WIENER-HOPF FACTORIZATION FOR TIME-INHOMOGENEOUS MARKOV CHAINS AND STATISTICAL INFERENCE FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

BY

YICONG HUANG

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Applied Mathematics
in the Graduate College of the
Illinois Institute of Technology

Approved ____________________
Advisor

Approved ____________________
Co-Advisor

Chicago, Illinois
May 2018

ACKNOWLEDGMENT

I would like to express my sincere thanks to my advisors Professor Tomasz R. Bielecki and Professor Igor Cialenco. Since my first day in the department, I have been on an unbelievable journey of doing research in mathematics under their guidance. I am so grateful to be their student. They taught me the foundations of stochastic analysis, stochastic PDEs and enormous skills in research. Their smart ideas enlighten me, their rigorous attitude influences me and their excellent guidance improves me.

I would like to thank Professor Kevin Cassel, Professor Fred Hickernell, and Professor Ruoting Gong for serving on my comprehensive exam and thesis committee. I am thankful to them for their great support and advice.

I am thankful to the Department of Applied Mathematics for the financial support, the excellent courses and so many wonderful research talks. I gained an incredible amount of knowledge in mathematics and finance from these courses. I listened to lots of introductory and cutting edge research talks in all fields of mathematics.

I also wish to thank my group mates, Yu-Sin Chang, Tao Chen, Ziteng Cheng, and Shibi Feng. I benefited a lot from the discussions with them.

I want to thank the Student Health and Wellness Center at IIT for health advice during the preparation of my thesis.

Last but not least, I would like to express my deepest gratitude to my parents and brothers. I would not be able to complete the thesis and the study without their love and endless support.

I dedicate this thesis to my parents and brothers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol	Description
X	Finite state Markov chain
$\mathbf{E}$	State space of X
Q	Generator matrix of X
$(\Omega, \mathscr{F}, \{\mathscr{F}_t\})_{t\geq 0}, \mathbb{P}$	Stochastic basis satisfying the usual assumptions
H	Separable Hilbert Space
W^Q	Q-cylindrical Brownian motion
θ	Drift coefficient
σ	Volatility coefficient
$\mathbb{R}_+$	Set of positive real numbers

ABSTRACT

The thesis consists of two major parts, and it contributes to two topics in stochastic analysis – Wiener-Hopf factorization (WHf) for Markov chains and statistical inference for Stochastic Partial Differential Equations (SPDEs).

The first part deals with Wiener-Hopf factorization for finite state time inhomogeneous Markov chains. To the best of our knowledge, this study is the first attempt to investigate the WHf for time-inhomogeneous Markov chains. In this work we only deal with a special class of time-inhomogeneous Markovian generators, namely piece-wise constant, which allows to derive the corresponding WHf by using an appropriately tailored randomization technique. Besides the mathematical importance of the WHf methodology, there is also an important computational aspect: it allows for efficient computation of important functionals of Markov chains. In this work, we also provide an efficient algorithm to compute the quantities in the Wiener-Hopf factorization for time-inhomogeneous Markov chains. Finally, we provide a comparison (based on numerical simulations) between our algorithm and the brute-force Monte Carlo simulations.

The second part is dedicated to statistical inference for Stochastic Partial Differential Equations (SPDEs). First, we study the problem of estimating the drift/viscosity coefficient for a large class of linear, parabolic SPDEs driven by an additive space-time noise. We propose a new class of estimators, called trajectory fitting estimators (TFEs). The estimators are constructed by fitting the observed trajectory with an artificial one, and can be viewed as an analog to the classical least squares estimators from the time-series analysis. As in the existing literature on statistical inference for SPDEs, we take a spectral approach, and assume that we observe the first N Fourier modes of the solution, and we study the consistency and the asymptotic normality of the TFE, as $N \to \infty$. Next we consider a parameter estimation problem

for one dimensional stochastic heat equation, when data is sampled discretely in time or spatial component. We establish some general results on derivation of consistent and asymptotically normal estimators based on computation of the p-variations of stochastic processes and their smooth perturbations. We apply these results to the considered SPDEs, by using some convenient representations of the solutions. For some equations such representations were ready available, while for other classes of SPDEs we derived the needed representations along with their statistical asymptotic properties. We prove that the real valued parameter next to the Laplacian, and the positive parameter in front of the noise can be consistently estimated by observing the solution at a fixed time and on a discrete spatial grid, or at a fixed space point and at discrete time instances of a finite interval, assuming that the mesh size goes to zero.

CHAPTER 1

INTRODUCTION

This thesis consists of two major parts, and it contributes to two topics in stochastic analysis: (i) Wiener-Hopf factorization (WHf) for Markov chains and (ii) statistical inference for Stochastic Partial Differential Equations (SPDEs).

1.1 Wiener-Hopf Factorization for Markov Chains

The theory of WHf for Markov chains was originated in [BRW80]. Later, Kennedy and Williams [KW90] studied the so-called "noisy" WHf. This theory has been applied to many practical problems such as ruin problem [APU03], fluid flow models [Rog94, RS94, Asm95], biology [Hie14], and finance [JR06, JP08, MP11, JP12, HSZ16]. However, all these theoretical and applied developments of WHf are only done for time-homogeneous Markov chains. Clearly, the case of time-inhomogeneous chains is important, both from theoretical and application point of view. The main goal of this part of the thesis is to develop the WHf theory for *time-inhomogeneous* Markov chains.

First, we outline the original WHf for time-homogeneous Markov chains. Consider a time-homogeneous Markov chain X with finite-state space $\mathbf{E}$ and generator Q. Let $v : \mathbf{E} \to \mathbb{R} \setminus \{0\}$ be a function such that the sets $\mathbf{E}^+ := \{i \in \mathbf{E} \mid v(i) > 0\}$ and $\mathbf{E}^- := \{i \in \mathbf{E} \mid v(i) < 0\}$ are nonempty. In addition, let $\phi(t) := \int_0^t v(X_s)\,\mathrm{d}s$ and define the first passage times $\tau_t^+ := \inf\{s \geq 0 \mid \phi(s) > t\}$ and $\tau_t^- := \inf\{s \geq 0 \mid -\phi(s) > t\}$, for $t \geq 0$. One of the problems is to compute the following expectations,

$$\mathbb{E}\left(e^{-c\tau_0^+}\mathbb{1}_{\{X_{\tau_0^+}=j\}}|X_0 = i\right), \quad i \in \mathbf{E}^-, j \in \mathbf{E}^+, \tag{1.1}$$

$$\mathbb{E}\left(e^{-c\tau_t^+}\mathbb{1}_{\{X_{\tau_t^+}=j\}}|X_0 = i\right), \quad i \in \mathbf{E}^+, j \in \mathbf{E}^+, t > 0, \tag{1.2}$$

$$\mathbb{E}\left(e^{-c\tau_0^-}\mathbb{1}_{\{X_{\tau_0^-}=j\}}|X_0 = i\right), \quad i \in \mathbf{E}^+, j \in \mathbf{E}^-, \tag{1.3}$$

$$\mathbb{E}\left(e^{-c\tau_t^-}\mathbb{1}_{\{X_{\tau_t^-}=j\}}|X_0=i\right), \quad i\in\mathbf{E}^-, j\in\mathbf{E}^-, t>0, \tag{1.4}$$

where $c>0$ is a positive constant.

Next, for illustration purposes, let us consider as an example the classical fluid flow problem (cf. [Mit88] and [Rog94] for detailed discussion). Suppose that we have a large water tank with infinite capacity. On the top of the tank, there are $\mathcal{I}$ pipes, with I_t pipes being open at time t and each pipe pouring water into the tank at the same rate r^+. At the bottom of the tank, there are $\mathcal{O}$ taps, with O_t being open at time t, each tap allowing water to flow out of the tank at the same rate r^-. We assume that $\mathcal{I}$ and $\mathcal{O}$ are finite. Then, the volume ξ_t of water in the tank at time t satisfies the dynamics

$$\frac{\mathrm{d}\xi_t}{\mathrm{d}t}=r^+I_t-r^-O_t, \quad \text{if } 0<\xi_t<a.$$

Moreover, if $\xi_t=0$, i.e. if the tank is empty, then the outflow ceases. Let f be a real valued function on $\mathcal{I}\times\mathcal{O}$. We assume that $X_t:=f(I_t,O_t),\ t\geq 0$, is a (finite state) time-homogeneous Markov chain, and we denote by $\mathbf{E}$ the state space of X. Let us consider the function

$$v(x):=V(r^+,r^-,x), \quad x\in\mathbf{E},$$

that will model the water outflow/inflow, in terms of the states of X, so that

$$v(X_t)=V(r^+,r^-,f(I_t,O_t)), \quad t\geq 0$$

represents the water outflow/inflow at time t. Let $\mathbf{E}^+$ be the set of states of X such that the water tank has greater water inflow than outflow, and let $\mathbf{E}^-$ be the set of states of X such that the water tank has greater water outflow than inflow. Note that the integral $\varphi_t = \int_0^t v(X_u)\,\mathrm{d}u$ is not exactly the water content at time t, since we should take into account those periods of time when the tank is empty. However, understanding φ_t, $\tau_t^{\pm}$ and $X_{\tau_t^{\pm}}$ allows us to express the quantities of interest for ξ_t in terms of the Wiener-Hopf factorization, and to consequently numerically compute these. We now assume that the tank contains $\ell > 0$ amount of water at time $t = 0$. Thus, τ_ℓ^- represents the first time after $t = 0$ that the tank goes empty and τ_0^- represents the first time after $t = 0$ that the tank returns to ℓ amount of water. Expectations (1.3) and (1.4) are the Laplace transforms of the joint distributions of $(\tau_0^-, X_{\tau_0^-})$ and $(\tau_\ell^-, X_{\tau_\ell^-})$, respectively.

Barlow et al. [BRW80] showed that the time changed processes $X_{\tau_t^+}$ and $X_{\tau_t^-}$ are time homogeneous Markov chains with state spaces $\mathbf{E}^+$ and $\mathbf{E}^-$, respectively. Moreover, in [BRW80] it was proved that Q can be factorized uniquely in terms of generators Q^+ and Q^- of $X_{\tau_t^+}$ and $X_{\tau_t^-}$ respectively. This factorization was called the Wiener-Hopf factorization. While the result is algebraic, its proof is probabilistic. Furthermore, the expectations (1.1), (1.2), (1.3), and (1.4) can be expressed in terms of this factorization. Thus, the problem reduces to find the WHf for the generator Q. On one hand, there is no closed-form solution to the WHf, except for some trivial cases, and one may need to apply numerical techniques to find the WHf. On the other hand, one may argue that these expectations can be computed by using Monte Carlo methods. However, by its very nature, WHf methods provide faster and more accurate results (cf. [RS94], [Hie14]) than Monte Carlo methods.

To the best of our knowledge, there are no results on WHf for time inhomogeneous Markov chains, i.e. whose generators are time dependent. One naive way to address this problem is to factorize the time dependent generator according to WHf in [BRW80] at each fixed time point. However, such factorization does not have any probabilistic interpretation. In particular, the expectations (1.1)-(1.4) cannot be expressed in terms of this factorization. Alternatively, one can homogenize the original Markov chain (cf. [BÏ4]), and then apply the general Wiener-Hopf factorization derived in [Wil08]. However, it is still not clear how to compute the general WHf as in [Wil08] and then convert the results back to the original problem. Our aim is to develop a numerically tractable WHf for time inhomogeneous Markov chains. As a first attempt, in this work, we consider a time-inhomogeneous Markov chain with a generator that is piece-wise constant as a matrix-valued function of time. In Chapter 2, we propose a randomization method to construct a suitable time-homogeneous Markov chain based on the original chain. We then apply the Wiener-Hopf factorization from [BRW80] to the new chain, and it turns out that the Laplace transform sets up the connection between the factorization and the expectation we were interfered to compute. In addition, the special structure of the generator allows to establish an efficient algorithm for computation of the Wiener-Hopf factorization. The results presented in this chapter are based on the recent work [BCGH18].

Chapter 2 is organized as follows. In Section 2.2, we briefly review the Wiener-Hopf factorization for time-homogeneous Markov chains – the algebraic factorization and the probabilistic interpretation. In Section 2.3 we present the proposed WHf methodology for time-inhomogeneous Markov chains with a piece-wise constant generator. The randomization method for constructing a time-homogeneous Markov chain is addressed in Section 2.3.1. This time-homogeneous Markov chain has to components, and we prove that the first component is itself a Markov chain. Moreover, we construct a new measure under which the second component is a time-inhomogeneous Markov chain with the same generator as the original chain. In Section 2.3.2, we apply the Wiener-Hopf factorization in [BRW80] to the newly constructed time-homogeneous Markov chain and set up the connections through the inverse Laplace transform. Section 2.4 is devoted to the numerical study for the developed theory. We first introduce an algorithm to compute the factorization (Section 2.4.1), and then we study an application to fluid flow problems (Section 2.4.2).

The main contributions of this part of the thesis can be summarized as follows:

- We extend the theory of Wiener-Hopf factorization for time-homogeneous Markov chains to time-inhomogeneous Markov chains with a piece-wise constant generator.
- We provide an algorithm to compute the quantities in the Wiener-Hopf factorization for time-inhomogeneous Markov chains. For a particular example we give a comparison between our algorithm and the brute-force Monte Carlo simulations.

1.2 Statistical Inference for Stochastic Partial Differential Equations

Stochastic Partial Differential Equations (SPDEs) arise from various applied topics, such as nonlinear filtering, modeling of turbulent flows, population growth models, fixed income market models, etc. The general theory of SPDEs has been studied quite intensively during the past few decades, and we refer to the classical monographs [Roz90, DPZ92], as well as some recent textbooks [Cho07, Hai09, LR17] for a detailed discussion of the theory of SPDEs and their applications. Usually, the general form of the equation is derived from the fundamental laws of the underlying process. However, the model parameters, generally speaking, are not known a priori, and have to be determined typically by statistical methods based on some historical observations of the underlying model. While statistical inference for Stochastic Ordinary Differential Equations (SODEs) is well understood (cf. [Kut04]), the statistical inference for SPDEs is still in its developing stage. We refer to the survey paper [Cia18] for the recent developments in this field. Since the pioneering works [HKR93, HR95], many statistical estimation problems for SPDEs have been studied under the spectral approach, namely assuming that a finite number N of the Fourier coefficients of the solutions are observed over some finite interval of time $[0,T]$. The statistical inference problems in large time asymptotic regime $T \to \infty$ essentially becomes a statistical inference problem for SODEs, and hence are well understood. It turns out that the asymptotic regime of large number of Fourier modes $N \to \infty$ is a viable regime to study, and in many cases one can derive consistent and asymptotically normal estimators for the parameters of interest such as the drift or viscosity coefficient (the coefficient appearing in the $\mathrm{d}t$ term) and/or the volatility (the coefficient in the noise term). Usually, in the existing literature these estimators are derived as maximum likelihood estimators (MLEs). In this work we propose two novel methods of estimating the drift and volatility coefficients for some linear parabolic SPDEs.

First, within the spectral approach, we propose a new estimator for the drift coefficient, called trajectory fitting estimator (TFE), which can be viewed as an analog of the least squares estimator from the time series analysis. This type of estimator was originally introduced by [Kut91] (see also [Kut04]) in the context of estimating drift coefficient for ergodic diffusion processes in the large time asymptotic regime. We study the asymptotic properties of TFE for SPDEs when $N \to \infty$. The obtained results are based on [CGH18].

The second method goes beyond the spectral approach, by assuming that the input data is the measurements of the values of the solution at discrete points in time and/or space. Besides the fact that this sampling scheme is practically more important, in contrast to observing the Fourier coefficients, this study is among the few works on parameter estimation for discretely sampled SPDEs. The proposed estimators and their asymptotic properties are derived through the computations of the $p-$variation of some suitable stochastic processes. We study the asymptotic properties in two sampling regimes: when the number of spatial observations increases while time is fixed, and respectively when the time resolution is increasing and while the solution is observed at a fixed spatial point. The results are based on the recent work [CH17]. We use some techniques from Malliavin calculus to prove asymptotic normality of some of these estimators. In should be mentioned that in a recent work [BT17] the authors studied independently similar problems and derived similar estimators for the volatility coefficient by using the mixing theory of Gaussian time-series.

Chapter 3 is organized as follows. In Section 3.1, we introduce the general equation and briefly discuss the existence and the uniqueness of the solution, and we also state the parameter estimation problem, and review the relevant literature. Section 3.2 is devoted to the study of trajectory fitting estimators for diagonalizable parabolic equations. In particular, we prove the consistency of TFE in Section 3.2.2 and the asymptotic normality in Section 3.2.3. Section 3.2.4 is devoted to some examples. The discrete sampling scheme is investigated in Section 3.3, starting with SPDEs on whole space – Section 3.3.2, and continuing with SPDEs on bounded domain in Section 3.3.3. In Section 3.4 we present some numerical simulations results.

The main contributions of this part amount to:

- Within the spectral approach, we proposed the trajectory fitting estimator for SPDEs. This is an analog to the least squares estimator in the classical time-series analysis. We prove the consistency and asymptotic normality of the proposed estimators for a large class of linear parabolic SPDEs.

- Beyond the spectral approach, we assume the solution is directly observed at discrete time and/or space points and propose the $p-$variational type estimators for the drift and volatility parameters. We prove the consistency and asymptotic normality of the proposed estimators.

CHAPTER 2
WIENER-HOPF FACTORIZATION FOR MARKOV CHAINS

2.1 Introduction

In this part of the thesis we will derive the Wiener-Hopf factorization (WHf) for a finite-state *time-inhomogeneous* Markov chain, which is the first attempt to investigate the WHf for time-inhomogeneous chains. In this pioneering study we only deal with a special class of time-inhomogeneous Markovian generators, namely piecewise constant, which allows to use an appropriately tailored randomization technique as seen below.

The Wiener-Hopf factorization for finite state time-homogeneous Markov chains was originally proposed in seminal work [BRW80]; see Section 2.2 below, as well as [LMRW82, Wil91]. For the WHf in case of time-homogeneous Feller Markov processes we refer to [Wil08]. For some related applied work we refer to [APU03], which deals with the ruin problem, and to [Asm95, Rog94, RS94] that study fluid models. In addition, [KW90] studies the so called "noisy" Wiener-Hopf factorizations; for applications see [Asm95, Rog94, RS94, JR06, JP08, MP11, JP12, Hie14, HSZ16]. In all these applied problems there is no practical reason to assume that the Markov chain is time-homogeneous, rather than that the existing WHf methodology was available only for time-homogeneous chains. This was one of the main motivations to study the WHf for time-inhomogeneous Markov chains.

It needs to be stressed that even though the classical WHf technique [BRW80] can be applied to the time dependent generator matrix, say G_t, of a time inhomogeneous Markov chain X at every time t, the obtained factorizations in this case do not have any probabilistic meaning with regard to the process X. In particular, they are of no use for computing functionals such as (2.1)-(2.4) below. So, a relevant WHf for a time-inhomogeneous Markov chain requires a different approach than the one that would just directly apply the results of [BRW80] to each G_t, $t \geq 0$.

2.2 Wiener-Hopf Factorization for Time-Homogeneous Markov Chains

We briefly review the Wiener-Hopf factorization for finite-state time homogeneous Markov chains that was originally derived in [BRW80].

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $\mathbf{E}$ be a finite set. Without loss of generality we assume that $\mathbf{E} = \{1, \ldots, |\mathbf{E}|\}$, where $|\mathbf{E}|$ denotes the cardinality of the set $\mathbf{E}$. We will denote by $\mathcal{Q}(\mathbf{E})$ the set of all $|\mathbf{E}| \times |\mathbf{E}|$ matrices Q such that, for $i, j \in \mathbf{E}$,

$$\mathsf{Q}(i,j) \geq 0, \quad i \neq j, \quad \text{and} \quad \sum_{k \in \mathbf{E}} \mathsf{Q}(i,k) \leq 0.$$

We consider a function $v : \mathbf{E} \to \mathbb{R} \setminus \{0\}$ such that

$$\mathbf{E}^+ := \{i \in \mathbf{E} \,|\, v(i) > 0\} \quad \text{and} \quad \mathbf{E}^- := \{i \in \mathbf{E} \,|\, v(i) < 0\},$$

are nonempty, and we denote by V the diagonal $|\mathbf{E}| \times |\mathbf{E}|$ matrix, i.e. $\mathsf{V} = \mathrm{diag}\{v(i) | i \in \mathbf{E}\}$. We denote by $\mathsf{I}^{\pm}$ the identity $|\mathbf{E}|^{\pm} \times |\mathbf{E}|^{\pm}$ matrices. In what follows, let $\mathsf{Q} \in \mathcal{Q}(\mathbf{E})$ and $c > 0$ be a fixed real number. The next result provides the so-called Wiener-Hopf factorization for the matrix $\mathsf{V}^{-1}(\mathsf{Q} - c\mathsf{I})$.

Theorem 2.2.1 ([BRW80])**.** *There exists a unique pair (Π_c^+, Π_c^-), where Π_c^+ is an $|\mathbf{E}|^- \times |\mathbf{E}|^+$ matrix and Π_c^- is an $|\mathbf{E}|^+ \times |\mathbf{E}|^-$ matrix, such that, if*

$$\mathsf{S} = \begin{pmatrix} \mathsf{I}^+ & \Pi_c^- \\ \Pi_c^+ & \mathsf{I}^- \end{pmatrix},$$

then S *is invertible and*

$$\mathsf{S}^{-1}(\mathsf{V}^{-1}(\mathsf{Q}-c\mathsf{I}))\mathsf{S} = \begin{pmatrix} \widetilde{\mathsf{Q}}_c^{+} & 0 \\ 0 & -\widetilde{\mathsf{Q}}_c^{-} \end{pmatrix},$$

where $\widetilde{\mathsf{Q}}_c^{+} \in \mathscr{Q}(\mathbf{E}^{+})$ *and* $\widetilde{\mathsf{Q}}_c^{-} \in \mathscr{Q}(\mathbf{E}^{-})$. *Moreover,* Π_c^{+} *and* Π_c^{-} *are strictly sub-stochastic, thus, for* $i \in \mathbf{E}^{-}, j \in \mathbf{E}^{+}$, $\Pi_c^{+}(i,j) \geq 0$, *and* $\sum_{k\in\mathbf{E}^{+}} \Pi_c^{+}(i,k) < 1$.

On the one hand, Theorem 2.2.1 just gives an algebraic factorization of the matrix $\mathsf{V}^{-1}(\mathsf{Q}-c\mathsf{I})$. On the other hand, as we will explain below, using the probabilistic proof of this theorem we also obtain a natural interpretation of this algebraic factorization in terms of some important quantities related to the time-homogeneous Markov chain X with the state space $\mathbf{E}$ and generator matrix Q. Let us define the additive functional

$$\varphi_t := \int_0^t v(X_u)\,\mathrm{d}u, \quad t \geq 0,$$

and the first passage times

$$\tau_t^{+} := \inf\{r \geq 0 \,|\, \varphi_r > t\} \quad \text{and} \quad \tau_t^{-} := \inf\{r \geq 0 \,|\, \varphi_r < -t\}.$$

It turns out that the time changed processes $X_{\tau_t^{\pm}}$ are Markov chains with generator matrices $\widetilde{\mathsf{Q}}_c^{\pm}$, respectively. In addition, $\Pi_c^{\pm}$ describe the fluctuations between $\mathbf{E}^{+}$ and $\mathbf{E}^{-}$. We assume that $\mathbb{P}(X_0 = i) > 0$ for each $i \in \mathbf{E}$, and we let $\mathbb{P}^i$ be the probability measure on $(\Omega, \mathscr{F})$ defined by

$$\mathbb{P}^i(A) := \mathbb{P}(A \,|\, X_0 = i), \quad A \in \mathscr{F},$$

with $\mathbb{E}^i$ denoting the associated expectation. The next result summarizes the relationships between the algebraic factorization and the time changed Markov chains.

Theorem 2.2.2 ([BRW80]). *For $i \in \mathbf{E}^-$ and $j \in \mathbf{E}^+$,*

$$\mathbb{E}^i \left(e^{-c\tau_0^+} \mathbb{1}_{\{X_{\tau_0^+}=j\}} \right) = \Pi_c^+(i,j).$$

For $i \in \mathbf{E}^+, j \in \mathbf{E}^+$, and $t \geq 0$,

$$\mathbb{E}^i \left(e^{-c\tau_t^+} \mathbb{1}_{\{X_{\tau_t^+}=j\}} \right) = e^{t\widetilde{\mathsf{Q}}_c^+}(i,j).$$

The corresponding minus results follow, on replacing ϕ by $-\phi$.

2.3 Wiener-Hopf Factorization for Time-Inhomogeneous Markov Chains

In what follows we adopt the same notations as in the previous section. We will assume now that $X := (X_t)_{t\geq 0}$ is a *time-inhomogeneous* Markov chain on $(\Omega, \mathscr{F}, \mathbb{P})$ with state space $\mathbf{E}$ and generator function $\mathsf{G} = \{\mathsf{G}_t, t \geq 0\}$.

In this work, we assume that the generator G is piecewise constant, namely we assume that

$$\mathsf{G}_t = \begin{cases} \mathsf{G}_1, & \text{if } s_0 \leq t < s_1, \\ \mathsf{G}_2, & \text{if } s_1 \leq t < s_2, \\ \vdots & \\ \mathsf{G}_n, & \text{if } s_{n-1} \leq t < s_n, \\ \mathsf{G}_{n+1}, & \text{if } t \geq s_n, \end{cases}$$

for some $n \in \mathbb{N}$ and $0 = s_0 < s_1 < \ldots < s_n$. Without loss of generality we assume that $\mathsf{G}_1, \ldots, \mathsf{G}_{n+1}$ are Markovian (not sub-Markovian). That is, the sums of row elements of G_k are all zero, for any $k = 1, \ldots, n+1$. The results of this work carry over to the sub-Markovian case by the standard augmentation of the state space.

The main goal of this work is to develop the Wiener-Hopf factorization technique (see Section 2.3.1) to compute the following expectations,

$$\Pi_c^+(i,j;s_1,\ldots,s_n) := \mathbb{E}\left(e^{-c\tau_0^+}\mathbb{1}_{\{X_{\tau_0^+}=j\}}|X_0=i\right), \quad i \in \mathbf{E}^-,\ j \in \mathbf{E}^+, \tag{2.1}$$

$$\Psi_c^+(\ell,i,j;s_1,\ldots,s_n) := \mathbb{E}\left(e^{-c\tau_\ell^+}\mathbb{1}_{\{X_{\tau_\ell^+}=j\}}|X_0=i\right),\ i \in \mathbf{E}^+,\ j \in \mathbf{E}^+,\ \ell > 0, \tag{2.2}$$

$$\Pi_c^-(i,j;s_1,\ldots,s_n) := \mathbb{E}\left(e^{-c\tau_0^-}\mathbb{1}_{\{X_{\tau_0^-}=j\}}|X_0=i\right), \quad i \in \mathbf{E}^+,\ j \in \mathbf{E}^-, \tag{2.3}$$

$$\Psi_c^-(\ell,i,j;s_1,\ldots,s_n) := \mathbb{E}\left(e^{-c\tau_\ell^-}\mathbb{1}_{\{X_{\tau_\ell^-}=j\}}|X_0=i\right),\ i \in \mathbf{E}^-,\ j \in \mathbf{E}^-,\ \ell > 0. \tag{2.4}$$

We will focus on the computations of $\Pi_c^+(i,j;s_1,\ldots,s_n)$ and $\Psi_c^+(\ell,i,j;s_1,\ldots,s_n)$. By symmetry, analogous results can be obtained for the counterparts $\Pi_c^-(i,j;s_1,\ldots,s_n)$ and $\Psi_c^-(\ell,i,j;s_1,\ldots,s_n)$. To simplify the notations, we will frequently write $\Pi_c^+(i,j)$ and $\Psi_c^+(\ell,i,j)$ in place of $\Pi_c^+(i,j;s_1,\ldots,s_n)$ and $\Psi_c^+(\ell,i,j;s_1,\ldots,s_n)$, respectively.

2.3.1 A Randomization Method for Time Homogenization. In this section we construct a *time-homogeneous* Markov chain associated to X, by randomizing the discontinuity times $s_1,\ldots,s_n$ of the generator G. This key construction will allow us to compute the expectations (2.1), (2.2), (2.3) and (2.4) using analogous expectations corresponding to this time-homogeneous chain. The latter expectations can be computed using Wiener-Hopf factorization theory of [BRW80].

Define $\mathbb{N}_n := \{0,\ldots,n\}$, $\widetilde{\mathbf{E}} := \mathbb{N}_n \times \mathbf{E}$ and let $(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{\mathbb{P}})$ be a complete probability space. Next, let us consider a *time-homogeneous* Markov chain, say $Z = (N,Y) := (N_t,Y_t)_{t\geq 0}$, defined on $(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{\mathbb{P}})$, taking values in $\widetilde{\mathbf{E}}$ and with generator matrix $\widetilde{\mathsf{G}}\left((n_1,j_1),(n_2,j_2)\right)_{(n_1,j_1),(n_2,j_2)\in\widetilde{\mathbf{E}}}$ given as

$$\widetilde{\mathsf{G}} = \begin{array}{c} \\ \scriptstyle\{0\}\times\mathbf{E} \\ \scriptstyle\{1\}\times\mathbf{E} \\ \vdots \\ \scriptstyle\{n-1\}\times\mathbf{E} \\ \scriptstyle\{n\}\times\mathbf{E} \end{array} \begin{array}{c} \begin{array}{ccccc} \scriptstyle\{0\}\times\mathbf{E} & \scriptstyle\{1\}\times\mathbf{E} & \scriptstyle\cdots & \scriptstyle\{n-1\}\times\mathbf{E} & \scriptstyle\{n\}\times\mathbf{E} \end{array} \\ \begin{bmatrix} \mathsf{G}_1 - q_1\mathsf{I} & q_1\mathsf{I} & \cdots & 0 & 0 \\ 0 & \mathsf{G}_2 - q_2\mathsf{I} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathsf{G}_n - q_n\mathsf{I} & q_n\mathsf{I} \\ 0 & 0 & \cdots & 0 & \mathsf{G}_{n+1} \end{bmatrix} \end{array},$$

where $q_1, \ldots, q_n$ are positive constants and I is the identity matrix. For each $i \in \mathbf{E}$, we define the probability measure $\widetilde{\mathbb{P}}^i$ on $(\widetilde{\Omega}, \widetilde{\mathscr{F}})$ by

$$\widetilde{\mathbb{P}}^i(A) := \widetilde{\mathbb{P}}\left(A \,|\, Z_0 = (0,i)\right), \quad A \in \widetilde{\mathscr{F}}. \tag{2.5}$$

The next result regards the Markov property of process N.

Proposition 2.3.1. *For any $i \in \mathbf{E}$, the process N is a time-homogeneous Markov chain under $\widetilde{\mathbb{P}}^i$, with generator matrix given by*

$$\widetilde{\mathsf{G}}_N = \begin{array}{c} \\ \scriptstyle 0 \\ \scriptstyle 1 \\ \vdots \\ \scriptstyle n-1 \\ \scriptstyle n \end{array} \begin{array}{c} \begin{array}{ccccc} \scriptstyle 0 & \scriptstyle 1 & \scriptstyle\cdots & \scriptstyle n-1 & \scriptstyle n \end{array} \\ \begin{bmatrix} -q_1 & q_1 & \cdots & 0 & 0 \\ 0 & -q_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -q_n & q_n \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \end{array}.$$

Proof. We will proceed in three steps.

Step 1. We start by showing that

$$\sum_{j_2\in\mathbf{E}} \left(\widetilde{\mathsf{G}}^k\right)((n_1,j_1),(n_2,j_2)) = \left(\widetilde{\mathsf{G}}^k_N\right)(n_1,n_2), \tag{2.6}$$

for any $j_1 \in \mathbf{E}$, $k \in \mathbb{N}$, and $0 \le n_1, n_2 \le n$, In particular, note that the left-hand-side of (2.6) does not depend on j_1.

We will prove (2.6) by induction in k. Clearly (2.6) holds true for $k = 1$. Next, assume that (2.6) holds for some $k = \ell \in \mathbb{N}$. Now, for $\ell + 1$,

$$\begin{aligned}
\sum_{j_2\in\mathbf{E}} \left(\widetilde{\mathsf{G}}^{\ell+1}\right)((n_1,j_1),(n_2,j_2)) &= \sum_{j_2\in\mathbf{E}}\sum_{m=0}^{n}\sum_{j\in\mathbf{E}} \left(\widetilde{\mathsf{G}}^{\ell}\right)((n_1,j_1),(m,j))\,\widetilde{\mathsf{G}}\left((m,j),(n_2,j_2)\right)\\
&= \sum_{m=0}^{n}\sum_{j\in\mathbf{E}} \left(\widetilde{\mathsf{G}}^{\ell}\right)((n_1,j_1),(m,j)) \sum_{j_2\in\mathbf{E}}\widetilde{\mathsf{G}}\left((m,j),(n_2,j_2)\right)\\
&= \sum_{m=0}^{n}\sum_{j\in\mathbf{E}} \left(\widetilde{\mathsf{G}}^{\ell}\right)((n_1,j_1),(m,j))\,\widetilde{\mathsf{G}}_N(m,n_2)\\
&= \sum_{m=0}^{n} \left(\widetilde{\mathsf{G}}^{\ell}_N\right)(n_1,m)\widetilde{\mathsf{G}}_N(m,n_2) = \left(\widetilde{\mathsf{G}}^{\ell+1}_N\right)(n_1,n_2),
\end{aligned}$$

where we used the inductive assumptions for $k = 1$ and $k = \ell$ in the third and the fourth equalities, respectively. Hence, (2.6) is established.

Step 2. We will show that

$$\widetilde{\mathbb{P}}^i\left(N_{t+s} = n_2 \,|\, N_t = n_1\right) = \widetilde{\mathbb{P}}^i(N_{t+s} = n_2 \,|\, N_t = n_1, Y_t = j) = e^{s\widetilde{\mathsf{G}}_N}(n_1,n_2), \tag{2.7}$$

for any $t, s \ge 0$, $j \in \mathbf{E}$, and $0 \le n_1 \le n_2 \le n$. In particular, note that the left-hand side of (2.7), and thus $\widetilde{\mathbb{P}}^i(N_{t+s} = n_2|N_t = n_1)$, does not depend on t. We start by checking the second equality in (2.7). For any $t, s \geq 0$, $j \in \mathbf{E}$, and $0 \leq n_1 \leq n_2 \leq n$,

$$
\begin{aligned}
\widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2 \,|\, N_t=n_1, Y_t=j\right) &= \sum_{k\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2, Y_{t+s}=k \,|\, N_t=n_1, Y_t=j\right) \\
&= \sum_{k\in\mathbf{E}} e^{s\,\widetilde{\mathsf{G}}}\left((n_1,j),(n_2,k)\right) \\
&= \sum_{k\in\mathbf{E}} \sum_{\ell=0}^{\infty} \frac{s^\ell}{\ell!}\, \widetilde{\mathsf{G}}^\ell\left((n_1,j),(n_2,k)\right) \\
&= \sum_{\ell=0}^{\infty} \frac{s^\ell}{\ell!} \sum_{k\in\mathbf{E}} \widetilde{\mathsf{G}}^\ell\left((n_1,j),(n_2,k)\right) \\
&= \sum_{\ell=0}^{\infty} \frac{s^\ell}{\ell!}\, \widetilde{\mathsf{G}}_N^\ell(n_1,n_2) = e^{s\widetilde{\mathsf{G}}_N}(n_1,n_2),
\end{aligned}
$$

where we used the result of Step 1 in the last two equalities. In particular, $\widetilde{\mathbb{P}}^i(N_{t+s} = n_2|N_t = n_1, Y_t = j)$ does not depend on the choice of $j \in \mathbf{E}$.

As far as the first equality in (2.6), for any $t, s \geq 0$ and $0 \leq n_1 \leq n_2 \leq n$,

$$
\begin{aligned}
\widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2 \,|\, N_t=n_1\right) &= \frac{\widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2, N_t=n_1\right)}{\widetilde{\mathbb{P}}^i\left(N_t=n_1\right)} \\
&= \frac{\sum_{\ell\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2, N_t=n_1, Y_t=j\right)}{\sum_{j\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_t=n_1, Y_t=j\right)} \\
&= \frac{\sum_{j\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t+s}=n_2 \,|\, N_t=n_1, Y_t=j\right) \widetilde{\mathbb{P}}^i\left(N_t=n_1, Y_t=j\right)}{\sum_{j\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_t=n_1, Y_t=j\right)} \\
&= \frac{\sum_{j\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_t=n_1, Y_t=j\right)}{\sum_{j\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_t=n_1, Y_t=j\right)}\, e^{s\widetilde{\mathsf{G}}_N}(n_1,n_2) = e^{s\widetilde{\mathsf{G}}_N}(n_1,n_2).
\end{aligned}
$$

Step 3. We are ready to complete the proof of the proposition. Towards this end we observe that, for any $m \in \mathbb{N}$, $0 = t_0 \leq t_1 < \ldots < t_m$, and any $0 \leq n_1 \leq \ldots \leq n_m \leq n$,

$$
\widetilde{\mathbb{P}}^i(N_{t_m}=n_m \mid N_{t_{m-1}}=n_{m-1},\ldots,N_{t_1}=n_1) = \frac{\widetilde{\mathbb{P}}^i\left(N_{t_1}=n_1,\ldots,N_{t_m}=n_m\right)}{\widetilde{\mathbb{P}}^i\left(N_{t_1}=n_1,\ldots,N_{t_{m-1}}=n_{m-1}\right)}
$$

$$
\begin{aligned}
&= \frac{\sum_{j_1,\dots,j_m\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t_1}=n_1, Y_{t_1}=j_1;\dots;N_{t_m}=n_m, Y_{t_m}=j_m\right)}{\sum_{j_1,\dots,j_{m-1}\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t_1}=n_1, Y_{t_1}=j_1;\dots;N_{t_{m-1}}=n_{m-1}, Y_{t_m}=j_{m-1}\right)}\\
&= \frac{\sum_{j_1,\dots,j_m\in\mathbf{E}} \prod_{k=1}^{m} \widetilde{\mathbb{P}}^i\left(N_{t_k}=n_k, Y_{t_k}=j_k \,|\, N_{t_{k-1}}=n_{k-1}, Y_{t_{k-1}}=j_{k-1}\right)}{\sum_{j_1,\dots,j_{m-1}\in\mathbf{E}} \prod_{k=1}^{m-1} \widetilde{\mathbb{P}}^i\left(N_{t_k}=n_k, Y_{t_k}=j_k \,|\, N_{t_{k-1}}=n_{k-1}, Y_{t_{k-1}}=j_{k-1}\right)}\\
&= \sum_{j_m\in\mathbf{E}} \widetilde{\mathbb{P}}^i\left(N_{t_m}=n_m, Y_{t_m}=j_m \,|\, N_{t_{m-1}}=n_{m-1}, Y_{t_{m-1}}=j_{m-1}\right)\\
&= \widetilde{\mathbb{P}}^i\left(N_{t_m}=n_m \,|\, N_{t_{m-1}}=n_{m-1}, Y_{t_{m-1}}=j_{m-1}\right)\\
&= \widetilde{\mathbb{P}}^i\left(N_{t_m}=n_m \,|\, N_{t_{m-1}}=n_{m-1}\right) = e^{(t_m-t_{m-1})\widetilde{\mathsf{G}}_N}(n_{m-1},n_m),
\end{aligned}
$$

where we used the Markov property of $Z=(N,Y)$ under $\widetilde{\mathbb{P}}^i$ in the third equality, and the result of Step 2 in the last two equalities. The proof is complete. □

Let $\widetilde{\mathbb{F}}^Y=(\widetilde{\mathscr{F}}^Y_t)_{t\geq 0}$ be the filtration generated by Y, and let $\widetilde{\mathscr{F}}^Y_\infty=\sigma(\bigcup_{t\geq 0}\widetilde{\mathscr{F}}^Y_t)$. For each $i\in\mathbf{E}$, we will construct a probability measure $\overline{\mathbb{P}}^i$ on $(\widetilde{\Omega},\widetilde{\mathscr{F}}^Y_\infty)$ such that, the law of Y under $\overline{\mathbb{P}}^i$ is the same as the law of X under $\mathbb{P}^i$. Moreover, we will establish a connection between $\overline{\mathbb{P}}^i$ and $\widetilde{\mathbb{P}}^i$. For this purpose, we first let

$$
S_k := \inf\left\{t\geq 0 \,|\, N_t=k\right\}, \quad k=1,\dots,n.
$$

We will now derive the joint density of N, and $(S_1,\dots,S_n)$ under $\widetilde{\mathbb{P}}^i$. For that, we set

$$
T_1 := S_1, \quad T_k := S_k - S_{k-1}, \quad k=2,\dots,n. \tag{2.8}
$$

It is shown in [Sys92, Section 1.1.4] that T_k's are independent and that

$$
\widetilde{\mathbb{P}}^i\left(T_1>t_1,\dots,T_n>t_n\right)=\prod_{k=1}^{n} e^{-q_k t_k}, \quad t_1,\dots,t_n>0,
$$

which implies that the joint density of $(T_1, \ldots, T_n)$ is given by

$$f_{T_1,\ldots,T_n}(t_1, \ldots, t_n) = \prod_{k=1}^{n} q_k\, e^{-q_k t_k}, \quad t_1, \ldots, t_n > 0. \tag{2.9}$$

Combining (2.8) and (2.9), we deduce that

$$f_{S_1,\ldots,S_n}(s_1, \ldots, s_n) = \prod_{k=1}^{n} q_k\, e^{-q_k(s_k - s_{k-1})}, \quad 0 = s_0 < s_1 < \ldots < s_n.$$

Theorem 2.3.1. *For any $i \in \mathbf{E}$, any $0 < s_1 < \ldots < s_n$, and any cylinder set $A \in \widetilde{\mathscr{F}}^Y_\infty$ of the form*

$$A = \{(Y_{u_1}, \ldots, Y_{u_m}) \in B\}, \quad 0 \le u_1 < u_2 < \ldots < u_m, \quad B \subseteq \mathbf{E}^m, \quad m \in \mathbb{N},$$

the limit

$$\overline{\mathbb{P}}^i (A; s_1, \ldots, s_n) := \lim_{\Delta s_k \to 0, k=1,\ldots,n} \frac{\widetilde{\mathbb{P}}^i (A, s_k < S_k \le s_k + \Delta s_k, k = 1, \ldots, n)}{\widetilde{\mathbb{P}}^i (s_k < S_k \le s_k + \Delta s_k, k = 1, \ldots, n)}, \tag{2.10}$$

exists, and can be extended to a probability measure $\overline{\mathbb{P}}^i(\cdot\,; s_1, \ldots, s_n)$ on $(\widetilde{\Omega}, \widetilde{\mathscr{F}}^Y_\infty)$. Moreover, for any $A \in \widetilde{\mathscr{F}}^Y_\infty$, the function $\overline{\mathbb{P}}^i(A; \ldots)$ is Borel measurable on $\{(s_1, \ldots, s_n) \in \mathbb{R}^n \,|\, 0 < s_1 < \ldots < s_n\}$, and

$$\widetilde{\mathbb{P}}^i(A) = \int_0^\infty \cdots \int_{s_{n-1}}^\infty \overline{\mathbb{P}}^i(A; s_1, \ldots, s_n) \prod_{k=1}^{n} \left(q_k e^{-q_k(s_k - s_{k-1})}\right) \mathrm{d}s_n \cdots \mathrm{d}s_2\, \mathrm{d}s_1. \tag{2.11}$$

In the proof of the theorem we will use the following lemma.

Lemma 2.3.1. *Let us fix $i \in \mathbf{E}$, $0 < s_1 < \ldots < s_n$, and let $0 = k_0 < k_1 < \ldots < k_{n+1}$ be positive integers. In addition, let $0 = u_0 < u_1 < \ldots < u_{k_1} \le s_1 < u_{k_1+1} < \ldots < u_{k_2} \le s_2 < \ldots \le s_n < u_{k_n+1} < \ldots < u_{k_{n+1}}$, $i_0 = i$ and $i_1, \ldots, i_{k_{n+1}} \in \mathbf{E}$. Then, for* *any cylinder set $A \in \widetilde{\mathscr{F}}^Y_\infty$ of the form*

$$A = \bigcap_{j=0}^{n} \left\{ Y_{u_{k_j+1}} = i_{k_j+1}, \dots, Y_{u_{k_{j+1}}} = i_{k_{j+1}} \right\} \tag{2.12}$$

we have

$$\lim_{\Delta s_\ell \to 0,\, \ell=1,\dots,n} \frac{\widetilde{\mathbb{P}}^i \left(A,\, s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\, \ell = 1,\dots,n \right)}{\widetilde{\mathbb{P}}^i \left(s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\, \ell = 1,\dots,n \right)} = \prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\mathsf{G}_\ell}(i_{m-1}, i_m) \right)$$
$$\cdot \sum_{j_1,\dots,j_n \in \mathbf{E}} \prod_{\ell=1}^{n} e^{(s_\ell - u_{k_\ell})\mathsf{G}_{\ell-1}}(i_{k_\ell}, j_\ell) e^{(u_{k_\ell+1} - s_\ell)\mathsf{G}_\ell}(j_\ell, i_{k_\ell+1}). \tag{2.13}$$

In particular, for any $A \in \widetilde{\mathscr{F}}^Y_\infty$ of the form (2.12)*, the above limit is Borel measurable with respect to $(s_1, \dots, s_n)$ in $\Delta_n := \{(s_1, \dots, s_n) \in \mathbb{R}^n \,|\, 0 < s_1 < \dots < s_n\}$.*

Proof. For $\ell = 1, \dots, n$ choose $\Delta s_\ell > 0$ so that, $s_\ell + \Delta s_\ell \le u_{k_\ell+1}$. Then,

$$\begin{aligned}
&\widetilde{\mathbb{P}}^i \left(A,\, s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\, \ell = 1,\dots,n \right) \\
&= \widetilde{\mathbb{P}}^i \Big(Y_{u_{k_\ell+1}} = i_{k_\ell+1}, \dots, Y_{u_{k_{\ell+1}}} = i_{k_{\ell+1}},\, \ell = 0,\dots,n; \\
&\qquad\qquad\qquad N_{s_\ell} = \ell - 1, N_{s_\ell + \Delta s_\ell} = \ell,\, \ell = 1,\dots,n \Big) \\
&= \sum_{j_1,\dots,j_n,\, j'_1,\dots,j'_n \in \mathbf{E}} \widetilde{\mathbb{P}}^i \Big(Z_{u_{k_\ell+1}} = (\ell, i_{k_\ell+1}),\, \dots,\, Z_{u_{k_{\ell+1}}} = (\ell, i_{k_{\ell+1}}),\, \ell = 0,\dots,n; \\
&\qquad\qquad\qquad Z_{s_\ell} = (\ell-1, j_\ell),\, Z_{s_\ell + \Delta s_\ell} = (\ell, j'_\ell),\; \ell = 1,\dots,n \Big) \\
&= \sum_{j_1,\dots,j_n,\, j'_1,\dots,j'_n \in \mathbf{E}} \\
&\quad \left[\prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\widetilde{\mathsf{G}}}((\ell, i_{m-1}), (\ell, i_m)) \right) \right] \left(\prod_{\ell=1}^{n} e^{\Delta s_\ell \widetilde{\mathsf{G}}}((\ell-1, j_\ell), (\ell, j'_\ell)) \right) \\
&\quad \cdot \left(\prod_{\ell=1}^{n} e^{(s_\ell - u_{k_\ell})\widetilde{\mathsf{G}}} \left((\ell-1, i_{k_\ell}), (\ell-1, j_\ell) \right) e^{(u_{k_\ell+1} - s_\ell - \Delta s_\ell)\widetilde{\mathsf{G}}} \left((\ell, j'_\ell), (\ell, i_{k_\ell+1}) \right) \right).
\end{aligned}$$

In the above summation, the first product in the brackets provides the transition probabilities of the evolutions of Z between the times u_{k_ℓ} and $u_{k_{\ell+1}}$, $\ell = 0, \ldots, n$, the second product gives the transition probabilities of the evolutions of Z between the times s_ℓ and $s_\ell + \Delta s_\ell$, for each $\ell = 1, \ldots, n$, and the third product denotes the transition probabilities of the evolutions of Z between the times u_{k_ℓ} and s_ℓ, and between the times $s_\ell + \Delta s_\ell$ and $u_{k_\ell+1}$, for each $\ell = 1, \ldots, n$.

Next, for each $\ell = 1, \ldots, n$,

$$\lim_{\Delta s_\ell \to 0} \frac{1}{\Delta s_\ell} e^{\Delta s_\ell \widetilde{\mathsf{G}}} \left((\ell - 1, j_\ell), (\ell, j'_\ell)\right) = \widetilde{\mathsf{G}} \left((\ell - 1, j_\ell), (\ell, j'_\ell)\right) = \begin{cases} q_\ell, & \text{if } j_\ell = j'_\ell, \\ 0, & \text{otherwise.} \end{cases}$$

Hence,

$$\begin{aligned}
&\lim_{\Delta s_\ell \to 0,\, \ell = 1, \ldots, n} \frac{1}{\Delta s_1 \cdots \Delta s_n} \widetilde{\mathbb{P}}^i \left(A,\, s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\, \ell = 1, \ldots, n\right) \\
&= \prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\widetilde{\mathsf{G}}}((\ell, i_{m-1}), (\ell, i_m)) \right) \\
&\quad \cdot \sum_{j_1, \ldots, j_n \in \mathbf{E}} \prod_{\ell=1}^{n} \Big(q_\ell \, e^{(s_\ell - u_{k_\ell})\widetilde{\mathsf{G}}} \left((\ell - 1, i_{k_\ell}), (\ell - 1, j_\ell)\right) \\
&\qquad\qquad \cdot e^{(u_{k_\ell+1} - s_\ell)\widetilde{\mathsf{G}}} \left((\ell, j_\ell), (\ell, i_{k_\ell+1})\right) \Big).
\end{aligned} \tag{2.14}$$

Note that, for any $j_1, j_2 \in \mathbf{E}$, and any $k \in \mathbb{N}$,

$$\begin{aligned}
&\widetilde{\mathsf{G}}^k \left((\ell, j_1), (\ell, j_2)\right) = (\mathsf{G}_\ell - q_{\ell+1}\mathsf{I})^k (j_1, j_2), \quad \ell = 0, \ldots, n-1, \\
&\widetilde{\mathsf{G}}^k \left((n, j_1), (n, j_2)\right) = \mathsf{G}_n^k (j_1, j_2),
\end{aligned}$$

so that, for $t \ge 0$, we have

$$\begin{aligned}
&e^{t\,\widetilde{\mathsf{G}}} \left((\ell, j_1), (\ell, j_2)\right) = e^{t\,(\mathsf{G}_\ell - q_{\ell+1}\mathsf{I})}(j_1, j_2) = e^{-q_{\ell+1} t}\, e^{t\,\mathsf{G}_\ell}(j_1, j_2), \quad \ell = 0, \ldots, n-1, \\
&e^{t\,\widetilde{\mathsf{G}}} \left((n, j_1), (n, j_2)\right) = e^{t\,\mathsf{G}_n}(j_1, j_2).
\end{aligned}$$

This, together with (2.14), implies that

$$\begin{aligned}
&\lim_{\Delta s_\ell \to 0,\, \ell=1,\ldots,n} \frac{1}{\Delta s_1 \cdots \Delta s_n} \widetilde{\mathbb{P}}^i \left(A,\ s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\ \ell = 1,\ldots,n\right) \\
&= e^{-\sum_{\ell=1}^n q_\ell (u_{k_\ell} - u_{k_{\ell-1}})} \cdot \prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\mathsf{G}_\ell}(i_{m-1}, i_m) \right) \\
&\quad e^{-\sum_{\ell=1}^n q_\ell (s_\ell - u_{k_\ell})} e^{-\sum_{\ell=1}^{n-1} q_\ell (u_{k_\ell+1} - s_\ell)} \\
&\quad \sum_{j_1,\ldots,j_n \in \mathbf{E}} \prod_{\ell=1}^{n} \left(q_\ell e^{(s_\ell - u_{k_\ell})\mathsf{G}_{\ell-1}}(i_{k_\ell}, j_\ell) e^{(u_{k_\ell+1} - s_\ell)\mathsf{G}_\ell}(j_\ell, i_{k_\ell+1}) \right) \\
&= e^{-\sum_{\ell=1}^n q_\ell (s_\ell - s_{\ell-1})} \cdot \prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\mathsf{G}_\ell}(i_{m-1}, i_m) \right) \\
&\quad \cdot \sum_{j_1,\ldots,j_n \in \mathbf{E}} \prod_{\ell=1}^{n} \left(q_\ell\, e^{(s_\ell - u_{k_\ell})\mathsf{G}_{\ell-1}}(i_{k_\ell}, j_\ell)\, e^{(u_{k_\ell+1} - s_\ell)\mathsf{G}_\ell}(j_\ell, i_{k_\ell+1}) \right) \\
&= \left(\prod_{\ell=1}^{n} q_\ell\, e^{-q_\ell (s_\ell - s_{\ell-1})} \right) \cdot \left[\prod_{\ell=0}^{n} \left(\prod_{m=k_\ell+1}^{k_{\ell+1}} e^{(u_m - u_{m-1})\mathsf{G}_\ell}(i_{m-1}, i_m) \right) \right] \\
&\quad \cdot \sum_{j_1,\ldots,j_n \in \mathbf{E}} \prod_{\ell=1}^{n} \left(e^{(s_\ell - u_{k_\ell})\mathsf{G}_{\ell-1}}(i_{k_\ell}, j_\ell)\, e^{(u_{k_\ell+1} - s_\ell)\mathsf{G}_\ell}(j_\ell, i_{k_\ell+1}) \right).
\end{aligned}$$

Finally, in view of the above and the fact that

$$\begin{aligned}
&\lim_{\Delta s_\ell \to 0,\, \ell=1,\ldots,n} \frac{1}{\Delta s_1 \cdots \Delta s_n} \widetilde{\mathbb{P}}^i \left(s_\ell < S_\ell \le s_\ell + \Delta s_\ell,\ \ell = 1,\ldots,n\right) \\
&\qquad = \prod_{\ell=1}^{n} q_\ell\, e^{-q_\ell (s_\ell - s_{\ell-1})},
\end{aligned} \tag{2.15}$$

we obtain (2.13). The proof is complete. □

We are now ready to prove Theorem 2.3.1.

Proof of Theorem 2.3.1. Let $\mathcal{C}$ be the collection of all cylinder sets in $\widetilde{\mathscr{F}}_\infty^Y$ of the form

$$C = \{(Y_{u_1}, \ldots, Y_{u_m}) \in B\}, \quad 0 \le u_1 < u_2 < \ldots < u_m, \quad B \subseteq \mathbf{E}^m, \quad m \in \mathbb{N}.$$

Clearly, $\mathcal{C}$ is an algebra.

We first show that for any $C \in \mathcal{C}$ the limit in (2.10) exists and that an explicit formula for it can be derived. In fact, Lemma 2.3.1 shows that the limit in (2.10) exists, and belongs to $[0, 1]$, for all the cylinder sets of the form (2.12). Thus, for a cylinder set $C \in \mathcal{C}$ an explicit formula for the limit on the right-hand side of (2.10) can be obtained as follows. First, we refine the partition $0 \le u_1 < u_2 < \ldots < u_m$ so that each subinterval of the partition $0 < s_1 < \ldots < s_n$ contains at least one of the u_i's. Clearly, since B_m is finite, A can be decomposed into a finite union of disjoint cylinder sets of the form (2.12) on the refined partition. Moreover, (2.13) provides an explicit formula for the limit in (2.10) for each of those cylinder sets of the form (2.12) on the refined partition. Finally, taking the finite sum over all those limits, we obtain the limit in (2.10) for C. In particular, for every cylinder set C, the limit in (2.10) is Borel measurable with respect to $(s_1, \ldots, s_n)$ in Δ_n.

In the second step we will demonstrate that the limit in (2.10) can be extended to a probability measure on $\sigma(\mathcal{C}) = \widetilde{\mathscr{F}}^Y_\infty$. We start from verifying the countable additivity of $\overline{\mathbb{P}}^i(\cdot\,; s_1, \ldots, s_n)$ on $\mathcal{C}$ for any fixed $0 < s_1 < \ldots < s_n$.

Since $\mathbf{E}$ is a finite set, if $(C_k)_{k \in \mathbb{N}}$ is a sequence of disjoint cylinder sets in $\mathcal{C}$ such that their union also belongs to $\mathcal{C}$, then only finite many of them are non-empty. Therefore, it suffices to verify the finite additivity of $\overline{\mathbb{P}}^i(\cdot\,; s_1, \ldots, s_n)$ on $\mathcal{C}$. Let $C_1, \ldots, C_k \in \mathcal{C}$ be disjoint cylinder sets, then there exists $m \in \mathbb{N}$ and $0 \le u_1 < u_2 < \ldots < u_m$, such that

$$C_\ell = \{(Y_{u_1}, \ldots, Y_{u_m}) \in B_\ell\} \quad \text{for some} \;\; B_\ell \subseteq \mathbf{E}^m, \quad \ell = 1, \ldots, k.$$

Each $\overline{\mathbb{P}}^i(C_\ell\,; s_1, \ldots, s_n)$ can be represented as

$$\overline{\mathbb{P}}^i(C_\ell\,; s_1, \ldots, s_n) = \sum_{A_\ell \in \mathcal{C}_\ell} \overline{\mathbb{P}}^i(A_\ell\,; s_1, \ldots, s_n), \quad j = 1, \ldots, k,$$

where $\mathcal{C}_\ell$, $\ell = 1, \ldots, k$, are disjoint classes of disjoint simple cylinder sets. Therefore, we have

$$\begin{aligned}\sum_{\ell=1}^{k} \overline{\mathbb{P}}^i(C_\ell\,; s_1, \ldots, s_n) &= \sum_{\ell=1}^{k} \sum_{A_\ell \in \mathcal{C}_\ell} \overline{\mathbb{P}}^i(A_\ell\,; s_1, \ldots, s_n) \\ &= \sum_{A \in \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_k} \overline{\mathbb{P}}^i(A\,; s_1, \ldots, s_n) = \overline{\mathbb{P}}^i\left(\bigcup_{\ell=1}^{k} C_\ell\,; s_1, \ldots, s_n\right).\end{aligned}$$

Note that for any $0 < s_1 < \ldots < s_n$, $\overline{\mathbb{P}}^i(C\,; s_1, \ldots, s_n) \leq 1$ for all $C \in \mathcal{C}$. By the Carathéodory extension theorem, for any $0 < s_1 < \ldots < s_n$, $\overline{\mathbb{P}}^i(\cdot\,; s_1, \ldots, s_n)$ can be uniquely extended to a probability measure on $(\widetilde{\Omega}, \widetilde{\mathscr{F}}^Y_\infty)$.

Let $\Delta_n := \{(s_1, \ldots, s_n) \in \mathbb{R}^n \,|\, 0 < s_1 < \ldots < s_n\}$ and

$$\mathcal{D}_1 := \left\{ A \in \widetilde{\mathscr{F}}^Y_\infty \,\middle|\, \overline{\mathbb{P}}^i(A\,; \cdot, \cdots, \cdot) \text{ is Borel measurable on } \Delta_n \right\}.$$

We will show that $\mathcal{D}_1 = \widetilde{\mathscr{F}}^Y_\infty$. Towards this end, we first observe that (2.10) and (2.13) imply that, for any $A \in \mathcal{C}$, $\overline{\mathbb{P}}^i(A\,; \cdot, \cdots, \cdot)$ is Borel measurable with respect to $(s_1, \ldots, s_n)$ on Δ_n, and thus $\mathcal{D}_1 \supset \mathcal{C}$.

Next, we will show that $\mathcal{D}_1$ is a monotone class. For this, let $(A_k)_{k \in \mathbb{N}} \subset \mathcal{D}_1$ be an increasing sequence of events, so that, for any $0 < s_1 < \ldots < s_n$, we have

$$\overline{\mathbb{P}}^i\left(\bigcup_{k=1}^{\infty} A_k\,;\, s_1, \ldots, s_n\right) = \lim_{m \to \infty} \overline{\mathbb{P}}^i\left(A_m\,;\, s_1, \ldots, s_n\right).$$

Thus, $\overline{\mathbb{P}}^i(\cup_k A_k\,; \cdot, \cdots, \cdot)$, being a limit of a sequence of Borel measurable functions on

Δ_n, is Borel measurable on Δ_n, and hence $\cup_k A_k \in \mathcal{D}_1$. Similarly, one can show that if $(A_k)_{k\in\mathbb{N}} \subset \mathcal{D}_1$ is a decreasing sequence of events, then $\cap_k A_k \in \mathcal{D}_1$. Therefore, $\mathcal{D}_1$ is a monotone class, and by the monotone class theorem $\mathcal{D}_1 = \sigma(\mathcal{C}) = \widetilde{\mathscr{F}}^Y_\infty$.

It remains to show that (2.11) holds true. In view of (2.10) and (2.15), for any cylinder set $A \in \mathcal{C}$,

$$
\begin{aligned}
\overline{\mathbb{P}}^i(A\,; s_1, \ldots, s_n &= \lim_{\Delta s_k \to 0,\, k=1,\ldots,n} \frac{\widetilde{\mathbb{P}}^i\left(A,\ s_k < S_k \le s_k + \Delta s_k,\ k = 1, \ldots, n\right)}{\widetilde{\mathbb{P}}^i\left(s_k < S_k \le s_k + \Delta s_k,\ k = 1, \ldots, n\right)} \\
&= \frac{\lim_{\Delta s_k \to 0,\, k=1,\ldots,n} (\Delta s_1 \cdots \Delta s_n)^{-1}\, \widetilde{\mathbb{P}}^i\left(A,\ s_k < S_k \le s_k + \Delta s_k,\ k = 1, \ldots, n\right)}{\lim_{\Delta s_k \to 0,\, k=1,\ldots,n} (\Delta s_1 \cdots \Delta s_n)^{-1}\, \widetilde{\mathbb{P}}^i\left(s_k < S_k \le s_k + \Delta s_k,\ k = 1, \ldots, n\right)} \\
&= \frac{\partial^n}{\partial s_1 \cdots \partial s_n} \widetilde{\mathbb{P}}^i\left(A,\ S_k \le s_k,\ k = 1, \ldots, n\right) \cdot \left(\prod_{k=1}^n q_k\, e^{-q_k(s_k - s_{k-1})} \right)^{-1}.
\end{aligned}
$$

Hence, for any $A \in \mathcal{C}$,

$$
\begin{aligned}
&\int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty \overline{\mathbb{P}}^i(A; s_1, \ldots, s_n) \prod_{k=1}^n q_k\, e^{-q_k(s_k - s_{k-1})}\, \mathrm{d}s_1 \cdots \mathrm{d}s_n \\
&\quad = \int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty \frac{\partial^n}{\partial s_1 \cdots \partial s_n} \widetilde{\mathbb{P}}^i\left(A,\ S_k \le s_k,\ k = 1, \ldots, n\right) \mathrm{d}s_1 \cdots \mathrm{d}s_n = \widetilde{\mathbb{P}}^i(A),
\end{aligned}
$$

and thus $\mathcal{C} \subset \mathcal{D}_2$, where $\mathcal{D}_2 := \left\{ A \in \widetilde{\mathscr{F}}^Y_\infty \,\middle|\, (2.11) \text{ holds for } A \right\}$. Next, for any increasing sequence of events $(A_k)_{k\in\mathbb{N}} \subset \mathcal{D}_2$, we have that

$$
\begin{aligned}
\widetilde{\mathbb{P}}^i\left(\bigcup_{k=1}^\infty A_k \right) &= \lim_{k\to\infty} \widetilde{\mathbb{P}}^i(A_k) \\
&= \lim_{k\to\infty} \int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty \overline{\mathbb{P}}^i(A_k; s_1, \ldots, s_n) \prod_{\ell=1}^n q_\ell\, e^{-q_\ell(s_\ell - s_{\ell-1})}\, \mathrm{d}s_1 \cdots \mathrm{d}s_n \\
&= \int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty \overline{\mathbb{P}}^i\left(\bigcup_{k=1}^\infty A_k; s_1, \ldots, s_n \right) \prod_{\ell=1}^n q_\ell\, e^{-q_\ell(s_\ell - s_{\ell-1})}\, \mathrm{d}s_1 \cdots \mathrm{d}s_n,
\end{aligned}
$$

where the last equality follows from the dominated convergence theorem as well as the fact that $\overline{\mathbb{P}}^i(A_k; s_1, \ldots, s_n) \le 1$, for all $k \in \mathbb{N}$ and $0 < s_1 < \ldots < s_n$. Hence,

$\cup_k A_k \in \mathcal{D}_2$. Similarly, one can show that if $(A_k)_{k\in\mathbb{N}} \subset \mathcal{D}_2$ is a decreasing sequence, then $\cap_k A_k \in \mathcal{D}_2$. Therefore, $\mathcal{D}_2$ is a monotone class, and by the monotone class theorem $\mathcal{D}_2 = \sigma(\mathcal{C}) = \widetilde{\mathscr{F}}^Y_\infty$. This completes the proof. $\square$

Next, we will prove that the law of Y under $\overline{\mathbb{P}}^i$ is the same as that of X under $\mathbb{P}^i$. As usual, $\overline{\mathbb{E}}^i(\cdot\,; s_1,\ldots,s_n)$ will denote the expectation associated with $\overline{\mathbb{P}}^i(\cdot\,; s_1,\ldots,s_n)$, for $i \in \mathbf{E}$ and $0 < s_1 < \ldots < s_n$. In the sequel, if there is no ambiguity, we will omit the parameters $s_1,\ldots,s_n$ in $\overline{\mathbb{P}}^i$ and $\overline{\mathbb{E}}^i$.

Theorem 2.3.2. *For any $i \in \mathbf{E}$ and $0 < s_1 < \ldots < s_n$, under $\overline{\mathbb{P}}^i$, Y is a time-inhomogeneous Markov chain with generator $\mathsf{G} = \{\mathsf{G}_t, t \geq 0\}$. In particular, X and Y have the same law under respective probability measures $\mathbb{P}^i$ and $\overline{\mathbb{P}}^i$.*

Proof. Let $u_0, u_1, \ldots, u_m$ be such that

$$\begin{aligned} 0 = u_0 \leq u_1 < \ldots < u_{k_1} \leq s_1 < u_{k_1+1} < \ldots < u_{k_2} \\ \leq s_2 < \ldots \leq s_n < u_{k_n+1} < \ldots < u_{k_{n+1}} = u_m. \end{aligned}$$

By (2.13), for any $i_1,\ldots,i_m \in \mathbf{E}$,

$$\begin{aligned} \overline{\mathbb{P}}^i(Y_{u_m} = i_m \mid Y_{u_{m-1}} = i_{m-1},\ldots,Y_{u_1} = i_1) &= \frac{\overline{\mathbb{P}}^i\left(Y_{u_1} = i_1,\ldots,Y_{u_m} = i_m\right)}{\overline{\mathbb{P}}^i\left(Y_{u_1} = i_1,\ldots,Y_{u_{m-1}} = i_{m-1}\right)} \\ &= \frac{\prod_{\ell=0}^{n}\left(\prod_{p=k_\ell+1}^{k_{\ell+1}} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1},i_p)\right)}{\left[\prod_{\ell=0}^{n-1}\left(\prod_{p=k_\ell+1}^{k_{\ell+1}} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1},i_p)\right)\right]\left(\prod_{p=k_n+1}^{k_{n+1}-1} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1},i_p)\right)} \\ &= e^{(u_m-u_{m-1})\mathsf{G}_n}(i_{m-1},i_m). \end{aligned}$$

On the other hand, by (2.13) again,

$$\overline{\mathbb{P}}^i(Y_{u_m} = i_m \mid Y_{u_{m-1}} = i_{m-1}) = \frac{\overline{\mathbb{P}}^i\left(Y_{u_m} = i_m\, Y_{u_{m-1}} = i_{m-1}\right)}{\overline{\mathbb{P}}^i\left(Y_{u_{m-1}} = i_{m-1}\right)}$$

$$= \frac{\sum_{i_1,\ldots,i_{m-2}\in\mathbf{E}} \overline{\mathbb{P}}^i \left(Y_{u_1} = i_1, \ldots, Y_{u_m} = i_m\right)}{\sum_{i_1,\ldots,i_{m-2}\in\mathbf{E}} \overline{\mathbb{P}}^i \left(Y_{u_1} = i_1, \ldots, Y_{u_{m-1}} = i_{m-1}\right)}$$

$$= \frac{\sum_{i_1,\ldots,i_{m-2}\in\mathbf{E}} \prod_{\ell=0}^{n} \left(\prod_{p=k_\ell+1}^{k_{\ell+1}} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1}, i_p)\right)}{\sum_{i_1,\ldots,i_{m-2}\in\mathbf{E}} \left[\prod_{\ell=0}^{n-1} \left(\prod_{p=k_\ell+1}^{k_{\ell+1}} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1}, i_p)\right)\right] \left(\prod_{p=k_n+1}^{k_{n+1}-1} e^{(u_p-u_{p-1})\mathsf{G}_\ell}(i_{p-1}, i_p)\right)}$$

$$= e^{(u_m-u_{m-1})\mathsf{G}_n}(i_{m-1}, i_m).$$

Analogous argument carries for any $u_0 < u_1 < \ldots < u_m$, which completes the proof.

□

In analogy to φ_t and τ_t^+ we now define an additive functional ψ given as $\psi_t := \int_0^t v(Y_u)\,\mathrm{d}u,\ t \geq 0$, and we consider the following first passage time $\rho_t^+ := \inf\{r \geq 0 \,|\, \psi_r > t\},\ t \geq 0$.

We end this part of this section with the following corollary to Theorem 2.3.2.

Corollary 2.3.1. *For any $(s_1, \ldots, s_n)$ in Δ_n, $c > 0$, and $t > 0$,*

$$\Pi_c^+(i, j; s_1, \ldots, s_n) = \overline{\mathbb{E}}^i \left(e^{-c\rho_0^+} \mathbb{1}_{\{Y_{\rho_0^+}=j\}}; s_1, \ldots, s_n\right), \quad i \in \mathbf{E}^-,\ j \in \mathbf{E}^+, \tag{2.16}$$

$$\Psi_c^+(t, i, j; s_1, \ldots, s_n) = \overline{\mathbb{E}}^i \left(e^{-c\rho_t^+} \mathbb{1}_{\{Y_{\rho_t^+}=j\}}; s_1, \ldots, s_n\right), \quad i \in \mathbf{E}^+,\ j \in \mathbf{E}^+. \tag{2.17}$$

In particular, $\Pi_c^+(i, j; s_1, \ldots, s_n)$ and $\Psi_c^+(t, i, j; s_1, \ldots, s_n)$ are Borel measurable with respect to $(s_1, \ldots, s_n)$ in Δ_n.

2.3.2 Wiener-Hopf Factorization for Time Homogenized Process. This subsection is devoted to computing the expectations on the right-hand side in (2.16) and (2.17). This will be done by computing the corresponding expectations related to the *time-homogeneous* Markov chain $Z = (N, Y)$. The latter computation will be done using the classical Wiener-Hopf factorization results for finite state time-homogeneous Markov chains, originally derived in [BRW80].

We begin with a restatement of the classical Wiener-Hopf factorization applied to Z. Towards this end, we let $\widetilde{\mathbf{E}}^{+} := \mathbb{N}_n \times \mathbf{E}^{+}$ and $\widetilde{\mathbf{E}}^{-} := \mathbb{N}_n \times \mathbf{E}^{-}$, and $\widetilde{v} : \widetilde{\mathbf{E}} \to \mathbb{R}\backslash\{0\}$ be a function on $\widetilde{\mathbf{E}}$ such that $\widetilde{v}(k,i) = v(i)$, for all $(k,i) \in \widetilde{\mathbf{E}}$. Next, we define the additive functional $\widetilde{\varphi}$ and the corresponding first passage times as

$$\widetilde{\varphi}_t := \int_0^t \widetilde{v}(Z_u)\,\mathrm{d}u, \quad \widetilde{\tau}_t^{\pm} := \inf\left\{r \geq 0 \,|\, \pm\widetilde{\varphi}_r > t\right\}, \quad t \geq 0.$$

Let $\widetilde{V} := diag\{\widetilde{v}(k,i) : (k,i) \in \widetilde{\mathbf{E}}\}$ (a diagonal matrix). We denote by $\widetilde{\mathsf{I}}^{\pm}$ the identity matrix of dimension $|\widetilde{\mathbf{E}}^{\pm}|$. Finally, $\mathcal{Q}(m)$ will stand for the set of $m \times m$ generator matrices (i.e., matrices with non-negative off-diagonal entries and non-positive row sums), and $\mathcal{P}(m,\ell)$ will be the set of $m \times \ell$ matrices whose rows are sub-probability vectors.

Theorem 2.3.3. *[BRW80, Theorem 1 & 2] Fix $c > 0$. Then,*

(i) there exists a unique quadruple of matrices $(\widetilde{\Lambda}_c^{+}, \widetilde{\Lambda}_c^{-}, \widetilde{\mathsf{G}}_c^{+}, \widetilde{\mathsf{G}}_c^{-})$, where $\widetilde{\Lambda}_c^{+} \in \mathcal{P}(|\widetilde{\mathbf{E}}^{-}|, |\widetilde{\mathbf{E}}^{+}|)$, $\widetilde{\Lambda}_c^{-} \in \mathcal{P}(|\widetilde{\mathbb{E}}^{+}|, |\widetilde{\mathbf{E}}^{-}|)$, $\widetilde{\mathsf{G}}_c^{+} \in \mathcal{Q}(|\widetilde{\mathbf{E}}^{+}|)$, and $\widetilde{\mathsf{G}}_c^{-} \in \mathcal{Q}(|\widetilde{\mathbf{E}}^{-}|)$, such that

$$\widetilde{V}^{-1}\left(\widetilde{\mathsf{G}} - c\widetilde{\mathsf{I}}\right)\begin{pmatrix} \widetilde{\mathsf{I}}^{+} & \widetilde{\Lambda}_c^{-} \\ \widetilde{\Lambda}_c^{+} & \widetilde{\mathsf{I}}^{-} \end{pmatrix} = \begin{pmatrix} \widetilde{\mathsf{I}}^{+} & \widetilde{\Lambda}_c^{-} \\ \widetilde{\Lambda}_c^{+} & \widetilde{\mathsf{I}}^{-} \end{pmatrix}\begin{pmatrix} \widetilde{\mathsf{G}}_c^{+} & 0 \\ 0 & -\widetilde{\mathsf{G}}_c^{-} \end{pmatrix}; \quad (2.18)$$

(ii) the matrices $\widetilde{\Lambda}_c^{+}$, $\widetilde{\Lambda}_c^{-}$, $\widetilde{\mathsf{G}}_c^{+}$, and $\widetilde{\mathsf{G}}_c^{-}$, admit the following probabilistic representations,

$$\widetilde{\Lambda}_c^{+}((k,i),(\ell,j)) = \widetilde{\mathbb{E}}\left(e^{-c\widetilde{\tau}_0^{+}}\mathbf{1}_{\{Z_{\widetilde{\tau}_0^{+}}=(\ell,j)\}}\,\middle|\, Z_0 = (k,i)\right), \quad (k,i) \in \widetilde{\mathbf{E}}^{-}, (\ell,j) \in \widetilde{\mathbf{E}}^{+},$$

$$\widetilde{\Lambda}_c^-((k,i),(\ell,j)) = \widetilde{\mathbb{E}}\left(e^{-c\widetilde{\tau}_0^-}\mathbf{1}_{\{Z_{\widetilde{\tau}_0^-}=(\ell,j)\}}\,\Big|\, Z_0=(k,i)\right), \quad (k,i)\in\widetilde{\mathbf{E}}^+,\ (\ell,j)\in\widetilde{\mathbf{E}}^-,$$

$$e^{t\,\widetilde{\mathsf{G}}_c^+}((k,i),(\ell,j)) = \widetilde{\mathbb{E}}\left(e^{-c\widetilde{\tau}_t^+}\mathbf{1}_{\{Z_{\widetilde{\tau}_t^+}=(\ell,j)\}}\,\Big|\, Z_0=(k,i)\right), \quad (k,i)\in\widetilde{\mathbf{E}}^+,\ (\ell,j)\in\widetilde{\mathbf{E}}^+,$$

$$e^{t\,\widetilde{\mathsf{G}}_c^-}((k,i),(\ell,j)) = \widetilde{\mathbb{E}}\left(e^{-c\widetilde{\tau}_t^-}\mathbf{1}_{\{Z_{\widetilde{\tau}_t^-}=(\ell,j)\}}\,\Big|\, Z_0=(k,i)\right), \quad (k,i)\in\widetilde{\mathbf{E}}^-,\ (\ell,j)\in\widetilde{\mathbf{E}}^-,$$

for any $t\geq 0$.

In what follows we will use the "+" part of the above formulas and only for $k=0$. Accordingly, we define (recall (2.5))

$$\widetilde{\Pi}_c^+(i,j,\ell) := \widetilde{\Lambda}_c^+((0,i),(\ell,j)) = \widetilde{\mathbb{E}}^i\left(e^{-c\widetilde{\tau}_0^+}\mathbf{1}_{\{Z_{\widetilde{\tau}_0^+}=(\ell,j)\}}\right), \quad i\in\mathbf{E}^-,\ j\in\mathbf{E}^+,\ \ell\in\mathbb{N}, \tag{2.19}$$

$$\widetilde{\Psi}_c^+(t,i,j,\ell) := e^{t\,\widetilde{\mathsf{G}}_c^+}((0,i),(\ell,j)) = \widetilde{\mathbb{E}}^i\left(e^{-c\widetilde{\tau}_t^+}\mathbf{1}_{\{Z_{\widetilde{\tau}_t^+}=(\ell,j)\}}\right), \quad i,j\in\mathbf{E}^+,\ \ell\in\mathbb{N},\ t\geq 0. \tag{2.20}$$

Note that, for any $t\geq 0$, $\widetilde{v}(Z_t)=v(Y_t)$, which implies that $\widetilde{\varphi}_t=\psi_t$, and so $\rho_t^+=\widetilde{\tau}_t^+$, $\rho_t^-=\widetilde{\tau}_t^-$. Hence, by taking summations over all $\ell\in\mathbb{N}$ in (2.19) and (2.20), we obtain that

$$\widetilde{\mathbb{E}}^i\left(e^{-c\rho_0^+}\mathbb{1}_{\{Y_{\rho_0^+}=j\}}\right) = \sum_{\ell=0}^{n}\widetilde{\Pi}_c^+(i,j,\ell), \quad i\in\mathbf{E}^-,\ j\in\mathbf{E}^+, \tag{2.21}$$

$$\widetilde{\mathbb{E}}^i\left(e^{-c\rho_t^+}\mathbb{1}_{\{Y_{\rho_t^+}=j\}}\right) = \sum_{\ell=0}^{n}\widetilde{\Psi}_c^+(t,i,j,\ell), \quad i,j\in\mathbf{E}^+,\ t\geq 0. \tag{2.22}$$

Observe that, in view of (2.11), if $U:\widetilde{\Omega}\to\mathbb{R}$ is an $\widetilde{\mathscr{F}}_\infty^Y$-measurable bounded random variable, then for any $i \in \mathbf{E}$,

$$\widetilde{\mathbb{E}}^i(U) = \int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty \overline{\mathbb{E}}^i(U; s_1, \ldots, s_n) \prod_{k=1}^n \left(q_k \, e^{-q_k(s_k - s_{k-1})} \right) \mathrm{d}s_n \cdots \mathrm{d}s_2 \, \mathrm{d}s_1.$$

Therefore, in light of Corollary 2.3.1, (2.21) and (2.22), we have that

$$\begin{aligned} \widehat{\Pi}_c^+(i, j; q_1, \ldots, q_n) &:= \sum_{\ell=0}^n \widetilde{\Pi}_c^+(i, j, \ell) \\ &= \int_0^\infty \int_{s_1}^\infty \ldots \int_{s_{n-1}}^\infty \Pi_c^+(i, j; s_1, \ldots, s_n) \prod_{k=1}^n \left(q_k e^{-q_k(s_k - s_{k-1})} \right) \mathrm{d}s_n \cdots \mathrm{d}s_2 \, \mathrm{d}s_1. \end{aligned}$$

$$\begin{aligned} \widehat{\Psi}_c^+(t, i, j; q_1, \ldots, q_n) &:= \sum_{\ell=0}^n \widetilde{\Psi}_c^+(t, i, j, \ell) \\ &= \int_0^\infty \int_{s_1}^\infty \ldots \int_{s_{n-1}}^\infty \Psi_c^+(t, i, j; s_1, \ldots, s_n) \prod_{k=1}^n \left(q_k e^{-q_k(s_k - s_{k-1})} \right) \mathrm{d}s_n \cdots \mathrm{d}s_2 \, \mathrm{d}s_1. \end{aligned}$$

By change of variables, we obtain

$$\begin{aligned} \widehat{\Pi}_c^+(i, j; q_1, \ldots, q_n) &= \int_0^\infty \ldots \int_0^\infty \Pi_c^+(i, j; t_1, \ldots, t_1 + \ldots + t_n) \prod_{k=1}^n \left(q_k e^{-q_k t_k} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_n, \\ \widehat{\Psi}_c^+(i, j; q_1, \ldots, q_n) &= \int_0^\infty \ldots \int_0^\infty \Psi_c^+(i, j; t_1, \ldots, t_1 + \ldots + t_n) \prod_{k=1}^n \left(q_k e^{-q_k t_k} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_n. \end{aligned}$$

The above two equalities together with the argument in Appendix A, implies that

$$q_1^{-1} \cdots q_n^{-1} \widehat{\Pi}_c^+(i, j; q_1, \ldots, q_n), \qquad q_1^{-1} \cdots q_n^{-1} \widehat{\Psi}_c^+(i, j; q_1, \ldots, q_n)$$

are well-defined for $q_k \in \mathbb{C}^+ := \{z \in \mathbb{C} \mid \Re(z) > 0\}$, $k = 1, \ldots, n$, as being the Laplace transforms of $\Pi_c^+(i, j; t_1, \ldots, t_1 + \ldots + t_n)$ and $\Psi_c^+(i, j; t_1, \ldots, t_1 + \ldots + t_n)$, respectively.

2.3.3 Main Result. All the above leads to the following result, which is our main theorem, and where we make use of the inverse multivariate Laplace transform. We refer to the Appendix for the definition and the properties of the inverse multivariate Laplace transform relevant to our set-up.

Theorem 2.3.4. *We have that*

$$\Pi_c^+(i,j;s_1,\ldots,s_n) \\ = \mathcal{L}^{-1}\left(q_1^{-1}\cdots q_n^{-1}\widehat{\Pi}_c^+(i,j;q_1,\ldots,q_n)\right)(s_1,s_2-s_1,\ldots,s_n-s_{n-1}),$$

for any $i \in \mathbf{E}^-$, $j \in \mathbf{E}^+$, and

$$\Psi_c^+(t,i,j;s_1,\ldots,s_n) \\ = \mathcal{L}^{-1}\left(q_1^{-1}\cdots q_n^{-1}\widehat{\Psi}_c^+(t,i,j;q_1,\ldots,q_n)\right)(s_1,s_2-s_1,\ldots,s_n-s_{n-1}),$$

for any $t > 0$, $i,j \in \mathbf{E}^+$, where $\mathcal{L}^{-1}$ is the inverse multivariate Laplace transform.

Remark 2.3.1. *It needs to be stressed that we can compute the values of $\widehat{\Pi}_c^+(i,j;q_1,\ldots,q_n)$ and $\widehat{\Psi}_c^+(t,i,j;q_1,\ldots,q_n)$ only for positive values of q_i's. Thus, Theorem 2.3.4 may not be directly applied to compute $\Pi_c^+(i,j;s_1,\ldots,s_n)$ and $\Psi_c^+(t,i,j;s_1,\ldots,s_n)$. However, we can approximate these functions, as explained in Appendix A by using only the values of $\widehat{\Pi}_c^+(i,j;q_1,\ldots,q_n)$ and $\widehat{\Psi}_c^+(t,i,j;q_1,\ldots,q_n)$ for positive values of q_i's.*

2.4 A Computational Method for WHf

In this section we will illustrate our theoretical results with a simple, but telling example. We first describe a numerical method to approximate Π_c^+ and Ψ_c^+, and then we proceed with its application to a concrete example.

2.4.1 Approximation of Π_c^+ and Ψ_c^+. We only consider Π_c^+. The procedure to approximate Ψ_c^+ is analogous.

According to Theorem 2.3.4 and Appendix A, to approximate Π_c^+, we need to compute $\widehat{\Pi}_c^+(i,j;q_1,\ldots,q_n)$ for any $q_1,\ldots,q_n > 0$, and then to use the Gaver-Stehfest algorithm. Note that $\widehat{\Pi}_c^+(i,j;q_1,\ldots,q_n)$ can be computed by solving (2.18) directly using the diagonalization method of [RS94]. However, because of the special structure of $\widetilde{\mathsf{G}}$, we can simplify the calculation by working on matrices of smaller dimensions. Towards this end we observe that matrices in (2.18) can be written the block form as follows,

$$
\widetilde{\mathsf{G}} =
\begin{array}{cc}
 & \begin{array}{cccccccc} (0,\mathbf{E}^+) & (1,\mathbf{E}^+) & \cdots & (n,\mathbf{E}^+) & (0,\mathbf{E}^-) & (1,\mathbf{E}^-) & \cdots & (n,\mathbf{E}^-) \end{array} \\
\begin{array}{c} (0,\mathbf{E}^+) \\ (1,\mathbf{E}^+) \\ \vdots \\ (n-1,\mathbf{E}^+) \\ (n,\mathbf{E}^+) \\ (0,\mathbf{E}^-) \\ (1,\mathbf{E}^-) \\ \vdots \\ (n-1,\mathbf{E}^-) \\ (n,\mathbf{E}^-) \end{array} &
\left[\begin{array}{cccccccc}
\mathsf{A}_1 - q_1\mathsf{I}^+ & q_1\mathsf{I}^+ & \cdots & 0 & \mathsf{B}_1 & 0 & \cdots & 0 \\
0 & \mathsf{A}_2 - q_2\mathsf{I}^+ & \cdots & 0 & 0 & \mathsf{B}_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & q_n\mathsf{I}^+ & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \mathsf{A}_{n+1} & 0 & 0 & \cdots & \mathsf{B}_{n+1} \\
\mathsf{C}_1 & 0 & \cdots & 0 & \mathsf{D}_1 - q_1\mathsf{I}^- & q_1\mathsf{I}^- & \cdots & 0 \\
0 & \mathsf{C}_2 & \cdots & 0 & 0 & \mathsf{D}_2 - q_2\mathsf{I}^- & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & q_n\mathsf{I}^- \\
0 & 0 & \cdots & \mathsf{C}_{n+1} & 0 & 0 & \cdots & \mathsf{D}_{n+1}
\end{array}\right]
\end{array},
$$

$$\widetilde{\mathsf{V}} = \begin{array}{c} \\ {\scriptstyle (0,\mathbf{E}^+)} \\ {\scriptstyle (1,\mathbf{E}^+)} \\ \vdots \\ {\scriptstyle (n-1,\mathbf{E}^+)} \\ {\scriptstyle (n,\mathbf{E}^+)} \\ {\scriptstyle (0,\mathbf{E}^-)} \\ {\scriptstyle (1,\mathbf{E}^-)} \\ \vdots \\ {\scriptstyle (n-1,\mathbf{E}^-)} \\ {\scriptstyle (n,\mathbf{E}^-)} \end{array} \begin{array}{c} \begin{array}{cccccccc} {\scriptstyle (0,\mathbf{E}^+)} & {\scriptstyle (1,\mathbf{E}^+)} & {\scriptstyle \cdots} & {\scriptstyle (n,\mathbf{E}^+)} & {\scriptstyle (0,\mathbf{E}^-)} & {\scriptstyle (1,\mathbf{E}^-)} & {\scriptstyle \cdots} & {\scriptstyle (n,\mathbf{E}^-)} \end{array} \\ \begin{bmatrix} \mathsf{V}^+ & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \mathsf{V}^+ & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \mathsf{V}^+ & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \mathsf{V}^- & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \mathsf{V}^- & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \mathsf{V}^- \end{bmatrix} \end{array}, \quad (2.23)$$

$$\widetilde{\Lambda}_c^+ = \begin{array}{c} \\ {\scriptstyle (0,\mathbf{E}^-)} \\ {\scriptstyle (1,\mathbf{E}^-)} \\ \vdots \\ {\scriptstyle (n-1,\mathbf{E}^-)} \\ {\scriptstyle (n,\mathbf{E}^-)} \end{array} \begin{array}{c} \begin{array}{ccccc} {\scriptstyle (0,\mathbf{E}^+)} & {\scriptstyle (1,\mathbf{E}^+)} & {\scriptstyle \cdots} & {\scriptstyle (n-1,\mathbf{E}^+)} & {\scriptstyle (n,\mathbf{E}^+)} \end{array} \\ \begin{bmatrix} \widetilde{\Lambda}^+_{c,00} & \widetilde{\Lambda}^+_{c,01} & \cdots & \widetilde{\Lambda}^+_{c,0,n-1} & \widetilde{\Lambda}^+_{c,0n} \\ 0 & \widetilde{\Lambda}^+_{c,11} & \cdots & \widetilde{\Lambda}^+_{c,1,n-1} & \widetilde{\Lambda}^+_{c,1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \widetilde{\Lambda}^+_{c,n-1,n-1} & \widetilde{\Lambda}^+_{c,n-1,n} \\ 0 & 0 & \cdots & 0 & \widetilde{\Lambda}^+_{c,nn} \end{bmatrix} \end{array}$$

$$
\widetilde{\Lambda}_c^- = \begin{array}{c} \\ (0,\mathbf{E}^+) \\ (1,\mathbf{E}^+) \\ \vdots \\ (n-1,\mathbf{E}^+) \\ (n,\mathbf{E}^+) \end{array}
\begin{array}{c}
\begin{array}{ccccc} (0,\mathbf{E}^-) & (1,\mathbf{E}^-) & \cdots & (n-1,\mathbf{E}^-) & (n,\mathbf{E}^-) \end{array} \\
\begin{bmatrix}
\widetilde{\Lambda}_{c,00}^- & \widetilde{\Lambda}_{c,01}^- & \cdots & \widetilde{\Lambda}_{c,0,n-1}^- & \widetilde{\Lambda}_{c,0n}^- \\
0 & \widetilde{\Lambda}_{c,11}^- & \cdots & \widetilde{\Lambda}_{c,1,n-1}^- & \widetilde{\Lambda}_{c,1n}^- \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \widetilde{\Lambda}_{c,n-1,n-1}^- & \widetilde{\Lambda}_{c,n-1,n}^- \\
0 & 0 & \cdots & 0 & \widetilde{\Lambda}_{c,nn}^-
\end{bmatrix}
\end{array},
$$

$$
\widetilde{\mathsf{G}}_c^+ = \begin{array}{c} \\ (0,\mathbf{E}^+) \\ (1,\mathbf{E}^+) \\ \vdots \\ (n-1,\mathbf{E}^+) \\ (n,\mathbf{E}^+) \end{array}
\begin{array}{c}
\begin{array}{ccccc} (0,\mathbf{E}^+) & (1,\mathbf{E}^+) & \cdots & (n-1,\mathbf{E}^+) & (n,\mathbf{E}^+) \end{array} \\
\begin{bmatrix}
\widetilde{\mathsf{G}}_{c,00}^+ & \widetilde{\mathsf{G}}_{c,01}^+ & \cdots & \widetilde{\mathsf{G}}_{c,0,n-1}^+ & \widetilde{\mathsf{G}}_{c,0n}^+ \\
0 & \widetilde{\mathsf{G}}_{c,11}^+ & \cdots & \widetilde{\mathsf{G}}_{c,1,n-1}^+ & \widetilde{\mathsf{G}}_{c,1n}^+ \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \widetilde{\mathsf{G}}_{c,n-1,n-1}^+ & \widetilde{\mathsf{G}}_{c,n-1,n}^+ \\
0 & 0 & \cdots & 0 & \widetilde{\mathsf{G}}_{c,nn}^+
\end{bmatrix}
\end{array},
$$

and

$$
\widetilde{\mathsf{G}}_c^- = \begin{array}{c} \\ (0,\mathbf{E}^-) \\ (1,\mathbf{E}^-) \\ \vdots \\ (n-1,\mathbf{E}^-) \\ (n,\mathbf{E}^-) \end{array} \begin{array}{c} \begin{array}{ccccc} (0,\mathbf{E}^-) & (1,\mathbf{E}^-) & \cdots & (n-1,\mathbf{E}^-) & (n,\mathbf{E}^-) \end{array} \\ \begin{bmatrix} \widetilde{\mathsf{G}}_{c,00}^- & \widetilde{\mathsf{G}}_{c,01}^- & \cdots & \widetilde{\mathsf{G}}_{c,0,n-1}^- & \widetilde{\mathsf{G}}_{c,0n}^- \\ 0 & \widetilde{\mathsf{G}}_{c,11}^- & \cdots & \widetilde{\mathsf{G}}_{c,1,n-1}^- & \widetilde{\mathsf{G}}_{c,1n}^- \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \widetilde{\mathsf{G}}_{c,n-1,n-1}^- & \widetilde{\mathsf{G}}_{c,n-1,n}^- \\ 0 & 0 & \cdots & 0 & \widetilde{\mathsf{G}}_{c,nn}^- \end{bmatrix} \end{array}. \tag{2.24}
$$

Plugging (2.23)–(2.24) into (2.18) and then comparing all the block entries on both sides, we end up with the following procedure to compute the factorization recursively.

In accordance to Theorem 2.3.3, for any generator matrix H and any constant $c > 0$, we denote by

$$(\Lambda_c^+(\mathsf{H}), \Lambda_c^-(\mathsf{H}), \mathsf{G}_c^+(\mathsf{H}), \mathsf{G}_c^-(\mathsf{H}))$$

the unique quadruple constituting the classical Wiener-Hopf factorization (cf. [BRW80]) corresponding to H with killing rate c. In order to proceed, we let $c_k = q_k + c, \quad k \geq 1$.

We are now ready to describe the algorithm to compute the value of $q_1^{-1} \cdots q_n^{-1} \widehat{\Pi}_c^+(i, j; q_1, \ldots, q_n)$.

Step 1. **Compute the first diagonal:** for $k = 1, \ldots, n+1$, compute

$$\widetilde{\Lambda}_{c,k-1,k-1}^+ = \Lambda_{c_k}^+(\mathsf{G}_k),$$

using the diagonalization method in [RS94].

Step 2. **Compute the second diagonal:** for $k = 1, \ldots, n$, solve the following linear system

$$q_k \mathsf{I}^+ + \mathsf{B}_k \widetilde{\Lambda}^+_{c,k-1,k} = \mathsf{V}^+ \widetilde{\mathsf{G}}^+_{c,k-1,k},$$
$$[\mathsf{D}_k - c_k \mathsf{I}^-]\widetilde{\Lambda}^+_{c,k-1,k} + q_k \widetilde{\Lambda}^+_{c,kk} = \mathsf{V}^- \widetilde{\Lambda}^+_{c,k-1,k-1} \widetilde{\mathsf{G}}^+_{c,k-1,k} + \mathsf{V}^- \widetilde{\Lambda}^+_{c,k-1,k} \widetilde{\mathsf{G}}^+_{c,kk},$$

for $\widetilde{\Lambda}^+_{c,k-1,k}$ and $\widetilde{\mathsf{G}}^+_{c,k-1,k}$.

Step 3. **Compute the other diagonals:** for $r = 2, \ldots, n$, $k = 0, \ldots, n-r$, solve the linear system

$$\mathsf{B}_{k+1} \widetilde{\Lambda}^+_{c,k,k+r} = \mathsf{V}^+ \widetilde{\mathsf{G}}^+_{c,k,k+r},$$
$$[\mathsf{D}_{k+1} - c_{k+1} \mathsf{I}^-]\widetilde{\Lambda}^+_{c,k,k+r} + q_{k+1} \widetilde{\Lambda}^+_{c,k+1,k+r} = \mathsf{V}^- \sum_{j=0}^{r} \widetilde{\Lambda}^+_{c,k,k+j} \widetilde{\mathsf{G}}^+_{c,k+j,k+r},$$

for $\widetilde{\Lambda}^+_{c,k,k+r}$ and $\widetilde{\mathsf{G}}^+_{c,k,k+r}$.

Step 4. **Compute**

$$P^+(q_1, \ldots, q_n) := q_1^{-1} \cdots q_n^{-1} \widehat{\Pi}^+_c(i, j; q_1, \ldots, q_n) = q_1^{-1} \cdots q_n^{-1} \sum_{\ell=0}^{n} \widetilde{\Lambda}_{c,0\ell}.$$

for $q_1, \ldots, q_n > 0$.

Step 5. **Compute the approximate inverse Laplace transform of** $P^+(q_1, \ldots, q_n)$**:** use the method discussed in Appendix A.

Remark 2.4.1. *If $|\mathbf{E}^+| = |\mathbf{E}^-| = 1$, then the matrices in Steps 1-3 become numbers. Step 1 reduces to solving $n+1$ quadratic equations for a root in $[0, 1]$. In Step 2 and 3, for each loop, the system reduces to a system of two linear equations of two unknowns. Moreover, in this case, P^+ has a closed-form representation for $q_1, \ldots, q_n > 0$, and hence, for any $q_1, \ldots, q_n \in \mathbb{C}^+$, as mentioned in the previous sec-*

tion. This allows us to use general numerical inverse Laplace transform methods, not necessary the Gaver-Stehfest formula from Appendix A. In particular, one can use Talbot approximation formula (A.1) *presented in Appendix A, which is more efficient than the Gaver-Stehfest under fairly general assumptions (cf. [AW06]).*

2.4.2 Application to Fluid Flow Problems. In this section, we will apply our results to the time-inhomogeneous Markov chain fluid flow problem introduced in Section 1.1. We will compute the quantity

$$\Pi_c^-(i,j) = \mathbb{E}^i\left(e^{-c\tau_0^-}\mathbb{1}_{\{X_{\tau_0^-}=j\}}\right), \quad i \in \mathbf{E}^+,\, j \in \mathbf{E}^-.$$

Towards this end, we further assume that the tank has either an aggregate water inflow at rate v^+ or an aggregate water outflow at rate v^-. In other words,

$$\mathbf{E}^+ = \{e_+\}, \quad \mathbf{E}^- = \{e_-\}, \quad v(e_+) = v^+, \quad \text{and} \quad v(e_-) = v^-.$$

Moreover, we assume that the time-inhomogeneous Markov chain X has the generator

$$\mathsf{G}_t = \begin{cases} \mathsf{G}_1, & s_0 \le t < s_1, \\ \mathsf{G}_2, & s_1 \le t < s_2, \\ \mathsf{G}_3, & t \ge s_2, \end{cases}$$

where $0 < s_1 < s_2$.

We take the following inputs: $c = 0.5, v(e_+) = 2, v(e_-) = -3, s_1 = 2, s_2 = 8$,

$$\mathsf{G}_1 = \begin{array}{c c} & \begin{array}{cc} e_+ & e_- \end{array} \\ \begin{array}{c} e_+ \\ e_- \end{array} & \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} \end{array}, \quad \mathsf{G}_2 = \begin{array}{c c} & \begin{array}{cc} e_+ & e_- \end{array} \\ \begin{array}{c} e_+ \\ e_- \end{array} & \begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix} \end{array}, \quad \mathsf{G}_3 = \begin{array}{c c} & \begin{array}{cc} e_+ & e_- \end{array} \\ \begin{array}{c} e_+ \\ e_- \end{array} & \begin{bmatrix} -5 & 5 \\ 3 & -3 \end{bmatrix} \end{array}.$$

The following table compares our result and execution time with Monte-Carlo simulation (10000 paths). The program is implemented in MATLAB®, and is available from the author upon request.

Table 2.1. Numerical Comparison between WHf and Monte-Carlo

Method	Wiener-Hopf	Monte-Carlo
$\Pi_c^-(e_+, e_-)$	0.6501	0.6462
Execution time	$0.15\,s$	$3.12\,s$

As expected, we obtain similar values for $\Pi_c^-(e_+, e_-)$ by both methods, while clearly the WHf method is much faster than Monte-Carlo.

Remark 2.4.2. *One can also compute $\Pi_c^+(e_-, e_+)$, if it is the quantity of interest in the model. Note that if we change the labels of the states from $\{e_+, e_-\}$ to $\{e_-, e_+\}$ and modify the inputs accordingly, we can compute $\Pi_c^+(e_-, e_+)$ using the same algorithm that computes $\Pi_c^-(e_+, e_-)$.*

CHAPTER 3

STATISTICAL INFERENCE FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

3.1 Introduction

The main goal of this part of the thesis is to derive some consistent and asymptotically normal estimators for model parameters appearing in some linear parabolic Stochastic Partial Differential Equations (SPDEs). We will start by introducing the relevant notations, and the main objects of study. Throughout this chapter, we fix a stochastic basis $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ that satisfies the usual assumptions. Let H be a separable Hilbert space, with the corresponding inner product $(\,\cdot\,,\,\cdot\,)_H$ and norm $||\cdot||_H$. We consider the following stochastic evolution equation

$$\mathrm{d}u(t) + (\theta\mathcal{A}_1 + \mathcal{A}_0)\,u(t)\,\mathrm{d}t = (\mathcal{M}u(t) + \sigma)\,\mathrm{d}W^Q(t), \tag{3.1}$$

with initial condition $u(0) = u_0 \in H$, and where $\mathcal{A}_0, \mathcal{A}_1$ and $\mathcal{M}$ are operators in H or other suitable Hilbert spaces, $W^Q := \{W^Q(t)\}_{t\geq 0}$ is a Q-cylindrical Brownian motion in H, $\theta, \sigma \in \mathbb{R}_+ := (0, \infty)$.

The rigorous study of the existence and uniqueness of the solution to (3.1) is out of the scope of this work, and refer to [Roz90, DPZ92, Cho07, Hai09, LR17] for a detailed discussion on the analytical properties of the solution. We will review some of these properties later in this chapter.

3.1.1 Classes and Examples of SPDEs. SPDEs can be classified into different categories according to noise term. If Q is the identity operator, then the noise is called **space-time white noise**. Otherwise, the noise is referred as space-time **colored noise**. If $\mathcal{M} = 0$, then (3.1) is called an equation with **additive noise**. Otherwise, it is called an equation with **multiplicative noise**. In this work, we mainly study SPDEs driven by an additive noise, i.e. $\mathcal{M} = 0$.

Furthermore, the SPDEs with additive noise can be classified according to the operators $\mathcal{A}_0$ and $\mathcal{A}_1$. If the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ have only point spectra, and a common system of eigenfunctions $\{h_k\}_{k\in\mathbb{N}}$ that form a complete, orthonormal system in H, then (3.1) is called a **diagonalizable** equation. Otherwise it is a **non-diagonalizable** equation.

Finally, similar to the deterministic PDEs, if $H = L^2(\mathbb{R}^d)$, then (3.1) is called an equation on the **whole space**. On the other hand, if $H = L^2(G)$ and $G \subset L^2(\mathbb{R}^d)$ is a bounded domain, then (3.1) is called an equation on **bounded domain**. The parabolicity of SPDEs driven by additive noise is defined by analogy to parabolicity of PDEs in terms of the properties of the operators $\mathcal{A}_0$ and $\mathcal{A}_1$. For equations with multiplicative noise the parabolicity involves also the operator $\mathcal{M}$. We refer to [LR17] for a detailed discussion of classifications of SPDEs.

Next, we give several examples of SPDEs that will be discussed in the applications of the proposed theoretical developments of this section.

Example 3.1.1 (Stochastic heat equation on bounded domain)**.** *If* $H = L^2([0, \pi])$, $\mathcal{A}_1 = -\Delta$, $\mathcal{A}_0 = 0$, $\mathcal{M} = 0$ *and* Q *is identity, then we have a stochastic heat equation on* $[0, \pi]$,

$$\begin{aligned}
&\mathrm{d}u(t,x) = \theta u_{xx}(t,x)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t,x), \quad x \in (0,\pi), \quad t > 0,\\
&u(0,x) = 0, \quad x \in (0,\pi),\\
&u(t,0) = u(t,\pi) = 0, \quad t > 0.
\end{aligned}$$

Example 3.1.2 (Stochastic heat equation on the whole space)**.** *If* $H = L^2(\mathbb{R})$, $\mathcal{A}_1 = -\Delta$, $\mathcal{A}_0 = 0$, $\mathcal{M} = 0$ *and* Q *is identity, then we have a stochastic heat equation on* *the whole real line* $\mathbb{R}$*,*

$$\mathrm{d}u(t,x) = \theta u_{xx}(t,x)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t,x), \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(0,x) = 0, \quad x \in \mathbb{R}.$$

Example 3.1.3. *The following equation is a fractional stochastic heat equation driven by an additive noise, possibly colored in space,*

$$du(t,x) + \theta(-\Delta)^{\beta} u(t,x)\,dt = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x)\,dw_k(t), \quad t \in [0,T], \quad x \in G,$$

with initial condition $u(0,x) = u_0(x) \in H$*, where* $\theta > 0$*,* $\beta > 0$*,* $\gamma \geq 0$ *and* $\sigma \in \mathbb{R} \setminus \{0\}$ *are constants, and where* $\lambda_k := \sqrt{-\tau_k}$*,* $k \in \mathbb{N}$*.*

Example 3.1.4. *We consider the following evolution equation*

$$du(t,x) + (\Delta u(t,x) + \theta u(t,x))\,dt = \sum_{k=1}^{\infty} h_k(x)\,dw_k(t), \quad t \in [0,T], \quad x \in G,$$
$$u(0,x) = u_0(x) \in H,$$

Note that in this case the parameter θ *is in front of the lower order differential operator.*

Example 3.1.5 (Multiplicative Noise)**.** *If* $H = L^2(\mathbb{R})$*,* $\mathcal{A}_1 = -\Delta$*,* $\mathcal{A}_0 = 0$*, and* $\mathcal{M}$ *and* Q *are identity operators, then we have a stochastic heat equation on the whole real line* $\mathbb{R}$ *with multiplicative noise,*

$$\mathrm{d}u(t,x) = \theta u_{xx}(t,x)\,\mathrm{d}t + u(t,x)\,\mathrm{d}W(t,x), \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(0,x) = u_0(x), \quad x \in \mathbb{R}.$$

3.1.2 Statistical Inference for SPDEs. Starting with the seminal papers [HKR93, HR95], most of the existing literature on statistical inference for SPDEs is framed within the so-called spectral approach that explores the parameter estimation problems using the Maximum Likelihood Estimators (MLE). In this approach it is assumed that we observe the first N Fourier modes of the solution are continuously on the time-interval $[0, T]$. The main goal is to construct suitable estimators for the unknown parameters θ and/or σ, and investigate the asymptotic properties of the estimators, as $N \to \infty$.

For example, let us consider the equation (3.3) with $\gamma = 0, \sigma = 1$ and $u_0 = 0$. By direct evaluations one can show that (cf. [Lot09]) that the MLE $\widehat{\theta}_N$ of θ based on the observations $u_k(t), k = 1, \ldots, N, t \in [0, T]$, is given by

$$\widehat{\theta}_N = -\frac{\sum_{k=1}^{N} \nu_k \int_0^T (u_k(t)\,\mathrm{d}u_k(t) + \rho_k u_k^2(t)\,\mathrm{d}t)}{\sum_{k=1}^{N} \nu_k^2 \int_0^T u_k^2(t)\,\mathrm{d}t}, \tag{3.2}$$

Moreover, $\widehat{\theta}_N$ is a strongly consistent estimator of the true parameter θ, i.e.

$$\lim_{N\to\infty} \widehat{\theta}_N = \theta \quad \text{with probability one for all } \theta \in \Theta.$$

It is also asymptotically normal, i.e.

$$\lim_{N\to\infty} \sqrt{\sum_{k=1}^{N} \frac{\nu_k^2}{\mu_k(\theta)}} (\widehat{\theta}_N - \theta) \overset{\mathcal{D}}{=} \mathcal{N}(0, 2/T).$$

The estimation of σ under continuous time observation can be found exactly by employing quadratic variation arguments, similar to finding the volatility in SODEs. For more details on MLEs and their modifications we refer to the surveys [Lot09, Cia18], textbook [LR17] and the monograph [Bis08] on linear SPDEs. For adaptation of MLEs to a nonlinear setup see [CGH11]. In [CX14, CX15], the authors studied the hypothesis testing for stochastic fractional heat equation. Equations driven by fractional Gaussian noise are studied in [Cia10, CLP09]. Finally, see [MB95, PR97, Mar03, PR03, MPR02] for parameter estimation for SPDEs under discrete time sampling schemes and assuming spectral approach. For linear, diagonalizable SPDEs, the MLE can be computed explicitly, and of course, from a statistical point of view there is no need to study other type of estimators. In general, this statement does not hold true for non-diagonalizable equations such as nonlinear SPDEs or SPDEs driven by a multiplicative noise. While the parameter estimation problem for SODEs went way beyond the MLEs (cf. the monograph [Kut04]), there exist a limited number of works dedicated to the non-MLE statistical inference for SPDEs. For example, in [CL09] the authors explore the singularity of corresponding probability measures and derive a closed-form estimators for the drift coefficient for some linear parabolic SPDEs driven by a multiplicative noise (of special structure).

In this work, we propose two methods. First, in Section 3.2, we study a non-MLE estimator for diagonalizable parabolic equations under spectral approach. In second part, going beyond spectral approach, we propose p-variation type estimators which are constructed from discretely observed data, which are discussed in Section 3.3.

3.2 Trajectory Fitting Estimators (TFEs)

In this section, we study the estimator that is related to what is known in the literature the trajectory fitting estimators (TFEs). Using as observations the first N Fourier modes, we construct the TFE by analogy to the TFE for continuously observed finite dimensional ergodic diffusion processes first introduced by Y. A. Kutoyants [Kut91]; see also [Kut04, Section 1.3 & Section 2.3] and references therein. The TFE can be also viewed as an analog of the least squares estimators widely used in time-series analysis. We study the asymptotic properties of the TFE as $N \to \infty$, in contrast to the diffusion processes where the asymptotics are done for the large-time regime. Surely one can investigate the large-time asymptotics for SPDEs too, but we find this case to be too similar to the estimators for diffusion processes, and we omitted it here. In this study, we consider a fairly simple, although general, class of SPDEs: linear, parabolic, diagonalizable equations, driven by an additive space-time noise. The diagonalizable nature of these equations, allows us to derive explicit expressions for the considered estimators and for the asymptotics of their first two moments, and hence to investigate the rate of convergence of these estimators. While simple, these equations can be viewed as a good approximation of some more complicated and practically important models. On the other hand, the obtained results will serve as benchmarks for future studies of more complicated and realistic models which will be addressed in the sequel. Under some general structural assumptions we prove consistency and asymptotic normality of the proposed estimators.

We will study the following diagonalizable parabolic equations with additive noise,

$$\mathrm{d}u(t) + (\theta\mathcal{A}_1 + \mathcal{A}_0)\, u(t)\, \mathrm{d}t = \sigma\, \mathrm{d}W(t), \tag{3.3}$$

that satisfy the following conditions,

(i) The operators $\mathcal{A}_0$ and $\mathcal{A}_1$ have only point spectra, and a common system of eigenfunctions $\{h_k\}_{k\in\mathbb{N}}$ that form a complete, orthonormal system in H. We denote the corresponding eigenvalues of $\mathcal{A}_0$ and $\mathcal{A}_1$ by ρ_k and ν_k, $k\in\mathbb{N}$, respectively.

(ii) The sequence $\{\mu_k(\theta)\}_{k\in\mathbb{N}}$, where $\mu_k(\theta) := \theta\nu_k + \rho_k$, is such that

$$\lim_{k\to\infty} \mu_k(\theta) = +\infty,$$

where the convergence is uniform in $\theta \in \Theta$.

(iii) There exist universal constants $J \in \mathbb{N}$ and $c_0 > 0$ such that, for any $k \geq J$ and any $\theta \in \Theta$,

$$\frac{\nu_k}{\mu_k(\theta)} \leq c_0.$$

(iv) The sequence $\{\nu_k\}_{k\in\mathbb{N}}$ is such that $\lim_{k\to\infty} \nu_k = +\infty$.[1]

(v) The noise term W is a cylindrical Brownian motion in H, and has the following form

$$W(t) = \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k w_k(t), \quad t \geq 0,$$

for some $\gamma \geq 0$, where $\lambda_k := \nu_k^{1/(2m)}$, $k \in \mathbb{N}$, for some $m \geq 0$,[2] and where $w_k := \{w_k(t)\}_{t\geq 0}$, $k \in \mathbb{N}$, is a collection of independent standard Brownian motions.

Conditions (i)–(v) imply that the equation (3.3) is linear, diagonalizable, parabolic, and that the solution exists and is unique; this can be established by standard methods from theory of SPDEs, and we refer, for instance, to [Lot09, HLR97, HR95] for similar setup, or to [Roz90, Cho07, LR17] for a general theory. Of course, one class of operators $\mathcal{A}_0$ and $\mathcal{A}_1$ that satisfy the above conditions are pseudo-differential operators on bounded domains with appropriate boundary conditions, with $\mathcal{A}_0$ being subordinated to $\mathcal{A}_1$.

---

[1]Without loss of generality, we will assume that $\nu_k \geq 0$, for all $k \in \mathbb{N}$.

[2]Of course, one can consider at once just $\lambda_k = \nu_k$. Our choice to consider m is to put the results on par with the notations from the existing literature. As mentioned later, if $\mathcal{A}_0$ and $\mathcal{A}_1$ are some pseudo-differential operators, then it is convenient to denote by $2m$ the order of the leading order operator.

The unique solution to (3.3), with initial condition $u(0) = u_0$, is given by

$$u(t) = \sum_{k=1}^{\infty} u_k(t) h_k, \quad t \geq 0,$$

where, for each $k \in \mathbb{N}$, $u_k := \{u_k(t)\}_{t\geq 0}$ satisfies the following ordinary stochastic differential equation (SDE)

$$\mathrm{d}u_k(t) + \mu_k(\theta)\, u_k(t)\ \mathrm{d}t = \sigma \lambda_k^{-\gamma}\ \mathrm{d}w_k(t), \tag{3.4}$$

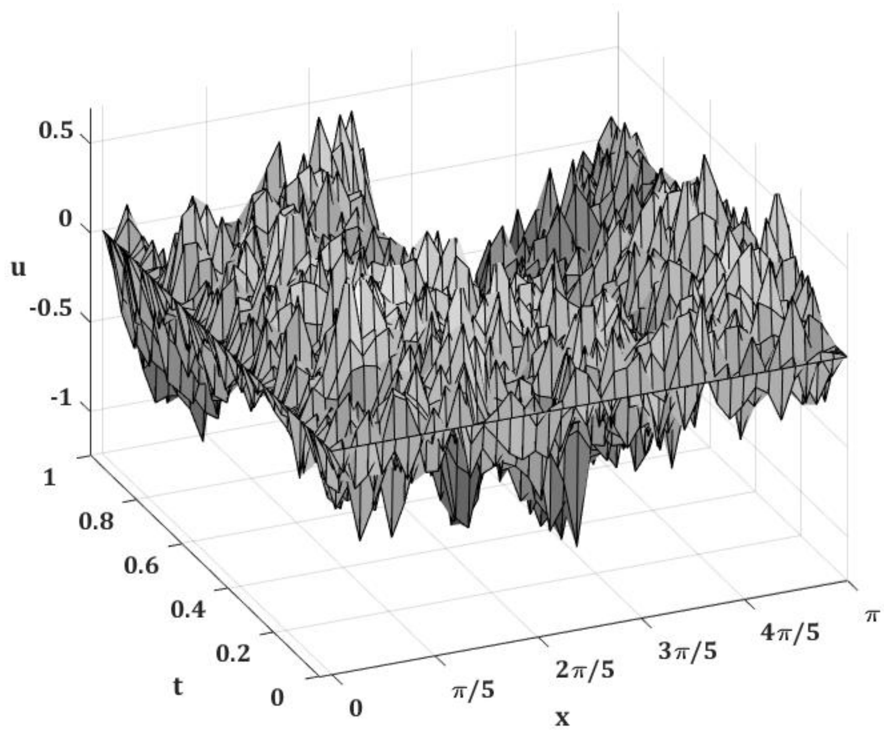
with initial condition $u_k(0) = (u_0, h_k)_H$. The stochastic processes u_k, $k \in \mathbb{N}$, are the Fourier modes of the solution u with respect to the basis $\{h_k\}_{k\in\mathbb{N}}$ of H, i.e., $u_k(t) = (u(t), h_k)_H$, $t \geq 0$, $k \in \mathbb{N}$. Note that the SDEs of the form (3.4), for $k \in \mathbb{N}$, provide an infinite system of decoupled/independent Ornstein–Uhlenbeck processes. As already mentioned, the decoupling nature of Fourier modes, or the diagonalizable property of the original equation, will play a critical role in our study, and it is essentially guaranteed by the assumptions (i) and (v). By Itô's formula, clearly we have

$$u_k(t) = e^{-\mu_k(\theta)t} u_k(0) + \sigma \lambda_k^{-\gamma} e^{-\mu_k(\theta)t} \int_0^t e^{\mu_k(\theta)s}\, dw_k(s), \quad t \geq 0, \quad k \in \mathbb{N}. \tag{3.5}$$

We plot the simulated solution for the following equation with additive space-time white noise,

$$\mathrm{d}u(t,x) - \Delta u(t,x)\, \mathrm{d}t = \mathrm{d}W(t,x), \quad t \geq 0, \quad x \in [0,\pi],$$

$$u(0,x) = 0, \quad x \in [0,\pi], \quad u(t,0) = u(t,\pi) = 0, \quad t \geq 0.$$



Next, we review the construction of TFEs for SODEs.

3.2.1 Trajectory Fitting Estimators for SODEs. The trajectory fitting estimators for continuous-time diffusion processes can be viewed as an analog of the least squares estimators widely used in time-series analysis. Following [Kut04, Section 1.3 & Section 2.3], we will briefly describe the TFEs for finite-dimensional diffusions. Assume that the observed process $S(\theta) := \{S(t;\theta)\}_{t\geq 0}$ follows the dynamics

$$dS(t;\theta) = b(\theta, S(t;\theta))dt + \sigma(S(t;\theta))\, dB(t), \tag{3.6}$$

where $B := \{B(t)\}_{t\geq 0}$ is a one-dimensional standard Brownian motion, and θ is the parameter of interest. We assume that the drift b and the volatility σ are known, and that the solution to (3.6) (with certain initial condition $S(0,\theta) = S_0$) exists and is unique, for any $\theta \in \Theta$. Let $F : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function.

By Itô's formula,

$$F(S(t;\theta)) = F(S_0) + \int_0^t \Big(F'(S(s;\theta)) b\left(\theta, S(s;\theta)\right) + \frac{1}{2} F''(S(s;\theta)) \sigma^2(S(s;\theta)) \Big)\, ds \\ + \int_0^t F'(S(s;\theta)) \sigma(S(s;\theta))\, dB(s).$$

For any $\theta \in \Theta$ and $t \in [0,T]$, let

$$\widetilde{F}(t;\theta) := F(S_0) + \int_0^t \Big(F'(S(s;\theta)) b\left(\theta, S(s;\theta)\right) + \frac{1}{2} F''(S(s;\theta)) \sigma^2(S(s;\theta)) \Big)\, ds.$$

The *trajectory fitting estimator*[3] $\widetilde{\theta}_T$ of θ is then defined as the solution to the minimization problem

$$\widetilde{\theta}_T := \operatorname*{arg\,inf}_{\theta \in \Theta} \int_0^T \Big(F(S(t;\theta)) - \widetilde{F}(t;\theta) \Big)^2 dt. \tag{3.7}$$

The choice of function F depends on the underlying models, and has to be taken such that the estimator satisfies the desired asymptotic properties (consistency, asymptotic normality, etc).

3.2.2 Construction of TFE for (3.3). We take $F(x) = x^2$ in (3.7). For each Fourier mode u_k, $k \in \mathbb{N}$, by Itô's formula, we have

$$u_k^2(t) = u_k^2(0) + \int_0^t \left(\sigma^2 \lambda_k^{-2\gamma} - 2\mu_k(\theta) u_k^2(s) \right) ds + 2\sigma \lambda_k^{-\gamma} \int_0^t u_k(s)\, dw_k(s), \quad t \geq 0. \tag{3.8}$$

By (3.7), one can easily construct a TFE for θ based on the trajectory on $[0,T]$, for some fixed horizon $T > 0$, of each Fourier mode u_k. The *long-time* asymptotic behavior of such estimators as $T \to \infty$ has been well investigated (cf. [Kut04]), and

[3]The terminology comes from the fact that the estimator is obtained by fitting the observed trajectory with the artificial one.

thus we will omit it here.

By analogy to the construction of maximum likelihood estimators for SPDEs (cf. [Lot09]), we will construct a TFE for the unknown parameter θ based on the trajectories of the first N Fourier modes of the solution. Moreover, for a fixed horizon $T > 0$, we will study the *large-space* asymptotic behavior of the TFE as the number of the Fourier modes increases, which is a distinguished feature for an infinite dimensional evolution system. Specifically, fix any $T > 0$, and for any $\theta \in \Theta$, let

$$V_k(t;\theta) := u_k^2(0) + \int_0^t \left(\sigma^2\lambda_k^{-2\gamma} - 2\mu_k(\theta)u_k^2(s)\right) ds, \quad k \in \mathbb{N}, \quad t \in [0,T]. \tag{3.9}$$

The TFE for the unknown parameter θ is defined as

$$\widetilde{\theta}_N = \widetilde{\theta}_N(\gamma, T, \sigma, m) := \arg\inf_{\theta\in\Theta} \sum_{k=1}^N \int_0^T \left(V_k(t;\theta) - u_k^2(t)\right)^2 dt.$$

We are interested in the asymptotic properties of $\widetilde{\theta}_N$, as $N \to \infty$, with T being fixed.

One advantage of the TFE is that it can be given by an explicit formula that does not contain a stochastic integral. Indeed, by (3.8) and (3.9),

$$\begin{aligned}&\sum_{k=1}^N \int_0^T \left(V_k(t;\theta) - u_k^2(t)\right)^2 dt \\ &= \sum_{k=1}^N \int_0^T \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t - 2\rho_k\xi_k(t) - u_k^2(t) - 2\theta\nu_k\xi_k(t)\right)^2 dt.\end{aligned}$$

The maximizer of the last expression, with respect to θ, can be computed simply by finding the root of the first-order derivative. Specifically, let

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\sum_{k=1}^N \int_0^T \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t - 2\rho_k\xi_k(t) - u_k^2(t) - 2\theta\nu_k\xi_k(t)\right)^2 dt$$

$$
\begin{aligned}
&= \frac{\mathrm{d}}{\mathrm{d}\theta} \sum_{k=1}^{N} \int_0^T \left(u_k^2(0) + \sigma^2 \lambda_k^{-2\gamma} t - 2\rho_k \xi_k(t) - u_k^2(t) \right)^2 dt \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}\theta} \sum_{k=1}^{N} \int_0^T 4\theta \nu_k \xi_k(t) \left(u_k^2(0) + \sigma^2 \lambda_k^{-2\gamma} t - 2\rho_k \xi_k(t) - u_k^2(t) \right) dt \\
&\quad + \frac{\mathrm{d}}{\mathrm{d}\theta} \sum_{k=1}^{N} \int_0^T \left(2\theta \nu_k \xi_k(t) \right)^2 dt \\
&= - \sum_{k=1}^{N} \int_0^T 4 \nu_k \xi_k(t) \left(u_k^2(0) + \sigma^2 \lambda_k^{-2\gamma} t - 2\rho_k \xi_k(t) - u_k^2(t) \right) dt \\
&\quad + \sum_{k=1}^{N} \int_0^T 8\theta \left(\nu_k \xi_k(t) \right)^2 dt \\
&= 0.
\end{aligned}
$$

This yields the following explicit expression for the TFE

$$
\widetilde{\theta}_N = \frac{\sum_{k=1}^{N} \int_0^T \nu_k \xi_k(t) \left(u_k^2(0) + \sigma^2 \lambda_k^{-2\gamma} t - 2\rho_k \xi_k(t) - u_k^2(t) \right) dt}{\sum_{k=1}^{N} \int_0^T 2 \left(\nu_k \xi_k(t) \right)^2 dt}.
$$

In what follows, we will make use of the following notations. For any $t \in [0, T]$, let

$$
\begin{aligned}
\xi_k(t) &:= \int_0^t u_k^2(s)\, ds, \quad X_k(t) := \int_0^t s \xi_k(s)\, ds, \\
Y_k(t) &:= \int_0^t \xi_k(s)\, ds, \quad Z_k(t) := \int_0^t \xi_k^2(s)\, ds.
\end{aligned}
\tag{3.10}
$$

Thus the TFE can be represented as

$$
\widetilde{\theta}_N = - \frac{\sum_{k=1}^{N} \nu_k \left(\frac{1}{2} \xi_k^2(T) - u_k^2(0) Y_k(T) - \sigma^2 \lambda_k^{-2\gamma} X_k(T) + 2\rho_k Z_k(T) \right)}{2 \sum_{k=1}^{N} \nu_k^2 Z_k(T)}. \tag{3.11}
$$

3.2.3 Auxiliary Results of Asymptotics. In what follows, we will denote by θ the true parameter. For notational simplicity, the T variable in $\xi_k(T)$, $X_k(T)$, $Y_k(T)$ and $Z_k(T)$ will be omitted from now on. Subtracting θ on the both sides of (3.11) leads to

$$
\begin{aligned}
\widetilde{\theta}_N - \theta &= -\frac{\sum_{k=1}^N \nu_k \left(\frac{1}{2}\xi_k^2 - u_k^2(0)Y_k - \sigma^2\lambda_k^{-2\gamma}X_k + 2\rho_k Z_k\right)}{2\sum_{k=1}^N \nu_k^2 Z_k} - \theta \\
&= -\frac{\sum_{k=1}^N \nu_k \left(\frac{1}{2}\xi_k^2 - u_k^2(0)Y_k - \sigma^2\lambda_k^{-2\gamma}X_k + 2\rho_k Z_k + 2\nu_k\theta Z_k\right)}{2\sum_{k=1}^N \nu_k^2 Z_k} \\
&= -\frac{\sum_{k=1}^N \nu_k \left(\frac{1}{2}\xi_k^2 - u_k^2(0)Y_k - \sigma^2\lambda_k^{-2\gamma}X_k + 2\mu_k(\theta) Z_k\right)}{2\sum_{k=1}^N \nu_k^2 Z_k} \\
&=: -\frac{\sum_{k=1}^N \nu_k A_k}{2\sum_{k=1}^N \nu_k^2 Z_k}. \qquad (3.12)
\end{aligned}
$$

As usual, for two sequences of positive numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, we will write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$, and will write $a_n \asymp b_n$ if there exist universal constants $K_2 > K_1 > 0$, such that $K_1 b_n \leq a_n \leq K_2 b_n$ for $n \in \mathbb{N}$ large enough.

We start with a technical result regarding the leading order terms of the means and variances of A_k and Z_k, as $k \to \infty$.

Proposition 3.2.1. *Let the assumptions* (i) - (v) *be satisfied. Then, as* $k \to \infty$,

$$
\begin{aligned}
\mathbb{E}(Z_k) &\sim \frac{u_k^4(0)T}{4\mu_k^2(\theta)} + \frac{\sigma^2\lambda_k^{-2\gamma}u_k^2(0)T^2}{4\mu_k^2(\theta)} + \frac{\sigma^4\lambda_k^{-4\gamma}T^3}{12\mu_k^2(\theta)}, \qquad (3.13)\\
\mathrm{Var}(Z_k) &\sim \frac{\sigma^2\lambda_k^{-2\gamma}u_k^6(0)T^2}{2\mu_k^5(\theta)} + \frac{2\sigma^4\lambda_k^{-4\gamma}u_k^4(0)T^3}{3\mu_k^5(\theta)} \\
&\quad + \frac{\sigma^6\lambda_k^{-6\gamma}u_k^2(0)T^4}{3\mu_k^5(\theta)} + \frac{\sigma^8\lambda_k^{-8\gamma}T^5}{15\mu_k^5(\theta)}, \qquad (3.14)\\
\mathbb{E}(A_k) &\sim \frac{\sigma^2\lambda_k^{-2\gamma}u_k^2(0)T}{\mu_k^2(\theta)} + \frac{\sigma^4\lambda_k^{-4\gamma}T^2}{2\mu_k^2(\theta)}, \qquad (3.15)\\
\mathrm{Var}(A_k) &\sim \frac{\sigma^2\lambda_k^{-2\gamma}u_k^6(0)T^2}{2\mu_k^3(\theta)} + \frac{2\sigma^4\lambda_k^{-4\gamma}u_k^4(0)T^3}{3\mu_k^3(\theta)} \\
&\quad + \frac{\sigma^6\lambda_k^{-6\gamma}u_k^2(0)T^4}{3\mu_k^3(\theta)} + \frac{\sigma^8\lambda_k^{-8\gamma}T^5}{15\mu_k^3(\theta)}. \qquad (3.16)
\end{aligned}
$$

Proof. Due to the nature of desired asymptotic results, the underlying computations are somehow extensive and tedious. Most of the evaluations were performed using symbolic computations in Mathematica®. For the simplicity of notation in computations, we denote

$$a := \mu_k(\theta), \quad b := \sigma\lambda_k^{-\gamma}, \quad , c := u_k(0).$$

For each $k \in \mathbb{N}$,

$$u_k(t) = e^{-at}c + be^{-at}\int_0^t e^{as}\, dw_k(s), \quad t \geq 0.$$

and with the notations introduced in (3.10) and (3.12), we get

$$A_k = \frac{1}{2}\xi_k^2 - c^2 Y_k - b^2 X_k + 2aZ_k.$$

Note that for any $t \in [0,T]$, $u_k(t)$ is a normal random variable with mean $e^{-at}c$ and variance $\frac{b^2}{2a}(1-e^{-2at})$. First, we compute the even moments of $u_k(t)$. By Itô's formula,

$$\begin{aligned}
\mathrm{d}u_k^2(t) &= (b^2 - 2au_k^2(t))\,\mathrm{d}t + 2bu_k(t)\,\mathrm{d}w_k(t),\\
\mathrm{d}u_k^4(t) &= (6b^2u_k^2(t) - 4au_k^4(t))\,\mathrm{d}t + 4bu_k^3(t)\,\mathrm{d}w_k(t),\\
\mathrm{d}u_k^6(t) &= (15b^2u_k^4(t) - 6au_k^6(t))\,\mathrm{d}t + 6bu_k^5(t)\,\mathrm{d}w_k(t),\\
\mathrm{d}u_k^8(t) &= (28b^2u_k^6(t) - 8au_k^8(t))\,\mathrm{d}t + 8bu_k^7(t)\,\mathrm{d}w_k(t).
\end{aligned}$$

Therefore, the expectations satisfy

$$\begin{aligned}
\mathbb{E}u_k^2(t) &= \int_0^t (b^2 - 2a\mathbb{E}u_k^2(s))\,\mathrm{d}s,\\
\mathbb{E}u_k^4(t) &= \int_0^t (6b^2\mathbb{E}u_k^2(s) - 4a\mathbb{E}u_k^4(s))\,\mathrm{d}s,\\
\mathbb{E}u_k^6(t) &= \int_0^t (15b^2\mathbb{E}u_k^4(s) - 6a\mathbb{E}u_k^6(s))\,\mathrm{d}s,\\
\mathbb{E}u_k^8(t) &= \int_0^t (28b^2\mathbb{E}u_k^6(s) - 8a\mathbb{E}u_k^8(s))\,\mathrm{d}s,
\end{aligned}$$

and to find these expectations, it is equivalent to solve the following ODEs,

$$
\begin{aligned}
\frac{\mathrm{d}\mathbb{E}u_k^2(t)}{\mathrm{d}t} &= b^2 - 2a\mathbb{E}u_k^2(t), \quad \mathbb{E}u_k^2(0) = c^2,\\
\frac{\mathrm{d}\mathbb{E}u_k^4(t)}{\mathrm{d}t} &= 6b^2\mathbb{E}u_k^2(t) - 4a\mathbb{E}u_k^4(t), \quad \mathbb{E}u_k^4(0) = c^4,\\
\frac{\mathrm{d}\mathbb{E}u_k^6(t)}{\mathrm{d}t} &= 15b^2\mathbb{E}u_k^4(t) - 6a\mathbb{E}u_k^6(t), \quad \mathbb{E}u_k^6(0) = c^6,\\
\frac{\mathrm{d}\mathbb{E}u_k^8(t)}{\mathrm{d}t} &= 28b^2\mathbb{E}u_k^6(t) - 8a\mathbb{E}u_k^8(t), \quad \mathbb{E}u_k^8(0) = c^8.
\end{aligned}
$$

We use Mathematica to solve these ODEs and get

$$
\begin{aligned}
\mathbb{E}u_k^2(t) &= \frac{b^2}{2a} - \frac{b^2e^{-2at}}{2a} + c^2e^{-2at},\\
\mathbb{E}u_k^4(t) &= \frac{3b^4}{4a^2} + \frac{3b^4e^{-4at}}{4a^2} - \frac{3b^2c^2e^{-4at}}{a} + c^4e^{-4at} - \frac{3b^4e^{-2at}}{2a^2} + \frac{3b^2c^2e^{-2at}}{a},\\
\mathbb{E}u_k^6(t) &= \frac{15b^6}{8a^3} - \frac{15b^6e^{-6at}}{8a^3} + \frac{45b^4c^2e^{-6at}}{4a^2} - \frac{15b^2c^4e^{-6at}}{2a} + c^6e^{-6at}\\
&\quad + \frac{45b^6e^{-4at}}{8a^3} - \frac{45b^4c^2e^{-4at}}{2a^2} + \frac{15b^2c^4e^{-4at}}{2a} - \frac{45b^6e^{-2at}}{8a^3} + \frac{45b^4c^2e^{-4at}}{4a^2},\\
\mathbb{E}u_k^8(t) &= \frac{105b^8}{16a^4} + \frac{105b^8e^{-8at}}{16a^4} - \frac{105b^6c^2e^{-8at}}{2a^3} + \frac{105b^4c^4e^{-8at}}{2a^2} - \frac{14b^2c^6e^{-8at}}{a}\\
&\quad + c^8e^{-8at} - \frac{105b^8e^{-6at}}{4a^4} + \frac{315b^6c^2e^{-6at}}{2a^3} - \frac{105b^4c^4e^{-6at}}{a^2} + \frac{14b^2c^6e^{-6at}}{a}\\
&\quad + \frac{315b^8e^{-4at}}{8a^4} - \frac{315b^6c^2e^{-4at}}{2a^3} + \frac{105b^4c^4e^{-4at}}{2a^2} - \frac{105b^8e^{-2at}}{4a^4} + \frac{105b^6c^2e^{-2at}}{2a^3}.
\end{aligned}
$$

In addition, we know

$$
\mathbb{E}\xi_k(t) = -\frac{b^2}{4a} + \frac{c^2}{2a} + \frac{b^2e^{-2at}}{4a^2} - \frac{c^2e^{-2at}}{2a} + \frac{b^2t}{2a}.
$$

We first verify (3.13). From the definitions in (3.10) we have

$$
\begin{aligned}
\mathbb{E}(Z_k) = \mathbb{E}\left(\int_0^T \xi_k^2(t)\,\mathrm{d}t\right) &= \int_0^T \mathbb{E}\left(\xi_k^2(t)\right)\mathrm{d}t,\\
&= 2\int_0^T\int_0^t \mathbb{E}\left(\xi_k(s)u_k^2(s)\right)\mathrm{d}s\,\mathrm{d}t.
\end{aligned} \tag{3.17}
$$

By Itô's formula,

$$\mathrm{d}\xi_k(t)u_k^2(t) = \left(u_k^4(t) + b^2\xi_k(t)\right)\mathrm{d}t - 2a\xi_k(t)u_k^2(t)\ \mathrm{d}t + 2bu_k(t)\xi_k(t)\,\mathrm{d}w_k(t).$$

Taking the integration and expectations on both sides above, we obtain that

$$\mathbb{E}\left(\xi_k(t)u_k^2(t)\right) = \int_0^t \left(\mathbb{E}u_k^4(s) + b^2\mathbb{E}\xi_k(s)\right)\mathrm{d}s - 2a\mathbb{E}\left(\xi_k(s)u_k^2(s)\right)\mathrm{d}s.$$

Thus $\mathbb{E}\left(\xi_k(t)u_k^2(t)\right)$ satisfies the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left(\xi_k(t)u_k^2(t)\right) = \left(\mathbb{E}u_k^4(t) + b^2\mathbb{E}\xi_k(t)\right)\mathrm{d}s - 2a\mathbb{E}\left(\xi_k(t)u_k^2(t)\right),$$

with zero initial condition. Therefore, we get

$$\begin{aligned}\mathbb{E}\left(\xi_k(t)u_k^2(t)\right) = {} & \frac{b^4}{8a^3} + \frac{b^2c^2}{4a^2} - \frac{3b^4e^{-4at}}{8a^3} + \frac{3b^2c^2e^{-4at}}{2a^2} \\ & - \frac{c^4e^{-4at}}{2a} + \frac{b^4e^{-2at}}{4a^3} - \frac{7b^2c^2e^{-2at}}{4a^2} + \frac{c^4e^{-2at}}{2a} \\ & + \frac{b^4t}{4a^2} - \frac{5b^4e^{-2at}t}{4a^2} + \frac{5b^2c^2e^{-2at}t}{2a}.\end{aligned}$$

Therefore, by (3.17),

$$\begin{aligned}\mathbb{E}Z_k = {} & \frac{35b^4}{64a^5} - \frac{9b^2c^2}{16a^4} - \frac{3c^4}{16a^3} - \frac{3b^4e^{-4aT}}{64a^5} + \frac{3b^2c^2e^{-4aT}}{16a^4} \\ & - \frac{c^4e^{-4aT}}{16a^3} - \frac{b^4e^{-2aT}}{2a^5} + \frac{3b^2c^2e^{-2aT}}{8a^4} + \frac{c^4e^{-2at}}{4a^3} - \frac{9b^4T}{16a^4} \\ & + \frac{b^2c^2T}{4a^3} + \frac{c^4T}{4a^2} - \frac{5b^4e^{-2aT}T}{8a^4} + \frac{5b^2c^2e^{-2aT}T}{4a^3} + \frac{b^4T^2}{8a^3} \\ & + \frac{b^2c^2T^2}{4a^2} + \frac{b^4T^3}{12a^2}.\end{aligned} \tag{3.18}$$

which leads to (3.13), since by the assumption (ii), we only keep the terms with denominators that are second order of a, and since $T > 0$ is a fixed constant.

Next, we study the asymptotic order of $\mathrm{Var}(Z_k)$, $k \to \infty$, given by (3.14). In light of (3.18), we are left to compute $\mathbb{E}(Z_k^2)$. By Itô's formula, and since the Itô integral terms have zero expectation, we have

$$\begin{aligned}\mathbb{E}(Z_k^2) &= 2\int_0^T \mathbb{E}\big(Z_k(t)\xi_k^2(t)\big)\,\mathrm{d}t \\ &= 2\int_0^T\int_0^t \mathbb{E}\big(\xi_k^4(s)\big)\mathrm{d}s\ \mathrm{d}t + 4\int_0^T\int_0^t \mathbb{E}\big(Z_k(s)\xi_k(s)u_k^2(s)\big)\mathrm{d}s\ \mathrm{d}t. \end{aligned} \tag{3.19}$$

To compute the first expectation in (3.19), by Itô's formula again, we have

$$\begin{aligned}\mathrm{d}\xi_k^4(t) &= 4\xi_k^3(t)u_k^2(t)\ \mathrm{d}t,\\ \mathrm{d}\big(\xi_k^3(t)u_k^2(t)\big) &= \big(3\xi_k^2(t)u_k^4(t) + \xi_k^3(t) - 2a\xi_k^3(t)u_k^2(t)\big)\,\mathrm{d}t + 2\xi_k^3(t)u_k(t)\ \mathrm{d}w_k(t),\\ \mathrm{d}\xi_k^3(t) &= 3\xi_k^2(t)u_k^2(t)\ \mathrm{d}t,\\ \mathrm{d}\big(\xi_k^2(t)u_k^4(t)\big) &= 2\left(\xi_k(t)u_k^6(t) + 3\xi_k^2(t)u_k^2(t) - 2a\xi_k^2(t)u_k^4(t)\right)\mathrm{d}t + 4\xi_k^2(t)u_k^3(t)\ \mathrm{d}w_k(t).\end{aligned}$$

We need to compute $\mathbb{E}(\xi_k(s)u_k^6(s))$ and $\mathbb{E}(\xi_k^2(s)u_k^2(s))$. Again by Itô's formula, we get

$$\begin{aligned}\mathrm{d}\xi_k(t)u_k^6(t) &= \big(u_k^8(t) + 15\xi_k(t)u_k^4(t) - 6a\xi_k(t)u_k^6(t)\big)\,\mathrm{d}t + 6\xi_k(t)u_k^5(t)\ \mathrm{d}w_k(t),\\ \mathrm{d}\xi_k^2(t)u_k^2(t) &= \big(2\xi_k(t)u_k^4(t) + \xi_k^2(t) - 2a\xi_k^2(t)u_k^2(t)\big)\,\mathrm{d}t + 2\xi_k^2(t)u_k(t)\ \mathrm{d}w_k(t),\\ \mathrm{d}\xi_k(t)u_k^4(t) &= \big(u_k^6(t) + 6\xi_k(t)u_k^2(t) - 4a\xi_k(t)u_k^4(t)\big)\,\mathrm{d}t + 4\xi_k(t)u_k^3(t)\ \mathrm{d}w_k(t).\end{aligned}$$

Then we find the following ODEs,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left(\xi_k(t)u_k^6(t)\right) &= \mathbb{E}u_k^8(t) + 15\mathbb{E}\left(\xi_k(t)u_k^4(t)\right) - 6a\mathbb{E}\left(\xi_k(t)u_k^6(t)\right),\\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left(\xi_k^2(t)u_k^2(t)\right) &= 2\mathbb{E}\left(\xi_k(t)u_k^4(t)\right) + \mathbb{E}\xi_k^2(t) - 2a\mathbb{E}\left(\xi_k^2(t)u_k^2(t)\right),\\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left(\xi_k(t)u_k^4(t)\right) &= \mathbb{E}u_k^6(t) + 6\mathbb{E}\left(\xi_k(t)u_k^2(t)\right) - 4\mathbb{E}\left(\xi_k(t)u_k^4(t)\right).\end{aligned}$$

Therefore, we get

$$\begin{aligned}\mathbb{E}\left(\xi_k(t)u_k^4(t)\right) = {} & \frac{9b^6}{16a^4} + \frac{3b^4c^2}{8a^3} + \frac{15b^6e^{-6at}}{16a^4} - \frac{45b^4c^2e^{-6at}}{8a^3} + \frac{15b^2c^4e^{-6at}}{4a^2} \\ & - \frac{c^6e^{-6at}}{2a} - \frac{21b^6e^{-4at}}{16a^4} + \frac{69b^4c^2e^{-4at}}{8a^3} - \frac{21b^2c^4e^{-4at}}{4a^2} + \frac{c^6e^{-4at}}{2a} \\ & - \frac{3b^6e^{-2at}}{16a^4} - \frac{27b^4c^2e^{-2at}}{8a^3} + \frac{3b^2c^4e^{-2at}}{2a^2} + \frac{3b^6t}{8a^3} + \frac{27b^6e^{-4at}t}{8a^3} \\ & - \frac{27b^4c^2e^{-4at}t}{2a^2} + \frac{9b^2c^4e^{-4at}t}{2a} - \frac{15b^6e^{-2at}t}{4a^3} + \frac{15b^4c^2e^{-2at}t}{2a^2},\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(\xi_k(t)u_k^6(t)\right) = {} & \frac{75b^8}{32a^5} + \frac{15b^6c^2}{16a^4} - \frac{105b^8e^{-8at}}{32a^5} + \frac{105b^6c^2e^{-8at}}{4a^4} - \frac{105b^4c^4e^{-8at}}{4a^3} \\ & + \frac{7b^2c^6e^{-8at}}{a^2} - \frac{c^8e^{-8at}}{2a} + \frac{15b^8e^{-6at}}{2a^5} - \frac{915b^6c^2e^{-6at}}{16a^4} + \frac{405b^4c^4e^{-6at}}{8a^3} \\ & - \frac{43b^2c^6e^{-6at}}{4a^2} + \frac{c^8e^{-6at}}{2a} - \frac{45b^8e^{-4at}}{16a^5} + \frac{585b^6c^2e^{-4at}}{16a^4} - \frac{30b^4c^4e^{-4at}}{a^3} \\ & + \frac{15b^2c^6e^{-4at}}{4a^2} - \frac{15b^8e^{-2at}}{4a^5} - \frac{105b^6c^2e^{-2at}}{16a^4} + \frac{45b^4c^4e^{-2at}}{8a^3} + \frac{15b^8t}{16a^4} \\ & - \frac{195b^8e^{-6at}t}{16a^4} + \frac{585b^6c^2e^{-6at}t}{8a^3} - \frac{195b^4c^4e^{-6at}t}{4a^2} + \frac{13b^2c^6e^{-6at}t}{2a} \\ & + \frac{405b^8e^{-4at}t}{16a^4} - \frac{405b^6c^2e^{-4at}t}{4a^3} + \frac{135b^4c^4e^{-4at}t}{4a^2} \\ & - \frac{225b^8e^{-2at}t}{16a^4} + \frac{225b^6c^2e^{-2at}t}{8a^3},\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left(\xi_k^2(t)u_k^2(t)\right) = {} & \frac{3b^6}{32a^5} + \frac{3b^4c^2}{8a^4} + \frac{b^2c^4}{8a^3} - \frac{15b^6e^{-6at}}{32a^5} + \frac{45b^4c^2e^{-6at}}{16a^4} - \frac{15b^2c^4e^{-6at}}{8a^3} \\ & + \frac{c^6e^{-6at}}{4a^2} - \frac{15b^6e^{-4at}}{32a^5} - \frac{3b^4c^2e^{-4at}}{2a^4} + \frac{23b^2c^4e^{-4at}}{8a^3} - \frac{c^6e^{-4at}}{2a^2} \\ & + \frac{27b^6e^{-2at}}{32a^5} - \frac{27b^4c^2e^{-2at}}{16a^4} - \frac{9b^2c^4e^{-2at}}{8a^3} + \frac{c^6e^{-2at}}{4a^2} + \frac{3b^6t}{8a^4} + \frac{b^4c^2t}{4a^3} \\ & - \frac{27b^6e^{-4at}t}{8a^4} + \frac{27b^4c^2e^{-4at}t}{2a^3} - \frac{9b^2c^4e^{-4at}t}{2a^2} - \frac{25b^4c^2e^{-2at}t}{4a^3} \\ & + \frac{5b^2c^4e^{-2at}t}{2a^2} + \frac{b^6t^2}{8a^3} - \frac{25b^6e^{02at}t^2}{8a^3} + \frac{25b^4c^2e^{-2at}t^2}{4a^2}.\end{aligned}$$

And then we get $\mathbb{E}\xi_k^3(t)$, $\mathbb{E}(\xi_k^2(t)u_k^4(t))$ and $\mathbb{E}(\xi_k^3(t)u_k^2(t))$, and finally $\mathbb{E}\xi_k^4(t)$ as follows.

$$\begin{aligned}\mathbb{E}\xi_k^3(t) = & -\frac{147b^6}{64a^6} + \frac{9b^4c^2}{32a^5} + \frac{9b^2c^4}{16a^4} + \frac{c^6}{8a^3} + \frac{15b^6e^{-6at}}{64a^6} - \frac{45b^4c^2e^{-6at}}{32a^5} \\ & + \frac{15b^2c^4e^{-6at}}{16a^4} - \frac{c^6e^{-6at}}{8a^3} + \frac{63b^6e^{-4at}}{64a^6} - \frac{45b^4c^2e^{-4at}}{32a^5} - \frac{21b^2c^4e^{-4at}}{16a^4} \\ & + \frac{3c^6e^{-4at}}{8a^3} + \frac{69b^6e^{-2at}}{64a^6} + \frac{81b^4c^2e^{-2at}}{32a^5} - \frac{3b^2c^4e^{-2at}}{16a^4} - \frac{3c^6e^{-2at}}{8a^3} \\ & + \frac{9b^6t}{32a^5} + \frac{9b^4c^2t}{8a^4} + \frac{3b^2c^4t}{8a^3} + \frac{81b^6e^{-4at}t}{32a^5} - \frac{81b^4c^2e^{-4at}t}{8a^4} \\ & + \frac{27b^2c^4e^{-4at}t}{8a^3} + \frac{75b^6e^{-2at}t}{16a^5} - \frac{15b^2c^4e^{-2at}t}{4a^3} + \frac{9b^6t^2}{16a^4} + \frac{3b^4c^2t^2}{8a^3} \\ & + \frac{75b^6e^{-2at}t^2}{16a^4} - \frac{75b^2c^2e^{-2at}t^2}{8a^3} + \frac{b^6t^3}{8a^3},\end{aligned}$$

$$\begin{aligned}\mathbb{E}(\xi_k^2(t)u_k^4(t)) = & \frac{69b^8}{64a^6} + \frac{15b^6c^2}{16a^5} + \frac{3b^4c^4}{16a^4} + \frac{105b^8e^{-8at}}{64a^6} - \frac{105b^6c^2e^{-8at}}{8a^5} + \frac{105b^4c^4e^{-8at}}{8a^4} \\ & - \frac{7b^2c^6e^{-8at}}{2a^3} + \frac{c^8e^{-8at}}{4a^2} + \frac{195b^6c^2e^{-6at}}{16a^5} - \frac{165b^4c^4e^{-6at}}{8a^4} + \frac{27b^2c^6e^{-6at}}{4a^3} \\ & - \frac{c^8e^{-6at}}{2a^2} - \frac{123b^8e^{-4at}}{32a^6} + \frac{111b^6c^2e^{-4at}}{16a^5} + \frac{141b^4c^4e^{-4at}}{16a^4} - \frac{4b^2c^6e^{-4at}}{a^3} \\ & + \frac{c^8e^{-4at}}{4a^2} + \frac{9b^8e^{-2at}}{8a^6} - \frac{111b^6c^2e^{-2at}}{16a^5} - \frac{3b^4c^4e^{-2at}}{2a^4} + \frac{3b^2c^6e^{-2at}}{4a^3} \\ & + \frac{15b^8t}{16a^5} + \frac{3b^6c^2t}{8a^4} + \frac{195b^8e^{-6at}t}{16a^5} - \frac{585b^6c^2e^{-6at}t}{8a^4} + \frac{195b^4c^4e^{-6at}t}{4a^3} \\ & - \frac{13b^2c^6e^{-6at}t}{2a^2} - \frac{135b^8e^{-4at}t}{16a^5} + \frac{513b^6c^2e^{-4at}t}{8a^4} - \frac{171b^4c^4e^{-4at}t}{4a^3} \\ & + \frac{9b^2c^6e^{-4at}t}{2a^2} - \frac{75b^8e^{-2at}t}{16a^5} - \frac{75b^6c^2e^{-2at}t}{8a^4} + \frac{15b^4c^4e^{-2at}t}{2a^3} + \frac{3b^8t^2}{16a^4} \\ & + \frac{243b^8e^{-4at}t^2}{16a^4} - \frac{243b^6c^2e^{-4at}t^2}{4a^3} + \frac{81b^4c^4e^{-4at}t^2}{4a^2} - \frac{75b^8e^{-2at}t^2}{8a^4} \\ & + \frac{75b^6c^2e^{-2at}t^2}{4a^3},\end{aligned}$$

$$\begin{aligned}\mathbb{E}(\xi_k^3(t)u_k^2(t)) = & -\frac{9b^8}{128a^7} + \frac{69b^6c^2}{64a^6} + \frac{15b^4c^4}{32a^5} + \frac{b^2c^6}{16a^4} - \frac{105b^8e^{-8at}}{128a^7} + \frac{105b^6c^2e^{-8at}}{16a^6} \\ & - \frac{105b^4c^4e^{-8at}}{16a^5} + \frac{7b^2c^6e^{-8at}}{4a^4} - \frac{c^8e^{-8at}}{8a^3} - \frac{75b^8e^{-6at}}{32a^7} + \frac{315b^6c^2e^{-6at}}{64a^6} \\ & + \frac{195b^4c^4e^{-6at}}{32a^5} - \frac{61b^2c^6e^{-6at}}{16a^4} + \frac{3c^8e^{-6at}}{8a^3} - \frac{27b^8e^{-4at}}{64a^7} - \frac{621b^6c^2e^{-4at}}{32a^5}\end{aligned}$$

$$
\begin{aligned}
&+\frac{111b^4c^4e^{-4at}}{32a^5}+\frac{39b^2c^6e^{-4at}}{16a^4}-\frac{3c^8e^{-4at}}{8a^3}+\frac{117b^8e^{-2at}}{32a^7}-\frac{183b^6c^2e^{-2at}}{64a^6}\\
&-\frac{111b^4c^4e^{-2at}}{32a^5}-\frac{7b^2c^6e^{-2at}}{16a^4}+\frac{c^8e^{-2at}}{8a^3}+\frac{69b^8t}{64a^6}+\frac{15b^6c^2t}{16a^5}+\frac{3b^4c^4t}{16a^4}\\
&-\frac{585b^8e^{-6at}t}{64a^6}+\frac{1755b^6c^2e^{-6at}t}{32a^5}-\frac{585b^4c^4e^{-6at}t}{16a^4}+\frac{39b^2c^6e^{-6at}t}{8a^3}\\
&-\frac{729b^8e^{-4at}t}{64a^6}+\frac{513b^4c^4e^{-4at}t}{16a^4}-\frac{27b^2c^6e^{-4at}t}{4a^3}+\frac{285b^8e^{-2at}t}{64a^6}\\
&-\frac{585b^6c^2e^{-2at}t}{32a^5}-\frac{75b^4c^4e^{-2at}t}{16a^4}+\frac{15b^2c^6e^{-2at}t}{8a^3}+\frac{15b^8t^2}{32a^5}+\frac{3b^6c^2t^2}{16a^4}\\
&-\frac{729b^8e^{-4at}t^2}{32a^5}+\frac{729b^6c^2e^{-4at}t^2}{8a^4}-\frac{243b^4c^4e^{-4at}t^2}{8a^3}-\frac{75b^8e^{-2at}t^2}{16a^5}\\
&-\frac{225b^6c^2e^{-2at}t^2}{16a^4}+\frac{75b^4c^4e^{-2at}t^2}{8a^3}+\frac{b^8t^3}{16a^4}-\frac{125b^8e^{-2at}t^3}{16a^4}\\
&+\frac{125b^6c^2e^{-2at}t^3}{8a^3},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\xi_k^4(t) = &-\frac{3519b^8}{256a^8}-\frac{9b^6c^2}{32a^7}+\frac{69b^4c^4}{32a^6}+\frac{5b^2c^6}{8a^5}+\frac{c^8}{16a^4}+\frac{105b^8e^{-8at}}{256a^8}\\
&-\frac{105b^6c^2e^{-8at}}{32a^7}+\frac{105b^4c^4e^{-8at}}{32a^6}-\frac{7b^2c^6e^{-8at}}{8a^5}+\frac{c^8e^{-8at}}{16a^4}+\frac{165b^8e^{-6at}}{64a^8}\\
&-\frac{75b^6c^2e^{-6at}}{8a^7}+\frac{2b^2c^6e^{-6at}}{a^5}-\frac{c^8e^{-6at}}{4a^4}+\frac{783b^8e^{-4at}}{128a^8}-\frac{27b^6c^2e^{-4at}}{16a^7}\\
&-\frac{123b^4c^4e^{-4at}}{16a^6}-\frac{3b^2c^6e^{-4at}}{4a^5}+\frac{3c^8e^{-4at}}{8a^4}+\frac{297b^8e^{-2at}}{64a^8}+\frac{117b^6c^2e^{-2at}}{8a^7}\\
&+\frac{9b^4c^4e^{-2at}}{4a^6}-\frac{b^2c^6e^{-2at}}{a^5}-\frac{c^8e^{-2at}}{4a^4}-\frac{9b^8t}{32a^7}+\frac{69b^6c^2t}{16a^6}+\frac{15b^4c^4t}{8a^5}\\
&+\frac{b^2c^6t}{4a^4}+\frac{195b^8e^{-6at}t}{32a^7}-\frac{585b^6c^2e^{-6at}t}{16a^6}+\frac{195b^4c^4e^{-6at}t}{8a^5}-\frac{13b^2c^6e^{-6at}t}{4a^4}\\
&+\frac{729b^8e^{-4at}t}{32a^7}-\frac{729b^6c^2e^{-4at}t}{16a^6}-\frac{135b^4c^4e^{-4at}t}{8a^5}+\frac{27b^2c^6e^{-2at}t}{4a^4}\\
&+\frac{765b^8e^{-2at}t}{32a^7}+\frac{285b^6c^2e^{-2at}t}{16a^6}-\frac{75b^4c^4e^{-2at}t}{8a^5}-\frac{15b^2c^6e^{-2at}t}{4a^4}+\frac{69b^8t^2}{32a^6}\\
&+\frac{15b^6c^2t^2}{8a^5}+\frac{3b^4c^4t^2}{8a^4}+\frac{729b^8e^{-4at}t^2}{32a^6}-\frac{729b^6c^2e^{-4at}t^2}{8a^5}+\frac{243b^4c^4e^{-4at}t^2}{8a^4}\\
&+\frac{525b^8e^{-2at}t^2}{16a^6}-\frac{75b^6c^2e^{-2at}t^2}{4a^5}-\frac{75b^4c^4e^{-2at}t^2}{4a^4}+\frac{5b^8t^3}{8a^5}+\frac{b^6c^2t^3}{4a^4}\\
&+\frac{125b^8e^{-2at}t^3}{8a^5}-\frac{125b^6c^2e^{-2at}t^3}{4a^4}+\frac{b^8t^4}{16a^4}.
\end{aligned}
$$

A similar argument leads to the computation of the second expectation in (3.19). To sum up, with the help of Mathematica, we obtain that

$$
\begin{aligned}
\mathrm{Var}(Z_k) = & \frac{-16917b^8}{512a^{10}} + \frac{1093b^6c^2}{32a^9} + \frac{303b^4c^4}{32a^8} + \frac{35b^2c^6}{48a^7} + \frac{3b^8e^{-8aT}}{128a^{10}} - \frac{3b^6c^2e^{-8aT}}{16a^9} \\
& + \frac{21b^4c^4e^{-8aT}}{128a^8} - \frac{b^2c^6e^{-8aT}}{32a^7} + \frac{79b^8e^{-6aT}}{128a^{10}} - \frac{3b^6c^2e^{-6aT}}{a^9} + \frac{19b^4c^4e^{-6aT}}{16a^8} \\
& + \frac{b^2c^6e^{-6aT}}{48a^7} + \frac{2953b^8e^{-4aT}}{512a^{10}} - \frac{447b^6c^2e^{-4aT}}{32a^9} - \frac{65b^4c^4e^{-4aT}}{128a^8} + \frac{15b^2c^6e^{-4aT}}{32a^7} \\
& + \frac{3409b^8e^{-2aT}}{128a^{10}} - \frac{17b^6c^2e^{-2aT}}{a^9} - \frac{165b^4c^4e^{-2aT}}{16a^8} - \frac{19b^2c^6e^{-2aT}}{16a^7} + \frac{1093b^8T}{32a^9} \\
& - \frac{659b^6c^2T}{32a^8} - \frac{155b^4c^4T}{16a^7} - \frac{9b^2c^6T}{8a^6} + \frac{45b^8e^{-6aT}T}{64a^9} - \frac{135b^6c^2e^{-6aT}T}{32a^8} \\
& + \frac{5b^4c^4e^{-6aT}T}{2a^7} - \frac{b^2c^6e^{-6aT}T}{4a^6} + \frac{1165b^8e^{-4aT}T}{128a^9} - \frac{953b^6c^2e^{-4aT}T}{32a^8} \\
& + \frac{127b^4c^4e^{-4aT}T}{32a^7} + \frac{3b^2c^6e^{-4aT}T}{4a^6} + \frac{2321b^8e^{-2aT}T}{64a^9} - \frac{1753b^6c^2e^{-2aT}T}{32a^8} \\
& - \frac{11b^4c^4e^{-2aT}T}{a^7} - \frac{659b^8T^2}{64a^8} - \frac{5b^6c^2T^2}{4a^7} + \frac{31b^4c^4T^2}{16a^6} + \frac{b^2c^6T}{2a^5} \\
& + \frac{53b^8e^{-4aT}T^2}{16a^8} - \frac{121b^6c^2e^{-4aT}T^2}{8a^7} + \frac{21b^4c^4e^{-4aT}T^2}{4a^6} - \frac{b^2c^6e^{-4aT}T^2}{4a^5} \\
& + \frac{71b^8e^{-2aT}T^2}{8a^8} - \frac{255b^6c^2e^{-2aT}T^2}{8a^7} - \frac{b^4c^4e^{-2aT}T^2}{4a^6} + \frac{b^2c^6e^{-2aT}T^2}{2a^5} \\
& - \frac{5b^8T^3}{12a^7} + \frac{23b^6c^2T^3}{12a^6} + \frac{2b^4c^4T^3}{3a^5} - \frac{5b^8e^{-4aT}T^3}{8a^7} + \frac{5b^6c^2e^{-4aT}T^3}{2a^6} \\
& - \frac{5b^4c^4e^{-4aT}T^3}{6a^5} - \frac{113b^8e^{-2aT}T^3}{24a^7} - \frac{7b^6c^2e^{-2aT}T^3}{12a^6} + \frac{7b^4c^4e^{-2aT}T^3}{2a^5} \\
& + \frac{23b^8T^4}{48a^6} + \frac{b^6c^2T^4}{3a^5} - \frac{5b^8e^{-2aT}T^4}{2a^6} + \frac{5b^6c^2e^{-2aT}T^4}{a^5} + \frac{b^8T^5}{15a^5}.
\end{aligned}
$$

which implies (3.14). The proof to (3.15) and (3.16) are done similarly through the symbolic computation in Mathematica®. We omit the details here. The code for the detailed computation can be obtained from the author upon request. □

Remark 3.2.1. *The above proposition implies that*

$$
\mathbb{E}(Z_k) \asymp \frac{1}{\mu_k^2(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^2, \tag{3.20}
$$

$$\operatorname{Var}(Z_k) \asymp \frac{\lambda_k^{-2\gamma}}{\mu_k^5(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^3, \tag{3.21}$$

$$\mathbb{E}(A_k) \asymp \frac{\lambda_k^{-2\gamma}}{\mu_k^2(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right), \tag{3.22}$$

$$\operatorname{Var}(A_k) \asymp \frac{\lambda_k^{-2\gamma}}{\mu_k^3(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^3. \tag{3.23}$$

These formulas will be used to obtain the exact asymptotic bias and the exact rate of convergence in the proof of asymptotic normality.

3.2.4 The Consistency of TFE. In this subsection we prove the large-space consistency of the TFE $\widetilde{\theta}_N$, as $N \to \infty$. The proof relies on the classical version of the Strong Law of Large Numbers (cf. [Shi96, Theorem IV.3.2]), which we state in the Appendix for sake of completeness. With this at hand, we now present the first main result of this section.

Theorem 3.2.1 (Consistency of TFE)**.** *Let the assumptions* (i) - (v) *be satisfied. Moreover, assume that*

$$\sum_{k=1}^{\infty} \nu_k^2 \, \mathbb{E}(Z_k) = \infty. \tag{3.24}$$

Then,

$$\lim_{N\to\infty} \widetilde{\theta}_N = \theta, \quad \mathbb{P} - a.\ s..$$

Proof. By (3.12),

$$\widetilde{\theta}_N - \theta = -\frac{\sum_{k=1}^N \nu_k A_k}{2\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} \cdot \frac{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}{\sum_{k=1}^N \nu_k^2 Z_k}. \tag{3.25}$$

We first study the second factor in (3.25). Clearly,

$$\begin{aligned}\sum_{N=1}^{\infty} \frac{\nu_N^4 \operatorname{Var}(Z_N)}{\left(\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)\right)^2} &\leq \frac{\operatorname{Var}(Z_1)}{(\mathbb{E}(Z_1))^2} + \sum_{N=2}^{\infty} \frac{\nu_N^4 \operatorname{Var}(Z_N)}{\sum_{k=1}^{N-1} \nu_k^2 \mathbb{E}(Z_k) \cdot \sum_{k=1}^{N} \nu_k^2 \mathbb{E} Z_k} \\ &= \frac{\operatorname{Var}(Z_1)}{(\mathbb{E}(Z_1))^2} + \sum_{N=2}^{\infty} \frac{\nu_N^2 \operatorname{Var}(Z_N)}{\mathbb{E}(Z_N)} \left(\frac{1}{\sum_{k=1}^{N-1} \nu_k^2 \mathbb{E}(Z_k)} - \frac{1}{\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)} \right). \end{aligned} \tag{3.26}$$

By (3.20), (3.21) and the assumption (iii), as $N \to \infty$,

$$\begin{aligned}\frac{\nu_N^2 \operatorname{Var}(Z_N)}{\mathbb{E}(Z_N)} &= O\left(\nu_N^2 \cdot \frac{\frac{\lambda_N^{-2\gamma}}{\mu_N^5(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right)^3}{\frac{1}{\mu_N^2(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right)^2} \right) \\ &= O\left(\frac{\lambda_N^{-2\gamma}}{\mu_N(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right) \right).\end{aligned}$$

Since $u_0 \in H$, we have

$$\lim_{N\to\infty} u_N^2(0) = 0. \tag{3.27}$$

Together with the assumptions (ii), (iv) and (v),

$$\lim_{N\to\infty} \frac{\lambda_N^{-2\gamma}}{\mu_N(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right) = \lim_{N\to\infty} \frac{1}{\mu_N(\theta) \nu_N^{\gamma/m}} \left(u_N^2(0) + \frac{\sigma^2 T}{\nu_N^{\gamma/m}} \right) = 0.$$

Hence, there exists a universal constant $C_1 > 0$ such that

$$\frac{\nu_N^2 \operatorname{Var}(Z_N)}{\mathbb{E}(Z_N)} \leq C_1 \quad \text{for all } N \in \mathbb{N},$$

which, together with (3.26), implies that

$$\begin{aligned}\sum_{N=1}^{\infty} \frac{\nu_N^4 \operatorname{Var}(Z_N)}{\left(\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)\right)^2} &\leq \frac{\operatorname{Var}(Z_1)}{(\mathbb{E}(Z_1))^2} + C_1 \sum_{N=2}^{\infty} \left(\frac{1}{\sum_{k=1}^{N-1} \nu_k^2 \mathbb{E}(Z_k)} - \frac{1}{\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)} \right) \\ &= \frac{\operatorname{Var}(Z_1)}{(\mathbb{E}(Z_1))^2} + \frac{C_1}{\nu_1^2 \mathbb{E}(Z_1)} < \infty. \end{aligned} \tag{3.28}$$

Combining (3.24) with (3.28), we conclude by Remark B.0.1 that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \nu_k^2 Z_k}{\sum_{k=1}^{N} \nu_k^2 \,\mathbb{E}(Z_k)} = 1, \quad \mathbb{P}-\text{a.s.}. \tag{3.29}$$

Next, we will analyze the asymptotic behavior of the first factor in (3.25). By (3.20), (3.23), (3.27) and the assumptions (ii), (iv) and (v), as $N \to \infty$, we get that

$$\begin{aligned}
\frac{\operatorname{Var}(A_N)}{\mathbb{E}(Z_N)} &= O\left(\frac{\frac{\lambda_N^{-2\gamma}}{\mu_N^3(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right)^3}{\frac{1}{\mu_N^2(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma}\right)^2} \right) \\
&= O\left(\frac{1}{\mu_N(\theta)\nu_N^{\gamma/m}} \left(u_N^2(0) + \frac{\sigma^2 T}{\nu_N^{\gamma/m}} \right) \right) \to 0.
\end{aligned}$$

An argument similar to that of (3.26) and (3.28) shows that, there exists a universal constant $C_2 > 0$, such that

$$\begin{aligned}
\sum_{N=1}^{\infty} \frac{\nu_N^2 \operatorname{Var}(A_N)}{\left(\sum_{k=1}^{N} \nu_k^2 \,\mathbb{E}(Z_k)\right)^2} &\le \frac{\operatorname{Var}(A_1)}{\nu_1^2 \left(\mathbb{E}(Z_1)\right)^2} \\
&\quad + \sum_{N=2}^{\infty} \frac{\operatorname{Var}(A_N)}{\mathbb{E}(Z_N)} \left(\frac{1}{\sum_{k=1}^{N-1} \nu_k^2 \,\mathbb{E}(Z_k)} - \frac{1}{\sum_{k=1}^{N} \nu_k^2 \,\mathbb{E}(Z_k)} \right) \\
&\le \frac{\operatorname{Var}(A_1)}{\nu_1^2 \left(\mathbb{E}(Z_1)\right)^2} + \frac{C_2}{\nu_1^2 \,\mathbb{E}(Z_1)} < \infty.
\end{aligned} \tag{3.30}$$

In view of Theorem B.0.1, (3.24) and (3.30) imply that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \nu_k \left(A_k - \mathbb{E}(A_k)\right)}{\sum_{k=1}^{N} \nu_k^2 \,\mathbb{E}(Z_k)} = 0, \quad \mathbb{P}-\text{a.s.}.$$

If the series $\sum_{k=1}^{N} \nu_k \,\mathbb{E}(A_k)$ converges, then by (3.24), we clearly have that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \nu_k \,\mathbb{E}(A_k)}{\sum_{k=1}^{N} \nu_k^2 \,\mathbb{E}(Z_k)} = 0, \tag{3.31}$$

On the other hand, if the series in the numerator of (3.31) diverges, then by Stolz–

Cesàro Theorem

$$\lim_{N\to\infty} \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} = \lim_{N\to\infty} \frac{\nu_N \, \mathbb{E}(A_N)}{\nu_N^2 \, \mathbb{E}(Z_N)},$$

and by (3.20), (3.22) and the assumption (iv), as $N \to \infty$, we deduce that

$$\begin{aligned} \frac{\mathbb{E}(A_N)}{\nu_N \, \mathbb{E}(Z_N)} &= O\left(\frac{\frac{\lambda_N^{-2\gamma}}{\mu_N^2(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma} \right)}{\frac{\nu_N}{\mu_N^2(\theta)} \left(u_N^2(0) + \sigma^2 T \lambda_N^{-2\gamma} \right)^2} \right) = O\left(\frac{1}{\nu_N \left(u_N^2(0) \lambda_N^{2\gamma} + \sigma^2 T \right)} \right) \\ &= O\left(\nu_N^{-1}\right) \to 0. \end{aligned}$$

Combining the above, we conclude that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^N \nu_k A_k}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} = 0, \quad \mathbb{P} - \text{a. s.}. \tag{3.32}$$

Finally, by (3.25), (3.29) and (3.32) we conclude the proof. □

Remark 3.2.2. *A note on condition* (3.24) *is in order. The divergence of the series* (3.24) *is needed for the Law of Large Numbers to hold true. In view of* (3.20)*, the condition* (3.24) *can be equivalently stated in terms of the known primary objects – the initial data* $u(0)$*, the asymptotics of the eigenvalues of* $\mathcal{A}_0$ *and* $\mathcal{A}_1$*, and the covariance structure of the noise (see Proposition 3.2.2 below). In particular, the consistency of the TFE does not depend on the regularity of the initial data.*

3.2.5 The Asymptotic Normality of TFE. In this subsection, we will investigate the asymptotic normality of the TFE $\widetilde{\theta}_N$. The proof is based on classical Central Limit Theorem (CLT) for independent random variables with the Lyapunov condition (which is a sufficient condition for the Lindeberg condition to hold). For convenience, we list this result in the Appendix, and the complete proof can be found, for instance, in [Shi96, Section III.4].

In what follows we will make use of the following technical lemma.

Lemma 3.2.1. *Let $\xi_k(t)$, $k \in \mathbb{N}$, $t \in [0,T]$, be defined as in (3.10). Then, for any $n \in \mathbb{N}$, there exist a constant $D_n = D_n(t) > 0$, depending only on n and t, such that, for every $k \in \mathbb{N}$,*

$$\mathbb{E}\left(\xi_k^n(t)\right) \leq D_n \left(\frac{u_k^2(0) + \sigma^2 t\, \lambda_k^{-2\gamma}}{\mu_k(\theta)}\right)^n, \quad t \in [0,T].$$

Proof. By (3.5) and Cauchy–Schwartz inequality, for any $0 \leq s \leq t \leq T$,

$$\begin{aligned} u_k^2(s) &= e^{-2\mu_k(\theta)s}\left(u_k(0) + \sigma\lambda_k^{-\gamma}\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2 \\ &\leq e^{-2\mu_k(\theta)s}\left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t\right)\left(1 + \frac{1}{t}\left(\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2\right). \end{aligned}$$

Hence, for any $t \in [0,T]$, and $n \in \mathbb{N}$,

$$\begin{aligned} \xi_k^n(t) &\leq \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t\right)^n \left\{\int_0^t e^{-2\mu_k(\theta)s}\left[1 + \frac{1}{t}\left(\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2\right] ds\right\}^n \\ &= \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t\right)^n \left[\frac{1 - e^{-2\mu_k(\theta)t}}{2\mu_k(\theta)} + \frac{1}{t}\int_0^t e^{-2\mu_k(\theta)s}\left(\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2 ds\right]^n \\ &\leq \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t\right)^n \\ &\quad \cdot \left\{\left(\frac{1 - e^{-2\mu_k(\theta)t}}{\mu_k(\theta)}\right)^n + \frac{2^n}{t^n}\left[\int_0^t e^{-2\mu_k(\theta)s}\left(\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2 ds\right]^n\right\}. \end{aligned}$$

By [Lot09, Theorem 2.1], there exists a constant $\widetilde{D}_n = \widetilde{D}_n(t) > 0$, such that

$$\mathbb{E}\left(\left[\int_0^t e^{-2\mu_k(\theta)s}\left(\int_0^s e^{\mu_k(\theta)r}\, dw_k(r)\right)^2 ds\right]^n\right) \leq \frac{\widetilde{D}_n}{\mu_k^n(\theta)}.$$

Therefore, for any $t \in [0,T]$ and $n \in \mathbb{N}$,

$$\mathbb{E}\left(\xi_k^n(t)\right) \leq \left(u_k^2(0) + \sigma^2\lambda_k^{-2\gamma}t\right)^n \left(\frac{1}{\mu_k^n(\theta)} + \frac{2^n\widetilde{D}_n}{t^n}\frac{1}{\mu_k^n(\theta)}\right) = D_n\left(\frac{u_k^2(0) + \sigma\lambda_k^{-2\gamma}t}{\mu_k(\theta)}\right)^n,$$

where $D_n = D_n(t) := 1 + 2^n t^{-n} \widetilde{D}_n$. $\square$

Now we present a version of the large-space asymptotic normality of the TFE $\widetilde{\theta}_N$.

Theorem 3.2.2 (Asymptotic Normality of TFE). *In addition to the conditions of Theorem 3.2.1, assume further that*

$$\sum_{k=1}^{\infty} \nu_k^2 \operatorname{Var}(A_k) = \infty. \tag{3.33}$$

Then, as $N \to \infty$,

$$\frac{\widetilde{\theta}_N - \theta + a_N}{b_N} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

where

$$a_N := \frac{\sum_{k=1}^{N} \nu_k \mathbb{E}(A_k)}{2\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)}, \quad b_N := \frac{\sqrt{\sum_{k=1}^{N} \nu_k^2 \operatorname{Var}(A_k)}}{2\sum_{k=1}^{N} \nu_k^2 \mathbb{E}(Z_k)},$$

and where $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution.

Proof. The proof is split in two steps.

Step 1. We will first show that the sequence $\{\nu_k A_k\}_{k\in\mathbb{N}}$ satisfies the Lyapunov condition (B.1) with $\delta = 2$. Clearly,

$$\begin{aligned}
&\mathbb{E}\left((A_k - \mathbb{E}(A_k))^4\right) \\
&= \mathbb{E}\left(A_k^4\right) - 4\,\mathbb{E}\left(A_k^3\right)\mathbb{E}(A_k) + 6\,\mathbb{E}\left(A_k^2\right)(\mathbb{E}(A_k))^2 - 3\,(\mathbb{E}(A_k))^4 \\
&= \mathbb{E}\left(A_k^4\right) - 4\,\mathbb{E}\left(A_k^3\right)\mathbb{E}(A_k) + 6\operatorname{Var}(A_k)(\mathbb{E}(A_k))^2 + 3\,(\mathbb{E}(A_k))^4.
\end{aligned} \tag{3.34}$$

We will estimate each term in the above expression separately. To begin with, for every $k \in \mathbb{N}$, let $\zeta_k := (\zeta_k(t))_{t\in[0,T]}$, where

$$\zeta_k(t) := \int_0^t u_k(s)\, dw_k(s), \quad t \in [0,T].$$

By (3.8) and (3.10), for any $k \in \mathbb{N}$ and $t \in [0,T]$,

$$u_k^2(t) = u_k^2(0) + \sigma^2 \lambda_k^{-2\gamma} t - 2\mu_k(\theta)\xi_k(t) + 2\sigma\lambda_k^{-\gamma}\zeta_k(t),$$

which, when multiplied by $\xi_k(t)$, and then integrated on $[0,T]$, leads to

$$A_k = 2\sigma\lambda_k^{-\gamma} \int_0^T \zeta_k(t)\xi_k(t)\, dt, \quad k \in \mathbb{N}.$$

Hence, by the Cauchy–Schwartz inequality and the definition of ξ_k, for any $k \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}\left(A_k^4\right) &\le 16\sigma^4\lambda_k^{-4\gamma}\, \mathbb{E}\left(\left(\int_0^T \zeta_k^2(t)\, dt \cdot \int_0^T \xi_k^2(t)\, dt\right)^2\right) \\
&\le 16\sigma^4\lambda_k^{-4\gamma} \left(\mathbb{E}\left(\left(\int_0^T \zeta_k^2(t) dt\right)^4\right) \mathbb{E}\left(\left(\int_0^T \xi_k^2(t) dt\right)^4\right)\right)^{1/2} \\
&\le 16\sigma^4\lambda_k^{-4\gamma} \left(T^2\, \mathbb{E}\left(\left(\int_0^T \zeta_k^4(t)\, dt\right)^2\right) \cdot T^4\, \mathbb{E}\left(\xi_k^8\right)\right)^{1/2} \\
&\le 16\sigma^4 T^{7/2}\lambda_k^{-4\gamma} \left(\mathbb{E}\left(\int_0^T \zeta_k^8(t)\, dt\right) \mathbb{E}\left(\xi_k^8\right)\right)^{1/2}.
\end{aligned}$$

Moreover, by the Burkholder–Davis–Gundy inequality, there exists a constant $C_1 = C_1(T) > 0$, depending only on T, such that

$$\mathbb{E}\left(\sup_{t\in[0,T]} \zeta_k^8(t)\right) \le C_1\, \mathbb{E}\left([\zeta_k, \zeta_k]^4(T)\right) = C_1\, \mathbb{E}\left(\xi_k^4\right).$$

Together with Lemma 3.2.1, we obtain that, for any $k \in \mathbb{N}$,

$$\mathbb{E}\left(A_k^4\right) \leq 16\, C_1 \sigma^4\, T^4 \lambda_k^{-4\gamma} \sqrt{\mathbb{E}\left(\xi_k^4\right) \mathbb{E}\left(\xi_k^8\right)} \leq C_2\, \frac{\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^8}{\mu_k^6(\theta)},$$

where $C_2 := 16\, C_1 \sqrt{D_4 D_8} T^2 > 0$ is a constant depending only on T.

Next, we will study the last three terms of (3.34). In view of (3.22) and (3.23), there exists a universal constant $C_3 > 0$, such that for any $k \in \mathbb{N}$,

$$\mathbb{E}(A_k) \leq C_3\, \frac{\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^2}{\mu_k^2(\theta)}, \quad \operatorname{Var}(A_k) \leq C_2\, \frac{\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^4}{\mu_k^3(\theta)}.$$

Hence, it suffices to estimate $\mathbb{E}(A_k^3)$. By the definition of A_k in (3.12),

$$-u_k^2(0) Y_k - \sigma^2 \lambda_k^{-2\gamma} X_k \leq A_k \leq \frac{1}{2}\xi_k^2 + 2\mu_k(\theta) Z_k.$$

Moreover, since $\xi_k(t)$ is increasing in t, we deduce that

$$Y_k = \int_0^T \xi_k(t)\, dt \leq T\xi_k, \quad X_k = \int_0^T t\xi_k(t)\, dt \leq T^2\xi_k, \quad Z_k = \int_0^T \xi_k^2(t)\, dt \leq T\xi_k^2,$$

and thus,

$$-T\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)\xi_k \leq A_k \leq \left(\frac{1}{2} + 2\mu_k(\theta) T\right)\xi_k^2.$$

Together with Lemma 3.2.1, we obtain that

$$-D_3 T \frac{\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^6}{\mu_k^3(\theta)} \leq \mathbb{E}(A_k^3) \leq D_6(2T+1) \frac{\left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^6}{\mu_k^3(\theta)}. \quad (3.35)$$

Combining (3.34)–(3.35), we conclude that there exists a constant $C_4 = C_4(T) >$

0, depending only on T, such that for any $k \in \mathbb{N}$,

$$\mathbb{E}\left((\nu_k A_k - \mathbb{E}(\nu_k A_k))^4\right) \leq C_4 \, \frac{\nu_k^4}{\mu_k^5(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^8 .$$

On the other hand, by (3.23) again, we can find a universal constant $C_5 > 0$, such that

$$\operatorname{Var}(\nu_k A_k) \geq C_5 \frac{\nu_k^2}{\mu_k^3(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^4, \quad \text{for all } \ k \in \mathbb{N}.$$

In view of the assumptions (ii)-(v), and since $\lim_{k\to\infty} u_k^2(0) = 0$, we deduce that there exists a constant $C_6 = C_6(c_0, \sigma, T) > 0$, depending on c_0, σ and T, such that

$$\mathbb{E}\left((\nu_k A_k - \mathbb{E}(\nu_k A_k))^4\right) \leq C_6 \operatorname{Var}(\nu_k A_k), \quad \text{for all } \ k \in \mathbb{N}.$$

Finally, by (3.33), we obtain that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^N \mathbb{E}\left((\nu_k A_k - \mathbb{E}(\nu_k A_k))^4\right)}{\left(\sum_{k=1}^N \operatorname{Var}(\nu_k A_k)\right)^2} \leq \lim_{N\to\infty} \frac{C_6}{\sum_{k=1}^N \operatorname{Var}(\nu_k A_k)} = 0.$$

Step 2: Note that

$$\frac{\widetilde{\theta}_N - \theta + a_N}{b_N} = -\frac{\sum_{k=1}^N \nu_k \left(A_k - \mathbb{E}(A_k)\right)}{2\, b_N \sum_{k=1}^N \nu_k^2 Z_k} - \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{2\, b_N \sum_{k=1}^N \nu_k^2 Z_k} + \frac{a_N}{b_N}. \tag{3.36}$$

For the first term in (3.36), by (3.29), Step 1 and Theorem B.0.2, as $N \to \infty$, we get

$$-\frac{\sum_{k=1}^N \nu_k \left(A_k - \mathbb{E}(A_k)\right)}{2\, b_N \sum_{k=1}^N \nu_k^2 Z_k} = -\frac{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}{\sum_{k=1}^N \nu_k^2 Z_k} \cdot \frac{\sum_{k=1}^N \nu_k \left(A_k - \mathbb{E}(A_k)\right)}{\sqrt{\sum_{k=1}^N \nu_k^2 \operatorname{Var}(A_k)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Moreover, for the last two terms in (3.36), we derive that

$$\frac{a_N}{b_N} - \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{2\, b_N \sum_{k=1}^N \nu_k^2 Z_k} = \frac{2\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}{\sqrt{\sum_{k=1}^N \nu_k^2 \, \mathrm{Var}(A_k)}} \left(\frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{2\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} - \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{2\sum_{k=1}^N \nu_k^2 Z_k} \right)$$

$$= \frac{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}{\sum_{k=1}^N \nu_k^2 Z_k} \cdot \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{\sqrt{\sum_{k=1}^N \nu_k^2 \, \mathrm{Var}(A_k)}} \cdot \frac{\sum_{k=1}^N \nu_k^2 \left(Z_k - \mathbb{E}(Z_k)\right)}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}. \tag{3.37}$$

In light of (3.29), we only need to show that the product of the last two factors above converges to zero in probability, as $N \to \infty$. Note that, by the independence of Z_k, $k \in \mathbb{N}$,

$$\mathbb{E}\left(\left(\frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{\sqrt{\sum_{k=1}^N \nu_k^2 \, \mathrm{Var}(A_k)}} \cdot \frac{\sum_{k=1}^N \nu_k^2 \left(Z_k - \mathbb{E}(Z_k)\right)}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} \right)^2 \right)$$
$$= \left(\frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} \right)^2 \frac{\sum_{k=1}^N \nu_k^4 \, \mathrm{Var}(Z_k)}{\sum_{k=1}^N \nu_k^2 \, \mathrm{Var}(A_k)}.$$

By (3.21), (3.23) and the assumption (iii), there exists a universal constant $C_7 > 0$, such that

$$\frac{\sum_{k=1}^N \nu_k^4 \, \mathrm{Var}(Z_k)}{\sum_{k=1}^N \nu_k^2 \, \mathrm{Var}(A_k)} \leq C_7 \frac{\sum_{k=1}^N \frac{\nu_k^4 \lambda_k^{-2\gamma}}{\mu_k^5(\theta)} (u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma})^3}{\sum_{k=1}^N \frac{\nu_k^2 \lambda_k^{-2\gamma}}{\mu_k^3(\theta)} (u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma})^3} \leq C_7 c_0,$$

Similarly, by (3.20), (3.22) and the assumption (iii),

$$\frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)} \leq C_8 \frac{\sum_{k=1}^N \frac{\nu_k \lambda_k^{-2\gamma}}{\mu_k^2(\theta)} (u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma})}{\sum_{k=1}^N \frac{\nu_k^2}{\mu_k^2(\theta)} (u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma})^2}, \tag{3.38}$$

where $C_8 > 0$ is some universal constant. Using (3.20) and (3.24), we conclude that the series in the denominator on the right-hand side of (3.38) diverges. Hence, the right-hand side of (3.38) converges to 0, as $N \to \infty$, if the series in the numerator on the right-hand side of (3.38) converges. Now assume that the numerator on the right-hand side of (3.38) diverges. By Stolz–Cesàro Theorem,

$$\begin{aligned}\lim_{N\to\infty} \frac{\sum_{k=1}^N \frac{\nu_k\lambda_k^{-2\gamma}}{\mu_k^2(\theta)}\left(u_k^2(0)+\sigma^2 T\lambda_k^{-2\gamma}\right)}{\sum_{k=1}^N \frac{\nu_k^2}{\mu_k^2(\theta)}\left(u_k^2(0)+\sigma^2 T\lambda_k^{-2\gamma}\right)^2} &= \lim_{N\to\infty} \frac{\frac{\nu_N\lambda_N^{-2\gamma}}{\mu_N^2(\theta)}\left(u_N^2(0)+\sigma^2 T\lambda_N^{-2\gamma}\right)}{\frac{\nu_N^2}{\mu_N^2(\theta)}\left(u_N^2(0)+\sigma^2 T\lambda_N^{-2\gamma}\right)^2} \\ &\le \lim_{N\to\infty} \frac{1}{\sigma^2 T\nu_N} = 0.\end{aligned}$$

Therefore, we have shown that

$$\frac{\sum_{k=1}^N \nu_k\,\mathbb{E}(A_k)}{\sqrt{\sum_{k=1}^N \nu_k^2\,\mathrm{Var}(A_k)}} \cdot \frac{\sum_{k=1}^N \nu_k^2\left(Z_k-\mathbb{E}(Z_k)\right)}{\sum_{k=1}^N \nu_k^2\,\mathbb{E}(Z_k)} \to 0 \quad \text{in } L^2(\Omega), \quad N\to\infty. \tag{3.39}$$

Combining (3.29), (3.36), (3.37) and (3.39) completes the proof. □

The next result provides some equivalent conditions, given explicitly in terms of the model coefficients, for (3.24) and (3.33) to hold. In particular, we note that the consistency and the asymptotic normality of the TFE do not depend on the regularity of the initial data.

Proposition 3.2.2. *Under the assumptions* (i) - (v)*,*

$$\sum_{k=1}^\infty \nu_k^2\,\mathbb{E}(Z_k) = \infty \quad \Leftrightarrow \quad \sum_{k=}^\infty \frac{\nu_k^2\lambda_k^{-4\gamma}}{\mu_k^2(\theta)} = \infty, \tag{3.40}$$

$$\sum_{k=1}^\infty \nu_k^2\,\mathrm{Var}(A_k) = \infty \quad \Leftrightarrow \quad \sum_{k=1}^\infty \frac{\nu_k^2\lambda_k^{-8\gamma}}{\mu_k^3(\theta)} = \infty. \tag{3.41}$$

Proof. We will only present the proof for (3.41), as (3.40) can be obtained similarly. Clearly, (3.23) implies the "$\Leftarrow$" direction in (3.41). Now assume that

$$\sum_{k=1}^\infty \nu_k^2\,\mathrm{Var}(A_k) = \infty,$$

which, by (3.23), is equivalent to

$$\infty = \sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-2\gamma}}{\mu_k^3(\theta)} \left(u_k^2(0) + \sigma^2 T \lambda_k^{-2\gamma}\right)^3 = \sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-2\gamma} u_k^6(0)}{\mu_k^3(\theta)} + 3\sigma^2 T \sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-4\gamma} u_k^4(0)}{\mu_k^3(\theta)} + 3\sigma^4 T^2 \sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-6\gamma} u_k^2(0)}{\mu_k^3(\theta)} + \sigma^6 T^3 \sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-8\gamma}}{\mu_k^3(\theta)}.$$

Hence, it suffices to show that the first three series on the right-hand side above are all convergent. We will only check the first series, and the other two can be verified using a similar argument. Indeed, by the assumptions (ii) - (v) and since $u(0) \in H$ (so that $\lim_{k\to\infty} u_k(0) = 0$), there exists a universal constant $C > 0$ such that

$$\sum_{k=1}^{\infty} \frac{\nu_k^2 \lambda_k^{-2\gamma} u_k^6(0)}{\mu_k^3(\theta)} \leq c_0^2\, C \sum_{k=1}^{\infty} u_k^2(0) < \infty.$$

This concludes the proof. □

We conclude this section by providing the asymptotically equivalent formulas for the sequences $\{a_N\}_{N\in\mathbb{N}}$ and $\{b_N\}_{N\in\mathbb{N}}$ in Theorem 3.2.2, given in terms of the model coefficients. In light of Proposition 3.2.2, the relations (3.24) and (3.33) imply that each of the last terms in (3.13)–(3.16) give the exact leading order term for $\mathbb{E}(Z_k)$, $\operatorname{Var}(Z_k)$, $\mathbb{E}(A_k)$ and $\operatorname{Var}(A_k)$, respectively. The following result follows immediately from Stolz–Cesàro Theorem.

Corollary 3.2.1. *Under the conditions of Theorem 3.2.2, as $N \to \infty$, we have*

$$a_N \sim \frac{3}{T} \frac{\sum_{k=1}^{N} \frac{\nu_k \lambda_k^{-4\gamma}}{\mu_k^2(\theta)}}{\sum_{k=1}^{N} \frac{\nu_k^2 \lambda_k^{-4\gamma}}{\mu_k^2(\theta)}}, \quad b_N \sim \sqrt{\frac{12}{5T}} \frac{\sqrt{\sum_{k=1}^{N} \frac{\nu_k^2 \lambda_k^{-8\gamma}}{\mu_k^3(\theta)}}}{\sum_{k=1}^{N} \frac{\nu_k^2 \lambda_k^{-4\gamma}}{\mu_k^2(\theta)}}.$$

3.2.6 Examples. In this part, we will present two examples that illustrate the theoretical results obtained before. Throughout this section, let $G \subseteq \mathbb{R}^d$ be a smooth and bounded domain, $H := L^2(G)$, and let Δ be the Laplace operator on G with zero boundary condition. It is known (cf. [Shu01]) that Δ has a complete orthonormal system of eigenfunctions $\{h_k\}_{k\in\mathbb{N}}$ in H. Moreover, the corresponding eigenvalues $\{\tau_k\}_{k\in\mathbb{N}}$ can be arranged such that $0 \leq -\tau_1 \leq -\tau_2 \leq \cdots$, and there exists a universal constant $c_1 > 0$ so that

$$\lim_{k\to\infty} |\tau_k| \cdot k^{-2/d} = c_1.$$

Example 3.2.1. *Consider the following fractional stochastic heat equation driven by an additive noise, possibly colored in space,*

$$du(t,x) + \theta(-\Delta)^\beta u(t,x)\, dt = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x)\, dw_k(t), \quad t \in [0,T], \quad x \in G,$$

with initial condition $u(0,x) = u_0(x) \in H$*, where* $\theta > 0$*,* $\beta > 0$*,* $\gamma \geq 0$ *and* $\sigma \in \mathbb{R}\setminus\{0\}$ *are constants, and where* $\lambda_k := \sqrt{-\tau_k}$*,* $k \in \mathbb{N}$*. In this case,* $\rho_k = 0$ *for all* $k \in \mathbb{N}$*, and*

$$\nu_k \sim c_1 k^{2\beta/d}, \quad \mu_k(\theta) \sim c_1\theta\, k^{2\beta/d}, \quad \lambda_k \sim \sqrt{c_1}\, k^{1/d}, \quad k \to \infty.$$

Together with Proposition 3.2.2, the conditions (3.24) *and* (3.33) *are equivalent to*

$$\frac{1}{c_1^{2\gamma}\theta^2} \sum_{k=1}^{\infty} \frac{1}{k^{4\gamma/d}} = \infty, \quad \text{and} \quad \frac{1}{c_1^{1+4\gamma}\theta^3} \sum_{k=1}^{\infty} \frac{1}{k^{(2\beta+8\gamma)/d}} = \infty,$$

respectively. Therefore, the consistency and the asymptotic normality hold for the TFE $\tilde{\theta}_N$ *given by* (3.11)*, whenever*

$$2\beta + 8\gamma \leq d.$$

Example 3.2.2. *Let us consider the following SPDE, with the parameter of interest*

θ in front of lower order differential operator,

$$du(t,x) + (\Delta u(t,x) + \theta u(t,x))\,dt = \sum_{k=1}^{\infty} h_k(x)\,dw_k(t), \quad t \in [0,T], \quad x \in G,$$

with initial condition $u(0,x) = u_0(x) \in H$. In this case, $\gamma = 0$, $\nu_k \equiv 1$ for all $k \in \mathbb{N}$, and

$$\rho_k \sim c_1\, k^{2/d}, \quad \mu_k(\theta) \sim \theta + c_1\, k^{2/d}, \quad k \to \infty.$$

Together with Proposition 3.2.2, the conditions (3.24) and (3.33) are equivalent to

$$\sum_{k=1}^{\infty} \frac{1}{\left(\theta + c_1 k^{2/d}\right)^2} = \infty \quad \textit{and} \quad \sum_{k=1}^{\infty} \frac{1}{\left(\theta + c_1 k^{2/d}\right)^3} = \infty,$$

respectively. Therefore, in order for the consistency and the asymptotic normality of $\tilde{\theta}_N$ to hold true, we need to have $d \geq 6$.

3.3 p-Variation Type Estimators

Apart from spectral approach, the literature on parameter estimation for SPDEs is limited, and only few papers are devoted to discretely sampled SPDEs [PR97, Mar03, PvsT07]. Of course, one way to deal with discretely sampled data, is to discretize or approximate the MLEs using the available discrete data, and show that the statistical properties are preserved. On the other hand, if we assume that the solution itself is observed at some space-time grid points, one needs to approximate additionally the Fourier modes. To best of our knowledge, a rigorous asymptotic analysis of this idea still has to be done. Finally, it needs to be mentioned, that by its very nature, the Fourier decomposition has to be performed with respect to the basis formed by the eigenfunctions of the operator $\mathcal{A}_1$. Usually, $\mathcal{A}_1$ is a differential operator, and thus essentially one has to deal with bounded domains.

In this section, we study the parameter estimation problem for simple parabolic SPDEs (3.51) and (3.55), when data is sampled discretely. Namely, we consider the stochastic heat equation, one dimensional, driven by an additive or multiplicative space-time noise, either on bounded domain or whole space, and when the solution u is observed at some discrete space-time points. As such, we do not rely on spectral approach, but rather use some suitable representations of the solution to derive the corresponding estimators. The key idea of the proposed method relies on an intuitively simple observation: the p-variation of a stochastic process is invariant with respect to smooth perturbations. Hence, if the p-variation of a process X can be computed by an explicit formula, and the parameter of interest enters non-trivially into this formula, one can derive consistent estimators of this parameter (similar to estimating the volatility through quadratic variation). However, since the p-variation of the perturbed process $X + Y$ remains the same, given that Y is smooth enough, then the same estimator remains consistent assuming that $X + Y$ is observed. Analogous arguments remain valid for asymptotic normality property. According to Theorem 3.3.1 and Theorem 3.3.2, using the p-variation idea described above, both θ and σ can be estimated in either time or space sampling regime. Hence, to construct a consistent, and asymptotically normal estimator for θ or σ it is enough to observe the solution at one time instant and discretely on a spacial grid of a finite interval, with mesh diameters going to zero. By the same token, it is sufficient to observe the solution just at one spacial point, and over a time-grid interval. We focus our study on these two sampling schemes. It should be mentioned that similar estimators, and same sampling schemes were studied in [PvsT07], where the authors considered the heat equation on $\mathbb{R}$ driven by a multiplicative noise. The methods of proof in [PvsT07] are different from ours. As already mentioned, there are no ready available results on the representations of the solution in time at a fixed spacial point for bounded domains is more delicate. We prove that the solution can be represented as a sum of a smooth process and a zero-mean Gaussian process with known finite fourth variation. Moreover, using some elements of Malliavin calculus, as well as a version of the central limit theorem from [NOL08], we establish a central limit type theorem for the fourth variation of the solution. Consequently, we derive weakly consistent estimators for θ and σ, and prove their asymptotic normality. The results on the representation of the solution are of independent interest, and could be used beyond statistical inference problems. It would be fair to note that a similar methodology of using Malliavin calculus techniques to establish central limit theorem can be found in [Cor12], albeit applied to similar processes but with a simpler covariance structure.

As usual, everywhere below, all equalities and inequalities between random variables, unless otherwise noted, will be understood in the $\mathbb{P}$-a.s. sense. The notations $\xrightarrow{\mathcal{D}}$ will be used for convergence in distribution, while $\xrightarrow{\mathbb{P}}$ or $\mathbb{P}-\lim$ will stand for convergence in probability.

We assume that $\theta \in \Theta \subset (0, +\infty)$ and $\sigma \in \mathbf{S} \subset (0, +\infty)$ are the (unknown) parameters of interest. In this work we focus on two sampling schemes[4]:

(A) *Fixed time and discrete space.* For a fixed instant of time $t > 0$, and given interval $[a, b] \subset G$, the solution u is observed at points (t, x_j), $j = 1, \dots, m$, with $x_j = a + (b - a)j/m, \quad j = 0, 1, \dots, m.$

(B) *Fixed space and discrete time.* For a fixed x from the interior of G, and given time interval $[c, d] \subset (0, +\infty)$, the solution u is observed at points $\{(t_i, x),\ i = 1, \dots, n\}$, where $t_i := c + (d - c)i/n,\ i = 0, 1, \dots, n.$

The main goal of this section is to derive consistent estimators for the parameters θ

[4]For simplicity of writing, we assume that the sampling points form a uniform grid. Generally speaking all the results hold true assuming only that the mesh size of the grid goes to zero.

and σ under these sampling schemes, and to study the asymptotic properties of these estimators. Our method is based on p-variation of the stochastic processes. We first review its definition and prove a result that smooth perturbation does not affect the p-variation of the original process.

3.3.1 p-Variation of a Stochastic Process. In what follows, we will use the notation $\Upsilon^m(a,b) = \{a_j \mid a_j = a + (b-a)j/m,\ j = 0, 1, \ldots, m\}$ for the uniform partition of size m of a given interval $[a,b] \subset \mathbb{R}$. For a given stochastic process X on some interval $[a,b]$, and $p \geq 1$, we will denote by $\mathsf{V}^p_m(X;[a,b])$ the sum

$$\mathsf{V}^p_m(X;[a,b]) := \sum_{j=1}^{m} |X(t_j) - X(t_{j-1})|^p,$$

where $t_j \in \Upsilon^m(a,b)$. Correspondingly,

$$\mathsf{V}^p(X;[a,b]) := \lim_{m\to\infty} \mathsf{V}^p_m(X;[a,b]), \quad \mathbb{P}-\text{a.s.},$$
$$\mathsf{V}^p_{\mathbb{P}}(X;[a,b]) := \mathbb{P}-\lim_{m\to\infty} \mathsf{V}^p_m(X;[a,b]),$$

will denote the p-variation of X on $[a,b]$, in $\mathbb{P}$-a.s. sense and respectively in probability. If no confusions arise, we will simply write $\mathsf{V}^p(X)$, and $\mathsf{V}^p_m(X)$ instead of $\mathsf{V}^p(X;[a,b])$ and $\mathsf{V}^p_m(X;[a,b])$; same applies to $\mathsf{V}^p_{\mathbb{P}}(X)$.

As already mentioned, the estimators proposed in this work are derived using the p-variation of some suitable processes. The next result shows that the 'quadratic variation type arguments' of estimating the diffusion coefficient are invariant with respect to smooth perturbations.

Proposition 3.3.1. *Let $X(t), Y(t)$, $t \in [a,b]$, be stochastic processes with continuous paths, and assume that the process Y has $C^1[a,b]$ sample paths, and there exists $p > 1$,* *such that* $0 < \mathsf{V}^p(X) < \infty$. *Then,*

$$\mathsf{V}^p(X+Y;[a,b]) = \mathsf{V}^p(X;[a,b]). \tag{3.42}$$

Similarly, if $0 < \mathsf{V}^p_{\mathbb{P}}(X) < \infty$, *then*

$$\mathsf{V}^p_{\mathbb{P}}(X+Y;[a,b]) = \mathsf{V}^p_{\mathbb{P}}(X;[a,b]). \tag{3.43}$$

If in addition, there exist $\alpha, \sigma_0 > 0$ *such that,* $\alpha + 1/p < 1$,

$$n^{\alpha}\left(\mathsf{V}^p_n(X;[a,b]) - \mathsf{V}^p(X;[a,b])\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma_0^2), \tag{3.44}$$

then

$$n^{\alpha}\left(\mathsf{V}^p_n(X+Y;[a,b]) - \mathsf{V}^p(X;[a,b])\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma_0^2). \tag{3.45}$$

Moreover, if Y *has* $C^2[a,b]$ *sample paths, and* (3.44) *holds for* $p = 2$ *and* $\alpha = 1/2$, *then* (3.45) *holds true too, with* $p = 2, \alpha = 1/2$.

Proof. First we prove (3.42). It should be noted that a similar result is proved in [CNW06, Corollary 2]. For completeness, we outline our proof too. All 'p-variations' below are on the fixed interval $[a,b]$, and as agreed above, we will omit writing their dependence on $[a,b]$. By Minkowski's inequality, we have that

$$\left|\left(\mathsf{V}^p_n(X)\right)^{1/p} - \left(\mathsf{V}^p_n(Y)\right)^{1/p}\right| \le \left(\mathsf{V}^p_n(X+Y)\right)^{1/p} \le \left(\mathsf{V}^p_n(X)\right)^{1/p} + \left(\mathsf{V}^p_n(Y)\right)^{1/p}. \tag{3.46}$$

Since Y has $C^1[a,b]$ sample paths, we have $\lim_{n\to\infty} \mathsf{V}^p_n(Y) = 0$. Hence, passing to the limit in (3.46), the identity (3.42) follows. As far as (3.43), note that in view of

(3.46), for any $\epsilon > 0$,

$$\begin{aligned}
&\left\{\left|(\mathsf{V}_n^p(X+Y))^{1/p} - (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p}\right| \geq \epsilon\right\} \\
&\qquad = \left\{(\mathsf{V}_n^p(X+Y))^{1/p} \geq (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p} + \epsilon\right\} \cup \left\{(\mathsf{V}_n^p(X+Y))^{1/p} \leq (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p} - \epsilon\right\} \\
&\qquad \subset \left\{(\mathsf{V}_n^p(X))^{1/p} + (\mathsf{V}_n^p(Y))^{1/p} \geq (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p} + \epsilon\right\} \\
&\qquad\qquad \cup \left\{\left|(\mathsf{V}_n^p(X))^{1/p} - (\mathsf{V}_n^p(Y))^{1/p}\right| \leq (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p} - \epsilon\right\} \\
&\qquad \subset \left\{\left|(\mathsf{V}_n^p(X))^{1/p} + (\mathsf{V}_n^p(Y))^{1/p} - (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p}\right| \geq \epsilon\right\} \\
&\qquad\qquad \cup \left\{\left|(\mathsf{V}_n^p(X))^{1/p} - (\mathsf{V}_n^p(Y))^{1/p} - (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p}\right| \geq \epsilon\right\} \\
&\qquad = \left\{\left|(\mathsf{V}_n^p(X))^{1/p} - (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p}\right| \geq \epsilon/2\right\} \cup \left\{(\mathsf{V}_n^p(Y))^{1/p} \geq \epsilon/2\right\} \qquad (3.47)
\end{aligned}$$

Due to the continuity of $x^{1/p}$, based on our initial assumptions, we have that $\mathbb{P} - \lim_{n\to\infty} (\mathsf{V}_n^p(X))^{1/p} = (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p}$, and $\mathbb{P} - \lim_{n\to\infty} (\mathsf{V}_n^p(Y))^{1/p} = 0$. Thus, by (3.47), we get at once that

$$\mathbb{P} - \lim_{n\to\infty} (\mathsf{V}_n^p(X+Y))^{1/p} = (\mathsf{V}_{\mathbb{P}}^p(X))^{1/p},$$

which consequently implies (3.43).

In view of Slutsky's Theorem, to prove (3.45), it is enough to show that

$$\lim_{n\to\infty} n^{\alpha} (\mathsf{V}_n^p(X+Y) - \mathsf{V}_n^p(X)) = 0.$$

By (3.46) and by mean-value theorem, we have

$$\begin{aligned}
\mathsf{V}_n^p(X+Y) &\leq \left((\mathsf{V}_n^p(X))^{1/p} + (\mathsf{V}_n^p(Y))^{1/p}\right)^p \\
&= \mathsf{V}_n^p(X) + p\left((\mathsf{V}_n^p(X))^{1/p} + \eta_{1,n}(\mathsf{V}_n^p(Y))^{1/p}\right)^{p-1} (\mathsf{V}_n^p(Y))^{1/p}, \qquad (3.48)
\end{aligned}$$

for some $\eta_{1,n} \in [0,1]$. Since Y has $C^1[a,b]$ sample paths, denoting $M = \sup_{a \le t \le b} \mid Y'(t) \mid$, and again by mean-value theorem, we get

$$\mathsf{V}_n^p(Y) = \sum_{j=1}^{n} |Y(t_j) - Y(t_{j-1})|^p = \sum_{j=1}^{n} |(t_j - t_{j-1})Y'(\zeta_j)|^p \le n(M/n)^p. \quad (3.49)$$

Therefore, by (3.48), and since $\alpha + 1/p < 1$, we conclude that

$$n^\alpha \left(\mathsf{V}_n^p(X+Y) - \mathsf{V}_n^p(X)\right) \le p\left(\left(\mathsf{V}_n^p(X)\right)^{1/p} + \eta_1 \left(\mathsf{V}_n^p(Y)\right)^{1/p}\right)^{p-1} n^{\alpha+1/p-1} M \underset{n\to\infty}{\longrightarrow} 0.$$

Similarly, we have that

$$n^\alpha \left(\mathsf{V}_n^p(X+Y) - \mathsf{V}_n^p(X)\right) \ge -p\left(\left(\mathsf{V}_n^p(X)\right)^{1/p} - \eta_2 \left(\mathsf{V}_n^p(Y)\right)^{1/p}\right)^{p-1} n^{\alpha+1/p-1} M \underset{n\to\infty}{\longrightarrow} 0,$$

and therefore, (3.45) is proved.

Now suppose that Y has $C^2[a,b]$ sample paths, and assume that (3.44) holds true for $p = 2, \alpha = 1/2$. To show that (3.45) also holds true, it is enough to prove that

$$\lim_{n\to\infty} n^{1/2} \left(\mathsf{V}_n^2(X+Y) - \mathsf{V}_n^2(X)\right) = 0. \quad (3.50)$$

Note that,

$$\mathsf{V}_n^2(X+Y) - \mathsf{V}_n^2(X) = 2\sum_{j=1}^{n} \left(X(t_j) - X(t_{j-1})\right)\left(Y(t_j) - Y(t_{j-1})\right) + \mathsf{V}_n^2(Y).$$

Using (3.49), we have $n^{1/2}\mathsf{V}_n^2(Y) \le n^{3/2}(M/n)^2 \to 0$.

By mean value theorem,

$$
\begin{aligned}
& n^{1/2} \sum_{j=1}^{n} \left(X(t_j) - X(t_{j-1})\right) \left(Y(t_j) - Y(t_{j-1})\right) \\
& = n^{-1/2}(b-a) \sum_{i=1}^{n} \left(X(t_j) - X(t_{j-1})\right) \left(Y'(\zeta_j) - Y'(t_{j-1})\right) \\
& \qquad + n^{-1/2}(b-a) \sum_{i=1}^{n} \left(X(t_j) - X(t_{j-1})\right) Y'(t_{j-1}) \\
& =: K_1 + K_2.
\end{aligned}
$$

Applying Cauchy-Schwartz inequality, we get

$$
\begin{aligned}
|K_1| & \leq n^{-3/2}(b-a)^2 \sum_{i=1}^{n} \left| \ \left(X(t_j) - X(t_{j-1})\right) \max_{a \leq t \leq b} \mid Y''(t) \mid \ \right| \\
& \leq n^{-1}(b-a)^2 \max_{a \leq t \leq b} \mid Y''(t) \mid \sqrt{\mathsf{V}_n^2(X)} \underset{n \to \infty}{\longrightarrow} 0.
\end{aligned}
$$

We rewrite K_2 as

$$
K_2 = n^{-1/2}(b-a) \left(X(b)Y'(b) - X(a)Y'(a) - \sum_{j=1}^{n} X(t_j) \left(Y'(t_j) - Y'(t_{j-1})\right) \right).
$$

Since, $\lim_{n\to\infty} \sum_{j=1}^{n} X(t_j) \left(Y'(t_j) - Y'(t_{j-1})\right) = \int_a^b X(t) dY'(t) = \int_a^b X(t) Y''(t) dt$, we have at once that

$$
\lim_{n \to \infty} K_2 = \lim_{n \to \infty} n^{-1/2}(b-a) \left(X(b)Y'(b) - X(a)Y'(a) - \int_a^b X(t)Y''(t)dt \right) = 0.
$$

Combining the above, (3.50) is proved.

This concludes the proof. $\square$

This result allows us to construct directly consistent and asymptotically normal estimators for some parameter entering the true law of the perturbed process

$X+Y$, given that the p-variation $\mathsf{V}^p(X;[a,b])$ of the unperturbed process X depends non-trivially on the parameter of interest, and this dependence can be computed explicitly.

For example, let B be a two-sided Brownian motion, and Y be a process with a $C^2(\mathbb{R})$ version, and consider the stochastic process

$$Z(x) = \sqrt{\beta}B(x) + Y(x), \quad x \in \mathbb{R},$$

where β is a positive, unknown parameter. Assume that Z is observed at grid points $\Upsilon^m(a,b)$, for some interval $[a,b] \subset \mathbb{R}$. In view of (3.42),

$$\mathsf{V}^2(Z;[a,b]) = \mathsf{V}^2(\sqrt{\beta}B;[a,b]) = \beta(b-a).$$

Consequently, the estimator

$$\widehat{\beta}_m = \frac{1}{b-a}\sum_{j=1}^{m}\left(Z(x_j) - Z(x_{j-1})\right)^2,$$

is a consistent estimator of β, namely $\lim_{m\to\infty}\widehat{\beta}_m = \beta$, $\mathbb{P}$-a.s.. Moreover, it is well-known (cf. [Nou08, AES16]) that

$$\sqrt{m}(\mathsf{V}^2_m(B,[a,b]) - (b-a)) \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0, 2(b-a)^2),$$

and thus, by Proposition 3.3.1, the estimator $\widehat{\beta}_m$ is asymptotically normal, with the convergence

$$\sqrt{m}(\widehat{\beta}_m - \beta) \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0, 2\beta^2).$$

Similarly, let B^H be a fractional Brownian Motion (fBM) with Hurst index

$H = \frac{1}{4}$, and Y be a process with continuously differentiable paths in $(0, +\infty)$. Assume that η is the parameter of interest, and suppose that the process

$$Z^H(t) = \eta^{1/4} B^H(t) + Y(t), \quad t > 0.$$

is sampled at grid points $t_i \in \Upsilon^n(c,d), i = 0, 1, \ldots, n$, with $[c,d] \subset (0, \infty)$. Then,

$$\widehat{\eta}_n = \frac{1}{3(d-c)} \sum_{i=1}^{n} \left(Z^H(t_i) - Z^H(t_{i-1})\right)^4,$$

is a consistent estimator of η, since an fBM with Hurst index H has a finite, non-zero $p = 1/H$-variation. The asymptotic normality of $\mathsf{V}_n^4(B^H; [c,d])$ is established in Theorem B.0.3, and Corollary B.0.1, and hence, by (3.45), $\widehat{\eta}_n$ is also asymptotically normal, and satisfying

$$\sqrt{n}(\widehat{\eta}_n - \eta) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \frac{1}{9}\breve{\sigma}^2\eta^2).$$

where $\breve{\sigma}^2$ is an explicit constant given in Corollary B.0.1.

As already mentioned, our method is based on p-variation that is illustrated above. To apply this method, we need some proper representation of the solution. In the next section, we will discuss the equations of interest and decompositions of their solutions.

3.3.2 Stochastic Heat Equations. First, we consider the stochastic heat equation in $\mathbb{R}$,

$$\begin{aligned} &\mathrm{d}u(t,x) = \theta u_{xx}(t,x)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t,x), \quad x \in \mathbb{R}, \quad t > 0, \qquad (3.51) \\ &u(0,x) = 0, \quad x \in \mathbb{R}, \end{aligned}$$

where W denotes a space-time white noise on $\mathbb{R}^2$, θ, σ are some positive constants. Denote

$$p_t(x) := \frac{1}{\sqrt{4\pi\theta t}} e^{-\frac{x^2}{4\theta t}}, \quad t > 0, \quad x \in \mathbb{R},$$

which is the Gaussian density function that corresponds to the heat operator

$$\mathcal{H} := \frac{\partial}{\partial t} - \theta \frac{\partial^2}{\partial x^2}.$$

It is shown in [Kho14, Section 3] that

$$u(t,x) = \sigma \int_{(0,t)\times\mathbb{R}} p_{t-s}(y-x)\,\mathrm{d}W(s,y), \tag{3.52}$$

is the unique mild solution to (3.51). The following decomposition results of the solution (3.52) are the keys to later discussion of parameter estimation. The detailed description of the solution and the proofs can be found in [Kho14, Section 3].

Theorem 3.3.1 ([Kho14])**.** *Let $u(t,x)$ be defined in* (3.52)*. Then we have,*

(a) For every fixed $t > 0$, there exist a two-sided Brownian motion $B(x)$ and a Gaussian process $X(x)$ with a $C^\infty(\mathbb{R})$ version, such that

$$u(t,x) = \frac{\sigma}{\sqrt{2\theta}} B(x) + X(x), \quad x \in \mathbb{R}. \tag{3.53}$$

(b) For every fixed $x \in \mathbb{R}$, there exists a fractional Brownian motion $B^H(t)$ with Hurst index $H = 1/4$ and a Gaussian process $Y(t)$ that is continuous on $\mathbb{R}_+$ and infinitely differentiable on $(0,\infty)$, such that

$$u(t,x) = \frac{\sigma}{(\theta\pi)^{1/4}} B^H(t) + Y(t), \quad t > 0. \tag{3.54}$$

Next, we consider the stochastic heat equation on the bounded domain $G \subset \mathbb{R}$. For simplicity, let $G = [0, \pi]$.

$$\begin{aligned}
\mathrm{d}u(t,x) &= \theta u_{xx}(t,x)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t,x), \quad x \in (0,\pi), \quad t > 0, \\
u(0,x) &= 0, \quad x \in (0,\pi), \\
u(t,0) &= u(t,\pi) = 0, \quad t > 0,
\end{aligned} \tag{3.55}$$

In this case, the Laplace operator $\Delta = \partial_{xx}$ has only discrete spectrum, with eigenvalues $\lambda_k = -k^2$, $k \in \mathbb{N}$, and corresponding eigenfunctions $h_k(x) = \sqrt{2/\pi}\sin(kx)$, $k \in \mathbb{N}$. Moreover, the functions $\{h_k,\, k \in \mathbb{N}\}$ form a complete orthonormal system in $L^2(G)$, and the noise term can be conveniently written as

$$W(t,x) = \sum_{k\geq 1} w_k(t)h_k(x),$$

where $w_k, k \in \mathbb{N}$, are independent standard Brownian motions. Then by the discussion in Section 3.2, the solution of this equation admits a Fourier series decomposition,

$$u(t,x) = \sum_{k\geq 1} u_k(t)h_k(x), \quad t > 0, \quad x \in (0,\pi),$$

where each Fourier mode $u_k(t)$ is an Ornstein–Uhlenbeck process of the form

$$\begin{aligned}
\mathrm{d}u_k(t) &= -\theta k^2 u_k(t)\,\mathrm{d}t + \sigma\,\mathrm{d}w_k(t), \quad t > 0, \\
u_k(0) &= 0.
\end{aligned}$$

Equivalently, we have that

$$u_k(t) = \sigma \int_0^t e^{-\theta k^2(t-s)}\,\mathrm{d}w_k(s). \tag{3.56}$$

Clearly, $u_k(t) \sim \mathcal{N}(0, \frac{(1-e^{-2\theta k^2 t})\sigma^2}{2\theta k^2})$, and $u_k, k \in \mathbb{N}$, are independent. We prove the following counterpart of the representation (3.53).

Theorem 3.3.2. *For every fixed $t > 0$, there is a Brownian motion $B(x)$ on $[0, \pi]$, and a Gaussian process $R(x)$, $x \in [0, \pi]$ with a $C^\infty(0, \pi)$ version, such that*

$$u(t, x) = \frac{\sigma}{\sqrt{2\theta}} B(x) + R(x), \quad x \in [0, \pi].$$

Proof. Let's fix $t > 0$ and compare (3.59) and the solution

$$u(t, x) = \sum_{k \geq 1} u_k(t) h_k(x). \tag{3.57}$$

with

$$\xi_k = \sqrt{\frac{2\theta k^2}{(1 - e^{-2\theta k^2 t})\sigma^2}} u_k(t)$$

are i.i.d. standard normal random variables. Rewrite (3.57) as

$$u(t, x) = \frac{\sigma x}{\sqrt{2\theta\pi}} \xi_0 + \frac{\sigma}{\sqrt{2\theta}} \sum_{k \geq 1} \frac{1}{k} \xi_k h_k(x) - \frac{\sigma x}{\sqrt{2\theta\pi}} \xi_0 + \frac{\sigma}{\sqrt{2\theta}} \sum_{k \geq 1} \frac{\lambda_k - 1}{k} \xi_k h_k(x),$$

where

$$\lambda_k = \sqrt{1 - e^{-2\theta k^2 t}}.$$

By Lemma 3.3.1,

$$\frac{x}{\sqrt{\pi}} \xi_0 + \sum_{k \geq 1} \frac{1}{k} \xi_k h_k(x)$$

is a standard Brownian motion, denoted by W_x. Let

$$R(x) = -\frac{\sigma x}{\sqrt{2\theta\pi}}\xi_0 + \frac{\sigma}{\sqrt{2\theta}} \sum_{k\geq 1} \frac{\lambda_k - 1}{k} \xi_k h_k(x). \tag{3.58}$$

It follows from Lemma 3.3.2 below, $R(x)$ is infinitely differentiable with respect to x for $x \in (0,\pi)$. This completes the proof. □

Lemma 3.3.1. *Let W_t be a Brownian motion on $[0,\pi]$. Then W_t admits the following expansion*

$$\begin{aligned} W_t &= \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}} \sum_{k\geq 1} \frac{1}{k}\xi_k \sin kt, \quad t \in [0,\pi], \\ &= \frac{t}{\sqrt{\pi}}\xi_0 + \sum_{k\geq 1} \frac{1}{k}\xi_k h_k(t), \quad t \in [0,\pi], \end{aligned}$$

where $\xi_k, k \geq 0$ are iid standard normal random variables on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and

$$h_k(t) = \sqrt{\frac{2}{\pi}} \sin kt, \quad k \geq 1.$$

Proof. Denote

$$W_t^n := \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}} \sum_{k=1}^{n} \frac{1}{k}\xi_k \sin kt,$$

then W_t^n is a Gaussian process for each $n \geq 1$. Moreover, for $n < m$, and any $t \in [0,\pi]$,

$$\begin{aligned} \mathbb{E}\left(W_t^n - W_t^m\right)^2 &= \mathbb{E}\left(\sqrt{\frac{2}{\pi}} \sum_{k=n+1}^{m} \frac{1}{k}\xi_k \sin kt\right)^2 \\ &= \frac{2}{\pi} \sum_{k=n+1}^{m} \mathbb{E}\left(\frac{1}{k}\xi_k \sin kt\right)^2 \\ &= \frac{2}{\pi} \sum_{k=n+1}^{m} \frac{1}{k^2} \sin^2 kt \to 0, \end{aligned} \tag{3.59}$$

as $n, m \to \infty$. Therefore, $(W_t^n)_{n\geq 1}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence

$$W_t = \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}}\sum_{k\geq 1}\frac{1}{k}\xi_k \sin kt$$

belongs to $L^2(\Omega, \mathcal{F}, \mathbb{P})$. In addition, we have the following expression

$$\begin{aligned}W_t &= \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}}\sum_{k\geq 1}\xi_k \int_0^\pi \mathbf{1}_{[0,t]}(s)\cos ks\mathbf{d}s \\ &= \sum_{k\geq 0}\xi_k(\mathbf{1}_{[0,t]}(x), \phi_k(x)).\end{aligned}$$

where $\{\phi_0(x) = \frac{1}{\sqrt{\pi}}, \phi_k(x) = \sqrt{\frac{2}{\pi}}\cos kx, k \geq 1\}$ is also a CONS of $L^2([0,\pi])$. For $t, s \geq 0$,

$$\begin{aligned}\mathbb{E}\,(W_t W_s) &= \sum_{k\geq 0}(\mathbf{1}_{[0,t]}(x), \phi_k(x))(\mathbf{1}_{[0,s]}(x), \phi_k(x)) \\ &= (\mathbf{1}_{[0,t]}(x), \mathbf{1}_{[0,s]}(x)) \\ &= t \wedge s.\end{aligned}$$

And we have

$$\mathbb{E}e^{iuW_t} = \lim_{k\to\infty}\mathbb{E}e^{iuW_t^{n_k}} = e^{-\frac{1}{2}u^2 t}.$$

That is, W_t is a standard Brownian motion.

$\square$

Lemma 3.3.2. *$R(x)$ defined in* (3.58) *is infinitely differentiable with respect to x for $x \in (0, \pi)$.*

Proof. We only need to show that

$$\sum_{k\geq 1} k(\lambda_k - 1)\xi_k h_k(x)$$

converges for $x \in (0, \pi)$ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, that is, to show

$$\sum_{k=n+1}^{m} (k(\lambda_k - 1)h_k(x))^2 \to 0, \tag{3.60}$$

as $m, n \to \infty$, for $x \in (0, \pi)$. Note that

$$\begin{aligned}
(\lambda_k - 1)^2 &= (\sqrt{1 - e^{-2\theta k^2 t}} - 1)^2 \\
&= \frac{e^{-4\theta k^2 t}}{(\sqrt{1 - e^{-2\theta k^2 t}} + 1)^2} \\
&\leq e^{-4\theta k^2 t},
\end{aligned}$$

then (3.60) is obvious, since $|h_k(x)| \leq \sqrt{2/\pi}$.

□

There is no ready available representation similar to (3.54). We will present a close result later.

3.3.3 SPDEs on the Whole Space. Assume that $t > 0$ is a fixed time instant, and consider the partition $\Upsilon^m(a, b)$ of the fixed interval $[a, b] \subset \mathbb{R}$. Suppose that the solution u of (3.51) is observed at the grid points $\{(t, x_j) \mid x_j \in \Upsilon^m(a, b), j = 1, \ldots, m\}$. Consider the following estimators for θ and σ^2 respectively

$$\widehat{\theta}_{m,t} := \frac{(b - a)\sigma^2}{2\sum_{j=1}^{m}(u(t, x_j) - u(t, x_{j-1}))^2}, \tag{3.61}$$

$$\widehat{\sigma}^2_{m,t} := \frac{2\theta}{b - a}\sum_{j=1}^{m}(u(t, x_j) - u(t, x_{j-1}))^2. \tag{3.62}$$

Clearly, (3.61) assumes that σ is known, while (3.62) assumes that θ is known. The following results show that these estimators are consistent and asymptotically normal.

Theorem 3.3.3. *Assuming that σ is known, the estimator* (3.61) *of θ is:*

(i) consistent, that is $\lim_{m\to\infty}\widehat{\theta}_{m,t} = \theta,\ \mathbb{P}-a.s.,$

(ii) asymptotically normal,

$$\sqrt{m}(\widehat{\theta}_{m,t} - \theta) \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0, 2\theta^2). \tag{3.63}$$

Proof. Using the representation (3.53), and in view of Proposition 3.3.1, consistency of $\widehat{\theta}_{m,t}$ follows at once. In addition, we also have that

$$\sqrt{m}\left(\sum_{j=1}^{m}(u(t,x_j) - u(t,x_{j-1}))^2 - \frac{(b-a)\sigma^2}{2\theta}\right) \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0, \frac{(b-a)^2\sigma^4}{2\theta^2}).$$

Consequently, a direct application of Delta-Method yields (3.63), and this concludes the proof. $\square$

Similarly, employing again Proposition 3.3.1, one has the following result.

Theorem 3.3.4. *Assuming that θ is known, the estimator* (3.62) *is a consistent and asymptotically normal estimator of σ^2, with*

$$\sqrt{m}(\widehat{\sigma}^2_{m,t} - \sigma^2) \xrightarrow[m\to\infty]{\mathcal{D}} \mathcal{N}(0, 2\sigma^4). \tag{3.64}$$

In this section we assume that the solution u of (3.55) is observed at the grid points $\{(t_i, x) : i = 1, \ldots, n\}$, where $x \in \mathbb{R}$ is a fixed spatial point, and $0 < c < d < \infty$. We consider the following estimators for θ and σ^2 respectively,

$$\widehat{\theta}_{n,x} := \frac{3(d-c)\sigma^4}{\pi\sum_{i=1}^{n}(u(t_i,x) - u(t_{i-1},x))^4}, \tag{3.65}$$

$$\widehat{\sigma}^2_{n,x} := \sqrt{\frac{\theta\pi}{3(d-c)}\sum_{i=1}^{n}(u(t_i,x)-u(t_{i-1},x))^4}. \tag{3.66}$$

Similar the previous section, the following results about asymptotic properties of these estimators hold.

Theorem 3.3.5. *Given that σ is known, we have that*

$$\lim_{n\to\infty}\widehat{\theta}_{n,x} = \theta, \quad \mathbb{P}-a.s.$$
$$\sqrt{n}(\widehat{\theta}_{n,x}-\theta) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\frac{1}{9}\theta^2\check{\sigma}^2).$$

Assuming that θ is known, we have that

$$\lim_{n\to\infty}\widehat{\sigma}^2_{n,x} = \sigma^2, \quad \mathbb{P}-a.s.$$
$$\sqrt{n}(\widehat{\sigma}^2_{n,x}-\sigma^2) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\frac{1}{36}\sigma^4\check{\sigma}^2).$$

where $\check{\sigma}^2$ is the constant given in (B.3).

The proof is analogous to the proofs of Theorems 3.3.3 and 3.3.4 and is omitted here.

3.3.4 SPDEs on a Bounded Domain. In view of Theorem 3.3.2, and similar to the Theorem 3.3.3, the estimator with spatial sampling at a fixed time instant is studied in the following result. The proof is analogous and is omitted here.

Theorem 3.3.6. *Let u be the solution to* (3.55)*, and assume that u is sampled at discrete points $\{(t,x_j) \mid x_j \in \Upsilon^m(a,b)\}$, for some fixed $t>0$ and $a,b\in(0,\pi)$. Then, assuming σ is known, $\widehat{\theta}_{m,t}$ given by* (3.61) *is a consistent and asymptotically normal estimator for θ, satisfying* (3.63)*. Respectively, if θ is known, then $\widehat{\sigma}^2_{m,t}$ in* (3.62) *is a consistent and asymptotically normal estimator of σ^2, satisfying* (3.64).

The case of sampling the solution in time at a fixed spatial point for bounded domains is more delicate, primarily since there is no ready available representation similar to (3.54). In [Wal81] the author proved that for a similar SPDE at $x = 0$ the 4-variation (in time) of the solution converges to a constant. We start by proving that the 4−variation converges to a constant at any fixed space point x. In addition, we also establish the asymptotic normality property of the 4-variation.

Proposition 3.3.2. *Let $x \in (0, \pi)$ be a fixed space point. Then, the solution $u(t, x)$ of the equation* (3.55) *admits the following decomposition*

$$u(t,x) = \frac{\sigma}{(\pi\theta)^{1/4}} v(t) + S(t), \quad t > 0, \tag{3.67}$$

where v and S are zero-mean Gaussian processes such that:

(a) $S(t)$ is continuous on $[0, +\infty)$, and infinitely differentiable on $(0, \infty)$;

(b) $v(t)$ has finite 4−variation (with convergence in probability)

$$\mathbb{P} - \lim_{n\to\infty} \mathsf{V}_n^4(v; [c, d]) = 3(d - c). \tag{3.68}$$

(c) the 4-variation admits the asymptotic normality property

$$\sqrt{n}\left(\frac{\mathsf{V}_n^4(v;[c,d])}{n\sigma_n^4} - 3\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \bar\sigma_2^2 + \bar\sigma_4^2), \tag{3.69}$$

where

$$\sigma_n^2 = \frac{2}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2}(1 - e^{-(d-c)\theta k^2/n}),$$

$$\bar\sigma_2^2 = 72 + 144 \lim_{n\to\infty} \sum_{j=1}^{n-1} (1 - \frac{j}{n}) |\frac{F(j)}{\sigma_n^2}|^2, \quad \bar\sigma_4^2 = 24 + 48 \lim_{n\to\infty} \sum_{j=1}^{n-1} (1 - \frac{j}{n}) |\frac{F(j)}{\sigma_n^2}|^4,$$

and

$$F(j) = \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(2e^{-j(d-c)\theta k^2/n} - e^{-(j+1)(d-c)\theta k^2/n} - e^{-(j-1)(d-c)\theta k^2/n}\right).$$

Moreover,

$$\sqrt{n}\left(\frac{\pi\theta \mathsf{V}_n^4\left(u(\cdot,x);[c,d]\right)}{n\sigma_n^4\sigma^4} - 3\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\bar{\sigma}_2^2+\bar{\sigma}_4^2). \tag{3.70}$$

where $\sigma_n^2, \bar{\sigma}_2^2$ *and* $\bar{\sigma}_4^2$ *are given above.*

Before we prove Proposition 3.3.2, we first prove the following key limit.

Lemma 3.3.3. *For any* $x \in (0,\pi)$ *and* $\theta > 0$, *the following holds true*

$$\lim_{n\to\infty} \sqrt{n} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2}\left(1 - e^{-\theta k^2/n}\right) = \frac{\sqrt{\pi\theta}}{2}. \tag{3.71}$$

Proof. Note that

$$\sin^2(kx) = \frac{1}{2} - \frac{\sin((2k+1)x) - \sin((2k-1)x)}{4\sin x},$$

and therefore,

$$\begin{aligned}
&\sqrt{n}\sum_{k\geq 1}\frac{\sin^2(kx)}{k^2}\left(1-e^{-\theta k^2/n}\right) \\
&\quad \sqrt{n}\sum_{k\geq 1}\frac{1}{2k^2}\left(1-e^{-\theta k^2/n}\right) - \sqrt{n}\sum_{k\geq 1}\frac{\sin((2k+1)x)-\sin((2k-1)x)}{4k^2\sin x}\left(1-e^{-\theta k^2/n}\right) \\
&\quad =: L_n^1 - L_n^2.
\end{aligned}$$

To prove (3.71), we will show that $L_n^1 \to \sqrt{\pi\theta}/2$, and $L_n^2 \to 0$.

It is straightforward to check that for any $\varepsilon > 0$, the function $(1 - e^{-\epsilon x})/x$, $x > 0$, is decreasing. It is also easy to show that

$$\int_0^\infty \frac{1 - e^{-z^2}}{z^2} dz = \sqrt{\pi}.$$

Using these, we obtain

$$\begin{aligned}
L_n^1 &= \sqrt{n} \sum_{k \geq 1} \int_{k-1}^k \frac{1}{2k^2} \left(1 - e^{-\theta k^2/n}\right) \mathrm{d}z \leq \sqrt{n} \sum_{k \geq 1} \int_{k-1}^k \frac{1}{2z^2} \left(1 - e^{-\theta z^2/n}\right) \mathrm{d}z \\
&= \frac{\sqrt{n}}{2} \int_0^\infty \frac{1}{z^2} \left(1 - e^{-\theta z^2/n}\right) \mathrm{d}z = \frac{\sqrt{n}}{2} \int_0^\infty \frac{1}{y^2 n/\theta} \left(1 - e^{-y^2}\right) \mathrm{d}y \sqrt{n/\theta} \\
&= \frac{\sqrt{\theta}}{2} \int_0^\infty \frac{1}{y^2} \left(1 - e^{-y^2}\right) \mathrm{d}y = \frac{\sqrt{\pi\theta}}{2}.
\end{aligned} \tag{3.72}$$

On the other hand,

$$\begin{aligned}
L_n^1 &= \sqrt{n} \sum_{k \geq 1} \int_k^{k+1} \frac{1}{2k^2} \left(1 - e^{-\theta k^2/n}\right) \mathrm{d}z \geq \sqrt{n} \sum_{k \geq 1} \int_k^{k+1} \frac{1}{2z^2} \left(1 - e^{-\theta z^2/n}\right) \mathrm{d}z \\
&= \frac{\sqrt{n}}{2} \int_1^\infty \frac{1}{z^2} \left(1 - e^{-\theta z^2/n}\right) \mathrm{d}z = \frac{\sqrt{n}}{2} \int_{\sqrt{\theta/n}}^\infty \frac{1}{y^2 n/\theta} \left(1 - e^{-y^2}\right) \mathrm{d}y \sqrt{n/\theta} \\
&= \frac{\sqrt{\theta}}{2} \int_{\sqrt{\theta/n}}^\infty \frac{1}{y^2} \left(1 - e^{-y^2}\right) \mathrm{d}y \underset{n \to \infty}{\longrightarrow} \frac{\sqrt{\pi\theta}}{2}.
\end{aligned} \tag{3.73}$$

Combing (3.72) and (3.73), we conclude that $L_n^1 \to \sqrt{\pi\theta}/2$.

Denote by

$$f_k := \frac{1 - e^{-\theta k^2/n}}{k^2}, \quad k \geq 1,$$

and as above, one can show that $\{f_k, k \in \mathbb{N}\}$ is a decreasing sequence. By simple rearrangement of terms, we get

$$L_n^2 = \sqrt{n} \sum_{k \geq 2} \sin((2k-1)x) \left(f_{k-1} - f_k\right) - \sqrt{n} \sin x f_1.$$

Thus,

$$\begin{aligned}
|L_n^2| &\leq \sqrt{n} \sum_{k \geq 2} \Big| \sin((2k-1)x) \Big| \left(f_{k-1} - f_k\right) + \sqrt{n} \sin x f_1 \\
&\leq \sqrt{n} \sum_{k \geq 2} \left(f_{k-1} - f_k\right) + \sqrt{n} f_1 \leq 2\sqrt{n} f_1 = 2\sqrt{n} \left(1 - e^{-\theta/n}\right) \\
&\leq 2\sqrt{n} \frac{\theta}{n} = 2\frac{\theta}{\sqrt{n}} \underset{n \to \infty}{\longrightarrow} 0.
\end{aligned}$$

The proof is complete. $\square$

Now let us prove Proposition 3.3.2.

Proof of Proposition 3.3.2. Assume that $x \in (0, \pi)$ is fixed. We start by constructing the Gaussian processes S, v. Let $\{\eta_k, k \in \mathbb{N}\}$ be a sequence of i.i.d. standard normal random variables, independent of $\{u_k, k \in \mathbb{N}\}$, and let

$$\begin{aligned}
S_k(t) &:= \frac{\sigma}{\sqrt{2\theta}k} e^{-\theta k^2 t} \eta_k, \qquad k \in \mathbb{N},\ t \geq 0, \\
S(t) &:= \sum_{k=1}^{\infty} S_k(t) h_k(x), \qquad t \geq 0.
\end{aligned}$$

Consequently, we put

$$\begin{aligned}
v_k(t) &:= \frac{(\theta\pi)^{1/4}}{\sigma} \left(u_k(t) - S_k(t)\right), \qquad k \in \mathbb{N},\ t \geq 0, \\
v(t) &:= \sum_{k \geq 1} v_k(t) h_k(x), \quad t \geq 0, \quad x \in (0, \pi).
\end{aligned}$$

Clearly, S and v are zero-mean Gaussian processes that satisfying (3.67).

(a) It is straightforward to check that S is continuous on $[0,+\infty)]$ and infinitely differentiable on $(0,\infty)$. Moreover,

$$\mathbb{E}\left|S_k(t+\epsilon)-S_k(t)\right|^2=\frac{\sigma^2}{2\theta k^2}e^{-2\theta k^2 t}\left(1-e^{-\theta k^2\epsilon}\right)^2,\qquad k\in\mathbb{N},\ t\geq 0. \tag{3.74}$$

(b) By direct computations, using (3.56), one can show that

$$\mathbb{E}\left|u_k(t+\epsilon)-u_k(t)\right|^2=\frac{\sigma^2}{2\theta k^2}(1-e^{-\theta k^2\epsilon})\left(2-(1-e^{-\theta k^2\epsilon})e^{-2\theta k^2 t}\right), \tag{3.75}$$

for $t\geq 0,\ \varepsilon>0,\ k\in\mathbb{N}$. Combining (3.74), (3.75) and the independence between S_k and u_k, we deduce that

$$\mathbb{E}\left|v_k(t+\epsilon)-v_k(t)\right|^2=\frac{\sqrt{\pi}}{\sqrt{\theta}k^2}(1-e^{-\theta k^2\epsilon}),\qquad k\in\mathbb{N},\ t\geq 0.$$

Consequently, we have that

$$\mathbb{E}\left|v(t+\epsilon)-v(t)\right|^2=\sum_{k\geq 1}\mathbb{E}\left|v_k(t+\epsilon)-v_k(t)\right|^2 h_k^2(x)=\frac{2}{\sqrt{\pi\theta}}\sum_{k\geq 1}\frac{\sin^2(kx)}{k^2}(1-e^{-\theta k^2\epsilon}).$$

We will prove (3.68) by showing that

$$\lim_{n\to\infty}\mathbb{E}\left(\mathsf{V}_n^4(v;[c,d])\right)=3(d-c), \tag{3.76}$$

$$\lim_{n\to\infty}\operatorname{Var}\left(\mathsf{V}_n^4(v;[c,d])\right)=0. \tag{3.77}$$

Denote by

$$\sigma_n^2:=\mathbb{E}\left|v(t_j)-v(t_{j-1})\right|^2=\frac{2}{\sqrt{\pi\theta}}\sum_{k\geq 1}\frac{\sin^2(kx)}{k^2}(1-e^{-(d-c)\theta k^2/n}),\quad n\in\mathbb{N}.$$

In view of Lemma 3.3.3,

$$\lim_{n\to\infty} \sqrt{n}\sigma_n^2 = \sqrt{d-c}. \tag{3.78}$$

Since v is a zero-mean Gaussian process, we have

$$\mathbb{E}\left|v(t_j) - v(t_{j-1})\right|^4 = 3\left(\mathbb{E}\left|v(t_j) - v(t_{j-1})\right|^2\right)^2 = 3\sigma_n^4,$$

therefore

$$\lim_{n\to\infty} \mathbb{E}\left(\mathsf{V}_n^4(v;[c,d])\right) = \lim_{n\to\infty} \sum_{j=1}^{n} \mathbb{E}\left|v(t_j) - v(t_{j-1})\right|^4 = \lim_{n\to\infty} 3n\sigma_n^4 = 3(d-c),$$

and hence (3.76) is proved. Next, note that

$$\begin{aligned}
&\operatorname{Var}\left(\mathsf{V}_n^4(v;[c,d])\right) = \mathbb{E}\left(\mathsf{V}_n^4(v;[c,d]) - \mathbb{E}\left(\mathsf{V}_n^4(v;[c,d])\right)\right)^2 \\
&= \sum_{j=1}^{n} \mathbb{E}\left(\left|v(t_j,x) - v(t_{j-1},x)\right|^4 - 3\sigma_n^4\right)^2 \\
&+ 2\sum_{i<j} \mathbb{E}\left(\left|v(t_i,x) - v(t_{i-1},x)\right|^4 - 3\sigma_n^4\right)\left(\left|v(t_j,x) - v(t_{j-1},x)\right|^4 - 3\sigma_n^4\right) \\
&=: J_1 + J_2.
\end{aligned}$$

According to (3.78), we deduce that

$$J_1 = \sum_{j=1}^{n} \mathbb{E}\left(\left|v(t_j,x) - v(t_{j-1},x)\right|^8\right) - 9n\sigma_n^8 = 96n\sigma_n^8 \underset{n\to\infty}{\longrightarrow} 0. \tag{3.79}$$

As far as J_2, for $j \geq 1$, we put

$$\begin{aligned}
F(j) &:= \mathbb{E}\left(v(t_i,x) - v(t_{i-1},x)\right)\left(v(t_{i+j},x) - v(t_{i+j-1},x)\right) \\
&= \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2}\left(2e^{-j(d-c)\theta k^2/n} - e^{-(j+1)(d-c)\theta k^2/n} - e^{-(j-1)(d-c)\theta k^2/n}\right)
\end{aligned}$$

$$= G_j - G_{j-1},$$

where

$$G_j := \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(e^{-j(d-c)\theta k^2/n} - e^{-(j+1)(d-c)\theta k^2/n} \right), \quad j \geq 0,$$

and also put $F(0) := \sigma_n^2$. Since $F(j) < 0$, we have that $G_j < G_{j-1}$. Using the property of joint normal distributions, we continue

$$\begin{aligned} J_2 &= 2 \sum_{i<j} \mathbb{E} \left(|v(t_i, x) - v(t_{i-1}, x)|^4 - 3\sigma_n^4 \right) \left(|v(t_j, x) - v(t_{j-1}, x)|^4 - 3\sigma_n^4 \right) \\ &= 2 \sum_{i<j} \left(24F^4(j-i) + 72F^2(j-i)\sigma_n^4 \right). \end{aligned}$$

From here, since $|F(j-i)| \leq \sigma_n^2$, we deduce that

$$\begin{aligned} J_2 &\leq 2 \sum_{i<j} \left(24|F(j-i)|\sigma_n^6 + 72|F(j-i)|\sigma_n^6 \right) = 192 \sum_{i<j} |F(j-i)|\sigma_n^6 \\ &= 192\sigma_n^6 \sum_{j=1}^{n-1} (n-j) \left(G_{j-1} - G_j \right). \end{aligned}$$

Note that $\sum_{j=1}^{n-1}(n-j)\left(G_{j-1} - G_j\right) = nG_0 - \sum_{j=0}^{n-1} G_j$, and since

$$\begin{aligned} \sum_{j=0}^{n-1} G_j &= \sum_{j=0}^{n-1} \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(e^{-j(d-c)\theta k^2/n} - e^{-(j+1)(d-c)\theta k^2/n} \right) \\ &= \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(1 - e^{-(d-c)\theta k^2} \right) = \frac{1}{2}\sigma_1^2, \end{aligned}$$

and $G_0 = \frac{1}{2}\sigma_n^2$, we conclude that

$$J_2 \leq 192\sigma_n^6 \left(n\frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(1 - e^{-(d-c)\theta k^2/n}\right) - \frac{1}{\sqrt{\pi\theta}} \sum_{k\geq 1} \frac{\sin^2(kx)}{k^2} \left(1 - e^{-(d-c)\theta k^2}\right)\right)$$

$$= 192\sigma_n^6 \left(\frac{n}{2}\sigma_n^2 - \frac{1}{2}\sigma_1^2\right) \overset{n\to\infty}{\longrightarrow} 0. \tag{3.80}$$

according to (3.78). Combining (3.79) and (3.80), (3.77) is proved. Consequently, by (3.76) and (3.77), we also have that $\mathsf{V}_n^4(v;[c,d])$ converges to $3(d-c)$, both in L^2 and in probability.

(c) We will apply Theorem B.0.4, by showing that (B.4) and along with condition (N1) are satisfied. We begin by establishing the following estimates

$$\sum_{j=-l}^{r} |F(|j|)|^m \leq 2\sigma_n^{2m}, \tag{3.81}$$

for any $m \geq 1$, $\ell, r \in \mathbb{N}$. Since $m \geq 1$,

$$\begin{aligned}
\sum_{j=1}^{r} |F(j)|^m = \sum_{j=1}^{r} |F(j)|^{m-1}\,|F(j)| &\leq \sum_{j=1}^{r} \sigma_n^{2(m-1)}|F(j)| \\
&= \sum_{j=1}^{r} \sigma_n^{2(m-1)}(G_{j-1} - G_j) = \sigma_n^{2(m-1)}(G_0 - G_{r-1}) \\
&\leq \sigma_n^{2(m-1)} G_0 = \frac{1}{2}\sigma_n^{2m},
\end{aligned}$$

where we used the fact that $G_j \geq 0$ and $G_0 = \frac{1}{2}\sigma_n^2$. Therefore,

$$\begin{aligned}
\sum_{j=-l}^{r} |F(|j|)|^m &= (\sigma_n^2)^m + \sum_{j=1}^{r} |F(j)|^m + \sum_{j=1}^{l} |F(j)|^m \\
&\leq \sigma_n^{2m} + \frac{1}{2}\sigma_n^{2m} + \frac{1}{2}\sigma_n^{2m} = 2\sigma_n^{2m}.
\end{aligned}$$

With slight abuse of notations, just in this proof, we denote by $\Delta v_j^n := v(t_j, x) - v(t_{j-1}, x)$. Let $\mathcal{H}$ be the closed subspace of $L^2(\Omega, \mathscr{F}, \mathbb{P})$ generated by the random variables $\frac{\Delta v_j^n}{\sigma_n}$, $1 \leq j \leq n$; $j, n \in \mathbb{N}$. Then,

$$\begin{aligned}\left|\frac{\Delta v_j^n}{\sigma_n}\right|^4 - 3 &= \left(\left|\frac{\Delta v_j^n}{\sigma_n}\right|^4 - 6\left|\frac{\Delta v_j^n}{\sigma_n}\right|^2 + 3\right) + 6\left(\left|\frac{\Delta v_j^n}{\sigma_n}\right|^2 - 1\right) \\ &= H_4\left(\frac{\Delta v_j^n}{\sigma_n}\right) + 6H_2\left(\frac{\Delta v_j^n}{\sigma_n}\right) = I_4\left[\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}\right] + 6I_2\left[\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}\right].\end{aligned}$$

Therefore,

$$\begin{aligned}&\sqrt{n}\left(\frac{\mathsf{V}_n^4(v;[c,d])}{n\sigma_n^4} - 3\right) \\ &= I_4\left[\frac{1}{\sqrt{n}}\sum_{j=1}^n \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}\right] + I_2\left[\frac{6}{\sqrt{n}}\sum_{j=1}^n \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}\right]\end{aligned} \tag{3.82}$$

Let

$$f_n^{(2)} := \frac{6}{\sqrt{n}}\sum_{j=1}^n \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}, \qquad f_n^{(4)} := \frac{1}{\sqrt{n}}\sum_{j=1}^n \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}, \tag{3.83}$$

and consider the sequence of two-dimensional random vectors $F_n := \left(I_2(f_n^{(2)}), I_4(f_n^{(4)})\right)$, $n \in \mathbb{N}$, to which we will apply Theorem B.0.4. Using the properties of Wiener integral, we obtain that

$$\lim_{n\to\infty} \mathbb{E}\left(I_2(f_n^{(2)}) I_4(f_n^{(4)})\right) = 0,$$

and hence (B.4) is satisfied.

Next, we move to verification of condition (N1), which in this case becomes

$$\lim_{n\to\infty} \|f_n^{(m)} \otimes_r f_n^{(m)}\|^2_{H^{2\otimes(m-r)}} = 0. \tag{3.84}$$

for $m = 2, 4$, and $1 \leq r \leq m-1$.

Using the linearity of the inner products and the properties of the tensor products of Hilbert spaces, we obtain

$$\begin{aligned}
\mathbb{E}\left(I_2(f_n^{(2)})\right)^2 &= 2\langle f_n^{(2)}, f_n^{(2)}\rangle_{\mathcal{H}^{\otimes 2}} = \frac{72}{n}\Big\langle \sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}, \sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}\Big\rangle_{\mathcal{H}^{\otimes 2}} \\
&= \frac{72}{n}\sum_{i,j=1}^{n}\Big\langle \left(\frac{\Delta v_i^n}{\sigma_n}\right)^{\otimes 2}, \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 2}\Big\rangle_{\mathcal{H}^{\otimes 2}} = \frac{72}{n}\sum_{i,j=1}^{n}\Big\langle \frac{\Delta v_i^n}{\sigma_n}, \frac{\Delta v_j^n}{\sigma_n}\Big\rangle_{\mathcal{H}}^2 \\
&= \frac{72}{n}\sum_{i,j=1}^{n}\left[\mathbb{E}\left(\frac{\Delta v_i^n}{\sigma_n}\cdot\frac{\Delta v_j^n}{\sigma_n}\right)\right]^2 = \frac{72}{n}\sum_{i,j=1}^{n}\frac{|F(|j-i|)|^2}{\sigma_n^4} \\
&= \frac{72}{n\sigma_n^4}\left(\sum_{j=1}^{n}|F(0)|^2 + 2\sum_{i<j}|F(j-i)|^2\right) \\
&= \frac{72}{n\sigma_n^4}\left(n\sigma_n^4 + 2\sum_{j=1}^{n-1}(n-j)|F(j)|^2\right) \\
&= 72 + \frac{144}{\sigma_n^4}\sum_{j=1}^{n-1}(1-\frac{j}{n})|F(j)|^2 = 72 + 144\sum_{j=1}^{n-1}(1-\frac{j}{n})\left|\frac{F(j)}{\sigma_n^2}\right|^2.
\end{aligned}$$

In view of (3.81), we have that

$$\sum_{j=1}^{n-1}(1-\frac{j}{n})\left|\frac{F(j)}{\sigma_n^2}\right|^2 \leq \sum_{j=1}^{\infty}\left|\frac{F(j)}{\sigma_n^2}\right|^2 < \infty,$$

and thus

$$\bar{\sigma}_2^2 := \lim_{n\to\infty}\mathbb{E}\left(I_2(f_n^{(2)})\right)^2 = 72 + 144\lim_{n\to\infty}\sum_{j=1}^{n-1}(1-\frac{j}{n})\left|\frac{F(j)}{\sigma_n^2}\right|^2 < \infty.$$

Similarly,

$$\begin{aligned}
\mathbb{E}\left(I_4(f_n^{(4)})\right)^2 &= 24\Big\langle f_n^{(4)}, f_n^{(4)}\Big\rangle_{\mathcal{H}^{\otimes 4}} = \frac{24}{n}\Big\langle \sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}, \sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}\Big\rangle_{\mathcal{H}^{\otimes 4}} \\
&= \frac{24}{n}\sum_{i,j=1}^{n}\Big\langle \left(\frac{\Delta v_i^n}{\sigma_n}\right)^{\otimes 4}, \left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes 4}\Big\rangle_{\mathcal{H}^{\otimes 4}} = \frac{24}{n}\sum_{i,j=1}^{n}\Big\langle \frac{\Delta v_i^n}{\sigma_n}, \frac{\Delta v_j^n}{\sigma_n}\Big\rangle_{\mathcal{H}}^4
\end{aligned}$$

$$
\begin{aligned}
&= \frac{24}{n}\sum_{i,j=1}^{n}\left[\mathbb{E}\left(\frac{\Delta v_i^n}{\sigma_n}\cdot\frac{\Delta v_j^n}{\sigma_n}\right)\right]^4 = \frac{24}{n}\sum_{i,j=1}^{n}\frac{|F(|j-i|)|^4}{\sigma_n^8} \\
&\le 24 + 48\sum_{j=1}^{n-1}\left|\frac{F(j)}{\sigma_n^2}\right|^4 \le 24 + 48\sum_{j=1}^{\infty}\left|\frac{F(j)}{\sigma_n^2}\right|^4 < \ \infty,
\end{aligned}
$$

and consequently,

$$
\bar{\sigma}_4^2 := \lim_{n\to\infty}\mathbb{E}\left(I_4(f_n^{(4)})\right)^2 = 24 + 48\lim_{n\to\infty}\sum_{j=1}^{n-1}(1-\frac{j}{n})\left|\frac{F(j)}{\sigma_n^2}\right|^4 < \infty.
$$

Let $a_2 = 6, a_4 = 1$. Then,

$$
\begin{aligned}
&\|f_n^{(m)}\otimes_r f_n^{(m)}\|^2_{H^{2\otimes(m-r)}} = \|\frac{a_m}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes m}\otimes_r \frac{a_m}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes m}\|^2_{H^{\otimes 2(m-r)}} \\
&= \|\frac{a_m^2}{n}\sum_{i,j=1}^{n}\left(\frac{\Delta v_i^n}{\sigma_n}\right)^{\otimes m}\otimes_r\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes m}\|^2_{H^{\otimes 2(m-r)}} \\
&= \|\frac{a_m^2}{n}\sum_{i,j=1}^{n}\langle\frac{\Delta v_i^n}{\sigma_n},\frac{\Delta v_j^n}{\sigma_n}\rangle_H^r\left(\frac{\Delta v_i^n}{\sigma_n}\right)^{\otimes(m-r)}\otimes\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes(m-r)}\|^2_{H^{\otimes 2(m-r)}} \\
&= \|\frac{a_m^2}{n}\sum_{i,j=1}^{n}\frac{|F(|j-i|)|^r}{\sigma_n^{2r}}\left(\frac{\Delta v_i^n}{\sigma_n}\right)^{\otimes(m-r)}\otimes\left(\frac{\Delta v_j^n}{\sigma_n}\right)^{\otimes(m-r)}\|^2_{H^{\otimes 2(m-r)}} \\
&= \frac{a_m^4}{n^2\sigma_n^{4m}}\sum_{i,j,i',j'=1}^{n}|F(|j-i|)|^r|F(|j'-i'|)|^r|F(|i'-i|)|^{m-r}|F(|j'-j|)|^{m-r} \\
&\le \frac{a_m^4}{n^2\sigma_n^{4m}}\sum_{i,j,i',j'=1}^{n}|F(|j-i|)F(|j'-i'|)F(|i'-i|)F(|j'-j|)|\,\sigma_n^{4m-8} \\
&= \frac{a_m^4}{n^2\sigma_n^{8}}\sum_{i,j,i',j'=1}^{n}|F(|j-i|)F(|j'-i'|)F(|i'-i|)F(|j'-j|)| \\
&= O_1 + 2O_2,
\end{aligned}
$$

where

$$
O_1 := \frac{a_m^4}{n^2\sigma_n^8}\sum_{i',j'=1}^{n}\sum_{i=1}^{n}\left|\ F(0)F(|j'-i'|)F(|i'-i|)F(|j'-i|)\ \right|,
$$

$$O_2 := \frac{a_m^4}{n^2\sigma_n^8} \sum_{i',j'=1}^{n} \sum_{i<j} \left| F(|j-i|)F(|j'-i'|)F(|i'-i|)F(|j'-j|) \right|.$$

First note that, by direct computations and using (3.81), we have

$$\begin{aligned}
O_1 &= \frac{a_m^4}{n^2\sigma_n^6} \sum_{i',j'=1}^{n} \sum_{i=1}^{n} \left| F(|j'-i'|)F(|i'-i|)F(|j'-i|) \right| \\
&\le \frac{a_m^4}{n^2\sigma_n^6} \sum_{i',j'=1}^{n} \sum_{i=1}^{n} \mid F(|j'-i'|) \mid \frac{F(|i'-i|)^2 + F(|j'-i|)^2}{2} \\
&\le \frac{a_m^4}{n^2\sigma_n^6} \sum_{i',j'=1}^{n} \mid F(|j'-i'|) \mid \frac{2\sigma_n^4 + 2\sigma_n^4}{2} \le \frac{2a_m^4}{n^2\sigma_n^2} \sum_{i',j'=1}^{n} \mid F(|j'-i'|) \mid \\
&\le \frac{2a_m^4}{n^2\sigma_n^2} \left(\sum_{j=1}^{n} |F(0)| + 2 \sum_{i<j} \mid F(j-i) \mid \right) \\
&\le \frac{2a_m^4}{n} + \frac{4a_m^4}{n^2\sigma_n^2} \sum_{j=1}^{n-1} (n-j) \mid F(j) \mid \; = \frac{2a_m^4}{n} + \frac{4a_m^4}{n} \sum_{j=1}^{n-1} (1 - \frac{j}{n}) \mid \frac{F(j)}{\sigma_n^2} \mid \\
&\underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Similarly,

$$\begin{aligned}
O_2 &= \frac{a_m^4}{n^2\sigma_n^8} \sum_{i',j'=1}^{n} \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \left| F(|i+k-i|)F(|j'-i'|)F(|i'-i|)F(|j'-i-k|) \right| \\
&= \frac{a_m^4}{n^2\sigma_n^8} \sum_{i',j'=1}^{n} \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \left| F(k)F(|j'-i'|)F(|i'-i|)F(|j'-i-k|) \right| \\
&\le \frac{a_m^4}{n^2\sigma_n^8} \sum_{i',j'=1}^{n} \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \left| F(|j'-i'|)F(|i'-i|) \right| \frac{F(k)^2 + F(|j'-i-k|)^2}{2} \\
&\le \frac{2a_m^4}{n^2\sigma_n^4} \sum_{i',j'=1}^{n} \sum_{i=1}^{n-1} \left| F(|j'-i'|)F(|i'-i|) \right| \le \frac{4a_m^4}{n^2\sigma_n^2} \sum_{i',j'=1}^{n} \left| F(|j'-i'|) \right| \\
&\underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Thus, (3.84) holds true. Therefore, (N2) from Theorem B.0.4 holds true, namely, we have that

$$F_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \begin{pmatrix} \bar{\sigma}_2^2 & 0 \\ 0 & \bar{\sigma}_4^2 \end{pmatrix}\right). \tag{3.85}$$

Consequently, (3.69) follows from (3.82),(3.83) and (3.85). Finally, (3.69) implies (3.70), by using that

$$\sqrt{n}\left(\frac{\pi\theta \mathsf{V}_n^4\left(u(\cdot, x); [c,d]\right)}{n\sigma_n^4\sigma^4} - \frac{\mathsf{V}_n^4(v; [c,d])}{n\sigma_n^4}\right) \to 0, \quad \text{in } L^2 \text{ and in probability} \tag{3.86}$$

The proof of (3.86) follows by similar arguments as in proof of Proposition 3.3.1 and we omit it here.

The proof is complete.

□

Next, we present the main results of this subsection on consistency and asymptotic normality of the estimators (3.65) and (3.66).

Theorem 3.3.7. *Let u be the solution to* (3.55)*, and assume that u is sampled at discrete points $\{(t_i, x) \mid t_i \in \Upsilon^n(c,d)\}$, for some fixed $x \in (0,\pi)$, and $0 < c < d < \infty$. Then, assuming σ is known, $\widehat{\theta}_{n,x}$ given by* (3.65) *is a weakly consistent estimator for θ, that is*

$$\mathbb{P}\text{-}\lim_{n\to\infty} \widehat{\theta}_{n,x} = \theta. \tag{3.87}$$

Respectively, if θ is known, then $\widehat{\sigma}^2_{n,x}$ in (3.66) *is a weakly consistent estimator of σ^2. Moreover, $\widehat{\theta}_{n,x}$ and $\widehat{\sigma}^2_{n,x}$ satisfy the following central limit type convergence*

$$\sqrt{n}\left(\widehat{\theta}_{n,x} - \frac{(d-c)\theta}{n\sigma_n^4}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \theta^2\left(\bar{\sigma}_2^2 + \bar{\sigma}_4^2\right)), \tag{3.88}$$

$$\sqrt{n}\left(\widehat{\sigma}^2_{n,x} - \frac{\sqrt{n}\sigma_n^2}{\sqrt{d-c}}\sigma^2\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \frac{1}{36}\sigma^4\left(\bar{\sigma}_2^2 + \bar{\sigma}_4^2\right)). \tag{3.89}$$

Proof. Consistency is a direct consequence of Proposition 3.3.2.(a)-(b) and (3.43) from Proposition 3.3.1.

Combining (3.65) and (3.70), we have

$$\sqrt{n}\left(\frac{3(d-c)\theta}{\widehat{\theta}_{n,x} n\sigma_n^4} - 3\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \bar{\sigma}_2^2 + \bar{\sigma}_4^2). \tag{3.90}$$

Due to (3.87), and by Slutsky's theorem, multiply by $\widehat{\theta}_{n,x}/3$ on the left side of (3.90), (3.88) follows at once.

Combining (3.66) and (3.70), we have

$$\sqrt{n}\left(\frac{3(d-c)\widehat{\sigma}^4_{n,x}}{n\sigma_n^4\sigma^4} - 3\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \bar{\sigma}_2^2 + \bar{\sigma}_4^2). \tag{3.91}$$

According to (3.78), we have

$$\lim_{n\to\infty} n\sigma_n^4 = d - c,$$

and thus by Slutsky's theorem, multiply $n\sigma_n^4/(d-c)$ on the left side of (3.91) will not affect the convergence in distribution, that is,

$$\sqrt{n}\left(\frac{3\widehat{\sigma}^4_{n,x}}{\sigma^4} - \frac{3n\sigma_n^4}{d-c}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \bar{\sigma}_2^2 + \bar{\sigma}_4^2).$$

Note that

$$\frac{3\widehat{\sigma}^4_{n,x}}{\sigma^4} - \frac{3n\sigma_n^4}{d-c} = \left(\frac{\widehat{\sigma}^2_{n,x}}{\sigma^2} - \frac{\sqrt{n}\sigma_n^2}{\sqrt{d-c}}\right)\left(\frac{3\widehat{\sigma}^2_{n,x}}{\sigma^2} + \frac{3\sqrt{n}\sigma_n^2}{\sqrt{d-c}}\right),$$

and

$$\frac{3\sigma_{n,x}^2}{\sigma^2} + \frac{3\sqrt{n}\sigma_n^2}{\sqrt{d-c}} \to 6 \quad \text{as } n \to \infty,$$

we have

$$\sqrt{n}\left(\frac{\widehat{\sigma}_{n,x}^2}{\sigma^2} - \frac{\sqrt{n}\sigma_n^2}{\sqrt{d-c}}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \frac{1}{36}(\bar{\sigma}_2^2 + \bar{\sigma}_4^2)).$$

Multiply σ^2 gives (3.89). This completes the proof. □

3.4 Numerical Simulation of Estimators

We conclude this chapter with numerical simulations of the solutions and estimators for the stochastic heat equation (3.55),

$$\mathrm{d}u(t,x) = \theta\Delta u(t,x)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t,x), \quad t \geq 0, \quad x \in (0,\pi),$$
$$u(0,x) = 0, \quad x \in (0,\pi), \quad u(t,0) = u(t,\pi) = 0, \quad t \geq 0.$$

where

$$W(t,x) = \sum_{k\geq 0} \sqrt{2/\pi}\sin(kx)w_k(t),$$

and $w_k, k \geq 0$ are independent standard Brownian motions. The simulations and plots in this section are implemented in MATLAB®, the programs are available from the author upon request.

3.4.1 Simulation of Fourier Modes and Solution. We know that the solution is

$$u(t,x) = \sum_{k\geq 1} u_k(t)h_k(x), \quad t > 0, \quad x \in (0,\pi), \tag{3.92}$$

where each Fourier mode $u_k(t)$ is an Ornstein–Uhlenbeck process of the form

$$\mathrm{d}u_k(t) = -\theta k^2 u_k(t)\,\mathrm{d}t + \sigma\,\mathrm{d}w_k(t), \quad t > 0, \quad u_k(0) = 0. \tag{3.93}$$

We can simulate each Fourier mode $u_k(t)$ for $k = 1, 2, \ldots, N$ and then approximate (3.92) by

$$u^N(t,x) = \sum_{k=1}^{N} u_k(t)h_k(x). \tag{3.94}$$

To simulate u_k, we first prove the following lemma.

Lemma 3.4.1. *Suppose that $f(t)$ is some nonzero deterministic integrable function. Let*

$$g(t) = \int_0^t [f(s)]^2\,\mathrm{d}s.$$

If the process Y is the solution to the SDE

$$\mathrm{d}Y(t) = f(t)\,\mathrm{d}B(t), Y(0) = 0$$

where B is a standard Brownian motion, then

$$Y(t) = W(g(t)),$$

for some standard Brownian motion W.

Proof. Define $\tau_s := \inf\{t \geq 0 : g(t) > s\}$. It is easy to see that τ_s is a stopping time w.r.t the natural filtration of B. And moreover, τ_s is strictly increasing w.r.t s. Applying Itô's formula to e^{ivY_t}, here $i = \sqrt{-1}$,

$$\mathrm{d}e^{ivY(t)} = ive^{ivY(t)}f(t)\,\mathrm{d}B(t) - \frac{1}{2}v^2e^{ivY(t)}f^2(t)\,\mathrm{d}t \tag{3.95}$$

Integrating (3.95) from τ_{s_1} to τ_{s_2}, $s_1 < s_2$,

$$e^{ivY(\tau_{s_2})} - e^{ivY(\tau_{s_1})} = \int_{\tau_{s_1}}^{\tau_{s_2}} ive^{ivY(t)} f(t)\,\mathrm{d}B(t) - \frac{1}{2}v^2 \int_{\tau_{s_1}}^{\tau_{s_2}} e^{ivY(t)} f^2(t)\,\mathrm{d}t$$

Consider the change of variable $t = \tau_s$, then

$$\int_{\tau_{s_1}}^{\tau_{s_2}} e^{ivY(t)} f^2(t)\,\mathrm{d}t = \int_{s_1}^{s_2} e^{ivY(\tau_s)}\,\mathrm{d}s$$

Therefore

$$e^{ivY(\tau_{s_2})} = e^{ivY(\tau_{s_1})} + \int_{\tau_{s_1}}^{\tau_{s_2}} ive^{ivY(t)} f(t)\,\mathrm{d}B(t) - \frac{1}{2}v^2 \int_{s_1}^{s_2} e^{ivY(\tau_s)}\,\mathrm{d}s$$

Define $Z(s, s_1) := \mathbb{E}\Big(e^{iv(Y(\tau_s) - Y(\tau_{s_1}))} | \mathcal{G}_{s_1}\Big)$, where $\mathcal{G}_s = \mathcal{F}_{\tau_s}$, with $\mathcal{F}_t, t \geq 0$ is the natural filtration generated by B. Then

$$Z(s_2, s_1) = 1 - \frac{1}{2}v^2 \int_{s_1}^{s_2} Z(s, s_1)\,\mathrm{d}s, \quad Z(s_1, s_1) = 1.$$

Solving the above integral equation,

$$Z(s_2, s_1) = e^{-\frac{1}{2}v^2(s_2 - s_1)},$$

which essentially means $Y(\tau_{s_2}) - Y(\tau_{s_1}) \sim N(0, s_2 - s_1)$. And $Y(\tau_{s_2}) - Y(\tau_{s_1})$ is independent of $\mathcal{G}_{s_1}$. In addition, $\mathbb{E}(Y(\tau_{s_2}) - Y(\tau_{s_1}))^4 = 3(s_2 - s_1)^2$. Thus, by Kolmogorov's Criterion, $Y(\tau_s)$ is continuous. In conclusion, $Y(\tau_s)$ is a standard Brownian motion, denoted by W. Then changing the variable, we get

$$Y(t) = W(g(t)).$$

$\square$

By (3.93),

$$\mathrm{d}e^{\theta k^2 t}u_k(t) = \sigma e^{\theta k^2 t}\,\mathrm{d}w_k(t).$$

Let $f(t) := \sigma e^{\theta k^2 t}$, $g(t) := \int_0^t [f(s)]^2 ds$, then by Lemma 3.4.1,

$$e^{\theta k^2 t}u_k(t) - e^{\theta k^2 s}u_k(s) = \widetilde{w}_k(g(t)) - \widetilde{w}_k(g(s)),$$

for some standard Brownian motion $\widetilde{w}_k$. Moreover, we know

$$g(t) = \sigma^2 \cdot \frac{e^{2\theta k^2 t} - 1}{2\theta k^2}.$$

If we have a partition $0 = t_0 < t_1 < \cdots < t_n = T$ of $[0, T]$ with $t_j = j\Delta t, j = 0, 1, \cdots, n$ and $\Delta t = T/n$, then we have a numerical scheme for approximating u_k,

$$\widetilde{u}_k(t_j) = e^{-\theta k^2 \Delta t}\widetilde{u}_k(t_{j-1}) + \sigma\sqrt{\frac{1 - e^{-2\theta k^2 \Delta t}}{2\theta k^2}}\xi_j,$$

where $\xi_j, j = 1, 2, \cdots, n$ are iid standard normal random variables. Once we simulate the Fourier modes, we can approximate the solution using (3.94).

3.4.2 Simulation of MLE and TFE. Once we simulate the first N Fourier modes, we are ready to compute MLE and TFE from these modes using (3.2) and (3.11), respectively. Here we give a plot of the convergence for these estimators as N increases.

3.4.3 Simulation of p-Variation Type Estimator. Once we simulate the solution $u(t, x)$, we are ready to compute the estimators (3.61), (3.62), (3.65), and (3.66). Here we give the plots of the convergence for these estimator as number of observations increase.

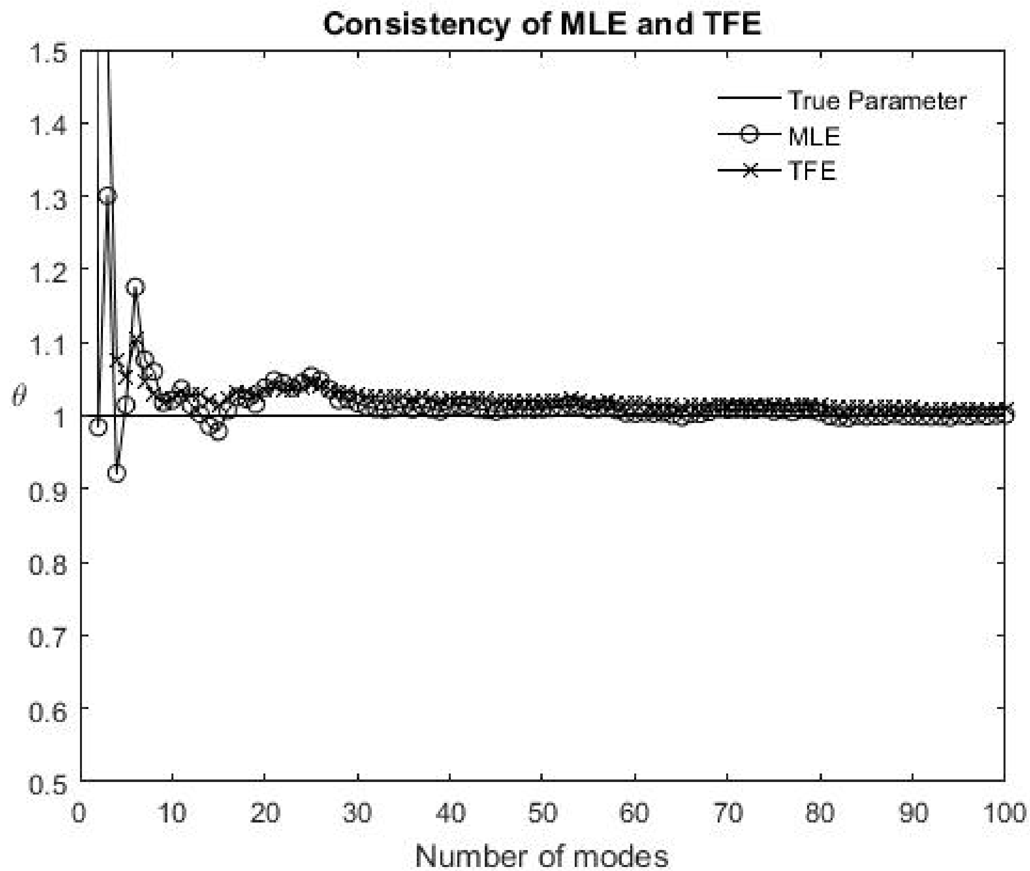


Figure 3.1. Convergence of MLE and TFE as $N \to \infty$

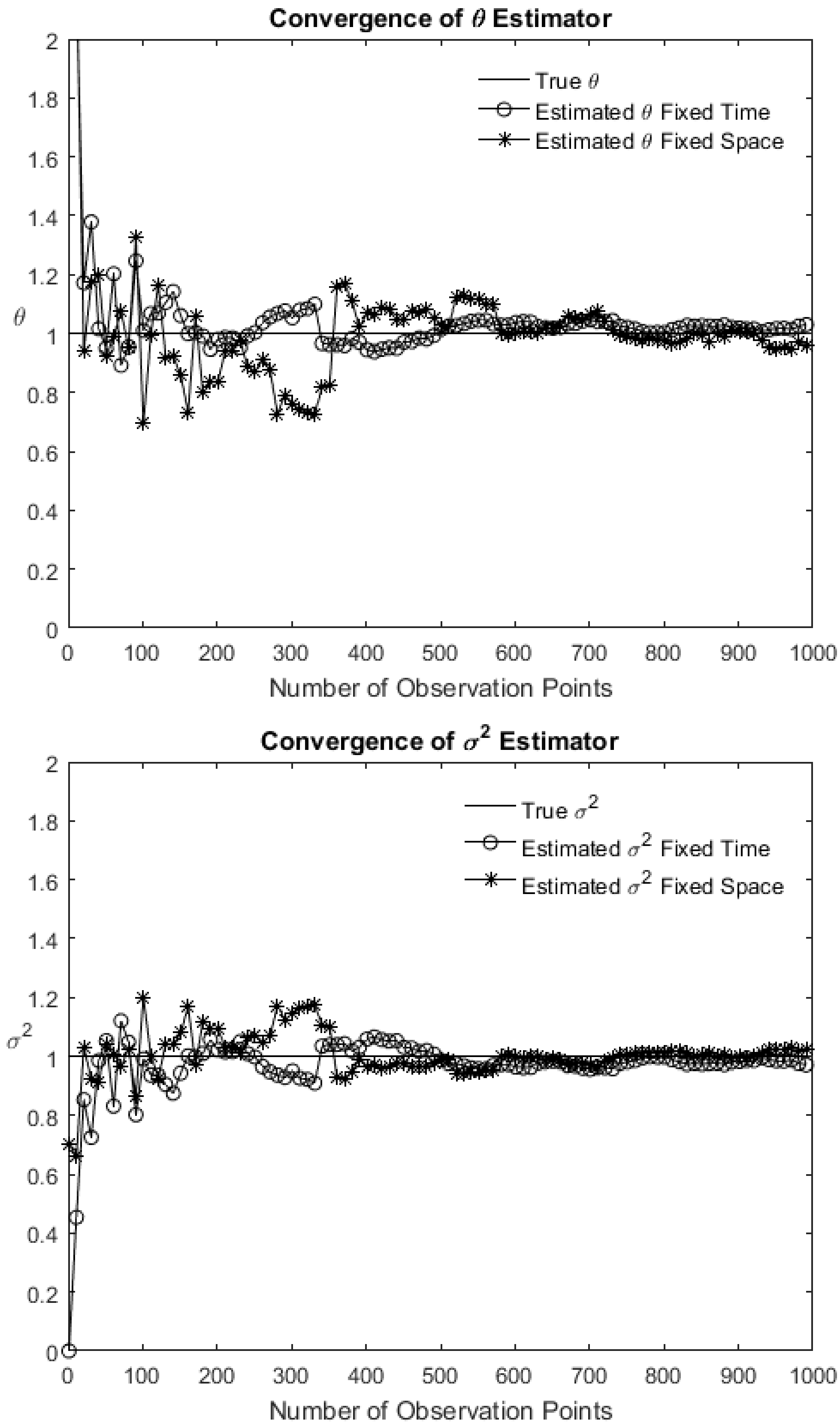


Figure 3.2. Convergence of the estimators for θ and σ^2

CHAPTER 4

FUTURE WORK

In Chapter 2, we have studied the Wiener-Hopf factorization for finite-state time-inhomogeneous Markov chains with a piece-wise constant generator matrix. We plan to address the following problems as part of future research.

1. We will start by developing the theory of Wiener-Hopf factorization for finite state Markov chains with general time dependent generators that are non necessarily piece-wise constant matrices.

2. Next we will study the theory of Wiener-Hopf factorization for time-inhomogeneous Markov processes with general state space.

3. We will investigate the so-called "noisy" Wiener-Hopf factorization where the additive functional is perturbed by an independent standard Brownian motion and the Markov chain is time-inhomogeneous.

4. We will study the connection between the Wiener-Hopf factorization for Markov processes developed herein and the existing Winer-Hopf factorization theory for Lévy processes. Consequently, we plant to develop the Wiener-Hopf factorization for time-inhomogeneous Lévy processes.

In Chapter 3, we have studied the statistical inference for SPDEs mostly driven by an additive noise. One of the major future task is to study the parameter estimation problem for SPDEs with multiplicative noise.

1. We will study the estimators for SPDEs with multiplicative noise within spectral approach.

2. Beyond spectral approach, we also plan to study the discrete sampling for stochastic heat equation with multiplicative noise. The consistency of the estimators is already studied, while the asymptotic normality remains an open question.

3. Since the solution to a stochastic heat equation with additive space-time white noise is a Gaussian random field, we know the characteristics of this random field, namely the mean and the covariance structure. Thus, the discretely sampled data has a multivariate normal distribution, and we plan to study the statistical inference problem using the methods from classical statistics for Gaussian fields.

4. Finally, we will investigate problems in statistical inference for SPDEs driven by non-Gaussian noises, such as Lévy noise.

APPENDIX A

NUMERICAL INVERSE LAPLACE TRANSFORM

For the convenience of the reader, we will briefly recall the basics of Laplace transform and its inverse. Then, we will proceed with an important result regarding the approximation of the multivariate inverse Laplace transform.

Let $f : [0,\infty)^n \to [0,\infty)$ be a Borel measurable function such that

$$\int_0^\infty \cdots \int_0^\infty f(t_1,\dots,t_n) \prod_{k=1}^n e^{-q_k t_k}\, \mathrm{d}t_1 \cdots \mathrm{d}t_n$$

exists for any $q_1,\dots,q_n > 0$. Then, the multivariate Laplace transform $\widehat{f}$ of f, defined by

$$\widehat{f}(q_1,\dots,q_n) = \mathcal{L}(f)(q_1,\dots,q_n) := \int_0^\infty \cdots \int_0^\infty f(t_1,\dots,t_n) \prod_{k=1}^n e^{-q_k t_k}\, \mathrm{d}t_1 \cdots \mathrm{d}t_n,$$

is well-defined for any $q_k \in \mathbb{C}^+$, $k = 1,\dots,n$, where[5] $\mathbb{C}^+ := \{z \in \mathbb{C} \mid \Re(z) > 0\}$ with $\Re(z)$ denoting the real part of $z \in \mathbb{C}$. The inverse multivariate Laplace transform of function $g : (\mathbb{C}^+)^n \to \mathbb{C}$, is the function $\check{g}$, such that $\mathcal{L}(\check{g}) = g$. We will also write $\check{g} = \mathcal{L}^{-1}(g)$. The existence and uniqueness of the inverse Laplace transform is a well understood subject (cf. [Wid41]). Although there are explicit formulas of the inverse Laplace transform for many functions, generally speaking, in many practical situations the inverse Laplace transform of a function is computed by numerical approximation techniques. We refer the reader to [AW06], and the references therein, for a unified framework for numerically inverting the Laplace transform. For sake of completeness, we present here one such method – the Talbot inversion formula – for one and two dimensional case; the multidimensional case is done by analogy.

Assume that $\widehat{f}$ is the Laplace transform of a function $f : (0,+\infty) \to \mathbb{C}$. The

[5] We will denote by $\Re(z)$ the real part of $z \in \mathbb{C}$, and $\mathrm{i} = \sqrt{-1}$ will be used to denote the imaginary unit.

Talbot inversion formula to approximate f is given by

$$f_M^b(t) = \frac{2}{5t} \sum_{k=0}^{M-1} \Re\left(\gamma_k \widehat{f}(\frac{\delta_k}{t})\right), \tag{A.1}$$

where

$$\begin{aligned} &\delta_0 = \frac{2M}{5}, \quad \delta_k = \frac{22k\pi}{5}(\cot(\frac{k\pi}{M}) + \mathrm{i}), \quad 0 < k < M, \\ &\gamma_0 = \frac{1}{2}e^{\delta_0}, \quad \gamma_k = \left(1 + \mathrm{i}\frac{k\pi}{M}(1 + \cot^2(\frac{k\pi}{M})) - \mathrm{i}\cot(\frac{k\pi}{M})\right) e^{\delta_k}, \quad 0 < k < M. \end{aligned} \tag{A.2}$$

Analogously, given a Laplace transform $\widehat{g}$ of a complex-valued function g of two non-negative real variables, the Talbot inversion formula to compute $g(t_1, t_2)$ numerically is given by

$$g_M^b(t_1, t_2) = \frac{2}{25t_1t_2} \sum_{k_1=0}^{M-1} \Re\left\{\gamma_{k_1} \sum_{k_2=0}^{M-1} \left[\gamma_{k_2}\widehat{g}\left(\frac{\delta_{k_1}}{t_1}, \frac{\delta_{k_2}}{t_2}\right) + \bar{\gamma}_{k_2}\widehat{g}\left(\frac{\delta_{k_1}}{t_1}, \frac{\bar{\delta}_{k_2}}{t_2}\right)\right]\right\},$$

where $\delta_k, \gamma_k, 0 \leq k < M$, are given in (A.2).

Next let us consider a function $f : [0, \infty) \to [0, \infty)$ and its Laplace transform $\widehat{f}(q)$, for $q \in \mathbb{C}^+$. It turns out that the inverse Laplace transform of f can be approximated numerically by using only values of the function $\widehat{f}$ on the positive real line. One such approximation is the Gaver-Stehfest formula

$$f_n(t) = \frac{n \log 2}{t} \binom{2n}{n} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \widehat{f}\left(\frac{(n+k)\log 2}{t}\right). \tag{A.3}$$

For other methods and the comparison of their speeds of convergence we refer to [AW06]. Consecutive application of (A.3) leads to the multivariate Gaver-Stehfest formula.

APPENDIX B

AUXILIARY RESULTS FROM STATISTICAL INFERENCE FOR SPDES

In this section, we present several results that are used in Chapter 3. The first important result is the Strong Law of Large Numbers in which the random variables are not necessarily identically distributed.

Theorem B.0.1 ([Shi96]). **(Strong Law of Large Number)** *Let* $\{\eta_n\}_{n\in\mathbb{N}}$ *be a sequence of independent random variables, and let* $\{b_n\}_{n\in\mathbb{N}}$ *be a sequence of non-decreasing positive numbers such that* $\lim_{n\to\infty} b_n = \infty$*. If*

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var}(\eta_n)}{b_n^2} < \infty,$$

then

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} (\eta_k - \mathbb{E}(\eta_k)) = 0, \quad \mathbb{P} - a.\ s..$$

Remark B.0.1. *As an immediate corollary, if* $\{\eta_n\}_{n\in\mathbb{N}}$ *is a sequence of independent non-negative random variables with*

$$\sum_{n=1}^{\infty} \mathbb{E}(\eta_n) = \infty \quad and \quad \sum_{n=1}^{\infty} \frac{\operatorname{Var}(\eta_n)}{\left(\sum_{k=1}^{n} \mathbb{E}(\eta_k)\right)^2} < \infty,$$

then

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \eta_k}{\sum_{k=1}^{n} \mathbb{E}(\eta_k)} = 1, \quad \mathbb{P} - a.\ s..$$

Next we discuss three central limit theorems in which the random variables are not necessarily identically distributed and/or independent.

Theorem B.0.2 ([Shi96]). **(Lyapunov Central Limit Theorem)** *Let* $\{\eta_n\}_{n\in\mathbb{N}}$ *be a sequence of independent random variables with finite second moments. If there exists* *some* $\delta > 0$, *such that*

$$\lim_{n\to\infty} \frac{1}{\left(\sum_{k=1}^{n} \mathrm{Var}(\eta_k)\right)^{2+\delta}} \sum_{k=1}^{n} \mathbb{E}\left(|\eta_k - \mathbb{E}(\eta_k)|^{2+\delta}\right) = 0, \tag{B.1}$$

then

$$\frac{\sum_{k=1}^{n} (\eta_k - \mathbb{E}(\eta_k))}{\sqrt{\sum_{k=1}^{n} \mathrm{Var}(\eta_k)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad n \to \infty.$$

The next result is a central limit theorem for the summation of Hermite polynomials of stationary Gaussian increments. In this result, the summands are not necessarily independent.

Theorem B.0.3. *Let* $\{X_t, t \geq 0\}$ *be a Gaussian process with the following properties*

(i) $X_0 = 0$, *and* $\mathbb{E}X_t = 0, \quad t \geq 0$.

(ii) $X_{t+s} - X_t \sim \mathcal{N}(0, \sigma^2(s))$, *where* $\sigma(s)$ *is a deterministic function of* s.

(iii) There exists a constant $\gamma > 0$ *such that* $(X_{\alpha t}, t \geq 0) \overset{\mathcal{D}}{=} \alpha^{\gamma} (X_t, t \geq 0)$, *for any* $\alpha > 0$.

(iv) For any $t \geq 0, \Delta t > 0$, *the sequence* $X_{t+n\Delta t} - X_{t+(n-1)\Delta t}$, $n \in \mathbb{N}$ *is stationary. In particular,* $Y_n = \frac{X_n - X_{n-1}}{\sigma(1)}, \quad n \in \mathbb{N}$, *is a zero mean and stationary Gaussian sequence with unit variance.*

(v) Let r *be the covariance function of* Y, $r(n) = \mathbb{E}Y_m Y_{m+n}$, *and assume that for some positive integer* k, $\sum_{n\geq 1} r^k(n) < \infty$.

Then,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} H\left(\frac{n^{\gamma}}{\sigma(1)} \left(X_{j/n} - X_{(j-1)/n}\right); k\right) \xrightarrow[n\to\infty]{\mathcal{D}} \check{\sigma}\mathcal{N}(0,1), \tag{B.2}$$

where

$$\check{\sigma}^2 = \sum_{l=k}^{\infty} c_l^2 l! \check{\sigma}_l^2, \qquad \check{\sigma}_l^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} r^l(|i-j|).$$

Proof. By [BM83, Theorem 1], applied to the sequence Y, we immediately get

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} H(Y_j; k) \xrightarrow[n\to\infty]{\mathcal{D}} \check{\sigma}\mathcal{N}(0,1),$$

where

$$\check{\sigma}^2 = \sum_{l=k}^{\infty} c_l^2 l! \check{\sigma}_l^2, \qquad \check{\sigma}_l^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} r^l(|i-j|).$$

Since

$$(X_{j/n} - X_{(j-1)/n}, j = 1, 2, \ldots, n) \overset{\text{law}}{=} \frac{1}{n^\gamma}(X_j - X_{j-1}, j = 1, 2, \ldots, n),$$

we conclude that (B.2) holds. □

The following result is an immediate consequence of Theorem B.0.3.

Corollary B.0.1. *Let B^H be a fractional Brownian motion with Hurst parameter $H = 1/4$. Then,*

$$\sqrt{n}\left(\mathsf{V}_n^4(B^H; [a,b]) - 3(b-a)\right) \xrightarrow[n\to\infty]{\mathcal{D}} (b-a)\check{\sigma}\mathcal{N}(0,1),$$

where

$$\check{\sigma}^2 = 72\check{\sigma}_2^2 + 24\check{\sigma}_4^2, \qquad \check{\sigma}_l^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} r^l(|i-j|). \tag{B.3}$$

The last central limit theorem we discuss here is a result from [NOL08], used in the proof of Proposition 3.3.2. For most of this part, we will use the standard notations from [Nua06] and [NOL08]. We will denote by $H(x;k)$ a polynomial with Hermite rank k, that is, H can be expanded in the form

$$H(x;k) = \sum_{j=k}^{\infty} c_j H_j(x),$$

where H_j is the jth Hermite polynomial (with leading coefficient 1), and $c_k \neq 0$. Let H be a separable Hilbert space. For every $n \geq 1$, the notation $H^{\otimes n}$ will stand for the nth tensor product of H, and $H^{\odot n}$ will denote the nth symmetric tensor product of H, endowed with the modified norm $\sqrt{n!}\|\cdot\|_{H^{\otimes n}}$. Suppose that $X = \{X(h), h \in H\}$ is an isonormal Gaussian process on H, on some fixed probability space, say $(\Omega, \mathscr{F}, \mathbb{P})$, and assume that $\mathscr{F}$ is generated by X.

For every $n \geq 1$, let $\mathcal{H}_n$ be the nth Wiener chaos of X, that is, the closed linear subspace of $L^2(\Omega, \mathscr{F}, \mathbb{P})$ generated by the random variables $\{H_n(X(h)), h \in H, \|h\|_H = 1\}$, where H_n is the nth Hermite polynomial. We denote by $\mathcal{H}_0$ the space of constant random variables. The mapping $I_n(h^{\otimes n}) = H_n(X(h))$, for $n \geq 1$, provides a linear isometry between $H^{\odot n}$ and $\mathcal{H}_n$. For $n = 0$, we have that $\mathcal{H}_0 = \mathbb{R}$, and take I_0 to be the identity map. It is well known that any square integrable random variable $F \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ admits the following expansion

$$F = \sum_{n=0}^{\infty} I_n(f_n),$$

where $f_0 = \mathbb{E}F$, and the $f_n \in H^{\odot n}$ are uniquely determined by F.

Let $\{e_k, k \geq 1\}$ be a complete orthonormal system in H. Given $f \in H^{\odot n}$ and $g \in H^{\odot m}$, for $\ell = 0, \ldots, n \wedge m$, the contraction of f and g of order ℓ is the element of

$H^{\otimes(n+m-2\ell)}$ defined by

$$f \otimes_\ell g = \sum_{i_1,\ldots,i_\ell} \langle f, e_{i_1} \otimes \cdots \otimes e_{i_\ell} \rangle_{H^{\otimes l}} \otimes \langle g, e_{i_1} \otimes \cdots \otimes e_{i_\ell} \rangle_{H^{\otimes l}}$$

Theorem B.0.4 ([NOL08]). *For $d \geq 2$, fix d natural numbers $1 \leq n_1 \leq \cdots \leq n_d$. Let $\{F_k\}_{k\in\mathbb{N}}$ be a sequence of random vectors of the form*

$$F_k = (F_k^1, \ldots, F_k^d) = (I_{n_1}(f_k^1), \ldots, I_{n_d}(f_k^d)),$$

where $f_k^i \in H^{\odot n_i}$ and I_{n_i} is the Wiener integral of order n_i, such that, for every $1 \leq i, j \leq d$,

$$\lim_{k\to\infty} \mathbb{E}\left[F_k^i F_k^j\right] = \delta_{ij}. \tag{B.4}$$

The following two[6] *statements are equivalent.*

(N1) For all $1 \leq i \leq d, 1 \leq \ell \leq n_i - 1$, $\|f_k^{(i)} \otimes_\ell f_k^{(i)}\|^2_{H^{2\otimes(n_i-\ell)}} \to 0$, as $k \to \infty$.

(N2) The sequence $\{F_k\}_{k\in\mathbb{N}}$, as $k \to \infty$, converges in distribution to a d-dimensional standard Gaussian vector $\mathcal{N}_d(0, I_d)$.

Finally, we recall the BDG inequality that is used in the proof of Theorem 3.2.2.

Theorem B.0.5 ([Cho07]). *Let $M_t, t \in [0,T]$ be any continuous real-valued martingale with $M_0 = 0$ and $\mathbb{E}|M_T|^p < \infty$. Then for any $p > 0$, there exist two positive constants c_p and C_p such that*

$$c_p\mathbb{E}\left([M]_T^{p/2}\right) \leq \mathbb{E}\left(\sup_{t\in[0,T]} |M_t|^p\right) \leq C_p\mathbb{E}\left([M]_T^{p/2}\right).$$

[6]The original result [NOL08, Theorem 7] contains six equivalent conditions; we list only those two that we use in this presentation.

BIBLIOGRAPHY

[AES16] S. Aazizi and K. Es-Sebaiy. Berry-Esseen bounds and almost sure CLT for the quadratic variation of the bifractional Brownian motion. *Random Oper. Stoch. Equ.*, 24(1):1–13, 2016.

[APU03] F. Avram, M. R. Pistorius, and M. Usabel. The two barriers ruin problem via a Wiener Hopf decomposition approach. *An. Univ. Craiova Ser. Mat. Inform.*, 30(1):38–44, 2003.

[Asm95] S. Asmussen. Stationary distributions for fluid flow models with or without Brownian noise. *Comm. Statist. Stochastic Models*, 11(1):21–49, 1995.

[AW06] J. Abate and W. Whitt. A unified framework for numerically inverting laplace transforms. *INFORMS Journal on Computing*, 18(4):408–421, 2006.

[BÏ4] B. Böttcher. Feller evolution systems: generators and approximation. *Stoch. Dyn.*, 14(3):1350025, 15, 2014.

[BCGH18] T. R. Bielecki, I. Cialenco, R. Gong, and Y. Huang. Wiener-hopf factorization for time-inhomogeneous markov chains and its application. *arXiv preprint arXiv:1801.05553*, 2018.

[Bis08] J. P. N. Bishwal. *Parameter estimation in stochastic differential equations*, volume 1923 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008.

[BM83] P. Breuer and P. Major. Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.*, 13(3):425–441, 1983.

[BRW80] M. T. Barlow, L. C. G. Rogers, and D. Williams. Wiener-hopf factorization for matrices. *Sminaire de probabilits de Strasbourg*, 14:324–331, 1980.

[BT17] M. Bibinger and M. Trabs. Volatility estimation for stochastic pdes using high-frequency observations. *arXiv preprint arXiv:1710.03519*, 2017.

[CGH11] I. Cialenco and N. Glatt-Holtz. Parameter estimation for the stochastically perturbed Navier-Stokes equations. *Stochastic Process. Appl.*, 121(4):701–724, 2011.

[CGH18] I. Cialenco, R. Gong, and Y. Huang. Trajectory fitting estimators for spdes driven by additive noise. *Statistical Inference for Stochastic Processes*, 21(1):1–19, Apr 2018.

[CH17] I. Cialenco and Y. Huang. A note on parameter estimation for discretely sampled (spde)s. *arXiv preprint arXiv:1710.01649*, 2017.

[Cho07] P. Chow. *Stochastic partial differential equations*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[Cia10] I. Cialenco. Parameter estimation for SPDEs with multiplicative fractional noise. *Stoch. Dyn.*, 10(4):561–576, 2010.

[Cia18] I. Cialenco. Statistical inference for SPDEs: an overview. *Statistical Inference for Stochastic Processes*, Feb 2018.

[CL09] I. Cialenco and S. V. Lototsky. Parameter estimation in diagonalizable bilinear stochastic parabolic equations. *Stat. Inference Stoch. Process.*, 12(3):203–219, 2009.

[CLP09] I. Cialenco, S. V. Lototsky, and J. Pospíšil. Asymptotic properties of the maximum likelihood estimator for stochastic parabolic equations with additive fractional Brownian motion. *Stoch. Dyn.*, 9(2):169–185, 2009.

[CNW06] J. M. Corcuera, D. Nualart, and J. H. C. Woerner. Power variation of some integral fractional processes. *Bernoulli*, 12(4):713–735, 2006.

[Cor12] J. M. Corcuera. New central limit theorems for functionals of Gaussian processes and their applications. *Methodol. Comput. Appl. Probab.*, 14(3):477–500, 2012.

[CX14] I. Cialenco and L. Xu. A note on error estimation for hypothesis testing problems for some linear SPDEs. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2(3):408–431, 2014.

[CX15] I. Cialenco and L. Xu. Hypothesis testing for stochastic PDEs driven by additive noise. *Stochastic Process. Appl.*, 125(3):819–866, March 2015.

[DPZ92] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[Hai09] M. Hairer. *Introduction to Stochastic PDEs*. Unpublished lecture notes, 2009.

[Hie14] P. Hieber. First-passage times of regime switching models. *Statist. Probab. Lett.*, 92:148–157, 2014.

[HKR93] M. Huebner, R. Khasminskii, and B. L. Rozovskii. Two examples of parameter estimation for stochastic partial differential equations. In *Stochastic processes*, pages 149–160. Springer, New York, 1993.

[HLR97] M. Huebner, S. V. Lototsky, and B. L. Rozovskii. Asymptotic properties of an approximate maximum likelihood estimator for stochastic PDEs. In *Statistics and control of stochastic processes (Moscow, 1995/1996)*, pages 139–155. World Sci. Publishing, 1997.

[HR95] M. Huebner and B. L. Rozovskii. On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. *Probab. Theory Related Fields*, 103(2):143–163, 1995.

[HSZ16] D. Hainaut, Y. Shen, and Y. Zeng. How do capital structure and economic regime affect fair prices of bank's equity and liabilities? *Annals of Operations Research*, Apr 2016.

[JP08] Z. Jiang and M. R. Pistorius. On perpetual American put valuation and first-passage in a regime-switching model with jumps. *Finance Stoch.*, 12(3):331–355, 2008.

[JP12] Z. Jiang and M. R. Pistorius. Optimal dividend distribution under Markov regime switching. *Finance Stoch.*, 16(3):449–476, 2012.

[JR06] A. Jobert and L. C. G. Rogers. Option pricing with Markov-modulated dynamics. *SIAM J. Control Optim.*, 44(6):2063–2078, 2006.

[Kho14] D. Khoshnevisan. *Analysis of stochastic partial differential equations*, volume 119 of *CBMS Regional Conference Series in Mathematics*. the American Mathematical Society, Providence, RI, 2014.

[Kut91] Y. A. Kutoyants. Minimum-distance parameter estimation for diffusion-type observations. *C. R. Acad. Sci. Paris Sér. I Math.*, 312(8):637–642, 1991.

[Kut04] Y. A. Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London Ltd., London, 2004.

[KW90] J. Kennedy and D. Williams. Probabilistic factorization of a quadratic matrix polynomial. *Math. Proc. Cambridge Philos. Soc.*, 107(3):591–600, 1990.

[LMRW82] R. R. London, H. P. McKean, L. C. G. Rogers, and D. Williams. A martingale approach to some Wiener-Hopf problems. I, II. In *Seminar on Probability, XVI*, volume 920 of *Lecture Notes in Math.*, pages 41–67, 68–90. Springer, Berlin-New York, 1982.

[Lot09] S. V. Lototsky. Statistical inference for stochastic parabolic equations: a spectral approach. *Publ. Mat.*, 53(1):3–45, 2009.

[LR17] S. V. Lototsky and B. L. Rozovsky. *Stochastic partial differential equations*. Universitext. Springer, Cham, 2017.

[Mar03] B. Markussen. Likelihood inference for a discretely observed stochastic partial differential equation. *Bernoulli*, 9(5):745–762, 2003.

[MB95] M. N. Mishra and J. P. N. Bishwal. Approximate maximum likelihood estimation for diffusion processes from discrete observations. *Stochastics and Stochastic Reports*, 52(1-2):1–13, 1995.

[Mit88] D. Mitra. Stochastic theory of a fluid model of producers and consumers coupled by a buffer. *Adv. in Appl. Probab.*, 20(3):646–676, 1988.

[MP11] A. Mijatović and M. R. Pistorius. Exotic derivatives under stochastic volatility models with jumps. In *Advanced mathematical methods for finance*, pages 455–508. Springer, Heidelberg, 2011.

[MPR02] M. N. Mishra and B. L. S. Prakasa Rao. Approximation of maximum likelihood estimator for diffusion processes from discrete observations. *Stochastic Anal. Appl.*, 20(6):1309–1329, 2002.

[NOL08] D. Nualart and S. Ortiz-Latorre. Central limit theorems for multiple stochastic integrals and malliavin calculus. *Stochastic Processes and their Applications*, 118(4):614 – 628, 2008.

[Nou08] I. Nourdin. Asymptotic behavior of weighted quadratic and cubic variations of fractional Brownian motion. *Ann. Probab.*, 36(6):2159–2175, 2008.

[Nua06] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[PR97] L. I. Piterbarg and B. L. Rozovskii. On asymptotic problems of parameter estimation in stochastic PDE's: discrete time sampling. *Math. Methods Statist.*, 6(2):200–223, 1997.

[PR03] B. L. S. Prakasa Rao. Estimation for some stochastic partial differential equations based on discrete observations. II. *Calcutta Statist. Assoc. Bull.*, 54(215-216):129–141, 2003.

[PvsT07] J. Pospí šil and R. Tribe. Parameter estimates and exact variations for stochastic heat equations driven by space-time white noise. *Stoch. Anal. Appl.*, 25(3):593–611, 2007.

[Rog94] L. C. G. Rogers. Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains. *Ann. Appl. Probab.*, 4(2):390–413, 1994.

[Roz90] B. L. Rozovskii. *Stochastic evolution systems*, volume 35 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990. Linear theory and applications to nonlinear filtering.

[RS94] L. C. G. Rogers and Z. Shi. Computing the invariant law of a fluid model. *J. Appl. Probab.*, 31(4):885–896, 1994.

[Shi96] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996.

[Shu01] M. A. Shubin. *Pseudodifferential operators and spectral theory*. Springer-Verlag, Berlin, second edition, 2001.

[Sys92] R. Syski. *Passage times for Markov chains*, volume 1 of *Studies in Probability, Optimization and Statistics*. IOS Press, Amsterdam, 1992.

[Wal81] J. B. Walsh. A stochastic model of neural response. *Adv. in Appl. Probab.*, 13(2):231–281, 1981.

[Wid41] D.V. Widder. *The Laplace Transform*. Princeton mathematical series. Princeton University Press, 1941.

[Wil91] D. Williams. Some aspects of Wiener-Hopf factorization. *Philos. Trans. Roy. Soc. London Ser. A*, 335(1639):593–608, 1991.

[Wil08] D. Williams. A new look at 'Markovian' Wiener-Hopf theory. In *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 349–369. Springer, Berlin, 2008.