ATHURULIYE LIYANA ARACHCHIGE, ROMESH RUWAN THANUJA, Ph.D.
THE WAVELET-GALERKIN METHOD ON GLOBAL RANDOM PROCESSES. (2021).
Directed by Dr. Haimeng Zhang. 109 pp.

One of the main usages of covariance function or kernel is to capture the spatial or temporal dependency of a random process or random field. Covariance functions have been widely used in many areas such as environmental statistics, economics, machine learning, atmospheric sciences, imaging analysis, etc. Hence, the understanding of a covariance function is crucial to the modeling, estimation, and prediction of a random process. In this dissertation, based on Mercer's theorem, we first modify an existing algorithm that uses the Wavelet-Galerkin method to approximate real-valued stationary covariance functions. We then apply the algorithm to approximate real-valued and complex-valued nonstationary covariance functions. In particular, we demonstrate the validity of the algorithm in the approximation of complex-valued covariance functions. In the second part of this dissertation, we apply the proposed algorithm to implement the Karhunen-Loéve expansion for studying axially symmetric Gaussian random processes on the sphere. The convergence of the truncated Karhunen-Loéve expansion to approximate axially symmetric Gaussian processes is established, and an expression for L^2 error bound of the above approximation is derived. Also, we propose an efficient algorithm to generate axially symmetric Gaussian data on the sphere with a given covariance structure, and we demonstrate that our method is comparable with the classical data generation method.

THE WAVELET-GALERKIN METHOD ON GLOBAL RANDOM PROCESSES

by

Romesh Ruwan Thanuja Athuruliye Liyana Arachchige

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2021

Approved by

_____
Committee Chair

To my family.

APPROVAL PAGE

This dissertation written by Romesh Ruwan Thanuja Athuruliye Liyana Arachchige has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Haimeng Zhang

Committee Members _____
Sat Gupta

_____
Scott Richter

_____
Xiaoli Gao

_____
Somya Mohanty

_____
Nicholas Bussberg

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGMENTS

These words cannot express my gratitude to my supervisor, Dr. Haimeng Zhang. I would not be where I am today without him. I greatly appreciate his support and kindness. I also would like to offer my deepest gratitude to the members of the committee, Dr. Gupta, Dr. Richter, Dr. Gao, Dr. Mohanty, and Dr. Bussberg for the wonderful suggestions to make improvements to this dissertation. My thanks are extended to Dr. Shivaji for the helpful guidance over the past five years. Furthermore, I wish to acknowledge the University of North Carolina at Greensboro, particularly, the Department of Mathematics and Statistics for providing me the financial support.

I would like to convey my heartfelt gratitude to my parents Nandasiri & Shanthi, my wife Bhashini, my sister Monisha, and my brother Dinuka, for always being there for me even in the darkest days of my life. Finally, I would like to thanks all of my school (Mahinda College, Sri Lanka) and university's (University of Peradeniya, Sri Lanka and University of North Carolina at Greensboro, USA) teachers and friends. I could not have made it this far without their support.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

In this section we will introduce basic concepts and important results in spatial statistics, including stationarity, intrinsic stationarity, and mean-squared continuity, etc. We will also give a brief discussion on stationary processes on a circle and axially symmetric processes on the sphere. Finally, the gridded data structure and the cross-covariance and cross-variogram estimators will be provided and with their elementary properties will be presented.

1.1 Spatial Random Fields

We define various random processes $\{X(t), t \in T\}$ as a collection of random variables based on the values of t taking from a specific domain T. Here are some of the commonly used random processes corresponding to certain choices of T.

- $T = \mathbb{N}$, a sequence of integers, $X(t)$ is a time series.
- $T = \mathbb{R}^1$, $X(t)$ is a stochastic process.
- $T = \mathbb{R}^d$, $d > 1$, $X(t)$ is a spatial process.
- $T = \mathbb{S}^2$, $X(t)$ is a random process on the sphere.
- $T = \mathbb{R}^d(\mathbb{S}^2) \times \mathbb{R}$, $X(t)$ is a spatio-temporal process on the sphere.

When $T = D \subset \mathbb{R}^d$ $(d > 0), \{X(t); t \in T\}$ is called a spatial process with $t = (t_1, t_2, \ldots, t_d)^\intercal \in \mathbb{R}^d$, being the spatial index.

In spatial statistics, to fully characterize a random process, one needs the finite dimensional distribution.

Definition 1.1. For any $n > 0$, the cumulative distribution function of a random vector $(X(t_1), \ldots, X(t_n))^\intercal$ on a n-tuple of spatial locations indexed by $(t_1, t_2, \ldots, t_n)$ is given by

$$F_{t_1,t_2,\ldots,t_n}(x_1, x_2, \ldots, x_n) = P\{X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n\},$$

where $x_1, x_2, \ldots, x_n \in \mathbb{R}$.

The Gaussian process is one of the most commonly used processes, where its finite dimensional distribution is multivariate normal. Note that a multivariate normal distribution is uniquely determined by its mean and variance-covariance matrix. Therefore, the Gaussian process is uniquely determined by its mean function and variance-covariance function.

1.1.1 Stationarity Concepts

A random field $X(t), t \in T$ is called to be strictly stationary if for any n, $t_1, t_2, \ldots, t_n \in T$, and random vectors $(X(t_1 + h), X(t_2 + h), \ldots, X(t_n + h))^\intercal$ and $(X(t_1), X(t_2), \ldots, X(t_n))^\intercal$ have the same joint distribution for any lag distance h. The assumption of strict stationarity might be too strong to be satisfied in real-world situation. A weaker assumption is weak stationarity (simply, stationarity). More specifically, a stationary random process satisfies following properties.

- $\forall t \in T$, $\mathrm{E}(X(t)) = \mu$ a constant.
- $\mathrm{E}(X^2(t)) < \infty \ \forall t \in T$, that is, the second moment of $X(t)$ is finite.

- The covariance (kernel) function $C(\cdot)$ only depends on the displacement (h) between two spatial locations. $C(h) = Cov(X(t), X(t+h))$.

A continuous covariance function $C(\cdot,\cdot)$ on $\mathbb{R}^d$ is called positive semi-definite if and only if

$$\sum_{i,j=1}^{n} C(t_i, t_j)\ a_i\ a_j \geq 0, \tag{1.1}$$

for any $n > 0$, a finite number of spatial locations $\{t_i, i = 1, 2, \ldots, n\}$ on $\mathbb{R}^d$, and any real numbers $\{a_i, i = 1, 2, \ldots, n\}$. If the process $\{X(t) : t \in T\}$ is strongly stationary and has finite second moment, $\{X(t) : t \in T\}$ is weakly stationary, but not vice versa (see Appendix A for an example). However, since a weakly stationary Gaussian random field is characterized by its mean and covariance function, weak stationarity implies strict stationarity.

If $C(\cdot)$: $\mathbb{R}^d \to \mathbb{R}$ is a covariance function of a stationary random field, it has following properties:

- $C(0) \geq 0$
- $C(t) = C(-t)\ , \forall t \in \mathbb{R}^d$
- $\mid C(t) \mid \leq C(0)\ , \forall t \in \mathbb{R}^d$
- $C(\cdot)$ satisfies the positive semi-definite as given by Definition (1.1).

1.1.2 Mean-Squared Continuity

For a sequence of random variables $X_1, X_2, \ldots$ and a random variable X defined on some common probability space, define $X_n \xrightarrow{L^2} X$ as $\mathrm{E}(X_n - X)^2 \to 0$, $n \to \infty$

and $\mathrm{E}(X^2) < \infty$. We say X_n converges in L^2 if there exists X such that $X_n \xrightarrow{L^2} X$. Suppose $X(t)$ is a random field on $\mathbb{R}^d$. We say $X(t)$ is mean squared continuous at s if

$$\lim_{t \to s} \mathrm{E}\{X(t) - X(s)\}^2 = 0.$$

For a weakly stationary process $X(t)$ with covariance function $C(\cdot)$, $\mathrm{E}\{X(t)-X(s)\}^2 = 2\left(C(0) - C(s-t)\right)$, [Mic98].

1.1.3 Construction of Covariance Functions

Suppose $C : \mathbb{R}^d \to \mathbb{R}$ is a continuous function, then C is the covariance function of some stationary Gaussian random field if and only if

$$C(h) = \int_{\mathbb{R}^d} e^{i<w,h>} d\mu(w), \text{ where } < w, h >= \sum_{i=1}^{d} w_i h_i, \tag{1.2}$$

for some finite symmetric Borel measure μ on $\mathbb{R}^d$. The measure μ is called spectral measure of the random field. In particular, if μ admits an even density $f : \mathbb{R}^d \to \mathbb{R}^+$ and with (1.2), then

$$C(h) = \int_{\mathbb{R}^d} e^{i<w,h>} f(w) dw = \int_{\mathbb{R}^d} f(w) \cos(< w, h >) dw. \tag{1.3}$$

By the inverse Fourier formula, $f(w)$ can be obtained as,

$$f(w) = \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} e^{-i<w,h>} C(h) dh, \tag{1.4}$$

which is the so-called spectral density. For instance, here are some stationary covariance functions and corresponding spectral densities.

(1) Matern class

$$f(w) = \frac{1}{\phi(\alpha^2 + w^2)^{v+\frac{1}{2}}} \text{ where } \phi, v, \alpha > 0,$$

$$C(h) = \frac{\pi^{1/2}\phi}{2^{v-1}\Gamma(v+1/2)\alpha^{2v}}(\alpha|h|)^v Y_v(\alpha|t|),$$

where Y_v is the modified Bessel function of second kind, Γ is the gamma function and α and v are non-negative parameters of the covariance function.

(2) Gaussian spectral density $f(w)$ and Gaussian covariance function $C(h)$,

$$f(w) = \frac{1}{2\sqrt{\pi\alpha}} c e^{\frac{-w^2}{4\alpha}},$$

$$C(h) = ce^{-\alpha h^2},$$

where $\alpha, c > 0$ and $\omega \in \mathbb{R}$.

1.1.4 The Semi-Variogram

Semi-variogram can be defined by

$$\gamma(h) = \frac{1}{2} var(X(t+h) - X(t)),$$

where h is the displacement. Note that for weakly stationary random fields, the following relationship holds

$$\gamma(h) = C(0) - C(h).$$

The variogram function $2\gamma(\cdot)$ is conditionally negative semi-definite; that is

$$\sum_{i,j=1}^{N} 2\gamma(t_i - t_j)\ a_i\ a_j \leq 0,\ \forall t_i, t_j \in \mathbb{R}^d,$$

for any finite number of spatial locations $\{t_i, i = 1, 2, \ldots, N\}$ on $\mathbb{R}^d$ and any real number $\{a_i, i = 1, 2, \ldots, N\}$ with constraint $\sum_{i=1}^{N} a_i = 0$.

1.1.5 Intrinsically Stationary Process

A process is called an intrinsically stationary process if it has a finite constant mean and the variogram function is only a function of displacement (h).

- $E(X(t)) = \mu,\ \forall t \in \mathbb{R}^d$.
- $Var(X(t+h) - X(t)) = 2\gamma(h)$.

Note that second order stationarity implies intrinsic stationarity, but the converse may not be true. For instance, consider the following linear semivariogram function.

$$\gamma(h) = \begin{cases} \tau^2 + \sigma^2 h, & \text{if } h > 0, \tau^2 > 0, \sigma^2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\gamma(h) \to \infty$ when $h \to \infty$ and so this semi-variogram does not correspond to a weakly stationary process (although it is intrinsically stationary).

1.2 Circularly Symmetric Gaussian Random Vectors

A complex-valued random variable Z is circularly symmetric if $e^{i\phi}Z$ has the same probability distribution as for Z for all ϕ. Let $\vec{Z} = (Z_1, Z_2, \ldots, Z_n)^\intercal$ be a complex jointly-Gaussian vector. We say $Z = Z^R + iZ^I$ is complex Gaussian if both Z^R and Z^I are real-valued random variables and jointly Gaussian. Then $\vec{Z}$ can be expressed as a set of $2n$ jointly-Gaussian real-valued random variables such that $Z_k = (Z_k^R, Z_k^I)^\intercal, k = 1, 2, \ldots, n$. In addition, covariance matrix of $\vec{Z}$, $K_{\vec{Z}}$ and the pseudo-covariance matrix $M_{\vec{Z}}$ are given by

$$K_{\vec{Z}} = \mathrm{E}[\vec{Z}\vec{Z}^*],$$
$$M_{\vec{Z}} = \mathrm{E}[\vec{Z}\vec{Z}^\intercal],$$

where $\vec{Z}^* = \vec{\bar{Z}}^\intercal$ is the conjugate transpose of $\vec{Z}$. Furthermore, according to the Theorem 1 in [Gal08], If we assume that Z is a complex jointly-Gaussian random vector, then Z is circularly symmetric if and only if $M_{\vec{Z}} = 0$.

1.3 Circulant and Block Circulant Matrices

A square matrix in which each row (after the first) has the elements of the previous row shifted cyclically one place to the right, is called a circulant matrix.

[Dav79] denotes it as

$$A = circ(a_0, a_1, \ldots, a_{n-1}) = \begin{pmatrix} a_0 & a_1 & a_2 & \ldots & \ldots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \ldots & \ldots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & a_0 & & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & & \vdots & \vdots \\ a_1 & a_2 & a_3 & \ldots & \ldots & a_{n-1} & a_0 \end{pmatrix},$$

for $\vec{a} = (a_0, a_1, \ldots, a_{n-1})^\intercal \in \mathbb{R}^n$.

Eigenvalues of A are given by

$$\lambda_j = \sum_{k=0}^{n-1} a_k w_j^k, \text{ where } w_j^k = e^{i2\pi jk/n}, \ j = 0, 1, \ldots, n-1.$$

If A is also symmetric, that is $a_i = a_{n-i}$, then eigenvalues are real. The eigenvector Ψ_j is given by,

$$\vec{\Psi_j} = \frac{1}{\sqrt{n}}(1, w_j, w_j^2, \ldots, w_j^{n-1})^\intercal.$$

A block circulant matrix was first proposed by [Mui1920]. $A \in \mathbb{R}^{np\times np}$ with $[A_k]_{p\times p},\ k = 0,\ldots,n-1$ sub block matrices is given below.

$$A = bcirc(A_0, A_1, \ldots, A_{n-1}) = \begin{pmatrix} A_0 & A_1 & A_2 & \ldots & \ldots & A_{n-2} & A_{n-1} \\ A_{n-1} & A_0 & A_1 & \ldots & \ldots & A_{n-3} & A_{n-2} \\ \vdots & \vdots & A_0 & & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & & \vdots & \vdots \\ A_1 & A_2 & A_3 & \ldots & \ldots & A_{n-1} & A_0 \end{pmatrix}.$$

According to [CL88], matrix A can be decomposed as

$$A = P\text{diag}(S_1, S_2, \ldots, S_n)P^*,$$

where $P = \frac{1}{\sqrt{n}}(E_1, E_2, \ldots, E_n)$ with

$$E_j = \begin{pmatrix} I_p \\ w_j I_p \\ w_j^2 I_p \\ \vdots \\ w_j^{n-1} I_p \end{pmatrix}, j = 1, 2, \ldots, n \text{ and } w_j^k = e^{i2\pi jk/2},$$

and $S_j = \sum_{k=0}^{n-1} w_j^k A_m$. More details about eigenvalues and eigenvectors can be found in [Gar15].

1.4 Covariance and Variogram Estimators on a Circle

It is important to investigate expressions of the variance (variogram) and the covariance as dependencies between two angles on the circle are dominated by these two parameters. In this section we will introduce random processes, stationary concepts, covariance, and variogram estimators on the circle.

1.4.1 Random Process on a Circle

Let $\{X(t),\ t \in \mathbb{S}^1\}$ be a random process on a unit circle that has a finite second moment and is continuous in quadratic mean. Then $X(t)$ can be represented as a Fourier series with convergence of the series in quadratic mean

$$X(t) = A_0 + \sum_{n=1}^{\infty}(A_n \cos(nt) + B_n \sin(nt)),\ t \in [0, 2\pi],$$

where random variables A_0, A_n and B_n $(n = 1, 2, \ldots)$ can be expressed by

$$\begin{aligned} A_0 &= \frac{1}{2\pi}\int_0^{2\pi} X(t)\ dt, \\ A_n &= \frac{1}{\pi}\int_0^{2\pi} X(t)\cos(nt)\ dt, \\ B_n &= \frac{1}{\pi}\int_0^{2\pi} X(t)\sin(nt)\ dt. \end{aligned}$$

1.4.2 Stationary and Intrinsically Stationary Processes on the Circle

A process is called to be stationary on the circle if the process has constant mean and the covariance function only depends on the angular difference.

$$C(\theta) = cov(X(t+\theta), X(t)),\ \theta \in [0, \pi].$$

Note that the stationary covariance function has the following spectral decomposition $C(\theta) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k\theta)$, where $a_k \geq 0,\ k = 0, 1, 2, \ldots$, and $\sum_{k=0}^{\infty} a_k < \infty$.

We can find a_0 and $a_k, k \geq 1$ by considering the orthogonality of $\cos(k\theta)$

$$a_0 = \frac{1}{\pi} \int_0^{\pi} C(\theta)\, d\theta, \quad a_k = \frac{2}{\pi} \int_0^{\pi} C(\theta) \cos(k\theta)\, d\theta, \quad k \geq 1.$$

Note that if a random process $X(t)$ is intrinsically stationary, $\mathrm{E}(X(t)) = \mu$ (constant), and the variogram function depends only on the angular distance θ. The variogram function on the circle has the following spectral decomposition

$$\gamma(\theta) = \sum_{k=1}^{\infty} a_k (1 - \cos(k\theta)),$$

where $a_k \geq 0,\ k = 0, 1, 2, \ldots$ and $\sum a_k < \infty$.

1.4.3 Covariance and Variogram Estimator on the Circle

Let $\vec{X} = (X_1, X_2, \ldots, X_n)^{\intercal}$ with $X_j = X((j-1)\delta)$ is a collection of gridded observations on the circle. Let $\Delta\lambda = k\delta$ for $k = 0, 1, 2, \ldots, N = n/2$, where n is the number of data points on the circle and $\delta = \frac{2\pi}{N}$ (here for simplicity, $n = 2N$ even). Then we have the following Method of Moment (MOM) covariance estimator.

$$\begin{aligned} \hat{C}(\Delta\lambda) &= \frac{1}{n} \sum_{i=1}^{n} (X((i-1)\delta + \Delta\lambda) - \bar{X})(X(i-1)\delta) - \bar{X}) \\ &= \frac{1}{n} \sum_{i=1}^{n} (X((i-1)\delta + \Delta\lambda)(X(i-1)\delta) - (\bar{X})^2, \text{where } \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i. \end{aligned}$$

According to the [Ada17], the above expression of the covariance estimator can be converted to a quadratic form,

$$\hat{C}(\Delta\lambda) = \vec{X}^{\intercal} M(\Delta\lambda)\vec{X},$$

where

$$M(\Delta\lambda) = \frac{1}{n} circ(-\frac{1}{n}, -\frac{1}{n}, \ldots, -\frac{1}{n}, \ldots, -\frac{1}{n}).$$

As stated in [Ada17], we can define the variogram estimator on the circle as follows. Let $\vec{X}_{(k+1)} = (X_{k+1}, \ldots, X_n, X_1, \ldots, X_{k-1}, X_k)^{\intercal}$ be the rotated vector $\vec{X}$ by k^{th} position. Hence, we have, for $\Delta\lambda = k\delta$, $k = 0, 1, 2, \ldots, N$.

$$\begin{aligned}\hat{\gamma}(\Delta\lambda) &= \frac{1}{2n}\sum_{i=1}^{n}(X((i-1)\delta + \Delta\lambda) - (X(i-1)\delta))^2 \\ &= \frac{1}{2n}(\vec{X}_{(k+1)} - \vec{X})^{\intercal}(\vec{X}_{(k+1)} - \vec{X}).\end{aligned}$$

We write $\hat{\gamma}(\Delta\lambda)$ as the following quadratic form

$$\hat{\gamma}(\Delta\lambda) = \vec{X}^{\intercal} A(\Delta\lambda)\vec{X},$$

where,

$$\begin{aligned}A(0) &= 0; \\ A(k\delta) &= \frac{1}{2n}\text{circ}(2, 0, 0, \ldots, -1, 0, \ldots, -1, 0, \ldots, 0),\ 1 \leq k \leq N-1,\end{aligned}$$

where -1's are placed at $(k+1)^{th}$ and $(n-k+1)^{th}$ positions, respectively.

$$A(N\delta) = \frac{1}{2n} circ(2, 0, 0, \ldots, -2, \ldots, 0, \ldots, 0),$$

where -2 is placed at $(N+1)^{th}$ position.

1.5 Random Processes on the Sphere

Let $\{X(P), P = (\phi, \lambda) \in \mathbb{S}^2\}$ be a possibly complex-valued random process with longitude $\lambda \in [0, 2\pi)$ and latitude $\phi \in [0, \pi]$. $X(P)$ can be expressed as a combination of spherical harmonics with quadratic mean convergence. If we can assume for every location the process $X(P)$ is continuous in quadratic mean and has a finite second moment [LN97]

$$X(P) = X(\lambda, \phi) = \sum_{v=0}^{\infty} \sum_{m=-\nu}^{\nu} Z_{\nu,m} e^{im\lambda} P_{\nu}^{m}(\cos\phi),$$

where P_{ν}^{m} is a normalized associated Legendre polynomial and indices ν and m (which are integers) are considered as the degree and order of the associated Legendre polynomial, respectively. $Z_{\nu,m}$ coefficients can be obtained by

$$Z_{\nu,m} = \int_{\mathbb{S}^2} X(P) e^{-im\lambda} P_{v}^{m}(\cos\phi)\, dP.$$

Without loss of generality (WLOG), we can set $E(X(P)) = 0$, then the covariance function between two locations $P = (\phi_P, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$ would become (for

example, [HZR12])

$$R(P,Q) = \mathrm{E}(X(P)\overline{X(Q)})$$
$$= \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{m=-\nu}^{\nu}\sum_{n=-\mu}^{\mu} \mathrm{E}(Z_{\nu,m}\overline{Z_{\mu,n}})e^{im\lambda_P}P_\nu^m(\cos\phi_P)e^{-in\lambda_Q}P_\mu^n(\cos\phi_Q), \quad (1.5)$$

where $\overline{Z_{\nu,m}}$ denotes the complex conjugate of $Z_{\nu,m}$.

1.5.1 Homogeneous (Stationary) Processes on the Sphere

If the underlying process is homogeneous then the covariance function only depends on the spherical distance between two locations. That is, [Yag61]

$$R(P,Q) = R(\theta(P,Q)) = \sum_{\nu=0}^{\infty}\frac{(2\nu+1)f_\nu}{2}P_\nu(\cos\theta(P,Q)),$$

where $\cos\theta(P,Q) = (\sin\phi_P\sin\phi_Q + \cos\phi_P\cos\phi_Q\cos(\lambda_P-\lambda_Q))$, $\theta(P,Q)$ is called the spherical angle, $P_\nu(\cdot)$ is the Lengendre polynomial of order ν, $f_\nu \geq 0$, and $\sum_{\nu=0}^{\infty}(2\nu+1)f_\nu < \infty$.

1.5.2 Axially Symmetric Processes on the Sphere

$\{X(P), P \in \mathbb{S}^2\}$ is said to be an axially symmetric process on the sphere, if the covariance function of the process only depends on the longitudinal difference. Consider the two location $P = (\phi_P, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$, then the covariance function between P and Q can be written as

$$R(\phi_Q, \phi_Q, \lambda_P, \lambda_Q) = R(\phi_P, \phi_Q, \lambda_P - \lambda_Q).$$

Under the assumption of axial symmetry [Jon63], we can simplify equation (1.5) as follows;

$$\mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) = \delta_{n,m} f_{\nu,\mu,m},$$

where $\delta_{n,m} = 1,$ when $n = m,$ and 0 otherwise. Hence, the covariance function is of the form (for example, [Ste07])

$$\begin{aligned} R(P,Q) &= R(\phi_P, \phi_Q, \lambda_P - \lambda_Q) \\ &= \sum_{m=-\infty}^{\infty} \sum_{\nu=|m|}^{\infty} \sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} e^{im(\lambda_P - \lambda_Q)} P_\nu^m(\cos\phi_P) P_\mu^m(\cos\phi_Q). \end{aligned} \tag{1.6}$$

For each fixed integer m, the matrix $F_m(N) = \{f_{\nu,\mu,m}\}_{\nu,\mu=|m|,|m|+1,\ldots,N}$ must be positive definite for all $N \geq |m|$. In equation (1.6), for each $m = 0, \pm 1, \ldots$, we let

$$C_m(\phi_P, \phi_Q) = \sum_{\nu=|m|}^{\infty} \sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} P_\nu^m(\cos\phi_P) P_\mu^m(\cos\phi_Q).$$

According to [HZR12], $C_m(\phi_P, \phi_Q)$ is Hermitian and positive definite by considering the properties of $f_{\nu,\mu,m}$ and $F_m(N)$.

As mentioned in [HZR12], a continuous covariance function defined on axially symmetric processes has the following structure

$$R(P,Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q), \tag{1.7}$$

where $\Delta\lambda \in [-\pi, \pi]$, and $C_m(\phi_P, \phi_Q)$ is Hermitian and positive definite with $\sum_{m=-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$.

1.5.3 Longitudinally Reversible Processes on the Sphere

A longitudinally reversible process is a special case of an axially symmetric process, where the covariance function satisfies the following additional property.

$$R(\phi_P, \phi_Q, \lambda_Q - \lambda_P) = R(\phi_P, \phi_Q, \lambda_P - \lambda_Q).$$

Under this assumption of reversibility, the above equation yields $C_{-m}(\phi_P, \phi_Q) = C_m(\phi_P, \phi_Q)$, then the equation (1.7) can be simplified as

$$R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=0}^{\infty} C_m(\phi_P, \phi_Q) \cos(m\Delta\lambda).$$

1.5.4 Cross Covariance and Variogram Functions on the Sphere

If $R(\phi_P, \phi_Q, \Delta\lambda)$ is the cross-covariance function between two locations $P = (\phi_P, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$ with $\Delta\lambda = \lambda_P - \lambda_Q$, then the cross-variogram function $\gamma(\phi_P, \phi_Q, \Delta\lambda)$ can be expressed as

$$\gamma(\phi_P, \phi_Q, \Delta\lambda) = R(\phi_P, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)). \tag{1.8}$$

When the process is longitudinally reversible then the above equation (1.8) becomes

$$\gamma(\phi_P, \phi_Q, \Delta\lambda) = R(\phi_P, \phi_Q, 0) - R(\phi_P, \phi_Q, \Delta\lambda). \tag{1.9}$$

1.5.5 Construction of Gridded Data Structure

Table 1 illustrates how gridded data points are distributed throughout the entire sphere. Let $\delta = \frac{2\pi}{n_L} = \frac{2\pi}{n}$ be a fixed common length on a latitude. The longitude difference, $\Delta\lambda$, between two points $P = (\phi_p, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$ can only be taken the following sequence of values. WLOG, consider n as an even number. $\Delta\lambda = k\delta, k = \pm 1. \pm 2, \ldots, \pm N,$ where $N\delta = \pi$ and $n = n_l = 2N$.

Table 1. Gridded data structure on the sphere ($\overline{X_{\phi_i}} = \frac{1}{n}\sum_{j=1}^{n_L} X_{ij}$)

Latitudes	Observed data on each latitude				Means of Latitudes
	0	$\boldsymbol{\delta}$	$\boldsymbol{\ldots}$	$\boldsymbol{(n_L-1)\delta}$	
ϕ_1	X_{11}	X_{12}	$\ldots$	X_{1n_L}	$\overline{X_{\phi_1}}$
ϕ_2	X_{21}	X_{22}	$\ldots$	X_{2n_L}	$\overline{X_{\phi_2}}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
ϕ_{n_l}	X_{n_l1}	X_{n_l2}	$\ldots$	$X_{n_ln_L}$	$\overline{X_{\phi_{n_l}}}$

When $n_l = 2$, that is there are only two latitudes denoted as ϕ_P and ϕ_Q. All observations on latitudes ϕ_P and ϕ_Q can be commonly written as $\{X(\phi_P, k\delta)\}_{k=0}^{n-1}$ and $\{X(\phi_Q, k\delta)\}_{k=0}^{n-1}$, respectively. Moreover, we form the following vector $\vec{X} = (X_{11}, X_{21}, , X_{12}, \ldots, X_{1n}, X_{2n})^\intercal$, where $X_{1k} = X(\phi_P, k\delta)$ and $X_{2k} = X(\phi_Q, k\delta)$.

1.5.6 Cross Covariance and Variogram Estimator on the Sphere

The MOM (Method of Moments) based cross covariance estimator between two points for an axially symmetric process can be expressed as follows [Ada17].

$$\begin{aligned}
\hat{R_{12}}(\Delta\lambda) &= \hat{R}(\phi_1, \phi_2, \Delta\lambda) \\
&= \frac{1}{n}\sum_{i=1}^{n}(X(\phi_1, (i-1)\delta + \Delta\lambda) - \bar{X_{\phi_1}})(X(\phi_2, (i-1)\delta) - \bar{X_{\phi_2}}) \\
&= \frac{1}{n}\sum_{i=1}^{n}(X(\phi_1, (i-1)\delta + \Delta\lambda))(X(\phi_2, (i-1)\delta)) - \bar{X_{\phi_1}}\bar{X_{\phi_2}} \\
&= \vec{X}^{\intercal} A(\Delta\lambda)\vec{X},
\end{aligned}$$

where,

$$A(\Delta\lambda) = \frac{1}{n} circ\left(\begin{pmatrix} 0 & \frac{-1}{n} \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1-\frac{1}{n} \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & \frac{-1}{n} \\ 0 & 0 \end{pmatrix}\right),$$

$A(\Delta\lambda)$ is a block-circulant matrix.

We can apply unitary diagonalization to $A(\Delta\lambda)$ which gives the following.

$$A(\Delta\lambda) = P\ \mathrm{diag}(S_1^{(A)}, S_2^{(A)}, \ldots, S_n^{(A)})\ P^*,$$

with

$$S_j^{(A)} = \sum_{m=0}^{n-1} w_j^m A_m,$$

where $A_m,\ m = 0, 1, \ldots, n-1$ represents the 2×2 blocks in $A(\Delta\lambda)$.

$$S_j^{(A)} = \frac{1}{n} w_j^k \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \text{if } 2 \leq j \leq n,$$

and

$$S_1^{(A)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ \text{if } j = 1.$$

1.5.7 Cross Variogram Estimator on the Sphere

For an axially symmetric process we can define the MOM cross variogram estimator as given in [Ada17].

$$\hat{\gamma_{12}}(\Delta\lambda) = \hat{\gamma}(\phi_1, \phi_2, \Delta\lambda) = \frac{1}{2n} \sum_{i=1}^{n} (X(\phi_1, (i-1)\delta + \Delta\lambda) - X(\phi_1, (i-1)\delta)$$
$$(X(\phi_2, (i-1)\delta + \Delta\lambda) - X(\phi_2, (i-1)\delta) = \vec{X}^{\intercal} A(\Delta\lambda) \vec{X}, \quad (1.10)$$

where $A(\Delta\lambda)$ is a block-symmetric circulant matrix.

For instance, let's consider $n = 6$, the following are the obtained matrices for $\Delta\lambda = 0, \pi/3, 2\pi/3$ values.

$A(0) = \mathbf{0}_{12\times 12}$

$$A(\pi/3) = \tfrac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} \right),$$

$$A(2\pi/3) = \tfrac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right),$$

$$A(\pi) = \tfrac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right).$$

Also, $A(\Delta\lambda)$ has a spectral decomposition as follows.

$$A(\Delta\lambda) = P\ \text{diag}(S_1^{(A)}, S_2^{(A)}, \ldots, S_n^{(A)})\ P^*$$

with

$$S_j^{(A)} = (1 - \cos((j-1)\Delta\lambda)))\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\ j = 1, 2, \ldots, n.$$

Hence,

$$A(\Delta\lambda) = P\ \text{diag}(1 - \cos((j-1)\Delta\lambda)))\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\ P^*.$$

CHAPTER II

LITERATURE REVIEW

2.1 Research Studies on Spatial Data

Spatial data is used to explain information at locations on the earth's surface. Such data first appeared in the form of data maps. For instance, in 1686, Halley mapped the direction of trade winds and monsoons onto a map of land forms. In general, spatial data are mainly classified into three types based on their characteristics (i.e., how data are available in nature). Geostatistical data is collected from regularly spaced or irregularly spaced spatial locations (for example, percentage of coal ash sampled on a grid in a mine). Lattice data are gathered from regularly spaced $\mathbb{R}^2$ that is linked to nearest neighborhoods, for example, temperature data that are measured from small rectangular areas of the earth's surface by satellites. In point patterns processes, data are collected at fixed locations and regions (for example, positions of some particular trees in an Island) or sometimes information of locations where "events" happen randomly. However, in my dissertation, I focus only on geostatistical data. According to [Cre93], many geostatistics applications can be found in a wide range of areas such as hydrology, soil science, public health, and climate predictions.

Modeling spatial data first appeared in a study of Student in 1907, apparently after a few centuries of mapping spatial data. Modeling spatial data and making predictions are challenging compared to traditional data modeling and predictions

in statistics. In order to make predictions we need to capture the local dependency between locations accurately. In spatial statistics, the dependency between locations is captured by covariance functions. Since every point itself is a random variable and only one observation is available at each location, we cannot construct covariance functions without any assumptions on those random variables. In addition, researchers often need to deal with high-dimensionality problem when the data set is large. For instance, 55 parameters need to be estimated to construct the variance-covariance matrix to predict a values of the unobserved spatial location from the 10 observed locations.

The covariance function of a random process is crucial for modeling, making inferences, and predictions. There has been a rich study of covariance functions that are valid on $\mathbb{R}^d$. However, with the availability of global-scale data in recent years, the study of covariance functions on global scale processes has received increasing attention. Note that due to the difference of topological structures of $\mathbb{R}^d$ and the sphere $\mathbb{S}^2$, the covariance functions that are valid in $\mathbb{R}^d$ might not be valid on $\mathbb{S}^2$ ([HZR11] and [Gne13]). In this dissertation, based on Mercer's theorem, we first develop an existing algorithm that uses the Wavelet-Galerkin method [PHQ02] to approximate real-valued nonstationary covariance functions. Galerkin method is a type of expansion method where any function can be written as a linear combination of finite basis functions when we solve the Fredlohom homogeneous equation. I use wavelets as the basis functions because of their capability of capturing the inhomogeneity of the approximating functions. Furthermore, we propose an algorithm using Wavelet-Galerkin method to approximate possibly nonstationary complex-valued co-

variance functions that are widely used in geostatistics, engineering, and computer science applications, etc. (see [KM11] and [SCI17]).

Over the past few decades with the development of technology, remote sensing data are widely used in many areas. These types of data are collected by satellites or remote sensing networks and they often display high-dimensionality. Gridded level 3 Total Ozone Spectrometer (TOMS) data measured by Nimbus-7 satellite daily from November 1, 1978, to May 1993, is a well-known example for the global data. There are some missing values in the data set and data points above 73^0S are not available in the image because the instrument used was back-scattered sunlight during the time of collecting data.

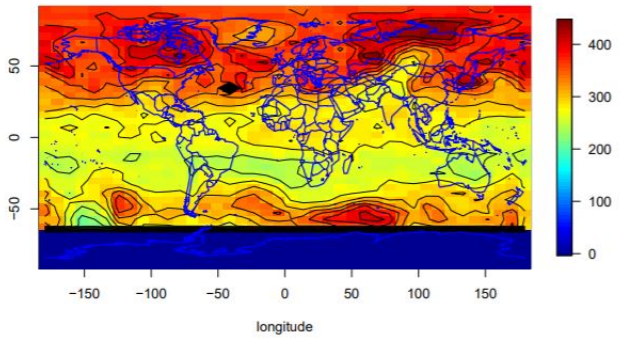

Figure 1. TOMS data; data resolution spatial 1^0 Latitude $\times$ 1.25^0 Longitudes in May, 1-6 1990.

1

[1]NASA reference publication, Nimbus–7 Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide, 1996.

According to the given spatial resolution 1^0 latitudes $\times$ 1.25^0 longitudes of the image (Figure 1) possibly there are more than 20,000 data points (after subtracting all missing data points from $\frac{100^0}{1^0} \times \frac{360^0}{1.25^0}$ data points) can be taken into the analysis [Ste07]. That means the variance-covariance matrix will be a high dimensional because of the number of parameters that are required to estimate for making predictions on unknown spatial locations. [Ste07] used 170 parameters in the variance-covariance matrix to model the TOMS data. However, he was not able to fully capture the global dependency.

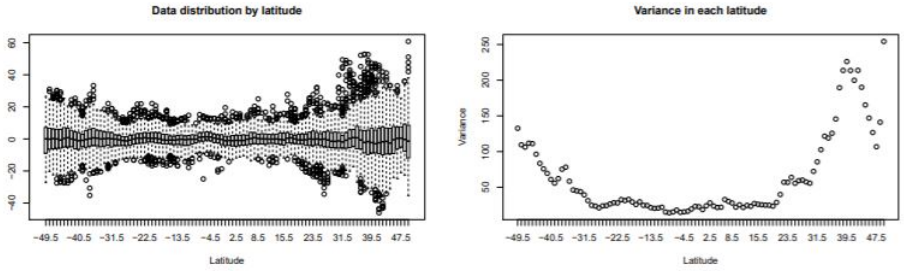

Figure 2. Distribution at each latitude (left) and variance at each latitude (right) between 50^0S and 50^0N

Before making any further assumptions of the random processes defined on the earth's surface, researchers tend to visualize the characteristics of the distributions along with the directions of latitudes and longitudes to investigate the important properties of the underlying processes. From the above Figure 2, one can observe that although the mean value of each latitude fluctuates around zero, variances are comparatively higher in the latitudes near the north and south poles. Homogeneity

assumption of the underlying random process may not be valid in this case. Therefore, we need to define alternative model assumptions to handle these types of global data.

A wide variety of random process assumptions other than the homogeneity assumptions have been proposed in the literature over the last few decades. For example, some properties and modeling using isotropic covariance functions on the sphere have been studied in [GF15]. [HS11] have investigated properties of an existing class of models for Gaussian processes on the sphere that are invariant to shifts in longitudes (anisotropic). [LRL11] used Gaussian Markov random fields and stochastic partial differential equations to analyze global data with non-stationary covariance models. A flexible family of non-stationary processes called intrinsic random functions (IRFs) is formally defined on the sphere by [HZRS19], under which the homogeneity can be achieved through a truncation operation. In research work, [BUS20] applied the IRF approach to propose a universal kriging procedure.

Due to the earth's inclination and rotation, many random processes such as environments and geophysical sciences exhibit a heterogeneous spatial dependency across latitudes and stationarity across longitudes. Therefore, finding statistical methods to model and analyze these types of random processes has received special attention in the recent literature. [Jon63] named these types of random processes as the axially symmetric random process on $\mathbb{S}^2$. In 2007, [Ste07] verified that these assumptions of the axially symmetry can be applied to global data by using the TOMS data set. [HZR12] proposed a simplified representation of covariance function for

axially symmetric random processes on the sphere. Since then researchers have been exploring covariance functions that characterize the dependency of axially symmetric random processes on the sphere, see [PAC18] for a complete review on the recent development.

There are limited studies in the literature on data generation of axially symmetric processes on the sphere. [Yan13] proposed convolution methods to generate random fields with a class of Matérn type kernel function. Recent work by [VWZ19] demonstrated that the axially symmetric data on the sphere can be decomposed into a Fourier series on circles, where the Fourier random coefficients can be expressed as circularly-symmetric complex random vectors. The Karhunen-Loeve expansion has been applied to a wide range of areas such as functional data analysis [RS97], finance [Sch04], pattern recognition [KS90], and machine learning [RW05], etc. In this dissertation I apply Karhunen-Loeve expansion to geostatistics, especially when the random process on the sphere is Gaussian axially symmetric. The convergence of the truncated Karhunen-Loéve expansion to approximate axially symmetric Gaussian random processes are established and a expression for L^2 error bound of the above approximation is derived, under which we propose an efficient algorithm to generate axially symmetric Gaussian data on the sphere with the given covariance structure.

2.2 Outline of This Dissertation

This dissertation is organized as follows. In Chapter 3, I briefly introduce Mercer's theorem, Haar wavelet, Fredlohom homogeneous equation, and the Galerkin method. Then an algorithm is proposed to numerically find solutions to Fredlohom

homogeneous equation. All of these concepts and methods are merged to numerically approximate three different types of covariance functions, stationary, real-valued nonstationary, and complex-valued nonstationary covariance functions. In Chapter 4, the Karhunen Loéve expansion defined on a complex domain is introduced to propose a computationally efficient algorithm to generate Gaussion axially symmetric global data. A convergence study is carried out for generating data, and the theoretical L^2 error bound is obtained for such an approximation. In addition, a validation study is conducted via simulations to compare our proposed model with the classical method on data generation. Finally a summary of this research as well as future work directions are provided in Chapter 5.

CHAPTER III
WAVELET-GALERKIN METHOD AND COVARIANCE FUNCTIONS APPROXIMATION

3.1 Introduction

Approximating real-valued positive definite covariance functions on $\mathbb{R}^d \times \mathbb{R}^d, d \geq 1$, has become a central part of geospatial statistics because an accurate approximation of the covariance (kernel) function is crucial for parametric inferences and optimal spatial prediction [Ste99]. In this chapter, we propose an approach using orthogonal basis functions to approximate possibly complex-valued covariance functions that are widely used in geostatistics, engineering, and computer science applications, etc. (see [KM11] and [SCI17]).

Different types of wavelets and their properties have been widely studied in recent years. Wavelet basis methods have been developed in areas such as regression, density and function estimation, factor analysis, modeling and forecasting in time series analysis, and spatial statistics [Vid99]. Wavelets basis functions are capable of capturing the local information of the approximating functions specially when the eigenfunction is inhomogeneous. The Galerkin method is a type of an expansion method where any function can be written as a linear combination of finite basis functions when we solve the Fredlohom homogeneous equation. Frequently, wavelets are considered as bases, but the choice of the basis function used in the Galerkin method solely depends on the application. There is a significant advantage of Gar-

lekin method over integration method. The accuracy of the solutions is determined not only by the number of terms in the expansion but also by the shape of the approximating functions. As stated in [PHQ02], the conventional Galerkin method (which use trigonometrics and polynomials as bases) would create a computational burden for large-scale data sets because this approach leads to a dense matrix.

In this Chapter, a brief introduction of Mercer's theorem and Wavelet-Galerkin method is first given. These concepts and methods are combined to numerically approximate three different types of covariance functions, stationary, real-valued nonstationary covariance functions, and complex-valued nonstationary covariance functions. Finally, the performance of the proposed algorithm will be illustrated with examples.

3.2 Introduction to Mercer's Theorem and the Wavelet-Galerkin Method

3.2.1 Mercer's Theorem

Mercer's theorem states that any symmetric, continuous, and positive semi definite kernel function $C(t,s)$ defined in the square interval $a \leq t \leq b$ and $a \leq s \leq b$ can be written as;

$$C(t,s) = \sum_{i=1}^{\infty} \eta_i \; f_i(t) \; f_i(s), \quad a \leq t, s \leq b,$$

where $\eta_i \geq 0$, $i = 1, 2, \ldots$ are the eigenvalues satisfying $\sum_{i=1}^{\infty} \eta_i < \infty$ and $f_i(\cdot)$ are orthonormal and square integratable eigenfunctions of the covariance function $C(\cdot, \cdot)$,

$$\int_a^b \; f_i(t) \; f_j(t) \; dt = \delta_{ij} \;\; (\delta_{ij} = 1 \text{ if } i = j, \text{and } 0 \text{ otherwise}).$$

In general, the leading terms in Mercer's theorem capture the main features of the covariance function because only the first few eigenvalues ($\eta_1 \geq \eta_2 \geq \ldots$) of the expansion make a significant impact on the true covariance function. So that true covariance function can be approximated using truncated covariance function. Truncated expansion yields a reasonable reduced rank approximation of the covariance function as follows

$$C_N(t,s) = \sum_{i=1}^{N} \eta_i \; f_i(t) \; f_i(s), \quad a \leq t, s \leq b. \tag{3.1}$$

It can be shown that if $C(t,s)$ is continuous and bounded on $[a,b] \times [a,b]$, then $C_N(t,s) \to C(t,s)$ uniformly converge on $[a,b] \times [a,b]$ as $n \to \infty$.

By solving the following Fredlohom homogeneous integral equation of second order, we can find a set of orthonormal squared integrable eigenfunctions $f_i(\cdot)$ and a set of eigenvalues η_i of the covariance function,

$$\eta_i \; f_i(t) = \int_a^b C(t,s) f_i(s) \; ds, \; a \leq s \leq b. \tag{3.2}$$

Many of the covariance functions may not have closed-form analytical solutions. It is very challenging to solve the above integral equation analytically, especially when the covariance function is defined over a complex domain. This strain motivated us to fabricate an algorithm to find solutions numerically. [WS01] used the Nyström method to solve the integral equations. However, in the next section, the Wavelet-Galerkin method is implemented to obtain accurate numerical solutions to the Fredlohom homogeneous integral equation.

3.2.2 Wavelet-Galerkin Method

Intuitively, a function is said to be inhomogeneous if the function has a sharp jump at some point but is very smooth otherwise. Approximating these kind of functions can be complicated using commonly use trigonometrics and polynomials bases [Dau92]. However, one way to estimate inhomogeneous functions is to use multiresolutional local bisquare wavelet bases. In this work we start with a particular wavelet, namely, Haar wavelets. Haar wavelets are the simplest basis form of Daubechies family, see [Nie13] for more details. The Haar wavelets defined on [a,b] are given as follows:

$$\psi(t) = \begin{cases} 1, & a \leq t < \frac{a+b}{2}; \\ -1, & \frac{a+b}{2} \leq t < b; \\ 0, & \text{otherwise.} \end{cases}$$

Haar wavelets are localized but they are not smooth. A family of orthogonal Haar wavelets can be constructed by shifting and scaling the mother wavelets:

$$\psi_{j,k}(t) = \alpha_j \psi(2^j t - k), \quad \text{for } j, k \in \mathbb{Z},$$

where j and k are dilation and transnational positive integer constants respectively, and α_j is the amplitude of the function. We can construct orthogonal functions by

letting $\alpha_j = 2^{\frac{j}{2}}$. Alternatively, the above expression can also be represented as:

$$\psi_{j,k}(t) = \begin{cases} 1, & (a+k)2^{-j} < t < (a+b)2^{j-1} + k2^{-j}, \\ -1, & (a+b)2^{j-1} + k2^{-j} \leq t < (b+k)2^{-j}, \\ 0, & \text{otherwise.} \end{cases}$$

One can check the orthogonality of the Haar wavelet bases by deriving

$$\int_a^b \psi_{j,k}(t)\psi_{m,n}(t)\, dt = \frac{(b-a)}{2^j}\delta_{j,m}.$$

where $\delta_{j,m} = 1$ if $j = m$ and 0 otherwise. For simplicity, let $\psi_0(t) = 1$ and $\psi_i(t) = \psi_{j,k}(t)$ and $i = 2^j + k; k = 0, 1, \ldots, 2^j - 1; j = 0, 1, \ldots, m-1$, where m is the maximum wavelet level.

According to the Galerkin method, any function (in our case the k^{th} eigenfunction of the covariance function $C_N(\cdot,\cdot)$) can be approximated through a linear combination of the finite number of Haar wavelets bases which is given by

$$f_k(t) = \sum_{i=0}^{N-1} d_{k,i}\psi_i(t) = \Psi(\vec{t})^\intercal \vec{D_v}. \tag{3.3}$$

Here $d_{v,i}$ are the wavelet coefficients and

$$\Psi(\vec{t})^\intercal = (\psi_1(t), \psi_2(t), \ldots, \psi_{N-1}(t)) \text{ with } N = 2^m,$$
$$\vec{D_v^\intercal} = (d_{v,1}, d_{v,2}, \ldots, d_{v,N-1}).$$

From equations (3.1) and (3.3) we have

$$C_N(t,s) = \Psi(\vec{t})^\intercal D^\intercal \Lambda D \Psi(\vec{s}), \tag{3.4}$$

where $D^\intercal = (\vec{D_1^\intercal}, \vec{D_2^\intercal}, \ldots, \vec{D_N^\intercal})$ and

$$\Lambda = \begin{pmatrix} \eta_1 & 0 & \ldots & 0 & 0 \\ 0 & \eta_2 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & 0 & \eta_N \end{pmatrix}$$

is a diagonal matrix. From (3.4), it is sufficient to find D and Λ. We implement the process with following two steps.

(i) Let $A = D^T \Lambda D$. Therefore, following [New93]

$$C_N(t,s) = \Psi(\vec{t})^\intercal A \vec{\Psi}(s) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} A_{ij} \psi_i(t) \psi_j(s), \tag{3.5}$$

where $A = (A_{ij})$ and $A_{ij} = \frac{1}{h_i h_j} \int_a^b \int_a^b C_N(t,s) \psi_i(t) \psi_j(s)\, dtds$.

A_{ij} is a $N \times N$ matrix which is two-dimensional wavelet discrete transform of the matrix $C_N(t,s)$. The algorithm to find A using the In-place fast Haar transform can be found in Algorithm 3.1 (Pseudo-code, Page 40).

(ii) Note that $A = D^T \Lambda D$ is not a singular value decomposition (SVD) for A since D may not be an orthogonal matrix. Therefore, we calculate them through the

Fredlohom homogeneous equation with $C(t,s)$ replaced with $C_N(t,s)$.

$$\int_a^b C_N(t,s) f_v(t) dt = \eta_v f_v(s).$$

From (3.3) and (3.4), we have

$$\Psi(\vec{s})^\intercal A \int_a^b \Psi(\vec{t})\Psi(\vec{t})^\intercal \vec{D}_v dt = \eta_k \Psi(\vec{s})^\intercal \vec{D}_v.$$

Note that the following orthogonality for $\Psi(t)$ holds.

$$\int_a^b \Psi(\vec{t})\Psi(\vec{t})^\intercal dt = H,$$

where

$$H = \begin{pmatrix} h_1 & 0 & \dots & 0 & 0 \\ 0 & h_2 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & h_N \end{pmatrix},$$

is a diagonal matrix with $h_i = \frac{(b-a)}{2^j}, i = 2^j + k;\ k = 0, 1, \dots, 2^j - 1;\ j = 0, 1, \dots, m-1$. We have

$$\Psi(\vec{s})^\intercal A H \vec{D}_v = \eta_v \Psi(\vec{s})^\intercal \vec{D}_v,$$

implying that

$$A H \vec{D}_v = \eta_v \vec{D}_v, \tag{3.6}$$

after considering the coefficients of both sides for $\Psi(s)$. Let $\hat{A} = H^{1/2}AH^{1/2}$ and $\hat{D_v} = H^{1/2}D_v$. Equation (3.6) becomes

$$\hat{A}\vec{\hat{D}}_v = \eta_k \vec{\hat{D}}_v. \tag{3.7}$$

Now we can see that finding Λ and D is equivalent to solving the eigenvalue problem (3.7) with the SVD method, with $\vec{\hat{D}}_v = H^{-1/2}\vec{D}_v$. Therefore, the v^{th} eigenfunction can be obtained by

$$f_v(t) = \Psi(t)^\intercal H^{-1/2}\vec{\hat{D}}_v.$$

Note that above algorithm is a modified version of the algorithm given in [PHQ02]. Now we illustrate the above procedure with a simple example. First we introduce the in-place transform. In the literature, there are two ways of calculating two-dimensional discrete wavelet transformation; one is called in-place transform and the other one is ordered transform. Although both methods give the same result, the in-place method is used in our study because it has a lower number of operations count than of the ordered method. Here we demonstrate how an in-place algorithm can be performed for an array of size 8 (one dimensional).

Initialization $\vec{s}^{(3)} = (s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8)$

First sweep

$$\begin{aligned}\vec{s}^{(2)} &= \left(\frac{s_1+s_2}{2}, \frac{s_1-s_2}{2}, \frac{s_3+s_4}{2}, \frac{s_3-s_4}{2}, \frac{s_5+s_6}{2}, \frac{s_5-s_6}{2}, \frac{s_7+s_8}{2}, \frac{s_7-s_8}{2}\right) \\ &= (a_1^{(2)}, c_1^{(2)}, a_2^{(2)}, c_2^{(2)}, a_3^{(2)}, c_3^{(2)}, a_4^{(2)}, c_4^{(2)}).\end{aligned}$$

Second sweep

$$\begin{aligned}\vec{s}^{(1)} &= \left(\frac{a_1^{(2)}+a_2^{(2)}}{2}, c_1^{(2)}, \frac{a_1^{(2)}-a_2^{(2)}}{2}, c_2^{(2)}, \frac{a_3^{(2)}+a_4^{(2)}}{2}, c_3^{(2)}, \frac{a_3^{(2)}-a_4^{(2)}}{2}, c_4^{(2)}\right) \\ &= (a_1^{(1)}, c_1^{(2)}, c_1^{(1)}, c_2^{(2)}, a_3^{(2)}, c_3^{(2)}, c_2^{(1)}, c_4^{(2)}).\end{aligned}$$

Last sweep

$$\begin{aligned}\vec{s}^{(0} &= (\frac{a_1^{(1)}+a_3^{(2)}}{2}, c_1^{(2)}, c_1^{(1)}, c_2^{(2)}, \frac{a_1^{(1)}-a_3^{(2)}}{2}, c_3^{(2)}, c_2^{(1)}, c_4^{(2)}) \\ &= (a_1^{(1)}, c_1^{(2)}, c_1^{(1)}, c_2^{(2)}, c_1^{(0)}, c_3^{(2)}, c_2^{(1)}, c_4^{(2)}).\end{aligned}$$

Now we consider the following example to illustrate our two dimensional Haar-Wavelet discrete transformation (2WDT) method.

Example 3.1: We consider a stationary covariance function $C(t,s) = e^{-|t-s|},\ 0 \leq t,s\ \leq \pi$ evaluated at $t,s = 0.392, 1.178, 1.963,$ and 2.748, which gives the following matrix

$$B(t,s) = \begin{bmatrix} 1 & 0.455 & 0.208 & 0.095 \\ 0.456 & 1 & 0.456 & 0.208 \\ 0.208 & 0.456 & 1 & 0.456 \\ 0.095 & 0.208 & 0.456 & 1 \end{bmatrix}.$$

We first the apply in-place one-dimensional wavelet transformation on each row of $B(t,s)$ to yield

$$B_{\text{new}}(t,s) = \begin{bmatrix} 0.439 & 0.288 & 0.272 & 0.056 \\ 0.529 & 0.198 & -0.272 & 0.124 \\ 0.529 & -0.198 & -0.124 & 0.272 \\ 0.439 & -0.288 & -0.056 & -0.272 \end{bmatrix}.$$

First take the transpose of the above matrix, then apply the in-place one-dimensional Wavelet transformation on rows again to obtain

$$A = \begin{bmatrix} 0.485 & 0 & -0.045 & 0.045 \\ 0 & 0.243 & 0.045 & 0.045 \\ -0.405 & 0.045 & 0.272 & -0.034 \\ 0.045 & 0.045 & -0.337 & 0.272 \end{bmatrix}.$$

Then $\hat{A}$ matrix can be derived by $\hat{A} = H^{1/2}AH^{1/2}$,

$$\hat{A} = \begin{bmatrix} 1.523 & 0 & -0.100 & -0.100 \\ 0 & 0.764 & 0.100 & 0.100 \\ -0.100 & 0.100 & 0.427 & -0.052 \\ 0.100 & 0.100 & -0.052 & 0.427 \end{bmatrix}, H = \begin{bmatrix} \pi & 0 & 0 & 0 \\ 0 & \pi & 0 & 0 \\ 0 & 0 & \frac{\pi}{2} & 0 \\ 0 & 0 & 0 & \frac{\pi}{2} \end{bmatrix}.$$

Then we solve the eigenvalue problem $\hat{A}\vec{\hat{D}}_v = \eta_v\vec{\hat{D}}_v$ to find $\Lambda = diag(\eta_v, v = 1, 2, \ldots, N)$.

$$\Lambda = \begin{bmatrix} 1.541 & 0 & 0 & 0 \\ 0 & 0.810 & 0 & 0 \\ 0 & 0 & 0.461 & 0 \\ 0 & 0 & 0 & 0.328 \end{bmatrix}, \hat{D} = \begin{bmatrix} 0.992 & 0 & 0.132 & 0 \\ 0 & -0.951 & 0 & -0.310 \\ -0.094 & -0.219 & 0.701 & 0.672 \\ 0.094 & -0.219 & -0.701 & 0.672 \end{bmatrix}.$$

Next we calculate, $D = H^{-1/2}\hat{D} = (\vec{D_1^\intercal}, \vec{D_2^\intercal}, \vec{D_3^\intercal}, \vec{D_4^\intercal})$.
Here, $\vec{D_1^\intercal} = (0.559, 0, -0.074, 0.074), \vec{D_2^\intercal} = (0, -0.536, -0.175, -0.175)$,
$\vec{D_3^\intercal} = (0.074, 0, 0.559, -0.559)$ and $\vec{D_4^\intercal} = (0, -0.175, 0.536, 0.536)$.

Now we recover some selected entities of $B(t, s)$ matrix using the expression given by (3.4), when $\vec{\Psi^\intercal}(0.392) = (1, 1, 1, 0), \vec{\Psi^\intercal}(1.178) = (1, 1, -1, 0), \vec{\Psi^\intercal}(1.963) = (1, -1, 0, 1)$,

$\vec{\Psi}^{\intercal}(2.748) = (1, -1, 0, -1)$.

$$B_{11}(0.392, 0.392) = \vec{\Psi}(0.392)^{\intercal} D^{\intercal} \Lambda D \vec{\Psi}(0.392) = 1.032,$$
$$B_{23}(1.178, 1.963) = \vec{\Psi}(1.178)^{\intercal} D^{\intercal} \Lambda D \vec{\Psi}(1.963) = 0.488,$$
$$B_{34}(1.963, 2.748) = \vec{\Psi}(1.963)^{\intercal} D^{\intercal} \Lambda D \vec{\Psi}(2.748) = 0.488.$$

Please note that obtained values are deviated from the actual values due to the impact of using a low number of bases in the calculations. However, if we utilize more bases (dimension), for instance, when $N = 128$ we can accurately recover all variance-covariance matrix entries.

Algorithm 3.1 (finding 2DWT) is computationally very efficient because the Mallat's tree [PHQ02] is the core concept of the DWT. The operation count is approximately N for 2DWT while the Fast Fourier Transformation (FFT) consumes $N \log_2 N$ operations. FFT only works for trigonometric functions which would produce a dense covariance matrix [New93].

Algorithm 3.1 (Pseudo-code)

- **Step 1** : First, an one-dimensional In-place wavelet transform is performed on each row (explain below) of the covariance matrix B which is evaluated over $N \times N$ grid.

  $B = \{b^1, b^2, \ldots, b^l, \ldots, b^N\}^\intercal$, where $b^l_{1\times N}$ represent l^{th} row of the matrix $[B]_{N\times N}$.

  $\overrightarrow{b^l} = \{b^l_1, b^l_2, \ldots, b^l_N\}$

  START

  $I = 1, J = 2$

  N=2^n where n is the maximum wavelet level

  FOR L = 1, ..., n do

  $M = M/2$

  FOR K = $0, \ldots, M-1$ do

  $a^{n-l}_k = (b^l[J.K] + b^l[J.K + I])/2$

  $c^{n-l}_k = (b^l[J.K] - b^l[J.K + I])/2$

  $b^l[J.K] = a^{(n-l)}_k$, $b^l[J.K] = c^{(n-l)}_k$

  END $I = J$, $J = 2 * J$

  END STOP

  RESULTS : $\overrightarrow{b^l_{\text{new}}} = (a^{(n-n)}_0, c^{(n-1)}_0, c^{(n-2)}_0, c^{(n-1)}_0, \ldots)$.

- **Step 2** : Restore outputs of the above step in a new matrix,

  $B_{\text{new}} = \{b^1_{\text{new}}, b^2_{\text{new}}, \ldots, b^l_{\text{new}}, \ldots, b^N_{\text{new}}\}^\intercal$.

- **Step 3** : Take the transpose of the step 2 matrix, $B_{\text{trans}} = B^\intercal_{\text{new}}$.

- **Step 4** : Then apply one dimensional in-place Haar wavelet transform again (Step 1) for the each row of the matrix B_{trans} to obtain A.

3.3 Comparison of Analytical and Numerical Results of the Wavelet-Galerkin Method

This section focuses on applying our Wavelet-Galerkin approach to approximate true eigenvalues and eigenfunctions of a covariance function. The approach will be evaluated by comparing approximations on each of the real-valued stationary and non-stationary covariance functions with true function values using numerical examples. The approximation to a complex-valued non-stationary covariance function will be given in Section 3.4.

3.3.1 Ornstein-Uhlenbeck Stationary Covariance Model

Although stationary processes are not common in spatial random fields, they are simple and convenient in terms of modeling and making predictions. Here we consider one of the well-known stationary Gauss-Markov processes in many random fields- the Ornstein-Uhlenbeck process-which has the following form of the covariance structure

$$C(t,s) = \frac{\rho^2}{2\beta} \exp(-\beta \mid t-s \mid),\ s,t \in [0,1]. \tag{3.8}$$

For ease of manual computation, we set spatial correlation parameter ($\beta = 1$) and variance component ($\rho = \sqrt{2}$) to covert covariance function into exponential covariance function $C(t,s) = \exp(- \mid t-s \mid)$. In order to find therotical eigenvalue-eigenfunction pairs we need to solve the Fredholm integral equation manually,

$$\int_0^1 \exp(-|t-s|) f_i(s)\, ds = \eta_i f_i(t) \tag{3.9}$$

where $f_i(\cdot)$ is an orthogonal eigenfunction and η_i is the corresponding eigenvalue of the exponential covariance function (3.8). After integrating (3.9) twice with respect to t we have

$$f''(t) + \omega_i^2 f(t) = 0, \text{ where } \omega_i^2 = \frac{2 - \eta_i}{\eta_i}. \tag{3.10}$$

Since (3.10) has the structure of second order differential equation, a general solution to the eigenfunction can be obtained.

$$f_i(t) = A_i \cos(\omega_i t) + B_i \sin(\omega_i t) \tag{3.11}$$

with the boundary conditions

$$f_i'(0) - f_i(0) = 0,$$
$$f_i'(\pi) + f_i(\pi) = 0.$$

Applying basic mathematics to above equations, ω_i values can be obtained by solving

$$\cot(\omega_i) = \frac{\omega_i^2 - 1}{2\omega_i}. \tag{3.12}$$

The Newton-Raphson method is the most suitable way to find optimal values of ω_i. Once we find out ω_i value, then the corresponding eigenvalue can be easily found from $\eta_i = \frac{2}{1+\omega_i^2}$. In addition, we can determine the unknown terms A_i and B_i

by considering the orthogonality of eigenfunctions. Thus, we have

$$A_i = \sqrt{\frac{2\omega_i^2}{3+\omega_i^2}}, B_i = \frac{A_i}{\omega_i} = \sqrt{\frac{2}{3+\omega_i^2}}.$$

Therefore, when $i \geq 1$, the general solution for the eigenfunction is given by

$$f_i(t) = \sqrt{\frac{2\omega_i^2}{3+\omega_i^2}}\cos(\omega_i t) + \sqrt{\frac{2}{3+\omega_i^2}}\sin(\omega_i t).$$

We now apply the Wavelet-Galerkin procedure that was proposed in Section 3.2 to find the approximate eigenvalues and eigenfunctions of $C(t,s)$. As a next step, we will explore how well numerical eigenvalues of the exponential covariance function approximate analytic ones for different numbers of Haar wavelets basis. Here, N denotes the number of wavelet bases we considered in the algorithm.

Table 2. Comparison between analytical and numerical eigenvalues for the exponential covariance function.

Analytical	**N=8**	**N=16**	**N=32**	**N=64**	**N=128**	**N=256**
0.73881	0.74812	0.73981	0.73906	0.73887	0.73882	0.73881
0.13800	0.14123	0.13880	0.13820	0.13805	0.13802	0.13800
0.04508	0.04800	0.04580	0.04526	0.04513	0.04509	0.04508
0.02133	0.02424	0.02201	0.02149	0.02137	0.02132	0.02133
0.01227	0.01531	0.01296	0.01244	0.01232	0.01229	0.01227

One big advantage of Wavelet-Galerkin method is the ability to compute a large number of eigen-solutions accurately and rapidly. Table 2 shows that when the number of wavelets basis used increases, the numerical eigenvalues obtained from the

Wavelet-Galerkin method are getting closer to analytical counterparts. Furthermore, we can notice that high order eigenvalues are significantly small compared to first few eigenvalues. Next let's focus on how well this method approximates eigenfunction when a finite number ($N = 256$) of wavelet bases are used.

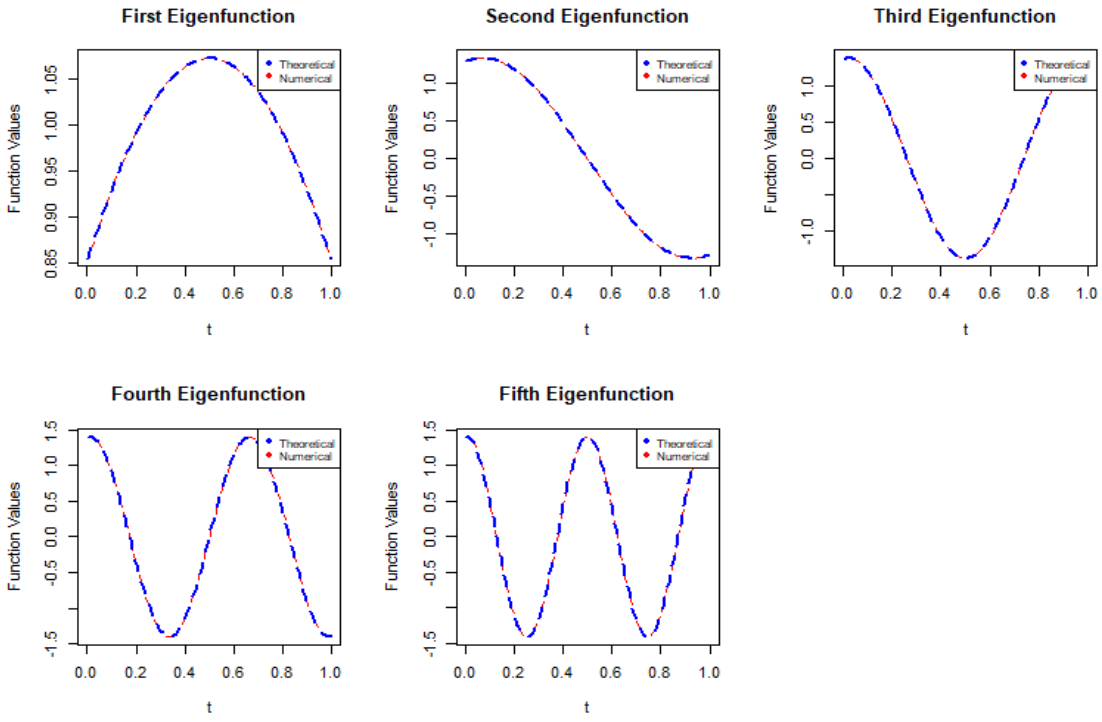

Figure 3. The comparison of first five eigenfunction of analytical and numerical stationary covariance function.

We can clearly notice that from Figure 3 Wavelet (Haar)-Galerkin method performs quite well (both theoretical and numerical eigenfunctions). We can also verify in a similar way that other analytical eigenfunctions are accurately approximated

by the corresponding numerical eigenfunctions. For future approximations, $N = 256$ wavelets will be considered as the benchmark.

[NWR02] proposed a technique to approximate one-dimensional stationary covariance models with continuous wavelets (the W-transform). However, in this section we demonstrate how well multiresolutional covariance model can be approximated (3.8) with square-shaped discrete Haar wavelets.

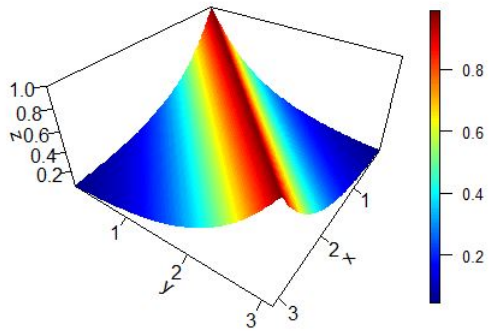

Figure 4. True covariance matrix values evaluated at 128 equally spaced grid points on [0,π]

Figures 4 and 5 represent the reconstruction of the covariance reconstruction using Haar wavelets. Obviously there is no significant difference between true and truncated Wavelet-Galerkin approximation, in particular. The approximation performs better in the diagonal, which is corresponding to the peak.

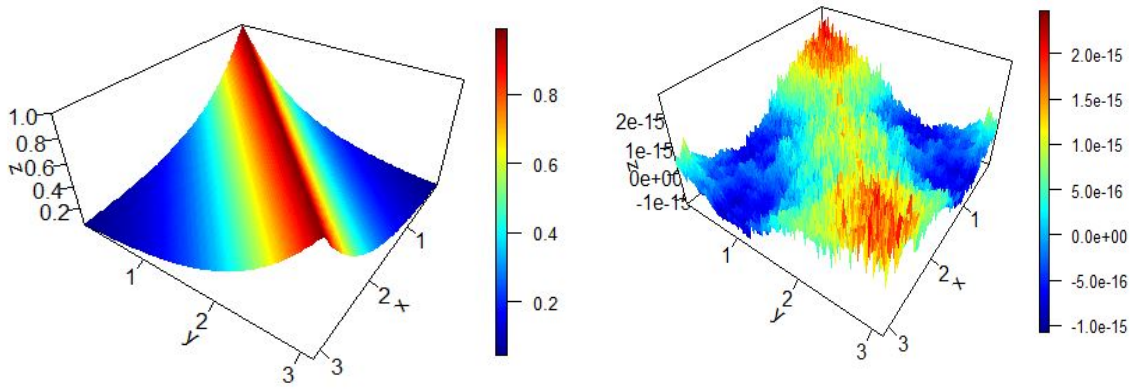

Figure 5. Haar wavelet approximation of a two-dimensional stationary covariance function (left) and the difference between true and approximate stationary covariance functions (right).

3.3.2 Non-Stationary Parametric Covariance Model

It is important to approximate non-stationary covariance functions as in many environmental and geo-physical applications one often sees non-stationary covariance structures. In this section, we will demonstrate how our Wavelet-Galerkin method and 2DWT method can be used to approximate non-stationary covariance functions. Here we use the non-stationary covariance function that was proposed by [VWZ19], which is given by

$$C(x,y) = C_2 - C(x) - C(y) + C(x-y) \;\text{ with }\; C_2 \geq \int_{-\infty}^{\infty} f(\omega)d\omega > 0.$$

where $f(\omega) \geq 0$ is the spectral density of a stationary covariance function $C(\cdot)$ through Bochner's theorem

$$C(x) = \int_{-\infty}^{\infty} e^{-ix\omega} f(\omega) d\omega.$$

We consider the stationary covariance function

$$C(x) = Ce^{-a|x|},$$

which implies

$$C(x, y) = C_1(C_2 - e^{-a|x|} - e^{-a|y|} + e^{-a|x-y|}),\ C_1 > 0,\ x, y \in \mathbb{R}.$$

In our study, the covariance functions of the following structure is being used (and for notation simplicity):

$$C(t, s) = (1 - e^{-|t|} - e^{-|s|} + e^{-|t-s|}) \text{ where } s, t \in [0, \pi]. \tag{3.13}$$

A method of finding analytical eigenvalues and eigenfunctions for $C(t, s)$ has been derived in Appendix. Here we provide a summary of the results.

Proposition 3.1. *The eigenvalues and eigenfunctions of the covariance function (*3.13*) are given by the following.*

(1) *When $\eta_i < 2$, η_i satisfies the following equation*

$$\left(\frac{2-\eta_i}{\eta_i}\right)^{\frac{1}{2}}(\eta_i - 1 - \pi)\sin(\sqrt{\frac{2-\eta_i}{\eta_i}}\pi) + \frac{1}{\eta_i}(-\eta_i^2 + \pi\eta_i + 2\eta_i - 2\pi - 4) \cos(\sqrt{\frac{2-\eta_i}{\eta_i}}\pi) + 2 = 0. \quad (3.14)$$

The corresponding eigenfunction is given by

$$f_i(t) = -C_{i3} + k_i C_{i3}\sin(\sqrt{\omega_i}t) + C_{i3}\cos(\sqrt{\omega}_i t),$$

where

$$k_i = \frac{\left(\omega_i + \sqrt{\omega_i}\sin(\sqrt{\omega_i}\pi) - \cos(\sqrt{\omega_i}\pi) + 1\right)}{\left(\sqrt{\omega_i}\cos(\sqrt{\omega_i}\pi) + \sin(\sqrt{\omega_i}\pi)\right)} \ , \ \omega_i = \frac{2-\eta_i}{\eta_i},$$

and

$$C_{i3}^2 = \frac{1}{\int_0^\pi \left(-1 + k_i\sin(\sqrt{\omega}_i t) + \cos(\sqrt{\omega}_i t)\right)^2 dt}.$$

(2) *When $\eta_i \geq 2$, η_i satisfies the following equation*

$$\left(\frac{\eta_i - 2}{\eta_i}\right)^{\frac{1}{2}}(\eta_i - 1 - \pi)\sinh(\sqrt{\frac{\eta_i-2}{\eta_i}}\pi) + \frac{1}{\eta_i}(\eta_i^2 - \pi\eta_i - 2\eta_i + 2\pi + 4) \cosh(\sqrt{\frac{\eta_i-2}{\eta_i}}\pi) - 2 = 0. \quad (3.15)$$

The corresponding eigenfunction is given by

$$g_i(t) = -A_{i3} + k_i' A_{i3}\ \sinh(\sqrt{\omega_i}t) + A_{i3}\cosh(\sqrt{\omega}_i t),$$

where

$$k_i' = \frac{(\omega_i - \sqrt{-\omega_i}\sinh(\sqrt{-\omega_i}\pi) - \cosh(\sqrt{-\omega_i}\pi) + 1))}{(\sqrt{-\omega_i}\cosh(\sqrt{-\omega_i}\pi) + \sinh(\sqrt{-\omega_i}\pi))} \ , \ \omega_i = \frac{\eta_i - 2}{\eta_i},$$

and

$$A_{i3}^2 = \frac{1}{\int_0^\pi \left(-1 + k_i' \sinh(\sqrt{-\omega}_i t) + \cosh(\sqrt{-\omega}_i t)\right)^2 dt}.$$

It would not be possible to obtain the closed-form analytic values for both eigenvalues and eigenfunctions. Therefore, we apply the Newton-Raphson method to obtain the first four eigenvalues and hence corresponding eigenfunctions.

Table 3. First four eigenvalues and eigenfunctions of the nonstaionary function

Eigenvalue	Eigenfunctions
2.947779	$g_i(t) = -1.384 - 1.203\sinh(\sqrt{0.321}t) + 1.384\cosh(\sqrt{0.321}t)$
0.5310727	$f_1(t) = -0.024 - 0.765\ \sin(\sqrt{2.766}t) + 0.024\ \cos(\sqrt{2.766}t)$
0.2625776	$f_2(t) = -0.030 + 0.791\ \sin(\sqrt{6.616}t) + 0.030\ \cos(\sqrt{6.617}t)$
0.144951	$f_3(t) = -0.009 - 0.789\ \sin(\sqrt{12.798}t) + 0.009\ \cos(\sqrt{12.798}t)$

Now we apply our Wavelet-Galerkin method that was given in Section 3.2 to obtain the corresponding eigenvalues and eigenfunctions. We present the following graph to demonstrate the accuracy of our approximation for the first five eigenfunctions. From Figure 6, we can see that difference between numerical eigenfunctions (red line) the theoretical eigenfunctions (blue line) values are very small for all $\phi = t$ values.

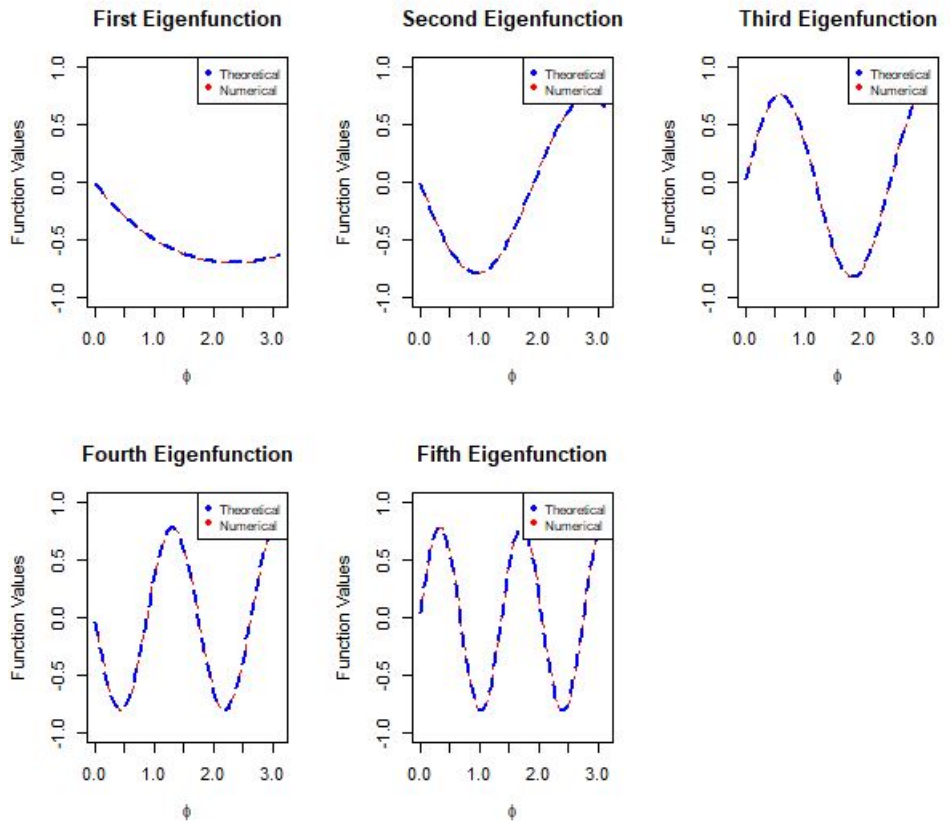

Figure 6. The comparison of the first five eigenfunctions between analytical and numerical solutions.

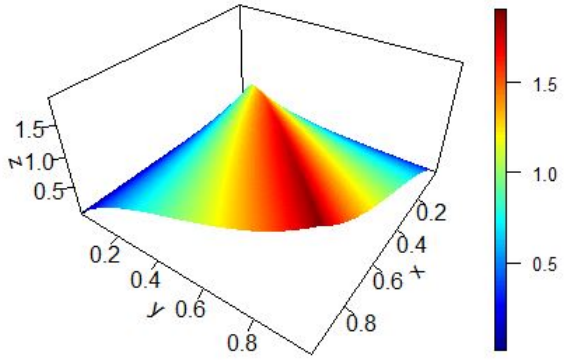

Figure 7. True covariance function values evaluated at 128 equally spaced grid points on $[0,\pi]$.

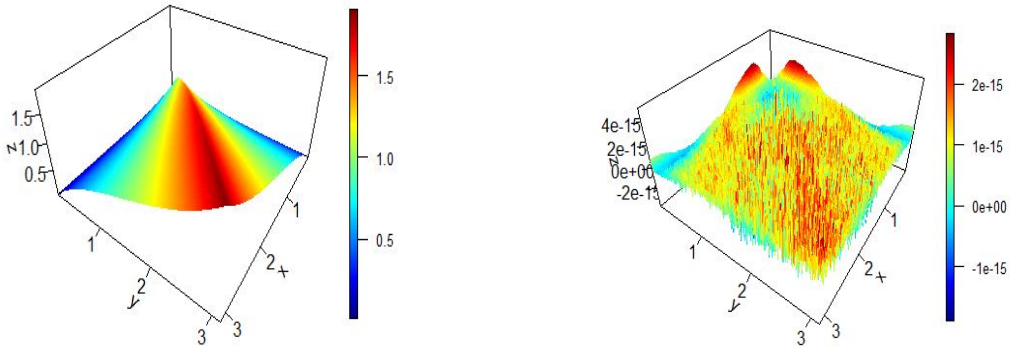

Figure 8. Left graph shows Wavelet-Galerkin approximation of real-valued nonstationary covariance function. The difference between true and approximate nonstationary covariance function is represented on the right graph.

The results of the Figures 7 and 8 illustrate that Wavelet-Galerkin method can be applied to approximate nonstastionary covariance functions.

3.4 Approximating Complex-valued Covariance Functions using the Wavelet-Galerkin Method

Investigating structures and properties of complex-valued covariance (kernel) functions have received special attention in many research fields, for instance, in engineering, complex-valued kernels for complex-values signals [TSRF15] and in spatial statistics, complex-valued covariance functions for axially symmetric random processes on the sphere [HZR12].

In this section, we consider a complex-valued covariance function $C(t,s)$ with $t,s \in [a,b]$. We denote $C(t,s) = C_R(t,s) + iC_I(t,s)$, where $C_R(t,s)$ is the real part of $C(s,t)$ and $C_I(s,t)$ is the imaginary part. Note that $C(s,t)$ is positive definite, that is, for any n, and any complex constants $C_1, C_2, \ldots, C_n$, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} C_i \bar{C}_j C(t,s) \geq 0, \quad \forall \;\; s,t \in [a,b].$$

According to Mercer's theorem, a bounded continuous complex-valued covariance function can be written as follows

$$C(t,s) = \sum_{i=1}^{\infty} \eta_i \; f_i(t) \; \overline{f_i(s)}, \quad a \leq t,s \leq b.$$

Here $\overline{f_i(s)}$ denotes the complex conjugate of $f_i(s)$. Now we consider a truncated function $C_N(t,s) = \sum_{i=1}^{N} \eta_i f_i(t)\overline{f_i(s)}$ to approximate $C(t,s)$. First we have the following results.

Proposition 3.2. *The function $C_N(t,s)$, $0 \le t,s \le b$ is Hermitian and positive definite.*

Proof: For any $0 \le t,s \le b$,

$$\begin{aligned}\overline{C_N(t,s)} &= \overline{\sum_{i=1}^{N} \eta_i f_i(t)\overline{f_i(s)}} \\ &= \sum_{i=1}^{N} \eta_i f_i(s)\overline{f_i(t)} \\ &= C_N(s,t).\end{aligned}$$

For the positive definiteness, we take any n, any complex constants $a_i, i = 1,2,\ldots,n$, and any $0 \le t_1, t_2, \ldots, t_n \le b$,

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}C_N(t_i,t_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}\sum_{k=1}^{N}\eta_k f_k(t_i)\overline{f_k(t_j)} \\ &= \sum_{k=1}^{N}\eta_k\sum_{i=1}^{n}\sum_{j=1}^{n} a_i f_k(t_i)\overline{a_j f_k(t_j)} \\ &= \sum_{k=1}^{N}\eta_k\left[\sum_{i=1}^{n} a_i f_k(t_i)\right]^2 \ge 0,\end{aligned}$$

concluding the proof.

Now we mimic the approach used in Section 3.3. The k^{th} eigenfunction of a truncated covariance function $C_N(t,s) = \sum_{i=1}^{N} \eta_i \; f_i(t) \; \overline{f_i(s)}$ can be written as,

$$f_k(t) = \sum_{i=0}^{N-1} d_i^k \psi_i(t) = \vec{\Psi}^{\intercal}(t)\vec{D_k}, \;\; t \in \mathbb{R}, \; d_i^k \in \mathbb{C} \text{ and } k \in 1, \ldots, N-1, \tag{3.16}$$

where, $\vec{D_k}^{\intercal} = (d_1^k, d_2^k, \ldots, d_N^k) \; \in \mathbb{C}^N$ is the set of complex-valued wavelet coefficients and $\vec{\Psi}^{\intercal}(t) = (\psi_1(t), \psi_2(t), \ldots, \psi_N(t)) \in \mathbb{R}^N$ real-valued Haar wavelets. Hence, a truncated complex-valued covariance function $C_N(t,s)$ can be written as;

$$C_N(t,s) = \vec{\Psi^{\intercal}}(t) A \Psi(\vec{s}),$$

where A, a possibly complex-valued matrix, can be obtained via Algorithm 3.1 in Section 3.2 as

$$A_{ij} = \frac{1}{h_i h_j} \int_a^b \int_a^b C_N(t,s) \psi_i(t) \psi_j(s) dt ds. \tag{3.17}$$

Proposition 3.3. *$A = (A_{ij})$ with A_{ij} given by (3.17) is a Hermitian, positive definite matrix.*

Proof: To prove that A is Hermitian, it is sufficient to show that

$$\overline{A_{ij}} = A_{ji}.$$

Now,

$$\begin{aligned}\overline{A_{ij}} &= \overline{\frac{1}{h_i h_j}\int_a^b\int_a^b C_N(t,s)\Psi_i(t)\Psi_j(s)dtds} \\ &= \frac{1}{h_i h_j}\int_a^b\int_a^b \overline{C_N(t,s)}\Psi_i(t)\Psi_j(s)dtds \\ &= \frac{1}{h_i h_j}\int_a^b\int_a^b C_N(t,s)\Psi_i(t)\Psi_j(s)dtds = A_{ji}\end{aligned}$$

Next for any n and any complex constants $a_1, a_2, \ldots, a_n$, we want to prove that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}A_{ij} \geq 0.$$

Now,

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}A_{ij} &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}\frac{1}{h_i h_j}\int_a^b\int_a^b C_N(t,s)\Psi_i(t)\Psi_j(s)dtds \\ &= \int_a^b\int_a^b C_N(t,s)dtds\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}\frac{1}{h_i h_j}\Psi_i(t)\Psi_j(s) \\ &= \int_a^b\int_a^b C_N(t,s)dtds\sum_{i=1}^{n}\frac{a_i}{h_i}\Psi_i(t)\overline{\sum_{j=1}^{n}\frac{a_j}{h_j}\Psi_j(s)}.\end{aligned}$$

Let $\xi(t) = \sum_{i=1}^{n}\frac{a_i}{h_i}\Psi_i(t)$, then above equation becomes

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\bar{a_j}A_{ij} = \int_a^b\int_a^b C_N(t,s)\xi(t)\overline{\xi(s)}dtds.$$

It is now sufficient to prove that

$$\int_a^b\int_a^b C_N(t,s)\xi(t)\overline{\xi(s)}dtds \geq 0.$$

Now we consider the definition of the above double integral. Take the gridded region $(t_i, t_{i+1}) \times (s_j, s_{j+1})$ where $t_i = \frac{b-a}{n}(i-1) + a$ and $s_j = \frac{b-a}{n}(j-1) + a$.

Therefore,

$$\begin{aligned}\int_a^b \int_a^b C_N(t,s)\xi(t)\overline{\xi(s)}dtds &= \lim_{n\to\infty} \sum_{j=1}^{n}\sum_{i=1}^{n} C_N(t_i,s_j)\xi(t_i)\overline{\xi(s_j)}(\frac{b-a}{n})^2 \\ &= \lim_{n\to\infty} (\frac{b-a}{n})^2 \sum_{j=1}^{n}\sum_{i=1}^{n} C_N(t_i,s_j)\xi(t_i)\overline{\xi(s_j)} \geq 0,\end{aligned}$$

since $C_N(t,s)$ is positive definite from Proposition 3.2. This concludes the proof.

Now we illustrate the above calculations with an example.

Example 3.2. We consider the following covariance function:

$$\begin{aligned}C(t,s) &= e^{-|t-s|}e^{i(t-s)} = e^{-|t-s|}(\cos((t-s)) + i\sin((t-s))) \\ &= C_R(t,s) + iC_I(t,s), \;\; t.s \in (0,1).\end{aligned}$$

It is easy to see that $C(t,s)$ is a covariance function since $e^{-|t-s|}$ and $e^{i(t-s)}$ are covariance functions. We evaluate $C(t,s)$ at $t, s = 0.25, 0.75$ to obtain

$$C = \begin{bmatrix} 1 & 0.5322 - 0.2907i \\ 0.5322 + 0.2907i & 1 \end{bmatrix}.$$

We want to use the Wavelet-Galerkin method from Section 3.2 to approximate C. We first use equation (3.17) to calculate A analytically.

$$\begin{aligned}
A_{00} &= 2 * \int_0^{1/2} \int_0^{1/2} 1 \; dtds + 2 \int_0^{1/2} \int_{1/2}^{1} 0.5322 \; dtds \\
&= 0.7661, \\
A_{01} &= -i \int_0^{1/2} \int_{1/2}^{1} 0.2907 \; dtds - i \int_{1/2}^{1} \int_0^{1/2} 0.2907 \; dtds \\
&= -0.1454i = -\bar{A_{10}}, \\
A_{11} &= 2 * \int_0^{1/2} \int_0^{1/2} 1 \; dtds - \int_0^{1/2} \int_{1/2}^{1} -0.5322 \; dtds \\
&= 0.2339.
\end{aligned}$$

Hence, we have A given below

$$A = \begin{bmatrix} 0.7661 & -0.1454i \\ 0.1454i & 0.2339 \end{bmatrix}.$$

This matrix matches perfectly with the matrix we obtained from the 2WDT numerical method (Algorithm 3.1).

Now for the given A, we want to find the eigenvalues η_k and the complex-valued vector $\vec{D_k}$. We proceeed with the same steps as before, that is, from the Fredlohom homogeneous equation of second kind

$$\int_a^b C_N(t,s)\vec{\Psi^\intercal}(t)\vec{D_k}dt = \eta_k \vec{\Psi^\intercal}(s)\vec{D_k}.$$

Hence

$$\int_a^b \vec{\Psi}^\intercal(s) A \vec{\Psi}(t) \vec{\Psi}^\intercal(t) \vec{D}_k dt = \eta_k \vec{\Psi}^\intercal(s) \vec{D}_k.$$

Substituting $\hat{A} = H^{1/2} A H^{1/2}$ and $\vec{\hat{D}}_k = H^{1/2} \vec{D}_k$, we have

$$\hat{A} \vec{\hat{D}}_k = \eta_k \vec{\hat{D}}_k, \tag{3.18}$$

with

$$\int_a^b \vec{\Psi}(t) \vec{\Psi}^\intercal(t) dt = H,$$

where

$$H = \begin{pmatrix} h_1 & 0 & \dots & 0 & 0 \\ 0 & h_2 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & h_N \end{pmatrix} \text{ is a diagonal matrix with } h_i = \frac{(b-a)}{2^j},$$

$i = 2^j + k; k = 0, 1, \dots, 2^j - 1; j = 0, 1, \dots, m-1$, where m is the maximum wavelet level. We first present the following result.

Proposition 3.4. *$\hat{A}$ is Hermitian and positive definite.*

Proposition 3.4 can be easily obtained from Proposition 3.3. Now we explore the calculations of η_k and $\vec{\hat{D}_k}$. We write

$$\begin{aligned} \hat{A} &= Re(\hat{A}) + iIm(\hat{A}), \text{ and} \\ \vec{\hat{D}_k} &= \vec{d_{1,k}} + i\vec{d_{2,k}}. \end{aligned}$$

Then we have

$$(Re(\hat{A}) + iIm(\hat{A}))(\vec{d_{1,k}} + i\vec{d_{2,k}}) = \eta_k(\vec{d_{1,k}} + i\vec{d_{2,k}}),$$

which implies that

$$\begin{aligned} Re(\hat{A})\vec{d_{1,k}} - Im(\hat{A})\vec{d_{2,k}} &= \eta_k\vec{d_{1,k}}, \\ Im(\hat{A})\vec{d_{1,k}} + Re(\hat{A})\vec{d_{2,k}} &= \eta_k\vec{d_{2,k}}. \end{aligned}$$

The above system of equations becomes the following.

$$\begin{pmatrix} Re(\hat{A}) & -Im(\hat{A}) \\ Im(\hat{A}) & Re(\hat{A}) \end{pmatrix} \begin{pmatrix} \vec{d_{1,k}} \\ \vec{d_{2,k}} \end{pmatrix} = \eta_k \begin{pmatrix} \vec{d_{1,k}} \\ \vec{d_{2,k}} \end{pmatrix}. \tag{3.19}$$

This is the same system as that we saw in the equation (3.7). Hence, eigenvalues and eigenvectors can be determined by solving the equation (3.19). We first give the following proposition about the system (3.19).

Proposition 3.5. *[Gal08] Let*

$$U = \begin{pmatrix} Re(\hat{A}) & -Im(\hat{A}) \\ Im(\hat{A}) & Re(\hat{A}) \end{pmatrix}. \tag{3.20}$$

If λ is the eigenvalue of $\hat{A}$ with the corresponding eigenvector $\vec{v}$, then λ is the eigenvalue (with a repetition order of 2) of U and $q = (Re(\vec{v}), Im(\vec{v}))$ is the corresponding eigenvectors of U. On the other hand, if λ (with an order 2) is the eigenvalue of U with corresponding eigenvector $\vec{v} = (\vec{v_1}, \vec{v_2})^\intercal$ ($\vec{v_1} = n \times 1$ and $\vec{v_2} = n \times 1$), then $\vec{q} = \vec{v_1} + i\vec{v_2}$ is the eigtenvector of $\hat{A}$ corresponding to eigenvalue λ.

We know that from mercer's theorem $C_N(t, s)$ can be decomposed as follows

$$C_N(t, s) = F(t)^\intercal \Lambda F(s),$$

where $F(t)^\intercal = (f_1(t), f_2(t), \ldots, f_N(t))$ and Λ is a diagonal matrix with N number of eigenvalue entries. Therefore, by expanding equation (3.16), k^{th} complex-valued orthonormal eigenfunction can be calculated by $f_k(t) = \Psi(\vec{t})^\intercal H^{-1/2} \vec{\hat{D}}_k$ and the covariance function can be approximated by

$$C_N(t, s) = \Psi(\vec{t})^\intercal D^* \Lambda D \Psi(\vec{s}),$$

where D^* is the conjugate transpose of D. Now we revisit Example 3.2.

Example 3.2 (continued). $\hat{A}$ matrix can be derived by $\hat{A} = H^{1/2}AH^{1/2}$,

$$\hat{A} = \begin{bmatrix} 0.7661 & -0.1454i \\ 0.1454i & 0.2338 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then we solve eigenvalue problem $\hat{A}\vec{\hat{D}}_k = \eta_v \vec{\hat{D}}_k$ to find $\Lambda = diag(\eta_1, \eta_2)$ and $\vec{\hat{D}}_k$:

$$\Lambda = \begin{bmatrix} 0.8032 & 0 \\ 0 & 0.1967 \end{bmatrix}, \hat{D} = \begin{bmatrix} -0.968 & -0.247 \\ -0.247i & -0.968i \end{bmatrix}.$$

In this case both $\hat{D}$ and D are the same because H is an identity matrix. At last, we recover some selected entities of C matrix. When $\vec{\Psi}^{\intercal}(0.25) = (1, 1)$ and $\vec{\Psi}^{\intercal}(0.75) = (1, -1)$,

$$C_{11}(0.25, 0.25) = \vec{\Psi}^{\intercal}(0.25)D^*\Lambda D\vec{\Psi}(0.25) = 1,$$
$$C_{12}(0.25, 0.75) = \vec{\Psi}^{\intercal}(0.25)D^*\Lambda D\vec{\Psi}(0.75) = 0.5322 - 0.2908i.$$

We can see that these values are exactly equal to corresponding entries of C matrix.

Now we approximate a more general complex-valued covariance function given by $C_R(t, s) = (1 - e^{-|t|} - e^{-|s|} + e^{-|t-s|})\cos(t - s)$ and $C_I(t, s) = (1 - e^{-|t|} - e^{-|s|} + e^{-|t-s|})\sin(t - s)$, where $s, t \in [0, \pi]$. Here, we show the results of the wavelet approach when approximating the truncated series of $C_R(t, s)$ and $C_I(t, s)$ respectively. From the figures 9 and 10 given below, we can observe that the error between approximations and true values are very small in terms of their magnitude.

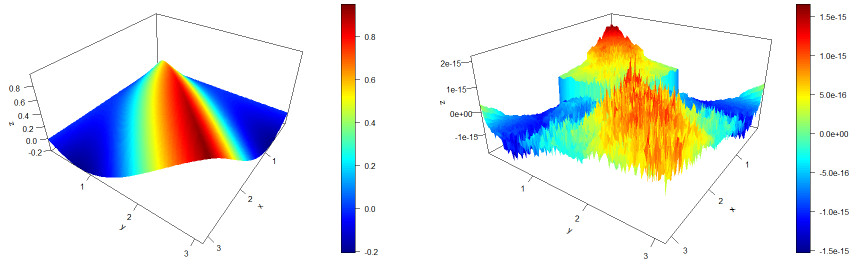

Figure 9. Approximated $C_R(t,s)$ using wavelet approach (left) and the difference between analytical and approximated $C_R(t,s)$ (right).

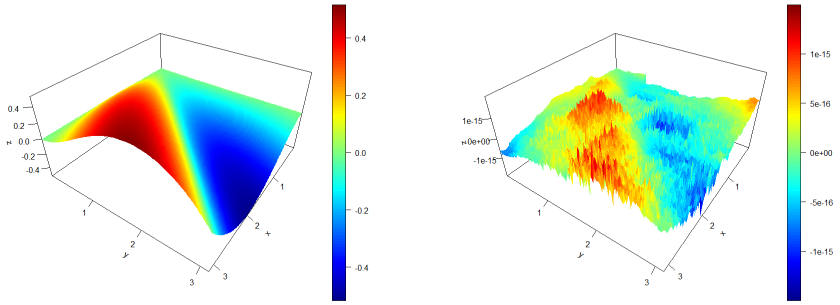

Figure 10. Approximated $C_I(t,s)$ using wavelet approach (left) and the difference between analytical and approximated $C_I(t,s)$ (right).

CHAPTER IV

KARHUNEN-LOÉVE EXPANSION FOR GAUSSIAN AXIALLY SYMMETRIC RANDOM PROCESSES ON THE SPHERE

When a random process is periodic, Fourier series would be the most appropriate representation to use. However, when the underlying random process is non-periodic, under which Fourier random coefficients can be correlated. In order to address this drawback, Karhunen [Kar46] and Loéve [Loe55] independently proposed a expansion which is the so-called "Karhunen-Loéve expansion" (or simply the K-L expansion).

One can consider the K-L expansion as a generalization (extension) of a Fourier series. Yet, there is a big difference in the representation of these two processes. The K-L expansion is a stochastic process which can be represented as with various types of orthogonal functions. However, in the Fourier series, the process can be approximated only by a summation of trigonometric functions.

[PHQ01] provided some discussions on Karhunen Loéve expansion. Their study mainly focuses on measuring and comparing the efficiency of the K-L expansion using various approaches. The K-L expansion has been applied to a wide range of areas such as functional data analysis [RS97], finance [Sch04], pattern recognition [KS90], and machine learning [RW05], etc.

In this chapter we apply K-L expansion to geostatistics, especially when the random process on the sphere is Gaussian axially symmetric. This chapter is organized as follows. In Section 1, the convergence of truncated K-L expansions for axially symmetric random processes is studied. In Section 2, an algorithm has proposed to generate axially symmetric random processes on the sphere using a truncated K-L expansion. In the last section, performance of our proposed method is illustrated through a simulation based data validation study.

4.1 A Study of Convergence for Truncated Karhunen-Loéve Expansion

In this section, we investigate the convergence and accuracy of the K-L expansion under the assumption of axially symmetric processes on the sphere. According to [PHQ01], there are many factors directly associated with the convergence rate of a random processes such as the structure of a covariance function, method of solving eigenvalue-eigenfunction pair, stationary criteria of the covariance function and the magnitude of the parameters in the covariance function as well as how a study of convergence is conducted. In our study, we focus on the complex form of parametric nonstationary covariance function on the sphere where eigen-solutions to that covariance function are determined numerically by the Wavelet-Garlekin method. The convergence study of a truncated process is important to evaluate the accuracy of the proposed random process generation method. Furthermore, the L^2 error of the approximation is calculated both analytically and numerically. The accuracy of our results are demonstrated via simulations.

As stated in [HZR12], let $\{X(P), P = (\lambda, \phi) \in \mathbb{S}^2\}$ be a zero-mean continuous real-valued Gaussian random process defined on a unit sphere with longitude $\lambda \in [0, 2\pi]$ and latitude $\phi \in [0, \pi]$. Then for each fixed point $P = (\lambda, \phi)$,

$$X(P) = X(\lambda, \phi) = \sum_{m=-\infty}^{\infty} W_m(\phi) e^{im\lambda}, \tag{4.1}$$

with $W_m(\phi) = \overline{W_m(\phi)}, m = 1, 2, \ldots$. Here, the $W_m(\phi)$ are independent Gaussian complex-valued circular symmetric random processes defined on $\phi \in [0, \pi]$ with

$$\mathrm{E}(W_m(\phi_P)\overline{W_n(\phi_Q)}) = \delta_{m,n} C_m(\phi_P, \phi_Q),$$

where

$$C_m(\phi_P, \phi_Q) = \sum_{v=1}^{\infty} \eta_{m,v} f_{m,v}(\phi_P) \overline{f_{m,v}(\phi_Q)},$$

$\eta_{m,v} > 0, \sum_{i=v}^{\infty} \eta_{m,i} < \infty$ and $f_{m,v}(\phi) = f_{m,v,R}(\phi) + i f_{m,v,I}(\phi)$ are the eigenvalues and complex orthogonal eigenfunctions such that

$$\int_0^{\pi} f_{m,i}(\phi)\ \overline{f_{n,j}(\phi)}\ d\phi = \delta_{m,n}\delta_{i,j}.$$

Here, $\overline{a}$ represents the conjugate of a complex number a.

We can determine the v^{th} eigenfunction and eigenvalue of the covariance function

$C_m(\phi_P, \phi_Q)$ by solving Fredlohom homogeneous equation

$$\int_0^\pi C_m(\phi_P, \phi_Q) f_{m,v}(\phi_Q) d\phi_Q = \eta_{m,v} f_{m,v}(\phi_P).$$

Therefore, $W_m(\phi)$ can be expressed in terms of the zero-mean K-L expansion,

$$W_m(\phi) = \sum_{v=1}^{\infty} \sqrt{\eta_{m,v}}\ \xi_{m,v}\ f_{m,v}(\phi), \quad \phi \in [0, \pi]. \tag{4.2}$$

Here the set of uncorrelated $\xi_{m,v}$ random coefficients is given by

$$\xi_{m,v} = \frac{1}{\sqrt{\eta_{m,v}}} \int_0^\pi W_m(\phi)\ \overline{f_{m,v}(\phi)}\ d\phi.$$

When $W_{-m}(\phi) = \overline{W_m(\phi)}$, the axially symmetric process $X(\phi, \lambda)$ in (4.1) can be expressed as

$$X(\phi, \lambda) = \sum_{v=1}^{\infty} \sqrt{\eta_{0,v}} f_{0,v}(\phi) \epsilon_{0,v} + 2 \sum_{m=1}^{\infty} \sum_{v=1}^{\infty} \sqrt{\eta_{m,v}} \epsilon_{m,v} (f_{m,v,R}(\phi) \cos(m\lambda) - f_{m,v,I}(\phi) \sin(m\lambda)). \tag{4.3}$$

For ease of notation we let

$$g_{m,v}(\phi, \lambda) = (f_{m,v,R}(\phi) \cos(m\lambda) - f_{m,v,I}(\phi) \sin(m\lambda)).$$

Now we consider the truncated K-L expansion $X_{N,M}(\phi, \lambda)$ where $N,\ M \in \mathbb{N}$,

$$X_{N,M}(\phi, \lambda) = \sum_{v=1}^{M} \sqrt{\eta_{0,v}} f_{0,v}(\phi) \epsilon_{0,v} + 2 \sum_{m=1}^{N} \sum_{v=1}^{M} \sqrt{\eta_{m,v}} \epsilon_{m,v} g_{m,v}(\phi, \lambda). \tag{4.4}$$

The error of the approximation can be obtained by taking the difference of (4.3) and (4.4), which gives

$$e_{NM}(\phi,\lambda) = \underbrace{\sum_{v\geq M+1}^{\infty}\sqrt{\eta_{0,v}}f_{0,v}(\phi)\epsilon_{0,v}}_{e_{0_{NM}}(\phi,\lambda)} + \underbrace{2\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty}\sqrt{\eta_{m,v}}\epsilon_{m,v}g_{m,v}(\phi,\lambda)}_{e_{1_{NM}}(\phi,\lambda)} + \underbrace{2\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\sqrt{\eta_{m,v}}\epsilon_{m,v}g_{m,v}(\phi,\lambda)}_{e_{2_{NM}}(\phi,\lambda)}. \quad (4.5)$$

We observe that the error $e_{n,M}(\phi,\lambda)$ can be split into three independent parts as follows,

$$e_{NM}(\phi,\lambda) = e_{0_{NM}}(\phi,\lambda) + e_{1_{NM}}(\phi,\lambda) + e_{2_{NM}}(\phi,\lambda).$$

Consider the absolute square error term for each component,

$$|e_{0_{NM}}(\phi,\lambda)|^2 = \sum_{v\geq M+1}^{\infty}\sum_{j\geq M+1}^{\infty}\sqrt{\eta_{0,v}}\sqrt{\eta_{0,j}}f_{0,v}(\phi)\overline{f_{0,j}(\phi)}\epsilon_{0,v}\epsilon_{0,j},$$

$$|e_{1_{NM}}(\phi,\lambda)|^2 = 4\sum_{m=1}^{N}\sum_{n=1}^{N}\sum_{v\geq M+1}^{\infty}\sum_{j\geq M+1}^{\infty}\sqrt{\eta_{m,v}}\sqrt{\eta_{n,j}}\epsilon_{m,v}\epsilon_{n,j}g_{m,v}(\phi,\lambda)g_{n,j}(\phi,\lambda),$$

$$|e_{2_{NM}}(\phi,\lambda)|^2 = 4\sum_{m\geq N+1}^{\infty}\sum_{n\geq N+1}^{\infty}\sum_{v=1}^{\infty}\sum_{j=1}^{\infty}\sqrt{\eta_{m,v}}\sqrt{\eta_{n,j}}\epsilon_{m,v}\epsilon_{n,j}g_{m,v}(\phi,\lambda)g_{n,j}(\phi,\lambda).$$

Since the ϵ_{ij} are uncorrelated Gaussian random variables with mean zero and unit variance, we have $Cov(\epsilon_{m,v},\epsilon_{m,v'}) = \delta_{v,v'}$, $Cov(\epsilon_{n,j},\epsilon_{n,j'}) = \delta_{j,j'}$, and $Cov(\epsilon_{m,v},\epsilon_{n,j}) = \delta_{m,n}\delta_{v,j}$. All the error terms are uncorrelated with each other because they do not

share the same running indices. Thus, we can express square error terms as follows,

$$\mathrm{E}(|e_{NM}(\phi,\lambda)|^2) = \mathrm{E}|e_{0_{NM}}(\phi,\lambda)|^2) + \mathrm{E}(|e_{1_{NM}}(\phi,\lambda)|^2) + E(|e_{2_{NM}}(\phi,\lambda)|^2),$$

where,

$$\mathrm{E}(|e_{0_{NM}}(\phi,\lambda)|^2) = \sum_{v\geq M+1}^{\infty} \eta_{0,v}|f_{0,v}(\phi)|^2, \tag{4.6}$$

$$\mathrm{E}(|e_{1_{NM}}(\phi,\lambda)|^2) = 4\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty} \eta_{m,v}\ |g_{m,v}(\phi,\lambda)|^2, \tag{4.7}$$

$$\mathrm{E}(|e_{2_{NM}}(\phi,\lambda)|^2) = 4\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty} \eta_{m,v}\ |g_{m,v}(\phi,\lambda)|^2. \tag{4.8}$$

Therefore, the L^2 error bound can be written as,

$$\begin{aligned}\|X(\phi,\lambda) - X_{N,M}(\phi,\lambda)\|^2 = \mathrm{E}(|e_{NM}(\phi,\lambda)|^2) = \sum_{v\geq M+1}^{\infty} \eta_{0,v}|f_{0,v}(\phi)|^2 \\ + 4\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty} \eta_{m,v}\ |g_{m,v}(\phi,\lambda)|^2 + 4\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty} \eta_{m,v}\ |g_{m,v}(\phi,\lambda)|^2.\end{aligned} \tag{4.9}$$

Theorem 4.1. *Let $X(\phi,\lambda)$ be a zero-mean continuous axially symmetric random process on the sphere. Then we have the expected L^2 error in terms of individual mean error term. $\|X(\phi,\lambda) - X_{N,M}(\phi,\lambda)\|^2$ is given by (4.9). Hence, when $N, M \to \infty$, $\|X(\phi,\lambda) - X_{N,M}(\phi,\lambda)\|^2 \to 0$. That is, the truncated K-L expansion approximates to the true process with mean-squared convergence.*

Proof. It is sufficient to prove that each of the terms in (4.9) converges to zero. Recall that for any $\phi_P, \phi_Q \in [0, \pi]$,

$$\sum_{m=0}^{\infty} C_m(\phi_P, \phi_Q) = \sum_{m=0}^{\infty} \left| \sum_{v=1}^{\infty} \eta_{m,v} f_{m,v}(\phi_P) \overline{f_{m,v}(\phi_Q)} \right| < \infty.$$

Hence, if we take $\phi_P = \phi_Q \in [0, \pi]$, we have

$$C_0(\phi_P, \phi_P) = \sum_{v=1}^{\infty} \eta_{0,v} f_{0,v}(\phi_P) \overline{f_{0,v}(\phi_Q)} = \sum_{v=1}^{\infty} \eta_{0,v} |f_{0,v}(\phi_P)|^2 < \infty,$$

implying that when $M \to \infty$,

$$\mathrm{E}(|e_{0_{NM}}(\phi, \lambda)|^2) = \sum_{v \geq M+1}^{\infty} \eta_{0,v} \, |f_{0,v}(\phi)|^2 \to 0.$$

In parallel,

$$\begin{aligned} \sum_{m=0}^{\infty} |C_m(\phi, \phi)| &= \sum_{m=0}^{\infty} \left| \sum_{v=1}^{\infty} \eta_{m,v} f_{m,v}(\phi) \overline{f_{m,v}(\phi)} \right| = \sum_{m=0}^{\infty} \left| \sum_{v=1}^{\infty} \eta_{m,v} |f_{m,v}(\phi)|^2 \right| \\ &= \sum_{m=0}^{\infty} \sum_{v=1}^{\infty} \eta_{m,v} |f_{m,v}(\phi)|^2 < \infty, \end{aligned}$$

since $\eta_{m,v} \geq 0$ and $|f_{m,v}(\phi)|^2 \geq 0$ for all m, v.

Therefore,

$$\sum_{m=0}^{\infty} \sum_{v=1}^{\infty} \eta_{m,v} |f_{m,v,R}|^2 < \infty, \tag{4.10}$$

and

$$\sum_{m=0}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|f_{m,v,I}|^2<\infty, \tag{4.11}$$

which implies that

$$\sum_{m=0}^{\infty}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,R}(\phi)|^2\to 0, \tag{4.12}$$

and

$$\sum_{m=0}^{\infty}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,I}(\phi)|^2\to 0, \tag{4.13}$$

as $M\to\infty$. Note that

$$\begin{aligned}|g_{m,v}(\phi,\lambda)|^2&=|f_{m,v,R}(\phi)\cos(m\lambda)-f_{m,v,I}\sin(m\lambda)|^2\leq\left[|f_{m,v,R}(\phi)|+|f_{m,v,I}(\phi)|\right]^2\\&\leq 2\left(|f_{m,v,R}(\phi)|^2+|f_{m,v,I}(\phi)|^2\right).\end{aligned}$$

Hence

$$\begin{aligned}&\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|g_{m,v}(\phi,\lambda)|^2\leq 2\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty}\eta_{m,v}\left(|f_{m,v,R}(\phi)|^2+|f_{m,v,I}(\phi)|^2\right)\\&=2\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,R}(\phi)|^2+2\sum_{m=1}^{N}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,I}(\phi)|^2\\&\leq 2\sum_{m=1}^{\infty}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,R}(\phi)|^2+2\sum_{m=1}^{\infty}\sum_{v\geq M+1}^{\infty}\eta_{m,v}|f_{m,v,I}(\phi)|^2\to 0\end{aligned}$$

as $M \to \infty$ from (4.12) and (4.13), giving $\mathrm{E}(|e_{1_{NM}}(\phi,\lambda)|^2) \to 0$ as $M \to \infty$.

In addition, from (4.10) and (4.11), we have $\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|f_{m,v,R}(\phi)|^2 \to 0$, $\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|f_{m,v,I}(\phi)|^2 \to 0$ as $N \to \infty$. Hence

$$\begin{aligned}\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|g_{m,v}(\phi,\lambda)|^2 \leq\ & 2\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|f_{m,v,R}(\phi)|^2 \\ & +2\sum_{m\geq N+1}^{\infty}\sum_{v=1}^{\infty}\eta_{m,v}|f_{m,v,I}(\phi)|^2 \to 0\end{aligned}$$

as $N \to \infty$, implying that $\mathrm{E}(|e_{2_{NM}}(\phi,\lambda)|^2) \to 0$ as $N \to \infty$. This concludes the proof of Theorem 4.1.

□

Theorem 4.2. *Let $X(\phi,\lambda)$ be a zero-mean continuous axially symmetric random process on the sphere. We further assume that for every $m = 0, 1, 2, \ldots,$*

$$C_m(\phi_P,\phi_Q) = C_0(\phi_P,\phi_Q)\xi_m e^{imu(\phi_P-\phi_Q)}, \quad \phi_P,\phi_Q \in [0,\pi],$$

where $u \geq 0$, $\xi_0 = 1$ (for simplicity), $\xi_m \geq 0, m = 1, 2, \ldots,$ and $\sum_{m=0}^{\infty}\xi_m < \infty$. $C_0(\phi_P,\phi_Q)$ is symmetric positive definite real function. Then we have

(1) *$C_m(\phi_P,\phi_Q)$ is Hermitian and positive definite.*

(2) *If $C_0(\phi_P,\phi_Q)$ has the following eigen-decomposition based on Mercer's Theorem*

$$C_0(\phi_P,\phi_Q) = \sum_{v=1}^{\infty}\eta_{0,v}f_{0,v}(\phi_P)f_{0,v}(\phi_Q),$$

with $\eta_{0,v} \geq 0$ *and* $f_{0,v}(\phi)$ *being an orthonormal real-valued basis in* $L^2([0,\pi])$*, Then if* $C_m(\phi_P, \phi_Q)$ *has the following decomposition:*

$$C_m(\phi_P, \phi_Q) = \sum_{v=0}^{\infty} \eta_{0,v} f_{m,v}(\phi_P) \overline{f_{m,v}(\phi_Q)},$$

we have

$$\eta_{m,v} = \xi_m \eta_{0,v}, \qquad f_{m,v}(\phi) = f_{0,v}(\phi) e^{imu\phi}.$$

(3) *With the above decomposition for* $C_m(\phi_P, \phi_Q)$*, the truncated K-L expansion of (*4.3*) is of the form*

$$X_{N,M}(\phi, \lambda) = \sum_{v=1}^{M} \sqrt{\eta_{0,v}} f_{0,v}(\phi) \epsilon_{0,v} + 2 \sum_{m=1}^{N} \sum_{v=1}^{M} \sqrt{\eta_{m,v}} \epsilon_{m,v} f_{0,v}(\phi) \cos m(u\phi + \lambda),$$

and we have

$$\|X(\phi, \lambda) - X_{N,M}(\phi, \lambda)\|^2 \leq \left(1 + 4 \sum_{m=1}^{N} \xi_m\right) \sum_{m \geq M+1} \eta_{0,v} f_{0,v}^2(\phi) + 4C_0(\phi, \phi) \sum_{m \geq N+1} \xi_m. \quad (4.14)$$

Moreover, when $N, M \to \infty$*,* $\|X(\phi, \lambda) - X_{N,M}(\phi, \lambda)\|^2 \to 0$*.*

Proof. (1) We first prove that $C_m(\phi_P, \phi_Q)$ is Hermitian and positive definite. For every $m \geq 0$,

$$\begin{aligned} C_m(\phi_Q, \phi_P) &= C_0(\phi_Q, \phi_P)\xi_m e^{imu(\phi_Q - \phi_P)} = C_0(\phi_P, \phi_Q)\xi_m \overline{e^{imu(\phi_P - \phi_Q)}} \\ &= \overline{C_m(\phi_P, \phi_Q)}. \end{aligned}$$

That is, $C_m(\phi_P, \phi_Q)$ is Hermitian. For positive definiteness, we take an arbitrary n, any n complex numbers $c_i, i = 1, 2, \cdots, n$, and any $\phi_i \in [0, \pi], i = 1, 2, \cdots, n$, considering

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{n} c_i\overline{c_j}C_m(\phi_i, \phi_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i\overline{c_j}C_0(\phi_i, \phi_j)\xi_m e^{imu(\phi_i - \phi_j)} \\ &= \xi_m \sum_{i=1}^{n}\sum_{j=1}^{n} c_i e^{imu\phi_i}\overline{c_j}e^{-imu\phi_j}C_0(\phi_i, \phi_j) \\ &= \xi_m \sum_{i=1}^{n}\sum_{j=1}^{n} c_i e^{imu\phi_i}\overline{c_j e^{imu\phi_j}}C_0(\phi_i, \phi_j) \geq 0, \end{aligned}$$

by the positive definiteness of $C_0(\phi_P, \phi_Q)$. Hence $C_m(\phi_P, \phi_Q)$ is the (nonstationary) covariance function of a possibly complex-valued random process on $[0, \pi]$.

(2) Now we prove the second result. As given before, we have

$$C_0(\phi_P, \phi_Q) = \sum_{v=1}^{\infty} \eta_{0,v} f_{0,v}(\phi_P) f_{0,v}(\phi_Q) < \infty,$$

and

$$\int_0^\pi C_0(\phi_P, \phi_Q) f_{0,v}(\phi_Q) d\phi_Q = \eta_{0,i} f_{0,v}(\phi_P).$$

Now considering the both sides of the above equation

$$\begin{aligned}
\eta_{0,v} f_{0,v}(\phi_P) &= \int_0^\pi C_0(\phi_P, \phi_Q) f_{0,v}(\phi_Q) d\phi_Q \\
\xi_m \eta_{0,v} f_{0,v}(\phi_P) e^{imu\phi_P} e^{-imu\phi_P} &= \int_0^\pi \xi_m C_0(\phi_P, \phi_Q) f_{0,v}(\phi_Q) e^{imu\phi_Q} e^{-imu\phi_Q} d\phi_Q \\
\xi_m \eta_{0,v} f_{0,v}(\phi_P) e^{imu\phi_P} &= \int_0^\pi \xi_m C_0(\phi_P, \phi_Q) e^{imu(\phi_P - \phi_Q)} \\
f_{m,v}(\phi_Q) e^{imu\phi_Q} d\phi_Q & \\
&= \eta_{m,v} f_{m,v}(\phi_P),
\end{aligned} \tag{4.15}$$

implying that

$$\eta_{m,v} = \xi_m \eta_{0,v}, \qquad f_{m,v}(\phi) = f_{0,v}(\phi) e^{imu\phi}, \qquad m = 1, 2, \ldots.$$

(3) With the given truncated K-L expansion

$$X_{N,M}(\phi, \lambda) = \sum_{v=1}^{M} \sqrt{\eta_{0,v}} f_{0,v}(\phi) \epsilon_{0,v} + 2 \sum_{m=1}^{N} \sum_{v=1}^{M} \sqrt{\eta_{m,v}} \epsilon_{m,v} g_{m,v}(\phi, \lambda),$$

we notice that

$$\eta_{m,v} = \xi_m \eta_{0,v},$$

$$
\begin{aligned}
g_{m,v}(\phi,\lambda) &= f_{m,v,R}(\phi)\cos(m\lambda) - f_{m,v,I}(\phi)\sin(m\lambda) \\
&= f_{0,v}(\phi)\cos(m\phi)\cos(m\lambda) - f_{0,v}(\phi)\sin(m\phi)\sin(m\lambda) \\
&= f_{0,v}(\phi)\cos m(u\phi+\lambda).
\end{aligned}
$$

That is

$$
X_{N,M}(\phi,\lambda) = \sum_{v=1}^{M}\sqrt{\eta_{0,v}}f_{0,v}(\phi)\epsilon_{0,v} + 2\sum_{m=1}^{N}\sum_{v=1}^{M}\sqrt{\xi_m\eta_{0,v}}\epsilon_{m,v}f_{0,v}(\phi)\cos m(u\phi+\lambda).
$$

Hence,

$$
\begin{aligned}
\|X_{N,M}(\phi,\lambda) - X(\phi,\lambda)\|^2 \leq & \sum_{v\geq M+1}\eta_{0,v}f_{0,v}^2(\phi) + 4\sum_{m=1}^{N}\xi_m\sum_{v\geq M+1}\eta_{0,v}f_{0,v}^2(\phi) + \\
& 4\sum_{v\geq N+1}\xi_m\sum_{v=1}^{\infty}\eta_{0,v}|f_{0,v}(\phi)|^2 \\
= & \left(1+4\sum_{m=1}^{N}\xi_m\right)\sum_{v\geq M+1}\eta_{0,v}f_{0,v}^2(\phi) + 4C_0(\phi,\phi)\sum_{v\geq N+1}\xi_m,
\end{aligned}
$$

which tends to 0 when $N, M \to \infty$ since $C_0(\phi,\phi) < \infty$ and $\sum_{i=0}^{\infty}\xi_m < \infty$.

□

Now consider a parametric nonstationary covariance function proposed in [VWZ19], as listed as model 1 in their paper.

$$
C_m(\phi_P,\phi_Q) = C_0(\phi_P,\phi_Q)p^m e^{imb} \tag{4.16}
$$

where $p \in (0,1)$, $b = u(\phi_P - \phi_Q)$ with $u \geq 0$, and

$$C_0(\phi_P, \phi_Q) = C_1(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P - \phi_Q|}),$$

for $C_1 > 0$, $C_2 \geq 1$ and $a > 0$. Note that when $u = 0$, the process becomes a longitudinally reversible process, under which we have

$$C_m(\phi_P, \phi_Q) = C_0(\phi_P, \phi_Q)p^m.$$

From Corollary 4.3, we have $\xi_m = p^m, m = 0, 1, 2, \ldots$. To justify the accuracy of our approximation with the K-L expansion, we consider the L^2 error bound of the approximation. Without further assumptions on the covariance function, it is difficult to provide the error bound. For our simulation, we consider the covariance function given by (4.16). In particular, we set $C_1 = C_2 = a = 1$, so that

$$C_0(\phi_P, \phi_Q) = 1 - e^{-\phi_P} - e^{-\phi_Q} + e^{-|\phi_P - \phi_Q|}, \tag{4.17}$$

which gives $C_0(\phi, \phi) = 2(1 - e^{-\phi}), \quad \phi \in [0, \pi]$. Therefore, we present the following corollary.

Corollary 4.3. *Under the covariance function (*4.16*) with the special case (*4.17*), the L^2 error bound on the K-L approximation is given as*

$$\|X_{N,M}(\phi, \lambda) - X(\phi, \lambda)\|^2 = \left(1 + 4\frac{p(1-p^N)}{1-p}\right) \sum_{v \geq M+1} \eta_{0,v} f_{0,v}^2(\phi) + 8\frac{p^{N+1}}{(1-p)}(1 - e^{-\phi}). \tag{4.18}$$

We will conduct simulations in Section 4.3 to demonstrate the accuracy of the empirical errors to the theoretical errors given by Corollary 4.3.

4.2 Global Data Generation on the Sphere

The generation of random data is a fundamental element in statistical research. However, there are only very few studies available in the literature for data generation on the sphere. It should be noted that with a data generation algorithm, researchers can reproduce samples that mimic real data sets to investigate the performance of a proposed statistical method. Therefore, it is critical to provide an efficient and accurate algorithm for data generation. There is limited research in the literature on data generation of axially symmetric processes on the sphere. For example, convolution methods to generate random fields with a class of Matérn type kernel function was proposed by [Yan13]. In their recent work [VWZ19] demonstrated that axially symmetric data on the sphere can be decomposed into a Fourier series on circles, where the Fourier random coefficients can be expressed as circularly-symmetric complex random vectors. In this section we develop an accurate and simple algorithm to generate axially symmetric data on $\mathbb{S}^2$.

According to (4.1), with $W_{-m}(\phi) = \overline{W_m(\phi)}$ we can write truncated axially symmetric Gaussian random process on the sphere as

$$X_N(P) = W_0(\phi) + 2\sum_{m=1}^{N}[W_{m,R}(\phi)\cos(m\lambda) - W_{m.I}(\phi)\sin(m\lambda)] \tag{4.19}$$

where,

$$W_m(\phi) = W_{m,R}(\phi) + iW_{m,I}(\phi) = \sum_{v=1}^{M} \sqrt{\eta_{m,v}}\epsilon_{m,v} f_{m,v,R}(\phi) + i\sum_{v=1}^{M} \sqrt{\eta_{m,v}}\epsilon_{m,v} f_{m,v,I}(\phi). \tag{4.20}$$

Under the assumption of a longitudinally reversible process, $W_m(\phi) = W_{-m}(\phi)$. Then the global data can be generated by simplifying equation (4.1) as follows

$$\begin{aligned} X_L(\lambda,\phi) &= W_0(\phi) + \sum_{m=1}^{\infty} W_m(\phi)(e^{im\lambda} + e^{-im\lambda}), \\ &= W_0(\phi) + 2\sum_{m=1}^{\infty} W_m(\phi)\cos(m\lambda), \text{ where } m = 0, 1, \ldots \end{aligned} \tag{4.21}$$

With the Wavelet-Galerkin method and Algorithm 3.1 from Section 3.2, we propose the following algorithm for generating axially symmetric data on the sphere.

Algorithm 4.1 (Pseudo-code)

- Choose a cross covariance function $C_m(\cdot,\cdot)$ with given parameters (C_1, C_2, a, u, p)
- Specify the set of latitudes $\phi_1, \phi_2, \ldots, \phi_{nl}$ and the set of longitudes $\lambda_1, \lambda_2, \ldots, \lambda_{nL}$.
- For each m, $m = 0, 1, \ldots, N$
  (1) For the covariance function $C_m(\cdot,\cdot)$, construct the covariance matrix evaluated over $M \times M$ gridded matrix
  (2) Apply the 2-D in-place Wavelet discrete transformation (Algorithm 3.1) on resulting covariance matrix to obtain A (see (3.17)).
  (3) Find the set of eigenvalues and eigenfunctions for $\hat{A}$ (see equation (3.18))
  (4) Obtain $W_m(\phi_i)$, $i = 1, 2, \ldots, n_l$ from equation (4.20).
- For each $i = 1, 2, \ldots, n_l$ apply equation (4.19) to generate data $\{X(\phi_i, \lambda_i)\}_{i=1,j=1}^{n_l, n_L}$.

4.3 Simulations and Results

In this section, we will present the results of two simulation studies. In the first simulation study, theoretical L^2 error bounds (as mentioned in Section 4.1) are compared to empirical L^2 error bounds to investigate the convergence rate of the generated random processes. Here we consider L^2 error for different number of terms and different parameter values in the approximation. The second simulation study illustrates a validation study for the global data generation algorithm (proposed in section 4.2) through simulations.

4.3.1 Convergence Study by Simulation

First we perform simulations to demonstrate how well Wavelet-Galerkin method based empirical error approximate the true L^2 error bound. More explicitly, we consider the L^2 error bound given in (4.18) when the covariance function is (4.16) with $C_0(\phi_P, \phi_Q)$ given by equation (4.17).

Simulation steps:

- Choose any values of $\phi \in [0, \pi]$, $\lambda \in [0, 2\pi]$. In this case, $\phi = \lambda = \frac{\pi}{18}$ is being used. Consider different values of $M = 4, 8, 16, 32, 64, 128,$ and 256 with fixed $N = 10$ and choose $p = 0.5$ (Note that: we choose $N = 10$ because there is a monotonic decay in the function of N. Other cases of N and p are also considered but the results are similar). Use (4.19) with the desired M value to generate Gaussian axially symmetric random processes on the sphere $(\hat{X}_{10,M}(\phi, \lambda))$ for a given location. The numerical eigenvalue-eigenfunction pairs are obtained from the Wavelet-Galerkin method (see Section 3.4).

- We use theoretical eigenvalues and eigenfunctions (as mentioned in Proposition 3.1) in the expression of (4.19) to calculate $X_{512,512}(\phi, \lambda)$.

- The empirical truncated square error for each i^{th} repetition can be calculated by

$$L^2 \text{ error}[i] = [X_{512,512}(\phi, \lambda)[i] - \hat{X}_{10,M}(\phi, \lambda)[i]]^2.$$

- Then we calculate the average empirical L^2 error out of n repetitions.

$$L^2 \text{ error} = \frac{1}{n} \sum_{i=1}^{n} L^2 \text{ error}[i].$$

- Finally, we compare those empirical error values with theoretical L^2 error values given by (4.18), where theoretical eigenvalues and eigenfunctions are used.

Note: We have obtained similar results for other sets of parameter values and locations as well, which are not presented here.

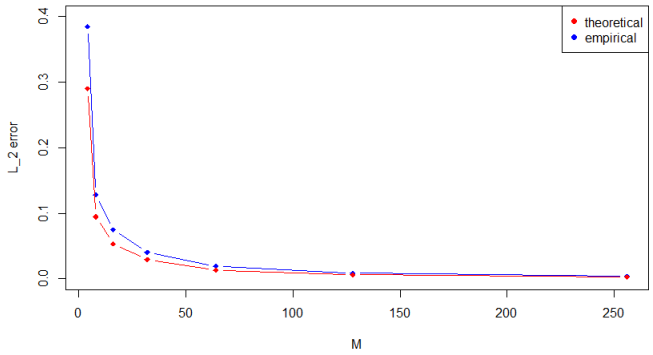

Figure 11. The L^2 error as a function of M. The theoretical error bound is represented by the red solid line and the average empirical L^2 error bound by blue line with 1000 repetitions.

The comparison between theoretical and simulated L^2 errors is given by Figure 11. When we increase the number of terms M in the truncated random process for a given location, the truncated process smoothly converges to the true value.

In addition, we explore how parameters of a covariance function would affect to the rate of convergence. Figure 12 shows how $p \in (0, 1)$ values in the covariance model (4.16) affect the L^2 convergence rate. We notice that the following two points: the convergence seems to be faster when p gets small, on the other hand, regardless of the p value, the L^2 error gets closer to zero when the number of terms M in the truncated expansion increases.

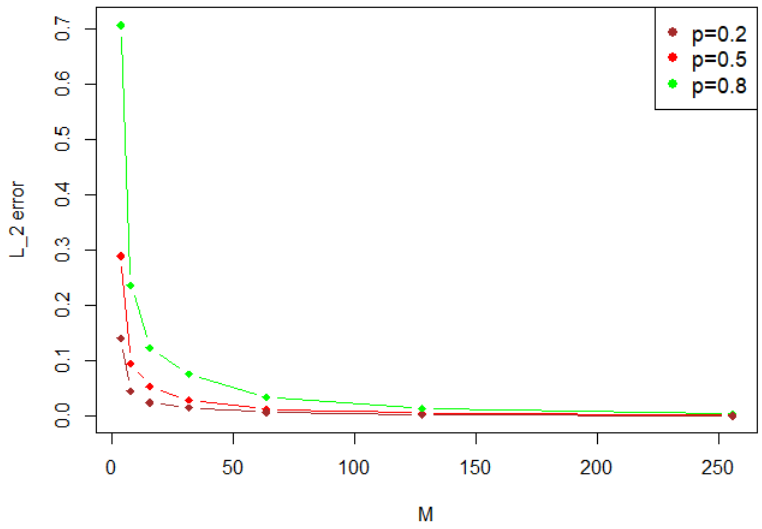

Figure 12. The average empirical L^2 error as a function of M for different p values

4.3.2 Simulation Study for Data Generation and Validation for Axially Symmetric Processes on $\mathbb{S}^2$

In this section, we perform data validation through simulations where the axially symmetric random data are generated based on the proposed algorithm in Section 4.2. Note that a Gaussian process on the sphere is uniquely determined by its mean and covariance function. Hence with the assumption of zero mean (without loss of generality), we need to validate the covariance function from the generated data to match the true covariance function. However, as indicated by [Ada17] and [Van16], the classical Method of Moment (MOM) covariance estimator based on axially symmetric processes may not be estimable, and hence we will use the unbiased MOM variogram estimator for data validation. More explicitly, we compute the

cross-variogram empirical estimator from the generated data and then compare them with theoretical cross-variogram value through simulation studies. Furthermore, we also compare the accuracy of our proposed data generation method with the classical data generation method (see [VWZ19] for more details) in terms of biasness and mean squared errors (MSE). Note that the data generation approach based on circular symmetric processes introduced in [VWZ19] performs almost the same as the classical method. Hence, we do not include in this analysis.

In our simulation, we consider two axially symmetric covariance models that were considered in [VWZ19], which are given below.

Model 1:

$$R(P,Q,\Delta\lambda) = \tilde{C}(\phi_P,\phi_Q)\frac{1-p^2}{1-2p\cos(\Theta)+p^2}. \tag{4.22}$$

Model 2:

$$R(P,Q,\Delta\lambda) = \tilde{C}(\phi_P,\phi_Q)\log\frac{1}{1-2p\cos(\Theta)+p^2}, \tag{4.23}$$

where $\tilde{C}(\phi_P,\phi_Q) = C_1(C_2 - e^{a|\phi_P|} - e^{a|\phi_Q|} + e^{-a|\phi_P-\phi_Q|})$, $\Theta = \Delta\lambda + u(\phi_P - \phi_Q) \in [0, 2\pi]$ with $u \geq 0$ and $p \in (0,1)$. When $u = 0$, $R(P,Q,\Delta\lambda)$ becomes a longitudinal-reversible covariance function with the inverse Fourier transformation given by (1.7), which can be obtained as (also see [VWZ19]).

Model 1:

$$C_m(\phi_P, \phi_Q) = \begin{cases} \tilde{C}(\phi_P, \phi_Q), & m = 0, \\ \tilde{C}(\phi_P, \phi_Q) p^m e^{imb}, & m \neq 0. \end{cases}$$

Model 2:

$$C_m(\phi_P, \phi_Q) = \begin{cases} 0, & m = 0, \\ \tilde{C}(\phi_P, \phi_Q) \frac{p^m}{m} e^{imb}, & m \neq 0 \end{cases}$$

where $m = \pm 1, \pm 2, \ldots, b = u(\phi_P - \phi_Q)$.

In our simulation study, we also follow the same sets of parameter values (see Table 4) that are given in [VWZ19]

Table 4. Parameter values of the covariance functions

	Parameter Values
Set 1	$C_1 = 1, C_2 = 1, a = 1, u = 1, p = 0.5$
Set 2	$C_1 = 1, C_2 = 2, a = 3, u = 1, p = 0.6$

In addition we have the following remarks:

- Axially symmetric random processes on the sphere are generated using Algorithm 4.1 for 100 longitudes (λ) on a fixed pair of latitudes (ϕ). We set M=256

and we consider two sets of parameter values mentioned in the above table to construct covariance functions.

- Then we calculate the empirical cross variogram estimator using the expression (1.10) for generated data and compare with the theoretical cross variogram function (1.8) of the desired model (either (4.22) or (4.23)).

We perform simulations with a repetition of 4000 times. For comparison purposes, we also include the simulation results based on the classical method. Following are the results for both covariance models evaluated over different pair of latitudes with two parameter sets.

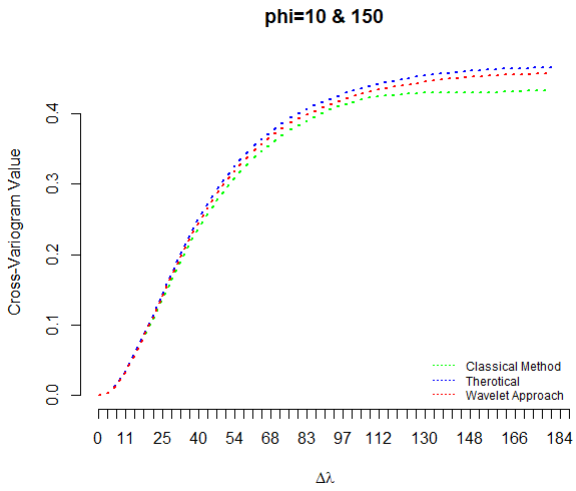

Figure 13. Cross variogram estimator comparison of Model 1 under the longitudinally reversible assumption ($u = 0$) for the parameter set 1.

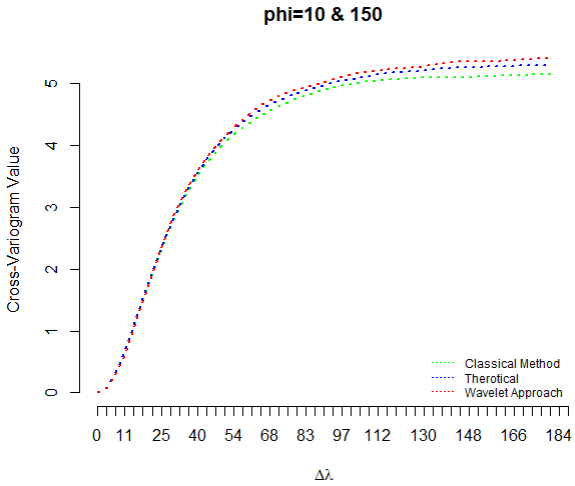

Figure 14. Cross variogram estimator comparison of Model 1 under the longitudinally reversible assumption ($u = 0$) for the parameter set 2.

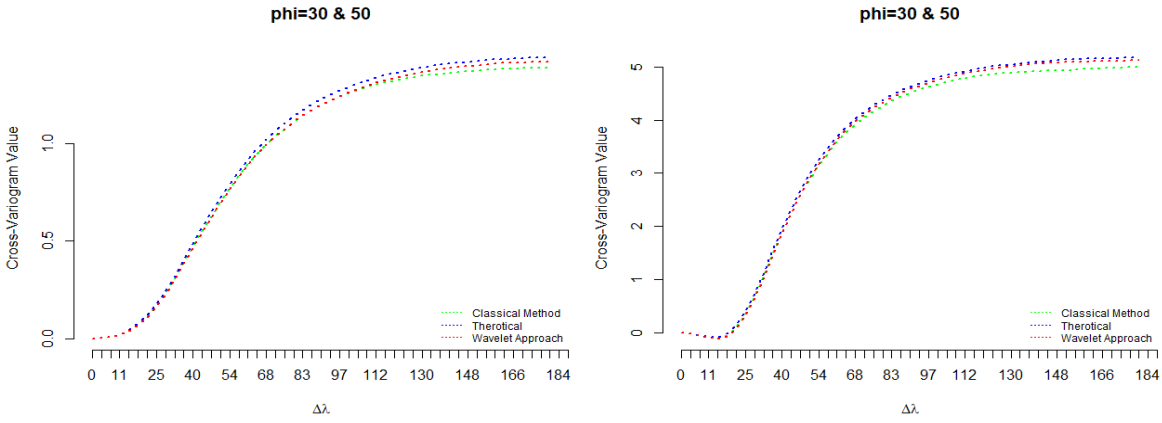

Figure 15. Cross variogram estimator comparison of Model 1 under the axially symmetric assumption for the parameter set1 (left) and set 2 (right).

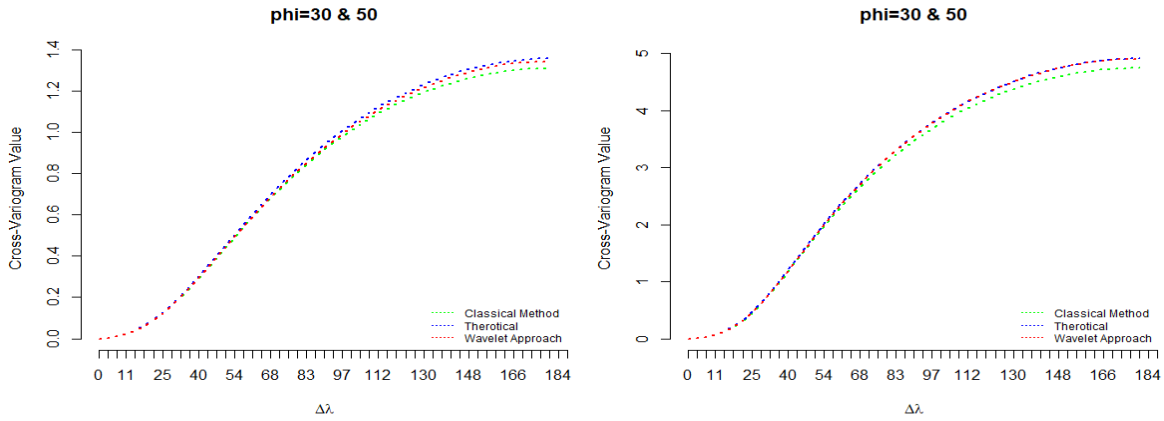

Figure 16. Cross variogram estimator comparison of Model 2 under the axially symmetric assumption for the parameter set1 (left) and set 2 (right).

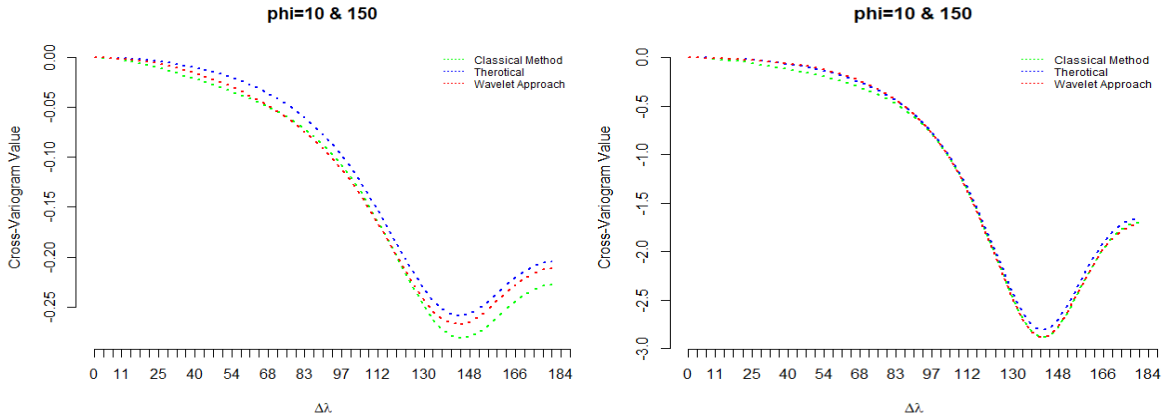

Figure 17. Cross variogram estimator comparison of Model 1 under the axially symmetric assumption for the parameter set1 (left) and set2 (right).

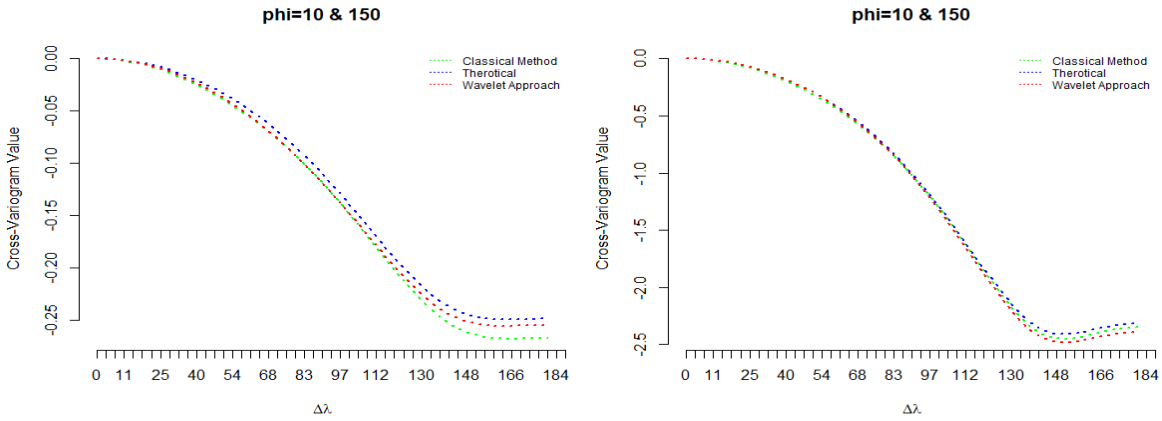

Figure 18. Cross variogram estimator comparison of Model 2 under the axially symmetric assumption for the parameter set1 (left) and set 2 (right).

From Figure 13 to Figure 18, we can see that empirical MOM cross-variogram values based on both the Wavelet-Galerkin method and classical method are very close to the theoretical counterparts, regardless the choice of latitudes pairs, the parameter sets used, the covariance models, and the random process (either longitudinally reversible or axially symmetric processes). However, the biasness of the empirical MOM cross-variogram values from the Wavelet-Galerkin method to the theoretical counterparts is overall less than that from the classical method.

4.3.3 Comparison of Method of Moments (MOM) Variogram Estimator Values between the Wavelet Approach and Classical Approach

In this section, we conduct a comparison study to evaluate the performance of the wavelet approach over the classical approach. We calculate the mean squared error (MSE) for both approaches. Here the MSE of MOM variogram estimator can be calculated using the following formula.

$$\begin{aligned} MSE_{j\Delta} &= Var_{j\Delta} + (Bias_{j\Delta})^2 \\ &= \frac{1}{nn}\sum_{i=1}^{nn}\left(\hat{\gamma}_i(j\Delta) - \overline{\hat{\gamma}(j\Delta)}^2\right) + (\gamma(j\Delta) - \overline{\hat{\gamma}(j\Delta)}^2), \end{aligned}$$

where $\Delta = \frac{2\pi}{n_L}$, n_L and nn are the number of longitudes and number of repeated simulations, respectively.

From both figures 19 and 20, we can notice that for all $\Delta\lambda$ values, bias is low for the wavelet approach compared to the classical method. However, the MSE values from the Wavelet-Galerkin approach are comparable with those from classical method when $\Delta\lambda$ is small, but tend to be larger when $\Delta\lambda$ is getting larger. It should be noted that the classical method involves the block circular matrix decomposition, and the computation cost might be expensive when the dimension is big, which is in contrast to the low dimensional matrix manipulation in the Wavelet-Galerkin approach. In summary, our method provides a comparable approach for axially symmetric random data generation on the sphere with lower computational cost.

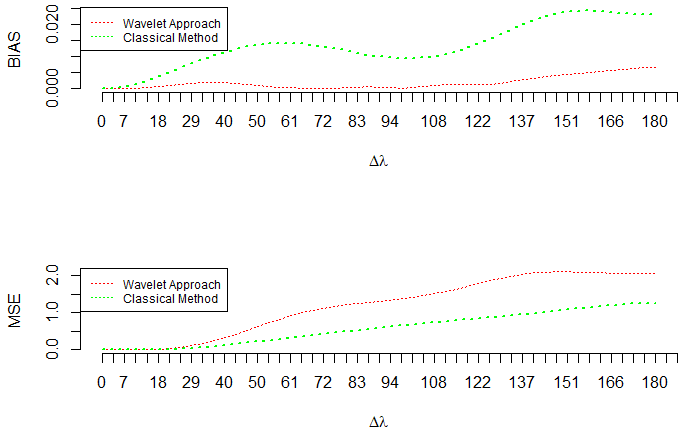

Figure 19. Bias and MSE comparison of the wavelet approach and classical method for a fixed pair of latitudes $(80^0S, 60^0N)$ over $\Delta\lambda \in (0, \pi)$ for Model 1 with set 1 parameter values.

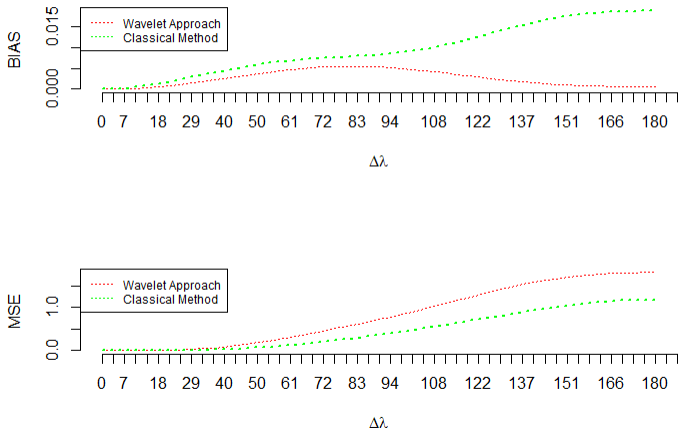

Figure 20. Bias and MSE comparison of the wavelet approach and classical method for a fixed pair of latitudes $(80^0S, 60^0N)$ over $\Delta\lambda \in (0, \pi)$ for Model 2 with set 1 parameter values.

CHAPTER V

SUMMARY AND FUTURE WORK

5.1 Summary

In this dissertation, we successfully developed an algorithm to perform a two-dimensional discrete wavelet transformation for any given complex-valued covariance functions. Moreover, the Wavelet-Galerkin numerical method has been extended to numerically determine the eigenvalues and eigenfunctions of real-valued and complex-valued non-stationary covariance functions defined on $\mathbb{R}^2$. Then we have conducted a convergence study for the truncated K-L expansion of an axially symmetric Gaussian random process on $\mathbb{S}^2$. Furthermore, we proved that the truncated process approximates to the true axially symmetric process value with convergence in L^2 (mean squared convergence). Theoretical L^2 error bound was derived for such an approximation. In addition, we have demonstrated how the values of the model parameters affect the convergence rate. Finally, we proposed an algorithm using the K-L expansion to generate global data when the underlying random process is Gaussian axially symmetric. The validation through simulation was conducted to verify the performance of our proposed method with the classical method.

5.2 Future Work

Our proposed Wavelet-Galerkin method relies on the known covariance function evaluated at the given locations. Therefore, it is important to obtain such covariance matrices when modeling. Note that with a given real data, one can obtain a nonparametric estimate of the axially symmetric covariance function $R(\phi_P, \phi_Q, \Delta\lambda))$

through the MOM approach, which in turn would provide a nonparametric estimator $C_m(\phi_P, \phi_Q)$ through the inverse of Fourier transform from (1.7). Therefore, we will explore how the proposed wavelet-Galerkin approach can be used to investigate the important features of possibly complex-valued covariance functions $C_m(\phi_P, \phi_Q)$, which would help us better understand the dependency of random processes on a global scale.

There has been limited research on axially symmetric covariance models in literature. Therefore, as our next research project, we will be working on constructing axially symmetric covariance models in $\mathbb{S}^2$ that could be practically useful in analyzing massive global data sets. We will also enforce sparsity to covariance function approximations to convert full rank variance-covariance matrices to reduced rank matrices which will be more beneficial in handling large data sets.

Finally, following [Van16] which has shown the non-consistency of the MOM variogram estimator on the circle, we will investigate the non-consistency of the MOM estimator on the sphere first for longitudinally reversible processes, and we will then extend the idea to prove the non-consistency for an axially symmetric process on the sphere.

REFERENCES

[Ada17] Adaramola, B.O., *Circulant matrices on global data analysis. Master's thesis, Department of Mathematics and Statistics, University of North Carolina at Greensboro, 2017.*

[BUS20] Bussberg, N. (2020). *Environmeental applications of temporal, spatial, and spatial-temporal statistics.* PhD Dissertation, Indiana University, Bloomington, IN, 2020.

[CL88] Claeyssen, R.C.J., Leal, L.A.D.S., *Diagonalization and Spectral Decomposition of Factor Block Circulant Matrix.* Elsevier Science Publishing, 1988.

[Cre93] Cressie, N.A.C, *Statistics for spatial data, 1993.*

[Dau92] Daubechies, I., 1992. *Ten lectures of Wavelets.* Springer-Verlag, 1992.

[Dav79] Davis, P. R., 1979. *Circulant Matrices.* John Wiley, New York.

[Gal08] Gallager, R. G., *Circularly-symmtric gaussion random vectors.*

[Gar15] Garry, J. T., *Eigenvectors of block circulant and alternating circulant matrices.* Res. Lett. Inf. Math. Sci., 2005, Vol.8, pp123-142, 2015.

[Gne13] Gneiting, T., *Strictly and non-strictly positive definite functions on spheres. Bernouli 19, no.4, 1347-1349, 2013.*

[GF15] Guinness, J., Fuentes, M., *Isotropic covariance functions on sphere: Some properties and modeling considerations. Journal of Multivariate Analysis, 2015.*

[Hao04] Hao, Z., *Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. No. 456, 250-261, 2004*

[HS11] Hitczenko, M., Stein, M.L., *Some theory for anisotropic processes on the sphere. Statistical Methodology, 2011*

[HJ11] Huiyan, S., Jianhua Z.H., *A full-scale approximation of covariance functions for large spatial data sets, 2011.*

[HZR11] Huang, C., Zhang, H., Robeson, S. M., *On the validity of commonly used covrince and variogram functions n the sphere.* Mathematical Geosciences 43, no. 6, 721-733, 2011.

[HZR12] Huang, C., Zhang, H., Robeson, S. M., *A simplified representation of the covariance structure of axially symmetric processes on the sphere.* Statistics and Probability Letters 82(7), 1346-1351, 2012.

[HZRS19] Huang, C., Zhang, H., Robeson, S. M., Shields, J. *Intrinsic Random Functions on the Sphere.* Statistics and Probability Letters, 146, 7-14.

[Jon63] Jones, R.H., *Stochastic processes on a sphere. The Annals of Mathematical Statistics 1963, Vol. 2, No. 1, 213-218, 1963.*

[JS08] Jun, M., Steal M.L., *Nonstionary covariance models for global data. The Annals of Applied Statistics, Vol. 2, No. 4, 1271–1289, 2008*

[Kar46] Karhunen, K., *'Zur Spektraltheorie Stochastischer Prozesse'. Ann. Acad. Sci. Fennicae, (1946), 37*

[KM11] Keng-Pei, L., Ming-Syan, C., *Efficient kernal approximation for large-scale support vector machine. Society for Industrial and Applied Mathematics Journal, 2011.*

[KS90] Kirby, M., Sirovich, L., *Application of the karhunen-loeve expansion procedure for the characterization of human faces. Pattern Analysis and Machine Intelligence, IEEE, Transcation on 12, Issue 1:103-108, 1990.*

[LN97] Li, T., North, G.R., *Aliasing effects and sampling theorems of spherical random fields when sampled on a finite grid. Ann. Inst.Statist. Math, Vol. 49, No 2, 341-354, 1997.*

[LRL11] Lindgren, F., Rue, H., Lindstorm, J. *An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. A journal of the Royal Statistical Society: Series B (Statistical Methodology, 73(4):423-498).*

[Loe55] Loève, M.M., *'Probability Theory', Princeton, N.J.: VanNostrand, 1955.*

[Mic98] Micheal, S. L., *Interpolation of Spatial Data: Some Theory for Kriging* . Springer Series in Statistics, 1999.

[New93] Newland, D. E., *An introduction to random vibration, spectral and Wavelet analysis.* New York: Wiley; 1993.

[Mui1920] Muir, T., *The Theory of Determinants in the Historical Order of Development (4 volumes).* Macmillan & Co. London: 1920

[Nie13] Nievergelt, Y., *Wavelets Made Easy.* Birkhauser, 2013.

[NWR02] Nychka, D., Wikle, C., Royle, A.J., *Multi-resolution models for non-stationary spatial covariance function.* Statistical Modelling (2002) 315-331, 2002.

[PHQ01] Phoon, K. K., Huang, S. T., Quek, S. T., *Convergence study of the truncated Karhunen-Loeve exapnsion for simulation of Stochastics processes.* International Journal for Numerical Methods in Engineering, 52:1029-1043, 2012.

[PHQ02] Phoon, K. K., Huang, S. T., Quek, S. T., *Implementation of Karhunen-Loeve expansion for simulation using a Wavelet-Galerkin scheme.* Probabilistic Engineering Mechanics 17 293-303, 2002.

[PAC18] Porcu, E. , Algeria, A., Crippa, P., *Axially symmetric models fo global data: A journey between geostatistics and stochastic generators.* Enviromentrices 30 (1), e2555, 2018.

[Sch04] Schmidt, T., *Credit Risk modeling Gaussian random fields. Working paper, 2004*

[Sch42] Schoenberg, I.J., *Positive definite functions on spheres.* Duke Math J.9, no. 1, 96-108, 1924.

[SCI17] Si, S., Cho-Jui, H., Inderjit, S.D., *Memory efficient kernel approximation. Journal of Machine Learning, Efficient kernal approximation for large-scale support vector machine, 2017.*

[Ste99] Stein, M., *Statistical Interpolation of Spatial Data: Some Theory for Kriging,. Springer, New York, 1999*

[Ste07] Stein, M.L., *Spatial variation of total column ozone on a global scale.* Annals of Applied Statistics, 191-210, 2007.

[TSRF15] Tortosa, R.B., Somet, J.P., Reyna, E.A., Fuentes J.J.M., *Complex kernels for proper complex-valued signals.* 23rd European Signal Processing Conference (EUSIPCO), 2015.

[Tri] Triacca, U., *Lesson 4: Stationary Stochastic Processes.*

[RS97] Ramsay, J.O., Silverman, B.W, *Functional data analysis. Spring-Verlag,1997*

[RW05] Rasmussen, C.E., Williams, C.K.I, *Gaussian Processes for machine learning. The MIT Press, Cambridge, Massachusetts, 2005*

[VWZ19] Vanlengenberg, C.D., Wang, W., Zhang, H., 2019. *Data generation for axially symmetric processes on the sphere.* Communication in Statistics-Simulation and Computation In Press.

[Van16] Vanlengenberg, C.D., *Data generation and estimation for axially symmetric processes on the sphere.* Ph.D Dissertation. UNC Greensboro, NC.

[Vid99] Vidakovic, B., *Statistical Modeling by Wavelets. Wiley, New York*

[WS01] Williams, C., Seeger, M., *Using the nystrom method to speed up kernel machines. Advances in Neural Information Processing Systems 13, pp. 682-688. MIT Press*

[Yag61] Yanglom, A. M., *Second-order homogeneous random fields.. In Proc. 4th Berkeley Sympos. Math Statist. and Prob, Volume 2 pages 593-622*

[Yan13] Yang, L., *Non-parametric and semi-parametric estimation of spatial covariance function.*

APPENDIX A

PROOF 1

Here is an example to demonstrate that a weakly stationary process may not necessarily be strictly stationary. Let $\{X_t; t \in \mathbb{Z}\}$ be a stochastic random process defined by

$$X_t = \begin{cases} Y_t, & \text{if } t \text{ is even} \\ \frac{1}{\sqrt{2}}(Y_t^2 - 1), & \text{if } t \text{ is odd} \end{cases}$$

where $Y_t \sim iid\ N(0,1)$, implying $E(Y_t) = 0$, $\text{var}(Y_t) = \text{E}(Y_t^2) = 1$, and $\text{var}(Y_t^2) = 2$ since $Y_t^2 \sim {\chi_1}^2$. Let's calculate the mean and the variance of the process X_t

$$\text{E}(X_t) = \begin{cases} \text{E}(Y_t) = 0, & \text{if } t \text{ is even} \\ \frac{1}{\sqrt{2}}\text{E}(Y_t^2 - 1) = 0, & \text{if } t \text{ is odd} \end{cases}$$

and

$$\text{var}(X_t) = \begin{cases} \text{var}(Y_t) = 1, & \text{if } t \text{ is even} \\ \frac{1}{2}\text{var}(Y_t^2 - 1) = 1, & \text{if } t \text{ is odd.} \end{cases}$$

Further, since for any $k \neq 0$ Y_t and Y_{t-k} are independent random variables, so is X_t and X_{t-k}. Hence we have

$$\text{cov}(X_t, X_{t-k}) = 0 \ \ \forall k.$$

Therefore, the process X_t is staionary with $X_t \sim WN(0, 1)$.

Now we know that, $P(X_t \leq 0) = P(Y_t \leq 0) = 0.5$ for t even and $P(X_t \leq 0) = P(\frac{1}{\sqrt{2}}(Y_t^2 - 1) \leq 0) = 0.6826$ for t odd. Hence, X_t is not identically distributed. This implies that the process is not strictly stationary (see [Tri] for more details).

APPENDIX B

PROOF 2

In this section, we will provide detailed calculations for Proposition 3.1. Here we assume the covariance function is given by (3.13). From (3.2), we have

$$\int_0^\pi (1 - e^{-|t|} - e^{-|s|} + e^{-|t-s|}) f_i(s) ds = \eta_i f_i(t),\ 0\ \leq t \leq \pi.$$

That is,

$$\int_0^\pi f_i(s)ds - \int_0^\pi e^{-|t|} f_i(s)ds - \int_0^\pi e^{-|s|} f_i(s)ds + \int_0^\pi e^{-|t-s|} f_i(s)ds = \eta_i f_i(t), \quad (2.1)$$

where $f_i(\cdot)$ and η_i denotes $i^{th}(i = 1, 2, \ldots)$ eigenfunction and eigenvalue of the covariance function, respectively. We can consider each term in (2.1).

$$\int_0^\pi e^{-|t-s|} f_i(s)ds = e^{-t} \int_0^t e^s f_i(s)ds + e^t \int_t^\pi e^{-s} f_i(s)ds$$
$$\int_0^\pi e^{-|t|} f_i(s)ds = e^{-t} \int_0^\pi f_i(s)ds,\ \text{since } t \in [0, \pi],$$
$$\int_0^\pi e^{-|s|} f_i(s)ds = \int_0^\pi e^{-s}\ f_i(s)ds.$$

Now (2.1) becomes

$$\int_0^\pi f_i(s)ds - e^{-t} \int_0^\pi f_i(s)ds - \int_0^\pi e^{-s} f_i(s)ds$$
$$+ e^{-t} \int_0^t e^s f_i(s)ds + e^t \int_t^\pi e^{-s} f_i(s)ds = \eta_i f_i(t). \quad (2.2)$$

Evaluating the first derivative with respect to t on both sides of the equation, we obtain

$$e^{-t}\int_0^{\pi} f_i(s)ds - e^{-t}\int_0^{t} e^s f_i(s)ds + e^{t}\int_t^{\pi} e^{-s} f_i(s)ds = \eta_i f_i'(t). \tag{2.3}$$

Differentiating both sides of equation (2.3) with respect to t for the second time, we obtain

$$-e^{-t}\int_0^{\pi} f_i(s)ds + e^{-t}\int_0^{t} e^s f_i(s)ds + e^{t}\int_t^{\pi} e^{-s} f_i(s)ds - 2f_i(t) = \eta_i f_i''(t). \tag{2.4}$$

From (2.2), (2.4) can be simplified as following.

$$(\eta_i - 2)f_i(t) - \int_0^{\pi} f_i(s)ds + \int_0^{\pi} e^{-s} f_i(s)ds = \eta_i f_i''(t). \tag{2.5}$$

Taking the derivative with respect to t on both sides of (2.5) for the third time, we obtain, noting $\eta_i > 0$,

$$\begin{aligned} (\eta_i - 2)f_i'(t) &= \eta_i f_i'''(t), \\ f_i'''(t) + \omega_i\ f_i'(t) &= 0, \text{ where } \omega_i = \frac{(2-\eta_i)}{\eta_i}. \end{aligned} \tag{2.6}$$

We have obtained a homogeneous third order differential equation. Based on the sign of ω_i we need to consider two cases of ω_i which will lead to two sets of eigensolutions.

Case 1 : $\omega_\mathbf{i} > \mathbf{0}$,

The characteristics equation for (2.6) becomes

$$r^3 + \omega_i r = 0,$$

resulting in three characteristics roots $r = 0, \pm\sqrt{\omega_i}i$. Thus, general solutions to (2.6) can be written as follows.

$$f_i(t) = C_{i1} + C_{i2}\sin(\sqrt{\omega_i}t) + C_{i3}\cos(\sqrt{\omega_i}t), \tag{2.7}$$

where C_{i1}, C_{i2} and C_{i3} are constants to be determined. Constructing adequate boundary conditions are required to determine the values of $C_{i1}, C_{i2},$ andC_{i3}. First explore the behavior of the function on the boundaries 0 and π by considering the function $f_i(t)$ given by (2.1).

When $t = 0$,

$$0 = \eta_i f_i(0) \implies f_i(0) = 0, \text{ since } \eta_i \neq 0. \tag{2.8}$$

When $t = \pi$,

$$\int_0^\pi f_i(s)ds - e^{-\pi}\int_0^\pi f_i(s)ds - \int_0^\pi e^{-s}f_i(s)ds + e^{-\pi}\int_0^\pi e^s f_i(s)ds = \eta_i f_i(\pi). \tag{2.9}$$

Considering equation (2.3), we have the following

When $t = 0$,

$$\int_0^\pi f_i(s)ds + \int_0^\pi e^{-s} f_i(s)ds = \eta_i f_i'(0). \tag{2.10}$$

When $t = \pi$,

$$e^{-\pi}\int_0^\pi f_i(s)ds - e^{-\pi}\int_0^\pi e^{s} f_i(s)ds = \eta_i f_i'(\pi). \tag{2.11}$$

Considering the equation (2.5), we obtain that

when $t = 0$,

$$-\int_0^\pi f_i(s)ds + \int_0^\pi e^{-s} f_i(s)ds = \eta_i f_i''(0), \tag{2.12}$$

when $t = \pi$,

$$-e^{-\pi}\int_0^\pi f_i(s)ds + e^{-\pi}\int_0^\pi e^{s} f_i(s)ds - 2f_i(\pi) = \eta_i f_i''(\pi). \tag{2.13}$$

From (2.7) and (2.8) one can easily see that

$$f_i(0) = 0, \quad f_i'(0) = C_{i2}\sqrt{\omega_i}, \quad f_i''(0) = -C_{i3}\omega_i.$$

For the case of $t = \pi$, from (2.7) we have

$$f_i(\pi) = C_{i1} + C_{i2}\sin(\sqrt{\omega_i}\pi) + C_{i3}\cos(\sqrt{\omega_i}\pi). \tag{2.14}$$

$$f_i'(\pi) = C_{i2}(\sqrt{\omega_i}\ \cos(\sqrt{\omega_i}\pi)) - C_{i3}(\sqrt{\omega_i}\ \sin(\sqrt{\omega_i}\pi)). \tag{2.15}$$

$$f_i''(\pi) = -C_{i2}(\omega_i\ \sin(\sqrt{\omega_i}\pi)) - C_{i3}(\omega_i\ \cos(\sqrt{\omega_i}\pi)). \tag{2.16}$$

From (2.14), (2.15), and (2.16), we have the following relations.

$$\eta_i f_i''(0) - \eta_i f_i''(\pi) + 2f_i(\pi) = \eta_i f_i(\pi),$$
$$f_i''(0) - f_i''(\pi) = \frac{2-\eta_i}{\eta_i} f_i(\pi) = \omega_i f_i(\pi).$$

However, if one plugs in the expression for $f_i(\cdot)$ and $f_i''(\cdot)$ and evaluates them at $t = 0$ and $t = \pi$, the above equation is always true. This implies that one of the equations (2.9), (2.12), and (2.13) is redundant. Here we remove equation (2.9) for further consideration.

Now we consider equation (2.12) and (2.13). We have

$$-\eta_i f_i'(\pi) - 2f_i(\pi) = \eta_i f_i''(\pi) \Rightarrow -\eta_i f_i'(\pi) - (2-\eta_i) f_i(\pi) - \eta_i f_i(\pi) = \eta_i f_i''(\pi),$$

or equivalently, after dividing η_i on both sides,

$$f_i'(\pi) + (\omega_i + 1) f_i(\pi) = -f_i''(\pi).$$

However, $\omega_i f_i(\pi) + f_i''(\pi) = f_i''(0) = -C_{3i}\omega_i$, hence

$$f_i'(\pi) + f_i(\pi) = C_{i3}\omega_i. \tag{2.17}$$

Equation (2.12) subtracting (2.10) gives

$$-2\int_0^{\pi} f_i(s)ds = \eta_i(f_i''(0) - f_i'(0)). \tag{2.18}$$

Adding (2.12) and (2.10) gives

$$2\int_0^{\pi} e^{-s} f_i(s)ds = \eta_i(f_i''(0) + f_i'(0)). \tag{2.19}$$

The three equations (2.17), (2.18), and (2.19) are considered to be the boundary conditions for our eigenvalue and eigenfunctions for the ordinary differential equation problem.

Our first aim is to find the explicit form for eigenvalues. We first convert (2.17) and (2.18) into following forms.

$$C_{i2}\left(\sqrt{\omega_i}\cos(\sqrt{\omega_i}\pi) + \sin(\sqrt{\omega_i}\pi)\right) = C_{i3}\left(\omega_i + \sqrt{\omega_i}\sin(\sqrt{\omega_i}\pi) - \cos(\sqrt{\omega_i}\pi) + 1\right), \tag{2.20}$$

$$C_{i2}(\eta_i\sqrt{\omega_i} - 2\int_0^{\pi}\sin(\sqrt{\omega_i}s)ds) = C_{i3}(-\eta_i\omega_i \ + 2\int_0^{\pi}(\cos(\sqrt{\omega_i}s) - 1)ds). \tag{2.21}$$

Dividing (2.20) by (2.21) we obtain (3.14). Note that by solving (2.20) and (2.21) homogeneous system of equations we can expect an infinite number of solutions for C_{i2} and C_{i3}. Here we explore the orthogonality of eigenfunctions to determine explicit non-zero values of $C_{i1}, C_{i2},$ and C_{i3}.

$$\int_0^{\pi} f_i^2(t)dt = \int_0^{\pi}\left(C_{1i} + C_{i2}\sin(\sqrt{\omega_i}t) + C_{i3}\cos(\sqrt{\omega}_i)\right)^2 dt = 1.$$

When $f_i(0) = 0$ we have

$$-C_{i1} = C_{i3}.$$

From equation (2.20) we obtain

$$C_{i2} = k_i \, C_{i3},$$

where

$$k_i = \frac{\left(\omega_i + \sqrt{\omega_i}\sin(\sqrt{\omega_i}\pi) - \cos(\sqrt{\omega_i}\pi) + 1\right)}{\left(\sqrt{\omega_i}\cos(\sqrt{\omega_i}\pi) + \sin(\sqrt{\omega_i}\pi)\right)} \text{ and } \omega_i = \frac{2 - \eta_i}{\eta_i}.$$

Now we have

$$\int_0^\pi f_i^2(t)d(t) = \int_0^\pi \left(-C_{3i} + k_i \;\; C_{i3}\sin(\sqrt{\omega_i}t) + C_{i3}\cos(\sqrt{\omega_i}t)\right)^2 d(t) = 1, \qquad (2.22)$$

giving

$$C_{i3}^2 = \frac{1}{\int_0^\pi \left(-1 + k_i\sin(\sqrt{\omega}_i t) + \cos(\sqrt{\omega}_i t)\right)^2 d(t)}. \qquad (2.23)$$

Hence,

$$f_i(t) = -C_{i3} + k_i C_{i3}\sin(\sqrt{\omega_i}t) + C_{i3}\cos(\sqrt{\omega}_i t).$$

This proves the first part of Proposition 3.1.

Case 2 : $\omega_i < 0$.

In this case, the following characteristic equation

$$r^3 + \omega_i r = 0,$$

results in three characteristics roots $r = 0, \pm\sqrt{-\omega_i}$. Thus, general solutions can be written as

$$f_i(t) = A_{i1} + A_{i2}\sinh(\sqrt{-\omega_i}t) + A_{i3}\cosh(\sqrt{-\omega_i}t). \tag{2.24}$$

Taking the first derivative of $f_i(t)$ gives

$$f_i'(t) = A_{i2}\sqrt{-\omega_i}\cosh(\sqrt{-\omega_i}t) - A_{i3}\sqrt{-\omega_i}\sinh(\sqrt{-\omega_i}t).$$

Taking the second derivative with respect to t , we obtain

$$f_i''(t) = A_{i2}(-\omega_i)\sinh(\sqrt{-\omega_i}t) + A_{i3}(-\omega_i)\cosh(\sqrt{-\omega_i}t).$$

We can notice that

$$f_i(0) = A_{i1} + A_{i3} = 0, \quad f_i'(0) = A_{i2}\sqrt{-\omega_i}, \quad f_i''(0) = A_{i3}(-\omega_i).$$

We can adjust same boundary conditions that we used in Case 1 to get

$$A_{i2}\left(\sqrt{-\omega_i}\cosh(\sqrt{-\omega_i}\pi) + \sinh(\sqrt{-\omega_i}\pi)\right) = A_{i3}\left(\omega_i - \sqrt{-\omega_i}\sinh(\sqrt{-\omega_i}\pi) - \cosh(\sqrt{-\omega_i}\pi) + 1\right). \tag{2.25}$$

$$A_{i2}(\eta_i\sqrt{-\omega_i} - 2\int_0^{\pi}\sinh(\sqrt{-\omega_i}s)ds) = A_{i3}(-\eta_i\omega_i\ + 2\int_0^{\pi}(\cosh(\sqrt{-\omega_i}s) - 1)ds). \tag{2.26}$$

Dividing (2.25) by (2.26), then we obtain (3.15). A similar argument under the case $\omega_i > 0$ can be adopted to the case $\omega_i < 0$, and with the property of the normalized eigenfunctions, we have

$$\int_0^{\pi} g_i^2(t) = \int_0^{\pi}\left(A_{i1} + A_{i3}\sinh(\sqrt{-\omega_i}t) + A_{i3}\cosh(\sqrt{-\omega}_i t)\right)^2 = 1. \tag{2.27}$$

Based on the boundary conditions we can derive

$$-A_{i1} = A_{i3}$$
$$A_{i2} = k_i'\ A_{i3},$$

where

$$k_i' = \frac{(\omega_i - \sqrt{-\omega_i}\sinh(\sqrt{-\omega_i}\pi) - \cosh(\sqrt{-\omega_i}\pi) + 1))}{(\sqrt{-\omega_i}\cosh(\sqrt{-\omega_i}\pi) + \sinh(\sqrt{-\omega_i}\pi))} \text{ and } \omega_i = \frac{2-\eta_i}{\eta_i}.$$

Therefore, equation (2.27) can be written as,

$$\int_0^{\pi}\left(-A_{i3} + k_i'\ A_{i3}\sinh(\sqrt{-\omega}t) + A_{i3}\cosh(\sqrt{-\omega_i}t)\right)^2 = 1,$$

with

$$A_{i3}^2 = \frac{1}{\int_0^{\pi}\left(-1 + k_i'\sinh(\sqrt{-\omega}_i t) + \cosh(\sqrt{-\omega}_i t)\right)^2}.$$

Hence,

$$g_i(t) = -A_{i3} + k' A_{i3} \sinh(\sqrt{\omega_i} t) + A_{i3} \cosh(\sqrt{\omega_i} t).$$

This concludes the proof of Proposition 3.1.