A FINITE SAMPLING PLAN, CENTRAL LIMIT THEOREM,
AND BOOTSTRAP ALGORITHM FOR A HOMOGENEOUS AND
ISOTROPIC RANDOM FIELD ON THE 3-DIMENSIONAL SPHERE

(Dissertation)

by

Jason J. Brown

May 1995

Minco Series #1997 T

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina
MIMEO Jason T. Brown
SERIES A FINITE SAMPLING PLAN,
2097T CENTRAL LIMIT THEOREM, &
BOOTSTRAP ALGORITHM FOR A
HOMOGENEOUS AND ISOTROPIC RANDOM
FIELD ON THE 3DIMENSIONAL SPHERE

NAME DATE

Department of Statistics Library
A Finite Sampling Plan, Central Limit Theorem, and Bootstrap Algorithm for a Homogeneous and Isotropic Random Field on the 3-dimensional Sphere

Jason J. Brown

A dissertation submitted to the faculty of the University of North Carolina in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics

Chapel Hill, 1993

Approved by:

__________________  Advisor

__________________  Reader

__________________  Reader
Let \( \{ X_{\vec{p}} \} \) be a real-valued, homogeneous, and isotropic random field indexed in \( \mathbb{R}^3 \); this means that the distribution of any finite number of random elements of the random field is rotation, reflection, and translation invariant. When we are restricted to those indices \( \vec{p} \) with \( || \vec{p} || \), the Euclidean length of \( \vec{p} \), equal to \( r \) (a positive constant), then the random field \( \{ X_{\vec{p}} \} \) resides on the surface of a sphere of radius \( r \). Given a finite sampling design \( \mathcal{R} \) on the sphere \( \{ \vec{p} : || \vec{p} || = r \} \), define \( \{ X_{\vec{p}} : \vec{p} \in \mathcal{R} \} \) to be a realization of the random process and \( |\mathcal{R}| \) to be the cardinality of \( \mathcal{R} \). Without specifying the dependence structure of \( \{ X_{\vec{p}} \} \) nor the marginal distribution of the \( X_{\vec{p}} \)'s, conditions for asymptotic normality of the standardized sample mean

\[
\frac{1}{\sqrt{|\mathcal{R}|}} \sum_{\vec{p} \in \mathcal{R}} (X_{\vec{p}} - \mathbb{E}\{X_{\vec{p}}\})
\]

are given. The conditions, in terms of \( \alpha \)-mixing and a nearly uniform sampling design \( \mathcal{R} \), are motivated by ideas and results for dependent stationary sequences (i.e., time-series). With this sampling design, a bootstrap algorithm is presented and conditions for strong uniform consistency of the bootstrap cumulative distribution function of the standardized sample mean are given. As an application, the sampling design and bootstrap algorithm are applied to global land-area and coastline data.
Contents

1 Introduction ............................................. 1
   1.1 Asymptotic Normality ................................. 3
       1.1.1 Asymptotic Normality for IID Data ................. 4
       1.1.2 Asymptotic Normality for Stationary Sequences in $\mathbb{R}^1$ ................. 4
       1.1.3 Asymptotic Normality for Stationary Random Fields .................. 6
       1.1.4 Asymptotic Normality for Homogeneous and Isotropic Random Fields on the Sphere ............................................. 7
   1.2 Resampling Methods ..................................... 9
       1.2.1 Resampling Methods for Independent Data ................. 9
       1.2.2 Resampling Methods for Dependent Data .................. 12
       1.2.3 Resampling Methods for the Homogeneous and Isotropic Sphere ............. 15
   1.3 Tiling the Sphere ..................................... 16
       1.3.1 Polygons and Polyhedra .......................... 17
       1.3.2 Spherical Tessellations .............................. 19
       1.3.3 A General Tiling of the Sphere .................. 22

2 Sampling Plan ............................................. 24
2.1 Introduction ................................................................. 24

2.2 Stratified Spherical Sampling Plan .................................... 25
   2.2.1 Approximation of the Latitude-Longitude Grid to $Z^2$ .......... 28

3 Asymptotic Normality ......................................................... 32
   3.1 Introduction ............................................................. 32
   3.2 Spatial Dependence .................................................... 32
   3.3 Tile Variance ........................................................... 34
   3.4 Asymptotic Normality .................................................. 38
      3.4.1 Intuition Underlying the Proof ................................. 39
      3.4.2 Big-Tile, Little Tile ........................................... 39
      3.4.3 Asymptotic Normality of the Standardized Sample Mean .... 43

4 Bootstrap Algorithm ......................................................... 46
   4.1 Introduction ............................................................. 46
   4.2 Stratified - Tilewise Bootstrap Algorithm ......................... 46
   4.3 Bounds on Moments ................................................... 49
   4.4 Convergence of the Bootstrap Estimate of Variance of the Sample Mean ........................................... 52
   4.5 Strong Uniform Convergence of the Bootstrap Cumulative Distribution Function of the Standardized Sample Mean ..................... 54

5 Applications ............................................................... 56
   5.1 Analysis of the Earth's Land ....................................... 56
   5.2 A Nearly Uniform Stratified Spherical Sampling Plan for the Earth .... 57
5.3 Stratified - Tilewise Bootstrap Algorithm .......................... 59

5.4 Bootstrap Algorithms with Linking ............................... 61
  5.4.1 Inner and Outer Linking Algorithms ......................... 62
  5.4.2 Distance Scaling ........................................... 63
  5.4.3 Closeness – Nearest Neighbor Systems ..................... 63
  5.4.4 Iteration and the Initial Random Field ..................... 64
  5.4.5 Stratified - Tilewise Outer Link Bootstrap Algorithm .... 65

5.5 Applications to Earth’s Land-Area and Coastline Data ............ 73

A Proofs of Lemmas and Theorems .................................... 78
  A.1 Sampling Design: Proofs of Lemmas and Theorems ............... 78
  A.2 Asymptotic Normality: Proofs of Lemmas and Theorems .......... 80
  A.3 Bootstrap Algorithm: Proofs of Lemmas and Theorems .......... 93

Bibliography .......................................................... 111
List of Figures

2.1 Spherical Dissection and Observation Locations .............................................. 26

3.1 Type l tiles .................................................................................. 40

5.1 Original Global Data ........................................................................... 58

5.2 Stratified - Tilewise Bootstrap Algorithm ..................................................... 60

5.3 Boundary Schematic ............................................................................ 62

5.4 Inner Linking Mechanism ...................................................................... 66

5.5 Nonparametric Metropolis Method .......................................................... 67

5.6 Outer Linking Mechanism ...................................................................... 69

5.7 Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 1 .... 70

5.8 Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 2 .... 71

5.9 Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 3 .... 72

5.10 Kernel Density Estimates and EDFs for Coastline Statistics .................... 75

5.11 Kernel Density Estimate and EDF for $\bar{X}$ ........................................... 76

5.12 Bivariate Plots of Statistics ................................................................... 77
List of Tables

1.1 Regular Polyhedra ............................................. 19

1.2 Orders of Fundamental Regions .............................. 22

2.1 Example of a Stratified Spherical Sampling Plan .......... 31

3.1 Sizes of Different Tile Types ................................. 41

4.1 Specific Nearly Uniform Stratified Spherical Sampling Plans ........ 55

5.1 Wafer Angles and Number of Observations .................. 57

5.2 Bootstrap Estimates of Variance for Different Tile Sizes .... 59

5.3 Univariate Information for Bootstrap Statistics .......... 73

5.4 Correlation Matrix of Bootstrap Statistics .................. 74

A.1 Horizontal and Vertical Separation Angles ................. 85

A.2 Number of Observations for Different Tile Types .......... 87

vii
Chapter 1

Introduction

There are many fields where spherical data arises. One of these is kriging, named after D.G Krige in the early 1950's, and it is the field of study relating to minimizing the estimation variance of some statistic of a random process that usually resides on a fixed finite grid of points on the Earth (Journel [Jou77] and Cressie [Cre86]). Among the fields in which spherical data and kriging have been applied are: oil and gas exploration (Djafarov [Dja89]), soil mapping (Burgess and Webster [BW80]), groundwater geochemical data (Myers, Begovitch, Butz, and Kane [MBBK82]), tillage trials (Cressie and Horton [CH86]), public health data (Cressie and Read [CR86]), and mining estimation (Bárdossy and Bárdossy [BB84]). Other fields include meteorology, forest surveys, and cartography.

When the topic of spherical data does arise, the type of spherical data must first be ascertained. There are two basic types: directional data and random fields which reside on the surface of the sphere. Directional data has been researched at length (Mardia [Mar71], Embleton et al. [EFL87] and Watson [Wat83]). Directional data are actually random directions (i.e., angles) in 2 or 3 dimensions and are depicted as points on the circle or sphere, respectively. For directional data it is the (random) placement of the point that is most important, although the value at this direction can also be taken into account by considering the marginal or conditional distribution for a given magnitude. For example, in 2-dimensions (circle) one might take readings of wind direction from the top of an oil platform and only consider wind speeds in the range of 20 – 30 mph; in 3-dimensions, one might record locations of earthquakes for Richter scale readings of more than 2.0.

A random field that resides on the surface of a sphere is also called spherical data. But in this setting, it is the (random) value at a point of the field that is most important. In fact, it is the random values throughout the space that define the random field (Ji [Ji90]). When a
realization of the random process is observed, the placement of the points helps to determine
the relationship between the random values. The observation sights (the points) are
non-random and comprise the finite sampling design. Since the sampling design helps to
determine the relationship between random values, it should be chosen with care. The field of
kriging uses this type of spherical data.

There are many types of random fields, each with a different way of describing the relationship
between points. A strictly stationary sequence, indexed by the integers \( Z = \{\ldots, -1, 0, 1, \ldots\} \),
for example, has the quality that the distribution of every finite set of observations is
translation invariant. That is, for positive integers \( k \) and \( t \), the distribution of the finite set of
observations \( \{X_{t_1}, X_{t_2}, \ldots, X_{t_k}\} \) is equivalent to the distribution of the set of observations
\( \{X_{t_1+i}, X_{t_2+i}, \ldots, X_{t_k+i}\} \) for \( i_j \in Z, \forall 1 \leq j \leq k \). A homogeneous and isotropic (in the
narrow sense) random field has the additional property that the distribution of any finite set of
points is also rotation and reflection invariant. This relationship can be made precise with the
following definition from Yaglom [Yag57]:

**Definition 1.1** Let \( \mathcal{G} \) be the group of all orthogonal (rotation and reflection) transformations
in \( \mathbb{R}^n \). A random field \( \{X_{\vec{F}}\} \) is called homogeneous and isotropic in the narrow sense if
it is distributionally translation invariant and in addition, \( \forall g \in \mathcal{G} \) and \( \forall \) fixed \( k \) the
distribution of

\[
(X_{\vec{F}_1}, X_{\vec{F}_2}, \ldots, X_{\vec{F}_k}) \overset{D}{\equiv} (X_{g(\vec{F}_1)}, X_{g(\vec{F}_2)}, \ldots, X_{g(\vec{F}_k)}).
\]

Notice that the mention of group is important so that the set has closure in the multiplicative
sense (if we rotate multiple times, it is as if we had only done it once). Recall that in addition
to closure, a group also has the properties of an identity element (no rotation, reflection, nor
translation), and an inverse relation.

For any two elements of a homogeneous and isotropic random field, \( X_{\vec{F}} \) and \( X_{\vec{G}} \), each with
mean 0, let \( \mathbb{E} \{X_{\vec{F}} \cdot X_{\vec{G}}\} = B(\tau_{\vec{F}, \vec{G}}) \) where \( \tau_{\vec{F}, \vec{G}} = ||\vec{p} - \vec{q}|| \) is the Euclidean distance between \( \vec{p} \)
and \( \vec{q} \). That is, the covariance between any two points depends only on the Euclidean distance
between them.

Necessarily, the class of homogeneous and isotropic random fields contains the class of strictly
stationary random fields. For example, a strictly stationary sequence indexed by the integers
has

\[
(X_1, X_2, X_4) \overset{D}{\equiv} (X_5, X_6, X_8).
\] (1.1)

For a homogeneous and isotropic random field indexed by the integers, not only does (1.1)
hold but in addition

\[
(X_1, X_2, X_4) \overset{D}{\equiv} (X_{13}, X_{12}, X_{10}).
\] (1.2)

That is, it is distributionally equivalent to its reflective partner about index 7. The stronger
condition of a homogeneous and isotropic random field on the sphere will be necessary in order
to "exchange" distributionally equivalent pieces of the sphere, as is required of the blocks in model-free bootstrap resampling algorithms for stationary sequences.

The homogeneous and isotropic random field has been studied in depth. Rozanov and Volkonski [RV59] study extreme value theory problems and asymptotic normality of additive functions on a continuously indexed, homogeneous and isotropic random field on an interval; in a follow-up paper Rozanov [Roz60] continues the study of asymptotic normality of additive functions. Popov and Yadrenko [PY69] consider linear extrapolation on a continuously indexed random field on a sphere. And later Yadrenko [Yad79] investigates optimal linear estimates of the sample mean (when \( E \{ X_T \} = \mu \) is unknown and the covariance function is known) and regression coefficients (when both \( \mu \) and the covariance function are known) for a homogeneous and isotropic random field on the boundary of a circle that is either observed continuously or on a set of vertices of a rectilinear polygon inscribed within the circle.

It is natural to consider some function of the random process. Let \( t_n \) be a function of \( n \) real variables (i.e., \( t_n(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^1 \)). In section 1.1, asymptotic normality of the general statistic, \( t_n \), is discussed in a variety of settings: (a) when the \( X_T \)'s are independent and identically distributed (iid), (b) when the random field is strictly stationary in \( \mathbb{R}^d \) for \( d \geq 1 \), and (c) when the random field is homogeneous and isotropic on the sphere in \( \mathbb{R}^3 \).

In section 1.2, a review of resampling methods is undertaken. These methods are outlined for iid data and then modified versions of these algorithms are applied to stationary sequences in \( \mathbb{R}^1 \). The resampling methods for stationary processes in \( \mathbb{R}^1 \) suggest how to resample the sphere in \( \mathbb{R}^3 \). In chapter 2, the ideas for resampling methods and asymptotic normality are used to create a finite sampling plan for the sphere. In chapter 3, a central limit theorem for the standardized sample mean based upon the suggested finite sampling design is given. In chapter 4, a bootstrap algorithm for the sphere based upon the suggested finite sampling design is given and strong uniform consistency of the bootstrap cumulative distribution function of the standardized sample mean is proved. In chapter 5, the proposed spherical bootstrap algorithm and a linked bootstrap algorithm are applied to global land-area and coastline data of the Earth. All of the proofs are given in the Appendix.

### 1.1 Asymptotic Normality

Suppose that for a mean 0 random process, \( \{ X_T \} \), one observes a realization, \( (X_{t_1}, X_{t_2}, \ldots, X_{t_m}) = \vec{X}_m \) for some \( m \geq 1 \). Some characteristic, \( \Theta \), of the process is usually of interest. An estimate of this characteristic can be found by using the realization and some statistic. In particular, let

\[
t_m = t_m(\vec{X}_m)
\]
be the statistic that estimates $\Theta$. A natural question to ask is, what is the limiting distribution of $t_n$ as $m \to \infty$? Is it asymptotically normal or does it converge at all? In this section, results for asymptotic normality of a statistic are discussed for both iid and dependent stationary data.

### 1.1.1 Asymptotic Normality for IID Data

When considering iid data, $\{X_1, X_2, \ldots, X_n\} = \bar{X}_n$, one usually turns attention to the central limit theorem (CLT). Loève [Loe55] and Billingsley [Bil86], have documented the CLT, by giving numerous necessary and sufficient conditions in order for the distribution of the standardized sample mean,

$$\sqrt{n} \bar{X}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \overset{D}{\to} N(0, \sigma^2)$$

as $n \to \infty$ for some $\sigma^2 > 0$. In proving that $\sqrt{n} \bar{X}_n$ satisfies the CLT, the Lindeberg condition or the Lyapounov condition is frequently used for sufficiency and Feller's theorem is used for necessity. Hartigan [Har75] modified some of these conditions in order to include statistics $t_n$ which were not the sample mean, but were sufficiently "mean-like".

### 1.1.2 Asymptotic Normality for Stationary Sequences in $\mathbb{R}^1$

For situations in which the data are dependent, new types of conditions must be found in order to obtain asymptotic normality of $t_n$ or else a new approach must be found in order to "neglect" the dependence of the data at large lags.

Hoeffding and Robbins [HR48] study asymptotic normality of $\sqrt{n} \bar{X}_n$ when the data are $m$-dependent. The condition of $m$-dependence maintains that the covariance between elements at lags larger than $m$ vanishes. The "big block - little block" idea of Bernstein [Ber26] was developed to allow the dependence to exist for large lags, but still be small enough to "neglect". In his paper, Bernstein breaks up the dependent sequence, $\bar{X}_n$, into large blocks of data separated by small blocks of data. His idea is that when $n$ is very large, and if the small blocks are small in size compared to the big blocks, but still large enough to separate the big blocks by a substantial amount, then the big blocks act almost independently while the small blocks are negligible compared with the big blocks.

Rosenblatt [Ros56] borrows this idea for stationary sequences and uses asymptotics, as $n \to \infty$, to solidify the asymptotic normality of $\sqrt{n} \bar{X}_n$. He quantifies the dependence using the idea of $\alpha$-mixing. Define $\mathcal{F}_k^+ = \sigma(X_i(\omega) : i \geq k)$ and $\mathcal{F}_j^- = \sigma(X_i(\omega) : i \leq j)$ to be the
sigma fields generated by \( \{X_k(\omega), X_{k+1}(\omega), \ldots\} \) and \( \{\ldots, X_{j-1}(\omega), X_j(\omega)\} \), respectively. Then the \( \alpha \)-mixing coefficient can be defined as

\[
\alpha(k) = \sup \{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_k^+, B \in \mathcal{F}_0^- \}.
\] (1.3)

A sequence \( \{X_i(\omega) : i \in \mathbb{Z}\} \) is called \( \alpha \)-mixing if \( \alpha(k) \to 0 \) as \( k \to \infty \). What this means intuitively is that the distance between two events is inversely related to their dependence. As the two events are separated by infinite distance, their dependence vanishes (i.e., they act as if they are independent). Using the concepts of \( \alpha \)-mixing and the big block – little block technique, Ibragimov and Linnik [IL71] prove asymptotic normality of \( \sqrt{n} \tilde{X}_n \) for stationary \( \alpha \)-mixing sequences.

**Theorem 1.1** Let \( \{X_t : t \in \mathbb{Z}\} \) be an \( \alpha \)-mixing, strictly stationary sequence with \( \alpha \)-mixing coefficient \( \alpha(k) \). In order for \( \sqrt{n} \tilde{X}_n \) to satisfy the central limit theorem, it is sufficient that

1. \( V \left( \sum_{i=1}^{n} X_i \right) = \sigma_n^2 = O(n) \)

2. \( \exists \) a pair of sequences \( p = p_n \) and \( q = q_n \) such that
   
   (a) \( p \to \infty, q \to \infty, \) and \( q = o(p) \) as \( n \to \infty \),
   
   (b) \( \lim_{n \to \infty} \frac{n \alpha(q)}{p} = 0, \)
   
   (c) \( \forall \beta > 0, \lim_{n \to \infty} \frac{n^{1-\beta}q^{1+\beta}}{p^2} = 0, \) and
   
   (d) \( \forall \epsilon > 0, \) the distribution function \( \tilde{F}(z) = P \left\{ \sum_{i=1}^{l} X_i < z \right\} \), satisfies

\[
\lim_{n \to \infty} \frac{n}{p \sigma^2} \int_{|z| \geq \sigma_n} z^2 d\tilde{F}_p(z) = 0.
\]

Condition (1) requires that the variance of a sum of observations is of the order of the number of elements in the sum. Condition (2a) characterizes the "big-block, little-block" technique with the big blocks having length \( p \) and the little blocks having length \( q \). Condition (2b) uses \( \alpha \)-mixing in order to approximate both the big and little block sums by independent block sums that have the same marginal distribution yet are independent. Condition (2c) is used to ensure the negligibility of the little block sums and condition (2d) is nothing more than a generalization of the Lindeberg condition for the created independent big blocks. Their idea formalizes Bernstein's intuitive feeling that the little blocks are negligible with respect to the big blocks, while the big blocks are asymptotically independent and is used as the motivation for the proof of asymptotic normality of the standardized sample mean on the 3-dimensional sphere.
Carlstein [Car86a] generalizes this result for a variety of "mean-like" statistics, \( t_n \). His results are in the same flavor as Hartigan [Har75], except that he is operating on dependent stationary sequences of random variables. Gastwirth and Rubin [GR75] use different conditions in order to prove asymptotic normality of the statistic, \( t_n \). Sherman [She92] extends these ideas to random fields observed on a \( d \)-dimensional lattice.

### 1.1.3 Asymptotic Normality for Stationary Random Fields

In two and more dimensions the methods of obtaining asymptotic normality of a statistic vary slightly. Some continue on with Bernstein's approach, while others investigate a new approach developed by Stein [Ste73]. Stein uses the pure theoretical approach of tightness of the partial sums of the random variables in order to prove asymptotic normality. Since in two dimensions the size of the field has increased by a power, (i.e., \( \mathbb{R}^2 \)), the notion of \( \alpha \)-mixing, sometimes, gets modified. The definition changes to include the cardinalities of the sets from which the \( \sigma \)-fields are defined. For example, in \( \mathbb{R}^1 \), the cardinalities of the sets which the \( \sigma \)-fields were defined have infinite cardinality.

Before the formal definition of \( \alpha \)-mixing in \( d \)-dimensions is given, let us first describe the notions of a boundary of a set and a boundary condition for this set. Let \( \{X_x\} \) be a stationary random field in \( \mathbb{R}^d \). Let \( \Lambda_n \subset \mathbb{R}^d \) and denote its boundary set by

\[
\partial \Lambda_n = \left\{ \vec{p} \in \Lambda_n : \exists \vec{q} \in \Lambda^c \text{ such that } d(\vec{p}, \vec{q}) = \max_{1 \leq j \leq d} |\vec{p}_j - \vec{q}_j| = 1 \right\}.
\]

If we denote cardinality by \( | \cdot | \) then the boundary condition is

\[
\lim_{n \to \infty} \frac{|\partial \Lambda_n|}{|\Lambda_n|} = 0. \tag{1.4}
\]

This means that the cardinality of the boundary of the set is negligible with respect to the cardinality of the entire set. The modified \( \alpha \)-mixing coefficient is then (Bolthausen [Bol82])

**Definition 1.2** For any two sets \( \Lambda_1, \Lambda_2 \subset \mathbb{R}^d \). Let

\[
d(\Lambda_1, \Lambda_2) = \inf \{ d(\vec{\lambda}, \vec{\lambda}') : \vec{\lambda} \in \Lambda_1 \text{ and } \vec{\lambda}' \in \Lambda_2 \}.
\]

Define the \( \alpha \)-mixing coefficient

\[
\alpha_{k,l}(n) = \sup_{\Lambda_1, \Lambda_2} \left\{ |P \{ A_1 \cap A_2 \} - P \{ A_1 \} - P \{ A_2 \} | \right\}
\]

where \( A_j \in A_{\Lambda_j} = \sigma(X_x : \vec{x} \in \Lambda_j) \) for \( j = 1, 2 \), \( |\Lambda_1| \leq k, |\Lambda_2| \leq l \), and \( d(\Lambda_1, \Lambda_2) \geq n \).

Bolthausen [Bol82] proves asymptotic normality for the standardized sample mean

\[
\frac{1}{\sqrt{|\Lambda_n|}} \sum_{\vec{x} \in \Lambda_n} X_{\vec{x}}
\]
by assuming (a) \( \Lambda_\alpha \) satisfies the boundary condition (1.4), (b) moment conditions on the \( X_f \)'s, and (c) the following \( \alpha \)-mixing conditions

\[
\sum_{m=1}^{\infty} m^{d-1} \alpha_{k,l}(m) < \infty \quad \text{for } k + l \leq 4
\]

\[
\alpha_{1,\infty}(m) = o(m^{-d}).
\]

His proof uses Stein's [Ste73] idea of the tightness of the distribution of the partial sums of the \( X_f \)'s and uniformity conditions on its characteristic function.

Neaderhouser [Nea78], on the otherhand, uses Bernstein's idea with a twist. Usually when using the big block – little block idea, the blocks are rectangles or their higher dimensional analogs. She adapts his idea to spheres of increasing radii, \( r_i \), where alternating annuli are "near" each other; in a sense, her method can be called the "big annuli – little annuli" idea. She investigates the asymptotic normality of the of the standardized sample mean by looking at the standardized sample means of the random variables whose indices fall into each of the big annuli.

1.1.4 Asymptotic Normality for Homogeneous and Isotropic Random Fields on the Sphere

Most of the work done on the homogeneous and isotropic random field on the sphere uses a slightly different set-up; it is usually assumed that the the random field is homogeneous and isotropic in the wide sense. This means that the first and second moments agree, but not necessarily that the distributions of sets of finite elements agree with their rotated and reflected counterparts. Also, since the distributional condition is relaxed, another condition is put in its place, that of the random process being mean-square continuous.

With this alternative set-up on the homogeneous and isotropic random process, \( \{X_f\} \), where \( ||p|| = r \) a constant, the correlation function has a very specific form, as was first shown by Schoenberg [Sch42] and then again by Yaglom [Yag61]. If one is considering a sphere in \( \mathbb{R}^n \) then the form of the correlation function is given by

\[
B(r) = \int_0^\infty Y_n(\lambda r) d\Phi(\lambda)
\]

where \( r \) is the Euclidean distance between \( \tilde{p} \) and \( \tilde{q} \), \( \Phi(\lambda) \) is a bounded, non-decreasing function on \([0, \infty)\),

\[
Y_n(\lambda r) = \left( \frac{2}{\lambda r} \right)^{n-1} \Gamma \left( \frac{n}{2} \right) J_{n-2}(\lambda r),
\]

and \( J_\nu(\cdot) \) is the Bessel function of the first kind defined by

\[
J_\nu(z) = \left( \frac{z}{2} \right)^\nu \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(\nu + m + 1)} \left( \frac{z}{2} \right)^m.
\]
Therefore, in our setting, \( n = 3 \), using formulas from Carrier et al. [CKP65],

\[
B(r_{\theta, \phi}) = \int_0^{\infty} \frac{1}{\lambda r_{\theta, \phi}} \sin(\lambda r_{\theta, \phi}) d\Phi(\lambda).
\]

Notice that the correlation function is completely characterized by the generalized distribution function \( \Phi(\cdot) \). Jones [Jon63] gives the correlation function, defined on the unit sphere in \( \mathbb{R}^3 \) as

\[
B(\gamma_{\theta, \phi}) = \sum_{\nu=0}^{\infty} \frac{2\nu + 1}{4\pi} f_{\nu} P_{\nu}(\cos(\gamma_{\theta, \phi}))
\]

(1.5)

where \( \gamma_{\theta, \phi} \) is the angle between \( \vec{p} \) and \( \vec{q} \), \( P_{\nu}(\cdot) \) is the Legendre polynomial of degree \( \nu \), and \( f_{\nu} \) are some constants to be determined. (This (1.5) is actually the original form given by Schoenberg [Sch42] and then refined by Yaglom [Yag61] using the fact that a correlation function in \( \mathbb{R}^n \) is also a correlation function on the sphere in \( n \)-dimensions and the addition theorem of spherical harmonics. See Yaglom [Yag87] for more details.) Jones [Jon63] estimates \( \Phi(\cdot) \) by estimating the \( f_{\nu}'s \) using the spectral representation of a homogeneous and isotropic random field and a realization of the random process.

Leonenko and Yadrenko [LY79b] give, without proof, sufficient conditions to ensure asymptotic normality of the standardized mean. Let the centered sphere of radius \( r \) in \( n \)-dimensions be denoted \( \mathcal{S}_r^{n-1} = \{ \vec{p} \in \mathbb{R}^n : ||\vec{p}|| = r \} \) and

\[
\eta_r = \int_{\mathcal{S}_r^{n-1}} X_{\vec{p}} \ dm_{\nu}^{(n)}(\vec{p})
\]

where \( m_{\nu}^{(n)}(\cdot) \) is Lebesgue measure on \( \mathcal{S}_r^{n-1} \). Their claim is that the standardized mean

\[
r(\tau - 2n + 3)^{1/2} \eta_r
\]

is asymptotically normal with variance \( \sigma^2(>0) \) as \( r \to \infty \) in \( \mathbb{R}^n \) if \( \Phi(\lambda) = \lambda^r g(\lambda) \) where

1. \(-1 < \tau < n \),
2. \( g(\lambda) \) is continuous in some neighborhood of 0,
3. \( g(0) \neq 0 \), and
4. \( g(\lambda) \) is bounded on \([0, \infty)\).

In their follow-up paper, Leonenko and Yadrenko [LY79a] prove this result. They also give many alternative sufficient conditions to ensure asymptotic normality of the standardized sample mean. These conditions range from \( \alpha \)-mixing with no cardinality restrictions, to smoothness conditions of the generalized distribution function \( \Phi(\cdot) \). One of these is:
Theorem 1.2 Let $S^{n-1}$ be the centered sphere in $\mathbb{R}^n$ of radius $r$. Let $\{X_{\vec{p}} : \vec{p} \in S^{n-1}\}$ be a real-valued random field that is measurable, mean-square continuous, and homogeneous and isotropic in the wide sense. For $C \subset S^{n-1}$, let $\mathcal{A}(C) = \sigma(X_{\vec{p}} : \vec{p} \in C)$. Define the $\alpha$-mixing coefficient
\[
\alpha(\mathcal{A}(C), \mathcal{A}(D)) = \sup \{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{A}(C), B \in \mathcal{A}(D) \} 
\leq \alpha(d(C, D))
\]
Suppose that

1. $-1 < \tau < n - 1$ and $n > 1$,
2. $E \left\{ |X_{\vec{p}}|^{2+\tau} \right\} < \infty$ and $\alpha(d) = O(\lambda^{-1})$ for some $q > 0$ and $p > 2/q$, and
3. $\exists \delta$ such that $0 < \delta < 2(n - 1)$ where
\[
\lim_{\lambda \to 0} \lambda^{-\delta}(\Phi(\lambda) - \Phi(0)) = \Psi_\delta > 0.
\]
Then the standardized sample mean $\bar{X}_n$ converges in distribution to a $\mathcal{N}(0, (2\pi)^n g(0) \gamma_{n-2, \tau})$ random variable as $r \to \infty$, where
\[
\gamma_{n, \tau} = \frac{\Gamma(n - \tau)\Gamma(\frac{n+1}{2})}{2^{n-\tau}\Gamma(\frac{n-\tau+1}{2})\Gamma(\frac{2n-\tau+1}{2})}.
\]

1.2 Resampling Methods

In order to obtain confidence intervals and do hypothesis testing on the statistic, $t_n$, of a set of random variables $\{X_{r_1}, X_{r_2}, \ldots, X_{r_n}\}$, it is necessary to know the sampling distribution of $t_n$. The field of resampling tries to find the distribution by reusing the data at hand, creating "new" samples from which the statistic is again calculated. There are many different types of resampling methods and what follows is but a brief list of those available (for a complete list of time-series resampling methods see Carlstein [Car92]).

1.2.1 Resampling Methods for Independent Data

The following methods are for iid data, and provide an intuitive foundation for the methods used on dependent stationary data. Throughout this section, assume that $\{X_1, X_2, \ldots, X_n\}$ are independent and identically distributed mean-zero random variables and that the statistic of interest, $t_n$, is symmetric in its $n$ ($\geq 1$) arguments.
Jackknife

The jackknife method, first introduced by Quenouille [Que49], "generates" replicates of the statistic, $t_n$, by deleting one observation from the sample, \{X_1, X_2, \ldots, X_n\}, and then computes the statistic on the remaining data. The $i^{th}$ jackknife replicate is given by

$$t_n^{(i)} = t_{n-1}(X_1, X_2, \ldots, X_{i-1}, X_{i+1}, X_{i+2}, \ldots, X_n)$$

for $1 \leq i \leq n$. An estimate of the variance of $t_n$, the jackknife estimate of variance (Tukey [Tuk58]), is

$$\hat{V}_J(t_n) = \sum_{i=1}^{n} \frac{(t_n^{(i)} - \bar{t}_n^{(\cdot)})^2}{n} \cdot (n - 1)$$

where

$$\bar{t}_n^{(\cdot)} = \sum_{i=1}^{n} \frac{t_n^{(i)}}{n}.$$

This jackknife estimate of variance uses the variability among the jackknife replicates to model the true sampling variability of $t_n$. The factor of $n - 1$ in the variance expression is the inflationary price for the jackknife replicates sharing so many observations.

In order for the jackknife to be asymptotically unbiased and consistent, it is necessary that the statistic, $t_n$, have a limiting normal distribution (van Zwet [vZ91]). Nevertheless, the jackknife resampling method can still "fail" when the limiting distribution of the statistic is normal, but not "smooth" (e.g., consider $t_n$ being the sample median). Wu [Wu90] shows that the jackknife can be modified so that it works in this instance if the deletion is of $d > 1$ observations.

Typical Values

Hartigan [Har69] introduced the typical-value principle for constructing nonparametric confidence intervals for an unknown parameter $\theta$. A collection of random variables \{V_1, V_2, \ldots, V_k\} are "typical values for $\theta$" if each of the $k + 1$ intervals between the ordered random variables

$$-\infty \equiv V_0 \leq V_1 \leq V_2 \leq \cdots \leq V_k \leq V_{k+1} \equiv \infty$$

contains $\theta$ with equal probability $\frac{1}{k+1}$.

This procedure is applied to resampling by considering a statistic $t_n$ that estimates the unknown parameter $\theta$. For any subset $S$ of the data indices $\{1, 2, \ldots, n\}$, say $S = \{i_1, i_2, \ldots, i_m\}$ for $m \leq n$, $1 \leq i_j \leq n$, compute the corresponding subset "replicate"

$$t_m(S) = t_m(X_{i_1}, X_{i_2}, \ldots, X_{i_m}).$$
In many situations, the collection of the subset replicates, \( \{ t_m(S_1), t_m(S_2), \ldots, t_m(S_k) \} \) for some \( k > 1 \), are actually typical-values for \( \theta \). This is true, for example, whenever the distribution, \( F \), of each \( X_i \) is continuous and symmetric with center of symmetry \( \theta \), the statistic is the sample mean, and the collection of \( S_k \)'s is the power set of the indices (without the empty set). Notice that \( F \) may be Cauchy (making \( t_m(S_k) \) Cauchy when \( t_m \) is the sample mean). Typical-values can be obtained from subset replicates of statistics \( t_n \) other than the power set of the indices and sample mean, respectively (Hartigan [Har75]).

Bootstrap

The bootstrap method ([Efr79]) uses the empirical distribution function, \( \hat{F}_n \), of the random sample \( \{X_1, X_2, \ldots, X_n\} = \bar{X}_n \) to generate replicates of the statistic, \( t_n \). Given \( \hat{F}_n \), which assigns mass \( \frac{1}{n} \) to each observation \( X_i \) for \( 1 \leq i \leq n \), a bootstrap sample of size \( m > 1 \), \( \{X_1^*, X_2^*, \ldots, X_m^*\} = \bar{X}_m^* \) is generated by iid sampling from \( \hat{F}_n \). The corresponding bootstrap replicate of \( t_m \) is given by \( t_m^* = t_m(\bar{X}_m^*) \). For fixed data, \( \{X_1, X_2, \ldots, X_n\} \), is it possible to generate arbitrarily many bootstrap replicates of \( t_n \). The conditional distribution (given the original data) of the bootstrap replicates is used to model the true sampling distribution of \( t_n \).

An estimate of the variance of \( t_n \) is given by the bootstrap estimate of variance, \( \hat{V}_B(t_n) = E^* \left\{ \left( t_m^* - E^* \{ t_m^* | \bar{X}_n \} \right)^2 | \bar{X}_n \right\} \), \( (1.7) \)

where \( E^* \{ \cdot | \bar{X}_n \} \) is expectation with respect to \( \hat{F}_n \).

In many situations, the bootstrap replicates are used to estimate the entire distribution of \( t_n \), not just moments. To obtain a valid estimate, choices must be made for the bootstrap sample size \( m = m_n \) and for standardizations \( (a_n, b_n) \) and \( (a_m^*, b_m^*) \) of \( t_n \) and \( t_m^* \), respectively.

Specifically, the bootstrap estimate of \( P \{ t_n \leq x \} \) is \( P \{ \tilde{t}_n^* \leq x | \bar{X}_n \} \) where \( \tilde{t}_n = a_n(t_n - b_n) \) and \( \tilde{t}_m^* = a_m^*(t_m^* - b_m^*) \). This leads to the following definition of convergence of the bootstrap distribution of a statistic to that of the actual asymptotic distribution of the statistic.

Definition 1.3 The bootstrap cumulative distribution function of a statistic \( t_n \) is said to be strongly uniformly consistent if

\[ \sup_{x \in \mathbb{R}} \left| P \{ t_m^* \leq x | \bar{X}_n \} - P \{ \tilde{t}_n \leq x \} \right| \overset{\text{a.s.}}{\to} 0 \text{ as } n \to \infty. \]

When the statistic is the sample mean, then strong uniform consistency of the bootstrap cumulative distribution function has been investigated and substantiated by Bickel and Freedman [BF81]. They also discuss statistics that are von Mises functionals and prove strong uniform consistency. Singh [Sin81] also investigates strong uniform consistency of the bootstrap cumulative distribution function for statistics that are asymptotically normal and
then shows that for the sample mean, the bootstrap cumulative distribution function is second order correct (i.e., the convergence holds with a rate of $\sqrt{n}$). Swanepoel [Swa86] proves strong uniform consistency of the cumulative distribution function for a modified bootstrap when the statistic of interest is the sample maximum (having a limiting exponential distribution).

1.2.2 Resampling Methods for Dependent Data

When the data are dependent, but stationary, there are two scenarios that can occur: either the dependence structure of \( \{X_1, X_2, \ldots, X_n\} \) is known or it is not. If the dependence structure is known, then that mechanism should be incorporated into the resampling method to generate replicates of the statistic \( t_n \). If the dependence structure is not known, then a "new" method of generating subsamples needs to be found.

Model-based Resampling Methods

First consider the case when the dependence structure is known. Since the dependence is known, this knowledge should be used in setting up the resampling algorithm. "Model-based" resampling usually comes down to using the data to obtain an estimate of the model and then using this model to obtain residuals. Then iid resampling is applied to the residuals.

Consider the situation where the data is autoregressive with lag \( k \) dependence (Freedman [Fre84]). That is, the model is assumed to be

\[ X_t = g(X_{t-1}, X_{t-2}, \ldots, X_{t-k}; \tilde{\beta}) + \epsilon_t \]

where \( g(\cdot) \) is a known function, \( \tilde{\beta} \) is a vector of unknown parameters, and the additive unobservable errors \( \{\epsilon_i : i \in \mathbb{Z}\} \) are iid with unknown distribution function \( F_i \) having mean zero.

A model-based bootstrap resampling method for this situation, begins by obtaining estimates of \( \tilde{\beta} \), say \( \hat{\beta} \), from the data. Next, the residuals are calculated by substituting these estimates into the model. That is, the residuals are found from the recursion

\[ \hat{\epsilon}_i = X_i - g(X_{i-1}, X_{i-2}, \ldots, X_{i-k}; \hat{\beta}) \]

for \( k + 1 \leq i \leq n \). The residuals are then treated as being iid and resampled via the empirical distribution function \( \hat{F}_i \) which puts mass \( \frac{1}{n-k} \) on each of the centered residuals

\[ \tilde{\epsilon}_i = \hat{\epsilon}_i - \sum_{j=k+1}^{n} \frac{\hat{\epsilon}_j}{n - k} \]
for \( k + 1 \leq i \leq n \). The bootstrap sample, \( \{X_1^*, X_2^*, \ldots, X_m^*\} \), is then generated from the estimated dependence model

\[
X_i^* = g(X_{i-1}^*, X_{i-2}^*, \ldots, X_{i-k}^*; \tilde{\beta}) + \epsilon_i^*
\]

for \( 1 \leq i \leq m \), where the \( \epsilon_i^* \) are independently drawn from \( \tilde{F}_\epsilon \). Finally, the bootstrap statistic, \( t_m^* \), is calculated by substituting the bootstrap sample for the actual data in the statistic \( t_m \).

A second method occurs in situations where the dependence structure is known to come from a first order Markovian process with unknown transition density \( f(x_1|z_0) \). This method of resampling (Rajarshi [Raj90]) begins by computing the bivariate kernel density estimate \( \hat{f}(|\cdot|) \) of \( f(x_1|z_0) \) from the data and selecting \( X_1^* \) at random from \( \hat{F}_n \). Successively generate \( X_2^*, X_3^*, \ldots \) by using the estimated transition density. Substitute this bootstrap sample into the statistic, \( t_n \), to form the bootstrap statistic, \( t_n^* \). Rajarshi [Raj90] proved that this resampling algorithm is strongly uniformly consistent when \( t_n \) is the sample mean for suitably chosen \( \hat{f}(|\cdot|) \).

Both of the above methods are modifications of the bootstrap. The following algorithm was proposed by Brillinger [Bri66] for jackknifing time-series data. He suggests that when the \( i^{th} \) observation is deleted, it should be considered a “missing value” and should be replaced by interpolation. In this manner, he obtains \( n \) “complete” jackknife replicates, \( t_n^{(i)} \). When the statistic, \( t_n \), itself is expressible as a function of ARMA residuals (e.g., \( \sum_{j=1}^{n} \frac{\epsilon_j^2}{n} \)), then Davis [Dav77] considers applying the iid jackknife directly by deleting residuals.

Model-free Resampling Methods

“Model-free” resampling does not try to estimate the dependence structure. On the contrary, it resembles the approach of iid resampling methods, by creating and resampling “objects” which are identically distributed. The objects, for a strictly stationary sequence, are blocks or subseries of the sequence, which retain the dependence structure of the series and are guaranteed to be identically distributed by strict stationarity. The blocks are not independent, but their dependence can be quantified with \( \alpha \)-mixing (definition given above).

In this section, we will use the following notation: for a sequence \( \{X_1, X_2, \ldots, X_n\} \), denote a block of observations of length \( l \), starting at position \( i + 1 \), by \( \hat{X}_i^l \). That is, for \( 0 \leq i < i + l \leq n, \hat{X}_i^l = (X_{i+1}, X_{i+2}, \ldots, X_{i+l}) \). In order for (a) the blocks to eventually reflect the dependence structure of the entire time series, (b) the blocks to be asymptotically independent (in the \( \alpha \)-mixing sense), and (c) the number of blocks to increase without bound, it is necessary to assume that

1. \( l = l_n \rightarrow \infty \) as \( n \rightarrow \infty \) and
2. \([p_n] = k_n \to \infty\) as \(n \to \infty\)

where \([\cdot]\) denotes integer part.

The most simplistic way to generate replicates of \(t_n\) is by calculating the statistic on the blocks themselves. Carlstein [Car86b] proposes using the non-overlapping blocks 
\[
\{\tilde{X}_{li}^i : 0 \leq i \leq k_n - 1, 1 \leq l \leq n\}
\]
with corresponding statistics 
\[
\{t_{il}^i = t_i(\tilde{X}_{li}^i) : 0 \leq i \leq k_n - 1, 1 \leq l \leq n\}.
\]
The blocks of length \(l\) are asymptotically independent, except for the two neighboring blocks, and also reflect the whole process as \(n \to \infty\). Carlstein [Car89] makes this typical values idea precise by saying that \((t_{il}^i + t_{il}^j)/2\) for \(1 \leq i \leq k_n\) are typical values. Estimates of moments and other qualities of the sampling distribution of \(t_n\) can then be found by using empirical estimates based on the subseries replicates. This method has been extended to spatial processes by Possolo [Pos91].

A second method for resampling blocks has been introduced by Künsch [Kun89] and is similar to the iid jackknife. The “blockwise jackknife” generates replicates of \(t_n\) by deleting blocks of length \(l\). The \(i^{th}\) jackknife replicate is
\[
t_{n-i}^{\langle i \rangle} = t_{n-i}(X_1, X_2, \ldots, X_i, X_{i+l+1}, X_{i+l+2}, \ldots, X_n)
\]
for \(0 \leq i \leq n - l\). The resulting estimate of the variance of \(t_n\) is then
\[
\hat{V}_{BJ}(t_n) = \sum_{i=0}^{n-l} \frac{(t_{n-i}^{\langle i \rangle} - \overline{t_n^{\langle \cdot \rangle}})^2}{n-l+1} \cdot c_{n,i}
\]
where
\[
\overline{t_n^{\langle \cdot \rangle}} = \sum_{i=0}^{n-l} \frac{t_{n-i}^{\langle i \rangle}}{n-l+1}
\]
and \(c_{n,i}\) is an appropriate standardizing constant. He shows that \(\hat{V}_{BJ}(t_n)\) is consistent when \(t_n\) belongs to a certain class of asymptotically normal functional statistics (including the sample mean).

Along with the blockwise jackknife, Künsch [Kun89], also introduced the “blockwise bootstrap”. One difference between this resampling algorithm and the subseries resampling algorithm of Carlstein [Car86b] is that, here, many blocks of data overlap and hence are not asymptotically independent. His method is as follows: for a fixed block length, \(l\), construct the “empirical \(l\)-dimensional marginal distribution”, \(\hat{F}_{n,l}\), that puts mass \(\frac{1}{n-l+1}\) on each block \(\tilde{X}_i^l\) for \(0 \leq i \leq n - l\). Assuming that \(k = k_n = \frac{n}{l}\) is an integer, generate \(k\) bootstrap blocks, each denoted by \(\tilde{X}_{ij}^* = (X_{j-1+l+1}, X_{j-1+l+2}, \ldots, X_{j,l})\), by iid resampling from \(\hat{F}_{n,l}\). The blockwise bootstrap sample, \(\tilde{X}_{n}^*\), is then constructed by appending the blocks together; that is, 
\[
\tilde{X}_{n}^* = (\tilde{X}_{1,1}^*, \tilde{X}_{1,2}^*, \ldots, \tilde{X}_{1,k}^*).
\]
The bootstrap statistic of \(t_n\) is then given by \(t_{n}^* = t_n(\tilde{X}_{n}^*)\) and its
variance is given according to (1.7) with the expectation being with respect to $\hat{F}_{n,t}$. He shows that when $t_n$ is the sample mean, the blockwise bootstrap cumulative distribution function is strongly uniformly consistent. In addition, his method leads to a smaller variance of the estimate of variance of $t_n$ (when $t_n$ is the sample mean) than Carlstein [Car86b], by a factor of $\frac{1}{3}$.

The final type of model-free resampling method, to be discussed here, is called the "linked-blockwise bootstrap" (see Carlstein [Car92]). Developed by Künsch and Carlstein [KC90], the linked-blockwise bootstrap modifies the blockwise bootstrap method by altering the incorrect dependence structure near the block endpoints. The main difference in their algorithm comes when the new blocks are drawn from the empirical $l$-dimensional marginal distribution, $\hat{F}_{n,t}$. Instead of resampling from all of the blocks, they only sample from the blocks that are the "$p$ nearest neighbors" to the previous block.

That is, when the first bootstrap block is drawn from $\hat{F}_{n,t}$, $\bar{X}_{i+1}^*$, look at the last observation, $X_i^*$, and then find its $p$ nearest neighbors from among the original data, $\{X_1, X_2, \ldots, X_n\}$. Randomly select one of these $p$ nearest neighbors, say $X_{i'}$. The second bootstrap block will be the block that starts immediately after this observation, namely $\bar{X}_{i+2}^* = \bar{X}_{i'}^*$. Continue this linking process until $k$ bootstrap blocks have been sampled. The linked-blockwise bootstrap guarantees a "smoother" transition from block to block. Notice that as $p$ gets smaller the transition gets smoother.

1.2.3 Resampling Methods for the Homogeneous and Isotropic Sphere

To date, there is no resampling method for the sphere. By using the ideas developed from previously mentioned resampling algorithms, a resampling plan for the sphere needs to be found. In particular, since the estimation of the dependence model can lead to error, and results from using an incorrect model are devastating, a model-free resampling algorithm for the sphere needs to be determined.

Since all of the model-free methods involve breaking up the random process into blocks, a way to break the sphere up into "blocks" (hereafter called tiles) needs to be found. These tiles must have the following characteristics:

1. The sampling design on each tile must be identical. This condition on a homogeneous and isotropic field (in the narrow sense) will allow topologically equivalent tiles, which have the same dependence structure, to be exchanged. (We will have classes of tiles where this quality holds.)
2. The number of observations on each tile must asymptotically grow without bound. This allows each tile to preserve the dependence structure of the entire process, since the dependence structure only depends upon distance. (This is analogous to $t_n \to \infty$.)

3. The size of the tile compared to the size of the sphere must be small enough so that the sphere may be broken up into a large number of tiles. (This is analogous to $\frac{t_n}{R} \to \infty$.)

4. It must be possible to put a separating strip between tiles. This allows the spherical analogy of Bernstein's idea to work.

5. And finally, it is most important to require that the resampled tiles form a sphere, since the statistic, $t_n$, could depend on the curvature of the sphere.

Characteristic (5) is also visually appealing, but this automatically rules out using the jackknife ("leave-one-out") method since the reconstructed object would not be a complete sphere. A method similar to the jackknife interpolation method of Brillinger [Bri66] could be employed, but this would mean that the dependence structure of the model would have to be estimated.

Therefore consider a model-free bootstrap type resampling method. In the next section, methods of covering the sphere that satisfy all of the conditions given above are investigated. The covering, in fact, helps to determine the sampling design on each individual tile which is given in Chapter 2.

1.3 Tiling the Sphere

Since most of the model-free resampling plans for strictly stationary sequences break the random process into blocks, in order to maintain the dependence structure of the entire random process, a way of breaking the sphere up into tiles needs to be found. This theory is called tiling.

In an intuitive setting, the iid bootstrap treats every observation the same; that is, it assigns a probability of $\frac{1}{n}$ to each observation. For a strictly stationary sequence, the blockwise bootstrap treats each block of $l$ observations, $X_l$, the same; assigning probability $\frac{1}{n-1+l}$ to each. The reasoning behind this, is that each object, either the observation or the block, is interchangeable with any other; and each object has the same marginal distribution.

Therefore, an analogous approach for tiling the sphere is to have all of the tiles with the same shape, size, and number of observations, while arranging the observations identically on each tile. This set-up combined with the quality of a homogeneous and isotropic random field in the narrow sense (from now on, any homogeneous and isotropic random field will be assumed to be in the narrow sense), allows tiles with the same marginal distribution to be interchanged.
As is shown in section 1.3.2, the sphere can be broken up into at most 120 reflectively equivalent spherical triangles, which satisfy all of the requirements in section 1.2.3, except for condition (3). There are tilings of the sphere into infinitely many topologically equivalent spherical triangles. Unfortunately, they all meet at a single point, violating condition (4). These triangles are called lunes and semi-lunes.

Lunes are degenerate spherical triangles and are wedges from the north pole to the south pole. They are degenerate because one of their angles has measure \( \pi \). Subdividing lunes at the equator yields semi-lunes. These are not degenerate triangles, but they have the same unfortunate (in the dependence sense) property that on each hemisphere all of the semi-lunes meet at one point, namely a pole (lunes have this quality at both poles, for a complete definition see section 1.3.2). So, if the sphere were to be tiled with either lunes or semi-lunes and Bernstein's approach was employed, then all of the tiles would be dependent near the poles, since there could be no small separating tile. Also, there is no obvious way to take observations in these regions so that the points are "uniformly dispersed". Nevertheless, if we could come up with a uniform sampling plan for these regions so that the number of observations near the poles is negligible, then these two tiling methods are plausible.

In section 1.3.3, a general tiling of the sphere that eliminates the weighting problem at the poles is introduced. In fact, the semi-lune tiling is a special case of the general tiling. This general tiling breaks the sphere up into an unlimited number of spherical rectangles (except at the poles) which are interchangeable within certain strata (although tiles across strata are not interchangeable).

1.3.1 Polygons and Polyhedra

We begin the discussion on tiling by defining regularity and basic structures in 2-dimensions, extending them to regular and non-regular three-dimensional objects, and finally showing how they apply for tiling the sphere. All of the results of this section can be found in Coxeter [Cox63a], Coxeter [Cox63b], and Melzak [Mel83].

Regular Objects in 2- and 3-dimensions

Since the sphere is tiled with spherical polygons, the definition of the polygon is a good starting point. This definition is then refined for a regular polygon.

Definition 1.4 A \( p \)-gon, polygon, is a circuit of \( p \) line segments \( A_1A_2, A_2A_3, \ldots A_pA_1 \), joining consecutive pairs of \( p \) points \( A_1, A_2, \ldots A_p \). The segments and points are called sides and vertices, respectively.
Definition 1.5 A p-gon in 2-dimensions is said to be regular if it is both equilateral and equiangular and is denoted \( \{p\} \). Each exterior angle of \( \{p\} \) is \( \frac{2\pi}{p} \), and the interior angle is the supplement, \( (1 - \frac{2}{p})\pi \).

Examples of regular polygons are: equilateral triangle, square, and the equilateral and equiangular pentagon. From these definitions, a simple translation to 3-dimensions can be made and the definition of a regular polyhedron may be given.

Definition 1.6 A p-gon in 3-dimensions, polyhedron, is said to be regular if its faces are regular and equal, while its vertices are all surrounded alike. If its faces are \( \{p\} \)'s, \( q \) surrounding each vertex, the regular polyhedron is denoted by the Schlälfi symbol \( \{p, q\} \).

For a regular polyhedron \( \{p, q\} \), the solid angle at a vertex has \( q \) face-angles, each \( (1 - \frac{2}{p})\pi \). These \( q \) angles must total less that \( 2\pi \) or else they would lie in a plane. Hence, \( q(1 - \frac{2}{p})\pi < 2\pi \).

That is,

\[
\frac{1}{p} + \frac{1}{q} > \frac{1}{2}
\]

or \( (p - 2)(q - 2) < 4 \); thus \( \{p, q\} \) cannot have any values other than \( \{3, 3\} \), \( \{3, 4\} \), \( \{4, 3\} \), \( \{3, 5\} \), and \( \{5, 3\} \). These are known as the Platonic solids and have the names tetrahedron, octahedron, cube, icosahedron, and the dodecahedron, respectively. (The same type of "angular sum" argument will be employed in obtaining all of the reflective and rotational spherical tilings.)

The definition of a regular polyhedron is confusing since it is difficult to determine what is meant by "its vertices are all surrounded alike" and what the solid angle is. The definition can be simplified using the notion of vertex figures.

Definition 1.7 The vertex figure at the vertex \( \bar{O} \) of a polygon is the segment joining the midpoints of the two sides through \( \bar{O} \). The vertex figure at the vertex \( \bar{O} \) of a polyhedron is the polygon whose sides are the vertex figures of all the faces that surround \( \bar{O} \).

For example, the vertex figures of the tetrahedron, cube, and octahedron are a triangle, triangle, and square, respectively. The regular polyhedron can now be defined by only considering vertex figures and the faces of the polyhedron.

Definition 1.8 A polyhedron is said to be regular if its faces and vertex figures are both regular. For a regular polyhedron \( \{p, q\} \), each face is a \( \{p\} \) and each vertex figure is a \( \{q\} \).

For any regular polyhedron \( \{p, q\} \), the number of vertices \( N_0 \), the number of edges \( N_1 \), and the number of faces \( N_2 \) can be quantified. Coxeter [Cox83a] gives the following relationships between \( p, q \) and \( N_0, N_1, \) and \( N_2 \).

\[
N_0 = \frac{4p}{4 - (p - 2)(q - 2)}
\]
\[ N_1 = \frac{2pq}{4 - (p - 2)(q - 2)} \]
\[ N_2 = \frac{4q}{4 - (p - 2)(q - 2)} \]

Table 1.1 gives an enumeration of these quantities for all regular polyhedra or Platonic solids.

<table>
<thead>
<tr>
<th>Regular Polyhedron</th>
<th>Schläfi Symbol</th>
<th>(N_0)</th>
<th>(N_1)</th>
<th>(N_2)</th>
<th>Vertex Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>tetrahedron</td>
<td>{3,3}</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>triangle</td>
</tr>
<tr>
<td>octahedron</td>
<td>{3,4}</td>
<td>6</td>
<td>12</td>
<td>8</td>
<td>square</td>
</tr>
<tr>
<td>cube</td>
<td>{4,3}</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>triangle</td>
</tr>
<tr>
<td>icosahedron</td>
<td>{3,5}</td>
<td>12</td>
<td>30</td>
<td>20</td>
<td>pentagon</td>
</tr>
<tr>
<td>dodecahedron</td>
<td>{5,3}</td>
<td>20</td>
<td>30</td>
<td>12</td>
<td>triangle</td>
</tr>
</tbody>
</table>

Table 1.1: Regular Polyhedra

Non-Regular Objects in 3-dimensions

Not all polyhedra are regular. In fact, there is one type of non-regular polyhedron which deserves mention, called the zonohedron. A zonohedron is a polyhedron with all of its faces being identical \(p\)-gons. Their faces are not regular in the sense that they are not equiangular nor equilateral, but all of the faces are topologically equivalent.

The class of zonohedrons where the \(p\)-gons are parallelograms is of particular interest. This is because for a regular polyhedron, there are only five Platonic solids, but there are infinitely many parallelogram zonohedrons. That is, there are parallelogram zonohedrons for an ever increasing number of faces. The unfortunate thing about this type of zonohedron is that they are very difficult to construct once the number of faces gets large. Nevertheless, a variation of the parallelogram zonohedron will be used to create a general tiling of the sphere.

1.3.2 Spherical Tessellations

A tessellation is simply a set of polygons fitting together to cover the whole space just once; so that every side of each polygon belongs to one and only one other polygon. A spherical tessellation is a tessellation which covers the sphere and the polygons are spherical polygons.

To obtain a clear understanding of a spherical polygon, assume that the radius of the sphere is 1. If the circumcircle of a regular polygon \(\{p\}\) is regarded as lying on the sphere, each edge of \(\{p\}\) can be replaced by a great circle arc joining the same vertices and obtain a regular spherical polygon, denoted \(<p>\).
In particular, if the circumcircle itself is a great circle, then all of the edges of \(<p>\) are arcs on this same great circle, all of length \(\frac{2\pi}{p}\). This great circle decomposes the sphere into two hemispheres and hence is a spherical tessellation. Each hemisphere is called a **monogon**, denoted \(<1>\); the monogon is the most basic type of spherical triangle with all angles equal to \(\pi\).

A **digon**, \(<2>\), is a lune consisting of two antipodal vertices (like the north and south poles on the globe) joined by two distinct great circles (like two distinct meridians of the globe). The angle, \(\phi\), formed by the two great circles at the pole can be anywhere from 0 to \(\pi\), with \(\pi\) being a monogon. If digons, having the same two vertices and same angle, are “stacked” together to cover the entire sphere, then using the fact that the total surface area of the sphere is \(4\pi\), the area of a digon is found to be \(2\phi\). This fact is used to deduce that the area of a spherical triangle with angles \(\alpha\), \(\beta\), and \(\gamma\) is

\[
\Delta = \alpha + \beta + \gamma - \pi,  \tag{1.9}
\]

which one can see is the “excess” of the sum of the angles with \(\pi\); and is commonly referred to as the spherical excess. (Notice, that for a triangle in the plane, this reduces to a point at \(\Delta = 0\) .) The number of tiles, or the **order** of the tiling, of a tessellation will be determined by using (1.9).

With the ideas of regular spherical polygons and the area of a spherical triangle in hand, the notion of a regular spherical tessellation can be considered. One way to create a spherical tessellation is by projecting the edges of the regular polyhedron \(<p,q>\) from its center to its concentric sphere; a regular spherical tessellation, denoted \(<p,q>\), is obtained. The spherical polygons are \(<p>\)'s each of interior angle \(\frac{2\pi}{q}\). This tessellation has the property that all of the edges of \(<p,q>\) form great circles on the sphere and that the faces, \(<p>\)'s, are rotationally and reflectorily equivalent. In a tessellation such as this, the faces are called fundamental regions of the tessellation (Coxeter [Cox63a]).

**Definition 1.9** For any group of reflection and rotation transformations of a plane (or space), a **fundamental region** is a region whose transformations just cover the plane (or space), without overlapping and without interstices.

A different way to create spherical tessellations is by using the idea of a kaleidoscope or virtual mirrors, all of the mirrors going through the center of the sphere. For this setting, let a spherical \(n\)-gon be a fundamental region of a spherical tessellation and tile the sphere with its reflections and rotations. It is important to realize the fundamental region need not be a regular \(n\)-gon (e.g., semi-lunes) and that only spherical triangles, or 3-gons, need to be considered as fundamental regions.

Using virtual mirrors, we only need to consider 3-gons as fundamental regions because if the
fundamental region were a spherical n-gon, with the sides of possibly unequal length and 
\( n \geq 4 \), then the interior angle sum is greater than the interior angle sum of a planar n-gon, 
which is \((n - 2)\pi\). This would mean that at least one of the interior angles has angle more 
than \((1 - \frac{2}{n})\pi\); and for \( n \geq 4 \) this means that this angle is obtuse. By considering virtual 
mirrors, Coxeter [Cox63a] deduces that for any fundamental region, all of the interior angles 
must be submultiples of \( \pi \) (i.e., \( \frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4}, \ldots \); none of them obtuse). Therefore, only spherical 
triangles need be considered as fundamental regions.

Denote the spherical triangle with angles \( \frac{\pi}{p}, \frac{\pi}{q}, \) and \( \frac{\pi}{r} \) as \((p, q, r)\) where \( p, q, \) and \( r \) are integers 
and \( p \geq q \geq r \). The most basic fundamental region is the monogon \(<1>\) or \((1, 1, 1)\). Obviously, 
it only takes 2 monogons to tile the sphere, so \((1, 1, 1)\) is said to have order 2.

The next logical region is the digon with angle \( \frac{\pi}{2} \) at the poles. This digon, denoted \((2, 2, 1)\), is 
of order 4. Clearly, the angle \( \frac{\pi}{2} \) can be replaced by any submultiple of \( \pi \) and hence in general 
we have the digon \((p, p, 1)\) with angle \( \frac{\pi}{p} \) of order \( 2p \), for \( p \geq 2 \). Therefore, the sphere can be 
tiled with an unlimited number of digons or lunes. Notice that the digon is a degenerate 
triangle in the sense that one on the angles is \( \pi \), implying that it has only two sides.

Similar claims can be made about semi-lunes \((p, 2, 2)\). Since they have half the area of a digon 
\((p, p, 1)\), clearly, it has order \( 4p \). So again, there is a spherical tiling by semi-lunes of unlimited 
order. Notice that there are 2 right angles and 3 sides for any semi-lune.

Next consider the case when the spherical triangle has only one right angle, \((p, q, 2)\). 
Substituting these values into the formula for spherical excess \((1.9)\), it is found that

\[
\frac{1}{p} + \frac{1}{q} > \frac{1}{2}
\]

or \((p - 2)(q - 2) < 4\). (The same equation for finding Platonic solids.) This leads to the 
spherical triangles \((3, 3, 2)\), \((4, 3, 2)\), and \((5, 3, 2)\). (Recall that \( p \geq q \geq r \).) Notice that from the 
formula for spherical excess, it is impossible to have \( p \geq q \geq r > 2 \) because then the angle sum 
would be too small.

Notice that when you reflect a spherical triangle about a mirror, the reflected image is 
typically not topologically equivalent to the original object; that is, all of the tiles are not 
interchangeable. Nevertheless, they do provide a way of determining a uniform dispersal of 
points on the sphere.

The order of these fundamental regions is simply the ratio of the area of the total sphere, \( 4\pi \), 
to the area of the spherical triangle. This is because the elements of the reflection group are 
simply transforms of the fundamental regions and the areas are translation invariant.
<table>
<thead>
<tr>
<th>Fundamental Region</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1,1,1)$</td>
<td>2</td>
</tr>
<tr>
<td>$(1,p,p)$</td>
<td>$2p$</td>
</tr>
<tr>
<td>$(2,2,p)$</td>
<td>$4p$</td>
</tr>
<tr>
<td>$(3,3,2)$</td>
<td>24</td>
</tr>
<tr>
<td>$(4,3,2)$</td>
<td>48</td>
</tr>
<tr>
<td>$(5,3,2)$</td>
<td>120</td>
</tr>
</tbody>
</table>

Table 1.2: Orders of Fundamental Regions

Therefore, the order is given by

$$\text{Order}((p, q, r)) = \frac{4pqr}{pq + pr + qr - pqr}.$$  \hfill (1.10)

Table 1.2 enumerates the triangular spherical tessellations using (1.10) and the notion of virtual mirrors.

Putting these together with the spherical tessellations generated by the blown-up Platonic solids with order equal to $N_2$ (see Table 1.1), it is evident that the highest order of a spherical tessellation is $4p$ for semi-lunes. Recall that digons and semi-lunes have the unattractive property that the tiles meet at two points and a single point, respectively. Therefore, the maximum order of spherical tessellation without this quality is 120. The general tiling approach described next combines the tiling of the semi-lunes and that of the zonohedra into one spherical tiling. Each of the tiles in this approach is not a fundamental region for the entire sphere, but it is for a section of the sphere. Using this tiling, we will be able to “neglect” the points where the semi-lunes meet.

1.3.3 A General Tiling of the Sphere

From the previous section, we have found that there is no tiling of the sphere into an unlimited number of topologically equivalent tiles that do not meet at a single point or pair of points. We make a generalization of the lune tiling so that the requirements given above are satisfied on sections of the sphere, called wafers.

The following procedure allows for the tiling of the sphere into an unlimited number of spherical rectangles (except at the poles), with all the spherical rectangles in a given wafer being interchangeable. As long as the number of tiles in each wafer can grow without bound, then we have an acceptable resampling plan (i.e., for the blockwise bootstrap the number of blocks increases without bound). The general tiling is as follows:

- Slice the sphere up into wafers, cut parallel to the “equator”, with equatorially reflective
wafers having the same width. (Equatorial reflection means that the object is reflected about the equator.) These cuts are like the latitudinal lines on the globe.

• Partition each of the two topologically equivalent wafers into an equal number of topologically equivalent tiles. This is done, locally, like the longitudinal lines on the globe. The angle between these longitudinal lines is the same within a wafer, but vary from wafer to wafer.

• Arrange the observations, from the random process, on each tile of a given wafer in the same manner. This is the sampling plan of a tile within a wafer and is discussed in chapter 2.

This spherical tessellation breaks the sphere up into spherical rectangles (except at the poles) which are interchangeable within the two topologically equivalent, equatorially symmetric wafers. Only these two wafers are interchangeable. Since this type of approach allows the sphere to be broken up into as many wafers as needed and each wafer can be sliced into as many tiles as necessary, it should make asymptotic theory possible. This tessellation is made precise in chapter 2.
Chapter 2

Sampling Plan

2.1 Introduction

From a practical point of view, there are several fundamental differences between data on a
sphere and data on a line or plane; one of these is the sampling plan. In order to take
advantage of the stationarity, the sampling plan should be uniform in the space spanned by
the process. For example, the typical sampling plan for a time-series is uniform on the line;
that is, a time-series is usually observed at integer intervals. Sampling plans for a planar
random field are usually the integer lattice or the hexagonal lattice (Matérn [Mat86]).

As a result of the geometrical properties of the sphere, however, there is no uniform spherical
sampling plan $\mathcal{R}$ when $|\mathcal{R}|$ is large. The idea of finding a finite sampling plan is directly
related to the question of tiling and has been studied from both the pure mathematical view
(Grünbaum and Shephard [GS87]) and applied geoscience view (Baumgardner and
Frederickson [BF85]). Since there is no generally accepted spherical sampling plan, our first
objective is to determine a suitable one. Some characteristics that it should have are

1. The points $\mathcal{P} \in \mathcal{R}$ should be symmetrically and approximately uniformly dispersed over
   $S^3$ (to use the isotropy of the random field),

2. There should be a natural way to group the data into similar "blocks" (in order to use
   the "big block - little block" technique [section 1.1.3]) such that

   (a) The number of observations in a block grows without bound as the radius of the
   sphere increases and

   (b) The number of blocks that the sphere is partitioned into also grows without bound
as the radius of the sphere increases, and

3. There should be a way to reconcile the differences between 3- and 2-dimensional Euclidean distances between neighboring points in the sampling plan.

### 2.2 Stratified Spherical Sampling Plan

The stratified spherical sampling plan is one way of sampling the sphere so that all of the conditions given above hold within certain ranges. (The portions of the sphere where these conditions do not hold are shown to be negligible [Lemma 3.6 and its corollary].) The generation of this sampling plan can be described in words by the following algorithm:

- **Slice the sphere up into wafers**, cut parallel to the "equator", with equatorially reflective wafers having the "same height". These cuts are like the latitudinal lines on the globe. Notice that since the random field is rotation invariant, the determination of the "true" equator is arbitrary.

- On each wafer create a latitude-longitude grid. Each grid can be generated by a unique set of latitudinal and longitudinal angles specific to each wafer. The random process is then observed at the vertices of this grid. By the equatorial reflectiveness of the design, the sampling plan on the "northern" hemisphere is identical to that on the "southern" hemisphere.

- The pieces of the sphere that are not cut into wafers, the two spherical caps, should also be sampled in a uniform manner, which can be accomplished by a hexagonal sampling plan.

This methodology maintains similar interpoint distances by using a **different** latitude-longitude grid on each wafer. As the radius of the sphere grows the latitudinal and longitudinal angles are chosen so as to make the sampling plan look (locally) like the planar integer lattice. See Figure 2.1 for a schematic diagram of this sampling plan.

Each stratified spherical sampling plan is made up of 5 parts: the northern cap $C_N(r)$, the southern cap $C_S(r)$, the northern hemisphere $H_N(r)$, the southern hemisphere $H_S(r)$, and the equatorial region $E(r)$. The northern and southern caps and the equatorial region are used as little blocks and separate the two hemispheres that drive the distribution theory. They can be explicitly calculated by using functions $\gamma_1(r), \phi(r), \theta_w(r)$, and integer sequences $J_r$ and $v_r$, where $\theta_w(r)$ and $\phi(r)$ are the horizontal and vertical generating angles of the latitude-longitude grid on wafer $w; J_r$, and $v_r$ are the number of $\phi(r)$ vertical angular
increments in each wafer and equatorial region, respectively, and \( \gamma_1(r) \) is used to calculate the top of the first wafer. From these quantities, we can calculate \( W_r \), the number of wafers that the sphere is partitioned into, \( n_{w,r} \), the number of \( \theta_w(r) \) angles that go around wafer \( w \), and \( \gamma_w(r) \), the vertical angle to the top of wafer \( w \). Since each stratified spherical sampling plan is symmetric about the equator, we only need to define quantities for \( 1 \leq w \leq W_r/2 \) and the equatorial region, denoted with an \( \text{"E"} \), so for convenience we let \( W_r = \{ w : 1 \leq w \leq W_r/2 \text{ or } E \} \) and \( W_r^+ = \{ w : 1 \leq w \leq W_r \text{ or } E \} \).

Denote a point \( P \) on a sphere of radius \( r \) by its spherical coordinates \( P = (r, \theta, \phi) \) where \( \theta \) is the angle between the positive z-axis and the ray from the origin to \( P^* \), the projection of \( P \) onto the \( xy \)-plane, and \( \phi \) is the angle between the positive z-axis and the ray from the origin to \( P \).

**Algorithm 1** Given functions \( \gamma_1^*(r), \phi(r), \theta_w(r), \) and integer sequences \( J_r \) and \( v_r \), calculate \( W_r, n_{w,r}, \) and \( \gamma_w(r) \), mathematically, by first calculating

\[
U_r(\gamma_1^*(r), \phi(r), J_r, v_r) = \frac{1}{\phi(r)J_r} \cdot \{ \pi - 2\gamma_1^*(r) \} - \frac{v_r}{J_r} \tag{2.1}
\]

and then put

\[
W_r = U_r(\gamma_1^*(r), \phi(r), J_r, v_r) - 2z_r^* \tag{2.2}
\]

where \( z_r^* \in [0, 1) \) is chosen such that \( W_r \) is an even integer. Then define the vertical angle to the top of the first wafer as

\[
\gamma_1(r) = \gamma_1^*(r) + z_r^*J_r\phi(r). \tag{2.3}
\]
For \(1 \leq w \leq W_c/2\), define the vertical angle to the top of wafer \(w\), \(\gamma_w(r)\), and the number of \(\theta_w(r)\) angles that go around wafer \(w\) as

\[
\gamma_w(r) = \gamma_1(r) + (w - 1)J_r \phi(r), \quad (2.4)
\]
\[
\gamma_{w_r+1-w}(r) = \pi - \gamma_w(r), \quad \text{and}
\]
\[
n_{w, r} = n_{w_r+1-w, r} = \left[ \frac{2\pi}{\theta_w(r)} \right]. \quad (2.6)
\]

and for the equatorial region, let

\[
\gamma_E(r) = \gamma_1(r) + \frac{1}{2} W_r J_r \phi(r) \quad \text{and}
\]
\[
n_{E, r} = \left[ \frac{2\pi}{\theta_E(r)} \right]. \quad (2.8)
\]

With these quantities, we can explicitly give the sampling design for a stratified spherical sampling plan \(\mathcal{R}\). For the hemispherical regions, \(H_N(r)\) and \(H_S(r)\), we sample at the vertices of the wafer-specific, latitude – longitude grid described in section 2.2 and create a stratified spherical sampling plan. Since the sampling plan in these two regions is approximately uniform, we would also like to have this quality in \(C_N(r)\), \(C_S(r)\), and \(E(r)\). For \(E(r)\) this is quite simple, just extend the definition of a wafer and its associated grid to this region. That is, let \(E(r)\) be described by a latitude – longitude grid where there are \(v_r\) latitude angles at vertical angle increments \(\phi(r)\), and \(n_{E, r}\) longitude angles at the horizontal angle increments \(\theta_E(r)\). Since the shape of each cap is topologically different that that of the wafers, a different type of sampling plan needs to be employed. In this instance, a natural choice is the hexagonal sampling plan (Matérn [Mat86]), which provides circular symmetry within the cap. The sampling plan is given by the following algorithm:

**Algorithm 2** Define

\[
H_N(r) = \bigcup_{w=1}^{W_r/2 J_r-1} \bigcup_{i=0}^{n_{w, r}-1} P_{i, j}^W(r) \quad \text{and} \quad (2.9)
\]
\[
H_S(r) = \bigcup_{w=1}^{W_r/2 J_r-1} \bigcup_{i=0}^{n_{w, r}-1} P_{i, j}^{W_r+1-w}(r) \quad (2.10)
\]

where in this range for \(w\),

\[
P_{i, j}^W(r) = (r, i \theta_w(r), \gamma_w(r) + j \phi(r)) \quad \text{and} \quad (2.11)
\]
\[
P_{i, j}^{W_r+1-w}(r) = (r, i \theta_w(r), \pi - (\gamma_w(r) + j \phi(r))). \quad (2.12)
\]

Define

\[
E(r) = \bigcup_{i=0}^{v_r-1} \bigcup_{j=0}^{n_{E, r}-1} P_{i, j}^E(r) \quad (2.13)
\]

27
where
\[ P_{i,j}^E(r) = (r, i\theta_E(r), \gamma_E(r) + j\phi(r)). \] (2.14)

Define
\[ C_N(r) = (r, 0, 0) \cup \left( \bigcup_{j=1}^{\lceil \gamma_N(r)/\phi(r) \rceil - 1} \bigcup_{i=1}^{6j-1} P_{i,j}^N(r) \right) \text{ and} \] (2.15)
\[ C_S(r) = (r, 0, \pi) \cup \left( \bigcup_{j=1}^{\lceil \gamma_N(r)/\phi(r) \rceil - 1} \bigcup_{i=1}^{6j-1} P_{i,j}^S(r) \right). \] (2.16)

where
\[ P_{i,j}^N(r) = \left( r, \frac{i\pi}{3j}, j\phi(r) \right) \text{ and} \] (2.17)
\[ P_{i,j}^S(r) = \left( r, \frac{i\pi}{3j}, \pi - j\phi(r) \right). \] (2.18)

A stratified spherical sampling plan \( R \) can now be given by
\[ R = C_N(r) \cup H_N(r) \cup E(r) \cup H_S(r) \cup C_S(r). \] (2.19)

Notice that on each wafer \( w \), the latitude-longitude grid is constructed by the latitudinal cuts of the sphere at \( \phi \) angles
\[ \gamma_w(r), \gamma_w(r) + \phi(r), \ldots, \gamma_w(r) + (J_r - 1)\phi(r) \]
and the longitudinal cuts at \( \theta \) angles
\[ 0, \theta_w(r), 2\theta_w(r), \ldots, (n_w,r - 1)\theta_w(r). \]

The same type of grid is also constructed in \( E(r) \), except that there are only \( n_r \) longitudinal cuts, half of them on each side of the equator. Notice that from this definition the whole sphere is covered since
\[ \pi = 2\gamma_1(r) + W_r J_r \phi(r) + n_r \phi(r). \]

### 2.2.1 Approximation of the Latitude-Longitude Grid to \( \mathbb{Z}^2 \)

In order to reconcile the difference between 2- and 3-dimensional Euclidean distances, we need to calculate the 3-dimensional distances. To ensure the negligibility of the spherical caps and equatorial region, constraints must be placed upon \( \phi(r), \theta_w(r), \gamma_w(r), J_r, \) and \( n_r \). Lemma 2.1 gives a pure analytic result that describes what is meant by “the sampling plan looks (locally) like the integer lattice”. It provides insight into how big the spherical blocks can actually be, yet still be approximated by planar blocks.
We require that 3-dimensional Euclidean distances between any two "local" points on a wafer should converge uniformly to their planar counterpart distance in $\mathbb{Z}^2$. That is, the Euclidean distance between the $(i, j)$th and $(i', j')$th points on the latitudinal – longitudinal grid starting at some arbitrary angle (let this angle be 0 for convenience) in a single wafer is uniformly approximated by its planar counterpart distance between points $(i, j)$ and $(i', j')$ in $\mathbb{Z}^2$. We will define local to mean that $|i - i'| \leq I_r$ and $|j - j'| \leq J_r$ where $I_r$ and $J_r$ grow with $r$.

Since the Euclidean distance between two points $P_{i,j} = (r, \phi_i, \phi_j)$ and $P_{i',j'} = (r, \phi_{i'}, \phi_{j'})$ is

$$
\|P_{i,j} - P_{i',j'}\| = 2r \left\{ \sin^2 \left( \frac{\phi_i - \phi_{i'}}{2} \right) + \sin(\phi_i) \sin(\phi_{i'}) \sin^2 \left( \frac{\theta_i - \theta_{i'}}{2} \right) \right\}^{1/2}
$$

(2.20)

then, for $w \in \mathcal{W}_r$, specializing (2.20) for points and $P_{i,j}^w(r)$ and $P_{i',j'}^w(r)$ defined by Algorithm 2, we have

$$
\|P_{i,j}^w(r) - P_{i',j'}^w(r)\| = 2r \left\{ \sin^2 \left( \frac{(j - j')\phi(r)}{2} \right) + \sin(\gamma_w(r) + j\phi(r)) \right. \cdot \sin(\gamma_w(r) + j'\phi(r)) \sin^2 \left( \frac{(i - i')\theta_w(r)}{2} \right) \right\}^{1/2}.
$$

(2.21)

Therefore, the neighboring interpoint horizontal wafer distance (i.e., when $i' < i$ and $j' = j$) is

$$
\|P_{i,j}^w(r) - P_{i',j}^w(r)\| = 2r \sin(\gamma_w(r) + j\phi(r)) \sin \left( \frac{(i - i')\theta_w(r)}{2} \right)
$$

(2.22)

and the interpoint vertical distance (i.e., when $i' = i$ and $j' < j$)

$$
\|P_{i,j}^w(r) - P_{i',j'}^w(r)\| = 2r \sin \left( \frac{(j - j')\phi(r)}{2} \right).
$$

(2.23)

We require that asymptotically these neighboring interpoint wafer distances converge to $(i - i')$ and $(j - j')$, respectively; as they are exactly in $\mathbb{Z}^2$. It follows from (2.22) that the top of the first wafer $\gamma_1(r)$ cannot be equal to 0, since if it is, then all of the horizontal interpoint distances on the top of this wafer would be zero (i.e., $j = j' = \gamma_1(r) = 0$). Nevertheless, it is possible to have $\gamma_1(r)$ converge to 0 as $r$ tends to $\infty$. Also note that the vertical distance is independent of the points' horizontal placement on the sphere (2.23). This is taken into account in our sampling plan by letting the vertical angular increments, $\phi(r)$, on the wafer grid be a function of $r$ only and not of $w$. It is further evident that the horizontal distance is dependent upon the points' height on the sphere (2.22) and hence this is incorporated into our plan by defining this horizontal angular increment of the wafer grid to be $\theta_w(r)$. Notice that $\theta_w(r)$ is inversely related to the height of the top of wafer $w$. That is,

$$
\theta_w(r) > \theta_{w+1}(r) \quad \text{and} \quad \gamma_w(r) < \gamma_{w+1}(r).
$$

This illustrates why the same latitude-longitude grid cannot be used over the entire sphere. Further note, that since we require that interpoint distance converge to $(i - i')$ and $(j - j')$,
both \( \phi(r) \) and \( \theta_w(r) \) are inversely related to \( r \); that is, \( \phi(r) \to 0 \) and for each \( w \in \mathcal{W}_r, \theta_w(r) \to 0 \) as \( r \to \infty \).

The following lemma gives us conditions on \( I_r, J_r, v_r, \gamma_w(r), \gamma_1^*(r), \phi(r), \) and \( \theta_w(r) \) such that the local 3-dimensional Euclidean distances on each wafer can be approximated by their planar counterpart distances in \( \mathbb{R}^2 \). Since the sampling plan is symmetric about the equator we only consider \( w \in \mathcal{W}_r \). We write \( l_r \approx m_r \) if as \( r \to \infty, l_r \to \infty, m_r \to \infty \), and \( \frac{l_r}{m_r} \to c \in (0, \infty) \).

**Lemma 2.1** Let functions \( \phi(r), \theta_w(r), \) and \( \gamma_1^*(r) \) and integer sequences \( \{I_r\}, \{r\}, \) and \( \{v_r\} \) be given. Define \( U_r(\gamma_1^*(r), \phi(r), J_r, v_r), \mathcal{W}_r, \gamma_1(r) \), and for \( w \in \mathcal{W}_r \), define \( \gamma_w(r), n_{w,r}, \) and \( P_{i,j}^w(r) \), according to Algorithms 1 and 2. If

1. \( I_r \to \infty, J_r \to \infty, I_r \approx J_r, v_r \to \infty, u_r \to 0, \)
2. \( \gamma_1^*(r) \to 0, \)
3. \( |\pi \phi(r) - 1| = O \left( \frac{J_r}{r} \right), \)
4. \( \sup_{w \in \mathcal{W}_r} [\theta_w(r) r \sin(\gamma_w(r)) - 1] = O \left( \frac{J_r}{r \sin(\gamma_1(r))} \right), \) and
5. \( \frac{J_r^3}{r \sin(\gamma_1(r))} \to 0, \)

then

\[
\sup \left\| P_{i,j}^w(r) - P_{i',j'}^w(r) \right\|^2 - [(i - i')^2 + (j - j')^2] = O \left( \frac{J_r^3}{r \sin(\gamma_1(r))} \right) = o(1) \tag{2.24}
\]

where the supremum is taken over all \( w \in \mathcal{W}_r \), and 0 \( \leq i', i \leq n_{w,r} - 1, 0 \leq j, j \leq J_r - 1 \), and 0 \( \leq |i - i'| \leq J_r \) and 0 \( \leq |j - j'| \leq v_r \) when \( w = E \) and 0 \( \leq |i - i'| \leq I_r \) and 0 \( \leq |j - j'| \leq J_r \) otherwise.

Condition (1) ensures that the maximal height and length of each tile is of the same order while growing larger and larger. Condition (2) forces the top of the first wafer to get closer and closer to the pole. Conditions (3) and (4) classify the robustness of the latitude - longitude grid. Condition (5) is the rate at which the error of the approximation decreases to 0; it is a function of \( r \) only. Notice that the angles \( \phi(r) = r^{-1} \) and \( \theta_w(r) = (r \sin(\gamma_w(r)))^{-1} \) give exactly zero conditions in (3) and (4). All of the conditions and (2.3) guarantee that \( \gamma_1(r) \) and \( \gamma_1^*(r) \) are of the same order.

The following is an example of a stratified spherical sampling plan. For some \( 0 < \eta, \epsilon, \zeta < 1 \), where \( 0 < 5\pi + \epsilon < \eta \) define \( \phi(r), \theta_w(r), \gamma_1^*(r), J_r, \) and \( v_r \) by Table 2.1. \( \mathcal{W}_r, \gamma_1^*(r), \gamma_w(r), \) and \( n_{w,r} \) are then calculated from Algorithm 1 and let \( I_r = J_r \). This choice of \( \theta_w(r) \) gives a
<table>
<thead>
<tr>
<th>\phi(r)</th>
<th>\theta_w(r)</th>
<th>\gamma^s_w(r)</th>
<th>\gamma^c_w(r)</th>
<th>\nu_w</th>
</tr>
</thead>
<tbody>
<tr>
<td>\frac{1}{r}</td>
<td>\frac{2\pi}{[2\pi r \sin(\gamma_w(r))]}</td>
<td>\frac{1}{r^c}</td>
<td>[r^s]</td>
<td>[r^c]^c</td>
</tr>
</tbody>
</table>

Table 2.1: Example of a Stratified Spherical Sampling Plan

A stratified spherical sampling plan where exactly $n_{w,r}$ angles of $\theta_w(r)$ fit into each wafer $w$. The number of observations taken on the entire sphere, denoted $n_r = n(r) = |\mathcal{R}|$, is found by using Algorithm 2 and is given by

$$n_r = 2|C_N(r)| + 2|H_N(r)| + |E(r)| = \mathcal{O}\left(J_r \sum_{w=1}^{W_r} n_{w,r}\right) = \mathcal{O}\left(r^2\right).$$ \hspace{1cm} (2.25)
Chapter 3

Asymptotic Normality

3.1 Introduction

In this chapter we prove the asymptotic normality of the standardized sample mean for an isotropic and homogeneous random field on the surface of a sphere under mild regularity conditions. The ideas utilize the big block – little block technique of Bernstein. In section 3.2, we describe spatial dependence. Then in section 3.3, we give additional conditions on the stratified spherical sampling plan and the random process itself that allow us to calculate the variance of a standardized sum of observations on sections of a wafer. And finally, in sections 3.4–3.4.3, we give a central limit theorem for the standardized sample mean on the sphere for a modified stratified spherical sampling plan called a nearly uniform stratified spherical sampling plan.

3.2 Spatial Dependence

For a random field in a plane it is natural to require translation invariance, whereas when considering a random field on a sphere rotational invariance is most natural. The spatial dependence structure of our spherical random field has 2 additional properties: $\alpha$-mixing (large-distance feature) and smoothness of the covariance function (short-distance feature). To quantify the dependence between the observations in different regions, it is natural to use the notion of $\alpha$-mixing. In order to explain and extend the $\alpha$-mixing idea to a multi-dimensional random field, it is necessary to describe what is meant by the distance between two sets. Define the distance between two sets in $\mathbb{R}^d$ to be the minimal Euclidean distance between any two elements, one in each of the two sets. In other words, let $\Lambda_1$ and $\Lambda_2$ be two subsets of $\mathbb{R}^d$. 
and define the distance between $\Lambda_1$ and $\Lambda_2$

$$d(\Lambda_1, \Lambda_2) = \inf \left\{ \| \vec{x}_1 - \vec{x}_2 \| : \vec{x}_1 \in \Lambda_1 \text{ and } \vec{x}_2 \in \Lambda_2 \right\}$$  \hspace{1cm} (3.1)$$

where $\| \cdot \|$ means Euclidean distance. We now extend the definition of $\alpha$-mixing first given by Rosenblatt [Ros56].

**Definition 3.1** Let $\{X_x\}$ be an isotropic and homogeneous, mean-zero random process indexed in $\mathbb{R}^d$. Let $\Lambda_1$ and $\Lambda_2$ be two subsets of $\mathbb{R}^d$ and define $\sigma(\Lambda_1)$ and $\sigma(\Lambda_2)$ to be the sigma fields generated by $\Lambda_1$ and $\Lambda_2$, respectively. The $\alpha$-mixing coefficient of $\{X_x\}$ can be defined as

$$\alpha(k) = \sup \{|\mathbb{P}(A \cap B) - \mathbb{P}(A) \cdot \mathbb{P}(B)|\}$$

where the supremum is taken over all subsets $\Lambda_1$ and $\Lambda_2$ of $\mathbb{R}^d$ such that $d(\Lambda_1, \Lambda_2) \geq k$ and all events $A$ and $B$ of $\sigma(\Lambda_1)$ and $\sigma(\Lambda_2)$, respectively. $\{X_x\}$ is said to satisfy the strong mixing condition with $\alpha$-mixing coefficient $\alpha(k)$ if $\alpha(k) \to 0$ as $k \to \infty$.

Lemma 3.1, Corollary 3.1, and Lemma 3.2 provide bounds for the covariance function in terms of the $\alpha$-mixing coefficient $\alpha(k)$ for a random process in $\mathbb{R}^3$. Lemma 3.1 and its corollary apply to bounded random variables that are real-valued and complex-valued, respectively, while Lemma 3.2 applies to unbounded, real-valued random variables. All of them are extensions of theorems, lemmas, and remarks of Ibragimov and Linnik [IL71].

**Lemma 3.1** Let $\{X_x\}$ be an isotropic and homogeneous, mean-zero random process indexed in $\mathbb{R}^3$ satisfying the strong mixing condition with $\alpha$-mixing coefficient $\alpha(k)$. If $\zeta$ and $\eta$ are real-valued, measurable random variables with respect to $\sigma(\Lambda_1)$ and $\sigma(\Lambda_2)$, respectively, such that $d(\Lambda_1, \Lambda_2) \geq k$, and if

$$|\zeta| \leq c_1, \quad \text{and} \quad |\eta| \leq c_2,$$

then

$$|\mathbb{E}\{\zeta \eta\} - \mathbb{E}\{\zeta\} \mathbb{E}\{\eta\}| \leq 4c_1 c_2 \alpha(k).$$

**Corollary 3.1** Let $\{X_x\}$ be an isotropic and homogeneous, mean-zero random process indexed in $\mathbb{R}^3$ satisfying the strong mixing condition with $\alpha$-mixing coefficient $\alpha(k)$. If $\zeta$ and $\eta$ are complex-valued, measurable random variables with respect to $\sigma(\Lambda_1)$ and $\sigma(\Lambda_2)$, respectively, such that $d(\Lambda_1, \Lambda_2) \geq k$, and if

$$|\zeta| \leq c_1, \quad \text{and} \quad |\eta| \leq c_2,$$

then

$$|\mathbb{E}\{\zeta \eta\} - \mathbb{E}\{\zeta\} \mathbb{E}\{\eta\}| \leq 16c_1 c_2 \alpha(k).$$

However, we would like to deal with random variables that are not bounded (e.g., normal random variables) and so Lemma 3.1 must be extended to this class.
Lemma 3.2 Let \( X_F \) be an isotropic and homogeneous, mean-zero random process indexed in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \). If \( \zeta \) and \( \eta \) are real-valued, measurable random variables with respect to \( \sigma(\Lambda_1) \) and \( \sigma(\Lambda_2) \), respectively, such that \( \text{d}(\Lambda_1, \Lambda_2) \geq k \), and if for some \( \epsilon > 0 \)

\[
\mathbb{E} \left\{ |\zeta|^{2+\epsilon} \right\} < c_1 < \infty \quad \text{and} \quad \mathbb{E} \left\{ |\eta|^{2+\epsilon} \right\} < c_2 < \infty
\]

then for \( \beta = (2 + \epsilon)^{-1} \),

\[
|\mathbb{E} \{\zeta \eta\} - \mathbb{E} \{\zeta\} \mathbb{E} \{\eta\}| \leq \left\{ 4 + 3 \left( c_1^{\beta} c_2^{1-\beta} + c_1^{1-\beta} c_2^\beta \right) \right\} \alpha(k)^{1-2\beta}.
\]

3.3 Tile Variance

Using Lemma 2.1, we can find further conditions on \( \phi(r) \), \( \theta_w(r) \), \( \gamma_w(r) \), \( J_r \), \( I_r \), and \( v_r \) so that the variance of a sum on a subset (or subtile) of a wafer is of the order of the number of elements in the subtile. Lemma 3.3 shows this result for bounded random variables and Corollary 3.2 gives this result for unbounded random variables with a finite \((2 + \delta)^{th}\) moment for some \( \delta > 0 \). Let us first describe what is meant by a subtile of a wafer.

Let a wafer, \( w \), be a group of points on a sphere with a top latitudinal angle of \( \gamma_w(r) \) and a bottom latitudinal angle of \( \gamma_w(r) + (J_r - 1)\phi(r) \). Let \( wS_{s,t}^{l,m} \) denote a subtile of wafer \( w \), that begins at the horizontal angle \( \theta_w(r) \) and the vertical angle \( \gamma_w(r) + t\phi(r) \) and continues in horizontal increments of angle \( \theta_w(r) \) for a length of \( l \) units and in vertical increments of angle \( \phi(r) \) for a height of \( m \) units. Specifically, for \( 1 \leq w \leq W_r/2 \), define for \( 0 \leq s \leq s + l \leq n_{w,r} \) and \( 0 \leq t \leq t + m \leq J_r \):

\[
\begin{align*}
\text{wS}_{s,t}^{l,m} &= \{ P_{i,j}^w(r) : s \leq i \leq s + l - 1 \quad \text{and} \quad t \leq j \leq t + m - 1 \} \quad (3.2) \\
\text{wS}_{s,t}^{l,m} &= \{ P_{i,j}^{w+1-w}(r) : s \leq i \leq s + l - 1 \quad \text{and} \quad t \leq j \leq t + m - 1 \} \quad (3.3)
\end{align*}
\]

where \( P_{i,j}^w(r) \) and \( P_{i,j}^{w+1-w}(r) \) are defined by (2.11) and (2.12), respectively, and for the equatorial region, define for \( 0 \leq s \leq s + l \leq n_{E,r} \) and \( 0 \leq t \leq t + m \leq v_r \):

\[
\text{E}_{s,t}^{l,m} = \{ P_{i,j}^E(r) : s \leq i \leq s + l - 1 \quad \text{and} \quad t \leq j \leq t + m - 1 \} \quad (3.4)
\]

where \( P_{i,j}^E(r) \) is given by (2.14). Notice that a subtile cannot "spill over" into a neighboring wafer; that is, it is completely contained in wafer \( w \), denoted \( wS_{0,0}^{w,r} \).

A tile of wafer \( w \) beginning at horizontal angle \( \theta_w(r) \), \( wS_{i,j}^{l,r,j} \), is simply a subtile of wafer \( w \) of length \( I_r \) and height \( J_r \) and thus all points within a tile are considered as being local. In order for the number of observations on each tile to grow without bound, yet be small enough so
that the sphere may be broken up into an ever increasing number of tiles, we require, as in condition (1) of Lemma 2.1, that

\[ I_r \to \infty, \frac{I_r}{r} \to 0, J_r \to \infty, \text{ and } \frac{J_r}{r} \to 0. \]

As mentioned previously, for an isotropic random field the covariance between any two points is a function of their Euclidean distance only. Since the Euclidean distance between two points on the sphere is quite cumbersome (see 2.20), it is necessary to approximate the covariance at a 3-dimensional Euclidean distance by the covariance at the Euclidean planar distance of analogous points in \( Z^2 \). This approximation requires the covariance function be "smooth" in some sense. To classify what is meant by smooth, use the following preliminary definitions: for any \( X, N, \epsilon, s \in \mathbb{R} \), let \( X^N = X \cdot 1_{\{|X| \geq N\}}, X^N = X - N, Z = \{0, 1, 2, \ldots\} \), and \( B(s, \epsilon) = \{ t \in \mathbb{R} : |t - s| \leq \epsilon \} \), and for any two subsets \( A \) and \( B \) of \( \mathbb{R} \), let \( |B| \) denote the cardinality of \( B \) and \( A \setminus B = \{ t \in \mathbb{R} : t \in A \text{ and } t \notin B \} \). The meaning of smoothness of the covariance function of a homogeneous and isotropic random process can now be defined.

**Definition 3.2** Let \( \{X_{\vec{q}}\} \) be an isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \). Given a sequence \( \{N_k : k \geq 0\} \) such that \( N_k \uparrow \infty \) and \( N_0 \equiv 0 \), define the truncated covariance function

\[ B_k(||\vec{q}||) = C \{ N_k \vec{X}_0, N_k \vec{X}_{\vec{q}} \} . \]

Let \( V = \{ s \in \mathbb{R} : s = ||\vec{q}|| \text{ for some } \vec{q} \in Z^2 \} \). The covariance function of \( \{X_{\vec{q}}\} \) is smooth if \( \exists \{N_k : k \geq 0\} \) such that for each \( k \geq 0, \exists c_k, c_k > 0, \) where \( \forall s \in V \) and \( \forall t \in B(s, \epsilon_k) \)

\[ |B_k(s) - B_k(t)| \leq c_k |s - t|. \]

The smoothness condition requires that for each \( k \) the truncated covariance function is locally continuous for points in \( V \).

Notice that this definition applies to the upper truncated random variables. Hence when \( N_0 = 0 \), then this corresponds to the covariance function for the untruncated random variables. In particular, if the random variables are bounded by some finite number \( N \), then there exists some \( k_0 \) such that for every \( k > k_0, N_k > N \), and hence \( B_k(s) \equiv 0 \). If for some \( \xi > 0 \) we set \( N_0 = 0, N_1 = N + \xi \) and then \( N_k = N_{k-1} + 1 \) for all \( k > 1 \), the condition is satisfied if the covariance function for the entire random variable is uniformly Lipschitz.

**Example 1** M-dependent, bounded, continuous random field

As for time-series, one can construct a set of random variables that are \( m \)-dependent from independent ones. Begin with a continuously indexed random field \( \{X_{\vec{q}}\} \) on the sphere such that the random variables are independent and bounded. (See Dorea [Dor72] for existence of this random field.) Given \( m \in \mathbb{R} \setminus V \), and a continuous, bounded function,
\( f(\cdot) : \mathbb{R}^1 \to \mathbb{R}^1 \), let \( C_{m \phi(r)} \) be the spherical cap on the sphere at point \( \hat{p} \) that has radial angle \( m \phi(r) \). Define a new random variable \( Y_{\hat{p}} \) by

\[
Y_{\hat{p}} = \frac{1}{w_{r,m}} \int_{C_{m \phi(r)}} f(X_{\hat{p}})d\hat{q}
\]

where \( w_{r,m} \) is the surface area element of \( C_{m \phi(r)} \), which is constant for all \( \hat{p} \). This creates a bounded, \( m \)-dependent, homogeneous and isotropic random field on the sphere. Using the \( m \)-dependence of the random process and the fact (Yadrenko [Yad83, Lemma 1, page 3]) that for any homogeneous and isotropic random field in \( \mathbb{R}^3 \), the correlation function of the random process is differentiable, then the smoothness condition is satisfied.

**Example 2 General bounded, homogeneous and isotropic random field**

An isotropic random field admits a decomposition into spherical harmonics. For a random field \( \{X_{\hat{p}}\} \) on the unit sphere, each element \( X_{\hat{p}} \) for \( \hat{p} = (\phi, \theta) \) is given by

\[
X_{\hat{p}} = \sum_{m=0}^{\infty} \sum_{l=-m}^{m} b_m \delta_m^l \sin^l \phi \frac{d^{m+l}}{d \cos(\theta)^{m+l}} (\cos^2(\theta) - 1)^m
\]

where \( \delta_m^l \) are uncorrelated random variables and \( C \left\{ \delta_m^l, \delta_m^l \right\} = b_m \delta_m^m, \delta_m^l, \delta_r^m \) is the Dirac delta function, the \( b_m \)'s are non-negative constants such that

\[
\sum_{m=0}^{\infty} (2m + 1) b_m < \infty
\]

\( S_m^l(\theta, \phi) = e^{il\phi} P_m^l(\cos(\theta)) \) are the usual 3-dimensional spherical harmonics, and

\( P_m^l(\cos(\theta)) = \frac{(m-l)!}{2m!(m+l)!} \sin^l(\theta) \frac{d^{m+l}}{d \cos(\theta)^{m+l}} (\cos^2(\theta) - 1)^m \)

are the associated Legendre polynomials. (See Yaglom [Yag87], Jeffreys [JS66], and Yadrenko [Yad83] for more details). From this we can see that, if we insist that for some finite positive \( M \), \( b_m \equiv 0 \) for all \( m > M \), then the random field can be broken down into a finite number of uncorrelated random variables. If we assume that these random variables are bounded, then again this random field is bounded by some finite number \( N \). Applying the result of Yadrenko [Yad83] implies that this covariance function is continuous.

With the definitions of a subtile and smoothness of the covariance function, we can find further conditions on \( \phi(r), \theta_u(r), \gamma^r(r), J_r, I_r \), and \( \nu_r \) so that the variance of a sum on a subtile is of the order of the number of elements in the subtile.

**Lemma 3.3** Let \( \{X_{\hat{p}}\} \) be a bounded, isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth
covariance function. Given $I_r, J_r, v_r, \gamma_1^r(r), \phi(r)$, and $\theta_\omega(r)$ define $W_r, \gamma_1(r), \gamma_\omega(r)$, and $n_{w, r}$ according to Algorithms 1 and 2 for $w \in W_r$ under the conditions of Lemma 2.1. For integer sequences $\{s_r\}, \{t_r\}, \{l_r\}, \{m_r\}$ such that for each $1 \leq w \leq W_r/2$,

1. $0 \leq s_r \leq a_r \leq n_{w, r}$ and $0 \leq t_r \leq b_r \leq l_r$ and

2. $l_r \to \infty, m_r \to \infty, \frac{l_r}{J_r} \to c_1 \in [0, 1], \frac{m_r}{v_r} \to c_2 \in [0, 1],$

and for $w = E_r$, (1) and (2) are replaced by

1'. $0 \leq s_r \leq a_r \leq n_{E, r}$ and $0 \leq t_r \leq b_r \leq l_r$ and

2'. $l_r \to \infty, m_r \to \infty, \frac{l_r}{J_r} \to c_1 \in [0, 1], \frac{m_r}{v_r} \to c_3 \in [0, 1],$

define for $w \in W_r$, $w S_{s_r, t_r}^{l_r, m_r}$ by (3.2) and (3.4). If

3. $\sum_{j=1}^{\infty} j^\alpha(j) < \infty$ and

4. $\frac{J_r^5}{r \sin(\gamma_1(r))} \rightarrow 0$

then

$$\sup_{w \in W_r} \left| \mathbb{V} \left\{ \sum_{\substack{S_{s_r, t_r}^{l_r, m_r} \in S_{s_r, t_r}^{l_r, m_r}}} X_{\mathbb{S}} \right\} / \left| w S_{s_r, t_r}^{l_r, m_r} \right| - \sigma^2 \right| \rightarrow 0$$

where

$$\sigma^2 = \sum_{\mathbb{S} \in \mathbb{Z}^2} C \left\{ X_{\mathbb{S}}, X_{\mathbb{S}} \right\}.$$

Notice that the asymptotic variance for the standardized subtile is the same as that for the planar integer lattice. This is due to the stratified spherical sampling plan subject to the conditions of Lemma 2.1. This result can be extended to the case when the random field is unbounded.

**Corollary 3.2** Let $\{X_{\mathbb{S}}\}$ be an isotropic and homogeneous, mean-zero random process in $\mathbb{R}^3$ satisfying the strong mixing condition with $\alpha$-mixing coefficient $\alpha(k)$ and having a smooth covariance function. Given $I_r, J_r, v_r, \gamma_1^r(r), \phi(r)$, and $\theta_\omega(r)$ define $W_r, \gamma_1(r), \gamma_\omega(r)$, and $n_{w, r}$ according to Algorithms 1 and 2 for $w \in W_r$ under the conditions of Lemma 2.1. If, for integer sequences $\{s_r\}, \{t_r\}, \{l_r\}, \{m_r\}$, the conditions of Lemma 3.3 are satisfied with condition (3) replaced by

$$\frac{J_r^5}{r \sin(\gamma_1(r))} \rightarrow 0$$

then

$$\sup_{w \in W_r} \left| \mathbb{V} \left\{ \sum_{\substack{S_{s_r, t_r}^{l_r, m_r} \in S_{s_r, t_r}^{l_r, m_r}}} X_{\mathbb{S}} \right\} / \left| w S_{s_r, t_r}^{l_r, m_r} \right| - \sigma^2 \right| \rightarrow 0$$

where

$$\sigma^2 = \sum_{\mathbb{S} \in \mathbb{Z}^2} C \left\{ X_{\mathbb{S}}, X_{\mathbb{S}} \right\}.$$
3a. \( \exists \delta > 0 \) such that \( \mathbb{E}\left\{ |X_p|^2 + \delta \right\} < \infty \), and

3b. \( \sum_{j=1}^{\infty} j \alpha(j) j^{4/(2+q)} < \infty \)

then, for \( \sigma^2 \) given by (3.5),

\[
\sup_{v \in \mathcal{W}_r} \left| V \left\{ \sum_{\phi \in S_{s, m_r}^r} X_{\phi} \right\} \right| \frac{w^{s, m_r}}{|w^{s, m_r}|^2} \to 0.
\]

Lemma 2.1, condition (4) of Lemma 3.3, and Corollary 3.2 can be used to completely characterize the class of “nearly uniform” stratified spherical sampling plans.

**Definition 3.3** A stratified spherical sampling plan \( \mathcal{R} = \mathcal{R}(\gamma_1(\rho), \phi(\rho), \theta_\omega(\rho), I_\rho, J_r, v_r) \) is said to be nearly uniform if the conditions of Lemma 2.1 are satisfied with condition (5) replaced by

\[
5') \quad \frac{J_5^s}{r \sin(\gamma_1(\rho))} \to 0
\]

and for every \( r \) and for each \( w \in \mathcal{W}_r \), \( 2\pi/\theta_\omega(\rho) \) is an integer.

Notice that the specific stratified spherical sampling plan given in Table 2.1 is actually a nearly uniform stratified spherical sampling plan, since for large enough \( r \), \( \exists \) a constant \( C \) such that

\[
\frac{J_5^s}{r \sin(\gamma_1(\rho))} = \left( \frac{1}{r^0} \right)^5 \left( \frac{\gamma_1(\rho)}{\sin(\gamma_1(\rho))} \right) \left( \frac{1}{r^4 \gamma_1(\rho)} \right) \left( \frac{r^{5\eta}}{r^{1 - \epsilon}} \right) \leq C r^{-(1 - (5\eta + \epsilon))} \to 0.
\]

### 3.4 Asymptotic Normality

In this section, the ideas used to prove the central limit theorem are presented. In section 3.4.1, the method of proving asymptotic normality of the standardized sample mean for the sphere is outlined. This method is similar to that of the time-series case, with one exception: we use an approximation of the variance of the spherical block sums by the variance of block sums of the same cardinality on the planar integer lattice (Lemma 3.3). The remainder of this section gives the Lemmas (their proofs are given in Appendix A) which lead to the final result.
3.4.1 Intuition Underlying the Proof

Theorem 1.1 by Ibragimov and Linnik [IL71], pertaining to the asymptotic normality of the standardized sample mean for a strongly mixing, strictly stationary sequence observed at integer intervals with \( \alpha \)-mixing coefficient \( \alpha(k) \), is used as motivation for our proof and formally brings together the concepts of \( \alpha \)-mixing and the big-block, little-block technique.

We adapt this same argument to the sphere; that is, break the sphere up into big tiles and little tiles and apply an \( \alpha \)-mixing condition to get approximately independent tile sums. In section 3.3, conditions were found that ensure that the variance of a tile sum was of the order of the number of observations on the tile. In section 3.4.2, the whole tiles are formally arranged on the sphere, and each tile is partitioned into four subtiles: three of them acting as little tiles and one acting as the big tile. The correct \( \alpha \)-mixing condition is given (Lemma 3.5) to ensure that the spherical tile sums can be approximated by independent spherical tile sums of the same cardinality and the negligibility of the little tile sums is then shown (Lemma 3.6 and Corollary 3.3). Finally, in section 3.4.3, a modified Lindeberg condition for the asymptotically independent big spherical tile sums is shown to be satisfied, and the asymptotic normality results (Lemma 3.7 and Theorem 3.1) follow.

3.4.2 Big-Tile, Little Tile

In this section, the sphere is formally partitioned into 8 different regions: the equatorial region, the spherical caps, and 5 regions that pertain to each wafer. If we define \( k_{w,t} = \lfloor n_{w,t}/b \rfloor \), then for each wafer \( w \) given by a nearly uniform stratified spherical sampling plan, \( \mathcal{P} \), there are exactly \( k_{w,t} \) whole tiles \( T_{w,k}(r) = wS_{(k-1)t}^{r / J_r} \) and \( n_{w,t} - k_{w,t} \) horizontal angular increments of \( \theta_w(r) \) remaining, with each increment having \( J_r \) vertical angular increments of \( \phi(r) \) associated with it. These remaining observation sites are put together and called a remaining tile, \( R_w(r) = wS_{k_{w,t}+1}^{r / J_r} \). Notice that from the definition of a nearly uniform stratified spherical sampling plan, for every \( w, k_{w,t}, \rightarrow \infty \) by condition (5').

Each of the \( k_{w,t} \) whole tiles is partitioned into 4 different parts (or types) : the vertical, corner, and horizontal corridor tiles and the main (big) tile, denoted as type 1, 2, 3, and 4 tiles, respectively. (See Figure 3.1 for a schematic diagram of the sphere and the tile types.) For \( \phi(r), \theta_w(r), \gamma^r_w(r), I_r, J_r, \) and \( v_r \) defined by any nearly uniform stratified spherical sampling plan \( \mathcal{P} \), calculate for \( w \in W_r, \gamma_w(r), W_r, \) and \( n_{w,r} \) according to Algorithm 1 and define

\[
T_{w,k}(r) = wS_{(k-1)t}^{r / J_r}, \quad \text{and} \quad T_{w,k+1}(r) = w^{r+1}S_{(k-1)t}^{r / J_r},
\]

as the \( k^{th} \) tiles on wafers \( w \) and \( W_r + 1 - w \), respectively, where \( 1 \leq k \leq k_{w,t} \). Define the type 1 tiles, \( T_{w,k}(r) \) and \( T_{w,k+1}(r) \), the remaining tiles, \( R_w(r) \) and \( R_E(r) \), and the
Figure 3.1: Type I tiles

pseudo-equatorial region, $E^*(r)$, for $1 \leq w \leq W_r/2$, as

$$
1_{w,k}(r) = \frac{w S_{w,r}^{v_r,F_r}}{S_{(k-1)f_r}^{v_r,F_r}} \\
2_{w,k}(r) = \frac{w S_{w,F_r}}{S_{(k-1)f_r}^{v_r,F_r}} \\
3_{w,k}(r) = \frac{w S_{w-r,F_r}}{S_{(k-1)f_r}^{v_r,F_r}} \\
4_{w,k}(r) = \frac{w S_{w-r,k,w,F_r}}{S_{(k-1)f_r}^{v_r,F_r}} \\
R_w = \frac{w S_{k,w,F_r}}{S_{(k-1)f_r}^{v_r,F_r}} \\
E^*(r) = \frac{E_{S_{k,F_r}}^{v_r,F_r}}{S_{(k-1)f_r}^{v_r,F_r}}
$$

Notice that Lemma 3.3 and Corollary 3.2 apply to these 4 tile types. That is, for each of the tile types we can relate like terms in the lemma and corollary by Table 3.1.

The set of points that contain the type I($1 \leq l \leq 4$) observation sites, the remaining wafer observation sites, and the equatorial region are given by,

$$
1_{w,k}(r) = \bigcup_{w=1}^{W_r} \bigcup_{k=1}^{k,w,F_r} 1_{w,k}(r)
$$

(3.7)
<table>
<thead>
<tr>
<th>Type 1</th>
<th>Type 2</th>
<th>Type 3</th>
<th>Type 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_r$</td>
<td>0</td>
<td>$v_r$</td>
<td>$v_r$</td>
</tr>
<tr>
<td>$l_r$</td>
<td>$v_r$</td>
<td>$v_r$</td>
<td>$L_r - v_r$</td>
</tr>
<tr>
<td>$t_r$</td>
<td>0</td>
<td>$J_r - v_r$</td>
<td>$J_r - v_r$</td>
</tr>
<tr>
<td>$m_r$</td>
<td>$J_r - v_r$</td>
<td>$u_r$</td>
<td>$J_r - u_r$</td>
</tr>
</tbody>
</table>

Table 3.1: Sizes of Different Tile Types

\[
R(r) = \left( \bigcup_{w=1}^{W} R_w(r) \right) \cup R_E = R^*(r) \cup R_E \tag{3.8}
\]

\[
E(r) = E^*(r) \cup R_E \tag{3.9}
\]

Given any nearly uniform stratified spherical sampling plan, the “big-block, little-block” technique can be employed, using the two caps, the remaining region, the equatorial region, and the type 1, 2, and 3 tiles as the little blocks, and the type 4 tiles as the big blocks.

By the construction of any nearly uniform stratified spherical sampling plan, each different type of corridor tile sum is distributionally equivalent to the same type of corridor tile sum within the same wafer, but is not distributionally equivalent to the same type of corridor tile sum on other wafers. We show that the corridor tile sums can be approximated by independent corridor tile sums, which in turn are approximated by planar corridor block sums. The independence approximation requires a telescoping argument for the characteristic functions of the collections of the tile sums for each of the 4 types throughout the sphere, even though these tile sums are not exactly identically distributed. Lemma 3.4 finds conditions for arbitrary random variables that are separated by some distance to be asymptotically independent.

Lemma 3.5 then applies this result to each collection of different tile types, using the planar approximation given by Lemma 2.1. Finally, Lemma 3.6 gives conditions for the negligibility of the corridor tile sums, the remaining tile region, the equatorial region, and both spherical caps.

**Lemma 3.4** Let \( \{M_n\} \) be a sequence for which \( M_n \to \infty \) as \( n \to \infty \). For each \( m \geq 1 \), let \( K_m(n) \to \infty \) as \( n \to \infty \). For each \( n \geq 1, m \geq 1, \) and \( k \geq 1 \), let \( Y_{m,k}^n \) be a random variable, measurable with respect to \( \sigma(\Lambda_{m,k}^n) \), where \( \Lambda_{m,k}^n \subset \mathbb{R}^3 \), and the underlying random field is homogenous and isotropic satisfying the strong \( \alpha \)-mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and define \( \bar{Y}_{m,k}^n = Y_{m,k}^n \) where \( \{\bar{Y}_{m,k}^n\} \) are independent. Assume that for each \( n \geq 1 \) and \( m \geq 1 \), \( Y_{m,k}^n \sim F_m^n \) for all \( k \geq 1 \). If

\[
\sum_{m=1}^{M_n} K_m(n) \sum_{k=1}^{K_m(n)} \alpha(d_{m,k}^n) \to 0 \text{ as } n \to \infty
\]

then

\[
\left| \left( \exp \left( \sum_{m=1}^{M_n} K_m(n) \sum_{k=1}^{K_m(n)} \bar{Y}_{m,k}^n \right) \right) - \left( \exp \left( \sum_{m=1}^{M_n} K_m(n) \sum_{k=1}^{K_m(n)} Y_{m,k}^n \right) \right) \right| \to 0
\]
where
\[
d_{m,k}^n = \min_{(m',k') \in \mathcal{A} \setminus \{(m,k)\}} d(\Lambda_{m,k}^n, \Lambda_{m',k'}^n).
\]

Let \( t(A) = \sum_{T \in A} T \). For each nearly uniform stratified spherical sampling plan, the sums on each tile type, \( t\left( T_{w,k}(r) \right) \), act as a \( Y_{m,k}^n \) and the set of points that comprise each tile part, \( T_{w,k}(r) \), is \( \Lambda_{m,k}^n \). The distance between the \( T_{w,k}(r) \)'s can be explicitly calculated from each nearly uniform stratified spherical sampling plan. Lemma 3.5 proves that the type 1, 2, 3, and 4 tiles can be asymptotically approximated by independent tile sums of the same cardinality and Lemma 3.6 and Corollary 3.3 show the asymptotic negligibility of all of the corridor tile sums when the random process is bounded and unbounded, respectively.

**Lemma 3.5** Let \( \{X_T\} \) be an isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth covariance function. Observe the random process according to a nearly uniform stratified spherical sampling plan \( \mathcal{R} \) with \( n_r = |\mathcal{R}| \) and assume that \( \exists m \geq 1 \) such that

1. \( \left( \frac{1}{v_r} \right)^{m+1} \left( \frac{1}{f_r \varphi(r)} \right)^2 = O(1) \) and
2. \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty. \)

Define for \( 1 \leq l \leq 4, 1 \leq w \leq W_r, 1 \leq k \leq k_{w,l,r}, \tilde{t}(T_{w,k}(r)) \) \( \overset{D}{=} t(T_{w,k}(r)) \) where \( \{\tilde{t}(T_{w,k}(r))\} \) are independent. Then as \( r \to \infty, \)

\[
\left| \mathbb{E} \left\{ \exp \left\{ is \left( \frac{\tilde{t}(T(r))}{\sqrt{n_r}} \right) \right\} \right\} - \mathbb{E} \left\{ \exp \left\{ is \left( \frac{t(T(r))}{\sqrt{n_r}} \right) \right\} \right\} \right| \to 0.
\]

This result is the driving force in determining the \( \alpha \)-mixing condition for approximating each tile type sum by independent tile type sums of the same cardinality. Conditions (1) and (2), together, describe the dependence structure of the random process. Condition (1) explicitly places a rate on the size of the sphere and its angles, while condition (2) gives the dependence structure, in terms of an \( \alpha \)-mixing constraint, for the spherical random field. Now we apply Lemma 3.5 to each corridor tile and use the Markov inequality to get the negligibility of the corridor tile sums, the equatorial region, and spherical cap regions over the sphere.

**Lemma 3.6** Let \( \{X_T\} \) be a bounded, isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth covariance function. Observe the random process according to a nearly uniform stratified spherical sampling plan \( \mathcal{R} \) with \( n_r = |\mathcal{R}| \) and assume that \( \exists m \geq 1 \) such that

42
1. \( \left( \frac{1}{n_r} \right)^{m+1} \left( \frac{1}{J_r \phi(r)} \right)^2 = O(1) \) and

2. \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty. \)

Then, for \( \sigma^2 \) given by (3.5), as \( r \to \infty, \)

\[
\frac{1}{\sigma \sqrt{n_r}} t \left( \mathcal{P} \setminus \mathcal{T}(r) \right) F \to 0.
\]

Corollary 3.3 Let \( \{X_r\} \) be an isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth covariance function. Observe the random process according to a nearly uniform stratified spherical sampling plan \( \mathcal{P} \) with \( n_r = |\mathcal{P}| \) and assume that

1. \( \exists \delta > 0 \) such that \( \mathbb{E} \left\{ |X_r|^{2+\delta} \right\} < \infty \) and

2. \( \sum_{j=1}^{\infty} j \alpha(j)^{\delta/(2+\delta)} < \infty, \) and

3. \( \exists m \geq 1 \) such that

   (a) \( \left( \frac{1}{n_r} \right)^{m+1} \left( \frac{1}{J_r \phi(r)} \right)^2 = O(1) \)

   (b) \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty. \)

Then, for \( \sigma^2 \) given by (3.5), as \( r \to \infty, \)

\[
\frac{1}{\sigma \sqrt{n_r}} t \left( \mathcal{P} \setminus \mathcal{T}(r) \right) F \to 0.
\]

These results prove that the corridor tile sums are negligible. Conditions (1) and (2) assure that the tile variance is finite. Together conditions (3a) and (3b) describe the dependence structure of the random process. Condition (3a) explicitly places a rate on the size of the sphere and its angles, while condition (3b) gives the dependence structure, in terms of an \( \alpha \)-mixing constraint, for the spherical random field.

### 3.4.3 Asymptotic Normality of the Standardized Sample Mean

With all of these pieces of our proof at hand, we only need to prove that the Lindeberg condition is satisfied for the associated independent type 4 tile sums for any nearly uniform
stratified spherical sampling plan. Lemma 3.7 proves asymptotic normality for a bounded random process and then this result is extended to Theorem 3.1 for unbounded random variables.

Lemma 3.7 Let \( \{X_x\} \) be a bounded, isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth covariance function. Observe the random process according to a nearly uniform stratified spherical sampling plan \( \mathcal{R} \) and assume that \( \exists \gamma \geq 1 \) such that

1. \( \left( \frac{1}{v_r} \right)^{m+1} \left( \frac{1}{J_r \phi(r)} \right)^{2} = O(1) \)

2. \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty. \)

Then for \( \sigma^2 \) given by (3.5), as \( r \to \infty \),

\[
\frac{1}{\sigma \sqrt{\text{vol}(S)}} \sum_{x \in \mathcal{R}} X_x \xrightarrow{D} N(0,1).
\]

It is interesting to note that the Lindeberg condition is trivially satisfied in this case simply because of the definition of a nearly uniform stratified spherical sampling plan. This means that there is probably a better sampling plan that gives asymptotic normality of the standardized sample mean in a "tighter" sense.

Theorem 3.1 for unbounded random variables can then be proved by truncating the random variables, \( X_x \), in \( \tilde{T}(r) \) at \( N_k \). Using the smoothness condition, we apply Lemma 3.7 to the truncated random variables and prove that the sum of the truncated parts is negligible, giving the desired result.

Theorem 3.1 Let \( \{X_x\} \) be an isotropic and homogeneous, mean-zero random process in \( \mathbb{R}^3 \) satisfying the strong mixing condition with \( \alpha \)-mixing coefficient \( \alpha(k) \) and having a smooth covariance function. If the random process is observed via a nearly uniform stratified spherical sampling plan \( \mathcal{R} \), and

1. \( \exists \delta > 0 \) such that \( \mathbb{E} \left\{ |X_x|^{2+\delta} \right\} < \infty \)

2. \( \sum_{j=1}^{\infty} j \alpha(j)^{4/(2+\delta)} < \infty \), and

3. \( \exists \gamma \geq 1 \) such that

44
(a) \( \left( \frac{1}{v_r} \right)^{m+1} \left( \frac{1}{J_r \phi(r)} \right)^2 = O(1) \)

(b) \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty. \)

Then for \( \sigma^2 \) given by (9.5), as \( r \to \infty \),

\[
\frac{1}{\sigma \sqrt{|P|}} \sum_{\beta \in P} X_{\beta} \overset{D}{\to} \mathcal{N}(0, 1).
\]
Chapter 4

Bootstrap Algorithm

4.1 Introduction

In this chapter we formally introduce the bootstrap algorithm and prove strong uniform consistency of the bootstrap cumulative distribution function of the standardized sample mean. In section 4.2, a bootstrap algorithm for the sphere, the stratified-tilewise bootstrap algorithm, is given along with some simple consequences when it is applied to the standardized sample mean. In section 4.3, some bounds on the moments of certain $\alpha$-mixing spatial random processes are given. These bounds are then used in section 4.4 to prove that the bootstrap estimate of variance converges both in probability and almost surely to its true value of $\sigma^2$. Finally in section 4.5, strong uniform consistency of the bootstrap distribution of the standardized sample mean is proved.

4.2 Stratified - Tilewise Bootstrap Algorithm

As has been noted before, the original sampling plan plays a major role in the development of a bootstrap algorithm. A poorly designed sampling plan would not allow for a bootstrap algorithm which can resample pieces that have the same dependence structure and also overlap the tiles in an efficient manner. Fortunately, the sampling plan that was given in Chapter 2 does allow us to create a bootstrap algorithm which satisfies all of the conditions set forth for a prospective bootstrap algorithm given in section 1.2.3.

Necessarily for each $1 \leq w \leq W_r/2$ and for every $0 \leq i, i' \leq n_{w,r} - 1$, the subtiles $wS_{i,0}^{i',r}$ and $wS_{i',0}^{i,r}$ are identically distributed. In addition, the subtiles of the equatorial region,
$E_{i,0}^{r}$ are also identically distributed, for every $0 \leq i \leq n_{B,r} - 1$. In what follows define for a nearly uniform stratified spherical sampling plan $P$, $X_r = \{X_P : P \in P\}$ to be the realization of the isotropic and homogeneous random field observed via $P$.

Recall that there are $k_{w,l_r}$ tiles of length $l_r$ that are completely contained in wafer $w$ and one tile of length $m_{w,l_r} = n_{w,r} - k_{w,l_r} l_r$ that is left over. Since each wafer $w$ wraps around on itself, there are $n_{w,r}$ tiles of each of these lengths on each wafer and also $n_{w,r}$ tiles of these lengths on its symmetric counterpart wafer, $W_r + 1 - w$. So, for each $1 \leq w \leq W_r / 2$, let $\hat{F}_{l_r}$ and $\hat{F}_{m_{w,l_r}}$ be the empirical distribution functions on wafer $w$ which assigns probability $\frac{1}{2n_{w,r}}$ to each tile of length $l_r$ and $m_{w,l_r}$, respectively. Similarly let $E \hat{F}_{l_r}$ and $E \hat{F}_{m_{w,l_r}}$ be the empirical distribution function on the equatorial region which assigns probability $\frac{1}{2n_{w,r}}$ to each tile of length $l_r$ and $m_{E,l_r}$, respectively, of the equatorial region. Using this setup, we can now give the stratified-tilewise bootstrap algorithm:

**Algorithm 3** For $w \in W_r$, define $\hat{F}_{l_r}$, $\hat{F}_{m_{w,l_r}}$, $E \hat{F}_{l_r}$, and $E \hat{F}_{m_{w,l_r}}$ as above. For each $w \in W_r$, sample $k_{w,l_r} = \lfloor n_{w,r} / l_r \rfloor$ subtiles from $\hat{F}_{l_r}$ and one subtile from $\hat{F}_{m_{w,l_r}}$ to complete the wafer. Call these tiles $\hat{S}_1^w$, $\hat{S}_2^w$, ..., $\hat{S}_{k_{w,l_r}}^w$. Then repeat the same for its symmetric counterpart wafer $W_r + 1 - w$. Call these tiles $\hat{S}_1^w$, $\hat{S}_2^w$, ..., $\hat{S}_{k_{w,l_r}}^w$. When $w = E$ sample $k_{E,l_r}$ tiles from $E \hat{F}_{l_r}$ and one from $E \hat{F}_{m_{E,l_r}}$ and call these tiles $E \hat{S}_1^w$, $E \hat{S}_2^w$, ..., $E \hat{S}_{k_{E,l_r}}^w$. Since the spherical caps have a different sampling plan, we will simply leave them as they are; that is, sample the spherical caps as themselves. The union of all of these resampled tiles and the two caps comprise the bootstrap sphere $\hat{X_r}$.

The statistic of interest, $t_{n}$, can then be calculated on the bootstrap sphere to create a single bootstrap statistic $t_n^* = t_n(\hat{X}_r)$. Repeat this process a large number $B$ times.

Notice that we could re-sample the caps by randomly interchanging them. In addition, since the caps have circular symmetry we could also randomly rotate them at angles of $\frac{\pi}{2}$.

In particular, suppose the statistic of interest is the standardized sample mean. Algorithm 3 is equivalent to the following setup. Define $G_{n^*(r)} \hat{X}$, $G_{n^*(r)} \hat{X}$, $E \hat{X}$, and $W \hat{X}$ to be the average of the random process on the northern cap, southern cap, equatorial region, and the rest of the sphere, respectively. For each $w \in W_r \setminus \{E\}$, define

\[
\hat{X}_{l_r}^w = \frac{1}{l_r J_r} \sum_{i=0}^{l_r-1} \sum_{j=0}^{J_r} X_{P_{i,j}}^{w^*}(r),
\]

\[
\hat{X}_{m_{w,l_r}}^w = \frac{1}{m_{w,l_r} J_r} \sum_{i'=0}^{m_{w,l_r}-1} \sum_{j=0}^{J_r} X_{P_{i',j}}^{w^*}(r),
\]

\[
E \hat{X}_{l_r}^w = \frac{1}{l_r v_r} \sum_{i'=0}^{l_r-1} \sum_{j=0}^{v_r-1} X_{P_{i',j}}^{w^*}(r)
\] and

\[
E \hat{X}_{m_{w,l_r}}^w = \frac{1}{m_{w,l_r} v_r} \sum_{i'=0}^{m_{w,l_r}-1} \sum_{j=0}^{v_r-1} X_{P_{i',j}}^{w^*}(r)
\]
\[
\hat{X}_{m_{E, m_{E, r}}} = \frac{1}{m_{E, m_{E, r}}} \sum_{i=0}^{i+m_{E, m_{E, r}}-1} \sum_{j=0}^{v_{E, m_{E, r}}-1} X_{p_{i, j}}(r).
\]

For \(1 \leq k \leq m_{w, m_{w, r}}\), let

\[
\begin{align*}
\hat{X}_{l_{r}} &= \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{i+m_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r), \\
E_{k} U_{l_{r}} &= \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{i+m_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r), \\
E_{k} U_{m_{w, m_{w, r}}} &= \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{i+m_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r), \\
C_{N}(r) \hat{X} &= \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{i+m_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r), \\
C_{S}(r) \hat{X} &= \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{i+m_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r)
\end{align*}
\]

We then have the following relationships

\[
\hat{X} = \frac{1}{n_{E, m_{E, r}}} \left\{ \left| C_{N}(r) \frac{C_{S}(r)}{C_{N}(r)} \hat{X} + C_{S}(r) \hat{X} \right| + |E(r)| E_{k} U_{l_{r}} + N_{E} \hat{X} \right\} \quad \text{and}
\]

\[
\hat{X}^* = \frac{1}{n_{E, m_{E, r}}} \left\{ \left| C_{N}(r) \frac{C_{S}(r)}{C_{N}(r)} \hat{X}^* + C_{S}(r) \hat{X}^* \right| + |E(r)| E_{k} U_{m_{w, m_{w, r}}} + N_{E} \hat{X}^* \right\}
\]

where

\[
N_{E} = J_{r} \left( \sum_{w=1}^{W_{r}} n_{w, r} \right),
\]

\[
E_{k} \hat{X}^* = \frac{1}{n_{E, m_{E, r}}} \left\{ \left| C_{N}(r) \frac{C_{S}(r)}{C_{N}(r)} \hat{X} + C_{S}(r) \hat{X} \right| + |E(r)| E_{k} U_{l_{r}} + m_{E, m_{E, r}} v_{E, m_{E, r}} E_{k} U_{m_{E, m_{E, r}}} \right\}, \quad \text{and}
\]

\[
W_{k} \hat{X}^* = \frac{1}{N_{E}} \left\{ \sum_{w=1}^{W_{k}/2} \left( \left| C_{N}(r) \frac{C_{S}(r)}{C_{N}(r)} \hat{X} + C_{S}(r) \hat{X} \right| + |E(r)| E_{k} U_{l_{r}} + m_{w, m_{w, r}} v_{w, m_{w, r}} E_{k} U_{m_{w, m_{w, r}}} \right) \right\}
\]

With the statistic of interest for \(t_{n}(\cdot)\) in Algorithm 3 as the sample mean, we can find a form for the bootstrap estimate of variance of the sample mean, \(\hat{\sigma}_{\text{boot}}^{2}\). For \(w \in W_{r} \setminus \{E\}\), let

\[
w \hat{X} = \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{n_{w, m_{w, r}}-1} \sum_{j=0}^{v_{w, m_{w, r}}-1} X_{p_{i, j}}(r)
\]

\[
w \hat{X}^* = \frac{1}{2} \left( w \hat{X} + \left( w \hat{X} + 1 \right) \hat{X} \right),
\]

\[
w \hat{X}^2 = \frac{1}{n_{w, m_{w, r}}} \sum_{i=0}^{n_{w, m_{w, r}}-1} \left( w \hat{X}_{i} - w \hat{X} \right)^{2}
\]

48
\[
\begin{align*}
w_{s_{l_r}} & = \frac{1}{2} \left( w_{s_{1_r}} + w_{s_{r+1}} - w_{s_{l_r}} \right), \text{ and} \\
E_{s_{l_r}}^2 & = \frac{1}{n_{E,r}} \sum_{i=0}^{n_{E,r}-1} \left( E_{l_r} \hat{X}_{l_r} - E \hat{X} \right)^2.
\end{align*}
\]

We obtain the following consequences for the expectation, \( \mathbb{E}^* \{ \cdot | \hat{X}_r \} \), and variance, \( \mathbb{V}^* \{ \cdot | \hat{X}_r \} \), under the stratified-tilewise bootstrap algorithm:

\[
\begin{align*}
\mathbb{E}^* \left\{ C_N(r) \hat{X}^* | \hat{X}_r \right\} & = C_N(r) \hat{X} \\
\mathbb{E}^* \left\{ C_S(r) \hat{X}^* | \hat{X}_r \right\} & = C_S(r) \hat{X} \\
\mathbb{E}^* \left\{ E \mathbb{U}_{l_r} | \hat{X}_r \right\} & = E \hat{X} \\
\mathbb{E}^* \left\{ E \mathbb{U}_{m_{l_r}, l_r} | \hat{X}_r \right\} & = E \hat{X} \\
\mathbb{E}^* \left\{ \mathbb{W}_{l_r} | \hat{X}_r \right\} & = \mathbb{W} \hat{X} \\
\mathbb{E}^* \left\{ \mathbb{W}_{m_{l_r}, l_r} | \hat{X}_r \right\} & = \mathbb{W} \hat{X} \\
\mathbb{V}^* \left\{ C_N(r) \hat{X}^* | \hat{X}_r \right\} & = 0 \\
\mathbb{V}^* \left\{ C_S(r) \hat{X}^* | \hat{X}_r \right\} & = 0 \\
\mathbb{V}^* \left\{ E \mathbb{U}_{l_r} | \hat{X}_r \right\} & = E_{s_{l_r}}^2 \\
\mathbb{V}^* \left\{ E \mathbb{U}_{m_{l_r}, l_r} | \hat{X}_r \right\} & = E_{s_{m_{l_r}, l_r}}^2 \\
\mathbb{V}^* \left\{ \mathbb{W}_{l_r} | \hat{X}_r \right\} & = \mathbb{W}_{s_{l_r}}^2 \\
\mathbb{V}^* \left\{ \mathbb{W}_{m_{l_r}, l_r} | \hat{X}_r \right\} & = \mathbb{W}_{s_{m_{l_r}, l_r}}^2.
\end{align*}
\]

Using these results and the definition of \( \hat{X}^* \), we find that,

\[
\mathbb{E}^* \left\{ \hat{X}^* | \hat{X}_r \right\} = \hat{X} \text{ and} \\
\mathbb{V}^* \left\{ \hat{X}^* | \hat{X}_r \right\} = \frac{1}{\left| \mathcal{I} \right|^2} \sum_{w \in \mathcal{W}_r^\bot} \left( k_{w, l_r} (h_{w, l_r})^2 w_{s_{l_r}}^2 + (m_{w, l_r} h_{w, r})^2 w_{s_{m_{l_r}, l_r}}^2 \right) = \hat{\sigma}_{\text{boot}}^2 (4.1)
\]

where \( h_{w, r} \) is \( j_r \) or \( v_r \) depending on whether \( w \neq E \) or \( w = E \) respectively. So unlike the blockwise bootstrap of Künsch [Kun89], the natural estimator, \( \mathbb{E}^* \left\{ \hat{X}^* | \hat{X}_r \right\} \), is unbiased for \( \mu \); that is, \( \mathbb{E} \left\{ \mathbb{E}^* \left\{ \hat{X}^* | \hat{X}_r \right\} \right\} = \mathbb{E} \{ \hat{X} \} = \mu \). This is because each wafer “wraps around on itself” and the time-series does not.

### 4.3 Bounds on Moments

We must now find the expected value and variance of \( \hat{\sigma}_{\text{boot}}^2 \). In order to accomplish this, we must first derive some bounds on moments. Lemma 4.1 is given so that we can use a common constant throughout each proof. The next 5 corollaries extend Lemma 3.2 to obtain an expression for the bound on the absolute value of eighth-order moments. All of these corollaries pertain to the sums on a given wafer and in particular to a random process observed on the boundary of a circle at equiangular increments, as it is for each wafer in our nearly uniform stratified spherical sampling plan \( \mathcal{I} \) at each latitudinal cut.

**Lemma 4.1** If \( \mathbb{E} \{ |X_{\mathcal{I}}|^{(2q+4)} \} \leq c_q(\delta) < \infty, \mathbb{E} \{ X_{\mathcal{I}} \} = 0, \text{ and } \{ X_{\mathcal{I}} \} \text{ is a strictly stationary random process, then for integer } r \text{ such that } 1 < r \leq q,

\[
\mathbb{E} \left\{ |X_{r_1} X_{r_2} \cdots X_{r_r}|^{2q+4} \right\} \leq \max \{ c_q(\delta), 1 \} = \tilde{c}_q(\delta).
\]
We are interested in random variables that are observed on the boundary of a circle at equiangular intervals, like the observations on any wafer at any vertical phi angle. For this reason, we need to describe the the meaning of a circular random process and also give its "ordering".

Let the circular process \( \{ X_{\bar{\phi}_i} : 0 \leq i < k \} \) be a homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \) such that for some integer \( k \), \( k \xi = 2\pi \). We will assume that \( \bar{\phi}_i \) is observed at angle \( i \xi \) and hence the points \( \bar{\phi}_1, \bar{\phi}_2, \ldots, \bar{\phi}_m \) are ordered if \( 0 \leq i_1 < i_2 < \cdots < i_m \leq k \). We let \( \bar{\phi} = (\bar{\phi}_1, \bar{\phi}_2, \ldots, \bar{\phi}_m) \) for any positive integer \( m \). Since we always need a moment higher than the first to exist, we may center the process. Therefore, without loss of generality, we assume that \( \mathbb{E} \{ X_{\bar{\phi}} \} = 0 \).

**Corollary 4.1** Let \( \{ X_{\bar{\phi}_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( \mathbb{E} \{ X_{\bar{\phi}} \} = 0 \) and \( \mathbb{E} \{ |X_{\bar{\phi}}|^{2+\delta} \} \leq c(\delta) < \infty \), then for any two ordered points \( \bar{\phi}_{i_1}, \bar{\phi}_{i_2} (1 \leq j \leq 2) \) such that \( i_2 - i_1 \leq k/2 \),

\[
\Delta_2(\bar{\phi}_{i_1, i_2}) = \mathbb{E} \left\{ X_{\bar{\phi}_{i_1}} X_{\bar{\phi}_{i_2}} \right\} \leq \tilde{c}_1(\delta) \left( \alpha(m(\bar{\phi}_{i_1, i_2})) \right)^{4/(2+\delta)}
\]

where \( \tilde{c}_1(\delta) \) is some constant depending on \( \delta \) and \( m(\bar{\phi}_{i_1, i_2}) \) is defined by

\[
m(\bar{\phi}_{i_1, i_2}) = \max_{1 \leq i \leq n} \left\{ \| \bar{\phi}_n - \bar{\phi}_{i-1} \| \right\}.
\]

**Corollary 4.2** Let \( \{ X_{\bar{\phi}_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( \mathbb{E} \{ X_{\bar{\phi}} \} = 0 \) and \( \mathbb{E} \{ |X_{\bar{\phi}}|^{2+\delta} \} \leq c(\delta) < \infty \), then for any three ordered points \( \bar{\phi}_{i_1}, \bar{\phi}_{i_2}, \bar{\phi}_{i_3} (1 \leq j \leq 3) \) such that \( i_3 - i_1 \leq k/2 \),

\[
\Delta_3(\bar{\phi}_{i_1, i_2, i_3}) = \mathbb{E} \left\{ X_{\bar{\phi}_{i_1}} X_{\bar{\phi}_{i_2}} X_{\bar{\phi}_{i_3}} \right\} \leq \tilde{c}_2(\delta) \left( \alpha(m(\bar{\phi}_{i_1, i_2, i_3})) \right)^{6/(4+\delta)}
\]

where \( m(\bar{\phi}_{i_1, i_2, i_3}) \) is defined by (4.2) and \( \tilde{c}_2(\delta) \) is some constant depending on \( \delta \).

**Corollary 4.3** Let \( \{ X_{\bar{\phi}_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( \mathbb{E} \{ X_{\bar{\phi}} \} = 0 \) and \( \mathbb{E} \{ |X_{\bar{\phi}}|^{6+\delta} \} \leq c(\delta) < \infty \), then for any four ordered points \( \bar{\phi}_{i_1}, \bar{\phi}_{i_2}, \bar{\phi}_{i_3}, \bar{\phi}_{i_4} (1 \leq j \leq 4) \) such that \( i_4 - i_1 \leq k/2 \),

\[
\Delta_4(\bar{\phi}_{i_1, i_2, i_3, i_4}) = \left| \mathbb{E} \left\{ \prod_{j=1}^{4} X_{\bar{\phi}_{i_j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\bar{\phi}_{i_j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{\phi}_{i_j}} \right\} \right| \\
\leq \tilde{c}_3(\delta) \left( \alpha(m(\bar{\phi}_{i_1, i_2, i_3, i_4})) \right)^{4/(6+\delta)}
\]

where \( m(\bar{\phi}_{i_1, i_2, i_3, i_4}) \) is defined by (4.2) and \( \tilde{c}_3(\delta) \) is some constant depending on \( \delta \).

**Corollary 4.4** Let \( \{ X_{\bar{\phi}_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( \mathbb{E} \{ X_{\bar{\phi}} \} = 0 \) and
\[ E \left\{ |X_{\varphi}^{1+4+4}| \right\} \leq c(\delta) < \infty, \text{ then for any five ordered points } \vec{v}_{ij} (1 \leq j \leq 5) \text{ such that } \delta - \delta_i \leq k/2, \]

\[ \Delta_5 (\vec{v}_{1,5}) = \left| E \left\{ \prod_{j=1}^{5} X_{\varphi_{ij}} \right\} - \sum_{k=2}^{3} E \left\{ \prod_{j=1}^{k} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=k+1}^{5} X_{\varphi_{ij}} \right\} \right| \]

\[ \leq c_4(\delta) [\alpha (m (\vec{v}_{1,5}))]^{4/(8+4)} \]

where \( m (\vec{v}_{1,5}) \) is defined by (4.2) and \( c_4(\delta) \) is some constant depending on \( \delta \).

Corollary 4.5 Let \( \{X_{\varphi_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( E \{X_{\varphi} \} = 0 \) and \( E \left\{ |X_{\varphi}^{1+4+4}| \right\} \leq c(\delta) < \infty, \text{ then for any six ordered points } \vec{v}_{ij} (1 \leq j \leq 6) \text{ such that } \delta - \delta_i \leq k/2, \]

\[ \Delta_6 (\vec{v}_{1,6}) = \left| E \left\{ \prod_{j=1}^{6} X_{\varphi_{ij}} \right\} - \sum_{k=2}^{4} E \left\{ \prod_{j=1}^{k} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=k+1}^{6} X_{\varphi_{ij}} \right\} + \right. \]

\[ \left. + E \left\{ \prod_{j=1}^{2} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=3}^{4} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=5}^{6} X_{\varphi_{ij}} \right\} \right| \]

\[ \leq c_6(\delta) [\alpha (m (\vec{v}_{1,6}))]^{4/(10+4)} \]

where \( m (\vec{v}_{1,6}) \) is defined by (4.2) and \( c_6(\delta) \) is some constant depending on \( \delta \).

Corollary 4.6 Let \( \{X_{\varphi_i} : 0 \leq i < k \} \) be a circular homogeneous and isotropic random field observed on the boundary of a circle at equiangular increments of angle \( \xi \). If \( E \{X_{\varphi} \} = 0 \) and \( E \left\{ |X_{\varphi}^{1+4+4+4}| \right\} \leq c(\delta) < \infty, \text{ then for any eight ordered points } \vec{v}_{ij} (1 \leq j \leq 8) \text{ such that } \delta - \delta_i \leq k/2, \]

\[ \Delta_8 (\vec{v}_{1,8}) = \left| E \left\{ \prod_{j=1}^{8} X_{\varphi_{ij}} \right\} - \sum_{k=2}^{6} E \left\{ \prod_{j=1}^{k} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=k+1}^{8} X_{\varphi_{ij}} \right\} + \right. \]

\[ \left. \sum_{k=2}^{6} E \left\{ \prod_{j=1}^{k} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=k+1}^{8} X_{\varphi_{ij}} \right\} E \left\{ \prod_{j=k+i+1}^{8} X_{\varphi_{ij}} \right\} \right| \]

\[ \leq c_8(\delta) [\alpha (m (\vec{v}_{1,8}))]^{4/(14+4)} \]

where \( m (\vec{v}_{1,8}) \) is defined by (4.2) and \( c_8(\delta) \) is some constant depending on \( \delta \).
4.4 Convergence of the Bootstrap Estimate of Variance of the Sample Mean

With these results in hand, we are now able to investigate the consistency of the bootstrap estimate of variance of the sample mean.

**Theorem 4.1** Consider a nearly uniform stratified spherical sampling plan \( \mathcal{P} \) and a homogeneous and isotropic random field \( \{X_p\} \). If

1. \( l_r \ll J_r \)
2. \( \mathbb{E}\left\{ |X_p|^{2+\delta} \right\} < \infty \) for some \( \delta > 0 \) and
3. \( \sum_{j=1}^{\infty} j^m \alpha(j) j^{(2+\delta)} < \infty \) for \( m \) such that \( \frac{J^2}{v_{m+1}} = O(1) \)

then for \( \sigma^2 \) given by (3.5),

\[
\mathbb{E}\left\{ |\mathcal{P}| \hat{\sigma}_{\text{boot}}^2 \right\} - \sigma^2 \xrightarrow{r} 0.
\]

Condition (1) is used to keep the height and the width of each resampled tile approximately the same. Conditions (2) and (3) are used to create approximately independent tiles and bound the tile sums. Notice that the correct standardization is \( |\mathcal{P}| \), which is the same standardization that is used in Künsch [Kun89]. This is a direct consequence of the fact that the variance of a tile sum is of the order of the number of elements in the sum.

In order to find the rate that the bootstrap estimate of variance of the standardized sample mean converges to \( \sigma^2 \), we need to state a result on the standardized sum of a column of each tile. This result uses the corollaries pertaining to the bounding of the \( n \)th-order \((2 < n \leq 8)\) moments given previously. In particular, let \( \Psi S_1 \) be the sum of observations on wafer \( w \) beginning at a theta angle of \( i\theta_w(r) \) and ending at angle \( (i + l - 1)\theta_w(r) \) and beginning at phi angle \( \gamma_w(r) \) and ending at angle \( \gamma_w(r) + (J_r - 1)\phi(r) \). So \( \Psi S_1 \) is the column sum of observations at angle \( 0 \) on wafer \( w \).

**Theorem 4.2** Consider a nearly uniform stratified spherical sampling plan \( \mathcal{P} \) and a homogeneous and isotropic random field \( \{X_p\} \). If \( \mathbb{E}\{X_p\} = 0 \) and for some \( \delta > 0 \)

\[
\mathbb{E}\left\{ |X_p|^{14+\delta} \right\} < \infty \quad \text{and} \quad \sum_{j=1}^{\infty} j^3 \alpha(j) j^{4(14+\delta)} < \infty
\]

then for some \( \epsilon > 0 \),

\[
\sup_r \sup_{w \in W_r \setminus E} \mathbb{E}\left\{ \frac{|\psi S_1}{\sqrt{J_r}} \right\}^{6+\epsilon} < \infty.
\]
The conditions of this theorem are necessary in order to apply the results of the previous corollaries. Theorem 4.2 guarantees that Lemma 3.2 can be applied to tile sums, not just to individual observations. We can now show that the variance of the standardized bootstrap estimate of variance decays at a very quick rate, and in fact, this rate guarantees that it is almost surely consistent (using the Borel-Cantelli lemma).

**Theorem 4.3** Consider a nearly uniform stratified spherical sampling plan $\mathcal{P}$ and a homogeneous and isotropic random field $\{X_\mathcal{P}\}$. If

1. $l_r \asymp J_r$,

2. $\sum_{j=1}^{\infty} j^2 \alpha^{1/4}(j) < \infty$,

3. for some $\delta > 0$, $\mathbb{E}\left\{ |X_\mathcal{P}|^{14+\delta} \right\} < \infty$, $\sum_{j=1}^{\infty} j^3 \alpha(j)^{5/(14+\delta)} < \infty$, and

4. $\sum_{j=1}^{\infty} j^m \alpha(j)^{5/(2+\delta)} < \infty$ for $m$ such that $\frac{r}{J^m_{\mathcal{P}}} = o(1)$ and $\frac{J^2_{\mathcal{P}}}{v^{m+1}_{\mathcal{P}}} = O(1)$

then

$$\mathbb{V}\{ |\mathcal{P}| \hat{\sigma}^2_{\text{boot}} \} = O\left( \frac{l_r J_r}{|\mathcal{P}|} \right)^r \rightarrow 0$$

and hence for $\sigma^2$ given by (3.5)

$$|\mathcal{P}| \hat{\sigma}^2_{\text{boot}} \overset{a.s.}{\rightarrow} \sigma^2.$$

Condition (1) keeps the tile length the same order as the tile height. Condition (2) is used so that we may treat the standardized vector sums as individual observations. Condition (3) allows us to apply Theorem 4.2. Condition (4) describes the dependence structure of the random process relative to the nearly uniform stratified spherical sampling plan. That is, it gives a relationship between the wafer thickness and the radius of the sphere, and the thickness of the separating strips relative to the thickness of each wafer.

In relation to the nearly uniform stratified sampling plan given in Table 2.1, this means that

$$\frac{1}{m-2} < \eta < \frac{1}{5}$$

and

$$\frac{2}{m+1} < \zeta < 1.$$

So if $m > \inf_{\eta}(2 + \frac{1}{\eta}) = 7$, then both of the conditions are satisfied.
4.5 Strong Uniform Convergence of the Bootstrap Cumulative Distribution Function of the Standardized Sample Mean

All that is left to be proved is that the bootstrap cumulative distribution function of the standardized sample mean on the sphere is strongly uniformly consistent. This comes down to showing that the Lindeberg condition is satisfied. The following result corresponding to the size of the supremum of the tile sums is fundamental step in this argument.

**Lemma 4.2** Consider a nearly uniform stratified spherical sampling plan \( P \) and a homogeneous and isotropic random field \( \{ X_\theta \} \). If \( \mathbb{E} \{ X_\theta \} = \mu < \infty \) and for some \( q > 0 \) we have \( X_\theta = \mathcal{O}(1) \) and \( \mathbb{E} \{ |X_\theta|^q \} < \infty \), then for \( l_r \approx J_r \)

\[
\sup_{u,t} \left| \sum_{i=1}^{l_r-1} \sum_{j=1}^{J_r} (X_{P_{(r)}(r)} - \mu) \right| \overset{a.s.}{=} o \left( \sqrt{R} \right).
\]

The conditions of this lemma are rather mild. With all the tools at hand, the strong uniform consistency of the bootstrap cumulative distribution function of the standardized sample mean is given.

**Theorem 4.4** Consider a nearly uniform stratified spherical sampling plan \( P \) and a homogeneous and isotropic random field \( \{ X_\theta \} \). If

1. \( l_r \approx J_r \),
2. \( \sum_{j=1}^{\infty} j^2 \alpha^{1/4}(j) < \infty \),
3. for some \( \delta > 0 \), \( \mathbb{E} \{ |X_\theta|^{14+\delta} \} < \infty \) and \( \sum_{j=1}^{\infty} j^3 \alpha(j)^{6/(14+\delta)} < \infty \),
4. \( \sum_{j=1}^{\infty} j^m \alpha(j)^{4/(14+\delta)} < \infty \) for \( m \) such that \( \frac{r}{j_r} = o(1), \ \left( \frac{1}{l_r^2} \right)^{m+1} \left( \frac{1}{j_r \phi(r)} \right)^2 = \mathcal{O}(1) \),
   and \( \frac{j_r^2}{l_r^2} = \mathcal{O}(1) \)
5. for some \( q > 0 \), \( \mathbb{E} \{ |X_\theta|^q \} < \infty \) and \( \frac{r}{j_r} = \mathcal{O}(1) \),

then

\[
\sup_{\sigma} \left| \mathbb{P} \left\{ \frac{\sqrt{|P|} (\bar{X} - \mu)}{\sigma} \leq x \right\} - \mathbb{P}^* \left\{ \frac{\sqrt{|P|} (\bar{X}^* - \bar{X})}{\sqrt{|P|} \bar{\sigma}_{boot}} \leq x | \bar{X}_r \right\} \right| \overset{a.s.}{=} 0.
\]

54
where $\sigma^2$ is given by (3.5) and $\tilde{\sigma}_{Kopt}^2$ is given by (4.1).

In particular, for the nearly uniform stratified spherical sampling plan given in Table 2.1, we need

$$\max \left\{ \frac{1}{m - 2}, \frac{1}{q} \right\} < \eta < \frac{1}{5}$$
$$\frac{2(1/\eta - 1)}{m + 1} \leq \zeta < 1$$
$$0 < \epsilon < 1 - 5\eta.$$  

There are many combinations of $m, q, \eta, \zeta$, and $\epsilon$ which satisfy the above conditions, a few are listed in Table 4.1. Recall that we still need the $\alpha$-mixing conditions and the moment conditions to be satisfied.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$q$</th>
<th>$\eta$</th>
<th>$\zeta$</th>
<th>$\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6</td>
<td>11/60</td>
<td>99/100</td>
<td>1/13</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>1/6</td>
<td>19/20</td>
<td>1/7</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>1/10</td>
<td>9/10</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Table 4.1: Specific Nearly Uniform Stratified Spherical Sampling Plans
Chapter 5

Applications

5.1 Analysis of the Earth's Land

One of the most obvious applications of the spherical plans given so far is to the globe. Given a nearly uniform sampling plan, we can sample anything on the globe at the prescribed points, assuming that the data arise from a homogeneous and isotropic random process. For example, we could take readings of the drilling depth to reach oil, of carbon monoxide levels, or of radiation levels. However, one of the most straightforward observations would be to take a reading of a random variable that is either 0 or 1.

Given a nearly uniform sampling plan $\mathcal{R}$, at each sampling point on the globe, we are either on land or on water. We let the random variable $X_p$ be

$$X_p = \begin{cases} 
1 & \text{if } p \text{ is on land} \\
0 & \text{if } p \text{ is on water}
\end{cases}$$

for $p \in \mathcal{R}$. Notice that this data has a nice local dependence structure. That is, if you are in water, you are most likely surrounded by water, and if you are on land then you are most likely surrounded by land. As you move further away from a point of water or land, you have no idea which type of surface you should be near. In order to gather data and apply the stratified-tilewise bootstrap algorithm, we must determine a suitable nearly uniform stratified spherical sampling plan.
5.2 A Nearly Uniform Stratified Spherical Sampling Plan for the Earth

Under the conditions of the sampling plan given in Chapter 2, a nearly uniform stratified spherical sampling plan for the globe is determined by angles $\phi(r), \theta_w(r), \gamma_w^r(r)$ and tile lengths $J_r$ and $n_r$. Since we would like to try various size combinations of the tiles and wafers, we need to choose these values carefully.

So that we may use the latitude and longitude markings on any global map, we decided to use a value of $\phi(r) = 1^\circ$. We also decided that an original tile size of $I_r = J_r = 10$ would be sufficient to allow us to make the tiles of three sizes, namely, $2 \times 2, 5 \times 5$, and $10 \times 10$. We further decided that the top of the first wafer should be $\gamma_w^1(r) = 15^\circ$, since planar global maps rarely go that high in the northern hemisphere and the polar maps usually coincide with our hexagonal sampling plan in that region.

Using this setup, we determine the tops of the wafers with a tile size of 10. Table 5.1 gives these angle values. Notice that if the effective tile size is smaller, say $I_r = 5$, then there are 31 wafers instead of 15 ($7 + 7 +$ equator) and we use the intermediate angle as the top of wafer $\gamma_w^r(r)$. For example, the top of wafer 4 in this setting is $\gamma_4^4(r) = 30^\circ$. As you can see from the table, the ratio of the horizontal to vertical distance is close to the target value of 1, indicating that the sampling plan is nearly uniform on the sphere.

Sampling along this sampling plan we get the globe picture as given in figure 5.1. Views 1–4 in the figure are all rotated at $90^\circ$ increments starting at the $0^\circ$ longitude of Greenwich, England. Each point of the sampling plan is therefore depicted twice in the full set of four views.

<table>
<thead>
<tr>
<th>Wafer</th>
<th>$\gamma_w(r)$</th>
<th>$\gamma_{w+1-w}(r)$</th>
<th>$\theta_w(r)$</th>
<th>$n_w,r$</th>
<th>distance ratio $h/v \approx 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15$^\circ$</td>
<td>165$^\circ$</td>
<td>$2\pi/90$</td>
<td>90</td>
<td>1.0351</td>
</tr>
<tr>
<td>2</td>
<td>25$^\circ$</td>
<td>155$^\circ$</td>
<td>$2\pi/144$</td>
<td>144</td>
<td>1.0565</td>
</tr>
<tr>
<td>3</td>
<td>35$^\circ$</td>
<td>145$^\circ$</td>
<td>$2\pi/240$</td>
<td>240</td>
<td>0.8604</td>
</tr>
<tr>
<td>4</td>
<td>45$^\circ$</td>
<td>135$^\circ$</td>
<td>$2\pi/240$</td>
<td>240</td>
<td>1.0606</td>
</tr>
<tr>
<td>5</td>
<td>55$^\circ$</td>
<td>125$^\circ$</td>
<td>$2\pi/360$</td>
<td>360</td>
<td>0.8192</td>
</tr>
<tr>
<td>6</td>
<td>65$^\circ$</td>
<td>115$^\circ$</td>
<td>$2\pi/360$</td>
<td>360</td>
<td>0.9063</td>
</tr>
<tr>
<td>7</td>
<td>75$^\circ$</td>
<td>105$^\circ$</td>
<td>$2\pi/360$</td>
<td>360</td>
<td>0.9659</td>
</tr>
<tr>
<td>8 (Equator)</td>
<td>85$^\circ$</td>
<td></td>
<td>$2\pi/360$</td>
<td>360</td>
<td>0.9962</td>
</tr>
</tbody>
</table>

Table 5.1: Wafer Angles and Number of Observations
Figure 5.1: Original Global Data
5.3 Stratified - Tilewise Bootstrap Algorithm

If the statistic of interest is the sample mean \((i.e., \text{proportion of land})\) using the stratified-tilewise bootstrap algorithm, we do not actually have to resample any data. The bootstrap estimates of variance and expectation can be calculated explicitly using the original data and the formulae given in Chapter 4:

\[
\begin{align*}
\mathbb{E}^* \{ \tilde{X}^* | \tilde{X}_r \} &= \bar{X} \\
\mathbb{V}^* \{ \tilde{X}^* | \tilde{X}_r \} &= \frac{1}{|\mathcal{R}|^2} \sum_{w \in \mathcal{W}_r^*} \left( k_{w,r} (h_{w,r}, l_r)^2 w_{l_r}^2 + (m_{w,r} h_{w,r})^2 w_{m_{w,r}}^2 \right).
\end{align*}
\]

When we set \(l_r = J_r\), we can calculate these values explicitly. Table 5.2 gives the exact bootstrap estimate of variance of the standardized sample mean depending on the tile size. (The bootstrap estimate of the mean does not depend on the tile size and is simply the sample mean of 0.2895.)

| Tile Size | \(\sqrt{\mathbb{V}^*} | \tilde{\sigma}_{\text{boot}}\) | Direct Estimate \((B = 100)\) |
|-----------|---------------------------------|---------------------|
| \(l_r = 2\) | 0.84969 0.74755 | 0.86303 |
| \(l_r = 5\) | 1.96355 1.90840 | 0.86303 |
| \(l_r = 10\) | 3.68573 3.94930 | 0.86303 |

Table 5.2: Bootstrap Estimates of Variance for Different Tile Sizes

As you can see from this table, the estimate of \(\sigma^2\) is still growing as the tile size increases. This is because the dependence is not wholly contained within a tile of size 2 or 5 or 10. In fact, using a tilesizes of 2, \(|\mathcal{R}| \tilde{\sigma}_{\text{boot}}^2\) is actually estimating \(\mathbb{V} \{ X_0 \} + 2 \mathbb{C} \{ X_0, X_{1,0} \} + \mathbb{C} \{ X_0, X_{1,1} \} \approx \mathbb{V} \{ X_0 \} + 3 \mathbb{C} \{ X_0, X_{1,0} \} \). Using the usual direct estimates of \(\mathbb{V} \{ X_0 \} (= 0.20568)\) and \(\mathbb{C} \{ X_0, X_{1,0} \} (= 0.17971)\), one arrives at the value of 0.86303 in column 4 of the table. Notice that for the case when \(l_r = 2\), the exact value of the bootstrap estimate of variance is estimating the variance quite well. Bootstrap estimates are also given where the number of bootstrap replicates is \(B = 100\). Notice that they are near the target levels of the exact bootstrap.

To get a visual representation of what the algorithm does to the actual sphere, we ran the stratified-tilewise bootstrap algorithm with tile sizes 2, 5, and 10. View 3 from the original data which the bootstrap algorithm is trying to emulate, along with View 3 for each tile size is displayed in Figure 5.2.

As you can see, the resampled spheres do not look anything like the original globe. The local dependence structure is guaranteed to asymptotically reflect the dependence structure of the
Figure 5.2: Stratified - Tilewise Bootstrap Algorithm
original sphere, but in the finite case the broader dependence structure is not preserved. If the information of a statistic is wholly contained in the tiles, then this method is reasonable. This is reflected in the fact that the bootstrap estimate of variance for a tile size of 2 was near its estimated value (see Table 5.2). For example, the sample mean (proportion of land) requires no knowledge of neighboring tiles in order to be calculated. However, if we are interested in the "coastline" statistic of the resampled object, these would not be viable resamples of the Earth. One way to overcome these dependence difficulties is to impose a linking mechanism at the boundaries of the tiles. This is similar to the idea of Künsch and Carlstein in the time-series case.

5.4 Bootstrap Algorithms with Linking

The idea of linking is, for each tile that is to be replaced, create a class of possible tiles from which to resample and then separate the class into two distinct sets: ones that are "close" to the one being replaced and ones that are not. Once we have identified the set of close tiles, we replace the tile in question with a tile from the set of close tiles via some random mechanism. The notion of closeness is quantified by many different aspects. Our random mechanism is just to select one of the close tiles at random, each with equal probability of being selected.

First we need to be able to characterize a tile. This is done by associating a tile with its boundary. We can either consider the inner boundary of a tile (its actual border) or its outer boundary (values just outside its border). Figure 5.3 shows a graphical depiction of these two types of borders. In the figure, the tile on wafer \( w \) itself is comprised of the values at locations 17–25. The tile’s inner boundary is comprised of elements 17–24, while its outer boundary is comprised of elements 1–16 which are outside of the tile. Elements 1–5 are from wafer \( w - 1 \), elements 9–13 are from wafer \( w + 1 \), elements 6–8 are from the tile on wafer \( w \) directly to the right of the tile and elements 14–16 are from the tile on wafer \( w \) directly to the left of the tile. The actual boundary of a tile is a vector of random values that the random process attains at the locations of the inner or outer boundary. So, for the global land data, the values in the boundary are either 0's or 1's.

We compare boundaries in a element-wise fashion. That is, if the tile to be replaced has a boundary of \((0, 0, 0, 0, 1, 1, 1, 0, 0, 0)\) and we compare it with a prospective tile from the resample class with a boundary of \((0, 0, 0, 1, 1, 0, 0, 0, 0, 0)\), then the boundaries are a “distance” of 3 apart. We calculate all of the distances from the tile to be replaced with tiles in the resampled class. We then separate the resample class into the set of tiles which are "close" and the set which are not. We then randomly select one of the tiles from the close set and replace the tile in question with the selected tile. We continue this process until the entire globe is resampled. We will then iterate this process until we have a stable resampled globe.
This will constitute a resampled bootstrap sphere.

Obviously, many things need to be addressed: which boundary should we use, how shall we define "distance", how shall we define "close", and how long should we iterate. A "distance" measure must be used in order to judge the "closeness" of tile boundaries. The "distance" is a measure of numerical difference, while the "closeness" is a rank-ordering corresponding to \( p \), as in "select a tile from its \( p \) nearest neighbors". The next four sections describe these concepts in greater detail.

### 5.4.1 Inner and Outer Linking Algorithms

For both the inner and outer boundary links, the class of prospective resampled tiles is comprised of the original tiles from a given wafer and its symmetric counterpart wafer. For each tile that is to be resampled, there are \( 2n_{w,r} \) tiles in the resample class, except for the equator where there are \( n_{E,r} \) tiles. We use the original tiles since they retain the exact dependence structure of the globe. If we were to use a resampled sphere to create our resampling pool, then we would not retain the true local dependence structure across two resampled tiles.

The inner link guarantees a better match within the tile, but has no mechanism for retaining the dependence structure across tiles. The inner link compares the observations on the inner boundary of a given tile with the observations on the inner boundary from each tile in the resample class. This method guarantees a better match within each tile, but since no
information about neighboring tiles and wafers is used, there is little hope that there will be a smooth transition among the surrounding wafers and tiles.

The outer link, on the other hand, provides for a smoother transition in retaining the dependence structure across tiles and wafers, while still maintaining the dependence within the original tiles. As mentioned previously, the outer boundary for each tile is created from the wafer strips directly above and below the tile and also from the tiles strips of the horizontally neighboring tiles. This method guarantees a dependence between wafers and tiles. Also, since the link is on the outer boundaries, the resampled tile boundary should be similar to the original tile boundary and hence the tile contents and its across-tile dependence should be similar.

5.4.2 Distance Scaling

"Distance scaling" places varied emphasis on the elementwise differences in the boundaries. For instance, if we go from water (0) to water (0), or land (1) to land (1), then no error is made and a scale of 0 is used. If we go from water (0) to land (1), we have made one type of error, given an arbitrary scale of 1. If we go from land (1) to water (0), we have made a different type of error; denote its scale by \( v > 0 \). Notice that if \( v = 1 \) then we have "no scale" or an "equal scale"; this is the method that was used in the example above. The scaling is applied elementwise to the boundary of each tile to be resampled, relative to the boundaries of the tiles in the prospective resample class for each tile and wafer; the sum of all of these scaled boundary difference values is regarded as the distance between the two tiles.

For the example given previously, the tile to be replaced has a boundary of (0, 0, 0, 1, 1, 0, 0, 0) and it is compared with a prospective tile from the resample class with a boundary of (0, 0, 0, 1, 1, 0, 0, 0). The distance between these two tiles is \( 1 + 2v \). Hence with no scaling (i.e., \( v = 1 \)) this distance is 3.

5.4.3 Closeness – Nearest Neighbor Systems

"Value", "number", and "percentage" are the three basic ways to describe nearest neighbors or closeness. All depend on the type of distance measure that is chosen. Assume that we have decided on a distance measure and have calculated the set of distances between a given tile and its prospective resample class. We are interested in finding the \( p \) nearest neighbors to this tile.

The \( p \) nearest neighbors value method can be thought of in two ways: 1) consider only the tiles in the prospective resample class that have a distance less than or equal to the value of \( p \)
or 2) starting from the minimum actual distance of the set, \( q \), consider only the tiles in the resample class that have a distance less than or equal to \( q + p \). The first method is just a special case of the second with \( q = 0 \). Notice that the first way can fail, if there are no tiles that have a value less than or equal to \( p \). For this reason, the \( p \) nearest neighbor value method is defined to be the second option.

The \( p \) nearest neighbors number method considers only the nearest \( p \) tiles in distance among the prospective resample class. If there are ties, then we take all of the values at this distance. For example, suppose that \( p = 10 \) and that for a given tile we find that the 10\(^{th}\) closest tile in the resample class has a distance of 3. If there are 15 tiles with this distance, then we consider all the tiles in the resample class that have a distance less than or equal to 3. There are usually ties, so this method can inflate the number of nearest neighbors beyond the nominal \( p \). This type of method is not consistent across wafers since \( p = 50 \) is more than 25% (50/180) of tiles in the resample class for wafer 1, but is 50/720 of the tiles in the resample class for wafers near the equator.

The \( p \) nearest neighbors percentage method accounts for the number of tiles in the resample class and alleviates the previous problem, creating an equal percentage of tiles from the resample class from which to choose. We consider \( p \) percent of the tiles in the resample class, and then use the number method as given above. For example, if we are on a wafer where the number of tiles in the resample class is 480 (240 x 2) and \( p = 10 \), then we take at least 48 (.10 x 480) nearest neighbors. If we are on a wafer with 720 (360 x 2) tiles in the resample class, then we take at least 72 (.10 x 720) nearest neighbors.

### 5.4.4 Iteration and the Initial Random Field

Iteration and the initial random field are important concepts in the resampling process. Given a starting random field, the more times we iterate using a resampling scheme, the further we go from the starting random field. If the dependence structure, along with the linking, is handled properly, then as we iterate more, we should converge to the correct limiting distribution.

The question of when to stop iterating is an interesting one. When does the iteration process not yield a substantial improvement? We have no rigorous answer to this question; rather we make an ad hoc judgement by looking at the iterations of the generated random fields.

Notice that if our initial random field for resampling is the original one which we sampled, then as we iterate, we go further from it. Nevertheless, as we repeat this resampling process to create \( B \) bootstrap realizations, we are always starting from the same position, possibly yielding correlated realizations. For this reason, it is desirable to start with a different initial random field for each bootstrap realization and iterate from there.

64
One place to start is from a stratified-tilewise bootstrap algorithm sphere (e.g., Figure 5.2 with tile size = 5). If the dependence structure on the random field is such that the tiles that are separated by a tile are independent, then the stratified-tilewise bootstrap algorithm initial starting field is plausible. That is, start with a random field that is resampled via the stratified-tilewise bootstrap algorithm and then iterate away from it using the dependence structure of the original data. This will give rise to completely different starting random fields which, nevertheless, have the correct dependence structure after the iteration process has run.

5.4.5 Stratified - Tilewise Outer Link Bootstrap Algorithm

There are many things that we must *simultaneously* investigate:

1. Which linking mechanism (inner or outer) is better in keeping the dependence structure of the sphere intact?
2. What is a good choice of tile size?
3. What nearest neighbor system should be used?
4. Which distance scaling value should be used?
5. How long should we iterate?
6. Where should we start from?

As one would suspect, question 1 can be answered quite quickly and efficiently. The inner link mechanism does not preserve the dependence structure of the sphere across wafers and tiles, whereas the outer linking mechanism does. The phenomenon continues even if we iterate the process or if we introduce the scaling mechanism. Figure 5.4 shows the inner link applied to the global data for various tile sizes with the View 3 from the original data in the upper left hand corner. (For all of them we used a percentage method with $p = 40$ and iteration of 3.)

Notice that no matter what the tile size is, one is able to determine North America and South America. The “new” realizations are not really new, but slight modifications of the original. Further note that the global dependence structure is not preserved, since the continents are breaking up into islands and new continents are not forming.

It is important to note that the value method does not work in this case, since the distribution of the distances is bi-modal at the endpoints of 0 and 4, 16, and 36 for tiles of size 2, 5, and 10 respectively. This is because a majority of the tiles are either all land or all water, reflecting the fact that there is little coastline relative to land and water.

65
Figure 5.4: Inner Linking Mechanism
From now on, we only consider the outer link bootstrap algorithm. Since the outer link is affected by the other wafers and tiles, we must be careful as to the order in which we resample. Following the lead of Hastings [Has70], we will employ a version of the Metropolis method [MRM+53]. Hasting's modification uses the conditional distribution of a random variable with respect to its neighbors (similar to the Gibbs sampler).

Besag [Bes74] modified this algorithm so that we could use the information from the original data. This method uses the checkerboard layout depicted in Figure 5.5 where one observation is located in each white and black square.

![Lattice and Sphere](image)

**Figure 5.5: Nonparametric Metropolis Method**

Besag used the phrase "auto-model" to describe a random field in which the conditional distribution of each observation only depends on its neighbors. Given a 1-step auto-model and the random field observed on a lattice, conditional on the black squares, the white squares are independent. Also, conditional on the white squares, the black squares are independent.

Lele [Le88] combined the checkerboard idea and the Metropolis method to create a nonparametric Metropolis method. Unfortunately, we cannot apply Lele's method directly because, conditional on the "black tiles", the "white tiles" are not independent since on neighboring wafers the tiles have different lengths. In particular, on sections of the sphere where the horizontal wafer angle, $\theta_w(r)$, changes between wafers, the "black tiles" and "white tiles" begin to overlap from wafer to wafer as is depicted in Figure 5.5. Nevertheless, this is still a reasonable approach since on a majority of the sphere (near the equator), the tiles will not overlap ($\theta_w(r)$ is identical across wafers in this region) and we get a true checkerboard effect.
When we try this approach, beginning with the original sphere and iterating away from it, we see in Figure 5.6 that the sphere is starting to break up and loses land mass. Further note that we are still able to determine North America, so that we are not going far away from the initial random field. The reason that the land disintegrates is possibly due to the fact that we are using a 1-step outer link. Since the dependence structure seems larger than one-step (as was indicated by the estimates of variance of the standardized sample mean), we need to weigh the tiles in the resample class accordingly; this is where the scaling comes in. If we penalize the resampling of water tiles to replace land tiles, then we may retain a balance of land and water in the correct proportion.

Question 2 can be answered by considering Figure 5.6. Here we notice that when the tiles are of size 10, the resampled globe is "too blocky", whereas when the tiles are size 2, then the resampled globe fails to retain the original local dependence structure and we get many small islands. For these reasons, we will only consider tiles of size 5.

All of the other questions are tied together. They relate to the following two phenomena, that we must simultaneously keep under control: preserving the global dependence structure of the Earth and maintaining the correct proportion of land mass.

We hypothesize that the Earth started with one, maybe two, large land masses, and that over time, the land broke up into continents as we know them (see [OMA93] for an account). If this process continues, we may come to a point when the Earth's land is simply comprised of little islands. If we can work backwards from islands to continents, using the original dependence structure, we may be able to reconstruct "new" Earths.

The stratified-tilewise bootstrap algorithm creates a resampled globe with the correct land mass, but the long-range dependence is wrong. Incorporating a "pseudo-parametric" linked-bootstrap algorithm that fixes the land mass proportion around the true sample land mass while using the original global dependence structure should address both of these phenomena. In this setting, the resampled globes look completely different than the original globe, yet retain the correct dependence structure.

So, the idea is to start from a stratified-tilewise bootstrap resampled Earth (e.g., Figure 5.2 for Tile size = 5) and then iterate using the linked-bootstrap until the Earth has coagulated into continents. This requires finding an appropriate value of \( v \) to preserve the land mass proportion and also an idea of a continent. The former is done by trial and error and the latter is done by eye. After various simulations, we conclude that a \( p \) nearest neighbor value method with \( p = 3 \) and a scale of \( v = 1.20 \) is the appropriate combination. Also, we found that if we iterate more than 10 times, we gain no more continental information. We call this method the stratified-tilewise outer link bootstrap algorithm. Three realizations of this type of bootstrap algorithm applied to the Earth are illustrated in Figures 5.7 – 5.9. As you can see, all three
Figure 5.6: Outer Linking Mechanism
Figure 5.7: Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 1
Figure 5.8: Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 2
Figure 5.9: Stratified - Tilewise Bootstrap Algorithm with Outer Link - Realization 3
<table>
<thead>
<tr>
<th></th>
<th>Coast1</th>
<th>Coast2</th>
<th>$\bar{X}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0573</td>
<td>0.20997</td>
<td>0.268617</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0076</td>
<td>0.02128</td>
<td>0.051340</td>
</tr>
<tr>
<td>Median</td>
<td>0.0572</td>
<td>0.2079</td>
<td>0.2654</td>
</tr>
<tr>
<td>Target</td>
<td>0.0549</td>
<td>0.1894</td>
<td>0.2895</td>
</tr>
</tbody>
</table>

Table 5.3: Univariate Information for Bootstrap Statistics

bootstrap realizations preserve the long-range dependence of the original globe. We have "continents", "islands", "oceans", and "bays". There is quite a bit of variability in the realizations, but all have roughly the same proportion of land. Since these realizations do have continents, we are able to calculate statistics that pertain to the long-range dependence of the globe.

### 5.5 Applications to Earth's Land-Area and Coastline Data

The ability to create new Earths that have the correct dependence structure, yet look completely different, is necessary if we are to investigate the sampling distribution of statistics that depend upon shapes or global structure. If we tried to use the stratified-tilewise bootstrap algorithm with this type of statistic, we would not get valid results. For example, if the statistic of interest were the coastline over the globe, then the stratified-tilewise bootstrap algorithm would give wildly incorrect results since we would get many islands and hence much too much coastline (see Figure 5.2).

We classify a point as coastline if it is land and any one of its 4 neighbors is water. Using the original data, we find that approximately 5.5% ($Coast_1 = \#land points/|\mathcal{P}|$) of Earth is coastline (of course this estimate depends on the grid resolution). Another way to describe the coast is by saying what the coastline is relative to the total land area. In this case, we find the coastline to be 18.9% ($Coast_2 = \#coast points/\#land points$). Notice that the sample mean is simply $\bar{X} = Coast_1/Coast_2$. For the coastline statistics, however, we have no way, except to use the bootstrap, to estimate their standard errors, nor their sampling distributions.

Using the stratified-tilewise outer link bootstrap algorithm, we can create new realizations of the Earth, calculating the coastline statistics for each. We then use the distribution of these bootstrap coastline statistics to model the true sampling distribution of the statistics. Table 5.3 gives the univariate information for the three statistics after observing $B = 100$ bootstrap realizations. The standard errors in the table are simply the standard errors from the bootstrap realizations. For $\bar{X}$, the bootstrap estimate of standard error under the
stratified-tilewise outer link bootstrap algorithm compared with the bootstrap estimate of standard error under the stratified-tilewise bootstrap algorithm (\(3.68573 / \sqrt{100} \approx 0.02\)) is close. This is quite remarkable since we have significantly changed the resampling mechanism.

As you can see, the averages of the statistics are near their target levels (as determined from the original data). With the mean and the median near the same value, we might expect to have symmetric distributions for all of the statistics.

We now turn our attention to the bootstrap distribution of the statistics and their empirical cumulative distribution functions (EDF). Figures 5.10 - 5.11 show that the three distributions are actually near normally distributed with Coast2 being the furthest from normality. This is justified by the fact that Coast2 is actually a ratio statistic of two random quantities.

As one would suspect, these statistics are highly correlated. Table 5.4 reveals that Coast1 and \(\bar{X}\) are positively correlated while Coast2 and \(\bar{X}\) are negatively correlated. This is because as there is more land there is more possibility of coastline, but less coast relative to total land. The fact that Coast1 and Coast2 are negatively correlated is from a combination of these factors. Figure 5.12 graphically illustrates these relationships.
Figure 5.10: Kernel Density Estimates and EDFs for Coastline Statistics
Figure 5.11: Kernel Density Estimate and EDF for $\bar{X}$
Figure 5.12: Bivariate Plots of Statistics
Appendix A

Proofs of Lemmas and Theorems

A.1 Sampling Design: Proofs of Lemmas and Theorems

Proof of Lemma 2.1

To facilitate the proof, recall the following facts:

1. If \( \{x_n : x_n \geq 1 \forall n\} \) and \( x_n \to \infty \) as \( n \to \infty \), then
   
   \begin{align*}
   & (a) \quad [x_n] \to \infty \text{ as } n \to \infty \text{ where } [\cdot] \text{ denotes integer part of, and } \\
   & (b) \quad x_n - [x_n] = x_*^n \in [0, 1) \Rightarrow 1 \leq \frac{x_n}{[x_n]} < 2 \text{ and } 0 < \frac{[x_n]}{x_n} \leq 1.
   \end{align*}

2. Taylor series expansion:

   Suppose \( h(\cdot) \) has a finite \( n^{th} \) derivative everywhere in \( (a, b) \) and \( h^{(n-1)}(\cdot) \) is continuous in \( [a, b] \). Let \( c \in [a, b] \) and then \( \forall x \in [a, b] \setminus \{c\}, \exists x^* = x^*(c, x) \) between \( c \) and \( x \) such that

   \[ h(x) = \sum_{k=0}^{n-1} \frac{h^{(k)}(c)}{k!}(x-c)^k + \frac{(x-c)^n}{n!}h^{(n)}(x^*). \]

   (a) For \( h(x) = \sin^2(x) \), when \( [a, b] = [0, \frac{\pi}{2}], c = 0, 0 < x < \frac{\pi}{2} \) and
       \( n = 4 \), \( \exists x^* = x^*(0, x) \in (0, x) \) such that

       \[ \frac{\sin^2(x)}{x^2} = 1 - \frac{x^2}{3} \cos(2x^*). \]

   (b) For \( h(x) = \sin(x) \), when \( [a, b] = [0, \frac{\pi}{2}], c = \gamma_w(r), x = \gamma_w(r) + y, y \geq 0, n = 1 \), and for some \( x^* = x^*(r, w, y) \in (\gamma_w(r), \gamma_w(r) + y) \)

       \[ \frac{\sin(\gamma_w(r) + y)}{\sin(\gamma_w(r))} = 1 + \frac{y}{\sin(\gamma_w(r))} \cos(x^*). \]
First note that from conditions (1), (3), (4), and (5), \( \forall w \in \mathcal{W}_r \),
\[
\frac{J_r}{r \sin(\gamma_w(r))} \to 0, \quad \frac{J_r}{r} \to 0, \quad \frac{I_r}{r \sin(\gamma_w(r))} \to 0, \quad \text{and} \quad \frac{I_r}{r} \to 0.
\]

Since the stratified spherical sampling plan is equatorially reflective, we need only consider distances within wafers of the northern hemisphere. First consider, \( w \in \mathcal{W}_r \setminus \{E\} \), \( 0 \leq \nu' \leq i \leq n_{w,r} - 1 \), and \( 0 \leq j' \leq j < J_r \), such that \( 0 < |i - \nu'| < I_r \) and \( 0 < |j - j'| < J_r \). Letting
\[
A_{j,j'}(r, \phi(r)) = 1 - \left( r \phi(r) \frac{\sin \left( \frac{(i-j')\phi(r)}{2} \right)}{\sin \left( \frac{(i-j)\phi(r)}{2} \right)} \right)^2,
\]
and
\[
B_{(i,j), (i', j')}(r, \theta_w(r), \phi(r)) =
\]
\[
= 1 - \left( \frac{\sin[\gamma_w(r) + j\phi(r)]}{\sin(\gamma_w(r))} \right) \left( \frac{\sin[\gamma_w(r) + j'\phi(r)]}{\sin(\gamma_w(r))} \right) \left( r \sin[\gamma_w(r)] \theta_w(r) \frac{\sin \left( \frac{(i-i')\theta_w(r)}{2} \right)}{\sin \left( \frac{(i-i')\theta_w(r)}{2} \right)} \right)^2,
\]
we find that the difference in squared distance between two points of wafer \( w \), \( P_{i,j}^w(r) \) and \( P_{i', j'}^w(r) \), and their associated points in \( Z^2 \), is
\[
\left\| P_{i,j}^w(r) - P_{i', j'}^w(r) \right\|^2 - [(i - \nu')^2 + (j - j')^2] \leq
\]
\[
\leq (j - j')^2 A_{j,j'}(r, \phi(r)) + (i - \nu')^2 B_{(i,j), (i', j')}(r, \theta_w(r), \phi(r))
\]
\[
\leq J_r^2 A_{j,j'}(r, \phi(r)) + I_r^2 B_{(i,j), (i', j')}(r, \theta_w(r), \phi(r)).
\]

(A.1)

Notice that from (2.21), when \( j' = j \) then we need not consider \( A_{j,j'}(r, \phi(r)) \) and when \( i' = i \) we need not consider \( B_{(i,j), (i', j')}(r, \theta_w(r), \phi(r)) \). Now using condition (3) and the Taylor series expansion of \( \sin^2(\cdot) \), for some \( x^* = x^*(r, j, j') \in (0, (j - j')\phi(r)/2) \) and large enough \( r \),
\[
A_{j,j'}(r, \phi(r)) = 1 - (r \phi(r))^2 \left( 1 - \frac{\cos(2x^*)}{12} ((j - j') \phi(r))^2 \right) \leq 3 |1 - r \phi(r)| + \frac{\cos(2x^*)}{12} (r \phi(r))^2 (J_r \phi(r))^2 = O \left( \frac{J_r}{r} \right).
\]

Also, for some \( y_1^* = y_1^*(r, w, j) \in (\gamma_w(r), \gamma_w(r) + j\phi(r)) \),
\( y_2^* = y_2^*(r, w, j') \in (\gamma_w(r), \gamma_w(r) + j'\phi(r)) \), and \( y_3^* = y_3^*(r, w, i, i') \in (0, (i - \nu') \theta_w(r)/2) \) and large enough \( r \),
\[
B_{(i,j), (i', j')}(r, \theta_w(r), \phi(r)) = 1 - (r \sin[\gamma_w(r)] \theta_w(r))^2 \left( 1 + \cos(y_1^* - \frac{j\phi(r)}{\sin[\gamma_w(r)]}) \right) \left( 1 + \cos(y_2^* - \frac{j'\phi(r)}{\sin[\gamma_w(r)]}) \right) \left( 1 - \frac{\cos(2y_3^*)}{12} ((i - \nu') \theta_w(r))^2 \right).
\]
Since \(0 < \gamma_1(r) < \gamma_2(r) < \cdots < \gamma_E(r) < \frac{\pi}{2}\), then
\[
\infty > \frac{1}{\sin(\gamma_1(r))} > \frac{1}{\sin(\gamma_2(r))} \cdots > \frac{1}{\sin(\gamma_E(r))} > 1.
\]

Now, expanding and collecting terms and using conditions (1) and (3), for large enough \(r\), we find that for some constant \(c_1\),
\[
B_{(i,j),(i',j')}^w(r, \theta_w(r), \phi(r)) \leq
1 - \theta_w(r) r \sin[\gamma_w(r)] \cdot [1 + \theta_w(r) r \sin[\gamma_w(r)] + \frac{c_1 \theta_w(r) r \sin[\gamma_w(r)]}{r} \left( \frac{J_r}{r \sin(\gamma_1(r))} \right) \cdot \left[ 1 + \left( \frac{J_r}{r \sin(\gamma_1(r))} \right) (\theta_w(r) r \sin[\gamma_w(r)]^2) \right].
\]

From condition (4)
\[
\sup_{w \in \mathcal{W}_r} |\theta_w(r) r \sin[\gamma_w(r)]| \leq c_r \to 1
\]
as \(r \to \infty\), where \(c_r\) depends only on \(r\). Taking the supremum over \(w \in \mathcal{W}_r\), we have using conditions (1), (3), and (4),
\[
\sup_{w \in \mathcal{W}_r \setminus \{E\}} |B_{(i,j),(i',j')}^w(r, \theta_w(r), \phi(r))| = O \left( \frac{J_r}{r \sin(\gamma_1(r))} \right).
\]
Hence,
\[
\sup_{w \in \mathcal{W}_r \setminus \{E\}} \left| \|P_{(i,j)}^w(r) - P_{(i',j')}^w(r)\|^2 - [(i - i')^2 + (j - j')^2] \right| = O \left( \frac{J_r^3}{r \sin(\gamma_1(r))} \right).
\]
Therefore using condition (5), the spherical distances converge to their planar counterpart distances uniformly within a local region of "height" \(J_r \phi(r)\) and "length" \(J_r \theta_w(r)\) over all wafers on the entire sphere. The equatorial region is similarly handled, noting that the bound in (A.1) is now
\[
\left( \frac{v_r}{J_r} \right)^2 J_r^2 A_{j',j}(r, \phi(r)) + L_r^2 B_{(i,j),(i',j')}^w(r, \theta_w(r), \phi(r)).
\]
Since \(v_r = o(J_r)\), for large enough \(r\), \([v_r J_r^{-1}] < 1\). Apply the rest of the proof. \(\blacksquare\)

### A.2 Asymptotic Normality: Proofs of Lemmas and Theorems

Throughout the remainder of the proofs, arbitrary constants are assigned symbols \(c_i\) where \(i \geq 1\). The constants are renumbered at the beginning of each proof. If the constants depend
upon some parameter then the parameter is incorporated into the constant. For example if the constant depends upon parameter $\delta$, then the constants are numbered $c_i(\delta)$ where $i \geq 1$. Note that $c_0$ and $c_k$ have a unique meaning in the definition of smoothness of a covariance function.

**Proof of Lemma 3.1**

$$|E(\zeta \eta) - E(\zeta) E(\eta)| = |E(\zeta [E(\eta|\sigma(A_1)) - E(\eta)])|$$

$$\leq E(|\zeta [E(\eta|\sigma(A_1)) - E(\eta)])|$$

$$\leq c_1 E(|E(\eta|\sigma(A_1)) - E(\eta)|).$$

Define $\zeta_1 = \text{sign}(E(\eta|\sigma(A_1)) - E(\eta))$ and $\eta_1 = \text{sign}(E(\zeta_1|\sigma(A_2)) - E(\zeta_1))$. Then

$$|E(\zeta \eta) - E(\zeta) E(\eta)| \leq c_1 E(|\zeta_1 [E(\eta|\sigma(A_1)) - E(\eta)])|$$

$$\leq c_1 |E(\zeta_1 \eta) - E(\zeta_1) E(\eta)|$$

$$\leq c_1 c_2 |E(\zeta_1 \eta_1) - E(\zeta_1) E(\eta_1)|.$$

Define events $E = \{\omega : \zeta_1(\omega) = 1\} \in \sigma(A_1)$ and $F = \{\omega : \eta_1(\omega) = 1\} \in \sigma(A_2)$. Then

$$|E(\zeta_1 \eta_1) - E(\zeta_1) E(\eta_1)| \leq |P(EF) - P(E) P(F)| + |P(E^c F^c) - P(E^c) P(F^c)| +$$

$$+ |P(EP^c) - P(E) P(F^c)| + |P(E^c F) - P(E^c) P(F)|$$

$$\leq 4\alpha(k).$$

**Proof of Corollary 3.1**

Let $\zeta = \zeta_1 + i\zeta_2$ and $\eta = \eta_1 + i\eta_2$ where $\zeta_1, \zeta_2, \eta_1$, and $\eta_2$ are real-valued. Then

$$|E(\zeta \eta) - E(\zeta) E(\eta)| \leq |E(\zeta_1 \eta_1) - E(\zeta_1) E(\eta_1)| + |E(\zeta_1 \eta_2) - E(\zeta_1) E(\eta_2)| +$$

$$+ |E(\zeta_2 \eta_1) - E(\zeta_2) E(\eta_1)| + |E(\zeta_2 \eta_2) - E(\zeta_2) E(\eta_2)|.$$

Now apply Lemma 3.1 to each part. ■

**Proof of Lemma 3.2**

If $\alpha(k) = 0$, then the result is trivial since $\eta$ and $\zeta$ would be independent, so assume that $\alpha(k) > 0$. For $N > 0$, define the truncated random variables: $\eta_N, \zeta_N, \eta_N \zeta_N, \zeta_N \eta_N, N\zeta, N\zeta N\eta, N\zeta N\eta N\eta$. Then

$$|E(\zeta \eta) - E(\zeta) E(\eta)| \leq |E(N \zeta N \eta) - E(N \zeta) E(N \eta)| +$$

$$+ |E(N \zeta N \eta)| + |E(N \zeta N \eta)| + |E(N \zeta) E(N \eta)| +$$

$$+ |E(N \zeta) E(N \eta)| + |E(N \zeta) E(N \eta)|.$$

Applying Hölder's inequality, we obtain the following facts:
\[ |E \{N \zeta\}| \leq \min \left\{ \left( \frac{1}{N}\right)^\beta c_1, c_1^\beta \right\}, \quad |E \{N \eta\}| \leq \min \left\{ \left( \frac{1}{N}\right)^\beta c_2, c_2^\beta \right\}, \quad |E \{N\zeta\}| \leq c_1^\beta, \quad |E \{N\eta\}| \leq c_2^\beta, \]

\[ E \left\{ |N\zeta|^{1/(1-\beta)} \right\} \leq \left( \frac{1}{N}\right)^{\beta/(1-\beta)} c_1, \quad \text{and} \quad E \left\{ |N\eta|^{1/(1-\beta)} \right\} \leq \left( \frac{1}{N}\right)^{\beta/(1-\beta)} c_2. \]

Using the above facts, Lemma 3.1, Hölder's inequality, and noting that \(|N\zeta| < N\), we have
\[
|E \{\zeta\} - E \{\zeta\} E \{\eta\}| \leq 4N^2 \alpha(k) + 3 \left( \frac{1}{N}\right)^\beta \left( c_1^\beta c_2 + c_1^\beta c_2^\beta \right).
\]

Letting \(N = [\alpha(k)]^\beta\) gives the desired result. \(\square\)

**Proof of Lemma 3.3**

For \(w \in \mathcal{W}_r\) and let
\[ G_r = G(s_r, t_r, l_r, m_r) = \{(i, j) : s_r \leq i \leq s_r + l_r - 1 \text{ and } t_r \leq i \leq t_r + m_r - 1\} \]
be a set of indices and \(X_{(a,b)}\) be a random variable of the random field observed at \((a, b) \in \mathbb{Z}^2\). Then
\[
V \left\{ \sum_{(i,j) \in G_r} X_{(i,j)} \right\} = \sum_{(i,j) \in G_r} \sum_{(i',j') \in G_r} E \{X_{(i,j)}X_{(i',j')}\} + \sum_{(i,j) \in G_r} \sum_{(i',j') \in G_r} \left[ E \{X_{P_{i,j}^r(r)}X_{P_{i',j'}^r(r)}\} - E \{X_{(i,j)}X_{(i',j')}\} \right].
\]

So then since \(\{X_{(i,j)}\}\) is a mean-zero random process and \(w_{s_r, t_r, l_r, m_r} = |G_r| = l_r m_r\),
\[
A_r = l_r m_r \left( 1 + \sum_{|i| < l_r, |j| < m_r} \left( 1 - \frac{|i|}{l_r} \right) \left( 1 - \frac{|j|}{m_r} \right) E \{X_{(i,j)}X_{(i,j)}\} \right).
\]

But by Lemma 3.1,
\[
|C_r| \leq \sum_{|i| < l_r, |j| < m_r} |E \{X_{(i,j)}X_{(i,j)}\}| \leq c_1^2 \sum_{|i| < l_r, |j| < m_r} \alpha(\sqrt{i^2 + j^2}). \tag{A.2}
\]

Now since \(\alpha(j) \downarrow 0\), if \(j \leq k\) then \(\alpha(k) \leq \alpha(j)\). For each \(k \geq 1\), we would like to find the number of pairs of \((i,j)\) that are contained in the annulus with inner radius \(k\) and outer radius \(k + 1\) (i.e., \(k \leq \sqrt{i^2 + j^2} < k + 1\)). We need only consider the first quadrant of \(\mathbb{Z}^2\), and then multiply by 4 to find the total number of pairs.
Since each grid square is $1 \times 1$, then if we create an annulus with outer radius $k + 1 + \sqrt{2}$ and inner radius of $k - \sqrt{2}$, then this annulus covers the annulus in question, including any grid square that touches it. Since the area of this annulus in the first quadrant is $\frac{1}{2}(1 + 2\sqrt{2})(2k + 1)$, then the most grid squares that can fit into it is less than $10k$. So with four vertices to each grid square, there are less than $40k$ vertices in the first quadrant or less than $160k$ vertices on the lattice that are contained in the annulus in question. Therefore by condition (3),

$$|C_r| \leq c_2 \sum_{j=1}^{\max(l_r, m_r)} j\alpha(j) = O \left( \sum_{j=1}^{\infty} j\alpha(j) \right) < \infty. \quad (A.3)$$

Hence, by the dominated convergence theorem,

$$\lim_{r \to \infty} C_r = \sum_{\vec{r} \in \mathbb{Z}^2 \setminus \vec{0}} C \{X_{\vec{r}}, X_{\vec{r}}\},$$

thus making $\lim_{r \to \infty} A_r / l_r m_r = \sigma^2$. In addition,

$$|B_r^w| \leq \sum_{(i,j) \neq (i',j') \in \mathcal{G}_r} \left| B_0 \left\{ \| P_{i,j}^w(r) - P_{i',j'}^w(r) \| \right\} - B_0 \left\{ \sqrt{(i-i')^2 + (j-j')^2} \right\} \right|. \quad (A.4)$$

By Lemma 2.1 and the definition of $V$, we have, uniformly for $w \in \mathcal{W}_r$, $\forall (i,j) \neq (i',j') \in \mathcal{G}_r$ such that $v = \sqrt{(i-i')^2 + (j-j')^2} \in V$ and as $r \to \infty$,

$$\| P_{i,j}^w(r) - P_{i',j'}^w(r) \| \to \sqrt{(i-i')^2 + (j-j')^2}.$$

So then for large enough $r$, uniformly for $w \in \mathcal{W}_r$, $\| P_{i,j}^w(r) - P_{i',j'}^w(r) \| \in B(v, c_0)$. For this $r$ and using condition (4), the smoothness of the covariance function, and Lemma 2.1,

$$\sup_{w \in \mathcal{W}_r} |B_r^w| \leq \sup_{w \in \mathcal{W}_r} \sum_{(i,j) \neq (i',j') \in \mathcal{G}_r} c_0 \left\| P_{i,j}^w(r) - P_{i',j'}^w(r) \right\|^2 - [(i-i')^2 + (j-j')^2]$$

$$\quad = o(l_r m_r).$$

So then putting it all together gives

$$\sup_{w \in \mathcal{W}_r} \left| V \left\{ \sum_{\vec{r} \in \mathcal{G}_r} X_{\vec{r}} \right\} / l_r m_r \right| - \sigma^2 \to 0 \quad \text{as} \quad r \to \infty.$$

Notice that there is no special case for when $w = E$, since we only invoke Lemma 2.1 which is uniform $\forall w \in \mathcal{W}_r$ and the definition of $m_r$ changes in the equatorial region (see conditions (1') and (2')).

Proof of Corollary 3.2

The boundedness of the $X_{\vec{r}}$'s was only used in (A.2) when we applied Lemma 3.1. Now, since the $X_{\vec{r}}$'s could be unbounded, apply Lemma 3.2 and hence the summability in (A.3) should be
of \( j\alpha(j)^{l/(2+4)} \). Apply condition (3) of the corollary and the rest of the proof remains identical. \( \blacksquare \)

**Proof of Lemma 3.4**

Let \( \Phi_m(t) = E \left\{ \exp(itY^n_{m,1}) \right\} \) be the characteristic function of \( Y^n_{m,1} \) for \( 1 \leq m \leq M_n \). Then by Corollary 3.1, we have

\[
L_{M_n} = \left| E \left\{ \exp \left( \sum_{m=1}^{M_n} \sum_{k=1}^{K_m(n)} itY^n_{m,k} \right) - \prod_{m=1}^{M_n} [\Phi_m^n(t)]^{K_m(n)} \right\} \right|
\leq \left| E \left\{ \exp \left( \sum_{m=1}^{M_n-1} \sum_{k=1}^{K_m(n)} itY^n_{m,k} + \sum_{k=1}^{K_{M_n}(n)-1} itY^n_{M_n,k} \right) \right\} \exp \left\{ itY^n_{M_n,K_{M_n}(n)} \right\} \right| +
- E \left\{ \exp \left( \sum_{m=1}^{M_n-1} \sum_{k=1}^{K_m(n)} itY^n_{m,k} + \sum_{k=1}^{K_{M_n}(n)-1} itY^n_{M_n,k} \right) \right\} \Phi_m^n(t) \right| +
+ |\Phi_m^n(t)| \left| E \left\{ \exp \left( \sum_{m=1}^{M_n-1} \sum_{k=1}^{K_m(n)} itY^n_{m,k} + \sum_{k=1}^{K_{M_n}(n)-1} itY^n_{M_n,k} \right) \right\} \right|
- \left[ \Phi_m^n(t) \right]^{K_{M_n}(n)-1} \prod_{m=1}^{M_n-1} [\Phi_m^n(t)]^{K_m(n)}
\leq 16\alpha(d_{M_n,K_{M_n}(n)}^n) + E \left\{ \exp \left( \sum_{m=1}^{M_n-1} \sum_{k=1}^{K_m(n)} itY^n_{m,k} + \sum_{k=1}^{K_{M_n}(n)-1} itY^n_{M_n,k} \right) \right\} +
- \left[ \Phi_m^n(t) \right]^{K_{M_n}(n)-1} \prod_{m=1}^{M_n-1} [\Phi_m^n(t)]^{K_m(n)}
\leq \vdots
\leq 16 \sum_{k=1}^{K_{M_n}(n)} \alpha(d_{M_n,k}^n) + E \left\{ \exp \left( \sum_{m=1}^{M_n-1} \sum_{k=1}^{K_m(n)} itY^n_{m,k} \right) \right\} - \prod_{m=1}^{M_n-1} [\Phi_m^n(t)]^{K_m(n)}
\leq \vdots
\leq 16 \sum_{m=1}^{M_n} \sum_{k=1}^{K_m(n)} \alpha(d_{m,k}^n) \rightarrow 0. \blacksquare

**Proof of Lemma 3.5**

Given a nearly uniform sampling plan \( \mathcal{P}_r \), calculate \( \gamma_w(r) \), \( W_r, \eta_w, \kappa_w, \iota_r, \iota T_w, k(r), \iota T_{w+1-w,k}(r) \), and \( \iota T(r) \) as in Algorithms 1–2, and the definitions in section 3.4.2. We must first determine \( \iota d_w(r) \), the minimal distance between the type \( l \) tiles of wafer \( w \) with the type \( l \).
tiles over the entire sphere. That is,

\[ l'd_w(r) = \min_{1 \leq k \leq \ell_{w,r}} d(lT_{w,k}(r), T(r)) \theta_{w}(r) \]

\[ = 2r \min \left\{ \sin [\gamma_w(r)] \sin \left( \frac{l's_h(r) \theta_w(r)}{2} \right), \sin \left( \frac{l's_v(r) \phi(r)}{2} \right) \right\} \]

where \( l's_h(r) \) and \( l's_v(r) \) are the number of separation angles between type \( l \) tiles in the horizontal and vertical directions, respectively, which are independent of \( w \). These are given in table A.1. Using the following facts,

<table>
<thead>
<tr>
<th></th>
<th>( l = 1 )</th>
<th>( l = 2 )</th>
<th>( l = 3 )</th>
<th>( l = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( l's_h(r) )</td>
<td>( I_r - v_r )</td>
<td>( I_r - v_r )</td>
<td>( v_r )</td>
<td>( v_r )</td>
</tr>
<tr>
<td>( l's_v(r) )</td>
<td>( v_r )</td>
<td>( J_r - v_r )</td>
<td>( J_r - v_r )</td>
<td>( v_r )</td>
</tr>
</tbody>
</table>

Table A.1: Horizontal and Vertical Separation Angles

1. as \( z \to 0, \frac{1}{2} \leq \frac{\sin(z)}{z} \leq 1 \),
2. \( \forall z \sin(z) \leq z \), and
3. if \( a \leq b \) and \( c \leq d \), then \( \min(a, c) \leq \min(b, d) \),

and applying the definition of a nearly uniform sampling plan, for large \( r \),

\[ l'd_w(r) \geq 2r \min \left\{ \sin [\gamma_w(r)] \sin \left( \frac{v_r \theta_w(r)}{2} \right), \sin \left( \frac{v_r \phi(r)}{2} \right) \right\} \]

\[ \geq \frac{v_r}{2} \min \{ r \sin [\gamma_w(r)] \theta_w(r), r \phi(r) \} \geq \frac{v_r}{4} = d(r). \]

Since \( \alpha(\cdot) \) is monotonic decreasing then \( l'd_w(r) \geq l'd(r) \Rightarrow \alpha(l'd_w(r)) \leq \alpha(l'd(r)) \). Also, if there exists an integer \( m > 1 \) such that \( \sum_{j=1}^{\infty} j^m \alpha(j) < \infty \), then

\[ \alpha(n) \leq \left( \frac{2}{n} \right)^{m+1} \sum_{j=n/2}^{n} j^m \alpha(j) = o \left( \left( \frac{1}{n} \right)^{m+1} \right). \quad (A.4) \]

Hence, by Lemma 3.4, if we let \( Y_{m,r} = t (l'T_{w,k}(r)) / \sqrt{n_r}, \) then for \( t (l'T_{w,k}(r)) / \sqrt{n_r} = t (l'T_{w,k}(r)) / \sqrt{n_r} \) where \( \{ t (l'T_{w,k}(r)) / \sqrt{n_r} \} \) are independent, we can study \( \{ t (l'T_{w,k}(r)) / \sqrt{n_r} \} \) if \( \exists m > 1 \) such that

\[ \sum_{w=1}^{W_r} k_{w,l} \alpha(l'd_w(r)) \leq \left( \frac{1}{v_r} \right)^{m+1} \left( \sum_{w=1}^{W_r} k_{w,l} \right) o(1) \to 0. \]

85
So then we must find $m$ such that $v_r^{-m+1} \sum k_{w,i_r} = \mathcal{O}(1)$. Using the definitions of $k_{w,i_r}$, $n_{w,r}, \theta_w(r), W_r,$ and $\mathcal{W}_r,$ condition (1), and the facts that

1. for any $x$, $x - 1 \leq [x] \leq x$ and
2. for every $0 < z \leq \pi/2$, $2/\pi \leq \sin(x)/x \leq 1$,

we have for large enough $r$, $\sum_{w=1}^{W_r} k_{w,i_r} \asymp W_r^2$ since

$$\sum_{w=1}^{W_r} k_{w,i_r} \leq \frac{4r\pi}{I_r} \sum_{w=1}^{W_r} \sin[\gamma_w(r)] \leq c_1 W_r^2$$

and

$$\sum_{w=1}^{W_r} k_{w,i_r} \geq \frac{4r\pi}{3I_r} \sum_{w=1}^{W_r} \sin[\gamma_w(r)] - 2W_r \geq c_2 W_r^2.$$ 

Therefore, we have

$$\left(\frac{1}{v_r}\right)^{m+1} \sum_{w=1}^{W_r} k_{w,i_r} = \mathcal{O}(1).$$

Further note that since for large enough $r$, $|C_N(r)| \leq c_0(\beta_1(r)/\phi(r))^2$ and for every $w \in \mathcal{W}_r$, $k_{w,i_r}I_r \leq n_{w,r} \leq (k_{w,i_r} + 1)I_r$, then $n_r \asymp r^2$. □

**Proof of Lemma 3.6**

Given a nearly uniform sampling plan $P$, calculate $\gamma_w(r), W_r, n_{w,r}, k_{w,i_r}, 'T_{w,k}(r), \ 'T(r), R^*(r),$ and $E(r)$ as in Algorithms 1 - 2, (3.6), (3.7), (3.8), and (3.9). Notice that with $t(\cdot)$ being the summation, then for any 2 sets $A$ and $B$, $t(A \cup B) = t(A) + t(B)$. Then

$$t \left( \frac{R \setminus T(r)}{\sqrt{n_r}} \right) = \left[ \sum_{i=1}^{3} t \left( 'T(r) \right) + \sum_{j=3}^{N} t \left( C_j(r) \right) + t \left( R^*(r) \right) + t \left( E(r) \right) \right] / \sqrt{n_r}. \quad (A.5)$$

We only need to show that each term on the RHS of (A.5) converges in probability to 0. We begin with the corridor tiles, $t \left( 'T(r) \right)$.

From the given conditions and Lemma 3.5, it is enough to study $\tilde{t} \left( 'T_{w,k}(r) \right) / \sqrt{n_r}$ where $\tilde{t} \left( 'T_{w,k}(r) \right) / \sqrt{n_r} \overset{p}{\rightarrow} t \left( 'T_{w,k}(r) \right) / \sqrt{n_r}$ and the $\{\tilde{t} \left( 'T_{w,k}(r) \right) / \sqrt{n_r}\}$ are independent. It is sufficient to show that $\forall \epsilon > 0$,

$$\lim_{r \to \infty} P \left\{ \left| \tilde{t} \left( 'T(r) \right) \right| \geq \epsilon \sigma \sqrt{n_r} \right\} = 0.$$

Using Chebyshev’s inequality and Lemma 3.3,

$$P \left\{ \left| \tilde{t} \left( 'T(r) \right) \right| \geq \epsilon \sigma \sqrt{n_r} \right\} \leq \frac{\mathcal{V} \left\{ \tilde{t} \left( 'T(r) \right) \right\}}{\epsilon^2 \sigma^2 n_r} \leq \frac{\mathcal{V} \left\{ T_{1,1}(r) \right\}}{\epsilon^2 \sigma^2 n_r} \sum_{w=1}^{W_r} k_{w,i_r}.$$

86
Using \( n_r \geq |T_{1,1}(r)| \sum k_{w,l_r} \) and using Table A.2, then for \( 1 \leq l \leq 3 \),

\[
P \left\{ \left| \hat{T}(r) \right| \geq \varepsilon \sigma \sqrt{n_r} \right\} \leq \mathcal{O}(1) \frac{|T_{1,1}(r)|}{|T_{1,1}(r)|} \to 0.
\]

Before we proceed to the other parts, note that from their definitions, we have

\[
|C_N(r)| = \mathcal{O}\left((\gamma_1(r))^2\right), \quad |T(r)| = \mathcal{O}(I_r V_r \sum k_{w,l_r}),
\]

\[
|E(r)| = \mathcal{O}(v_r n_{E,r}), \text{ and} \quad |R^*(r)| = \mathcal{O}(I_r J_r W_r).
\]

We have for \( 1 \leq l \leq 3 \), since \( I_r J_r \sum k_{w,l_r} \approx n_r \approx r^2 \) from Lemma 3.5,

\[
\frac{|C_N(r)|}{I_r J_r \sum k_{w,l_r}} = \mathcal{O}\left(\frac{r^2}{r}\right) \to 0, \quad \frac{|T(r)|}{I_r J_r \sum k_{w,l_r}} = \mathcal{O}\left(\frac{v_r}{I_r}\right) \to 0,
\]

\[
\frac{|R^*(r)|}{I_r J_r \sum k_{w,l_r}} = \mathcal{O}\left(\frac{J_r}{r}\right) \to 0, \quad \frac{|E(r)|}{I_r J_r \sum k_{w,l_r}} = \mathcal{O}\left(\frac{v_r}{I_r}\right) \to 0.
\]

We continue on with \( R^*(r) \) and \( E(r) \). Note that from (3.8) and (3.9), \( t(R^*(r)) + t(E(r)) = t(R(r)) + t(E^*(r)) \). We consider \( R(r) \) and \( E^*(r) \). Here, we also break up into smaller regions and use Lemma 3.5 to approximate the regional tile sums by independent planar block sums.

Define for \( u \in W_r \setminus \{E\} \) and \( 1 \leq k \leq k_{w,l_r} \),

\[
1_{R_u}(r) = \sum_{k_{w,l_r}}^n, \quad 1_{R_{w,+1-u}}(r) = \sum_{k_{w,l_r}}^n
\]

\[
2_{R_u}(r) = \sum_{k_{w,l_r}}^n, \quad 2_{R_{w,+1-u}}(r) = \sum_{k_{w,l_r}}^n
\]

\[
1_{E_k}(r) = \sum_{(k-1)I_r,0}^{k_{w,l_r}} \quad 2_{E_k}(r) = \sum_{(k-1)I_r,0}^{k_{w,l_r}}
\]

Further, define the remaining and equatorial collection of type \( l \) tiles, for \( 1 \leq l \leq 2 \), as

\[
1_{R}(r) = \bigcup_{u=1}^{W_r/2} (1_{R_u}(r) \cup 1_{R_{w,+1-u}}(r)) \cup 1_{R_E}(r) \quad \text{and} \quad 1_{E^*}(r) = \bigcup_{k=1}^{k_{w,l_r}} 1_{E_k}(r).
\]

So we have for \( u \neq u' \in W_r \setminus \{E\} \) and \( 1 \leq k \neq k' \leq k_{w,l_r} \),

87
\[
\begin{align*}
\quad d^{(1)} R_{w}(r), & \quad \geq 2r \sin \left( \frac{v_r \phi (c)}{2} \right) \geq \frac{v_r}{2} \\
\quad d^{(1)} R_{w}(r), & \quad R_{E}(r) \quad \geq 2r \sin \left( \frac{v_r \phi (c)}{2} \right) \geq \frac{v_r}{2} \\
\quad d^{(2)} R_{w}(r), & \quad R_{w}(r) \quad \geq 2r \sin \left( \frac{v_r \phi (c)}{2} \right) \geq \frac{v_r}{2} \\
\quad d^{(1)} E_{k}(r), & \quad R_{E}(r) \quad \geq 2r \sin[\gamma_{E}(r)] \sin \left( \frac{v_r \phi (c)}{2} \right) \geq \frac{v_r}{2}, \text{ and} \\
\quad d^{(2)} E_{k}(r), & \quad E_{k}(r) \quad \geq 2r \sin[\gamma_{E}(r)] \sin \left( \frac{(L_{r} - v_{x}) \phi (c)}{2} \right) \geq \frac{L_{r} - v_{x}}{2} \geq \frac{v_r}{2}.
\end{align*}
\]

And hence, applying Lemma 3.4 to the remainder and equatorial type \( l \) tiles, we can study the independent versions of the planar block sums on the remainder and equatorial regions, \( \bar{t}^{(1)} R_{w}(r) \) and \( \bar{t}^{(1)} E_{k}(r) \), respectively. First note that from Lemma 3.5,

\[
\sum_{w=1}^{W_{r}} k_{w,I_{r}} \geq W_{r}^2 \quad \text{and} \quad k_{E,I_{r}} = \mathcal{O} \left( \frac{r}{L_{r}} \right).
\]

Using Lemma 3.3, we have for \( l = 1, 2 \)

\[
P \left\{ \left| t^{(1)} R_{r}(r) \right| \geq \varepsilon \sigma \sqrt{n_{r}} \right\} \leq \frac{V \left\{ \bar{t}^{(1)} R_{r}(r) \right\}}{\varepsilon^2 \sigma^2 n_{r}} \leq \frac{(W_{r} + 1)L_{r} \varepsilon n_{r}(1 + \Theta(1))}{\sigma^2 \varepsilon^2 n_{r}} \to 0
\]

and

\[
P \left\{ \left| t^{(1)} E_{r}(r) \right| \geq \varepsilon \sigma \sqrt{n_{r}} \right\} \leq \frac{V \left\{ \bar{t}^{(1)} E_{r}(r) \right\}}{\varepsilon^2 \sigma^2 n_{r}} \leq \frac{k_{E,I_{r}} \varepsilon \varepsilon n_{r}(1 + \Theta(1))}{\sigma^2 \varepsilon^2 n_{r}} \to 0.
\]

Using the hexagonal sampling plan on the cap \( C_{N}(r) \), we can get a bound for the variance of the sum in this region. (Since \( C_{S}(r) \) is distributionally equivalent to \( C_{N}(r) \), we only need to consider the variance of \( C_{N}(r) \).) Let \( L_{r} = [\gamma_{N}(r) / \phi (r)] \) and note that the Euclidean distance from any point to the pole is a “vertical” distance. Using Algorithm 2, (2.23), Lemma 3.1, and an argument similar to the distance calculations in Lemma 3.5, we have

\[
V \left\{ \sum_{\rho \in C_{N}(r)} X_{\rho} \right\} \leq |C_{N}(r)| \left( V \left\{ \mathcal{X}_{\delta} \right\} + 6 \sum_{k=1}^{2L_{r}} k \left| \mathcal{C} \left\{ \mathcal{X}_{(r,0,0)}(r,k \phi (c),0) \right\} \right| \right)
\]

\[
\leq |C_{N}(r)| \left( V \left\{ \mathcal{X}_{\delta} \right\} + c_{1} \sum_{k=1}^{2L_{r}} k \alpha \left( 2r \sin \left( \frac{k \phi (r)}{2} \right) \right) \right)
\]

\[
\leq |C_{N}(r)| \left( V \left\{ \mathcal{X}_{\delta} \right\} + c_{2} \left[ \alpha \left( \frac{1}{3} \right) + \sum_{k=1}^{\infty} k \alpha (k) \right] \right) = \mathcal{O} \left( |C_{N}(r)| \right) . \quad (A.6)
\]

Hence by recalling the size of \( C_{N}(r) \) given earlier,

\[
P \left\{ \left| t(C_{N}(r)) \right| \geq \varepsilon \sigma \sqrt{n_{r}} \right\} \leq \frac{1}{\varepsilon^2 \sigma^2 n_{r}} V \left\{ \sum_{\rho \in C_{N}(r)} X_{\rho} \right\} \leq \frac{\mathcal{O} \left( |C_{N}(r)| \right)}{\varepsilon^2 n_{r}} \to 0. \]

88
Proof of Corollary 3.3

Immediate upon noting where Lemma 3.3 was used and substituting Corollary 3.2. Also in the calculation of the variance of the spherical cap sum, use Corollary 3.2, and hence (A.6) and (A.7) incorporate $\alpha(j)^{4/(3+4)}$. The same argument still applies, the constants simply change. ■

Proof of Lemma 3.7

From Lemma 3.5 and the definition of $W_r$, we have: $W_r \propto r/J_r$ and $\sum k_{w,l_r} \propto W_r^2$, and write

$$\frac{1}{\sigma \sqrt{n_r}} \sum_{\vec{f} \in \mathcal{R}} X_\vec{f} \propto \frac{t^{t(r)}}{\sigma \sqrt{n_r}} + \frac{t^{(\mathcal{R} \setminus \mathcal{A})\{T(r)}{\sigma \sqrt{n_r}}.$$

From the given conditions and Lemma 3.6 we have $t^{(\mathcal{R} \setminus \mathcal{A})\{T(r)}/(\sigma \sqrt{n_r}) \xrightarrow{L} 0$. Hence by Slutsky, all we need is

$$\frac{t^{t(r)}}{\sigma \sqrt{n_r}} \xrightarrow{D} \mathcal{N}(0,1). \quad (A.8)$$

Using Lemma 3.5, we can study the independent type 4 tile sums. Define $Y_{w,k}(r) = t^{-t(r)}W_{w,k}(r)/\sqrt{n_r}$, then

$$\mathbb{E}\{Y_{w,k}(r)\} = 0 \quad \text{and} \quad \mathbb{V}\{Y_{w,k}(r)\} = I_r/J_r \sigma^2 (1 + o(1))/n_r$$

and (A.8) holds, if $\forall \varepsilon > 0$,

$$\lim_{r \to \infty} \sum_{w=1}^{W_r} \sum_{k=1}^{k_{w,l_r}} \int_{|s| \geq \varepsilon r} z^2 dP(Y_{w,k}(r) \leq z) = 0$$

where

$$s^2 = \sum_{w=1}^{W_r} \sum_{k=1}^{k_{w,l_r}} \mathbb{V}\{Y_{w,k}(r)\} = I_r/J_r \sigma^2 \left( \sum_{w=1}^{W_r} k_{w,l_r} (1 + o(1))/n_r \right) \xrightarrow{r \to \infty} \sigma^2.$$

Using the boundedness of $X_\vec{f}$, $n_r \propto I_r/J_r \sum k_{w,l_r} \propto r^2$, and the definition of a nearly uniform stratified spherical sampling plan, we have

$$\sum_{w=1}^{W_r} \sum_{k=1}^{k_{w,l_r}} \int_{|s| \geq \varepsilon r} z^2 dP(Y_{w,k}(r) \leq z) \leq \left( \frac{n_r(1 + o(1))}{I_r/J_r \sum k_{w,l_r} \sigma^2} \right)^2 \sum_{w=1}^{W_r} k_{w,l_r} \mathbb{E}\{(Y_{w,1}(r))^4\} \leq O \left( \frac{(I_r/J_r)^3}{n_r} \right) \to 0. \quad \blacksquare$$

Proof of Theorem 3.1

As in the proof of Lemma 3.7, write

$$\frac{1}{\sigma \sqrt{n_r}} \sum_{\vec{f} \in \mathcal{R}} X_\vec{f} \propto \frac{t^{t(r)}}{\sigma \sqrt{n_r}} + \frac{t^{(\mathcal{R} \setminus \mathcal{A})\{T(r)}{\sigma \sqrt{n_r}}.$$
From the given conditions and Corollary 3.3, we have $\frac{\tilde{t}(\mathcal{R}\setminus T(r))}{(\sigma\sqrt{n_r})} \xrightarrow{D} 0$. Using Slutsky's theorem, all we must show is (A.8). Applying Lemma 3.5, we can study the independent versions of the type 4 tile sums, $\tilde{t}(4T_{w,k}(r))$. We can then define

$$\tilde{t}(4T(r))/(\sigma\sqrt{n_r}) = \sum_{w=1}^{W_r} \sum_{k=1}^{k_{w,f}} \tilde{t}(4T_{w,k}(r))/(\sigma\sqrt{n_r})$$

$$= \tilde{Y}(r,N_k) + \bar{\tilde{Y}}(r,N_k)$$

where

$$\tilde{Y}(r,N_k) = \frac{\sigma(N_k)}{\sigma} \frac{1}{\sqrt{n_r\sigma(N_k)}} \sum_{\rho \in 4T(r)} \left[ N_k \bar{X}_\rho - \mathbb{E}\left\{ N_k \bar{X}_\rho \right\} \right] \text{ and}$$

$$\bar{\tilde{Y}}(r,N_k) = \frac{1}{\sqrt{n_r\sigma}} \sum_{\rho \in 4T(r)} \left[ N_k \bar{X}_\rho - \mathbb{E}\left\{ N_k \bar{X}_\rho \right\} \right] \text{ where}$$

$$\sigma^2(N_k) = \sum_{\rho \in \mathbb{Z}^2} C\left\{ N_k X_\rho, N_k X_\rho \right\},$$

and the $\bar{X}_\rho$'s are jointly distributed as the $X_\rho$'s within each tile $4T_{w,k}(r)$, but are independent between tiles $4T_{w,k}(r)$.

By separating the positive and negative parts of $N_k X_\rho$ and $N_k X_\rho N_k X_\rho$, we can use the Monotone Convergence Theorem to show that as $k \to \infty$, $\mathbb{E}\left\{ N_k X_\rho \right\} \to \mathbb{E}\left\{ X_\rho \right\}$ and $\mathbb{E}\left\{ N_k X_\rho N_k X_\rho \right\} \to \mathbb{E}\left\{ X_\rho X_\rho \right\}$. Applying the Dominated Convergence Theorem and condition 2, we have as $k \to \infty$, $\sigma^2(N_k) \to \sigma^2$. So then by Lemma 3.7, $\forall$ fixed $k > 0$ as $r \to \infty$,

$$\frac{\sigma}{\sigma(N_k)} \tilde{Y}(r,N_k) \xrightarrow{D} N(0,1).$$

Furthermore,

$$\mathbb{E}\left\{ \left| \tilde{Y}(r,N_k) \right|^2 \right\} \leq \frac{1}{\sigma^2 n_r} \sum_{w=1}^{W_r} k_{w,f} \mathbb{V}\left\{ \sum_{\rho \in 4T_{w,1}(r)} N_k X_\rho \right\}$$

$$\leq \frac{1}{\sigma^2} \left\{ \frac{1}{n_r} \sum_{w=1}^{W_r} k_{w,f} \left| 4T_{w,1}(r) \right| \mathbb{V}\left\{ N_k X_\rho \right\} + \right.$$  

$$+ \frac{1}{n_r} \sum_{w=1}^{W_r} k_{w,f} \sum_{\rho \notin 4T_{w,1}(r)} \left| C\left\{ N_k X_\rho, N_k X_\rho \right\} \right| \right\}.$$
In what follows all \( o(\cdot) \) and \( O(\cdot) \) are with respect to \( r \) as \( r \to \infty \) for some fixed \( k \). Since \( E \{ |X|^{2+\lambda} \} \leq c_1 < \infty \), we have \( V \{ N_k X \} \leq c_1 / N_k^k \) and since from Lemma 3.5, \( n_r \approx I_r J_r \sum k_{w,I_r} \), we have

\[
A_{r,k} = \frac{V \{ N_k X \}}{n_r} \left( (I_r - v_r)(J_r - v_r) \sum_{w=1}^{W_r} k_{w,I_r} \right) = (1 + o(1)) \frac{c_1}{N_k^k}.
\]

Letting \( G_r = G_r(0, I_r, 0, J_r) = \{(i,j) : 0 \leq i < I_r, 0 \leq j < J_r \} \), we get

\[
B_{r,k} \leq \frac{1}{n_r} \sum_{w=1}^{W_r} k_{w,I_r} \left\{ \sum_{(i,j) \neq (i',j') \in G_r} \left| C \left\{ N_k X_{(i,j)}, N_k X_{(i',j')} \right\} \right| + \sum_{(i,j) \neq (i',j') \in G_r} \left| C \left\{ N_k X_{(i,j)}, N_k X_{(i',j')} \right\} - C \left\{ N_k X_{(i,j)}, N_k X_{(i',j')} \right\} \right| \right\}.
\]

Applying Lemma 3.2 and an \( \alpha \)-mixing argument similar to that in Lemma 3.3,

\[
B_{r,k} = \sum_{(i,j) \neq (i',j') \in G_r} \left| C \left\{ N_k X_{(i,j)}, N_k X_{(i',j')} \right\} \right| + \sum_{(i,j) \neq (i',j') \in G_r} \left| C \left\{ N_k X_{(i,j)}, N_k X_{(i',j')} \right\} \right| - \sum_{j \geq \lfloor N_k^k \rfloor + 1} j \left[ \alpha(j) \right]^{\delta/(2+\lambda)}.
\]

Using the smoothness condition on the covariance function, we have uniformly in \( w \) and uniformly in \( (i,j) \) and \( (i',j') \in G_r \), \( \| P_{i,j}^w (r) - P_{i',j'}^w (r) \| \to \sqrt{(i-i')^2 + (j-j')^2} \). So for large enough \( r \geq r_k \), uniformly in \( w \), \( \| P_{i,j}^w (r) - P_{i',j'}^w (r) \| \in B(v, c_k) \).

\[
B_{2,r,k}^w = \sum_{(i,j) \neq (i',j') \in G_r} \left| B_k \left\{ \| P_{i,j}^w (r) - P_{i',j'}^w (r) \| \right\} - B_k \left\{ \sqrt{(i-i')^2 + (j-j')^2} \right\} \right| \leq c_k O \left( \frac{I_r J_r^2}{\sin \left( \gamma_1 (r) \right)} \right).
\]
Therefore for \( r \geq r_k \),

\[
E \left\{ \left| \bar{Y}(r, N_k) \right|^2 \right\} \leq \frac{(1 + o(1))}{\sigma^2} \left\{ \frac{2c_1}{N_k^{4/3}} + c_2 \sum_{j=\lfloor N_k^{4/3} \rfloor + 1}^{\infty} j[\alpha(j)]^{6/(2+\varepsilon)} + c_3 \mathcal{O}\left( \frac{J_5}{r \sin[\gamma_1(r)]} \right) \right\}.
\]

So then for any \( s > 0, \eta > 0 \), find \( k = k(s, \eta) \) such that

1. \( \left| 1 - \left( \frac{\sigma(N_k)}{\sigma} \right) \right|^2 \leq \frac{\eta}{2s^2} \),

2. \( \frac{2c_1}{\sigma^2 N_k^{4/3}} \leq \frac{\eta^2}{96s^2} \), and

3. \( \frac{c_2}{\sigma^2} \sum_{j=\lfloor N_k^{4/3} \rfloor + 1}^{\infty} j[\alpha(j)]^{6/(2+\varepsilon)} \leq \frac{\eta^2}{96s^2} \).

We know that as \( r \to \infty, (R \setminus T(r))/(\sigma \sqrt{\pi}) \overset{P}{\to} 0 \) and that for any fixed \( k \),

\( \bar{Y}(r, N_k) \overset{P}{\to} N \left( 0, (\sigma(N_k)/\sigma)^2 \right) \). With \( k \) fixed at the value found above, find \( r \) so that,

4. \( \left| E \left\{ \exp \left\{ is\bar{Y}(r, N_k) \right\} - \exp \left\{ -\frac{s^2}{2} \left( \frac{\sigma(N_k)}{\sigma} \right)^2 \right\} \right\} \right| \leq \frac{\eta}{4} \) and

5. \( \frac{c_3 \mathcal{O}\left( \frac{J_5}{r \sin[\gamma_1(r)]} \right)}{\sigma^2} \leq \frac{\eta^2}{96s^2} \).

We can use these in the expression for \( E \left\{ \left| \bar{Y}(r, N_k) \right|^2 \right\} \), so that \( E \left\{ \left| \bar{Y}(r, N_k) \right|^2 \right\} \leq \frac{s}{4\pi} \). We turn to the characteristic function of \( t(T(r))/(\sigma \sqrt{n_r}) \) and find that

\[
E \left\{ e^{ist(T(r))/(\sigma \sqrt{n_r})} - e^{-s^2/2} \right\} \leq
\]

\[
\leq \left| E \left\{ \exp \left\{ ist(T(r))/(\sigma \sqrt{n_r}) \right\} \right\} - E \left\{ \exp \left\{ is\bar{Y}(r, N_k) \right\} \right\} \right| +
+ \left| E \left\{ \exp \left\{ is\bar{Y}(r, N_k) \right\} - e^{-s^2/2} \right\} \right| +
+ \left| E \left\{ \bar{Y}(r, N_k) \right\} \left( \exp \left\{ is\bar{Y}(r, N_k) \right\} - 1 \right) \right| \]

\[
\leq \left| E \left\{ \exp \left\{ ist(T(r))/(\sigma \sqrt{n_r}) \right\} \right\} - E \left\{ \exp \left\{ is\bar{Y}(r, N_k) \right\} \right\} \right| +
+ \left| \exp \left\{ -\frac{s^2}{2} \left( \frac{\sigma(N_k)}{\sigma} \right)^2 \right\} - e^{-s^2/2} \right| +
+ \left| E \left\{ \exp \left\{ is\bar{Y}(r, N_k) \right\} \right\} \right| - \left| \exp \left\{ -\frac{s^2}{2} \left( \frac{\sigma(N_k)}{\sigma} \right)^2 \right\} \right| \]

92.
\begin{align*}
&+ E \left\{ \exp \left\{ i \tilde{y} (r, N_k) \right\} - 1 \right\} \\
&\leq \eta.
\end{align*}

Letting $\eta \downarrow 0$ gives the desired result. ■

A.3 Bootstrap Algorithm: Proofs of Lemmas and Theorems

Proof of Lemma 4.1

By repeated application of the Hölder inequality, we have

\begin{align*}
E \left\{ |X_{\theta_1} X_{\theta_2} \cdots X_{\theta_r}|^{2+e} \right\} &\leq \\
&\leq E^{1/4} \left\{ |X_{\theta_1}|^{2+4e/4} \right\} E^{(r-1)/4} \left\{ |X_{\theta_2} X_{\theta_3} \cdots X_{\theta_r}|^{2+4e/(r-1)} \right\} \\
&\leq c_q^{1/4} (\delta) E^{1/4} \left\{ |X_{\theta_1}|^{2+4e/4} \right\} E^{(r-2)/4} \left\{ |X_{\theta_2} X_{\theta_3} \cdots X_{\theta_r}|^{2+4e/(r-2)} \right\} \\
&\vdots \\
&\leq c_q^{(r-1)/4} (\delta) E^{(r-r+1)/4} \left\{ |X_{\theta_1}|^{2+4e/(r-r+1)} \right\}.
\end{align*}

If $r = q$, then the result is proved. If $r < q$, then repeat the above process one more time. ■

Proof of Corollary 4.1 Simply apply Lemma 3.2 when the random variables have mean zero and the same bounding constant. ■

Proof of Corollary 4.2

Case 1: $m(\tilde{\theta}_1, 3) = \|\tilde{\theta}_1 - \tilde{\theta}_2\|$ then let $\eta_1 = X_{\theta_1}$, $\eta_2 = \prod_{j=2}^3 X_{\theta_j}$. Notice that when $\epsilon = \delta/2$ then $E \left\{ |\eta_2|^{2+4\epsilon} \right\} \leq E \left\{ |X_{\theta}|^{4+2\epsilon} \right\} < \infty$ and $E \left\{ |\eta_2| \right\} \leq c_1 (\delta)$. Applying Lemma 3.2 using $\epsilon = \delta/2$ gives the desired result.

Case 2: $m(\tilde{\theta}_1, 3) = \|\tilde{\theta}_1 - \tilde{\theta}_2\|$ then let $\eta_2 = X_{\theta_1}$, and $\eta_1 = \prod_{j=1}^2 X_{\theta_j}$. Applying Lemma 3.2 using $\epsilon = \delta/2$ gives the desired result. ■

Proof of Corollary 4.3

Case 1: $m(\tilde{\theta}_1, 4) = \|\tilde{\theta}_1 - \tilde{\theta}_4\|$. Using Lemma 4.1

\begin{align*}
\Delta_4 (\tilde{\theta}_1, 4) \leq \left\| E \left\{ \prod_{j=1}^4 X_{\theta_j} \right\} \right\| + c_1 (\delta) \left\| E \left\{ \prod_{j=1}^2 X_{\theta_j} \right\} \right\|.
\end{align*}
For the first term on the right-hand side (RHS), let \( \eta_1 = X_{\tilde{r}_1} \) and \( \eta_2 = \prod_{j=2}^{4} X_{\tilde{r}_j} \).

Noting that we need \( \mathbb{E}\left\{ \left| \prod_{j=2}^{4} X_{\tilde{r}_j} \right|^{2+\epsilon} \right\} < \infty \), apply Lemma 4.1 and Lemma 3.2 using \( \epsilon = \delta/3 \) to give the desired result. For the second term on the RHS, let \( \eta_1 = X_{\tilde{r}_1} \) and \( \eta_2 = X_{\tilde{r}_2} \). Apply Lemma 3.2 using \( \epsilon = \delta/3 \) to give the desired result.

Case 2: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). Let \( \eta_1 = \prod_{j=1}^{2} X_{\tilde{r}_j} \) and \( \eta_2 = \prod_{j=3}^{4} X_{\tilde{r}_j} \). Applying Lemma 4.1 and Lemma 3.2 using \( \epsilon = \delta/3 \) to \( \Delta_4(\tilde{p}_{1,s}) \) gives the desired result.

Case 3: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). Similar to case 1, except we isolate \( \eta_1 = X_{\tilde{r}_1} \) and let \( \eta_2 \) be the product of the remaining random variables in both terms noting that

\[
\Delta_4(\tilde{p}_{1,s}) \leq \mathbb{E}\left\{ \prod_{j=1}^{4} X_{\tilde{r}_j} \right\} + c_2(\delta) \mathbb{E}\left\{ \prod_{j=3}^{4} X_{\tilde{r}_j} \right\} .
\]

**Proof of Corollary 4.4**

Case 1: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). Using Lemma 4.1 treat each summand separately and note that \( \mathbb{E}\left\{ \left| \prod_{j=2}^{5} X_{\tilde{r}_j} \right|^{2+\epsilon} \right\} \leq c_1(\delta) < \infty \) where \( \epsilon = \delta/4 \). Notice that

\[
\Delta_5(\tilde{p}_{1,s}) \leq \mathbb{E}\left\{ \prod_{j=1}^{5} X_{\tilde{r}_j} \right\} + c_2(\delta) \left[ \mathbb{E}\left\{ \prod_{j=1}^{2} X_{\tilde{r}_j} \right\} + \mathbb{E}\left\{ \prod_{j=3}^{4} X_{\tilde{r}_j} \right\} \right] .
\]

In all the terms on the RHS, let \( \eta_1 = X_{\tilde{r}_1} \) and \( \eta_2 \) be the product of the remaining random variables. Applying Lemma 3.2 using \( \epsilon = \delta/4 \) gives the desired result.

Case 2: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). Using Lemma 4.1

\[
\Delta_5(\tilde{p}_{1,s}) \leq \mathbb{E}\left\{ \prod_{j=1}^{5} X_{\tilde{r}_j} \right\} - \mathbb{E}\left\{ \prod_{j=1}^{2} X_{\tilde{r}_j} \right\} \mathbb{E}\left\{ \prod_{j=3}^{5} X_{\tilde{r}_j} \right\} + c_3(\delta) \mathbb{E}\left\{ \prod_{j=1}^{3} X_{\tilde{r}_j} \right\} .
\]

In the first term on the RHS, let \( \eta_1 = \prod_{j=1}^{2} X_{\tilde{r}_j} \) and \( \eta_2 = \prod_{j=3}^{5} X_{\tilde{r}_j} \). Now apply Lemma 3.2 using \( \epsilon = \delta/4 \). For the second term on the RHS, apply Corollary 4.2 to give the desired result.

Case 3: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). Similar to case 2, except we isolate \( \prod_{j=1}^{5} X_{\tilde{r}_j} \) since

\[
\Delta_5(\tilde{p}_{1,s}) \leq \mathbb{E}\left\{ \prod_{j=1}^{5} X_{\tilde{r}_j} \right\} - \mathbb{E}\left\{ \prod_{j=1}^{3} X_{\tilde{r}_j} \right\} \mathbb{E}\left\{ \prod_{j=4}^{5} X_{\tilde{r}_j} \right\} + c_4(\delta) \mathbb{E}\left\{ \prod_{j=1}^{3} X_{\tilde{r}_j} \right\} .
\]

Case 4: \( m(\tilde{p}_{1,s}) = ||\tilde{r}_1 - \tilde{r}_s|| \). As in case 1, we treat each summand separately, but we isolate \( X_{\tilde{r}_1} \) and let \( \eta_2 \) be the product of the remaining random variables. \( \blacksquare \)
Proof of Corollary 4.5

Case 1: \( m(\tilde{p}_{1,s}) = \|\tilde{p}_{1,s} - \tilde{p}_{4,s}\| \). Using Lemma 4.1 treat each summand separately and note that \( E\left\{ \prod_{j=2}^{2+4} X_{\tilde{p}_{1}} \right\} \leq c_1(\delta) < \infty \) where \( \epsilon = \delta/5 \). Apply Lemma 3.2 with \( \epsilon = \delta/5 \), always letting \( \eta_1 = X_{\tilde{p}_{1,s}} \) and let \( \eta_2 \) be the product of the remaining random variables.

Case 2: \( m(\tilde{p}_{1,s}) = \|\tilde{p}_{2,s} - \tilde{p}_{4,s}\| \). Using Lemma 4.1,

\[
\Delta_6(\tilde{p}_{1,s}) \leq \left| E\left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1}} \right\} - E\left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1}} \right\} E\left\{ \prod_{j=3}^{6} X_{\tilde{p}_{1}} \right\} \right| + c_2(\delta) \left| E\left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1}} \right\} \right| +
\]

\[
+c_3(\delta) \left| E\left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1}} \right\} - E\left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1}} \right\} E\left\{ \prod_{j=3}^{4} X_{\tilde{p}_{1}} \right\} \right|.
\]

Now apply Lemma 3.2 with \( \epsilon = \delta/5 \) to each term on the RHS, letting \( \eta_1 = \prod_{j=1}^{2} X_{\tilde{p}_{1}} \) and \( \eta_2 \) being the product of the remaining random variables.

Case 3: \( m(\tilde{p}_{1,s}) = \|\tilde{p}_{3,s} - \tilde{p}_{4,s}\| \). Using Lemma 4.1,

\[
\Delta_6(\tilde{p}_{1,s}) \leq \left| E\left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1}} \right\} - E\left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1}} \right\} E\left\{ \prod_{j=4}^{6} X_{\tilde{p}_{1}} \right\} \right| +
\]

\[
+c_4(\delta) \left| E\left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1}} \right\} \right| + \left| E\left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1}} \right\} \right| + c_5(\delta) \left| E\left\{ \prod_{j=3}^{4} X_{\tilde{p}_{1}} \right\} \right|.
\]

Now apply Lemma 3.2 with \( \epsilon = \delta/5 \) to each term on the RHS always selecting \( \eta_1 \) and \( \eta_2 \) in such a way as to separate them between \( X_{\tilde{p}_{3,s}} \) and \( X_{\tilde{p}_{4,s}} \).

Case 4: \( m(\tilde{p}_{1,s}) = \|\tilde{p}_{4,s} - \tilde{p}_{4,s}\| \). Similar to case 2, except we isolate \( \prod_{j=5}^{2} X_{\tilde{p}_{1}} \), since

\[
\Delta_6(\tilde{p}_{1,s}) \leq \left| E\left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1}} \right\} - E\left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1}} \right\} E\left\{ \prod_{j=5}^{6} X_{\tilde{p}_{1}} \right\} \right| +
\]

\[
+c_7(\delta) \left| E\left\{ \prod_{j=3}^{6} X_{\tilde{p}_{1}} \right\} - E\left\{ \prod_{j=3}^{4} X_{\tilde{p}_{1}} \right\} E\left\{ \prod_{j=5}^{6} X_{\tilde{p}_{1}} \right\} \right|.
\]

Case 5: \( m(\tilde{p}_{1,s}) = \|\tilde{p}_{3,s} - \tilde{p}_{4,s}\| \). As in case 1, we treat each summand on the RHS separately, except we isolate \( X_{\tilde{p}_{3,s}} \) and let \( \eta_2 \) be the product of the remaining random variables.

Proof of Corollary 4.6

95
Case 1: \( m(\tilde{p}_{1,8}) = ||\tilde{p}_{1,1} - \tilde{p}_{1,2}|| \). Using Lemma 4.1 treat each summand separately noting that 
\[
\mathbb{E} \left\{ \left| \prod_{j=2}^{8} X_{\tilde{p}_{1,j}} \right|^{2 + \epsilon} \right\} \leq c_1(\delta) < \infty \text{ where } \epsilon = \delta/7. \text{ Apply Lemma 3.2 with } \epsilon = \delta/7 \text{ to the term that has } X_{\tilde{p}_{1,1}} \text{ in it, always letting } \eta_1 = X_{\tilde{p}_{1,1}} \text{ and let } \eta_2 \text{ be the product of the remaining random variables.}
\]

Case 2: \( m(\tilde{p}_1,8) = ||\tilde{p}_3,8 - \tilde{p}_{1,8}|| \). Using Lemma 4.1,
\[
\Delta_8(\tilde{p}_{1,8}) \leq \left| \mathbb{E} \left\{ \prod_{j=1}^{8} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\tilde{p}_{1,j}} \right\} + c_2(\delta) \left| \mathbb{E} \left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1,j}} \right\} \right| + 
\right.
\]
\[
+ c_3(\delta) \left| \mathbb{E} \left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ \left| \mathbb{E} \left\{ \prod_{j=1}^{5} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{5} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ \left| \mathbb{E} \left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{6} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ c_4(\delta) \left| \mathbb{E} \left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\tilde{p}_{1,j}} \right\} \right| .
\]

Now apply Lemma 3.2 with \( \epsilon = \delta/7 \) to each term on the RHS, letting \( \eta_1 = \prod_{j=1}^{2} X_{\tilde{p}_{1,j}} \) and \( \eta_2 \) being the product of the remaining random variables.

Case 3: \( m(\tilde{p}_1,8) = ||\tilde{p}_4 - \tilde{p}_{1,8}|| \). Using Lemma 4.1,
\[
\Delta_8(\tilde{p}_{1,8}) \leq \left| \mathbb{E} \left\{ \prod_{j=1}^{8} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=4}^{8} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ c_5(\delta) \left| \mathbb{E} \left\{ \prod_{j=1}^{5} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=4}^{5} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ \left| \mathbb{E} \left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1,j}} \right\} - \mathbb{E} \left\{ \prod_{j=1}^{3} X_{\tilde{p}_{1,j}} \right\} \mathbb{E} \left\{ \prod_{j=4}^{6} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ \left| \mathbb{E} \left\{ \prod_{j=1}^{7} X_{\tilde{p}_{1,j}} \right\} \right| + \left| \mathbb{E} \left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ c_6(\delta) \left| \mathbb{E} \left\{ \prod_{j=1}^{6} X_{\tilde{p}_{1,j}} \right\} \right| + \left| \mathbb{E} \left\{ \prod_{j=1}^{4} X_{\tilde{p}_{1,j}} \right\} \right| + \left| \mathbb{E} \left\{ \prod_{j=3}^{5} X_{\tilde{p}_{1,j}} \right\} \right| + 
\]
\[
+ c_7(\delta) \left| \mathbb{E} \left\{ \prod_{j=3}^{6} X_{\tilde{p}_{1,j}} \right\} \right| .
\]

96
Apply Lemma 3.2 with $\epsilon = \delta/7$ to the first three terms on the RHS with $\eta_1 = \prod_{j=1}^{3} X_{\tilde{p}_{i,j}}$ and $\eta_2$ being the product of the remaining random variables. For the last six terms on the RHS, apply Lemma 3.2 with $\epsilon = \delta/7$ and $\eta_1 = X_{\tilde{p}_{i,s}}$ or $X_{\tilde{p}_{i,a}}$ (whichever is on the end) and $\eta_2$ being the product of the remaining random variables.

Case 4: $m(\tilde{p}_{1,s}) = ||\tilde{p}_{i,s} - \tilde{p}_{i,s}||$. Using Lemma 4.1,

$$\Delta_{\delta}(\tilde{p}_{1,s}) \leq \left| \mathbb{E}\left\{ \prod_{j=1}^{8} X_{\tilde{p}_{i,j}} \right\} - \mathbb{E}\left\{ \prod_{j=1}^{4} X_{\tilde{p}_{i,j}} \right\} \mathbb{E}\left\{ \prod_{j=5}^{8} X_{\tilde{p}_{i,j}} \right\} + \right.$$ 

$$+ c_{9}(\delta) \left[ \left| \mathbb{E}\left\{ \prod_{j=1}^{6} X_{\tilde{p}_{i,j}} \right\} - \mathbb{E}\left\{ \prod_{j=1}^{4} X_{\tilde{p}_{i,j}} \right\} \mathbb{E}\left\{ \prod_{j=5}^{6} X_{\tilde{p}_{i,j}} \right\} \right| + \right.$$ 

$$+ c_{10}(\delta) \left| \mathbb{E}\left\{ \prod_{j=3}^{6} X_{\tilde{p}_{i,j}} \right\} - \mathbb{E}\left\{ \prod_{j=3}^{4} X_{\tilde{p}_{i,j}} \right\} \mathbb{E}\left\{ \prod_{j=5}^{6} X_{\tilde{p}_{i,j}} \right\} \right| + \right.$$ 

$$+ c_{11}(\delta) \left[ \left| \mathbb{E}\left\{ \prod_{j=6}^{8} X_{\tilde{p}_{i,j}} \right\} + \mathbb{E}\left\{ \prod_{j=4}^{5} X_{\tilde{p}_{i,j}} \right\} \right| \right].$$

Apply Lemma 3.2 with $\epsilon = \delta/7$ to the first two terms on the RHS with $\eta_1 = \prod_{j=1}^{6} X_{\tilde{p}_{i,j}}$, to the third and fourth terms on the RHS with $\eta_1 = \prod_{j=3}^{4} X_{\tilde{p}_{i,j}}$, and to the last terms on the RHS either with $\eta_1 = X_{\tilde{p}_{i,s}}$ or $X_{\tilde{p}_{i,a}}$ (whichever is on the end) and $\eta_2$ being the product of the remaining random variables.

Case 5: $m(\tilde{p}_{1,s}) = ||\tilde{p}_{i,s} - \tilde{p}_{i,s}||$. Similar to case 3, except we isolate $\prod_{j=6}^{8} X_{\tilde{p}_{i,j}}$ or $X_{\tilde{p}_{i,a}}$.

Case 6: $m(\tilde{p}_{1,s}) = ||\tilde{p}_{i,s} - \tilde{p}_{i,s}||$. Similar to case 2, except we isolate $\prod_{j=7}^{8} X_{\tilde{p}_{i,j}}$.

Case 7: $m(\tilde{p}_{1,s}) = ||\tilde{p}_{i,s} - \tilde{p}_{i,s}||$. Similar to case 1, except we isolate $X_{\tilde{p}_{i,a}}$ and let $\eta_2$ be the product of the remaining random variables.■

**Proof of Theorem 4.1** Since

$$\mathbb{E}\left\{ |\mathcal{R}_{eps} \dot{\theta}_{smart}^{2} | \right\} = \frac{1}{|\mathcal{R}_{eps}|} \sum_{w \in \mathcal{W}_{eps}^{+}} \left( k_{w, r} \left( h_{w, r} + 0 \right) \mathbb{E}\left\{ w s_{i,s}^{2} \right\} + \left( m_{w, r} - h_{w, r} \right) \mathbb{E}\left\{ w s_{m_{w, r}}^{2} \right\} \right),$$

where $h_{w, r}$ is either $v_{r}$ or $J_{r}$ depending on whether $w$ is the equator or not, we must

investigate $\mathbb{E}\left\{ w s_{i,s}^{2} \right\}$ and $\mathbb{E}\left\{ w s_{m_{w, r}}^{2} \right\}$ for all $w \in \mathcal{W}_{eps}^{+}$. Notice that we only need to consider,
\( w \in \mathcal{W}_r \) since \( \mathbb{E} \{ w_{s_r}^2 \} = \mathbb{E} \{ w_{s_r}^{W_r+1-w_{s_r}^2} \} \). For \( w \in \mathcal{W}_r \setminus \{E\} \), we have

\[
\mathbb{E} \{ w_{s_r}^2 \} = \frac{V \{ \tilde{X}_r \} + \frac{1}{2} V \{ \tilde{X} \} - C \{ \tilde{X}_r, w \tilde{X} \} + \text{A}_{w}^{w} \frac{V \{ \tilde{X}_r \} + \frac{1}{2} V \{ \tilde{X} \} - C \{ \tilde{X}_r, w \tilde{X} \}}{\text{B}_{w}^{w}} \frac{1}{2} C \{ \tilde{X}_r, W_r+1-w \tilde{X} \}}{\text{C}_{w}^{w}} + \frac{1}{2} C \{ \tilde{X}, W_r+1-w \tilde{X} \}.
\]

Note that for \( w = E \), we only need to consider \( A_{w}^{w}, B_{w}^{w}, \) and \( C_{w}^{w} \) since there is no reflective counterpart wafer. Applying Corollary 3.2, we have \( \sup_{w \in \mathcal{W}_r} |A_{w}^{w} I_r - \sigma^2| = o(1) \). If we define for \( 1 \leq k \leq k_{w,l_r} \)

\[
S_{k}^{w} = \sum_{i=0}^{l_r-1} \sum_{j=0}^{l_r-1} X_{P_{k-1,l_r+i,j}(r)} \quad \text{and} \quad S_{k_{w,l_r}+1}^{w} = \sum_{i=0}^{k_{w,l_r}} \sum_{j=0}^{k_{w,l_r}+1} X_{P_{k_{w,l_r}+1,l_r+i,j}(r)}
\]

then \( \tilde{X} = (n_{w,r} J_r)^{-1} \sum_{k=1}^{k_{w,l_r}+1} S_{k}^{w} \). With this setup,

\[
(n_{w,r} J_r)^2 V \{ \tilde{X} \} = \frac{k_{w,l_r} V \{ S_{l_r}^{w} \} + V \{ S_{k_{w,l_r}+1}^{w} \}}{\text{B}_{r,1}^{w}} + \frac{2 \sum_{k=1}^{k_{w,l_r}} C \{ S_{k}^{w}, S_{k_{w,l_r}+1}^{w} \}}{\text{B}_{r,2}^{w}} \]

\[
+ \prod_{1 \leq k \neq k_{w,l_r} \leq k_{w,l_r}} C \{ S_{k}^{w}, S_{k}^{w} \}.
\]

Using the definitions of and conditions on \( l_r, n_{w,r} \) and \( m_{w,l_r} \), we find that as \( r \to \infty \)

\[
\sup_{w \in \mathcal{W}_r} |B_{r,1}^{w} / n_{w,r} J_r - \sigma^2| = o(1) \quad \text{and} \quad \sup_{w \in \mathcal{W}_r} |B_{r,2}^{w} / n_{w,r} J_r| = o(1).
\]

Before we proceed to the other terms, we first need to determine the covariance between neighboring tile sums. We begin by finding

\[
|C \{ S_{i}^{w}, S_{j}^{w} \}| \leq \sum_{i=0}^{l_r-1} \sum_{j=0}^{l_r-1} \sum_{i'=0}^{l_r-1} \sum_{j'=0}^{l_r-1} |C \{ X_{P_{i,j}(r)}, X_{P_{i'+j',r}(r)} \}|
\]

\[
\leq \sum_{i=1}^{l_r-1} \sum_{j=0}^{l_r-1} \sum_{i'=0}^{l_r-1} \sum_{j'=0}^{l_r-1} |C \{ X_{P_{i,j}(r)}, X_{P_{i',j'}(r)} \}|
\]

\[
+ \sum_{i=1}^{2l_r-1} \sum_{j=0}^{l_r-1} \sum_{i'=0}^{l_r-1} |C \{ X_{P_{i,j}(r)}, X_{P_{i,j'}(r)} \}|
\]

\[
+ \sum_{i=1}^{l_r} \sum_{j=0}^{l_r} \sum_{i'=0}^{l_r} (2l_r - i) |C \{ X_{P_{i,j}(r)}, X_{P_{i',j}(r)} \}|
\]

\[
98
\]
Using the Kronecker Lemma, condition 3, definition of a smooth covariance function, and Lemma 2.1,

\[
\sup_{w \in W_r} \{P_{r,1}^w\} \leq \\
\leq \sum_{i=1}^{I_r-1} \sum_{j=0}^{J_r-1} \sum_{j'=0}^{J_r-1} i |C \{X_{(0,j)}, X_{(i,j')}\}| + \\
+ \sum_{i=1}^{I_r-1} \sum_{j=0}^{J_r-1} \sum_{j'=0}^{J_r-1} \sup_{w \in W_r} i |C \{X_{P_{r,j}(r)}, X_{P_{r,j'}(r)}\} - C \{X_{(0,j)}, X_{(i,j')}\}| \\
\leq c_1 J_r^2 \sum_{i=0}^{I_r-1} \sum_{j=0 \leq j' < J_r} (i/I_r) |C \{X_{(0,0)}, X_{(i,j)}\}| + \\
+ c_2 J_r \sum_{w \in W_r} \sup_{0 \leq j < J_r} |B_0 \{\sqrt{i^2 + (j-j')^2}\} - B_0 \{\|P_{0,j}(r) - P_{0,j'}(r)\|}\} \\
\leq J_r^2 \left( o(1) + O \left(\frac{j^5}{r \sin(\gamma_1(r))}\right) \right) = o(J_r^2).
\]

We use the following facts about the arbitrary Euclidean distance between two points on a circle of radius \(y\) separated by an angle of \(\eta\): 1) their Euclidean distance is \(2y \sin(\eta/2)\) and 2) for \(0 < z < \pi/2, 2z/\pi < \sin(z) < z\). So then on a wafer with top angle \(\gamma_w(r)\) and two points on the top of the wafer separated by an angle of \(m\theta_w(r)\), then for sufficiently large \(r\), the Euclidean distance between these two points is more than \(2r \sin(\gamma_w(r))m\theta_w(r)/\pi\) and hence using the definition of \(\theta_w(r)\) is more than \(m/2\) uniformly in \(w\). Therefore for large enough \(r\) and using the monotonicity of \(\alpha(-)\), we have

\[
\sup_{w \in W_r} \{P_{r,2}^w\} \leq c_3 J_r^3 \sum_{i=1}^{I_r-1} \sup_{w \in W_r} [\alpha (\|P_{0,0}^w(r) - P_{i,0}^w(r)\|)]^{i/2+4} \\
\leq c_4 J_r^2 \sum_{i=1}^{I_r-1} \alpha (i/2) \leq o(J_r^2).
\]

So putting it all together gives \(\sup_{w \in W_r} |C \{S_1^w, S_2^w\}| = o(J_r^2)\). Furthermore,

\[
\sum_{k=2}^{k_{w,1\leq}} \sup_{w \in W_r} |C \{S_{k_{w,1\leq}}^w, S_{k_{w,1\leq}+1}^w\}| \leq c_5 J_r^4 \sum_{i=2}^{k_{w,1\leq}/2} \sup_{w \in W_r} [\alpha (\|P_{0,0}^w(r) - P_{(i-1)I_r,0}^w(r)\|)]^{i/(2+4)} \\
\leq c_6 J_r^4 \sum_{i=2}^{k_{w,1\leq}/2} \alpha (I_r (i-1)/2)^{i/(2+4)} \\
\leq c_7 J_r^{k_{w,1\leq}/2-1} \sum_{j=I_r/2}^{k_{w,1\leq}/2} \alpha (j)^{k_{w,1\leq}/2-1}\]

99
The final inequality holds because \( j \geq I_r/2 \) for each summand. Using these facts, 
\[
\sup_{w \in W_r} |B_{r,3}^{w}| = o(J_r^2). 
\]
If we define,

\[
m_{w,1, r}(x) = (k_{w,1, r} + x)I \{x \leq 0\} + xI \{0 < x \leq k_{w,1, r}\} + (x - k_{w,1, r})I \{x > k_{w,1, r}\}
\]

then using the same argument as in \( B_{r,3}^{w} \), we have for large enough \( r \),

\[
\sup_{w \in W_r} \left| \frac{B_{r,4}^{w}}{n_{w, r} J_r} \right| \leq \sup_{w \in W_r} \left[ \sum_{k=1}^{k_{w, r}} \left( \sum_{i=1}^{k_{w, r}/2} \left[ \frac{C \left\{ S_{i,1}^{w}, S_{i,1, r}^{w} \right\}^{(k+1)}}{n_{w, r} J_r} + \frac{C \left\{ S_{i,1}^{w}, S_{i,1, r}^{w} \right\}^{(k-1)}}{n_{w, r} J_r} \right] \right) \right] + \sup_{w \in W_r} \left[ \sum_{k=3}^{k_{w, r}/2-1} \left[ \frac{C \left\{ S_{i,1}^{w}, S_{i,1, r}^{w} \right\}^{(k+1)}}{n_{w, r} J_r} + \frac{C \left\{ S_{i,1}^{w}, S_{i,1, r}^{w} \right\}^{(k-1)}}{n_{w, r} J_r} \right] \right] \]

\[
\leq O(J_r^{-2}) \left( \sup_{w \in W_r} |C \left\{ S_{1}^{w}, S_{2}^{w} \right\}| + \sup_{w \in W_r} \sum_{k=3}^{k_{w, r}/2-1} |C \left\{ S_{i}^{w}, S_{k}^{w} \right\}| \right) \]

\[
= o(1),
\]

which implies that as \( r \to \infty \), \( \sup_{w \in W_r} |n_{w, r} J_r C_r^{w} - \sigma^2| \to 0 \). Let \( \mathbb{F}_r \) be the sum of observations on wafer \( w \) beginning at a theta angle of \( i \theta_w(r) \) and ending at angle \( (i + 1 - 1) \theta_w(r) \) and beginning at phi angle \( \gamma_w(r) \) and ending at angle \( \gamma_w(r) + (J_r - 1) \phi(r) \). Now then following the same argument as in \( B_r^w \),

\[
\sup_{w \in W_r} |n_{w, r} J_r C_r^{w} - \sigma^2| \leq \left( \frac{1}{I_{r} J_{r}} \right) \left\{ \sup_{w \in W_r} \left| \mathbb{F}_r \{ w \} \mathbb{F}_r \{ w \} - \sigma^2 \right| + 2 \sup_{w \in W_r} |C \left\{ w S_{i,1}, w S_{i,1} \right\}| + \sup_{w \in W_r} |C \left\{ w S_{i,1,1}, w S_{i,1, r} \right\}| \right\}.
\]

Using

\[
|C \left\{ w S_{i,1,1}, w S_{i,1, r} \right\}| \leq c_{10} J_r \left( \alpha \left( \left\| P_{i,1}^{w}(r) - P_{i,1}^{w}(r) \right\|^{2+4} \right) \right) \sum_{i=1}^{(n_{w, r}+1)J_r/2} \alpha \left( \left\| P_{i,1}^{w}(r) - P_{i,1}^{w}(r) \right\|^{2+4} \right)
\]

\[
\leq c_{11} J_r \left( \frac{1}{2} \sum_{j=I_{r}/2}^{\infty} j \alpha(j)^{2+4} \right) = o(J_r^2),
\]

we have, \( \sup_{w \in W_r} |n_{w, r} J_r C_r^{w} - \sigma^2| = o(1) \). For the other terms,

\[
\sup_{w \in W_r} |D_r^{w}| \leq \sup_{w \in W_r} \left[ \left\{ \sum_{i=0}^{I_{r} J_{r}} \sum_{j=0}^{I_{r} J_{r}} \sum_{i' = 0}^{I_{r} J_{r}} \sum_{j' = 0}^{I_{r} J_{r}} \left| C \left\{ X_{P_{i,j}^{w}(r)} X_{P_{i',j'}^{w}(r)} \right\} \right| \right\} \right].
\]
\[
\leq c_{12} \left[ \alpha \left( 2r \sin(v_r \phi(r)/2) \right) \right]^{d/(2+\delta)} = o(v_r^{-(m+1)}) \text{ and } \\
\sup_{w \in W_r} E_{w_r}^w \leq o(v_r^{-(m+1)}).
\]

Putting it all together and using condition 3, we have

\[
\sup_{w \in W_r} \left| J_r \cdot E \left\{ w_{s, r}^2 \right\} - \sigma^2 \right| = o(1) \left( 1 + \frac{l_r}{n_{w, r}} + l_r J_r v_r^{-(m+1)} \right) = o(1).
\]

We can similarly find \( E \left\{ E_{s, r}^w \right\} \) by above arguments, restricting ourselves to only \( A_r, B_r \), and \( C_r \) since the equatorial region does not have a reflective partner. We find that

\[
\sup_{w \in W_r} \left| l_r v_r \cdot E \left\{ E_{s, r}^w \right\} - \sigma^2 \right| = o(1) \left( 1 + \frac{l_r}{n_{E, r}} \right) = o(1).
\]

We similarly find that

\[
\sup_{w \in W_r} \left| m_{w, l_r} J_r \cdot E \left\{ w_{m, s, r}^2 \right\} - \sigma^2 \right| = o(1) \text{ and } \\
\sup_{w \in W_r} \left| m_{E, l_r} v_r \cdot E \left\{ E_{m, s, r}^w \right\} - \sigma^2 \right| = o(1).
\]

Joining all of this and recalling the definition of \( h_{w, r} \), we have

\[
E \left\{ |\hat{b}_n|^2 \right\} = \frac{1}{|\mathcal{B}|} \sum_{w \in W_r} \left( k_{w, l_r} (h_{w, r, l_r})^2 E \left\{ w_{s, r}^2 \right\} + (m_{w, l_r} h_{w, r, l_r})^2 E \left\{ w_{m, s, r}^2 \right\} \right)
\]

\[
= \frac{\sigma^2}{|\mathcal{B}|} \sum_{w \in W_r^+} k_{w, l_r} h_{w, r, l_r} (1 + o(1))
\]

\[
= \sigma^2 (1 + o(1)) \rightarrow \sigma^2. \Box
\]

Proof of Theorem 4.2

Note that

\[
\left( \sum_{i=1}^n z_i \right)^8 \leq 8! \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_8 \leq n} \left( \prod_{j=1}^8 z_{i_j} \right).
\]

Let \( m(\tilde{p}_{i, k}) = \max_{1 \leq k \leq 8} \left\{ \| \tilde{p}_{i, k} - \tilde{p}_{i, k-1} \| \right\} \) where \( \tilde{p}_{i, k} = P^n_{i, k}(r) \), we have

\[
E \left\{ \left( \frac{\hat{S}_1}{\sqrt{J_r}} \right)^8 \right\} \leq c_{11} J_r^{-4} \left\{ \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_8 \leq J_r} \left[ \alpha \left( m(\tilde{p}_{i, k}) \right) \right]^{d/(14+\delta)} + \right\}
\]

\[
+ \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_8 \leq J_r} \left\{ \sum_{k=2}^6 \sum_{j=1}^{k-1} E \left\{ X_{\tilde{p}_{i, j}} \right\} \right\} \mathbf{E} \left\{ \prod_{j=1}^8 X_{\tilde{p}_{i, j}} \right\} + \]
\[- \sum_{k=2}^{4} \sum_{l=2}^{6-k} \mathbb{E} \left\{ \prod_{j=1}^{k} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=k+1}^{k+l} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=k+l+1}^{8} X_{\bar{p}_{ij}} \right\} + \]
\[+ \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=7}^{8} X_{\bar{p}_{ij}} \right\} \}.\]

Consider the first term in the sum on the expectations, \( \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\bar{p}_{ij}} \right\} \). For the first expectation, we do not need any new terms. For the second term, we utilize Corollary 4.5 and hence need to add terms. Noting that for large enough \( r \) using the arbitrary Euclidean distance argument in Theorem 4.1, \( \| p_0 - p_{i_k} \| \geq i_k/3 \), then we have, uniformly in \( w \),

\[
\left| \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\bar{p}_{ij}} \right\} \right| \leq \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\bar{p}_{ij}} \right\} - \sum_{k=4}^{6} \mathbb{E} \left\{ \prod_{j=3}^{k} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=k+1}^{8} X_{\bar{p}_{ij}} \right\} + \]
\[+ \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=7}^{8} X_{\bar{p}_{ij}} \right\} + \]
\[+ \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \left[ \mathbb{E} \left\{ \prod_{j=5}^{8} X_{\bar{p}_{ij}} \right\} - \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=7}^{8} X_{\bar{p}_{ij}} \right\} \right] + \]
\[+ \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \left[ \mathbb{E} \left\{ \prod_{j=7}^{8} X_{\bar{p}_{ij}} \right\} + \mathbb{E} \left\{ \prod_{j=3}^{5} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=6}^{8} X_{\bar{p}_{ij}} \right\} \right] + \]
\[+ \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\bar{p}_{ij}} \right\} \left[ \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} - \mathbb{E} \left\{ \prod_{j=5}^{3} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=6}^{8} X_{\bar{p}_{ij}} \right\} \right] + \]
\[+ \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \left[ \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \right] + \]
\[+ \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \left[ \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \right] + \mathbb{E} \left\{ \prod_{j=3}^{4} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=5}^{6} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=7}^{8} X_{\bar{p}_{ij}} \right\} \] \[\leq \Delta_6 (\bar{p}_{3,8}) + \Delta_2 (\bar{p}_{3,4}) \Delta_4 (\bar{p}_{5,8}) + \Delta_2 (\bar{p}_{5,4}) \Delta_2 (\bar{p}_{5,6}) \Delta_2 (\bar{p}_{7,8}) + \Delta_3 (\bar{p}_{5,3}) \Delta_3 (\bar{p}_{6,8}) + \Delta_4 (\bar{p}_{5,3}) \Delta_2 (\bar{p}_{7,8}).\]

We use the fact that

\[
\sum_{1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq J_r} \Delta_k (\bar{p}_{1,k}) \leq c_2 J_r \sum_{j=1}^{J_r} j^{k-2} \alpha(j)^4 / (3^{(k-1)+4}).
\]

We then obtain

\[
\sum_{1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq J_r} \left| \mathbb{E} \left\{ \prod_{j=1}^{2} X_{\bar{p}_{ij}} \right\} \mathbb{E} \left\{ \prod_{j=3}^{8} X_{\bar{p}_{ij}} \right\} \right| \leq .
\]
\[
C_3 \left\{ J_r^2 \left( \sum_{j=1}^{J_r} \alpha(j)^{\beta/(2+\delta)} \right) \left( \sum_{j=1}^{J_r} j^4 \alpha(j)^{\beta/(10+\delta)} \right) + 
+ J_r^3 \left( \sum_{j=1}^{J_r} \alpha(j)^{\beta/(2+\delta)} \right)^2 \left( \sum_{j=1}^{J_r} j^2 \alpha(j)^{\beta/(6+\delta)} \right) + J_r^4 \left( \sum_{j=1}^{J_r} \alpha(j)^{\beta/(2+\delta)} \right)^4 + 
+ J_r^5 \left( \sum_{j=1}^{J_r} \alpha(j)^{\beta/(2+\delta)} \right) \left( \sum_{j=1}^{J_r} j \alpha(j)^{\beta/(4+\delta)} \right)^2 \right\}.
\]

Using the conditions of the theorem and taking the supremum over \( r \) and \( w \), we see that this term is finite. Follow the same argument as above for all of the other terms in the sum of the expectation by continuing to add and subtract terms in order to utilize Corollaries 4.2 – 4.6. ■

**Proof of Theorem 4.3**

Let \( h_{w,r} \) be the height of wafer \( w \) (either \( u_r \) or \( J_r \), depending on whether \( w \) is the equator or not) and \( W_r^* = \{1, 2, \ldots, W_r/2, E, W_r/2 + 1, \ldots, W_r\} \), then

\[
\sigma_{\text{boot}}^2 = |R| \sigma_{\text{boot}}^2 - E \{ |R| \sigma_{\text{boot}}^2 \} = \\
\frac{1}{|R|} \sum_{w \in W_r^*} \left[ \frac{h_{w,r} (h_{w,r} r_{w})^2}{c_{w,r}} \left( s_{w,r}^2 - E \{ w_{w,r}^2 \} \right) \right] \\
+ \frac{(h_{w,r} m_{w,r})^2}{d_{w,r}} \left( s_{m,w,r}^2 - E \{ w_{m,r}^2 \} \right).
\]

We have,

\[
V \{ |R| \sigma_{\text{boot}}^2 \} \leq \frac{4}{|R|^2} \left[ E \left\{ \left( \sum_{w \in W_r^*} c_{w,r} Y_{w,r} \right)^2 \right\} + E \left\{ \left( \sum_{w \in W_r^*} d_{w,r} Z_{w,r} \right)^2 \right\} \right].
\]

Starting with \( A_r \), we have

\[
A_r = \sum_{w \in W_r^*} c_{w,r}^2 \sum_{Y_{w,r}} + \sum_{w \in W_r^* \setminus \{E\}} c_{w,r} c_{w,r+1-w,r} C \{ Y_{w,r}, Y_{w+1-w,r} \} +
\]

\[
+ \sum_{w \in W_r^* \setminus \{W_r/2,E,W_r\}} c_{w,r} c_{w+1,r} C \{ Y_{w,r}, Y_{w+1,r} \} +
\]

\[
103
\]
\[
\begin{align*}
&\sum_{w \in W_{r} \setminus \{1, E, W_{r}/2+1\}} c_{w,l_{r}} c_{w-1,l_{r}} C \{ Y_{w,l_{r}}, Y_{w-1,l_{r}} \} + \\
&\sum_{w \in W_{r} \setminus \{W_{r}/2, E, W_{r}\}} c_{w,l_{r}} c_{W_{r}-w,l_{r}} C \{ Y_{w,l_{r}}, Y_{W_{r}-w,l_{r}} \} + \\
&\sum_{w \in W_{r} \setminus \{1, E, W_{r}/2+1\}} c_{w,l_{r}} c_{W_{r}+2-w,l_{r}} C \{ Y_{w,l_{r}}, Y_{W_{r}+2-w,l_{r}} \} + \\
&\left[ \text{all other pairs } w, w' \right] c_{w,l_{r}} c_{w',l_{r}} C \{ Y_{w,l_{r}}, Y_{w',l_{r}} \}.
\end{align*}
\]

If we let \( E \{ X \} = \mu < \infty \), then using the definitions of \( \bar{X} \) and \( c_{w,l_{r}} \), and properties of the variance of a sum, we have

\[
A_{r,1} \leq \sum_{w \in W_{r}} c_{l_{r}} (n_{w,r} h_{w,r})^2 (h_{w,r} l_{r})^2 \left[ V \left\{ \frac{1}{n_{w,r}} \sum_{i=0}^{n_{w,r}-1} \left( \bar{X}_{l_{r}} - \mu \right)^2 \right\} + \\
\left. + V \left( \bar{X} - \mu \right)^2 \right\} + V \left( \bar{X} - \mu \right)^2 \right].
\]

Using the circularity of the wafers and the fact that the sphere is homogeneous and isotropic, and defining

\[
\Delta_{r}^{w} = 2 \sum_{i=0}^{n_{w,r}/2} \sum_{i=0}^{l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \left\{ C \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} + \\
- E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} + \\
- E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} \right\} \text{ and}
\]

\[
\Gamma_{r}^{w} = 2 \sum_{i=0}^{n_{w,r}/2} \sum_{i=0}^{l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \sum_{i=0}^{i+l_{r}-1} \left\{ E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} + \\
+ E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} E \left\{ (\bar{X}_{l_{r}} - \mu) (\bar{X}_{l_{r}} - \mu) \right\} \right\}.
\]

104
we have
\[ A_{r,1}^w \leq \frac{2}{n_{w,r}} \sum_{i=0}^{n_{w,r}/2} C \left\{ (\langle y, X_i \rangle - \mu)^2, (\langle y, X_i \rangle - \mu)^2 \right\} \leq \Delta_r^w + \Gamma_r^w. \]

However,
\[ \Gamma_r^w = \frac{2^2}{n_{w,r}} \left( \frac{1}{h_{w,r} \cdot l_r} \right) \sum_{i=0}^{n_{w,r}/2} [C \{ \delta S_{i}, y S_{i} \}]^2 \]
and we isolate the cases when \( i = 0,0 < i < l_r + l_r, \) and \( i \geq l_r + l_r. \) When \( i = 0, \) we simply have \( V \{ \delta S_{i} \} = O(l_r h_{w,r}) \) by Corollary 3.2. For the other two cases a little care is needed, but nonetheless they follow the same basic setup: expand \( y S_{i} \) into a summation of \( y S_{1}, \) use the fact that under the conditions of the theorem, \( E \left\{ \left| \frac{y S_{1}}{\sqrt{r}} \right|^4 \right\} < \infty, \) and then upon applying Lemma 3.2 and the arbitrary Euclidean distance argument in Theorem 4.1, we have,
\[ \sup_{w \in W_{\Delta}^+} |C \{ \delta S_{i}, y S_{i} \} / (h_{w,r} l_r)| \leq c_1 \sum_{|i| < l_r} [\alpha((t + i)/2)]^{4/2 + 4}. \]
Using \( \sum_{j=1}^{\infty} \alpha(j)^{4/2 + 4} < \infty, \) we have for \( 0 < i < l_r + l_r, \)
\( \sup_{w \in W_{\Delta}^+} |C \{ \delta S_{i}, y S_{i} \} / (h_{w,r} l_r)| = O(1). \)
For \( i \geq l_r + l_r, \) use the fact that \( \sum_{j=1}^{\infty} j^m \alpha(j)^{4/2 + 4} < \infty \) which implies that
\( \sup_{w \in W_{\Delta}^+} |C \{ \delta S_{i}, y S_{i} \} / (h_{w,r} l_r)| = o(1). \)
Putting this all together and using the conditions of the theorem yields
\[ \sup_{w \in W_{\Delta}^+} \left| \Gamma_r^w \right| / \left( \frac{1}{n_{w,r} h_{w,r} l_r} \right) = O(1). \]
As for \( \Delta_r^w, \) if we let \( m_w(s, t, u, v) = \max \{ \| P_{w,0} - P_{w,0} \|, \| P_{w,0} - P_{t,0} \|, \| P_{t,0} - P_{r,0} \| \}, \)
using the fact that under the conditions of the theorem \( E \left\{ \left| \frac{y S_{1}}{\sqrt{r}} \right|^4 \right\} < \infty, \) and the arbitrary distance argument, we have
\[ \Delta_r^w \leq \frac{c_2}{n_{w,r}} \left( \frac{1}{h_{w,r} l_r} \right)^{2} \left[ l_r \sum_{0 \leq s \leq t \leq u < l_r} [\alpha(m_w(s, t, u, v))]^{1/4} + \sum_{i=2l_r}^{n_{w,r}/2} \left[ \alpha(m_w(i, t, u, v))]^{1/4} \right] \right] \]
\[ \leq c_3(\delta) \frac{1}{n_{w,r} h_{w,r} l_r} \left[ \frac{1}{l_r} \sum_{j=1}^{3l_r - 1} j^2 \alpha(j)^{1/4} + \sum_{j=l_r}^{\infty} j^2 \alpha(j)^{1/4} \right]. \]
Thus
\[ \sup_{w \in W_{\Delta}^+} \left| \Delta_r^w \right| / \left( \frac{1}{n_{w,r} h_{w,r} l_r} \right) = o(1). \]
As for \( A_{w,2}, \) we use the Cauchy-Schwartz Inequality and \( \sum_{j=1}^{\infty} j^2 \alpha(j)^{1/4} < \infty. \) Break the sum on the wafer into four quarter wafers and use the following fact: for two mean - zero random
variables $U$ and $V$, \( V \{ UV \} \leq \sqrt{V \{ U^2 \} V \{ V^2 \}} + 2V \{ U \} V \{ V \} \), to arrive at

\[
A_{r,2}^w = V \left\{ \left( \frac{w}{0} \bar{X}_{n_w,r} - \mu \right)^2 \right\} \leq c_0 \frac{1}{n_{0,r} J_r^2}.
\]

Applying the above fact for mean zero random variables to $A_{r,3}^w$ and using the results that $n_{w,r} J_r V \{ w \bar{X} \} = \mathcal{O}(1)$ from Theorem 4.1, then

\[
A_{r,3}^w \leq c_0 \frac{1}{n_{w,r} J_r^2}.
\]

Putting it all together, yields $|A_{r,1}| = \mathcal{O}(J_r l_r |\mathcal{R}|)$. As for $A_{r,2}$, using the fact that $V \{ Y_{w,l_r} \} \leq c n_{w,r} J_r^2 l_r$, we have

\[
|A_{r,2}| \leq \sum_{w \in W_r^* \setminus \{ E \}} \left( n_{w+1,r} J_r^2 l_r \right)^2 V \{ Y_{w,l_r} \} + V \{ Y_{w+1,l_r} \} = \mathcal{O}(J_r l_r |\mathcal{R}|).
\]

Notice that $A_{r,3} = A_{r,4}$. For large enough $r$ and using properties of $n_{w,r}, \theta_w(r)$, and $\phi(r)$, then

\[
|A_{r,3}| \leq \sum_{w \in W_r^* \setminus \{ W_r/2,E,W_r \}} \left( n_{w+1,r} J_r^2 l_r \right)^2 V \{ Y_{w,l_r} \} = c_0 \sum_{w \in W_r^* \setminus \{ W_r/2,E,W_r \}} n_{w+1,r} J_r^2 l_r \left( \frac{\theta_w(r)}{\theta_{w+1}(r)} \right) \leq \mathcal{O}(J_r l_r |\mathcal{R}|).
\]

As for $A_{r,5} = A_{r,6}$, apply the same logic as in $A_{r,3}$,

\[
|A_{r,5}| \leq c_1 \max \left\{ V \{ Y_{w,l_r} \}, V \{ Y_{w+1,l_r} \} \right\} = \mathcal{O}(l_r J_r |\mathcal{R}|).
\]

And then,

\[
|A_{r,7}| \leq 4 \left( n_{E,r} J_r^2 l_r \right)^2 = o(J_r l_r |\mathcal{R}|).
\]

Using results from Chung [Chu74, page 48], the Minkowski inequality, the Hölder inequality, Jensen's inequality and $E \left\{ |X|^4 \right\} < \infty$, we find that for all other pairs $w, w'$ in $A_{r,8}$,

\[
E \left\{ |Y_{w,l_r}|^{2+4} \right\} \leq E \left\{ \left( \frac{w}{0} \bar{X}_{l_r} - w \bar{\hat{X}} \right)^2 - E \left\{ \left( \frac{w}{0} \bar{X}_{l_r} - w \bar{\hat{X}} \right)^2 \right\}^{2+4} \right\}
\]

\[
\leq 2^{2+4} \frac{1}{n_{w,r} J_r^2 l_r} \left\{ X_{P_{E}\delta(r)} - \mu \right\}^{2+4} < \infty.
\]

And hence, using the fact that $\sum_{j=1}^{\infty} j^m \alpha(j)^{4/(2+4)} < \infty$ implies that $\alpha(j)^{4/(2+4)} = o(j^{-m+1})$ (see A.4), $1 < \frac{|w-w'| J_r \phi(r)/2}{\sin(|w-w'| J_r \phi(r)/2)} < \frac{\pi}{2}$, the definitions of $n_{w,r}$ and $l_r$, we have for large enough $r$

\[
|A_{r,8}| \leq \sum_{\text{all other pairs } w, w'} (n_{w,r} J_r^2 l_r) (n_{w',r} J_r^2 l_r) C \{ Y_{w,l_r}, Y_{w',l_r} \}.
\]

106
\[ \leq c_{0}J_{r}^{2} \sum_{\text{all other pairs } w, w'} n_{w, r}n_{w', r} \left[ \alpha \left( 2r \sin \left( \frac{|w - w'| - 1}{2}J_{r}\phi(r) \right) \right) \right]^{4/(2+3)} \]
\[ \leq c_{0}l_{r}J_{r} |P| \sum_{\text{all other pairs } w, w'} \frac{n_{w, r}n_{w', r}}{|P|J_{r}^{m-2}} \left( \frac{1}{|w - w'| - 1} \right)^{m+1} \]
\[ \leq c_{10}l_{r}J_{r} |P| \left( \frac{r}{|P|J_{r}^{m-2}} \right) \sum_{w=1}^{W_{r}} k_{w, l_{r}} = o(J_{r}l_{r} |P|). \]

And therefore, \( A_{r} = O(J_{r}l_{r} |P|). \) As for \( B_{r}, \) we follow the same argument as above, and so for large enough \( r, \)
\[ B_{r} \leq c_{11}l_{r}J_{r}^{2}W_{r} \sum_{w \in \mathcal{W}_{r}} h_{w, r}^{2}m_{w, l_{r}}^{2} \mathbb{V} \left\{ s_{w, m_{w, l_{r}}}^{2} \right\} \]
\[ \leq c_{12}l_{r}J_{r}^{2}W_{r} \sum_{w \in \mathcal{W}_{r}} \frac{m_{w, l_{r}}}{n_{w, r}} \]
\[ \leq O(l_{r}J_{r} |P|). \]

And hence, as \( r \to \infty \)
\[ \mathbb{V} \left\{ |P| \phi_{\text{boot}}^{2} \right\} = O \left( \frac{J_{r}l_{r}}{|P|} \right) \to 0. \]

Applying the Borel-Cantelli lemma to the sum of the variances of \( |P| \phi_{\text{boot}}^{2} \) shows that for a nearly uniform stratified spherical sampling plan, since \( J_{r}^{2}/r \to 0, \)
\[ \sum_{r=1}^{\infty} \mathbb{V} \left\{ |P| \phi_{\text{boot}}^{2} \right\} = c_{13} \sum_{r=1}^{\infty} \frac{J_{r}^{2}}{r} < \infty \]
and hence, \( |P| \phi_{\text{boot}}^{2} \xrightarrow{a.s.} \sigma^{2}. \)

**Proof of Lemma 4.2**

Notice that each \( \omega \in \Omega, \) completely determines the sequence of observations for the entire process. That is, for each \( \omega \in \Omega, \) the first \( |P_{1}| \) observations belong to the sphere of radius \( r = 1, \) the next \( |P_{2}| \) observations belong to the sphere of radius \( r = 2, \) and so on.

Using the Borel - Cantelli lemma, \( \mathbb{E} \{ |X_{k} - \mu|^{q} \} < \infty \) implies that for the ordering given above, there exists a \( k \) such that \( X_{k} = X_{P_{r_k}(r_k)}, \) and given \( \omega \in \Omega \setminus N, \) there exists an \( m^{*}(\omega) \) such that for all \( k \geq m^{*}(\omega), \) \( |X_{k} - \mu| \leq k^{1/q}. \) So since \( m^{*}(\omega) \leq \sum_{r=1}^{q^{*}(\omega)} |P| \) for some \( r^{*}(\omega), \) then for \( r \geq r^{*}(\omega), \) \( |X_{k} - \mu| \leq (r |P|)^{1/q}. \) Hence under the given conditions,
\[ \max_{w, t} \left| \sum_{i=0}^{t-1} \sum_{j=0}^{J_{r}-1} (X_{P_{r_k}(r_k)} - \mu) \right|^{q} \xrightarrow{a.s.} \frac{l_{r}J_{r} (r |P|)^{1/q}}{\sqrt{|P|}} = o \left( \sqrt{|P|} \right). \]
Furthermore, notice that if we break a centered sample mean on wafer $w$, $w\bar{X} - \mu$, into its $k_{w,l_r} + 1$ blocks of length less than or equal to $l_r$, then we also get,

$$\sup_{w \in W^+_r} \ell_r J_r |w\bar{X} - \mu|^2 = o(\sqrt{|R|}).$$

**Proof of Theorem 4.4**

If we let $E(X_F) = \mu < \infty$, then under the conditions of the theorem as $r \to \infty$,

$$\frac{1}{\sigma \sqrt{|R|}} \sum_{\beta \in R} (X_{\beta} - \mu) \overset{D}{\rightarrow} N(0,1).$$

Hence, we must show that for almost all $\{\bar{X}_r = \{X_{\beta} : \beta \in R\}, r \geq 1\}$,

$$\frac{1}{\hat{\sigma}_{\text{boot}}} (\bar{X}^* - \bar{X}) \overset{D}{\rightarrow} N(0,1).$$

For $w \in W_r$ and $1 \leq k \leq k_{w,l_r}$, let

$$w_{k-1,l_r}^* \bar{X}_{l_r} = \frac{w_{k-1}^* U_{l_r}}{\sqrt{|R|}},$$

$$w_0 \bar{X}_{m_{w,l_r}} = \frac{w_{0} U_{m_{w,l_r}}}{\sqrt{|R|}},$$

$$R_r^* = \frac{1}{|R|} \left\{ |E(r)| \bar{E} \bar{X}^* + |C_N(r)| \bar{C}_N(r) \bar{X}^* + |C_S(r)| \bar{C}_S(r) \bar{X}^* + \sum_{w=1}^{W_r} m_{w,l_r} J_r w \bar{X}_{m_{w,l_r}} \right\},$$

$$R_r = \frac{1}{|R|} \left\{ |E(r)| \bar{E} \bar{X} + |C_N(r)| \bar{C}_N(r) \bar{X} + |C_S(r)| \bar{C}_S(r) \bar{X} + \sum_{w=1}^{W_r} m_{w,l_r} J_r w \bar{X} \right\},$$

$$\bar{X}^* = \frac{1}{|R|} \sum_{w=1}^{W_r} \sum_{k=0}^{k_{w,l_r} - 1} J_{k,l_r} w \bar{X}_{l_r}^* + R^*_r,$$

$$\bar{X} = \frac{1}{|R|} \sum_{w=1}^{W_r} \sum_{k=0}^{k_{w,l_r} - 1} J_{k,l_r} w \bar{X} + R_r.$$

If we denote

$$Y_{w,k}(r) = \frac{l_r J_r}{\sigma \sqrt{|R|}} \left( w_{k-1,l_r}^* \bar{X}_{l_r}^* - w \bar{X} \right)$$

then we see that for each $w$, there are $2k_{w,l_r}$ $Y_{w,k}$'s that are conditionally independent and identically distributed with

$$\mathbb{E}^* \left\{ Y_{w,k}(r) | \bar{X}_r \right\} = 0 \quad \text{and} \quad \text{Var}^* \left\{ Y_{w,k}(r) | \bar{X}_r \right\} = \frac{l_r J_r^2}{2\sigma^2 |R|} \left( w_{k-1,l_r}^2 + W_r + 1 - w_{k-1,l_r}^2 \right).$$

So if we define

$$s^2_r = \sum_{w=1}^{W_r} \sum_{k=0}^{k_{w,l_r} - 1} \text{Var}^* \left\{ Y_{w,k}(r) | \bar{X}_r \right\} = \frac{l_r J_r^2}{\sigma^2 |R|} \sum_{w=1}^{W_r} k_{w,l_r} w_{k-1,l_r}^2,$$

(A.9)
then writing $|\mathcal{R}| \hat{\sigma}^2_{\text{boot}} = \sigma^2 s^2_r + C_r$ and using the fact that under the conditions of the theorem $|\mathcal{R}| \hat{\sigma}^2_{\text{boot}} \overset{a.s.}{\rightarrow} \sigma^2$ and $s^2_r \overset{a.s.}{\rightarrow} 1$ by Theorem 4.3, then $C_r \overset{a.s.}{\rightarrow} 0$. Defining $\mathcal{G} = \{ \omega \in \Omega : s^2_r \rightarrow 1 \text{ and } |\mathcal{R}| \hat{\sigma}^2_{\text{boot}} \rightarrow \sigma^2 \}$, then we must show that for every fixed $\omega \in \mathcal{G}$,

1. $A_r = \frac{1}{|\mathcal{R}|} (R^*_r - R^*_r) \overset{P}{\rightarrow} 0$, and

2. $B_r = \frac{1}{|\mathcal{R}|} \sum_{w=1}^{W_r} \sum_{k=0}^{k_{w,0} - 1} \left( \delta_0^w \tilde{X}^{*,\omega}_{r,0} - \delta_0^w \tilde{X}_r \right) \overset{P}{\rightarrow} N(0,1)$.

First note that

$$A_r = \frac{1}{|\mathcal{R}|} \left\{ |\mathcal{R}| \left( \langle \tilde{E} \tilde{X}^* - E \tilde{X} \rangle + \sum_{w=1}^{W_r} J_r m_{w,0} \left( \delta_0^w \tilde{X}^{*,\omega}_{m,w,0} - \delta_0^w \tilde{X}_r \right) \right) \right\}.$$

Using the definitions of $\tilde{E} \tilde{X}$, $\delta_0^w \tilde{X}^{*,\omega}_{m,w,0}$, and $\delta_0^w \tilde{X}$ and Chebyshev's inequality, we find that $\forall \epsilon > 0$,

$$P^* \left\{ |A_r| > \epsilon |\tilde{X}_r| \right\} \leq \frac{1}{|\mathcal{R}|} \epsilon^2 \left\{ \frac{1}{\epsilon^2} \left( \left( \frac{1}{\epsilon^2} \sum_{w=1}^{W_r} J_r m_{w,0} \left( \delta_0^w \tilde{X}^{*,\omega}_{m,w,0} - \delta_0^w \tilde{X}_r \right) \right) + \sum_{w=1}^{W_r} J_r m_{w,0} \left( \delta_0^w \tilde{X}^{*,\omega}_{m,w,0} - \delta_0^w \tilde{X}_r \right) \right) \right\} = C_r / \epsilon^2.$$

Using Theorem 4.3, we have $A_r \overset{P}{\rightarrow} 0$ as $r \rightarrow \infty$ almost surely. So then (2.) holds if the following Lindeberg condition holds $\forall \epsilon > 0$,

$$\lim_{r \rightarrow \infty} \sum_{w=1}^{W_r} k_{w,0} J_r \frac{1}{|\mathcal{R}|} J_r \mathbb{E}^* \left\{ \left( \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right)^2 \mathbb{I} \left\{ \left| \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right| \geq \sigma \epsilon |\sqrt{|\mathcal{R}|}| \frac{|\mathcal{R}|}{J_r} \right\} \right\} |\tilde{X}_r| = 0.$$

So if $\forall \epsilon > 0, \sup_{w} |C^*_r l_r | \sigma^r = o(1)$ then (2.) holds. Using the properties of the indicator function and the definition of $s^2_r$, for large enough $r$, we can bound $C^*_r$ by

$$C^*_r \leq \mathbb{E}^* \left\{ \left( \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right)^2 \mathbb{I} \left\{ \left| \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right| \geq \sigma \epsilon \frac{|\mathcal{R}|}{4 l_r J_r} \right\} \right\} + \mathbb{E}^* \left\{ \left( \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right)^2 \mathbb{I} \left\{ \left| \delta_0^w \tilde{X}^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right| \geq \sigma \epsilon \frac{|\mathcal{R}|}{4 l_r J_r} \right\} \right\}.$$

Since the conditions of Lemma 4.2 are satisfied, then for large enough $r \geq r^*(\omega, \epsilon)$,

$$I \left\{ \max_{w, I} \left( \frac{1}{\epsilon^2} \sum_{i=0}^{l-1} \sum_{j=0}^{J_r - 1} \left( \delta_0^w X^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right) \left( \delta_0^w X^{*,\omega}_{l,r} - \delta_0^w \tilde{X}_r \right) \right) \geq \sigma \epsilon \frac{|\mathcal{R}|}{4 l_r J_r} \right\} \equiv 0.$$
Hence $\sup_{w \in W_r} C^w_{r,1} \equiv 0$ for large enough $r$. Also breaking up $w \bar{X}$ into the two wafer averages and applying the result at the end of Lemma 4.2, we eventually have

$$1 \left\{ \left[ \frac{l_r d_r}{\sqrt{r}} \left( w \bar{X} - \mu \right) \right]^2 \geq \left( \frac{c_\sigma}{4} \right)^2 \right\} \equiv 0.$$  

Hence $\sup_{w \in W_r} C^w_{r,2} \equiv 0$ for large enough $r$. Putting these two together, we have $\sup_{w \in W_r} C^w_r \equiv 0$ for large enough $r$, and hence (2.) holds.

Strong uniform convergence now follows by adding and subtracting the cumulative distribution function of the standard normal and separating the bootstrap part and the non-bootstrap part. □
Bibliography


