République algérienne démocratique et populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université Frères Mentouri Constantine 1

Faculté des sciences exactes

Département de mathématiques

N° d'ordre : 116/DS/2022

N° de série : 06/MATH/2022

THÈSE

PRÉSENTÉE POUR L'OBTENTION DU DIPLÔME DE DOCTORAT EN SCIENCES EN MATHÉMATIQUES

Inférence Statistique Asymptotique dans les Processus ARMA Fractionnaire

Présentée par

LAALA Barkahoum

Spécialité

Probabilité et Statistique

Devant le jury :

Président	Mme H. MERABET	Prof.	Université frères Mentouri, Constantine 1.
Encadreur de thèse	Mme S. BELALOUI	Prof.	Université frères Mentouri, Constantine 1.
Examinateur	Mr A. M. ELSAWAH	Prof.	BNU-HKBU United International College, China.
Examinateur	Mr A. MEMOU	Prof.	Université Mohamed Boudiaf, Msila.
Examinatrice	Mme M. CHIKHI	M.C.A.	Université frères Mentouri, Constantine 1.
Examinateur	Mr M. BOUKALOUA	M.C.A.	École Nationale Polytechnique, Constantine.

Soutenue le : 22 / 12 /2022

Contents

List of Figures

List of Tables

Nul ne peut cacher les rayons du soleil.

No one can hide the sun's rays.

Acknowledgments

First of all I would like to thank " God" who gave me the strength to complete this work, Elhamedou Ellah.

Most importantly, I am grateful for my family's unconditional, unequivocal, and loving support.

A big thank to "Ahmed ELSAWAH" who helped me with all his strenght during in the long nights and gave me the certainy of the quality of this work. He helped me with his patience and his encouragement to persevere in the development of this work and its content and to get it out in the best form. It is a great pleasure to thank :

• Prof. S. Belaloui, for her help and guidance throughout the course of this work;

• Prof. Ameur Memou, who gave me the decision to engage in this way;

• I also wish to thank, Mss CHIKHI and MERABET, who accepted to examine my thesis, without forgetting Mr. BOUKALOUA.

I believe the idea of Jean-Jaques Roussou

" La connaissance est bien un devoir qu'il faut rendre"

List of Notation

$\mathbb{C}$ set of complex numbers

$\mathbb{N}$ set of natural numbers

$\mathbb{R}$ set of real numbers

$\mathbb{Z}$ set of integers

$[x]$ largest integer smaller or equal to $x \neq 0, x \in \mathbb{R}$

$E(.)$ expectation operator

$\bar{Z}$ complex conjugate of Z

$var(.)$ variance operator

$cov(.,.)$ covariance operator

$\mathcal{A}(h)$ autocovariance at lag h

$\mathcal{N}(\mu, \sigma^2)$ normal distribution with mean μ and variance σ^2

$\chi^2(n)$ chi-square distribution with n degrees of freedom

$\xrightarrow{D}$ convergence in distribution

$a \stackrel{D}{=} b$ means a and b are equal in distribution

$\simeq$ approximately equal

$a := b$ a is defined to equal b

$a_n \sim b_n$ a_n/b_n converges to 1 as $n \to \infty$

$\widehat{\theta}$ estimator of parameter θ

Acronyms

AR Auto-Regressive

ACF Auto-Covariance Function

ARFIMA Auto-Regressive Fractionally Integrated Moving Average

ARMA Auto-Regressive Moving Average

CLT Central Limit Theorem

DFT Discrete Fourier Transform

DOM Degree Of Memory

DSGP Discrete Stationary Gaussian Process

GP Gaussian Process

LWE Lag Window Estimator

MA Moving Average

i.i.d Independently and Identically Distributed

MSE Mean Squared Error

OLS Ordinary Least Squares

PE Periodogram Estimator

SDF Spectral Density Function

REM Regression method

rvs random variables

WEM local Whittle method

WN White Noise

Abstract 1. *Gaussian process (GP) is a stochastic process that has been successfully applied in finance, black-box modeling of biosystems, machine learning, geostatistics, multitask learning or robotics and reinforcement learning. Effectively estimating the spectral density function (SDF) and degree of memory (DOM) of a long-memory stationary GP (LMSGP) is a significant hard problem investigators may face. This thesis gives some new sufficient conditions (NSCs) for improving the lag window estimators (LWEs) of the SDF and DOM for LMSGPs. A comparison study among the behavior of the LWEs under the NSCs, the LWEs without the NSCs and the existing widely used periodogram estimators (PEs) is given. The theoretical and computational justifications show that: the LWEs under the NSCs are better than the LWEs without the NSCs; the LWEs under the NSCs are better than the PEs; the LWEs under the NSCs are asymptotically unbiased and consistent; the asymptotic distributions of the LWEs under the NSCs of the SDF and DOM under the NSCs are chi-square and normal, respectively; the LWE of the DOM under the NSCs has a fast vanishing variance under the regression method; and the LWEs under the NSCs improve the finite sample properties for the regression and local Whittle estimation methods.*

Key words: Gaussian process; spectral density; degree of memory; lag window; periodogram; local Whittle method; regression method.

Résumé 1. *Le processus gaussien (GP) est un processus stochastique qui a été appliqué avec succès dans le domaine de la finance, la modélisation en boîte noire des biosystèmes, l'apprentissage automatique, la géostatistique, l'apprentissage multitâche ou de la robotique et l'apprentissage par renforcement. L'estimation efficace de la fonction de densité spectrale (SDF) et du degré de mémoire (DOM) d'un GP stationnaire à longue mémoire (LMSGP) est un problème important et difficile auquel les chercheurs sont confrontés. Ce travail donne quelques nouvelles conditions suffisantes (NSC) pour améliorer les estimateurs de fenêtre de décalage (LWE) de la SDF et du DOM pour les LMSGP. Une étude comparative entre le comportement des LWE sous et sans les NSC, et le périodogramme (PE), qui est largement utilisé, est faite en détail. Les justifications théoriques et numériques montrent que: les LWE sous NSC sont meilleurs que les LWE sans NSC et le PE; les LWE sous les NSC sont asymptotiquement sans biais et consistants; les distributions asymptotiques des LWE de la SDF et DOM sous les NSC sont respectivement du chi deux et de la normale; le LWE du DOM sous les NSC a une variance à disparition rapide selon la méthode de régression; et les LWE sous les NSC améliorent les propriétés de l'échantillon pour les méthodes de régression et d'estimation de locale Whittle.*

Mots clés: Processus Gaussien; densité spectrale; degré de mémoire; fenêtre de décalage; periodogramme; méthode de locale Whittle ; méthode de régression.

Introduction

The purpose of statistical studies is to make judgments about a population based on observations drawn from it. Series of observations which are indexed by time and can be observed continuously or discretely, they are called a time series. Financing models, for example, are often based on a continuous assumption. Because, in the stock market, the transactions appear very close to each other. Contrariwise, macroeconomic data is usually observed at a discrete time, in a period of a month, a quarter or even a year. As a title of an example in the economic and social world: inflation, unemployment, production, exports, birth rate, immigration, education, housing, etc [37].

Time series, when it is not deterministic, are seen as the realization of a stochastic process. Whereas, the stochastic process is a collection of random variables on a probability space. That is, if we take any realization of such a process and divide it up into a number of time intervals, the various section of realization look pretty much the same. We express this type of behaviour more precisely by saying that, in such cases, the statistical properties of the process do not change over time, they are the same at all time point. Processes which have this property are called stationary process, the opposite case is the non stationary, its main is to facilitate the extraction mathematical theories (see [42, 10]). Analysis of stationary process is carried out in two complementary approaches (domains), temporal and spectral (see [26, 30]). Spectral analysis reveals the correspondence between the spectral domain and the temporal domains for a stochastic process, which includes the spectral density function (SDF) and the autocovariance function(ACF) respectively. A stationary process with slowly decaying ACF is therefore called a stationary process with long memory or long-range dependence or strong dependence (in contrast to processes with summable correlations which are also called processes with short memory or short-range correlations or weak dependence). Briefly, we examine the dependence structure of these time series by considering plots (in log-log coordinates) of the sample ACF against the lag, the variance of the sample against the length of series, or the sample SDF against the frequencies. Yearly minimal water levels, of the Nile River, are the typical example how displays the strong dependence structure.

Gaussian process (GP) is a useful stochastic process, it used in different fields and contain less conditions. Therefore, it is necessary in the treatment of many real cases.

The main objective addressed in this thesis is the estimation of the SDF and the degree of memory (DOM) of a stationary long-memory second-order Gaussian process (LMDSGP) which has long remained an interesting and challenging problem to a long time, it posed in the field of inferential statistics.

Therefore, for well explain the previous ideas, the thesis has been divided into: an introduction, four chapters, a conclusion, an appendix and bibliographical references.

• Chapter 1 gives the basic statistical and analytical tools which are used to explain, study, and develop statistical inferences from long memory discrete stationary Gaussian processes (LMDSGPs). Moreover, to rigorously develop the asymptotic theory, we have introduced definitions on a slowly evolving functions, summability techniques and estimation methods.

• Chapter 2 contains an in-depth discussion of stationarity and its role in defining the ACF and the SDF. Sequentially, we deal the DSGPs and their properties, as a processes that is easy to interpret its parameters and essential in our studies (to be seen). We end this chapter by distinguishing the type of process dependency, short memory discrete stationary Gaussian process (SMDSGP) and LMDSGP in the time domain, ACF, and the spectral domain, SDF. The SMDSGP is separated from the LMDSGP with a decay rate of ACF towards zero, all this is illustrated by examples.

• Chapter 3 focuses on statistical inference of LMDSGPs in the spectral domain. We start by introducing lag window estimator (LWE) and its characteristics, then we present some theoretical justifications giving an overview of the behavior of estimation of the spectral density function by lag window estimator (LWE-SDF) and the estimation of degree of memory using the lag window estimator (LWE-DOM). Then, a clear improvement is brought on the estimators in question and this by introducing sufficient conditions. The chapter ends with a comparative study between the behaviors of LWE, under and without these conditions and of periodogram (PE), to evaluate the beneficial effect of these conditions on the estimators mentioned.

• Chapter 4 concretizes the theoretical results of the optimality of LWE, under established conditions, by simulation. The programs used in the simulation are algorithms directly implemented on the methods cited in the work in the MATALB language. The end of the thesis contains an appendix where the demonstrations are detailed.

Chapter 1

Preliminaries

1.1 Context

Optimal data collection and efficient data analysis are the two significant key issues to understand the behavior of complex experiments for real-life phenomena, industrial applications and scientific investigations [16, 17]. Efficient methods are able to capture maximum valuable (accurate) information about the behavior of a given experiment and thus more significant unknown parameters can be estimated without bias and with minimum variance, whereas non-efficient methods cannot produce accurate information nor provide good estimators for the unknown parameters [18, 19]. The practice demonstrated that effectively estimating unknown parameters is a significant hard problem experimenters may face in many real-life experiments. Even though there are several techniques provided to estimate unknown parameters, the challenge faced by the experimenters is still daunting. Gaussian process is a stochastic process that has been successfully used in finance, black-box modelling of biosystems [3], machine learning [45], geostatistics [12], multitask learning or robotics [2] and reinforcement learning [13].

In this work, it must define two concept, spectral analysis and dependence structure, that they help us to meet the main objective of this thesis.

1.1.1 Spectral Analysis

The spectral analysis is a main components in the studied data, it reveals hidden periodicities, which are to be associated with cyclic behavior or recurring processes. It considers the problem of determining the spectral analysis (i.e., the distribution over frequency) of a process from a finite set of measurements, either nonparametric or parametric techniques. Spectral analysis is divided into two major areas, a Fourier transform or a spectral density function (SDF).

(1) When the data contains no random effects or noise, it is called deterministic. In this case, one computes a Fourier transform.

(2) One computes a SDF when random effects obscure the desired underlying phenomenon.

The spectral analysis gives information and other characteristics of the process under study, it finds in many divers fields:

- i) In economics, meteorology, astronomy and several other fields, the spectral analysis may reveal "hidden periodicities" in the studied data, which are to be associated with cyclic behavior or recurring processes.
- ii) In radar and sonar systems, the spectral contents of the received signals provide information on the location of the sources (or targets) situated in the field of view.
- iii) In medicine, spectral analysis of various signals measured from a patient, such as electrocardiogram (ECG) or electroencephalogram (EEG) signals, can provide useful material for diagnosis.
- iv) In seismology, the spectral analysis of the signals recorded prior to and during a seismic event (such as a volcano eruption or an earthquake) gives useful information on the ground movement associated with such events and may help in predicting them. Seismic spectral estimation is also used to predict subsurface geologic structure in gas and oil exploration.
- v) In control systems, there is a resurging interest in spectral analysis methods as a means of characterizing the dynamical behavior of a given system, and ultimately synthesizing a controller for that system.

The history of spectral analysis as an established discipline started more than a century ago with the work by Schuster on detecting cyclic behavior in time series. An interesting historical perspective on the developments in this field can be found in Marple (1987) [38]. This reference notes that the word "spectrum" was apparently introduced by Newton in relation to his studies of the decomposition of white light into a band of light colours, when passed through a glass prism. This word appears to be a variant of the Latin word "spectre" which means "ghostly apparition". The contemporary English word that has the same meaning as the original Latin word is "spectre".

1.1.2 Measure of Dependence Structure

A stochastic process is a sequence of random variables that are dependent in time. The long memory (or long-range dependence) process has become a rapidly developing subject. Because of the diversity of applications, the literature on the topic is broadly scattered in a large number of works.A particular measure of dependence structure is the covariance which specifies the linear part of the dependence. The variables X and Y are said to be correlated if their covariance function:

$$cov(X,Y) := E(XY) - E(X)E(Y) \neq 0,$$

and they are uncorrelated if their covariance function vanishes:

$$X \text{ and } Y \text{ are uncorrelated} \Longrightarrow cov(X,Y) = 0.$$

The opposite statement is not necessarily true, i.e.,

$$cov(X,Y) = 0 \nRightarrow X \text{ and } Y \text{ are uncorrelated.}$$

Therefore, we showed as follows

$$cov(X,Y) \neq 0 \Longrightarrow X \text{ and } Y \text{ are correlated} \Longrightarrow \text{dependent},$$

$$X \text{ and } Y \text{ are independent} \Longrightarrow \text{uncorrelated} \Longrightarrow cov(X,Y) = 0.$$

Independent random variables $(X_t)_{t \in T}$ with identical distribution function are referred to as independent and identically-distributed random variables (i.i.d). There exist various characteristics describing the dependence structure of a stochastic process, which can be placed between the two extreme scenarios.

- Scenario 1: The observations are dependent, due to the nature of the observed phenomenon and/or the way observations are taken. It is a completely dependent sequence $\{X_j \equiv X, j \in \mathbb{Z}\}$, where X is a given random variable, which allows for only trivial inference.

- Scenario 2: The observations are expected to be (more or less) independent. It corresponds to an i.i.d. sequence $\{\varepsilon_j\} = \{\varepsilon_j, j \in \mathbb{Z}\}$ with mean μ and variance σ^2, i.e., $\{\varepsilon_j\} \sim \text{i.i.d}(\mu, \sigma^2)$.

Between these two extremes are many other stochastic processes, there is a mixing process.

Definition 1. *For stochastic processes, "mixing" means "asymptotically independent": that is, the statistical dependence between $X(t_1)$ and $X(t_2)$ goes to zero as $|t_1 - t_2|$ increases. To make this precise, we need to specify how we measure the dependence between $X(t_1)$ and $X(t_2)$.*

In particular, α-mixing and m-dependent processes defined as follows. Let

$$\alpha(k) := sup\left\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_0^-(X), B \in \mathcal{F}^+{}_k(X)\right\},$$

where $\mathcal{F}_k^-(X)$ and $\mathcal{F}_k^+(X)$ are the σ-fields generated by the "past information" $X_s,\ s \leq k$ and the "future information" $X_s, s > k$, respectively.

Definition 2. *A stochastic process $\{X_j\}$ is said to be α-mixing if*

$$\alpha(k) \to 0, \quad as \quad k \to \infty.$$

Given a positive integer m, a stationary process $\{X_j\}$ is called m-dependent if

$$\alpha(k) = 0, \quad \forall k > m.$$

In many cases, α-mixing processes $\{X_j\}$ have asymptotically similar properties as autoregressive moving average (ARMA) process and Markov process, including the fast decay of dependence and the correlation between observations X_j and X_k, as the distance $|j - k|$ in time increases.

Under m-dependence, the collections of variables $\{X_s, s \leq k\}$ and $\{X_s, s > k + m\}$ are independent for any $k \in \mathbb{Z}$, i.e., independence of the times series up to time m.

- A simple example of an m-dependent process is a moving-average process (MA)

$$X_j = a_0\varepsilon_j + \cdots + a_m\varepsilon_{j-m} = \sum_{i=0}^{m} a_i\varepsilon_{j-i}, \quad \text{where } \{\varepsilon_j\} \sim \text{i.i.d}(0, \sigma^2).$$

- It is clear that, an i.i.d. sequence is 0-dependent.

The rate of decay of the mixing coefficients $\alpha(k)$, as $k \to \infty$, characterizes the degree of dependence between "past" and "future" and the distant observations in time. However, it does not impose any additional assumptions on the structure of the process $\{X_j\}$. Various other measures of dependence and classes of mixing processes have been introduced in the literature, for more details see[23]. As a rule, mixing conditions are not easy to verify and for concrete classes of processes they may be too restrictive. The dependence structure of the time series, by considering plots (in log-log coordinates), can be determined by:

(1) The variance of $\bar{X_\gamma}$ against the γ, where $\bar{X_\gamma} = \gamma^{-1}\sum_{t=1}^{\gamma} X_t$,

(2) The estimator of ACF $\widehat{\mathcal{A}(h)}$ against the lag h, or

(3) The estimator of SDF $\widehat{\mathcal{S}(\eta_i)}$ against at Fourier frequencies $\eta_i = \frac{2\pi i}{\gamma}$.

The meaning of the relationship between these properties is explained as follows

$$var(\bar{X_\gamma}) = var\left(\gamma^{-1}\sum_{t=1}^{\gamma} X_t\right) = \gamma^{-2}\sum_{t,k=1}^{\gamma} cov(X_t, X_k). \tag{1.1}$$

(a) If the series are independent (so it's uncorrelated), i.e., $cov(X_t, X_k) = \begin{cases} 0 & \text{if } t \neq k, \\ var(X_t) & \text{for } t = k. \end{cases}$

Then for (1.1),

$$var(\bar{X_\gamma}) = \gamma^{-2}\sum_{t=1}^{\gamma} var(X_t) = \gamma^{-1} var(X_\gamma).$$

(b) In case of a dependent process $X = \{X_t,\ t \in \mathbb{Z}\}$, if the series are correlated and the ACF, $cov(X_t, X_k)$, depend only on the lag, $|t-k|$, then (1.1) can be simplified by

$$var(\bar{X}\gamma) = \gamma^{-2}\sum_{h=0}^{\gamma-1}(\gamma - h)\, cov(X_0, X_h) = \gamma^{-1}\left(var(X_\gamma) + 2\sum_{h=1}^{\gamma-1}\left(1 - \frac{h}{\gamma}\right)\widehat{\mathcal{A}(h)}\right). \tag{1.2}$$

These formulas will be well detailed in the following sections.

1.1.3 Empirical Example

Discrete stationary Gaussian processes (DSGPs) are widely used in many real-life applications, including hydrology, geophysics, economics, econometrics, ecology and telecommunication traffic (cf. [4, 22, 47, 48, 60]).

The famous example in this field is the yearly minimal water levels of the Nile River for the years 622-1281, whose water level has been measured at the Roda Gauge near Cairo in Egypt(Tousson, 1925, pp. 366-385), it is one of the real-life applications of the long memory DSGP (LMDSGP). This set of data describes the long periods of dryness which followed by long periods of yearly returning floods.

Historically, The analysis of this and several similar time series led to the discovery of the so-called Hurst effect (Hurst 1951). It exhibits a long-term behavior that might give an explanation of the seven good years and seven bad years described in Genesis. There were long periods where the maximal level tended to stay high. On the other hand, there were long periods with low levels. Overall, the series looks stationary. When one only looks at short time periods, then there seem to be cycles or local trends. However, looking at the whole series, there is no apparent persisting cycle. It rather seems that cycles of (almost) all frequencies occur, superimposed and in random sequence. Also, there is no global trend. In reference to the biblical seven years of great abundance and seven years of famine, Mandelbrot called this behavior the Joseph effect (Mandelbrot 1977, 1983a, Mandelbrot and Wallis 1968a,b, Mandelbrot and van Ness, 1968). Motivated by Hurst's empirical findings, Mandelbrot and co-workers later introduced fractional Gaussian noise (FGN) as a statistical model with long memory (see, e.g., Mandelbrot and Wallis 1968 a, 1969 a, Mandelbrot and van Ness 1968).

The presence of long memory in this data set, it can be indicate that the variance of $\bar{X}_n$ converges to zero at a slower rate than n^{-1}. Or, due to the slow decay of the ACF against the lag which corresponds to the property (2.12). The slope of the ln(ACF) relative to ln(lag) is represented by a single parameter which is the degree of memory (DOM) (cf. Figure 1.4a in [6]). For more details and more real-world examples, the interested reader can refer to [6] and [28].

1.2 Related Work and Problem

In this thesis, we are mainly interested in a class of LMDSGP whose ACF varies regularly to infinity with the DOM parameter. We are particularly interested in the estimation of the SDF, $\mathcal{S}(.)$, and the DOM, d. There are several methods to estimate them. The estimation of the SDF and DOM of a LMDSGPs are the most significant hard problem in this regard. The periodogram technique (PT) is the classical technique for estimating the SDF. The practice demonstrated that, the PT has some defects, such as (cf. [6, 42])

a satisfactory estimator is not guaranteed; the resultant periodogram estimator (PE) is inconsistent; and the PE has an erratic and wild fluctuating form. Lag window technique (LWT) consists of windowing the autocorrelation coefficients prior to estimating the SDF that is shown to be a special case of smoothing a SDF estimator by giving decreasing weight to the autocovariances as the lag increases. The weighting function is known as the lag window (kernel) and leads to a smoothed SDF estimator (SDE). Some improvements of the LWT for effectively estimating the SDF and DOM are investigated in this study (cf. Section 3).

1.2.1 SDF Estimation

The goal of spectral estimation is to describe the distribution (over frequency) of the power contained in a process, based on a finite set of data. Estimation of SDF (the power spectra) is useful in a variety of applications, including the detection of signals buried in wide-band noise. In nonparametric methods, concerning the SDF estimation, are based on the discrete Fourier transform (DFT). In these methods no need to obtain the parameters of the time series. All these methods have the advantage of possible implementation using the DFT, but with the disadvantage in the case of short data lengths of limited frequency resolution. Parametric methods on the other hand can provide high resolution in addition to being computationally efficient. The most common parametric approach is to derive the spectrum from the parameters of an autoregressive model of the process ([43]). The most nonparametric methods include :

(1) The PE is the classical technique for estimating the SDF. The practice demonstrated that, the PE has some defects, such as (cf. [6, 42]) a satisfactory estimator is not guaranteed; the resultant PE is inconsistent; and it has an erratic and wild fluctuating form.

(2) Bartlett method divide the signal into blocks, find their PEs and average to get the SDE (The data segments are non-overlapping). The final effect is true SDF convoluted with a window. Due to windowing (leakage frequency due to side lobes) the frequency resolution is low [5, 51].

(3) Welch Method can be overlapping and window the data segments before computing PE (we may use different windows for each segment). This method has better precision but less frequency resolution than Bartlett method([52]).

(4) Blackman-Tukey Method windowed the ACF and take Fourier transform to get SDF estimator, in effect we smooth out the PE. It has better variance (even at large lags) and better precision than above two methods, but frequency resolution is less than the others. The estimates are based entirely on a finite record of data, the frequency resolution is equal to the spectral width of rectangular window of length N, which is approximately $1/N$. The estimates are computed at discrete frequencies [51].

(5) The LWT consists of windowing the autocovariance coefficients prior to estimating the SDF that is shown to be a special case of smoothing a SDE by giving decreasing weight to the autocovariances as the lag increases. The weighting function is known as the lag window (kernel) and leads to a smoothed SDF estimator. The LWT is the same as the Blackman-Tukey Method, but for various windows[42, 51, 10].

1.2.2 DOM Estimation

The semi-parametric methods allow to estimate the DOM without completely specifying the distribution of the process (and in particular, the SDF of the process). The idea is to consider the estimation of the DOM as the estimation of a parameter of interest, in the weight of a prior infinite-dimensional nuisance parameter, which is the spectrum of the "short-memory" part of the process.

(1) The most classic and well-known method is that of the regression of the log-periodogram introduced by [21], which consists in regressing the log-periodogram with respect to the logarithm of the frequency normalized in a neighborhood of zero frequency. The asymptotic normal and the consistency of this estimator were established by Robinson (1995a) [47] for LMDSGP.

(2) Another commonly used estimator of the DOM is the local Whittle estimator proposed by Künsch (1987)[34]. Instead of regressing the logarithm of the periodogram to zero frequency, this method consists in using a local approximation of the Gaussian likelihood. The consistency and asymptotic normality of this estimator were established in [48] for LMDSGP.

Bibliographical Notes

We summarize some research work concerning the SDF and DOM estimators applied in LMDSGP and short memory DSGP (SMDSGP)

• Hunt et al. (2003) [31] are derived an approximation bias, in the lag window estimator (LWE), of the DOM in ARFIMA models. By simulation, the expression obtained is compared with the observed bias, for the various windows.

• For long memory process, the asymptotic behavior of the DFTs has received a good amount of attention in recent years. Among other reasons, this may be attributed to the important role played by the DFTs in the semiparametric estimation of the DOM parameter d. See, for example, Geweke and Porter-Hudak (1983)[21], Robinson (1995)[50], Hurvich, Deo and Brodsky (1998)[33] and the references therein.

• For processes $\{X_t\}$ having the SDF defined by (2.13) with $d \in (0, 1/2)$ and with a bounded $\mathcal{V}(.)$, Yajima (1989) and Pham and Guégan (1994) established asymptotic normality and asymptotic independence of the DFTs at a finite set of ordinates that are asymptotically distant. In an important work, Robinson (1995)[50] proved that for a stationary process $\{X_t\}$ having SDF $\mathcal{S}$ of the form (2.11).

• For short memory process, Beran[6] and Priestly[42] have treated the the asymptotic behavior of the SDF estimator using the LWT.

• In Hunt[31] made the simulation study for LWT for the long memory process and in Hassler[28] has collected everything related to the long memory process

• Lahiri (2003)[35] provide a characterization of asymptotic independence of the DFTs in terms of the distance between their arguments under both cases, SMDSGP and LMDSGP. We represent the above works with a diagram to clearly the researches. Such as,η is the frequency, the Fourier frequency defined by $\eta_j = \frac{2\pi j}{\gamma}$, n_γ the length of window and γ is the sample size.

1.3 Basic Concept

1.3.1 Slowly Varying Functions

A slowly varying function is a function of a real variable whose behaviour at infinity is in some sense similar to the behaviour of a function converging at infinity.

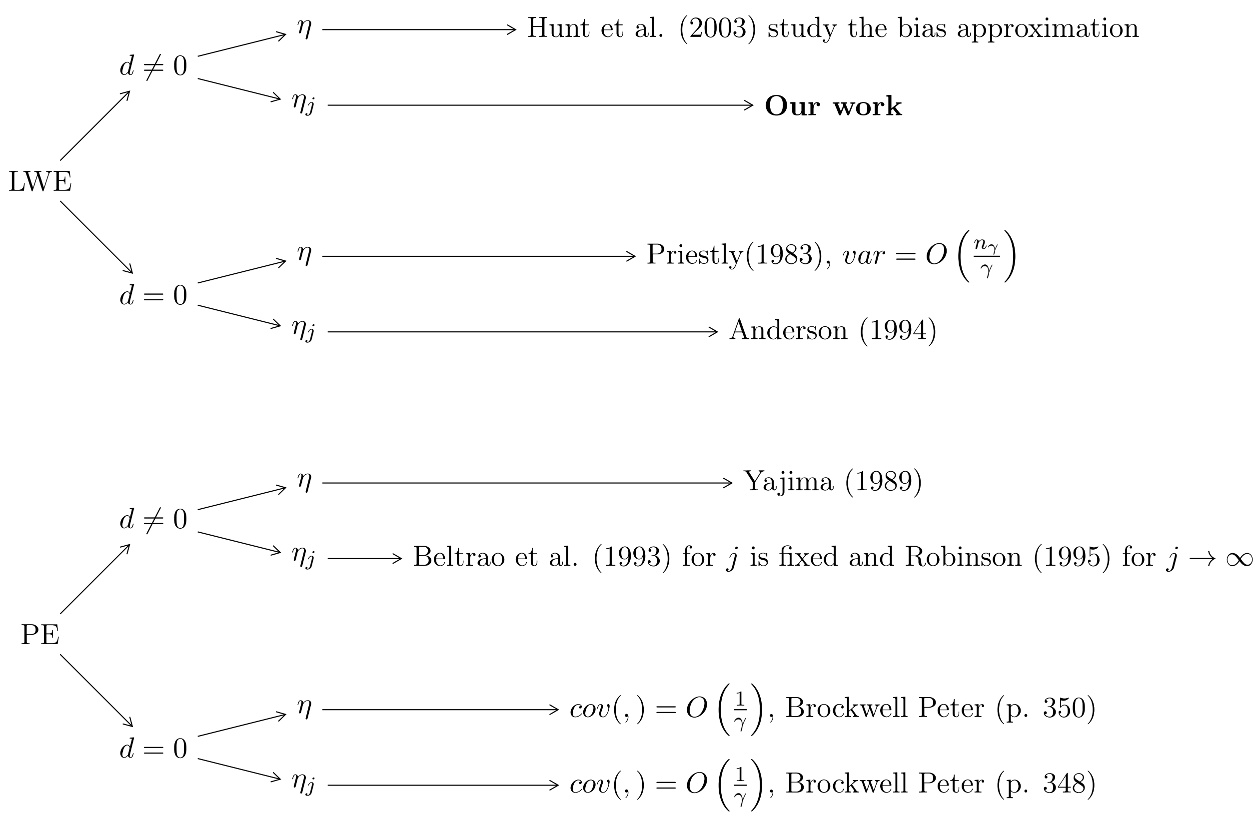

Figure 1.1: Illustrative diagram of previous works concerning spectral studies

Definition 3. *A measurable function $L : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is called slowly varying (at infinity),*

$$\forall a > 0, \ \lim_{x \to \infty} \frac{L(ax)}{L(x)} = 1.$$

Note

- A function $L : (0, \infty) \longrightarrow \mathbb{R}_+$ is called a regularly varying if,

$$\forall a > 0, \quad \lim_{x \to \infty} \frac{L(ax)}{L(x)} = g(a) < \infty.$$

- Every regularly varying function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is of the form, $f(x) = x^{\beta} L(x)$ where, $\beta \in \mathbb{R}$ and L is a slowly varying function.

- L is slowly varying at the origin and $\widetilde{L}$ is slowly varying at infinity, if $\widetilde{L}(x) = L(x^{-1})$.

Examples

The constant function is trivially slowly varying. Moreover, any function with a strictly positive finite limit is slowly varying, i.e., if L has a limit, $\lim_{x \to \infty} L(x) = b \in \mathbb{R}_+$, is a slowly varying function. More interesting examples are:

$$* \log x; \quad * \ \log \log x; \quad * \ \exp(\log x)^b, \quad b \in (0; 1).$$

The function $L(x) = x$ is not slowly varying, neither is $L(x) = x^b$ for any real $b \neq 0$. However, these functions are regularly varying. Among its advantages. Let finally $f(t)$ be integrable in every finite interval and let

$$T(x) = \int_0^{\infty} f(t) L(xt) dt.$$

The study of the asymptotic behavior of $T(x)$, or more precisely of the relation

$$T(x) \simeq L(x) \int_0^{\infty} f(t) dt, \quad \text{as } x \to 0, \quad \text{or as } x \to \infty.$$

For more informations, see (cf, [53]).

1.3.2 Periodic Function

A periodic function is a function that repeats its values at regular intervals. For example, the trigonometric functions, which repeat at intervals of 2π, are periodic functions. Periodic functions are used throughout science to describe oscillations, waves, and other phenomena that exhibit periodicity.

Definition 4. *A function f is said to be periodic if, for some nonzero constant P, it is the case that* $f(x+P) = f(x)$, *for all values of x in the domain.*

In parallel, the periodic cycles function are trigonometric functions like

$$f(t) = a\cos(\eta t) + b\sin(\eta t) = a\cos(\eta(t+P)) + b\sin(\eta(t+P)) = f(t+\eta P),$$

where P stands for the period. The period P and the frequency η are inversely related. By the properties of the sine and cosine, we have

$$\eta P = 2\pi \implies P = 2\pi/\eta, \quad \text{as long as} \quad \eta \neq 0. \tag{1.3}$$

In particular, we will focus on the so-called harmonic frequencies, also called Fourier frequencies $\eta_1 = \frac{2\pi}{\gamma}, \quad \eta_2 = 2\eta_1, \ldots, \eta_j = j\eta_1 = \frac{2\pi j}{\gamma}$, where γ is the sample length.
The first Fourier frequency, η_1, is also called the fundamental with period $P_1 = \frac{2\pi}{\eta_1} = \gamma$. Clearly, a cycle of a longer period (i.e. smaller frequency) cannot be observed from a sample of length γ. Similarly, frequencies larger than π are not considered since they correspond to periods $2\pi/\eta$ shorter than 2, which is not observable in discrete time with $t = 1, 2, \ldots$. [30] Hence, the set of Fourier frequencies typically consists of $\eta_j = \frac{2\pi j}{\gamma}, j = 1, 2, \ldots, M = \left[\frac{\gamma-1}{2}\right]$. In the case of an odd or even sample size, respectively, we hence have

$$M = \begin{cases} \frac{\gamma-1}{2}, & \text{if } \gamma \text{ is odd} \\ \gamma/2 - 1, & \text{if } \gamma \text{ is even.} \end{cases}$$

The main result, that we obtain, says that any given time series can be decomposed into the sum of (weighted) trigonometric functions evaluated at $\eta_j h$, where h is the lag between the successive value of the time series, which we will be explained in the spectral study.

1.3.3 Technical Results on Summability

In this part, we will collect some technical results on the summability of real sequences. Let $\{c_j\}_{j\in\mathbb{N}}$ be a sequence of real numbers. Then the following holds:

- If $\{c_j\}$ is summable, then

$$\sum_{j=0}^{\infty} c_j < \infty \implies \lim_{j\to\infty} c_j = 0 \implies \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} c_j = 0,$$

- If $\{|c_j|\}$ is summable, then $\{c_j\}$ is said to be absolutely summable. If $\{c_j^2\}$ is summable, then $\{c_j\}$ is said to be square summable. Hence, absolute summability implies summability and square summability:

$$\sum_{j=0}^{\infty} |c_j| < \infty \Longrightarrow \begin{cases} \sum_{j=0}^{\infty} c_j < \infty, & (a) \\ \sum_{j=0}^{\infty} c_j^2 < \infty. & (b) \end{cases}$$

  The opposite does not hold: (a) shows that summability is not sufficient for absolute summability, and (b) shows a square summable sequence is not necessarily (absolutely) summable. The most explicit example is $c_j = \frac{(-1)^j}{j}$. For these and further results, see ([28], section 3.2).

1.3.4 Asymptotic Notation

Big O provides an upper bound on the rate of growth, however the little o means a loose upper bound. An infinite sequence $\{a_n,\ n \in \mathbb{N}^*\}$ is $O(1)$ if it is bounded, i.e.,

$$\exists\ c \in \mathbb{R}_+^*, \quad \text{such that } |a_n| \leq c,\ \forall n \geq 1.$$

And, the sequence $\{a_n,\ n \in \mathbb{N}^*\}$ is $o(1)$ if $a_n \to 0$ as $n \to \infty$. The concept can be generalized. Let $\{b_n,\ n \in \mathbb{N}^*\}$ be a positive infinite sequence. We say

$$a_n = O(b_n) \quad \text{if the sequence} \quad a_n/b_n \quad \text{is bounded for} \quad n \in \mathbb{N}^*.$$

$$\text{And we say} \quad a_n = o(b_n) \quad \text{if } a_n/b_n \to 0 \quad \text{as} \quad n \to \infty.$$

The simple lemma below gives an alternative expression.

Lemma 1. $a_n = O(b_n) \iff a_n = b_n O(1)$ *and* $a_n = o(b_n) \iff a_n = b_n o(1)$.

Properties: Below are some simple facts and rules of operation for O and o.

- (i) If $a_n = o(b_n) \Longrightarrow a_n = O(b_n)$.
- (ii) If $a_n = O(b_n)$ and $b_n = o(c_n) \Longrightarrow a_n = o(c_n)$.
- (iii) If $a_n = o(b_n)$ and $b_n = O(c_n) \Longrightarrow a_n = o(c_n)$.
- (iv) If $a_n = o(b_n)$ and $b_n = o(c_n) \Longrightarrow a_n = o(c_n)$.
- (v) If $a_n = o(b_n) \Longrightarrow \forall p > 0, |a_n|^p = o(b_n^p)$.

1.3.5 Estimation Methods

Parametric Method

Parametric methods are based on parametric models of a time series, such as autoregression (AR) models, moving average (MA) models, and autoregressive-moving average (ARMA) models. Therefore, parametric methods also are known as model-based methods. To estimate the SDF or DOM of a time series with parametric methods, you need to obtain the model parameters of the time series first.

Nonparametric Method

The nonparametric method does not require the population under study to meet particular assumptions or specific parameters to characterize the observations, as is the case with parametric methods. To illustrate, such as a case of probability density estimation using a Histogram, the spectral density estimation using a PE.

Semiparametric Methods

A semiparametric model is a statistical model that has parametric and nonparametric components. It is considered to be "smaller" than a completely nonparametric model because we are often interested only in the finite-dimensional component of the parameter. That is, the infinite-dimensional component (the non-parametric component) is regarded as a nuisance parameter. Sachs as a nuisance parameter is any parameter which is not of immediate interest but which must be accounted for in the analysis of those parameters which are of interest. Based to this, we define the semiparametric estimation.

Definition 5. *Semiparametric estimation methods are used to obtain estimators of the parameters of interest, typically the coefficients of an underlying regression function, without a complete parametric specification of the conditional distribution of the dependent variable given the explanatory variables.*

As the name suggests, semiparametric method are models that are part parametric and part nonparametric. An example is a partially linear regression model of the form

$$Y = bX + f(z) + \varepsilon,$$

where Y is a responce vector, X is a matrix which is known as explanatory variables and ε is i.i.d. The function f is an arbitrary function of z, it is the nonparametric part and b

is an unknown vector, it is the parametric part of the model. The theory of inference for such models can become very complex. Under appropriate conditions and if $\widehat{f}$ is chosen carefully, this will lead to good estimate of b.

1.3.6 Bias-Variance Tradeoff

Why the best estimator is selected via smallest a mean square error (MSE)?
The MSE is the trade-off between bias and variance and is the best approach to select the preferable estimator via the smallest MSE.
Let $\widehat{f(x)}$ be an estimate of a function $f(x)$. The squared error (or L2) loss function is

$$L(f) = \left(\widehat{f(x)} - f(x)\right)^2.$$

The average of this loss is called the Risk or MSE and is denoted by:

$$\text{MSE} = E\left(L(f)\right) = E\left(\left(\widehat{f(x)} - f(x)\right)^2\right).$$

The random variable in this equation is the function $\widehat{f}$ which implicitly depends on the observed data. We will use the terms risk or MSE, such as a simple calculation shows that

$$\text{MSE} := Bias^2 + variance,$$

where, $Bias(f(x)) = E\left(\widehat{f(x)} - f(x)\right)$ is the bias of $\widehat{f}$ and $var(\widehat{f})$ is the variance of $\widehat{f}$. The main challenge in smoothing is to determine how much smoothing to do. When the data are over smoothed, the bias term is large and the variance is small. When the data are under smoothed the opposite is true; see Figure 1.2. This is called the bias–variance tradeoff. Minimizing risk corresponds to balancing bias and variance. The bias increases and the variance decreases with the amount of smoothing. The optimal amount of smoothing, indicated by the vertical line, minimizes the Risk=MSE (graph 1.2).

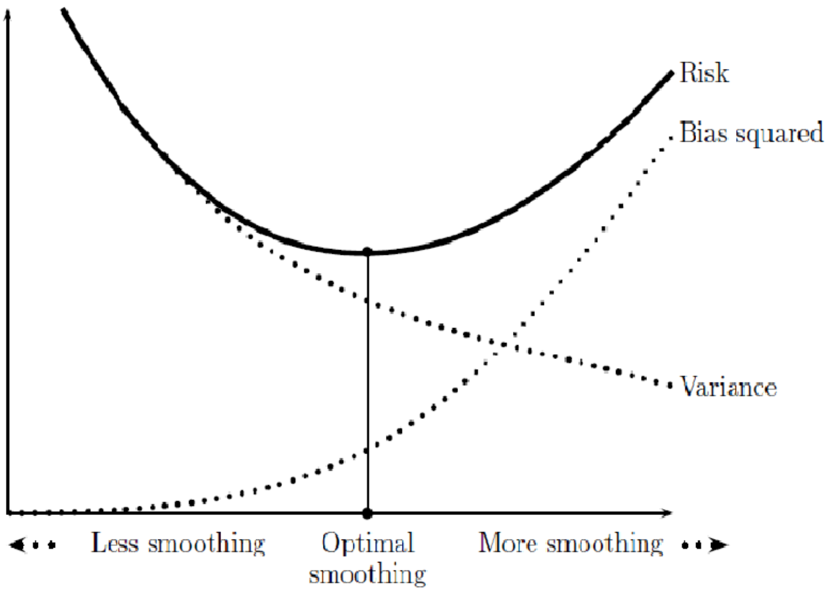

Figure 1.2: The bias - variance tradeoff.

Chapter 2

Long Memory Gaussian Stationary Process

2.1 Stationary Stochastic Process

2.1.1 Stochastic Process

Definition 6. *(Stochastic) The word stochastic originates from the Greek stochazesthai meaning "to aim at" or "to guess at". It is used in the sense of random in contrast to deterministic. While in a deterministic model the outcome is completely determined by the equations and the input (initial conditions), in a stochastic model no exact values are determined but probability distributions. In that sense, a stochastic model can be understood as a means to guess at something.*

The choice between a deterministic and a stochastic model is basically one of what information is to be included in the equations describing the system. On the one hand information can be limited simply by the lack of knowledge. On the other hand it might not be benefiting the modelling objective to include certain information.

Definition 7. *(Stochastic Process). A stochastic process is a family of random variables* $\{X_t,\ t \in T\}$ *defined on a probability space* $(\Omega, \mathcal{F}, P)$.

Definition 8. *(Realizations of Stochastic Process). The family* $\{x_t = X_t(\omega),\ t \in T,\ \omega \in \Omega\}$ *is known as the realizations of the process* $\{X_t,\ t \in T\}$.

Definition 9. *(Time Series) A time series is a set of observations (realization of a stochastic process)* x_t*, each one being recorded at a specified time* t*. The time series*

$\{x_t,\ t \in T_o\}$ is then a realization of the family of random variables $\{X_t,\ t \in T_o\}$, where T_o is a given set. These considerations suggest modelling the data as a realization (or part of a realization) of a stochastic process $\{X_t,\ t \in T\}$ where $T_o \subseteq T$.

• Time series is divided in two parts

1. A discrete-time series (the type to which this work is primarily devoted) is one in which the set T_o of times at which observations are made is a discrete set, as is the case for example when observations are made at fixed time intervals.

2. Continuous-time series are obtained when observations are recorded continuously over some time interval, e.g. when $T_o = [0, 1]$. We shall use the notation $x(t)$ rather than x_t, if we wish to indicate specifically that observations are recorded continuously.

• Objectives of the analysis and the study of time series can be divided into two main components:

1- Describe and understand the chain production mechanism, which includes metadata analysis and modelling, then

2- Forecasting future values and estimating maximum risks.

• General problem is to construct adequate mathematical models for the data (time series), i.e., the selection of a suitable mathematical model (or class of models) for the data. To allow for the possibly unpredictable nature of future observations it is natural to suppose that each observation x_t is a realized value of a certain random variable X_t.

2.1.2 Stationary Process

Stationary is an invariant property which means that statistical characteristics of the time series do not change over time. For example, the yearly rainfall may vary year by year, but the average rainfall in two equal length time intervals will be roughly the same as would the number of times the rainfall exceeds a certain threshold. But, over long periods of time this assumption may not be plausible, so it guides us to the non-stationary notion. Also, for example, the climate change that we are currently experiencing is causing changes in the overall weather patterns. However in many situations, including short time intervals, the assumption of stationary is quite a plausible. Indeed often the statistical analysis of a time series is done under the assumption that a time series is stationary.

There are two cases of stationary, the strict stationary and weak stationary. Whose the weak stationary concerns only the covariance of a process and the strict stationary which is a much stronger condition and supposes the distributions are invariant over time.

Definition 10. *(Strict Stationary) The time series $\{X_t\}$ is said to be strictly stationary if for any finite sequence of integers $t_1, ..., t_k$ and shift h the distribution of $(X_{t_1}, ..., X_{t_k})$ and $(X_{t_1+h}, ..., X_{t_k+h})$ are the same.*

The above assumption is often considered to be rather strong (and given a data it is very hard to check). Often it is possible to work under a weaker assumption called weak or second order stationary.

Definition 11. *(Weak Stationary) The process $\{X_t\}$ is said to be second order stationary or weak stationary if:*

- *The mean is constant for all t,i.e., $E(X_t) = m, \ \forall t \in \mathbb{R}$; and*
- *For any $(t, k) \in \mathbb{R}^2$, the covariance between X_t and X_{t+k} only depends on the lag difference k. In other words there exists a function*

$$\mathcal{A} : \mathbb{Z}^2 \longrightarrow \mathbb{R}$$
$$(t, h) \to \mathcal{A}(h, t) = cov(X_t, X_{t+h}),$$

such that for all t and k we have $\mathcal{A}(h) = cov(X_t, X_{t+h})$.

A second-order (weak) stationary process is a process whose statistics of order less than three, such as mean and covariance, do not change over time, however other statistics of order greater than two, such as k^{th} moments for $k > 2$, are free to change with the time. To show that strict stationarity (with $E(X_t^2) < \infty$) implies second order stationarity, suppose that $\{X_t\}$ is a strictly stationary process, then

$$\begin{aligned} cov(X_t, X_{t+k}) &= E(X_t X_{t+k}) - E(X_t)E(X_{t+k}) \\ &= \int xy \left(P_{X_t,X_{t+k}}(dx, dy) - P_{X_t}(dx)P_{X_{t+k}}(dy)\right) \\ &= \int xy \left(P_{X_0,X_k}(dx, dy) - P_{X_0}(dx)P_{X_k}(dy)\right) = cov(X_0, X_k), \end{aligned}$$

where $P_{X_t,X_{t+k}}$ is the joint distribution and $P_{X_t}, \ P_{X_{t+k}}$ are the marginal distribution of $X_t, \ X_{t+k}$ respectively. The above shows that $cov(X_t, X_{t+k})$ does not depend on t and $\{X_t\}$ is second order stationary.

Remark 1. *(i) If a process is strictly stationary and $E(X_t^2) < \infty$, then it is also second order stationary. But the converse is not necessarily true.*

$$\textit{strictly stationary} \xrightarrow{E(X_t^2)<\infty} \textit{second order stationary.}$$

(ii) If a process is strictly stationary but the second moment is not finite, $E(X_t^2) = \infty$, then it is not second order stationary (Cauchy's distribution).

$$\textit{strictly stationary} \overset{E(X_t^2)=\infty}{\nrightarrow} \textit{second order stationary.}$$

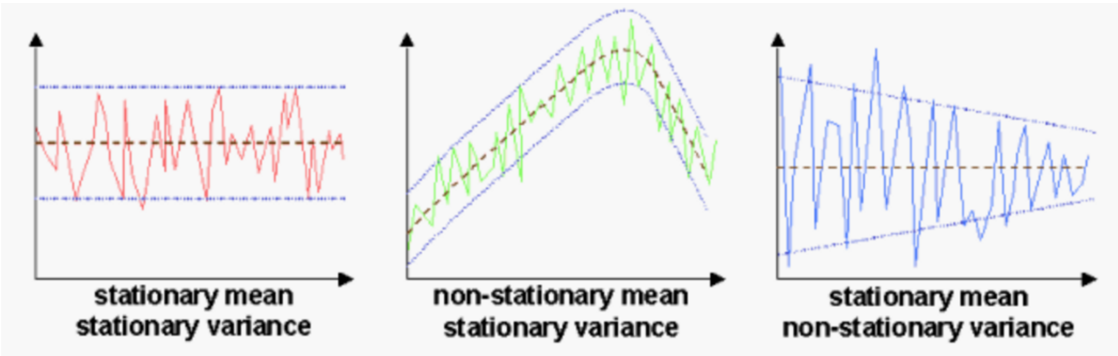

Figure 2.1: The stationarity in the mean and variance.

2.1.3 Autocovariance Function

Stochastic process $\{X_t\}_{t\in\mathbb{Z}}$ has the mean function $E(X_t)$, his ACF is given by

$$\mathcal{A}(t,s) := cov(X_t, X_s) = E(X_t X_s) - E(X_t)E(X_s), \quad \forall(t,s) \in \mathbb{Z}^2.$$

In addition, let $\{X_t\}_{t\in\mathbb{Z}}$ be a discrete (second-order) stationary process, so

$$\mathcal{A}(h) = cov(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)^2, \quad \forall(t,h) \in \mathbb{Z}^2.$$

• From the stationarity process, the ACF has the following basic properties, $\forall h \in \mathbb{Z}$:

$$(i)\ \mathcal{A}(h) = \mathcal{A}(-h); \quad (ii)\ |\mathcal{A}(h)| \leq \mathcal{A}(0) = var(X_t); \quad (iii)\ |\mathcal{A}(h)| < \infty.$$

• The empirical ACF (i.e., the estimator of ACF) of an observed dataset $\{X_1, \ldots, X_\gamma\}$ from a stationary process $\{X_t\}_{t\in\mathbb{Z}}$ is defined as follows

$$\widehat{\mathcal{A}}_\gamma(h) = \begin{cases} \dfrac{1}{\gamma} \displaystyle\sum_{j=1}^{\gamma-|h|} (X_j - \bar{X})(X_{j+|h|} - \bar{X}), & \text{if} \quad |h| \leq \gamma - 1; \\ 0, & \text{if} \quad |h| \geq \gamma. \end{cases} \tag{2.1}$$

where $\bar{X}$ is the sample mean, $\bar{X} = \gamma^{-1} \sum_{t=1}^{\gamma} X_t$.

2.1.4 Spectral Density Function

Spectral analysis breaks down a process into a periodic function that identifies cycles or periodic. Moreover, it is convenient to map the correlation of stochastic processes from the time domain to the so-called frequency domain. The following Theorem characterizes that the ACF can be written in the form of integral by bounded distribution function F with mass concentrated on $[-\pi, \pi]$.

Theorem 1. *Herglotz's Theorem*
A complex valued function $\mathcal{A}(.)$ defined on the integers is non-negative definite if and only if

$$\mathcal{A}(h) = \int_{-\pi}^{\pi} e^{ih\eta} dF(\eta), \tag{2.2}$$

where $F(.)$ is a bounded, increasing, continuous function over $[-\pi, \pi]$ and $F(-\pi) = 0$. F is called a spectral distribution function of $\mathcal{A}(.)$.

Herglotz's theorem shows that the ACF of the stationary process, $\mathcal{A}(.)$, are Fourier coefficients of a measure over the interval $[-\pi, \pi]$.

Remark 2. *If we put $h = 0$, in the previous reltion (2.2, we see that*

$$\mathcal{A}(0) = var(X_t) = \int_{-\pi}^{\pi} dF(\eta). \tag{2.3}$$

Therefore, $F()$ represents a distribution of the variance of X_t over the interval $[-\pi, \pi]$. We can also deduce, from (2.2) and (2.3), a new presentation of the random process X_t called spectral representation.

If there exists a function $\mathcal{S} : (-\pi, \pi] \to \mathbb{R}_+$ integrable such that for all $\eta \in (-\pi, \pi]$,

$$F(\eta) = \int_{-\pi}^{\eta} \mathcal{S}(\theta) d\theta, \quad -\pi < \eta \leq \pi,$$

then $\mathcal{S}$ is called spectrum or the SDF. In analogy to formula (2.2), we have: $\mathcal{A}(h) = \int_{-\pi}^{\pi} \mathcal{S}(\eta) e^{ih\eta} d\eta$. To evaluated the role of the SDF, $\mathcal{S}(\eta)$, and for (2.3), we get

$$var(X_\gamma) = \mathcal{A}(0) = \int_{-\pi}^{\pi} \mathcal{S}(\eta) d\eta.$$

Therefore, $\mathcal{S}(\eta)$ represents the variance value of X_t of the oscillation of the pulse η in the interval $[-\pi, \pi]$.

Theorem 2. *Wiener-Khinchin theorem*
The spectral density function of a stationary stochastic process is the Fourier transform of its autocorrelation function.

From the Wiener-Khinchin theorem, a necessary and sufficient condition that the function $\mathcal{A}(h)$ be an ACF for a process DSGP , $\{X_t\}_{t\in\mathbb{Z}}$, is that there exists a SDF defined on $(-\pi, \pi]$ as follows

$$\mathcal{S}(\eta) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \mathcal{A}(h)e^{-ih\eta}, \quad \eta \in (-\pi, \pi], \tag{2.4}$$

where η is the frequency. It is obvious that:

(1) SDF is the Fourier transformation of the ACF;

(2) SDF is a non-negative function: $\mathcal{S}(\eta) \geq 0$;

(3) SDF is an even function: $\mathcal{S}(\eta) = \mathcal{S}(-\eta)$;

(4) SDF is a periodic function of period $P = 2\pi$: $\mathcal{S}(\eta) = \mathcal{S}(\eta + 2\pi)$; and

(5) We can calculate the variance of the process $var(X_t) = \mathcal{A}(0) = 2\int_0^\pi S(\eta)\mathrm{d}\eta$.

$\mathcal{S}(\eta)$ measures how strongly the frequency η or period P contributes to the variance of the process. For the formula 1.3 , $\eta \to 0$ it holds that $P = 2\pi/\eta \to \infty$. This may be interpreted the following way: A cycle with infinite period is no longer cyclic (periodic); it is rather a trend. Hence, SDF at the origin, $\mathcal{S}(0)$, measures how strongly a trending behavior affects the variance of the process. Examples of spectra given below will support this intuition from Brockwell (1991)[10] (Corollary 4.3.1,Theorem 5.7.2, and Theorem 5.8.1).

2.2 Gaussian Stationary Process

2.2.1 Definition

A GP is a stochastic process that is in general a collection of random variables indexed by time or space. Its special property is that any finite collection of these variables follows a multivariate Gaussian distribution (or jointly Gaussian). Thus, the GP is a distribution over infinitely many variables and, therefore, a distribution over functions with a continuous domain.

Definition 12. *(Gaussian Process). A real-valued stochastic process* $\{X_t,\ t \in T\}$, *where* T *is an index set, is a GP if all the finite-dimensional distributions have a multivariate normal distribution. That is, for any choice of distinct values* $t_1, ..., t_k \in T$, *the random vector* $X_t = \{X_{t_1}, ..., X_{t_k}\}'$ *has a multivariate normal distribution with:*
• *mean vector* $m_t = E(X_t)$, *and*
• *covariance matrix* $\Sigma(t, s) = cov(X_t, X_s)$,
which will be denoted by $X_t \sim \mathcal{N}(m_t, \Sigma(t, s))$.

2.2.2 Properties

i) For Gaussian Process

The popularity of GP stems primarily from two essential properties:
• First, a GP is completely determined by its mean and covariance functions. This property facilitates model fitting as only the first and second order moments of the process require specification.
• Second, solving the prediction problem is relatively straightforward. The best predictor of a GP at an unobserved location is a linear function of the observed values and, in many cases, these functions can be computed rather quickly using recursive formulas.
• Three, equivalence between independence and uncorrelated

$$X \text{ and } Y \text{ are independent} \overset{\text{GP}}{\Longleftrightarrow} \text{uncorrelated} \overset{\text{GP}}{\Longleftrightarrow} cov(X, Y) = 0.$$

ii) For Stationary Gaussian Process

More of the above properties, for any stationary Gaussian process (SGP) $X = \{X_t\}_{t \in T}$, we have

(1) $E(X_t) = m_t = m$ is independent of t;

(2) $cov(X_{t+h}, X_t) = \mathcal{A}(t + h, t) = \mathcal{A}(h)$ is independent of t for all h. It is conventional to express the ACF $\mathcal{A}$ as a function on T instead of on $T \times T$;

(3) $X \sim \mathcal{N}(m, \mathcal{A}(0))$ for all t;

(4) $(X_{t+h}, X_t)'$ has a bivariate normal distribution with covariance matrix

$$\Sigma(t, t + h) = \Sigma(h) = \begin{pmatrix} \mathcal{A}(0) & \mathcal{A}(h) \\ \mathcal{A}(h) & \mathcal{A}(0) \end{pmatrix};$$

(5) It should be noted that the first and second order moments (covariances) of a GP determine its full probability structure. This implies that the probability structure of a weakly stationary GP is invariant (does not change) with time shifts. A process whose probability structure is invariant under time shifts is called a strictly stationary process. Thus, for GP, a weakly stationary is also strictly stationary too, this is the only case where weakly stationary implies strictly stationary;

$$\textit{weakly stationary} \overset{\text{GP}}{\Longleftrightarrow} \textit{strictly stationary}$$

(6) We also remark that one can define some non Gaussian stationary processes in terms of Gaussian stationary processes by taking a non-linear function of a Gaussian stationary process and its shifts. If $X = (X_n)_{n\in\mathbb{Z}}$ is a stationary Gaussian process, a simple example of such a non Gaussian stationary process, $Y_{(n,k)}$, is given by

$$Y_{(n,k)} = g(X_n, X_{n-1}, ..., X_{n-k}),$$

where g is a polynomial $g(U_0, ..., U_k)$ in the variables $U_n, ..., U_k$.

2.3 Memory of Process

2.3.1 Definitions

For some statistical analysis, such as deriving an expression for the variance of an estimator, the covariance is often sufficient as a measure, because it is considered a measure of a linear dependence.

- What is the memory of the process?

Definition 13. *The concept of memory refers to how strongly the past can influence the future in a given process. Or, the degree of dependence between the values taken during a given period.*

- Is there a reason why long memory is such a widely observed phenomenon in a variety of empirical applications?

For long memory process, analysing the past would be really useful because it can provide information about what is going to happen in the future. long memory processes have higher dependent than short memory processes. A series of different generating mechanisms has been suggested to explain the feature of long memory theoretically, for the special case of realized volatility process (see, [28]). The major justification is the ubiquity of long memory in empirical time series as a stylized fact from many different fields of science and practice.

2.3.2 Types of the Memory

To distinguish the types of the memory of a process, one studies it in two domains, temporal domain and spectral domain.

a) In Temporal Domain

Next, we want to address a stronger result than convergence of the sample mean $\bar{X}_\gamma$. It is on the variance of the sample mean, $var(\bar{X}_\gamma)$, under the more restrictive assumptions, that the ACFs are absolutely summable. We label such processes as short memory. In particular, the DSGPs can be divided into three types based on the behavior of the ACF as follows (cf. [40, 25])

Definition 14. *(Negative, Short and long memory) A stationary process $\{X_t\}$ is said to display long memory if the sequence of autocovariances $\mathcal{A}(.)$ dies out so slowly that it is not absolutely summable:*

- *Short-memory DSGP (SMDSGP), if the ACF is absolutely summable with positive sum, i.e.,*

$$\sum_{h\in\mathbb{Z}} |\mathcal{A}(h)| < \infty \text{ and } \sum_{h\in\mathbb{Z}} \mathcal{A}(h) > 0. \tag{2.5}$$

- *Negative-memory DSGP (NMDSGP), if the ACF is absolutely summable with zero sum, i.e.,*

$$\sum_{h\in\mathbb{Z}} |\mathcal{A}(h)| < \infty \text{ and } \sum_{h\in\mathbb{Z}} \mathcal{A}(h) = 0. \tag{2.6}$$

- *Long-memory DSGP (LMDSGP), if the ACF is not absolutely summable, i.e.,*

$$\sum_{h\in\mathbb{Z}} |\mathcal{A}(h)| = \infty. \tag{2.7}$$

This definition of long memory coincides with the Definition 3.1.2 in Giraitis et al. [23]. While long memory is defined in terms of absolute ACFs, it will turn out in Proposition 2.3 ([23]), that it makes sense to consider the summation over $\{\mathcal{A}(h)\}$.

b) In Spectral Domain

Moreover, the memory behaviors of a process $\{X_t\}_{t\in\mathbb{Z}}$ equivalently to (2.5), (2.6) and (4.4) can be investigated via the SDF as follows

- SMDSGP, if the SDF is bounded and does not vanish at the zero, i.e.,

$$0 < \mathcal{S}(0) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \mathcal{A}(h) < \infty. \tag{2.8}$$

- NMDSGP, if the SDF vanishes at zero, i.e.,

$$\mathcal{S}(0) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \mathcal{A}(h) = 0. \tag{2.9}$$

- LMDSGP, if the SDF is generally unbounded at the zero, i.e.,

$$\mathcal{S}(0) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \mathcal{A}(h) = \infty. \tag{2.10}$$

c) Relationship between ACF and SDF

A DSGP $\{X_t\}_{t\in\mathbb{Z}}$ can be characterized by its degree of memory (DOM, d). The DOM controls the shape of the SDF near to the zero frequency and the decay rate of its ACF. The SDF of a DSGP $\{X_t\}_{t\in\mathbb{Z}}$ can be approximated in the neighborhood of the zero frequency in terms of the DOM as follows (cf. [63])

$$\mathcal{S}(\eta)\simeq\alpha\,|\eta|^{-2d}, \quad \text{as} \quad \eta \to 0,\ |d| < \frac{1}{2},\ \ 0 < \alpha < \infty. \tag{2.11}$$

The approximate behavior of the SDF of a DSGP $\{X_t\}_{t\in\mathbb{Z}}$ is equivalent to the following asymptotic behavior of the ACF in terms of the DOM

$$\mathcal{A}(h)\simeq\beta|h|^{2d-1}, \quad \text{as} \ \ h \to \infty,\ |d| < \frac{1}{2},\ \ 0 < \beta < \infty. \tag{2.12}$$

Further, $\mathcal{S}(\eta)$ is continuous on $[0,\pi]$ if $\{X_t\}$ has short memory

The memory behavior of a DSGP $\{X_t\}_{t\in\mathbb{Z}}$ can be investigated via the DOM. For any DSGP $\{X_t\}_{t\in\mathbb{Z}}$, the SDF in (2.11) can be generalized by including a non-negative function slowly varying at infinity $\mathcal{V}_{\mathcal{S}}(.)$ as follows

$$\mathcal{S}(\eta) = |\eta|^{-2d}\mathcal{V}_{\mathcal{S}}\left(\frac{1}{|\eta|}\right),\ -\pi < \eta \leq \pi. \tag{2.13}$$

From [7], the ACF in (2.12) can be generalized by the same technique of the corresponding SDF as follows

$$\mathcal{A}(h) = |h|^{2d-1}\mathcal{V}_{\mathcal{A}}\left(\frac{1}{|h|}\right), \quad \mathcal{V}_{\mathcal{A}}\left(\frac{1}{|h|}\right) = 2\Gamma(1-2d)\sin(\pi d)\mathcal{V}_{\mathcal{S}}\left(|h|\right),\ h \in \mathbb{Z}, \tag{2.14}$$

where Γ denotes the Gamma function and $\mathcal{V}_{\mathcal{A}}(.)$ is a slowly varying function at the origin. A DSGP $\{X_t\}_{t\in\mathbb{Z}}$ is said to be NMDSGP, SMDSGP, or LMDSGP depending on whether $-\frac{1}{2} < d < 0$, $d = 0$, or $0 < d < \frac{1}{2}$, respectively.

Autoregressive moving average (ARMA) processes (cf. [9, 10]) are a type of SMDSGPs, while the autoregressive fractional integral moving average (ARFIMA) processes are LMDSGPs (cf. [6]).

2.3.3 Examples of Long Memory Models

In this subsection, we exhibit the examples of long memory process according to its properties. We shown successively a continuous non-stationary process, fractional Brownian motion (FBM); a discrete stationary process, fractional Gaussian noise (FGN); and a linear process, ARFIMA process which is the simplest long memory models.

a) Fractional Brownian Movement (FBM)

The Fractional Brownian Movement (FBM) denoted by $B_H(t)$ and defined over $\mathbb{R}$ is a non stationary Gaussian process with continuous paths centred and with an ACF $\mathcal{A}(.)$

$$\mathcal{A}(t-s) = E\left[B_H(t)B_H(s)\right] = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \quad \forall t, s \in \mathbb{R},$$

where $H \in (0, 1)$, is the Hurst parameter and σ^2 the innovation variance. The value of H determines the nature of the FBM. Thereby, if

- $H = 1/2$, the process is standard Brownian motion;
- $H > 1/2$, the increments of the process are positively correlated; and
- $H < 1/2$, the increments of the process are negatively correlated.

This process was introduced by Mandelbrot and Van Ness [1968] to explain the Hurst phenomenon (persistence of periods of high and low flows of the Nile river) observed by Hurst [1951] on Nile data.

b) Fractional Gaussian Noise (FGN)

A process $X = \{X_k, k \in \mathbb{Z}\}$ is an FGN if it is defined as the increments of an FBM at integer times, namely

$$X_k = B_H(k+1) - B_H(k), \quad k \in \mathbb{Z},$$

where $(B_H(t))$ is a FBM with Hurst parameter $H \in [0, 1)$. It is a stationary central Gaussian process whose ACF $\mathcal{A}(h) = E(X_k X_{k+h})$ is given by

$$\mathcal{A}(h) = \frac{\sigma^2}{2}\left(|h+1|^{2H} + |h-1|^{2H} - 2|h|^{2H}\right), \quad \forall h \in \mathbb{Z},$$

where H and σ^2 are defined as above. The asymptotic behavior of the auto-covariance function $\mathcal{A}(.)$ of an FGN is given by the relation

$$\mathcal{A}(h) \sim \sigma^2 H(2H-1)h^{2H-2}, \quad as\ h \to \infty,$$

when $H \neq 1/2$. If we denote by d the memory parameter, then, $d = H - 1/2$; hence in the case of the FGN, $d \in (-1/2, 1/2)$ and in an equivalent way, the last formula is rewritten

$$\mathcal{A}(h) \sim \sigma^2 d(2d+1)|h|^{2d-1}\ as\ \ h \to \infty.$$

c) Autoregressive Fractionally Integrated Moving Average (ARFIMA)

Another example of a very widespread long memory process is the ARFIMA(p, d, q). Where, "p" represents the number of coefficients of the part AR, "d" the memory parameter and "q" the number of coefficients of the part MA.

ARMA models integrated of order d are a standard tool for time series analysis, where typically $d \in \{0, 1, 2\}$. The integrated ARMA (ARIMA) model of order d means that a time series has to be differenced d times in order to obtain a stationary and invertible ARMA representation. The papers by Granger and Joyeux (1980)[22] and Hosking (1981) [29] extended the ARIMA model with integer d to the so-called fractionally integrated model, where d takes on non integer values, often restricted to $|d| < 1/2$. In particular, the case of $0 < d < 1/2$ corresponds to a stationary model with long memory, where the latter case means that the autocovariances die out slowly, that they are not absolutely summable. This process has become very popular since its introduction at the beginning of the 1980s, which remain the predominant model for the analysis of linear chronological series.

If $d > 1/2$, the process is non-stationary and the process is not invertible if $d < -1/2$. For a broader explanation without going into details, we have $1/2 \leq d < 1$, the fractionally integrated model bridges the gap from stationarity to the so-called unit root behavior ($d = 1$), where past shocks have a permanent effect on the present and values of $d > 1$ allow for even more extreme persistence [28].

An ARFIMA process (p, d, q), Y_t, can be defined by the (fractional) difference equation

$$\phi(B)Y_t = \theta(B)(I-B)^{-d}\varepsilon_t, \tag{2.15}$$

with $\phi(B) = 1 + \phi_1 B + \ + \phi_p B^p \quad and \ \phi(B) = 1 + \theta_1 B + \ + \theta_q B^q$ are the autoregressive and moving average operators, respectively; $\phi(B)$ and $\theta(B)$ have no common zeros and all have their roots outside the unit circle. $(I-B)^{-d}$ is a fractional operator of differentiation or of integration with $d \in [0, 1/2)$, it defined by the binomial expansion

$$(I-B)^{-d} = \sum_{j=0}^{\infty} \frac{(j+d)}{\Gamma(j+1)\Gamma(d)},$$

for $d < 1/2, B$ is the backward shift operator, $BX_t = X_{t-1}$ and (ε_t) is white noise of finite variance. The term white is used because white light is thought of as composed equally of light from the whole visible spectral range and all frequencies are equally represented, and the "noise" is because there's no pattern, just a random variation (cf, [51]).

Note that the ARMA or ARIMA processes can be considered as special cases of the ARFIMA processes with $d = 0$ and $d = 1, 2, ...$ respectively. Palma [2007] establishes the existence and uniqueness of a stationary solution of the ARFIMA process defined by (2.15), as well as its causality and its reversibility. For definitions of causal and invertible process see ([10]).

Chapter 3

Statistical Inference for Long Memory Gaussian Stationary Process

3.1 Main Contribution of the Work

For SMDSGPs, the theoretical justifications of the behavior of the LWE-SDF have been studied by many researchers. For instance, Priestly[42] presented two sufficient conditions (cf. SC1 and SC2 in Section 3) under which an asymptotically unbiased consistent LWE-SDF is given and Rashid et al.[44] investigated the best lag window for the LWE-SDF of a law order moving average process. For LMDSGP, the advantages of the LWE-DOM are investigated via computational studies (simulations) only without realizing the theoretical justifications of the findings (cf. [11, 41, 44]). For instance, Chen et al.[11] developed a regression type of the LWE-DOM and Hunt et al.[31] derived an approximate bias of the LWE-DOM in fractionally integrated time series models. It is a hard problem to get some theoretical justifications for the advantages of the LWEs, especially for the LMDSGPs.

This paper provides some theoretical justifications for the powerful of the LWEs, which give an in-depth look at the behavior of the LWE-SDF and LWE-DOM. This closer look at the behavior of the LWE-SDF and LWE-DOM gives the sufficient conditions under which the LWE-SDF and LWE-DOM can be improved. Based on theoretical and computational justifications, the precision; the convergence rate of the bias and variance; and the asymptotic distributions of the improved LWE-SDF and LWE-DOM under the new sufficient conditions are investigated. A comparison study (in the following section) among the behaviors of the LWEs under the new conditions, the LWEs without the new conditions and the PEs is given to investigate the significance of the new conditions.

The theoretical and computational justifications show the following main results:

- (i) the LWEs under the new conditions are better than the LWEs without the new conditions;
- (ii) the LWEs under the new conditions are better than the PEs;
- (iii) the LWEs under the new conditions are asymptotically unbiased and consistent;
- (iv) the asymptotic distributions of the LWE-DSF and LWE-DOM under the new conditions are chi-square and normal, respectively;
- (v) the LWE-DOM under the new conditions has a fast vanishing variance under the regression estimation method (REM); and
- (vi) the LWEs under the new conditions improve the finite sample properties for the REM and the local Whittle estimation method (WEM).

3.1.1 Disadvantage of the PE and the reliability of the LWE

The PE of the SDF (PE-SDF) is the widely used classical estimator of the SDF, which is given based on the empirical ACF in (2.1) as follows

$$\widehat{\mathcal{S}}_p(\eta) := \frac{1}{2\pi} \sum_{h=-(\gamma-1)}^{\gamma-1} \widehat{\mathcal{A}}_\gamma(h) e^{-ih\eta} = \frac{1}{2\pi\gamma} \Big| \sum_{t=1}^{\gamma} X_t e^{-i\eta t} \Big|^2, \quad \eta \in (-\pi, \pi]. \tag{3.1}$$

An analysis of the difficulty indicates the following:

• In addition, if $\widehat{\mathcal{A}(k)}$ converges slowly to zero, then the PE will be biased. When the argument $|h|$ is large, i.e., near $\gamma - 1$ ($|h| \simeq \gamma$), the random variables $\widehat{\mathcal{A}(h)}$ are averages of a relatively small number of the products $X_{t+|h|}X_t$, see the formula 2.1 . $\widehat{\mathcal{A}(h)}$ will be a poor estimate of $\mathcal{A}(h)$ since $\widehat{\mathcal{A}(h)}$ is the sum of only a few lag products that are divided by γ (see 2.1). Thus, $\widehat{\mathcal{A}(h)}$ will be much closer to zero than $\mathcal{A}(h)$:

$$E(\widehat{\mathcal{A}(h)}) = \frac{(\gamma - |h|)}{\gamma} \mathcal{A}(h),$$

and the bias is significant for $|h| \simeq \gamma$ if $\mathcal{A}(h)$ is not close to zero in this region.

• If $\widehat{\mathcal{A}(h)}$ decays rapidly to zero, i.e., h be small. The bias, $E(\widehat{\mathcal{A}(h)}) - \mathcal{A}(h) = \frac{|h|}{\gamma}\mathcal{A}(h)$,

will be small and will not contribute significantly to the total error in PE.
This accounts for the fact that the PE has variance which never approaches zero with increasing γ. This analysis indicates, the PE-SDF is not satisfactory for the following two main reasons:

- (i) it is an inconsistent estimator which does not converge to the SDF $\mathcal{S}(\eta)$, i.e., $var(\widehat{\mathcal{S}}_p(\eta)) \nrightarrow 0$ (cf. [42]); and
- (ii) it has an erratic form, however the SDF $\mathcal{S}(\eta)$ is a smooth continuous function and thus an estimator that shares the same property is expected to be a good estimator.

Moreover, the asymptotic behavior of the PE-SDF depends on the process and the type of its frequency, fixed frequencies or Fourier frequencies. For example:

- (i) for any DSGP with a fixed frequency, the PE-SDF is an asymptotically unbiased estimator (cf. [62]); and
- (ii) for any LMDSGP with Fourier frequency, the PE-SDF has an asymptotic relative bias (cf. [33]) if the number of the frequencies is held fixed.

The above characteristics of the PE-SDF give us an indication that there is a significant need to develop a more appropriate estimator. To overcome these defects of the PE-SDF, omitting some ACF's terms and introducing a smooth version in the sum (3.1), the formula is written as a weighted sum are the expected form of a good estimator. Therefore, the LWE-SDF is highly recommended for such situations. However,

(1) we know that the ACF $\mathcal{A}(h) \to 0$ as $h \to \infty$, and hence if we omit only those terms with correspondent to the "tail" of the estimator of the ACF. These ideas suggest that we might consider as an estimate of $\mathcal{S}(\eta)$ an expression of the form

$$\widehat{\mathcal{S}}_{w'}(\eta) := \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \widehat{\mathcal{A}}_\gamma(h) e^{-i\eta h}, \ \eta \in (-\pi, \pi]. \tag{3.2}$$

(2) Due the formula 3.2, $var(\widehat{\mathcal{S}}_{w'}) \to 0$, but hopefully the bias will be affected too seriously. A possible approach for eliminating the inconsistency of the PE-SDF is to introduce a weight function $\mathcal{W}(.)$, called lag window function (LWF). Grenander and Rosenblatt[24] were the first to use the notion of periodogram smoothing to improve the estimator of the SDF. The term window was introduced by Blackman

and Tucky[8] to give a view of the PE-SDF through a narrow window. This weighting windows have an essential place in the spectral analysis methods. Their main role is to better control the influence of the side lobes of these spectral estimators. This weight function has the same properties as the ACF, i.e., it must be defined and continuous over $[-1, 1]$, such that

C1: $\mathcal{W}(x) = \mathcal{W}(-x) \leq \mathcal{W}(0) = 1$;

C2: $\forall x \in (0, 1)$, $\mathcal{W}(x)$ is not an increasing function. It decreases to zero at an appropriate rate according to the choice of a weight function with $\mathcal{W}(1) = 0$;

C3: $\forall\ x_1, x_2 \in [-1, 1]$, $\exists k > 0$, such that $|\mathcal{W}(x_1) - \mathcal{W}(x_2)| \leq k\,|x_1 - x_2|$.

Various Lag Window Function

Several types of lag window are available in the statistical literature, proposed by famous statisticians (Bartlett, Parzen, Blackman-Tukey, Daniell, ...). The following table 3.1 contain the most important Lag window.

Windows	The form of $\mathcal{W}(\mathbf{x})$	
Rectangular	1	
Triangular	$1 - \lvert x\rvert$	$\lvert x\rvert \leq 1$
Blackman	$0.42 - 0.5\cos(\pi x) + 0.08\cos(2\pi x)$	$\lvert x\rvert \leq 1$
Hanning	$0.5 - 0.5\cos(\pi x)$	$\lvert x\rvert \leq 1$
Hamming	$0.54 - 0.46\cos(\pi x)$	$\lvert x\rvert \leq 1$
Bartlett	$2x$	$x \leq \frac{1}{2}$
	$2(1 - x)$	$\frac{1}{2} \leq x \leq 1$
J. Parzen	$1 - 6x^2 + 6\lvert x\rvert^3$	$\lvert x\rvert \leq \frac{1}{2}$
	$2(1 - \lvert x\rvert)^3$	$\frac{1}{2} \leq \lvert x\rvert \leq 1$

Table 3.1: Some Common Window Functions

The Hamming, Hanning, Bartlet, Blachman by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.2. For more detail of the LWFs and their characteristics can be found in [42]. Based on the above conditions of the LWF, a truncation point t_γ and the formula (3.2), the LWE-SDF is given as follows:

$$\widehat{\mathcal{S}}_w(\eta) := \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \widehat{\mathcal{A}}_\gamma(h) e^{-i\eta h} \mathcal{W}\left(\frac{h}{t_\gamma}\right), \ \eta \in (-\pi, \pi]. \tag{3.3}$$

From (3.1) and the Fourier transformation, the empirical ACF of an observed dataset in

(2.1) can be rewritten as follows:

$$\widehat{\mathcal{A}}_\gamma(h) = \int_{-\pi}^{\pi} \widehat{\mathcal{S}}_p(\theta) e^{ih\theta} d\theta, \qquad |h| \leq \gamma - 1. \tag{3.4}$$

From (3.4), the LWE-SDF in (3.3) can be rewritten as follows:

$$\widehat{\mathcal{S}}_w(\eta) = \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \left[\int_{-\pi}^{\pi} \widehat{\mathcal{S}}_p(\theta) e^{ih\theta} d\theta \right] e^{-i\eta h} \mathcal{W}\left(\frac{h}{t_\gamma}\right) = \int_{-\pi}^{\pi} \widehat{\mathcal{S}}_p(\theta) \mathcal{U}_{t_\gamma}(\eta - \theta) d\theta, \tag{3.5}$$

where

$$\mathcal{U}_{t_\gamma}(\eta) = \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \mathcal{W}\left(\frac{h}{t_\gamma}\right) e^{-ih\eta} \tag{3.6}$$

is the Fourier transform of the LWF, called a spectral window function (SWF). Based on the conditions (**C1-C3**) of the LWF, the SWF has the following properties

P1: $\mathcal{U}_{t_\gamma}(\eta)$ is a real valued even function, thus $\int_{-\pi}^{\pi} \eta \, \mathcal{U}_{t_\gamma}(\eta) d\eta = 0$. Therefore, for (3.6),

$$\mathcal{U}_{t_\gamma}(-\eta) = \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \mathcal{W}\left(\frac{h}{t_\gamma}\right) e^{ih\eta} = \mathcal{U}_{t_\gamma}(\eta)$$

$$\begin{aligned}
\int_{-\pi}^{\pi} \eta \mathcal{U}_{t_\gamma}(\eta) d\eta \quad & \overset{(3.6)}{=} \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \mathcal{W}\left(\frac{h}{t_\gamma}\right) \int_{-\pi}^{\pi} \eta e^{-ih\eta} d\eta \\
& = \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \mathcal{W}\left(\frac{h}{t_\gamma}\right) \left(\int_{-\pi}^{0} \eta e^{-ih\eta} d\eta + \int_{0}^{\pi} \eta e^{-ih\eta} d\eta \right) \\
& = \frac{1}{2\pi} \sum_{h=-t_\gamma}^{t_\gamma} \mathcal{W}\left(\frac{h}{t_\gamma}\right) \left(-\int_{0}^{\pi} \eta e^{ih\eta} d\eta + \int_{0}^{\pi} \eta e^{-ih\eta} d\eta \right) = 0.
\end{aligned}$$

P2: $\mathcal{U}_{t_\gamma}(\eta)$ becomes more concentrated around zero;

P3: $\mathcal{U}_{t_\gamma}(\eta)$ is a periodic function with period 2π, thus $\int_{-\pi}^{\pi} \mathcal{U}_{t_\gamma}(\eta) d\eta \overset{(3.6)}{=} 1$.

Therefore,

$$\begin{aligned}\int_{-\pi}^{\pi}\mathcal{U}_{t_\gamma}(\zeta)d\zeta &\overset{(3.6)}{=} \frac{1}{2\pi}\sum_{h=-t_\gamma}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\int_{-\pi}^{\pi}e^{-ih\zeta}d\zeta\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}d\zeta+\frac{1}{2\pi}\sum_{h=1}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\int_{-\pi}^{\pi}(e^{-ih\zeta}+e^{ih\zeta})d\zeta\\ &= 1+\frac{1}{2\pi}\sum_{h=1}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\int_{-\pi}^{\pi}(2\cos(h\zeta))d\zeta\\ &= 1+\frac{1}{\pi}\sum_{h=1}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\frac{1}{h}(\sin(h\pi)-\sin(-h\pi))\\ &= 1+\frac{1}{\pi}\sum_{h=1}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\frac{2}{h}\underbrace{\sin(h\pi)}_{=0,\ \forall h\in\mathbb{N}}=1.\end{aligned}$$

3.1.2 New Sufficient Conditions for Improving the LWEs

From the above discussions, the accuracy of the LWE-SDF depends on the following three main parameters: the truncation point t_γ (called the lag number); the number of Fourier frequencies n_γ; and the LWF $\mathcal{W}(x)$. To improve the accuracy of the LWE-SDF in (3.3) or (3.5) at Fourier frequencies, the following four sufficient conditions are considered, concerning the smooth behavior of the SDF in the neighborhood of the singularity located at zero.

SC1: The lag number t_γ is a sequence of integers satisfy $t_\gamma = c_1\gamma^a$, where c_1 is a constant and $0 < a < 1$. That is, $\frac{1}{t_\gamma}+\frac{t_\gamma}{\gamma}\to 0$, when $\gamma\to\infty$;

SC2: The number of frequencies n_γ is a sequence of integers called a spectral bandwidth parameter corresponding to the number of the Fourier frequencies $\eta_j = \frac{2\pi j}{\gamma}$, $j = 1,\ldots,n_\gamma$, which satisfy $n_\gamma = c_2\gamma^b$, where c_2 is a constant and $0 < b < 1$;

SC3: The sequences t_γ and n_γ satisfy $\frac{\gamma}{t_\gamma n_\gamma}\to 0$ as $\gamma\to\infty$;

SC4: For $\eta\neq 0$, the SDF $\mathcal{S}(\eta)$ is twice differentiable with a bounded second derivative. Precisely, in a neighborhood of the origin, $\mathcal{S}(\eta)$ is differentiable, i.e., $\forall\eta\in(\epsilon,\pi)$ with $\epsilon > 0$, $\frac{d^2}{d\eta^2}\mathcal{S}(\eta) = O(\eta^{-2d-2})$.

Remark 3. *It is worth mentioning that, the existing smoothed periodogram estimation techniques, such as the averaged periodogram estimator (called also the Welch's estimator) (cf. Otis [39]), the Bartlett and Parzen estimator (BPE) (cf. Bartlett [5]) and the* *integrated periodogram estimator (IPE) (cf. Lang and Azais [36] and Rosenblatt [51]), have the same principles as the LWE with the obvious changes in the SWF or/and the LWF. Therefore, the discussions in this paper for improving the LWE can be simply extended to improve the existing smoothed periodogram estimation techniques.*

3.1.3 Working Mechanism of the New Sufficient Conditions

The above mentioned conditions **SC1-SC4** are technical conditions concerning the smooth behavior of the SDF in the neighborhood of the singularity located at zero to get asymptotically unbiased consistent LWEs of the SDF and DOM. Each condition has the following reason to be given:

(1) The first condition **SC1**, which is the main idea of the LWT, is given to minimize the variance of the LWE-SDF and the variance of the LWE-DOM and make them tends to zero as follows:

For any DSGP with zero mean and finite variance, the first two moments of the empirical ACF in (2.1) are given by simple calculations as follows

$$E\left(\widehat{\mathcal{A}}_\gamma(h)\right) = \frac{1}{\gamma}\sum_{j=1}^{\gamma-|h|} E\left(X_j X_{j+|h|}\right) = \frac{1}{\gamma}(\gamma-|h|)\mathcal{A}(h) = \left(1-\frac{|h|}{\gamma}\right)\mathcal{A}(h). \quad (3.7)$$

$$E\left(\widehat{\mathcal{A}}_\gamma(h)\right)^2 = \frac{1}{\gamma^2}\sum_{i,j=1}^{\gamma-|h|} E\left(X_j X_{j+|h|} X_i X_{i+|h|}\right). \quad (3.8)$$

From (3.7) and (3.8), the variance of the empirical ACF can be given by the same technique of [42] as follows

$$var\left(\widehat{\mathcal{A}}_\gamma(h)\right) \simeq \frac{1}{\gamma}\sum_{m\in\mathbb{Z}}\left(\mathcal{A}^2(m)+\mathcal{A}(m+h)\mathcal{A}(m-h)\right) = O\left(1/\gamma\right). \quad (3.9)$$

Since $\forall\ |h| \leq \gamma-1,\ \widehat{\mathcal{A}}_\gamma(h)$ are asymptotically uncorrelated, the variance of the PE-SDF in (3.1) and the variance of the LWE-SDF in (3.3) can be given from (3.9) as follows

$$var\left(\widehat{\mathcal{S}_p}(\eta)\right) = \frac{1}{(2\pi)^2}\sum_{h=-(\gamma-1)}^{\gamma-1} e^{i\eta h} var(\widehat{\mathcal{A}}_\gamma(h)) \simeq \gamma\ var\left(\widehat{\mathcal{A}}_\gamma(h)\right) = O(1) \neq 0. \quad (3.10)$$

$$var\left(\widehat{\mathcal{S}_w}(\eta)\right) \overset{(3.2)}{=} \frac{1}{(2\pi)^2}\sum_{h=-(t_\gamma)}^{t_\gamma} e^{i\eta h} var\left(\widehat{\mathcal{A}}_\gamma(h)\right) \simeq t_\gamma\ var\left(\widehat{\mathcal{A}}_\gamma(h)\right) = O\left(\frac{t_\gamma}{\gamma}\right). \quad (3.11)$$

Under the first condition **SC1**, it is obvious from (3.11) that $var\left(\widehat{\mathcal{S}_w}(\eta)\right) \xrightarrow[\gamma\to\infty]{SC1} 0$. Therefore, under **SC1** the variance of the LWE-SDF tends to zero unlike the variance of the PE-SDF and the rate of decrease depends on the number of lags t_γ. Thus, the application of the window is to smooth the erratic and wild fluctuating form of the PE (cf. [42]). This behavior of the LWE-SDF under **SC1** helps to minimize the variance of the LWE-DOM and make it tends to zero. For more details, see the discussions in Sections 4.

(2) The second condition **SC2** is designed to facilitate the calculations in the spectral domain by using the Fourier frequencies, where the Fourier frequencies have the following useful orthogonality property

$$\begin{cases} \sum_{t=1}^{\gamma} e^{it(\eta_j \pm \eta_k)} = 0, & \text{for } 1 \leq\ j \neq k \leq n_\gamma; \\ \sum_{t=1}^{\gamma} e^{it(\eta_j - \eta_k)} = \gamma, & \text{for } 1 \leq\ j = k \leq n_\gamma. \end{cases}$$

Moreover, under **SC2** we get $\eta_j \xrightarrow[\gamma\to\infty]{} 0$. This behavior is necessary for the following two main reasons:

(2,1) Since the SDF has a pole at zero frequency, the necessary asymptotic behavior of the LWE-SDF close to zero frequency ($\eta_j \xrightarrow[\gamma\to\infty]{} 0$) can be studied under **SC2** as given in the next section .

(2,2) The second condition **SC2** ($\eta_j \xrightarrow[\gamma\to\infty]{} 0$) helps to approximately rewritten the LWE-SDF as a linear regression model and thus use the ordinary least squares estimator (OLSE) of the DOM (OLSE-DOM) to simply study the LWE-DOM as given in next section.

(3) The third condition **SC3** is given for improving the LWE-SDF and the LWE-DOM of a LMDSGP, where these estimators will be unbiased estimators under **SC3** as given in Sections 3.3. Moreover, the LWE-DOM is a better estimator compared to the PE-DOM under **SC3** for the simulation study as given in Chapter 4.

(4) The fourth condition **SC4** is provided to find the bias of the normalized LWE-SDF as given in the proof of Theorem 1 (cf. the appendix). Moreover, the behavior of normalized LWE-SDF under **SC4** will be used to study the behavior of the LWE-DOM.

3.2 Asymptotic Properties of LWE-SDF under the New Conditions

To obtain the semi-parametric spectral estimation of the DOM in the next section, the asymptotic behavior of the normalized LWE-SDF for the frequencies $\eta_j = \frac{2\pi j}{\gamma} \neq 0$ and any DSGP $\{X_t\}_{t\in\mathbb{Z}}$ with $0 < |d| < \frac{1}{2}$ under the above sufficient conditions need to be investigated. For this purpose, the LWE-SDF in (3.5) can be approximately given in terms of the discrete sum over the Fourier frequencies $\theta_k = \frac{2\pi k}{\gamma}$ as follows

$$\widehat{\mathcal{S}}_w(\eta) \simeq \frac{2\pi}{\gamma} \sum_{k=-[\frac{\gamma-1}{2}]}^{[\frac{\gamma}{2}]} \widehat{\mathcal{S}}_p(\theta_k)\mathcal{U}_{t_\gamma}(\eta - \theta_k). \tag{3.12}$$

Based on the LWE-SDF in (3.12), the asymptotic behaviors of the bias, variance and distribution of the normalized LWE-SDF are investigated as follows

3.2.1 Asymptotic Consistency

Theorem 3. *For any DSGP $\{X_t\}_{t\in\mathbb{Z}}$ with SDF $\mathcal{S}(\eta)$ and ACF $\mathcal{A}(h)$ given in terms of a slowly varying function in (2.13) and (2.14), respectively. If the sufficient conditions* **SC1-SC4** *are fulfilled, the bias of the normalized LWE-SDF has the following asymptotic behavior*

$$Bias\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) \simeq \begin{cases} O\left(\frac{1}{t_\gamma \eta_j}\right)^2, & \textit{if} \quad 0 < d < \frac{1}{2}, \ \textit{i.e., for LMDSGP;} \\ \\ O\left(\frac{1}{j}\right), & \textit{if} \quad -\frac{1}{2} < d < 0, \ \textit{i.e., for NMDSGP.} \end{cases} \tag{3.13}$$

Theorem 4. *For any LMDSGP or SMDSGP $\{X_t\}_{t\in\mathbb{Z}}$ with SDF $\mathcal{S}(\eta)$ given in terms of a slowly varying function at the infinity given by (2.13). If the conditions* **SC1** *and* **SC2** *are fulfilled, the variance of the normalized LWE-SDF has the following asymptotic behavior*

$$var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq O\left(\frac{t_\gamma}{\gamma}\right), \quad 0 < |d| < \frac{1}{2}.$$

Corollary 1. *From Theorems 1 and 2, the bias and variance of the normalized PE-SDF have the following asymptotic behaviors*

$$Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \simeq \begin{cases} O\left(\frac{\ln j}{j}\right), & \textit{if} \quad 0 < d < \frac{1}{2}, \ \textit{i.e., for LMDSGP;} \\ \\ O\left(\frac{1}{j}\right), & \textit{if} \quad -\frac{1}{2} < d < 0, \ \textit{i.e., for NMDSGP.} \end{cases}$$

$$var\left(\frac{\widehat{\mathcal{S}}_P(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq O\left(1\right), \quad 0 < |d| < \frac{1}{2}.$$

Corollary 2. *From Theorems 1 and 2 and Corollary 1, we can show that*

- *The normalized LWE-SDF is an asymptotic unbiased and consistent estimator under the above sufficient conditions, where*

$$Bias\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) \xrightarrow[j\to\infty]{SC3} 0 \text{ and } var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) \xrightarrow[j\to\infty]{SC1} 0.$$

- *The normalized PE-SDF is an asymptotic unbiased but non-consistent estimator, where*

$$Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \xrightarrow[j\to\infty]{} 0 \text{ and } var\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \nrightarrow 0.$$

- *Therefore, the asymptotic behavior of the (normalized) LWE-SDF under the sufficient conditions is much better than the asymptotic behavior of the (normalized) PE-SDF.*

3.2.2 Asymptotic Distribution

The asymptotic distributions of the PE-SDF for SMDSGPs and LMDSGPs have been studied in [42] and [35], respectively. In this paper, those results are extended to the LWE-SDF as follows

Theorem 5. *For any LMDSGP or SMDSGP $\{X_t\}_{t\in\mathbb{Z}}$ with SDF $\mathcal{S}(\eta)$ given in terms of a slowly varying function at the infinity given by (2.13), let $\{\eta_{j\gamma}\}_{j=1}^{\varphi\geq 2}$ be a sequence of asymptotically distant frequencies, i.e.,*

$$|\gamma(\eta_{j\gamma}-\eta_{i\gamma})| \xrightarrow[1\leq i\neq j\leq\varphi]{} \infty \quad \text{with} \quad \eta_{j\gamma} \xrightarrow[\gamma\to\infty]{} \eta_j.$$

If the sufficient conditions **SC1-SC4** *are fulfilled and $\gamma|\eta_{i\gamma}| \xrightarrow[\gamma\to\infty]{} \infty$, the asymptotic distribution of the normalized LWE-SDF follows the following chi-square distribution*

$$\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}_(\eta_j)} \xrightarrow[\gamma\to\infty]{D} \frac{1}{\nu}\chi^2(\nu),\ \ \nu = \frac{2\gamma}{t_\gamma\int_{-1}^{1}\mathcal{W}^2(x)dx},$$

where $X \xrightarrow{D} Y$ means that X and Y have the same distribution and ν is the degree of freedom of χ^2.

Corollary 3. *From Theorem 5, the $(1-\alpha)$ confidence interval of the SDF $\mathcal{S}(\eta_j)$ for $\eta_j \neq 0$ is given as follows*

$$\frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{l_2} \leq \mathcal{S}(\eta_j) \leq \frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{l_1},$$

where l_1 and l_2 are obtained from the table of the chi-square distribution satisfying

$$P(\chi^2(\nu)\leq l_2) = P(\chi^2(\nu)\geq l_1) = \alpha/2.$$

Remark 4. *It is worth mentioning that the result in Theorem 5 is compatible with Theorems 1 and 2. For particular use, here are some special windows:*

- *(i) For the Bartlett window, we have*
  $\mathcal{W}(x) = (1 - |x|)\mathbb{1}_{[-1,1]}(x)$, $\int_{-1}^{1} \mathcal{W}^2(x)dx = 2\int_0^1 (1-x)^2 dx = \frac{2}{3}$, *and* $\nu = \frac{3\gamma}{t_\gamma}$.

- *(ii) For the Parzen window, we have*
  $\mathcal{W}(x) = (1 - x^2)\mathbb{1}_{[-1,1]}(x)$, $\int_{-1}^{1} \mathcal{W}^2(x)dx = 2\int_0^1 (1-x^2)^2 dx = \frac{16}{15}$, *and* $\nu = 1.87\,\frac{\gamma}{t_\gamma}$.

Moreover, the results of the normalized PE-SDF can be obtained as special cases of Theorem 5 and Corollary 3 by taking $\nu = 2$.

3.3 Asymptotic Properties of LWE-DOM under the New Conditions

The definition of LMSDGP, in the frequency domain, only specify the behavior of the SDF close to the zero frequency, the case (2.11) or (2.13). Similarly, in the time domain, the definitions only specify the behavior of the autocovariances at long lags, (2.12) and (2.14).

Hence, the relevant estimation procedures are those that employ sample information only in a neighborhood of the zero frequency (2.11 or 2.13), or at long lags (2.12 and 2.14). All these are called semiparametric procedures (Robinson, 1994a; 1994c; 1995)[47, 48, 49].

Semiparametric estimation of the DOM, in LMSDGP, is appealing in empirical work, due to its dependence on the treatment of the short-memory component in the estimation. Two common statistical procedures in this class, the regression method (REM) (cf, [47, 31]) and the Local Whittle method (WEM) (cf, [49, 56]. The WEM is known to be more efficient than the REM in the stationary case ($|d| < 1/2$), although numerical optimization methods are needed in the calculation (cf, [54]).

The main feature of the semiparametric approach is to employ a bandwidth number. In the frequency domain, a spectral bandwidth number n_γ, this number reflects the highest frequency, $\eta_{n_\gamma} = \frac{2\pi n_\gamma}{\gamma}$, where γ is the sample size, at which the statistics used to estimate the SDF are evaluated. In the time domain, the bandwidth number, t_γ, reflects the lowest sample autocovariance employed (see Robinson, [48]). In order to develop the asymptotic theory, n_γ has to tend to infinity as γ tends to infinity, but in such a way that their ratio tends to zero. The two procedures based on the estimation of the SDF.

3.3.1 Regression Method

The most applied semiparametric one in the literature, it has been the REM, it is introduced by Geweke and Porter- Hudak (1983) [21] and improved by Robinson (1995)[48]. This estimate is easily computed (it just involves ordinary least squares). For any DSGP $\{X_t\}_{t\in\mathbb{Z}}$ with SDF given by (2.13), let $\widehat{d}_p$ and $\widehat{d}_w$ be the PE-DOM and LWE-DOM, respectively. The REM provides a general method for estimating the DOM without assuming a parametric model. The form of the SDF in (2.13) is asymptotically equivalent to the following form

$$\mathcal{S}(\eta) \simeq |1 - e^{-i\eta}|^{-2d}\mathcal{S}^*(\eta), \tag{3.14}$$

where

$$|1 - e^{-i\eta}|^{-2d} = |\eta|^{-2d}(1 + o(1)) \quad \text{as} \quad \eta \to 0$$

and the function $\mathcal{S}^*(.)$ satisfies the same conditions as $\mathcal{V}_{\mathcal{S}}(.)$ in (2.13). After taking logarithms and adding $\ln\widehat{\mathcal{S}}(\eta)$ to both sides of (3.14), where $\widehat{\mathcal{S}}(\eta)$ is any estimator of the SDF, and evaluating at Fourier frequencies $\eta_j = \frac{2\pi j}{\gamma}$, $j = 1, \ldots, n_\gamma$, we get

$$\ln\widehat{\mathcal{S}}(\eta_j) = \ln\mathcal{S}^*(0) - d\,\ln\left(4\sin^2\left(\frac{\eta_j}{2}\right)\right) + \ln\left(\frac{\widehat{\mathcal{S}}(\eta_j)}{\mathcal{S}(\eta_j)}\right) + \ln\left(\frac{\mathcal{S}^*(\eta_j)}{\mathcal{S}^*(0)}\right). \tag{3.15}$$

If n_γ satisfies the sufficient condition **SC2**, the last term in (3.15) can be negligible. Therefore, (3.15) can be approximately rewritten as the following regression form

$$v_j \simeq a + b\,u_j + e_j, \tag{3.16}$$

where $v_j = \ln\widehat{\mathcal{S}}(\eta_j)$, $u_j = \ln\left(4\sin^2\left(\frac{\eta_j}{2}\right)\right)$, $e_j = \ln\left(\frac{\widehat{\mathcal{S}}(\eta_j)}{\mathcal{S}(\eta_j)}\right)$, $a = \ln\mathcal{S}^*(0)$ and $b = -d$. Then, the OLSE-DOM is given as follows

$$\widehat{d}_o := -\frac{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})(v_j - \bar{v})}{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2}. \quad \text{Then, we get,} \quad \widehat{d}_o - d = \frac{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})e_j}{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2}.$$

The bias and variance of the OLSE-DOM are given as follows

$$\begin{cases} Bias(\widehat{d}_o) = E(\widehat{d}_o) - d = \frac{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})E(e_j)}{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2}, \\ \\ var(\widehat{d}_o) = \frac{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2 var(e_j)}{\left(\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2\right)^2}. \end{cases} \tag{3.17}$$

Remark 5. *It is worth mentioning that the OLSE-DOM in (3.17) depends on the estimator of the SDF. If the PE-SDF is used, the OLSE-DOM will be the PE-DOM. If the LWE-SDF is used, the OLSE-DOM will be the LWE-DOM.*

Geweke and Porter-Hudak [21] discussed the REM (i.e., the form in (3.16)) based on the PE-SDF by taking

$$v_j = \ln \widehat{\mathcal{S}}_p(\eta_j),\ a = \ln \mathcal{S}^*(0) - \zeta \text{ and } e_j = \ln\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right) + \zeta,$$

where $\zeta = 0.5772$ is the Euler constant. When $\gamma \to \infty$, $-\ln\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right)$, $j = 1, ..., n_\gamma$ are independent identically distributed (i.i.d) random variables (rvs) follow standard Gumbel distribution (cf. Hassler[27, 28]). Therefore, we get

$$E\left(\ln\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \xrightarrow[\gamma\to\infty]{} -\zeta \quad \text{and} \quad var\left(\ln\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \xrightarrow[\gamma\to\infty]{} \frac{\pi^2}{6}. \tag{3.18}$$

From (29), the asymptotic mean and variance of the e_j in the REM based on the PE-SDF are given as follows

$$E(e_j) = E\left(\ln\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right) + \zeta\right) \xrightarrow[\gamma\to\infty]{} 0 \quad \text{and} \quad var(e_j) \xrightarrow[\gamma\to\infty]{} \frac{\pi^2}{6}. \tag{3.19}$$

From Geweke and Porter-Hudak [21], we get

$$\sum_{j=1}^{n_\gamma} (u_j - \bar{u})^2 \simeq n_\gamma. \tag{3.20}$$

Since $\{e_j\}_{j=1}^{n_\gamma}$ are i.i.d rvs, we get the following asymptotic bias and variance of the PE-DOM by combining (28), (30) and (31)

$$Bias(\widehat{d}_p) \xrightarrow[\gamma\to\infty]{} E(e_j) = 0 \quad \text{and} \quad var(\widehat{d}_p) \xrightarrow[\gamma\to\infty]{} \frac{\pi^2}{6n_\gamma}. \tag{3.21}$$

From (3.21), the PE-DOM has the following asymptotic normal distribution (cf. Theorem 1 in Hassler[27])

$$\sqrt{n_\gamma}(\widehat{d}_p - d) \xrightarrow[\gamma\to\infty]{D} \mathcal{N}(0, \frac{\pi^2}{6}). \tag{3.22}$$

The estimator of the DOM via the REM proposed by Geweke and Porter-Hudak [21] is asymptotically biased if n_γ used in the REM is held fixed as $\gamma \to \infty$ (cf. [33]) (i.e., the absence of the condition **SC2**). Then, improving the estimator of the DOM via the REM is needed. In this paper, the mean and variance of the LWE-DOM under the sufficient conditions are investigated as follows

Theorem 6. *For any LMDSGP $\{X_t\}_{t\in\mathbb{Z}}$ satisfies the sufficient conditions **SC1-SC4**, the asymptotic behavior of the bias and variance of the LWE-DOM for the REM are given from the OLSE-DOM in (3.17) and from the Theorem 3 as follows*

$$Bias(\widehat{d}_w) \xrightarrow[\gamma\to\infty]{} E\left(\ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \quad \textit{and} \quad var(\widehat{d}_w) \simeq \frac{t_\gamma}{\gamma n_\gamma}\int_{-1}^{1} \mathcal{W}^2(x)dx, \ \textit{as } \gamma \to \infty.$$

Corollary 4. *From Theorems 3 and 4, for any LMDSGP* $\{X_t\}_{t\in\mathbb{Z}}$ *the LWE-DOM has asymptotic normal distribution as follows*

$$\sqrt{\frac{\gamma n_\gamma}{t_\gamma}}(\widehat{d}_w - d) \xrightarrow[\gamma\to\infty]{D} \mathcal{N}\left(0, \int_{-1}^{1} \mathcal{W}^2(x)dx\right).$$

Corollary 5. *From the above discussions and results, it is obvious that*

- *The LWE-DOM is an asymptotic unbiased and consistent estimator under the above sufficient conditions, where*

$$Bias(\widehat{d}_w) \xrightarrow[\gamma\to\infty]{\mathbf{SC3}} 0 \quad and \quad var(\widehat{d}_w) \xrightarrow[\gamma\to\infty]{\mathbf{SC1},\mathbf{SC2}} 0.$$

- *The approximate value of the variance of the LWE-DOM is less than the approximate value of the variance of the PE-DOM, i.e.,*

$$var(\widehat{d}_w) \simeq \frac{t_\gamma}{\gamma n_\gamma}\int_{-1}^{1} \mathcal{W}^2(x)dx < var(\widehat{d}_p) \simeq \frac{\pi^2}{6n_\gamma}.$$

- *Therefore, the asymptotic behavior of the LWE-DOM under the sufficient conditions is much better than the asymptotic behavior of the PE-DOM.*

Remark 6. *The REM has two main drawbacks:*

• *In order to derive a Gaussian asymptotic distribution, Robinson needed to assume the Gaussianity (which is very restrictive and probably not valid for most of financial series) and,*

• *furthermore, he needed to introduce an additional user-chosen number to trim out frequencies very close to zero, it is the* $\ln \mathcal{S}^*(0)$ *in 3.15.*

3.3.2 Local Whittle Method

The Local Whittle name seems to be widely spread nowadays, although Robinson (1995b)[49] called this estimator Gaussian semiparametric. In fact, Künsch (1987, p. 71)[34] was the first to suggest this estimator, however, without establishing any statistical properties. Unlike the Whittle estimator, the Local Whittle estimator is a semiparametric method, in that only specifies the parametric form of the SDF when the frequency is close to zero, which gives a model close to Model (2.13) and (2.11). To define Local Whittle estimator, first consider the approximate log-likelihood of a Gaussian process with SDF $\mathcal{S}$:

$$L(\mathcal{S}) := \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\log \mathcal{S}(\eta) + \frac{\widehat{\mathcal{S}}(\eta)}{\mathcal{S}(\eta)}\right)d\eta,$$

the above approximation of the quadratic form, $\frac{\widehat{\mathcal{S}}(\eta)}{\mathcal{S}(\eta)}$, in the exponent of Gaussian density (see Beran [6]) is similar to Whittle's estimate. Therefore, the SDF has a pole at zero frequency when $d > 0$ for the LMDSGP. The Local Whittle estimator is based on the SDE, using the PE $\widehat{\mathcal{S}_p}$ or the LWE $\widehat{\mathcal{S}_w}$, and its computation involves the number of Fourier frequencies, n_γ, satisfying the **SC2**, i.e

$$1/n_\gamma + n_\gamma/\gamma \to 0 \quad \text{as;} \quad \gamma \to \infty.$$

The above estimator is further simplified by the following considerations:

i) Replacing $\mathcal{S}(\eta)$ by the approximate formula $\alpha|\eta|^{-2d}$ defined in 2.11 , $L(\mathcal{S})$ becomes

$$L(\alpha; d) := \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\log(\alpha|\eta|^{-2d}) + \frac{\widehat{\mathcal{S}}(\eta)}{\alpha|\eta|^{-2d}}\right) d\eta.$$

ii) Since the behavior of $\mathcal{S}$ near $\eta = 0$ is only important (we want to estimate, asymptotic, the parameters α and d only). Restricting the integral to low frequencies

$$|\eta| < \eta_{n_\gamma} := \frac{2\eta n_\gamma}{\gamma}, n_\gamma \to \infty, \quad n_\gamma/\gamma \to 0,$$

so, we take the Fourier frequencier, $\eta_j = \frac{2\pi j}{\gamma}, j = 1, ..., n_\gamma$.

iii) Replacing integration by summation over $1 \leq j < n_\gamma$. For a SDF satisfying (2.11), the discrete analogue of the function $L(,)$ called objective function (see Künsch (1987) [34]) in the Whittle estimator is

$$L(\alpha; d) = \frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}\left(\log(\alpha|\eta_j|^{-2d}) + \frac{\widehat{\mathcal{S}}(\eta_j)}{\alpha|\eta_j|^{-2d}}\right). \tag{3.23}$$

iv) The Local Whittle estimate minimizes the above objective function $L(\alpha; d)$, the resulting approximate log-likelihood is

$$(\widehat{\alpha}; \widehat{d}) := \underset{0<\alpha,\ d\in(-1/2,\ 1/2)}{\text{argmin}} L(\alpha; d).$$

v) For fixed d, the minimum of $L(\alpha; d)$ is achieved by

$$\widehat{\alpha} := \frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}|\eta_j|^{2d}\widehat{\mathcal{S}}(\eta_j)$$

vi) Replacing in (3.23) the constant α by its estimate, one obtains

$$L(d) := L(\widehat{\alpha}, d) - 1 = \log\left(\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}|\eta_j|^{2d}\widehat{\mathcal{S}}(\eta_j)\right) - 2d\,\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}\log(\eta_j), \tag{3.24}$$

and the local Whittle estimator $\widehat{d}_{LW}$ of the fractional parameter d is defined by

$$\widehat{d}_{LW} := \underset{d\in(-1/2,1/2)}{\text{argmin}} L(d). \tag{3.25}$$

- **Using the Periodogram**

  Under the relationship (3.25), we establisheb the following formula,

$$\widehat{d}_{LW_P} := \underset{d\in(-1/2,1/2)}{\text{argmin}} \; L_P(d), \tag{3.26}$$

  where

$$L_P(d) := \log\left(\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}|\eta_j|^{2d}\widehat{\mathcal{S}}_P(\eta_j)\right) - 2d\,\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}\log(\eta_j).$$

  Furthermore, the value of $\widehat{d}_{LW_P}$ that minimizes $L_P(d)$ converges in probability to the actual value of d under certain assumptions. Robinson (1995b) [49] and Shimotsu [54] were established the finite distribution theory, they have the following result

$$\sqrt{n_\gamma}\,(\widehat{d}_{LW_P} - d) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1/4), \quad \text{as} \quad \gamma \to \infty.$$

  The local Whittle estimator was developed by Robinson (1995b) [49] and was further studied, along with the Whittle and Aggregated Whittle estimators, by Taqqu and Teverovsky (1996)[56] and Shimotsu [54].

- **Using the Lag Window**

  As the same of the regression method, we replace the PE $\widehat{\mathcal{S}_p}$ by the LWE $\widehat{\mathcal{S}}_w$ given in (3.3), in the objective function (3.23) and (4.3). The resulting estimator for d is

$$\widehat{d}_{LW_w} := \underset{d\in(-1/2,1/2)}{\text{argmin}} \; L_w(d), \tag{3.27}$$

  where

$$L_w(d) := \log\left(\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}|\eta_j|^{2d}\widehat{\mathcal{S}}_w(\eta_j)\right) - 2d\,\frac{1}{n_\gamma}\sum_{j=1}^{n_\gamma}\log(\eta_j).$$

The asymptotic normality of the Local Whittle estimator of $d \in (0,\ 1/2)$, the approximate MSE and the corresponding optimal bandwidth, will be part of the future perspectives.

Chapter 4

Simulation Study of the Accuuracy of LWT

4.1 Description of Simulation

The above theoretical justifications demonstrated that the LWE-DOM is better than the PE-DOM for the REM under the new sufficient conditions. Therefore, there are the following three logical scenarios to check the significance of the new conditions for improving the accuracy of the LWE over the PE:

- **First logical scenario:** Simulation (numerical) justifications are needed to support the correctness of the proved theoretical justifications, i.e., the LWE-DOM is better than the PE-DOM for the REM under the new sufficient conditions.

- **Second logical scenario:** Even though the theoretical justification of the efficiency of the LWT under the new conditions is proved based on the REM only, the logical question is that what is the performance of the LWT under the new conditions based on other models, such as the WEM (cf. Frederiksen et al.[20]), i.e., can the new sufficient conditions be used for improving the performance of the LWT for also different models?

- **Third logical scenario:** After investigating the first two logical scenarios, the logical third question to check the significance of the new sufficient conditions for improving the performance of the LWT is that: what is the performance of the LWT without the new sufficient conditions?

The programs used in the simulation are algorithms directly implemented on the above methods in the MATALB language. this language allows us to study the above mentioned three scenarios, which are given in the following two preparation steps

Preparation step 1: Selecting the parameters to test the sufficient conditions: The ARFIMA(p, d, q) models with a normally distributed $\mathcal{N}(0, 1)$ white noise process, $0 \leq p, q \leq 2$ and DOM $0 \leq d < \frac{1}{2}$ are used in the following simulation study. It is obvious that, the SDFs of these models satisfy the fourth sufficient condition **SC4**. The truncation points t_γ is selected as $t_\gamma = \gamma^a,\ 0 < a < 1$ to satisfy the first sufficient condition **SC1**. The spectral bandwidth parameter n_γ is selected as $n_\gamma = \gamma^b,\ 0 < b < 1$ to satisfy the second sufficient condition **SC2**. Therefore, the three sufficient conditions **SC1**, **SC2** and **SC4** are satisfied in all of the following simulation results. Thus, only the third condition **SC3** need to be tested. It is obvious that, under $n_\gamma = \gamma^b,\ 0 < b < 1$ and $t_\gamma = \gamma^a,\ 0 < a < 1$ the third condition **SC3** is satisfied if and only if $a + b > 1$, otherwise it is not satisfied.

Preparation step 2: Selecting the optimal window for the LWT: The selection of the optimal window (i.e., LWF) for the LWT is investigated in this step as follows. The three widely used LWFs, Hamming LWF, Bartlett LWF and Blackman LWF, are tested under the new sufficient conditions. The selection of the preferred window is based on the smallest MSE over 1000 replications for the corresponding LWE-DOM. The data are generated from long memory ARIMA$(0, d, 0)$ models with DOM $0 < d < \frac{1}{2}$, sample size $\gamma \in \{700, 1000, 2000\}$, spectral bandwidth parameter $n_\gamma = \gamma^{0.65}$ and truncation points $t_\gamma \in \{\gamma^{0.396}, \gamma^{0.496}\}$. It is obvious from the **Preparation step 1** that, these selections of the parameters satisfy all of the four sufficient conditions **SC1-SC4**. The MSE of an estimator is given as follows

$$\text{MSE} = \frac{1}{1000} \sum_{k=1}^{1000} \left(\hat{d}_w^{(k)} - d \right)^2, \text{ where } \hat{d}_w^{(k)} \text{ is the } k\text{th LWE of } d.$$

The MSE values of the LWE-DOMs using the Hamming LWF, Bartlett LWF and Blackman LWF are given in Table 4.2.

• Figure 4.2 gives a clear comparison study among the MSE values of the LWE-DOMs under these three LWFs. From Table 4.2 and Figure 4.2, we can show that the MSE values of the LWE-DOMs using the Hamming LWF are less than the MSE values of the LWE-DOMs using Bartlett LWF and Blackman LWF. Therefore, the

Hamming LWF is the optimal window that will be used to get the LWE-DOMs in this study.

4.2 Numerical Results

Table 4.1: MSE of the LWE-DOMs under the new conditions via the different LWF

Via the Hamming LWF, Bartlett LWF and Blackman LWF using spectral bandwidth $n_\gamma = \gamma^{0.65}$ and different values of sample size γ and DOM d

$\gamma = 700$	LWF	$d = 0.05$	$d = 0.10$	$d = 0.15$	$d = 0.20$	$d = 0.25$	$d = 0.30$	$d = 0.35$	$d = 0.4$	$d = 0.45$
	Hamming	0.0172	0.0359	0.0339	0.0614	0.0651	0.0286	0.0350	0.0315	0.0894
$t_\gamma = \gamma^{0.396}$	Bartlett	0.0172	0.0396	0.0344	0.0669	0.0730	0.0452	0.0508	0.0573	0.1101
	Blackman	0.0194	0.0387	0.0361	0.0651	0.0700	0.0411	0.0525	0.0481	0.1071
	Hamming	0.0460	0.0614	0.1112	0.1142	0.1395	0.1501	0.1833	0.1916	0.2324
$t_\gamma = \gamma^{0.496}$	Bartlett	0.0531	0.0633	0.1202	0.1170	0.1392	0.1459	0.1726	0.1952	0.2396
	Blackman	0.0498	0.0660	0.1168	0.1221	0.1454	0.1643	0.1988	0.2193	0.2644
$\gamma = 1000$	LWF	$d = 0.05$	$d = 0.10$	$d = 0.15$	$d = 0.20$	$d = 0.25$	$d = 0.30$	$d = 0.35$	$d = 0.4$	$d = 0.45$
	Hamming	0.0289	0.0208	0.0423	0.0245	0.0320	0.0540	0.0258	0.0508	0.0553
$t_\gamma = \gamma^{0.396}$	Bartlett	0.0296	0.0224	0.0466	0.0268	0.0366	0.0654	0.0411	0.0489	0.0529
	Blackman	0.0299	0.0213	0.0489	0.0257	0.0383	0.0648	0.0408	0.0651	0.0674
	Hamming	0.0329	0.0178	0.0485	0.0275	0.0385	0.0363	0.0225	0.0252	0.0603
$t_\gamma = \gamma^{0.496}$	Bartlett	0.0330	0.0186	0.0494	0.0313	0.0411	0.0432	0.0298	0.0366	0.0687
	Blackman	0.0330	0.0173	0.0480	0.0290	0.0406	0.0388	0.0218	0.0275	0.0658
$\gamma = 2000$	LWF	$d = 0.05$	$d = 0.1$	$d = 0.15$	$d = 0.20$	$d = 0.25$	$d = 0.30$	$d = 0.35$	$d = 0.40$	$d = 0.45$
	Hamming	0.0191	0.0152	0.0233	0.0292	0.0358	0.0307	0.0194	0.0484	0.0132
$t_\gamma = \gamma^{0.396}$	Bartlett	0.0200	0.0161	0.0271	0.0361	0.0371	0.0468	0.0368	0.0573	0.0102
	Blackman	0.0208	0.0182	0.0278	0.0366	0.0432	0.0440	0.0330	0.0629	0.0209
	Hamming	0.0 239	0.0406	0.0147	0.0246	0.0191	0.0248	0.0118	0.0316	0.0255
$t_\gamma = \gamma^{0.496}$	Bartlett	0.0246	0.0408	0.0157	0.0256	0.0235	0.0336	0.0116	0.0345	0.0259
	Blackman	0.0245	0.0410	0.0150	0.0247	0.0200	0.264	0.0192	0.0350	0.0294

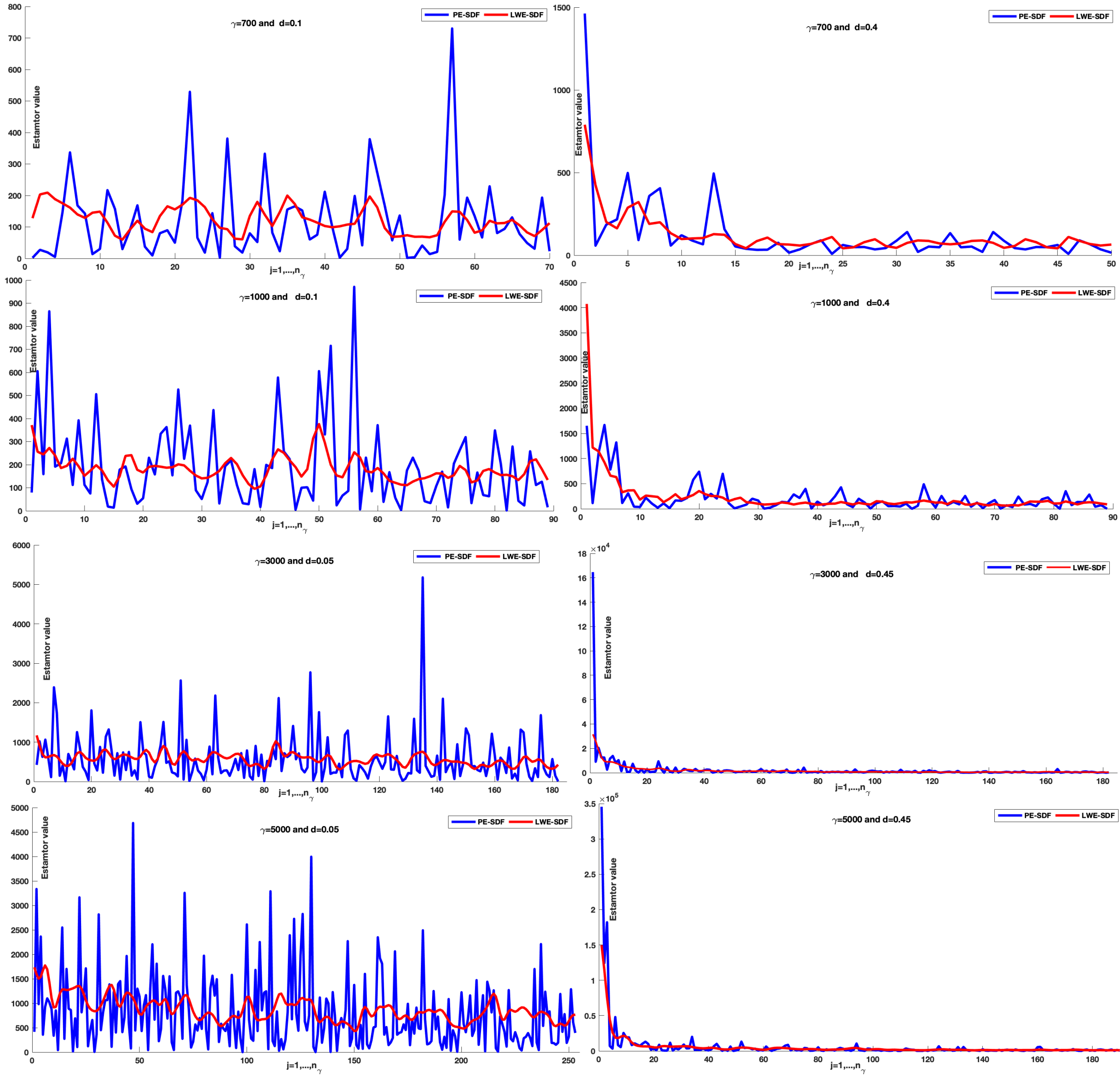

Figure 4.1: Comparative graph of SDF for PE and LWE with **SC3**

A comparison study among the PE-SDF and the LWE-SDF for ARFIMA$(0, d, 0)$ under the new conditions by taking $n_\gamma = \gamma^{0.65}$, $t_\gamma = \gamma^{0.696}$ and different values of sample size γ and DOM d

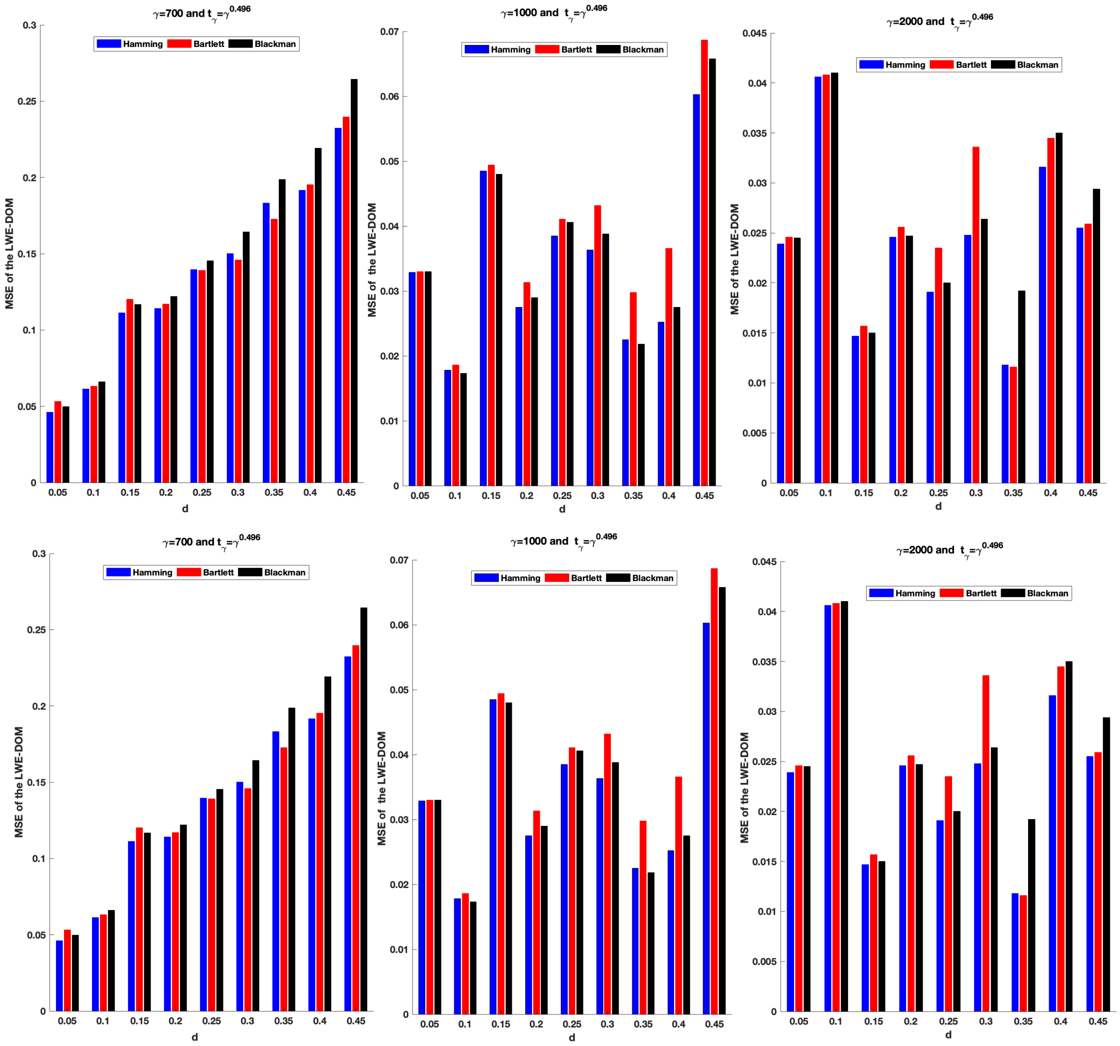

Figure 4.2: Choise of LWF

A comparison study among the MSE values of the LWE-DOMs via the Hamming LWF, Bartlett LWF and Blackman LWF in Table 4.2

Table 4.2: MSE of DOM for ARFIMA (0,d,0) using the REM with **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(0, d, 0)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.0590	0.0551	0.0760	0.1061	0.1379	0.1328	0.2139	0.2885	0.2623	0.2604
	LWE	0.0436	0.0332	0.0454	0.0605	0.1006	0.0932	0.1524	0.1636	0.1897	0.1921
700	PE	0.0590	0.0809	0.1127	0.1165	0.1114	0.1357	0.1979	0.2315	0.2309	0.2585
	LWE	0.0480	0.0549	0.0536	0.0634	0.0640	0.0863	0.1077	0.1030	0.1115	0.1474
1000	PE	0.0328	0.0351	0.0574	0.0921	0.1073	0.1462	0.1789	0.2106	0.2020	0.2828
	LWE	0.0232	0.0357	0.0218	0.0380	0.0436	0.0685	0.0591	0.0795	0.0779	0.0926
1500	PE	0.0321	0.0498	0.0523	0.1013	0.1217	0.1345	0.1952	0.1894	0.2114	0.2594
	LWE	0.0226	0.0219	0.0391	0.0494	0.0462	0.0328	0.0624	0.0479	0.0566	0.0470
2000	PE	0.0215	0.0178	0.0577	0.0835	0.1213	0.1324	0.1486	0.1940	0.2245	0.2474
	LWE	0.0079	0.0057	0.0326	0.0205	0.0218	0.0415	0.0232	0.0547	0.0427	0.0692
2500	PE	0.0284	0.0257	0.0678	0.0913	0.1142	0.1635	0.1569	0.1836	0.2141	0.2186
	LWE	0.0385	0.0116	0.0104	0.0157	0.0275	0.0255	0.0301	0.0173	0.0437	0.0398
3000	PE	0.0215	0.0382	0.0417	0.0753	0.1125	0.1165	0.1639	0.2101	0.2337	0.2427
	LWE	0.0123	0.0259	0.0171	0.0228	0.0412	0.0189	0.0295	0.0229	0.0434	0.0327
3500	PE	0.0386	0.0363	0.0745	0.0924	0.1198	0.1511	0.1712	0.1948	0.2067	0.2500
	LWE	0.0153	0.0100	0.0171	0.0422	0.0331	0.0318	0.0282	0.0468	0.0216	0.0296
4000	PE	0.0240	0.0392	0.0541	0.0724	0.0836	0.1169	0.1564	0.1737	0.1973	0.2273
	LWE	0.0131	0.0170	0.0165	0.0260	0.0209	0.0356	0.0161	0.0234	0.0288	0.0378
4500	PE	0.0215	0.0275	0.0508	0.0845	0.1051	0.1249	0.1374	0.1895	0.2279	0.2418
	LWE	0.0101	0.0078	0.0135	0.0201	0.0188	0.0175	0.0222	0.0222	0.0291	0.0248
5000	PE	0.0226	0.0360	0.0683	0.0882	0.1123	0.1274	0.1609	0.1986	0.2008	0.2216
	LWE	0.0207	0.0174	0.0386	0.0201	0.0198	0.0399	0.0226	0.0197	0.0148	0.0443

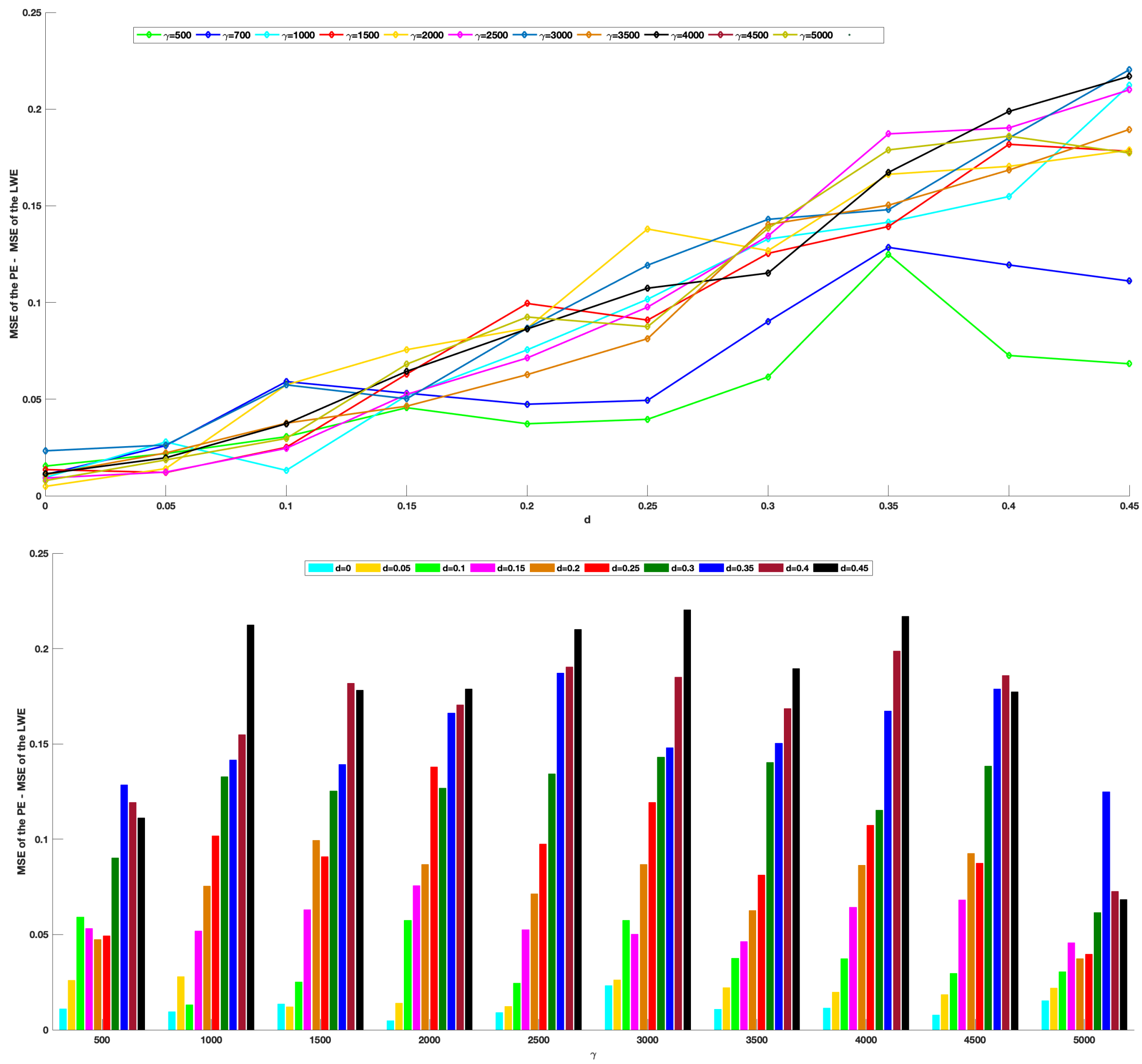

Figure 4.3: MSE of DOM for ARFIMA (0,d,0) using the REM with **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA $(0, d, 0)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.2

Table 4.3: MSE of DOM for ARFIMA (1,d,1) using the REM with **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(1, d, 1)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.0635	0.0720	0.1068	0.0612	0.1451	0.1717	0.1994	0.1912	0.2701	0.2848
	LWE	0.0248	0.0535	0.0702	0.0588	0.1042	0.1101	0.1585	0.1479	0.1899	0.1962
700	PE	0.0477	0.0531	0.0593	0.0845	0.1379	0.1300	0.1549	0.2344	0.2273	0.2482
	LWE	0.0228	0.0434	0.0459	0.0729	0.0923	0.0750	0.0888	0.1108	0.1214	0.1381
1000	PE	0.0373	0.0569	0.0683	0.1346	0.1131	0.1467	0.1611	0.1824	0.1907	0.2512
	LWE	0.0111	0.0125	0.0235	0.0817	0.0536	0.0673	0.0553	0.0648	0.0628	0.0859
1500	PE	0.0210	0.0455	0.0534	0.0863	0.1220	0.1331	0.1665	0.1758	0.2362	0.2692
	LWE	0.0157	0.0162	0.0199	0.0265	0.0370	0.0332	0.0487	0.0455	0.0665	0.0747
2000	PE	0.0250	0.0521	0.0777	0.0976	0.0754	0.1426	0.1611	0.1555	0.2167	0.2180
	LWE	0.0086	0.0185	0.0417	0.0182	0.0315	0.0336	0.0365	0.0618	0.0585	0.0552
2500	PE	0.0191	0.0262	0.0581	0.0733	0.1081	0.1400	0.1720	0.2199	0.1966	0.2343
	LWE	0.0078	0.0221	0.0259	0.0135	0.0435	0.0428	0.0440	0.0634	0.0323	0.0265
3000	PE	0.0304	0.0482	0.0433	0.0809	0.0877	0.1630	0.1333	0.1915	0.2138	0.2559
	LWE	0.0274	0.0324	0.0130	0.0331	0.0166	0.0272	0.0199	0.0266	0.0361	0.0563
3500	PE	0.1022	0.1241	0.1262	0.1488	0.1710	0.1823	0.2178	0.2397	0.2549	0.2711
	LWE	0.0193	0.0094	0.0282	0.0225	0.0111	0.0185	0.0152	0.0363	0.0288	0.0278
4000	PE	0.0240	0.0392	0.0541	0.0724	0.0836	0.1169	0.1564	0.1737	0.1973	0.2273
	LWE	0.0815	0.0666	0.0555	0.0840	0.0977	0.0976	0.0997	0.1363	0.1281	0.1868
4500	PE	0.0215	0.0275	0.0508	0.0845	0.1051	0.1249	0.1374	0.1895	0.2279	0.2418
	LWE	0.0101	0.0078	0.0135	0.0201	0.0188	0.0175	0.0222	0.0222	0.0291	0.0248
5000	PE	0.0226	0.0360	0.0683	0.0882	0.1123	0.1274	0.1609	0.1986	0.2008	0.2216
	LWE	0.0207	0.0174	0.0386	0.0201	0.0198	0.0399	0.0226	0.0197	0.0148	0.0443

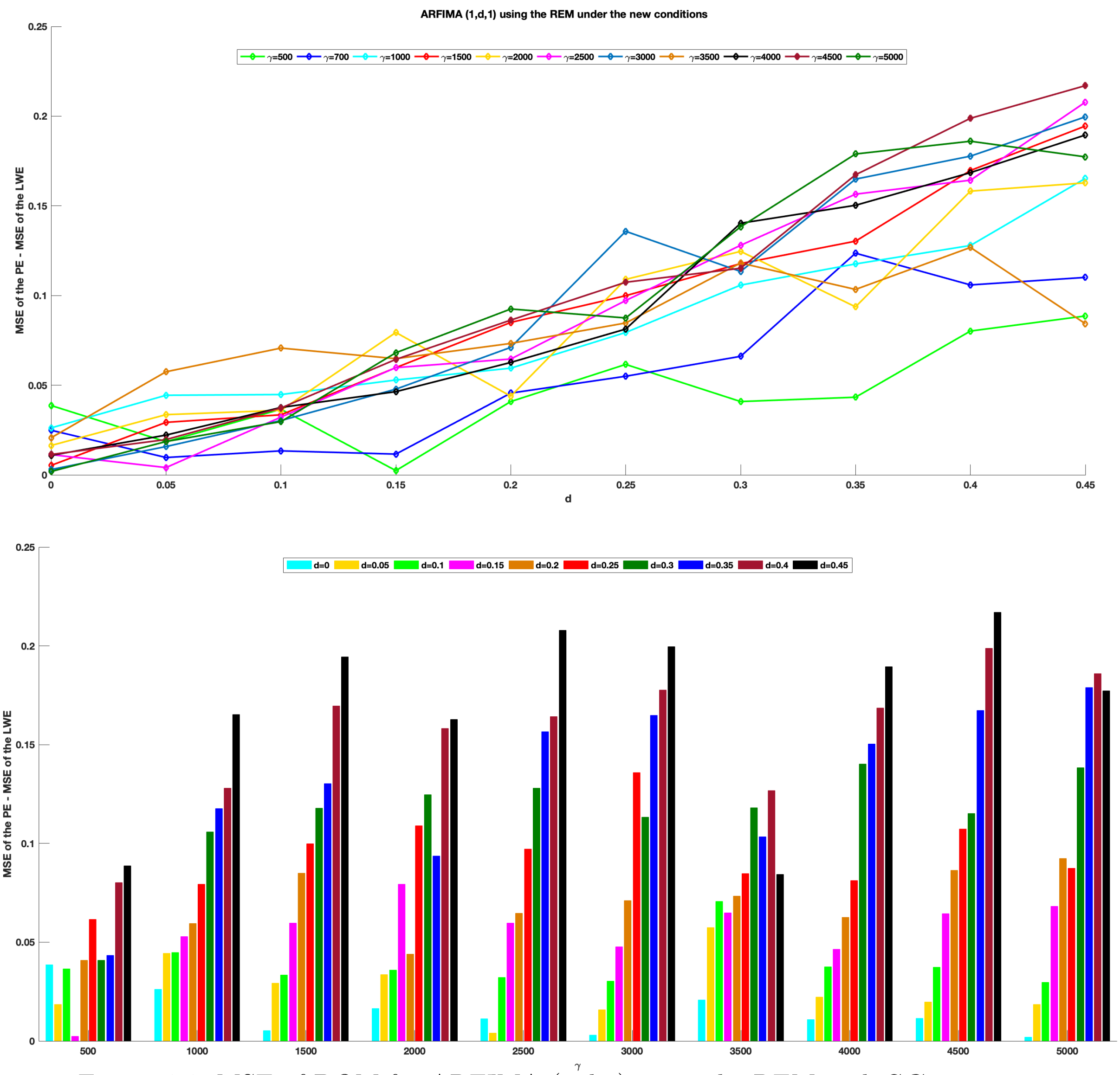

Figure 4.4: MSE of DOM for ARFIMA (1,d,1) using the REM with **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA$(1, d, 1)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.3

Table 4.4: MSE of DOM for ARFIMA (2,d,2) using the REM with **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(2, d, 2)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.0650	0.0827	0.1160	0.1758	0.1895	0.2567	0.2837	0.3652	0.3913	0.4535
	LWE	0.0482	0.0775	0.1119	0.1611	0.2021	0.2490	0.2772	0.3516	0.3890	0.4449
700	PE	0.0807	0.0839	0.0697	0.1083	0.1390	0.1674	0.1858	0.2130	0.2220	0.2658
	LWE	0.0509	0.0591	0.0551	0.0829	0.0975	0.1068	0.1407	0.1270	0.1303	0.1629
1000	PE	0.0412	0.0677	0.1014	0.1513	0.1965	0.2415	0.2920	0.3363	0.3797	0.4297
	LWE	0.0178	0.0537	0.0959	0.1464	0.1943	0.2345	0.2846	0.3277	0.3763	0.4212
1500	PE	0.0405	0.0559	0.1072	0.1437	0.1895	0.2333	0.2836	0.3225	0.3701	0.4226
	LWE	0.0170	0.0505	0.1009	0.1395	0.1884	0.2298	0.2764	0.3166	0.3657	0.4104
2000	PE	0.0313	0.0527	0.0897	0.1414	0.1818	0.2293	0.2807	0.3237	0.3696	0.4172
	LWE	0.0180	0.0497	0.0893	0.1378	0.1789	0.2228	0.2711	0.3157	0.3566	0.4009
2500	PE	0.0274	0.0536	0.0906	0.1420	0.1788	0.2275	0.2784	0.3244	0.3680	0.4175
	LWE	0.0206	0.0482	0.0896	0.1375	0.1764	0.2191	0.2637	0.3070	0.3512	0.3938
3000	PE	0.0266	0.0542	0.0922	0.1361	0.1822	0.2271	0.2719	0.3160	0.3635	0.4065
	LWE	0.0170	0.0520	0.0899	0.1288	0.1753	0.2149	0.2590	0.3003	0.3444	0.3845
3500	PE	0.0235	0.0438	0.0897	0.1398	0.1833	0.2213	0.2696	0.3170	0.3642	0.4070
	LWE	0.0182	0.0507	0.0879	0.1290	0.1733	0.2112	0.2565	0.2960	0.3419	0.3840
4000	PE	0.0233	0.0503	0.0878	0.1382	0.1851	0.2290	0.2678	0.3222	0.3630	0.4023
	LWE	0.0192	0.0466	0.0876	0.1340	0.1731	0.2107	0.2549	0.2969	0.3364	0.3775
4500	PE	0.0269	0.0484	0.0867	0.1362	0.1760	0.2203	0.2657	0.3118	0.3495	0.3969
	LWE	0.0255	0.0470	0.0829	0.1263	0.1637	0.2050	0.2466	0.2894	0.3272	0.3690
5000	PE	0.0424	0.0666	0.0807	0.1043	0.1333	0.1768	0.1661	0.2007	0.2197	0.2693
	LWE	0.0555	0.0784	0.0723	0.0574	0.0673	0.0792	0.0592	0.0579	0.0639	0.0698

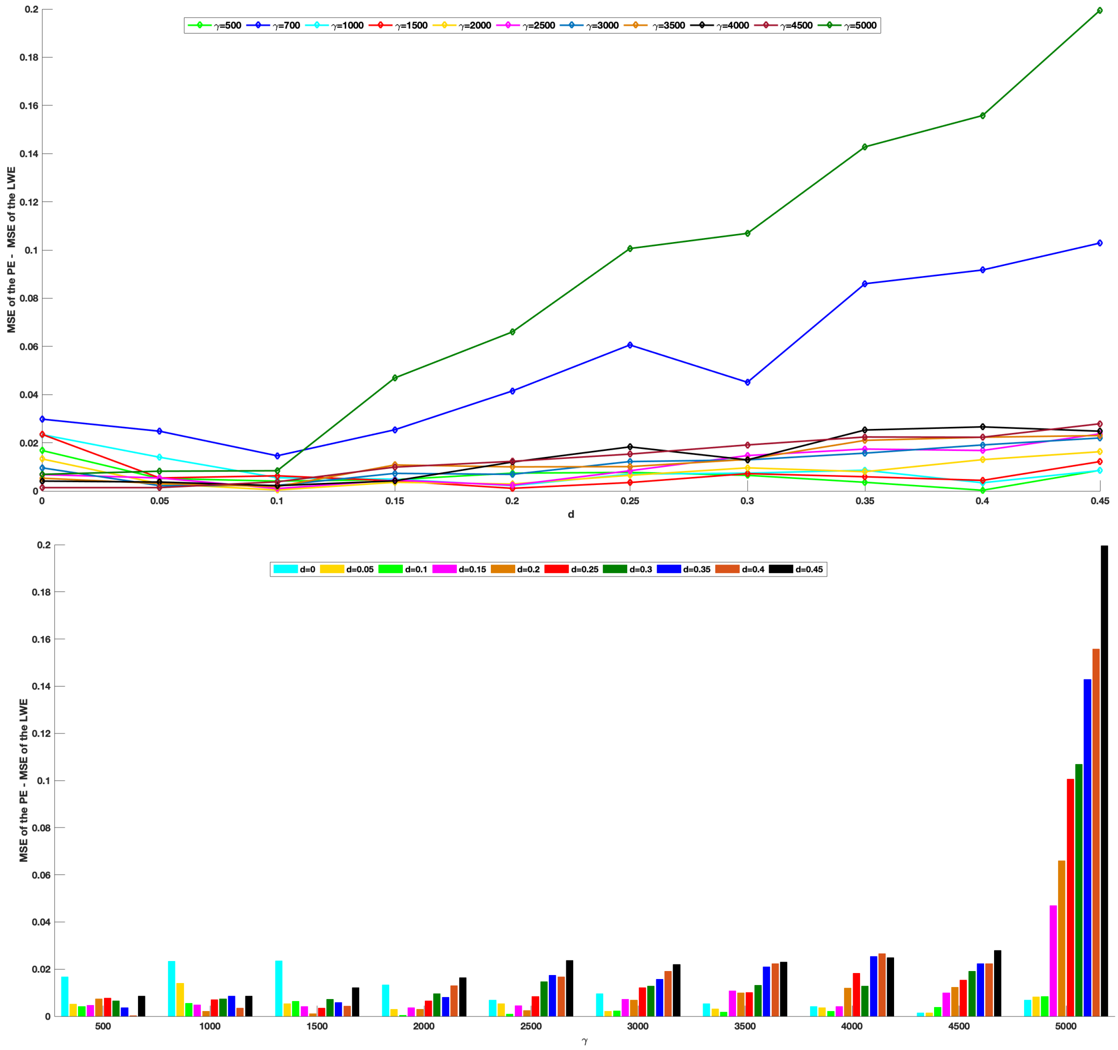

Figure 4.5: MSE of DOM for ARFIMA (2,d,2) using the REM with **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA$(2, d, 2)$ using the REM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.4

Table 4.5: MSE of DOM for ARFIMA (0,d,0) using the WEM with **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(0, d, 0)$ using the WEM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.0308	0.0437	0.0476	0.0259	0.0552	0.0432	0.0994	0.1135	0.1432	0.0377
	LWE	0.0155	0.0424	0.0921	0.1383	0.1883	0.2358	0.2822	0.3240	0.3655	0.4195
700	PE	0.0590	0.0809	0.1127	0.1165	0.1114	0.1357	0.1979	0.2315	0.2309	0.2585
	LWE	0.0480	0.0549	0.0536	0.0634	0.0640	0.0863	0.1077	0.1030	0.1115	0.1474
1000	PE	0.0814	0.0218	0.1123	0.0111	0.0557	0.0713	0.0205	0.0056	0.0494	0.0727
	LWE	0.0079	0.0425	0.0915	0.1348	0.1823	0.2302	0.2766	0.3238	0.3697	0.4181
1500	PE	0.0078	0.0993	0.0243	0.0338	0.0294	0.1055	0.0621	0.0266	0.0833	0.0267
	LWE	0.0001	0.0260	0.0178	0.0571	0.0245	0.0701	0.0464	0.0554	0.0582	0.0511
2000	PE	0.0468	0.0734	0.0019	0.0186	0.0228	0.0049	0.0083	0.0668	0.0081	0.0681
	LWE	0.0269	0.0339	0.0163	0.0280	0.0379	0.0352	0.0350	0.0477	0.0344	0.0511
2500	PE	0.0119	0.0474	0.0473	0.0384	0.0191	0.0575	0.0288	0.0505	0.0173	0.0253
	LWE	0.0023	0.0240	0.0216	0.0137	0.0179	0.0320	0.0321	0.0395	0.0294	0.0398
3000	PE	0.0334	0.0413	0.0125	0.0191	0.0118	0.0338	0.0182	0.0047	0.0247	0.0423
	LWE	0.0099	0.0182	0.0209	0.0149	0.0309	0.0305	0.0143	0.0353	0.0279	0.0348
3500	PE	0.0252	0.0359	0.0279	0.0102	0.0032	0.0013	0.0924	0.0526	0.0374	0.0305
	LWE	0.0011	0.0202	0.0276	0.0154	0.0175	0.0166	0.0381	0.0273	0.0458	0.0277
4000	PE	0.0157	0.0296	0.0387	0.0445	0.0193	0.0480	0.0663	0.0155	0.0538	0.1238
	LWE	0.0001	0.0184	0.0281	0.0177	0.0225	0.0260	0.0327	0.0132	0.0503	0.0307
4500	PE	0.0146	0.0728	0.0372	0.0506	0.0196	0.0131	0.0482	0.0380	0.0126	0.0051
	LWE	0.0071	0.0415	0.0270	0.0313	0.0185	0.0133	0.0325	0.0229	0.0154	0.0267
5000	PE	0.0306	0.0397	0.0144	0.0121	0.0228	0.0219	0.0621	0.0134	0.0681	0.0692
	LWE	0.0022	0.0213	0.0110	0.0091	0.0163	0.0112	0.0340	0.0210	0.0428	0.0362

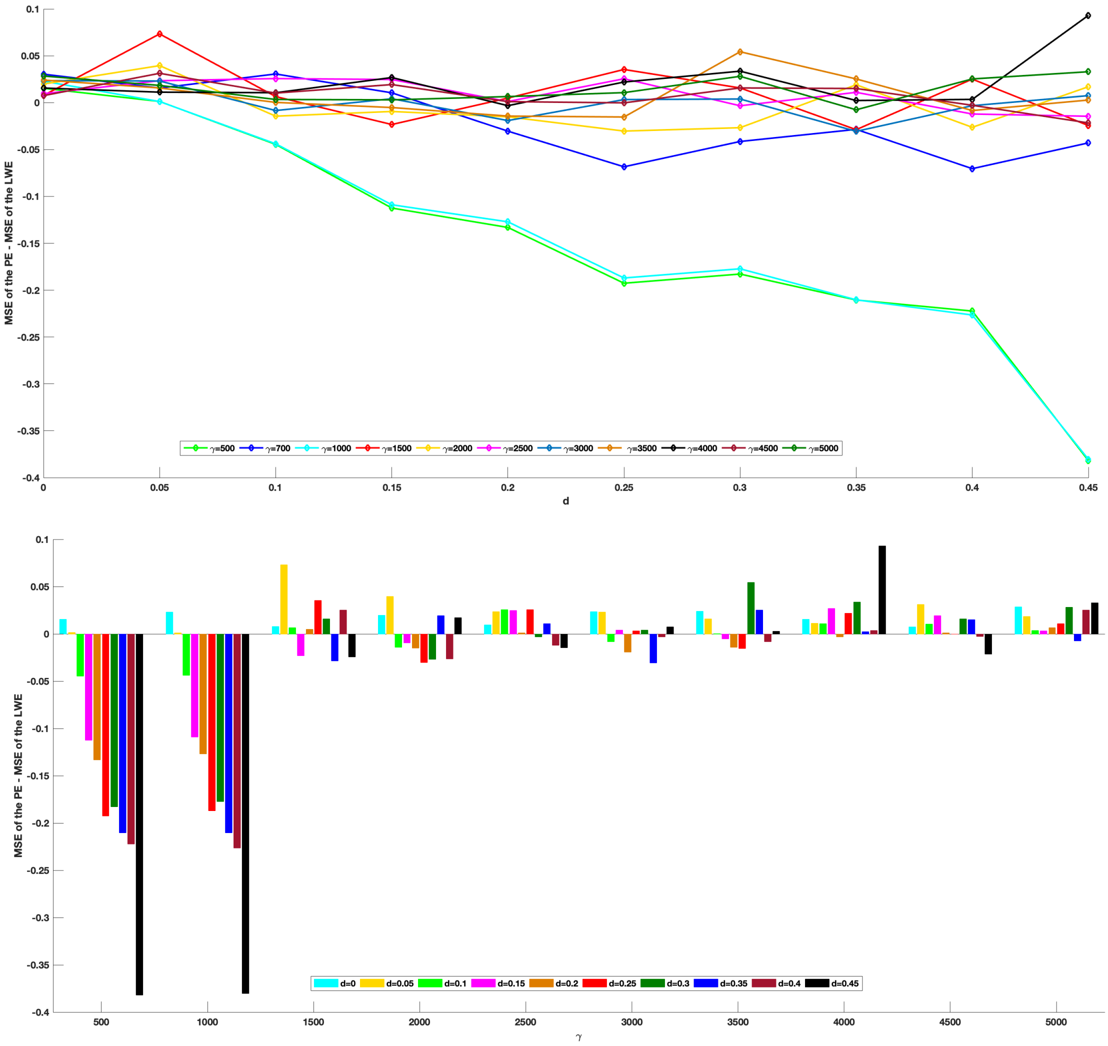

Figure 4.6: MSE of DOM for ARFIMA (0,d,0) using the WEM with **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA$(0, d, 0)$ using the WEM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.5

Table 4.6: MSE of DOM for ARFIMA (0,d,0) using the REM without **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(0, d, 0)$ using the REM without the new conditions by taking $n_\gamma = t_\gamma = \gamma^{1/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.2637	0.1675	0.1015	0.1496	0.2030	0.1994	0.1860	0.2402	0.2118	0.2876
	LWE	0.0154	0.0429	0.0586	0.1076	0.1272	0.1410	0.1675	0.2134	0.2139	0.2422
700	PE	0.1529	0.1792	0.1933	0.0967	0.2328	0.1572	0.2089	0.2748	0.2202	0.2679
	LWE	0.0038	0.0397	0.0568	0.1011	0.1231	0.1727	0.2116	0.2719	0.2732	0.3046
1000	PE	0.0577	0.1358	0.1034	0.1663	0.1644	0.2300	0.1299	0.1841	0.2697	0.2239
	LWE	0.0131	0.0435	0.0782	0.1043	0.1432	0.2119	0.1817	0.2200	0.2633	0.2890
1500	PE	0.1431	0.1561	0.0835	0.0721	0.1428	0.1016	0.1763	0.2694	0.3405	0.2745
	LWE	0.0150	0.0337	0.0798	0.1058	0.1582	0.1685	0.2137	0.2543	0.3093	0.3362
2000	PE	0.1831	0.1155	0.1262	0.1436	0.1535	0.1512	0.1943	0.2335	0.1980	0.3739
	LWE	0.0123	0.0544	0.0873	0.1126	0.1798	0.1971	0.2334	0.2787	0.3168	0.3651
2500	PE	0.2040	0.0878	0.1379	0.2160	0.1244	0.1830	0.2236	0.1325	0.2197	0.2463
	LWE	0.0025	0.0378	0.0782	0.1260	0.1614	0.1956	0.2403	0.2852	0.3210	0.3736
3000	PE	0.1000	0.1125	0.0765	0.1006	0.2055	0.1844	0.2133	0.2515	0.2415	0.2877
	LWE	0.0139	0.0439	0.0760	0.1172	0.1672	0.2089	0.2429	0.2831	0.3389	0.3822
3500	PE	0.0943	0.0833	0.1713	0.0880	0.1177	0.2045	0.3022	0.2971	0.2206	0.2068
	LWE	0.0058	0.0452	0.0902	0.1294	0.1671	0.2135	0.2602	0.3038	0.3395	0.3734
4000	PE	0.0962	0.1100	0.1659	0.1455	0.2405	0.1544	0.2423	0.2781	0.2952	0.2029
	LWE	0.0043	0.0375	0.0968	0.1355	0.1772	0.2137	0.2588	0.3025	0.3498	0.3764
4500	PE	0.0010	0.0476	0.0863	0.1327	0.1729	0.2271	0.2624	0.2987	0.3491	0.3938
	LWE	0.0165	0.0416	0.0897	0.1360	0.1732	0.2103	0.2588	0.3112	0.3455	0.3960
5000	PE	0.0910	0.1761	0.0832	0.1840	0.1625	0.2586	0.1391	0.1612	0.2091	0.2295
	LWE	0.0010	0.0476	0.0863	0.1327	0.1729	0.2271	0.2624	0.2987	0.3491	0.3938

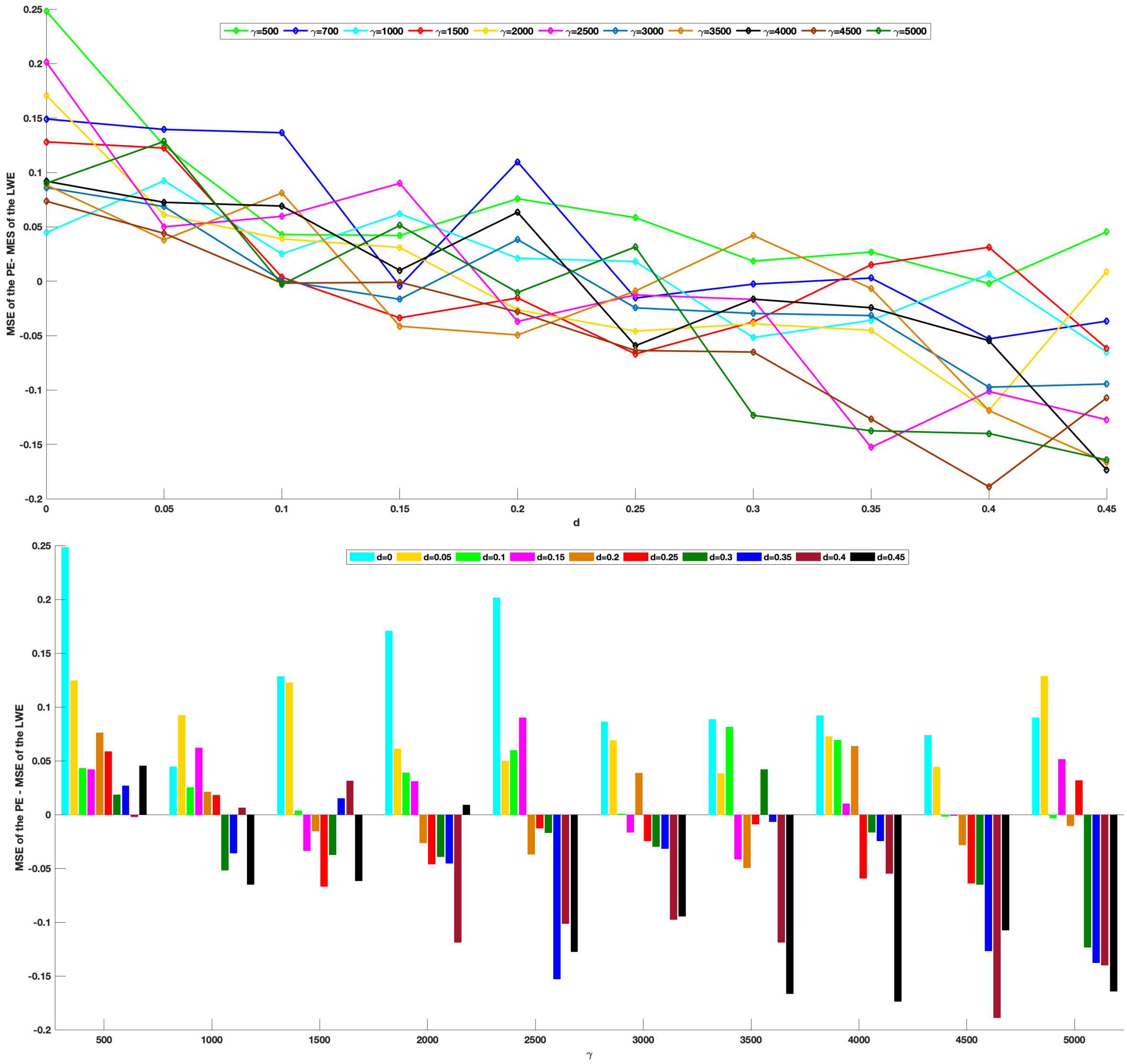

Figure 4.7: MSE of DOM for ARFIMA (0,d,0) using the REM without **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA$(0, d, 0)$ using the WEM with the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.6

Table 4.7: MSE of DOM for ARFIMA (1,d,1) using the WEM with **SC3**

MSE values of the PE-DOMs and LWE-DOMs for ARFIMA$(1, d, 1)$ using the WEM under the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d

γ	estimators	d values									
		0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
500	PE	0.0590	0.0551	0.0760	0.1061	0.1379	0.1328	0.2139	0.2885	0.2623	0.2604
	LWE	0.0436	0.0332	0.0454	0.0605	0.1006	0.0932	0.1524	0.1636	0.1897	0.1921
700	PE	0.0590	0.0809	0.1127	0.1165	0.1114	0.1357	0.1979	0.2315	0.2309	0.2585
	LWE	0.0480	0.0549	0.0536	0.0634	0.0640	0.0863	0.1077	0.1030	0.1115	0.1474
1000	PE	0.0328	0.0351	0.0574	0.0921	0.1073	0.1462	0.1789	0.2106	0.2020	0.2828
	LWE	0.0232	0.0357	0.0218	0.0380	0.0436	0.0685	0.0591	0.0795	0.0779	0.0926
1500	PE	0.0321	0.0498	0.0523	0.1013	0.1217	0.1345	0.1952	0.1894	0.2114	0.2594
	LWE	0.0226	0.0219	0.0391	0.0494	0.0462	0.0328	0.0624	0.0479	0.0566	0.0470
2000	PE	0.0215	0.0178	0.0577	0.0835	0.1213	0.1324	0.1486	0.1940	0.2245	0.2474
	LWE	0.0079	0.0057	0.0326	0.0205	0.0218	0.0415	0.0232	0.0547	0.0427	0.0692
2500	PE	0.0284	0.0257	0.0678	0.0913	0.1142	0.1635	0.1569	0.1836	0.2141	0.2186
	LWE	0.0385	0.0116	0.0104	0.0157	0.0275	0.0255	0.0301	0.0173	0.0437	0.0398
3000	PE	0.0215	0.0382	0.0417	0.0753	0.1125	0.1165	0.1639	0.2101	0.2337	0.2427
	LWE	0.0123	0.0259	0.0171	0.0228	0.0412	0.0189	0.0295	0.0229	0.0434	0.0327
3500	PE	0.0386	0.0363	0.0745	0.0924	0.1198	0.1511	0.1712	0.1948	0.2067	0.2500
	LWE	0.0153	0.0100	0.0171	0.0422	0.0331	0.0318	0.0282	0.0468	0.0216	0.0296
4000	PE	0.0240	0.0392	0.0541	0.0724	0.0836	0.1169	0.1564	0.1737	0.1973	0.2273
	LWE	0.0131	0.0170	0.0165	0.0260	0.0209	0.0356	0.0161	0.0234	0.0288	0.0378
4500	PE	0.0215	0.0275	0.0508	0.0845	0.1051	0.1249	0.1374	0.1895	0.2279	0.2418
	LWE	0.0101	0.0078	0.0135	0.0201	0.0188	0.0175	0.0222	0.0222	0.0291	0.0248
5000	PE	0.0226	0.0360	0.0683	0.0882	0.1123	0.1274	0.1609	0.1986	0.2008	0.2216
	LWE	0.0207	0.0174	0.0386	0.0201	0.0198	0.0399	0.0226	0.0197	0.0148	0.0443

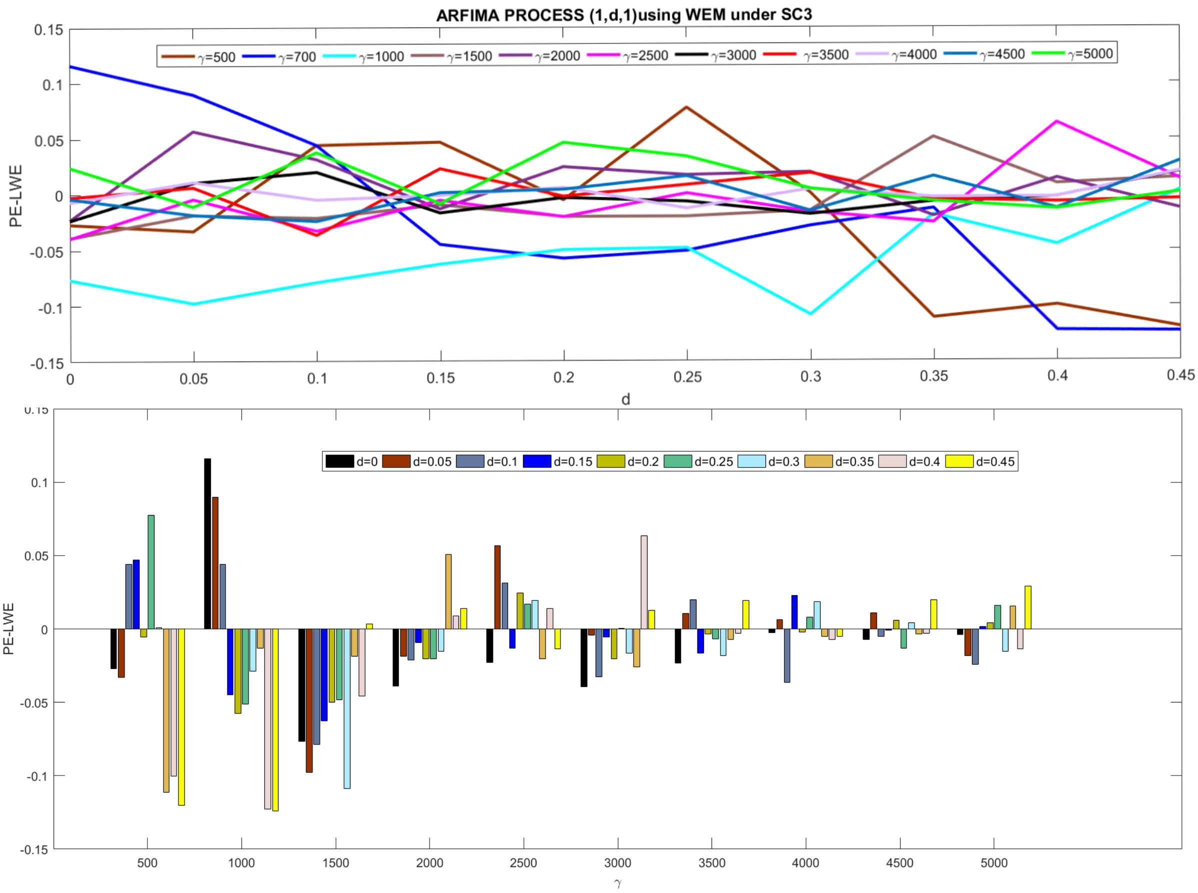

Figure 4.8: MSE of DOM for ARFIMA (1,d,1) using the WEM with **SC3**

The differences among the MSE values of the PE-DOMs and the LWE-DOMs for ARFIMA$(1, d, 1)$ using the WEM with the new conditions by taking $n_\gamma = t_\gamma = \gamma^{2/3}$ and different values of sample size γ and DOM d in Table 4.7

4.3 Interpretation of Results

Based on the **Preparation step 1** and **Preparation step 2**, the above-mentioned three scenarios are investigated as follows:

Scenario 1: The accuracy of the LWT for the REM under the new conditions:

The theoretical justifications prove that the LWE-DOM is better than the PE-DOM for the REM under the new sufficient conditions. This scenario investigates the numerical justifications for the given theoretical justifications. Tables 6.18, 4.4 and 4.5 give the MSE values of the LWE-DOMs by using the Hamming LWF as given in the **Preparation step 2** and the PE-DOMs based on the REM for ARFIMA$(0, d, 0)$, $0 \leq d < \frac{1}{2}$, ARFIMA$(1, d, 1)$, $0 \leq d < \frac{1}{2}$ and ARFIMA$(2, d, 2)$, $0 \leq d < \frac{1}{2}$, respectively. The parameters are selected to satisfy the sufficient conditions **SC1-SC4** as given in the **Preparation step 1** by taking $t_\gamma = n_\gamma = \gamma^{\frac{2}{3}}$ using the following different sample sizes $500 \leq \gamma \leq 5000$. The differences among the MSE values of the PE-DOMs and the MSE values of the LWE-DOMs (i.e., MSEs of the PEs - MSEs of the LWEs) in Table 2, 3 and 4 are given in Figures 6.18, 4.4 and 4.5, respectively.

• From Tables 6.18-4.5 and Figures 6.18-4.5, we can show that for short memory ARFIMA$(p, 0, q)$ and long memory ARFIMA$(p, 0 < d < \frac{1}{2}, q)$ the MSE values of the LWE-DOMs under the new conditions are less than the MSE values of the PE-DOMs for all the values of the DOMs $0 \leq d < \frac{1}{2}$. That is, the LWT is better than the PT for the REM under the new sufficient conditions.

Scenario 2: The accuracy of the LWT for the WEM under the new conditions:

Even though the theoretical justification of the efficiency of the LWT is proved based on the REM only, this scenario investigates the performance of the LWT using the new conditions for other methods, such as the WEM. The WEM is a semi-parametric method which estimates the DOM on a specific asymptotic form of the SDF in (2.11) and it depends on the efficiency of the estimator of the SDF (cf. Künsch[34]) as follows

$$\widehat{d} := \underset{-\frac{1}{2}<d<\frac{1}{2}}{\operatorname{argmin}} \left[\ln \left(\frac{1}{n_\gamma} \sum_{j=1}^{n_\gamma} \eta_j^{2d} \widehat{\mathcal{S}}(\eta_j) \right) - 2d \frac{1}{n_\gamma} \sum_{j=1}^{n_\gamma} \ln(\eta_j) \right].$$

Table 4.6 gives the MSE values of the LWE-DOMs by using the Hamming LWF as given in the **Preparation step 2** and the PE-DOMs based on the WEM for

ARFIMA$(0, d, 0)$, $0 \leq d < \frac{1}{2}$. The parameters are selected to satisfy the sufficient conditions **SC1-SC4** as given in the **Preparation step 1** by taking $t_\gamma = n_\gamma = \gamma^{\frac{2}{3}}$ using the following different sample sizes $500 \leq \gamma \leq 5000$. The differences among the MSE values of the PE-DOMs and the MSE values of the LWE-DOMs (i.e., MSEs of the PEs - MSEs of the LWEs) in Table4.6 are given in Figure 4.6.

• From Table 4.6 and Figure 4.6, we can show that for short memory ARFIMA$(p, 0, q)$ the MSE values of the LWE-DOMs under the new conditions are less than the MSE values of the PE-DOMs for all the cases based on the WEM, i.e., the LWT is better than the PT for the WEM under the sufficient conditions **SC1-SC4** for short memory. However, for long memory ARFIMA$(p, 0 < d < \frac{1}{2}, q)$ the MSE values of the LWE-DOMs under the new conditions are less than the MSE values of the PE-DOMs for many cases, i.e., the LWT is better than the PT based on the REM under the sufficient conditions **SC1-SC4** for many cases of DOMs $0 < d < \frac{1}{2}$ and sample sizes γ, especially for large DOMs and sample sizes.

Scenario 3: The accuracy of the LWT for the REM without the new conditions: This scenario investigates the performance of the LWT without the new sufficient conditions. Table 4.7 gives the MSE values of the LWE-DOMs by using the Hamming LWF as given in the **Preparation step 2** and the PE-DOMs based on the REM for ARFIMA$(0, d, 0)$, $0 \leq d < \frac{1}{2}$. The parameters are selected to satisfy the sufficient conditions **SC1**, **SC2** and **SC4** and not satisfy the sufficient condition **SC3** as given in the **Preparation step 1** by taking $t_\gamma = n_\gamma = \gamma^{\frac{1}{3}}$ using the following different sample sizes $500 \leq \gamma \leq 5000$. The differences among the MSE values of the PE-DOMs and the MSE values of the LWE-DOMs (i.e., MSEs of the PEs - MSEs of the LWEs) in Table 4.7 are given in Figure 4.7.

From Table 4.7 and Figure 4.7, we can show that for long memory ARFIMA$(p, 0 < d < \frac{1}{2}, q)$ the MSE values of the LWE-DOMs without the new condition **SC3** are larger than the MSE values of the PE-DOMs for many cases. That is, the PT is better than the LWT based on the REM if any of the new sufficient conditions **SC1-SC4** are not satisfied for many cases of DOMs $0 < d < \frac{1}{2}$ and sample sizes γ, especially for large DOMs and sample sizes.

On the other hand, Figure 4.1 gives a comparison study between the shape of the variation of the LWE-SDF under the new conditions and the shape of the variation of the PE-SDF for different sample sizes $\gamma \in \{700, 1000, 3000, 5000\}$ and DOMs $d \in \{0.05, 0.1, 0.4, 0.45\}$. From Figure 4.1, we can show that:

- (i) the PE-SDF has many so sharply peaks and unsmoothed wild fluctuation form;
- (ii) the LWE-SDF under the new conditions is more smoothed than the PE-SDF and it converges to zero for large sample sizes and DOMs.

Therefore, the LWE-SDF under the new conditions is a good representative of the SDF according to its characteristics compared the PE-SDF.

Remark 7. *From the above discussions and numerical results, it is obvious that the new sufficient conditions that are given in this work are needed to improve the efficiency of the LWT for estimating the SDFs and DOMs of SMDSGPs and LMDSGPs.*

Chapter 5

Conclusion and Future Work

5.1 Conclusion

Estimating the SDF and the DOM of a Gaussian process is needed for many real-world problems. This thesis gives sufficient conditions for improving the efficiency of the lag window technique (LWT) that is one of the widely techniques for estimating the SDF and the DOM. The asymptotic behaviors of the resulting estimators under the new conditions are investigated. A comparison study between the LWT under the new conditions, the LWT without the new conditions and periodogram technique (PT) that is the classical widely used technique, is given theoretically and computationally. The main results show that the LWT under the new conditions is better than the PT and the LWT without the new conditions. Therefore, the new conditions are needed to improve the performance of the LWT for many models.

5.2 Future Work

After reading this work, some interesting ideas for further study can arise. We are working on these ideas and some interesting results are obtained. However, we cannot give any conclusion at this stage and the results will be given in our future works.

- The results in this research work only concentrate on a specific type of stochastic process, the stationary Gaussian process. Then, future research will concentrate on generalizing these results for the case of a general stochastic process.

- Although good estimators for the SDF and the DOM are obtained in this work, the problem is still difficult due to the limited use of these estimators (due to the sufficient conditions **SC1-SC4**). The future work will try to relax the sufficient conditions **SC1-SC4** to get good estimators in the general case.

- The theoretical justifications with sufficient conditions for obtaining good estimators are investigated based on the REM only. Then, future research will concentrate on generalizing the theoretical justifications and conditions for other methods, such as the WEM.

- Machine learning and its techniques, such as neural networks, are widely used recently for detecting outliers and estimating and forecasting parameters in time series and stochastic processes (cf. [58, 59]). The future work will try to provide new efficient estimation techniques using the power of the neural networks.

Long memory is still a very active field of time series research. There are many further interesting and important issues, that it is far from complete.

Chapter 6

Appendix

Proof of Theorem 3. From (2.4) and (3.3) evaluate at Fourier frequencies, we have

$$Bias(\widehat{\mathcal{S}}_w(\eta_j)) = \alpha_1 + \alpha_2,$$

where

- $\alpha_1 = E(\widehat{\mathcal{S}}_w(\eta_j)) - \mathcal{S}_1(\eta_j)$ the bias due the periodogram itself,
- $\alpha_2 = \mathcal{S}_1(\eta_j) - \mathcal{S}(\eta_j)$ the bias due to the smoothing,
- $\mathcal{S}_1(\eta_j) = \frac{1}{2\pi}\sum_{|h|\leq t_\gamma} \mathcal{A}(h)e^{-i\eta_j h}\mathcal{W}\left(\frac{h}{t_\gamma}\right).$

Therefore, we have the following cases based on the DOM values.

Case 1: For NMDSGP, $-\frac{1}{2} < d < 0$. From (2.1) and autocovariance $\mathcal{A}(h) = E(X_{j+h}X_j)$, we have $E(\widehat{\mathcal{A}}(h)) = \left(1 - \frac{|h|}{\gamma}\right)\mathcal{A}(h)$ and thus,

$$\begin{aligned}
\alpha_1 &= \frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\Big(E(\widehat{\mathcal{A}}(h)) - \mathcal{A}(h)\Big)e^{-i\eta_j h}\mathcal{W}\left(\frac{h}{t_\gamma}\right) \\
&= -\frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\frac{|h|}{\gamma}\mathcal{A}(h)e^{-i\eta_j h}\mathcal{W}\left(\frac{h}{t_\gamma}\right) \\
&= -\frac{1}{\pi\gamma}\sum_{h=1}^{t_\gamma} h\mathcal{A}(h)\cos(\eta_j h)\mathcal{W}\left(\frac{h}{t_\gamma}\right). \qquad (6.1)
\end{aligned}$$

For $t_\gamma \to \infty$, we have

$$\mathcal{A}(h) \simeq \beta|h|^{2d-1} \text{ and } \mathcal{W}\left(\frac{h}{t_\gamma}\right) \to 1. \qquad (6.2)$$

From Zygmund [64], the sum of two consecutive products, with one decreasing (d is negative) and the other trigonometric, is given as follows

$$\sum_{h=1}^{\infty} h^{2d}\cos(\eta_j h) \simeq -\Gamma(1+2d)\sin(\pi d)\eta_j^{-(2d+1)}. \tag{6.3}$$

Combining (6.1)-(6.3) and $\eta_j = \frac{2\pi j}{\gamma}$ for $j \neq 0$

$$\alpha_1 \simeq O\left(\frac{1}{\gamma}\eta_j^{-(2d+1)}\right) \simeq O\left(\frac{\gamma\eta_j^{-2d}}{j\gamma}\right) \simeq O\left(\frac{\eta_j^{-2d}}{j}\right). \tag{6.4}$$

The form α_2 can be written as follows

$$\begin{aligned}
\alpha_2 &= \frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\mathcal{W}\left(\frac{h}{t_\gamma}\right) - \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\mathcal{A}(h)e^{-i\eta_j h} \\
&= \frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\mathcal{W}\left(\frac{h}{t_\gamma}\right) - \frac{1}{2\pi}\left(\sum_{|h|>t_\gamma}\mathcal{A}(h)e^{-i\eta_j h} + \sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\right) \\
&= \frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\left(\mathcal{W}\left(\frac{h}{t_\gamma}\right)-1\right) - \frac{1}{2\pi}\sum_{|h|>t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}. \qquad (6.5)
\end{aligned}$$

From (4.7) and Tikhonov[57], we get

$$\sum_{|h|>t_\gamma}\mathcal{A}(h)e^{-i\eta_j h} \leq \sum_{|h|>t_\gamma}\mathcal{A}(h) \simeq O\left(t_\gamma^{2d}\right). \tag{6.6}$$

From $\mathcal{W}(0) \overset{\mathbf{C1}}{=} 1$ and (2.14), we have

$$\begin{aligned}
\frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\left(1-\mathcal{W}\left(\frac{h}{t_\gamma}\right)\right) &\leq \frac{1}{\pi}\sum_{1\leq h\leq t_\gamma}\left(k\frac{h}{t_\gamma}\right)\mathcal{A}(h)\cos(\eta_j h) \\
&\simeq \frac{k}{\pi t_\gamma}\mathcal{V}_{\mathcal{A}}\left(\frac{1}{t_\gamma}\right)\sum_{1\leq h\leq t_\gamma}h^{2d}\cos(\eta_j h). \qquad (6.7)
\end{aligned}$$

It is obvious that

$$\left|\sum_{1\leq h\leq t_\gamma}h^{2d}\cos(\eta_j h)\right| \leq \sum_{1\leq h\leq t_\gamma}h^{2d} \simeq \frac{t_\gamma^{2d+1}}{2d+1}. \tag{6.8}$$

Combining (6.7) and (6.8), we obtain that

$$\frac{1}{2\pi}\sum_{|h|\leq t_\gamma}\mathcal{A}(h)e^{-i\eta_j h}\left(1-\mathcal{W}\left(\frac{h}{t_\gamma}\right)\right) \leq \frac{kt_\gamma^{2d}}{2\pi(d+1)}\mathcal{V}_{\mathcal{A}}\left(\frac{1}{t_\gamma}\right) \overset{(2.12)}{\simeq} O\left(t_\gamma^{2d}\right). \tag{6.9}$$

Combining (6.5), (6.6) and (6.9), we get

$$\alpha_2 \simeq O\left(t_\gamma^{2d}\right). \tag{6.10}$$

From (6.4) and (6.10), we have

$$Bias(\widehat{\mathcal{S}}_w(\eta_j)) \simeq O\left(t_\gamma^{2d}\right) + O\left(\frac{\eta_j^{-2d}}{j}\right). \tag{6.11}$$

Case 2: From LMDSGP, $0 < d < \frac{1}{2}$. Using the same arguments for $\widehat{\mathcal{S}}_w(\eta)$ to obtain expression (3.5) and by its Riemann approximating sum (3.12), $\mathcal{S}_1(\eta)$ can be rewritten as follows

$$\mathcal{S}_1(\eta) = \int_{-\pi}^{\pi} \mathcal{S}(\theta)\mathcal{U}_\gamma(\eta-\theta)\,d\theta \simeq \frac{2\pi}{\gamma}\sum_{k=-[\frac{\gamma-1}{2}]}^{[\frac{\gamma}{2}]} \mathcal{S}(\eta_k)\mathcal{U}_{t_\gamma}(\eta-\eta_k), \tag{6.12}$$

where [.] means the integer part. From (3.12) and (6.12), we get

$$\begin{aligned}
\alpha_1 &= \frac{2\pi}{\gamma}\sum_{k=-[\frac{\gamma-1}{2}]}^{[\frac{\gamma}{2}]} \Big(E\left(\widehat{\mathcal{S}}_p(\eta_k)\right) - \mathcal{S}(\eta_k)\Big)\mathcal{U}_{t_\gamma}(\eta_j-\eta_k) \\
&= \frac{2\pi}{\gamma}\sum_{k=-[\frac{\gamma-1}{2}]}^{[\frac{\gamma}{2}]} \Big[E\left(\widehat{\mathcal{S}}_p(\eta_j-\eta_k)\right) - \mathcal{S}(\eta_j-\eta_k)\Big]\mathcal{U}_{t_\gamma}(\eta_k) \\
&\overset{\mathbf{P1}}{\leq} \max_{|k|\leq[\gamma/2]} \Big|E\left(\widehat{\mathcal{S}}_p(\eta_j-\eta_k)\right) - \mathcal{S}(\eta_j-\eta_k)\Big| \\
&\overset{\eta_j=2\pi j/\gamma,\ \eta_k=2\pi k/\gamma}{=} \max_{|k|\leq[\gamma/2]} \Big|E\left(\widehat{\mathcal{S}}_p(\eta_{j-k})\right) - \mathcal{S}(\eta_{j-k})\Big| \\
&= \max_{j-[\gamma/2]\leq l\leq j+[\gamma/2]} \Big|E\left(\widehat{\mathcal{S}}_p(\eta_j)\right) - \mathcal{S}(\eta_l)\Big|.
\end{aligned} \tag{6.13}$$

From (6.13) and the arguments developed in Beran[7] (cf. also proof of Theorem 2 in Robinson [49]), we have

$$\alpha_1 \simeq O\left(\frac{\ln j}{j\eta_j^{2d}}\right). \tag{6.14}$$

From (6.12) and the fact that $\int_{-\pi}^{\pi} \mathcal{U}_{t_\gamma}(\theta)d\theta = 1$, the form α_2 can be written as follows

$$\alpha_2 = \int_{-\pi}^{\pi} \mathcal{S}(\eta_j-\theta)\mathcal{U}_{t_\gamma}(\theta)d\theta - \mathcal{S}(\eta_j)\int_{-\pi}^{\pi} \mathcal{U}_{t_\gamma}(\theta)d\theta = \int_{-\pi}^{\pi}\Big(\mathcal{S}(\eta_j-\theta)-\mathcal{S}(\eta_j)\Big)\mathcal{U}_{t_\gamma}(\theta)d\theta. \tag{6.15}$$

The second-order Taylor expansion of $\mathcal{S}(\eta)$ at the neighborhood of η_j ($\epsilon \overset{\mathbf{(SC4)}}{<} \eta_j$) is

$$\mathcal{S}(\eta)\simeq\mathcal{S}(\eta_j) + (\eta-\eta_j)\mathcal{S}'(\eta_j) + \frac{(\eta-\eta_j)^2}{2}\mathcal{S}''(\eta_j) + o((\eta-\eta_j)^2).$$

By taking $\eta = \eta_j - \theta$ and intervening the spectral window in the preceding formula with the properties P1 and P3, (6.15) can be written as follows

$$\alpha_2 \leq \int_{-\pi}^{\pi}\Big[-\theta\mathcal{S}'(\eta_j) + \frac{\theta^2}{2}\mathcal{S}''(\eta_j) + o(\theta^2)\Big]\mathcal{U}_{t_\gamma}(\theta)d\theta \simeq \frac{\mathcal{S}''(\eta_j)}{2}\int_{-\pi}^{\pi}\theta^2\mathcal{U}_{t_\gamma}(\theta)d\theta. \tag{6.16}$$

From (3.6), we get

$$\mathcal{U}_{t_\gamma}(\theta) = \frac{1}{2\pi}\sum_{|h|\le t_\gamma}\mathcal{W}\Big(\frac{h}{t_\gamma}\Big)e^{-i\theta h} \overset{(x=\frac{h}{t_\gamma})}{\simeq} \frac{t_\gamma}{2\pi}\int_{-1}^{1}\mathcal{W}(x)e^{-ixt_\gamma\theta}dx = t_\gamma\mathcal{U}(t_\gamma\theta). \quad (6.17)$$

Consequently,

$$\begin{aligned}
\mathcal{U}(t_\gamma\theta) = \frac{1}{2\pi}\int_{-1}^{1}\mathcal{W}(x)e^{-ix(t_\gamma\theta)}dx \quad &\Longrightarrow \quad \mathcal{U}(\theta) = \frac{1}{2\pi}\int_{-1}^{1}\mathcal{W}(x)e^{-ix\theta}dx \\
&\overset{\text{IFT}}{\Longleftrightarrow} \quad \mathcal{W}(x) = \int_{-\infty}^{\infty}\mathcal{U}(\theta)e^{ix\theta}d\theta \\
&\Longrightarrow \quad \mathcal{W}''(x) = -\int_{-\infty}^{\infty}\theta^2\mathcal{U}(\theta)e^{ix\theta}d\theta \\
&\Longrightarrow \quad \mathcal{W}''(0) = -\int_{-\infty}^{\infty}\theta^2\mathcal{U}(\theta)d\theta < \infty. \quad (6.18)
\end{aligned}$$

where IFT is the inverse Fourier transformation. Combining (6.16)-(6.18), we get

$$\begin{aligned}
\alpha_2 \le \frac{\mathcal{S}''(\eta_j)}{2}\int_{-\pi}^{\pi}\theta^2\mathcal{U}_{t_\gamma}(\theta)d\theta \quad &\overset{(6.17)}{=} \quad \frac{\mathcal{S}''(\eta_j)}{2}\int_{-\pi}^{\pi}\theta^2 t_\gamma\mathcal{U}(t_\gamma\theta)d\theta \\
&\overset{(6.18),\ u=t_\gamma\theta}{\simeq} \quad \frac{-\mathcal{S}''(\eta_j)}{2t_\gamma^2}\int_{-\infty}^{\infty}u^2\mathcal{U}(u)du. \quad (6.19)
\end{aligned}$$

From (6.18), (6.19) and (**SC4**), we get

$$\alpha_2 \simeq O\left(\frac{\eta_j^{-2d-2}}{t_\gamma^2}\right). \quad (6.20)$$

Combining (6.14) and (6.20), we have

$$Bias(\widehat{\mathcal{S}}_w(\eta_j)) \simeq O\left(\frac{\ln j}{j\eta_j^{2d}}\right) + O\left(\frac{1}{t_\gamma^2\eta_j^{2d+2}}\right). \quad (6.21)$$

From LMDSGP (6.21), α_1 is smaller then α_2 which is similar to SMDSGP (cf. Priestley [42] section 6), but for NMDSGP (6.11) it's the opposite. Then, we get

$$Bias\left(\widehat{\mathcal{S}}_w(\eta_j)\right) \overset{(6.11,6.21)}{\simeq} \begin{cases} O\left(\frac{1}{t_\gamma\eta_j^{d+1}}\right)^2, & \text{if} \quad 0<d<\frac{1}{2}, \text{ i.e., for LM-DSGP;} \\ O\left(\frac{1}{\sqrt{j}\eta_j^{d}}\right)^2, & \text{if} \quad -\frac{1}{2}<d<0, \text{ i.e., for NM-DSGP.} \end{cases}$$

For approximate formula (2.11) of SDF for LM and NM-DSGP, we have (3.13).■

Proof of Theorem 4. From the definition of variance, we have

$$var(\widehat{\mathcal{S}}_w(\eta)) = \left(\frac{1}{2\pi}\right)^2\sum_{h=-t_\gamma}^{t_\gamma}\sum_{g=-t_\gamma}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\mathcal{W}\left(\frac{g}{t_\gamma}\right)e^{-i\eta(h-g)}\,Cov\left(\widehat{\mathcal{A}}(h),\widehat{\mathcal{A}}(g)\right). \quad (6.22)$$

For any DSGP $\{X_t\}_{t\in\mathbb{Z}}$ by using the relation (5.3.26) with the arguments developed in Priestley[42] and for (2.13),

$$cov\left(\widehat{\mathcal{A}}(h),\widehat{\mathcal{A}}(g)\right)\simeq\frac{2\pi}{\gamma}\int_{-\pi}^{\pi}\left(e^{i\theta(g-h)}+e^{i\theta(h+g)}\right)\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\theta|}\right)|\theta|^{-4d}\,d\theta. \tag{6.23}$$

Combining (6.22) and (6.23), we get

$$\begin{aligned}
var\big(\widehat{\mathcal{S}}_w(\eta_j)\big) &\simeq \frac{1}{2\pi\gamma}\int_{-\pi}^{\pi}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\theta|}\right)|\theta|^{-4d}\sum_{h,g=-t_\gamma}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\mathcal{W}\left(\frac{g}{t_\gamma}\right)e^{-i\eta_j(h-g)}\left(e^{i\theta(g-h)}+e^{i\theta(h+g)}\right)d\theta\\
&\simeq \frac{1}{2\pi\gamma}\int_{-\pi}^{\pi}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\theta|}\right)|\theta|^{-4d}\sum_{h,g=-t_\gamma}^{t_\gamma}\mathcal{W}\left(\frac{h}{t_\gamma}\right)\mathcal{W}\left(\frac{g}{t_\gamma}\right)e^{-ig(\eta_j-\theta)}\left(e^{ih(\eta_j-\theta)}+e^{-ih(\eta_j+\theta)}\right)d\theta\\
&\overset{\mathbf{P2}}{\simeq} \frac{2\pi}{\gamma}\int_{-\pi}^{\pi}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\theta|}\right)|\theta|^{-4d}\mathcal{U}_{t_\gamma}^2(\eta_j-\theta)d\theta\simeq\frac{2\pi}{\gamma}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\eta_j|}\right)|\eta_j|^{-4d}\int_{-\pi}^{\pi}\mathcal{U}_{t_\gamma}^2(\theta)d\theta. \qquad (6.24)
\end{aligned}$$

By Parseval's theorem,

$$2\pi\int_{-\pi}^{\pi}\mathcal{U}_{t_\gamma}^2(\theta)d\theta=\sum_{|h|\le t_\gamma}\mathcal{W}^2\Big(\frac{h}{t_\gamma}\Big),$$

(6.24) will become

$$\begin{aligned}
var\big(\widehat{\mathcal{S}}_w(\eta_j)\big) &\simeq \frac{1}{\gamma}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\eta_j|}\right)|\eta_j|^{-4d}\sum_{|h|\le t_\gamma}\mathcal{W}^2\Big(\frac{h}{t_\gamma}\Big)\\
&\simeq \frac{t_\gamma}{\gamma}\mathcal{V}_{\mathcal{S}}^2\left(\frac{1}{|\eta_j|}\right)|\eta_j|^{-4d}\int_{-1}^{1}\mathcal{W}^2(x)\,dx. \qquad (6.25)
\end{aligned}$$

Combining (2.13) and (6.25), we get

$$var\big(\widehat{\mathcal{S}}_w(\eta_j)\big)\simeq(\mathcal{S}(\eta_j))^2\,\frac{t_\gamma}{\gamma}\int_{-1}^{1}\mathcal{W}^2(x)\,dx.$$

From **C1-C3**, we obtain

$$\int_{-1}^{1}\mathcal{W}^2(x)\,dx<\infty.$$

Then for the above formula and $\eta_j\neq 0$, it is obvious that

$$var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\simeq\frac{t_\gamma}{\gamma}\int_{-1}^{1}\mathcal{W}^2(x)\,dx\simeq O\Big(\frac{t_\gamma}{\gamma}\Big). \tag{6.26}$$

The proof is completed.■

Proof of Corollary 1. For (6.21), (6.11) and (2.11)

$$Bias\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right)\simeq\begin{cases}Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right)+O\left(\frac{1}{t_\gamma\eta_j}\right)^2, & \text{if}\quad 0<d<\frac{1}{2},\ \text{i.e., for LM-DSGP;}\\[2ex] Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right)\simeq O\left(\frac{1}{j}\right), & \text{if}\quad -\frac{1}{2}<d<0,\ \text{i.e., for NM-DSGP.}\end{cases} \tag{6.27}$$

As $\widehat{\mathcal{S}}_p(\eta_j)$ is a special case of $\widehat{\mathcal{S}}_w(\eta_j)$, $\mathcal{W} \equiv 1$ and $t_\gamma = \gamma$. Then, for (6.4) and (6.14),

$$Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \overset{(\alpha_2=0)}{\simeq} \begin{cases} O\left(\frac{\ln j}{j}\right), & \text{if} \quad 0 < d < \frac{1}{2}, \text{ i.e., for LMDSGP;} \\ \\ O\left(\frac{1}{j}\right), & \text{if} \quad -\frac{1}{2} < d < 0, \text{ i.e., for NMDSGP.} \end{cases}$$

For (6.26) with $\mathcal{W} \equiv 1$ and $t_\gamma = \gamma$, $var\left(\frac{\widehat{\mathcal{S}}_P(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq O(1), \quad 0 < |d| < \frac{1}{2}$. ■

Proof of Corollary 2.

- For a LM-DSGP, $0 < d < \frac{1}{2}$, j large enough and for (3.13)

$$Bias\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) \simeq O\left(\frac{\gamma}{t_\gamma j}\right)^2 = O\left(\frac{\gamma}{t_\gamma n_\gamma}\left(\frac{n_\gamma}{j}\right)\right)^2 \simeq O\left(\frac{\gamma}{t_\gamma n_\gamma}\right)^2. \tag{6.28}$$

  From **SC3**, for $j \to \infty$ as $\gamma \to \infty$ we get

$$O\left(\frac{\gamma}{t_\gamma n_\gamma}\right) \xrightarrow[j\to\infty]{SC3} 0. \tag{6.29}$$

  From (3.13), (6.28) and (6.29), for $0 < d < \frac{1}{2}$ we get

$$Bias\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) = E\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)} - 1\right) \xrightarrow[j\to\infty]{SC3} 0,$$

  and for (6.26),

$$var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{S(\eta_j)}\right) \xrightarrow[j\neq 0]{SC1} 0.$$

- The normalized PE-SDF is an asymptotic unbiased;

$$Bias\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \xrightarrow[j\to\infty]{\text{Corollary 1}} 0,$$

  but an inconsistent, $\forall j \neq 0$;

$$var\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{S(\eta_j)}\right) \simeq O(1) \nrightarrow 0.$$

  The proof can be completed. ■

Proof of Theorem 5. It is useful to rewrite the PE-SDF as following with $E(X_t) = 0$

$$\widehat{\mathcal{S}}_p(\eta) := \frac{1}{2\pi\gamma}\left(\mathcal{C}^2_\gamma(\eta) + \mathcal{H}^2_\gamma(\eta)\right), \tag{6.30}$$

where

$$\mathcal{C}_\gamma(\eta) = \sum_{t=1}^{\gamma} X_t \cos(\eta t), \quad \mathcal{H}_\gamma(\eta) = \sum_{t=1}^{\gamma} X_t \sin(\eta t)$$

and

$$E\left(\mathcal{C}_\gamma(\eta)\right) = E\left(\mathcal{H}_\gamma(\eta)\right) = 0.$$

It is tempting to conclude from the results of [33, 49, 35] that the normalized LWE-SDF are asymptotically independently and identically distributed (i.i.d.) for;

- (i) a stationary GP (independent is equivalent to uncorrelated);
- (ii) under the formula (2.11), of
- (iii) Fourier frequencies and
- (iv) $0 < d < 1/2$.

Hurvich and Beltrao [33] showed that for n_γ fixed (**SC2** is not verified), the normalized PE-SDF asymptotically, cannot be i.i.d. Robinson [49] established the asymptotically independent of the normalized PE-SDF under **SC2**, for any such frequencies

$$\eta_j < \eta_k \quad \text{satisfying} \quad (\ln k)/j \xrightarrow[\gamma\to\infty]{} 0.$$

Lahiri [35] defined a sequences of discrete frequencies $\{\eta_{j\gamma}\}_{j=1}^{\varphi}$, $\forall \varphi \geq 2$, converging to the Fourier frequencies, i.e.,

$$\eta_{j\gamma} \xrightarrow[\gamma\to\infty]{} \eta_j \in [0, \pi],$$

then treated the asymptotically i.i.d, by defining the class of admissible sequences of discrete frequencies $\{\eta_{j\gamma}\}$ converge to η,

- For $\eta \in (0, \pi]$, $\mathcal{F}_\eta = \left\{ \{\eta_{j\gamma}\} : \eta_{j\gamma} \in \Omega_\gamma,\ \eta_{j\gamma} \xrightarrow[\gamma\to\infty]{} \eta \in (0, \pi] \right\}$, where $\Omega_\gamma = \left\{ \frac{2\pi j}{\gamma} : j = 1, \ldots, \gamma/2 \right\}$.
- For $\eta = 0$, we put $\mathcal{F}_0 = \mathcal{F}_{01} \cup \mathcal{F}_{02} \cup \mathcal{F}_{03}$ with

$$\mathcal{F}_{01} = \left\{ \{\eta_{j\gamma}\} : \eta_{j\gamma} \in \Omega_\gamma,\ \eta_{j\gamma} + |\gamma\eta_{j\gamma}|^{-1} \xrightarrow[\gamma\to\infty]{} 0 \right\},$$

$$\mathcal{F}_{02} = \left\{ \{\eta_{j\gamma}\} : \eta_{j\gamma} \in \Omega_\gamma,\ \forall\gamma \text{ and } \gamma\eta_\gamma \xrightarrow[\gamma\to\infty]{} 2\pi l,\ l \in \mathbb{Z} \right\}$$

and

$$\mathcal{F}_{03} = \left\{ \{\eta_{j\gamma}\} : \eta_{j\gamma} \equiv 0 \right\}.$$

Then, the converging class of the sequences $\{\eta_{j\gamma}\}$ is $\mathcal{F} = \cup_{\eta\in[0,\pi]}\mathcal{F}_\eta$. Define the variables

$$Y_{j\gamma} = \frac{\mathcal{C}_\gamma(\eta_{j\gamma}) - E(\mathcal{C}_\gamma(\eta_{j\gamma}))}{\psi_\gamma(\eta_{j\gamma})} \text{ and } Z_{j\gamma} = \frac{\mathcal{H}_\gamma(\eta_{j\gamma}) - E(\mathcal{H}_\gamma(\eta_{j\gamma}))}{\psi_\gamma(\eta_{j\gamma})}, \tag{6.31}$$

where

$$\psi_\gamma^2(\eta) = \mathcal{S}(\eta)\gamma \overset{(2.13)}{\simeq} \gamma|\eta|^{-2d}\mathcal{V}\left(\frac{1}{|\eta|}\right).$$

Combining (6.30) and (6.31), we have

$$\widehat{\mathcal{S}}_p(\eta_{j\gamma}) = \frac{1}{2\pi\gamma}\left(Y_{j\gamma}^2 + Z_{j\gamma}^2\right)\psi_\gamma^2(\eta_{j\gamma}). \tag{6.32}$$

Thus,

- (i) if $|\gamma(\eta_{j\gamma} - \eta_{i\gamma})| \xrightarrow[1\le i\ne j\le\varphi]{} \infty$, $\mathcal{H}_\gamma(\eta_{j\gamma})$ and $\mathcal{C}_\gamma(\eta_{j\gamma})$ are asymptotically independent, so $Y_{j\gamma}$ and $Z_{j\gamma}$ are also asymptotically independent.
- (ii) If $\gamma|\eta_{j\gamma}| \xrightarrow[\gamma\to\infty]{} \infty$, the components $(\mathcal{H}_\gamma(\eta_{j\gamma}), \mathcal{C}_\gamma(\eta_{j\gamma}))$ between them are asymptotically independent and each one has an asymptotic normal distribution, as same as $(Y_{j\gamma}, Z_{j\gamma})$, i.e., (cf. Theorem 2.2. of Lahiri[35]),

$$\begin{cases} Y_{j\gamma} \xrightarrow[\gamma|\eta_{j\gamma}|\to\infty]{D} \mathcal{N}(0,\pi), \\ \\ Z_{j\gamma} \xrightarrow[\gamma|\eta_{j\gamma}|\to\infty]{D} \mathcal{N}(0,\pi). \end{cases} \Longrightarrow \begin{cases} Y_{j\gamma}^* = \frac{1}{\sqrt{\pi}}Y_{j\gamma} \xrightarrow{D} \mathcal{N}(0,1), \\ \\ Z_{j\gamma}^* = \frac{1}{\sqrt{\pi}}Z_{j\gamma} \xrightarrow{D} \mathcal{N}(0,1). \end{cases} \tag{6.33}$$

For (6.32), $\{\widehat{\mathcal{S}}_p(\eta_{j\gamma})\}$ are asymptotically independent an immediate consequence of (i) and (ii). Then for (3.12), $\{\widehat{\mathcal{S}}_w(\eta_{j\gamma})\}$ are also asymptotically independent. Combining (6.31)-(6.33),

$$\frac{2\gamma}{\psi_\gamma^2(\eta_{j\gamma})}\widehat{\mathcal{S}}_p(\eta_{j\gamma}) = \frac{2}{\mathcal{S}(\eta_{j\gamma})}\widehat{\mathcal{S}}_p(\eta_{j\gamma}) = \frac{1}{\pi}\left(Y_{j\gamma}^2 + Z_{j\gamma}^2\right) = Y_{j\gamma}^{*2} + Z_{j\gamma}^{*2} \xrightarrow{d} \chi^2(2). \tag{6.34}$$

Combining (3.12) and (6.34), we get $\widehat{\mathcal{S}}_w(\eta_{j\gamma})/\mathcal{S}(\eta_{j\gamma})$ approximated as a weighted linear combination of independent χ^2 variables. That is

$$\frac{\widehat{\mathcal{S}}_w(\eta_{j\gamma})}{\mathcal{S}(\eta_{j\gamma})} \xrightarrow[\gamma\to\infty]{} \frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)} \sim \delta\chi^2(\nu), \tag{6.35}$$

$X \sim Y$ means that the X and Y have the same distribution. The constants δ and ν in (6.35) determined that the mean and the variance of $\{\delta\chi^2(\nu)\}$ are the same as the asymptotic mean and variance of $\{\widehat{\mathcal{S}}_w(\eta_j)/\mathcal{S}(\eta_j)\}$. From the Theorem 3 and 4, we get

$$E\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right) \xrightarrow[\gamma\to\infty]{} 1, \quad var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq \frac{t_\gamma}{\gamma}\int_{-1}^{1}\mathcal{W}^2(x)\,dx. \tag{6.36}$$

Combining (6.26), (6.35) and (6.36), we get

$$\begin{cases} E(\delta\chi^2(\nu)) = \delta\nu \simeq 1, \\ var\left(\delta\chi^2(\nu)\right) = 2\delta^2\nu \simeq \frac{t_\gamma}{\gamma}\int_{-1}^{1}\mathcal{W}^2(x)\,dx. \end{cases} \Longrightarrow \begin{cases} \delta \simeq \frac{1}{\nu}, \\ \nu \simeq \frac{2\gamma}{t_\gamma\int_{-1}^{1}\mathcal{W}^2(x)dx}. \end{cases} \tag{6.37}$$

From (6.37), the proof is completed.■

Proof of Corollary 3.

$$P\left(l_1 \leq \frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)} \leq l_2\right) \simeq 1-\alpha, \text{ i.e., } P\left(\frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{l_2} \leq \mathcal{S}(\eta_j) \leq \frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{l_1}\right) \simeq 1-\alpha,$$

where l_1 and l_2 are obtained from the table of the chi-square distribution satisfying

$$P(\chi^2(\nu) \leq l_2) = P(\chi^2(\nu) \geq l_1) = \alpha/2.$$

Therefore, the $(1-\alpha)$ confidence interval of the SDF $\mathcal{S}(\eta_j)$ is given as follows

$$\frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{a} \leq \mathcal{S}(\eta_j) \leq \frac{\nu\widehat{\mathcal{S}}_w(\eta_j)}{b} \quad \text{and} \quad \frac{\widehat{\mathcal{S}}_p(\eta_j)}{a} \leq \mathcal{S}(\eta_j) \leq \frac{\widehat{\mathcal{S}}_p(\eta_j)}{b}.$$

As we have already mentioned, $\widehat{\mathcal{S}}_p(\eta_j)$ find by $\mathcal{W} \equiv 1$ and $t_\gamma = \gamma$ in formula (3.3). So for (6.37) and (6.26), $\delta = \nu = 1$;

$$E\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq E(\delta\chi^2(\nu)) = 1$$

and

$$var\left(\frac{\widehat{\mathcal{S}}_p(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq var\left(\delta\chi^2(\nu)\right) = 2.$$

■

Proof of Theorem 4.5. The mean and variance of the LWE-DOM $\widehat{d}_w$ for a LMDSGP can be obtained by substituting the LWE-SDF $\widehat{\mathcal{S}}_w$ in (3.3) in the regression formula (3.16), where

$$v_j = \ln \widehat{\mathcal{S}}_w(\eta_j) \quad \text{and} \quad \tilde{e}_j = \ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right).$$

From Theorem 5 (cf. (6.36) and (6.26)),

$$E\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right) \xrightarrow[\gamma\to\infty]{} 1 \quad \text{and} \quad var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right) \simeq \frac{t_\gamma}{\gamma}\int_{-1}^{1} \mathcal{W}^2(x)dx, \ as\ \gamma \to \infty.$$

From Theorem 3, the variable $\tilde{e}_j$ is i.i.d., formula (3.17) becomes

$$Bias(\widehat{d}_w) = E(\tilde{e}_j), \qquad var(\widehat{d}_w) = \frac{var(\tilde{e}_j))}{\sum_{j=1}^{n_\gamma}(u_j - \bar{u})^2} \simeq \frac{var(\tilde{e}_j)}{n_\gamma}. \tag{6.38}$$

For the transformed variable (6.35) and Lemma 4 in Hassler[27], we have

$$\begin{cases} E\left(\tilde{e}_j\right) = E\left(\ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) = \Psi\left(\frac{\gamma}{t_\gamma\int_{-1}^{1}\mathcal{W}^2(x)dx}\right) - \ln\left(\frac{\gamma}{t_\gamma\int_{-1}^{1}\mathcal{W}^2(x)dx}\right), \\ \\ var\left(\tilde{e}_j\right) = var\left(\ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) = \sum_{n_\gamma=0}^{\infty}\left(\frac{\gamma}{t_\gamma\int_{-1}^{1}\mathcal{W}^2(x)dx} + n_\gamma\right)^{-2}. \end{cases} \tag{6.39}$$

Due to the difficulty of calculating, using the approximation in [42] as same as Lemma 1 in [41], with $j \to \infty$

$$\begin{cases} E\left(\tilde{e}_j\right) = E\left(\ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \simeq \ln\left(E\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \xrightarrow[\gamma\to\infty]{\text{Theorem 1}} 0, \\ var\left(\tilde{e}_j\right) = var\left(\ln\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)\right) \simeq \frac{var\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)}{E\left(\frac{\widehat{\mathcal{S}}_w(\eta_j)}{\mathcal{S}(\eta_j)}\right)^2} \overset{\text{Theorem 2}}{\simeq} \frac{t_\gamma}{\gamma}\int_{-1}^{1}\mathcal{W}^2(x)dx. \end{cases} \tag{6.40}$$

Combining (6.38) and (6.40), we get the mean and variance of the LWE-DOM $\widehat{d}_w$.■

Proof of Corollary 4. The proof is obvious from the sufficient conditions **SC1-SC3** and Theorem 4.5 with same reasoning as (3.22). ■

Proof of Corollary 4.6. The proof is obvious from Theorem4 and Corollary4. ■

Bibliography

[1] Agiakloglou, C., Newbold, P., and Wohar, M. 1993. Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis* 14: 235-246.

[2] Alvarez, M. A., Rosasco, L. and Lawrence, N.D., 2012. Kernels for vector-valued functions: A review. *Foundations and Trends in Machine learning* 4(3), 195-266.

[3] Azman, K. and Kocijan, K., 2007. Application of Gaussian processes for black-box modelling of biosystems. *ISA Transactions* 46, 443-457.

[4] Baillie, R. T., 1996. Long-memory processes and fractional integration in econometrics. *J. Econometrics* 73, 5-59.

[5] Bartlett, M. S, 1948. Smoothing Periodograms from Time Series with Continuous Spectra. *Nature, 161, 686-687.*

[6] Beran, J., 1994. Statistics for Long Memory Processes.*Chapman and Hall, New York.*

[7] Beran, J., Kulik, R., Ghosh, S. and Feng, Y., 2013. Long-memory processes: Probabilistic properties and statistical methods. *Springer Science and Business Media, London.*

[8] Blackman, R. B. and Tukey, J. W., 1959. The measurement of power spectra from the point of view of communications engineering- Part I. *Dover, New York.*

[9] Box, G. E. P. and Jenkins, G. M., 1976. Time series analysis: forecasting and control. *Holden Day, San Francisco.*

[10] Brockwell, P. J. and Davis, R. A., 1991. Time Series: Theory and Methods, 2nd Edn. *Springer Verlag, New York.*

[11] Chen, G., Abraham, B. and Peiris, S., 1993. Lag window estimation of the degree of differencing in fractionally integrated time series models. *J. Time Ser. Anal.* 15 (5), 473-487.

[12] Cressie, N., 1992. Statistics for Spatial Data. *Terra Nova* 4(5), 613-617.

[13] Deisenroth, M. P. D. and Rasmussen, C. E., 2015. Gaussian Processes for Data-efficient Learning in Robotics and Control. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37(2), 408-423.

[14] Fay, G. and Soulier, P. 2001. The periodogram of an i.i.d. sequence. *Stochastic Process. Appl., Vol. 92, No. 2, 315-343.*

[15] Fox, R. and Taqqu, M. S., 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* 14(2), 517-532.

[16] Elsawah, A. M., 2021a. Multiple doubling: a simple effective construction technique for optimal two-level experimental designs. *Statist Papers* 62, 2923-2967.

[17] Elsawah, A. M., 2021b. An appealing technique for designing optimal large experiments with three-level factors. *J. Comput. Appl. Math.* 384, 113164.

[18] Elsawah, A. M., 2021c. Designing optimal large four-level experiments: a new technique without recourse to optimization softwares. *Communications in Mathematics and Statistics* https://doi.org/10.1007/s40304-021-00241-y.

[19] Elsawah, A. M., 2020. Building some bridges among various experimental designs. *J Korean Stat Soc* 49, 55-81.

[20] Frederiksen, P., Nielsen, F. S. and Nielsen, M., 2012. Local polynomial Whittle estimation of perturbed fractional processes. *J. Econometrics.*

[21] Geweke, J. and Porter-Hudak, S., 1983. The estimation and application of long memory time series models. *J. Time Ser. Anal.* 4(4), 221-238.

[22] Granger, C. W. J. and Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* 1, 15-29.

[23] Giraitis, L., Koul, H.L., and Surgailis, D. (2012). Large Sample Inference for Long Memory Processes. *Imperial College Press.*

[24] Grenander, U. and Rosenblatt, M., 1953. Statistical spectral analysis of time series arising from stationary stochastic processes. *Ann. Math. Statist.* 24(4), 537-558.

[25] Gušgan, D., 2003. Les chaos finance: approche statistique. *Economica, Paris.*

[26] Hamilton, J. B., 1994. Time Series Analysis. *Princeton University Press.*

[27] Hassler, U., 1993. Regression of spectral estimator with fractionally integrated time series. *J. Time Ser. Anal.* 14(4), 369-380.

[28] Hassler, U., 2019. Time Series Analysis with Long Memory in View. *John Wiley and Sons, Inc.*

[29] Hosking, J. R. M., 1981. Fractional differencing. *Biometrika* 68(1), 165-176.

[30] Koopmans, L. M., 1995. The Spectral Analysis of Time Series. *Academic press, INC.* ISBN 0080541569, 9780080541563

[31] Hunt, R. L., Peiris, M. S. and Weber, N. C., 2003. The bias of lag window estimators of the fractional difference parameter. *J. Appl. Math. Comput.* 12, 67-79.

[32] Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Trans. Am. Soc. Civil Engineers, 116, 770-799.

[33] Hurvich, C. and Beltrao, K., 1993. Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *J. Time Ser. Anal.* 14 (5), 455-472.

[34] Künsch, H. R., 1987. Statistical aspects of self-similar processes. *Proceedings of the First World Congress of the Bernoulli Society.*

[35] Lahiri, S. N., 2003. A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short-and long-range dependence. *Ann. Statist.* 31(2), 613-641.

[36] Lang, G. and Azais, J. M., 1999. Non-parametric estimation of the long-range dependence exponent for Gaussian processes. *Journal of Statistical Planning and Inference. 80, 59-80.*

[37] Lo, A. 1991. Long-term memory in stock market prices. *Econometrica 59.1279-1313.*

[38] Marple, L. (1987). Digital Spectral Analysis with Applications. Englewood Cliffs, NJ:Prentice Hall.

[39] Otis M. Solomon, Jr. 1991. PSD Computations Using Welch¡s Method. *Sandia National Laboratories Albuquerque.*

[40] Parzen, E., 1981. Time series model identification and prediction variance horizon in Applied Time Series Analysis II, Findley Ed., Academic Press, New York, pp 415-447.

[41] Peiris, M. S. and Court, J. R., 1993. A note on the estimation of degree of differencing in long memory time series analysis. *Probab. Math. Statist.* 14(2), 223-229.

[42] Priestley, M. B., 1981. Spectral analysis and Time Series. *Academic Press, New York. School of Mathematics and Statistics.*

[43] Rahi, P. K. and Rajesh Mehra, R., 2014. Analysis of Power Spectrum Estimation Using Welch Method for Various Window Techniques. *International Journal of Emerging Technologies and Engineering (IJETE) ISSN: 2348-8050*

[44] Rashid, A. S., Allami, M. J. H. and Mutasher, A. K., 2020. Best Lag Window for Spectrum Estimation of Law Order MA Process, *Abstract and Applied Analysis,* ID 9352453.

[45] Rasmussen, C. E. and Williams, C. K. I., 2006. Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning. *MIT Press*

[46] Reisen, V.A., 1994. Estimation of the fractional difference parameter in the ARIMA(p, d, q) model using the smoothed periodogram. *J. Time Ser. Anal.* 15(3), 335-350

[47] Robinson, P. M., 1994a. Time series with strong dependence. In: Sims, C.A. (Ed.), Proceedings of the Sixth World Congress on Advances in Econometrics. Cambridge University Press, Cambridge.

[48] Robinson, P. M., 1994b. Semiparametric analysis of long-memory time series. *Ann. Statist.* 22(1), 515-539.

[49] Robinson, P. M., 1995. Log-periodogram regression of time series with long range dependence. *Ann. Statist.* 23(3), 1048-1072.

[50] Robinson, P. M., 1995b. Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* 23 (5), 1630-1661

[51] Rosenblatt, M., 1985. Stationary Sequences and Random Fields. *Boston; Basel; Stuttgart: Birkhauser, ISBN -13: 978-0-8176-3264-9.*

[52] Same, M. H., Gandubert, G., Gleeton, G., Ivanov, P. and Landry, R., Jr. 2021. Simplified Welch Algorithm for Spectrum Monitoring. Appl. Sci., 11, 86. https://dx.doi.org/10.3390/ app11010086

[53] Seneta, E., 2019. Slowly varying functions in the Zygmund sense and generalized regular variation. *Journal of Mathematical Analysis and Applications*, 475, 2, 1647-1657. https://doi.org/10.1016/j.jmaa.2019.03.036.

[54] Shimotsu, K. and Phillips, P. C. B. 2005. Exact Local Whittle Estimation of Fractional Integration. *The Annals of Statistics*, 2005, Vol. 33, No. 4, 1890-1933.

[55] Surgailis, D., Koul. H and Giraitis, L., 2012. Large Sample Inference For Long Memory Processes. *World Scientific Publishing Company*

[56] Taqqu, M. S. and Teverovsky, V. 1996. Semi-parametric graphical estimation techniques for long-memory data. *Conference on Applied Probability and Time Series Analysis*, 420-432.

[57] Tikhonov, S., 2007. Trigonometric series with general monotone coefficients. *J. Maths. anal. appl.* 326, 721-735.

[58] Vishwakarma, G.K., Paul, C and Elsawah, A.M., 2020. An algorithm for outlier detection in a time series model using backpropagation neural network. *J. King Saud University-Science* 32, 3328šC3336.

[59] Vishwakarma, G.K., Paul, C and Elsawah, A.M., 2021. A hybrid feedforward neural network algorithm for detecting outliers in non-stationary multivariate time series. *Expert Systems with Applications* 184, 115545.

[60] Willinger, W., Taqqu, M. S., Leland, W.E. and Wilson, D.V., 1995. Self-similarity in high-speed packet traffic: analysis and modelling of Ethernet traffic measurements. *Statistical Science* 10, 67-85.

[61] Wold, H., 1938. A Study in the Analysis of Stationary Time Series. 2nd Edition, Almqristand Witsett, Stockhol

[62] Yajima, Y., 1989. A central limit theorem of Fourier transforms of strongly dependent stationary processes. *J. Time Series Anal.* 10, 375-83.

[63] Yong, C. H., 1974. Asymptotic behaviour of trigonometric series. Chinese University of Hong Kong.

[64] Zygmund, A., 1977. Trigonometric Series. Cambridge University Press.