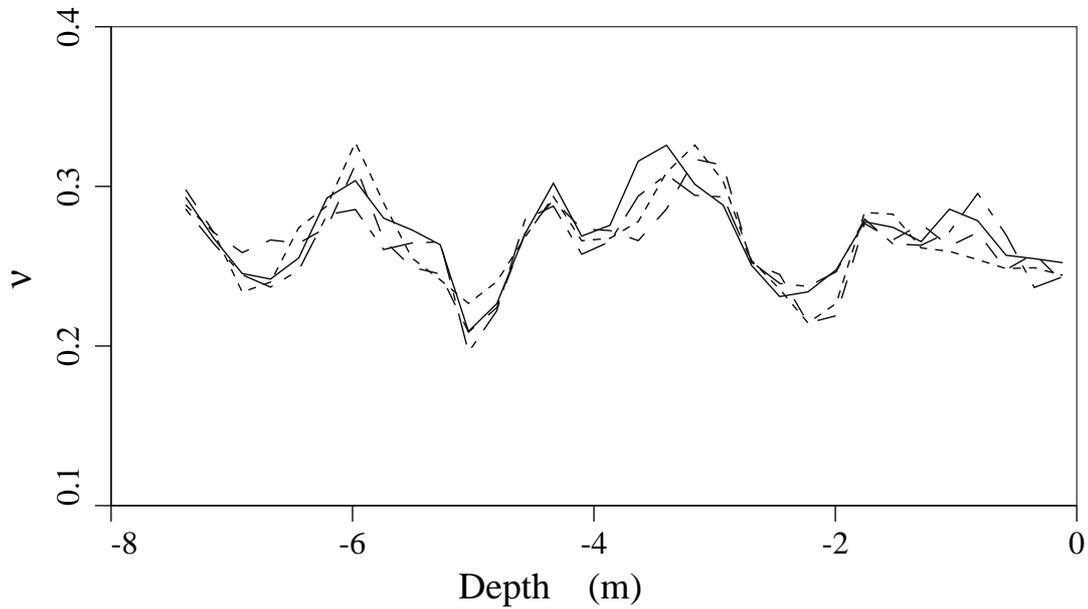


Figure 7.8 Typical realization of soil Poisson ratios over depth ($\theta_v = 1$ m). Four adjacent columns of the soil model are shown.

7.3.4 Dilation Angle

The relationship between the dilation angle, Φ , and the friction angle, ϕ , determines whether the soil subsequently dilates or compacts upon shaking. If the ratio Φ/ϕ exceeds 1.0 then only compaction occurs, otherwise initial compaction is followed by dilation. Since dilation does not result in increasing pore water pressures, this ratio is of considerable interest in the analysis of a liquefiable medium. To reduce the number of random variables considered, the friction angle is assumed fixed at the values given in Table 7.1 and the dilation angle is assumed to be random following a lognormal distribution with means and variances listed in Table 7.4. Figure 7.9 shows a typical realization of the dilation angle over four adjacent columns of the field. A summary of the random soil properties appears in Table 7.5.

Table 7.4 Parameters used in the modeling of soil dilation angle, Φ , at the Wildlife site.

Depth (m)	μ_Φ	$\mu_{\ln \Phi}$	$\sigma_{\ln \Phi}^2$
0.0 - 1.2	21.3°	2.95	0.2
1.2 - 2.5	20.0°	2.90	0.2
2.5 - 3.5	19.0°	2.84	0.2
3.5 - 6.8	18.0°	2.77	0.3
6.8 - 7.5	5.0°	1.51	0.2

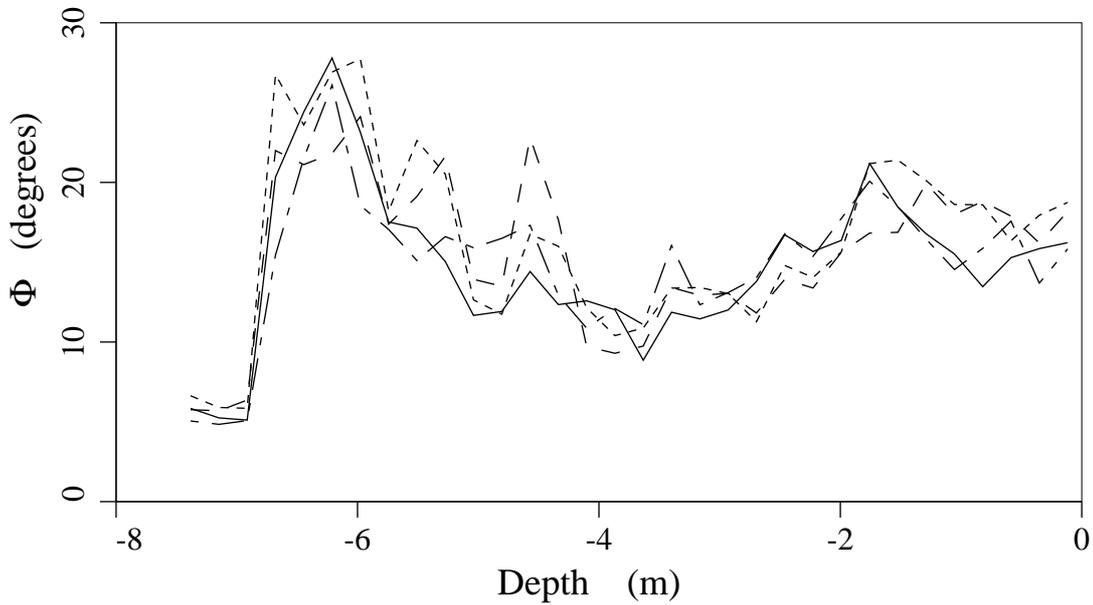


Figure 7.9 Typical realization of soil dilation angles over depth ($\theta_v = 1 \text{ m}$). Four adjacent columns of the soil model are shown.

Table 7.5 Summary of random soil properties used in the soil model.

Property	Distribution	Description
Permeability k	lognormal	mean and variance functions of depth
Porosity n	bounded	constant mean and variance, correlated with permeability
Elastic Modulus E	lognormal	mean and variance functions of depth
Poisson's Ration ν	bounded	constant mean and variance
Dilation Angle Φ	lognormal	mean and variance functions of depth

7.4 Finite Element Model

Realizations of the soil mass are excited by earthquake motions applied at the base of each soil column and analyzed using a one-dimensional finite element model developed by Prevost [55] called DYNA1D which represents the current state-of-the-art. The soil mass is divided up into 64 columns arranged on an 8×8 grid and each column consists of 32 elements (33 nodes) vertically. The size of each element is 0.23 m vertically and $10 \times 10 \text{ m}$ horizontally. Realizations of the random soil properties are obtained by column-wise extraction from a set of $32 \times 32 \times 32$ realizations of the random fields as discussed in the previous section. Which columns are to be extracted depends on the horizontal scale of

fluctuation to be used, however for the example being run here, an adjacent 8×8 group of columns is extracted and stretched horizontally to achieve the desired lateral scale of fluctuation. The choice of an 8×8 grid is arbitrary but a larger grid size will substantially increase the required computer time.

Soil columns are then analyzed individually by DYNA1D, implying that 64 runs of DYNA1D are required for each realization of the soil mass. The runs are made independently and so the only link between the soil columns is through their correlated properties. It is unknown how the coupling between columns in a fully three-dimensional dynamic analysis would effect the determination of liquefaction potential, however it is believed that the analysis proposed herein represents the best approximation to the fully three-dimensional analysis at this time.

The finite element model DYNA1D employs multiple yield level elasto-plastic constitutive theory to take into account the non-linear, anisotropic, and hysteretic stress-strain behavior of the soil as well as the effects of the transient flow of the pore water through the soil media [55]. The code was written so that the material constitutive parameters are obtained internally from conventional soil properties, such as those discussed in the previous section. Each finite element is assigned soil properties obtained from realizations of the random properties and from the set of deterministic properties. Below the water table, the element nodes have four degrees of freedom: two for the solid phase and two for the fluid phase, to accommodate both the vertical and lateral motion of each phase. Above the water table, the soil is assumed to be dry and only two degrees of freedom are needed to fully describe the motion of the soil.

DYNA1D was ported to the CYBER 205 supercomputer and optimized for use both on the supercomputer as well as within the framework of a Monte-Carlo type analysis. Although the run times were considerably improved through explicit optimization, DYNA1D still consumed some 90 to 95 percent of the total run time. Thus it is essential to reduce the work performed by DYNA1D without degrading the accuracy. These reductions will be noted in the following paragraphs.

Earthquake motions applied to the base of each soil column are realizations of a space-time process generated according to the procedures given in Chapter 5. Only one component of motion is used, modeled after the North-South component of the Superstition Hills event. Preliminary results by Keane and Prevost (not published) indicate that including the East-West and vertical components makes little difference to their computed site response (the North-South component had much larger amplitudes), and using it alone, they obtain remarkably good agreement with the observed site response. The original recorded N-S component had an apparent duration of about 40 seconds and a time increment of 0.02 seconds. In an effort to reduce the computational overhead, 1024 time steps are used here with a

time increment of 0.02 seconds giving a duration of 20.48 seconds. The four power spectral density functions shown in Figure 5.6 are employed to govern the simulated time histories, each spanning a time window of 5.12 seconds. To investigate the effect of earthquake intensity, a second case is considered in which the original ground motions are scaled by a factor of 0.7.

Figure 7.10 shows a typical set of input acceleration time histories at each of the four corner columns of the 8×8 grid. The space-frequency correlation function is (see 5.28)

$$\rho_{\omega_k}(\mathbf{r}_{ij}) = \exp \left\{ \frac{-\omega_k |\mathbf{r}_{ij}|}{2\pi c s} \right\}, \quad (7.16)$$

with a shear wave velocity $c = 130 \text{ m/sec}$ (see Layer 5, Table 7.1) and scale parameter $s = 5$. It can be seen that over the relatively short distances between the corner points, the input motions are very similar.

The surface response obtained from the DYNA1D analysis of a realization of a single column of soil is shown in Figure 7.11. For this particular example, the element at a depth of about 2.2 m liquefied after about 13 seconds of motion. In the surface response, this is evident from the drastic reduction in acceleration as the liquefied layer absorbs the motion. The spurious high frequencies apparent in the post-liquefied response are artifacts of the DYNA1D analysis – the convergence tolerance was set at a relatively high value (0.1) to reduce computations and the maximum number of iterations was limited to 20. Prior to liquefaction these choices result in negligible error. For the unstable conditions at or near liquefaction the algorithm obtains poor solutions which tend to fluctuate unpredictably. As the interest here is primarily in the onset of liquefaction, the details of the post-liquefaction response are considered unimportant.

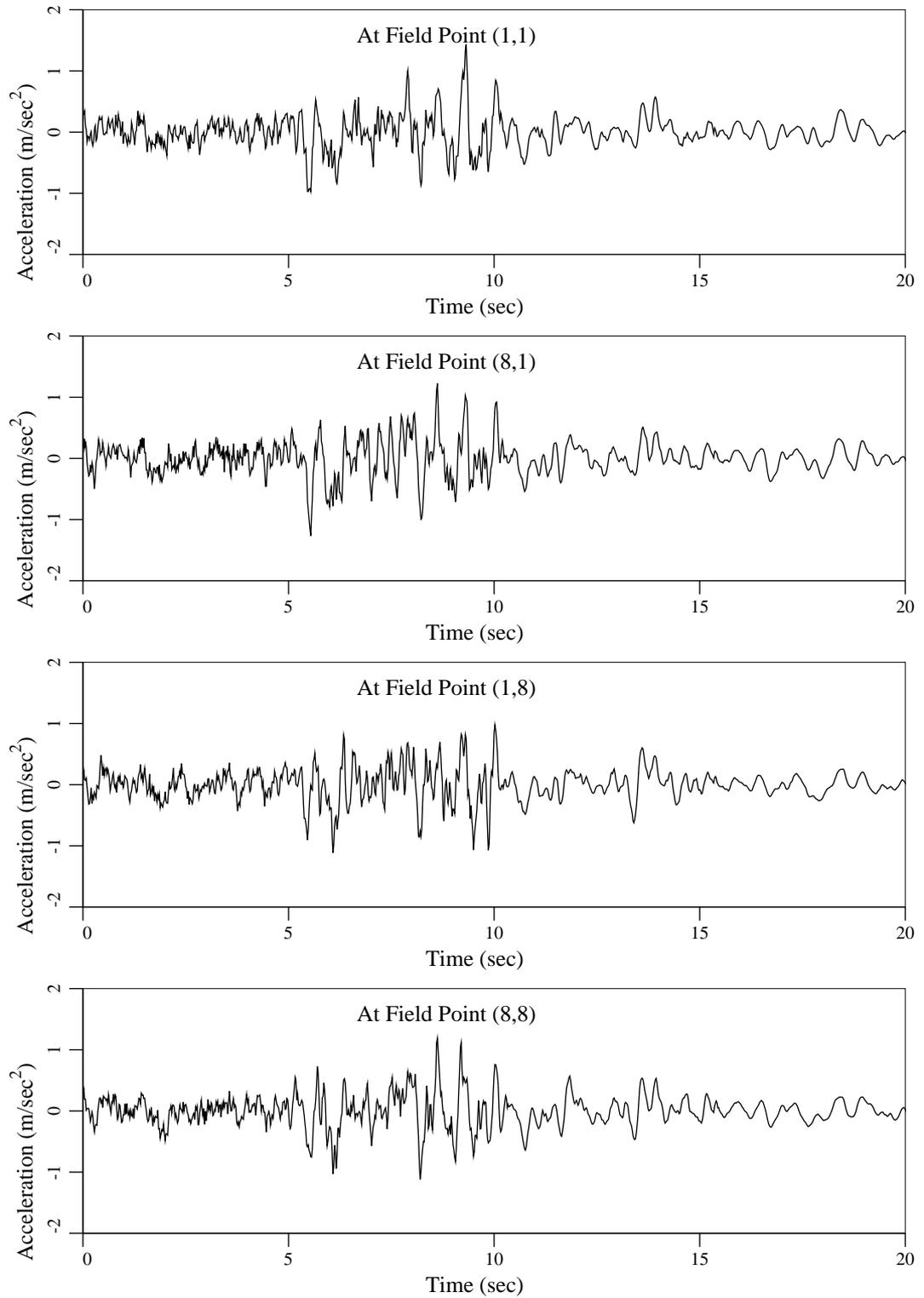


Figure 7.10 Realization of base acceleration time histories at the four corner columns of the 80×80 m soil model.

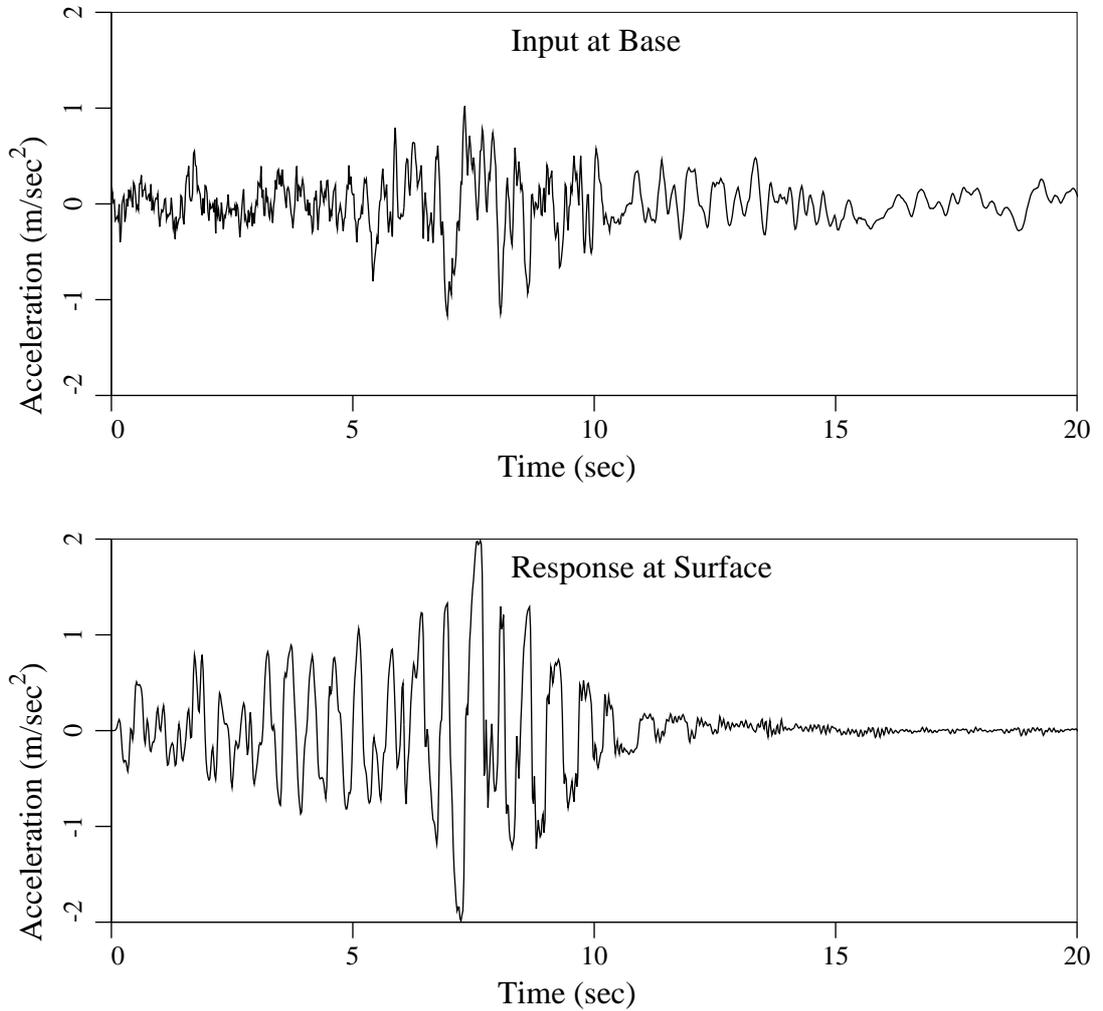


Figure 7.11 Surface response computed by DYNA1D for a particular soil column realization.

7.5 Definition of Liquefaction

From the one-dimensional finite element analysis discussed above, the excess pore water pressure, u_i , is obtained for each element i as a function of time. The ratio $q_i = u_i/\sigma'_{oi}$, where σ'_{oi} is the initial vertical effective stress in the i^{th} element, is commonly thought of as the parameter governing the occurrence of liquefaction [20] and will be referred to herein as the liquefaction index. When q_i reaches a value of 1, the pore water is carrying the load so that soil particles become free to slip and liquefaction takes place. It is also possible for liquefaction to take place at values of q_i slightly less than 1, it only being necessary that most of the lateral strength or bearing capacity is lost. Fardis and Veneziano [19] suggest that the liquefied fraction of the i^{th} element of soil, η_i , be calculated as

$$\eta_i = P \left[\frac{u_i}{\sigma'_{oi}} \geq 0.96 \right] \quad (7.17)$$

for undrained and partially drained effective stress models. The probability on the right hand side can be evaluated through a sequence of simulations. Fardis then went on to evaluate the risk of liquefaction, L , as the probability that the maximum of η_i over the depth z is close to 1,

$$L = \text{P} \left[\max_z(\eta_i) \approx 1 \right]. \quad (7.18)$$

For individual soil columns where interaction with adjacent soil is ignored, such an approach is reasonable since the occurrence of liquefaction at a given layer will result in the loss of lateral resistance at the surface. Ohtomo and Shinozuka [61] have a slightly different approach involving summing the liquefaction indices q over depth to obtain the vertically averaged liquefaction index Q ,

$$Q = \frac{1}{h} \int_0^h \frac{u(z)}{\sigma'_o(z)} dz, \quad (7.19)$$

where h is the total depth of the modeled column. In this way the effect of the vertical extent of a liquefied region can be incorporated into a risk analysis. The question to be answered is how important is the vertical extent of liquefaction? While it certainly has some bearing on the liquefaction risk, it is easy to imagine a situation in which a thin layer some distance below the surface becomes completely liquefied while adjoining layers above and below remain stable. Such a condition would yield a relatively low value of Q even though all lateral stability at the surface may be lost. On the other hand, the vertical extent of liquefied regions may be more important to the occurrence of sand boils and vertical settlement.

In the three-dimensional situation, neither approach was deemed entirely suitable. The occurrence of high q_i indices within an individual column will not necessarily imply liquefaction at the site if adjacent columns retain sufficient strength. Likewise if a particular layer is found to have high q values over a significant lateral extent, then the liquefaction risk could be high even though the average for the site (and the layer) may be low. In this study, the lateral spatial extent of liquefied regions is assumed to be the more important factor in the determination of liquefaction risk for a site. For each realization, the analysis proceeds as follows

- 1) compute the ratio $q_{ij}(t_\ell) = u_i/\sigma'_{oi}$ for each element i in the j^{th} column at each time step t_ℓ and repeat for all the columns,

- 2) compute the sum

$$Q_{i\ell} = \frac{1}{A} \sum_{j=1}^{n_c} q_{ij}(t_\ell) \Delta A_j$$

where A is the total area of the site model, ΔA_j is the area of the j^{th} column and n_c is the number of columns; $Q_{i\ell}$ is the i^{th} layer average stress ratio at each time step t_ℓ .

- 3) determine the indices i^* and ℓ^* which maximize $Q_{i\ell}$. The index i^* now represents the depth of the plane with the maximum likelihood of liquefying at the time t_{ℓ^*} and $q_{i^*j}(t_{\ell^*})$ is the corresponding two-dimensional field of stress ratios (indexed by j).
- 4) determine the excursion areas defined by

$$A_q = \sum_{j=1}^{n_c} I_A \left(q_{i^*j}(t_{\ell^*}) - q \right) \Delta A_j$$

for a variety of levels $q \in (0, 1)$ along with other statistical measures discussed in Chapter 6. The indicator function $I_A(\cdot)$ is defined by (6.2).

Repeating the above steps for a number of realizations allows the estimation of the spatial statistics of the liquefaction indices q on the horizontal plane of maximum liquefaction likelihood. In particular, the excursion areas A_q will be evaluated for $q = \{0.1, 0.2, \dots, 0.9\}$.

To reduce computational costs, DYNA1D analysis of an individual column is discontinued if any of the ratios $q_i = u_i/\sigma'_{oi}$ exceed 0.96. There are a number of reasons for this: First for q_i approaching 1, DYNA1D requires more iterations to converge on a solution (as might be expected since the shear resistance is approaching zero). Secondly, once ratios above about 0.96 are exceeded at a level, that level begins to absorb all the motion and liquefaction indices above that level no longer increase. As also mentioned by Fardis and Veneziano [19], this appears to be a reasonable cutoff point.

In summary, liquefaction of a column will be defined as occurring when the liquefaction index q_i exceeds 0.96 in some element (at which point the analysis of that column is discontinued). The horizontal plane having the highest average liquefaction index is then found and the statistics of those indices determined. This plane will be referred to as the maximal plane. It is recognized that when liquefaction does take place it is not likely to be confined to a horizontal plane of a certain thickness. At the very least the plane could be inclined, but more likely the liquefaction would follow a undulating surface. This level of sophistication is beyond the scope of this initial study and so we confine ourselves herein to the assumption that liquefaction will occur along horizontal planes.

7.6 Monte-Carlo Analysis

Figure 7.12 illustrates two realizations of the maximal plane. Contours are drawn at q indices of 0.3, 0.5, 0.7, and 0.9. In both examples a significant portion of the area has q indices exceeding 0.9 and thus the likelihood of liquefaction is deemed to be relatively high.

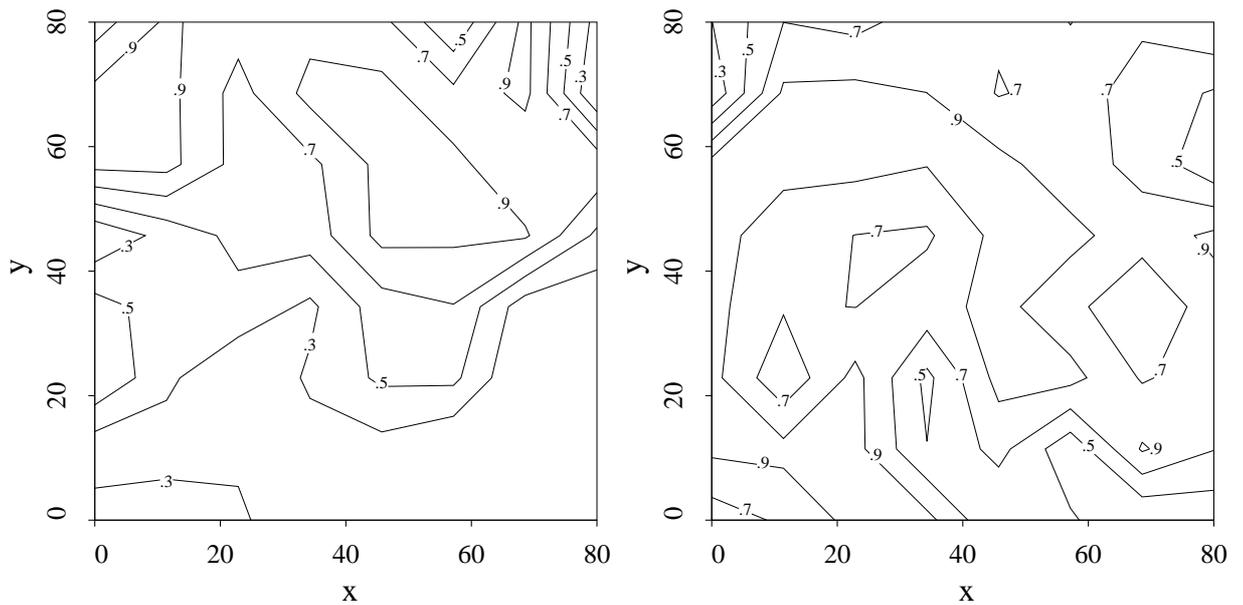


Figure 7.12 Contour maps of the planes having the highest average stress ratio q drawn from two realizations of the soil mass and analyzed by DYNA1D. Contours are drawn at q values of 0.3, 0.5, 0.7, and 0.9. Notice the significant areas with $q > 0.9$.

The four cases considered are summarized in Table 7.6. In the following, we will refer to the first set of simulated ground motions as Event 1 and to the ground motions scaled by a factor of 0.7 as Event 2.

Table 7.6 Ground motion scaling factors and vertical scales of fluctuation considered in the probabilistic liquefaction analysis.

Case	Input Motion Scaling Factor	Scale of Fluctuation	Number of Realizations
1	1.0	1.0	9
2	1.0	5.0	9
3	0.7	1.0	9
4	0.7	5.0	9

The average liquefaction index for the maximal plane, $Q_{i^*l^*}$, is derived from each realization and again averaged over 9 realizations to obtain \bar{Q} . Table 7.7 presents these results

for the 4 cases along with the standard deviation of $Q_{i^*\ell^*}$ (σ_Q) and the average depth to the maximal plane (\bar{D}). Even for Event 1, the average liquefaction index, \bar{Q} , is quite low indicating that it is not the best measure of liquefaction risk (given the knowledge that the Wildlife site is highly liquefiable). It may have some meaning when compared to \bar{Q} values obtained for other sites and earthquakes, but only as a relative measure.

Table 7.7 Average liquefaction indices $Q_{i^*\ell^*}$ of the maximal planes are averaged over the nine realizations to obtain \bar{Q} . The standard deviation of $Q_{i^*\ell^*}$, σ_Q , is estimated for each case and shown along with the average depth, \bar{D} , to the maximal plane.

Event	Scale	\bar{Q}	σ_Q	\bar{D}
1	1.0	0.481	0.252	2.90
1	5.0	0.478	0.227	3.02
2	1.0	0.173	0.129	3.14
2	5.0	0.184	0.113	3.57

The average area of the maximal plane exceeding a threshold liquefaction index q , \bar{A}_q , is shown in Figure 7.13 and the standard deviation of the A_q estimates is shown in Figure 7.14. The excursion area \bar{A}_q is obtained by averaging the A_q values over the nine realizations for each case. At the threshold $q = 0.9$, the coefficient of variation of excursion areas is quite high (1.0) but the trend in Figure 7.13 is evident;

- 1) the scale of fluctuation has little effect on the average excursion areas \bar{A}_q ,
- 2) the intensity of the input motion has a significant effect on the excursion areas. A 30% reduction in input motion intensity reduced the liquefaction index corresponding to $\bar{A}_q = 0.15A_T$ from 0.9 to about 0.3, a three-fold reduction.

It appears that the likelihood of liquefaction of Event 2 is quite low. To some extent, this is substantiated by the fact that the Wildlife site did not liquefy during the Elmore Ranch event ($M_s = 6.2$ compared to the Superstition Hills event, $M_s = 6.6$) [34]. Figure 7.13 suggests a possible approach to the evaluation of liquefaction risk which is again based on the knowledge that the Wildlife site is highly liquefiable: determine the average area of the maximal planes which exceed a liquefaction index of 0.9 and associate a high risk of liquefaction to the site if this area exceeds 10 to 15% of the total area. Such a rule of thumb needs to be substantiated and improved through the analysis of other sites and earthquakes.

Figure 7.15 shows the average number of isolated excursions above the liquefaction index q for each case study and Figure 7.16 shows the corresponding cluster measure. The

cluster measures associated with very low thresholds ($q \leq 0.2$) have been ignored (see Section 6.9) since the excursion areas in these cases are essentially the same as the domain area. Both figures exhibit a much more pronounced effect arising from changes in the scale of fluctuation. Typically, the scale of fluctuation $\theta = 5$ halved the average number of excursions and substantially increased the cluster measure. This implies that for the same total area exceeding a certain index q , the regions show higher clustering at higher scales of fluctuation. In turn higher clustering implies a higher likelihood of liquefaction since there are fewer pockets of ‘resistance’ within the excursion region. Notice that Event 2 typically has higher mean values of Ψ since it has fewer excursions at high thresholds (a single excursion, or no excursions, corresponds to $\Psi \simeq 1$). The likelihood of liquefaction thus cannot depend on the cluster measure alone; it must also take into consideration the total excursion area above a high threshold.

Given the fact that the Wildlife site was known to have liquefied during the Superstition Hills event, the following tentative summary of the results of this chapter can be made;

- 1) The likelihood of liquefaction appears to be most easily quantified by the total area of the domain whose liquefaction indices exceed some threshold index q^* . For this example it appears that the threshold index should be taken as $q^* \simeq 0.9$ and a high likelihood of liquefaction associated with mean total excursion areas A_{q^*} in excess of about 10 to 15% of the total domain area. This criteria quantifies the effect of the earthquake intensity.
- 2) The likelihood of liquefaction can be modified by the cluster measure – as the cluster measure decreases, the liquefiable regions become separated by pockets of resistance and the likelihood of liquefaction decreases. This correction incorporates the effect of the scale of fluctuation on the likelihood of liquefaction.

Both of these measures of the likelihood of liquefaction need considerable empirical calibration before adoption can be considered. However the analysis of a single site and the layout of the methodology is an important start.

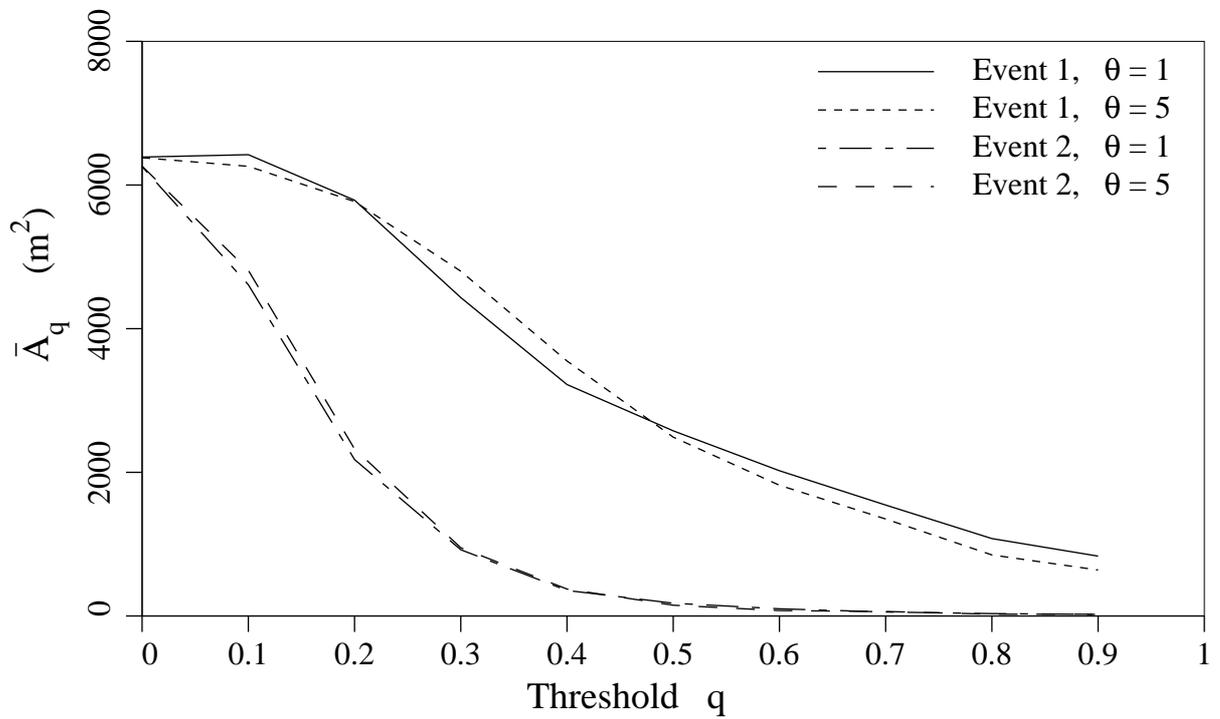


Figure 7.13 Average area of the maximal plane, \bar{A}_q , having liquefaction indices in excess of the indicated q thresholds. Averaging is performed over 9 realizations for each line.

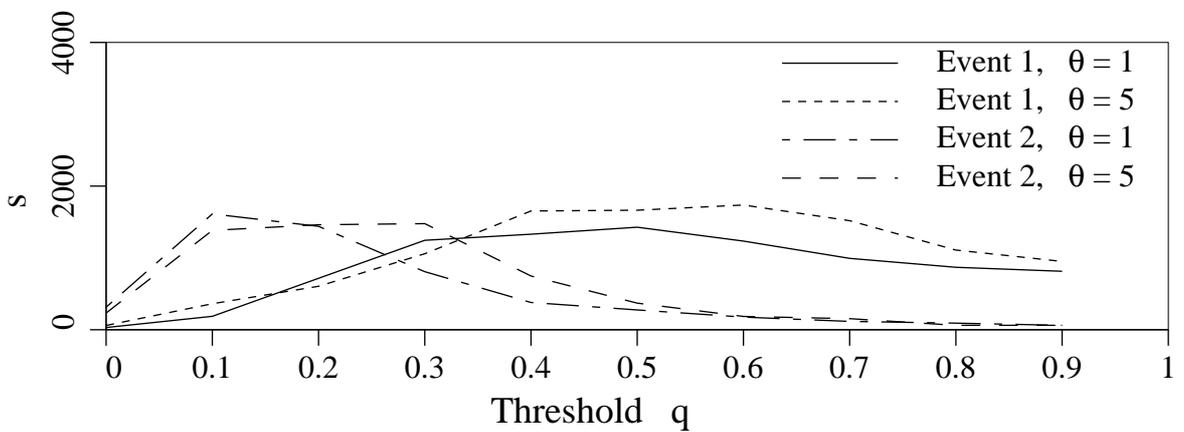


Figure 7.14 Estimated standard deviations, s , of the excursion areas A_q .

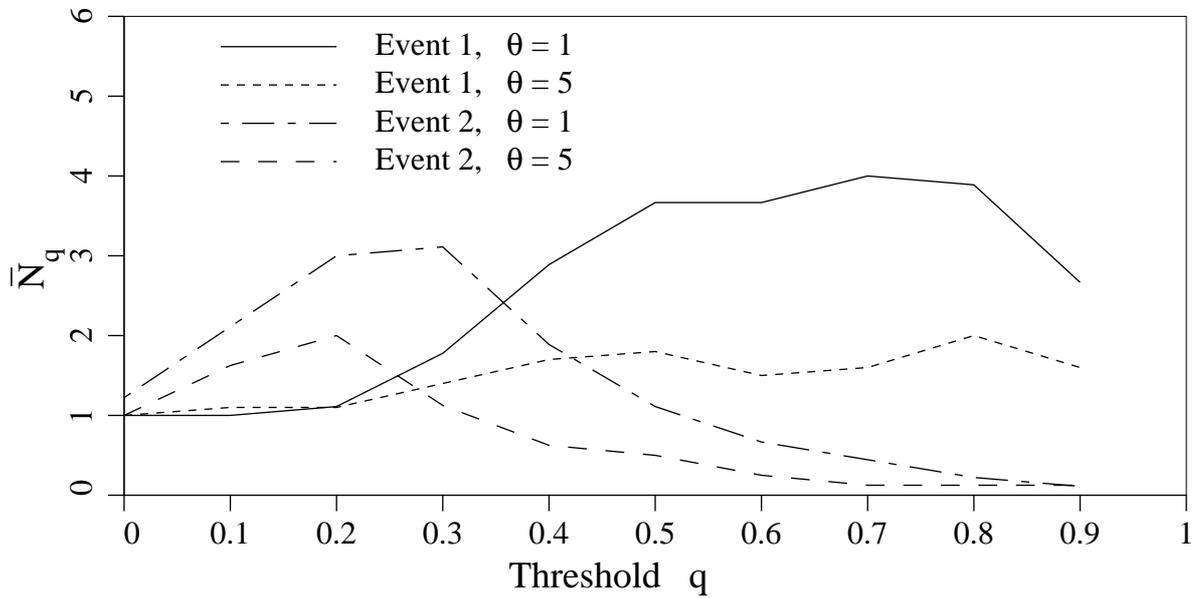


Figure 7.15 Estimated number of isolated excursion areas, \bar{N}_q , above the liquefaction index thresholds q . Each line represents the average over 9 realizations and the coefficient of variation of the estimate averages about 1.0.

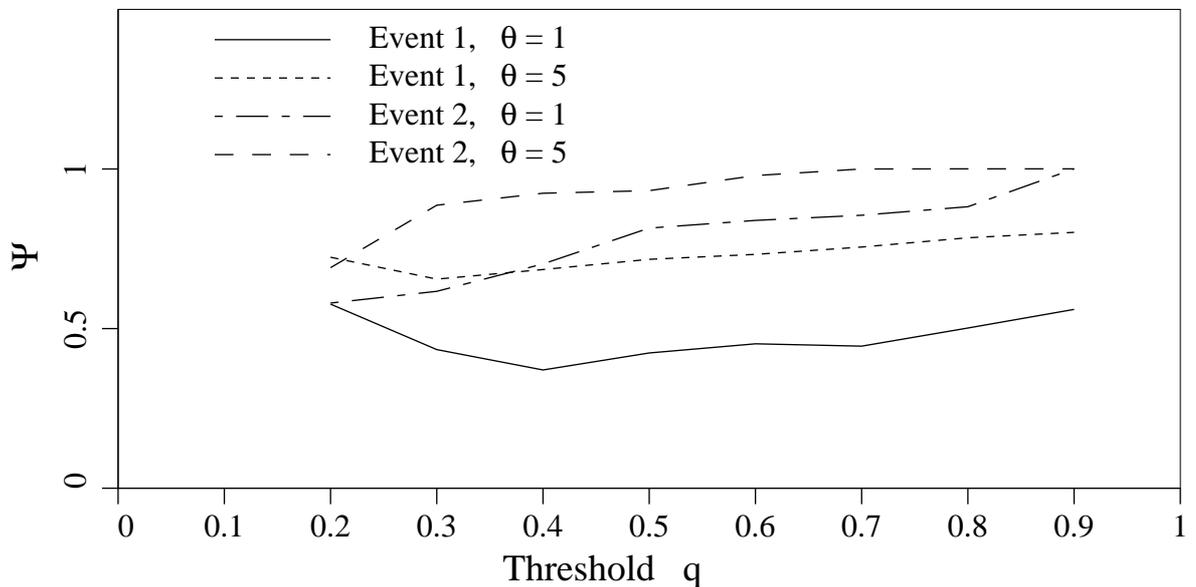


Figure 7.16 Estimated cluster measure, Ψ , of the excursion areas above the liquefaction index thresholds q . Each line represents the average over 9 realizations and the coefficient of variation of the estimate averages about 0.5.

Chapter 8

Conclusions and Recommendations

Concerning existing simulation techniques, which are evaluated in Chapters 2 and 3, the following conclusions can be drawn;

- 1) if sufficient care is taken in the discretization of the field, the FFT (spectral formulation) method is accurate, computationally efficient, and simple to use. The major drawback to the method is that realizations have symmetric covariances, a problem which can be alleviated by generating fields larger than required and ignoring the excess.
- 2) the proposed FFT covariance formulation represents an attempt to reduce errors arising from poor discretization schemes. Although the approach improves the variance estimate under such conditions, it introduces a streaked appearance to the realizations and still leads to a symmetric covariance structure. It is recommended that the use of the FFT method be restricted to properly discretized fields.
- 3) when used in higher dimensions, the Moving Average (MA) method suffers from difficulties in determining suitable weighting functions for arbitrary covariance functions and from computational inefficiency.
- 4) the Turning Bands Method (TBM) requires an equivalent one-dimensional covariance or spectral function which is obtained through the solution of an integral equation. While solutions exist for a number of processes, this requirement renders the method difficult to use in the general case. In addition, realizations produced by TBM have a streaked appearance if an insufficient number of turning bands are used. In three dimensions, the minimum number of bands to use has not been determined (visualization problems hinder such a determination).

As an alternative and complement to these methods, a technique called Local Average Subdivision (LAS) is introduced in Chapter 4 which produces realizations of locally averaged random processes in one, two, or three dimensions. The main features of the LAS method are

- 1) it is easily conditioned both on known data and on portions of the simulated field. The latter allows for changes of resolution in sub-regions of the field, an ability useful

in detailed investigation of local areas and in the use of the field as input to a finite element model.

- 2) it generates local averages of the underlying process that are scale dependent and thus exhibit the correct first- and second-order statistical nature at any resolution. This contrasts with traditional methods where the discrete values represent point samples drawn from the continuous process. Although such point samples are often assumed to be constant over the interval, changes in resolution are not reflected in the statistical nature of the samples. The local averaging property of the LAS method renders its realizations ideal for use as input to finite element codes using efficient, low-order interpolation functions.
- 3) knowledge of the variance function for the process considerably simplifies the algorithm and improves its accuracy. This function can be obtained through the integration of the covariance function (rather than through an integral equation).
- 4) the LAS method is found to run faster than the FFT approach (uncorrected for symmetric covariance) in one dimension and about 1.5 to 2 times slower in two and three dimensions.
- 5) as implemented, the method is restricted to homogeneous, isotropic, Gaussian random fields. Second-order statistics estimated from LAS realizations show the correct convergence to the exact over the entire field and realizations have the desired appearance.

In Chapter 6, the LAS method is employed in a simulation-based study of the statistics of excursions and extrema of two-dimensional Gauss-Markov processes. The simulation-based estimates of the mean total excursion area, mean number and area of isolated excursions, and the integral geometric characteristic of the excursion field are compared with existing theories and matched to semi-empirical relationships. A cluster measure is introduced that is a means of quantifying the degree to which excursions are clustered within the domain. Concerning extrema statistics, a Type I extreme value distribution function with empirically derived parameters is found to match very well the estimated extrema distribution obtained by simulation. The methodology developed in Chapter 6 shows considerable promise both as a means of obtaining useful semi-empirical results and as a guide to the discovery of exact theories. The approach paves the way for a much more detailed study in which other common processes are examined over a wider range of parameters.

Best linear estimation techniques in the frequency domain are incorporated in a new approach to the simulation of optionally conditioned stationary or non-stationary space-time processes and applied to earthquake ground motion simulation in Chapter 5. This method is used along with the LAS algorithm in a Monte-Carlo analysis of site liquefaction risk in

Chapter 7. The soil is modeled as a three-dimensional stochastic medium and input ground motions come from a space-time random field. Four cases are considered corresponding to two different levels of input motion intensity and two scales of fluctuation of the random soil properties. Nine realizations are evaluated for each case. The study is largely concerned with the spatial distribution of the liquefaction indices over horizontal planes deemed to be the most likely to liquefy. Three excursion characteristics of the liquefaction indices are estimated; the mean total area of excursions above a range of threshold indices, the mean number of isolated excursions, and the corresponding cluster measures. The mean total area (or area fraction) of excursions above some critical liquefaction index may be a useful indicator of liquefaction potential at a site. An improved prediction of liquefaction risk could be made by considering the cluster measure – a low value would indicate that liquefied ‘pockets’ are separated by more resistant regions and so the likelihood of liquefaction decreases. Conversely, a high value implies that the liquefied regions are clumped together, raising the likelihood of liquefaction.

A number of similar liquefaction studies of different sites and different earthquakes are required before the relationship between the spatial distribution of liquefaction indices and liquefaction risk can be confidently stated. The procedures described in Chapter 7 lay the groundwork for such future applications. It should be noted that the Monte-Carlo type analysis used in this probabilistic study is ideally suited to parallel architecture computers with little modification of the code. Each realization could be analyzed on separate CPU’s, vastly decreasing the time required to carry out the study (which was a significant limitation here).

References

1. ABRAHAMSON, N.A. and BOLT, B.A., The Spatial Variation of the Phasing of Seismic Strong Ground Motion, *Bulletin of the Seismological Society of America*, **75**(5), 1985, 1247-1264.
2. ADLER, R.J., *The Geometry of Random Fields*, John Wiley & Sons, New York, NY, 1981.
3. ALBRECHT, K.A., LOGSDON, S.D., PARKER, J.C. and BAKER, J.C., Spatial Variability of Hydraulic Properties in the Emporia Series, *Soil Science Society of America Journal*, **49**(6), 1985, 1498-1501.
4. ANDERSON, N., On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis, *Geophysics*, **39**, 1974, 69-72.
5. BENNETT, M.J., McLAUGHLIN, P.V., SARMIENTO, J.S. and YOUND, T.L., *Geotechnical Investigation of Liquefaction Sites, Imperial County, California*, United States Department of the Interior Geological Survey, Open-File Report 84-252, Menlo Park, California, 1984.
6. BLACKMAN, R.B. and TUKEY, J.W., *The Measurement of Power Spectra*, Dover Publications, New York, NY, 1958.
7. BOLT, B.A., ABRAHAMSON, N.A. and YEH, Y.T., The Variation of Strong Ground Motion over Short Distances, in , Vol. 2, Proceedings of the Eighth World Conference on Earthquake Engineering, San Francisco, California, 1984, 183-189.
8. BRILLINGER, D.R. and TIAO, G.C., eds., *Directions in Time Series*, (Proceedings of the IMS Special Topics on Time Series Analysis), Institute of Mathematical Statistics, Iowa State University, May 1-3, 1978.
9. BUIJS, H.L., Implementation of a Fast Fourier Transform (FFT) for Image Processing Applications, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-22**, 1974, 420-424.
10. BURG, J.P., The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra, *Geophysics*, **37**(2), 1972, 375-376.
11. BURG, J.P., *Maximum Entropy Spectral Analysis*, Ph.D. Thesis, Stanford University , Department of Electrical Engineering, Stanford, California, 1975.
12. CARPENTER, L., Computer Rendering of fractal curves and surfaces, in , ACM, New York, SIGGRAPH 80 Proceedings, Seattle, Washington, July 1980, 108-120.

13. CASSEL, D.K., Spatial and Temporal Variability of Soil Physical Properties Following Tillage of Norfolk Loamy Sand, *Soil Science Society of America Journal*, **47**(2), 1983, 196-200.
14. CHAMEAU, J.L. and CLOUGH, G.W., Probabilistic Pore Pressure Analysis for Seismic Loading, *ASCE Journal of Geotechnical Engineering*, **109**(4), 1983, 507-524.
15. COLLINS, M.E. and FENTON, T.E., Statistical Modeling of the Variability of Selected Colo Soil Properties, *Soil Science Society of America Journal*, **48**(5), 1984, 1107-1114.
16. COOLEY, J.W. and TUKEY, J.W., An Algorithm for the Machine Calculation of Complex Fourier Series, *Mathematics of Computation*, **19**(90), 1965, 297-301.
17. CRAMER, H. and LEADBETTER, M.R., *Stationary and Related Stochastic Processes*, John Wiley & Sons, New York, NY, 1967.
18. DUDGEON, D.E. and MERSEREAU, R.M., *Multi-Dimensional Digital Signal Processing*, Prentice-Hall Signal Processing Series, Englewood Cliffs, NJ, 1984.
19. FARDIS, M.N. and VENEZIANO, D., Probabilistic Analysis of Deposit Liquefaction, *ASCE Journal of Geotechnical Engineering*, **108**(3), Mar. 1982, 395-418.
20. FINN, W.D.L., Soil Liquefaction Studies in the People's Republic of China, in *Soil Mechanics – Transient and Cyclic Loads*, Chap. 22, Pande, G.N. and Zienkiewicz, O.C., eds., John Wiley & Sons, New York, NY, 1982.
21. FOURNIER, A., FUSSELL, D. and CARPENTER, L., Computer rendering of stochastic models, *Commun. ACM*, **25**(6), 1982, 371-384.
22. GÓMEZ-HERNÁNDEZ, J.J. and GORELICK, S.M., Effective Groundwater Model Parameter Values: Influence of Spatial Variability of Hydraulic Conductivity, Leakance, and Recharge, *Water Resources Research*, **25**(3), 1989, 405-419.
23. GELHAR, L.W., Stochastic Analysis of Flow in Heterogeneous Porous Media, in *Advances in Transport Phenomena in Porous Media*, NATO Advanced Study Institute on Fundamentals of Transport Phenomena in Porous Media, Bear, J. and Corapcioglu, M.Y., eds., Dordrecht, Boston, MA, July 1985, 675-717.
24. GREMINGER, P.J., SUD, Y.K. and NIELSEN, D.R., Spatial Variability of Field-measured Soil-water Characteristics, *Soil Science Society of America Journal*, **49**(5), 1985, 1075-1081.
25. HARICHANDRAN, R.S. and VANMARCKE, E.H., *Space-Time Variation of Earthquake Ground Motion*, Massachusetts Institute of Technology, Department of Civil Engineering, R84-12, Cambridge, Massachusetts, 1984.

26. HARICHANDRAN, R.S. and VANMARCKE, E.H., *Space-Time Variation of Earthquake Ground Motion*, Massachusetts Institute of Technology, Department of Civil Engineering, R84-12, Cambridge, Massachusetts, 1984.
27. HARICHANDRAN, R.S. and VANMARCKE, E.H., Stochastic Variation of Earthquake Ground Motion in Space and Time, *ASCE Journal of Engineering Mechanics*, **112**(2), 1986, 154-174.
28. HARRIS, D.B., MCCLELLAN, J.H., CHAN, D.S.K. and SCHUESSLER, H.W., Vector Radix Fast Fourier Transform, *Proceeding of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 1977, 548-551.
29. HOLZER, T.L., YOUNG, T.L. and BENNETT, M.J., *In situ* Measurement of Pore Pressure Build-Up During Liquefaction, in , issued by NIST, Jan 1989, Proceedings of the 20th Joint Meeting of United States-Japan Panel on Wind and Seismic Effects, Gaithersburg, MD, 1988, 118-130.
30. IEEE ACOUSTICS, SPEECH, AND SIGNAL PROCESSING SOCIETY, IEEE PRESS, *Programs for Digital Signal Processing*, Digital Signal Processing Committee, Montvale, NJ, 1979.
31. IEEE TRANSACTIONS ON COMPUTERS, *Special Issue on Two-Dimensional Digital Signal Processing*, C-21, 1972.
32. JOURNEL, A.G. and HUIJBREGTS, CH.J., *Mining Geostatistics*, Academic Press, New York, NY, 1978.
33. KACHANOSKI, R.G., ROLSTON, D.E. and DE JONG, E., Spatial Variability of a Cultivated Soil as Affected by Past and Present Microtopography, *Soil Science Society of America Journal*, **49**(5), 1985, 1082-1086.
34. KEANE, C.M. and PREVOST, J.H., An Analysis of Earthquake Data Observed at the Wildlife Liquefaction Array Site, Imperial County, California, in , Technical Report NCEER-89-0032, Proceedings of the Second U.S.-Japan Workshop on Liquefaction, Large Ground Deformations and Their Effects on Lifelines, Buffalo, NY, 1989, 176-192.
35. KHINCHIN, A.Y., Korrelationstheorie der Stationären Stochastischen Prozesse, *Math. Ann.*, **109**, 1934, 604.
36. KOLMOGOROV, A.N., Interpolation and Extrapolation of Stationary Random Sequences, *Izv. Akad. Nauk SSSR, Ser. Mat.*, **5**(3), 1941, .
37. KRIGE, D.G., Two-Dimensional Weighted Moving Averaging Trend Surfaces for Ore Evaluation, in , Proceedings of the Symposium on Mathematical Statistics

- and Computer Applications for Ore Evaluation, Johannesburg, South Africa, 1966, 13-38.
38. LEIBLEIN, J., *A New Method of Analysing Extreme-Value Data*, NACA, Technical Note 3053, 1954.
 39. LEWIS, J.P., Generalized Stochastic Subdivision, *ACM Transactions on Graphics*, **6**(3), 1987, 167-190.
 40. LUMB, P., The Variability of Natural Soils, *ASCE Journal of Geotechnical Engineering*, **3**(2), 1966, 74-97.
 41. MANDELBROT, B.B. and NESS, J.W., Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review*, **10**(4), 1968, 422-437.
 42. MANDELBROT, B.B., *The Fractal Geometry of Nature*, W. H. Freeman and Co., New York, NY, 1982.
 43. MANTOGLOU, A. and WILSON, J.L., *Simulation of Random Fields with the Turning Bands Method*, Massachusetts Institute of Technology, Report #264, Cambridge, MA, 1981.
 44. MARSILY, G., Spatial Variability of Properties in Porous Media: A Stochastic Approach, in *Advances in Transport Phenomena in Porous Media*, NATO Advanced Study Institute on Fundamentals of Transport Phenomena in Porous Media, Bear, J. and Corapcioglu, M.Y., eds., Dordrecht, Boston, MA, 1985, 719-769.
 45. MATHERON, G., The Intrinsic Random Functions and their Applications, *Advances in Applied Probability*, **5**, 1973, 439-468.
 46. MATHERON, G., *Eléments Pour une Théorie des Milieux Poreux*, Masson et Cie, Paris, 1967.
 47. MCCLELLAN, J.H. and RADER, C.M., *Number Theory in Digital Signal Processing*, Prentice-Hall Signal Processing Series, Englewood Cliffs, NJ, 1979.
 48. MOHR, D.L., *Modeling Data as a Fractional Gaussian Noise*, Ph.D. Thesis, Princeton University, Department of Statistics, Princeton, New Jersey, USA, June 1981.
 49. NAGANUMA, T., DEODATIS, G. and SHINOZUKA, M., An ARMA Model for Two-Dimensional Processes, *ASCE Journal of Engineering Mechanics*, **113**(2), 1987, 234-251.
 50. NEWLAND, D.E., *An Introduction to Random Vibrations and Spectral Analysis*, (2nd edition), Longman, New York, NY, 1984.

51. NOBILE, A. and ROBERTO, V., Efficient Implementation of Multi-Dimensional Fast Fourier Transforms on a CRAY X-MP, *Computer Physics Communications*, **40**, 1986, 189-201.
52. NUSSBAUMER, H.J., *Fast Fourier Transform and Convolution Algorithms*, Springer-Verlag, New York–Heidelberg–Berlin, 1981.
53. PEITGEN, H-O. and SAUPE, D., eds., *The Science of Fractal Images*, Springer-Verlag, New York, NY, 1988.
54. POLGE, R.J., BHAGAVAN, B.K. and CARSWELL, J.M., Fast Combinational Algorithms for Bit Reversal, *IEEE Transactions on Computers*, **C-23**, 1974, 1-9.
55. PREVOST, J.H., *A Computer Program for Nonlinear Seismic Response Analysis*, National Center for Earthquake Engineering Research, NCEER-89-0025, Buffalo, NY, 1989.
56. PRIESTLEY, M.B., *Spectral Analysis and Time Series*, Vol. Volume 2: Multivariate Series, Prediction and Control, Academic Press, New York, NY, 1981.
57. PUGACHEV, V.S., *Probability Theory and Mathematical Statistics for Engineers*, (Translated by I.V. Sinitsyna), Eykhoff, P., ed., Pergamon Press, New York, NY, 1984.
58. RALSTON, A. and WILF, H.S., *Mathematical Methods for Digital Computers*, (Volumes 1 and 2), John Wiley & Sons, New York, NY, 1967.
59. RIVARD, G.K., Direct Fast Fourier Transform of Bivariate Functions, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-25**(3), 1977, 250-252.
60. RODRIGUEZ-ITURBE, I., Scale of Fluctuation of Rainfall Models, *Water Resources Research*, **22**(9), 1986, 15S-37S.
61. SHINOZUKA, M. and OHTOMO, K., Spatial Severity of Liquefaction, *Proceedings of the Second U.S.-Japan Workshop on Liquefaction, Large Ground Deformations and Their Effects on Lifelines*, 1989.
62. SHINOZUKA, M. and JAN, C.M., Digital Simulation of Random Processes and it Applications, *Journal of Sound and Vibration*, **25**(1), 1972, 111-128.
63. SKINNER, D.P., Pruning the Decimation-in-Time FFT Algorithm, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-24**, 1976, 193-194.
64. TOMPSON, A.F.B., ABABOU, R. and GELHAR, L.W., Implementation of the Three-Dimensional Turning Bands Random Field Generator, *Water Resources Research*, **25**(10), 1989, 2227-2243.

65. VANMARCKE, E.H., Spatial Averaging of Earthquake Ground Motion: The Effect on Response Spectra, in , Proceedings of the EPRI Workshop on Engineering Characterization of Small-Magnitude Earthquakes, Jan. 1987.
66. VANMARCKE, E.H., Reliability of Earth Slopes, *ASCE Journal of Geotechnical Engineering*, **103**(GT11), 1977, 1247–1265.
67. VANMARCKE, E.H., *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Massachusetts, 1984.
68. VOSS, R., Random Fractal Forgeries – From Mountains to Music, in , London, England, Conference on Science and Uncertainty, Mar. 1984.
69. WHITTLE, P., On the Variation of Yield Variance with Plot Size, *Biometrika*, **43**, 1956, 337-343.
70. WIENER, N., *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, The MIT Press, Cambridge, Massachusetts, 1949.
71. YAGLOM, A.M., *An Introduction to the Theory of Stationary Random Functions*, Dover, Mineola, NY, 1962.
72. YAMAZAKI, F. and SHINOZUKA, M., Digital Generation of Non-Gaussian Stochastic Fields, in *Stochastic Mechanics, Vol. I*, Shinozuka, M., ed., Columbia University, New York, NY, June 1987, 209-234.
73. YOUNG, L.S., *Computer Programs for Minimum Cost FFT Algorithms*, Rockwell International, Rep. T77-986/501, Anaheim, California, 1977.

Appendix A

Some Additional Results for Multi-Dimensional FFT Simulation

A.1 2-D Covariance Formulation

Use of the known covariance function,

$$B_{k\ell} = \text{E} [Z_{ij} Z_{i+k,j+\ell}] \quad (\text{A.1})$$

in (2.48) and (2.49) leads to

$$\begin{aligned} \text{E} [\mathcal{A}_{mn}]^2 = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1,\ell-1} \sum_{i=1}^{K_1-k+1} \sum_{j=1}^{K_2-\ell+1} (C_{im,jn} C_{i+k-1,m,j+\ell-1,n} \\ + C_{im,j+\ell-1,n} C_{i+k-1,m,jn}) \quad (\text{A.2}) \end{aligned}$$

$$\begin{aligned} \text{E} [\mathcal{B}_{mn}]^2 = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1,\ell-1} \sum_{i=1}^{K_1-k+1} \sum_{j=1}^{K_2-\ell+1} (S_{im,jn} S_{i+k-1,m,j+\ell-1,n} \\ + S_{im,j+\ell-1,n} S_{i+k-1,m,jn}) \quad (\text{A.3}) \end{aligned}$$

where,

$$C_{im,jn} = \cos \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right) \quad (\text{A.4})$$

$$S_{im,jn} = \sin \left(\frac{2\pi(i-1)(m-1)}{K_1} + \frac{2\pi(j-1)(n-1)}{K_2} \right) \quad (\text{A.5})$$

$$\delta_{k\ell} = \begin{cases} \frac{1}{2} & \text{when } k = \ell = 1 \\ 1 & \text{when } k = 1 \text{ or } \ell = 1 \\ 2 & \text{otherwise} \end{cases}$$

Defining the following four functions

$$X_{mn} = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1,\ell-1} (K_1 - k + 1)(K_2 - \ell + 1) C_{mk}^1 C_{n\ell}^2 \quad (\text{A.6})$$

$$Y_{mn} = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1,\ell-1} S_{mk}^1 S_{n\ell}^2 \quad (\text{A.7})$$

$$V_{mn} = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1,\ell-1} (K_1 - k + 1) C_{mk}^1 S_{n\ell}^2 \quad (\text{A.8})$$

$$W_{mn} = \frac{1}{(K_1 K_2)^2} \sum_{k=1}^{K_1} \sum_{\ell=1}^{K_2} \delta_{k\ell} B_{k-1, \ell-1} (K_2 - \ell + 1) S_{mk}^1 C_{n\ell}^2 \quad (\text{A.9})$$

which can be evaluated using a series of Fast Fourier transforms, where

$$C_{mk}^\alpha = \cos\left(\frac{2\pi(m-1)(k-1)}{K_\alpha}\right) \quad (\text{A.10})$$

$$S_{mk}^\alpha = \sin\left(\frac{2\pi(m-1)(k-1)}{K_\alpha}\right), \quad (\text{A.11})$$

then the variances of the Fourier coefficients can be expressed as

$$\text{E}[A_{mn}]^2 = R_{mn} X_{mn} - \frac{P_m^1 C_n^2}{S_n^2} V_{mn} - \frac{P_n^2 C_m^1}{S_m^1} W_{mn} + I_{mn} Y_{mn} \quad (\text{A.12})$$

$$\text{E}[B_{mn}]^2 = \bar{R}_{mn} X_{mn} + \frac{P_m^1 C_n^2}{S_n^2} V_{mn} + \frac{P_n^2 C_m^1}{S_m^1} W_{mn} - I_{mn} Y_{mn} \quad (\text{A.13})$$

where,

$$R_{mn} = \begin{cases} 2 & \text{if } m = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_2}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$\bar{R}_{mn} = \begin{cases} 0 & \text{if } m = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_2}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$I_{mn} = \begin{cases} \frac{C_m^1 C_n^2}{S_m^1 S_n^2} - 1 & \text{if } m = 2, 3, \dots, \frac{K_1}{2} \text{ and } n = 2, 3, \dots, \frac{K_2}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$P_m^\alpha = \begin{cases} 1 & \text{if } m = 1 \text{ or } 1 + \frac{K_\alpha}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$C_m^\alpha = \cos(2\pi(m-1)/K_\alpha) \quad (\text{A.14})$$

$$S_m^\alpha = \sin(2\pi(m-1)/K_\alpha) \quad (\text{A.15})$$

A.2 3-D Spectral Formulation

In Chapter 2 the procedure for obtaining the spectral formulation variances of the Fourier coefficients was explored. Following a similar procedure for the 3-D FFT simulation,

$$Z_{ijk} = \sum_{\ell=1}^{K_1} \sum_{m=1}^{K_2} \sum_{n=1}^{K_3} \left\{ \mathcal{A}_{\ell mn} C_{i\ell, jm, kn} + \mathcal{B}_{\ell mn} S_{i\ell, jm, kn} \right\} \quad (\text{A.16})$$

where,

$$Z_{ijk} = Z\left((i-1)\Delta x_1, (j-1)\Delta x_2, (k-1)\Delta x_3\right) \quad (\text{A.17})$$

$$C_{i\ell,jm,kn} = \cos \left(\frac{2\pi(i-1)(\ell-1)}{K_1} + \frac{2\pi(j-1)(m-1)}{K_2} + \frac{2\pi(k-1)(n-1)}{K_3} \right) \quad (\text{A.18})$$

$$S_{i\ell,jm,kn} = \sin \left(\frac{2\pi(i-1)(\ell-1)}{K_1} + \frac{2\pi(j-1)(m-1)}{K_2} + \frac{2\pi(k-1)(n-1)}{K_3} \right) \quad (\text{A.19})$$

The variances of the independent, zero-mean, normally distributed Fourier coefficients $\mathcal{A}_{\ell mn}$ and $\mathcal{B}_{\ell mn}$ can now be expressed as

$$\text{E} [\mathcal{A}_{\ell mn}]^2 = \frac{1}{16} \delta_{\ell mn}^A \Delta\omega \left(G_{\ell mn}^d + G_{\ell mN}^d + G_{\ell Mn}^d + G_{Lmn}^d + G_{\ell MN}^d + G_{LmN}^d + G_{LMn}^d + G_{LMN}^d \right) \quad (\text{A.20})$$

$$\text{E} [\mathcal{B}_{\ell mn}]^2 = \frac{1}{16} \delta_{\ell mn}^B \Delta\omega \left(G_{\ell mn}^d + G_{\ell mN}^d + G_{\ell Mn}^d + G_{Lmn}^d + G_{\ell MN}^d + G_{LmN}^d + G_{LMn}^d + G_{LMN}^d \right) \quad (\text{A.21})$$

where,

$$\Delta\omega = \Delta\omega_1 \Delta\omega_2 \Delta\omega_3$$

$$L = K_1 - \ell + 2$$

$$M = K_2 - m + 2$$

$$N = K_3 - n + 2$$

$$G_{\ell mn}^d = G^d(\omega_{1\ell}, \omega_{2m}, \omega_{3n}) \quad (\text{see Equation 2.52})$$

$$\delta_{\ell mn}^A = \begin{cases} 2 & \text{if } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } m = 1 \text{ or } 1 + \frac{K_2}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$\delta_{\ell mn}^B = \begin{cases} 0 & \text{if } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } m = 1 \text{ or } 1 + \frac{K_2}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 1 & \text{otherwise} \end{cases}$$

A.3 3-D Covariance Formulation

Using the inverse relationships for the 3-D FFT leads to the determination of the Fourier coefficients $\mathcal{A}_{\ell mn}$ and $\mathcal{B}_{\ell mn}$ in terms of the known covariance function

$$B_{ijk} = \text{E} [Z_{pqr} Z_{p+i,q+j,r+k}],$$

as follows

$$\text{E} [\mathcal{A}_{\ell mn}]^2 = \frac{1}{(K_1 K_2 K_3)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_3} \sum_{p=1}^{K_1} \sum_{q=1}^{K_2} \sum_{r=1}^{K_3} B_{p-i,q-j,r-k} C_{i\ell,jm,kn} C_{pl,qm,rn} \quad (\text{A.22})$$

$$\text{E} [\mathcal{B}_{\ell mn}]^2 = \frac{1}{(K_1 K_2 K_3)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_3} \sum_{p=1}^{K_1} \sum_{q=1}^{K_2} \sum_{r=1}^{K_3} B_{p-i,q-j,r-k} S_{i\ell,jm,kn} S_{pl,qm,rn} \quad (\text{A.23})$$

where C and S are defined by (A.18) and (A.19) respectively. If the field is assumed to be quadrant symmetric then (A.22) and (A.23) reduce to

$$\text{E} [\mathcal{A}_{\ell mn}]^2 = \frac{1}{(K_1 K_2 K_3)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_3} \delta_{ijk} B_{ijk} h^A(i, j, k, \ell, m, n) \quad (\text{A.24})$$

$$E [\mathcal{B}_{\ell mn}]^2 = \frac{1}{(K_1 K_2 K_3)^2} \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \sum_{k=1}^{K_3} \delta_{ijk} B_{ijk} h^B(i, j, k, \ell, m, n) \quad (\text{A.25})$$

where h^A and h^B are functions of trigonometric functions,

$$\begin{aligned} h^A = & R_{\ell mn}^a (K_1 - i + 1)(K_2 - j + 1)(K_3 - k + 1) C_{i\ell}^1 C_{jm}^2 C_{kn}^3 \\ & - P_{\ell mn}^a (K_1 - i + 1)(K_2 - j + 1) C_{i\ell}^1 C_{jm}^2 S_{kn}^3 \\ & - P_{\ell mn}^b (K_1 - i + 1)(K_3 - k + 1) C_{i\ell}^1 S_{jm}^2 C_{kn}^3 \\ & - P_{\ell mn}^c (K_2 - j + 1)(K_3 - k + 1) S_{i\ell}^1 C_{jm}^2 C_{kn}^3 \\ & + Q_{\ell mn}^a (K_1 - i + 1) C_{i\ell}^1 S_{jm}^2 S_{kn}^3 \\ & + Q_{\ell mn}^b (K_2 - j + 1) S_{i\ell}^1 C_{jm}^2 S_{kn}^3 \\ & + Q_{\ell mn}^c (K_3 - k + 1) S_{i\ell}^1 S_{jm}^2 C_{kn}^3 \\ & + R_{\ell mn}^b S_{i\ell}^1 S_{jm}^2 S_{kn}^3 \end{aligned} \quad (\text{A.26})$$

$$\begin{aligned} h^B = & \bar{R}_{\ell mn}^a (K_1 - i + 1)(K_2 - j + 1)(K_3 - k + 1) C_{i\ell}^1 C_{jm}^2 C_{kn}^3 \\ & + P_{\ell mn}^a (K_1 - i + 1)(K_2 - j + 1) C_{i\ell}^1 C_{jm}^2 S_{kn}^3 \\ & + P_{\ell mn}^b (K_1 - i + 1)(K_3 - k + 1) C_{i\ell}^1 S_{jm}^2 C_{kn}^3 \\ & + P_{\ell mn}^c (K_2 - j + 1)(K_3 - k + 1) S_{i\ell}^1 C_{jm}^2 C_{kn}^3 \\ & - Q_{\ell mn}^a (K_1 - i + 1) C_{i\ell}^1 S_{jm}^2 S_{kn}^3 \\ & - Q_{\ell mn}^b (K_2 - j + 1) S_{i\ell}^1 C_{jm}^2 S_{kn}^3 \\ & - Q_{\ell mn}^c (K_3 - k + 1) S_{i\ell}^1 S_{jm}^2 C_{kn}^3 \\ & - R_{\ell mn}^b S_{i\ell}^1 S_{jm}^2 S_{kn}^3 \end{aligned} \quad (\text{A.27})$$

and the functions C^α and S^α are defined by (A.14) and (A.15). The coefficients in (A.26) and (A.27) are given by

$$\delta_{ijk} = \begin{cases} \frac{1}{2} & \text{when } i = j = k = 1 \\ 1 & \text{when } i = j = 1 \text{ or } i = k = 1 \text{ or } j = k = 1 \\ 2 & \text{when } i = 1 \text{ or } j = 1 \text{ or } k = 1 \\ 4 & \text{otherwise} \end{cases}$$

$$R_{\ell mn}^a = \begin{cases} 2 & \text{when } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } m = 1 \text{ or } 1 + \frac{K_2}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$\bar{R}_{\ell mn}^a = \begin{cases} 0 & \text{when } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } m = 1 \text{ or } 1 + \frac{K_2}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$R_{\ell mn}^b = F_\ell^1 + F_m^2 + F_n^3 - F_\ell^1 F_m^2 F_n^3$$

$$F_p^\alpha = \begin{cases} 0 & \text{when } p = 1 \text{ or } 1 + \frac{K_\alpha}{2} \\ \frac{\cos(2\pi(p-1)/K_\alpha)}{\sin(2\pi(p-1)/K_\alpha)} & \text{otherwise} \end{cases}$$

$$P_{\ell mn}^a = \begin{cases} F_n^3 & \text{when } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } m = 1 \text{ or } 1 + \frac{K_2}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$P_{\ell mn}^b = \begin{cases} F_m^2 & \text{when } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$P_{\ell mn}^c = \begin{cases} F_\ell^1 & \text{when } m = 1 \text{ or } 1 + \frac{K_2}{2} \text{ and } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$Q_{\ell mn}^a = \begin{cases} F_m^2 F_n^3 - 1 & \text{when } \ell = 1 \text{ or } 1 + \frac{K_1}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$Q_{\ell mn}^b = \begin{cases} F_\ell^1 F_n^3 - 1 & \text{when } m = 1 \text{ or } 1 + \frac{K_2}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$Q_{\ell mn}^c = \begin{cases} F_\ell^1 F_m^2 - 1 & \text{when } n = 1 \text{ or } 1 + \frac{K_3}{2} \\ 0 & \text{otherwise} \end{cases}$$

Making use of these relationships, (A.24) and (A.25) can be calculated through a series of Fast Fourier transforms in much the same manner as was illustrated for the 2-D case.

Appendix B

Spectral Analysis of One-Dimensional LAS Processes

B.1 Spectral Analysis

The spectral analysis of realizations produced by the one-dimensional LAS method will be accomplished through a Fast Fourier Transform of the sequence to obtain the complex Fourier coefficients \mathcal{X}_k as in (2.18),

$$\mathcal{X}_k = \frac{1}{K} \sum_{j=1}^K \mathcal{Z}_j e^{-i2\pi(k-1)(j-1)/K}, \quad (B.1)$$

and then taking the spectral estimate of the single realization, \hat{G} , at the frequency ω_k to be

$$\hat{G}(\omega_k) = |\mathcal{X}_k|^2 \quad (B.2)$$

for $k = 1, 2, \dots, K$. Since the analysis is performed on simulations, no smoothing of the Fourier estimate will be made. The variance of the spectral estimate will be reduced by averaging over N_f realizations to obtain

$$\bar{G}(\omega_k) = \frac{1}{N_f} \sum_{j=1}^{N_f} \hat{G}(\omega_k), \quad (B.3)$$

where \bar{G} is the desired estimate at the FFT frequencies

$$\omega_k = \frac{2\pi(k-1)(K-1)}{KL}, \quad (B.4)$$

and L is the physical length of the process.

An alternate method of estimating the spectral power is the so-called Maximum Entropy (ME) approach, details of which can be found in Burg [11] and Anderson [4]. This method will also be used as a comparison in the following analyses.

B.2 Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process has a simple exponential covariance function,

$$B(\tau) = \sigma^2 \exp \left\{ -\frac{2|\tau|}{\theta} \right\}, \tag{B.5}$$

and corresponding spectral density function

$$G(\omega) = \frac{4\sigma^2\theta}{\pi(4 + \theta^2\omega^2)}. \tag{B.6}$$

Vanmarcke [25] expresses the spectral density function of the local average process, $G_D(\omega)$, in terms of $G(\omega)$

$$G_D(\omega) = G(\omega) \left[\frac{\sin(\omega D/2)}{(\omega D/2)} \right]^2, \tag{B.7}$$

where D is the averaging dimension. Figure B.1 shows the estimated spectral density functions using the FFT and ME methods versus the exact given by (B.7). The simulated process has a physical length of 5 subdivided into 256 intervals giving an averaging length of $D = 5/256$. A scale of fluctuation of $\theta = \frac{1}{2}$ was used and the spectral estimates were averaged over 500 realizations. Both the FFT and ME estimates show excellent agreement with the exact.

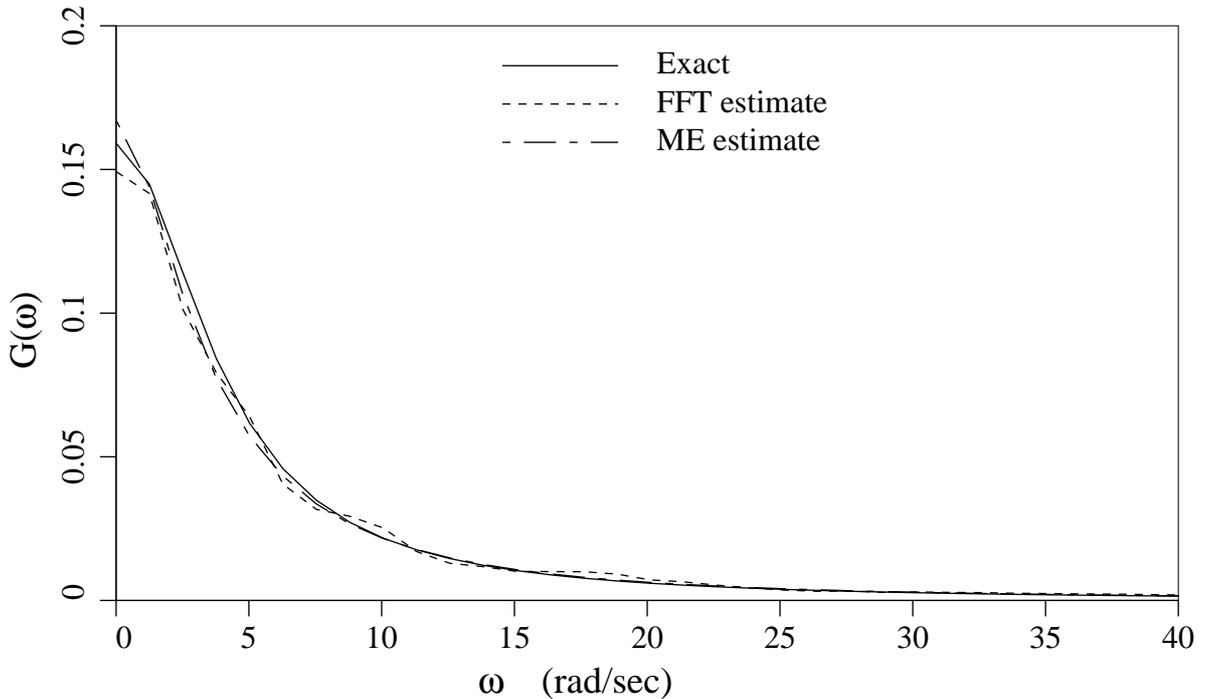


Figure B.1 Spectral estimates of the locally averaged Ornstein-Uhlenbeck process for $\theta = \frac{1}{2}$ averaged over 500 realizations.

B.3 Damped Oscillatory Noise

The damped oscillatory noise used to investigate the required neighborhood size in Chapter 4 has an oscillatory covariance function of the form

$$B(\tau) = \sigma^2 \cos(\lambda\tau) \exp\left\{-\frac{2|\tau|}{\theta}\right\}, \quad (B.8)$$

where we used $\lambda = 8$. The corresponding exact spectral density function of the locally averaged process is

$$G_D(\omega) = \frac{2\theta\sigma^2}{\pi} \left[\frac{1}{4 + \theta^2(\lambda - \omega)^2} + \frac{1}{4 + \theta^2(\lambda + \omega)^2} \right] \left[\frac{\sin(\omega D/2)}{(\omega D/2)} \right]^2. \quad (B.9)$$

The estimated and exact spectras are compared in Figures B.2 and B.3 for neighborhood sizes of 3 and 5 respectively. Again averaging of the 256 point process ($L = 5$) was performed over 500 realizations although in this case a scale of fluctuation of $\theta = 4$ was used (smaller scales of fluctuation are less interesting since the oscillations in the covariance are too rapidly damped out). As with the covariance estimate, a neighborhood of 5 gives much better agreement between the exact and estimated functions. Notice that for this process, the FFT estimate seems to be preferable to the ME estimate.

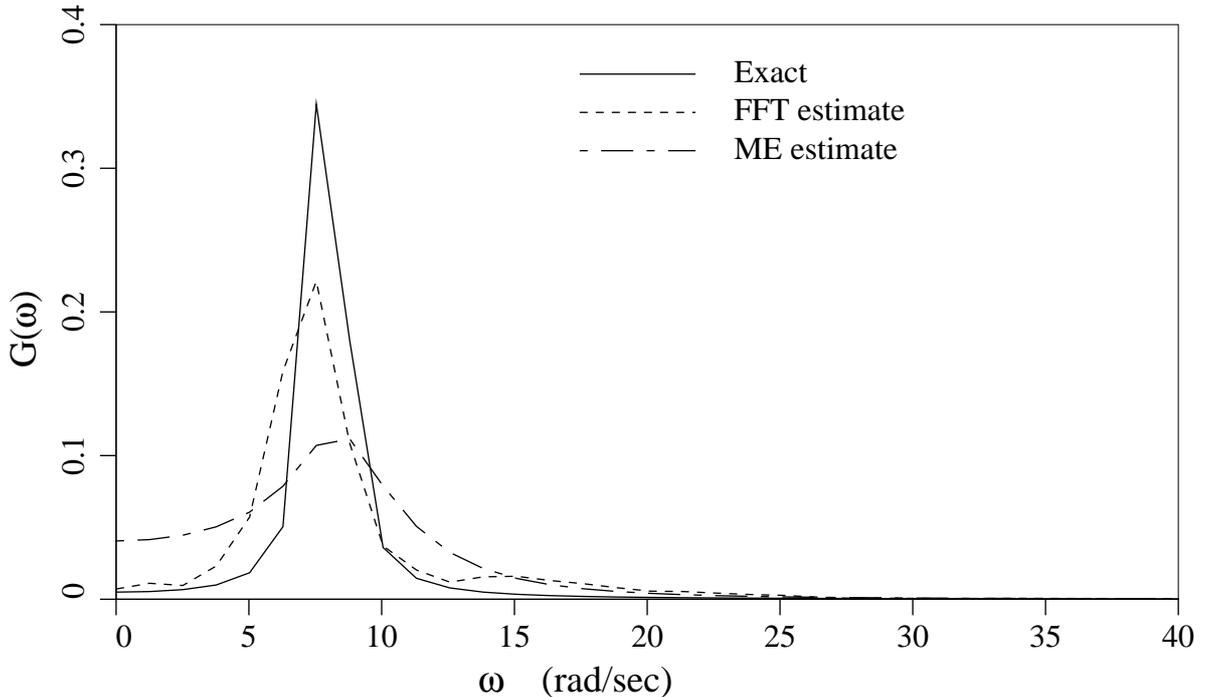


Figure B.2 Spectral estimates of the damped oscillatory noise for $\lambda = 8$ and $\theta = 4$ using a neighborhood size of 3.

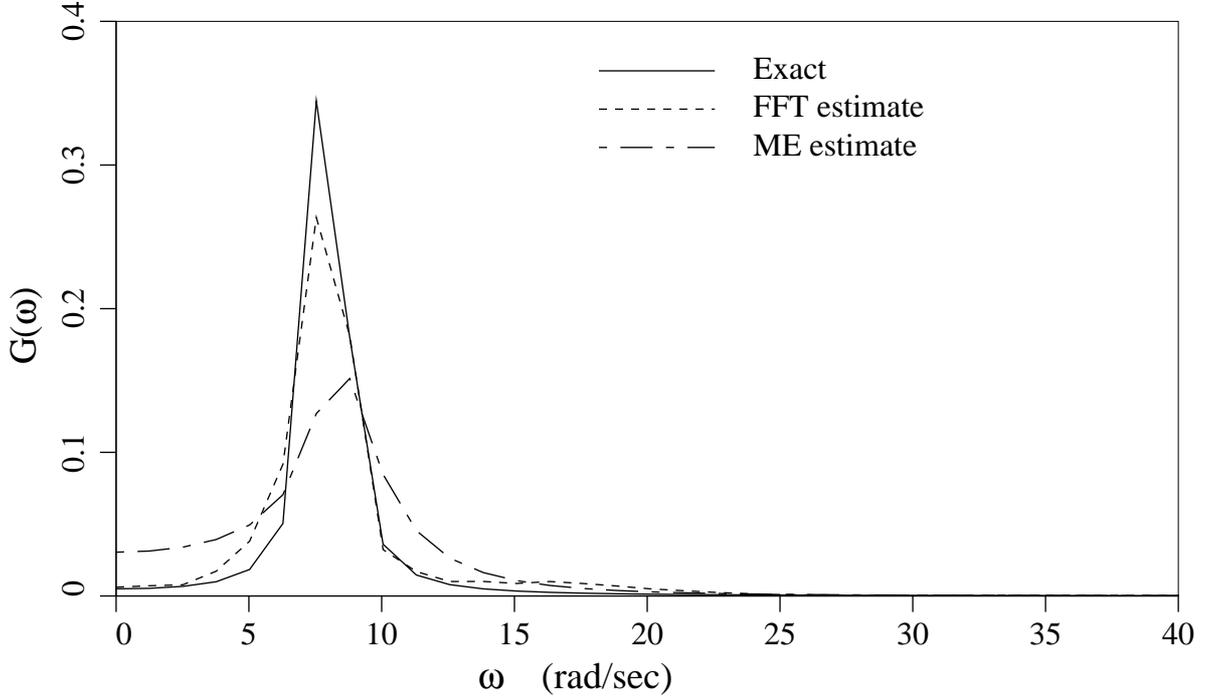


Figure B.3 Spectral estimates of the damped oscillatory noise for $\lambda = 8$ and $\theta = 4$ using a neighborhood size of 5.

B.4 Fractional Gaussian Noise

Perhaps the most interesting process to obtain spectral estimates of is that of fractional Gaussian noise as defined by Mandelbrot and van Ness [41]. The process itself is obtained by performing a small amount of averaging of fractional Gaussian motion over some distance δ and then taking the derivative. The resulting covariance function is

$$B(\tau) = \frac{\sigma^2}{2\delta^{2H}} \left[|\tau + \delta|^{2H} - 2|\tau|^{2H} + |\tau - \delta|^{2H} \right], \quad (B.10)$$

in which the parameter H is the self-similarity parameter. This covariance function is defined for $0 < H < 1$ (in fact for $H = 1$, it loses positive-definiteness). The corresponding spectral density function was derived by Mandelbrot and van Ness for small $\delta\omega$ to be

$$G_D(\omega) = 2V_H \left(\frac{\pi(2H-1)H}{\Gamma(2-2H)\cos\pi(H-1)} \right) (2\pi\omega)^{1-2H} \left[\frac{\sin(\omega D/2)}{(\omega D/2)} \right]^2, \quad (B.11)$$

where $\Gamma(\cdot)$ is the gamma function and V_H is defined by

$$V_H = \frac{1}{[\Gamma(H + \frac{1}{2})]^2} \left[\frac{1}{2H} + \int_{-\infty}^0 \left[(1-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right]^2 ds \right], \quad (B.12)$$

which, unfortunately, is a divergent integral for $H > 0.5$. Monin and Yaglom's [] work leads to the following spectral density function corrected for the local average process,

$$G_D(\omega) = \frac{\Gamma(1 + 2H) \sin \pi H}{\pi \omega^{2H-1}} \left[\frac{\sin(\omega D/2)}{(\omega D/2)} \right]^2, \quad (B.13)$$

which is tractable and the relationship we will take as the exact. For a process with physical length of 5 ($D = 5/256$), Figure B.4 shows the exact versus estimated spectral density function averaged over 500 realizations of fractional Gaussian noise with $H = 0.95$. The spectra is plotted on a log-log scale to demonstrate the $\omega^{-\beta}$ type noise ($\beta = 2H - 1$). The slope of the line is about -0.9 as expected. Figure B.5 illustrates the similar nature of the process when the physical length is taken to be 500 rather than 5. Now the size of the LAS cell is $D = 500/256$ and the frequency range over which the FFT estimate is made is much narrower. Nevertheless, if anything, the FFT spectral estimates are closer to the exact. The curve in the tail of the exact spectral density arises due to the local averaging correction factor which becomes significant for wavelengths approaching the size of the LAS cell.

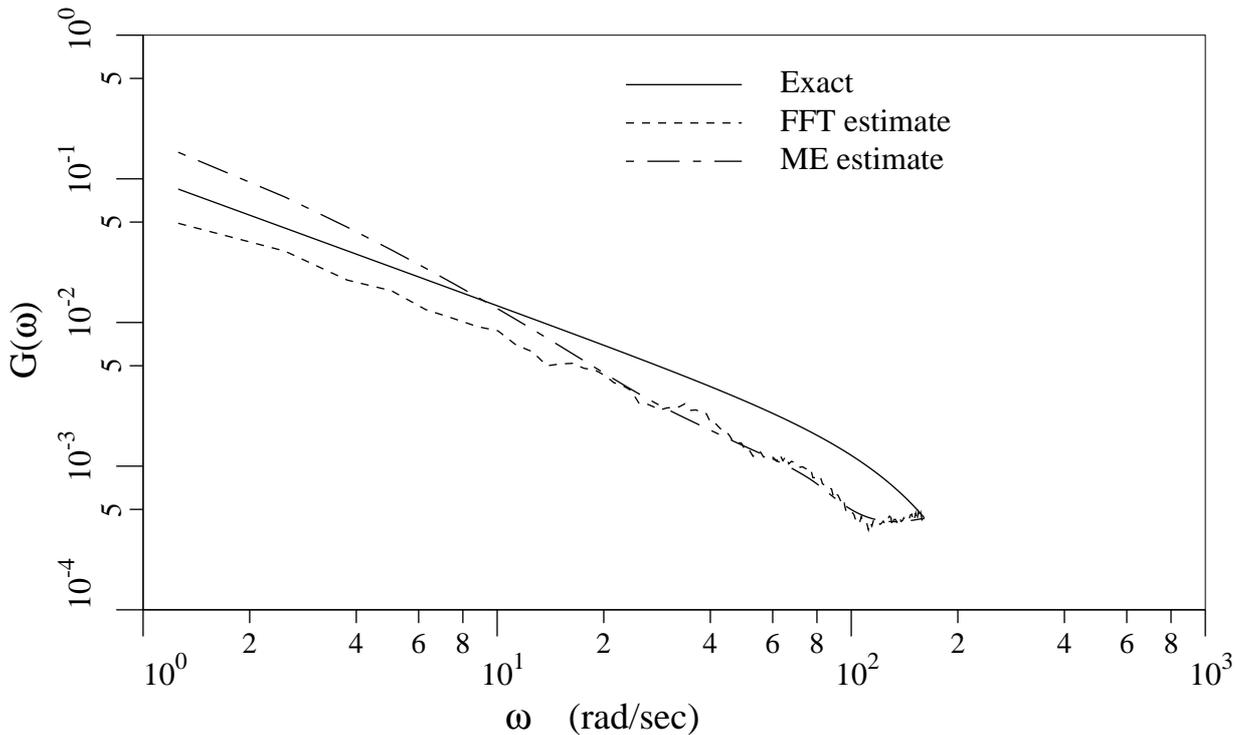


Figure B.4 Spectral estimates of fractional Gaussian noise with self-similarity parameter $H = 0.95$ and a physical process length of 5 divided into 256 intervals.

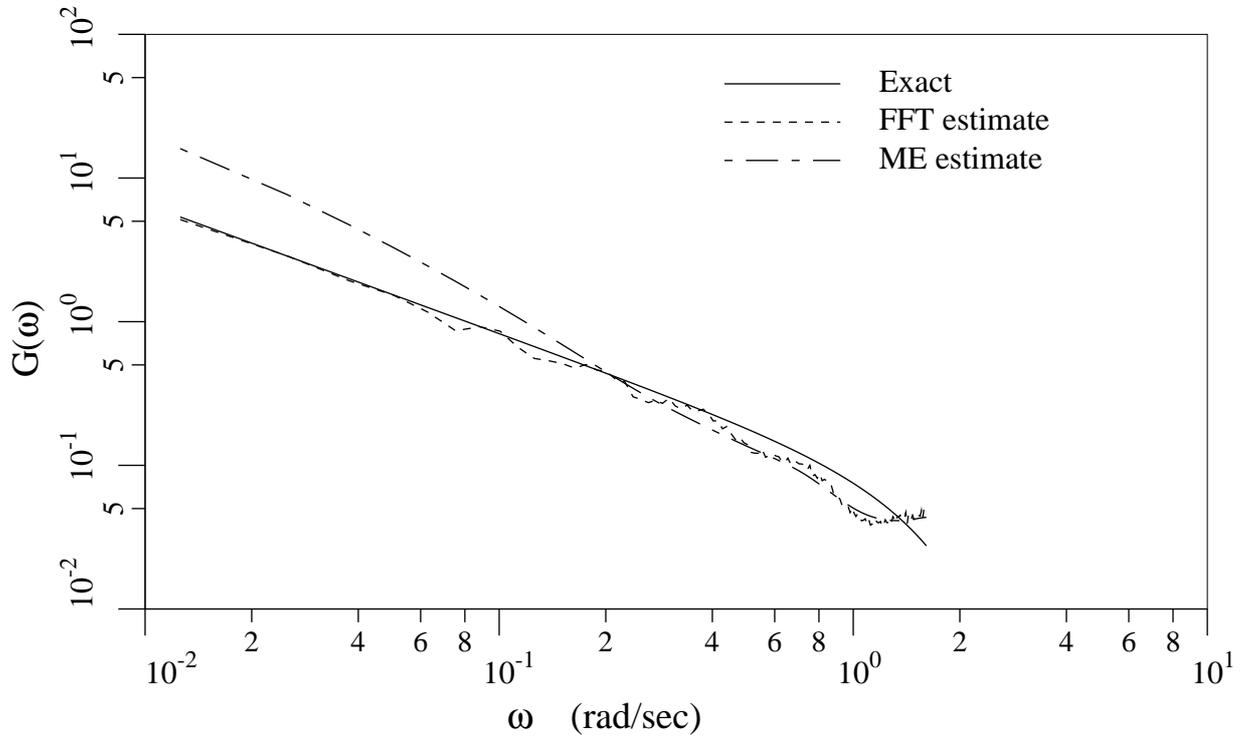


Figure B.5 Spectral estimates of fractional Gaussian noise with self-similarity parameter $H = 0.95$ and a physical process length of 500 divided into 256 intervals.