1996

A Poisson-driven Stationary Process Model In Spectroscopy

Alfred Geoffrey Benn

Follow this and additional works at: https://ir.lib.uwo.ca/digitizedtheses

Recommended Citation
https://ir.lib.uwo.ca/digitizedtheses/2679

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.
A Poisson-Driven Stationary Process Model in Spectroscopy

by

Alfred G. Benn

Department of Statistical and Actuarial Sciences

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
June 1996

© Alfred Geoffrey Benn 1996
The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L’auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L’auteur conserve la propriété du droit d’auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-15026-7
Abstract

A variety of physical and biological phenomena can be modeled by filtered Poisson processes, that is, by Poisson-driven stationary real-valued processes

\[ X(t) = \int_0^t h(s, \tau; u)N(ds), \quad t \in \mathcal{R}, \]

where \( \tau \) denotes an arbitrary occurrence point and \( u \) the associated mark of a marked Poisson process. In these models, the output or response process can be expressed as a sum of the individual responses to each marked occurrence point.

The present study is concerned with specific filtered Poisson process models of two biochemical spectroscopy imaging techniques, Scanning Fluorescence Correlation Spectroscopy (S-FCS) and its extension, Image Correlation Spectroscopy (ICS). These techniques are used in studying certain aspects of cell membrane physiology, specifically the rate or extent of membrane receptor protein clustering or aggregation. This phenomenon has been observed directly by electron microscopy. These measurements constitute an important step towards understanding how such receptors influence or interact with other cell membrane components.

The data obtained from S-FCS and ICS are measurements of fluorescence intensity as a function of position. The intensity process is modeled as

\[ X(t) = \int_{S_k} f_k(x,y) Y_{(x,y)} N(dx\,dy) + \gamma \varepsilon(t), \quad t \in \mathcal{Z}^d \]

where \( d = 1 \) (S-FCS) or \( d = 2 \) (ICS),

\[ f_k(x,y) = \exp\left(-\frac{1}{\omega^2}\left((x-t_1\delta)^2 + (y-t_2\delta)^2\right)\right), \]
$S_{2\delta}$ is a disc centred at $(t_1\delta, t_2\delta)$ of radius $\frac{3\delta}{\sqrt{2}}$. $Y_{(x,y)}$ is the nonnegative integer-valued random mark associated with the occurrence point $(x, y)$ and $N(\cdot)$ is the spatial Poisson counting process. The term $\gamma(t)\epsilon$ is a Gaussian (background) noise process which represents experimental recording error. The observed data constitute a realization of a finite-range (or $m-$) dependent stationary process (S-FCS) or random field (ICS).

Certain second-order moment functionals of $X(t)$ contain important model parameters related to protein cluster sizes and densities. Specifically, in the absence of additive noise ($\gamma = 0$), the quantity

$$R \propto \frac{1}{N}$$

where $R = \frac{\text{Var}(X(t))}{\mu_X^2}$ and $N$ is the average number of fluorescent protein clusters in the observed volume. The ratio $R$ is a function of two other parameters, $\theta_1 = \lambda \mu_{2Y}$ and $\theta_2 = \mu$ where $\lambda$ is the Poisson intensity parameter and $\mu_{2Y}$ is the second moment of the marking distribution. The main purpose of this study is the consistent interval estimation of these parameters. Nonlinear least squares methods are proposed involving sample autocovariances (S-FCS) and periodograms (ICS). An asymptotic log likelihood procedure is also discussed for ICS.

The estimators are shown to obey central limit theorems as a consequence of the finite-range or mixing dependence of $\{X(t)\}$. Thus, one is able to obtain approximate confidence intervals for important model parameters such as $R$ provided consistent estimators of limiting variances are obtained. Much of this work is concerned with this variance estimation problem and an effective method is proposed.

An extensive Monte Carlo simulation study is undertaken to investigate the
asymptotics of the least squares estimator $\hat{\theta}$ and the ratio estimator $\hat{R}$ and to implement a variance estimation method. Due to the computationally intensive nature of the ICS calculations, these are performed on a MasPar MP-2 massively parallel (2K) computer. Finally, several data sets are analyzed for both S-FCS and ICS.
Acknowledgements

I would like to generously thank Reg Kulperger, without whose supervision, guidance and insight, this thesis would not have been possible. His dedication and leadership were vital to my statistical development.

John Braun was instrumental in the early stages of this work. I thank him for our many conversations. I thank Nils Petersen for proposing and encouraging this interesting collaborative venture as well as suggesting the use of the MasPar computer. Thanks also to those in his lab, in particular Paul St-Pierre and Mamta Srivastava for providing me with detailed information of these spectroscopic techniques as well as actual experimental datasets. I thank Lifu Guo for his tireless and unselfish assistance with the MasPar. Without him, the task of using the MasPar would have been considerably more difficult if not impossible. I thank Alicia Pleasence for her expert assistance in the preparation of this thesis. Only through her unsurpassed LaTeX skills was the final form of this thesis realized.

Finally I thank my family and friends for their generous support and encouragement through the years.

I gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada for their years of postgraduate scholarship support. I also thank the faculty of the Department of Statistical and Actuarial Sciences at the University of Western Ontario and Dr Bruce Smith of Dalhousie University for providing many helpful comments and suggestions.

I dedicate this thesis to my grandfather, Alfred W.H. Benn.
# Table of Contents

Certificate of Examination ........................................ ii
Abstract ...................................................................... iii
Acknowledgements ...................................................... vi
Table of Contents ....................................................... vii
List of Figures .......................................................... xi
List of Tables .......................................................... xii

1 Introduction ............................................................. 1
   1.1 Background ..................................................... 1
   1.2 Scanning Fluorescence Correlation Spectroscopy ............ 3
   1.3 Image Correlation Spectroscopy .................................. 8
   1.4 A Stochastic Model for S-FCS and ICS ......................... 12
   1.5 Summary ........................................................ 14

2 Mathematical Preliminaries ........................................ 18

3 Statistical Inference for Scanning Fluorescence Spectroscopy 30
   3.1 Introduction ..................................................... 30
   3.2 A Mathematical Model for S-FCS ............................. 34
   3.3 The Least Squares Estimators .................................. 36
   3.4 Proofs ............................................................. 39
      3.4.1 Proof of Theorem 3.1 .................................... 39
      3.4.2 Proof of Theorem 3.2 ................................. 42
3.4.3 Proof of Theorem 3.3 ........................................ 43
3.4.4 Proof of Theorem 3.4 ........................................ 45
3.5 Asymptotic Results for $X_k$ Defined Over $\mathbb{R}^2$ ........ 45
3.6 Moment Calculations ........................................... 48
3.7 Components of $\Sigma$ ........................................... 52
   3.7.1 $\Sigma_{11}$ .................................................. 52
   3.7.2 $\Sigma_{12}$ .................................................. 52
   3.7.3 $\Sigma_{22}$ .................................................. 54
3.8 Tuning Parameters ............................................. 57
3.9 Numerical Results ............................................ 62
   3.9.1 Point Estimates ........................................... 64
   3.9.2 Variance Estimation ...................................... 66
3.10 S-FCS Data .................................................. 71
3.11 Discussion and Summary ................................... 78

4 Statistical Inference for Image Correlation Spectroscopy ........ 81
4.1 Introduction .................................................. 81
4.2 A Mathematical Model for ICS ................................ 85
4.3 Model Parameter Estimation ................................... 88
4.4 Least Squares Estimation ................................... 89
4.5 Quasi-Maximum Likelihood Estimation ......................... 92
4.6 Proofs ....................................................... 95
   4.6.1 Proof of Theorems 4.1 and 4.5 ......................... 99
4 6.2 Proof of Theorems 4.2 and 4.6: Consistency of Asymptotic m.i.e. \( \hat{\theta} \) ........................................... 114

4.6.3 Proof of Theorem 4.3 ........................................... 118

4.6.4 Proof of Theorem 4.4 ........................................... 119

4.7 Cumulant Calculations ........................................... 120

4.8 Summary ......................................................... 126

5 Massively Parallel Computation ..................................... 128

5.1 MasPar MP-2 ..................................................... 128

5.2 MasPar Hardware ................................................ 130

5.3 MasPar Software ................................................ 130

5.4 Introduction ...................................................... 130

5.5 Calculation of Limiting Covariances .............................. 133

5.5.1 Riemann Sum Approximation .................................. 134

5.5.2 Monte Carlo Integration ....................................... 137

5.6 Simulation and Estimation ....................................... 141

5.6.1 Simulation of Poisson-Driven Random Field .................. 142

5.6.2 Least Squares Estimation ..................................... 144

5.6.3 Variance Estimation .......................................... 146

5.7 ICS Data ......................................................... 161

5.8 Discussion and Summary ........................................ 162

6 Concluding Remarks and Further Work ................................ 164

Bibliography ..................................................... 168
List of Figures

1.2.1 Confocal image of large fluorescent beads .................................................. 6
1.3.2 Confocal scanning laser microscope image (512 × 512 pixels) of small beads .......................................................... 10
1.3.3 Confocal scanning laser microscope image (512 × 512 pixels) of large beads .......................................................... 11
3.8.1 Analytic Variance of $\hat{\theta}_1$ Versus Lag for various $\lambda$ and $\delta$, with $\beta = 0.5$ ........................................... 59
3.8.2 Analytic Variance of $\hat{R}$ Versus Lag for various $\lambda$ and $\delta$, with $\beta = 0.5$ ........................................... 60
3.9.3 Histograms of $\hat{R}$, Truncated at 3 Std Deviations of $\hat{\theta}_2$, and limiting Normal Approximation .................................................. 65
3.10.4 Petersen Data e5c330, $\hat{R} = .0125$ .................................................. 75
4.1.1 Confocal microscope image and resulting fluorescence intensity image as a function of pixel position .......................... 83
5.2.1 MasPar MP-2 architecture .................................................. 131
5.2.2 Two-dimensional representation of MasPar PE Array .................................. 132
5.6.3 Histograms of $X$, $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{R}$ .................................................. 147
5.6.4 MasPar X-Net communication network .................................................. 148
List of Tables

3.1 Optimal Lags ................................................. 61
3.2 Parameter Cases for Simulations ......................... 63
3.3 Average Estimators, No Trimming, \( w = .235 \) .................. 67
3.4 Average Estimators, Trimming at 3 Standard Deviations of \( \hat{\theta}_2 \), \( w = .235 \) .................................................. 68
3.5 Standard deviation estimator for \( \theta_1 \) divided by true standard deviation. \( \lambda = 1, \theta_2 = .235 \) ........................................ 72
3.6 Standard deviation estimator for \( \theta_2 \) divided by true standard deviation. \( \lambda = 1, \theta_2 = .235 \) ........................................ 73
3.7 Standard deviation estimator for \( R \) divided by true standard deviation. \( \lambda = 1, \theta_2 = .235 \) ........................................ 74
3.8 Analysis of Petersen's Data, at lag = 78 ....................... 77
3.9 e5c332 Estimates of \( \theta_2 \) and standard errors ................... 78
Chapter 1

Introduction

1.1 Background

The following background information, taken from [24] and [32], summarizes the biological characteristics of the spectroscopy models considered in this work.

Many biological processes occur in the cell membrane. It is through such processes that the cell interacts with its surrounding environment. Cell membrane receptor molecules control the flow of information into the cell through a process known as transmembrane signalling. These receptors bind to external agents such as hormones and in doing so activate the cell to perform a certain function.

Much of the biochemistry of this receptor activation is well known. However, aspects of the mechanisms governing the binding of a certain external agent to its cell surface receptor are not fully understood. The activity of enzymes and other membrane proteins is regulated by receptors possibly through direct interaction in the cell membrane. When an external agent (ligand) binds to these receptors, this interaction is altered.

It is believed that the cell membrane is in a fluid state and that most mem-
brane proteins are free to diffuse. However, this diffusion is restricted in two ways. A fraction appears to have their movement restricted completely while another moves freely but more slowly than should be the case in a fluid membrane. Current models, which examine the mechanism whereby protein diffusion is restricted or prevented, suggest a dependence upon interactions of proteins with other membrane components and the spatial distribution of such cell surface components. The phenomenon may also depend upon the particular type of cell and surface protein.

Various theoretical models have addressed the dynamics of ligand binding to cell surface receptors. In summary, the process involves the diffusion of the ligand in the area surrounding the cell surface followed by the binding to a specific receptor. In the presence of a large number of cell surface receptors, the binding rate should depend upon the rate of ligand diffusion and the receptor density or distribution.

However, receptor-receptor interactions on the cell surface may also influence this binding process as evidence exists of the micro clustering of membrane protein receptors upon binding of ligands (hormones) such as growth factors and insulin. This protein-protein aggregation or clustering may be necessary for intracellular biological responses and also may account for observed restricted protein diffusion. It would thus seem that quantitative measurements of protein receptor distributions on cell surfaces is an important aspect in understanding receptor diffusion, receptor-receptor clustering and ligand-receptor interaction.

Membrane protein receptor aggregation or clustering has been directly observed by electron microscopy. However, this technique has been unable to provide firm
evidence of an association between ligand concentrations and binding rates due to inherent difficulties in controlling sampling resolutions. Fluorescence energy transfer experiments have provided evidence for protein clustering but have been unable to measure cluster sizes. In general, quantitative measurements of the extent of protein clustering have been difficult to obtain.

1.2 Scanning Fluorescence Correlation Spectroscopy

In Petersen [24], a new experimental technique was introduced for making quantitative measurements of the distribution of fluorescently labelled proteins on the surface of living cells. The method is called Scanning Fluorescence Correlation Spectroscopy (S-FCS) and to date has provided the highest quality measurements of cluster sizes and densities.

The experimental technique known as spectroscopy refers to the measurement of the absorption or the emission of particles from a specific source material when these particles are excited by an external radiation source. The degree of absorption or emission is a function of the particles’ energy. In its most common form, spectroscopy involves electromagnetic radiation such as visible light where the particles involved are photons. The photon energy is inversely proportional to the radiation wavelength. This wavelength varies depending upon the specific spectroscopic application considered. The entities in a biological specimen, for example, that absorb light, are known as chromophores. These may consist of small individual units such as the individual amino acids comprising a protein molecule.
When a chromophore absorbs a photon, it is put into a higher or excited energy state. This state usually lasts for only a fraction of a second and when it starts to decay, this extra energy is lost. This energy decay can involve the emission of light in a process known as fluorescence.

S-FCS is based on Concentration Correlation Spectroscopy and was initially used to measure the molecular weights of DNA fragments. Using fluorescence as a measure of concentration, it was adapted by Magde et al. [20] to Fluorescence Correlation Spectroscopy (FCS) and was used to measure diffusion or directed flow rates of proteins through cell membranes. It was through this application that the potential of the technique for measuring protein aggregation was proposed.

In FCS, concentration is measured by fluorescence intensity $i(t)$. However, the protein diffusion rate in cell membranes was too slow to measure changes in fluorescence intensity on any practical time scale.

In S-FCS, these fluorescence intensity measurements are not obtained in real time but are instead determined by the position of a focused laser beam on the cell. The cell is translated through the beam in a fluorescence microscope resulting in intensity data as a function of position across the cell. A typical microscope image is depicted in Figure 1.2.1. Certain correlation-based quantities are then determined which estimate the protein cluster density (number of clusters per unit area) as well as the average number of protein monomers in the cluster in a given area. The observed fluctuations in fluorescence intensity arise from the random variation of the number and sizes of protein molecules in the region illuminated by the beam.
The primary objective of S-FCS is to obtain measurements of the surface density and aggregation sizes of cell surface receptors. In theory, protein aggregate densities could be determined by counting the total number of receptors per cell and dividing this result by the cell surface area. Alternatively, one could determine the number of receptors in discs of fixed area over certain sections of the cell surface and average these results [32]. It is this latter method to which S-FCS corresponds. The discs referred to above represent the area of the cell surface illuminated by the laser beam while focused at that position. The fluorescence intensity measurement is essentially the count of the number of receptors within the disc after allowing for local variation in the laser beam intensity or power.

In practice, the analysis of S-FCS data is based on correlation functions. Of primary interest is the magnitude of the zero-lag correlation or variance of the fluorescence intensity as this is associated with the average number of molecules in the area illuminated by the beam. The correlation function decay is proportional to the laser beam intensity profile and is used to fit empirical correlation functions, via a least squares method, in order to obtain correlation zero-lag values. This fitting procedure is necessary as the fluorescence intensity variance cannot be estimated directly from experimental data due to photon counting noise, which causes a sharp peak at the zero-lag correlation.

In the notation of Petersen [24], let \( i(t) \) denote the fluorescence intensity reading obtained at position \( t \Delta x \) on the cell surface where \( \Delta x \) is the known sampling interval. It has been shown that \( Var(i(t)) \) is inversely related to the mean number of molecules (aggregates) \( \bar{N}_a \), in the observation area. The variance is the zero-lag
Figure 1.2.1: Confocal image of large fluorescent beads
value $G(0)$ of the intensity autocorrelation function and

$$\frac{G(0)}{i^2} = \frac{1}{N_a}$$

where $G(\tau) = <\delta i(t) \delta i(t + \tau)>$, $\delta i(t) = i(t) - \bar{i}$, $\bar{i}$ is the mean intensity and $<\cdot>$ denotes the average or expected value. Let $g(0) = G(0)/i^2$. It is also shown in Petersen [24] that $g(0) \cdot \bar{i}$ is proportional to the mean number of monomers per aggregate. These facts justify the importance of determining $g(0)$ from S-FCS data as this quantity directly measures protein cluster sizes and densities.

In practice, the determination of $g(0)$ proceeds by a least squares fitting procedure. The theoretical correlation function is fit to empirical correlation data and the $g(0)$ value is obtained by extrapolation. The theoretical function is determined by a Gaussian laser beam intensity profile. This Gaussian assumption is standard in laser optics applications such as S-FCS. Also, the width of the beam, $w$, can be measured by a technique known as the knife-edge method which employs gold-plated microslides (St-Pierre [31]). As such, the beam width is assumed known. Given the S-FCS fluorescence intensity record, the empirical correlations are calculated. The actual correlation function is derived to be

$$g(\alpha) = g(0)e^{-\left(\frac{\alpha}{w}\right)^2},$$  \hspace{1cm} (1.1)

assuming an infinite data record length. As the sample correlations are based upon finite intensity records, (1.1) is modified to

$$g(\alpha) = g(0)e^{-\left(\frac{\alpha}{w}\right)^2} + g_0,$$  \hspace{1cm} (1.2)

where $g_0$ is added to compensate for the empirical correlations not obtaining zero for suitably large lag values. The model (1.2) is fit to the observed correlation data.
by least squares involving the three parameters: $g(0), g_0$ and $w$. Although $w$ has been previously measured, its estimation serves as an internal consistency check for the model. Only those data sets with an estimated beam $w$ that is within 30% of the actual measured width are retained. Extreme $w$ values indicate very large noise levels or clusters that are too large compared to the beam size.

1.3 Image Correlation Spectroscopy

S-FCS is an excellent experimental method for measuring protein receptor aggregation. However, the measurements tend to be slow requiring 2–4 min/scan per cell and several hours of prescanning cell preparation, and while the measurements tend to be quite precise, a desired level of accuracy is achieved only after multiple measurements on many different cells have been averaged. Also, only a few laboratories around the world have been able to make these measurements, as S-FCS requires expensive, specialized equipment.

Recently, Petersen et.al [25] introduced an extension of S-FCS called Image Correlation Spectroscopy (ICS) a technique which is based on a quantitative analysis of confocal scanning laser microscopy (CSLM) images, depicted in Figures 1.3.2 and 1.3.3. This method of measuring receptor distributions improves upon certain aspects of S-FCS.

Firstly, these microscopes can collect image data in a matter of seconds from an entire cell or a collection of cells, providing much more information than would be obtained from several one-dimensional scans. Also, confocal microscopes are being used in a larger number of laboratories, increasing the availability of a technique
for measuring aggregation. The disadvantage of ICS is an increase in sampling variability leading to a subsequent loss of some precision. Nonetheless, it is still believed that because of the increase in accuracy and speed, ICS maintains some advantage over S-FCS.

As with S-FCS, the analysis of the spatial (lattice) intensity data obtained from ICS proceeds from the calculation of correlation functions (Petersen et. al. [24])

\[ g(\xi, \eta) = \langle \delta i(x, y) \delta i(x + \xi, y + \eta) \rangle \]

where

\[ \delta i(x, y) = \frac{i(x, y) - \langle i(x, y) \rangle}{\langle i(x, y) \rangle} \]

and \( \langle \cdot \rangle \) denotes average value as before. Also, in an ideal system where background photon counting noise is absent,

\[ g(0, 0) = \frac{1}{N_a} \]

that is, the variance of the observed fluorescence intensities is inversely proportional to the average number of fluorescent aggregates \((N_a)\) in the value being observed. The other fundamental relation involving aggregate sizes

\[ g(0, 0) = \langle i(x, y) \rangle \propto N_m \]

where \(N_m\) is the mean number of monomers per cluster, likewise carries over from S-FCS.

In practice, \(g(0, 0)\) contains a sharp peak at \((0, 0)\) due to background noise and so must be estimated by the least squares extrapolation method described earlier,
Figure 1.3.2: Confocal scanning laser microscope image (512 × 512 pixels) of small beads
Figure 1.3.3: Confocal scanning laser microscope image (512 × 512 pixels) of large beads
involving the theoretical autocorrelation function

\[ g(\xi, \eta) = g(0,0)e^{-\frac{\xi^2 + \eta^2}{w^2}}. \]

Empirical correlations are determined by

\[ G(\xi, \eta) = \mathcal{F}^{-1}([\mathcal{F}(i(x, y))]\mathcal{F}^*(i(x, y))) \]

where \( \mathcal{F} \) is the fast Fourier transform with complex conjugate transpose \( \mathcal{F}^* \) and

\[ G(\xi, \eta) = \langle i(x, y)i(x + \xi, y + \eta) \rangle. \]

The normalized correlation \( g(\xi, \eta) \) is

\[ g(\xi, \eta) = \frac{G(\xi, \eta)}{< i(x, y) >^2} - 1. \]

Then \( g(0,0) \) is estimated by fitting

\[ g(\xi, \eta) = g(0,0)e^{-\frac{\xi^2 + \eta^2}{w^2}} + g_0 \]

to the observed correlation data as with S-FCS.

Henceforth in this thesis, the quantity \( g(\cdot) \) is denoted by \( R \). Later, a stochastic model-based estimation procedure is proposed which leads to consistent large-sample interval estimates of \( R \).

### 1.4 A Stochastic Model for S-FCS and ICS

The biophysical characteristics associated with these spectroscopy techniques suggest that a point process based model should provide a reasonable mathematical approximation to the mechanisms involved in generating the observed fluorescence.
intensity data obtained in these experiments. From these characteristics, a mechanistic model can be derived.

Specific external agents or ligands are fluorescently labeled on the cell surface. Certain membrane protein receptors, upon binding to these ligands, become observable when the cell surface is translated under a fluorescence or confocal scanning laser microscope.

The laser beam is assumed to follow a two-dimensional Gaussian intensity profile. This is a common assumption in laser optics as a surface viewed through a small aperture produces an image filtered by a local (near 0) sine curve. It is this sine curve which is well approximated by the Gaussian kernel [32]. When the beam is focused at position \( t\delta \), \( t \in \mathbb{Z}^2 \), a certain region of the cell surface is illuminated. This region is represented by a disc centred at \( t\delta \) of finite radius. This radius is a function of the laser beam width.

Within these illuminated disc regions a random number of fluorescently labeled protein receptors are visible in the microscope. These protein clusters are of varying size depending upon the number of individual units or monomers which comprise the aggregate.

The concentration measurement made at position \( t\delta \) is the fluorescence intensity at that site. Large clusters closest to the beam centre make the greatest contribution to the intensity reading. The overall contribution of a given aggregate in the observation region is weighted according to the Gaussian kernel.

The above comments suggest that a sensible model for the aggregate process is a spatial marked Poisson process. Each occurrence point represents the location
of the receptor while a nonnegative random integer mark represents the number of monomers comprising the aggregate. To incorporate the laser intensity, each point event must be filtered by a Gaussian kernel. Thus, the fluorescence intensity measurement at position $\xi \delta$, $X(\xi)$, may be represented by

$$X(\xi) = \sum_{i=1}^{N(\xi_\xi \xi)} f_\xi(x_i, y_i)Y_{(x_i, y_i)} + \gamma \xi(t)$$ (1.3)

where $f_\xi(\cdot)$ is the Gaussian kernel, $Y_{(x_i, y_i)}$ is the random mark associated with the occurrence point $(x_i, y_i)$ and $\gamma(\xi)$ is an uncorrelated noise process with parameter $\gamma$ representing photon counting background noise. Equation (1.3) is a version of a Poisson-driven real-valued stationary process known as a filtered Poisson process. A more formal discussion of these processes may be found in Chapter 2. This model is consistent with one proposed by correlation spectroscopy experimentalists (N.O. Petersen).

### 1.5 Summary

This section provides an outline of the contents and main results of the thesis.

Chapter 2 collects together the principal mathematical results required in the thesis. A formal discussion of filtered Poisson processes and random counting integrals is provided. These are integrals defined with respect to random counting measures such as point processes. Certain technical results are required for the derivation of the asymptotic distributions of specific estimators. These include the delta method, the Cramer-Wold device and Slutsky’s Theorem. The central limit theorems in Chapters 3 and 4 are based on finite-range or $m$-dependent
limit theorems for stationary processes. Statements of these are included. The calculations of the limiting variances in Chapter 4 involve the cumulant calculus of Leonov and Shiryaev [19] and joint cumulants of finite Fourier transforms for stationary random fields given in Brillinger [5].

Chapter 3 presents the first statistical model-based estimation methodology for S-FCS. A filtered Poisson process is proposed as a model of the observed fluorescence intensity process. Intensity data obtained from S-FCS experiments are considered to be a realization of this stochastic process and constitute an ordinary discrete-time series. Covariance functionals contain important model parameters related to cluster sizes and densities as well as laser beam width. A nonlinear least squares method is used to estimate these parameters. The least squares function involves sample and theoretical autocovariances of the data. The ratio parameter $R$ is a function of the above model parameters and its estimation (by $\hat{R}$) is discussed. Central limit theorems for the estimators, particularly for $\hat{R}$, are obtained for the first time, allowing the construction of asymptotic confidence intervals. The least squares function involves a lag parameter $L$. A method of obtaining an optimal value for $L$ with respect to a minimal limiting variance for $\hat{R}$ is numerically implemented. No such method was previously available. A Monte Carlo simulation study of the fluorescence intensity process is undertaken and the normal approximations of the estimators are studied. To obtain confidence intervals for $R$, the asymptotic variance of $\hat{R}$ must be estimated. A method based on the mixing properties of the data sequence is implemented and is shown to provide good results. Several data sets from actual S-FCS experiments are then analyzed. Certain of
the model assumptions are also checked and the results indicate that the Poisson model provides a good approximation to the fluorescence intensity process.

Chapter 4 discusses statistical inference for ICS. As with S-FCS, the fluorescence intensity data is modeled by a filtered Poisson process. The observed image however, is a realization of a stationary random field or spatial series, indexed by the lattice of integers. The same model parameters are estimated specifically $R$. The least squares equation is now based upon the spectral density and periodogram of the process. Central limit theorems are obtained for the estimators, for the first time and several remarks are made concerning the proofs of these results. Alternatively, the parameters may be estimated by an asymptotic or quasi-likelihood function with corresponding limiting results for the estimators. The limiting variance of $\hat{R}$ involves multi-dimensional integrals of spectral densities of second- to fourth-order. These are approximated by a Monte Carlo integration method. The spatial fluorescence intensity process is simulated and the least squares method is used to estimate the model's parameters based on the simulated image data. Two variance estimation techniques are examined. The first method is the extension to lattice data of the method used for S-FCS data. The second technique is based upon estimators proposed by Taniguchi [34]. Here they are generalized to the random field case and numerically implemented. Numerical results from these variance estimators are not commonly available. The Taniguchi method essentially involves replacing spectral densities by their corresponding periodograms. These ICS calculations are computationally intensive with some being of the order $O(T^8)$ when the sample size is $T^2$. They are performed on a MasPar MP-2 massively par-
allel (2K) computer, demonstrating an innovative application of this leading-edge technology, to the statistical sciences.
Chapter 2
Mathematical Preliminaries

This chapter gives the main mathematical results and tools that are required in the thesis. These include a formal discussion of filtered Poisson processes and the various classical limit theorems which are used to obtain asymptotic results for the estimators in Chapters 3 and 4.

**Definition 2.1** Suppose $\mathcal{X}$ is a complete separable metric space and $N(A)$ (the number of points in the set $A$) is defined and finite for every bounded set $A$ in the Borel $\sigma$-field $\mathcal{B}_\mathcal{X}$ generated by the open spheres of $\mathcal{X}$ i.e. the trajectories $N(\cdot)$ are boundedly finite almost surely. A multidimensional Poisson process $\{N(A) : A \subseteq \mathcal{X}\}$ with intensity $\lambda(x)$ has the following properties:

a) for each $A \subseteq \mathcal{X}$, $N(A)$ has the Poisson distribution with parameter

$$\lambda(A) = \int_A \lambda(x) dx ;$$

b) if $A_1, A_2, \ldots, A_k, \ldots$ is a sequence of disjoint subsets of $\mathcal{X}$, then

$N(A_1), N(A_2), \ldots, N(A_k) \ldots$ are independent random variables and
\[ N(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} N(A_k) . \]

The space \( \mathcal{X} \) of interest for S-FCS and ICS is the plane \( \mathbb{R}^2 \). The boundedly finite measure \( \Lambda(\cdot) \) is called the parameter measure of the process. Property a) assumes the absolute continuity of \( \Lambda(\cdot) \) with respect to Lebesgue measure. The point process \( N \) is formally defined as a random counting measure, that is, \( N \) is a measurable mapping from a probability space \( (\Omega, \mathcal{F}, P) \) into \( (\hat{\mathcal{N}}_*, \mathcal{B}(\hat{\mathcal{N}}_*)) \) where \( \hat{\mathcal{N}}_* \) is the space of all boundedly finite, integer-valued measures \( N \) defined on \( \mathcal{B}_\mathcal{X} \). \( \mathcal{B}(\hat{\mathcal{N}}_*) \) is the smallest \( \sigma \)-algebra with respect to which the mappings \( N \to N(A) \) are measurable for each \( A \in \mathcal{B}_\mathcal{X} \). Property b) then implies the countable additivity of the counting measure \( N \). The next definition follows immediately from the fact that each realization of \( N \equiv N(\cdot, w) \) is a boundedly finite measure.

**Definition 2.2** Let \( f \) be a Borel measurable function on \( \mathcal{X} \). Then for each realization of \( N \), the integral

\[ I_f = \int_{\mathcal{X}} f(x) N(dx) \]

exists as a Lebesgue-Stieltjes integral.

Integrals of the type encountered in Definition 2.2 are frequently referred to as random counting integrals. Their evaluation is discussed below.

A marked Poisson process is a Poisson process where, attached to each point event (or random atom of \( N \)) \( x_i \), is a random variable or mark \( U_i \) taking values
in some complete separable metric space $\mathcal{K}$ (the nonnegative integers $\mathbb{Z}_0^+$ for example). A realization of such a process would consist of the sequence of pairs $(x_1, u_1), \ldots, (x_n, u_n)$. An associated process is a compound Poisson process.

**Definition 2.3** Suppose

a) $\{N(A) : A \subseteq \mathcal{X}\}$ is a nonhomogeneous Poisson process with intensity $\lambda(x)$, $x \in \mathcal{X}$;

b) $\{U_i\}$ is a sequence of independent, identically distributed random variables that are independent of $N$;

The compound Poisson process $\{X(A) : A \subseteq \mathcal{X}\}$ is defined by

$$X(A) = \sum_{i=1}^{N(A)} U_i.$$

The process $X$ is sometimes referred to as a mark accumulator process. In general, marked Poisson processes are formally defined as follows.

**Definition 2.4** A marked Poisson process, with occurrence points in $\mathcal{X}$ and marks in $\mathcal{K}$, is a Poisson process on $\mathcal{X} \times \mathcal{K}$ such that the marginal process $\{N(A \times \mathcal{K}) : A \in \mathcal{B}_\mathcal{X}\}$ is itself a Poisson process (denoted by $N_X$).

As stated earlier, this work involves an adaptation of the compound Poisson process called a filtered (or smoothed) Poisson process. These may be thought of as Poisson-driven real-valued stationary processes. Filtered Poisson process models are used to characterize a variety of physical and biological phenomena where the output or response process can be expressed as a sum of individual responses
to each marked point event. Applications include shot noise modeling of anode electron currents, telephone network studies as well as models of photoconductivity, atmospheric noise and radar signals [28]. As before, let the complete separable metric spaces \( X \) and \( K \) denote the spaces of point locations and marks respectively.

**Definition 2.5** Let \( x_n \) and \( u_n \) denote the \( n \)th occurrence point and mark of a marked Poisson process. Then a filtered Poisson process is defined by the following superposition:

\[
Y(A) = \begin{cases} 
0 & \text{if } N(A) = 0 \\
\sum_{n=1}^{N(A)} f_A(x_n, u_n) & \text{if } N(A) \geq 1
\end{cases}
\]

where \( A \in \mathcal{B}_X \). The function \( f_A(\cdot) \) is called the weight or impulse response function.

The response of the output process \( Y \) to the \( n \)th marked point \( x_n \in A \) is measured by \( f_A(\cdot) \). The filtered Poisson process (2.1) may be formally defined by a random counting integral,

\[
Y(A) = \int_A f_A(x, u) N(dx), \ A \in \mathcal{B}_X.
\]

This representation facilitates certain moment computations. The integral is evaluated as

\[
\sum_{n=1}^{N(A)} f_A(x_n, u_n).
\]

These integrals and their evaluation merit further discussion.

Let \( X \) and \( K \) have their usual meaning and take \( A \in \mathcal{B}_X \) and \( B \in \mathcal{B}_K \). Let

\[
\{ N(A \times B) : A \subseteq X, \ B \subseteq K \}
\]
be a marked Poisson process on $\mathcal{X} \times \mathcal{K}$. From the independent increments property of the Poisson process, disjoint subsets $(A \times B)$ of $\mathcal{X} \times \mathcal{K}$ are independent and $N(A \times B)$ has a Poisson distribution with parameter measure

$$\int_A \int_B \lambda(t) p(u) dt du,$$

where we assume that the intensity function of $N$ factors into the product of the intensity function of the marginal process of locations and the probability density or mass function of the marking random variable. The marginal process of point locations is obtained by

$$N(A \times \mathcal{K}) = \int_A \int_{\mathcal{K}} N(dx \times du) = \int_A N(dx \times \mathcal{K})$$  \hspace{1cm} (2.3)

and may be denoted by $N_X(A)$. This process counts the number of points in $A$ independently of their marks. In a technical sense, the integral in (2.2) involves this marginal process, i.e.

$$Y(A) = \int_A f_A(x, u) N_X(dx).$$

There should be no confusion, however, if the subscript on $N$ is suppressed.

In the special case where $\mathcal{X} = \mathcal{R}$ and $\mathcal{K} = \mathcal{R}^d$, the Ito [23] theory applies to the random counting integral

$$I_b = \int_{t_0}^t \int_B b(s, u) N(ds \times du)$$  \hspace{1cm} (2.4)

for a Borel measurable random function $b(\cdot)$ on $\mathcal{R} \times \mathcal{R}^d$ where $B \in \mathcal{B}(\mathcal{R}^d)$. It can be shown that (2.4) exists and has the evaluation

$$I_b = \begin{cases} 0 & \text{if } N(t) = 0 \\ \sum_{n=1}^{N(t)} b(s_n, u_n) & \text{if } N(t) \geq 1 \end{cases}$$
This situation does not directly apply, however, to the spectroscopy application
considered here as the proposed model involves $\mathcal{X} = \mathcal{R}^2$ (and $\mathcal{K} = \mathcal{Z}_0^+$). Instead, let $\mu$ be a measure on $\mathcal{X}$ and call $x \in \mathcal{X}$ an atom of $\mu$ if $\mu(\{x\}) > 0$. The measure defined on $\mathcal{B}_\mathcal{X}$ by
\[
\delta_x(A) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{otherwise}
\end{cases}
\]
for $A \in \mathcal{B}_\mathcal{X}$ is called the Dirac measure (at $x$).

Atoms are called random if $P(\mu(\{x\}) > 0) = 0$. A measure $\delta_x(\cdot)$ with only atoms is called purely atomic. Proposition 7.1.II of Daley and Vere-Jones [12] states that the Poisson counting process $N$ is expressible as $N = \sum \delta_{x_i}$, where $\{x_i\}$ is a collection of the atoms of the Dirac measures $\delta_{x_i}$. This implies that $N$ is a purely atomic random measure. For each realization $N(\cdot, w)$ generating the collection of random atoms $\{x_i\}$, $i = 1, \ldots, n$, $N$ is a boundedly finite, integer-valued measure and so from Definition 2.2 (with the choice $f_A = fI_A$)
\[
\int_A f_A(x, u)N(dx) = \sum_{i=1}^{N(A)} f_A(x_i, u_i).
\]

With these results, a formal model of the fluorescence intensity process of S-
FCS and ICS may be developed. Let $\mathcal{X} = \mathcal{R}^2$ be the space of point locations and $\mathcal{K} = \mathcal{Z}_0^+$ the associated mark space. Denote the disc centred at $(t_1 \delta, t_2 \delta)$, $(t_1, t_2) \in \mathcal{Z}^2$ by
\[
S_\delta : (x - t_1 \delta)^2 + (y - t_2 \delta)^2 \leq r^2
\]
where $\delta$ is a known stepsize or resolution parameter. Suppose

(i) $\{N(A) : A \subseteq \mathcal{R}^2\}$ is a homogeneous Poisson process with intensity $\lambda$. 
(ii) \( \{Y(x,y)\} \) is a sequence of independent, identically distributed random variables that are also independent of \( N \) where \((x,y)\) is an occurrence point of \( N \).

The point event \((x,y)\) represents the location of a protein cluster while the associated mark represents the number of monomers comprising the cluster. The laser beam intensity profile is modeled by the Gaussian kernel

\[
f_{\xi}(x,y) = f_0(x - t_1 \delta, y - t_2 \delta), \quad \text{where}\]
\[
f_0(x,y) = e^{-\frac{1}{2\sigma^2}(x^2+y^2)}.
\]

The radius \( r \) of \( S_{\xi} \) is chosen to be \( r = \frac{3\sigma}{\sqrt{2}} \). The (deterministic) impulse response or weight function previously discussed takes the form

\[
f_{S_{\xi}}((x,y), Y(x,y)) = f_{\xi}(x,y)Y(x,y).
\]

Photon counting white (background) noise is represented by the independent, identically distributed collection of Gaussian random variables \( \{\gamma_\epsilon(\xi)\} \) with \( E(\epsilon(\xi)) = 0 \), \( E(\epsilon^2(\xi)) = 1 \) and \( \gamma \) is the noise parameter.

The fluorescence intensity process \( \{X(\xi) : \xi \in Z^2\} \) is thus represented by the filtered Poisson process with additive Gaussian noise:

\[
X(\xi) = \int_{S_{\xi}} f_{\xi}(x,y) Y(x,y) N(dx dy) + \gamma_\epsilon(\xi). \tag{2.5}
\]

Specifically (2.5) refers to ICS fluorescence intensity measurements. Those obtained in S-FCS correspond to the special case \( \xi = (t_1,0) \). The process \( \{X(\xi) : \xi \in Z^2\} \) is a stationary finite-range dependent random field. This follows from the independent increments property of \( N \) and the decay of the Gaussian kernel \( (w) \) as the disc \( S_0 \) and \( S_\xi \) are disjoint for sufficiently large \( \xi \).
It is of interest to estimate certain parameters of the model (2.5). The resulting estimators are shown to be asymptotically normal in Chapters 3 and 4. The proofs require the following results.

**Proposition 2.1 (The δ-method)**

Suppose the i.i.d. central limit theorem applies to the sequence \( \{X_n\} \) so that

\[
\sqrt{n}\left(\frac{X_n - c}{\sigma}\right) \overset{D}{\rightarrow} N
\]

where \( E(X_n) = c \) and \( \text{Var}(X_n) = \sigma^2 < \infty \). If \( f'(c) \neq 0 \) then

\[
\sqrt{n}\left(\frac{f(X_n) - f(c)}{\sigma f'(c)}\right) \overset{D}{\rightarrow} N.
\]

This result may be generalized as follows.

**Proposition 2.2** Suppose that

\[
v_n(X_n - \alpha, Y_n - \beta) \overset{D}{\rightarrow} (X, Y)
\]

where \( v_n \to \infty \) and suppose \( f(x, y) \) is differentiable at \( (\alpha, \beta) \). Then

\[
v_n(f(X_n, Y_n) - f(\alpha, \beta)) \overset{D}{\rightarrow} \frac{\partial f}{\partial x}(\alpha, \beta)X + \frac{\partial f}{\partial y}(\alpha, \beta)Y.
\]

**Theorem 2.1 (Cramér-Wold)**

For random vectors \( X_n = (X_{n1}, \ldots, X_{nk}) \) and \( Y = (Y_1, \ldots, Y_k) \), a necessary and sufficient condition for \( X_n \overset{D}{\rightarrow} Y \) is that \( \sum_{u=1}^{k} t_u X_{nu} \overset{D}{\rightarrow} \sum_{u=1}^{k} t_u Y_u \) for each \( (t_1, \ldots, t_k) \in \mathbb{R}^k \).

**Theorem 2.2 (Slutsky)**

Suppose \( \{a_n\} \) and \( \{b_n\} \) are sequences of random variables such that \( a_n \overset{P}{\rightarrow} a \) and
\( b_n \overset{P}{\rightarrow} b \) where \( a \) and \( b \) are constants. If

\[ X_n \overset{D}{\rightarrow} X \]

then

\[ a_n X_n + b_n \overset{D}{\rightarrow} aX + b. \]

**Definition 2.6 (\( m \)-dependence)**

A strictly stationary sequence of random variables \( \{X_t\} \) is said to be \( m \)-dependent (where \( m \) is a nonnegative integer) if for each \( t \) the two sets of random variables \( \{X_j : j \leq t\} \) and \( \{X_j : j \geq t + m + 1\} \) are independent.

**Theorem 2.3 (The Central Limit Theorem for Stationary \( m \)-Dependent Sequences)**

If \( \{X_t\} \) is a strictly stationary \( m \)-dependent sequence of random variables with \( E(X_t) = \mu \) and autocovariance function \( \gamma(\cdot) \) and if \( \nu_m = \gamma(0) + 2 \sum_{j=1}^{m} \gamma(j) \neq 0 \), then

(i) \( \lim_{n \to \infty} n \text{ Var } (X_n) = \nu_m \) and

(ii) \( \sqrt{n}(X_n - \mu) \overset{D}{\rightarrow} N(0, \nu_m) \).

The random field model of ICS in Chapter 4 requires the following results of Rosenblatt [27]. They are somewhat stronger than what is actually required. Let \( X = \{X(t) : t \in \mathbb{Z}^2\} \) be a strictly stationary random field. Let \( S \) and \( S' \) be two sets of indices where

\[ \mathcal{B}(S) = \mathcal{B}(X(\tau), \tau \in S) \]

and

\[ \mathcal{B}(S') = \mathcal{B}(X(\tau), \tau \in S') \]
are the generated $\sigma$-fields. Suppose $d(S, S')$ is the minimum Euclidean distance between the points of the sets $S$ and $S'$.

**Definition 2.7 (Rosenblatt [27])**

The random field $X$ is said to be strongly mixing if

$$\sup_{A \in \mathcal{B}(S), B \in \mathcal{B}(S')} |P(A \cap B) - P(A)P(B)| \leq \varphi(d(S, S'))$$

with $\varphi$ a function such that $\varphi(d) \to 0$ as $d \to \infty$.

The definition discusses a form of asymptotic independence. An $m$-dependent stationary random field is strongly mixing. The following theorem generalizes a central limit theorem (Rosenblatt, [27]) for strongly mixing stationary sequences to stationary random fields.

**Theorem 2.4 (Rosenblatt, 1985)**

Let $X(t), t \in \mathbb{Z}^d$, be a strictly stationary random field with $E(X(t)) = 0$ and $E|X(t)|^{2+\delta} < \infty$ for some $\delta > 0$. Also assume that $E|\sum_{t_i = a_i}^b X(t)|^2 = h(b - a) \to \infty$ as $b_i - a_i \to \infty$, $i = 1, \ldots, d$ and that $h(\alpha) = O(h(\beta))$ if $\alpha_i \to \infty$, $\alpha_i = O(\beta_i)$, $i = 1, \ldots, d$ but for some $j$, $\alpha_j = o(\beta_j)$. Further, if

$$E|\sum_{t_i = a_i}^b X(t)| = O(h(b - a))^{1+\delta/2}$$

as $b_i - a_i \to \infty$, $i = 1, \ldots, d$, then

$$\sum_{t_i = 1}^{\alpha_i T} X(t)$$

when appropriately normalized is asymptotically $N(0, 1)$ as $T \to \infty$ for each fixed $\alpha_i > 0$, $i = 1, \ldots, d$. 
In addition, certain cumulant summability conditions are sufficient to ensure that one normalizes certain functionals of \( X(t) \) by \( T^{d/2} \).

**Theorem 2.5 (Rosenblatt, 1985)**

Let \( \{X(t)\} \) be a strongly mixing \( d \)-dimensional strictly stationary random field with \( E(X(t)) = 0 \) and \( E(X(t)^8) < \infty \). Assume that the cumulants of \( \{X(t)\} \) up to eighth-order are absolutely summable. It then follows that \( T^{d/2}(\hat{c}(u) - c(u)) \) for a fixed number of lags \( u \) are jointly asymptotically normal as \( T \to \infty \) with mean 0 and covariances

\[
C_{u,v} = \sum_{\alpha} (c(\alpha)c(\alpha + v - u) + c(\alpha + v)c(\alpha - u) + c_4(u, \alpha, \alpha + v)).
\]

**Theorem 2.6 (Rosenblatt, 1985)**

Let \( \{X(t)\} \) be an ergodic strictly stationary random field that satisfies the assumptions of Theorem 2.5. Consider

\[
\int_0^{2\pi} f(T)(\lambda)A_j(\lambda)d\lambda, \ j = 1, \ldots, s
\]

which are quadratic forms in \( X(t) \) with real-valued weight functions \( A_j(\lambda) \) square integrable. The smoothed periodograms then are asymptotically normal with means

\[
\int_0^{2\pi} f(\lambda)A_j(\lambda)d\lambda, \ j = 1, \ldots, s
\]

and limiting covariances

\[
(2\pi)^d \{2 \int_0^{2\pi} A_j(\lambda)A_k(\lambda)f^2(\lambda)d\lambda + \int_0^{2\pi} \int_0^{2\pi} f_4(\lambda, \mu, \mu)A_j(\lambda)A_k(\mu)d\lambda d\mu\}.
\]

The following two results are taken from Brillinger [5] and are required to establish certain of the technical results of Chapter 4.
Theorem 2.7  Consider a two-way array of random variables $X_{ij}; \ j = 1, \ldots, N_i; \ i = 1, \ldots, M$. Consider the $M$ random variables

$$Y_i = \prod_{j=1}^{N_i} X_{ij}, \ i = 1, \ldots, M.$$  

The joint cumulant $\text{cum}(Y_1, \ldots, Y_M)$ is then given by

$$\sum_{\nu} \text{cum} (X_{ij}; \ ij \in \nu_1) \ldots \text{cum} (X_{ij}; \ ij \in \nu_p)$$

where the summation is over all indecomposable partitions $\nu = \nu_1 \cup \ldots \cup \nu_p$ of the two-way table:

$$
\begin{array}{ccc}
(1, 1) & \cdots & (1, N_1) \\
\vdots & & \vdots \\
(M, 1) & \cdots & (M, N_M).
\end{array}
$$

If the rows of the above table are denoted by $R_1, \ldots, R_M$, then a partition $\nu_1, \ldots, \nu_p$ is indecomposable if and only if there exists no sets $\nu_{m_1}, \ldots, \nu_{m_q}, \ (q < p)$ and rows $R_{i_1}, \ldots, R_{i_\ell}, \ (\ell < M)$ with

$$\nu_{m_1} \cup \ldots \cup \nu_{m_q} = R_{i_1} \cup \ldots \cup R_{i_\ell}.$$  

Theorem 2.8 Let $X(t), \ t \in \mathbb{Z}^d$ be a strictly stationary random field all of whose moments exist and satisfy:

$$\sum_{\mathbf{u}^{(1)}} \cdots \sum_{\mathbf{u}^{(k)}} (||\mathbf{u}^{(1)}|| + \cdots + ||\mathbf{u}^{(k-1)}||) c_k(u^{(1)}, \ldots, u^{(k-1)}) < \infty$$

where $||\mathbf{u}|| =< \mathbf{u}, \mathbf{u}>^{\frac{1}{2}}, \ k = 1, 2, \ldots$. Then as $T \to \infty$,

$$\text{cum}_k\{d^{(T)}(\lambda^{(1)}), \ldots, d^{(T)}(\lambda^{(k)})\}$$

$$= \Delta^{(T)}(\sum_{j=1}^{k-1} \lambda^{(j)})(2\pi)^{d(k-1)} f_k(\lambda^{(1)}, \ldots, \lambda^{(k-1)}) + O(T^{d-1})$$

where $\Delta^{(T)}(\lambda) = \sum_{t_0}^{T-1} e^{-i\lambda t_0}$. The function $\Delta^{(T)}(\lambda)$ has the properties: $\Delta^{(T)}(\lambda) = T^2$ for $\lambda = 0$ (mod $2\pi$), and $\Delta^{(T)}(\frac{2\pi}{T} s) = 0$ for $s \in \mathbb{Z}^d$ with $s \neq 0$ (mod $T$).
Chapter 3

Statistical Inference for Scanning Fluorescence Correlation Spectroscopy

3.1 Introduction

Protein–protein aggregation or clustering may be necessary for intracellular biological responses. Quantitative measurements of protein receptor distributions on cell surfaces are an important aspect in understanding receptor–receptor clustering and ligand–receptor interactions.

Membrane protein receptor aggregation has been directly observed by electron microscopy and inferred from fluorescence energy transfer experiments. However, these techniques have been unable to accurately measure cluster sizes. Quantitative measurements of the extent of protein clustering have, in general, been difficult to obtain. In Petersen [24], a new experimental technique was introduced for making quantitative measurements of the distribution of fluorescently labeled proteins on the surfaces of living cells. The method is called Scanning Fluorescence Correlation Spectroscopy (S–FCS) and to date has provided the highest quality measurements
of protein cluster sizes and densities. S-FCS is based on Concentration Correlation Spectroscopy and was initially used to measure the molecular weights of DNA fragments. Using fluorescence as an index of concentration, it was adapted by Magde et al. [20] to Fluorescence Correlation Spectroscopy (FCS) and was used to measure diffusion or directed flow rates of proteins through cell membranes. Although the potential of the technique for measuring protein aggregation was recognized, the protein diffusion rate in cell membranes was too slow to measure changes in fluorescence intensity on any practical time scale.

In S-FCS, these fluorescence intensity measurements are not obtained in real time but are instead determined by the position of a focused laser beam on the cell. The cell is translated through the beam in a fluorescence microscope resulting in intensity data as a function of position across the cell surface. Areas of the cell surface illuminated by the laser beam are represented by discs and fluorescence intensities are counts of the number of receptors within these discs, after allowing for local variation in the laser beam intensity (power). The primary objective of S-FCS is to determine (estimate) the protein cluster density (average number of clusters per unit area) as well as the average number of protein monomers per cluster in a given area. It is known that the ratio (denoted by $R$) of the fluorescence intensity variance over the square of the mean intensity is directly related to these quantities [24, 32]. However, the variance or zero-lag correlation cannot be estimated directly from the experimental data due to photon counting noise which causes a sharp peak at this value. In practice, a least squares fitting procedure involving empirical correlation functions is used to obtain correlation
zero-lag values.

The biophysical characteristics associated with FCS suggest that a point process model should provide a reasonable mathematical approximation to the observed fluorescence intensity process. Specific external agents or ligands are fluorescently labeled on the cell surface. Membrane protein receptors, upon binding to these ligands, become observable when the cell is translated under a fluorescence or confocal scanning laser microscope. The beam is assumed to have a two-dimensional Gaussian intensity profile. This is a common assumption in laser optics as a surface viewed through a small aperture produces an image filtered by a local (near 0) sine curve, which is well-approximated by the Gaussian kernel [32]. When the centre of the laser is focused at a position \( \theta \), \( t \) an integer, a disc region of the cell surface is illuminated. The radius of the disc is a function of the laser beam width parameter. Within each illuminated disc, a random number of fluorescently labeled protein receptors of varying size, are visible in the microscope. The concentration measurement obtained at position \( \theta \) is the overall fluorescence intensity at that site where a given protein aggregate's contribution is weighted by the Gaussian kernel. These comments suggest that a sensible mechanistic model for the aggregate process would be a spatial marked Poisson process. Each occurrence point would represent the location of a receptor while a nonnegative random integer mark represents the number of monomers comprising the aggregate. To model the effect of the variable laser intensity, each point event is filtered by the Gaussian kernel. Thus, the fluorescence intensity process is represented by a Poisson-driven real-valued stationary process known as a filtered (or smoothed) Poisson process.
Snyder and Miller [28] discuss a variety of filtered Poisson modeling applications including atmospheric noise, photoconductivity and radar clutter.

This chapter discusses the estimation of specific parameters associated with a filtered Poisson process model of S-FCS. The model is formally discussed in Section 3.2. The fluorescence intensity autocovariances are functions of the identifiable parameters and their estimation via a nonlinear least squares technique is discussed in Section 3.3. The laser beam width parameter can be measured experimentally [32]. Its estimation serves as an internal consistency check for the model with unusually small or large values indicating excessively noisy or contaminated data. The parameter $R$ directly measures cluster sizes and densities. It is a smooth function of the above parameters. The asymptotic properties of the estimators are also obtained. The least squares equation involves a lag parameter $L$. Currently, S-FCS experimentalists do not have a formal method for choosing an optimal value for $L$. Section 3.8 proposes a method for choosing an optimal value for this and other design (or tuning) parameters based on minimizing the limiting variances of the estimators. The results of an extensive Monte Carlo simulation study are presented in Section 3.9. Of particular importance is the discussion of the method used to estimate the estimators' limiting variances. Asymptotic confidence intervals for the parameters (based on a single data set) can now be obtained where previously only point estimates were available and then only after averaging over a large number of data sets. Finally, Section 3.10 discusses the analysis of several S FCS data sets. Where possible, the consistency of the data with the proposed model is examined.
3.2 A Mathematical Model for S-FCS

The spatial distribution of protein aggregates is modeled by a marked homogeneous Poisson process. The Poisson process gives the locations of the centres of the aggregates; that is, each point occurrence of the Poisson process \((x, y)\) is the location of an aggregate. The Poisson process has intensity \(\lambda\). The size of the aggregate at \((x, y)\) is represented by a random mark \(Y(x, y)\), where the \(Y\)'s are independent of the Poisson process and are also independent and identically distributed (i.i.d.). The aggregate process is then

\[
\{Y(x, y) : (x, y) \in N\},
\]

where \(N\) is the Poisson process. The aggregate mass for a region \(A\) is defined by \(\int_{A} Y(x, y) N(dx, dy)\).

The intensity or kernel of the laser beam follows a Gaussian distribution

\[
f_0(x, y) = e^{-\frac{1}{2}(x^2 + y^2)}, w > 0.
\]

The effective radius of the beam is \(\frac{3w}{\sqrt{2}}\). When the laser is centred at \((k\delta, 0)\), the region effectively illuminated by the laser beam is

\[
S_k : (x - k\delta)^2 + y^2 \leq \left(\frac{3w}{\sqrt{2}}\right)^2.
\]

At this position, the beam has intensity kernel \(f_k(x, y) = f_0(x - k\delta, y)\).

The fluorescence intensity \(X_k\) is formally defined as:

\[
X_k = \int f_k(x, y)Y(x, y) N(dx, dy) + \gamma \epsilon_k
\]

\[
= \int e^{-\frac{1}{2}(x^2 + y^2)} Y(x, y) N(dx, dy) + \gamma \epsilon_k
\]
\[ \sum e^{-\frac{1}{2}((x-k\delta)^2+y^2)}Y(x,y) + \gamma \epsilon_k \]
\[ = Z_k + \gamma \epsilon_k, \]

for \( k = 0, 1, \ldots \) where the integral is over the region \( S_k \). The last integral equals the sum since the integral is a Lebesgue integral with respect to the random counting measure \( N \). The random variable \( X_k \) is a weighted integral of the aggregate process (3.1) with weight function \( f_k \) (3.2). The term \( \gamma \epsilon_k \) represents measurement device recording error due to photon counting noise as well as model error. \{\epsilon_k\} is a sequence of i.i.d., mean 0 and variance 1 random variables. \( \gamma \) is an error size parameter. If \( \gamma = 0 \) there is no measurement recording device error.

\{X_k\} is a sequence of stationary finite-range dependent real random variables. This is because the regions \( S_0 \) and \( S_k \) given by (3.3), \( k = 0, 1, \ldots \) are disjoint for large enough \( k \), depending of course upon \( w \), and the independent increments property of the Poisson process \( N \). The sequence \( X_0, \ldots, X_n \) is the observable data from the S-FCS experiment. As with many stochastic models, the process is only partially observable.

By using a truncation argument, one can consider the \( X_k \)'s to be defined by (3.4), but with the integration carried out over \( \mathbb{R}^2 \) instead of \( S_k \). This can be used to simplify some of the moment calculations. Central limit theorems for various moment functionals of \( X_k \), such as sample means and covariances will continue to hold.
3.3 The Least Squares Estimators

Certain functions of the model parameters can be estimated via a nonlinear least squares technique, involving the fluorescence intensity autocovariances. Specifically one can match the parameters in the theoretical covariance with the empirical covariances.

The least squares equation is given by

\[ H_n(\theta_1, \theta_2) = \sum_{k=1}^{L} (\hat{c}(k) - c(k; \theta_1, \theta_2))^2 \]  \hspace{1cm} (3.5)

where

\[ \hat{c}(k) = \frac{1}{n} \sum_{i=0}^{n-k-1} (X_i - X)(X_{i+k} - X) \]

is the empirical covariance, and

\[ c(k; \theta_1, \theta_2) = \text{cov}(Z_0, Z_k) \]
\[ = \lambda E(Y^2) \int_{S_0 \cap S_k} e^{-\frac{1}{2}(x^2 + (y-k)^2 +2y^2)} \, dx \, dy \]
\[ = \theta_1 \int_{S_0 \cap S_k} e^{-\frac{1}{2}(x^2 + (y-k)^2 +2y^2)} \, dx \, dy . \]

The parameter \( L \) refers to the number of lags used in the least squares equation. Then,

\[ \text{cov}(X_0, X_k) = \begin{cases} c(0; \theta) + \gamma^2 & \text{if } k = 0 \\ c(k; \theta) & \text{otherwise.} \end{cases} \]

The parameter that can be estimated is \( \theta = (\theta_1, \theta_2) = (\lambda E(Y^2), \mu) \). The main goal in the chemistry problem is the determination of \( \mu \equiv c(0; \theta_1, \theta_2)/\mu^2 \) where \( \mu = E(X_k) \) [30]. This can be estimated by \( \hat{\mu} = c(0; \hat{\theta}_1, \hat{\theta}_2)/X^2 \) where \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) are the least squares estimators.
The least squares estimators and \( \hat{R} \) are smooth functions of the empirical covariances \( \hat{c}(k) \), and \( X \). Theorem 3.1 below gives the joint asymptotic normality of \( \hat{c}(k) \) and \( X \). The result is obtained by a direct application of an \( m \)-dependent (finite range dependent) central limit theorem [10]. The stationary process \( \{X_k\} \) is said to be \( m \) dependent if random variables separated by more than \( m \) lags are independent. Let \( \theta_0 \) represent the true value of the parameter \( \theta \). The second theorem is the consistency of the least squares estimator. This follows from a lemma in Brillinger [4]. The remaining limit theorems in this section then follow from these two by Taylor's formula with remainder and Slutsky's Theorem.

**Theorem 3.1**

\[
\sqrt{n} \begin{bmatrix}
    \hat{X} - \mu \\
    \hat{c}(1) - c(1; \theta_0) \\
    \vdots \\
    \hat{c}(L) - c(L; \theta_0)
\end{bmatrix} \xrightarrow{D} N_{L+1}(0, \Sigma).
\]

where

\[
\Sigma = \begin{bmatrix}
    \Sigma_{11} & \Sigma_{12} \\
    \Sigma_{21} & \Sigma_{22}
\end{bmatrix},
\]

and the component matrices are given below.

\[
\Sigma_{11} = \gamma^2 + \sum_{j=-\infty}^{\infty} c(j; \theta_0),
\]

\[
\Sigma_{12} = \Sigma_{21} = \left[ \sum_{j=-\infty}^{\infty} \text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \right]_{1 \times L},
\]

and

\[
\Sigma_{22} = \left[ \sum_{j=-\infty}^{\infty} \text{cov}((X_j - \mu)(X_{j+k} - \mu), (X_0 - \mu)(X_L - \mu)) \right]_{L \times L}.
\]
Because of the finite-range dependence or mixing of \( \{X_k\} \), the infinite series are finite sums, and can be truncated in practice. For example, the summation for \( \Sigma_{11} \) can be truncated to a sum over \( j \) from \(-m\) to \( m\) with \( m = \left\lfloor \frac{6n}{\sqrt{2\pi}} \right\rfloor - 1 \) to a high degree of accuracy.

**Theorem 3.2** \( \hat{\theta} \xrightarrow{P} \theta_0 \).

**Theorem 3.3**

\[
\sqrt{n} \left[ \begin{array}{c}
\frac{X - \mu}{\hat{\theta}_1 - \theta_{1,0}} \\
\frac{X - \mu}{\hat{\theta}_2 - \theta_{2,0}}
\end{array} \right] \xrightarrow{D} N_3(0, C, \Sigma^*(C')).
\]

where

\[
\Sigma^* = \begin{bmatrix}
e_i e_1 & e_i e_2 & e_i e_3 \\
e_i e_2 & e_i e_1 & e_i e_3 \\
e_i e_3 & e_i e_2 & e_i e_1
\end{bmatrix},
\]

\[
C = \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}, \quad A = \begin{bmatrix} \partial c_i^* / \partial c_1 & \partial c_i^* / \partial c_2 \\ \partial c_i^* / \partial c_2 & \partial c_i^* / \partial c_1 \end{bmatrix}^{-1},
\]

\[
\partial c_i = \left[ \frac{\partial c_i}{\partial \theta_i} \right]_{k=1}^{L}, \quad i = 1, 2,
\]

\( e_i' = (1, 0, \ldots, 0)_{1 \times (L+1)} \) and \( \partial c_i^* = [c(1; \theta_0), \ldots, c(L; \theta_0)] \), \( i = 1, 2 \).

**Theorem 3.4**

\[
\sqrt{n} \left( \hat{R} - R \right) = \sqrt{n} \left( \frac{c(0; \hat{\theta})}{X^2} - \frac{c(0; \theta_0)}{\mu^2} \right) \xrightarrow{D} N(0, c' \Sigma^* c'),
\]

where

\[
c' = \begin{bmatrix} \partial / \partial \mu \left( \frac{c(0; \theta_0)}{\mu^2} \right) \\ \partial / \partial \theta_1 \left( \frac{c(0; \theta_0)}{\mu^2} \right) \\ \partial / \partial \theta_2 \left( \frac{c(0; \theta_0)}{\mu^2} \right) \end{bmatrix}.
\]

The least squares method has several tuning parameters, such as \( L \), the number of lags used, and the laser beam step resolution or spacing \( \delta \). Natural questions arise as to the optimal choice of these.
3.4 Proofs

3.4.1 Proof of Theorem 3.1

The proof will employ the Cramér-Wold device; consider an arbitrary linear combination of the elements of

\[
\begin{bmatrix}
    X - \mu \\
    \hat{c}(1) - c(1; \theta_0) \\
    \vdots \\
    \hat{c}(L) - c(L; \theta_0)
\end{bmatrix}
\]

\[
\begin{bmatrix}
    \sqrt{n} (X - \mu) \\
    \sqrt{n} (\hat{c}(1) - c(1; \theta_0)) \\
    \vdots \\
    \sqrt{n} (\hat{c}(L) - c(L; \theta_0))
\end{bmatrix}
\]

\[
[a_0, \ldots, a_L]
\]

\[
= a_0 \sqrt{n} (X - \mu) + \sqrt{n} \sum_{k=1}^{L} (\hat{c}(k) - c(k; \Theta_0)) a_k
\]

\[
= a_0 \sqrt{n} (X - \mu) + \sqrt{n} \sum_{k=1}^{L} (\hat{c}(k; \mu) - c(k; \Theta_0)) a_k
\]

\[
+ \sqrt{n} \sum_{k=1}^{L} (\hat{c}(k) - \hat{c}(k; \mu)) a_k
\]

where

\[
\hat{c}(k; \mu) = \frac{1}{n} \sum_{i=0}^{n-1} (X_i - \mu)(X_{i+k} - \mu)
\]

Let

\[
R_n = \sqrt{n} \sum_{k=1}^{L} (\hat{c}(k) - \hat{c}(k; \mu)) a_k
\]

The above expression (3.6) becomes

\[
\sqrt{n} \left( a_0 (X_1 - \mu) + \sum_{k=1}^{L} ((X_1 - \mu)(X_{1+k} - \mu) - c(k; \Theta_0)) a_k) + R_n \right) \xrightarrow{D} N(0, \sigma^2)
\]
by the $m$-dependent central limit theorem since $R_n = o_P(1)$. To see that the remainder term $R_n$ is negligible, write

\[
R_n = \frac{\sqrt{n}}{n} \sum_{k=1}^{L} \left( \sum_{i=0}^{n-k-1} (X_i - X)(X_{i+k} - X) - \sum_{i=0}^{n-1} (X_i - \mu)(X_{i+k} - \mu) \right) a_k \\
= \frac{\sqrt{n}}{n} \sum_{k=1}^{L} \left( -(X - \mu) \sum_{i=0}^{n-1} (X_i - \mu) - (X - \mu) \sum_{i=0}^{n-1} (X_{i+k} - \mu) + n(X - \mu)^2 \right) a_k \\
- \sum_{i=n-k}^{n-1} (X_i - X)(X_{i+k} - X) \right) a_k \\
= -\sqrt{n}(X - \mu) \frac{1}{n} \sum_{i=0}^{n-1} (X_i - \mu) \left( \sum_{k=1}^{L} a_k \right) \\
- \sqrt{n}(X - \mu) \frac{1}{n} \sum_{i=0}^{n-1} \left( \sum_{k=1}^{L} (X_{i+k} - \mu) a_k \right) \\
+ \sqrt{n}(X - \mu)(X - \mu) \left( \sum_{k=1}^{L} a_k \right) \\
- \frac{\sqrt{n}}{n} \sum_{k=1}^{L} a_k \left( \sum_{i=n-k}^{n-1} (X_i - X)(X_{i+k} - X) \right) .
\]

The first three terms above converge to zero in probability by the $m$-dependent central limit theorem and Slutsky’s Theorem. The last term can be expressed as

\[
-\sqrt{n} \left\{ \sum_{k=1}^{L} a_k \left( \frac{1}{n} \sum_{i=n-k}^{n-1} (X_i - X)(X_{i+k} - X) \right) \right\} \\
= -\frac{\sqrt{n}}{n} \sum_{k=1}^{L} a_k \sum_{j=0}^{k-1} (X_{n+j-k} - X)(X_{n+j} - X) \\
= -\frac{\sqrt{n}}{n} \sum_{k=1}^{L} a_k \sum_{j=0}^{k-1} ((X_{n+j-k} - \mu)(X_{n+j} - \mu) \\
- (X - \mu)(X_{n+j-k} - \mu) - (X - \mu)(X_{n+j} - \mu) + (X - \mu)^2) \\
= -\frac{\sqrt{n}}{n} \sum_{k=1}^{L} a_k \sum_{j=0}^{k-1} ((X_{n+j-k} - \mu)(X_{n+j} - \mu)) \\
+ \sqrt{n}(X - \mu) \frac{1}{n} \sum_{k=1}^{L} a_k (X_{n+j-k} - \mu)
\]
Consider the first term in this sum. It is shown that this term as well as the others goes to zero in probability by virtue of Chebyshev's inequality.

\[
\begin{align*}
P \left\{ \left| \frac{\sqrt{n}}{n} \sum_{k=1}^{L} \sum_{j=0}^{k-1} a_k (X_{n+j-k} - \mu)(X_{n+j} - \mu) \right| \geq \epsilon \right\} \\
\leq \frac{1}{\epsilon^2} E \left\{ \left[ \frac{\sqrt{n}}{n} \sum_{k=1}^{L} \sum_{j=0}^{k-1} a_k (X_{n+j-k} - \mu)(X_{n+j} - \mu) \right]^2 \right\} \\
= \frac{n}{\epsilon^2 n^2} E \left\{ \sum_{k=1}^{L} \sum_{l=1}^{L} \sum_{j=0}^{k-1} \sum_{i=0}^{l-1} a_k a_l (X_{n+j-k} - \mu)(X_{n+j} - \mu)(X_{n+i-l} - \mu)(X_{i} - \mu) \right\} \\
= \frac{1}{n\epsilon^2} \sum_{k=1}^{L} \sum_{l=1}^{L} \sum_{j=0}^{k-1} \sum_{i=0}^{l-1} a_k a_l E((X_{j-k} - \mu)(X_j - \mu)(X_{i-l} - \mu)(X_i - \mu))
\end{align*}
\]

which goes to zero as \( n \) goes to infinity. Similarly.

\[
\begin{align*}
P \left\{ \left| \frac{1}{n} \sum_{k=1}^{L} \sum_{j=0}^{k-1} a_k (X_{n+j-k} - \mu) \right| \geq \epsilon \right\} \\
\leq \frac{1}{\epsilon^2} E \left\{ \left[ \sum_{k=1}^{L} \sum_{j=0}^{k-1} a_k (X_{n+j-k} - \mu) \right]^2 \right\} \\
= \frac{1}{\epsilon^2} E \left\{ \sum_{k=1}^{L} \sum_{l=1}^{L} \sum_{j=0}^{k-1} \sum_{i=0}^{l-1} a_k a_l (X_{n+j-k} - \mu)(X_{n+i-l} - \mu) \right\} \\
= \frac{1}{\epsilon^2} \sum_{k=1}^{L} \sum_{l=1}^{L} \sum_{j=0}^{k-1} \sum_{i=0}^{l-1} a_k a_l E((X_{j-k} - \mu)(X_{i-l} - \mu))
\end{align*}
\]

which also goes to zero. The same argument holds for the other terms.
\[
\sigma^2 = \lim_{n \to \infty} n \text{Var} \left[ a_0(X - \mu) + \sum_{k=1}^L a_k(\hat{c}(k) - c(k; \theta_0)) \right] \\
= \lim_{n \to \infty} n \left[ a_0^2 \text{Var}(X) + 2 \sum_{k=1}^L a_0 a_k \text{cov}(\hat{c}(k), X) \\
+ \sum_{k=1}^L \sum_{l=1}^L a_k a_l \text{cov}(\hat{c}(k), \hat{c}(l)) \right] \\
= a_0^2(\gamma^2 + \sum_{j=-m}^m c(j; \theta_0)) \\
+ 2 \sum_{k=1}^L a_0 a_k \left( \sum_{j=-m-k}^m \text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \right) \\
+ \sum_{k=1}^L \sum_{l=1}^L a_k a_l \left( \sum_{i=-m-k}^{m+l} \text{cov}((X_i - \mu)(X_{i+k} - \mu), (X_0 - \mu)(X_{i} - \mu)) \right) \\
= \mathbf{a}' \Sigma \mathbf{a} \text{ where } \mathbf{a}' = (a_0, \ldots, a_L) \text{ and } \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.
\]

\[
\Sigma_{11} = \gamma^2 + \sum_{j=-m}^m c(j; \theta_0), \\
\Sigma_{12} = \Sigma'_{21} = \left[ \sum_{j=-m-k}^m \text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \right]_{1 \times L}, \text{ and } \\
\Sigma_{22} = \left[ \sum_{i=-m-k}^{m+l} \text{cov}((X_i - \mu)(X_{i+k} - \mu), (X_0 - \mu)(X_{i} - \mu)) \right]_{L \times L}.
\]

Theorem 3.1 now follows by the Cramér-Wold device.

### 3.4.2 Proof of Theorem 3.2

The consistency of the least squares estimator result follows by verifying the conditions of the general consistency result given in Brillinger [4]. Brillinger's lemma
is restated in Section 4.6 where the consistency of an asymptotic likelihood estimator, based on periodograms, is discussed. The proof of the consistency of the least squares estimator parallels the proof of Theorem 4.2, Section 4.6. The result follows from the fact that

\[ H_n(\theta) \overset{P}{\rightarrow} H(\theta) \]

uniformly on compact sets \( K \), where

\[ H(\theta) = \sum_{k=1}^{L} (c(k; \theta_0) - c(k; \theta))^2. \]

### 3.4.3 Proof of Theorem 3.3

By Taylor's Theorem one can write:

\[ \sqrt{n} \left[ \begin{array}{c} \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{array} \right] = 2 \begin{bmatrix} \frac{\partial}{\partial \theta_1} H_n(\theta^*) & \frac{\partial}{\partial \theta_1 \theta_2} H_n(\theta^*) \\ \frac{\partial}{\partial \theta_2 \theta_1} H_n(\theta^*) & \frac{\partial}{\partial \theta_2} H_n(\theta^*) \end{bmatrix}^{-1} \begin{bmatrix} -\frac{n}{2} \frac{\partial}{\partial \theta_1} H_n(\theta_0) \\ -\frac{n}{2} \frac{\partial}{\partial \theta_2} H_n(\theta_0) \end{bmatrix} \]

\[ = \Lambda_n \begin{bmatrix} \sqrt{n} \partial c_1'(\hat{\theta} - c(\theta_0)) \\ \sqrt{n} \partial c_2'(\hat{\theta} - c(\theta_0)) \end{bmatrix} \]

where \( \theta^* \) lies between \( \hat{\theta} \) and \( \theta_0 \),

\[ \partial c_i = \begin{bmatrix} \frac{\partial}{\partial \theta_i} (1; \theta_0) \\ \vdots \\ \frac{\partial}{\partial \theta_i} (L; \theta_0) \end{bmatrix}, \quad i = 1, 2, \quad \hat{\theta} = \begin{bmatrix} \hat{\theta}(1) \\ \vdots \\ \hat{\theta}(L) \end{bmatrix} \quad \text{and} \quad c(\theta_0) = \begin{bmatrix} c(1; \theta_0) \\ \vdots \\ c(L; \theta_0) \end{bmatrix}. \]

Now,

\[ \sqrt{n} \begin{bmatrix} X - \mu \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \Lambda_n \end{bmatrix} \begin{bmatrix} \sqrt{n}(X - \mu) \\ \sqrt{n} \partial c_1'(\hat{\theta} - c(\theta_0)) \\ \sqrt{n} \partial c_2'(\hat{\theta} - c(\theta_0)) \end{bmatrix} \]

where \( \Theta = (0, 0) \). The proof again employs the Cramér-Wold device. Consider:

\[ a_0 \sqrt{n}(X - \mu) + \sqrt{n} \sum_{k=1}^{L} (\hat{c}(k) - c(k; \theta_0))(a_1 \frac{\partial}{\partial \theta_1}(k; \theta_0) + a_2 \frac{\partial}{\partial \theta_2}(k; \theta_0)) \]
\[ = b_0 \sqrt{n}(X - \mu) + \sqrt{n} \sum_{k=1}^{1} (c(k) - c(k; \theta_0))b_k \]

\[ \overset{D}{\rightarrow} N(0, b'\Sigma b) \]

from Theorem 3.1.

\[ \mathbf{b} = a_0 \mathbf{e}_1 + a_1 \partial \mathbf{c}_1 + a_2 \partial \mathbf{c}_2 \quad \mathbf{e}'_1 = (1, 0, \ldots, 0)_{1 \times L+1} \text{ and} \]

\[ \partial \mathbf{c}_{i}' = (0, \frac{\partial c}{\partial \theta_i}(1; \theta_0), \ldots, \frac{\partial c}{\partial \theta_i}(L; \theta_0)), \quad i = 1, 2. \text{ Thus,} \]

\[ b'\Sigma b = (a_0 e_1 + a_1 \partial c_1 + a_2 \partial c_2)' \Sigma (a_0 e_1 + a_1 \partial c_1 + a_2 \partial c_2) \]

\[ = [a_0 \ a_1 \ a_2] \begin{bmatrix} e'_1 \Sigma e_1 & e'_1 \Sigma \partial c_1 & e'_1 \Sigma \partial c_2 \\ \partial c_1' \Sigma e_1 & \partial c_1' \Sigma \partial c_1 & \partial c_1' \Sigma \partial c_2 \\ \partial c_2' \Sigma e_1 & \partial c_2' \Sigma \partial c_1 & \partial c_2' \Sigma \partial c_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \]

\[ = a'\Sigma^* a. \]

Therefore,

\[ \begin{bmatrix} \sqrt{n}(X - \mu) \\ \sqrt{n} \partial \mathbf{c}_1'(\hat{c} - c(\theta_0)) \\ \sqrt{n} \partial \mathbf{c}_2'(\hat{c} - c(\theta_0)) \end{bmatrix} \overset{D}{\rightarrow} N_3(0, \Sigma^*). \]

The matrix \( A_n \) defined earlier converges in probability to a matrix \( A \) because

\[ \hat{c} \overset{p}{\rightarrow} c(\theta_0) \text{ and } \hat{\theta} \overset{p}{\rightarrow} \theta_0 \text{ (hence } \theta^* \overset{p}{\rightarrow} \theta_0): \]

\[ A = \begin{bmatrix} \partial \mathbf{c}_1' \partial \mathbf{c}_1 & \partial \mathbf{c}_1' \partial \mathbf{c}_2 \\ \partial \mathbf{c}_2' \partial \mathbf{c}_1 & \partial \mathbf{c}_2' \partial \mathbf{c}_2 \end{bmatrix}^{-1}. \]

By Slutsky's Theorem, one then has,

\[ \sqrt{n} \begin{bmatrix} X - \mu \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} \overset{D}{\rightarrow} N_3(0, C\Sigma^*(C') \quad \text{where} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}. \]
3.4.4 Proof of Theorem 3.4

By Taylor’s Theorem one can write

\[
\sqrt{n} \left( \frac{c(0; \theta)}{X^2} - \frac{c(0; \theta_0)}{\mu^2} \right) = \left[ \frac{\partial}{\partial \mu} \left( \frac{c(0; \theta^*)}{\mu^2} \right), \frac{\partial}{\partial \theta_1} \left( \frac{c(0; \theta^*)}{\mu^2} \right), \frac{\partial}{\partial \theta_2} \left( \frac{c(0; \theta^*)}{\mu^2} \right) \right] \left[ \frac{\sqrt{n}(X - \mu)}{\sqrt{n}(\hat{\theta}_1 - \theta_{1,0})}, \frac{\sqrt{n}(X - \mu)}{\sqrt{n}(\hat{\theta}_2 - \theta_{2,0})} \right]
\]

where \((\theta^*, \mu^*)\) lies between \((\hat{\theta}, \bar{X})\) and \((\theta_0, \mu)\). Now \((\hat{\theta}, \bar{X}) \xrightarrow{P} (\theta_0, \mu)\) implies \((\theta^*, \mu^*) \xrightarrow{P} (\theta_0, \mu)\) and hence the vector of partial derivatives above converges in probability to

\[
\left[ \frac{\partial}{\partial \mu} \left( \frac{c(0; \theta_0)}{\mu^2} \right), \frac{\partial}{\partial \theta_1} \left( \frac{c(0; \theta_0)}{\mu^2} \right), \frac{\partial}{\partial \theta_2} \left( \frac{c(0; \theta_0)}{\mu^2} \right) \right]
\]

which we shall denote by \(c'\).

Then by Slutsky’s Theorem and Theorem 3.3, one obtains

\[
\sqrt{n} \left( \frac{c(0; \hat{\theta})}{X^2} - \frac{c(0; \theta_0)}{\mu^2} \right) \xrightarrow{D} N(0, c' \Sigma^* C' c).
\]

3.5 Asymptotic Results for \(X_k\) Defined Over \(R^2\)

Let

\[
X_k = \int_{R^2} f_k(x, y) Y_{(x, y)} N(dxdy)
\]

and

\[
X_k^{(m)} = \int_{S_k} f_k(x, y) Y_{(x, y)} N(dxdy)
\]

where \(m\) is the finite-range dependence parameter determined by

\[
m = \left[ \frac{6w}{\delta \sqrt{2}} \right] - 1.
\]
To simplify this discussion, the noise parameter $\gamma$ is set to zero. The limit theorems in Section 3.3 involve linear and quadratic moment functionals of the sequence \( \{X_k^{(m)}\} \) where explicit dependence upon $m$ is now indicated. These limit theorems continue to hold for functionals of \( \{X_k\} \). The following discussion considers only the linear case, specifically the sample mean $X$ based on $X_0, \ldots, X_{n-1}$. The asymptotic normality of sample covariances is obtained by an analogous argument.

Let

\[
X = \frac{1}{n} \sum_{k=0}^{n-1} X_k
\]

and

\[
X^{(m)} = \frac{1}{n} \sum_{k=0}^{n-1} X_k^{(m)}
\]

be the respective sample means of the data sequences $X_0, \ldots, X_{n-1}$ and $X_0^{(m)}, \ldots, X_{n-1}^{(m)}$. Also, let $E(X_k) = \mu$ and $E(X_k^{(m)}) = \mu^{(m)}$. The disc region $S_k$ has diameter $(m+1)\delta$ for a fixed stepsize parameter $\delta$. This follows from the relation

\[
 m = \left\lfloor \frac{\text{diameter} (S_k)}{\delta} \right\rfloor - 1.
\]

Let

\[
\text{cov} (X_0, X_\ell) = c(\ell; \theta_0),
\]

\[
\text{cov} (X_0^{(m)}, X_\ell^{(m)}) = c^{(m)}(\ell; \theta_0)
\]

and

\[
\text{cov} (X_0, X_\ell^{(m)}) = \gamma^{(m)}(\ell; \theta_0),
\]

where

\[
c(\ell; \theta_0) = \lambda \mu_{2y} \int_{\mathbb{R}^2} f_0(x, y) f_\ell(x, y) dx dy,
\]
\[ c^{(m)}(\ell; \theta_0) = \lambda \mu_{2y} \int_{S_0 \cap S_0^y} f_0(x, y) f_\ell(x, y) \, dx \, dy \]

and

\[ \gamma^{(m)}(\ell; \theta_0) = \lambda \mu_{2y} \int_{S_0 \cap R_0^y} f_0(x, y) f_\ell(x, y) \, dx \, dy . \]

Then, one can write

\[ \sqrt{n}(X - \mu) = \sqrt{n}(X^{(m)} - \mu^{(m)}) + \sqrt{n}R^{(m)} , \]  

(3.9)

where

\[ R^{(m)} = (X - \bar{X}^{(m)}) - (\mu - \mu^{(m)}) . \]

Consider the first term in (3.9). From Theorem 3.1,

\[ \sqrt{n}(X^{(m)} - \mu^{(m)}) \xrightarrow{D} N(0, \sigma_m^2) \]

where

\[ \sigma_m^2 = \sum_{\ell = -\infty}^{\infty} c^{(m)}(\ell; \theta_0) = \sum_{\ell = -\infty}^{m} c^{(m)}(\ell; \theta_0) . \]

Let

\[ \sigma^2 = \sum_{\ell = -\infty}^{\infty} c(\ell; \theta_0) < \infty . \]

Then, by the dominated convergence theorem, \( \sigma_m^2 \to \sigma^2 \) as \( m \to \infty \).

Chebyshev's inequality can be applied to the second term in (3.9).

\[ P(\sqrt{n}R^{(m)} > \epsilon) \leq \frac{n}{\epsilon^2} \text{Var} (X - X^{(m)}) \]

where

\[ n \text{Var} (X - X^{(m)}) = \sum_{|j| < n} (1 - \frac{|j|}{n}) c(j; \theta_0) + \sum_{|j| < n} (1 - \frac{|j|}{n}) c^{(m)}(j; \theta_0) \]

\[ -2 \sum_{|j| < n} (1 - \frac{|j|}{n}) \gamma^{(m)}(j; \theta_0) . \]  

(3.10)
By dominated convergence, for fixed \( m \),

\[
n \operatorname{Var} (X - X^{(m)}) \to \sum_{j=\infty}^{\infty} e(j; \theta_0) \tag{3.11}
\]

\[
+ \sum_{j=\infty}^{\infty} e^{(m)}(j; \theta_0) - 2 \sum_{j=\infty}^{\infty} \gamma^{(m)}(j; \theta_0) . \quad \text{as} \quad n \to \infty .
\]

Also, as \( m \to \infty \), the second term in (3.11) converges to \( \sum_j e(j; \theta_0) \) while the third term converges to \( -2 \sum_j c(j; \theta_0) \). Thus, as \( n, m \to \infty \),

\[
n \operatorname{Var} (X - X^{(m)}) \to 0
\]

implying that \( \sqrt{n} R^{(m)} = o_P(1) \) as \( m \to \infty \). Combining this result with the convergence results concerning \( \sqrt{n}(X^m - \mu^{(m)}) \) and returning to (3.9), one obtains,

\[
\sqrt{n}(X - \mu) \xrightarrow{D} N(0, \sigma^2) ,
\]

where

\[
\sigma^2 = \sum_{\ell=\infty}^{\infty} r(\ell; \theta_0) , \quad \text{and}
\]

\[
c(\ell; \theta_0) = \lambda \mu_2 \nu \int_{R^2} f_0(x, y)f_1(x, y)dxdy .
\]

The same argument pertains to the joint asymptotic normality of \( X \) and the \( c(k) \)'s. One can thus replace the original \( m \)-dependent sequence with an asymptotically independent sequence and obtain the same limit theorems. This provides analytic expressions for moments of interest as discussed in Section 3.6.

### 3.6 Moment Calculations

The asymptotic covariance matrix \( \Sigma \) appearing in Theorem 3.1 involves various moments of the \( X_k \)'s, up to fourth order. For example, in the asymptotic covari-
ance matrix of the \( \hat{c}(k) \)'s \( \Sigma_{22} \), one has an expression of the form:

\[
\text{cov}((X_0 - \mu)(X_I - \mu), (X_J - \mu)(X_{J+k} - \mu))
= E((X_0 - \mu)(X_I - \mu)(X_J - \mu)(X_{J+k} - \mu)) - c(k; \theta_0)c(l; \theta_0).
\]

Recall now the definition of \( X_k \), \( X_k = Z_k + \gamma \varepsilon_k \) given in (3.4), where \( Z_k \) and \( \varepsilon_k \) are independent and

\[
Z_k = \int_{S_k} e^{-\frac{1}{2}(x-k)^2+\nu^2)} Y(x,y) N(\text{d}x \text{d}y).
\]

\[
E(Z_0 Z_I Z_J Z_{J+k})
= E\left\{ \int_{S_0} f_0(s)Y_0 N(\text{d}s) \int_{S_I} f_I(t)Y_I N(\text{d}t) \int_{S_J} f_J(u)Y_J N(\text{d}u) \right. \\
\left. \int_{S_{J+k}} f_{J+k}(v)Y_{J+k} N(\text{d}v) \right\}
= \lambda^4 \mu_Y^4 \int_{S_0} \int_{S_I} \int_{S_J} \int_{S_{J+k}} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_{J+k}} \int_{S_J} \int_{S_{J+k}} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_I} \int_{S_{J+k}} \int_{S_{0} \cap S_J} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_J} \int_{S_{J+k}} \int_{S_0 \cap S_{J+k}} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_{J+k}} \int_{S_I} \int_{S_0 \cap S_{J+k}} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_J} \int_{S_{J+k}} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_{J+k}} \int_{S_I} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^3 E(Y^2) \mu_Y^3 \int_{S_J} \int_{S_{J+k}} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^2 E(Y^2)^2 \int_{S_{J+k}} \int_{S_J} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^2 E(Y^2)^2 \int_{S_{J+k}} \int_{S_I} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v
+ \lambda^2 E(Y^2)^2 \int_{S_J} \int_{S_{J+k}} \int_{S_0} f_0(s)f_I(t)f_J(u)f_{J+k}(v)d\text{d}t \text{d}u \text{d}v

\[ + \lambda^2 E(Y^2)(Y^2) \int_{S_i \cap S_j} \int_{S_i} f_0(s)f_i(t)f_j(t)f_{j+k}(s)dsdt \]

\[ + \lambda^2 E(Y^3)\mu Y \int_{S_i} \int_{S_j} f_0(s)f_i(s)f_{j+k}(v)dsdv \]

\[ + \lambda^2 E(Y^3)\mu Y \int_{S_i \cap S_j} \int_{S_i \cap S_j} f_0(s)f_i(t)f_j(t)f_{j+k}(t)dsdt \]

\[ + \lambda^2 E(Y^3)\mu Y \int_{S_i \cap S_j} \int_{S_i \cap S_j} f_0(s)f_i(s)f_j(s)f_{j+k}(s)dsdt \]

\[ + \lambda^2 E(Y^3)\mu Y \int_{S_i \cap S_j} \int_{S_i \cap S_j} f_0(s)f_i(s)f_j(s)f_{j+k}(s)dsdu \]

\[ + \lambda E(Y^4) \int_{S_i \cap S_j} f_0(s)f_i(s)f_j(s)f_{j+k}(s)ds. \]

The derivation of the expressions for the other product moments is analogous to that above.

\[
E(X_0X_iX_jX_{j+k})
= E((Z_0 + \epsilon_0)(Z_i + \epsilon_i)(Z_j + \epsilon_j)(Z_{j+k} + \epsilon_{j+k}))
= E(Z_0Z_iZ_jZ_{j+k}) + E(Z_0Z_i)E(\epsilon_j\epsilon_{j+k}) + E(Z_0Z_j)E(\epsilon_i\epsilon_{j+k})
+ E(Z_0Z_{j+k})E(\epsilon_i\epsilon_j) + E(Z_0)E(\epsilon_j\epsilon_{j+k}) + E(Z_iZ_j)E(\epsilon_i\epsilon_{j+k})
+ E(Z_iZ_{j+k})E(\epsilon_0\epsilon_j) + E(Z_i)E(\epsilon_0\epsilon_{j+k}) + E(Z_jZ_{j+k})E(\epsilon_0\epsilon_i)
+ E(Z_j)E(\epsilon_0\epsilon_{j+k}) + E(Z_{j+k})E(\epsilon_0\epsilon_l) + E(\epsilon_0\epsilon_l\epsilon_{j+k}).
\]

To avoid the roundoff error associated with the numerical integration of such an expression and so facilitate the evaluation of the matrix \( \Sigma \), one can carry the integration over \( \mathbb{R}^2 \) instead of \( S_k \). This provides analytic expressions for the moments. Take

\[
Z_k = \int_{\mathbb{R}^2} e^{-\frac{1}{2}((x-k)^2+v^2)}Y(x,y)N(dxdy).
\]

As it is also representative of the other moments, consider \( E(Z_aZ_bZ_cZ_d) \).
Let
\[ f_a = f_a(x, y) = f_0(x - ab, y) = e^{-\frac{1}{2z}((x - ab)^2 + y^2)} \]
and
\[ c(a - b) = \lambda \mu_{2z} \int_{\mathbb{R}^2} f_a f_b dxdy \]
where
\[
\int_{\mathbb{R}^2} f_a f_b dxdy = \int_{\mathbb{R}^2} e^{-\frac{1}{2z}((x - ab)^2 + (x - bb)^2 + 2y^2)} dxdy
\]
\[
= \int_{\mathbb{R}} e^{-\frac{1}{\sqrt{2\sigma}}(x - \frac{1}{2}(a+b))^2} dx \cdot \int_{\mathbb{R}} e^{-\frac{1}{\sqrt{2\sigma}}y^2} dy \cdot e^{-\frac{b^2}{2\sigma}((a^2 + b^2) - (\frac{a+b}{2})^2)}
\]
\[
= \frac{\pi w^2}{2} e^{-\frac{b^2}{2\sigma}(a^2 + b^2) - (\frac{a+b}{2})^2}
\]
\[
= \frac{\pi w^2}{2} e^{-\frac{b^2}{2\sigma}(a-b)^2}.
\]

Then
\[ E(Z_a Z_b Z_c Z_d) \]
\[
= \lambda \mu_{2z} \int_{\mathbb{R}^2} f_a f_b f_c f_d dxdy + \mu Z \cdot \lambda \mu_{3z} \left\{ \int_{\mathbb{R}^2} f_b f_c f_d dxdy + \int_{\mathbb{R}^2} f_a f_c f_d dxdy \right\}
\]
\[
+ \{ c(a - b)c(c - d) + c(a - c)c(b - d) + c(a - d)c(b - c) \}
\]
\[
+ \mu_{2z} \{ c(a - b) + c(a - c) + c(a - d) + c(b - c) + c(b - d) + c(c - d) \}
\]
\[
+ \mu_{2z}^2,
\]
where \( \mu_Z = \mu_X = \mu = E(X_0) \). Similarly
\[
\int_{\mathbb{R}^4} f_a f_b f_c f_d dxdy
\]
\[
\begin{align*}
&= \int_{\mathbb{R}^2} e^{- \frac{1}{2} \left( \left( x - \frac{d}{\delta} (a+b+c+d) \right)^2 + 4y^2 + b^2 \left( a^2 + b^2 + c^2 + d^2 - \left( \frac{a+b+c+d}{\delta} \right)^2 \right) \right) } dx \, dy \\
&= \int_{\mathbb{R}} e^{- \frac{1}{2} \left( x - \frac{d}{\delta} (a+b+c+d) \right)^2 } dx \cdot \int_{\mathbb{R}} e^{- \frac{4y^2}{2\delta^2} } dy \cdot \int_{\mathbb{R}} e^{- \frac{b^2}{2\delta^2} \left( a^2 + b^2 + c^2 + d^2 - \left( \frac{a+b+c+d}{\delta} \right)^2 \right) } dy \\
&= \frac{\pi w^2}{4} e^{- \frac{d^2}{\delta^2} \left( a^2 + b^2 + c^2 + d^2 \right) - \left( \frac{a+b+c+d}{\delta} \right)^2 } \\
\end{align*}
\]

and

\[
\int_{\mathbb{R}^2} f_a f_b f_c dx = \frac{\pi w^2}{3} e^{- \frac{d^2}{\delta^2} \left( a^2 + b^2 + c^2 \right) - \left( \frac{a+b+c+d}{\delta} \right)^2 } .
\]

### 3.7 Components of \( \Sigma \)

#### 3.7.1 \( \Sigma_{11} \)

\[
\Sigma_{11} = \lim_{n \to \infty} n \, \text{Var}(X) = \gamma^2 + \sum_{j=-\infty}^{\infty} c(j; \theta_0) \\
= \gamma^2 + \sum_{j=-\infty}^{\infty} \lambda \mu_2 \gamma \frac{\pi w^2}{2} e^{- \frac{d^2}{\delta^2} } .
\]

#### 3.7.2 \( \Sigma_{12} \)

\[
\sigma_{12}(k) = \lim_{n \to \infty} n \, \text{cov}(X, \hat{c}(k)), \quad k = 1, \ldots, L \\
= \sum_{j=-\infty}^{\infty} \text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) ,
\]

Where,

\[
\text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \\
= E(X_0(X_j - \mu)(X_{j+k} - \mu)) - \mu E((X_0 - \mu)(X_k - \mu))
\]

and

\[
E(X_0(X_j - \mu)(X_{j+k} - \mu))
\]
\begin{align*}
&= E(X_0(X_j X_{j+k} - \mu X_j - \mu X_{j+k} + \mu^2)) \\
&= E(X_0 X_j X_{j+k}) - \mu E(X_0 X_j) - \mu E(X_0 X_{j+k}) + \mu^3 \\
&= E(Z_0 Z_j Z_{j+k}) + \mu E(\epsilon_j \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_j) + \mu E(\epsilon_0 \epsilon_{j+k}) \\
&\quad + \mu E(\epsilon_j \epsilon_{j+k}) - \mu (E(Z_0 Z_j) + E(\epsilon_0 \epsilon_j)) - \mu (E(Z_0 Z_{j+k}) + E(\epsilon_0 \epsilon_{j+k})) + \mu^3 \\
&= E(Z_0 Z_j Z_{j+k}) + \mu E(\epsilon_j \epsilon_{j+k}) + E(\epsilon_0 \epsilon_j \epsilon_{j+k}) \\
&\quad - \mu (E(Z_0 Z_j) + E(Z_0 Z_{j+k})) + \mu^3 \\
&= E(Z_0 Z_j Z_{j+k}) + \mu E(\epsilon_j \epsilon_{j+k}) + E(\epsilon_0 \epsilon_j \epsilon_{j+k}) - \mu (c(j) + c(j+k)) - \mu^3
\end{align*}

Therefore,

\begin{align*}
\text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \\
&= E(Z_0 Z_j Z_{j+k}) - \mu (c(j) + c(j+k)) - \mu^3 \\
&\quad - \mu E((X_0 - \mu)(X_k - \mu)) \\
&= E(Z_0 Z_j Z_{j+k}) - \mu (c(j) + c(j+k) + c(k)) - \mu^3 \\
&= \lambda \mu_3 \int \int f_0 f_j f_{j+k} \, dx \, dy \\
&\quad + \mu \{c(k) + c(j+k) + c(j)\} + \mu^3 \\
&\quad - \mu \{c(j) + c(j+k) + c(k)\} - \mu^3 \\
&= \lambda \mu_3Y \int \int f_0 f_j f_{j+k} \, dx \, dy \\
&= \lambda \mu_3Y \int \int f_0 f_j f_{j+k} \, dx \, dy \\
&= \lambda \mu_3Y \cdot \frac{\pi u^3}{3} e^{-\frac{2}{\sigma^2}(j^2 + (j+k)^2 - \frac{(2j+k)^2}{3})}.
\end{align*}

Thus,

\begin{align*}
\sigma_{12}(k) &= \lambda \mu_3Y \cdot \frac{\pi u^2}{3} \sum_{j=-\infty}^{\infty} e^{-\frac{2}{\sigma^2}(j^2 + (j+k)^2 - \frac{(2j+k)^2}{3})}.
\end{align*}
\[ \sigma_{22}(k,l) = \lim_{n \to \infty} n \text{ cov}(\hat{c}(k), \hat{c}(l)) \]

\[ = \sum_{j=-\infty}^{\infty} \text{ cov}((X_0 - \mu)(X_l - \mu), (X_j - \mu)(X_{j+k} - \mu)) \]

where

\[ \text{cov}((X_0 - \mu)(X_l - \mu), (X_j - \mu)(X_{j+k} - \mu)) \]

\[ = E((X_0 - \mu)(X_l - \mu)(X_j - \mu)(X_{j+k} - \mu)) - \text{cov}(X_0, X_l) \text{cov}(X_0, X_k). \]

Now,

\[ E((X_0 - \mu)(X_l - \mu)(X_j - \mu)(X_{j+k} - \mu)) \]

\[ = E(X_0 X_l X_j X_{j+k}) \]

\[ - \mu(E(X_0 X_l X_j) + E(X_0 X_l X_{j+k}) + E(X_0 X_j X_{j+k}) + E(X_l X_j X_{j+k})) \]

\[ + \mu^2 (E(X_0 X_l) + E(X_0 X_j) + E(X_0 X_{j+k}) + E(X_l X_j) \]

\[ + E(X_l X_{j+k}) + E(X_j X_{j+k})) - 3\mu^4 \]

\[ = E(Z_0 Z_l Z_j Z_{j+k}) + E(\epsilon_0 \epsilon_{j+k})(c(l) + \mu^2) \]

\[ + E(\epsilon_0 \epsilon_{j+k})(c(k) + \mu^2) + E(\epsilon_0 \epsilon_{j})(c(j + k) + \mu^2) \]

\[ + \mu E(\epsilon_0 \epsilon_{j+k}) + E(\epsilon_0 \epsilon_{j+k})(c(j - l) + \mu^2) \]

\[ + E(\epsilon_0 \epsilon_{j})(c(j + k - l) + \mu^2) + \mu E(\epsilon_0 \epsilon_{j+k}) \]

\[ + E(\epsilon_0 \epsilon_{j})(c(k) + \mu^2) + \mu E(\epsilon_0 \epsilon_{j+k}) \]

\[ + \mu E(\epsilon_0 \epsilon_{j}) + E(\epsilon_0 \epsilon_{j+k}) \]

\[ - \mu \{ E(Z_0 Z_l Z_j) + \mu E(\epsilon_0 \epsilon_{j}) + \mu E(\epsilon_0 \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_{j}) + E(\epsilon_0 \epsilon_{j+k}) \} \]
\[ + E(Z_0 Z_j Z_{j+k}) + \mu E(\epsilon_1 \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_l) \]
\[ + E(Z_0 Z_j Z_{j+k}) + \mu E(\epsilon_j \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_{j+k}) + \mu E(\epsilon_0 \epsilon_j) \]
\[ + E(Z_j Z_j Z_{j+k}) + \mu E(\epsilon_j \epsilon_{j+k}) + \mu E(\epsilon_1 \epsilon_{j+k}) + \mu E(\epsilon_1 \epsilon_j) + E(\epsilon_j \epsilon_{j+k}) \]
\[ + \mu^2 \left\{ c(l) + \mu^2 + c(j) + \mu^2 + c(j + k) + \mu^2 + c(j - l) + \mu^2 \right\} \]
\[ + E(\epsilon_0 \epsilon_{j+k}) + E(\epsilon_1 \epsilon_{j+k}) + E(\epsilon_j \epsilon_{j+k}) + E(\epsilon_l \epsilon_j) \}
\[ - 3 \mu^4 \]
\[ = \lambda \lambda_4 \int f_0 f_1 f_2 f_{j+k} dxdy + \mu \lambda_{3Y} \left\{ \int f_1 f_2 f_{j+k} dxdy + \int f_0 f_j f_{j+k} dxdy \right\} \]
\[ + \int f_0 f_1 f_{j+k} dxdy + \int f_0 f_1 f_j dxdy \left\{ c(l) c(k) + c(j) c(j + k - l) + c(j + k) c(j - l) + \mu^2 \left\{ c(l) + c(j) + c(j + k) + c(j - l) + c(j + k - l) + c(k) \right\} + \mu^4 \right\} \]
\[ + c(j - l) E(\epsilon_0 \epsilon_{j+k}) + c(j + k - l) E(\epsilon_0 \epsilon_j) + c(k) E(\epsilon_0 \epsilon_l) \]
\[ + E(\epsilon_0 \epsilon_1 \epsilon_j \epsilon_{j+k}) \]
\[ - \mu \left\{ \lambda \lambda_3 Y \int f_0 f_1 f_j dxdy + \mu (c(j - l) + c(j) + c(l)) + \mu^3 \right\} \]
\[ + \lambda \lambda_3 Y \int f_0 f_1 f_j dxdy + \mu (c(j + k - l) + c(j + k) + c(l)) + \mu^3 \right\} \]
\[ + \lambda \lambda_3 Y \int f_0 f_1 f_{j+k} dxdy + \mu (c(k) + c(j + k) + c(j)) + \mu^3 \right\} \]
\[ + \lambda \lambda_3 Y \int f_1 f_2 f_{j+k} dxdy + \mu (c(j + k - l) + c(j - l) + c(j + k - l) + c(k)) + \mu^3 \right\} \]
\[ + c(j - l) E(\epsilon_1 \epsilon_{j+k}) + c(l) c(k) + c(j) c(j + k - l) + c(j + k) c(j - l) + \mu^4 \right\} \]
\[ + c(j - l) E(\epsilon_1 \epsilon_{j+k}) + c(l) E(\epsilon_1 \epsilon_{j+k}) + c(j + k) E(\epsilon_1 \epsilon_j) \]
\[ + c(j - l) E(\epsilon_0 \epsilon_{j+k}) + c(j + k - l) E(\epsilon_0 \epsilon_j) + c(k) E(\epsilon_0 \epsilon_l) \]
\[ + E(\epsilon_0 \epsilon_i \epsilon_j \epsilon_{j+k}) \]
\[ - 4\mu^4 + 6\mu^4 - 3\mu^4. \]

Now

\[
\text{cov}(X_0, X_i) \text{cov}(X_0, X_k)
\]
\[ = (\text{cov}(Z_0, Z_i) + E(\epsilon_0 \epsilon_i))(\text{cov}(Z_0, Z_k) + E(\epsilon_0 \epsilon_k)) \]
\[ = c(l)c(k) + c(l)E(\epsilon_0 \epsilon_k) + c(k)E(\epsilon_0 \epsilon_i) + E(\epsilon_0 \epsilon_i)E(\epsilon_0 \epsilon_k). \]

Therefore

\[
\text{cov}((X_0 - \mu)(X_i - \mu), (X_j - \mu)(X_{j+k} - \mu))
\]
\[ = \lambda \mu_{4Y} \int_0^1 f_0 f_1 f_{j+k} dxdy + c(j)c(j + k - l) + c(j + k)c(j - l) \]
\[ + c(j)E(\epsilon_i \epsilon_{j+k}) + c(j + k)E(\epsilon_i \epsilon_j) \]
\[ + c(j - l)E(\epsilon_0 \epsilon_{j+k}) + c(j + k - l)E(\epsilon_0 \epsilon_j) \]
\[ + E(\epsilon_0 \epsilon_i \epsilon_j \epsilon_{j+k}). \]

Thus,

\[
\sigma_{22}(k, l) = \sum_{j=-\infty}^{\infty} \left\{ \lambda \mu_{4Y} \frac{\pi w^2}{4} e^{-\frac{\xi^2}{w^2} (l^2 + j^2 + (j+k)^2 + (j+l)^2)} \right\}
\]
\[ + c(j)c(j + k - l) + c(j + k)c(j - l) \]
\[ + c(j)E(\epsilon_i \epsilon_{j+k}) + c(j + k)E(\epsilon_i \epsilon_j) \]
\[ + c(j - l)E(\epsilon_0 \epsilon_{j+k}) + c(j + k - l)E(\epsilon_0 \epsilon_j) \]
\[ + E(\epsilon_0 \epsilon_i \epsilon_j \epsilon_{j+k}) \right\}, \; k, l = 1, \ldots, L. \]
3.8 Tuning Parameters

For given parameter values of the process, one can determine optimal choices for $\delta$ and the lag parameter $L$, that minimize the marginal variances of the least squares estimators and $\hat{R}$. This can then be done over an appropriate range of parameter values, corresponding to the actual quantities considered in the chemistry modeling problem. The typical parameter settings used in the numerical sections below, correspond to those observed in the laboratory of N.O. Petersen [24]. The optimization of $L$ can be done with respect to minimizing the limiting variance of any one of the three statistics, $\hat{\theta}_1$, $\hat{\theta}_2$ or $\hat{R}$. However, $R$ is the most important parameter [30]. The tuning parameter optimization is carried out with respect to a fixed length of the biological tissue sample. This implies that the number of fluorescence intensity measurements is proportional to $\delta^{-1}$.

In the case of no noise, the optimal $\delta$ turns out to be around .10 to .15, in the scales used here. When noise is present in the system, the optimal choice of $\delta$ is not achieved, but smaller values of $\delta$ produce smaller limiting variances. Because of finite precision digital computing, one cannot of course use $\delta$ too small. In the numerical section below $\delta$ is studied at .05, .025, and .005. $\delta = .025$ corresponds to the value used in Petersen’s laboratory. These results are consistent over the range of parameter settings studied.

The results for the optimal choice of $L$ are more interesting. For a given $\delta$, the limiting variances are calculated for $L = 2$ up to an appropriately large lag. Figure 3.8.1 plots, for particular parameters, the variance of $\hat{\theta}_1$ as a function of lag.
Figure 3.8.2 plots the variance of $\hat{R}$ as a function of lag $L$. It is clearly seen that a minimum variance is achieved. These plots are quite typical over the range of parameters studied. The figures are produced in Splus. Since Greek symbols could not be included in the text of these figures, equivalent symbols have been used. For example $\lambda = \lambda$, $\beta = \beta$, $\delta = \delta$, $\mu_Y = \mu_Y$, $\theta_1 = \theta_1$, and $R = \hat{R}$ in the figures.

The results over a range of parameters are summarized in Table 3.1. The column labeled $\beta$ indicates the noise level. Three levels of noise are examined, with the noise variance being proportional to $\epsilon(0) = Var(Z_0)$. Specifically, the label $\beta$ means that the noise variance is $\gamma^2 = \beta \epsilon(0)$. We study $\beta = 0$, no noise, $\beta = .5$ and $\beta = 2$. The column labeled $L_R$ refers to the optimal lag for the statistic $\hat{R}$ (see Theorem 3.4). Similarly, the other lag columns $L_{\theta_1}$ and $L_{\theta_2}$ have the obvious meaning.

For a given value of the Gaussian kernel parameter $w$, the optimal number of lags is quite consistent over the range of other parameters. When there is no noise ($\beta = 0$), $L_{\theta_1}$ and $L_{\theta_2}$ are small around 2 to 4 lags, and do not change with respect to $\delta$. The optimal lags are much larger when noise is present. They do change with respect to $\delta$, but not in proportion to $\delta^{-1}$. In all cases, the optimal lag number is quite consistent with respect to changes in $\lambda$ and $\mu_Y$, the average aggregate size. These lag numbers are close to those currently being used in practice in Petersen's laboratory.

The optimal lag $L_R$ is much larger than those corresponding to the estimation of $\theta_1$ and $\theta_2$. It is also more stable over both the noise level and parameter ranges,
Figure 3.8.1: Analytic Variance of $\hat{\theta}_1$ Versus Lag for various $\lambda$ and $\delta$, with $\beta = 0.5$
Figure 3.8.2: Analytic Variance of $\hat{R}$ Versus Lag for various $\lambda$ and $\delta$, with $\beta = 0.5$
### Table 3.1: Optimal Lags

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\mu_y$</th>
<th>$w$</th>
<th>$\delta$</th>
<th>$\beta$</th>
<th>$L_{\theta_1}$</th>
<th>$L_{\theta_2}$</th>
<th>$L_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.005</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>113</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.005</td>
<td>0.5</td>
<td>27</td>
<td>25</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.005</td>
<td>2.0</td>
<td>40</td>
<td>37</td>
<td>111</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.005</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>113</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.005</td>
<td>0.5</td>
<td>27</td>
<td>25</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.005</td>
<td>2.0</td>
<td>40</td>
<td>37</td>
<td>111</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.005</td>
<td>0.0</td>
<td>2</td>
<td>3</td>
<td>113</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.005</td>
<td>0.5</td>
<td>27</td>
<td>25</td>
<td>112</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.005</td>
<td>2.0</td>
<td>40</td>
<td>37</td>
<td>111</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.025</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.025</td>
<td>0.5</td>
<td>8</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.025</td>
<td>2.0</td>
<td>12</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.025</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.025</td>
<td>0.5</td>
<td>8</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.025</td>
<td>2.0</td>
<td>12</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.025</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.025</td>
<td>0.5</td>
<td>8</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.025</td>
<td>2.0</td>
<td>12</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.050</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>0.235</td>
<td>0.050</td>
<td>2.0</td>
<td>7</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.050</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.050</td>
<td>2.0</td>
<td>7</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.050</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>0.235</td>
<td>0.050</td>
<td>2.0</td>
<td>7</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.005</td>
<td>0.5</td>
<td>24</td>
<td>22</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.005</td>
<td>0.5</td>
<td>27</td>
<td>25</td>
<td>112</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.025</td>
<td>0.5</td>
<td>7</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.025</td>
<td>0.5</td>
<td>8</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.050</td>
<td>0.5</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>126</td>
<td>0.235</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.005</td>
<td>0.5</td>
<td>24</td>
<td>22</td>
<td>95</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.025</td>
<td>0.5</td>
<td>7</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.235</td>
<td>0.025</td>
<td>0.5</td>
<td>8</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.270</td>
<td>0.025</td>
<td>0.5</td>
<td>9</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.200</td>
<td>0.050</td>
<td>0.5</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.235</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>219</td>
<td>0.270</td>
<td>0.050</td>
<td>0.5</td>
<td>5</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>
for a given Gaussian kernel parameter $w$. For example, in the first three cases in Table 3.1, $L_R$ is 112 plus or minus 1 over the non noise case versus the intermediate and high noise levels. This is very consistent in comparison to $L_{\phi_1}$ or $L_{\phi_2}$, which vary from 2 to 40. A similar phenomenon occurs at each $\delta$ value. $L_R$ is also roughly proportional to $\delta^{-1}$, and depends only upon $\delta$ and $w$, where the kernel parameter value is determined at the outset of the experiment. The laser beam stepsize parameter $\delta$ is also set by the experimenter, and is therefore known. Thus, the optimal lag $L_R$ is known before the experiment is begun.

Since $R$ is the main parameter of interest [32, 30, 31], this is the optimal lag number that should be used. For example, if $\delta = .025$, the number of lags used should be around 20.

3.9 Numerical Results

S-FCS experiments are simulated using the proposed model discussed in Section 3.2. The parameter values are chosen to make the resulting data comparable to those observed in practice. The random number generator [35] is used. It involves three multiplicative congruential generators. Each generator uses a prime number for its modulus and a primitive root for its multiplier. The fractional part of the sum of the three results is taken, resulting in a pseudo-random $U(0,1)$ variate. The period of the algorithm exceeds $2.78 \times 10^{13}$. A Poisson $\lambda$ process over a large enough rectangular region is simulated to correspond to aggregate locations. The rectangle has length $n\delta + 6w/\sqrt{2}$ and width $6w/\sqrt{2}$, which is the effective region of observation for the fluorescence scanning experiment. Attached to each
Table 3.2: Parameter Cases for Simulations

<table>
<thead>
<tr>
<th>case</th>
<th>λ</th>
<th>w</th>
<th>β</th>
<th>δ</th>
<th>μ_Y</th>
<th>L_{opt}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.005</td>
<td>150</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.005</td>
<td>219</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.005</td>
<td>300</td>
<td>112</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.025</td>
<td>150</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.025</td>
<td>219</td>
<td>21</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.025</td>
<td>300</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.05</td>
<td>150</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.05</td>
<td>219</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0.235</td>
<td>0</td>
<td>0.05</td>
<td>300</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.005</td>
<td>150</td>
<td>112</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.005</td>
<td>219</td>
<td>112</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.005</td>
<td>300</td>
<td>112</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.025</td>
<td>150</td>
<td>21</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.025</td>
<td>219</td>
<td>21</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.025</td>
<td>300</td>
<td>21</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.05</td>
<td>150</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.05</td>
<td>219</td>
<td>10</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>0.235</td>
<td>0.5</td>
<td>0.05</td>
<td>300</td>
<td>10</td>
</tr>
</tbody>
</table>

point is a Poisson mark, with parameter μ_Y representing the average aggregate size. That is, μ_Y is the average number of monomers per aggregate. For a given δ, the \{X_k\} sequence (3.4), is then calculated.

The parameters used are given in Table 3.2. The sample size \( n \) depends upon the microscope stepsize δ. For the estimation procedure, the optimal \( L_R \) lag is used. The (δ, \( L_R, n \)) triples used in this simulation study are (.05, 10, 500), (.025, 21, 1000) and (.005, 112, 5000). The middle case most closely corresponds to that used in practice.
3.9.1 Point Estimates

The parameters $\theta_1 = \lambda E(Y^2)$ and $\theta_2 = \omega$ are estimated by Newton's method employing a rectangular grid search to find an appropriate starting value. The ratio estimator $\hat{R}$ is then calculated.

The parameter $\omega$ is estimated in these S FCS experiments as a check of consistency of the point estimates. It turns out to be useful, as explained later in this section.

For each parameter case in Table 3.3, 100 replications were made.

Figure 3.9.3 gives histograms of $\hat{R}$, for both noise and non noise cases. In a few of these parameter cases, some extremely large values of $\hat{R}$ occur. A few percent of the estimates $\hat{R}$ are far from their true values. For example, in parameter case 8, one point estimate of $R$ is 1200, about 400 times too large.

Recall that $\hat{\theta}_2$ is calculated by the lab for an internal consistency check. Upon investigation, it was found that extreme values of $\hat{\theta}_2$ always gave rise to extreme values of $\hat{R}$. Thus one can use $\hat{\theta}_2$ as a predictor of whether or not a particular experimental run should be removed as an outlier. In Figure 3.9.3, some large values of $\hat{R}$ are removed by truncating for extreme values of $\hat{\theta}_2$. A few outliers of $\hat{R}$ remain, but most have been removed by the $\hat{\theta}_2$ predictor. Figure 3.9.3 indicates that $\hat{R}$ is well approximated by the limiting normal distribution. Asymptotic confidence intervals for $R$ can thus be used.

The true standard deviation of $\hat{\theta}_2$ can be obtained from the variance Table 3.6 below, and is about .036 for each parameter case.

Table 3.3 gives averages of the various statistics with no trimming. In the
Figure 3.9.3: Histograms of \( \hat{R} \), Truncated at 3 Std Deviations of \( \hat{\theta}_2 \), and limiting Normal Approximation
column mean \( R \), the average \( \hat{R} \) is given. Cases 8 and 15 are very bad, and a few other cases are poor. In each of these cases, the empirical standard deviation of \( \hat{\theta}_2 \) is quite high, as seen in column Std dev(\( \hat{\theta}_2 \)). For example, in cases 8 and 15, the empirical standard deviations are .304 and .267, instead of the true value of .036.

Table 3.4 gives averages when trimming at three standard deviations of \( \hat{\theta}_2 \). Specifically, only those simulation runs with \( \hat{\theta}_2 \) being within three standard deviations of \( \theta_2 = .235 \) are averaged. For example in case 8, 97 simulation runs are averaged. This results in the empirical standard deviation of \( \hat{\theta}_2 \) dropping to .034, and the average of \( \hat{R} \) dropping to 2.90. Trimming in the other cases leads to some improvement in both \( \hat{R} \) and the standard deviation of \( \hat{\theta}_2 \). The amount of trimming is usually a little higher in the noise case, but extremes do occur even in the non noise case.

In actual S–FCS experiments, \( \hat{\theta}_2 \) is calculated for an internal consistency check. In conjunction with the variance estimation discussed in the next section, this internal consistency check can be done for a single S–FCS experiment.

### 3.9.2 Variance Estimation

The model and method would not be very useful in practice if one cannot obtain sample based estimates of the limiting variances. This section considers the problem of estimating the asymptotic or limiting variance in Theorem 3.3.

The limiting variance of \( \hat{R} \) is obtained from a matrix multiplication of \( \Sigma \) from Theorem 3.1 with some quantities that are known functions of \( (\theta_1, \theta_2) \). In the latter part, one can substitute the estimators \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) for the actual parameter
Table 3.3: Average Estimators. No Trimming. $w = .235$

<table>
<thead>
<tr>
<th>par case</th>
<th>$L$</th>
<th>$\theta_1$</th>
<th>$R$</th>
<th>mean($\hat{\theta}_1$)</th>
<th>mean($\hat{\theta}_2$)</th>
<th>Std dev($\hat{\theta}_2$)</th>
<th>mean($\hat{R}$)</th>
<th>n.rep</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112</td>
<td>22650</td>
<td>2.901</td>
<td>25049</td>
<td>0.222</td>
<td>0.031</td>
<td>2.904</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>112</td>
<td>48180</td>
<td>2.895</td>
<td>55900</td>
<td>0.223</td>
<td>0.044</td>
<td>3.171</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>112</td>
<td>90300</td>
<td>2.892</td>
<td>104131</td>
<td>0.225</td>
<td>0.052</td>
<td>3.333</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>22650</td>
<td>2.901</td>
<td>24815</td>
<td>0.224</td>
<td>0.029</td>
<td>3.172</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>48180</td>
<td>2.895</td>
<td>50518</td>
<td>0.243</td>
<td>0.090</td>
<td>4.445</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>90300</td>
<td>2.892</td>
<td>107108</td>
<td>0.228</td>
<td>0.047</td>
<td>3.323</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>22650</td>
<td>2.901</td>
<td>23373</td>
<td>0.225</td>
<td>0.040</td>
<td>3.058</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>48180</td>
<td>2.895</td>
<td>53943</td>
<td>0.257</td>
<td>0.304</td>
<td>14.765</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>90300</td>
<td>2.892</td>
<td>106512</td>
<td>0.225</td>
<td>0.040</td>
<td>3.474</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>112</td>
<td>22650</td>
<td>2.901</td>
<td>25229</td>
<td>0.228</td>
<td>0.046</td>
<td>3.228</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>112</td>
<td>48180</td>
<td>2.895</td>
<td>51253</td>
<td>0.239</td>
<td>0.127</td>
<td>4.963</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>112</td>
<td>90300</td>
<td>2.892</td>
<td>104450</td>
<td>0.230</td>
<td>0.102</td>
<td>4.173</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>22650</td>
<td>2.901</td>
<td>26753</td>
<td>0.224</td>
<td>0.033</td>
<td>3.125</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>21</td>
<td>48180</td>
<td>2.895</td>
<td>52572</td>
<td>0.234</td>
<td>0.095</td>
<td>4.048</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>90300</td>
<td>2.892</td>
<td>109349</td>
<td>0.255</td>
<td>0.267</td>
<td>9.062</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>22650</td>
<td>2.901</td>
<td>27499</td>
<td>0.216</td>
<td>0.035</td>
<td>3.162</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>48180</td>
<td>2.895</td>
<td>52847</td>
<td>0.229</td>
<td>0.070</td>
<td>3.468</td>
<td>100</td>
</tr>
<tr>
<td>18</td>
<td>10</td>
<td>90300</td>
<td>2.892</td>
<td>102636</td>
<td>0.225</td>
<td>0.047</td>
<td>3.745</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3.4: Average Estimators, Trimming at 3 Standard Deviations of $\hat{\theta}_2$, $w = .235$

<table>
<thead>
<tr>
<th>par case</th>
<th>$L$</th>
<th>$\theta_1$</th>
<th>$R$</th>
<th>mean($\hat{\theta}_1$)</th>
<th>mean($\hat{\theta}_2$)</th>
<th>std d $v(\hat{\theta}_2)$</th>
<th>mean($\hat{R}$)</th>
<th>$n_{rej}$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112</td>
<td>22650</td>
<td>2.901</td>
<td>25049</td>
<td>0.222</td>
<td>0.041</td>
<td>2.901</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>112</td>
<td>48180</td>
<td>2.895</td>
<td>55159</td>
<td>0.225</td>
<td>0.037</td>
<td>3.155</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>112</td>
<td>90300</td>
<td>2.892</td>
<td>100121</td>
<td>0.229</td>
<td>0.039</td>
<td>3.345</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>22650</td>
<td>2.961</td>
<td>24815</td>
<td>0.221</td>
<td>0.029</td>
<td>3.172</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>48180</td>
<td>2.895</td>
<td>50734</td>
<td>0.232</td>
<td>0.039</td>
<td>3.683</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>90300</td>
<td>2.892</td>
<td>105300</td>
<td>0.228</td>
<td>0.033</td>
<td>3.273</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>22650</td>
<td>2.901</td>
<td>23311</td>
<td>0.226</td>
<td>0.034</td>
<td>3.078</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>48180</td>
<td>2.895</td>
<td>53800</td>
<td>0.225</td>
<td>0.034</td>
<td>2.900</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>90300</td>
<td>2.892</td>
<td>105550</td>
<td>0.225</td>
<td>0.034</td>
<td>3.310</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>112</td>
<td>22650</td>
<td>2.901</td>
<td>25208</td>
<td>0.226</td>
<td>0.036</td>
<td>3.118</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>112</td>
<td>48180</td>
<td>2.895</td>
<td>50606</td>
<td>0.229</td>
<td>0.031</td>
<td>3.038</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>112</td>
<td>90300</td>
<td>2.892</td>
<td>104374</td>
<td>0.216</td>
<td>0.037</td>
<td>3.209</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>22650</td>
<td>2.901</td>
<td>26753</td>
<td>0.224</td>
<td>0.033</td>
<td>3.125</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>21</td>
<td>48180</td>
<td>2.895</td>
<td>51704</td>
<td>0.228</td>
<td>0.036</td>
<td>2.986</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>90300</td>
<td>2.892</td>
<td>106004</td>
<td>0.222</td>
<td>0.033</td>
<td>3.456</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>22650</td>
<td>2.901</td>
<td>26827</td>
<td>0.219</td>
<td>0.030</td>
<td>3.213</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>48180</td>
<td>2.895</td>
<td>51580</td>
<td>0.225</td>
<td>0.035</td>
<td>3.258</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>10</td>
<td>90300</td>
<td>2.892</td>
<td>101800</td>
<td>0.223</td>
<td>0.037</td>
<td>3.536</td>
<td>97</td>
<td></td>
</tr>
</tbody>
</table>
values. Thus, a consistent estimator of the limiting variance of \( \hat{R} \) can be obtained by finding a consistent estimator of \( \Sigma \).

Since \( \Sigma \) involves moments of \( X_k \) up to fourth order, one could substitute sample moments and obtain an estimator, say \( \hat{\Sigma} \). The estimator \( \hat{\Sigma} \) however, is numerically too unstable, and can lead to negative variance estimates. Even for the case when \( L = 2 \), \( \hat{\Sigma} \) has occasional negative eigenvalues. This problem worsens as \( L \) is increased. When \( L = 9 \), there were no observed simulation runs in which \( \hat{\Sigma} \) yields positive variance estimates. The lag value \( L = 9 \) is quite small in comparison to the optimal lag of \( L_R = 21 \) for \( \delta = .025 \) or \( L_R = 112 \) for \( \delta = .005 \). The moment estimator \( \hat{\Sigma} \) is unsuitable for this particular variance estimation problem.

Another procedure involves exploiting the mixing dependence of the sequence \( \{X_k\} \). Recall from Theorem 3.1

\[
\Sigma = \lim_{n \to \infty} n \text{cov} \left( X - \mu, \hat{c}(1) - c(1; \theta_0), \ldots, \hat{c}(L) - c(L; \theta_0) \right).
\]

Given \( N_0 \) independent samples of size \( n_1 \) of the process \( \{X_k\} \), one can calculate the value of the \( L + 1 \) vector

\[
\left( X, \hat{c}(1), \ldots, \hat{c}(L) \right)
\]

for each sample, and then calculate the empirical variance of these \( L + 1 \) vectors. The empirical variance will then be an estimate of \( n_1^{-1} \Sigma \).

For the S FCS model of Section 3.2, data segments (or blocks) are nearly independent once they are separated by more than twice the effective radius \( m \) of the Gaussian kernel. Here \( m = \left[ 6 \delta / \sqrt{2} \right] - 1 \).
However, additional information is available. If the method above is used starting at $X_0$ with the appropriate independent subsamples (blocks), an estimate of $n^{-1}_1 \Sigma$ is obtained. However one could also start at $X_{n_0}$ and then consider approximately independent subsamples. This offset can then be used to estimate $n^{-1}_1 \Sigma$. The covariance estimates from several offsets can then be averaged, giving a better estimate of $n^{-1}_1 \Sigma$. When multiplied by $n_1$, this yields an estimator of $\Sigma$. By construction this estimator is nonnegative definite. This variance estimation algorithm can be more formally described as follows:

(a) Divide the data $X_j$, $j = 0, \ldots, n$ into segments of lengths $n_1 + 2m$, starting at $j = j_0$.

(b) With $j_0 = 0$, obtain the sample variance estimator $\hat{\Sigma}^{(j_0)}$.

(c) Repeat steps 1 and 2 at offsets or starting values $j_0$ as multiples of $n/50$.

(d) Average $\hat{\Sigma}^{(j_0)}$ over the offsets, yielding the estimator $\hat{\Sigma}$.

The results are summarized in Table 3.5 3.7. To simplify the presentation, Table 3.5 gives the ratio of the estimate of the standard deviation of $\hat{\theta}_1$ divided by the true standard deviation, as estimated by the Monte Carlo results. Some quantiles of these ratios are given. Tables 3.6 and 3.7 give the same results for the statistics $\hat{\theta}_2$ and $\hat{R}$. As suggested by the previous section, extreme values of $\hat{\theta}_2$ are trimmed.

The variance estimator for $\hat{\theta}_1$ is median biased high, while for $\hat{\theta}_2$ it is median biased low. The variance estimator for $\hat{R}$ is generally in the right range. The estimates of the variance of $\hat{\theta}_1$ appear to be the most unreliable.
The estimated variance of $\hat{\theta}_2$ is generally too small. Since $\theta_2$ is a useful benchmark for trimming extreme $R$'s, this suggests that one trim at five estimated standard deviations instead of three.

The estimated standard deviations of $\hat{R}$ are generally quite good. Eighty percent of the estimates range from one third to 4 times their true values. In particular, this demonstrates that on a single experimental run, one can produce a confidence interval for $R$.

The block variance estimator described above is related to, though not directly based upon the estimators described in [9], [18] and [26]. In particular, [9] considers a variance estimation algorithm based upon nonoverlapping subseries of a mixing stationary process, although the method of using shifted or offset blocks is not discussed.

**3.10 S–FCS Data**

Ten S–FCS data sets from Petersen's lab are analyzed. These provide some strong evidence that the model assumptions are consistent with the data.

Figure 3.10 gives plots of experiment e5c330.scs. Plot (a) shows a moderately noisy data set $X_t$ plotted against position $t$. Plot (b) shows that the raw empirical lag covariance (cov.dat) plot is consistent with a signal plus independent additive noise. Plots (b) and (c) also show that the covariance function is of the form $c(k; \theta)$ as predicted by the model, indicating that the laser beam is well represented by a Gaussian surface and the aggregates follow a marked Poisson process. In plot (c), the y axis is cov.nils, representing the model based covariance function for lags
Table 3.5: Standard deviation estimator for $\theta_1$ divided by true standard deviation.
$\lambda = 1, \theta_2 = .235$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta_1$</th>
<th>StDev</th>
<th>$q(.05)$</th>
<th>$q(.25)$</th>
<th>$q(.5)$</th>
<th>$q(.75)$</th>
<th>$q(.95)$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.050</td>
<td>22650</td>
<td>10056.32</td>
<td>0.59</td>
<td>0.88</td>
<td>1.08</td>
<td>1.57</td>
<td>2.23</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>21347.75</td>
<td>0.62</td>
<td>0.84</td>
<td>1.16</td>
<td>1.50</td>
<td>2.19</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>39962.32</td>
<td>0.64</td>
<td>0.89</td>
<td>1.16</td>
<td>1.71</td>
<td>2.18</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>10070.58</td>
<td>0.62</td>
<td>0.89</td>
<td>1.19</td>
<td>1.48</td>
<td>1.98</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>21378.13</td>
<td>0.55</td>
<td>0.84</td>
<td>1.08</td>
<td>1.38</td>
<td>2.31</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>40019.33</td>
<td>0.63</td>
<td>0.92</td>
<td>1.19</td>
<td>1.57</td>
<td>2.13</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>10208.54</td>
<td>0.57</td>
<td>0.86</td>
<td>1.18</td>
<td>1.64</td>
<td>2.14</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>21672.18</td>
<td>0.62</td>
<td>0.91</td>
<td>1.16</td>
<td>1.55</td>
<td>2.45</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>40571.10</td>
<td>0.60</td>
<td>0.83</td>
<td>1.13</td>
<td>1.53</td>
<td>2.60</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>22650</td>
<td>10621.20</td>
<td>0.67</td>
<td>0.90</td>
<td>1.17</td>
<td>1.65</td>
<td>2.38</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>22551.66</td>
<td>0.53</td>
<td>0.85</td>
<td>1.11</td>
<td>1.39</td>
<td>2.02</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>42221.28</td>
<td>0.58</td>
<td>0.90</td>
<td>1.23</td>
<td>1.52</td>
<td>2.23</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>10342.60</td>
<td>0.63</td>
<td>0.85</td>
<td>1.14</td>
<td>1.50</td>
<td>2.36</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>21957.91</td>
<td>0.52</td>
<td>0.82</td>
<td>1.07</td>
<td>1.41</td>
<td>2.20</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>41107.23</td>
<td>0.58</td>
<td>0.85</td>
<td>1.10</td>
<td>1.42</td>
<td>2.08</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>10262.29</td>
<td>0.63</td>
<td>0.87</td>
<td>1.26</td>
<td>1.56</td>
<td>2.74</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>21786.76</td>
<td>0.62</td>
<td>0.84</td>
<td>1.03</td>
<td>1.46</td>
<td>2.38</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>40786.09</td>
<td>0.61</td>
<td>0.82</td>
<td>1.14</td>
<td>1.54</td>
<td>2.14</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Table 3.6: Standard deviation estimator for $\theta_2$ divided by true standard deviation. $
abla = 1, \ \theta_2 = .235$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta_1$</th>
<th>StDev</th>
<th>$q(.05)$</th>
<th>$q(.25)$</th>
<th>$q(.5)$</th>
<th>$q(.75)$</th>
<th>$q(.95)$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.050</td>
<td>22650</td>
<td>0.0369</td>
<td>0.56</td>
<td>0.71</td>
<td>0.89</td>
<td>1.19</td>
<td>1.89</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>0.0369</td>
<td>0.57</td>
<td>0.72</td>
<td>0.86</td>
<td>1.03</td>
<td>1.59</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>0.0369</td>
<td>0.45</td>
<td>0.66</td>
<td>0.85</td>
<td>1.11</td>
<td>1.39</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>0.0372</td>
<td>0.45</td>
<td>0.71</td>
<td>0.84</td>
<td>1.06</td>
<td>1.51</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>0.0372</td>
<td>0.49</td>
<td>0.69</td>
<td>0.86</td>
<td>1.15</td>
<td>1.77</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>0.0372</td>
<td>0.47</td>
<td>0.72</td>
<td>0.90</td>
<td>1.15</td>
<td>1.66</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>0.0381</td>
<td>0.49</td>
<td>0.72</td>
<td>0.89</td>
<td>1.05</td>
<td>1.46</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>0.0381</td>
<td>0.47</td>
<td>0.70</td>
<td>0.85</td>
<td>1.08</td>
<td>1.59</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>0.0381</td>
<td>0.51</td>
<td>0.72</td>
<td>0.94</td>
<td>1.27</td>
<td>1.77</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>22650</td>
<td>0.0398</td>
<td>0.50</td>
<td>0.69</td>
<td>0.85</td>
<td>0.99</td>
<td>1.30</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>0.0398</td>
<td>0.50</td>
<td>0.65</td>
<td>0.83</td>
<td>1.03</td>
<td>1.34</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>0.0398</td>
<td>0.48</td>
<td>0.68</td>
<td>0.84</td>
<td>1.06</td>
<td>1.53</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>0.0386</td>
<td>0.43</td>
<td>0.61</td>
<td>0.78</td>
<td>1.00</td>
<td>1.50</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>0.0386</td>
<td>0.47</td>
<td>0.65</td>
<td>0.78</td>
<td>0.98</td>
<td>1.53</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>0.0386</td>
<td>0.37</td>
<td>0.63</td>
<td>0.78</td>
<td>0.99</td>
<td>1.52</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>0.0384</td>
<td>0.56</td>
<td>0.76</td>
<td>0.90</td>
<td>1.12</td>
<td>1.48</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>0.0384</td>
<td>0.53</td>
<td>0.70</td>
<td>0.90</td>
<td>1.14</td>
<td>1.61</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>0.0384</td>
<td>0.36</td>
<td>0.63</td>
<td>0.85</td>
<td>1.05</td>
<td>1.77</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Table 3.7: Standard deviation estimator for $R$ divided by true standard deviation. 
$\lambda = 1, \theta_2 = .235$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta_1$</th>
<th>StDev</th>
<th>$q(.05)$</th>
<th>$q(.25)$</th>
<th>$q(.5)$</th>
<th>$q(.75)$</th>
<th>$q(.95)$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.050</td>
<td>22650</td>
<td>0.852</td>
<td>0.44</td>
<td>0.70</td>
<td>1.06</td>
<td>1.36</td>
<td>2.28</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>0.849</td>
<td>0.36</td>
<td>0.68</td>
<td>0.85</td>
<td>1.17</td>
<td>1.78</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>0.848</td>
<td>0.46</td>
<td>0.66</td>
<td>0.85</td>
<td>1.26</td>
<td>3.79</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>0.852</td>
<td>0.50</td>
<td>0.81</td>
<td>1.03</td>
<td>1.30</td>
<td>1.99</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>0.850</td>
<td>0.48</td>
<td>0.68</td>
<td>0.98</td>
<td>1.36</td>
<td>4.71</td>
<td>0.0</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>0.848</td>
<td>0.45</td>
<td>0.69</td>
<td>0.91</td>
<td>1.29</td>
<td>2.43</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>0.852</td>
<td>0.46</td>
<td>0.70</td>
<td>0.92</td>
<td>1.24</td>
<td>2.12</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>0.850</td>
<td>0.56</td>
<td>0.76</td>
<td>0.98</td>
<td>1.31</td>
<td>3.25</td>
<td>0.0</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>0.849</td>
<td>0.41</td>
<td>0.74</td>
<td>0.96</td>
<td>1.45</td>
<td>2.71</td>
<td>0.0</td>
</tr>
<tr>
<td>0.050</td>
<td>22650</td>
<td>0.927</td>
<td>0.44</td>
<td>0.71</td>
<td>0.99</td>
<td>1.29</td>
<td>2.71</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>48180</td>
<td>0.925</td>
<td>0.43</td>
<td>0.67</td>
<td>0.91</td>
<td>1.32</td>
<td>3.97</td>
<td>0.5</td>
</tr>
<tr>
<td>0.050</td>
<td>90300</td>
<td>0.923</td>
<td>0.44</td>
<td>0.66</td>
<td>0.94</td>
<td>1.31</td>
<td>2.91</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>22650</td>
<td>0.890</td>
<td>0.53</td>
<td>0.72</td>
<td>0.97</td>
<td>1.27</td>
<td>2.42</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>48180</td>
<td>0.888</td>
<td>0.44</td>
<td>0.69</td>
<td>0.90</td>
<td>1.19</td>
<td>2.03</td>
<td>0.5</td>
</tr>
<tr>
<td>0.025</td>
<td>90300</td>
<td>0.887</td>
<td>0.44</td>
<td>0.68</td>
<td>0.91</td>
<td>1.42</td>
<td>3.31</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>22650</td>
<td>0.860</td>
<td>0.47</td>
<td>0.81</td>
<td>1.03</td>
<td>1.37</td>
<td>2.24</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>48180</td>
<td>0.858</td>
<td>0.45</td>
<td>0.71</td>
<td>0.98</td>
<td>1.32</td>
<td>2.74</td>
<td>0.5</td>
</tr>
<tr>
<td>0.005</td>
<td>90300</td>
<td>0.856</td>
<td>0.46</td>
<td>0.69</td>
<td>0.91</td>
<td>1.23</td>
<td>2.53</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Figure 3.10.4: Petersen Data e5c330. $\hat{R} = .0125$
greater than zero. The empirical covariance is also overlaid on this plot. Plot (d) is a normal QQ plot, indicating that the noise is approximately Gaussian, as in our simulation study. The other experiments show that the raw data normal QQ plots are closer to straight lines with higher noise levels, giving further credibility to the assumption of additive Gaussian noise.

For these experiments, \( \delta = .023 \) and \( \theta_2 = w = .77 \). The analytic program for determining the optimal number of lags to be used was run with the appropriate values for \( w \), and with the appropriate noise levels. The noise level is estimated as described below, suggesting that \( \beta \) ranges from .5 to 7.5. The optimal number of lags needed for \( \theta_2 \) varies with the noise levels, ranging from 3 with no noise to 42 with the high noise level. The optimal number of lags for \( R \) is very stable near 78 over these noise levels. In addition, the analytic variance of \( \hat{R} \) does not vary substantially over lags 40 to 100.

The parameters are estimated by a grid search. A large initial grid region is searched to get into the right region, and then a finer grid search is done. These give a starting value for a Newton's iteration. The variances of the estimators are estimated using the block method as described in Section 3.9.2. Since a larger lag of 78 is used, the block sizes are suitably adjusted, giving 10 segments of length 100 each, and then using data samples of length 80 in a block. The optimal number of lags is computed using an analytic variance computation as done in Section 3.8.

Table 3.8 summarizes the experimental information, using the optimal 78 lags, for \( R \). The table gives the experiment number, the noise level, the 99% confidence interval for \( \theta_2 = w \), and the 95% confidence interval for \( R \). The noise terminology
Table 3.8: Analysis of Petersen’s Data, at lag = 78

<table>
<thead>
<tr>
<th>file</th>
<th>noise</th>
<th>$w$ 99.9% ci</th>
<th>$R$ 95% ci</th>
</tr>
</thead>
<tbody>
<tr>
<td>e5c330.scs</td>
<td>1.98</td>
<td>0.37</td>
<td>0.01029</td>
</tr>
<tr>
<td>e5c331.scs</td>
<td>1.53</td>
<td>0.71</td>
<td>0.01111</td>
</tr>
<tr>
<td>e5c332.scs</td>
<td>3.56</td>
<td>1.34</td>
<td>0.00618</td>
</tr>
<tr>
<td>e5c333.scs</td>
<td>1.51</td>
<td>0.87</td>
<td>0.02618</td>
</tr>
<tr>
<td>e5c334.scs</td>
<td>1.53</td>
<td>0.47</td>
<td>0.01357</td>
</tr>
<tr>
<td>e5c335.scs</td>
<td>2.16</td>
<td>1.08</td>
<td>0.00949</td>
</tr>
<tr>
<td>e5c336.scs</td>
<td>8.65</td>
<td>0.54</td>
<td>0.002</td>
</tr>
<tr>
<td>e5c337.scs</td>
<td>1.84</td>
<td>0.58</td>
<td>0.00882</td>
</tr>
<tr>
<td>e5c338.scs</td>
<td>4.13</td>
<td>0.89</td>
<td>0.00338</td>
</tr>
<tr>
<td>e5c339.scs</td>
<td>1.88</td>
<td>1.5</td>
<td>0.02291</td>
</tr>
</tbody>
</table>

is the same as described in Section 3.8. Here the noise level is defined as $(c(0) + \gamma^2)/c(0)$ and is estimated by $\hat{Var}(X_k)/c(0; \hat{\theta})$. This is an estimate of $\beta + 1$, where $\beta$ is the noise parameter used in the simulation study.

The results show that some of the data sets are consistent in the sense of agreeing with the known value of $w$, but some are not. Experiments e5c332.scs and e5c339.scs are not consistent with the known $w$. It also shows that the value of $R$ differs in some of the experimental runs. For example, experiments e5c330.scs and e5c331.scs have the same $R$ while experiment e5c336.scs has a different value of $R$.

In practice 50 lags have been used [32, p 57]. This is somewhat smaller than we use. If the data is analyzed using 50 lags, the point estimates do not change much (only a few percent), but the estimates of the standard deviations, and hence lengths of the confidence intervals change. Note that only point estimates are currently obtained in practice. Variance estimates have not been previously
Table 3.9: e5c332 Estimates of $\theta_2$ and standard errors

<table>
<thead>
<tr>
<th>lag</th>
<th>$\hat{\theta}_2$</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1.905</td>
<td>17.1</td>
</tr>
<tr>
<td>8</td>
<td>0.616</td>
<td>.444</td>
</tr>
<tr>
<td>20</td>
<td>0.787</td>
<td>.183</td>
</tr>
<tr>
<td>30</td>
<td>1.068</td>
<td>.188</td>
</tr>
<tr>
<td>50</td>
<td>1.644</td>
<td>.256</td>
</tr>
<tr>
<td>70</td>
<td>1.522</td>
<td>.082</td>
</tr>
<tr>
<td>78</td>
<td>1.488</td>
<td>.058</td>
</tr>
</tbody>
</table>

available. For the $\theta_2$ confidence intervals, the resulting lengths are 1.06 - 4.11 times the lengths when 78 lags are used. As indicated earlier, the estimation of $R$ is a more stable procedure, and the resulting confidence intervals are .96 - 1.18 times as large as the lag 78 confidence intervals. It is recommended that 78 lags, or the optimal number of lags for $R$ be used.

The estimation procedure is somewhat sensitive to the number of lags $L$ used in the least squares equation (3.5). Using other smaller lag values can give strange results. Particularly in the noisy cases, using too few lags can result in unstable estimates, as illustrated in Table 3.9. In addition it was learned from the simulation study that the standard error of $\hat{\theta}_2$ is quite stable and rarely exceeds .2. With too few lags we see unusually large estimated standard errors. The model based estimation is useful for determining the number of lags to use.

3.11 Discussion and Summary

This chapter presents a filtered Poisson process model of the biophysical chemistry imaging technique known as Scanning Fluorescence Correlation Spectroscopy.
Based on the observed fluorescence intensity process \( \{ X_k \} \), the parameter \( R \) can be identified and consistently estimated. \( R \) contains direct information about average protein cluster sizes and surface densities. The measurement of these quantities is the primary focus of \( S \)-FCS. The estimator \( \hat{R} \) is asymptotically normal. Based on the model, an optimal least squares lag parameter \( L \) can be chosen. The lag \( L \) is optimal in the sense of minimizing the limiting variance of \( \hat{R} \). There was previously no formal method for determining this optimal lag value. A method for obtaining consistent estimates of the limiting variance of \( \hat{R} \) is also presented. It exploits the finite-range dependence of the data sequence and should be applicable in settings other than that considered here. The variance estimation allows the construction of approximate confidence intervals for \( R \) based on a single set of data. Prior to this work, only point estimates of \( R \) were obtained from actual \( S \)-FCS experiments. In the current practice, a desired level of accuracy for these point estimates is achieved only by taking several measurements on many cells. The proposed confidence interval approach provides considerable time savings and efficiency as \( S \)-FCS requires approximately 2-4 min/scan on each cell, plus several hours of prescanning cell preparation [32].

The Monte Carlo simulation study indicates that the estimation procedures are stable over a wide range of parameter values, including those values which are encountered in practice. The Gaussian kernel parameter \( w \) determines the width of the laser beam. A procedure exists for measuring \( w \). However, its value can also be estimated from \( S \)-FCS data. A method for removing extreme outliers of \( \hat{R} \) can be based on a comparison of the estimated value and the measured value of \( w \).
The statistical methodology is applied to several S FCS data sets. The empirical covariances are consistent with the theoretical covariances predicted by the filtered Poisson process and additive Gaussian noise assumptions. The observed noise level is also estimated.
Chapter 4

Statistical Inference for Image Correlation Spectroscopy

4.1 Introduction

S-FCS has been an excellent experimental tool for measuring protein aggregation. However, the measurements tend to be slow requiring 2-4 min/scan per cell and, while the measurements tend to be quite precise, a desired level of accuracy is achieved only after multiple measurements on many different cells. Also, only a few laboratories around the world have been able to make these measurements, as S-FCS requires expensive, specialized equipment and several hours of prescanning preparation.

This chapter is concerned with an extension of S-FCS to the so-called Image Correlation Spectroscopy (ICS), a technique which is based on a quantitative analysis of confocal scanning laser microscopy (CSLM) images [25]. This technique addresses a number of the limitations of S-FCS.

This chapter presents a filtered Poisson process model for the data obtained from an ICS experiment and a subsequent method of statistical analysis for these
data. The model is of the same form as that of S FCS. The resulting data or image is a realization of a finite-range dependent stationary random field (or spatial series) \( X(t), t \in Z^2 \). A white noise process \( \epsilon(t) \) models measurement recording error due to background noise and is added to the fluorescence intensity measurement.

The Gaussian weight function form is quite general in this and related settings. It is a consequence of optics. Any image viewed through a small lens or viewing hole produces an image filtered according to a sine curve, which is well approximated by a scaled Gaussian kernel [32].

The identifiable parameters are the same as in the S FCS case. From a statistical point of view, the problems deal with the implementation of a tractable estimation method for a non-Gaussian random field.

The first parameter \( \theta_1 \) is the product of the Poisson process intensity \( \lambda \) and the second moment of the marring distribution and hence is a function of cluster size, while \( \theta_2 \) is the laser beam width parameter. See Section 4.3 for a more formal definition of \( (\theta_1, \theta_2) \).

Large-sample interval estimators of these parameters are obtained, where only point estimators were previously possible. Central to this goal is the asymptotic normality of the estimators and the consistent estimation of their limiting variances.

A nonlinear least squares equation involving the periodogram \( I^{(t)}(\cdot) \) and spectral density \( f(\cdot; \theta) \) of the random field \( X(t) \) is used to estimate \( \theta_1 \) and \( \theta_2 \). These estimates are shown to be asymptotically normal by expressing them, through the delta method, in terms of smoothed periodograms and then applying a result
Figure 4.1.1: Confocal microscope image and resulting fluorescence intensity image as a function of pixel position
of Rosenblatt [27]. The estimator of the ratio \( \hat{R} \) and its asymptotic normality subsequently follow.

The parameters may be alternatively estimated through a quasi-likelihood function of \( \theta_1 \) and \( \theta_2 \) as the periodograms \( I^{(T)}(\cdot) \) are asymptotically independent exponential random variables with mean \( f(\cdot ; \theta) \) for distinct frequencies. The same method of proof then applies to show these estimators are asymptotically normal.

In both the least squares and likelihood cases, the limiting variances of the normal distributions involve multi-dimensional integrals of the square of the spectral density and the fourth-order cumulant spectral density. These are evaluated at fixed parameter settings for both estimation methodologies and compared. In order to obtain confidence intervals, these limiting variances must be estimated and, using the least squares method, two different estimation techniques are implemented. The first method exploits the finite-range dependence of the process and involves splitting the data into approximately independent blocks as in Section 3.9.2. On each block, smoothed periodograms are calculated resulting in an approximately independent sample from which sample variances are obtained. When scaled appropriately, these provide consistent variance estimators. The second method was first discussed in Taniguchi [33], [34] and later generalized in Keenan [14] although not numerically implemented in either case. It involves essentially replacing spectral densities by periodograms (both second and fourth order) and integrals by their corresponding Riemann sums.

The calculations associated with the estimation of the ICS model parameters and variances are difficult and computationally intensive. There are more than
$O(T^6)$ computations, where the sample size is $T^2$. As such, they are performed on a MasPar MP-2 massively parallel computer consisting of 2048 separate RISC processors on an array connected to a DEC workstation. A complete description of this machine's architecture is included in Chapter 5. A simulation study for the least squares method is carried out and both the block and Taniguchi-Keenan variance estimation methods are implemented. Complete details are provided in Chapter 5 which discusses these computations.

### 4.2 A Mathematical Model for ICS

The spatial distribution of protein clusters (aggregates) is modeled by a marked Poisson process. A random point occurrence $(x, y)$ denotes the aggregate location. Suppose:

i) $\{N(A) : A \subseteq \mathbb{R}^2\}$ is a homogeneous Poisson process with intensity $\lambda$.

ii) $\{Y_{(x_i,y_i)}\}$ is a sequence of mutually independent, identically distributed random variables that are also independent of $N(\cdot)$, and represents the cluster size at $(x_i, y_i)$, a point event of $N(\cdot)$.

Then from Chapter 2, a marked Poisson process, with positions in $\mathbb{R}^2$ and marks in $\mathbb{Z}_0^+$, is a Poisson process on $\mathbb{R}^2 \times \mathbb{Z}_0^+$ such that the marginal process of locations

$$\{N(A \times \mathbb{Z}_0^+) : A \in \mathcal{B}_{\mathbb{R}^2}\}$$

is itself a Poisson process. This marginal process is denoted by $N_{\mathbb{R}^2}(\cdot)$. Also, $\mathcal{B}_{\mathbb{R}^2}$ is the Borel $\sigma$-field on $\mathbb{R}^2$ and $\mathbb{Z}_0^+$ is the set of nonnegative integers.
With this construction, the total protein aggregate mass for a region \( A \in \mathbb{B}_{2} \) is
\[
\int_{A} Y_{(x,y)} N(dx\,dy) \equiv M(A).
\]
This random integral reduces to the finite sum
\[
M(A) = \sum_{i=1}^{N(A)} Y_{(x_{i}, y_{i})}
\]
as \( N(\cdot) \) is a random counting measure and puts mass only on the random point locations \((x_{i}, y_{i})\). These assumptions imply that \( M(\cdot) \) is a compound Poisson process, sometimes termed a mark-accumulator process [28].

The intensity of the laser beam in the confocal scanning laser microscope is modeled by a Gaussian kernel (or weight function)
\[
f_{0}(x, y) = e^{-\frac{1}{w^{2}}(x^{2} + y^{2})}, \quad w > 0.
\]
The effective radius of the beam is taken to be \( \frac{3w}{\sqrt{2}} \) so that when the laser is positioned at \((t_{1} \delta, t_{2} \delta)\), \((t_{1}, t_{2}) \in \mathbb{B}^{2}\), the region effectively illuminated is
\[
S_{(t_{1}, t_{2})} : (x - t_{1} \delta)^{2} + (y - t_{2} \delta)^{2} \leq \left( \frac{3w}{\sqrt{2}} \right)^{2}.
\]
The known quantity \( \delta \) is a stepsize or resolution parameter determined by the total area scanned divided by the total number of data points obtained in an ICS experiment. Let \( t = (t_{1}, t_{2}) \). The laser beam intensity weight function at location \( t \delta \) is thus
\[
f_{g}(x, y) = f_{0}(x - t_{1} \delta, y - t_{2} \delta).
\]
From Chapter 2, the fluorescence intensity $X(t)$ at location $t \delta$ is formally defined as the filtered Poisson process:

$$ X(t) = \int_{S_\delta} f_\delta(x,y) Y(x,y) N(dx\,dy) + \gamma \epsilon(t) $$

$$ \equiv Z(t) + \gamma \epsilon(t) \, . $$

(4.1)

where $\epsilon(t)$ is a white noise process and $\gamma$ is an error or noise parameter. The term $\gamma \epsilon(t)$ represents experimental recording error due to background or photoncounting noise in the microscope image. In general, a filtered Poisson process is of the form

$$ Z(t) = \int_A h(t, \tau; u(\tau)) N(dt) $$

where $\tau$ and $u(\tau)$ represent point occurrences and associated marks [28]. For the fluorescence intensity process, $h(\cdot)$ takes the form

$$ h(t, (x,y); Y(x,y)) = f_\delta(x,y) Y(x,y) \, . $$

The fluorescence intensity $X(t)$ is defined in terms of a random counting integral. As before, the integral reduces to a sum since $N(\cdot)$ is a random counting measure. Thus,

$$ Z(t) = \sum_{i=1}^{N(S_\delta)} f_\delta(x_i,y_i) Y(x_i,y_i) \, . $$

The process $\{X(t)\}$ is a stationary finite range dependent real valued random field. This follows from the independent increments property of the Poisson process $N$ as the disc regions $S_0$ and $S_\delta$ are disjoint for sufficiently large $||t||$. 
4.3 Model Parameter Estimation

The fluorescence intensity data is modeled as

\[
X(t) = \int_{s_k} f_k(x,y)Y(x,y)N(dx\,dy) + \gamma e(t), \quad t \in \mathbb{Z}^2
\]

\[
\equiv Z(t) + \gamma e(t) .
\]

Certain model parameters of interest in ICS appear in the autocovariances of the process. Let \( c(k; \theta) \) and \( c_X(k; \theta) \) denote the autocovariance functions of \( Z(t) \) and \( X(t) \) respectively. Then, it can be shown that

\[
c(k; \theta) = \lambda \mu_{2Y} \int_{s_0 \cap s_k} f_0(x,y)f_k(x,y)dx\,dy
\]

where \( f_k(x,y) = e^{-\frac{1}{2}(y-k_2)^2} \) and \( \mu_{2Y} \) is the second moment of the marking distribution \( P(Y = y_i) \). The two parameters of initial interest are \( \theta_1 = \lambda \mu_{2Y} \) and \( \theta_2 = w \). Notice that \( \theta_1 \) is a function of protein aggregate density and size whereas \( \theta_2 \) is the laser beam width parameter. From (4.2), one obtains the autocovariance function for \( X(t) \) as,

\[
c_X(k; \theta) = \begin{cases} 
  c(0; \theta) + \gamma^2 & \text{if } k = 0 \\
  c(k; \theta) & \text{otherwise} .
\end{cases}
\]

The parameter of primary interest in ICS is \( R = c(0; \theta)/\mu^2 \), the variance of the nonnoise process \( \{Z(t)\} \) divided by the square of the mean, and for the same reason as in S FCS.

Again as in the S FCS case, the estimator of main interest is

\[
\hat{R} = c(0; \hat{\theta})/X^2
\]
where $X$ is the sample mean of the process and estimates $\mu$. Two methods of estimating $\theta$ are discussed in the following sections, least squares and quasi-likelihood. Both methods involve the spectral density $f(\lambda; \theta)$ and periodogram of the process $I^{(T)}(\lambda)$ where the explicit dependence of the spectral density upon $\theta$ (via $c_X(k; \theta)$) is denoted.

### 4.4 Least Squares Estimation

The observed fluorescence intensity data is modeled by the stationary random field $X(t)$ where $t \in \{0, \ldots, T-1\}^2$. Let $d^{(T)}(\lambda) = \sum_{t} X(t)e^{-\imath \lambda \cdot t}$ be the finite Fourier transform of $\{X(t)\}$ and $I^{(T)}(\lambda) = \frac{1}{(2\pi T)^2} |d^{(T)}(\lambda)|^2$ the associated periodogram. The spectral density of $X(t)$ is given by,

$$f(\lambda; \theta) = \frac{1}{(2\pi)^2} \sum_{k} c(k; \theta)e^{-\imath \lambda \cdot k}.$$ 

The parameter $\theta$ may be estimated by the nonlinear least squares function,

$$L^{(T)}(\theta) = \left( \frac{2\pi}{T} \right)^2 \sum_{\mathbf{s}} (I^{(T)}(\frac{2\pi}{T} \mathbf{s}) - f(\frac{2\pi}{T} \mathbf{s}; \theta))^2 \tag{4.6}$$

where $\mathbf{s} \in \{1, \ldots, T-1\}^2 = S_T$ say. Note that (4.6) is simply the Riemann sum approximation to

$$\int_{0}^{2\pi} (I^{(T)}(\lambda) - f(\lambda; \theta))^2 d\lambda.$$ 

An application of Taylor’s Theorem (via the delta method) gives,

$$\sqrt{T} \left[ \begin{array}{c} \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{array} \right] = 2 \left[ \begin{array}{cc} \frac{\partial^2}{\partial \theta_1^2} L^{(T)}(\theta^*) \\ \frac{\partial^2}{\partial \theta_2 \partial \theta_1} L^{(T)}(\theta^*) \\ \frac{\partial^2}{\partial \theta_2^2} L^{(T)}(\theta^*) \end{array} \right]^{-1} \left[ -\frac{T}{2} \frac{\partial}{\partial \theta_1} I^{(T)}(\theta_0) \\ -\frac{T}{2} \frac{\partial}{\partial \theta_2} I^{(T)}(\theta_0) \right]$$
where $\theta_0$ is the true value of the parameter and $\theta^*$ lies on the line segment joining $\hat{\theta}$ and $\theta_0$. Specifically,

$$-\frac{T}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0) = (\frac{2\pi}{T})^2 \sum_s (f^{(T)}(\frac{2\pi}{T}; s) - f(\frac{2\pi}{T}; \theta_0)) \frac{\partial}{\partial \theta_j} f(\frac{2\pi}{T}; \theta_0), \quad j = 1, 2,$$

where $s \in S_T$. This is a linear combination of periodograms as discussed in Brillinger [3] and Rosenblatt [27]. They have shown that under certain conditions, these smoothed periodograms have limiting normal distributions. This fact will form the basis of the asymptotic results that follow.

**Theorem 4.1** Suppose $X$ is the sample mean of the data $X(t)$, $t \in \{0, \ldots, T-1\}^2$ and $-\frac{1}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0) = (\frac{2\pi}{T})^2 \sum_s (f^{(T)}(\frac{2\pi}{T}; s) - f(\frac{2\pi}{T}; \theta_0)) \frac{\partial}{\partial \theta_j} f(\frac{2\pi}{T}; \theta_0), \quad j = 1, 2$, is the smoothed periodogram. Then,

$$\sqrt{T^2} \left[ \begin{array}{c} \bar{X} - \mu_0 \\ -\frac{1}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0) \\ -\frac{1}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0) \end{array} \right] \overset{D}{\rightarrow} N_3(\mu, \Sigma).$$

The limiting covariance structure $\Sigma = [\Sigma_{i,j}]$, $i, j = 0, \ldots, 2$, is given by,

$$
\Sigma_{0,0} = \lim_{T \to \infty} T^2 \text{Var}(\bar{X}) = \gamma^2 + \sum_k c(k; \theta_0),
$$

$$
\Sigma_{0,j} = \lim_{T \to \infty} T^2 \text{cov}(\bar{X}, -\frac{1}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0))
= (2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f_3(0, \alpha) d\alpha,
$$

and

$$
\Sigma_{j,k} = \lim_{T \to \infty} T^2 \text{cov}(-\frac{1}{2} \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0), -\frac{1}{2} \frac{\partial}{\partial \theta_k} L^{(T)}(\theta_0))
= 2(2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) f^2(\alpha; \theta_0) d\alpha
+ (2\pi)^2 \int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_k} f(\beta; \theta_0) f_4(\alpha, \beta, -\alpha) d\alpha d\beta.
$$
The limiting covariances in the above result contain integrals of the third- and fourth-order cumulant spectral densities. The $k$th-order cumulant spectral density is defined as

$$f_k(\lambda^{(1)}, \ldots, \lambda^{(k-1)}) = \frac{1}{(2\pi)^{2(k-1)}} \sum_{\mathbf{u}^{(1)}} \cdots \sum_{\mathbf{u}^{(k-1)}} c_k(\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k-1)}) e^{-i \sum_{j=1}^{k-1} \lambda^{(j)} \cdot u^{(j)}}.$$ 

Theorem 4.1 is obtained by applying a theorem of Rosenblatt [27] (see Section 4.6). The following results follow from a modification of our least squares function where $f(\cdot; \theta)$ has been replaced by $E(I^T(I^T/s))$ (see Remark 4.1).

Theorem 4.2, as well as being of independent interest, will be required later. It states that $\hat{\theta}$ is a consistent estimator of $\theta_0$. The result follows from the general consistency conditions outlined in Brillinger [4] as discussed in Kulperger [16].

**Theorem 4.2** $\hat{\theta} \overset{p}{\to} \theta_0$

Theorems 4.1 and 4.2 together with applications of the delta method and Slutsky's Theorem yield the following result.

**Theorem 4.3**

$$\sqrt{T} \begin{bmatrix} \hat{X} - \mu_0 \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} \overset{D}{\to} N_3(\mathbf{0}, \Sigma^*)$$

where $\Sigma^* = B\Sigma B'$, $B = \begin{bmatrix} 1 & 0 \\ 0 & A^{-1} \end{bmatrix}$ and the elements of $A = \{A_{k,j}\}$, $k, j = 1, 2$ are of the form $A_{k,j} = \int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) d\alpha$. The matrix $\Sigma$ is given in Theorem 4.1.

To prove this result, one uses a Taylor's approximation to express the vector of the sample mean and least squares estimators as a linear function of the vector of
the sample mean and smoothed periodograms. The convergence of the matrix coefficient to $B$ is easily demonstrated and the sample mean, smoothed periodogram vector is discussed in Theorem 4.1. An application of Slutsky’s Theorem gives the final result.

Theorem 4.4 is the main theorem concerning the asymptotic distribution of the ratio estimator $\hat{R}$.

**Theorem 4.4**

$$\sqrt{T^2(\hat{R} - R)} = \sqrt{T^2\left(\frac{c(0; \hat{\theta})}{X^2} - \frac{c(0; \theta_0)}{\mu_0^2}\right)}$$

$$\overset{D}{\longrightarrow} N(0, c'\Sigma c)$$

where $c' = [\frac{\partial R}{\partial \mu}, \frac{\partial R}{\partial \theta_1}, \frac{\partial R}{\partial \theta_2}]_{(\mu_0, \theta_0)}$.

The theorem is proved by appealing to Theorem 4.3 and as before, applying the delta method and Slutsky’s Theorem.

### 4.5 Quasi-Maximum Likelihood Estimation

The parameter $\theta_0$ may also be estimated by a quasi-maximum likelihood function because for finite collections of distinct frequencies $\lambda$, the periodograms $I^{(T)}(\lambda)$ are asymptotically independent exponential random variables with mean $f(\lambda; \theta_0)$ ([4]). The asymptotic loglikelihood function takes the form,

$$L^{(T)}(\theta) = -\left(\frac{2\pi}{T}\right)^2 \sum_s (\log f(\frac{2\pi}{T}s; \theta) + \frac{I^{(T)}(\frac{2\pi}{T}s)}{f(\frac{2\pi}{T}s; \theta)}), \ s \in S_T.$$  

As in Section 4.4,

$$\frac{\partial}{\partial \theta_j} L^{(T)}(\theta) = \left(\frac{2\pi}{T}\right)^2 \sum_s (I^{(T)}(\frac{2\pi}{T}s) - f(\frac{2\pi}{T}s; \theta) \frac{\partial}{\partial \theta_j} \frac{I^{(T)}(\frac{2\pi}{T}s)}{f^2(\frac{2\pi}{T}s; \theta)})$$
which is a linear combination of periodograms. This allows one to express the likelihood estimator as a linear function of these derivatives or equivalently, smoothed periodograms just as in the least squares case. The asymptotic results follow from the same arguments. The analogues of the theorems of the previous section are now given.

**Theorem 4.5** Suppose $X$ is the sample mean of the data $X(t)$, $t \in \{0, \ldots, T-1\}^2$ and

$$\frac{\partial}{\partial \theta_j} L^{(T)}(\theta) = (\frac{2\pi}{T})^2 \sum_s \left( I^{(T)}(\frac{2\pi}{T} s) - f(\frac{2\pi}{T} s; \theta) \right) \left( \frac{\partial}{\partial \theta_j} f(\frac{2\pi}{T} s; \theta) \right), \quad j = i, 2,$$

is a smoothed periodogram. Then,

$$\sqrt{T^2} \begin{bmatrix} \frac{\partial}{\partial \theta_0} L^{(T)}(\theta_0) \\ \frac{\partial}{\partial \theta_2} L^{(T)}(\theta_0) \end{bmatrix} \xrightarrow{D} N_2(\zeta, \Gamma).$$

The limiting covariance structure $\Gamma = [\Gamma_{i,j}]$, $i, j = 0, \ldots, 2$, is given by,

$$\Gamma_{0,0} = \lim_{T \to \infty} T^2 \text{Var}(X) = \gamma^2 + \sum_k c(k; \theta_0),$$

$$\Gamma_{0,j} = \lim_{T \to \infty} T^2 \text{cov}(X, \frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0)) = (2\pi)^2 \int_0^{2\pi} \left( \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \right) f_4(0, \alpha) d\alpha,$$

and

$$\Gamma_{j,k} = \lim_{T \to \infty} T^2 \text{cov}(\frac{\partial}{\partial \theta_j} L^{(T)}(\theta_0), \frac{\partial}{\partial \theta_k} L^{(T)}(\theta_0))$$

$$= 2 \cdot (2\pi)^2 \int_0^{2\pi} \left( \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \right) \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{f(\alpha; \theta_0)}{f^2(\alpha; \theta_0)} d\alpha$$

$$+ (2\pi)^2 \int_0^{2\pi} \int_0^{2\pi} \left( \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \right) \left( \frac{\partial}{\partial \theta_k} f(\beta; \theta_0) \right) \frac{f(\alpha; \theta_0)}{f^2(\alpha; \theta_0)} \frac{f(\beta; \theta_0)}{f^2(\beta; \theta_0)} f_4(\alpha, \beta, -\alpha) d\alpha d\beta.$$
The covariance expressions for $\Gamma_{0,j}$ and $\Gamma_{j,k}$ may be expressed as,

$$(2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} \log f(\alpha; \theta_0)(\frac{f_3(0, \alpha)}{f(\alpha; \theta_0)})d\alpha$$

and

$$2 \cdot (2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} \log f(\alpha; \theta_0)\frac{\partial}{\partial \theta_k} \log f(\alpha; \theta_0)d\alpha$$

$$+ (2\pi)^2 \int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_j} \log f(\alpha; \theta_0)\frac{\partial}{\partial \theta_k} \log f(\beta; \theta_0)(\frac{f_4(\alpha, \beta, -\alpha)}{f(\alpha; \theta_0)f(\beta; \theta_0)})d\alpha d\beta.$$

The approximate mle $\hat{\theta}$ is also consistent.

**Theorem 4.6** $\hat{\theta} \overset{P}{\to} \theta_0$

As before, Theorems 4.5 and 4.6, the delta method and Slutsky's theorem give

The next result.

**Theorem 4.7**

$$\sqrt{T^2} \begin{bmatrix} \hat{X} - \mu_0 \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} \overset{D}{\to} N_3(0, \Gamma^*)$$

where $\Gamma^* = D\Gamma D'$, $D = \begin{bmatrix} 1 & 0 \\ 0 & C^{-1} \end{bmatrix}$ and the elements of $C = [C_{j,k}]$, $j, k = 1, 2$, are of the form $C_{j,k} = \int_0^{2\pi} \frac{\partial}{\partial \theta_j} \log f(\alpha; \theta_0)\frac{\partial}{\partial \theta_k} \log f(\alpha; \theta_0)d\alpha$. The matrix $\Gamma$ is that of **Theorem 4.5**.

The main result now follows.

**Theorem 4.8**

$$\sqrt{T^2}(\hat{R} - R) = \sqrt{T^2}(\frac{c(0; \hat{\theta})}{X^2} - \frac{c(0; \theta_0)}{\mu_0^2})$$

$$\overset{D}{\to} N(0, c^T\Gamma^*c)$$

where

$$c' = [\frac{\partial R}{\partial \mu}, \frac{\partial R}{\partial \theta_1}, \frac{\partial R}{\partial \theta_2}]_{(\mu_0, \theta_0)}.$$
4.6 Proofs

The method of proof for the theorems in the least squares and likelihood sections are identical. As such, let

\[ J^{(T)}(A_k) = \left( \frac{2\pi}{T} \right)^2 \sum_s \left[ I^{(T)} \left( \frac{2\pi}{T} s \right) - f \left( \frac{2\pi}{T} s; \theta_0 \right) \right] A_k \left( \frac{2\pi}{T} s; \theta_0 \right), \quad s \in S_\tau, \quad k = 1, 2, \]

where

\[ A_k \left( \frac{2\pi}{T} s; \theta_0 \right) = \frac{\partial}{\partial \theta_k} f \left( \frac{2\pi}{T} s; \theta_0 \right) \]

for least squares and

\[ A_k \left( \frac{2\pi}{T} s; \theta_0 \right) = \frac{\partial}{\partial \theta_k} f \left( \frac{2\pi}{T} s; \theta_0 \right) \]

in the likelihood case. To prove Theorems 4.1 and 4.5 then, one must show that

\[ \sqrt{T} \left[ \begin{array}{c} \bar{X} - \mu \\ J^{(T)}(A_1) \\ J^{(T)}(A_2) \end{array} \right] \xrightarrow{D} N, \]

and the limiting normal random variable has the specified covariance structure.

Asymptotic normality results such as this involving smoothed periodograms in \( d = 1 \) are usually proved by showing that all cumulants tend to the corresponding cumulants of a normal variate, since the normal distribution is determined by its moments [5]. In particular, the variate must possess the correct first and second-order moment structure and have all third- and higher order cumulants tend to zero. This technique is standard in the time series literature and has been used extensively by Brillinger, Rosenblatt et al. [5] to show, for example, the asymptotic normality of estimators of spectral densities and spectral measures as well as sample autocovariances. Kulperger applies the technique to asymptotic
likelihood are least squares estimators in [15] and [16]. The algebraic details associated with the computation of polynomials of random variables (Leonov and Shiryaev [19]) are complicated and are discussed further in the Proof of Theorems 4.1 and 4.5 section. The cumulant technique is not applied here but see Remark 4.1. Instead, a theorem of Rosenblatt [27] is utilized.

Before proceeding to the Proof of Theorems 4.1 and 4.5, two theorems of Rosenblatt [27] are repeated from Chapter 2. Two corresponding corollaries are also presented. Theorems 4.1 and 4.5 then follow as a direct consequence of these results.

From [27], suppose one observes a stationary random field \( X(t) \), \( E(X(t)) = 0 \), over the rectangle of lattice points \( t_i = 0, \ldots, T - 1, \ i = 1, \ldots, d \). The sample covariances are taken to be

\[
\tilde{c}(u) = \frac{1}{T^2} \sum_{t_i=0}^{T-|u_i|-1} X(t)X(t+u), \ i = 1, \ldots, d.
\]

Let

\[
c(u) = \text{cov} (X(t), X(t+u))
\]

and assume cumulants up to eighth-order exist.

Further, let

\[
c_4(a, b, d) = \text{cum}_4(X(t), X(t+a), X(t+b), X(t+d))
\]

be the fourth cumulant.

**Theorem 4.9 (Rosenblatt, 1985)** Let \( \{X(t)\} \) be a strongly mixing \( d \)-dimensional strictly stationary random field with \( E(X(t)) = 0 \) and \( E(X(t)^8) < \infty \). Assume
that the cumulants of \( \{X(t)\} \) up to eighth-order are absolutely summable. It then follows that \( T^{d/2}(\hat{c}(u) - c(u)) \) for a fixed number of lags \( u \) are jointly asymptotically normal as \( T \to \infty \) with mean 0 and covariances

\[
C_{u,v} = \sum_{a}(c(a)c(a + v - u) + c(a + v)c(a - u) + c_4(u, a, a + v)) .
\]

To obtain this theorem, take \( \delta = 2 \) in Theorem 2.4, Chapter 2 and apply it to arbitrary linear combinations of \( \hat{c}(u) - c(u) \). Then, one is dealing with a random field \( Y(t) \) defined by the partial sum

\[
\sum_{s} \alpha_s X(t)X(t + u(s))
\]

up to an error term which goes to zero. The result follows by the Cramer-Wold device.

**Corollary 4.1** Under the same conditions as Theorem 4.9,

\[
T^{d/2} \begin{bmatrix} X \\ \hat{c}(u) - c(u) \end{bmatrix}
\]

for a fixed number of lags \( u \), is jointly asymptotically normal as \( T \to \infty \) with

\[
\lim_{T \to \infty} T^d \text{Var} (X) = \sum_{u} c(u) ,
\]

\[
\lim_{T \to \infty} T^d \text{cov} (X, \hat{c}(u) - c(u)) = \sum_{a} c_3(a, a + u)
\]

and

\[
\lim_{T \to \infty} T^d \text{cov} (\hat{c}(u) - c(u), \hat{c}(v) - c(v)) = C_{u,v}
\]

with \( C_{u,v} \) defined in Theorem 4.9.
Corollary 4.1 is obtained by a direct application of Theorem 4.9. The next theorem and corollary are related to Theorems 4.1 and 4.5.

Let
\[
I^{(T)}(\lambda) = \frac{1}{(2\pi)^d} d^{(T)}(\lambda) d^{(T)}(-\lambda) = \frac{1}{(2\pi)^d} \sum_{\|u\| \leq T-1} \hat{c}(u)e^{-i<\lambda,u>}, \quad j = 1, \ldots, d
\]
be the d-dimensional periodogram.

The next theorem involves smoothed periodograms
\[
\int_0^{2\pi} I^{(T)}(\lambda) A_j(\lambda) d\lambda, \quad j = 1, \ldots, s
\]
with weight functions \(A(\lambda)\). The fourth-order cumulant spectral density arises and is denoted by
\[
f_4(\lambda, \mu, \eta) = \frac{1}{(2\pi)^3} \sum_{a,b,d} c_4(a,b,d)e^{-(i<\lambda,a> + i<\mu,b> + i<d,\eta>)}
\]

**Theorem 4.10 (Rosenblatt, 1985)** Let \(\{X(t)\}\) be an ergodic strictly stationary random field that satisfies the assumptions of Theorem 4.9. Consider
\[
\int_0^{2\pi} I^{(T)}(\lambda) A_j(\lambda) d\lambda, \quad j = 1, \ldots, s
\]
which are quadratic forms in \(X(t)\) with real-valued weight functions \(A_j(\lambda)\) square integrable. The smoothed periodograms then are asymptotically normal with means
\[
\int_0^{2\pi} f(\lambda) A_j(\lambda) d\lambda, \quad j = 1, \ldots, s
\]
and limiting covariances
\[
(2\pi)^d \{2 \int_0^{2\pi} A_j(\lambda) A_k(\lambda) f^2(\lambda) d\lambda + \int_0^{2\pi} \int_0^{2\pi} f_4(\lambda, -\mu, \mu) A_j(\lambda) A_k(\mu) d\lambda d\mu\}.
\]
The result follows from approximating the integrals \( \int_0^{2\pi} I^{(T)}(\lambda)A_j(\lambda)d\lambda \) by

\[
\sum_{|u_j| \leq L} (\hat{c}(u) - c(u)) a_j(u) + R^{(T)}(L)
\]

where \( a_j(u) \) is the Fourier coefficient of \( A_j(\lambda) \),

\[
a_j(u) = \frac{1}{(2\pi)^d} \int_0^{2\pi} A_j(\lambda)e^{-i\lambda \cdot u}d\lambda
\]

and \( R^{(T)}(L) \) is a remainder term depending on \( T \) and \( L \) which goes to zero in probability. Theorem 4.9 is then directly applied to

\[
\sum_{|u_j| \leq L} (\hat{c}(u) - c(u)) a_j(u).
\]

**Corollary 4.2** Let \( \{X(t)\} \) satisfy the assumptions of Theorem 4.10. Then

\[
T^{d/2} \begin{bmatrix}
\bar{X} \\
\int_0^{2\pi} I^{(T)}(\lambda)A_1(\lambda)d\lambda \\
\vdots \\
\int_0^{2\pi} I^{(T)}(\lambda)A_d(\lambda)d\lambda
\end{bmatrix}
\]

is jointly asymptotically normal and

\[
\lim_{T \to \infty} T^d \text{cov} \left( \bar{X}, \int_0^{2\pi} I^{(T)}(\lambda)A_j(\lambda)d\lambda \right) = (2\pi)^2 \int_0^{2\pi} A_j(\lambda)f_3(0, \lambda)d\lambda
\]

in addition to the covariances specified in Theorem 4.10.

This result is obtained in the same manner as Theorem 4.10 and involves the direct application of Corollary 4.1.

**Proof of Theorems 4.1 and 4.5**

The details are worked out here as they are not readily available in the literature. This also leads to a remark concerning asymptotic bias (Remark 4.1).
Let

\[ J^{(T)}(A_k) = \left( \frac{2\pi}{T} \right)^2 \sum_{s, \tau=1}^{T-1} (J^{(T)}(\frac{2\pi}{T}\mathbf{s}) - f(\frac{2\pi}{T}\mathbf{s}; \theta_0)) A_k(\frac{2\pi}{T}\mathbf{s}; \theta_0) \]

and

\[ J(A_k) = \int_0^{2\pi} (J^{(T)}(\mathbf{\lambda}) - f(\mathbf{\lambda}; \theta_0)) A_k(\mathbf{\lambda}; \theta_0) d\mathbf{\lambda}. \]

It will be shown that

\[ T \begin{bmatrix} \bar{X} - \mu \\ J^{(T)}(A_1) \\ J^{(T)}(A_2) \end{bmatrix} \]

is asymptotically normal with the specified first- and second-order moments. Now

\[ T \begin{bmatrix} \bar{X} - \mu \\ J^{(T)}(A_1) \\ J^{(T)}(A_2) \end{bmatrix} = T \begin{bmatrix} \bar{X} - \mu \\ J(A_1) \\ J(A_2) \end{bmatrix} + T \begin{bmatrix} 0 \\ J^{(T)}(A_1) - J(A_1) \\ J^{(T)}(A_2) - J(A_2) \end{bmatrix}. \]

The first vector is that of Corollary 4.2 with \(d = 2\) and \(s = 2\) (Theorem 4.10). In showing its asymptotic normality, Theorem 4.10 and Corollary 4.2 are established for a random field \(\{X(t)\}\) in \(\mathcal{Z}^2\) with a two-dimensional parameter \(\theta\).

Consider an arbitrary linear combination

\[ T(\alpha_0(\bar{X} - \mu) + \alpha_1 J(A_1) + \alpha_2 J(A_2)) \].

Write \(J(A_k)\) as

\[ \frac{1}{(2\pi)^2} \int_0^{2\pi} \left( \sum_{|u_1| \leq T-1} \hat{c}(u) - \sum_{u_2 = -\infty}^{\infty} c(u) \right) e^{-i\mathbf{\lambda} \cdot \mathbf{u}} A(\mathbf{\lambda}) d\mathbf{\lambda} \]

\[ = \sum_{|u_j| \leq T-1} (\hat{c}(u) - c(u)) \left( \frac{1}{(2\pi)^2} \int_0^{2\pi} A(\mathbf{\lambda}) e^{-i\mathbf{\lambda} \cdot \mathbf{u}} d\mathbf{\lambda} \right) + R^{(T)}_1 \]

where

\[ a(u) = \frac{1}{(2\pi)^2} \int_0^{2\pi} A(\mathbf{\lambda}) e^{-i\mathbf{\lambda} \cdot \mathbf{u}} d\mathbf{\lambda} \]
is the Fourier coefficient of $A(\lambda)$. The above equality further becomes

$$\sum_{|u_j| \leq L} (\hat{c}(u) - c(u))a(u) + R_2^T(L) + R_1^T$$

where one may directly apply Theorem 4.10 to the first term.

Returning now to the linear combination above we have

$$T(\alpha_0(\bar{X} - \mu) + \sum_{|u_j| \leq L} (\hat{c}(u) - c(u))a(u)) + T(R_1^T + R_2^T(L))$$

where

$$a(u) = \frac{1}{(2\pi)^2} \int_0^{2\pi} (\alpha_1 A_1(\lambda) + \alpha_2 A_2(\lambda))e^{-i<\lambda, u>}d\lambda$$

is the Fourier coefficient of $\alpha_1 A_1(\lambda) + \alpha_2 A_2(\lambda)$. By the Cramer-Wold device and an application of Corollary 4.1,

$$T(\alpha_0(\bar{X} - \mu) + \sum_{|u_j| \leq L} (\hat{c}(u) - c(u))a(u))$$

is asymptotically normal with limiting variance $a'\Sigma a$ where $a' = (\alpha_0, a(-L), \ldots, a(L))$ and the matrix $\Sigma$ has the covariance structure of the aforementioned corollary. Now

$$a(u) = \frac{\alpha_1}{(2\pi)^2} \int_0^{2\pi} A_1(\lambda)e^{-i<\lambda, u>}d\lambda + \frac{\alpha_2}{(2\pi)^2} \int_0^{2\pi} A_2(\lambda)e^{-i<\lambda, u>}d\lambda$$

say. Therefore $a = \alpha_0 b_0 + \alpha_1 b_1 + \alpha_2 b_2$ where $b_0' = (1, 0, \ldots, 0)$ and $b_j' = (0, b_j(-L), \ldots, b_j(L))$, $j = 1, 2$.

Now rewrite $a'\Sigma a$ as $a'\Sigma^*a$ where $a' = (\alpha_0, \alpha_1, \alpha_2)$ and

$$\Sigma^* = \begin{bmatrix}
    b_0'\Sigma b_0 & b_0'\Sigma b_1 & b_0'\Sigma b_2 \\
    b_1'\Sigma b_0 & b_1'\Sigma b_1 & b_1'\Sigma b_2 \\
    b_2'\Sigma b_0 & b_2'\Sigma b_1 & b_2'\Sigma b_2
\end{bmatrix}.$$
Another application of Cramer-Wold gives

\[
T \left[ \begin{array}{c} X - \mu \\ \int(A_1) \\ \int(A_2) \end{array} \right] \overset{D}{\rightarrow} N_3(\mu, \Sigma^*),
\]

where the mean vector \( \mu \) is as yet unspecified. The elements of \( \Sigma^* \) are of the form:

\[
b'_0 \Sigma b_0 = \sum \limits_u c(u),
\]

\[
b'_j \Sigma b_j = \frac{1}{(2\pi)^2} \int_0^{2\pi} A_j(\lambda)(\sum \limits_{|u_x| \leq L} \sum \limits_a c_0(a, a + u)e^{-i\lambda \cdot u})d\lambda, \quad j = 1, 2
\]

and for \( k, \ell = 1, 2 \)

\[
b'_k \Sigma b_\ell = \frac{1}{(2\pi)^4} \sum \sum \sum \sum (c(a)c(a + v - u) + c(a + v)c(a - u))
\]

\[
\int_0^{2\pi} \int_0^{2\pi} A_k(\lambda)A_\ell(\beta)e^{-i\lambda \cdot u}e^{-i\beta \cdot v}d\lambda d\beta
\]

\[
= \frac{1}{(2\pi)^4} \int_0^{2\pi} \int_0^{2\pi} A_k(\lambda)A_\ell(\beta) \sum \sum (c(a)c(a + v - u))
\]

\[
+ c(a + v)c(a - u) + c(a, a + v))e^{-i\lambda \cdot u}e^{-i\beta \cdot v}d\lambda d\beta.
\]

The cumulant summability conditions are sufficient to ensure that these limiting covariances are finite. By computing the limiting covariances of \( X \) and the \( J(A_k) \) directly, they may be expressed more compactly in terms of integrals of second-, third- and fourth-order cumulant spectral densities. These computations require the following results of Leonov and Shiryaev [19], which are also given in Brillinger [5]. Brillinger has applied the first result frequently. Both theorems are recalled from Chapter 2.

**Theorem 4.11** Consider a two-way array of random variables \( X_{ij}; \quad j = 1, \ldots, N_i; \quad i = 1, \ldots, M \). Consider the \( M \) random variables

\[
Y_i = \prod_{j=1}^{N_i} X_{ij}, \quad i = 1, \ldots, M.
\]
The joint cumulant \( \text{cum}(Y_1, \ldots, Y_M) \) is then given by

\[
\sum_{\nu} \text{cum}(X_{ij}; \ ij \in \nu_1) \cdots \text{cum}(X_{ij}; \ ij \in \nu_p)
\]

where the summation is over all indecomposable partitions \( \nu = \nu_1 \cup \cdots \cup \nu_p \) of the two-way table

\[
(1, 1) \cdots (1, N_1) \\
\vdots \quad \vdots \\
(M, 1) \cdots (M, N_M)
\]

If the rows of the above table are denoted by \( R_1, \ldots, R_M \), then a partition \( \nu_1, \ldots, \nu_p \) is indecomposable if and only if there exist no sets \( \nu_{m_1}, \ldots, \nu_{m_q}, \ (q < p) \) and rows \( R_{i_1}, \ldots, R_{i_\ell}, \ (\ell < M) \) with

\[
\nu_{m_1} \cup \cdots \cup \nu_{m_q} = R_{i_1} \cup \cdots \cup R_{i_\ell}.
\]

**Theorem 4.12** Let \( X(t), \ t \in \mathbb{Z}^d \) be a strictly stationary random field all of whose moments exist and satisfy:

\[
\sum \cdots \sum (||u^{(1)}|| + \cdots + ||u^{(k-1)}||) c_k(u^{(1)}, \ldots, u^{(k-1)}) < \infty
\]

where \( ||u|| = \langle u, u \rangle^{\frac{1}{2}} \), \( k = 1, 2, \ldots \). Then as \( T \to \infty \),

\[
\text{cum}_k \{d^{(T)}(\lambda^{(1)}), \ldots, d^{(T)}(\lambda^{(k)})\} = \Delta^{(T)}(\sum_{j=1}^{k-1} \lambda^{(j)})(2\pi)^d (k-1) f_k(\lambda^{(1)}, \ldots, \lambda^{(k-1)}) + O(T^{d-1})
\]

where \( \Delta^{(T)}(\lambda) = \sum_{t=0}^{T-1} e^{-i < \lambda, t>} \). The function \( \Delta^{(T)}(\lambda) \) has the properties: \( \Delta^{(T)}(\lambda) = i^2 \) for \( \lambda = 0 \mod 2\pi \), and \( \Delta^{(T)}(2\pi s) = 0 \) for \( s \in \mathbb{Z}^d \) with \( s \neq 0 \mod T \).
The limiting covariance results are obtained by directly applying these theorems.

\[
\lim_{T \to \infty} T^2 \text{Var} (\bar{X}) = \lim_{T \to \infty} \sum_{u_1 = -T+1}^{T-1} \frac{(T - |u_1|)(T - |u_2|)}{T^2} c(u) = \sum_{u} c(u)
\]

after interchanging the order of summation.

\[
\lim_{T \to \infty} T^4 \text{cov} (\bar{X}, J(A_k)) = \lim_{T \to \infty} T^2 \text{cov} (\bar{X}, \int_0^{2\pi} (I^{(T)}(\lambda) - f(\lambda))A_k(\lambda)d\lambda) = \lim_{T \to \infty} T^2 \int_0^{2\pi} A_k(\lambda) \text{cov} (\bar{X}, I^{(T)}(\lambda))d\lambda.
\]

Now,

\[
\text{cov} (\bar{X}, I^{(T)}(\lambda)) = \frac{1}{(2\pi)^2 T^4} \text{cov} (d^{(T)}(0), d^{(T)}(\lambda)d^{(T)}(-\lambda)) = \frac{1}{(2\pi)^2 T^4} \text{cum}_3 (d^{(T)}(0), d^{(T)}(\lambda)d^{(T)}(-\lambda)) = \frac{1}{(2\pi)^2 T^4}((2\pi)^4 \Delta^{(T)}(0)f_3(0, \lambda) + O(T)) = \frac{(2\pi)^2}{T^2} f_3(0, \lambda) + O(T^{-1}).
\]

Therefore,

\[
\lim_{T \to \infty} T^2 \int_0^{2\pi} A_k(\lambda) \text{cov} (\bar{X}, I^{(T)}(\lambda))d\lambda = (2\pi)^2 \int_0^{2\pi} A_k(\lambda)f_3(0, \lambda)d\lambda.
\]

Similarly,

\[
\lim_{T \to \infty} T^2 \text{cov} (J(A_k), J(A_L))
\]
\[
\lim_{T \to \infty} T^2 \int_0^{2\pi} \int_0^{2\pi} A_k(\lambda)A_t(\beta) \text{cov} \left( I^{(T)}(\lambda), I^{(T)}(\beta) \right) d\lambda d\beta
\]

where

\[
\text{cov} \left( I^{(T)}(\lambda), I^{(T)}(\beta) \right) = \frac{1}{(2\pi T)^4} \text{cov} \left( d^{(T)}(\lambda) d^{(T)}(-\lambda), d^{(T)}(\beta) d^{(T)}(-\beta) \right)
\]

\[
= \frac{1}{(2\pi T)^4} \left\{ \text{cov} \left( d^{(T)}(\lambda), d^{(T)}(\beta) \right) \text{cov} \left( d^{(T)}(-\lambda), d^{(T)}(-\beta) \right)
\right.
\]

\[
+ \text{cov} \left( d^{(T)}(\lambda), d^{(T)}(-\beta) \right) \text{cov} \left( d^{(T)}(-\lambda), d^{(T)}(\beta) \right)
\]

\[
+ \text{cum}_4 \left( d^{(T)}(\lambda), d^{(T)}(-\lambda), d^{(T)}(\beta), d^{(T)}(-\beta) \right) \left\} \right.
\]

\[
= \frac{1}{(2\pi T)^4} \left\{ (2\pi)^4 |\Delta^{(T)}(\lambda + \beta)|^2 f^2(\lambda)
\right.
\]

\[
+ (2\pi)^4 |\Delta^{(T)}(\lambda - \beta)|^2 f^2(\lambda)
\]

\[
+ (2\pi)^6 T^2 f_4(\lambda, -\lambda, \beta) + O(T) \}
\]

Therefore,

\[
\lim_{T \to \infty} T^2 \int_0^{2\pi} \int_0^{2\pi} A_k(\lambda)A_t(\beta) \text{cov} \left( I^{(T)}(\lambda), I^{(T)}(\beta) \right) d\lambda d\beta
\]

\[
= (2\pi)^2 \left\{ 2 \int_0^{2\pi} A_k(\lambda)A_t(\lambda) f^2(\lambda) d\lambda
\right.
\]

\[
+ \int_0^{2\pi} \int_0^{2\pi} A_k(\lambda)A_t(\beta) f_4(\lambda, -\lambda, \beta) d\lambda d\beta \}
\]

Recall that the arbitrary linear combination of

\[
T \begin{bmatrix} X - \mu \\ J(A_1) \\ J(A_2) \end{bmatrix}
\]

takes the form

\[
T(\alpha_0(\bar{X} - \mu) + \sum_{|u_j| \leq L} (\hat{c}(u) - c(u))a(u)) + T(H_1 + H_2(L))
\]
One must now show that the remainder terms go to zero.

\[ |TR_1^{(T)}| = \left| \sum_{|u_1| \geq T} \sum_{u_2 = -\infty}^{\infty} Tc(u)a(u) + \sum_{|u_1| \leq T-1} \sum_{|u_2| \geq T} Tc(u)a(u) \right| \]

\[ \leq \sum_{|u_1| \geq T} \sum_{u_2 = -\infty}^{\infty} |u_1||c(u)||a(u)| + \sum_{|u_1| \leq T-1} \sum_{|u_2| \geq T} |u_2||c(u)||a(u)| \]

which is \( < \varepsilon \) for \( T \geq T_0(\varepsilon) \), by the cumulant summability conditions of Theorem 4.12 and \( |a(u)| \leq M \) for some \( M > 0 \).

Also,

\[ TR_2^{(T)}(L) = 2 \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = -T+1}^{T-1} T(\hat{c}(u) - c(u))a(u) \]

\[ + 2 \sum_{u_1 = -T+1}^{T-1} \sum_{u_2 = L+1}^{T-1} T(\hat{c}(u) - c(u))a(u). \]

Consider the first term and rewrite it as

\[ 2 \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = -T+1}^{T-1} T(\hat{c}(u) - c^{(T)}(u) - (c(u) - c^{(T)}(u)))a(u) \]  \hspace{1cm} (4.7)

where

\[ c^{(T)}(u) = E(\hat{c}(u)) = \frac{(T - |u_1|)(T - |u_2|)}{T^2} c(u). \]

Then (4.7) can be expressed as

\[ 2 \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = -T+1}^{T-1} T(\hat{c}(u) - c^{(T)}(u))a(u) - 2 \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = -T+1}^{T-1} (|u_1| + |u_2|)c(u)a(u) \]

\[ + 2 \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = -T+1}^{T-1} \frac{|u_1||u_2|}{T} c(u)a(u). \]

The cumulant summability conditions of Theorem 4.12 and the dominated convergence theorem imply that the last two terms above are negligible for some \( L \). Apply Chebychev's inequality to the first term.

\[ P\left( \sum_{u_1 = L+1}^{T-1} \sum_{u_2 = T+1}^{T-1} T(\hat{c}(u) - c^{(T)}(u))a(u) > \varepsilon \right) \]
\[ \leq \frac{T^2}{c^2} \sum_{u_1=L+1}^{T-1} \sum_{v_1=L+1}^{T-1} \sum_{u_2=-T+1}^{-1} \sum_{v_2=-T+1}^{-1} \text{cov} (\hat{c}(u), \hat{c}(v)) a(u)a(v) \]

where

\[ \lim_{T \to \infty} T^2 \text{cov} (\hat{c}(u), \hat{c}(v)) = C_{u,v} \]
\[ = \sum_{\alpha} (c(\alpha)c(\alpha + v - u) + c(\alpha + v)c(\alpha - u) + c_4(u, \alpha, \alpha + v)) \]

The cumulant summability conditions of Theorem 4.9 imply that there exists an \( L \) such that the above probability is negligible. By the same argument, the second term of \( TR_2^{(T)}(L) \) goes to zero in probability.

All that remains now is to demonstrate that
\[ T \begin{bmatrix} 0 \\ J^{(T)}(A_1) - J(A_1) \\ J^{(T)}(A_2) - J(A_2) \end{bmatrix} \xrightarrow{P} 0. \]

It is sufficient to show that
\[ T(J^{(T)}(A) - J(A)) \xrightarrow{P} 0. \]

\[ T(J^{(T)}(A) - J(A)) \]
\[ = T\left( \frac{(2\pi)^2}{T} \sum_{u_j=-T}^{T-1} (f^{(T)}(2\pi T s) - f^{(T)}(2\pi s)) A(2\pi T s) - \int_0^{2\pi} (f^{(T)}(\lambda) - f(\lambda)) A(\lambda) d\lambda \right) \]
\[ = T \left( (\sum_{u_j=-T+1}^{T-1} \hat{c}(u) - \sum_{u} c(u)) (\frac{2\pi}{T})^2 \frac{1}{(2\pi)^2} \sum_{u_j=1}^{T-1} A(\frac{2\pi}{T} s) e^{-i \frac{2\pi}{T} s \cdot u} \right. \]
\[ - \left. (\sum_{u_j=-T+1}^{T-1} \hat{c}(u) - \sum_{u} c(u)) \frac{1}{(2\pi)^2} \int_0^{2\pi} A(\lambda) e^{-i \lambda \cdot u} d\lambda \right) \]
\[ = T \left( \sum_{u_j=-T+1}^{T-1} (\hat{c}(u) - c(u)) b^{(T)}(u) \right) + TR_3^{(T)} \]
\[ = \sum_{|u_j| \leq L} (\hat{c}(u) - c(u)) b^{(T)}(u) + TR_3^{(T)} + TR_4^{(T)}(L) \]
where
\[
\delta^{(T)}(u) = \frac{1}{(2\pi)^2} \left( \left( \frac{2\pi}{T} \right)^2 \sum_{i,j=1}^{T-1} A\left( \frac{2\pi}{T} s \right) e^{-i\frac{2\pi}{T} s_u} - \int_0^{2\pi} A(\lambda) e^{-i\lambda u} d\lambda \right)
= O(T^{-1}) \text{ for each } u.
\]

The same arguments as before are used to show that \( T(R_3^{(T)} + R_4^{(T)}(L)) \overset{P}{\to} 0 \).

The term
\[
T \left( \sum_{|u_j| \leq L} (\hat{c}(u) - c(u))\delta^{(T)}(u) \right)
\]
becomes
\[
\sum_{|u_j| \leq L} (\hat{c}(u) - c(u))K(u)
\]
where \( K(u) \) is a bounded function of \( u \). This goes to zero in probability as a direct consequence of the asymptotic normality of \( T \sum_{|u_j| \leq L} (\hat{c}(u) - c(u))K(u) \). Hence the results of Theorems 4.1 and 4.5 follow.

**Remark 4.1** The first-order moment conditions are examined under the cumulant summability assumptions of Theorem 4.12.

It is known that,
\[
\lim_{T \to \infty} TE(\bar{X} - \mu) = 0.
\]
Let \( J(A) = J(A_k), \quad k = 1, 2 \). Then,
\[
\lim_{T \to \infty} TE(J(A)) = \lim_{T \to \infty} \int_0^{2\pi} (I^{(T)}(\lambda) - f(\lambda))A(\lambda)d\lambda
= \lim_{T \to \infty} T \int_0^{2\pi} (f^{(T)}(\lambda) - f(\lambda))A(\lambda)d\lambda
\]
where
\[
E(I^{(T)}(\lambda)) = f^{(T)}(\lambda)
\]
\[
\begin{align*}
&= \frac{1}{(2\pi)^2} \sum_{u_j = -T+1}^{T-1} \frac{(T - |u_1|)(T - |u_2|)}{T^2} c(u) e^{-i\lambda \cdot u} \\
&= \frac{1}{(2\pi)^2} \sum_{u_j = -T+1}^{T-1} \left( 1 - \frac{(|u_1| + |u_2|)}{T} + \frac{|u_1||u_2|}{T^2} \right) c(u) e^{-i\lambda \cdot u} , \quad j = 1, 2.
\end{align*}
\]

The spectral density can be expressed as

\[
f(\lambda) = \frac{1}{(2\pi)^2} \sum_{u_j = -T+1}^{T-1} c(u) e^{-i\lambda \cdot u}
\]

\[
+ \frac{1}{(2\pi)^2} \left( \sum_{|u_1| \geq T} \sum_{u_2 = -\infty}^{\infty} c(u) e^{-i\lambda \cdot u} + \sum_{|u_1| \leq T-1} \sum_{|u_2| \geq T} c(u) e^{-i\lambda \cdot u} \right)
\]

and so

\[
f^{(T)}(\lambda) - f(\lambda)
\]

\[
= \frac{1}{(2\pi)^2} \sum_{u_j = -T+1}^{T-1} \left( -\frac{(|u_1| + |u_2|)}{T} + \frac{|u_1||u_2|}{T^2} \right) c(u) e^{-i\lambda \cdot u} + o(T^{-1})
\]

\[
= o(1) \quad \text{uniformly in } \lambda, \text{ as } T \to \infty.
\]

However,

\[
\lim_{T \to \infty} T \int_0^{2\pi} (f^{(T)}(\lambda) - f(\lambda)) A(\lambda) d\lambda
\]

\[
= \lim_{T \to \infty} \left( \sum_{u_j = -T+1}^{T-1} -(|u_1| + |u_2|) c(u) a(u) + \sum_{u_j = -T+1}^{T-1} \frac{|u_1||u_2|}{T} c(u) a(u) + o(1) \right)
\]

\[
= -\sum_{u_j = -\infty}^{\infty} (|u_1| + |u_2|) c(u) a(u) , \quad j = 1, 2.
\]

Thus,

\[
\lim_{T \to \infty} T E(\int_0^{2\pi} (I^{(T)}(\lambda) - f(\lambda)) A(\lambda) d\lambda) = O(1)
\]

if \( I^{(T)}(\lambda) \) is based on the usual sample covariance

\[
\hat{c}(u) = \frac{1}{T^2} \sum_{0 \leq t_j + u_j \leq T-1} X(t)X(t + u) , \quad j = 1, 2.
\]
In $\mathbb{Z}^2$, Rosenblatt assumes that one observes a stationary random field over the square of lattice points $t_i = 0, \ldots, T - 1$, $i = 1, 2$. He considers the covariance estimator as given by $\hat{c}(\mathbf{u})$ above. However, the asymptotic normal distribution of $TI(A)$ will not have zero mean. It appears that Rosenblatt computes his covariance estimates over the square of lattice points $t_i = 0, \ldots, T - 1$ under the assumption that he in fact has data available on the larger square lattice $t_i = 0, \ldots, 2T - 1$.

His covariance estimator then is actually of the form

$$c^{(T)}(\mathbf{u}) = \frac{1}{T^2} \sum_{t_i=0}^{T-1} X(t) X(t + \mathbf{u})$$

which has expectation $c(\mathbf{u})$.

From an asymptotic point of view, this is not unreasonable. However, from a data-analytic perspective, this amounts to considering only a subset of the available data. In particular, the variance estimator $c^{(T)}(0)$ will be based upon only one-quarter of the data.

These observations seem to suggest that in practice, one should work with $f^{(T)}(\lambda)$ in the likelihood and least squares functions, that is

$$J^*(A) = \int_0^{2\pi} (I^{(T)}(\lambda) - f^{(T)}(\lambda)) A(\lambda) d\lambda$$

The same argument as that used in the proof of Theorem 4.1 then shows that

$$TJ^*(A) \xrightarrow{D} N(0, \sigma^2)$$

where $J^*(A)$ and $J(A)$ share the same limiting variance $\sigma^2$. From a computational perspective, the difference between the point estimates is $f^{(T)}(\lambda) = f(\lambda) + O(T^{-1})$
and the error term is uniform in \( \lambda \). However, centering the least squares or likelihood functions at \( f(\lambda; \theta_0) \) will lead to incorrect asymptotic confidence intervals for \( \theta_0 \) and the ratio parameter \( R \).

**Note**

These considerations do not arise for a stationary sequence \( X(t), t \in \mathbb{Z} \), where the bias is of smaller order than \( T^{-\frac{1}{2}} \).

**Remark 4.2** Theorem 4.1 is the random field extension of Theorem 5.10.1 in Brillinger [5] where the asymptotic normality of

\[
J^{(T)}(A) = \frac{2\pi}{T} \sum_{s=1}^{T-1} f^{(T)}(\frac{2\pi s}{T}) A(\frac{2\pi s}{T})
\]

for a stationary real-valued sequence \( X(t) \) is obtained. It is assumed that \( X(t) \) satisfies

\[
\sum_{u_1, \ldots, u_{k-1}} |u_j|^\ell |c_k(u_1, \ldots, u_{k-1})| < \infty \text{ for some } \ell \geq 0.
\]

Brillinger first computes first- and second-order moments before showing that higher-order cumulants \( \text{cum}\{J^{(T)}(A_{j_1}), \ldots, J^{(T)}(A_{j_k})\} \) have dominant terms of order \( T^{-k+1} \) so that when the \( J^{(T)}(A) \) are properly normalized by \( T^{\frac{1}{2}} \), these joint cumulants of order three and higher go to zero. The asymptotic normality follows from the fact that the normal distribution is determined by its moments.

The following result involving the joint cumulants of discrete Fourier transforms in conjunction with the rules of Leonov and Shiryaev (Theorem 4.11 in the Proof of Theorems 4.1 and 4.5 section) is essential to this argument.
Brillinger's Theorem 4.3.2

\[ \text{cum} \{ d^{(T)}(\lambda_1), \ldots, d^{(T)}(\lambda_k) \} = (2\pi)^{k-1} \Delta^{(T)}(\Sigma \lambda_j) f_k(\lambda_1, \ldots, \lambda_{k-1}) + O(1) \]

where \( \Delta^{(T)}(\lambda) = \sum_{t=0}^{T-1} e^{-i\lambda t} \).

It seems reasonable to expect the same argument to carry over into the stationary random field case \( X(t), t \in \mathbb{Z}^2 \) using Theorem 4.12 (Proof of Theorems 4.1 and 4.5 section). In \( \mathbb{Z}^2 \), Theorem 4.12 asserts

\[ \text{cum} \{ d^{(T)}(\lambda^{(1)}), \ldots, d^{(T)}(\lambda^{(k)}) \} = (2\pi)^{2(k-1)} \Delta^{(T)}(\Sigma \lambda^{(j)}) f_k(\lambda_1, \ldots, \lambda_{k-1}) + O(T). \]

However, as will be briefly outlined, it appears as if even the third-order cumulant of the \( T J^{(T)}(A) \) does not go to zero but is instead \( O(1) \) if we directly apply the result of Theorem 4.12 above.

Let

\[ J^{(T)}(A_k) = \left( \frac{2\pi}{T} \right)^2 \sum_{s_j=1}^{T-1} j^{(T)}(\frac{2\pi}{T} s) A_k(\frac{2\pi}{T} s). \]

Then,

\[ \text{cum}_3 \{ T J^{(T)}(A_{k_1}), T J^{(T)}(A_{k_2}), T J^{(T)}(A_{k_3}) \} \]

\[ = \frac{(2\pi)^2}{T^9} T^{-1} \sum_{s_j^{(1)}=1}^{T-1} \sum_{s_j^{(2)}=1}^{T-1} \sum_{s_j^{(3)}=1}^{T-1} \]

\[ \text{cum}_3 \left\{ d^{(T)}(\frac{2\pi}{T} s^{(1)}), d^{(T)}(-\frac{2\pi}{T} s^{(1)}), d^{(T)}(\frac{2\pi}{T} s^{(2)}), d^{(T)}(-\frac{2\pi}{T} s^{(2)}), \right. \]

\[ \left. d^{(T)}(\frac{2\pi}{T} s^{(3)}), d^{(T)}(-\frac{2\pi}{T} s^{(3)}) \right\}. \]

One applies the rules of Leonov and Shiryaev to evaluate this third cumulant and sum the products of cumulants of the \( d^{(T)}(\cdot) \) over all indecomposable partitions of the following table:
\( s^{(1)} - s^{(1)} \)
\( s^{(2)} - s^{(2)} \)
\( s^{(3)} - s^{(3)} \)

Consider the following partition of size 3,

\[ p = \{ \{ s^{(1)}, -s^{(2)} \}, \{ s^{(2)}, -s^{(3)} \}, \{ -s^{(1)}, s^{(3)} \} \} . \]

Considering this one partition only, the equation above becomes

\[
\frac{(2\pi)^2}{T^9} \sum_{s_j^{(1)}=1}^{T-1} \sum_{s_j^{(2)}=1}^{T-1} \sum_{s_j^{(3)}=1}^{T-1} \left( \text{cov} \left( d^{(T)} \left( \frac{2\pi}{T} s^{(1)} \right), d^{(T)} \left( -\frac{2\pi}{T} s^{(2)} \right) \right) \right.

\left. \text{cov} \left( d^{(T)} \left( \frac{2\pi}{T} s^{(2)} \right), d^{(T)} \left( -\frac{2\pi}{T} s^{(3)} \right) \right) \right.

\left. \text{cov} \left( d^{(T)} \left( -\frac{2\pi}{T} s^{(1)} \right), d^{(T)} \left( \frac{2\pi}{T} s^{(3)} \right) \right) \right)

= \frac{(2\pi)^2}{T^9} \sum_{s_j^{(1)}=1}^{T-1} \sum_{s_j^{(2)}=1}^{T-1} \sum_{s_j^{(3)}=1}^{T-1} \left( ((2\pi)^2\Delta^{(T)} \left( \frac{2\pi}{T} (s^{(1)} - s^{(2)}) \right) f \left( \frac{2\pi}{T} s^{(2)} \right) + O(T)) \right.

\left. ((2\pi)^2\Delta^{(T)} \left( \frac{2\pi}{T} (s^{(2)} - s^{(3)}) \right) f \left( \frac{2\pi}{T} s^{(3)} \right) + O(T)) \right.

\left. ((2\pi)^2\Delta^{(T)} \left( \frac{2\pi}{T} (s^{(3)} - s^{(2)}) \right) f \left( \frac{2\pi}{T} s^{(3)} \right) + O(T)) \right)

= O(1)

due to the remainder terms being \( O(T) \). If the remainder term was \( o(T) \), the third- and higher-order cumulants would in fact tend to zero as \( T \to \infty \).
Proof of Theorems 4.2 and 4.6: Consistency of Asymptotic m.l.e. $\hat{\theta}$

One can show that under suitable conditions, the approximate m.l.e. $\hat{\theta}$ is a consistent estimator of the true parameter value $\theta_0$. Without loss of generality, assume $\theta$ is uni-dimensional i.e. $\theta \in \mathcal{R}$. The consistency of the least squares estimator is obtained by the same argument and hence the details are omitted. Our discussion parallels that of Kulperger [15] and [16]. Further details may be found there. Kulperger considers the consistency of the m.l.e. $\hat{\theta}$ of the parameters of a simple branching diffusion immigration process.

The consistency of $\hat{\theta}$ follows by verifying the general conditions outlined in Brillinger [4].

**Brillinger’s Lemma**

Suppose

i) $\Theta$ is an open subset of $\mathcal{R}^k$,

ii) $(\Omega, \mathcal{F}, P)$ is a probability space with $\Omega$ a complete, separable, metric space,

iii) $L^{(T)}(\theta, \omega)$ is a real-valued, Borel measurable function for each $(\theta, \omega) \in \Theta \times \Omega$ and $T > 0$,

iv) $L(\theta)$ is a real-valued, Borel measurable, lower semi-continuous function with $L(\theta) < L(\theta_0)$ $\forall \theta \neq \theta_0$ (i.e. $L$ has a unique maximum at $\theta_0$),

v) $L^{(T)}(\theta_0, \omega) \xrightarrow{P} L(\theta_0)$ and $L^{(T)}(\theta, \omega) \leq L(\theta) + o_P(1)$ as $T \to \infty$, 
vi) given $\epsilon > 0$, $\eta > 0$, $\theta \neq \theta_0$, there exists a neighbourhood $U_1$ of $\theta_1$ and a number $T_0 > 0$ such that

\[ P(\omega : \sup_{U_1} L^{(T)}(\theta, \omega) - L^{(T)}(\theta_0, \omega) > \epsilon) < \eta \quad \forall T \geq T_0. \]

vii) there exists a compact set $K \subset \Theta$, $\theta_0 \in K$, $\hat{\theta} \in K$, such that

\[ L^{(T)}(\hat{\theta}, \omega) \geq L^{(T)}(\theta, \omega) \quad \forall \theta \in K. \]

Then there exists $\hat{\theta}$, a local maximum of $L^{(T)}(\theta, \omega)$ such that

\[ \hat{\theta} \overset{P}{\to} \theta_0 \quad \text{as} \ T \to \infty. \]

Let

\[ L(\theta) = -\int_0^{2\pi} \left( \log f(\lambda; \theta) + \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} \right) d\lambda. \]

$L(\cdot)$ exists as a Riemann integral as $f(\cdot; \theta)$ is a continuous, bounded function in $(\lambda, \theta)$. As such, consider

\[ L^{(T)}(\theta) = -\left( \frac{2\pi}{T} \right)^2 \sum_{s=1}^{T-1} \left( \log f \left( \frac{2\pi}{T} s; \theta \right) + \frac{f^{(T)} \left( \frac{2\pi}{T} s \right)}{f \left( \frac{2\pi}{T} s; \theta \right)} \right). \]

It can be shown that $L(\theta)$ has a unique maximum at $\theta_0$, the true parameter value. Consider,

\[ L(\theta) - L(\theta_0) = -\int_0^{2\pi} \left( \log f(\lambda; \theta) + \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} - 1 - \log f(\lambda; \theta_0) \right) d\lambda \]

\[ = \int_0^{2\pi} \left( \log \left( \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} \right) - \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} + 1 \right) d\lambda. \]

The integrand is of the form

\[ \log x - x + 1 \quad \text{where} \]


\[
\log x < 0 \text{ if } 0 < x < 1 , \\
\log x = 0 \text{ if } x = 1 \text{ and} \\
\log x > 0 \text{ if } x > 1 .
\]

Therefore, \( L(\theta) - L(\theta_0) \leq 0 \) and \( L(\theta) = L(\theta_0) \) if and only if \( f(\lambda; \theta) = f(\lambda; \theta_0) \) and hence \( \theta = \theta_0 \).

The conditions of Brillinger’s Lemma are now verified:

i) \( \Theta = \mathcal{R} \) which is a locally compact, complete separable metric space.

ii) This has already been assumed.

iii) for each \( \theta \) and \( T \), \( L^{(T)}(\theta) \) is a finite sum of random variables and \( f(\lambda; \theta) \) is a bounded continuous function of \( \theta \) for each \( \lambda \).

iv) \( L(\theta) \) is a continuous function of \( \theta \) and from above, \( L(\theta) < L(\theta_0) \) for \( \theta = \theta_0 \).

v) \( L^{(T)}(\theta) \xrightarrow{L} L(\theta) \) as a consequence of Theorem 4.5 as

\[
L^{(T)}(\theta) = -\left(\frac{2\pi}{T}\right)^2 \sum_{s_j = 1}^{T-1} (\log f\left(\frac{2\pi}{T}s; \theta\right) + I^{(T)}(\frac{2\pi}{T}s) f^{-1}\left(\frac{2\pi}{T}s; \theta\right)) \\
= -\left(\frac{2\pi}{T}\right)^2 \sum_{s_j = 1}^{T-1} \log f\left(\frac{2\pi}{T}s; \theta\right) \\
-\left(\frac{2\pi}{T}\right)^2 \sum_{s_j = 1}^{T-1} (I^{(T)}(\frac{2\pi}{T}s) - f\left(\frac{2\pi}{T}s; \theta_0\right)) f^{-1}\left(\frac{2\pi}{T}s; \theta\right) \\
-\left(\frac{2\pi}{T}\right)^2 \sum_{s_j = 1}^{T-1} f\left(\frac{2\pi}{T}s; \theta_0\right) f^{-1}\left(\frac{2\pi}{T}s; \theta\right) .
\]

From Theorem 4.5,

\[
TJ^{(T)}(A) = T\left(\frac{2\pi}{T}\right)^2 \sum_{s_j = 1}^{T-1} (I^{(T)}(\frac{2\pi}{T}s) - f\left(\frac{2\pi}{T}s; \theta_0\right)) (-f^{-1}(\frac{2\pi}{T}s; \theta))
\]
\[ \overset{D}{\to} N(0, \sigma^2) \]

where \( \sigma^2 < \infty \) and \( A(\frac{2\pi}{I} \mathbf{s}; \theta) = -f^{-1}(\frac{2\pi}{I} \mathbf{s}; \theta) \). Thus

\[
\left( \frac{1}{T} \right)^{T} f^{(T)}(A) \overset{P}{\to} 0 \,.
\]

From this it follows that for each \( \theta \),

\[
L^{(T)}(\theta) \overset{P}{\to} -\int_{0}^{2\pi} \left( \log f(\lambda; \theta) + \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} \right)d\lambda
\]

which is \( L(\theta) \). In particular

\[
L^{(T)}(\theta_0) \overset{P}{\to} L(\theta_0) \,.
\]

Proposition 4.2.3 of Kulperger [15] further establishes:

i) given \( \epsilon, \eta > 0 \), \( \exists \) a neighbourhood of \( \theta_0 \in \Theta \), \( N_{\theta_0}(\epsilon) \) and a positive integer \( T_0(\epsilon, \eta, \theta_0) \) such that

\[
P(\sup_{\theta \in N_{\theta_0}(\epsilon)} |L^{(T)}(\theta) - L(\theta)| > \epsilon) < \eta \forall T \geq T_0 \,.
\]

ii) Also, for each compact set \( K \subset \Theta \),

\[
P(\sup_{K} |L^{(T)}(\theta) - L(\theta)| > \epsilon) < \eta \forall T \geq T_0 \,.
\]

i.e. \( L^{(T)}(\theta) \overset{a}{\to} L(\theta) \) in probability on compact sets \( K \).

These results follow from Chebyshev's inequality and together with

\[
L^{(T)}(\theta_0) \overset{P}{\to} L(\theta_0)
\]

verify conditions v) and vi) of Brillinger's Lemma.
The remaining conditions of the Lemma and its assertion are established in [15, Theorem 4.2.4] and its subsequent corollary. For \( \theta_0 \in G \subset K \), \( G \) open and \( K \) compact, there exists \( \hat{\theta} \in K \) such that \( L^{(T)}(\hat{\theta}) \geq L^{(T)}(\theta) \) \( \forall \theta \in K \). Then under conditions i) and ii) above, \( P(\hat{\theta} \in K/G) \to 0 \) as \( T \to \infty \). The corollary asserts the existence of \( \hat{\theta} \) such that

\[
\frac{\partial}{\partial \theta} L^{(T)}(\theta) \bigg|_{\theta = \hat{\theta}} = 0 \quad \text{and} \quad \hat{\theta} \overset{P}{\to} \theta_0.
\]

**Proof of Theorem 4.3:**

Consider the least squares equation centred at \( E(I^{(T)}(\frac{2\pi}{T}s)) \),

\[
L^{(T)}(\theta) = \left( \frac{2\pi}{T} \right)^2 \sum_{s_j=1}^{T-1} \left( I^{(T)}(\frac{2\pi}{T}s) - f^{(T)}(\frac{2\pi}{T}s; \theta) \right)^2.
\]

By Taylor's Theorem,

\[
\sqrt{T^2} \left[ \frac{\hat{\theta}_1 - \theta_{1,0}}{\hat{\theta}_2 - \theta_{2,0}} \right] = 2 \left[ \begin{array}{cc}
\frac{\partial^2}{\partial \theta_1 \partial \theta_2} L^{(T)}(\theta^*) & \frac{\partial^2}{\partial \theta_1 \partial \theta_2} L^{(T)}(\theta^*) \\
\frac{\partial^2}{\partial \theta_2 \partial \theta_1} L^{(T)}(\theta^*) & \frac{\partial^2}{\partial \theta_2 \partial \theta_2} L^{(T)}(\theta^*)
\end{array} \right]^{-1} \left[ \begin{array}{c}
- \frac{T}{2} \frac{\partial}{\partial \theta_1} L^{(T)}(\theta_0) \\
- \frac{T}{2} \frac{\partial}{\partial \theta_2} L^{(T)}(\theta_0)
\end{array} \right]
\]

(4.8)

where \( \theta^* \) lies between \( \hat{\theta} \) and \( \theta_0 \). Denote the matrix above by \( A^{(T)} \) where

\[
\frac{1}{2} \frac{\partial^2}{\partial \theta_k \partial \theta_l} L^{(T)}(\theta^*) = \left( \frac{2\pi}{T} \right)^2 \sum_{s_j=1}^{T-1} \left( \left( \frac{\partial}{\partial \theta_k} f^{(T)}(\frac{2\pi}{T}s; \theta^*) \right) \frac{\partial}{\partial \theta_l} f^{(T)}(\frac{2\pi}{T}s; \theta^*) \right) + \left( f^{(T)}(\frac{2\pi}{T}s; \theta^*) - I^{(T)}(\frac{2\pi}{T}s; \theta_0) \right) \frac{\partial^2}{\partial \theta_k \partial \theta_l} f^{(T)}(\frac{2\pi}{T}s; \theta^*).
\]

Also

\[
- \frac{T}{2} \frac{\partial}{\partial \theta_k} L^{(T)}(\theta_0) = T \left( \frac{2\pi}{T} \right)^2 \sum_{s_j=1}^{T-1} \left( I^{(T)}(\frac{2\pi}{T}s) - f^{(T)}(\frac{2\pi}{T}s; \theta_0) \right) \frac{\partial}{\partial \theta_k} f^{(T)}(\frac{2\pi}{T}s; \theta_0).
\]
From (4.8),
\[
T \begin{bmatrix} X - \mu_0 \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A^{(T)} \end{bmatrix} T \begin{bmatrix} X - \mu_0 \\ - \frac{1}{2} \frac{\partial}{\partial \theta_1} L^{(T)}(\theta_0) \\ - \frac{1}{2} \frac{\partial}{\partial \theta_2} L^{(T)}(\theta_0) \end{bmatrix}.
\]

Denote the matrix above by $B^{(T)}$. From Theorem 4.1,
\[
T \begin{bmatrix} X - \mu_0 \\ - \frac{1}{2} \frac{\partial}{\partial \theta_1} L^{(T)}(\theta_0) \\ - \frac{1}{2} \frac{\partial}{\partial \theta_2} L^{(T)}(\theta_0) \end{bmatrix} \overset{D}{\sim} N_3(0, \Sigma).
\]

Also, the consistency of $\hat{\theta}$ and Theorem 4.1 imply that $B^{(T)} \overset{L}{\rightarrow} B$ where
\[
B = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A^{-1} \end{bmatrix}
\]

and the $(k, \ell)$th element of $A$ is of the form
\[
\int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\lambda; \theta_0) \frac{\partial}{\partial \theta_\ell} f(\lambda; \theta_0) d\lambda.
\]

This follows because
\[
\frac{1}{2} \frac{\partial^2}{\partial \theta_k \partial \theta_\ell} L^{(T)}(\theta^*) = \frac{2\pi}{T} (\frac{2\pi}{T})^2 \sum_{\ell=1}^{T-1} \frac{\partial}{\partial \theta_k} f^{(T)}(\frac{2\pi}{T} \cdot \theta^*) \frac{\partial}{\partial \theta_\ell} f^{(T)}(\frac{2\pi}{T} \cdot \theta^*)
\]
\[
+ (\frac{2\pi}{T})^2 \sum_{\ell=1}^{T-1} (f^{(T)}(\frac{2\pi}{T} \cdot \theta^*) - f^{(T)}(\frac{2\pi}{T} \cdot \theta^*)) (\frac{\partial^2}{\partial \theta_k \partial \theta_\ell} f^{(T)}(\frac{2\pi}{T} \cdot \theta^*)
\]
\[
= \int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\lambda; \theta_0) \frac{\partial}{\partial \theta_\ell} f(\lambda; \theta_0) d\lambda + o_p(1).
\]

Therefore, by Slutsky's Theorem,
\[
T \begin{bmatrix} X - \mu_0 \\ \hat{\theta}_1 - \theta_{1,0} \\ \hat{\theta}_2 - \theta_{2,0} \end{bmatrix} \overset{D}{\sim} N_3(0, BB')
\]

**Proof of Theorem 4.4:**
By Taylor's Theorem
\[
T(\hat{R} - R_0) = \left[ \frac{\partial}{\partial \mu} R, \frac{\partial}{\partial \theta_1} R, \frac{\partial}{\partial \theta_2} R \right]_{(\Theta^*, \mu^*)} T \left[ \frac{X - \mu_0}{\theta_1 - \theta_{1,0}}, \frac{X - \mu_0}{\theta_2 - \theta_{2,0}} \right]
\]
where \((\Theta^*, \mu^*)\) lies between \((\hat{\Theta}, X)\) and \((\Theta_0, \mu_0)\). Thus, by the consistency of \((\hat{\Theta}, \bar{X})\), Slutsky's Theorem and Theorem 4.3,
\[
T(\hat{R} - R_0) \xrightarrow{D} N(0, r'_0 B \Sigma B' r_0)
\]
where
\[
r'_0 = \left[ \frac{\partial}{\partial \mu} R, \frac{\partial}{\partial \theta_1} R, \frac{\partial}{\partial \theta_2} R \right]_{(\mu_0, \Theta_0)}.
\]
Theorems 4.7 and 4.8 follow from identical arguments to those above.

### 4.7 Cumulant Calculations

The limiting covariance matrices \(\Sigma\) and \(\Gamma\) from Theorems 4.1 and 4.5 involve integrals of spectral densities from second- to fourth-order. Expressions for these are now obtained. The spectral density
\[
f(\lambda; \Theta) = \frac{1}{(2\pi)^2} \sum_{a,j=-\infty}^{\infty} c_X(a; \Theta) e^{-i<\lambda,a>}
\]
\[
= \gamma^2 + \frac{1}{(2\pi)^2} \sum_{a,j=-\infty}^{\infty} c(a; \Theta) e^{-i<\lambda,a>}, \quad \lambda, \Theta \in \mathbb{R}^2, \quad j = 1, 2
\]
where
\[
c(a; \Theta) = \text{cov}(Z(0), Z(a)) \]
\[
= \theta_1 \int_{\mathbb{R}^2} e^{-\frac{1}{\theta_2}(x^2+(x-a_1)^2+y^2+(y-a_2)^2)}
\]
\[
= \frac{\pi \theta_1 \theta_2^2}{2} e^{-\frac{1}{\theta_2}(a_1^2+a_2^2)}
\]
by the same argument as that used in Section 3.6, Chapter 3.

Also,
\[ \frac{\partial}{\partial \theta_1} f(\lambda; \theta) = \frac{1}{(2\pi)^2} \sum_{a_j = -\infty}^{\infty} \frac{\partial}{\partial \theta_1} c(a; \theta)e^{-i\langle \lambda, a \rangle} \]

where
\[ \frac{\partial}{\partial \theta_1} c(a; \theta) = \frac{\pi \theta_1^2}{2} e^{-\frac{\theta_1^2}{2\theta_2^2}(a_1^2 + a_2^2)} . \]

and
\[ \frac{\partial}{\partial \theta_2} f(\lambda; \theta) = \frac{1}{(2\pi)^2} \sum_{a_j = -\infty}^{\infty} \frac{\partial}{\partial \theta_2} c(a; \theta)e^{-i\langle \lambda, a \rangle} \]

where
\[ \frac{\partial}{\partial \theta_2} c(a; \theta) = \pi \theta_1 \theta_2 e^{-\frac{\theta_1^2}{2\theta_2^2}(a_1^2 + a_2^2)} \left(1 + \frac{\delta^2(a_1^2 + a_2^2)}{2\theta_2^2}\right) . \]

The third-order cumulant spectrum is
\[ f_3(\lambda^{(1)}, \lambda^{(2)}) = \frac{1}{(2\pi)^4} \sum_{a_j = -\infty}^{\infty} \sum_{b_j = -\infty}^{\infty} c_3(a, b)e^{-i\langle \lambda^{(1)}, a \rangle + \langle \lambda^{(2)}, b \rangle} \]

where
\[ c_3(a, b) = \text{cum}_3\{X(\theta), X(a), X(b)\} . \]

In general,
\[
\text{cum}_3\{X(a), X(b), X(c)\} = (-1)^2 \cdot 2! \{E(X(a))E(X(b))E(X(c))
+ (-1)^1 \cdot 1! \{E(X(a))E(X(b))E(X(c))
+ E(X(a)X(c))E(X(b))
+ E(X(b)X(c))E(X(a))\}
+ E(X(a)X(b)X(c))\} .
\]
\[ = 2\mu^3 - \mu \{ E(X(a)X(b)) + E(X(a)X(c)) + E(X(b)X(c)) \} \\
+ E(X(a)X(b)X(c)) \]
\[ = 2\mu^3 - \mu \left\{ c(a - b) + \mu^2 + E(\epsilon(a)\epsilon(b)) \\
+ c(a - c) + \mu^2 + E(\epsilon(a)\epsilon(c)) \\
+ c(b - c) + \mu^2 + E(\epsilon(b)\epsilon(c)) \right\} \\
+ \lambda_{\mu_3Y} \frac{\pi w^2}{3} \exp \left[ -\frac{\delta^2}{w^2}(<a, a> + <b, b> + <c, c> \\
- \frac{1}{3} <a + b + c, a + b + c>) \right] \\
+ \mu \{ c(a - c); c(b - c) + c(a - b) \} + \mu^3 \\
+ \mu \epsilon^\gamma(\epsilon(a)\epsilon(b)) + E(\epsilon(a)\epsilon(c))E(\epsilon(b)\epsilon(c)) \}
\]
\[ = \lambda_{\mu_3Y} \frac{\pi w^2}{3} \exp \left[ -\frac{\delta^2}{w^2}(<a, a> + <b, b> + <c, c> \\
- \frac{1}{3} <a + b + c, a + b + c>) \right] \\
+ \mu \epsilon^\gamma(\epsilon(a)\epsilon(b)) + E(\epsilon(a)\epsilon(c))E(\epsilon(b)\epsilon(c)) \}
\]

Therefore,
\[
c_3(a, b) = \frac{\pi \lambda_{\mu_3Y} w^2}{3} \exp \left[ -\frac{\delta^2}{w^2}(<a, a> + <b, b> \\
- \frac{1}{3} <a + b, a + b>) \right].
\]

The above result uses the fact that
\[
E(X(a)X(b)X(c)) \\
= E(Z(a)Z(b)Z(c)) + \mu(E(\epsilon(a)\epsilon(b)) \\
+ E(\epsilon(a)\epsilon(c)) + E(\epsilon(b)\epsilon(c))) \\
+ E(\epsilon(a)\epsilon(b)\epsilon(c)) .
\]
where

\[
E(Z(a)Z(b)Z(c))
= \frac{\pi \lambda \mu_{3y} w^2}{3} \exp \left[ - \frac{b^2}{w^2} (\langle a \cdot a \rangle + \langle b \cdot b \rangle + \langle c \cdot c \rangle - \frac{1}{3} \langle a + b + c, a + b + c \rangle) \right] + \mu (c(a - c) + c(b - c) + c(a - b)) + \mu^3
\]

The corresponding expression for the fourth-order product moment will be required in the following discussion.

\[
E(X(a)X(b)X(c)X(d))
= E(Z(a)Z(b)Z(c)Z(d)) + E(Z(a)Z(b))E(\epsilon(c)\epsilon(d))
+ E(Z(a)Z(c))E(\epsilon(b)\epsilon(d)) + E(Z(a)Z(d))E(\epsilon(b)\epsilon(c))
+ E(Z(b)Z(c))E(\epsilon(a)\epsilon(d)) + E(Z(b)Z(d))E(\epsilon(a)\epsilon(c))
+ E(Z(c)Z(d))E(\epsilon(a)\epsilon(b))
+ \mu (E(\epsilon(a)\epsilon(b)\epsilon(c)) + E(\epsilon(a)\epsilon(b)\epsilon(d)))
+ E(\epsilon(b)\epsilon(c)\epsilon(d)) + E(\epsilon(a)\epsilon(c)\epsilon(d))
+ E(\epsilon(a)\epsilon(b)\epsilon(c)\epsilon(d))
\]

where

\[
E(Z(a)Z(b)Z(c)Z(d))
= \frac{\pi \lambda \mu_{4y} w^2}{4} \exp \left[ - \frac{b^2}{w^2} (\langle a \cdot a \rangle + \langle b \cdot b \rangle + \langle c \cdot c \rangle - \langle d \cdot d \rangle) \right]
\]
\[ -\frac{1}{4} \langle a + b + c + d, a + b + c + d \rangle \]
\[ + \mu (c^{(3)}(a, b, c) + c^{(3)}(b, c, d) + c^{(3)}(a, b, d)) \]
\[ + c(a - b)c(c - d) + c(a - c)c(b - d) + c(a - d)c(b - c) \]
\[ + \mu^2 (c(a - b) + c(a - c) + c(a - d) + c(b - c) \]
\[ + c(b - d) + c(c - d)) + \mu^4, \]

and

\[ c^{(3)}(a, b, c) \]
\[ = \frac{\pi \lambda \mu_3 Y w^2}{3} \exp \left[ -\frac{\delta^2}{w^2} (\langle a, a \rangle + \langle b, b \rangle + \langle c, c \rangle \right. \]
\[ \left. - \frac{1}{3} \langle a + b + c, a + b + c \rangle \right] \]
\[ = \text{cum}_3\{ X(a), X(b), X(c) \} \]

from before.

Now then,

\[ f_4(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}) = \frac{1}{(2\pi)^6} \sum_{a_j=-\infty}^{\infty} \sum_{b_j=-\infty}^{\infty} \sum_{d_j=-\infty}^{\infty} c_4(a, b, d) \exp(-\langle \lambda^{(1)}, a \rangle + \langle \lambda^{(2)}, b \rangle + \langle \lambda^{(3)}, d \rangle) \]

where

\[ c_4(a, b, d) = \text{cum}_4\{ X(0), X(a), X(b), X(d) \} \]

In general,

\[ \text{cum}_4\{ X(a), X(b), X(c), X(d) \} \]
\[ = -6\mu^4 \]
\[ + 2\mu^2 \{ E(X(a)X(b)) + E(X(b)X(c)) \]
\[ + E(X(a)X(c)) + E(X(a)X(d)) \\
+ E(X(b)X(d)) + E(X(c)X(d)) \]

\[ - \{ E(X(a)X(b))E(X(c)X(d)) \]

\[ + E(X(a)X(c))E(X(b)X(d)) \\
+ E(X(a)X(d))E(X(b)X(c)) \}

\[ - \mu \{ E(X(a)X(b)X(c)) + E(X(a)X(b)X(d)) \]

\[ E(X(a)X(c)X(d)) + E(X(b)X(c)X(d)) \}

\[ + E(X(a)X(b)X(c)X(d)) \]

\[ = -6\mu^4 \]

\[ + 2\mu^2 \{ c(a-b) + \mu^2 + E(\epsilon(a)\epsilon(b)) + c(b-c) + \mu^2 + E(\epsilon(b)\epsilon(c)) \\
+ c(a-c) + \mu^2 + E(\epsilon(a)\epsilon(c)) + c(a-d) + \mu^2 + E(\epsilon(a)\epsilon(d)) \\
+ c(b-d) + \mu^2 + E(\epsilon(b)\epsilon(d)) + c(c-d) + \mu^2 + E(\epsilon(c)\epsilon(d)) \}\]

\[ - \{ (c(a-b) + \mu^2 + E(\epsilon(a)\epsilon(b))) (c(c-d) + \mu^2 + E(\epsilon(c)\epsilon(d))) \\
+ (c(a-c) + \mu^2 + E(\epsilon(a)\epsilon(c))) (c(b-d) + \mu^2 + E(\epsilon(b)\epsilon(d))) \\
+ (c(a-d) + \mu^2 + E(\epsilon(a)\epsilon(d))) (c(b-c) + \mu^2 + E(\epsilon(b)\epsilon(c))) \}\]

\[ - \mu \{ E(Z(a)Z(b)Z(c)) + \mu(E(\epsilon(a)\epsilon(b)) + E(\epsilon(a)\epsilon(c)) \\
+ E(\epsilon(b)\epsilon(c)) + E(\epsilon(a)\epsilon(b)\epsilon(c)) \\
+ E(Z(a)Z(b)Z(d)) + \mu(E(\epsilon(a)\epsilon(b))) \\
+ E(\epsilon(a)\epsilon(d)) + E(\epsilon(b)\epsilon(d)) + E(\epsilon(a)\epsilon(b)\epsilon(d)) \\
+ E(Z(a)Z(c)Z(d)) + \mu(E(\epsilon(a)\epsilon(c))) \\
+ E(\epsilon(a)\epsilon(d)) + E(\epsilon(c)\epsilon(d)) + E(\epsilon(a)\epsilon(c)\epsilon(d)) \} \]
\begin{align*}
+ E(Z(b)Z(c)Z(d)) + \mu(E(c(b)c(c)) \\
+ E(c(b)c(d)) + E(c(c)c(d))) + E(c(b)c(c)(d))\} \\
+ E(Z(a)Z(b)Z(c)Z(d)) + (c(a - b) + \mu^2)E(c(c)c(d)) \\
+ (c(a - c) + \mu^2)E(c(b)c(d)) + (c(a - d) + \mu^2)E(c(b)c(c)) \\
+ \mu E(c(b)c(c)(d)) + (c(b - c) + \mu^2)E(c(a)c(d)) \\
+ (c(b - d) + \mu^2)E(c(a)c(c)) + \mu E(c(a)c(c)(d)) \\
+ (c(c - d) + \mu^2)E(c(a)c(b)) + \mu E(c(a)c(b)c(d)) \\
+ \mu E(c(a)c(b)c(c)) + E(c(a)c(b)c(c)(d)).
\end{align*}

After some algebraic cancellation,

\[
\text{cum}_4\{X(a), X(b), X(c), X(d)\} = \frac{\pi \lambda \mu_4 w^2}{4} \exp \left[ -\frac{\delta^2}{w^2} (\langle a, a \rangle + \langle b, b \rangle + \langle c, c \rangle + \langle d, d \rangle \\
- \frac{1}{4} \langle a + b + c + d, a + b + c + d \rangle) \right] \\
+ E(c(a)c(b)c(c)c(d)) \\
- (E(c(a)c(b))E(c(c)c(d)) + E(c(a)c(c))E(c(b)c(d)) \\
+ E(c(a)c(d))E(c(b)c(c))).
\]

4.8 Summary

This chapter presents a filtered Poisson process model of Image Correlation Spectroscopy. The stationary random field is an extension of the process considered in Chapter 3, and the principal model parameter is the ratio \( R \). Two different estimation methodologies are considered, nonlinear least squares and quasi-likelihood,
both involving the periodogram and spectral density of the random field \(X(t)\). The joint asymptotic normality of these least squares and likelihood estimators is obtained. The estimator \(\hat{R}\) is then shown to also have a limiting normal distribution, allowing the construction of asymptotic confidence intervals. Theorems 4.1 and 4.5 are based upon Theorem 2.6 [27]. Rosenblatt discusses how one would establish the result but the details are not provided and do not appear to be readily available in the literature. Of principal interest then is the proof of Theorems 4.1 and 4.5 and the remarks which consequently follow relating to the cumulant method of proof and asymptotic bias. The complete algebraic details involved with the derivation of the limiting covariance matrices are also provided. The calculation of these expressions is considered in Chapter 5.
Chapter 5
Massively Parallel Computation

5.1 MasPar MP-2

The ICS calculations from Chapter 4 are performed on a MasPar MP-2 parallel computer. This machine combines a massively parallel computer with an advanced data parallel software development environment. The following data parallel terminology is required in the following discussion:

(i) PROCESSOR ELEMENTS:

Parallel computers have multiple processors in addition to a CPU, called processor elements or PE’s.

(ii) PARALLEL COMPUTER:

This is combining processors together to solve certain problems in seconds that might take hours on a serial (sequential) computer.

(iii) MASSIVELY PARALLEL:

Parallel computing systems that typically have more than 1000 processors.
(iv) SIMD:

This is an acronym for a Single Instruction, Multiple Data parallel architecture. These systems are used when there is a large data set to be processed and a single processor for each item (or sub-collection of items) in the data set. The same set of instructions is applied to all the items in the data set.

(v) DATA PARALLEL COMPUTING:

A programming style used on SIMD massively parallel computers.

(vi) DPU:

An acronym for Data Parallel Unit, that part of the MasPar that does all the parallel processing. This unit contains the PE's, the Array Control Unit (ACU) and various communication mechanisms.

(vii) ACU:

The Array Control Unit, a processor which has its own data and instruction memory. The ACU controls the PE's, performs operations on scalar data and communicates with the front end.

(viii) PE ARRAY

The two-dimensional array of processor elements.

(ix) PLURAL VARIABLE:

Parallel variables which are replicated on each PE.
(x) SINGULAR VARIABLE

Ordinary variables which reside on the ACU or front end, not on the PE ARRAY.

5.2 MasPar Hardware

The hardware consists of a front end DEC graphics workstation and the DPU. It is depicted in Figure 5.2.1. A two-dimensional representation of the MP-2 architecture is given in Figure 5.2.2.

5.3 MasPar Software

The software system includes the high-level ULTRIX operating system on the front end, high performance programming languages, an interactive programming and debugging environment called MPPE as well as other tools and libraries. The MasPar programming language is known as MPL and is ANSI C based (a version of Fortran 90 is also available). MPL is a low level language which provides programmers with direct access to the DPU and hence the PE array, and requires a good understanding of the DPU’s parallel architecture.

5.4 Introduction

The MasPar MP-2 computer at the University of Western Ontario has a front end DEC’station 5000 ULTRIX workstation and a PE array of 64(nxproc) by 32(nyproc) elemental processors. It is referred to as a 2K machine (due to its 2048 processors). One may take (without loss of generality) the southwest corner
Figure 5.2.1: MasPar MP-2 architecture
Figure 5.2.2: Two-dimensional representation of MasPar PE Array
of the PE array to be PE(0,0), although in some settings it may be more convenient to take PE(0,0) to be the northwest corner, corresponding geometrically to an ordinary array. Its neighbour to the east is PE(0,1) and so on leading to PE(0,63). Similarly, the neighbour to the north of PE(0,0) is PE(1,0) leading to PE(31,0). An arbitrary PE may be indexed by PE(iyproc, ixproc) where iyproc = 0, ..., 63 and ixproc = 0, ..., 31, that is iyproc would correspond to the row and ixproc to the column of the PE array. In this way, the entire PE array is indexed so that one might refer to individual processor elements as one might elements of a two-dimensional array or matrix. The choice of the reference PE(0,0) is arbitrary. The southwest corner is chosen so that the PE array can be considered to correspond to the first quadrant of the Cartesian plane. In this way, one can topologize the PE array itself to have the same structure as the marked Poisson process region and subsequent random field. One may also consider the array to be some subregion of the frequency domain when objects such as periodograms and spectral densities are examined.

5.5 Calculation of Limiting Covariances

Recall that Theorem 4.1 of Chapter 4 involved the limiting covariance matrix $\Sigma = [\Sigma_{ij}]$ where

$$\Sigma_{0,0} = \gamma^2 + \sum_k c(k; \theta_0),$$

$$\Sigma_{0,j} = (2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0)f_3(0, \alpha)d\alpha, \quad j = 1, 2$$
and for \( j, k = 1, 2 \),

\[
\Sigma_{k,j} = 2(2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f^2(\alpha; \theta_0) d\alpha \\
+ (2\pi)^2 \int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_j} f(\beta; \theta_0) f_4(\alpha, \beta, -\alpha) d\alpha d\beta.
\]

This section discusses the computation of these integrals. The method of calculations carries over immediately to the limiting expressions of Section 4.5.

### 5.5.1 Riemann Sum Approximation

The calculation of \( \Sigma_{0,0} \) is straightforward and proceeds by evaluating

\[
\gamma^2 + \sum_{k_j = -m}^m c(k; \theta_0)
\]

where \( m \) is the finite-range parameter equal to the beam diameter divided by \( \delta \) minus one (i.e. \( m = \frac{\text{diameter}}{\delta} - 1 \)). This is a singular operation performed on the ACU.

Next, the double integral

\[
\int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f_3(0, \alpha) d\alpha
\]

is approximated. A reasonable first approximation would seem to be an ordinary Riemann sum,

\[
(\frac{2\pi}{T})^2 \sum_{s_j = 1}^{T-1} \frac{\partial}{\partial \theta_k} f(\frac{2\pi}{T} s; \theta_0) f_3(0, \frac{2\pi}{T} s).
\]

In this context then, the PE array is considered to correspond to the 512 \( \times \) 512 lattice of discrete Fourier frequencies \([\frac{2\pi}{T}(s_1, s_2)]\) where \( s_1, s_2 = 0, \ldots, 511 \). For example, on PE(0,0) one would have the sub-array of frequencies \([\frac{2\pi}{T}(s_1, s_2)]\), \( s_1 = 0, \ldots, 7, \ s_2 = 0, \ldots, 15 \) (recalling that \( nxproc = 64 \) and \( nyproc = 32 \)) and in
general PE \((iyproc,ixproc)\) would contain \(\left[\frac{2\pi}{T}(s_1, s_2)\right]\). \(s_1 = ixproc \times 8 + i. i = 0, \ldots, 7,\) and \(s_2 = iyproc \times 16 + j, j = 0, \ldots, 15.\) One must then evaluate the derivatives of the spectral density and the third-order cumulant spectrum at these discrete frequencies. The derivatives are calculated by

\[
\frac{\partial}{\partial \theta_k} f\left(\frac{2\pi}{T} s; \theta_0\right) = \frac{1}{(2\pi)^2} \sum_{a_j = -m}^{m} \frac{\partial}{\partial \theta_k} c(\alpha; \theta_0) e^{-i <\frac{2\pi}{T} s, \alpha>}
\]

and these values are then stored in plural arrays on each PE corresponding to the sub-array of frequencies for that PE. The same situation exists for

\[
f_3\left(0, \frac{2\pi}{T} s\right) = \frac{1}{(2\pi)^4} \sum_{a_j = -m}^{m} \sum_{b_j = -m}^{m} c_3(\alpha, b) e^{-i <\frac{2\pi}{T} s, b>}
\]

One then forms the appropriate \(\frac{\partial f(\cdot)}{\partial \theta_k}\) and \(f_3(\cdot)\) products and sums over the frequencies \(\left[\frac{2\pi}{T} s\right]_{8 \times 16}\) on each PE. The resulting plural sums are then added together using the MPL function \(reduceAddf(\ )\), and when multiplied by \((\frac{2\pi}{T})^2\), provide estimates of the spectral integrals.

The same algorithm can be applied to approximate

\[
\int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f^2(\alpha; \theta_0) d\alpha
\]

\[\quad + \int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_j} f(\beta; \theta_0) f_4(\alpha, \beta, -\alpha) d\alpha d\beta
\]

where the fourth-order spectral density is evaluated as

\[
f_4\left(\frac{2\pi}{T} s, \frac{2\pi}{T} s, -\frac{2\pi}{T} s\right) = \frac{1}{(2\pi)^6} \sum_{a_j = -m}^{m} \sum_{b_j = -m}^{m} \sum_{c_j = -m}^{m} c_4(a, b, c) e^{-i <\frac{2\pi}{T} s, \alpha> + <\frac{2\pi}{T} s, b> - <\frac{2\pi}{T} s, c>}
\]

Note that the Riemann sum approximation to the second integral in (5.1) is of the
form
\[ \left( \frac{2\pi}{T} \right)^4 \sum_{r_2=1}^{T-1} \sum_{s_2=1}^{T-1} \frac{\partial}{\partial \theta_k} f \left( \frac{2\pi}{T} r_2; \theta_0 \right) \frac{\partial}{\partial \theta_l} f \left( \frac{2\pi}{T} s_2; \theta_0 \right) f_4 \left( \frac{2\pi}{T} r_2, \frac{2\pi}{T} s_2, \frac{2\pi}{T} r_1, \frac{2\pi}{T} s_1 \right) \] (5.2)
which is \( O(T^4) \). One must evaluate the spectral density derivatives and \( f_4(\cdot) \) over four-dimensional sub-arrays of frequencies \( \left[ \frac{2\pi}{T} (r, s) \right] \). \( r_1, s_1 = 0, \ldots, 7 \) and \( r_2, s_2 = 0, \ldots, 15 \), on each PE.

The parallel summing operation involves nested for loops involving ten for statements. So even though (5.2) is evaluated in parallel, the computation becomes quickly infeasible with DPU times measured in terms of days (approximately 7 days for the parameter settings considered here). Not surprisingly, the length of time required for these calculations is highly dependent upon the value of the finite-range parameter \( m \).

**Remark 5.1** Truncating the spectral densities at lags \( +/- m \) may lead to negative spectral density values as the frequencies move radially away from frequencies in the neighborhood of zero (i.e. from PE(0,0) in the southwest corner of the array). As a result, the upper index of summation for the Riemann sum (5.2) is chosen to be some integer \( L < T - 1 \). This simply implies that in the least squares (or likelihood) equation, the summation is over the discrete frequencies \( \frac{2\pi}{T} (s_1, s_2) \). \( s_1, s_2 = 1, \ldots, L \) (or equivalently that the integrals are over the region \( [0, \lambda_0]^2 \) for some \( \lambda_0 < 2\pi \)). All asymptotic results remain valid. To perform this summation in parallel requires a conditional statement of the form if(ixproc < m && iyproc < n) for suitably chosen integers \( m \) and \( n \) depending upon \( L \).

The remark suggests that in practice, one cannot utilize the entire PE array to
compute these Riemann sums. This fact, coupled with the already unacceptable execution times suggests that another approach is required.

5.5.2 Monte Carlo Integration

The MasPar's architecture is ideal for approximating these spectral integrals by simple Monte Carlo integration. The basic idea is that for \( \mathbf{x}_1, \ldots, \mathbf{x}_N \) i.i.d. \( U(V) \) in the multidimensional volume \( V \), by the Strong Law of Large Numbers,

\[
\int f dV \approx \frac{|V|}{N} \sum_{i=1}^{N} f(\mathbf{x}_i) \pm |V| \sqrt{\frac{<f^2>-<f>^2}{N}}
\]

where \( <f>=\frac{1}{N} \sum_{i=1}^{N} f(\mathbf{x}_i) \) and \( <f^2>=\frac{1}{N} \sum_{i=1}^{N} f^2(\mathbf{x}_i) \). For example,

\[
\int_{0}^{\lambda_0} f(\mathbf{x}: \theta_0) d\alpha \approx \frac{\lambda_0^2}{N} \sum_{i=1}^{N} f(\mathbf{x}_i; \theta_0), \quad \lambda_0 \leq 2\pi
\]

where \( f(\cdot) \) is the spectral density function. As with the Riemann sums, the value of \( m \) dictates the choice of \( \lambda_0 \) due to possible truncation error.

The limiting variance of \( X \) (\( \Sigma_{0,0} \)) can be calculated in the same fashion as before.

The Monte Carlo algorithm requires uniform random variates. The pseudo-random number generator [25] is converted to a plural float function, i.e. each PE generates different \( U(0,1) \) random variables. To seed the generator, the MPI plural function \( p\text{-random}() \) is used to obtain three random integer seeds (\( p\text{-random}() \) returns pseudo-random integers in the range from 1 to \( 2^{31}-1 \); these are converted to integers in the range 1 to 30,000). The plural function \( sp\text{-srandom}(time(0)) \) is used to seed \( p\text{-random} \).
To approximate the double integrals

\[ \int_0^{\lambda_0} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f_3(0, \alpha) d\alpha \]

and

\[ \int_0^{\lambda_0} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) f^2(\alpha; \theta_0) d\alpha \]

then, one generates a \( U([0, \lambda_0]^2) \) random variate \((u_1, u_2)\) on each PE, evaluates the corresponding integrand and then uses the \textit{reduce Addf(\cdot)} function to sum the results over all PE's. The result, when multiplied by \( \lambda_0^2 \) and divided by \( \text{nproc} \) (an MPL constant for the number of processors, 2018), is the Monte Carlo integral based on a sample of size 2048.

The same algorithm applies to the approximation of the four-dimensional integral in (5.1), except now a \( U([0, \lambda_0]^4) \) random variate \((u_1, u_2, u_3, u_4)\) is generated and \( \frac{\partial}{\partial \theta_j} f(u_1, u_2; \theta_0) \) and \( f_4(u_1, u_2, u_3, u_4, -u_1, -u_2) \) are evaluated. As before, one can use \textit{reduce Addf(\ )} and, after dividing the result by 2048, multiply by \( \lambda_0^4 \) to obtain the approximation.

Recall that the matrix A Theorem 4.3. Chapter 4 also involves spectral integrals, of the form

\[ \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_k} f(\alpha; \theta_0) d\alpha . \]

These are also approximated by the Monte Carlo technique.

The time required for these calculations is a substantial improvement over that for the Riemann sum approximation, although DPU time is still highly dependent upon the value of \( m \). The evaluation of the fourth-order spectral density remains an expensive computation. This is why only one uniform random variate is typically
generated on each PE whereas in principle, one could obtain an arbitrarily large sample size $N$ simply by generating a suitably large number $N_p$ of uniform variates on each PE (where $N = N_p \times 2048$).

**Numerical Results**

A numerical comparison of the least squares and likelihood procedures is carried out. The comparison is based on the calculation of the limiting variance of $\hat{R}$ arising from each method. The results are summarized in the table below for specific parameter settings corresponding closely to those in actual ICS experiments. Average DPU times and approximate standard errors of the Monte Carlo method are also included.

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameters</th>
<th>Standard Dev. ($\hat{R}$)</th>
<th>Standard Error</th>
<th>DPU Time/mins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least Squares</td>
<td>$\lambda$</td>
<td>$\mu_y$</td>
<td>$w$</td>
<td>$\delta$</td>
</tr>
<tr>
<td>Likelihood</td>
<td>1.0</td>
<td>4.0</td>
<td>0.15</td>
<td>0.10</td>
</tr>
</tbody>
</table>

The parameter settings chosen for this study correspond closely to those observed or measured in ICS experiments. The values for the standard deviation of $\hat{R}$ are not significantly different although the likelihood value is calculated somewhat more precisely. It should be noted that the observed DPU times are highly dependent upon the process parameters, specifically the laser beam width parameter $w$ and the resolution $\delta$. For example, as $\delta$ decreases, the finite-range parameter $m$ increases leading to unreasonable DPU times. Notwithstanding, this numerical comparison is both interesting and useful. Direct comparisons of these estimation methodologies (in the random field setting) where the comparison is based upon asymptotic variances are not common in the random field literature due to the
computational complexity of the variance expressions. Although either method can be used, the simulation results in Section 5.6 pertain to the least squares fitting procedure which is used in practice to analyze ICS image data as described in Sections 1.2 and 1.3.

One can also examine the effect, upon the limiting variance of $\hat{R}$, of incorrectly assuming that the observed fluorescence intensities follow a Gaussian process. This assumption would fail to capture the necessary nonnegativity of the fluorescence intensity measurements. Also, certain protein cluster parametric information (via the marked Poisson process parameters) would be unavailable. Mathematically, the integrals involving third- and fourth-order spectral densities in Theorems 4.1 and 4.5 would no longer appear. Numerical results are presented in the following table.

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameters</th>
<th>Standard Dev. ($\hat{R}$)</th>
<th>Standard Error</th>
<th>DPU Time/secs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least Squares</td>
<td>$1.0$</td>
<td>$4.0$</td>
<td>$0.15$</td>
<td>$0.10$</td>
</tr>
<tr>
<td>Likelihood</td>
<td>$1.0$</td>
<td>$4.0$</td>
<td>$0.15$</td>
<td>$0.10$</td>
</tr>
</tbody>
</table>

The results show that although there is considerable improvement in DPU time, there is also a significant loss of precision in estimating $R$ with standard deviations of $\hat{R}$ exceeding the correct model values by approximately 40 percent.

For the ICS process, since there is no closed form for the cumulant spectra, it is not known a priori if the integrals involving the $f_3(\cdot)$ and $f_4(\cdot)$, and the derivatives of $f(\cdot)$, will be positive or negative. Numerical evaluations seem to indicate that the $f_3(\cdot)$ integrals are negative for the parameter values considered here.

In contrast, for an ARMA process, this Gaussian approximation yields asymptotic variances which are too small as can be seen from an examination of the
limiting variance in [17]

It is also interesting to compare the performance of the MPL Monte Carlo algorithm with its C language analogue. The DECstation 5000 running ULTRIX C is used to provide one such comparison. The performance differences (summarized below) are indicative of the advantages the parallel algorithm holds over the corresponding sequential algorithm, in this setting. The least squares variance expressions are used in the comparison with identical parameter settings.

<table>
<thead>
<tr>
<th>Machine</th>
<th>Language</th>
<th>Sample Size (N)</th>
<th>DPU/CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECstation</td>
<td>C</td>
<td>2048</td>
<td>13.67 hrs (\equiv) 820 mins</td>
</tr>
<tr>
<td>MasPar MP-2</td>
<td>MPL</td>
<td>2048</td>
<td>17 mins</td>
</tr>
</tbody>
</table>

### 5.6 Simulation and Estimation

The Poisson-driven random field corresponding to the observed data of ICS is simulated with parameter values that reasonably correspond to those used (or observed) in practice. The least squares estimation method of Section 4.4 is then used to estimate the model parameters based on the simulated data. The estimates obtained from repeated simulations can be used to justify the asymptotics of the least squares estimators as discussed in Section 4.4. Two different methods of estimating \(\Sigma\) (Theorem 4.1) based on the simulated data are implemented: the block method and the Taniguchi-Keenan method. These topics are discussed in more detail in the following sections.
5.6.1 Simulation of Poisson-Driven Random Field

The spatial marked Poisson process is simulated on the ACU by a singular (or sequential) algorithm using [35]. To do this in parallel requires substantial nearest-neighbour PE to PE communication (via the DPU communication network known as X-Net) due to the fact that $X(t)$ is a random sum over all points within some radius of a disc centred at $t\delta$ and the resulting random field is finite-range dependent (i.e. neighbouring discs overlap). This added complexity does not seem to justify the savings in overall run time.

The area under consideration is a square with dimensions $(511\delta + 2r) \times (511\delta + 2r)$ where $\delta$ is the stepsize parameter and $r$ is the beam radius. A random Poisson ($\lambda$) number of points $(x_i, y_i)$ are uniformly distributed across this region and a random nonnegative integer mark $Y(x_i, y_i)$ is attached to each occurrence point. For this simulation study, a Poisson marking distribution is used (as this corresponds to one of interest to Nils Petersen [24]). This, however, may be generalized to include other marking distributions such as the negative binomial or gamma. The resulting occurrence point, mark triples $(x_i, y_i, Y(x_i, y_i))$ are stored in a pointer $^*p$.

The simulated random field driven by this Poisson process is generated in parallel i.e. on the PE array. Essentially, each PE is responsible for obtaining an $8 \times 16$ sub-array of the entire $512 \times 512$ image array of points (recalling that the PE array consists of $nxproc = 64$ by $nyproc = 32$ processor elements).

Now if one chooses the noise parameter $\gamma$ to be zero for the moment, from
Section 4.2.

\[ X(t) = \int_{S_t} f_t(x, y) Y(x, y) \cdot \sum_{1}^{N(S_t)} f_t(x, y_t) Y(x, y_t) \]

where the disc region \( S_t \) is chosen as

\[ S(t_1, t_2) : (x - t_1 \delta - r)^2 + (y - t_2 \delta - r)^2 \leq r^2, \quad (t_1, t_2) \in \mathbb{Z}^2 \]

The point \((t_1 \delta + r, t_2 \delta + r)\) is taken to be the position of the laser beam centre while \( S_t \) is the effective area of illumination.

The algorithm for the calculation of \( X(0, 0) \) for example, then proceeds as follows:

i) determine those Poisson occurrence points \((x, y)\) that lie within the circle centred at \((r, r)\)

ii) form the products of the occurrence point weights and random marks \( f_0(x, y) Y(x, y) \)

iii) sum over all such points lying within the circle from i).

To calculate \( X(0, 1) \), the centre of the circular region is simply shifted by an amount \( \delta \) to \((r, r + \delta)\), and the above algorithm is once again employed. This is then repeated, giving rise to \( X(t_1, t_2), t_1, t_2 = 0, \ldots, 511 \). The spatial index then of the observed random field corresponds to the position of the laser beam centre (plus the additive constant \( r \)).

To compute \( X(t) \) in parallel, take each PE to correspond to the rectangular region which would enclose an \( 8 \times 16 \) collection of the overlapping discs \( S_t \) determining \( X(t) \). One can then map the entire Poisson process realization \*p to each
PE and employ the above algorithm. To determine $X(t_1, t_2)$ on $\text{PE}(ixproc, ixproc)$, sum over all points $(x_i, y_i)$ within the circle centred at $((t_1 + 8 \times ixproc)\delta + r, (t_2 + 16 \times ixproc)\delta + r)$, $ixproc = 0, \ldots, 63$, $iyproc = 0, \ldots, 31$. The PE array has been topologized to correspond to the square subregion of (without loss of generality) the first quadrant of the Cartesian plane containing the Poisson process realization while each PE corresponds to a rectangular subregion of this square enclosing the discs $S_t$. The simulated image data is then stored on each PE in plural arrays $X[8][16]$.

### 5.6.2 Least Squares Estimation

The model parameters can be estimated by a nonlinear least squares technique involving the random field periodograms. Newton's method is used to find a simultaneous solution to

$$\frac{\partial}{\partial \theta_1} L^{(T)}(\hat{\theta}) = 0$$

$$\frac{\partial}{\partial \theta_2} L^{(T)}(\hat{\theta}) = 0$$

after obtaining an appropriate starting value.

To find this starting value, a simple rectangular grid search is used. Basically, this requires evaluating the least squares function

$$L^{(T)}(\theta) = \left(\frac{2\pi}{T}\right)^2 \sum_{s_j=1}^{T-1} (T)^{(T)}(\frac{2\pi}{T} s) - f^{(T)}(\frac{2\pi}{T} s; \theta))^2$$

at each grid point $(\theta_1, \theta_2)$. As this sum is $O(T^2)$ it seems reasonable to evaluate it in parallel on the PE array which is considered to correspond to the lattice of discrete Fourier frequencies $\left[\frac{2\pi}{T}(s_1, s_2)\right]$. $s_1, s_2 = 0, \ldots, 511$, following the discussion
in Section 5.5.1. However, in practice, the upper limit in the least squares function is chosen to be some integer \( L < T - 1 \) due to spectral density truncation error (see Remark 5.1). This implies that only a subset of the available PE's should be used to perform this summation. This is still superior (in terms of run time) however, to a purely sequential calculation.

Newton's method involves the first- and second-order derivatives of the least squares function (5.3) and consequently, sums of a similar form. These are also evaluated in parallel over the same sub-lattice of discrete frequencies.

**Remark 5.2** To compute the 2-d periodograms, the MPL function `cfft2d()` is used. Basically, this function calculates the 2-d fast Fourier transform of a complex-valued data set (set the imaginary part to zero and on each PE and store \( X[8][32] \) as opposed to \( X[8][16] \)) spread over all processors. The original data is thus transformed from the spatial domain \((X(t_1, t_2), \text{ where } t_1 \text{ and } t_2 \text{ correspond to the centre of the laser beam})\) to the frequency domain \((X(s_1, s_2), \text{ where } s_1 \text{ and } s_2 \text{ correspond to the associated discrete Fourier frequency } \frac{2\pi}{T}(s_1, s_2))\).

**Numerical Results**

The results of 100 simulations are graphically summarized in Figure 5.6.3. Each simulation run requires approximately 5 minutes, leading to a total run time of approximately 8.3 hours. The parameter settings are identical to those of Section 5.5.2. The actual value of the ratio parameter here is 8.7 which falls within the asymptotic confidence interval depicted. The asymptotic normal density functions are also provided. The value for the background image noise parameter \( \beta \) is
chosen to be 2.0 for these simulations corresponding to a moderate noise level.

5.6.3 Variance Estimation

The limiting variance of \( \hat{R} \) is a quadratic form involving the matrix \( \Sigma^* = B \Sigma B' \) where \( \Sigma \) is the limiting covariance matrix of Theorem 4.1 and \( B \) involves integrals of spectral density derivatives (known functions of \( \theta_0 \)). Thus, in order to obtain approximate confidence intervals for \( R \), one must find a consistent estimator of \( \Sigma \) as \( B \) may be estimated by substituting the least squares (or likelihood) estimates \( \hat{\theta} \) for \( \theta_0 \).

Block Method

One method of estimating \( \Sigma \) (or \( \Gamma \)) is to exploit the finite-range dependence of the process \( X(t) \) and divide the data into approximately independent blocks \( B_1, \ldots, B_N \) of respective sizes \( T_0^2 \) (Section 3.9.2). Then on each block \( B_j \), sample means \( X_{B_j} \) and smoothed periodograms

\[
I_{B_j}^{(k)} = \left( \frac{2\pi}{T} \right)^2 \sum_{s=1}^{T_0-1} I^{(T_0)}(\frac{2\pi}{T_0} s) \frac{\partial}{\partial \theta_k} f(\frac{2\pi}{T_0} s; \hat{\theta}), \quad k, \ell = 1, 2 \quad \text{and} \quad j = 1, \ldots, N,
\]

can be calculated resulting in approximately independent samples (of size \( N \)) from which first- and second-order sample moments (variances) can be obtained.

Now dividing the data \( X(t) \) into blocks \( B_j \) is analogous to partitioning the PE array into sub-arrays of \( n_{xproc}/n_x \) by \( n_{yproc}/n_y \) processor elements. After some trial and error, it has been determined that reasonable values for \( n_x \) and \( n_y \) (given the model parameter values) were 16 and 16 respectively. That is, the blocks \( B_j \) consist of sub-arrays of \( 4 \times 2 \) PE's or equivalently, sub-arrays of \( X(t) \) of \( 32 \times 32 \) data
Histograms for ICS, from 100 simulations

Figure 5.6.3: Histograms of $X$, $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{R}$
Figure 5.6.4: MasPar X-Net communication network
points (recalling that plural arrays of $8 \times 16$ data points are stored on each PE).

To make these data blocks approximately independent, some of the data points on the north and east block boundaries may be set to zero. For example, consider the PE block \( \{ \text{PE} (iyproc, ixpro) : ixpro = 0, \ldots, 3, iyproc = 0, 1 \} \), i.e., those PE's in the southwest corner of the PE array. To set the appropriate data values to zero, one can use conditional statements of the form:

\[
\text{if } (ixpro \% 4 == 3)
\]

or

\[
\text{if } (iyproc \% 2 == 1),
\]

and set certain \( X(t) \) values to zero.

Denote the block discussed above by \( B_1 \). Now to compute sample means on each block, for example on \( B_1 \), one must use the MasPar PE communication network, \( X\text{-Net} \) (Figure 5.6.4). First, one sums all the data points on each PE in \( B_1 \) and stores the results in plural variables called \( xsum \). Then, add together all these sums and store the result (for convenience) on the lower southwest PE (PE(0, 0) in this case). This is accomplished by using statements of the following form. Suppose \( x \) is a plural variable. To update \( x \) on certain PE's say with \( ixproc \% 2 == 0 \) (where \% is integer remainder), by adding the value of \( x \) stored on the PE's immediately to the east, one uses a statement of the form

\[
\text{if } (ixpro \% 2 == 0)
\]

\[
x += x\text{netE}[1] \cdot x,
\]
where the direction (East) and distance away from the PE of interest (in this case, the first PE to the east) are explicitly denoted in the statement. In this way one can calculate the sum of all the data points in $B_j$ (or $B_j$, $j = 1, \ldots, N$) and obtain the sample means $X_{B_j}$.

To compute smoothed periodograms $I^{(i)}_{B_j}$ on each block $B_j$, $j = 1, \ldots, N$, first evaluate the spectral density derivatives $\frac{\partial}{\partial \theta} f(\frac{2\pi}{f_0}; \hat{\theta})$ at the discrete Fourier frequencies $\frac{2\pi}{f_0}(s_1, s_2)$, $s_1, s_2 = 1, \ldots, T_0 - 1$, (i.e. consider each PE block to correspond to the above sub-lattice of discrete frequencies). These derivative evaluations are the same on each block and so only the situation on $B_1$ is considered. Recall that $B_1$ is composed of the eight individual processor elements PE($i\text{typroc},i\text{txproc}$), $i\text{txproc} = 0, \ldots, 3$, $i\text{typroc} = 0, 1$, and one must evaluate $\frac{\partial}{\partial \theta} f(\cdot)$ at frequencies spread over these eight PE's. Consider PE(0,0). Evaluate $\frac{\partial}{\partial \theta} f(\cdot; \hat{\theta})$ over the frequencies $\frac{2\pi}{f_0}(s_1, s_2)$, $s_1 = 1, \ldots, 7$, $s_2 = 1, \ldots, 15$ using the conditional statement if ($i\text{typroc} \% 4 == 0$ && $i\text{typroc} \% 2 == 0$). Similarly, on PE(0,1) evaluate at $\frac{2\pi}{f_0}(s_1, s_2)$, $s_1 = 8, \ldots, 15$, $s_2 = 1, \ldots, 15$ using the statement if($i\text{txproc} \% 4 == 1$ && $i\text{txproc} \% 2 == 0$). Continue in this way for all eight PE's in $B_1$ (and hence $B_j$, $j = 1, \ldots, N$).

The next part of the program involves computing the 2-d periodograms of the data in each block $B_j$, $j = 1, \ldots, N$. This is easily accomplished using the MPL function cfft $m2d(\cdot, 32, 32, 16, 16, \cdot)$ which computes $16 \times 16$ 2-d complex fast Fourier transforms each of size $32 \times 32$ with the data spread across all processors. One then determines on each PE, the relevant periodogram-spectral derivative products and sums these over all discrete frequencies corresponding to that PE,
as determined by \( \partial \Theta \). Finally, the same X-Net communication algorithm discussed earlier in relation to \( X_{B_j} \) can be employed to sum these products over all eight processors comprising \( B_1 \) and to store the result on PE(0,0) (or the corresponding southwestern-most PE in \( B_j \)).

On each block \( B_j \) then, one can calculate the quantities of interest namely, \( X_{B_j} \), \( I_{B_j}^{(1)} \) and \( I_{B_j}^{(2)} \) and store them on PE's of the form PE(\( iyproc, ixproc \)). \( iyproc \% 4 == 0 \) and \( ixproc \% 2 == 0 \), giving an approximately independent sample of size \( N = 16^2 \). It is now a relatively easy exercise (using the statement \( \text{if}(ixproc \% 4 == 0 \& \& iyproc \% 2 == 0) \) and the \( \text{reduceAddf}(\cdot) \) function) to compute the relevant sample variances and covariances, which, when appropriately scaled (by multiplying by \( T_0^2 \), \( T_0 = 32 \)) yield consistent estimates of the elements of \( \Sigma \).

This procedure can also be enhanced by employing the block shifting (or offset idea of Section 3.9.2.) This allows a larger sample of independent replicates upon which to base a variance estimate. In this way, the data can be used more efficiently by creating different starting positions of the initial block.

**Taniguchi Method**

Another way of estimating the components of \( \Sigma \) (or \( \Gamma \)) (specifically the spectral integrals) involves a variation of a method initially proposed by Taniguchi ([33], [34]). In his 1982 paper, Taniguchi considers the estimation of

\[
H = \int_{-\tau}^{\tau} u(\lambda) \Phi(f(\lambda)) d\lambda
\]
where \( f(\lambda) \) is the spectral density of a stationary, mean zero process \( \{X(t)\}, t \in \mathcal{Z} \), and \( \Phi(\cdot) \) is an arbitrary function. The proposed estimator is of the form

\[
H^{(n)} = \int_{-\pi}^{\pi} \psi(\lambda) \mathcal{L}^{-1}\{\Phi(\frac{1}{\omega})\frac{1}{\omega}\}_{<f^{(n)}(\lambda)>} d\lambda
\]

where \( \mathcal{L}^{-1}\{F(\omega)\}_{<u>} \) denotes the inverse Laplace transform of \( F(\omega) \) at the argument \( u \) and \( f^{(n)}(\lambda) \) is the usual periodogram of \( X(t), t = 0, \ldots, n - 1 \). Under certain conditions, it is shown that \( H^{(n)} \) is consistent and asymptotically normal.

As an example of \( H^{(n)} \), consider \( \Phi(\omega) = \omega^2 \). Then,

\[
\mathcal{L}^{-1}\{\Phi(\frac{1}{\omega})\frac{1}{\omega}\}_{<u>} = \frac{u^2}{1(3)}.
\]

Thus to estimate

\[
\int_{-\pi}^{\pi} \psi(\lambda) f^2(\lambda) d\lambda
\]

one would use

\[
\int_{-\pi}^{\pi} \psi(\lambda) \frac{f^{(n)}(\lambda)^2}{2} d\lambda.
\]

Of course, the former integral appears as the first term of the limiting covariance of smoothed periodograms (Brillinger, [5]).

Taniguchi [34] considers the estimation of the second term of the above covariance, namely the double integral of the fourth-order cumulant spectral density,

\[
\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \psi(\alpha, \beta) f_4(\alpha, \beta, -\beta) d\alpha d\beta
\]

for a strictly stationary process \( \{X(t)\}, t \in \mathcal{Z} \). He obtains a consistent estimator based on summing certain products of discrete Fourier transforms over the usual discrete frequencies \( \frac{2\pi j}{n} \). Taniguchi also uses the spectral window (for the time
series $X(t), t = 0, \ldots, T - 1$

$$H_T(x) = B_T^{-1} H(B_T^{-1} x)$$

where $H(x)$ is an even, real-valued function of bounded variation with

$$\int_{-\pi}^{\pi} H(x) dx = 1, \quad H(x) = 0 \text{ for } x \notin [-\pi, \pi]$$

and $\{B_T\}$ is a sequence such that $B_T \to 0, B_T^2 T \to \infty$ as $T \to \infty$.

Keenan [14] discusses the same integral and proposes a modification of the Taniguchi estimator which is of a simpler form. In Keenan’s notation, let

$$F^{(4)}(A) = \int_A f_4(\lambda_1, \lambda_2, \lambda_3) d\lambda_1 d\lambda_2 d\lambda_3$$

be the fourth-order cumulant spectral distribution (measure) where $A$ is a Borel set in $S^3$, $S$ being the unit circle (with the order of integration reduced due to stationarity). Then the parameter of interest $\Theta(F^{(4)})$ may be denoted in a general sense by

$$\Theta(F^{(4)}) = \int_{S^3} h(\lambda_1, \lambda_2, \lambda_3) dF^{(4)}(\lambda_1, \lambda_2, \lambda_3)$$

where $h(\cdot)$ is of bounded variation. For example, the asymptotic covariance matrix discussed earlier contains integrals of the form

$$\int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_1} f(\lambda_1; \theta_0) \frac{i}{i\theta_2} f(\lambda_2; \theta_0) f_4(\lambda_1, \lambda_2, -\lambda_2) d\lambda_1 d\lambda_2$$

where

$$h(\lambda_1, \lambda_2, \lambda_3) = \frac{i}{i\theta_1} f(\lambda_1; \theta_0) \frac{i}{i\theta_2} f(\lambda_2; \theta_0)$$

and

$$dF^{(4)}(\lambda_1, \lambda_2, \lambda_3) = f_4(\lambda_1, \lambda_2, -\lambda_2) d\lambda_1 d\lambda_2.$$
Note that this differs in form from $\Theta(F^{(4)})$ in that the integration is carried out over $[0, 2\pi]^2$ instead of $[0, 2\pi]^3$, and it is to handle this extra dimension that Taniguchi [34] uses the spectral window.

Keenan further defines $F_T^{(4)}(\cdot)$ (based upon $X(t), t = 0, \ldots, T - 1$) as

$$F_T^{(4)}(\lambda_1, \lambda_2, \lambda_3) = \left(\frac{2\pi}{T}\right)^3 \sum_{2\pi r_1 \leq \lambda_1} I_T^{(4)}(\frac{2\pi}{T}r_1, \ldots, \frac{2\pi}{T}r_4)$$

where $I_T^{(4)}(\cdot)$ is the fourth-order periodogram (Brillinger [2])

$$I_T^{(4)}(\lambda_1, \ldots, \lambda_4) = \frac{(2\pi)^{-3}}{T} \eta\{\sum_{j=1}^{4} \lambda_j\} \prod_{j=1}^{4} d_T(\lambda_j).$$

$\eta\{\cdot\}$ is the Kronecker comb

$$\eta\{\alpha\} = \begin{cases} 1 & \text{if } \alpha \equiv 0 \mod 2\pi \\ 0 & \text{otherwise} \end{cases}$$

and $d_T(\lambda)$ is the finite Fourier transform.

Now let $\tilde{F}_T^{(4)}(\cdot)$ be defined as $F_T^{(4)}(\cdot)$ except with the sum being over only those $(\frac{2\pi}{T}r_1, \ldots, \frac{2\pi}{T}r_4)$ which are contained in $\Omega^{(4)}$. defined to be the collection of $(\lambda_1, \ldots, \lambda_4)$ in $S_4$ such that $\sum_{j=1}^{4} \lambda_j \equiv 0(\mod 2\pi)$ but is not contained in any further submanifold $\sum_{j=1}^{s} \lambda_{i_j} \equiv 0(\mod 2\pi)$ where $(i_j, j = 1, \ldots, s) \subset \{1, 2, 3, 4\}$.

Taniguchi's estimator involves three sums, the first of which induces bias due to the inclusion of submanifolds of the form above; as such, the latter sums remove the bias. Keenan's simplified estimator is the first of these sums with the summation restricted to points in $\Omega^{(4)}$:

$$\left(\frac{2\pi}{T}\right)^3 \sum_{\Omega^{(4)}} \psi(\lambda_1, \lambda_2, \lambda_3) H_T(\lambda_2 + \lambda_3) I_T^{(4)}(\lambda_1, \lambda_2, \lambda_3)$$

where $H_T(\cdot)$ is the corresponding spectral window.
In this work, these estimators of Taniguchi and Keenan are adapted to obtain those corresponding to a stationary random field \( \{ X(t) \} \), \( t \in \mathbb{Z}^2 \).

Turning next to the estimation of the components of \( \Sigma \), recall that

\[
\sigma_{00} = \sum_{k_j = -\infty}^{\infty} c_X(k; \theta_0) \\
= \gamma^2 + \sum_{k_j = -\infty}^{\infty} c(k; \theta_0) = (2\pi)^2 f(0; \theta_0).
\]

The infinite sums are in practice finite sums ranging from \(-m\) to \(m\) due to the finite-range mixing of \( \{ X(t) \} \), \( t \in \mathbb{Z}^2 \). A possible estimator of \( \sigma_{00} \) would be

\[
\sum_{k_j = -\infty}^{\infty} \hat{c}(k)
\]

where \( \hat{c}(k) \) is the usual sample autocovariance function based on \( X(t) \), \( t \in \{ 0, \ldots, T - 1 \}^2 \).

\[
\hat{c}(k) = \frac{1}{T^2} \sum_{t_j = 0}^{T - |k_j| - 1} (X(t) - X)(X(t + k) - X), \quad j = 1, 2.
\]

However, calculation of \( \hat{c}(k) \) on the MasPar's PE array is not as straightforward as it might first appear, due to the inherent asymmetry of \( \hat{c}(k) \) given that the observed data is stored across all processors. That is, to calculate \( \hat{c}(k) \) would require a substantial amount of bookkeeping and inter-PE communication via X Net. Also, to keep in the spirit of Taniguchi-Keenan, one should concentrate on estimating \( (2\pi)^2 f(0; \theta_0) \) directly based on the usual finite Fourier transform approach.

With this in mind, estimate \( (2\pi)^2 f(0; \theta_0) \) in a neighbourhood of \( 0 \) by

\[
(2\pi)^2 \hat{f}(0) = W^{(T)}(0) I^{(T)}(\frac{2\pi}{T}, \frac{2\pi}{T}) + 2 \sum_{k_1 = 1}^{m} \sum_{k_2 = 1}^{m} W^{(T)}(k) I^{(T)}(\frac{2\pi}{T} (k_1 + 1), \frac{2\pi}{T} (k_2 + 1)).
\]
with, for simplicity, a uniform weight function

\[ W^{(T)}(k) = \frac{1}{2m^2_T + 1} \]

and values of \(m_T\) of 2 or 3. The above is the random field extension of an estimator discussed in Brockwell and Davis [8].

To estimate the third-order spectral integrals of \(\sigma_{0,j}\), \(j = 1, 2\),

\[ (2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_j} f(\alpha; \theta_0) f_3(0, \alpha) d\alpha \, . \]

First consider (following Keenan [14]) estimators of the form

\[ (2\pi)^2 \frac{1}{T} \sum_{s=1}^{T-1} \frac{\partial}{\partial \theta_k} f(\frac{2\pi}{T} s; \hat{\theta}) f_3^{(T)}(\frac{2\pi}{T}; 1, \frac{2\pi}{T} s, -\frac{2\pi}{T} (s + 1)). \, j, k = 1, 2 \]

where

\[ f_3^{(T)}(\frac{2\pi}{T}; 1, \frac{2\pi}{T} s, -\frac{2\pi}{T} (s + 1)) \]

\[ = \frac{1}{(2\pi)^4 T^2} d^{(T)}(\frac{2\pi}{T}; 1) d^{(T)}(\frac{2\pi}{T} s) d^{(T)}(\frac{2\pi}{T} (s + 1)) \, . \]

The sum is evaluated in parallel with the frequencies \(\frac{2\pi}{T} s\) spread across the processors. One can improve the estimator by locally averaging over frequencies in the first argument of \(f_3^{(T)}(\cdot)\).

Now, the asymptotic covariances of the smoothed periodograms involve two terms, the first being the "Gaussian part" (Taniguchi [34]) or second-order spectral integrals.

\[ 2(2\pi)^2 \int_0^{2\pi} \frac{\partial}{\partial \theta_1} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_2} f(\alpha; \theta_0) f^2(\alpha; \theta_0) d\alpha \, . \]

Following Taniguchi [33] this is estimated by

\[ (2\pi)^2 \frac{1}{T} \sum_{s=1}^{T-1} \frac{\partial}{\partial \theta_1} f(\frac{2\pi}{T}; s; \hat{\theta}) \frac{\partial}{\partial \theta_2} f(\frac{2\pi}{T}; s; \hat{\theta})(f^{(T)}(\frac{2\pi}{T}; s))^2 \, . \]
The second term is the usual fourth-order spectral integral

$$(2\pi)^2 \int_0^{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \theta_1} f(\alpha; \theta_0) \frac{\partial}{\partial \theta_2} f(\beta; \theta_0) f_4(\alpha, \beta, -\alpha) d\alpha d\beta.$$  

One requires a corresponding two-dimensional spectral window $H_T(x, y)$. For convenience, choose $H(x, y)$ such that

$$\int_0^{2\pi} \int_0^{2\pi} H(x, y) dx dy = 1$$

to be the uniform probability density function

$$H(x, y) = \begin{cases} \frac{1}{4\pi^2} & \text{if } 0 \leq x, y \leq 2\pi \\ 0 & \text{otherwise} \end{cases}.$$

Now choose $B_T = T^n, -\frac{1}{2} < n < 0$, such that $B_T \to 0$ and $B_T^2 T \to \infty$ as $T \to \infty$. Then

$$H_T(x, y) = B_T^{-2n} H(B_T^{-1} x, B_T^{-1} y) = T^{-2n} H(T^{-n} x, T^{-n} y)$$

$$= \begin{cases} \frac{T^{-2n}}{4\pi^2} & \text{if } 0 \leq x, y \leq T^n 2\pi \\ 0 & \text{otherwise} \end{cases}$$

and the estimator becomes

$$(2\pi)^2 (\frac{2\pi}{T})^{6n} \sum_{a_j=1}^{T-1} \sum_{b_j=1}^{T-1} \sum_{c_j=1}^{T-1} \left\{ \frac{\partial}{\partial \theta_1} f(\frac{2\pi}{T}; a; \hat{\theta}) \frac{\partial}{\partial \theta_2} f(\frac{2\pi}{T}; b; \hat{\theta}) H_T(\frac{2\pi}{T} (b + c)) I_4(\frac{2\pi}{T}; a, \frac{2\pi}{T}; b, \frac{2\pi}{T}; c) \right\}. \tag{5.4}$$

where

$$I_4(\frac{2\pi}{T}; a, \frac{2\pi}{T}; b, \frac{2\pi}{T}; c) = \frac{1}{(2\pi)^6 T^2} d(T)(\frac{2\pi}{T}; a) d(T)(\frac{2\pi}{T}; b) d(T)(\frac{2\pi}{T}; c) d(T)(-\frac{2\pi}{T}; (a + b + c)).$$
The sum is subject to the following restrictions imposed by the submanifolds (as specified in $\Omega^{(4)}$):

\[
\begin{align*}
\mathbf{a} + \mathbf{b} & \neq \mathbf{T} \\
\mathbf{a} + \mathbf{c} & \neq \mathbf{T} \\
\mathbf{b} + \mathbf{c} & \neq \mathbf{T} \\
\mathbf{a} + \mathbf{b} + \mathbf{c} & \neq \mathbf{T}.
\end{align*}
\]

To fully exploit the data-parallel architecture of the MasPar, one would evaluate the sum (which is $O(T^6)$) in (5.4) on the PE array. That is, the PE array would be thought of as corresponding to the six-dimensional lattice of frequencies $\frac{2\pi}{T}(\mathbf{a}, \mathbf{b}, \mathbf{c}), (\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \{1, \ldots, T - 1\}^6$ and this lattice would then be partitioned into sub-lattices of the form $\frac{2\pi}{T}(\mathbf{a}, \mathbf{b}, \mathbf{c}), (\mathbf{a}, \mathbf{b}, \mathbf{c}) \in (\{0, \ldots, 7\}) \times \{0, \ldots, 15\}^3$. The sum evaluation then would be done in parallel with sums over the above sub-lattices of frequencies being calculated on each PE. The submanifold restrictions, however, make this data-parallel implementation of Keenan's estimator impractical as the conditions would have to be checked sequentially for each frequency $\frac{2\pi}{T}(\mathbf{a}, \mathbf{b}, \mathbf{c})$, in most algorithms (unless of course one determined all the submanifold cases in advance). So although (5.4) is of a particularly simple form, it is not directly compatible with the SIMD architecture of the MasPar. As such, in order to efficiently implement estimators of this type on a massively parallel architecture one is led back to the estimator suggested by Taniguchi [34]. Let

\[
u'\left(\frac{2\pi}{T} \mathbf{a}, \frac{2\pi}{T} \mathbf{b}\right) = \frac{\partial}{\partial \theta_1} f\left(\frac{2\pi}{T} \mathbf{a}; \theta_0\right) \frac{\partial}{\partial \theta_2} f\left(\frac{2\pi}{T} \mathbf{b}; \theta_0\right).
\]

Then Taniguchi's estimator (extended to the random field) is of the form
\[
\left(\frac{2\pi}{T}\right)^6 \sum_{a_j=1}^{T-1} \sum_{b_j=1}^{T-1} \sum_{c_j=1}^{T-1} \left\{ \psi\left(\frac{2\pi}{T} a, \frac{2\pi}{T} b\right) H_T\left(\frac{2\pi}{T} (b + c)\right) \right. \\
\left. \left(\frac{1}{(2\pi)^6 T^2}\right)d^{(T)}\left(\frac{2\pi}{T} a\right)d^{(T)}\left(\frac{2\pi}{T} b\right)d^{(T)}\left(\frac{2\pi}{T} c\right)d^{(T)}\left(-\frac{2\pi}{T} (a + b + c)\right) \right\} \\
- \frac{1}{2}\left(\frac{2\pi}{T}\right)^2 \sum_{a_j=1}^{T-1} \left\{ \psi\left(\frac{2\pi}{T} a, \frac{2\pi}{T} a\right) + \psi\left(\frac{2\pi}{T} a, -\frac{2\pi}{T} a\right) \right\} \left(\frac{1}{2\pi T}\right)^4 |d^{(T)}\left(\frac{2\pi}{T} a\right)|^4 \\
- \frac{1}{B_T}\left(\frac{2\pi}{T}\right)^4 \sum_{a_j=1}^{T-1} \sum_{b_j=1}^{T-1} H(0)\psi\left(\frac{2\pi}{T} a, \frac{2\pi}{T} b\right)\left(\frac{1}{2\pi T}\right)^4 |d^{(T)}\left(\frac{2\pi}{T} a\right)d^{(T)}\left(\frac{2\pi}{T} b\right)|^2 . \quad j = 1, 2,
\]

where \(B_T\), \(H(\cdot)\) and \(H_T(\cdot)\) are as discussed previously.

**Numerical Results**

The DPU time required to simulate the fluorescence intensity process and perform the required parameter estimation varies considerably due to its dependence upon the finite-range parameter \(m\). Using the frequency spacing \(\frac{2\pi}{T}\), the Taniguchi and Keenan algorithms run indefinitely, and it has been estimated that they might require in the order of years to complete their numerical evaluation. This is based on some smaller truncation runs, and extrapolation. A truncation at \(m\) terms due to the effective Gaussian kernel diameter results in \(m^2 O(T^6)\) computations. Extrapolating to the size \(m\) in this problem yields the estimate of years, clearly an infeasible length of time. One way to expedite the performance of the Taniguchi-Keenan algorithm is to sum over a sparser grid of frequencies in the periodogram sums, since the algorithm is essentially a Riemann approximation using higher-order periodograms in place of cumulant spectra. This suggests that one sum over frequencies, say \(\frac{2\pi Mk}{T}\) instead of summing over the entire grid of frequencies \(\frac{2\pi k}{T}\).
This coarser grid results in a fraction $1/M^6$ as many terms in the sum. However, more bias tends to occur in the estimation of the spectral integrals as the grid parameter $M$ increases.

For $M = 8$, the calculation requires about two minutes while $M = 4$ about 5.7 days. Due to the inherent difficulty in verifying the reliability of the $M = 4$ estimate (over several runs), the value is not presented. These DPU times extrapolate to $M = 1$ (full grid) requiring approximately 364 days or 1 year. Unfortunately, using the coarse grid ($M = 8$) can result in negative variance estimates. As such, an extensive numerical survey of the variance estimation techniques discussed above is somewhat limited. Nevertheless, results have been obtained and are presented in the table below. The following parameter settings are used:

$$\mu = 0.285, \quad \theta_1 = 20.00, \quad \theta_2 = 0.15, \quad R = 8.75 \quad \text{and} \quad \delta = 0.10 .$$

<table>
<thead>
<tr>
<th>Method</th>
<th>$M$</th>
<th>$\hat{\sigma}_R$</th>
<th>Range ($\hat{\sigma}_R$)</th>
<th>DPU times (approx.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>1</td>
<td>0.269</td>
<td>0.225 - -0.295</td>
<td>5 mins.</td>
</tr>
<tr>
<td>Taniguchi</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>364 days</td>
</tr>
<tr>
<td>Taniguchi</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>5 days</td>
</tr>
<tr>
<td>Taniguchi</td>
<td>8</td>
<td>0.15 - .20</td>
<td>-</td>
<td>2 mins</td>
</tr>
</tbody>
</table>

The noise level is $\beta = 2.0$. The actual analytic variance (as determined by Monte Carlo integration) is $0.328 \pm 0.020$. The range of values for the block method was over ten simulations.

The block method is seen to provide acceptable results with reasonably short DPU times and is based on the analysis of a single image dataset. The Taniguchi estimator is considerably more difficult to implement but can be used on a coarser grid. The estimates, however, are consistently too small. This approach may also give negative variance estimates. It should be noted, however, that this work rep-
resents one of the first attempts to numerically implement and study this estimator in the context of a two-dimensional random field. Due to its inherent simplicity and versatility, the block method seems to provide a superior variance estimator in this setting. It is expected that this superiority carries over to other random field and spatial problems.

5.7 ICS Data

Five ICS data sets from N. Petersen’s laboratory are analyzed using the least squares and block variance estimation methods. The actual confocal scanning laser microscope image (512 × 512 pixels) and associated fluorescence intensity plot would appear similar to Figure 4.1.1. The ratio \( R \) remains the parameter of primary focus. The results of the analysis are tabulated below.

<table>
<thead>
<tr>
<th>Data</th>
<th>( \hat{R} )</th>
<th>( \hat{\sigma}_R )</th>
<th>DPU times (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tr-fl1.bin</td>
<td>1390.85</td>
<td>3.53</td>
<td>34</td>
</tr>
<tr>
<td>tr-f2.bin</td>
<td>12 826.17</td>
<td>51.24</td>
<td>34</td>
</tr>
<tr>
<td>tr-f3.bin</td>
<td>148 197.53</td>
<td>26.98</td>
<td>34</td>
</tr>
<tr>
<td>tr-f4.bin</td>
<td>11 366.697</td>
<td>12.58</td>
<td>34</td>
</tr>
<tr>
<td>tr-f5.bin</td>
<td>2340.65</td>
<td>3.37</td>
<td>34</td>
</tr>
</tbody>
</table>

The laser beam width parameter \( w \) is essentially known to be 0.1885 while the pixel size is \( \delta = 0.04 \). The variability of the \( \hat{R} \) and \( \hat{\sigma}_R \) values is indicative of ICS data sets with significantly different protein cluster sizes and densities and different background noise levels. These background noise levels can also be estimated in a manner similar to that discussed in Section 3.10.
5.8 Discussion and Summary

The MasPar MP-2 represents the state-of-the-art in massively parallel computing. This chapter presents an innovative application of this technology to a computational problem in the statistical sciences.

The parallel architecture of the MP-2 is demonstrated to be ideally suited for most aspects of the spatial model of Image Correlation Spectroscopy considered here. Indeed, image analysis problems of this type would represent one of the major areas of application for this technology. However, many other computational problems which exhibit a data parallel structure can be examined on the MP-2.

This chapter has focused on the calculation and estimation of the asymptotic covariances arising in Chapter 4. Section 5.5.1 discusses the major issues which arise in the evaluation of the usual spectral integrals using a simple Riemann sum (numerical integration) approach. To address these limitations, a Monte Carlo integration technique is proposed and implemented, leading to an interesting comparative study of the nonlinear least squares and asymptotic likelihood methods of Chapter 4. The numerical results of such a comparison are infrequently found due to the inherent computational challenges posed by a stationary random field model of the type considered here. Section 5.6 begins with a discussion of the simulation of the Poisson-driven spatial process. Complete details of the parallel algorithm are provided. The implementation of the least squares method is then presented. The various issues encountered in its implementation are also addressed. The central component of the chapter is the discussion of the estimation of the asymptotic
variance of $\hat{R}$ in Section 5.6.3. An adaptation of the block method of Chapter 3 is introduced and is shown to provide good results. The method is compared to that of Taniguchi [34], which is implemented in the stationary random field case. Keenan's [14] generalization of the method is also discussed along with the infeasibility of its actual implementation on the MP-2. The Taniguchi estimator should provide similar results but only after several days of DPU time, demonstrating the superiority of the block method in this particular setting. The chapter concludes with the analysis of five actual ICS data sets demonstrating the practicality of the proposed estimation methods.

The parallel computing technology of the MasPar MP-2 is not commonly applied to problems in statistical computing. This chapter demonstrates the enormous potential of this modern technology in the solution of complex computational issues arising in the statistical sciences.
Chapter 6

Concluding Remarks and Further Work

Scanning Fluorescence and Image Correlation Spectroscopy are among the world's leading experimental techniques for making quantitative measurements on the surfaces of living cells. The techniques involve the analysis of confocal scanning laser microscope images resulting in estimates of cell membrane protein cluster sizes and surface densities. These estimates form an important part in the understanding of various aspects of cell membrane physiology.

This thesis work presents the first stochastic model-based approach to S-FCS and ICS. This statistical application serves to go beyond the empirical least squares fitting procedures currently used in laboratory practice. The advantages of a formal Poisson-driven stationary process model, which arises from the biophysical characteristics of these spectroscopy techniques, are demonstrated throughout this work. Specifically, the model suggests that asymptotic confidence intervals can be obtained for the physically meaningful ratio parameter $R$, based on a single fluorescence intensity data set. Previously, only after the time consuming analysis
of many data sets, could the precision of the point estimator \( \hat{R} \) be assessed.

A detailed study of statistical inference for S-FCS is presented in Chapter 3. It is shown that an optimal least squares lag parameter \( L \) can be obtained based on the criterion of minimizing the limiting variance of \( \hat{R} \). The block variance estimation method is also introduced. The method is shown to provide high quality consistent estimates of the asymptotic variance of \( \hat{R} \). A method is also discussed for removing extreme outliers of \( \hat{R} \) based on a comparison of the measured and estimated values of the laser beam width parameter \( \omega \). Finally, techniques for estimating background noise levels and assessing model fit are presented.

The theoretical details of the model’s implementation to ICS are described in Chapter 4. Both least squares and asymptotic likelihood estimation methods based on the periodograms of the random field process are proposed and the relevant asymptotic results are derived. The complete details of the proof of the asymptotic normality of smoothed periodograms are provided including some detailed remarks concerning Rosenblatt’s [27] statement of the result and its inappropriateness for a data analytic approach. Also discussed is the apparent nonapplicability to the random field setting of Brillinger’s [2] method of proof based on the limiting characteristics of cumulants of periodograms.

ICS has recently been extended to Image Cross-Correlation Spectroscopy (ICCS), a technique to measure the slow diffusion of fluorescent particles [29]. The method is based on the analysis of temporal fluctuations of the observed fluorescence intensity. The proposed model and estimating equations of Chapter 4 carry over directly to the analysis of ICCS data where at time \( t \) and pixel position \( s \), \( X(s; t) \)
would represent an observed fluorescence intensity measurement.

Due to the computational challenges presented by the random field model of ICS, the modern technology of massively parallel computation and its application are introduced in Chapter 5. The use of such technology is not commonly found within the context of statistical computing although a related discussion of the application of the MasPar’s SIMD architecture can be found in [13]. Several interesting numerical comparisons are undertaken including that generated by the limiting variances arising from both the least squares and likelihood method. The variance estimators of Taniguchi [34] and Keenan [14] are reviewed and numerically implemented on the MasPar MP-2. An extension of the block method is also studied and is shown to provide a superior variance estimation algorithm based on the excessive DPU times required by the Taniguchi estimator. A coarse grid version of the Taniguchi-Keenan algorithm is studied and is shown to provide poor results. The practicality of the least squares and block methods is demonstrated through their application to several real ICS data sets. The innovative application, significant potential as well as practical limitations of massively parallel computation to problems in statistical modeling are the central themes of this chapter.

Obtaining asymptotic confidence intervals for the important model parameters of S-FCS and ICS has been one of the principal goals of this work. Such intervals involve the calculation and estimation of various limiting variance expressions, a potentially computationally intensive exercise. One possible method of circumventing this difficulty involves the introduction of a bootstrapping technique based on a modified least squares equation involving higher-order periodograms (cumu-
lants) of the stochastic process. The use of higher-order periodograms and spectra in a related setting is discussed in [6] and [7]. In this way, the parameters of the underlying (unobservable) marking distribution can be identified and estimated leading to direct bootstrap interval estimates of parameters such as the ratio $R$. This idea is the subject of current research.
Bibliography


