# Mathematical modeling, estimation and application in finance

By

Yiying Cheng

Submitted to the Department of Mathematics and the
Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Committee members

Xuemin Tu, Chairperson

Yaozhong Hu , Co-Chair

David Nualart

Zhipeng Liu

Jie Han

Date defended: Feb 11, 2021

The Dissertation Committee for Yiying Cheng certifies
that this is the approved version of the following dissertation :

Mathematical modeling, estimation and application in finance

Xuemin Tu, Chairperson

Yaozhong Hu , Co-Chair

Date approved: Feb 11, 2021

# Abstract

Parameter estimation has wide applications in such fields as finance, oil deposit detection, etc. In this dissertation, we discuss the parameter estimation problems in a stochastic differential equation and a partial differential equation.

In chapter one, we provide a general moment estimator for the Ornstein-Uhlenbeck Process driven by $\alpha$-stable Lévy motion. When the noise is an $\alpha$-stable Lévy motion, the process does not have the second moment which makes the parametric estimation problem more difficult. In this case, there are limited papers dealing with the parametric estimation problem. In previous work, one can only estimate the drift parameter $\theta$ assuming the other parameters ($\alpha$, $\sigma$ and $\beta$) are known and under discrete observations. In most literature, one also needs to assume that the time step $h$ depends on $n$ and converges to 0 as $n$ goes to infinity. This means that a high frequency data must be available for the estimators to be effective. The main mathematical tool that we use is ergodic theory and sample characteristic functions so that we can estimate all the parameters simultaneously. We also obtain the strong consistency and asymptotic normality of the proposed joint estimators when the time step $h$ remains constant.

In chapter three, we describe how to use implicit sampling in parameter estimation problems where the goal is to find parameters of a partial differential equation, such that the output of the numerical model is compatible with data. We could generate independent samples, so that some of the practical difficulties one encounters with Markov Chain Monte Carlo methods, e.g. burn-in time or correlations among dependent samples, are avoided. We describe a new implementation of implicit sampling for parameter estimation problems that makes use of a class of overlapping Newton

Krylov-Schwarz algorithms to solve it. With a reasonably large overlap, the Newton Krylov-Schwarz method is scalable and capable of finding the solution with noise. The comparison with BFGS method demonstrates the superiority of our method. We also use the local Karhunan-Loève expansion to reduce the dimension of the parameter which enables the parallel and efficient computation of a possibly large number of dominant KL modes.

Another important topic considered in this dissertation is in chapter two, a novel approach for solving optimal price adjustment problems, when the underlying process is geometric Brownian motion process. Several countries use the administratively-set fuel prices close to their international free market counterparts. However, chasing a global market price of energy has the disadvantage that the domestic prices need to fluctuate daily. This creates uncertainties for households and firms and expose them to global price shocks. In chapter two, we offers a model of adjustment rule which is based on optimal lower and upper price barriers. Once the ratio of the domestic and global price hit the bounds, the domestic price will then be readjusted to the original desired level. We offer a procedure to use expected hitting time approach to solve the model, which does not require solving a PDE or running Monte-Carlo simulations. We characterize the optimal policy behavior as a function of underlying parameters and also compare the gains from adopting an optimal policy versus a mechanical policy.

# Acknowledgements

I want to take this opportunity to express my gratitude to all the people who have helped me during my Ph.D. study, and who made the six years at University of Kansas a lifestyle memory.

I would like to give my deepest gratitude to my advisors Professor Yaozhong Hu and Professor Xuemin Tu for their sagacious guidance and inspiring encouragement throughout my PhD study in Mathematics. They also helped me and gave me many suggestions for developing my career plan. Without their consistent help, nothing in this dissertation could be achieved.

I would like to thank Professor David Nualart helped me to extend my knowladge on stochastic calculus. I am greatly appreciate my committee members Professor Zhipeng Liu and Professor Jie Han for serving as my committee members.

I am grateful to Professor Weizhang Huang, Professor Terry Soo, Professor Agnieszka Miedlar for their teaching and valuable advices.

I would like to thank Professor Jila Niknejad, Professor Jack Porter sharing with me precious teaching experiences. Many thanks also go to Ms. Kate Pleskac, Ms. Lori Springs, Ms. Gloria Prothe, and Ms. Kerrie Brecheisen for their help with administrative arrangements.

I would also like to thank Jun Fu for her care and support, my friends Yanru Su, Panqiu Xia, Promit Kundu, Jinjin Zhang, Seyed Hamidreza Mofidi for their help in many ways.

Finally, I would like to thank my parents and my husband Chang Liu for their love and encouragement that support me all the time.

# Contents

# Chapter 1

# Generalized moment estimators for $\alpha$-stable Ornstein–Uhlenbeck motions from discrete observations

In this chapter, we study the parameter estimation problem for discretely observed Ornstein-Uhlenbeck processes driven by $\alpha$-stable Lévy motions. A method of moments via ergodic theory and via sample characteristic functions is proposed to estimate all the parameters involved in the Ornstein-Uhlenbeck processes. We obtain the strong consistency and asymptotic normality of the proposed joint estimators when the sample size $n \to \infty$ while the sampling time step h remains arbitrarily fixed. And the numerical simulation illustrate our result.

## 1.1 Preliminary

**Definition 1.1.** Generalized Ornstein-Uhlenbeck processes driven by Lévy processes satisfies the following stochastic Langevin equation

$$\begin{cases} dX_t = -\theta X_t dt + dL_t, \quad t \geq 0 \\ X_0 = x_0 \end{cases} \tag{1.1.1}$$

where $\theta$ is an unknown parameter, $\{L_t, t \geq 0\}$ is a one-dimensional Lévy process. Lévy processes are closely related to stable distributions.

**Definition 1.2.** A random variable $\eta$ is said to follow a stable distribution, denoted by $\eta \sim$

$S_\alpha(\sigma,\beta,\gamma)$, if its characteristic function has the following form:

$$\phi_\eta(u)=\mathbb{E}\left[e^{iu\eta}\right]=\begin{cases}\exp\left\{-\sigma^\alpha|u|^\alpha\left(1-i\beta\operatorname{sgn}(u)\tan\frac{\alpha\pi}{2}\right)+i\gamma u\right\}, & \text{if } \alpha\neq 1\\ \exp\left\{-\sigma|u|\left(1+i\beta\frac{2}{\pi}\operatorname{sgn}(u)\log|u|\right)+i\gamma u\right\}, & \text{if } \alpha=1\end{cases}$$

In the above definition $\alpha\in(0,2],\sigma\in(0,\infty),\beta\in[-1,1]$, and $\gamma\in(-\infty,\infty)$ are called the index of stability, the scale, skewness, and location parameters, respectively.

We shall assume $\gamma=0$. This means that we consider only strictly $\alpha$ -stable distribution. If in addition $\beta=0$, we call $\eta$ symmetric $\alpha$ -stable. Note that $\eta$ is strictly stable $(\alpha=1)$ if and only if $\beta=0$.

**Definition 1.3.** An $\mathscr{F}_t$ – adapted stochastic process $\{Z_t\}_{t>0}$ is called a standard $\alpha$ -stable Lévy motion if

(i) $Z_0=0$, a.s.;

(ii) $Z_f-Z_s\sim S_\alpha\left((t-s)^{1/\alpha},\beta,0\right),t>s\geq 0$

(iii) For any finite time points $0\leq s_0<s_1<\cdots<s_m<\infty$, the random variables $Z_{s_0},Z_{s_1}-Z_{s_0},\cdots,Z_{s_m}-Z_{s_{m-1}}$ are independent.

Suppose that $\{L_t,t\geq 0\}$ is a Lévy process generated by the triplet $(0,\rho,\lambda)$, i.e. the distribution of $L_t$ has characteristic function

$$\phi_{L_t}(u)=E\left[e^{iuL_t}\right]=\exp\left\{it\lambda u+t\int_{\mathbb{R}\setminus\{0\}}\left(e^{iux}-1-iux1_D(x)\right)\rho(dx)\right\},u\in\mathbb{R}$$

where $D=\{x:|x|\leq 1\}$ and $\rho$ is the Lévy measure given by

$$\rho(dx)=\frac{c_1}{x^{1+\alpha}}1_{(0,\infty)}(x)dx+\frac{c_2}{|x|^{1+\alpha}}1_{(-\infty,0)}(x)dx$$

where $1<\alpha<2,c_1\geq 0,c_2\geq 0$, and $c_1+c_2>0$. It is easy to see that (2.1) can be written as

$$\phi_{L_t}(u)=\exp\left\{it\left(\lambda+\int_{|x|>1}x\rho(dx)\right)u-t\sigma^\alpha|u|^\alpha\left[1-\beta\operatorname{sgn}(u)\tan\left(\frac{\pi\alpha}{2}\right)\right]\right\}$$

where $\sigma^{\alpha} = -(c_1+c_2)\Gamma(-\alpha)\cos(\pi\alpha/2)$ and $\beta = (c_1-c_2)/(c_1+c_2)$. Then, by the Itô-Lévy decomposition, we have

$$L_t = \lambda t + \int_0^t \int_{|x|<1} x\tilde{N}(ds,dx) + \int_0^t \int_{|x|\geq 1} xN(ds,dx)$$

where $N(dt,dx)$ is a Poisson random measure defined by

$$N((0,t],A) = \sum_{s\leq t} 1_A(\Delta L_s)$$

for $A \in \mathscr{B}(\mathbb{R}\backslash\{0\})$ and $\Delta L_s = L_s - L_{s-}$ denoting the jump of $L_s$ at time $s$, and the compensated Poisson random measure $\tilde{N}(dt,dx)$ is given by

$$\tilde{N}((0,t],A) = N((0,t],A) - t\rho(A)$$

with

$$\rho(A) = \int_A \rho(dx)$$

The Itô-Lévy decomposition can be rewritten as

$$\begin{aligned} L_t &= \lambda t + \int_0^t \int_{\mathbb{R}\backslash\{0\}} x\tilde{N}(ds,dx) + t\int_{|x|\geq 1} x\rho(dx) \\ &= \left(\lambda + \int_{|x|\geq 1} x\rho(dx)\right)t + \int_0^t \int_{\mathbb{R}\backslash\{0\}} x\tilde{N}(ds,dx) \end{aligned} \tag{1.1.2}$$

Let

$$m = \lambda + \int_{|x|>1} x\rho(dx)$$

Then

$$m = \lambda + \frac{c_1 - c_2}{\alpha - 1}$$

Denote

$$\tilde{Z}_t = \int_0^t \int_{\mathbb{R}\setminus\{0\}} x\tilde{N}(ds,dx)$$

Then $\tilde{Z}_t$ is a $\alpha$ -stable Lévy motion and $\tilde{Z}_t - \tilde{Z}_s \sim S_\alpha\left(\sigma(t-s)^{1/\alpha},\beta,0\right)$ for any $0 \leq s < t < \infty$. We can renormalize $\tilde{Z}_t$ and define $Z_t = \tilde{Z}_t/\sigma$. Then we can easily see that $\{Z_t, t \geq 0\}$ is a standard $\alpha$ -stable Lévy motion (see Janicki and Weron [22]) so that $Z_1$ has a stable distribution $S_\alpha(1,\beta,0)$. It is clear that $L_t = mt + \sigma Z_t$ and $E[L_t] = mt$. If we assume $\theta > 0$, then $X_t$ is ergodic and the solution of the SDE (1.1.1) can be written in the following way:

$$\begin{aligned} X_t &= e^{-\theta t}x + \int_0^t e^{-\theta(t-s)} dL_s \\ &= e^{-\theta t}x + m\int_0^t e^{-\theta(t-s)} ds + \sigma \int_0^t e^{-\theta(t-s)} dZ_s \end{aligned} \tag{1.1.3}$$

The general properties of generalized Ornstein-Uhlenbeck processes driven by Lévy processes have been comprehensively studied in the monograph of Sato [40] We shall use some important results in Sato [40] freely.

**Lemma 1.4.** The generalized Ornstein-Uhlenbeck processes $\{X_t, t \geq 0\}$ (generated by the triplet $(0,\rho,\lambda)$ ) have a unique invariant distribution $\mu_\infty$ which is self-decomposable and generated by the triplet $(0,\nu,\mu)$ with

$$\nu(B) = \frac{1}{\theta}\int_{\mathbb{R}} \rho(dy) \int_0^\infty 1_B\left(e^{-s}y\right) ds, B \in \mathscr{B}(R)$$

and

$$\mu = \frac{\lambda}{\theta} + \frac{1}{\theta}\int_{|y|>1} \frac{y}{|y|}\rho(dy)$$

*Proof.* By Theorem 17.5 of Sato [40], we only need to verify that the Lévy measure $\rho$ satisfies the following condition

$$\int_{|x|>2} \log|x|\rho(dx) < \infty$$

In fact, we have

$$\begin{aligned}
\int_{|x|>2} \log|x|\rho(dx) &= \int_{|x|>2} \log|x| \left(\frac{c_1}{x^{1+\alpha}} 1_{(0,\infty)}(x) + \frac{c_2}{|x|^{1+\alpha}} 1_{(-\infty,0)}(x)\right) dx \\
&= c_1 \int_2^{\infty} \frac{\log x}{x^{1+\alpha}} dx + c_2 \int_{-\infty}^{-2} \frac{\log(-x)}{(-x)^{1+\alpha}} dx \\
&= (c_1 + c_2) \int_2^{\infty} \log x \cdot x^{-1-\alpha} dx \\
&= \frac{c_1 + c_2}{\alpha 2^{\alpha}} \left(\log 2 + \frac{1}{\alpha}\right) < \infty
\end{aligned}$$

This completes the proof. □

. In the following, we consider the case when $m = 0$, i.e the Ornstein-Uhlenbeck processed driven by the $\alpha$-stable Lévy motion. We can easily find that $X_t$ converges weakly to a random variable

$$X_o = \sigma \int_0^{\infty} e^{-\theta s} dZ_s$$

**Theorem 1.5.** *(Ergodic theorem)(Luis Barreira, Yakov Pesin*[28]*)*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(X_{t_k}\right) = \mathbb{E} f\left(\tilde{X}_o\right)$$

*almost surely, where we recall that the distribution of $\tilde{X}_o$ is the invariant measure of the $\alpha$ -stable Ornstein-Uhlenbeck motion $X_t$.*

**Theorem 1.6.** *(A more sophisticated ergodic theorem)(Luis Barreira, Yakov Pesin*[28]*)*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(X_{t_k}, X_{t_{k+1}}\right) = \mathbb{E} f\left(\tilde{X}_0, \tilde{X}_{t_1}\right)$$

*where $\tilde{X}_t$ satisfies (1.1.1) with initial condition $\tilde{X}_0$ having the invariant measure (namely, $\tilde{X}_0$ and $\tilde{X}_o$ have the same probability measure) and being independent of the $\alpha$-stable motino $Z_t$*

Note that the explicit forms of the probability density functions of $\tilde{X}_o$ and the joint probability density function of $\tilde{X}_0$, $\tilde{X}_{t_1}$ are unknown except for some very special parameters. However, it is possible to find the explicit forms of the explicit forms of the characteristic functions of $\tilde{X}_o$ and that of $\tilde{X}_0$, $\tilde{X}_{t_1}$.

**Definition 1.7.** 1. Strong Consistency: For every $\varepsilon > 0 \quad X_t \overset{a.s.}{\rightarrow} \mu$ if

$$P\left(\omega; \cap_{m=1}^{\infty} \cup_{t=m}^{\infty} \{|X_t(\omega) - \mu| > \varepsilon\}\right) = 0$$

2. Asymptotic Normality. We say that $\hat{\theta}$ is asymptotically normal if

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \rightarrow^d N\left(0, \sigma_{\theta_0}^2\right)$$

where $\sigma_{\theta_0}^2$ is called the asymptotic variance of the estimate $\hat{\theta}$. Asymptotic normality says that the estimator not only converges to the unknown parameter, but it converges fast enough, at a rate $1/\sqrt{n}$.

## 1.2 Introduction

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a basic probability space equipped with a right continuous and increasing family of $\sigma$-algebras $(\mathscr{F}_t, t \geq 0)$ and let $(Z_t, t \geq 0)$ be a standard $\alpha$-stable Lévy motion with $Z_1 \sim S_\alpha(1, \beta, 0)$, where $\alpha$ is the stability index and $\beta \in [-1, 1]$ is the skewness parameter (we shall briefly recall the relevant definitions in next section). The so-called $\alpha$-stable Ornstein-Uhlenbeck motion $X = (X_t, t \geq 0)$, starting from a point $x_0 \in \mathbb{R}$ satisfies the following stochastic Langevin equation

$$dX_t = -\theta X_t dt + \sigma dZ_t\,, \quad t \in [0, \infty)\,, \quad X_0 = x_0, \tag{1.2.1}$$

where $\theta$, $\sigma$ are some constants. Recalll that if $\theta > 0$, $X_t$ is ergodic and it converges in law to the random variable $X_o = \sigma \int_0^\infty e^{-\theta s} dZ_s$. From the above definition we see that the $\alpha$-stable Ornstein-Uhlenbeck motion $X_t$ depends on the following parameters: the stability index $\alpha$, the skewness parameter $\beta$, the drift parameter $\theta$ and the dispersion parameter $\sigma$. In this work we assume that the values of these parameters are unknown but the $\alpha$-stable Ornstein-Uhlenbeck motion $(X_t, t \geq 0)$ can be observed at discrete time $t_k$ (For simplicity, we let $t_k = kh$ for some fixed $h > 0$). We want to use the available data $\{X_{t_k}, k = 1, 2, \cdots, n\}$ to estimate the parameters $\alpha, \beta, \theta$, and $\sigma$ simultaneously.

The parametric estimation problems for diffusion processes driven by a Lévy process such as compound Poisson process, gamma process, inverse Gaussian process, variance gamma process, normal inverse Gaussian process or some generalized tempered stable processes have been studied earlier. Let us mention the following works: Brockwell et al. [5], Masuda [31], Ogihara and Yoshida [33], Shimizu[41], Shimizu and Yoshida [42], Spiliopoulos [44], and Valdivieso et al [46]. In these works it is considered the quasi-maximum likelihood, least squares estimators, or trajectory-fitting estimator and it is also established the consistency and asymptotic normality for those estimators. Masuda [13] proposed a self-weighted least absolute deviation estimator for discretely observed ergodic Ornstein-Uhlenbeck processes driven by symmetric Lévy processes. For some recent developments on estimation of drift parameters for stochastic processes driven by small Lévy noises, we refer to Long et al. ([26], [25]) as well as related references therein.

However, all aforementioned papers did not cover the case that the noise is given by an $\alpha$-stable Lévy motion. When the noise is an $\alpha$-stable Lévy motion the process does not have the second moment which makes the parametric estimation problem more difficult. In this case there are limited papers dealing with the parametric estimation problem. Let us first summarize some relevant work. Hu and Long ([18], [17]) proposed the trajectory fitting estimator and least squares estimators for the drift parameter $\theta$ assuming other parameters $\alpha$, $\beta$, and $\sigma$ are known and under both continuous or discrete observations. They discovered that the limiting distributions are stable distributions which are different from the classical ones where asymptotic distributions are normal. Fasen [10] extended the results of Hu and Long [17] to high dimensions.

To deal with the discrete time observations, which is the common practice and the main focus of this paper, in most literature, one needs to assume that the time step $h$ depends on $n$ and converges to 0 as $n$ goes to infinity. This means that a high frequency data must be available for the estimators to be effective. In some situations such as in finance high frequency data collection is possible.

But in many other situations high frequency data collection may be impossible or too expensive. To construct estimators applicable to this situation, one has to find consistent estimators which allow $h$ to be an arbitrarily fixed constant. Along with this line, some progresses have been made in Hu and Song [19] and Hu et al.[16] for Ornstein-Uhlenbeck processes or reflected Ornstein-Uhlenbeck processes driven by Brownian motion or fractional Brownian motions as well as Zhang and Zhang [39] for Ornstein-Uhlenbeck processes driven by symmetric $\alpha$-stable motions. The idea is to use the ergodic theorems for the underlying Ornstein-Uhlenbeck processes to construct ergodic type estimators. The strong consistency and the asymptotic normality are proved when the time step $h$ remains constant (as the number of sample point $n$ goes to infinity). However, in the above papers, one can only estimate the drift parameter $\theta$. There have been no available estimators simultaneously for all parameters. The main goal of the present paper is to fill this gap. We want to simultaneously estimate all the parameters $\theta, \alpha, \sigma$ and $\beta$ in the $\alpha$-stable Ornstein-Uhlenbeck motion. We still use the generalized method of moments via ergodic theory. But since the $\alpha$-stable motion has no second or higher moments we shall use the sample characteristic functions. Namely, we use the following the ergodic theorem: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_{t_k}) = \mathbb{E} f(\tilde{X}_o)$ almost surely, where we recall that the distribution of $\tilde{X}_o$ is the invariant measure of the $\alpha$-stable Ornstein-Uhlenbeck motion $X_t$. However, this cannot be used to estimate all the parameters $\theta, \alpha, \sigma$ and $\beta$ since we cannot separate all the parameters in the stationary distribution of $\tilde{X}_o$ (see Remark 1.8). The idea is then to use a more sophisticated ergodic theorem: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_{t_k}, X_{t_{k+1}}) = \mathbb{E} f(\tilde{X}_0, \tilde{X}_{t_1})$, where $\tilde{X}_t$ satisfies (1.2.1) with initial condition $\tilde{X}_0$ having the invariant measure (namely, $\tilde{X}_0$ and $\tilde{X}_o$ have the same probability measure) and being independent of the $\alpha$-stable motion $Z_t$. Note that the explicit forms of the probability density functions of $\tilde{X}_o$ and the joint probability density function of $\tilde{X}_0, \tilde{X}_{t_1}$ are unknown except for some very special parameters. However, it is possible to find the explicit forms of the characteristic functions of $\tilde{X}_o$ and that of $\tilde{X}_0, \tilde{X}_{t_1}$. These characteristic functions will be used to construct estimators for $\theta, \alpha, \sigma$ and $\beta$.

To validate our approach we have done a number of simulations to illustrate our estimators. First, we simulate some data from 1.2 assuming some given values of $\alpha, \beta, \theta$ and $\sigma$. Then we apply our estimators to estimate these parameters. The numerical results show that our estimators are accurate and converge fast to all the true parameters. Our estimators work for all fixed $h > 0$ (even large $h$) although we list only $h = 0.5$ (which is already big enough). As discussed in Rosinski [20] and Zhang [38], the Euler scheme in simulating Ornstein-Uhlenbeck process driven by a Lévy process is seldom used. To save computation time we find a way to simulate the $\alpha$-stable Lévy motion $\{X_{kh}, k = 1, \cdots, n\}$ in a straightforward way without any extra computations.

We note that another method of estimating all the parameters for time series models is the ECF (empirical characteristic function) method discussed in Knight and Yu [23] and Yu[21]. They fit the ECF to the theoretical one continuously in frequency by minimizing a distance measure between the joint CF (characteristic function) and joint ECF. Under certain regularity conditions, they established consistency, asymptotic normality, and the asymptotic efficiency of the proposed ECF estimators. The i.i.d. case was discussed much earlier by Paulson et al. [35] and Heathcote [7], where they called it the integrated squared error method.

In this project, we employ the well-known generalized method of moments (GMM) for parameter estimation. GMM is referred to a class of estimators which can be constructed by utilizing the sample moment counterparts of population moments. It nests the classical method of moments, least squares method, and maximum likelihood method. GMM has been extensively studied and widely used in many applications since the seminal work of Hansen [27]. In particular, GMM has been successfully applied to parameter estimation and inference for stochastic models in finance including foreign exchange markets and asset pricing in Hansen and Hodrick [14], Hansen and Singleton [15]. For a comprehensive treatment of GMM, we refer to Hall[1]. For generalization and improvement on GMM, we refer to Qian and Schmidt [37] and Lynch and Wachter[29]

In Section 1.3, we recall some basic results for $\alpha$-stable Lévy motions which we need in our work. In Section 1.4, we construct estimators for all the parameters in the $\alpha$-stable Ornstein-Uhlenbeck motion by using ergodic theory and sample characteristic functions. The consistency of the estimators is established. The asymptotic normality of the joint estimators is obtained and the asymptotic covariance matrix is computed. The asymptotic covariance depends on the parameters we choose in the characteristic function. We also design a procedure of selecting the four grid points used for the parameter estimation in certain optimal way. Section 1.5 provides validation of our estimators from numerical simulations. The values of the (true) parameters are given and then they are used to simulate the $\alpha$-stable Ornstein-Uhlenbeck motion $X_t$. With these simulated values we compute our estimators and compare them with the true parameters. Numerical results show that our estimators converges fast to the true parameters. Finally, all the lemmas with their proofs, proof of Theorem 1.9, and the explicit expression of the crucial covariance matrix defined in Section 1.4 are provided in Section 1.6.

## 1.3 Limiting distributions of $\alpha$-stable Ornstein-Uhlenbeck motions

We first recall some basic definitions. A random variable $\eta$ is said to follow a stable distribution, denoted by $\eta \sim S_\alpha(\sigma,\beta,\gamma)$, if its characteristic function has the following form:

$$\phi_\eta(u) = \mathbb{E}[e^{iu\eta}] = \begin{cases} \exp\left\{-\sigma^\alpha |u|^\alpha \left(1 - i\beta \operatorname{sgn}(u) \tan\frac{\alpha\pi}{2}\right) + i\gamma u\right\}, & \text{if } \alpha \neq 1, \\ \exp\left\{-\sigma |u| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(u) \log|u|\right) + i\gamma u\right\}, & \text{if } \alpha = 1\,. \end{cases}$$

In the above definition $\alpha \in (0,2]$, $\sigma \in (0,\infty)$, $\beta \in [-1,1]$, and $\gamma \in (-\infty,\infty)$ are called the index of stability, the scale, skewness, and location parameters, respectively.

When $Z$ is an $\alpha$-stable Lévy motion, the stochastic Langevin equation (1.2.1) has a unique solution which is given explicitly by

$$X_t = e^{-\theta t} x_0 + \sigma \int_0^t e^{-\theta(t-s)} dZ_s. \tag{1.3.1}$$

It is known that the $\alpha$-stable Ornstein-Uhlenbeck motion $X_t$ has a limiting distribution which coincides with the distribution of $\tilde{X}_o = \sigma \int_0^\infty e^{-\theta s} dZ_s$. It is also well-known that $X_t$ is ergodic. This means that for any function $f:\mathbb{R}\xrightarrow{d}\mathbb{R}$ such that $\mathbb{E}|f(\tilde{X}_o)|<\infty$ we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}f(X_{t_j})=\mathbb{E}\left[f(\tilde{X}_o)\right] \tag{1.3.2}$$

almost surely. The explicit computation of the above right hand side is usually difficult for general function $f$ since the the explicit form of the probability density function of $\tilde{X}_o$ is not available. But when $f$ has some specific form (the characteristic function), it is explicit which is given below. The limiting random variable $\tilde{X}_o$ is $\alpha$-stable with distribution $(\frac{1}{\alpha\theta})^{1/\alpha}S_\alpha(\sigma,\beta,0)$ $=S_\alpha(\sigma(\frac{1}{\alpha\theta})^{1/\alpha},\beta,0)$ (via time change technique and self-similarity). So the characteristic function of $\tilde{X}_o$ in this one-dimensional case is given by

$$\phi(u)=\mathbb{E}[\exp(iu\tilde{X}_o)]=\begin{cases}\exp\left\{-\frac{\sigma^\alpha}{\alpha\theta}|u|^\alpha\left(1-i\beta\operatorname{sgn}(u)\tan\frac{\alpha\pi}{2}\right)\right\}, & \text{if } \alpha\neq 1,\\ \exp\left\{-\frac{\sigma}{\theta}|u|\left(1+i\beta\frac{2}{\pi}\operatorname{sgn}(u)\log|u|\right)\right\}, & \text{if } \alpha=1\,.\end{cases} \tag{1.3.3}$$

**Remark 1.8.** Since the probability distribution is uniquely determined by its characteristic function we see from the above expression (1.3.3) that the probability distribution function of $\tilde{X}_0$ is a function of $\frac{\sigma^\alpha}{\alpha\theta}$. We cannot separate $\alpha$, $\sigma$, and $\theta$. This further implies that for any measurable function $f$ the expectation $\mathbb{E}|f(\tilde{X}_o)|$ is also a function of $\frac{\sigma^\alpha}{\alpha\theta}$ when it is finite.

The ergodic theorem (1.3.2) can then be written as

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\exp(iuX_{t_j})=\phi(u),\quad u\in\mathbb{R},\quad a.s. \tag{1.3.4}$$

This identity will be used to construct statistical estimators of the parameters appeared in (1.2.1).

One may think to use the ergodic theorem (1.3.2) to estimate all the parameters: There are reasons to support this thought; one may choose $f$ differently to obtain sufficient large number of different equations, which may be used to obtain all the unknown parameters. However, this is impossible in our current situation since in the stationary distribution, as we can see from its characteristic function (1.3.3), one can only estimate $\frac{\sigma^\alpha}{\alpha\theta}$ as a whole. For example, one can not separate $\sigma$ and $\theta$ in the characteristic function $\phi(u)$ of $\tilde{X}_o$. This forces us to seek other possibilities. To this end we shall use the ergodic theorem for $X_{t_k} - X_{t_{k-1}}$. More precisely, from Theorem 1.1 of Billingsley [34], it follows

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \exp[iu(X_{t_k} - X_{t_{k-1}})] = \mathbb{E}[e^{iu(\tilde{X}_h - \tilde{X}_0)}] \quad \text{almost surely}, \tag{1.3.5}$$

for arbitrarily fixed $u \in \mathbb{R}$, where $\tilde{X}_t$ satisfies the Langevin equation (1.2.1) with $\tilde{X}_0 = \tilde{X}_o$. To make this formula (1.3.5) useful for the statistical estimation of the parameters, we need to find the explicit form of the characteristic function of $\tilde{X}_h - \tilde{X}_0$. From (1.2.1), we have

$$\tilde{X}_h = e^{-\theta h}\tilde{X}_0 + \sigma e^{-\theta h} \int_0^h e^{\theta s} dZ_s$$

and

$$\tilde{X}_h - \tilde{X}_0 = (e^{-\theta h} - 1)\tilde{X}_0 + \sigma e^{-\theta h} \int_0^h e^{\theta s} dZ_s\,.$$

Note that $\tilde{X}_0 = \tilde{X}_o$ and $\sigma e^{-\theta h} \int_0^h e^{\theta s} dZ_s \sim \sigma \left(\frac{1-e^{-\alpha\theta h}}{\alpha\theta}\right)^{1/\alpha} Z_1$. Note also that $\tilde{X}_0$ and $\sigma e^{-\theta h} \int_0^h e^{\theta s} dZ_s$ are independent. Therefore, we find

$$\begin{aligned} \psi(u) &:= \mathbb{E}[\exp\{iu(\tilde{X}_h - \tilde{X}_0)\}] \\ &= \mathbb{E}[\exp\{iu(e^{-\theta h} - 1)\tilde{X}_0)\}]\mathbb{E}[\exp\{iu\sigma e^{-\theta h} \int_0^h e^{\theta s} dZ_s\}] \\ &= \begin{cases} \exp\Big\{-\frac{\sigma^\alpha |u|^\alpha}{\alpha\theta}\left[(1-e^{-\theta h})^\alpha \left(1 + i\beta \operatorname{sgn}(u) \tan\frac{\alpha\pi}{2}\right)\right. \\ \qquad\qquad \left.+(1-e^{-\alpha\theta h})\left(1 - i\beta \operatorname{sgn}(u)\tan\frac{\alpha\pi}{2}\right)\right]\Big\}, & \text{if } \alpha \neq 1\,; \\ \exp\left\{-\frac{2\sigma(1-e^{-\theta h})}{\theta}|u|\right\}, & \text{if } \alpha = 1\,. \end{cases} \end{aligned} \tag{1.3.6}$$

## 1.4 Moment estimation of all parameters

In this section, we assume that all the parameters $\theta$, $\sigma$, $\alpha$ and $\beta$ involved in the $\alpha$-stable Ornstein-Uhlenbeck motion $(X_t, t \geq 0)$ are unknown and we want to estimate them based on the discrete time observations $\{X_{t_1}, \cdots, X_{t_n}\}$, where as in the last section $t_k = kh$ for some fixed time step $h$. As we explained in Remark 1.8 or paragraphs after that remark, we cannot use (1.3.4) alone to estimate all the parameters in the $\alpha$-stable Ornstein-Uhlenbeck motion $X_t$ given by (1.2.1). As indicated in Section 2, we shall use e.2.5a which motivates us to set $\frac{1}{n}\sum_{j=1}^{n} e^{iu(X_{t_j}-X_{t_{j-1}})} = \psi(u)$. We define the empirical characteristic functions $\hat{\phi}_n(u)$ and $\hat{\psi}_n(v)$ as follows:

$$\hat{\phi}_n(u) := \frac{1}{n}\sum_{j=1}^{n} \exp(iuX_{t_j}), \quad \hat{\psi}_n(v) := \frac{1}{n}\sum_{j=1}^{n} \exp\left[iv(X_{t_j} - X_{t_{j-1}})\right].$$

Motivated by (1.3.4) and (1.3.5), we can estimate all the parameters by matching the empirical characteristic functions $\hat{\phi}_n(u)$ and $\hat{\psi}_n(v)$ with the corresponding theoretical characteristic functions $\phi(u)$ and $\psi(v)$, respectively as given as follows

$$\hat{\phi}_n(u) = \phi(u); \tag{1.4.1}$$

$$\hat{\psi}_n(v) = \psi(v), \tag{1.4.2}$$

where $u, v$ are two constants to be appropriately chosen so that the parametric estimators for all parameters can be constructed.

### 1.4.1 Methodology of parameter estimation

Now we provide the details to obtain the estimators for the parameters in the order of $\alpha$, $\theta$, $\sigma$, and $\beta$. We shall first find the moment estimator for $\alpha$.

#### 1.4.1.1 Estimator for $\alpha$

Choose any arbitrarily two non-zero values $u_1$ and $u_2$ such that $u_1 \neq u_2$. Then, we have

$$\log(-\log|\phi(u_1)|^2) = \log\left(\frac{2\sigma^\alpha}{\alpha\theta}\right) + \alpha \log|u_1|, \tag{1.4.3}$$
$$\log(-\log|\phi(u_2)|^2) = \log\left(\frac{2\sigma^\alpha}{\alpha\theta}\right) + \alpha \log|u_2|, \tag{1.4.4}$$

where $\phi(u)$ is defined in e.2.1. Subtracting the equation e.3.6 from e.3.5, and replacing $\phi(u)$ with its estimated value $\hat{\phi}_n(u)$ as indicated in (1.4.1), we find an estimator of $\alpha$ as follows

$$\hat{\alpha}_n = \frac{\log\left(\frac{\log|\hat{\phi}_n(u_2)|}{\log|\hat{\phi}_n(u_1)|}\right)}{\log\frac{|u_2|}{|u_1|}}. \tag{1.4.5}$$

Since for any fixed $u \in \mathbb{R}$, $\hat{\phi}_n(u)$ converges to $\phi(u)$ almost surely, we see that $\hat{\alpha}_n$ converges to $\alpha$ almost surely.

#### 1.4.1.2 Estimator for $\theta$ given $\alpha$

To construct an estimator for $\theta$ (which depends on the estimation of $\alpha$), we need to use the characteristic function $\psi(u)$ of $\tilde{X}_{t_1} - \tilde{X}_0$. It is easy to verify from the expressions (1.3.3) and (1.3.6) of $\phi(u)$ and $\psi(u)$

$$\frac{\log|\psi(u)|^2}{\log|\phi(u)|^2} = (1-e^{-\theta h})^\alpha + 1 - e^{-\alpha\theta h}. \tag{1.4.6}$$

For any arbitrarily $u$, denote

$$\delta = \frac{\log|\psi(u)|^2}{\log|\phi(u)|^2} \tag{1.4.7}$$

and rewrite equation (1.4.6) as

$$(1-e^{-\theta h})^{\alpha}+1-e^{-\alpha\theta h}=\delta\,. \tag{1.4.8}$$

This is a nonlinear algebraic equation of $\theta$, when $\alpha$ and $\delta$ are considered as given. To simplify notation, we denote $\lambda=e^{-\theta h}$ and then $\theta$ is related to $\lambda$ via

$$\theta=-\log\lambda/h\,.$$

With this substitution, the equation (1.4.8) can be written as an equation for $\lambda$:

$$(1-\lambda)^{\alpha}+1-\lambda^{\alpha}=\delta. \tag{1.4.9}$$

Let $\zeta_\lambda(\alpha,\delta)$ denote the solution of the above equation. Then we can construct an estimator for $\theta$ by

$$\hat{\theta}_n=-\log\left(\hat{\lambda}_n\right)/h\,,\quad \text{where}\quad \hat{\lambda}_n=\zeta_\lambda(\hat{\alpha}_n,\hat{\delta}_n)\,. \tag{1.4.10}$$

Here $\hat{\alpha}_n$ is the estimator for $\alpha$ defined by (1.4.5) and

$$\hat{\delta}_n=\frac{\log|\hat{\psi}_n(u_3)|^2}{\log|\hat{\phi}_n(u_3)|^2} \tag{1.4.11}$$

with $\hat{\phi}_n(u_3)$ and $\hat{\psi}_n(u_3)$ being defined by (1.4.1) and (1.4.2) when $u=u_3\neq u_2\neq u_1$. Since $\hat{\alpha}_n\to\alpha$ a.s. and $\hat{\delta}_n\to\delta$ a.s., we have $\hat{\lambda}_n\to\lambda$ a.s. and $\hat{\theta}_n\to\theta$ a.s.

Our estimator $\hat{\theta}_n$ depends on the function $\zeta_\lambda(\alpha,\delta)$, which is the solution to (1.4.9). This is a simple algebraic equation. There are many methods to solve general algebraic equation numerically. Here we shall use the Newton's method. Denote

$$g(\lambda)=g(\lambda,\hat{\alpha}_n,\hat{\delta}_n)=(1-\lambda)^{\hat{\alpha}_n}+1-\lambda^{\hat{\alpha}_n}-\hat{\delta}_n\,.$$

For any fixed value of $\theta$, we can always choose $h$ fixed but small enough (e.g. $0 < h < \ln 2/\theta$ ) such that $\lambda = e^{-\theta h} \in (\frac{1}{2}, 1)$ and $0 < \hat{\delta}_n < 1$. Note that $g$ is decreasing with derivative $g^{'}(\lambda) = -\hat{\alpha}_n[\lambda^{\hat{\alpha}_n - 1} + (1-\lambda)^{\hat{\alpha}_n - 1}] < 0$ for $\lambda \in (\frac{1}{2}, 1)$, $g(\frac{1}{2}) = 1 - \hat{\delta}_n > 0$ and $g(1) = -\hat{\delta}_n < 0$. Hence there is a unique root for $g(\lambda)$ in the interval $(\frac{1}{2}, 1)$. Namely, there exists a unique $\hat{\lambda}_n \in (\frac{1}{2}, 1)$ such that $g(\hat{\lambda}_n) = 0$. Then, the Newton's method to approximate $\hat{\lambda}_n$ is as follows. First, we define $\lambda_{n,0} = \frac{1}{2}$. Then, we define

$$\lambda_{n,m+1} = \lambda_{n,m} - \frac{g(\lambda_{n,m})}{g^{'}(\lambda_{n,m})}, \quad m = 0, 1, 2, \cdots \tag{1.4.12}$$

Note that $g^{''}(\lambda) = \hat{\alpha}_n(\hat{\alpha}_n - 1)\frac{\lambda^{2-\hat{\alpha}_n} - (1-\lambda)^{2-\hat{\alpha}_n}}{\lambda^{2-\hat{\alpha}_n}(1-\lambda)^{2-\hat{\alpha}_n}} > 0$ if $1 < \hat{\alpha}_n < 2$ and $\lambda \in (1/2, 1)$. In this case, we have global convergence of the Newton's iterations $\{\lambda_{n,m}\}_{m=1}^{\infty}$. In fact, let the approximation error at the (m+1)-th interation be $\varepsilon_{n,m+1} = \lambda_{n,m+1} - \hat{\lambda}_n$. By e.3.13, we have

$$\varepsilon_{n,m+1} = \varepsilon_{n,m} - \frac{g(\lambda_{n,m})}{g^{'}(\lambda_{n,m})}. \tag{1.4.13}$$

Then by Taylor expansion we find that $\varepsilon_{n,m+1} = \frac{1}{2}\frac{g^{''}(\xi_{n,m})}{g^{'}(\lambda_{n,m})}\varepsilon_{n,m}^2 < 0$, where $\xi_{n,m}$ is between $\lambda_{n,m}$ and $\hat{\lambda}_n$. This implies that $\lambda_{n,m} < \hat{\lambda}_n$ for each $m \geq 1$. Since $g$ is decreasing, we have $g(\lambda_{n,m}) > g(\hat{\lambda}_n) = 0$. Thus $\varepsilon_{n,m+1} > \varepsilon_{n,m}$ and $\lambda_{n,m+1} > \lambda_{n,m}$ for each $m \geq 1$. Hence, the two sequences $\{\varepsilon_{n,m}\}_{m=1}^{\infty}$ and $\{\lambda_{n,m}\}_{m=1}^{\infty}$ are increasing and bounded from above. Thus there exist $\varepsilon_n^*$ and $\lambda_n^*$ such that

$$\lim_{m\to\infty} \varepsilon_{n,m} = \varepsilon_n^*, \quad \lim_{m\to\infty} \lambda_{n,m} = \lambda_n^*.$$

Thus, by e.3.14, it follows that

$$\varepsilon_n^* = \varepsilon_n^* - \frac{g(\lambda_n^*)}{g^{'}(\lambda_n^*)}. \tag{1.4.14}$$

This implies that $g(\lambda_n^*) = 0$ and consequently $\lambda_n^* = \hat{\lambda}_n$.

Now, when $0 < \hat{\alpha}_n < 1$, we can use similar arguments to show that Newton's method converges to the unique root $\hat{\lambda}_n$ of $g(\lambda)$ from any starting point (namely we have global convergence of the Newton's method).

#### 1.4.1.3 Estimator for $\sigma$ given $\alpha$ and $\theta$

Next we turn to the estimation of $\sigma$. Let $\tau = \frac{2\sigma^\alpha}{\alpha\theta}$ and $\sigma$ is related to $\tau$ by

$$\begin{aligned} \sigma &= \exp\left\{\frac{\log\tau + \log\alpha + \log\theta - \log 2}{\alpha}\right\} \quad \text{or} \\ \log\sigma &= \frac{\log\tau + \log\alpha + \log\theta - \log 2}{\alpha}. \end{aligned} \tag{1.4.15}$$

Thus, the estimation of $\sigma$ is reduced to the estimation of $\tau$ since we already have estimators for $\alpha$ and $\theta$.

To obtain an estimator for $\sigma$ (or for $\log\sigma$), we may use any one of the equations e.3.5 and e.3.6. However, we shall use both of these two equations in the following way, which will eliminate the explicit dependence on $\alpha$. Multiply equations e.3.5 by $\log|u_2|$ and multiply equations e.3.6 by $\log|u_1|$. Taking the difference yields

$$\log\tau = \frac{\log(|u_1|)\log\left(-\log|\phi(u_2)|^2\right) - \log(|u_2|)\log\left(-\log|\phi(u_1)|^2\right)}{\log\frac{|u_1|}{|u_2|}}. \tag{1.4.16}$$

From this identity, we construct the estimator for $\tau$ as follows

$$\log\hat{\tau}_n = \frac{\log(|u_1|)\log\left(-\log|\hat{\phi}_n(u_2)|^2\right) - \log(|u_2|)\log\left(-\log|\hat{\phi}_n(u_1)|^2\right)}{\log\frac{|u_1|}{|u_2|}}, \tag{1.4.17}$$

where $\hat{\phi}_n(u)$ is given by (1.4.1) . Thus, we can construct the estimator for $\sigma$ by

$$\hat{\sigma}_n = \exp\left\{\frac{\log\hat{\tau}_n + \log\hat{\alpha}_n + \log\hat{\theta}_n - \log 2}{\hat{\alpha}_n}\right\}. \tag{1.4.18}$$

Based on the almost sure convergence of $\hat{\phi}_n(u)$ to $\phi(u)$, we see easily that $\hat{\sigma}_n \to \sigma$ almost surely.

#### 1.4.1.4 Estimator for $\beta$ given $\alpha$, $\theta$, and $\sigma$

Finally, we discuss the estimation of the skewness parameter $\beta \in [-1,1]$. Note from (1.3.3) that for $\alpha \neq 1$, we have

$$\arctan\left(\frac{\Im(\phi(u))}{\Re(\phi(u))}\right) = \beta \frac{\sigma^\alpha}{\alpha\theta} \tan\left(\frac{\alpha\pi}{2}\right) |u|^\alpha \operatorname{sgn}(u)\,, \tag{1.4.19}$$

where $\Im(\phi(u))$ and $\Re(\phi(u))$ are the imaginary and real parts of the complex valued function $\phi(u)$, respectively. In order to make sure that the right hand side is in the range of arctan, choose $u = u_4$ in e.3.26 such that $-\frac{\pi}{2} < \frac{\sigma^\alpha}{\alpha\theta} \tan\left(\frac{\alpha\pi}{2}\right) |u|^\alpha \operatorname{sgn}(u) < \frac{\pi}{2}$ . Replacing $\phi(u_4)$, $\alpha$, $\theta$, and $\sigma$ by $\hat{\phi}_n(u_4)$,$\hat{\alpha}_n$, $\hat{\theta}_n$, and $\hat{\sigma}_n$, we can construct an estimator of $\beta$ as follows

$$\hat{\beta}_n = \frac{\hat{\alpha}_n \hat{\theta}_n \arctan[(\sum_{j=1}^n \sin u_4 X_{t_j})/(\sum_{j=1}^n \cos u_4 X_{t_j})]}{\hat{\sigma}_n^{\hat{\alpha}_n} \tan(\hat{\alpha}_n \pi/2) |u_4|^{\hat{\alpha}_n} \operatorname{sgn}(u_4)}\,. \tag{1.4.20}$$

When $\alpha = 1$, we have

$$\hat{\beta}_n = -\frac{\hat{\theta}_n \arctan[(\sum_{j=1}^n \sin u_4 X_{t_j})/(\sum_{j=1}^n \cos u_4 X_{t_j})]}{\hat{\sigma}_n \frac{2}{\pi} \log |u_4| \operatorname{sgn}(u_4)}\,. \tag{1.4.21}$$

By the almost sure convergence of $\hat{\alpha}_n$, $\hat{\theta}_n$, $\hat{\sigma}_n$ and $\hat{\phi}_n(u_4)$, we can easily get the almost sure convergence of $\hat{\beta}_n$ to $\beta$.

### 1.4.2 Joint asymptotic behavior of all the obtained estimators

In this subsection, we are going to study the joint behavior of the estimators of all the parameters $\alpha, \theta, \sigma$, and $\beta$. We let $\eta = (\alpha, \theta, \sigma, \beta)^T$ and $\hat{\eta}_n = (\hat{\alpha}_n, \hat{\theta}_n, \hat{\sigma}_n, \hat{\beta}_n)^T$. Our main task is to compute the asymptotic covariance of the estimators of all the parameters $\alpha, \theta, \sigma$, and $\beta$. We want to compute the covariance matrix of the limiting distribution of $\sqrt{n}(\hat{\eta}_n - \eta)$. Due to the difficulty that the $\alpha$-stable Ornstein-Uhlenbeck motion has no second moment, we shall discuss how to find the asymptotic covariance matrix of $\sqrt{n}(\hat{\eta}_n - \eta)$ in detail.

For any nice function $f$ denote

$$S_n(f) = \frac{1}{n}\sum_{j=1}^{n} f(X_{t_j}) \quad \text{and} \quad T_n(f) = \frac{1}{n}\sum_{j=1}^{n} f(X_{t_j} - X_{t_{j-1}})\,. \tag{1.4.22}$$

Let $F_u(x) = \cos(ux)$ and $G_u(x) = \sin(ux)$. Then $\hat{\phi}_n(u) = S_n(F_u) + iS_n(G_u)$ and $|\hat{\phi}_n(u)|^2 = S_n^2(F_u) + S_n^2(G_u)$. Let $V_{n1} = S_n(F_{u_1})$, $V_{n2} = S_n(G_{u_1})$, $V_{n3} = S_n(F_{u_2})$, $V_{n4} = S_n(G_{u_2})$, $V_{n5} = S_n(F_{u_3})$, $V_{n6} = S_n(G_{u_3})$, $V_{n7} = T_n(F_{u_3})$, $V_{n8} = T_n(G_{u_3})$, $V_{n9} = S_n(F_{u_4})$, $V_{n10} = S_n(G_{u_4})$. We need first to compute the asymptotic covariance matrix associated with

$$V_n = (V_{n1}, V_{n2}, V_{n3}, V_{n4}, V_{n5}, V_{n6}, V_{n7}, V_{n8}, V_{n9}, V_{n10})^T\,.$$

Then we shall use this computation to find the asymptotic covariance matrix of $\hat{\eta}_n$.

To compute the asymptotic covariance matrix associated with $V_n$ we consider the functional $S_n(f)$ and $T_n(f)$ as a special case of

$$R_n(f) = \frac{1}{n}\sum_{j=1}^{n} f(X_{t_{j-1}}, X_{t_j})\,,$$

where $f(x,y)$ is a function of two variables. It is well-known that for two functions $f(x,y)$ and

$g(x,y)$, the asymptotic covariance cov $(R_n(f), R_n(g))$ of $R_n(f)$ and $R_n(g)$ is given by

$$\sigma_{fg} := \lim_{n\to\infty} \text{cov } (R_n(f), R_n(g)) = \text{cov } (f(\tilde{X}_0, \tilde{X}_h), g(\tilde{X}_0, \tilde{X}_h)) \\ + \sum_{j=1}^{\infty} [\text{cov } (f(\tilde{X}_0, \tilde{X}_h), g(\tilde{X}_{jh}, \tilde{X}_{(j+1)h})) + \text{cov } (g(\tilde{X}_0, \tilde{X}_h), f(\tilde{X}_{jh}, \tilde{X}_{(j+1)h}))].$$

The asymptotic covariance matrix of $V_n$ will then be given by the covariance matrix

$$\Sigma_{10} := \lim_{n\to\infty} (\text{cov } (V_{nk}, V_{nl}))_{1\le k,l\le 10} = (\sigma_{g_k g_l})_{1\le k,l\le 10}\,, \tag{1.4.23}$$

where

$$\begin{cases} g_1(x,y) = F_{u_1}(x)\,,\ g_2(x,y) = G_{u_1}(x)\,,\ g_3(x,y) = F_{u_2}(x)\,, \\ g_4(x,y) = G_{u_2}(x)\,,\ g_5(x,y) = F_{u_3}(x),\ g_6(x,y) = G_{u_3}(x)\,, \\ g_7(x,y) = F_{u_3}(y-x)\,,\ g_8(x,y) = G_{u_3}(y-x)\,, \\ g_9(x,y) = F_{u_4}(x),\ \ g_{10}(x,y) = G_{u_4}(x)\,. \end{cases}$$

Let $v = (v_1, v_2, \ldots, v_{10})^T$, where $v_j = \mathbb{E}[g_j(\tilde{X}_0, \tilde{X}_h)], j = 1,2,\ldots,10$ . The explicit expressions of the elements in the covariance matrix $\Sigma_{10}$ will be provided in the Appendix. For $z = (z_1, \ldots, z_{10})^T$, we define the following functions

$$\begin{cases} \hat{\gamma}_1(z) = \log\left(-\log(z_1^2 + z_2^2)\right), \quad \hat{\gamma}_2(z) = \log\left(-\log(z_3^2 + z_4^2)\right), \\ \hat{\gamma}_3(z) = \frac{\log(z_7^2+z_8^2)}{\log(z_5^2+z_6^2)}, \quad \hat{\gamma}_4(z) = \arctan\left(\frac{z_{10}}{z_9}\right). \end{cases}$$

Then, basic calculation shows that

$$\begin{cases} \gamma_1(\eta) := \hat{\gamma}_1(v) = \log\left(\frac{2\sigma^\alpha}{\alpha\theta}\right) + \alpha \log|u_1|\,, \\ \gamma_2(\eta) := \hat{\gamma}_2(v) = \log\left(\frac{2\sigma^\alpha}{\alpha\theta}\right) + \alpha \log|u_2|\,, \\ \gamma_3(\eta) := \hat{\gamma}_3(v) = (1 - e^{-\theta h})^\alpha + 1 - e^{-\alpha\theta h}\,, \\ \gamma_4(\eta) := \hat{\gamma}_4(v) = \beta \frac{\sigma^\alpha}{\alpha\theta} \tan\left(\frac{\alpha\pi}{2}\right) |u_4|^\alpha \text{sgn}(u_4). \end{cases}$$

Let $\hat{\gamma}(z) = (\hat{\gamma}_1(z), \hat{\gamma}_2(z), \hat{\gamma}_3(z), \hat{\gamma}_4(z))^T$ for $z \in \mathbb{R}^{10}$, $\quad \hat{\gamma}^{(1)}(z) = \left(\frac{\partial \hat{\gamma}_j}{\partial z_k}\right)_{1\le j\le 4, 1\le k\le 10}$, $\quad$ and $\gamma(\eta) =$

$(\gamma_1(\eta), \gamma_2(\eta), \gamma_3(\eta), \gamma_4(\eta))^T$. We have

$$
\begin{aligned}
&\frac{\partial \gamma_1}{\partial z_1} = \frac{-2z_1}{(z_1^2+z_2^2)\log(z_1^2+z_2^2)}, \quad \frac{\partial \gamma_1}{\partial z_2} = \frac{-2z_2}{(z_1^2+z_2^2)\log(z_1^2+z_2^2)} \\
&\qquad \frac{\partial \gamma_1}{\partial z_3} = \cdots = \frac{\partial \gamma_1}{\partial z_{10}} = 0; \\
&\frac{\partial \gamma_2}{\partial z_3} = \frac{-2z_3}{(z_3^2+z_4^2)\log(z_3^2+z_4^2)}, \quad \frac{\partial \gamma_2}{\partial z_4} = \frac{-2z_4}{(z_3^2+z_4^2)\log(z_3^2+z_4^2)} \\
&\qquad \frac{\partial \gamma_2}{\partial z_1} = 0, \quad \frac{\partial \gamma_2}{\partial z_2} = 0, \quad \frac{\partial \gamma_2}{\partial z_5} = \cdots = \frac{\partial \gamma_2}{\partial z_{10}} = 0; \\
&\frac{\partial \gamma_3}{\partial z_5} = \frac{-2z_5\log(z_7^2+z_8^2)}{(z_5^2+z_6^2)\log^2(z_5^2+z_6^2)}, \quad \frac{\partial \gamma_3}{\partial z_6} = \frac{-2z_6\log(z_7^2+z_8^2)}{(z_5^2+z_6^2)\log^2(z_5^2+z_6^2)} \\
&\qquad \frac{\partial \gamma_3}{\partial z_7} = \frac{2z_7}{(z_7^2+z_8^2)\log(z_5^2+z_6^2)}, \quad \frac{\partial \gamma_3}{\partial z_8} = \frac{2z_8}{(z_7^2+z_8^2)\log(z_5^2+z_6^2)} \\
&\qquad \frac{\partial \gamma_3}{\partial z_1} = \cdots = \frac{\partial \gamma_3}{\partial z_4} = 0, \quad \frac{\partial \gamma_3}{\partial z_9} = \frac{\partial \gamma_3}{\partial z_{10}} = 0 \\
&\frac{\partial \gamma_4}{\partial z_9} = \frac{-z_{10}}{z_9^2+z_{10}^2}, \quad \frac{\partial \gamma_4}{\partial z_{10}} = \frac{z_9}{z_9^2+z_{10}^2}, \quad \frac{\partial \gamma_4}{\partial z_1} = \cdots = \frac{\partial \gamma_4}{\partial z_8} = 0.
\end{aligned}
$$

Let $\Phi_n(\eta) = (\Phi_{1,n}(\eta), \Phi_{2,n}(\eta), \Phi_{3,n}(\eta), \Phi_{4,n}(\eta))^T$, where $\Phi_{j,n}(\eta) = \hat{\gamma}_j(V_n) - \gamma_j(\eta),\ j = 1,2,3,4.$ Then, we know that $\hat{\eta}_n$ is the generalized moment estimator of $\eta$, which satisfies

$$
\Phi_n(\hat{\eta}_n) = 0. \tag{1.4.24}
$$

Basic calculation gives

$$
\begin{aligned}
&\frac{\partial\gamma_1}{\partial\alpha}=\log\sigma-\frac{1}{\alpha}+\log|u_1|,\quad \frac{\partial\gamma_1}{\partial\theta}=-\frac{1}{\theta},\quad \frac{\partial\gamma_1}{\partial\sigma}=\frac{\alpha}{\sigma},\quad \frac{\partial\gamma_1}{\partial\beta}=0;\\
&\frac{\partial\gamma_2}{\partial\alpha}=\log\sigma-\frac{1}{\alpha}+\log|u_2|,\quad \frac{\partial\gamma_2}{\partial\theta}=-\frac{1}{\theta},\quad \frac{\partial\gamma_2}{\partial\sigma}=\frac{\alpha}{\sigma},\quad \frac{\partial\gamma_2}{\partial\beta}=0;\\
&\frac{\partial\gamma_3}{\partial\alpha}=(1-e^{-\theta h})^{\alpha}\log(1-e^{-\theta h})+\theta h e^{-\alpha\theta h},\\
&\frac{\partial\gamma_3}{\partial\theta}=\alpha h e^{-\theta h}(1-e^{-\theta h})^{\alpha-1}+\alpha h e^{-\alpha\theta h},\\
&\frac{\partial\gamma_3}{\partial\sigma}=0,\quad \frac{\partial\gamma_3}{\partial\beta}=0;\\
&\frac{\partial\gamma_4}{\partial\alpha}=\frac{\beta\sigma^{\alpha}|u_4|^{\alpha}\mathrm{sgn}(u_4)}{\alpha\theta}\left[\log(\sigma|u_4|)\tan\left(\frac{\alpha\pi}{2}\right)-\alpha^{-1}\tan\left(\frac{\alpha\pi}{2}\right)+\frac{\pi}{2}\sec^2\left(\frac{\alpha\pi}{2}\right)\right],\\
&\frac{\partial\gamma_4}{\partial\theta}=-\beta\frac{\sigma^{\alpha}}{\alpha\theta^2}\tan\left(\frac{\alpha\pi}{2}\right)|u_4|^{\alpha}\mathrm{sgn}(u_4),\\
&\frac{\partial\gamma_4}{\partial\sigma}=\beta\frac{\sigma^{\alpha-1}}{\theta}\tan\left(\frac{\alpha\pi}{2}\right)|u_4|^{\alpha}\mathrm{sgn}(u_4),\\
&\frac{\partial\gamma_4}{\partial\beta}=\frac{\sigma^{\alpha}}{\theta}\tan\left(\frac{\alpha\pi}{2}\right)|u_4|^{\alpha}\mathrm{sgn}(u_4).
\end{aligned}
$$

Note that

$$\nabla_\eta\Phi_n(\eta)=-\nabla_\eta\gamma(\eta), \tag{1.4.25}$$

where

$$\nabla_\eta\gamma(\eta)=\begin{pmatrix}
\frac{\partial\gamma_1(\eta)}{\partial\alpha} & \frac{\partial\gamma_1(\eta)}{\partial\theta} & \frac{\partial\gamma_1(\eta)}{\partial\sigma} & \frac{\partial\gamma_1(\eta)}{\partial\beta}\\
\frac{\partial\gamma_2(\eta)}{\partial\alpha} & \frac{\partial\gamma_2(\eta)}{\partial\theta} & \frac{\partial\gamma_2(\eta)}{\partial\sigma} & \frac{\partial\gamma_2(\eta)}{\partial\beta}\\
\frac{\partial\gamma_3(\eta)}{\partial\alpha} & \frac{\partial\gamma_3(\eta)}{\partial\theta} & \frac{\partial\gamma_3(\eta)}{\partial\sigma} & \frac{\partial\gamma_3(\eta)}{\partial\beta}\\
\frac{\partial\gamma_4(\eta)}{\partial\alpha} & \frac{\partial\gamma_4(\eta)}{\partial\theta} & \frac{\partial\gamma_4(\eta)}{\partial\sigma} & \frac{\partial\gamma_4(\eta)}{\partial\beta}
\end{pmatrix}. \tag{1.4.26}$$

For convenience, let $I(\eta)=\nabla_\eta\gamma(\eta)$.

Finally we have the following main result.

**Theorem 1.9.** *Fix an arbitrary $h>0$. Denote $\eta=(\alpha,\theta,\sigma,\beta)^T$ and $\hat{\eta}_n=(\hat{\alpha}_n,\hat{\theta}_n,\hat{\sigma}_n,\hat{\beta}_n)^T$, where $\hat{\alpha}_n,\hat{\theta}_n,\hat{\sigma}_n,\hat{\beta}_n$ are given by* (1.4.5), (1.4.10), (1.4.18), (1.4.20) *and* (1.4.21)*, respectively. Then we* *have the following statements. (i) The ergodic estimators $\hat{\eta}_n$ converges to $\eta$ almost surely as $n \to \infty$. (ii) As $n \to \infty$ we have the following central limit type theorem:*

$$\sqrt{n}(\hat{\eta}_n - \eta) \xrightarrow{d} N(0, \Sigma_4), \tag{1.4.27}$$

*where*

$$\Sigma_4 = (I(\eta))^{-1} \hat{\gamma}^{(1)}(v) \Sigma_{10} (\hat{\gamma}^{(1)}(v))^T ((I(\eta))^{-1})^T.$$

Now we provide all the necessary lemmas with their proofs and the proof of Theorem 3.1.

Let $U = (U_1, U_2, \ldots, U_{10})^T \sim N(0, \Sigma_{10})$. Then, we have the following result:

**Lemma 1.10.** We have the CLT

$$\sqrt{n}(V_n - v) \xrightarrow{d} U. \tag{1.4.28}$$

*Proof.* Let $U = (U_1, U_2, \ldots, U_{10})^T$ be a normally distributed random vector with mean 0 and covariance matrix $\Sigma_{10}$. Then for any non-zero vector $a = (a_1, a_2, \ldots, a_{10})^T \in \mathbb{R}^{10}$, we have $a^T U \sim N(0, a^T \Sigma_{10} a)$. By the Cramer-Wold device (Theorem 29.4 of Billingsley 1995), it suffices to prove that

$$a^T \sqrt{n}(V_n - v) \xrightarrow{d} a^T U.$$

Define $K = a^T (g_1, g_2, \ldots, g_{10})^T$ and $\bar{K} = K - \mathbb{E}[K(\tilde{X}_0, \tilde{X}_h)] = a^T (\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_{10})^T$. Note that the underlying Ornstein-Uhlenbeck process is stationary and exponentially $\alpha$-mixing (see Theorem 2.6 of Masuda 2007). Then by the univariate CLT Theorem 18.6.3 of Ibragimov and Linnik (1971) for stationary process with $\alpha$-mixing condition, we have

$$a^T \sqrt{n}(V_n - v) = \sqrt{n} T_n(\bar{K}) \xrightarrow{d} N(0, \sigma_K^2), \tag{1.4.29}$$

where

$$\sigma_K^2 = \mathbb{E}_\mu[\bar{K}^2(\tilde{X}_0, \tilde{X}_h)] + 2 \sum_{j=1}^{\infty} \mathbb{E}_\mu[\bar{K}(\tilde{X}_0, \tilde{X}_h) \bar{K}(\tilde{X}_{jh}, \tilde{X}_{(j+1)h})] = a^T \Sigma_{10} a.$$

Therefore, we have $a^T\sqrt{n}(V_n - v) \xrightarrow{d} a^T U$ for any non-zero $a \in \mathbb{R}^{10}$. The proof is complete. □ □

**Lemma 1.11.** We have the following CLT

$$\sqrt{n}\Phi_n(\eta) \xrightarrow{d} \hat{\gamma}^{(1)}(v)U. \tag{1.4.30}$$

*Proof.* Note that $\sqrt{n}\Phi_n(\eta) = \sqrt{n}(\hat{\gamma}(V_n) - \hat{\gamma}(v))$. The result follows directly from Lemma 1.10 and the delta method (see, e.g., Lemma 5.3.3 of Bickel and Doksum 2001). □ □

Now we are ready to prove our main result Theorem 3.1.

**Proof of Theorem 1.9**. (i) It is obvious since each component of $\hat{\eta}_n$ converges to the corresponding component of $\eta$ almost surely as $n \to \infty$ as discussed in subsections 3.1.1-3.1.4.
(ii) By Taylor's formula, we have

$$\Phi_n(\hat{\eta}_n) - \Phi_n(\eta) = \int_0^1 \nabla_\eta \Phi_n(\eta + s(\hat{\eta}_n - \eta))ds \cdot (\hat{\eta}_n - \eta). \tag{1.4.31}$$

Let $I_n(\eta) = -\int_0^1 \nabla_\eta \Phi_n(\eta + s(\hat{\eta}_n - \eta))ds$ be invertible. Note that $\Phi_n(\hat{\eta}_n) = 0$. Then, we have

$$\sqrt{n}(\hat{\eta}_n - \eta) = (I_n(\eta))^{-1} \cdot \sqrt{n}\Phi_n(\eta). \tag{1.4.32}$$

Note that $(I_n(\eta))^{-1} \to (I(\eta))^{-1}$ a.s. since $\hat{\eta}_n \to \eta$ a.s. Therefore by using Lemma 1.11 and Slutsky's Theorem, we have

$$\sqrt{n}(\hat{\eta}_n - \eta) \xrightarrow{d} (I(\eta))^{-1}\hat{\gamma}^{(1)}(v)U.$$

The proof is complete. □

### 1.4.3 Optimal selection of the four grid points $\{u_1, u_2, u_3, u_4\}$

Following some ideas in Zhang and He [48], we shall discuss how to select the four grid points $\{u_1, u_2, u_3, u_4\}$ in certain optimal way. We first choose a relatively extensive grid set consisting of $K$ grid points defined by

$$\Delta_K = \left\{ \frac{ka}{K}, k = 1, 2, \cdots, K \right\},$$

where $a$ is a fixed positive number, and $K$ is a relatively large positive integer. For example, we can set $a = 5$ (or 8, 10 etc) and $K = 200$ (or 400, 500 etc). For a finite set $A$, we use $\min - \arg\min_{x \in A} f(x)$ to denote the minimal value of $x \in A$ that minimizes $f(x)$. Note that the values that minimize $f(x)$ are not always unique. We will use the following two steps to select four grid points $\{u_1, u_2, u_3, u_4\}$ optimally.

**Step 1**. We choose

$$\{u_1^*, u_2^*, u_3^*, u_4^*\} = \{\hat{u}_{1,n}^*, \hat{u}_{2,n}^*, \hat{u}_{3,n}^*, \hat{u}_{4,n}^*\} \subset \Delta_K$$

arbitrarily in an increasing order, i.e. $u_1^* < u_2^* < u_3^* < u_4^*$. Then we compute $\hat{\eta}_n = (\hat{\alpha}_n, \hat{\theta}_n, \hat{\sigma}_n, \hat{\beta}_n)$, $\Sigma_{4,n}^* = \Sigma_4(\hat{\eta}_n, \{\hat{u}_{1,n}^*, \hat{u}_{2,n}^*, \hat{u}_{3,n}^*, \hat{u}_{4,n}^*\})$ (which is the matrix $\Sigma_4$ computed by replacing $\eta$ with $\hat{\eta}_n$ in Theorem 1.9) as well as the closeness measure $m(\Sigma_{4,n}^*) = \mathrm{tr}(\Sigma_{4,n}^*)$ (namely the trace of $\Sigma_{4,n}^*$).

**Step 2**. Adjust the location of $\{u_1^*, u_2^*, u_3^*, u_4^*\}$ to $\{u_1^{**}, u_2^{**}, u_3^{**}, u_4^{**}\}$ by

$$\begin{aligned}
u_1^{**} = \hat{u}_{1,n}^{**} &= \min - \arg \min_{u \in \{u \in \Delta_K : u < \hat{u}_{2,n}^*, u \neq \hat{u}_{1,n}^*\}} m(\Sigma_4(\hat{\eta}_n, \{u, \hat{u}_{2,n}^*, \hat{u}_{3,n}^*, \hat{u}_{4,n}^*\})), \\
u_2^{**} = \hat{u}_{2,n}^{**} &= \min - \arg \min_{u \in \{u \in \Delta_K : \hat{u}_{1,n}^{**} < u < \hat{u}_{3,n}^*, u \neq \hat{u}_{2,n}^*\}} m(\Sigma_4(\hat{\eta}_n, \{\hat{u}_{1,n}^{**}, u, \hat{u}_{3,n}^*, \hat{u}_{4,n}^*\})), \\
u_3^{**} = \hat{u}_{3,n}^{**} &= \min - \arg \min_{u \in \{u \in \Delta_K : \hat{u}_{2,n}^{**} < u < \hat{u}_{4,n}^*, u \neq \hat{u}_{3,n}^*\}} m(\Sigma_4(\hat{\eta}_n, \{\hat{u}_{1,n}^{**}, \hat{u}_{2,n}^{**}, u, \hat{u}_{4,n}^*\})), \\
u_4^{**} = \hat{u}_{4,n}^{**} &= \min - \arg \min_{u \in \{u \in \Delta_K : u > \hat{u}_{3,n}^{**}, u \neq \hat{u}_{4,n}^*\}} m(\Sigma_4(\hat{\eta}_n, \{\hat{u}_{1,n}^{**}, \hat{u}_{2,n}^{**}, \hat{u}_{3,n}^{**}, u\})).
\end{aligned}$$

**Step 3**. Compute $m(\Sigma_{4,n}^{**})$, where

$$\Sigma_{4,n}^{**} = \Sigma_4(\hat{\eta}_n, \{u_1^{**}, u_2^{**}, u_3^{**}, u_4^{**}\}).$$

Then compute

$$\hat{\rho}_n = \frac{m(\Sigma_{4,n}^{*}) - m(\Sigma_{4,n}^{**})}{m(\Sigma_{4,n}^{*})}.$$

**Step 4**. If $\hat{\rho}_n > \varepsilon$ (a pre-specified error value like 0.001), then set $\{u_1^{**}, u_2^{**}, u_3^{**}, u_4^{**}\}$ to be $\{u_1^{*}, u_2^{*}, u_3^{*}, u_4^{*}\}$ and repeat steps 2-3; else stop and output

$$\{u_1, u_2, u_3, u_4\} = \{u_1^{**}, u_2^{**}, u_3^{**}, u_4^{**}\}.$$

Thus, we get our optimal selection of four grid points $\{u_1, u_2, u_3, u_4\}$ and the corresponding estimator $\hat{\eta}_n$ in terms of these four points.

The explicit expressions of the elements in the covariance matrix $\Sigma_{10}$ are given in this subsection.

### 1.4.4 Computation of the covariance matrix $\Sigma_{10}$

By using the characteristic function $\phi(u)$ given in (1.3.3), we define

$$\begin{aligned} A_0(u) &= \mathbb{E}(\cos u\tilde{X}_0) \\ &= \exp\left\{-\frac{\sigma^\alpha}{\alpha\theta}|u|^\alpha\right\}\cos\left(\frac{\sigma^\alpha}{\alpha\theta}|u|^\alpha\beta\sin(u)\tan\frac{\alpha\pi}{2}\right). \end{aligned} \tag{1.4.33}$$

$$\begin{aligned} B_0(u) &= \mathbb{E}(\sin u\tilde{X}_0) \\ &= \exp\left\{-\frac{\sigma^\alpha}{\alpha\theta}|u|^\alpha\right\}\sin\left(\frac{\sigma^\alpha}{\alpha\theta}|u|^\alpha\beta\sin(u)\tan\frac{\alpha\pi}{2}\right). \end{aligned} \tag{1.4.34}$$

*Computation of* $\sigma_{g_1g_1}$. From the definition of $g_1$ we have

$$\begin{aligned}
\sigma_{g_1g_1} &= \text{cov }(\cos u_1\tilde{X}_0, \cos u_1\tilde{X}_0), +2\sum_{j=1}^{\infty}[\text{cov }(\cos u_1\tilde{X}_0, \cos u_1\tilde{X}_{jh})] \\
&= \mathbb{E}((\cos u_1\tilde{X}_0)^2) - (\mathbb{E}(\cos u_1\tilde{X}_0))^2 \\
&\quad +2\sum_{j=1}^{\infty}\left\{\mathbb{E}(\cos u_1\tilde{X}_0 \cos u_1\tilde{X}_{jh}) - \mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\cos u_1\tilde{X}_{jh})\right\}. \qquad (1.4.35)
\end{aligned}$$

The first term in (1.4.35) is given by

$$\begin{aligned}
\mathbb{E}((\cos u_1\tilde{X}_0)^2) &= \mathbb{E}\left(\frac{\cos 2u_1\tilde{X}_0 + 1}{2}\right) = \frac{1}{2} + \frac{1}{2}\mathbb{E}(\cos 2u_1\tilde{X}_0) \\
&= \frac{1}{2} + \frac{1}{2}A_0(2u_1). \qquad (1.4.36)
\end{aligned}$$

To compute the second term in (1.4.35) one needs

$$\mathbb{E}(\cos u_1\tilde{X}_0) = A_0(u_1). \qquad (1.4.37)$$

Notice that $\mathbb{E}(\cos u_1\tilde{X}_{jh}) = \mathbb{E}(\cos u_1\tilde{X}_0)$ and then the second summand in the sum of (1.4.35) is also given by the above formula. We write

$$u\tilde{X}_0 + v\tilde{X}_{jh} = (u + ve^{-\theta jh})\tilde{X}_0 + v\sigma e^{-\theta jh}\int_0^{jh} e^{\theta s}dZ_s$$

and then we see

$$\begin{aligned}
&\mathbb{E}[\exp\{iu\tilde{X}_0 + iv\tilde{X}_{jh}\}] \\
&= \mathbb{E}[\exp\{i(u + ve^{-\theta jh})\tilde{X}_0\}]\mathbb{E}[\exp\{iv\sigma e^{\theta jh}\int_0^{jh} e^{\theta s}dZ_s\}] \\
&= \exp\left\{-\frac{\sigma^\alpha}{\alpha\theta}[|u + ve^{-\theta jh}|^\alpha(1 - i\beta\sin(u + ve^{-\theta jh})\tan\frac{\alpha\pi}{2})\right. \\
&\left. +|v|^\alpha(1 - e^{-\alpha\theta jh})(1 - i\beta\sin(v)\tan\frac{\alpha\pi}{2})]\right\}. \qquad (1.4.38)
\end{aligned}$$

Let

$$
\begin{aligned}
&A_j(u,v) \qquad\qquad (1.4.39)\\
&= \mathbb{E}(\cos(u\tilde{X}_0 + v\tilde{X}_{jh}))\\
&= \exp\left\{ -\frac{\sigma^\alpha}{\alpha\theta}[|u+ve^{-\theta jh}|^\alpha + |v|^\alpha(1-e^{-\alpha\theta jh})]\right\}\\
&\cos\left(\frac{\sigma^\alpha}{\alpha\theta}\beta\tan\frac{\alpha\pi}{2}[|u+ve^{\theta jh}|^\alpha \sin(u+ve^{-\theta jh}) + |v|^\alpha(1-e^{-\alpha\theta jh})\sin(v)]\right),
\end{aligned}
$$

$$
\begin{aligned}
&B_j(u,v) \qquad\qquad (1.4.40)\\
&= \mathbb{E}(\sin(u\tilde{X}_0 + v\tilde{X}_{jh}))\\
&= \exp\left\{ -\frac{\sigma^\alpha}{\alpha\theta}[|u+ve^{-\theta jh}|^\alpha + |v|^\alpha(1-e^{-\alpha\theta jh})]\right\}\\
&\sin\left(\frac{\sigma^\alpha}{\alpha\theta}\beta\tan\frac{\alpha\pi}{2}[|u+ve^{\theta jh}|^\alpha \sin(u+ve^{-\theta jh}) + |v|^\alpha(1-e^{-\alpha\theta jh})\sin(v)]\right).
\end{aligned}
$$

From this computation we have the following formula for the first summand in the sum of e.6.7.

$$
\begin{aligned}
\mathbb{E}(\cos u_1\tilde{X}_0 \cos u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}(\cos u_1(\tilde{X}_0 + \tilde{X}_{jh})) + \mathbb{E}(\cos u_1(\tilde{X}_0 - \tilde{X}_{jh}))}{2}\\
&= \frac{A_j(u_1,u_1) + A_j(u_1,-u_1)}{2}. \qquad (1.4.41)
\end{aligned}
$$

Substituting (1.4.36)-(1.4.37), (1.4.39), and (1.4.41) into (1.4.35) gives the computation for $\sigma_{g_1g_1}$.

*Computation of* $\sigma_{g_2g_2}$. From the definition of $g_2$ we have

$$
\begin{aligned}
\sigma_{g_2g_2} &= \text{cov }(\sin u_1\tilde{X}_0, \sin u_1\tilde{X}_0) + 2\sum_{j=1}^{\infty}[\text{cov }(\sin u_1\tilde{X}_0, \sin u_1\tilde{X}_{jh})]\\
&= \mathbb{E}((\sin u_1\tilde{X}_0)^2) - (\mathbb{E}(\sin u_1\tilde{X}_0))^2\\
&\quad +2\sum_{j=1}^{\infty}\mathbb{E}(\sin u_1\tilde{X}_0 \sin u_1\tilde{X}_{jh}) - \mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\sin u_1\tilde{X}_{jh}). \qquad (1.4.42)
\end{aligned}
$$

The first term in (1.4.42) is given by

$$\begin{aligned}\mathbb{E}((\sin u_1\tilde{X}_0)^2) &= \mathbb{E}\left(\frac{1-\cos 2u_1\tilde{X}_0}{2}\right)=\frac{1}{2}-\frac{1}{2}\mathbb{E}(\cos 2u_1\tilde{X}_0)\\ &= \frac{1}{2}-\frac{1}{2}A_0(2u_1). \end{aligned}\tag{1.4.43}$$

The other terms appeared in (1.4.42) are given by

$$\mathbb{E}(\sin u_1\tilde{X}_0)=B_0(u_1)\tag{1.4.44}$$

and

$$\begin{aligned}\mathbb{E}(\sin u_1\tilde{X}_0\sin u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}(\cos u_1(\tilde{X}_0-\tilde{X}_{jh}))-\mathbb{E}(\cos u_1(\tilde{X}_0+\tilde{X}_{jh}))}{2}\\ &= \frac{A_j(u_1,-u_1)-A_j(u_1,u_1)}{2}\end{aligned}\tag{1.4.45}$$

We can get $\sigma_{g_2g_2}$ from equation (1.4.42).

The method of getting $\sigma_{g_3g_3}$, $\sigma_{g_4g_4}$, $\sigma_{g_5g_5}$, $\sigma_{g_6g_6}$, $\sigma_{g_9g_9}$ and $\sigma_{g_{10}g_{10}}$ are essentially the same as $\sigma_{g_1g_1}$ and $\sigma_{g_2g_2}$ by simply changing the value of $u$.

*Computation of* $\sigma_{g_7g_7}$. From the definition of $g_7$ we have

$$\begin{aligned}\sigma_{g_7g_7} &= \text{cov}\ (\cos u_3(\tilde{X}_h-\tilde{X}_0),\cos u_3(\tilde{X}_h-\tilde{X}_0))\\ &\quad +2\sum_{j=1}^{\infty}[\text{cov}\ (\cos u_3(\tilde{X}_h-\tilde{X}_0),\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))]\\ &= \mathbb{E}((\cos u_3(\tilde{X}_h-\tilde{X}_0))^2)-(\mathbb{E}\cos u_3(\tilde{X}_h-\tilde{X}_0))^2\\ &\quad +2\sum_{j=1}^{\infty}\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\ &\quad -\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})).\end{aligned}\tag{1.4.46}$$

The first term in (1.4.46) is given by

$$\begin{aligned}
&\mathbb{E}((\cos u_3(\tilde{X}_h - \tilde{X}_0))^2) = \mathbb{E}\left(\frac{\cos 2u_3(\tilde{X}_h - \tilde{X}_0) + 1}{2}\right) \\
&= \frac{1}{2} + \frac{1}{2}\mathbb{E}(\cos 2u_3(\tilde{X}_h - \tilde{X}_0)) \\
&= \frac{1}{2} + \frac{1}{2}A_1(-2u_3, 2u_3). \qquad (1.4.47)
\end{aligned}$$

The second term in (1.4.46) is given by

$$\mathbb{E}(\cos u_3(\tilde{X}_h - \tilde{X}_0)) = A_1(-u_3, u_3). \qquad (1.4.48)$$

For any real numbers $u$ and $v$ we have

$$\begin{aligned}
&u(\tilde{X}_h - \tilde{X}_0) + v(\tilde{X}_{(j+1)h} - \tilde{X}_{jh}) \qquad (1.4.49) \\
&= [u(e^{-\theta h} - 1) + v(e^{-\theta(j+1)h} - e^{-\theta jh})]X_0 \\
&+ \int_0^\infty u\sigma e^{-\theta h} e^{\theta s} 1_{[0,h]}(s) dZ_s + \int_0^\infty v\sigma e^{-\theta(j+1)h} e^{\theta s} 1_{[0,(j+1)h]}(s) dZ_s \\
&- \int_0^\infty v\sigma e^{-\theta jh} e^{\theta s} 1_{[0,jh]}(s) dZ_s\,.
\end{aligned}$$

Therefore, we have

$$
\begin{aligned}
w_j(u,v) :=\ & \mathbb{E}[\exp\{iu(\tilde{X}_h - \tilde{X}_0) + iv(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})\}] \\
&= \mathbb{E}[\exp\{i[u(e^{-\theta h} - 1) + v(e^{-\theta(j+1)h} - e^{-\theta jh})]X_0] \\
&\mathbb{E}[\exp\{i(u\sigma e^{-\theta h}\int_0^\infty e^{\theta s}1_{[0,h]}(s)dZ_s - v\sigma e^{-\theta jh}\int_0^\infty e^{\theta s}1_{[0,jh]}(s)dZ_s \\
&\quad + v\sigma e^{-\theta(j+1)h}\int_0^\infty e^{\theta s}1_{[0,(j+1)h]}(s)dZ_s\}] \\
&= \exp\Big\{ -\frac{\sigma^\alpha}{\alpha\theta}[|u(e^{-\theta h} - 1) + v(e^{-\theta(j+1)h} - e^{-\theta jh})|^\alpha \\
&\quad (1 - i\beta(u(e^{-\theta h} - 1) + v(e^{-\theta(j+1)h} - e^{-\theta jh}))\tan\frac{\alpha\pi}{2}) \\
&\quad + |ue^{-\theta h} + ve^{-\theta(j+1)h} - ve^{-\theta jh}|^\alpha (e^{\alpha\theta h} - 1) \\
&\quad (1 - i\beta(ue^{-\theta h} + ve^{-\theta(j+1)h} - ve^{-\theta jh})\tan\frac{\alpha\pi}{2}) \\
&\quad + |v|^\alpha(1 - e^{-\theta h})^\alpha(1 - e^{-\alpha\theta(j-1)h})(1 + i\beta(v)\tan\frac{\alpha\pi}{2}) \\
&\quad + |v|^\alpha(1 - e^{-\alpha\theta h})(1 - i\beta(v)\tan\frac{\alpha\pi}{2})]\Big\}. \qquad (1.4.50)
\end{aligned}
$$

Then the first summand of the sum in (1.4.46) is given by

$$
\begin{aligned}
& \mathbb{E}(\cos u_3(\tilde{X}_h - \tilde{X}_0)\cos u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
= \ & \frac{1}{2}\Big[\mathbb{E}(\cos u_3((\tilde{X}_h - \tilde{X}_0) + (\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& + \mathbb{E}(\cos u_3((\tilde{X}_h - \tilde{X}_0) - (\tilde{X}_{(j+1)h} - \tilde{X}_{jh}))\Big] \\
= \ & \frac{1}{2}\Re\left[w_j(u_3,u_3) + w_j(u_3,-u_3)\right]. \qquad (1.4.51)
\end{aligned}
$$

Then we can get $\sigma_{g_7 g_7}$ from equation (1.4.46).

*Computation of* $\sigma_{g_8 g_8}$. From the definition of $g_8$ we have

$$\begin{aligned}
\sigma_{g_8 g_8} &= \text{cov}\,(\sin u_3(\tilde{X}_h-\tilde{X}_0),\sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
&\quad +2\sum_{j=1}^{\infty}[\text{cov}\,(\sin u_3(\tilde{X}_h-\tilde{X}_0),\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))] \\
&= \mathbb{E}((\sin u_3(\tilde{X}_h-\tilde{X}_0))^2)-(\mathbb{E}\sin u_3(\tilde{X}_h-\tilde{X}_0))^2 \\
&\quad +2\sum_{j=1}^{\infty}\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad -\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\,.
\end{aligned} \tag{1.4.52}$$

$$\begin{aligned}
\mathbb{E}((\sin u_3(\tilde{X}_h-\tilde{X}_0))^2) &= \mathbb{E}\left(\frac{1-\cos 2u_3(\tilde{X}_h-\tilde{X}_0)}{2}\right) \\
&= \frac{1}{2}-\frac{1}{2}\mathbb{E}(\cos 2u_3(\tilde{X}_h-\tilde{X}_0)) \\
&= \frac{1}{2}-\frac{1}{2}A_1(-2u_3,2u_3).
\end{aligned} \tag{1.4.53}$$

$$\mathbb{E}\sin u_3(\tilde{X}_h-\tilde{X}_0)=B_1(-u_3,u_3). \tag{1.4.54}$$

$$\begin{aligned}
&\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
=\;& \frac{1}{2}\left[\mathbb{E}(\cos(u_3(\tilde{X}_h-\tilde{X}_0)-u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})))\right. \\
&\left.-\,\mathbb{E}(\cos(u_3(\tilde{X}_h-\tilde{X}_0)-u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})))\right] \\
=\;& \frac{1}{2}\Re\left[w_j(u_3,-u_3)-w_j(u_3,u_3)\right].
\end{aligned} \tag{1.4.55}$$

Then we can get $\sigma_{g_8 g_8}$ from equation (1.4.52).

*Computation of* $\sigma_{g_1g_2}$. From the definition of $g_1$ and $g_2$ we have

$$\begin{aligned}\sigma_{g_1g_2} &= \text{cov}\,(\cos u_1\tilde{X}_0, \sin u_1\tilde{X}_0) + \sum_{j=1}^{\infty}[\text{cov}\,(\cos u_1\tilde{X}_0, \sin u_1\tilde{X}_{jh}) \\ &\quad + \text{cov}\,(\sin u_1\tilde{X}_0, \cos u_1\tilde{X}_{jh})] \\ &= \mathbb{E}(\cos u_1\tilde{X}_0 \sin u_1\tilde{X}_0) - \mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\sin u_1\tilde{X}_0) \\ &\quad + \sum_{j=1}^{\infty}[\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_1\tilde{X}_{jh}) - \mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\sin u_1\tilde{X}_{jh}) \\ &\quad + \mathbb{E}(\sin u_1\tilde{X}_0 \cos u_1\tilde{X}_{jh}) - \mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\cos u_1\tilde{X}_{jh})],\end{aligned} \tag{1.4.56}$$

where

$$\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_1\tilde{X}_0) = \frac{\mathbb{E}(\sin 2u_1\tilde{X}_0)}{2} = \frac{1}{2}B_0(2u_1), \tag{1.4.57}$$

$$\mathbb{E}(\cos u_1\tilde{X}_0) = A_0(u_1), \tag{1.4.58}$$

$$\mathbb{E}(\sin u_1\tilde{X}_0) = B_0(u_1), \tag{1.4.59}$$

$$\begin{aligned}\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}(\sin u_1(\tilde{X}_0 + \tilde{X}_{jh})) - \mathbb{E}(\sin u_1(\tilde{X}_0 - \tilde{X}_{jh}))}{2} \\ &= \frac{B_j(u_1, u_1) - B_j(u_1, -u_1)}{2},\end{aligned} \tag{1.4.60}$$

$$\begin{aligned}\mathbb{E}(\sin u_1\tilde{X}_0 \cos u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}(\sin u_1(\tilde{X}_0 + \tilde{X}_{jh})) + \mathbb{E}(\sin u_1(\tilde{X}_0 - \tilde{X}_{jh}))}{2} \\ &= \frac{B_j(u_1, u_1) + B_j(u_1, -u_1)}{2}.\end{aligned} \tag{1.4.61}$$

Similarly, we can get $\sigma_{g_3g_4}$, $\sigma_{g_5g_6}$, $\sigma_{g_9g_{10}}$ by changing $u_1$ to $u_2$, $u_3$ and $u_4$.

*Computation of* $\sigma_{g_1 g_3}$. From the definition of $g_1$ and $g_3$ we have

$$\begin{aligned}
\sigma_{g_1 g_3} &= \text{cov}\ (\cos u_1 \tilde{X}_0, \cos u_2 \tilde{X}_0) + \sum_{j=1}^{\infty} [\text{cov}\ (\cos u_1 \tilde{X}_0, \cos u_2 \tilde{X}_{jh}) \\
&\quad + \text{cov}\ (\cos u_2 \tilde{X}_0, \cos u_1 \tilde{X}_{jh})] \\
&= \mathbb{E}(\cos u_1 \tilde{X}_0 \cos u_2 \tilde{X}_0) - \mathbb{E}(\cos u_1 \tilde{X}_0)\mathbb{E}(\cos u_2 \tilde{X}_0) \\
&\quad + \sum_{j=1}^{\infty} \Big[ \mathbb{E}(\cos u_1 \tilde{X}_0 \cos u_2 \tilde{X}_{jh}) - \mathbb{E}(\cos u_1 \tilde{X}_0)\mathbb{E}(\cos u_2 \tilde{X}_{jh}) \\
&\quad + \mathbb{E}(\cos u_2 \tilde{X}_0 \cos u_1 \tilde{X}_{jh}) - \mathbb{E}(\cos u_2 \tilde{X}_0)\mathbb{E}(\cos u_1 \tilde{X}_{jh}) \Big], \qquad (1.4.62)
\end{aligned}$$

where

$$\begin{aligned}
&\mathbb{E}(\cos u_1 \tilde{X}_0 \cos u_2 \tilde{X}_0) \qquad (1.4.63) \\
&= \frac{1}{2}[\mathbb{E}(\cos(u_1+u_2)\tilde{X}_0) + \mathbb{E}(\cos(u_1-u_2)\tilde{X}_0)] \\
&= \frac{1}{2}[A_0(u_1+u_2) + A_0(u_1-u_2)], \\
&\mathbb{E}(\cos u_1 \tilde{X}_0) = A_0(u_1),\ \mathbb{E}(\cos u_2 \tilde{X}_0) = A_0(u_2), \qquad (1.4.64)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\cos u_1 \tilde{X}_0 \cos u_2 \tilde{X}_{jh}) &= \frac{\mathbb{E}\cos(u_1 \tilde{X}_0 + u_2 \tilde{X}_{jh}) + \mathbb{E}\cos(u_1 \tilde{X}_0 - u_2 \tilde{X}_{jh}))}{2} \\
&= \frac{A_j(u_1,u_2) + A_j(u_1,-u_2)}{2}, \qquad (1.4.65) \\
\mathbb{E}(\cos u_2 \tilde{X}_0 \cos u_1 \tilde{X}_{jh}) &= \frac{\mathbb{E}\cos(u_2 \tilde{X}_0 + u_1 \tilde{X}_{jh}) + \mathbb{E}\cos(u_2 \tilde{X}_0 - u_1 \tilde{X}_{jh}))}{2} \\
&= \frac{A_j(u_2,u_1) + A_j(u_2,-u_1)}{2}. \qquad (1.4.66)
\end{aligned}$$

Then we can get $\sigma_{g_1 g_3}$ from equation (1.4.62).

Similarly, we can get $\sigma_{g_1 g_5}$, $\sigma_{g_1 g_9}$, $\sigma_{g_3 g_5}$, $\sigma_{g_3 g_9}$, and $\sigma_{g_5 g_9}$.

*Computation of* $\sigma_{g_1g_4}$. From the definition of $g_1$ and $g_4$ we have

$$\begin{aligned}
\sigma_{g_1g_4} &= \text{cov}\ (\cos u_1\tilde{X}_0, \sin u_2\tilde{X}_0) + \sum_{j=1}^{\infty}[\text{cov}\ (\cos u_1\tilde{X}_0, \sin u_2\tilde{X}_{jh}) \\
&\quad + \text{cov}\ (\sin u_2\tilde{X}_0, \cos u_1\tilde{X}_{jh})] \\
&= \mathbb{E}(\cos u_1\tilde{X}_0 \sin u_2\tilde{X}_0) - \mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\sin u_2\tilde{X}_0) \\
&\quad + \sum_{j=1}^{\infty} \mathbb{E}(\cos u_1\tilde{X}_0 \sin u_2\tilde{X}_{jh}) - \mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\sin u_2\tilde{X}_{jh}) \\
&\quad + \sum_{j=1}^{\infty} \mathbb{E}(\sin u_2\tilde{X}_0 \cos u_1\tilde{X}_{jh}) - \mathbb{E}(\sin u_2\tilde{X}_0)\mathbb{E}(\cos u_1\tilde{X}_{jh}), \qquad (1.4.67)
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_2\tilde{X}_0) &= \frac{1}{2}[\mathbb{E}(\sin(u_1+u_2)\tilde{X}_0) - \mathbb{E}(\sin(u_1-u_2)\tilde{X}_0)] \\
&= \frac{1}{2}[B_0(u_1+u_2) - B_0(u_1-u_2)], \qquad (1.4.68)
\end{aligned}$$

$$\mathbb{E}(\cos u_1\tilde{X}_0) = A_0(u_1),\ \mathbb{E}(\sin u_2\tilde{X}_0) = B_0(u_2), \qquad (1.4.69)$$

$$\begin{aligned}
\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_2\tilde{X}_{jh}) &= \frac{\mathbb{E}\sin(u_1\tilde{X}_0 + u_2\tilde{X}_{jh}) - \mathbb{E}\sin(u_1\tilde{X}_0 - u_2\tilde{X}_{jh}))}{2} \\
&= \frac{B_j(u_1,u_2) - B_j(u_1,-u_2)}{2}, \qquad (1.4.70)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\sin u_2\tilde{X}_0 \cos u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}\sin(u_2\tilde{X}_0 + u_1\tilde{X}_{jh}) - \mathbb{E}\sin(u_2\tilde{X}_0 - u_1\tilde{X}_{jh}))}{2} \\
&= \frac{B_j(u_2,u_1) - B_j(u_2,-u_1)}{2}. \qquad (1.4.71)
\end{aligned}$$

Then we can get $\sigma_{g_1g_4}$ from equation (1.4.67).

Similarly, we can get $\sigma_{g_1g_6}$, $\sigma_{g_1g_{10}}$, $\sigma_{g_3g_2}$, $\sigma_{g_3g_6}$, $\sigma_{g_3g_{10}}$, $\sigma_{g_5g_2}$, $\sigma_{g_5g_4}$, $\sigma_{g_5g_{10}}$, $\sigma_{g_9g_2}$, $\sigma_{g_9g_4}$, and $\sigma_{g_9g_6}$ by changing the value of $u_1$ and $u_2$.

*Computation of* $\sigma_{g_2g_4}$. From the definition of $g_2$ and $g_4$ we have

$$\begin{aligned}
\sigma_{g_2g_4} &= \text{cov}\,(\sin u_1\tilde{X}_0, \sin u_2\tilde{X}_0) + \sum_{j=1}^{\infty}[\text{cov}\,(\sin u_1\tilde{X}_0, \sin u_2\tilde{X}_{jh}) \\
&\quad + \text{cov}\,(\sin u_2\tilde{X}_0, \sin u_1\tilde{X}_{jh})] \\
&= \mathbb{E}(\sin u_1\tilde{X}_0 \sin u_2\tilde{X}_0) - \mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\sin u_2\tilde{X}_0) \\
&\quad + \sum_{j=1}^{\infty} \mathbb{E}(\sin u_1\tilde{X}_0 \sin u_2\tilde{X}_{jh}) - \mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\sin u_2\tilde{X}_{jh}) \\
&\quad + \sum_{j=1}^{\infty} \mathbb{E}(\sin u_2\tilde{X}_0 \sin u_1\tilde{X}_{jh}) - \mathbb{E}(\sin u_2\tilde{X}_0)\mathbb{E}(\sin u_1\tilde{X}_{jh}), \qquad (1.4.72)
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}(\sin u_1\tilde{X}_0 \sin u_2\tilde{X}_0) &= \frac{1}{2}[\mathbb{E}(\cos(u_1-u_2)\tilde{X}_0) - \mathbb{E}(\cos(u_1+u_2)\tilde{X}_0)] \\
&= \frac{1}{2}\left[A_0(u_1-u_2) - A_0(u_1+u_2)\right], \qquad (1.4.73)
\end{aligned}$$

$$\mathbb{E}(\sin u_1\tilde{X}_0) = B_0(u_1),\ \mathbb{E}(\sin u_2\tilde{X}_0) = B_0(u_2), \qquad (1.4.74)$$

$$\begin{aligned}
\mathbb{E}(\sin u_1\tilde{X}_0 \sin u_2\tilde{X}_{jh}) &= \frac{\mathbb{E}\cos(u_1\tilde{X}_0 - u_2\tilde{X}_{jh}) - \mathbb{E}\cos(u_1\tilde{X}_0 + u_2\tilde{X}_{jh}))}{2} \\
&= \frac{A_j(u_1,-u_2) - A_j(u_1,u_2)}{2}, \qquad (1.4.75) \\
\mathbb{E}(\sin u_2\tilde{X}_0 \sin u_1\tilde{X}_{jh}) &= \frac{\mathbb{E}\cos(u_2\tilde{X}_0 - u_1\tilde{X}_{jh}) - \mathbb{E}\cos(u_2\tilde{X}_0 + u_1\tilde{X}_{jh}))}{2} \\
&= \frac{A_j(u_2,-u_1) - A_j(u_2,u_1)}{2}. \qquad (1.4.76)
\end{aligned}$$

Then we can get $\sigma_{g_2g_4}$ from equation (1.4.72).

Similarly, we can get $\sigma_{g_2g_6}$, $\sigma_{g_2g_{10}}$, $\sigma_{g_4g_6}$, $\sigma_{g_4g_{10}}$, and $\sigma_{g_6g_{10}}$ by changing the value of $u_1$ and $u_2$.

*Computation of* $\sigma_{g_7g_8}$. From the definition of $g_7$ and $g_8$ we have

$$\begin{aligned}
\sigma_{g_7g_8} &= \text{cov}\ (\cos u_3(\tilde{X}_h-\tilde{X}_0), \sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
&\quad + \sum_{j=1}^{\infty} [\text{cov}\ (\cos u_3(\tilde{X}_h-\tilde{X}_0), \sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad + \text{cov}\ (\sin u_3(\tilde{X}_h-\tilde{X}_0), \cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))] \\
&= \mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
&\quad - \mathbb{E}\cos u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}\sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
&\quad + \sum_{j=1}^{\infty} [\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad - \mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad + \mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad - \mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))], \qquad (1.4.77)
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_h-\tilde{X}_0)) &= \mathbb{E}\left(\frac{\sin 2u_3(\tilde{X}_h-\tilde{X}_0)}{2}\right) \\
&= \frac{1}{2}B_1(-2u_3, 2u_3), \qquad (1.4.78)
\end{aligned}$$

$$\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)) = A_1(-u_3, u_3), \qquad (1.4.79)$$

$$\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)) = B_1(-u_3, u_3), \qquad (1.4.80)$$

$$
\begin{aligned}
& \mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) && (1.4.81)\\
= \; & \frac{1}{2}\Big[\mathbb{E}(\sin(u_3(\tilde{X}_h-\tilde{X}_0))+u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\
& -\mathbb{E}(\sin(u_3(\tilde{X}_h-\tilde{X}_0)-u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\Big]\\
= \; & \frac{1}{2}\Im\left[w_j(u_3,u_3)-w_j(u_3,-u_3)\right],\\
& \mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) && (1.4.82)\\
= \; & \frac{1}{2}\Big[\mathbb{E}(\sin(u_3(\tilde{X}_h-\tilde{X}_0))+u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\
& +\mathbb{E}(\sin(u_3(\tilde{X}_h-\tilde{X}_0)-u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\Big]\\
= \; & \frac{1}{2}\Im\left[w_j(u_3,u_3)+w_j(u_3,-u_3)\right].
\end{aligned}
$$

Then we can get $\sigma_{g_7g_8}$ from equation (1.4.77).

*Computation of* $\sigma_{g_1g_7}$. From the definition of $g_1$ and $g_7$ we have

$$
\begin{aligned}
\sigma_{g_1g_7} = \; & \text{cov}\,(\cos u_1\tilde{X}_0,\cos u_3(\tilde{X}_h-\tilde{X}_0))\\
& +\sum_{j=1}^{\infty}[\text{cov}\,(\cos u_1\tilde{X}_0,\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\
& +\text{cov}\,(\cos u_3(\tilde{X}_h-\tilde{X}_0),\cos u_1\tilde{X}_{jh})]\\
= \; & \mathbb{E}(\cos u_1\tilde{X}_0\cos u_3(\tilde{X}_h-\tilde{X}_0))-\mathbb{E}\cos u_1\tilde{X}_0\mathbb{E}\cos u_3(\tilde{X}_h-\tilde{X}_0))\\
& +\sum_{j=1}^{\infty}[\mathbb{E}(\cos u_1\tilde{X}_0\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\
& -\mathbb{E}(\cos u_1\tilde{X}_0)\mathbb{E}(\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\\
& +\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\cos u_1\tilde{X}_{jh})\\
& -\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\cos u_1\tilde{X}_{jh}))]. && (1.4.83)
\end{aligned}
$$

Note that

$$
\mathbb{E}\cos u_3(\tilde{X}_h-\tilde{X}_0)=A_1(-u_3,u_3),\mathbb{E}\cos u_1\tilde{X}_0=A_0(u_1). \tag{1.4.84}
$$

We write

$$
\begin{aligned}
u\tilde{X}_0+v(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}) &= (u+ve^{-\theta(j+1)h}-ve^{-\theta jh})\tilde{X}_0 \\
&\quad +v\sigma e^{-\theta(j+1)h}\int_0^{(j+1)h} e^{\theta s}dZ_s - v\sigma e^{-\theta jh}\int_0^{jh} e^{\theta s}dZ_s\,.
\end{aligned}
$$

Let

$$
\begin{aligned}
\rho_j(u,v) &:= \mathbb{E}[\exp\{iu\tilde{X}_0+iv(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})\}] \\
&= \mathbb{E}[\exp\{i[u+v(e^{-\theta(j+1)h}-e^{-\theta jh})]X_0] \\
&= \mathbb{E}[\exp\{i(v\sigma e^{-\theta(j+1)h}\int_0^{\infty} e^{\theta s}1_{[0,(j+1)h]}(s)dZ_s \\
&\qquad -v\sigma e^{-\theta jh}\int_0^{\infty} e^{\theta s}1_{[0,jh]}(s)dZ_s\})] \\
&= \exp\Bigg\{-\frac{\sigma^\alpha}{\alpha\theta}[|u+v(e^{-\theta(j+1)h}-e^{-\theta jh})|^\alpha \\
&\qquad (1-i\beta(u+v(e^{-\theta(j+1)h}-e^{-\theta jh}))\tan\frac{\alpha\pi}{2}) \\
&\qquad +|v|^\alpha(1-e^{-\theta h})^\alpha(1-e^{-\alpha\theta jh})(1+i\beta(v)\tan\frac{\alpha\pi}{2}) \\
&\qquad +|v|^\alpha(1-e^{-\alpha\theta h})(1-i\beta(v)\tan\frac{\alpha\pi}{2})]\Bigg\}\,. \qquad (1.4.85)
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathbb{E}(\cos u_1\tilde{X}_0\cos u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\quad = \frac{1}{2}\Bigg[\mathbb{E}(\cos(u_1\tilde{X}_0+u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
&\qquad\quad +\mathbb{E}(\cos(u_1\tilde{X}_0-u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh}))\Bigg] \\
&\quad = \frac{1}{2}\Re\left[\rho_j(u_1,u_3)+\rho_j(u_1,-u_3)\right]. \qquad (1.4.86)
\end{aligned}
$$

We write

$$
\begin{aligned}
u\tilde{X}_{jh}+v(\tilde{X}_h-\tilde{X}_0) &= (ue^{-\theta jh}+v(e^{-\theta h}-1))\tilde{X}_0 \\
&+u\sigma e^{-\theta jh}\int_0^{jh} e^{\theta s}dZ_s+v\sigma e^{-\theta h}\int_0^{h} e^{\theta s}dZ_s\,.
\end{aligned}
$$

Let

$$
\begin{aligned}
\kappa_j(u,v) &:= \mathbb{E}[\exp\{iu\tilde{X}_{jh}+iv(\tilde{X}_h-\tilde{X}_0)\}] \\
&= \mathbb{E}[\exp\{i[ue^{-\theta jh}+v(e^{-\theta h}-1)]\tilde{X}_0] \\
&\quad \times\mathbb{E}[\exp\{i(u\sigma e^{-\theta jh}\int_0^{jh} e^{\theta s}dZ_s+v\sigma e^{-\theta h}\int_0^{h} e^{\theta s}dZ_s)] \\
&= \exp\Big\{-\frac{\sigma^\alpha}{\alpha\theta}[ue^{-\theta jh}+v(e^{-\theta h}-1)|^\alpha \\
&\quad (1-i\beta(ue^{-\theta jh}+v(e^{-\theta h}-1))\tan\frac{\alpha\pi}{2}) \\
&\quad +|ue^{-\theta jh}+ve^{-\theta h}|^\alpha(e^{\alpha\theta h}-1) \\
&\quad (1-i\beta(ue^{-\theta jh}+ve^{-\theta h})\tan\frac{\alpha\pi}{2}) \\
&\quad +|u|^\alpha(1-e^{-\alpha\theta(j-1)h})(1-i\beta(u)\tan\frac{\alpha\pi}{2})]\Big\}\,. \qquad (1.4.87)
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\cos u_1\tilde{X}_{jh}) &= \frac{1}{2}\Big[\mathbb{E}(\cos(u_1\tilde{X}_{jh}+u_3(\tilde{X}_h-\tilde{X}_0))) \\
&\quad +\mathbb{E}(\cos(u_1\tilde{X}_{jh}-u_3(\tilde{X}_h-\tilde{X}_0)))\Big] \\
&= \frac{1}{2}\Re\left[\kappa_j(u_1,u_3)+\kappa_j(u_1,-u_3)\right]. \qquad (1.4.88)
\end{aligned}
$$

Then we can get $\sigma_{g_1g_7}$ from equation (1.4.83). By changing the value of $u_1$, we can get $\sigma_{g_3g_7}$ ,$\sigma_{g_5g_7}$ , and $\sigma_{g_9g_7}$.

*Computation of* $\sigma_{g_1 g_8}$. From the definition of $g_1$ and $g_8$ we have

$$\begin{aligned}
\sigma_{g_1 g_8} &= \text{cov }(\cos u_1\tilde{X}_0, \sin u_3(\tilde{X}_h - \tilde{X}_0)) \\
&\quad + \sum_{j=1}^{\infty} [\text{cov }(\cos u_1\tilde{X}_0, \sin u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
&\quad + \text{cov }(\cos u_3(\tilde{X}_h - \tilde{X}_0), \sin u_1\tilde{X}_{jh})] \\
&= \mathbb{E}(\cos u_1\tilde{X}_0 \sin u_3(\tilde{X}_h - \tilde{X}_0)) - \mathbb{E}\cos u_1\tilde{X}_0 \mathbb{E}\sin u_3(\tilde{X}_h - \tilde{X}_0)) \\
&\quad + \sum_{j=1}^{\infty} [\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
&\quad - \mathbb{E}(\cos u_1\tilde{X}_0) E(\sin u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
&\quad + \mathbb{E}(\sin u_3(\tilde{X}_h - \tilde{X}_0) \cos u_1\tilde{X}_{jh}) \\
&\quad - \mathbb{E}(\sin u_3(\tilde{X}_h - \tilde{X}_0)) E(\cos u_1\tilde{X}_{jh}))], \qquad (1.4.89)
\end{aligned}$$

where

$$\mathbb{E}\cos u_1\tilde{X}_0 = A_0(u_1), \ \mathbb{E}\sin u_3(\tilde{X}_h - \tilde{X}_0) = B_1(-u_3, u_3). \qquad (1.4.90)$$

By equations (1.4.85) and (1.4.87), we get

$$\begin{aligned}
&\mathbb{E}(\cos u_1\tilde{X}_0 \sin u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \qquad (1.4.91) \\
&= \frac{1}{2}\Big[\mathbb{E}(\sin(u_1\tilde{X}_0 + u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
&\quad - \mathbb{E}(\sin(u_1\tilde{X}_0 - u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh}))\Big] \\
&= \frac{1}{2}\Im\left[\rho_j(u_1, u_3) - \rho_j(u_1, -u_3)\right], \\
&\mathbb{E}(\sin u_3(\tilde{X}_h - \tilde{X}_0) \cos u_1\tilde{X}_{jh}) \qquad (1.4.92) \\
&= \frac{1}{2}\Big[\mathbb{E}(\sin(u_1\tilde{X}_{jh} + u_3(\tilde{X}_h - \tilde{X}_0))) \\
&\quad - \mathbb{E}(\sin(u_1\tilde{X}_{jh} - u_3(\tilde{X}_h - \tilde{X}_0)))\Big] \\
&= \frac{1}{2}\Im\left[\kappa_j(u_1, u_3) - \kappa_j(u_1, -u_3)\right].
\end{aligned}$$

Then we can get $\sigma_{g_1g_8}$ from equation (1.4.89). By changing the value of $u_1$, we can get $\sigma_{g_3g_8}$ ,$\sigma_{g_5g_8}$, and $\sigma_{g_9g_8}$.

*Computation of* $\sigma_{g_2g_7}$. From the definition of $g_2$ and $g_7$ we have

$$\begin{aligned}
\sigma_{g_2g_7} \quad = \quad & \text{cov }(\sin u_1\tilde{X}_0, \cos u_3(\tilde{X}_h - \tilde{X}_0)) \\
& + \sum_{j=1}^{\infty} [\text{cov }(\sin u_1\tilde{X}_0, \cos u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& + \text{cov }(\cos u_3(\tilde{X}_h - \tilde{X}_0), \sin u_1\tilde{X}_{jh})] \\
& = \mathbb{E}(\sin u_1\tilde{X}_0 \cos u_3(\tilde{X}_h - \tilde{X}_0)) - \mathbb{E}\sin u_1\tilde{X}_0 \mathbb{E}\cos u_3(\tilde{X}_h - \tilde{X}_0)) \\
& + \sum_{j=1}^{\infty} [\mathbb{E}(\sin u_1\tilde{X}_0 \cos u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& - \mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\cos u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& + \mathbb{E}(\cos u_3(\tilde{X}_h - \tilde{X}_0) \sin u_1\tilde{X}_{jh}) \\
& - \mathbb{E}(\cos u_3(\tilde{X}_h - \tilde{X}_0))\mathbb{E}(\sin u_1\tilde{X}_{jh}))]. \qquad (1.4.93)
\end{aligned}$$

Note that

$$\mathbb{E}\sin u_1\tilde{X}_0 = B_0(u_1),\ \mathbb{E}\cos u_3(\tilde{X}_h - \tilde{X}_0) = A_1(-u_3, u_3). \qquad (1.4.94)$$

By equations (1.4.85) and (1.4.87), we get

$$\begin{aligned}
& \mathbb{E}(\sin u_1\tilde{X}_0 \cos u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
= \quad & \frac{1}{2}\Big[\mathbb{E}(\sin(u_1\tilde{X}_0 + u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& + \mathbb{E}(\sin(u_1\tilde{X}_0 - u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh}))\Big] \\
= \quad & \frac{1}{2}\Im\left[\rho_j(u_1, u_3) + \rho_j(u_1, -u_3)\right] \qquad (1.4.95)
\end{aligned}$$

and

$$
\begin{aligned}
& \mathbb{E}(\cos u_3(\tilde{X}_h-\tilde{X}_0)\sin u_1\tilde{X}_{jh}) \\
= \;& \frac{1}{2}\Big[\mathbb{E}(\sin(u_1\tilde{X}_{jh}+u_3(\tilde{X}_h-\tilde{X}_0))) \\
& +\mathbb{E}(\sin(u_1\tilde{X}_{jh}-u_3(\tilde{X}_h-\tilde{X}_0)))\Big] \\
= \;& \frac{1}{2}\Im\left[\kappa_j(u_1,u_3)+\kappa_j(u_1,-u_3)\right]. \qquad (1.4.96)
\end{aligned}
$$

Then we can get $\sigma_{g_2g_7}$ from equation (1.4.93). By changing the value of $u_1$, we can get $\sigma_{g_4g_7}$, $\sigma_{g_6g_7}$, and $\sigma_{g_10g_7}$.

*Computation of* $\sigma_{g_2g_8}$. From the definition of $g_2$ and $g_8$ we have

$$
\begin{aligned}
\sigma_{g_2g_8} = \;& \operatorname{cov}\left(\sin u_1\tilde{X}_0,\sin u_3(\tilde{X}_h-\tilde{X}_0)\right) \qquad (1.4.97)\\
& +\sum_{j=1}^{\infty}[\operatorname{cov}\left(\sin u_1\tilde{X}_0,\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})\right) \\
& +\operatorname{cov}\left(\sin u_3(\tilde{X}_h-\tilde{X}_0),\sin u_1\tilde{X}_{jh}\right)] \\
= \;& \mathbb{E}(\sin u_1\tilde{X}_0\sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
& -\mathbb{E}\sin u_1\tilde{X}_0\mathbb{E}\sin u_3(\tilde{X}_h-\tilde{X}_0)) \\
& +\sum_{j=1}^{\infty}[\mathbb{E}(\sin u_1\tilde{X}_0\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
& -\mathbb{E}(\sin u_1\tilde{X}_0)\mathbb{E}(\sin u_3(\tilde{X}_{(j+1)h}-\tilde{X}_{jh})) \\
& +\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0)\sin u_1\tilde{X}_{jh}) \\
& -\mathbb{E}(\sin u_3(\tilde{X}_h-\tilde{X}_0))\mathbb{E}(\sin u_1\tilde{X}_{jh}))].
\end{aligned}
$$

Note that

$$
\mathbb{E}\sin u_1\tilde{X}_0=B_0(u_1),\ \mathbb{E}\sin u_3(\tilde{X}_h-\tilde{X}_0)=B_1(-u_3,u_3). \qquad (1.4.98)
$$

By equations (1.4.85) and (1.4.87), we find

$$
\begin{aligned}
& \mathbb{E}(\sin u_1 \tilde{X}_0 \sin u_3 (\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
= \; & \frac{1}{2}\Big[\mathbb{E}(\cos(u_1 \tilde{X}_0 - u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh})) \\
& -\mathbb{E}(\cos(u_1 \tilde{X}_0 + u_3(\tilde{X}_{(j+1)h} - \tilde{X}_{jh}))\Big] \\
= \; & \frac{1}{2}\Re\left[\rho_j(u_1, -u_3) - \rho_j(u_1, u_3)\right].
\end{aligned} \tag{1.4.99}
$$

$$
\begin{aligned}
& \mathbb{E}(\sin u_3 (\tilde{X}_h - \tilde{X}_0) \sin u_1 \tilde{X}_{jh}) \\
& = \frac{1}{2}\left[\mathbb{E}(\cos(u_1 \tilde{X}_{jh} - u_3(\tilde{X}_h - \tilde{X}_0)))\right. \\
& \left. - \;\mathbb{E}(\cos(u_1 \tilde{X}_{jh} + u_3(\tilde{X}_h - \tilde{X}_0)))\right] \\
& = \frac{1}{2}\Re\left[\kappa_j(u_1, -u_3) - \kappa_j(u_1, u_3)\right].
\end{aligned} \tag{1.4.100}
$$

Then we can get $\sigma_{g_2 g_8}$ by equation (1.4.97). Similarly, we can get $\sigma_{g_4 g_8}$,$\sigma_{g_6 g_8}$,and $\sigma_{g_{10} g_8}$.

Thus, we have obtained the explicit expression of $\Sigma_{10} = (\sigma_{g_k g_l})_{1 \le k,l \le 10}$.

## 1.5 Simulation

In this section we shall validate our estimators discussed in Section 1.4. We consider the following specific $\alpha$-stable Ornstein-Uhlenbeck motion determined by (1.2.1) which we restate as follows:

$$
dX_t = -\theta X_t dt + \sigma dZ_t, \quad X_0 \quad \text{is given}. \tag{1.5.1}
$$

First we describe our approach to simulate the above process. There have been numerous schemes to simulate the above process. However, in all the existing schemes one needs to divide the interval $[0, T]$ into small intervals $0 = t_0 < t_1 < \cdots < t_N = T = n\tilde{h}$ such that the partition step size $t_{k+1} - t_k =$

$\tilde{h}$ goes to zero. This means that we would need to simulate $nh/\tilde{h}$ many random variables. As we need $n \to \infty$ and we allow $h$ to be a constant, this will require too large amount of computations. For this specific equation (1.5.1), we shall use the following scheme. This scheme may also be useful in other applications. For our scheme we can allow $\tilde{h} = h$. From (1.5.1) we see easily that

$$X_t = e^{-\theta(t-s)} X_s + \sigma \int_s^t e^{-\theta(t-r)} dZ_r.$$

Thus

$$X_{k+1} = e^{-\theta h} X_k + \sigma \int_{kh}^{(k+1)h} e^{-\theta((k+1)h-r)} dZ_r.$$

Since $f(r) = \sigma e^{-\theta((k+1)h-r)}$ is a deterministic function we see that

$$\sigma \int_{kh}^{(k+1)h} e^{-\theta((k+1)h-r)} dZ_r \stackrel{d}{=} \left( \int_{kh}^{(k+1)h} f^{\alpha}(t) dt \right)^{\frac{1}{\alpha}} DZ_k,$$

where $DZ_k$ are iid $\alpha$-stable random variables. Janicki and Weron (1994) proposed numerical simulation of independent $\alpha$-stable random variables. However, there is an error in Janicki and Weron (1994), which is corrected in Weron and Weron (1995). We shall use the following formula to simulate $DZ_k$:

$$DZ_k = D \sin(\alpha U_k + \alpha C) \left( \frac{cos(U_k - \alpha(U_k + C))}{W_k} \right)^{\frac{1-\alpha}{\alpha}} / \cos(U_k)^{\frac{1}{\alpha}}.$$

Here, $U_k$ are iid uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$, $W_k$ are iid exponentially distributed with mean 1, $D = \left(1 + \beta^2 \tan^2 \frac{\alpha\pi}{2}\right)^{\frac{1}{2\alpha}}$ and $C = \left(\arctan(\beta \tan \frac{\alpha\pi}{2})\right)/\alpha$

Then, we have the iteration as

$$X_{k+1} = e^{-\theta h} X_k + \sigma \frac{1}{(\theta\alpha)^{\frac{1}{\alpha}}} (1 - e^{-\alpha\theta h})^{\frac{1}{\alpha}} DZ_k.$$

To be specific we choose the following baseline parameter values and simulate the process in the interval $[0, T]$ with $nh = T = 10000$. We shall fix $h = 0.5$. For the four grid points $u_1$, $u_2$, $u_3$ and

$u_4$, we select them in a certain optimal way which is discussed in detail in the Subsection 3.3 and here we choose the $a = 12$, $K = 120$ and $\varepsilon = 10^{-3}$. Values of the four parameters used are given in Table 2.1. Here we use two sets of values.

**Table 1.1** Parameter for the following tables.

| Variable | $\beta$ | $\alpha$ | $\sigma$ | $\theta$ |
|---|---|---|---|---|
| Assumed Value | 0.4 | 1.7 | 0.2 | 2 |
| Assumed Value | -0.6 | 0.6 | 0.4 | 5 |

The following two tables give the mean and standard deviation of the estimators with the first set of assumed values as the value $n$ changing from a smaller value to a greater value. For the grid points, we are choosing them in the optimal way. So they are different for different sample paths, here we just list one set of values. The optimal grid points we got from one sample path is $\{4.97, 5.92, 6.04, 10.80\}$. We can see that as the value of $n$ is getting larger, the standard deviation will become smaller.

**Table 1.2** Mean the estimators $\hat{\alpha}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\beta}$ with $h = 0.5$ through 500 paths at different value of $n$. Case: $\alpha = 1.7$, $\theta = 2$, $\sigma = 0.2$, $\beta = 0.4$

| Mean | n ($\times 10^4$) | | | |
|---|---|---|---|---|
| | 0.8 | 1.2 | 1.6 | 2 |
| $\alpha$ | 1.7008 | 1.69458 | 1.6980 | 1.6994 |
| $\theta$ | 2.0158 | 2.0117 | 2.0087 | 2.0049 |
| $\sigma$ | 0.2007 | 0.1989 | 0.1997 | 0.1998 |
| $\beta$ | 0.3975 | 0.4063 | 0.4009 | 0.4029 |

**Table 1.3** Standard deviation of the estimators $\hat{\alpha}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\beta}$ with $h = 0.5$ through 500 paths at different value of $n$. Case: $\alpha = 1.7$, $\theta = 2$, $\sigma = 0.2$, $\beta = 0.4$

| Std | n ($\times 10^4$) | | | |
|---|---|---|---|---|
| | 0.8 | 1.2 | 1.6 | 2 |
| $\alpha$ | 0.0233 | 0.0229 | 0.0169 | 0.0162 |
| $\theta$ | 0.0716 | 0.0604 | 0.0492 | 0.0405 |
| $\sigma$ | 0.0069 | 0.0066 | 0.0060 | 0.0051 |
| $\beta$ | 0.0573 | 0.0435 | 0.0312 | 0.0278 |

The following two tables give the mean and standard deviation of the estimators with the second set of assumed values as the value $n$ changing from a smaller value to a greater value. In this case, $0 < \alpha < 1$ and $\beta < 0$. And the optimal grid points we got from one sample path is $\{0.2, 3.08, 6.05, 9.03\}$. It will be different for other paths.

**Table 1.4** Mean the estimators $\hat{\alpha}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\beta}$ with $h = 0.5$ through 500 paths at different value of $n$. Case: $\alpha = 0.6$, $\theta = 5$, $\sigma = 0.4$, $\beta = -0.6$

| Mean | n ($\times 10^4$) | | | |
|---|---|---|---|---|
| | 0.8 | 1.2 | 1.6 | 2 |
| $\alpha$ | 0.5926 | 0.5874 | 0.5907 | 0.5958 |
| $\theta$ | 5.0948 | 5.1334 | 5.1137 | 5.0479 |
| $\sigma$ | 0.3933 | 0.3888 | 0.3919 | 0.3925 |
| $\beta$ | $-0.6378$ | $-0.6560$ | $-0.6442$ | $-0.6018$ |

**Table 1.5** Standard deviation of the estimators $\hat{\alpha}$, $\hat{\theta}$, $\hat{\sigma}$, $\hat{\beta}$ with $h = 0.5$ through 500 paths at different value of $n$. Case: $\alpha = 0.6$, $\theta = 5$, $\sigma = 0.4$, $\beta = -0.6$

| Std | n ($\times 10^4$) | | | |
|---|---|---|---|---|
| | 0.8 | 1.2 | 1.6 | 2 |
| $\alpha$ | 0.0188 | 0.0185 | 0.0160 | 0.0118 |
| $\theta$ | 0.5300 | 0.3922 | 0.2761 | 0.2101 |
| $\sigma$ | 0.0547 | 0.0312 | 0.0279 | 0.0271 |
| $\beta$ | 0.0639 | 0.0627 | 0.0515 | 0.0446 |

# Chapter 2

# The Expected Hitting Time Approach to Optimal Price Adjustment Problems

In this chapter, we offer a novel approach for solving optimal price adjustment problems, when the underlying process is a Geometric Brownian Motion (GBM) process. Our approach relies on characterizing the cumulative cost of deviation and the cost of adjusting price until the hitting time of the lower or upper barriers. Using this approach, we are able to derive an analytical expression for the cost function, that does not require solving a PDE or running Monte-Carlo simulations. We apply our framework to the real world problem of adjusting domestic energy prices in countries that adopt administratively-set energy price rules.

## 2.1 Preliminary

**Definition 2.1.** A stochastic process $S_t$ is said to follow a GBM if it satisfies the following stochastic differential equation (SDE):

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where $W_t$ is a Wiener process, and $\mu$ (drift) and $\sigma$ (volatility) are constants.

Most economists prefer geometric Brownian motion as a simple model for market prices because it is everywhere positive (with probability 1), in contrast to Brownian Motion, or even Brownian Motion with drift. Furthermore, as we see from the stochastic differential equation for geometric Brownian motion, the relative change is combination of a deterministic term similar to the inflation or interest rate plus a normally distributed random term. For an arbitrary initial value

$S_0$, by a straightforward application of Itô's lemma the above SDE has the analytic solution:

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right)$$

The above solution $S_t$ (for any value of t) is a log-normally distributed random variable with the expected value and variance given by

$$\begin{aligned}
&\mathrm{E}\left(S_t\right) = S_0 e^{\mu t} \\
&\mathrm{Var}\left(S_t\right) = S_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right)
\end{aligned}$$

**Properties 2.1.** The product or quotion of two uncorrelated geometric Brownian motions remains geometric Brownian motion.

*Proof.* Let

$$dY/Y = adt + bdW_Y$$

$$dZ/Z = fdt + gdW_Z$$

where $\langle dW_Y, dW_Z\rangle = 0$, consider the process $U \equiv Y * Z$. By applying the Ito's lemma, we have

$$\begin{aligned}
dU =& ZdY + YdZ + dYdZ \\
=& ZY\left(adt + bdW_Y\right) + YZ\left(fdt + gdW_Z\right) \\
&+ YZ\left(adt + bdW_Y\right)\left(fdt + gdW_Z\right) \\
=& U(a+f)dt + UbdW_Y + UgdW_Z
\end{aligned}$$

Thus

$$\begin{aligned}
dU/U &= (a+f)dt + bdW_Y + gdW_Z \\
&= (a+f)dt + \sqrt{b^2+g^2}dW_U
\end{aligned}$$

□

**Definition 2.2.** (First Hitting Time) The mathematical definition for the first hitting time $H = H_{a,b}$ for the real valued stochastic process V to reach (or cross) the barrier $a$ or $b$, assuming that the process starts with $V(0) \leftarrow (a,b)$, is given by the equation:

$$H = H_{a,b} = \min\{s : V_s \notin (a,b)\}$$

## 2.2 Introduction

Several countries use the administratively-set fuel prices (which are offered in domestic currency) close to their international free market counterparts ([8], [24]). However, chasing a global market price of energy/fuel has the disadvantage that domestic prices need to fluctuate daily. This creates uncertainties for households and firms, and exposes them to global price shocks. A solution between the two extremes, one totally floating and one completely detached from global prices, is to fix prices for a certain time interval (e.g., a season or a year) and then re-adjust and announce new prices at the beginning of the new period.

The limitation of the adjustment on a fixed time interval is that if global prices deviate too much from the domestically announced prices, one side of the domestic market may bear large costs. For example, if the global price starts rising or the country's domestic currency devaluates quickly, the domestic prices may become too cheap (compared to the global benchmark) and the government needs to fill the gap by paying a large subsidy. On the other hand, if global prices drop significantly (similar to what happened after 2008 or in 2014), domestic prices will be more expensive than the global equilibrium price and consumers are forced to pay an extra price. Given the high level of volatility in global fuel markets, those deviations from the optimal level are likely and can include substantial costs.

We offer a model of adjustment rules which is based on optimal lower and upper price barriers. Instead of adjusting domestic prices in calendar time, the policy maker observes the dynamics of global prices and the exchange rate. More precisely, the policy maker observes the system up to a certain random time as opposed to a fixed time horizon, known as a sequential plan. The domestic price will be intact as long as the ratio of global and domestic prices (using the same currency) stays inside of the optimal range. Once the ratio of the two prices hits the lower or upper bound, the domestic price will then be re-adjusted to the original desired level. We formulate this intuitive rule as a simplified impulse control problem and solve it explicitly.

## 2.3 Model

We first present a general form of the model and its solutions and then discuss the specific application of domestic energy price setting in Section 2.5.

The regulator is controlling a stochastic process $R = (R_t : t \geq 0)$. We assume the process follows a Geometric Brownian Motion (GBM) structure:

$$\frac{dR}{R} = \mu dt + \sigma dW, \quad \mu \in (-\infty, \infty), \sigma \in (0, \infty), \tag{2.3.1}$$

where $W = (W_t : t \geq 0)$ is a standard Brownian motion on a filtered probability space $(\Omega, \mathscr{F}, P)$. The process $R = (R_t : t \geq 0)$ is assumed to start from 1 but through the time it can become bigger or smaller than 1.

We assume that when $R$ deviates from 1 (in either direction) the system is out of the desired condition and incurs some deviation costs. The controller's goal is to keep the process $R$ as close as possible to 1, while minimizing the recurring cost of adjusting the process.

$(s,S)$**-Policies** We assume that the regulator is committed to a $(s,S)$-type policy, in which the process is reset to $R = 1$ once it hits either the lower or upper barrier (see Figure 2.1). Several papers (e.g., [45]) have shown that the $(s,S)$-policy is the unique optimal for such impulse control problems. Thus, given a $(s,S)$-type policy as given, we focus on characterizing the optimal values of upper and lower barriers (henceforth, denoted as $U^*$ and $L^*$, respectively)

**Cost structure:** The regulator will choose to adjust the stochastic process only when the process $R$ deviates substantially away from 1. This is determined by two boundary values $L < 1 < U$. When

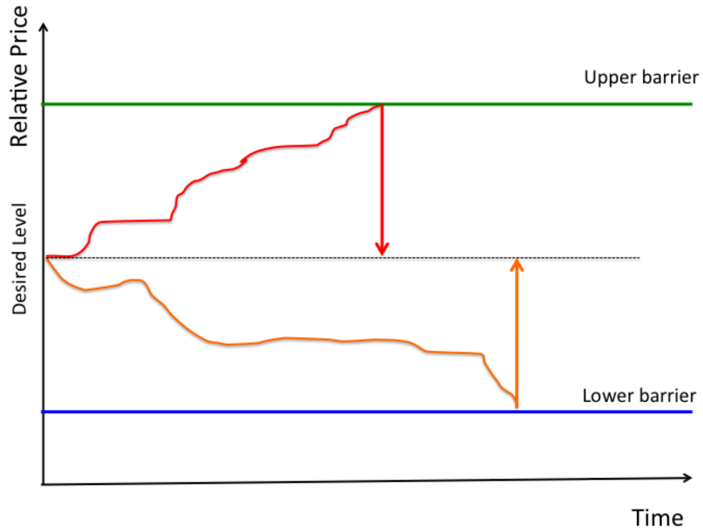


**Figure 2.1:** Conceptual Model of the Problem. The process is reset to one (the desired level) once it hits either the upper or the lower bound.

the process $R$ hits one of the boundaries $L$ or $U$ at time $t$, then the process is reset to its initial level at time $t$. The principle to choose the optimal lower and upper boundaries $L^*$ and $U^*$ is through the following three functions.

The cost function $C(R)$ maps the level of deviation from the equilibrium level to a penalty. A popular choice of $C(x) = c_1(x-1)^2$ is a quadratic function that penalizes positive and negative deviations in a symmetric manner. Since the process $R$ is always positive, when $R_t > 1$ is arbitrarily large the penalty can be arbitrarily large. However, when $R_t < 1$ is small, the penalty cannot be arbitrarily large since $R_t$ is always positive. For this reason, we use the cost function $C(x) = c_1(x-1)^2 + c_2(\frac{1}{x}-1)^2$. The part $(x-1)^2$ penalizes the deviation above the equilibrium price and the part $(\frac{1}{x}-1)^2$ penalizes the deviation below the equilibrium price. The penalty weighting constants $c_1$ and $c_2$ can be chosen differently.

### 2.3.1 Optimization Problem

Since the regulator brings the process back to $R = 1$ after every adjustment, we do not need to solve an infinite horizon problem. Our approach enables us to focus only on a single cycle of adjustment. In this way, we do not need to include a discount factor, which is essential to make sure the sum of infinite series of costs is a bounded function.

The cost function $Z(L,U)$ (for single cycle) is given by:

$$Z(L,U) \equiv \underbrace{\frac{1}{\mathbb{E}[H(L,U)]}\left\{\mathbb{E}\left[\int_0^{H(L,U)}\left[c_1(R_t-1)^2+c_2\left(\frac{1}{R_t}-1\right)^2\right]dt\right]\right\}}_{\text{Cost of Deviation}} \tag{2.3.2}$$

$$+\underbrace{\frac{K_L(L)}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=L}]}+\frac{K_U(U)}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=U}]}}_{\text{Cost of Adjusting}}, \tag{2.3.3}$$

where $H \equiv H(L,U)$ is the first exit time of the GBM $(R_t)$ from the region between lower and upper bounds $(L,U)$. The integral term in (2.3.2) captures the implicit costs of tolerating a deviation from the equilibrium price until price adjustment, whereas the terms in (2.3.3) represent the expected cost of adjusting the process. In (2.3.2), $c_1, c_2 \in (0,\infty)$ are given constants representing the weights for each associated unit deviation cost. Also, the constants $K_L(L), K_U(U) \in (0,\infty)$ in (2.3.3) represent the cost of adjusting $R$ from the lower boundary $L$ and upper boundary $U$, respectively. Our goal is to determine two optimally chosen adjustment barriers $L^*$ and $U^*$, such that the cost function is minimized.

$$\min_{L<1<U} Z(L,U). \tag{2.3.4}$$

The factor $\frac{1}{\mathbb{E}[H(L,U)]}$ in (2.3.2) and the terms $\frac{1}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=L}]}$, $\frac{1}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=U}]}$ in (2.3.3) divide deviation and adjustment costs by the length of the expected adjustment cycle. Thus, we express both components of the cost function as the stream of per unit of time (e.g., equivalent annualized costs if the time unit is one year) cost flow.

## 2.4 Analysis

Our problem can be classified as a special class of optimal stopping time problems if we replace the first hitting time $H(L,U)$ by a general stopping time. This problem has been studied by many researchers. It can be related to a variational inequality or free boundary problem ([3], [36], [43]).

However, as in many stochastic control problems it is usually impossible to find the explicit solution. In our model, we shall use the explicit computations for expectations of the hitting time $H(L,U)$ and associated cost functions to obtain the explicit formula for the objective function. Then we can minimize such a function to obtain the optimal barriers $L^*$ and $U^*$.

We decompose the cost function $Z(L,U)$ into multiple components. First, we shall deal with the cost of deviation terms. Note that since both the upper limit of integration and the integrand are random variables, calculating the cost of deviation terms requires knowledge about the joint distribution of the integral of the instantaneous deviation cost function and the first hitting time. In other words, we find the join probability distribution of $M_1$ and $M_2$, where $M_1 = [\int_0^X R_t^n dt], n \in Z$ and $M_2 = H(L,U)$.

If we expand the squares in the cost of deviation term in $Z(a,b)$, we will encounter the following expectations:

$$\begin{aligned}
&\text{(I) } \mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right], \qquad \text{(II) } \mathbb{E}\left[\int_0^{H(L,U)} R_t^2 dt\right], \text{ (III) } \mathbb{E}\left[H(L,U)\right], \\
&\text{(IV) } \mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t} dt\right], \qquad \text{(V) } \mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t^2} dt\right]. 
\end{aligned} \tag{2.4.1}$$

All the above quantities can be computed explicitly by using some special functions such as Bessel functions. Recall the GBM process in (2.3.1). Letting $\nu \equiv \frac{1}{\sigma^2}\left\{\mu - \frac{1}{2}\sigma^2\right\}$, we see that $R_t = \exp\left\{\sigma^2 \nu t + \sigma W_t\right\}$. Then, the cost function is explicitly given by:

$$\begin{aligned}
Z(L,U) &= \frac{1}{\mathbb{E}[H(L,U)]}\left[c_1 \mathbb{E}\int_0^H R_t^2 dt - 2c_1 \mathbb{E}\int_0^H R_t dt + c_2 \mathbb{E}\int_0^H \frac{1}{R_t^2} dt - 2c_2 \mathbb{E}\int_0^H \frac{1}{R_t} dt\right] \\
&\quad + (c_1 + c_2) + \frac{K_L(L)}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=L}]} + \frac{K_U(U)}{\mathbb{E}[H(L,U)\mathbb{1}_{R(H)=U}]} \\
&= \frac{1}{\text{(III)}}[c_1\text{(II)} - 2c_1\text{(I)} + c_2\text{(V)} - 2c_2\text{(IV)}] + (c_1 + c_2) - \frac{K_L(L)}{\mathscr{C}_1} - \frac{K_U(U)}{\mathscr{C}_2},
\end{aligned}$$

where

$$
\begin{aligned}
&(\mathrm{I}) = -L^{\nu}\mathscr{A}_1 - U^{\nu}\mathscr{A}_2, \quad (\mathrm{II}) = -L^{\nu}\mathscr{B}_1 - U^{\nu}\mathscr{B}_2, \quad (\mathrm{III}) = -\mathscr{C}_1 - \mathscr{C}_2, \\
&(\mathrm{IV}) = -L^{\nu}\mathscr{D}_1 - U^{\nu}\mathscr{D}_2, \quad (\mathrm{V}) = -L^{\nu}\mathscr{E}_1 - U^{\nu}\mathscr{E}_2.
\end{aligned} \tag{2.4.2}
$$

**(I):** Using [4, Formula 3.20.7(a), (b)] with $\alpha = 0$, $\beta = \frac{1}{2}$, and $\gamma^2$ replaced by $\sqrt{2\gamma}$, we have

$$
\begin{aligned}
& \mathbb{E}\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right) \\
= \;& \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right); R_{H(L,U)=L}\right\} + \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right); R_{H(L,U)=U}\right\} \\
= \;& \frac{S_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma}}{\sigma})}{L^{|\nu|-\nu} S_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma})} + \frac{S_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma})}{U^{|\nu|-\nu} S_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma})} \equiv g_1(\gamma),
\end{aligned} \tag{2.4.3}
$$

where the special functions $S_\nu(x,y)$, $I_\nu(x)$, $K_\nu(x)$ are defined as follows:

$$
\begin{cases}
S_\nu(x,y) \equiv (xy)^{-\nu}\left(I_\nu(x)K_\nu(y) - K_\nu(x)I_\nu(y)\right); & (2.4.4) \\
I_\nu(x) \equiv \displaystyle\sum_{k=0}^{\infty} \frac{(\frac{x}{2})^{\nu+2k}}{k!\Gamma(\nu+k+1)}; & (2.4.5) \\
K_\nu(x) \equiv \dfrac{\pi}{2\sin(\nu\pi)}\left(I_{-\nu}(x) - I_\nu(x)\right). & (2.4.6)
\end{cases}
$$

From the equation (.1.1), we see that

$$
\mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right] = -\frac{d}{d\gamma}\bigg|_{\gamma=0} g_1(\gamma).
$$

Thus, computation of the term $\mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right]$ can be given by computing $\frac{d}{d\gamma}\Big|_{\gamma=0} g_1(\gamma)$. To compute this derivative, we first substitute the expression of $K_\nu$ into the expression for $S_\nu$ in (.1.2) to obtain

$$
\begin{aligned}
S_\nu(x,y) &= \frac{\pi}{2\sin(\nu\pi)}(xy)^{-\nu}\Big(I_\nu(x)(I_{-\nu}(y)-I_\nu(y))-(I_{-\nu}(x)-I_\nu(x))I_\nu(y)\Big) \\
&= \frac{\pi}{2\sin(\nu\pi)}(xy)^{-\nu}(I_\nu(x)I_{-\nu}(y)-I_{-\nu}(x)I_\nu(y)) . \qquad (2.4.7)
\end{aligned}
$$

Then, we have

$$
g_1(\gamma) = L^\nu A_1(\gamma) + U^\nu A_2(\gamma) ,
$$

where

$$
A_1(\gamma) \equiv \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})}
$$

and

$$
A_2(\gamma) \equiv \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})} .
$$

To compute the derivative of $g_1(\gamma)$, we need first to compute the derivative of the modified Bessel function $I_\nu$ in (.1.3). We shall use the series expansion to compute the relevant derivatives. First, by the definition of the modified Bessel function $I_\nu$, we have

$$
\begin{aligned}
I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}) &= \sum_{k=0}^{\infty}\frac{(\frac{\sqrt{2\gamma U}}{\sigma})^{2|\nu|+2k}}{k!\Gamma(2|\nu|+k+1)} = (\frac{2\gamma U}{\sigma^2})^{|\nu|}\sum_{k=0}^{\infty}\frac{(\frac{2\gamma U}{\sigma^2})^k}{k!\Gamma(2|\nu|+k+1)} \\
&= (\frac{2\gamma U}{\sigma^2})^{|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(2|\nu|+2)} + o(\gamma^2)\right).
\end{aligned}
$$

Similarly, we have

$$
I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) = (\frac{2\gamma L}{\sigma^2})^{|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)} + \frac{2\gamma L}{\sigma^2\Gamma(2|\nu|+2)} + o(\gamma^2)\right).
$$

Again by the definition of the modified Bessel function of negative index, we have

$$
I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}) = (\frac{2\gamma U}{\sigma^2})^{-|\nu|}\left(\frac{1}{\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(-2|\nu|+2)} + o(\gamma^2)\right).
$$

Thus, we can compute the derivative of $A_1(\gamma)$ and $A_2(\gamma)$ as follows:

$$
\begin{aligned}
\mathscr{A}_1 \quad &\equiv \quad \left.\frac{d}{d\gamma}\right|_{\gamma=0} A_1(\gamma) = \left.\frac{A'_{11}(\gamma)A_{12}(\gamma) - A'_{12}(\gamma)A_{11}(\gamma)}{A^2_{12}(\gamma)}\right|_{\gamma=0} \\
&= \quad \Big\{\frac{2}{\sigma^2(-2|\nu|+1)}(U^{2|\nu|}L^{-|\nu|} - L^{|\nu|} - L^{-|\nu|}U - U^{2|\nu|}L^{-|\nu|+1} + L^{-|\nu|+1} + L^{|\nu|}U) \\
&\quad\quad + \frac{2}{\sigma^2(2|\nu|+1)}(U^{-2|\nu|}L^{|\nu|} - L^{|\nu|}U - L^{-|\nu|} + L^{-|\nu|}U + L^{|\nu|+1} - U^{-2|\nu|}L^{|\nu|+1})\Big\} \\
&\quad\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2, \qquad (2.4.8)
\end{aligned}
$$

and from noticing $A_{12} = A_{22}$, we have

$$
\begin{aligned}
\mathscr{A}_2 \quad &\equiv \quad \left.\frac{d}{d\gamma}\right|_{\gamma=0} A_2(\gamma) = \left.\frac{A'_{21}(\gamma)A_{22}(\gamma) - A'_{22}(\gamma)A_{21}(\gamma)}{A^2_{22}(\gamma)}\right|_{\gamma=0} \\
&= \quad \Big\{\frac{2}{\sigma^2(-2|\nu|+1)}(U^{-|\nu|}L^{2|\nu|} - U^{-|\nu|}L - U^{|\nu|} + U^{|\nu|}L + U^{-|\nu|+1} - U^{-|\nu|+1}L^{2|\nu|}) \\
&\quad\quad + \frac{2}{\sigma^2(2|\nu|+1)}(U^{|\nu|}L^{-2|\nu|} - U^{-|\nu|} - U^{|\nu|}L - U^{|\nu|+1}L^{-2|\nu|} + U^{|\nu|+1} + U^{-|\nu|}L)\Big\} \\
&\quad\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2. \qquad (2.4.9)
\end{aligned}
$$

Combining all of the above computations, we obtain the following explicit expression for the first term (I) in (2.4.1).

**Proposition 2.3.** The term (I) in (2.4.1) is given by the following explicit formula:

$$
\text{(I)} = \mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right] = -L^\nu \mathscr{A}_1 - U^\nu \mathscr{A}_2, \qquad (2.4.10)
$$

where $\mathscr{A}_1$ and $\mathscr{A}_2$ are defined by (.1.7) and (.1.8), respectively. We can further simplify the expression as

$$
\text{(I)} = \frac{2(U - L + U^{-2\nu}(L-1) + L^{-2\nu}(1-U))}{\sigma^2(2\nu+1)(U^{-2\nu} - L^{-2\nu})}.
$$

Similarly we can find the explicit formula for the other terms. And we can find the explicit computation in the appendix.

## 2.5 Application to Energy Pricing Adjustment

In this section we consider the problem of domestic fuel pricing adjustments to provide a showcase for applications of the model, and to estimate some empirically calibrated parameters to observe the behavior of the model.

We use historical time-series data of gasoline prices and also the exchange rate fluctuations of a typical oil-based economy to estimate the drift and volatility of the underlying stochastic processes.

### 2.5.1 Setup

The global price of fuel $P_G = (P_G^t, t \geq 0)$ is given by a geometric Brownian motion (GBM) process:

$$\frac{dP_G}{P_G} = \mu_G dt + \sigma_G dW_G\,, \tag{2.5.1}$$

and the exchange rate $E = (E^t, t \geq 0)$ follows a GBM process:

$$\frac{dE}{E} = \mu_E dt + \sigma_E dW_E\,, \tag{2.5.2}$$

where $W_G$ and $W_E$ are one-dimensional standard Brownian motions. We also assume that the shocks to global price and exchange rate are independent, namely, $\langle dW_G, dW_E \rangle = 0$.

We call the domestic price optimal when the domestic price is equal to international price multiplied by the current exchange rate. The optimal domestic price $P_D = (P_D^t, t \geq 0)$ is simply given by the domestic equivalent of the global price: $P_D = P_G E$. Since both $P_G$ and $E$ are GBM, it follows that $P_D$ is also a GBM process, following the dynamics of

$$\frac{dP_D}{P_D} = \mu_D dt + \sigma_D dW. \tag{2.5.3}$$

Supposing at time 0 the government sets the domestic price at its equilibrium price and defining the equilibrium domestic price at time $t = 0$ by $P_D^* = P_G^0 E_0$, the *gap* price will be the ratio of the administrative price (which is equal to $P_D^*$) to the hypothetical equilibrium price given by:

$$R_t = \frac{P_G^t E_t}{P_D^*}; \tag{2.5.4}$$

Denote

$$\sigma \equiv \sqrt{\sigma_G^2 + \sigma_E^2} \qquad \text{and} \qquad \nu \equiv \frac{1}{\sigma^2}\left\{\mu_G + \mu_E - \frac{1}{2}\left[\sigma_G^2 + \sigma_E^2\right]\right\}.$$

Then

$$R_t = R_0 \exp\left\{\sigma^2 \nu t + \sigma W_t\right\}, \quad t \geq 0, \tag{2.5.5}$$

is a GBM, where $W$ is a Brownian motion. We assume $R_0 = 1$. We can then use the ratio process as the underlying process and apply the results obtained in the previous section. The solution of the model enables us to characterize the optimal barriers for adjusting the prices and also to calculate the expected saving associated with intermittent adjustment of domestic prices, in comparison to a policy of no adjustment or continuous adjustment.

We use empirical data (collected from the publicly available sources of US Energy Information Administrative (EIA) (for fuel prices) and the Central Bank of Iran (for exchange rate)) on $P_E, P_G$ to estimate the set of model parameters $\sigma_E, \sigma_G, \mu_G, \mu_E$ and hence obtain estimates for mean $\mu_D$ and volatility parameter $\sigma_D$ of the $P_D$ process. The negative deviation cost parameters $c_2$ represents the subsidy costs of keeping fuel below the international prices. On the other hand, the positive deviation cost parameter $c_1$ represents the distortion in the economy because of setting fuel prices too high.

Since point estimates for $c_1, c_2, k_L, k_U$ are not available, we solve the problem for a wide range of values for these parameters. We explore the impacts of $\sigma_D$ on the expected optimized costs, the width of the adjustment band $(U^* - L^*)$, and the expected length of the adjustment period $\mathbb{E}[H(L,U)]$. We also examine the impact of unit costs $c_1$, $c_2$ and unit adjustment costs $k_L$, $k_U$ on the optimal barriers $L^*$ and $U^*$, and the expected time of adjustment period $\mathbb{E}[H(L,U)]$.

The baseline parameter values we considered are provided in Table 2.1.

| Parameter | Definition | Baseline Value | Source |
|---|---|---|---|
| $\mu_G$ | Drift of the gasoline price process | 0.002 | Empirical estimation |
| $\sigma_G$ | Volatility of the gasoline price process | 0.04 | Empirical estimation |
| $\mu_E$ | Drift of the exchange rate process | 0.012 | Empirical estimation |
| $\sigma_E$ | Volatility of the exchange rate process | 0.09 | Empirical estimation |
| $\mu_D$ | Drift of the $R$ process | 0.009 | Empirical estimation |
| $\sigma_D$ | Volatility of the $R$ process | 0.1 | Empirical estimation |
| $c_1$ | Cost of deviation | 8 | Assumption |
| $c_2$ | Cost of deviation | 12 | Assumption |
| $k_L$ | Fixed cost of upward adjustment | 100 | Assumption |
| $k_U$ | Fixed cost of downward adjustment | 100 | Assumption |

**Table 2.1** Key Benchmark Values of Parameter

### 2.5.2 Behavior of the Cost Function

Before solving the problem to find the optimal barriers, we look at the overall behavior of the cost function when *arbitrary lower and upper barriers are chosen. For our method to work it is important to observe a smooth and convex objective function, which has a unique minimum. The behavior of the cost function and the expected hitting times are depicted in Figure 2.2. In the next step we use a Matlab optimization solver (using the fminsearch()* function) to find the set of lower and upper barriers, which *minimize* the cost function for a given set of structural parameters.

### 2.5.3 Sensitivity Analysis of the Optimal Solution

After inspecting the overall behavior of the objective function, we examine the response of the optimal solution as well as the associated costs and expected adjustment period to changes in the underlying parameters. One goal of this exercise is to observe the quantitative and qualitative behavior of the model as a way to backtest the validity and the robustness of the solution approach.

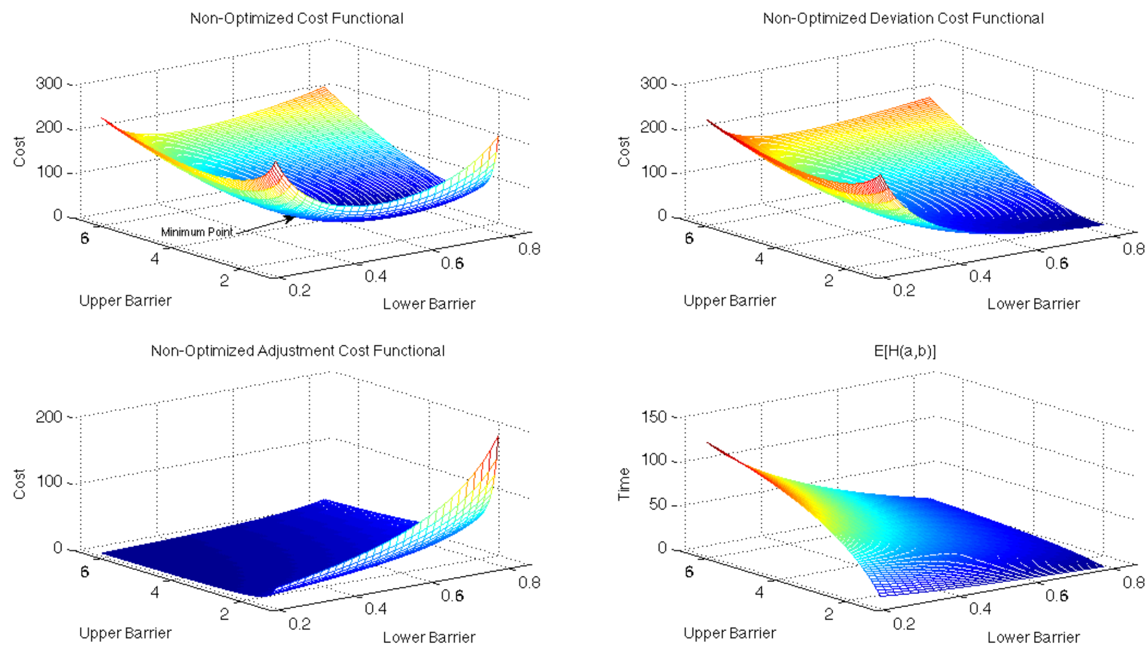


**Figure 2.2:** Graph of the cost function $(Z(L,U))$ for arbitrary (non-optimized) values of lower and upper bound. We observe a well behaved function, which contains a unique minimum. The optimal solution will pick values for the lower and upper bounds to minimize the cost function. As one moves lower and upper bounds closer to one, deviation costs converge to zero; however, adjustment costs become very large. As we move the barrier away from one, the expected heating time also significantly increases.

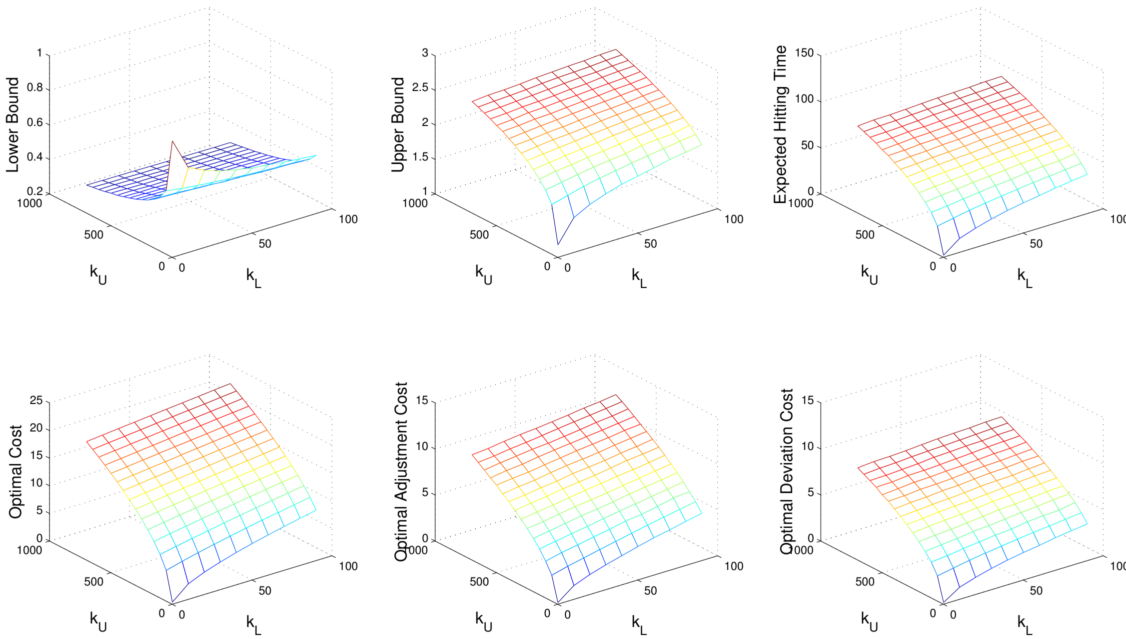


**Figure 2.3:** Impact of Adjustment Costs $(K_L, K_U)$ on the Optimal Behavior. We change the value of fixed costs associated with upward and downward adjustments and observe the response of the optimal solutions. When adjustment costs approach zero (i.e the optimal solution converges to instantaneous control) the barriers collapse to one. The response of optimized total costs to adjustment costs is concave; when the adjustment costs become too large, the problem chooses a boundary far from one. Since the likelihood of hitting a distant barrier is very small the expected adjustment costs behave in a concave fashion.

#### 2.5.3.1 Effect of Adjustment Costs

Figure 2.3 plots the behavior of the solution when the adjustment costs associated with upward and downward corrections are changed. As expected, a higher adjustment cost will further penalize the model for frequent adjustments. Therefore, the optimal solution tends to widen the barriers and reduces the frequency of hitting the barriers (of course at the cost of higher deviation costs).

#### 2.5.3.2 Effect of Volatility

Volatility is a key parameter for optimal adjustment policy. A higher volatility is equivalent to a more turbulent system, which typically requires a more frequent adjustment. In Figure 2.4 we plot the response of key variables to changes of the underlying process volatility.

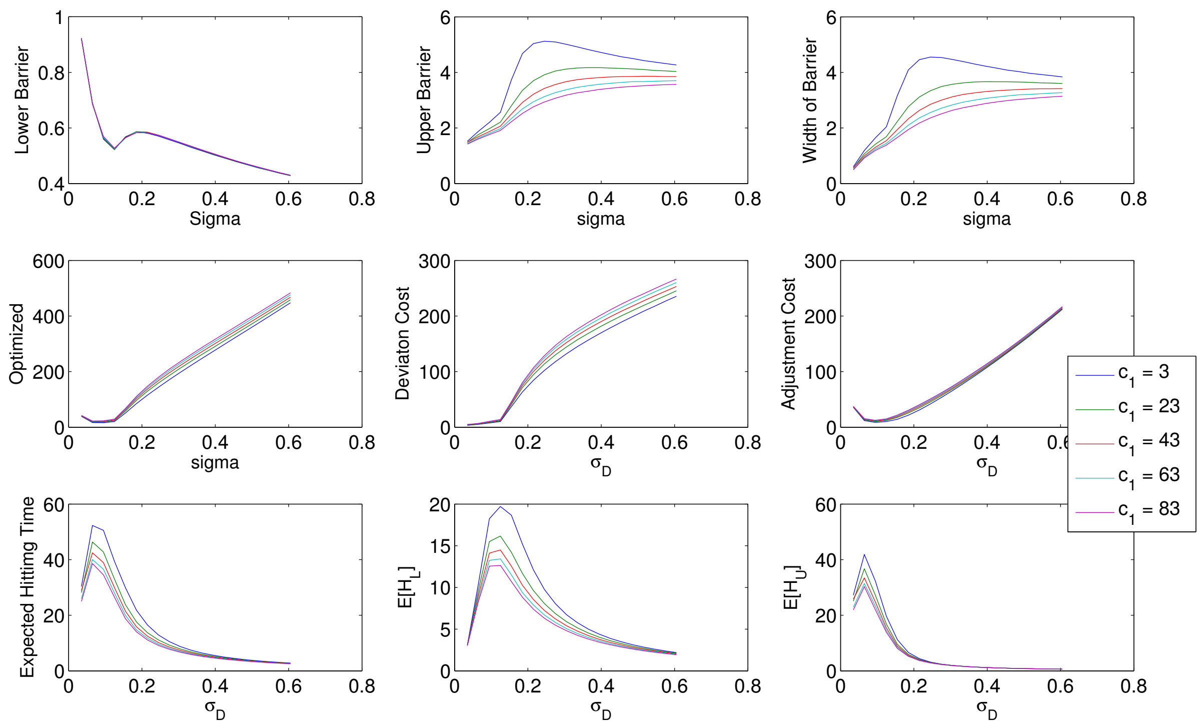


**Figure 2.4:** Effect of Volatility on the Optimal Solution for a Set of Upward Deviation Cost Coefficients. We observe a non-monotonic response of the width of the optimal barriers and the expected hitting times to increased volatility. An increase in the initial levels of volatility causes the expected hitting times to first increase and then decrease and converge toward zero.

**Observation** From the graphs, we find that the optimal barriers $L^*$, $U^*$, the width of the optimal control band, $(U^* - L^*)$, and the expected hitting time, and the minimum cost $Z$ are all non-monotone with respect to the volatility. This is in line with the intuition offered by [12], in which it shown that the probability of hitting barriers behaves non-monotonically as the volatility increases. /

### 2.5.4 Policy Implications

We compare the outcome of our method with a policy, that regulates the process in pre-determined time epochs. We refer to the former as the “optimal" policy and the latter as the “mechanical" policy. As a popular example of the mechanical policy, we consider an annual revision of energy prices and compare it to our $(s,S)$-type adjustment policy.

To estimate the cost of the mechanical policy, we use a Monte Carlo exercise and simulate 1000 paths of the global price process for 360 months (30 years). Deviation and adjustment costs for each path is recorded. The exercise is repeated 1000 times to produce a robust estimate. We assume that the mechanical policy adjusts fuel prices every 12 months (at the beginning of the fiscal year) regardless of the size of the deviation.

Figure 2.5 compares the cost behavior of the mechanical and optimal policies and also reports the gains from implementing an optimal policy. Figure 2.6 provides a more detailed comparison by breaking down the cost function to adjustment and deviation cost components. We notice that, for a certain range of the volatility parameter, the mechanical method entails a lower adjustment cost compared to the optimal one. However, the total cost of the optimal method is always substantially lower than the mechanical method.

## 2.6 Conclusion

We introduced a model of costly price adjustments and formulated it as a stochastic optimal control problem. The key contribution is to offer a new solution approach based on the expected hitting

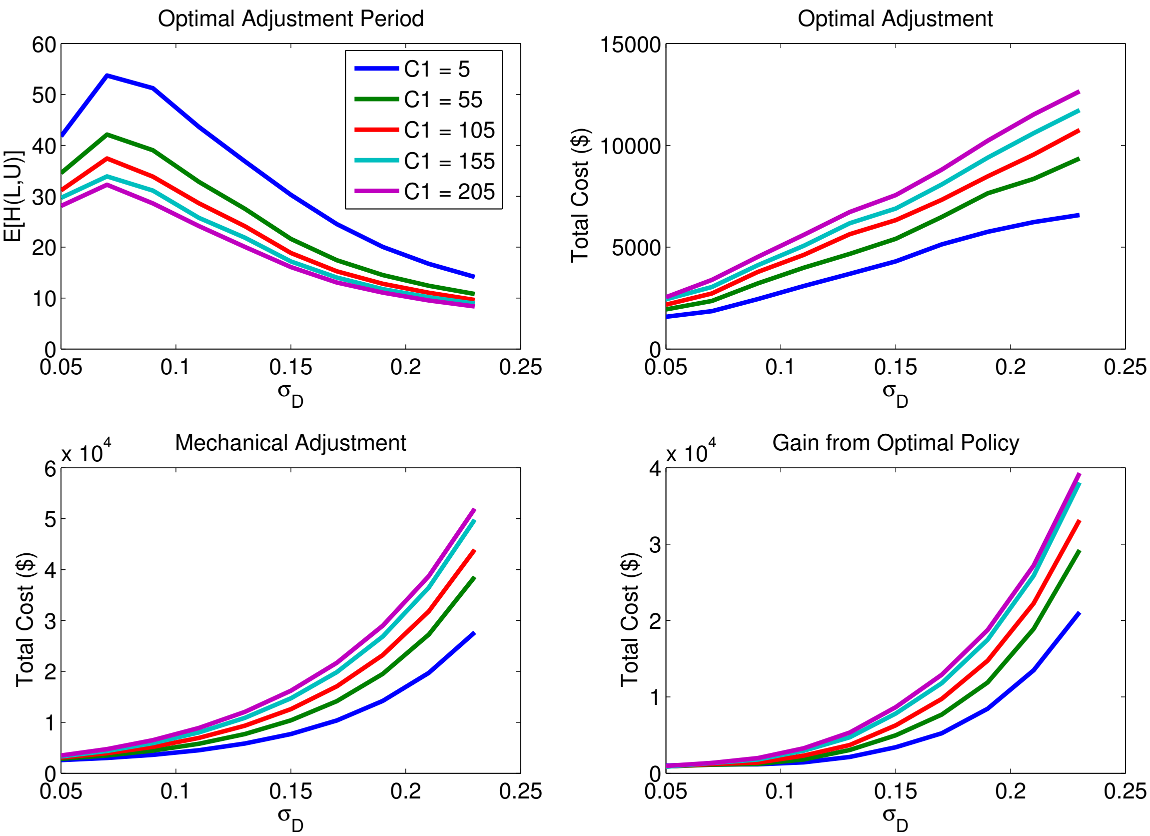


**Figure 2.5:** Gain from Implementing Optimal Policy. We compare the expected cost of a mechanical policy (adjusting every 12 months) with the optimal adjustment solution. When the volatility is very small, the difference between the two policies converges to zero. As the volatility of the underlying process increases, the expected cost of the mechanical policy increases in a convex fashion; whereas, the cost of the optimal solution (which adjusted the location of optimal barriers in response to changes in the volatility) only goes up linearly. The gains from adopting an optimal solution is convex in the volatility of the underlying process.

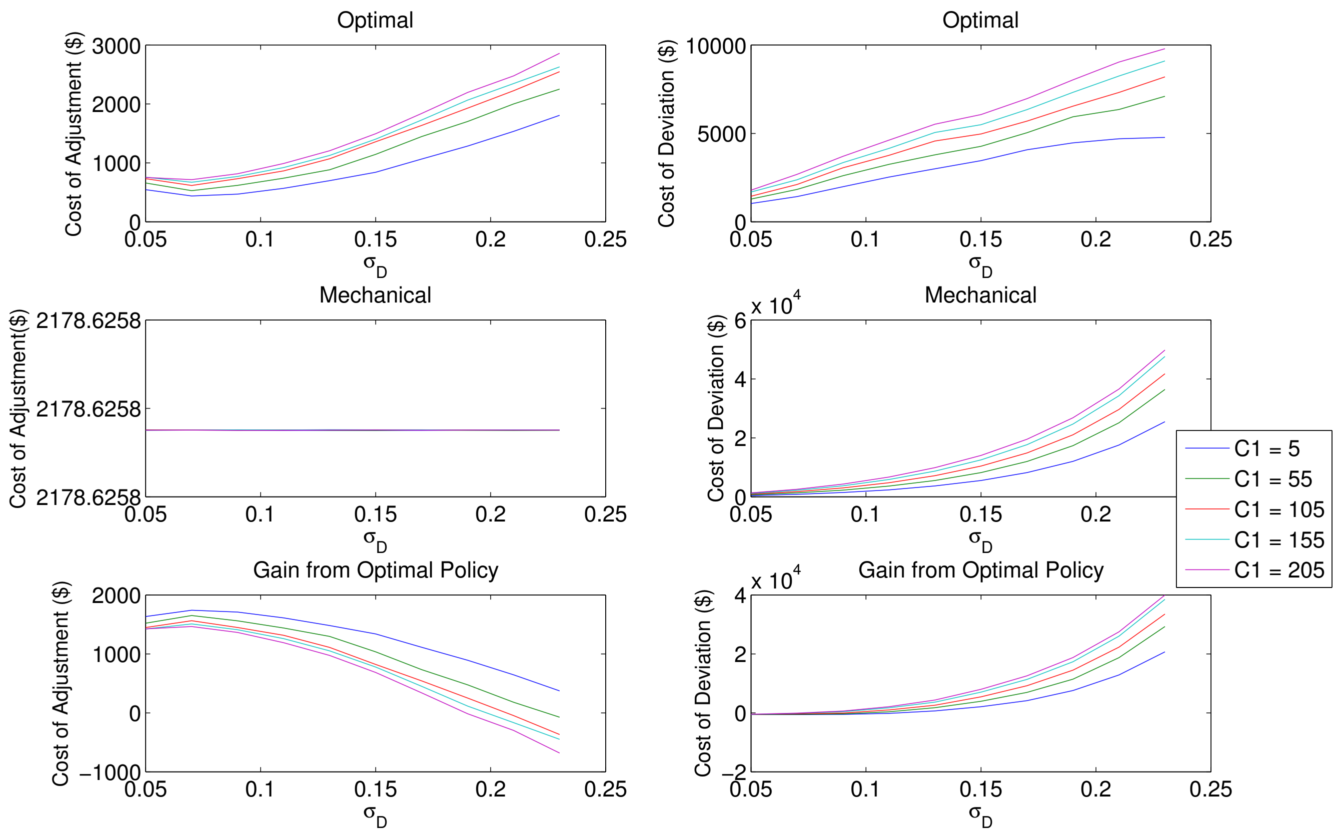


**Figure 2.6:** Decomposition of the Gain from Implementing Optimal Policy. As expected, when the volatility increases the adjustment cost of the optimal solution is larger than the mechanical policy because the latter does not change the frequency of adjusting; however, the gain from a lower deviation cost associated with the optimal policy dominates higher adjustment costs and the optimal policy provides a lower overall cost.

time to the barriers. This approach allows us to derive closed-form expressions of the cost function, which does not rely on solving functional equations or PDE representations.

To demonstrate the robustness of the model it is applied to a real-world case of optimal domestic energy price adjustment. We characterize the optimal policy behavior as a function of underlying parameters and also compare the gains from adopting an optimal policy versus a mechanical policy.

# Chapter 3

# Parameter estimation by implicit sampling

## 3.1 Introduction

In this chapter we describe how to use implicit sampling to find parameters of a PDE equation. We use the Bayesian approach in which the posterior probability density describes the probability of the parameter conditioned on data and compute an empirical estimate of this posterior with implicit sampling. The approach generates independent samples and avoid some issues encountered with Markov Chain Monte Carlo such as the estimation of burn-in and strong correlations among the samples. We describe a new implementation of implicit sampling for parameter estimation problems, the Newton-Krylov-Schwarz optimization method which is scalable compared with the BFGS methods that is used in. We also provide an example involves an elliptic PDE and discussed the global and local Karhunan-Loève expansion expansions by which we can get the finite dimensional approximation of the parameter that we plan to estimate.

## 3.2 Bayesian Framework

Estimating the parameters in a partial differential equation is a main problem in many applications. For example, we have a PDE that describes the subsurface flow where we want to know the subsurface structures from pressure measurements of flow through a porous medium. The uncertain quantity in this problem is the permeability which describes the subsurface structures we are interested in.

Let me introduce the forward model and inverse problem at first.

$$Z = F(\rho) + \varepsilon$$

where $F$ is the forward model, $\rho$ is the parameter, $\varepsilon$ is the random error which is assumed to be Gaussian and $Z$ is the observation data. We want to estimate the parameter $\rho$ and this is a nonlinear inverse problem. In applications, it can be derived from the discretization of PDEs.
Uncertainty may come from the measurement errors, model errors or the uncertainty of the prior information. We assume that all available information we already know before the experiment about the parameter is available and is summarized by the prior probability density function (pdf) $p_0(\theta)$. The Bayesian approach combines the prior information with the likelihood function $p(z|\theta)$ which measures how likely the data would be for a given parameter where $z$ denotes the observations to find the posterior density

$$p(\theta|z) \propto p_0(\theta)p(z|\theta)$$

The posterior pdf defines which parameters of the numerical model are compatible with the data $z$. It incorporates information from both the historical information and the observations. Our goal is to compute the posterior function. We do not need a full description of the posterior, instead, we can estimate and quantify the uncertainty of the parameters by the mean and variance of the posterior density. If the prior and likelihood are Gaussian, then the posterior is also Gaussian. We just need to compute the mean and covariance of $\theta|z$. The posterior mean and covariance are the minimizer and inverse of the Hessian of the negative logarithm of a Gaussian posterior pdf. In nonlinear and non-Gaussian problems, we can compute the posterior mode or the maximum a posterior (MAP) point by minimizing the negative logarithm of the posterior, and use the MAP point as an approximation of the parameter $\theta$. The inverse of the Hessian of the negative logarithm of the posterior can be used to measure the uncertainty of the approximation.

## 3.3 Markov Chain Monte Carlo

In the previous section, we introduced the Bayesian inference. Most of Bayesian inference is concerned with posterior

$$p(\theta|z) \propto p_0(\theta)p(z|\theta)$$

without knowing the constant of proportionality. This leads to the general sampling problem: Suppose we are given a distribution function

$$p(z) = \frac{1}{Z_p}\tilde{p}(z)$$

where $\tilde{p}(z) \geq 0$ is easy to compute but $Z_p$ is hard to compute. So in the Bayesian models where $\tilde{p}(\theta) = p(z|\theta)p(\theta)$ is easy to compute but $Z_p = p(z) = \int_\theta p(\theta)p(z|\theta)d\theta$ can be very difficult or impossible to compute. So the sampling problem is the problem of simulating from $p(z)$ without knowing the constant $Z_p$.

The most common choice for exploring the posterior is Markov chain Monte Carlo (MCMC), which generates a serial of samples to evaluate the statistical information. In this section we will introduce basics of MCMC and some popular algorithms which are widely used in practice. Monte Carlo is a technique for randomly sampling a probability distribution and from the samples that are drawn, we can then estimate the sum or the integral quantity as the mean or variance of the drawn samples. Markov chain is a systematic method for generating a sequence of random variables where the current value is probabilistically dependent on the value of the prior variable. Combining these two methods, Markov Chain Monte Carlo allows random sampling of high-dimensional probability distributions that honors the probabilistic dependence between samples by constructing a Markov Chain that comprise the Monte Carlo samples. MCMC is for estimating a quantity for probability distributions where independent samples from the distribution cannot be drawn easily. Samples are drawn from the probability distribution by constructing a Markov Chain, where the next sample that is drawn from the probability distribution is dependent upon the last sample that was drawn. So although the first sample may be generated from the prior the sequence is constructed so that successive samples are generated from distribution that get closer and closer to the desired posterior. In the following, we will see the design of Metropilis-Hastings MCMC algorithm.

Suppose we want to sample from a distribution $p(x) = \tilde{p}(x)/Z_p$. To do this we first construct a Markov chain as follows. Let $X_t = x$ be the current state. We then perform the following two steps repeatedly:

1. Generate $\mathbf{Y} \sim Q(\cdot|\mathbf{x})$ for some Markov transition matrix $Q$. Let $y$ be the generated value.
2. Set $X_{t+1} = y$ with probability $\alpha(y|x) = \min\{\frac{\tilde{p}(y)}{\tilde{p}(x)} \cdot \frac{Q(x|y)}{Q(y|x)}, 1\}$. Otherwise set $X_{t+1} = x$.

**Claim**: The resulting Markov chain is reversible with stationary distribution $p(x) = \tilde{p}(x)/Z_p$. We can therefore sample from $p(x)$ by running the algorithm until stationary is achieved and then using generated points as our samples.

*Proof.* We simply check that $p(x)$ satisfies the detailed balance equations. We have

$$\begin{aligned}
\underbrace{\alpha(\mathbf{y}|\mathbf{x})Q(\mathbf{y}|\mathbf{x})}_{P(\mathbf{y}|\mathbf{x})}p(\mathbf{x}) &= \min\left\{\frac{p(\mathbf{y})}{p(\mathbf{x})} \cdot \frac{Q(\mathbf{x}|\mathbf{y})}{Q(\mathbf{y}|\mathbf{x})}, 1\right\} Q(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \\
&= \min\{Q(\mathbf{x}|\mathbf{y})p(\mathbf{y}), Q(\mathbf{y}|\mathbf{x})p(\mathbf{x})\} \\
&= \min\left\{1, \frac{p(\mathbf{x})}{p(\mathbf{y})} \cdot \frac{Q(\mathbf{y}|\mathbf{x})}{Q(\mathbf{x}|\mathbf{y})}\right\} Q(\mathbf{x}|\mathbf{y})p(\mathbf{y}) \\
&= \underbrace{\alpha(\mathbf{x}|\mathbf{y})Q(\mathbf{x}|\mathbf{y})}_{P(\mathbf{x}|\mathbf{y})}p(\mathbf{y})
\end{aligned}$$

So the Markov chain is reversible and $p$ is therefore the stationary distribution of the Markov chain since we have

$$\sum_{\mathbf{x}} P(\mathbf{y}|\mathbf{x})p(\mathbf{x}) = \sum_{\mathbf{x}} P(\mathbf{x}|\mathbf{y})p(\mathbf{y}) = p(\mathbf{y})$$

□

However, there are still some practical issues of choosing the appropriate proposal distribution $Q(\cdot|\cdot)$ since it influences how much time is required to reach stationarity and it is difficult to provide a theoretical answer of when the stationarity is achieved.

## 3.4 Implicit Sampling

Implicit sampling method acts as a special formulation of importance sampling to improve sample performance by providing an important function based on optimization. The importance sampling method is a popular Monte Carlo method which generated independent samples without any Gaussian assumption. The idea is to draw samples from another easy-sampling importance function with a weight of each sample instead of drawing samples from the target distribution itself, which is usually difficult to explore directly. But if the variance of the weights is large, the effective sample size may be very small and the number of samples required can increase quickly with the dimension of problem. Suppose we wish to estimate an m-dimensional parameter vector $\theta$ from data. One can present the posterior by $M$ weighted samples, The samples $\theta_j,\ j = 1,\ldots,M$ are obtained from an importance function $\pi(\theta)$, and the $j-th$ sample is assigned the weight

$$\omega_j \propto \frac{p(\theta_j)p(z|\theta_j)}{\pi(\theta_j)}$$

The weights describes how likely the samples is in view of the posterior. The weighted samples $\{\theta_j, \omega_j\}$ form an empirical estimate of $p(\theta|z)$ so that for a smooth function $u$,

$$E_M(u) = \sum_{j=0}^{M} u(\theta_j)\hat{\omega}_j$$

where $\hat{\omega}_j = \omega_j / \sum_{j=0}^{M} \omega_j$, converges almost surly to the expected value of $u$ as $M \to \infty$.
Choosing important function appropriately is crucial to the implementation of the importance sample. A good choice of the importance density should have the following properties:

1. should be easy to simulate.

2. should be close to the posterior density.

For example, if we choose the importance function be the prior, then the weights are proportional to the likelihood. There are two scenarios in which the samples we draw from the prior have a low posterior probability so that the estimate of the posterior is inaccurate. The prior may have probability mass in a small region of the space in which the likelihood is small. And in second scenario, the prior may be broad while the likelihood is sharply peaked. Poor choices of the importance function may lead to huge amount of computational waste on samples that contribute little or even nothing (e.g. q and p are singular to each other) to the posterior density. And the number of samples required can increase dramatically with the dimension. So the importance sampling algorithm cannot be applied to the high-dimensional problem.

The implicit sampling method provides a general framework to construct the importance function for the importance sampling method, which has a significant overlap with the posterior density. Let us denote the negative logarithm of the posterior density by $F(\theta)$

$$F(\theta) = -\log(p(\theta)p(z|\theta))$$

The first step of implicit sampling is to locate the high region of posterior density by minimizing $F(\theta)$. we denote $\phi$ as the minimum value of $F(\theta)$ and $\mu$ as the minimizer and it is the same as finding the MAP (maximum a posterior) point. Our goal is to construct an importance function that assigns high probability to generate samples around the MAP point. We first pick up a reference random variable $\xi$ with pdf $g(\xi) \propto \exp(-G(\xi))$ and define $\phi_G = \min G$. The samples for $\theta$ are generated by first first drawing samples from $\xi$ and then solving the algebraic equation

$$F(\theta) - \phi_F = G(\xi) - \phi_G \tag{3.4.1}$$

By a change of variables, we can derive that the associated weights for the samples are given by $w_j \propto J(\theta)$ where $J$ is the Jacobian of the map from $\theta$ to $\xi$ provided that the map $\xi \to x$ is one-to-one and onto. Since the samples of $\xi$ are independent and close to the MAP point, the samples of $\theta$ will also be independent to each other and forced to lie near the MAP point $\mu$.

### 3.4.1 Solving the implicit equation

We want to solve for a Gaussian $\xi$ with mean 0 and covariance matrix $H^{-1}$, where $H$ is the Hessian of the function $F$ at the minimum. With this $\xi$, equation 3.4.1 becomes

$$F(\theta) - \phi_F = \frac{1}{2}\xi^T H \xi \tag{3.4.2}$$

Two strategies [`?linearmap`] for solving the equation are popular: random and linear map.

**Random map**. We seek the solution of 3.4.1 in the form

$$\theta = \mu + \lambda(\xi)\xi \tag{3.4.3}$$

Here $\lambda$ can be computed by substituting into and solving for the scalar $\lambda(\xi)$ with Newton's method. A formula for the Jacobian of the random map was derived in [??????]:

$$w \propto |J(\xi)| = \left| \lambda^{m-1} \frac{\xi^T H \xi}{\nabla_\theta F \cdot \xi} \right| \tag{3.4.4}$$

where $m$ is the number of nonzero eigenvalues of $H$.

**Linear map**. We first expand the $F(\theta)$ to the second order:

$$F(\theta) \approx \phi_F + \frac{1}{2}(\theta - \mu)^T H (\theta - \mu) = F_0(\theta)$$

where $H$ is the Hessian at $\mu$. We solve the equation

$$F_0(\theta) - \phi = \frac{1}{2}\xi^T H \xi \tag{3.4.5}$$

We solve the equation

$$F_0(\theta) - \phi_F = \frac{1}{2}\xi^T H \xi \tag{3.4.6}$$

We simply shift $\xi$ by the mode: $\theta = \mu + \xi$. The bias created by solving (3.4.6) instead of (3.4.1) can be removed by the weights

$$w \propto \exp\left(F_0(\theta) - F(\theta)\right)$$

### 3.4.2 Optimization

The first step in implicit sampling is to find the MAP point by minimizing $F$. This can be done numerically by Newton, quasi-Newton methods. We will introduce how to use BFGS and Newton-Krylov-Schwarz optimization in the example below. We also compare these two methods and find that the Newton-Krylov-Schwarz method is scalable which also enables us to do the parallel computing.

## 3.5 Application to subsruface flow problem

As a test of the performance of various sampling algorithms introduced in this chapter, we apply these methods to a subsurface flow problem, where we estimate subsurface subsurface structures from pressure measurements of flow through a porous medium. (see [9]) We consider the elliptic problem

$$-\nabla \cdot (\rho \nabla u) = f \tag{3.5.1}$$

on a domain $\Omega$, with Neumann boundary conditions, where $k$ is the permeability and describes the subsurface structures we are interested in, $\nabla u$ is the pressure gradient across the porous medium and $f$ represents externally prescribed inward or outward flow rates.

The uncertain quantity in this problem is the permeability parameter $\rho$, and we assume for each $\rho$, a unique solution of (3.5.1) exists. We want to estimate the $\rho$ on the basis of noisy measurements of the pressure at $n$ locations so that we have the equation

$$z = f(\rho(u), x, y) + r \tag{3.5.2}$$

We consider a 2D problem on the domain $\Omega = [0,1] \times [0,1]$ and discritize (3.5.1) with a piece-wise linear finite element method on a uniform $(N+1) \times (N+1)$ mesh of triangular elements.

$$Au = f$$

where $A$ is a $(N+1)^2 \times (N+1)^2$ matrix and $u$ and $f$ are $(N+1)^2$ vectors. We could first decompose the domain into smaller sundomains and then solve the subdomain interface problem. Then solving the PDE equivalents to solving the linear systerm.

In the numerical experiments, the data equation is

$$z = u + r$$

So we assume we have observations on every point and the data are perturbed with a Gaussian random variable $r \sim \mathcal{N}(0, R)$ where $R$ is a diagonal covariance matrix. We know the dimension of $\rho$ is mesh dependent. So if the mesh is fine, the dimension of $\rho$ is large. We will introduce two methods that help us reduce the dimension of $\rho$.

### 3.5.1 The log-normal prior, discretization, and dimensional reduction

We assume that the prior density function is log-normal with exponential covariance function

$$K(x_1, x_2, y_1, y_2) = \exp\left(-\frac{(x_1 - x_2)^2}{l_x^2} - \frac{(y_1 - y_2)^2}{l_y^2}\right) \tag{3.5.3}$$

where $(x_1, y_1)$ and $(x2, y2)$ are the points in the domain and $l_x$, $l_y$ are the correlation length. Then the elements of the covariance matrix $\Sigma$ is

$$\Sigma(i,j) = K\left(x_i, x_j, y_i, y_j\right), \quad i,j = 1,\ldots,N$$

where $N$ is the number of grid points in each direction. Then we perform a dimension reduction by Karhunen–Loève (KL) expansions [11] and use the resulting low rank approximation of the covariance matrix $\Sigma$ for all subsequent computations. We obtain the low-rank approximation for the covariance matrix on the grid from the SVD of the covariance:

$$\hat{\Sigma} = V^T \Lambda V$$

where $\Lambda$ is a diagonal matrix whose diagonal elements are the m largest eigenvalues of $\Sigma$ and $V$ is an $m \times N$ matrix whose columns are the corresponding eigenvectors.
Thus, in the reduced dimension on the grid, the prior is

$$\hat{\rho} \sim \log \mathcal{N}(\hat{\mu}, \hat{\Sigma})$$

With the linear change of the variables

$$\theta = V^T \Lambda^{-0.5} \hat{\rho}$$

the prior for the variable $\theta$ is

$$p(\theta) = N(\mu, I_m)$$

where $\mu = V^T \Lambda^{-0.5} \hat{\mu}$ Then $\rho$ can be written as $\rho = \exp(U\theta)$ where $U = V\Lambda^{0.5}$ and we will do the computation in the reduced coordinate $\theta$ instead of the original parameter $\rho$. This will reduce the effective dimension for the parameter $\rho$ from $N^2$ to $m$. **Local dimension reduction:**

We also use a local dimension reduction method [2] which is suitable to the parallel computation of the KL decomposition. It uses a domain decomposition approach to conventinently distribute the computational load among several processors and recast in a reduced eigenvalue problem. The computational domain is partitioned into smaller non-overlapping subdomains, over which indepedent local KL decompositions are performed to generate local bases which are subsequently used to discretize the global modes over the entire domain. Later We can see the number of the iterations are almost the same for the optimization no matter which KL expansion is used for the vector $\rho$.

### 3.5.2 Optimization

Implicit sampling requires the minimization of $F$. In this section, we introduce solve the optimization problem using the Newton-Krylov-Schwarz method and BFGS method. The optimization problem takes the form

$$J(\rho) = \frac{1}{2}\int_{\Omega}(u-z)^2 dx + \frac{1}{2}\int_{\Omega}|\rho|^2 dx \tag{3.5.4}$$

#### 3.5.2.1 BFGS method

Under the discretization, the cost function could be written as

$$\begin{aligned} J(\rho) &= \frac{1}{2}\int \left(\sum (u_i - z_i)\,\phi_i\right)^2 dx + \frac{1}{2}\int |\rho_i|^2 dx \\ &= \frac{1}{2}(u-z)^T M (u-z) + \frac{a}{2}\rho^T \rho \end{aligned}$$

where $M = \left(\int \phi_i^T \phi_j\right)_{i,j}$, $a$ is the area of each element, $\rho = \exp(U\theta)$. Now, let us deal with the gradient of the cost function $\nabla J$.

Since $Au = f$, where $A_{i,j} = \sum_\ell \rho_k \int \nabla\phi_j \nabla\phi_i$

$$\frac{\partial A}{\partial \rho}u + A\frac{\partial u}{\partial \rho} = 0$$

i.e

$$\frac{\partial u}{\partial \rho} = -A^{-1}\frac{\partial A}{\partial \rho}u$$

Then

$$\begin{aligned}\frac{\partial J}{\partial \rho_j} &= (u-z)^T M\frac{\partial u}{\partial \rho_j} + a_j * \rho_j \\ &= -(u-z)^T MA^{-1}\frac{\partial A}{\partial \rho_j}u + a_j * \rho_j \\ &= -\left(A^{-T}M^T(u-z)\right)^T \frac{\partial A}{\partial \rho_j}u + a_j * \rho_j\end{aligned}$$

where $a_j$ is the area of the element corresponding to $\rho_j$. Thus

$$\nabla J_\rho(\rho) = -\left(A^{-T}M^T(u-z)\right)^T \begin{bmatrix} \frac{\partial A}{\partial \rho_1}u \\ \vdots \\ \frac{\partial A}{\partial \rho_{N^2}}u \end{bmatrix} + a_j * \rho$$

Since $\rho = \exp(U\theta)$, we finally have

$$\nabla_\theta J(\theta) = -U^T * diag(\rho)\left(A^{-T}M^T(u-z)\right)^T \begin{bmatrix} \frac{\partial A}{\partial \theta_1}u \\ \vdots \\ \frac{\partial A}{\partial \theta_m}u \end{bmatrix} + a_j * U * diag(\rho) * \rho$$

The gradient is used in BFGS method with a cubic interpolation line search (see [32], [30]).

#### 3.5.2.2 Newton–Krylov-Schwarz method

Here, we introduce another method to solve the optimization problem [47]. Instead of solving the constraint optimization problems, we turn to solving the saddle point problems associated with the Lagrangian functional $\mathscr{L}$ :

$$\mathscr{L}(\rho,u,\lambda) = \frac{1}{2}\int_\Omega (u-z)^2 dx - \int_\Omega (\nabla\cdot\rho\nabla u + f)\lambda dx + \frac{1}{2}\int_\Omega |\rho|^2 dx$$

Find $(\rho, u, \lambda)$ such that

$$\begin{cases} \nabla_\rho \mathscr{L} = 0 \\ \nabla_u \mathscr{L} = 0 \\ \nabla_\lambda \mathscr{L} = 0 \end{cases}$$

Then we can get

$$\begin{cases} F^{(\rho)} \equiv \rho + \nabla u \cdot \nabla \lambda = 0 \\ F^{(u)} \equiv -\nabla \cdot (\rho \nabla \lambda) + (u - z) = 0 \\ F^{(\lambda)} \equiv -\nabla \cdot (\rho \nabla u) - f = 0 \end{cases}$$

with boundary conditions

$$\begin{cases} u = 0 \\ \lambda = 0 \\ \frac{\partial \rho}{\partial n} = 0 \end{cases}$$

By Galerkin's method, we can discretize the equations as

$$F_j^{(\rho)} \equiv \rho_{ne} \cdot a_j + \sum_{i=1}^{n} \sum_{k=1}^{n} u_i \lambda_k \int \nabla \phi_i \nabla \phi_k = 0$$

$$F_j^{(u)} \equiv \rho_{ne} \sum_{k=1}^{n} \lambda_k \int \nabla \phi_k \nabla \phi_j + \sum_{i=1}^{n} (u_i - z_i) \int \phi_i \phi_j = 0$$

$$F_j^{(\lambda)} \equiv \rho_{ne} \sum_{k=1}^{n} u_k \int \nabla \phi_k \nabla \phi_j - \int f \phi_j = 0$$

where $a_j$ is the area of the jth element, $\rho_{ne}$ is the value of $\rho$ in the element which corresponds to the $u$ or $\lambda$, $\phi_i$ is the basic functions.

**Ordering of unknowns:**

We use the so-called fully coupled ordering[47], by which we mean that all three variables defined at the same mesh point are always together throughout the calculations. Since here we use the piece-wise linear function and KL expansion for $\rho$, we cannot use this coupled ordering anymore. Here, the unknowns are ordered in the order $\theta$, $u_{ij}$, $\lambda_{ij}$, that is

$$U = \left(\theta_1, \theta_2, \dots \theta_N, \lambda_{11}, u_{11}, \lambda_{21}, u_{21}, \dots, \lambda_{n_x n_y}, u_{n_x n_y},\right)^T$$

And we order the functions in exactly the same order

$$F = \left(F_{11}^{(\theta)}, F_{12}^{(\theta)}, \dots F_{1N}^{(\theta)}, F_{11}^{(u)}, F_{11}^{(\lambda)}, F_{21}^{(u)}, F_{21}^{(\lambda)}, \dots, F_{nxny}^{(u)}, F_{n_x n_y}^{(\lambda)}\right)^T = 0$$

The Newton–Krylov–Schwarz (NKS) methods are a family of general-purpose parallel algorithms for solving systems of nonlinear NKS and it has three main components: (i) an inexact Newton method for the nonlinear system; (ii) restarted GMRES for the Jacobian systems; and (iii) an additive Schwarz type preconditioner. Newton iterations are as follows:

$$U_{k+1} = U_k - \lambda_k J\left(U_k\right)^{-1} F\left(U_k\right), k = 0, 1, \dots,$$

where $U_0$ is an ininitial approximation to the solution, $J\left(U_k\right) = F'\left(U_k\right)$ is the Jacobian at $U_k$ , and $\lambda_k$ is the steplength determined by a linesearch procedure $[9; 10]$. Here we do not solve the Jacobian systems exactly, the accuracy of the Jacobian solver is determined by

$$\left\|F\left(U_k\right) + J\left(U_k\right) s_k\right\| \leq \eta_k \left\|F\left(U_k\right)\right\|.$$

where $\eta_k \in [0, 1)$. The algorithm can be described as follows:

(1) Inexactly solve the linear system $J\left(U_k\right) s_k = -F\left(U_k\right)$ for $s_k$ using a preconditioned GMRES(30).

(2) Perform a full Newton step with $\lambda_0 = 1$ in the direction $s_k$ .

(3) If the full Newton step is unacceptable, we backtrack using the cubic back-tracking procedure until a new $\lambda$ is obtained that makes $U_+ + \lambda_k s_k$ an tracking procedure until a new $\lambda$ is obtained that makes $U_+ = U_k + \lambda_k s_k$ an acceptable step. (4) Set $U_{k+1} = U_+$ and return to step 1 unless a stopping

condition has been met. In step 1 above the vector $s_k$ is obtained by approximately solving the right preconditioned Jacobian system

$$J\left(U_k\right) M_k^{-1} s_k' = -F\left(U_k\right)$$

where $M_k^{-1}$ is a one-level additive Schwart preconditioner and $s_k = M_k^{-1} s_k'$

(4) Set $U_{k+1} = U_+$ and return to step 1 unless a stopping condition has been met.

**One-level additive Schwarz preconditioning**

We first partion the domain into nonverlapping ing subdomains $\Omega_l, l = 1, \ldots, N,$. In order to obtain an overlaping decomposition of the domain, we extend each subdomain $\Omega_t$ to a larger region $\Omega_l'$, that is, $\Omega_l \subset \Omega_l'$. On each extended subdomain $\Omega_l'$, we construct a subdomain preconditioner $B_l$ which is the discretization of the Frechet derivative taken at the current iteration,

$$J = \begin{pmatrix} \frac{\partial F^{(\theta)}}{\partial \theta} & \frac{\partial F^{(\theta)}}{\partial \lambda} & \frac{\partial F^{(\theta)}}{\partial u} \\ \frac{\partial F^{(u)}}{\partial \theta} & \frac{\partial F^{(u)}}{\partial \lambda} & \frac{\partial F^{(u)}}{\partial u} \\ \frac{\partial F^{(2)}}{\partial \theta} & \frac{\partial F^{(\lambda)}}{\partial \lambda} & \frac{\partial F^{(\lambda)}}{\partial u} \end{pmatrix}$$

Since $\rho = \exp(U * \theta)$, $F_\theta = \left(\frac{\partial \rho}{\partial \theta}\right)^T \cdot F_\rho$ where

$$\frac{\partial \rho}{\partial \theta} = \begin{pmatrix} \frac{\partial \rho_1}{\partial \theta_1} & \frac{\partial \rho_1}{\partial \theta_2} & \cdots & \frac{\partial \rho_2}{\partial \theta_N} \\ \frac{\partial \rho_2}{\partial \theta_1} & \frac{\partial \rho_2}{\partial \theta_2} & \cdots & \frac{\partial \rho_2}{\partial \theta_N} \\ \cdots & & & \\ \frac{\partial \rho_n}{\partial \theta_1} & \frac{\partial \rho_n}{\partial \theta_2} & \cdots & \frac{\partial \rho_n}{\partial \theta_N} \end{pmatrix}$$

Suppose

$$U = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nN} \end{pmatrix}$$

then

$$\frac{\partial\rho}{\partial\theta}=\begin{pmatrix} a_{11}\rho_1 & a_{12}\rho_1 & \cdots & a_{1N}\rho_1 \\ a_{21}\rho_2 & a_{22}\rho_2 & \cdots & a2N\rho_2 \\ \cdots & & & \\ a_{n1}\rho_n & a_{n2}\rho_n & \cdots & a_{nN}\rho_n \end{pmatrix}$$

$$\frac{\partial F_\theta}{\partial u}=(\frac{\partial\rho}{\partial\theta})^T\cdot\frac{\partial F_\rho}{\partial u}$$

$$\frac{\partial F_\theta}{\partial\lambda}=(\frac{\partial\rho}{\partial\lambda})^T\cdot\frac{\partial F_\rho}{\partial\lambda}$$

If we write $F_\theta$ explicitly, we have

$$\frac{\partial F_\theta}{\partial\theta}=(\frac{\partial\rho}{\partial\theta})^T\cdot\frac{\partial F_\rho}{\partial\rho}\cdot\frac{\partial\rho}{\partial\theta}+U^T*repmat(F_\rho,1,N)^T*\frac{\partial\rho}{\partial\theta}$$

$$\frac{\partial F_\lambda}{\partial\theta}=(\frac{\partial F_\lambda}{\partial\rho})^T\cdot\frac{\partial\rho}{\partial\theta}$$

$$\frac{\partial F_u}{\partial\theta}=(\frac{\partial F_u}{\partial\rho})^T\cdot\frac{\partial\rho}{\partial\theta}$$

In the test runs, we stop the Newton iteration if the following condition is satisfied

$$\|F(U_k)\|\leq\max\left\{10^{-6}\|F(U_0)\|,10^{-10}\right\}$$

For the Jacobian solver, the GMRES iteration is stopped if

$$\|F(U_k)+J(U_k)s\|\leq\max\left\{10^{-6}\|F(U_k)\|,10^{-10}\right\}$$

Once the minimization is completed, we could generate the samples using either the linear map or random map methods we described above to generate the samples.

### 3.5.3 Test case

We next describe our test case with the observation function

$$z(x,y) = \sin(\pi x)\sin(\pi y)$$

**Test**. we take $\Omega = (0,1) \times (0,1)$, and the right side $f$ is constructed such that $\rho = 1 + 6x^2 y(1-y)$ is the elliptic coefficient to be identified.

#### 3.5.3.1 Results

We can see as the dimension of $\theta$ getting larger, the number of iterations that required by BFGS method is increasing. For the Newton-Krylov-Schwarz method, as the dimension of $\theta$ increasing, the number of iterations almost stay the same. So we can see the Newton-Krylov method is more scalable and it can be easily implemented in parallel. Also the number of iterations of the local KL expansion and global KL expansion are almost the same. But the local KL expansion enables the efficient computation of a possibly large number of dominant KL modes.

| $\theta$ | 5 | 8 | 10 |
|---|---|---|---|
| BFGS | 39 | 63 | 75 |
| Function evaluation | 86 | 136 | 164 |

**Table 3.1** NVM=4, nvm=8, Global KL expansion, BFGS method

| $\theta$ | 5 | 8 | 10 |
|---|---|---|---|
| Newton | 5 | 6 | 6 |
| GMRES(200) | 44.4 | 71.67 | 75.83 |

**Table 3.2** NVM=4, nvm=8, Global KL expansion, Newton-Krylov method

| $\theta$ | 5 | 8 | 10 |
|---|---|---|---|
| Newton | 5 | 6 | 6 |
| GMRES(200) | 46.2 | 73.3 | 72.5 |

**Table 3.3** NVM=4, nvm=8, Local KL expansion, Newton-Krylov method

# References


[1] Hall A, *Generalized method of moments.*, Oxford University Press, Oxford. (2005).

[2] Olivier P. Andres A.C. Paul Mycek, *Parallel domain decomposition strategies for stochastic elliptic equations.*, Society for Industrial and Applied Mathematics (2017).

[3] Alain Bensoussan and J-L Lions, *Applications of variational inequalities in stochastic control*, Vol. 12, Elsevier, 2011.

[4] Andrei N Borodin and Paavo Salminen, *Handbook of Brownian motion-facts and formulae*, Birkhäuser, 2012.

[5] Davis R.A. Brockwell P. J. and Yang, *Estimation for non-negative lévy-driven ornstein-uhlenbeck processes*, J Appl Probab. (2007).

[6] Abel Cadenillas and Fernando Zapatero, *Optimal central bank intervention in the foreign exchange market*, Journal of Economic Theory (1999).

[7] Heathcote CR, *The integrated squared error estimation of parameters.*, Biometrika. (1977).

[8] Francisco Javier Arze del Granado, David Coady, and Robert Gillingham, *The unequal benefits of fuel subsidies: A review of evidence for developing countries*, World Development **40** (2012), no. 11, 2234–2248.

[9] A.C. Reynolds D.S. Oliver and N. Liu., *Inverse theory for petroleum reservoir characterization and history matching.*, Cambridge University Press. (2008).

[10] V. Fasen, *Statistical estimation of multivariate ornstein-uhlenbeck processes and applications to co-integration.*, J. Econometrics (2013).

[11] R. Ghanem and P. Spanos, *Stochastic finite elements: A spectral approach*, Dover. (2003).

[12] Hamed Ghoddusi and Arash Fahim, *Volatility can be detrimental to option values!*, Economics Letters (2016).

[13] Masuda H., *Approximate self-weighted lad estimation of discretely observed ergodic ornstein-uhlenbeck processes*, Electronic Journal of Statistics. (2010).

[14] Hodrick R Hansen LP, *Forward exchange rates as optimal predictors of future spot rates.*, J Polit Econ. (1980).

[15] SingletonK Hansen LP, *Generalized instrumental variable estimation of nonlinear rational expectations models.*, Econometrica. (1982).

[16] Lee C. Lee M. Hu Y. and J. Song, *Parameter estimation for re ected ornstein-uhlenbeck processes with discrete observations.*, Stat. Inference Stoch. Process. (2015).

[17] Y. Hu, *Least squares estimator for ornstein-uhlenbeck processes driven by $\alpha$-stable motions.*, Stochastic Process Appl. (2009).

[18] Y. Hu and H. Long, *Parameter estimation for ornstein-uhlenbeck processes driven by $\alpha$-stable l évy motions.*, Communications on Stochastic Analysis. (2007).

[19] Y. Hu and J. Song, *Parameter estimation for fractional ornstein-uhlenbeck processes with discrete observations.*, Malliavin Calculus and Stochastic Analysis, (2013).

[20] Rosinski J, *Series representations of lévy processes from the perspective of point processes.*, In: Barndorff-Nielsen OE,Mikosch T, Resnick S (eds) Lévy processes—theory and applications.Birkhäuser, Boston. (2002).

[21] Yu J, *Empirical characteristic function estimation and its applications.*, Econom Rev. (2004).

[22] A. Janicki and A Weron, *Simulation and chaotic behavior of -stable stochastic pro- cesses*, Marcel Dekker (1994).

[23] Yu J Knight JL, *Empirical characteristic function in time series estimation.*, Econom Theory. (2002).

[24] Masami Kojima, *Petroleum product pricing and complementary policies: Experience of 65 developing countries since 2009*, World Bank Policy Research Working Paper **6396** (2013).

[25] Ma C. Long H. and Y. Shimizu, *Least squares estimator for stochastic differential equations driven by small lévy noises*, Stochastic Processes and their Applications. (2017).

[26] Shimizu Y. Long H. and W. Sun, *Least squares estimators for discretely observed stochastic processes driven by small lévy noises*, Journal of Multivariate Analysis. (2013).

[27] Hansen LP, *Generalized instrumental variable estimation of nonlinear rational expectations models.*, Econometrica. (1982).

[28] Yakov Pesin Luis Barreira, *Introduction to smooth ergodic theory*, Graduate Studies in Mathematics, 2013.

[29] Wachter JA Lynch AW, *Using samples of unequal length in generalized method of moments estimation.*, J Financ Quant Anal. (2013).

[30] M. Ulrbich M. Hinze R. Pinnau and S. Ulbrich., *Optimization with pde constraints.*, Springer. (2009).

[31] H. Masuda, *Simple estimators for parametric markovian trend of ergodic processes based on sampled data*, J. Japan Statist. (2007).

[32] Jon Wilkening Matthias Morzfeld Xuemin Tu and Alexandre J. Chorin1., *Parameter estimation by implicit sampling.*, Commun. Appl. Math. Comput. Sci. (2015).

[33] T. Ogihara and N. Yoshida, *Quasi-likelihood analysis for the stochastic differential equation with jumps*, Stat. Inference Stoch. Process. (2011).

[34] Billingsley P, *Statistical inference for markov processes.*, The University of Chicago Press, Chicago. (1961).

[35] Leitch RA Paulson AS Holcomb EW, *The estimation of the parameters of the stable laws.*, Biometrika. (1975).

[36] Goran Peskir and Albert Shiryaev, *Optimal stopping and free-boundary problems*, Springer, 2006.

[37] Schmidt P Qian H, *Improved instrumental variables and generalized method of moments estimators.*, J Econom. (1999).

[38] Zhang S, *Transition-law based simulation of generalized inverse gaussian ornstein–uhlenbeck processes*, Methodol Comput Appl Probab. (2011).

[39] Zhang S and Zhang X, *A least squares estimator for discretely observed ornstein–uhlenbeck processes driven by symmetric $\alpha$-stable motions.*, Ann Inst Stat Math. (2013).

[40] K. I Sato, *L evy processes and infinitely divisible distributions.*, Cambridge University Press (1999).

[41] Y. Shimizu, *M-estimation for discretely observed ergodic diffusion processes with infinite jumps.*, Stat. Inference Stoch. Process. (2006).

[42] Y. Shimizu and N. Yoshida, *Estimation of parameters for diffusion processes with jumps from discrete observations.*, Stat. Inference Stoch. Process. (2006).

[43] Albert N Shiryaev, *Optimal stopping rules, volume 8 of stochastic modeling and applied probability*, Springer-Verlag, Berlin, 2008.

[44] K. Spiliopoulos, *Methods of moments estimation of ornstein-uhlenbeck processes driven by general lévy process.*, University of Maryland, College Park. (2008).

[45] Agnès Sulem, *A solvable one-dimensional model of a diffusion inventory system*, Mathematics of Operations Research (1986).

[46] Schoutens W. Valdivieso L. and F. Tuerlinckx, *Maximum likelihood estimation in processes of ornstein-uhlenbeck type.*, Stat. Infer. Stoch. Process. (2009).

[47] Si Liu Xiao-Chuan Cai and Jun Zou., *Parallel overlapping domain decomposition methods for coupled inverse elliptic problems.*, Commun. Appl. Math. Comput. Sci. (2009).

[48] He X Zhang S, *Inference based on adaptive grid selection of probability transforms.*, Statistics. (2016).

## .1 Appendix for: The Expected Hitting Time Approach to Optimal Price Adjustment Problems

In this section we are going to provide details on the explicit computations of all the terms (I)–(V) in Equation (2.4).

**(I):** Using [4, Formula 3.20.7(a), (b)] with $\alpha = 0$, $\beta = \frac{1}{2}$, and $\gamma^2$ replaced by $\sqrt{2\gamma}$, we have

$$\begin{aligned}
& \mathbb{E}\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right) \\
= \; & \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right); R_{H(L,U)=L}\right\} + \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t dt\right); R_{H(L,U)=U}\right\} \\
= \; & \frac{S_{2|\nu|}\left(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma}}{\sigma}\right)}{L^{|\nu|-\nu} S_{2|\nu|}\left(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma}\right)} + \frac{S_{2|\nu|}\left(\frac{2\sqrt{2\gamma}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma}\right)}{U^{|\nu|-\nu} S_{2|\nu|}\left(\frac{2\sqrt{2\gamma U}}{\sigma}, \frac{2\sqrt{2\gamma L}}{\sigma}\right)} \equiv g_1(\gamma), \qquad (.1.1)
\end{aligned}$$

where the special functions $S_\nu(x,y)$, $I_\nu(x)$, $K_\nu(x)$ are defined as follows:

$$\begin{cases}
S_\nu(x,y) \equiv (xy)^{-\nu}\left(I_\nu(x)K_\nu(y) - K_\nu(x)I_\nu(y)\right); & (.1.2) \\
I_\nu(x) \equiv \displaystyle\sum_{k=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{\nu+2k}}{k!\Gamma(\nu+k+1)}; & (.1.3) \\
K_\nu(x) \equiv \dfrac{\pi}{2\sin(\nu\pi)}\left(I_{-\nu}(x) - I_\nu(x)\right). & (.1.4)
\end{cases}$$

From the equation (.1.1), we see that

$$\mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right] = -\frac{d}{d\gamma}\bigg|_{\gamma=0} g_1(\gamma).$$

Thus, computation of the term $\mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right]$ can be given by computing $\frac{d}{d\gamma}\Big|_{\gamma=0} g_1(\gamma)$. To compute this derivative, we first substitute the expression of $K_\nu$ into the expression for $S_\nu$ in (.1.2) to obtain

$$
\begin{aligned}
S_\nu(x,y) &= \frac{\pi}{2\sin(\nu\pi)}(xy)^{-\nu}\Big(I_\nu(x)(I_{-\nu}(y)-I_\nu(y))-(I_{-\nu}(x)-I_\nu(x))I_\nu(y)\Big) \\
&= \frac{\pi}{2\sin(\nu\pi)}(xy)^{-\nu}(I_\nu(x)I_{-\nu}(y)-I_{-\nu}(x)I_\nu(y))\,. \qquad (.1.5)
\end{aligned}
$$

Then, we have

$$
g_1(\gamma) = L^\nu A_1(\gamma) + U^\nu A_2(\gamma)\,,
$$

where

$$
A_1(\gamma) \equiv \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})}
$$

and

$$
A_2(\gamma) \equiv \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma})}\,.
$$

We denote the numerators of $A_1$ and $A_2$ by $A_{11}(\gamma)$ and $A_{21}(\gamma)$ and the denominators of $A_1$ and $A_2$ by $A_{12}(\gamma)$ and $A_{22}(\gamma)$, respectively. Thus we can write

$$
A_1(\gamma) = \frac{A_{11}(\gamma)}{A_{12}(\gamma)}, \qquad A_2(\gamma) = \frac{A_{21}(\gamma)}{A_{22}(\gamma)}\,. \qquad (.1.6)
$$

[Notice that $A_{12}(\gamma) = A_{22}(\gamma)$.] To compute the derivative of $g_1(\gamma)$, we need first to compute the derivative of the modified Bessel function $I_\nu$ in (.1.3). We shall use the series expansion to compute the relevant derivatives. First, by the definition of the modified Bessel function $I_\nu$, we have

$$
\begin{aligned}
I_{2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}) &= \sum_{k=0}^{\infty}\frac{(\frac{\sqrt{2\gamma U}}{\sigma})^{2|\nu|+2k}}{k!\Gamma(2|\nu|+k+1)} = (\frac{2\gamma U}{\sigma^2})^{|\nu|}\sum_{k=0}^{\infty}\frac{(\frac{2\gamma U}{\sigma^2})^k}{k!\Gamma(2|\nu|+k+1)} \\
&= (\frac{2\gamma U}{\sigma^2})^{|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(2|\nu|+2)} + o(\gamma^2)\right).
\end{aligned}
$$

Similarly, we have

$$
I_{2|\nu|}(\frac{2\sqrt{2\gamma L}}{\sigma}) = (\frac{2\gamma L}{\sigma^2})^{|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)} + \frac{2\gamma L}{\sigma^2\Gamma(2|\nu|+2)} + o(\gamma^2)\right).
$$

Again by the definition of the modified Bessel function of negative index, we have

$$I_{-2|\nu|}(\frac{2\sqrt{2\gamma U}}{\sigma}) = (\frac{2\gamma U}{\sigma^2})^{-|\nu|}\bigg(\frac{1}{\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(-2|\nu|+2)} + o(\gamma^2)\bigg).$$

Combining the above computations, we have

$$\begin{aligned} A_{11}(\gamma) &= U^{|\nu|}\bigg(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} \\ &\quad + \frac{2\gamma}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) - U^{-|\nu|}\bigg(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} \\ &\quad + \frac{2\gamma}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) + o(\gamma^2). \end{aligned}$$

Its derivative is then

$$\begin{aligned} \frac{d}{d\gamma}\bigg|_{\gamma=0} A_{11}(\gamma) &= U^{|\nu|}\bigg(\frac{2U}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) \\ &\quad - U^{-|\nu|}\bigg(\frac{2}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2U}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) + o(\gamma). \end{aligned}$$

As for $A_{12}$, we have by a similar computation

$$\begin{aligned} A_{12}(\gamma) &= U^{|\nu|}L^{-|\nu|}\bigg(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} \\ &\quad + \frac{2\gamma L}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) - U^{-|\nu|}a^{|\nu|}\bigg(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} \\ &\quad + \frac{2\gamma L}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2\gamma U}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) + o(\gamma^2). \end{aligned}$$

Its derivative is

$$\begin{aligned} \frac{d}{d\gamma} A_{12}(\gamma) &= U^{|\nu|}L^{-|\nu|}\bigg(\frac{2U}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2L}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) \\ &\quad - U^{-|\nu|}L^{|\nu|}\bigg(\frac{2L}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} + \frac{2U}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\bigg) \\ &\quad + o(\gamma). \end{aligned}$$

We also compute the expression for $A_{21}(\gamma)$:

$$
\begin{aligned}
A_{21}(\gamma) \;\; = \;\; & L^{-|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}\right. \\
& \left.+\frac{2\gamma L}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\right)-L^{|\nu|}\left(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)}\right. \\
& \left.+\frac{2\gamma L}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\right)+o(\gamma^2)\,.
\end{aligned}
$$

Hence, the derivative takes the form:

$$
\begin{aligned}
\frac{d}{d\gamma}A_{21}(\gamma) \;\; = \;\; & L^{-|\nu|}\left(\frac{2}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2a}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\right) \\
& -L^{|\nu|}\left(\frac{2L}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\right) \\
& +o(\gamma)\,.
\end{aligned}
$$

Finally, we need the following expression for the denominator of the derivative:

$$
A_{12}^2(\gamma) \;\; = \;\; \left(U^{|\nu|}L^{-|\nu|}-U^{-|\nu|}L^{|\nu|}\right)^2\frac{1}{\Gamma(2|\nu|+1)^2\Gamma(-2|\nu|+1)^2}+o(\gamma^2)\,.
$$

Thus, we can compute the derivative of $A_1(\gamma)$ and $A_2(\gamma)$ as follows:

$$
\begin{aligned}
\mathscr{A}_1 \;\; \equiv \;\; & \left.\frac{d}{d\gamma}\right|_{\gamma=0}A_1(\gamma)=\left.\frac{A'_{11}(\gamma)A_{12}(\gamma)-A'_{12}(\gamma)A_{11}(\gamma)}{A_{12}^2(\gamma)}\right|_{\gamma=0} \\
= \;\; & \left\{\frac{2}{\sigma^2(-2|\nu|+1)}(U^{2|\nu|}L^{-|\nu|}-L^{|\nu|}-L^{-|\nu|}U-U^{2|\nu|}L^{-|\nu|+1}+L^{-|\nu|+1}+L^{|\nu|}U)\right. \\
& \left.+\frac{2}{\sigma^2(2|\nu|+1)}(U^{-2|\nu|}L^{|\nu|}-L^{|\nu|}U-L^{-|\nu|}+L^{-|\nu|}U+L^{|\nu|+1}-U^{-2|\nu|}L^{|\nu|+1})\right\} \\
& /[(U/L)^{|\nu|}-(L/U)^{|\nu|}]^2\,, \qquad (.1.7)
\end{aligned}
$$

and from noticing $A_{12} = A_{22}$, we have

$$\begin{aligned}
\mathscr{A}_2 &\equiv \frac{d}{d\gamma}\bigg|_{\gamma=0} A_2(\gamma) = \frac{A'_{21}(\gamma)A_{22}(\gamma) - A'_{22}(\gamma)A_{21}(\gamma)}{A^2_{22}(\gamma)}\bigg|_{\gamma=0} \\
&= \bigg\{\frac{2}{\sigma^2(-2|\nu|+1)}(U^{-|\nu|}L^{2|\nu|} - U^{-|\nu|}L - U^{|\nu|} + U^{|\nu|}L + U^{-|\nu|+1} - U^{-|\nu|+1}L^{2|\nu|}) \\
&\quad + \frac{2}{\sigma^2(2|\nu|+1)}(U^{|\nu|}L^{-2|\nu|} - U^{-|\nu|} - U^{|\nu|}L - U^{|\nu|+1}L^{-2|\nu|} + U^{|\nu|+1} + U^{-|\nu|}L)\bigg\} \\
&\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2 . \qquad (.1.8)
\end{aligned}$$

Combining all of the above computations, we obtain the following explicit expression for the first term (I) in (2.4.1).

The term (I) in (2.4.1) is given by the following explicit formula:

$$(\mathrm{I}) = \mathbb{E}\left[\int_0^{H(L,U)} R_t dt\right] = -L^\nu \mathscr{A}_1 - U^\nu \mathscr{A}_2 , \qquad (.1.9)$$

where $\mathscr{A}_1$ and $\mathscr{A}_2$ are defined by (.1.7) and (.1.8), respectively. We can further simplify the expression as

$$(\mathrm{I}) = \frac{2(U - L + U^{-2\nu}(L-1) + L^{-2\nu}(1-U))}{\sigma^2(2\nu+1)(U^{-2\nu} - L^{-2\nu})} .$$

**(II):** Now we compute the second term (II) in (2.4.1). Using [4, Formula 3.20.7(a), (b)] with $\alpha = 0$, $\beta = 1$, and $\frac{\gamma^2}{2}$ replaced by $\gamma$, we have

$$\begin{aligned}
& \mathbb{E}\exp\left(-\gamma\int_0^{H(L,U)} R_t^2 dt\right) \\
= \;& \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t^2 dt\right) ; R_{H(L,U)=L}\right\} + \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)} R_t^2 dt\right) ; R_{H(L,U)=U}\right\} \\
= \;& \frac{S_{|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma}, \frac{\sqrt{2\gamma}}{\sigma})}{L^{|\nu|-\nu} S_{|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma}, \frac{\sqrt{2\gamma}L}{\sigma})} + \frac{S_{|\nu|}(\frac{\sqrt{2\gamma}}{\sigma}, \frac{\sqrt{2\gamma}L}{\sigma})}{U^{|\nu|-\nu} S_{|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma}, \frac{\sqrt{2\gamma}L}{\sigma})} \equiv g_2(\gamma) .
\end{aligned}$$

Then, by (.1.5) we have that $g_2(\gamma) = L^{\nu}B_1(\gamma) + U^{\nu}B_2(\gamma)$, where

$$
\begin{aligned}
B_1 &\equiv \frac{I_{|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma})I_{-|\nu|}(\frac{\sqrt{2\gamma}}{\sigma}) - I_{-|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma})I_{|\nu|}(\frac{\sqrt{2\gamma}}{\sigma})}{I_{|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma})I_{-|\nu|}(\frac{\sqrt{2\gamma}L}{\sigma}) - I_{-|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma})I_{|\nu|}(\frac{\sqrt{2\gamma}L}{\sigma})} = \frac{B_{11}(\gamma)}{B_{12}(\gamma)}, \\
B_2 &\equiv \frac{I_{|\nu|}(\frac{\sqrt{2\gamma}}{\sigma})I_{-|\nu|}(\frac{\sqrt{2\gamma}L}{\sigma}) - I_{-|\nu|}(\frac{\sqrt{2\gamma}}{\sigma})I_{|\nu|}(\frac{\sqrt{2\gamma}a}{\sigma})}{I_{|\nu|}(\frac{\sqrt{2\gamma}b}{\sigma})I_{-|\nu|}(\frac{\sqrt{2\gamma}L}{\sigma}) - I_{-|\nu|}(\frac{\sqrt{2\gamma}U}{\sigma})I_{|\nu|}(\frac{\sqrt{2\gamma}L}{\sigma})} = \frac{B_{21}(\gamma)}{B_{22}(\gamma)}.
\end{aligned}
\tag{.1.10}
$$

Now we can compute $B_{11}$, $B_{12}$, $B_{21}$, and $B_{22}$ as follows:

$$
\begin{aligned}
B_{11}(\gamma) &= L^{|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} + \frac{\gamma U^2}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) - L^{-|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} + \frac{\gamma U^2}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) + o(\gamma^2),
\end{aligned}
$$

$$
\begin{aligned}
B_{12}(\gamma) &= U^{|\nu|}L^{-|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} + \frac{\gamma U^2}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma L^2}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) - U^{-|\nu|}L^{|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma L^2}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} + \frac{\gamma U^2}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) + o(\gamma).
\end{aligned}
$$

Also, it can be seen that

$$
\begin{aligned}
B_{21}(\gamma) &= L^{-|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} + \frac{\gamma}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma L^2}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) - L^{|\nu|}\Bigg(\frac{1}{\Gamma(|\nu|+1)\Gamma(-|\nu|+1)} \\
&\quad + \frac{\gamma L^2}{2\sigma^2\Gamma(|\nu|+2)\Gamma(-|\nu|+1)} + \frac{\gamma}{2\sigma^2\Gamma(-|\nu|+2)\Gamma(|\nu|+1)}\Bigg) + o(\gamma^2).
\end{aligned}
$$

By the same method of computing $g_1(\gamma)$, we have that

$$\begin{aligned}
\mathscr{B}_1 &\equiv \left.\frac{d}{d\gamma}\right|_{\gamma=0} B_1(\gamma) \\
&= \left\{\frac{1}{2\sigma^2(-|\nu|+1)}(U^{2|\nu|}L^{-|\nu|} - L^{-|\nu|}U^2 - L^{|\nu|} - U^{2|\nu|}L^{2-|\nu|} + L^{|\nu|}U^2 + L^{2-|\nu|})\right. \\
&\quad \left. + \frac{1}{2\sigma^2(|\nu|+1)}(U^{-2|\nu|}L^{|\nu|} - L^{-|\nu|} - L^{|\nu|}U^2 + L^{2+|\nu|} + L^{-|\nu|}U^2 - U^{-2|\nu|}L^{|\nu|+2})\right\} \\
&\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2, \qquad (.1.11)
\end{aligned}$$

$$\begin{aligned}
\mathscr{B}_2 &\equiv \left.\frac{d}{d\gamma}\right|_{\gamma=0} B_2(\gamma) \\
&= \left\{\frac{1}{2\sigma^2(-|\nu|+1)}(L^{2|\nu|}U^{-|\nu|} - U^{|\nu|} - U^{-|\nu|}L^2 + U^{2-|\nu|} + U^{|\nu|}L^2 - L^{2|\nu|}U^{2-|\nu|})\right. \\
&\quad \left. + \frac{1}{2\sigma^2(|\nu|+1)}(U^{|\nu|}L^{-2|\nu|} - U^{|\nu|}L^2 - U^{-|\nu|} - U^{2+|\nu|}L^{-2|\nu|} + U^{2+|\nu|} + U^{-|\nu|}L^2)\right\} \\
&\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2. \qquad (.1.12)
\end{aligned}$$

Then, we have the following explicit expression for the second term (II) in (2.4.1).

The term (II) in (2.4.1) has the following explicit formula:

$$(\mathrm{II}) = \mathbb{E}\left[\int_0^{H(L,U)} R_t^2 dt\right] = -L^{\nu}\mathscr{B}_1 - U^{\nu}\mathscr{B}_2, \qquad (.1.13)$$

where $\mathscr{B}_1$ and $\mathscr{B}_1$ are defined by (.1.11) and (.1.12), respectively. Further simplifying, we obtain

$$(\mathrm{II}) = \frac{U^2 - L^2 + U^{-2\nu}(L^2-1) + L^{-2\nu}(1-U^2)}{2\sigma^2(\nu+1)(U^{-2\nu} - L^{-2\nu})}.$$

**(III):** Using [4, Formula 3.0.5(a)] and [4, Formula 3.0.5(a)], we have

$$\begin{aligned}
\mathbb{E}\left\{e^{-\alpha H(L,U)}\right\} &= \mathbb{E}\left\{e^{-\alpha H(L,U)}\,; R_{H(L,U)=L}\right\} + \mathbb{E}\left\{e^{-\alpha H(L,U)}\,; R_{H(L,U)=U}\right\} \\
&\equiv C_1 + C_2 \equiv g_3(\alpha),
\end{aligned}$$

where

$$C_1(\alpha) \equiv L^{\nu} \frac{U^{\sqrt{\nu^2+2\alpha/\sigma^2}} - U^{-\sqrt{\nu^2+2\alpha/\sigma^2}}}{(U/L)^{\sqrt{\nu^2+2\alpha/\sigma^2}} - (L/U)^{\sqrt{\nu^2+2\alpha/\sigma^2}}}, \tag{.1.14}$$

and

$$C_2(\alpha) \equiv U^{\nu} \frac{L^{-\sqrt{\nu^2+2\alpha/\sigma^2}} - L^{\sqrt{\nu^2+2\alpha/\sigma^2}}}{(U/L)^{\sqrt{\nu^2+2\alpha/\sigma^2}} - (L/U)^{\sqrt{\nu^2+2\alpha/\sigma^2}}}. \tag{.1.15}$$

Then we have that

$$\begin{aligned} \mathscr{C}_1 &\equiv \left.\frac{d}{d\alpha}\right|_{\alpha=0} C_1(\alpha) \\ &= \frac{L^{\nu}}{|\nu|\sigma^2} \frac{2\ln U(L^{-|\nu|} - L^{|\nu|}) + \ln L[(U/L)^{|\nu|} + (U/L)^{-|\nu|}](U^{|\nu|} - U^{-|\nu|})}{[(U/L)^{|\nu|} - (U/L)^{-|\nu|}]^2}, \end{aligned} \tag{.1.16}$$

$$\begin{aligned} \mathscr{C}_2 &\equiv \left.\frac{d}{d\alpha}\right|_{\alpha=0} C_2(\alpha) \\ &= \frac{U^{\nu}}{|\nu|\sigma^2} \frac{2\ln L(U^{-|\nu|} - U^{|\nu|}) - \ln U[(U/L)^{|\nu|} + (U/L)^{-|\nu|}](L^{-|\nu|} - L^{|\nu|})}{[(U/L)^{|\nu|} - (U/L)^{-|\nu|}]^2}. \end{aligned} \tag{.1.17}$$

In particular, when $\nu < 0$,

$$\mathscr{C}_1 = \frac{L^{\nu}}{-\nu\sigma^2} \frac{2\ln U(L^{\nu} - L^{-\nu}) + \ln L[(U/L)^{-\nu} + (U/L)^{\nu}](U^{-\nu} - U^{\nu})}{[(U/L)^{-\nu} - (U/L)^{\nu}]^2}, \tag{.1.18}$$

$$\mathscr{C}_2 = \frac{U^{\nu}}{-\nu\sigma^2} \frac{2\ln L(U^{\nu} - U^{-\nu}) - \ln U[(U/L)^{-\nu} + (U/L)^{\nu}](L^{\nu} - L^{-\nu})}{[(U/L)^{-\nu} - (U/L)^{\nu}]^2}. \tag{.1.19}$$

Multiplying the numerator and denominator by $\frac{(UL)^{-2\nu}}{U^{-2\nu} - L^{-2\nu}}$, we have the following simplified expression known in the literature [6, Equation (56)]

$$-\mathscr{C}_1 - \mathscr{C}_2 = \frac{(\ln L)U^{-2\nu} - (\ln U)L^{-2\nu} + \ln U - \ln L}{\nu\sigma^2(U^{-2\nu} - L^{-2\nu})}. \tag{.1.20}$$

The third term (III) in (2.4.1) is given by the following explicit formula:

$$(\mathrm{III}) = \mathbb{E}[H(L,U)] = -\mathscr{C}_1 - \mathscr{C}_2\,, \tag{.1.21}$$

where $\mathscr{C}_1$ and $\mathscr{C}_2$ are given by (.1.18) and (.1.19) respectively.

In summary, the first term in the Cost of Deviation of (2.3.2) becomes

$$c_1\mathbb{E}\left[\int_0^{H(L,U)}(R_t-1)^2dt\right] \;=\; c_1\mathbb{E}\int_0^{H(L,U)}R_t^2dt - 2c_1\mathbb{E}\int_0^{H(L,U)}R_tdt + c_1\mathbb{E}H(L,U)\,,$$

where the above first, second and third expectations are given by (.1.13), (.1.9) and (.1.21), respectively.

**(IV):** We first use [4, Formula 3.19.7(a), (b)] with $\alpha = 0$, to obtain

$$\begin{aligned}
&\mathbb{E}\exp\left(-\gamma\int_0^{H(L,U)}\frac{1}{R_t}dt\right)\\
= \;& \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)}\frac{1}{R_t}dt\right),R_{H(L,U)=L}\right\} + \mathbb{E}\left\{\exp\left(-\gamma\int_0^{H(L,U)}\frac{1}{R_t}dt\right),R_{H(L,U)=U}\right\}\\
= \;& L^{|\nu|+\nu}\frac{S_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}},\frac{2\sqrt{2\gamma}}{\sigma})}{S_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}},\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})} + U^{|\nu|+\nu}\frac{S_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma},\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})}{S_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}},\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})} \equiv L^{\nu}D_1(\gamma) + U^{\nu}D_2(\gamma) \equiv g_4(\gamma)\,,
\end{aligned}$$

where

$$\begin{aligned}
D_1(\gamma) \;&\equiv\; \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})} = \frac{D_{11}(\gamma)}{D_{12}(\gamma)}\,,\\
D_2(\gamma) \;&\equiv\; \frac{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})}{I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}}) - I_{-2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{U}})I_{2|\nu|}(\frac{2\sqrt{2\gamma}}{\sigma\sqrt{L}})} = \frac{D_{21}(\gamma)}{D_{22}(\gamma)}\,.
\end{aligned}$$

We also need to compute the derivative of $g_4(\gamma)$ to find $(\mathrm{IV}) = \mathbb{E}\left[\int_0^{H(L,U)}\frac{1}{R_t}dt\right]$. The computation involves in computing the derivative of the modified Bessel function $I_\nu$ as we performed earlier. We just list some computation results and omit the details:

$$\begin{aligned} D_{11}(\gamma) &= U^{-|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2 U\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)-U^{|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2 U\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)+o(\gamma^2)\,; \end{aligned}$$

$$\begin{aligned} D_{22}(\gamma) &= U^{-|\nu|}L^{|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2 U\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2 L\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)-U^{|\nu|}L^{-|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2 L\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2 U\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)+o(\gamma^2)\,; \end{aligned}$$

and

$$\begin{aligned} D_{21}(\gamma) &= L^{|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2 L\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)-L^{-|\nu|}\Big(\frac{1}{\Gamma(2|\nu|+1)\Gamma(-2|\nu|+1)} \\ &\quad +\frac{2\gamma}{\sigma^2 L\Gamma(2|\nu|+2)\Gamma(-2|\nu|+1)}+\frac{2\gamma}{\sigma^2\Gamma(-2|\nu|+2)\Gamma(2|\nu|+1)}\Big)+o(\gamma^2)\,. \end{aligned}$$

Thus,

$$\begin{aligned} \mathscr{D}_1 &\equiv \frac{d}{d\gamma}\Big|_{\gamma=0} D_1(\gamma) \\ &= \Big\{\frac{2}{\sigma^2(-2|\nu|+1)}(U^{-2|\nu|}L^{|\nu|}-L^{-|\nu|}-L^{|\nu|}U^{-1}-U^{-2|\nu|}L^{|\nu|-1}+L^{|\nu|-1}+L^{-|\nu|}U^{-1}) \\ &\quad +\frac{2}{\sigma^2(2|\nu|+1)}(U^{2|\nu|}L^{-|\nu|}-L^{|\nu|}-L^{-|\nu|}U^{-1}+L^{-|\nu|-1}+L^{|\nu|}U^{-1}-U^{2|\nu|}L^{-|\nu|-1})\Big\} \\ &\quad /[(U/L)^{|\nu|}-(L/U)^{|\nu|}]^2\,; \end{aligned} \tag{.1.22}$$

$$
\begin{aligned}
\mathscr{D}_2 &\equiv \left.\frac{d}{d\gamma}\right|_{\gamma=0} D_2(\gamma) \\
&= \left\{ \frac{2}{\sigma^2(-2|\nu|+1)}(L^{-2|\nu|}U^{|\nu|} - U^{-|\nu|} - U^{|\nu|}L^{-1} + U^{-|\nu|}L^{-1} + U^{|\nu|-1} - U^{|\nu|-1}L^{-2|\nu|}) \right. \\
&\quad \left. + \frac{2}{\sigma^2(2|\nu|+1)}(L^{2|\nu|}U^{-|\nu|} - U^{-|\nu|}L^{-1} - U^{|\nu|} - L^{2|\nu|}U^{-|\nu|-1} + U^{-|\nu|-1} + U^{|\nu|}L^{-1}) \right\} \\
&\quad /[(U/L)^{|\nu|} - (L/U)^{|\nu|}]^2 . \qquad (.1.23)
\end{aligned}
$$

Summarizing the above, we obtain: The fourth term (IV) in (2.4.1) is given explicitly by the following expression:

$$
(\mathrm{IV}) = \mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t} dt\right] = -L^\nu \mathscr{D}_1 - U^\nu \mathscr{D}_2 , \qquad (.1.24)
$$

where $\mathscr{D}_1$ and $\mathscr{D}_2$ are defined by (.1.22) and (.1.23), respectively. Then, further simplification shows

$$
(\mathrm{IV}) = \frac{2(L^{-1} - U^{-1} + U^{-2\nu}(1 - L^{-1}) + L^{-2\nu}(U^{-1} - 1))}{\sigma^2(1-2\nu)(U^{-2\nu} - L^{-2\nu})}.
$$

**(V):** Lastly, in order to complete the computation of $Z(L,U)$ it remains to compute the fifth term (V) in (2.4.1). At first glance, there is no relevant formula in [4] to compute $\mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t^2} dt\right]$. This seems presenting a challenge for the explicit computation of $Z(a,b)$. However, we can overcome this difficulty by using the following facts:

(i) Letting $\tilde{R}_t \equiv \frac{1}{R_t} = \exp(-\sigma^2\nu - \sigma W_t)$, $\tilde{R}$ is also a GBM with parameter $\nu$ replaced by $-\nu$ (note that there is no need to replace $\sigma$ by $-\sigma$ since $-W$ is also a Brownian motion);

(ii) The first exit time $H(L,U)$ is the same as the exit time of $\tilde{R}_t$ from $(1/U, 1/L)$.

Thus, we conclude

$$
\mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t^2} dt\right] = \mathbb{E}\left[\int_0^{H(\frac{1}{U},\frac{1}{L})} \tilde{R}_t^2 dt\right]
$$

and we only need to replace the associated terms in the expression of $\mathbb{E}\left[\int_0^{H(L,U)} R_t^2 dt\right]$ with $\nu$ replaced by $-\nu$, $L$ by $\frac{1}{L}$, and $U$ by $\frac{1}{U}$. In this way, we can obtain the following result.

The fifth term $(\mathrm{V})$ in (2.4.1) is given explicitly by the following:

$$(\mathrm{V}) = \mathbb{E}\left[\int_0^{H(L,U)} \frac{1}{R_t^2} dt\right] \quad = -L^{\nu}\mathscr{E}_1 - U^{\nu}\mathscr{E}_2\,, \tag{.1.25}$$

where

$$\begin{aligned}
\mathscr{E}_1 \quad \equiv \quad & \left\{\frac{1}{2\sigma^2(-|\nu|+1)}(L^{-2|\nu|}U^{|\nu|} - U^{|\nu|}L^{-2} - U^{-|\nu|} - L^{-2|\nu|}U^{|\nu|-2} + U^{-|\nu|}L^{-2} + U^{|\nu|-2})\right. \\
+ \quad & \left.\frac{1}{2\sigma^2(|\nu|+1)}(L^{2|\nu|}U^{-|\nu|} - U^{|\nu|} - U^{-|\nu|}L^{-2} + U^{-2-|\nu|} + U^{|\nu|}L^{-2} - L^{2|\nu|}U^{-|\nu|-2})\right\} \\
& /[(U/L)^{|\nu|} - (U/L)^{-|\nu|}]^2
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{E}_2 \quad \equiv \quad & \left\{\frac{1}{2\sigma^2(-|\nu|+1)}(U^{-2|\nu|}L^{|\nu|} - L^{-|\nu|} - L^{|\nu|}U^{-2} + L^{|\nu|-2} + L^{-|\nu|}L^{-2} - U^{-2|\nu|}L^{|\nu|-2})\right. \\
+ \quad & \left.\frac{1}{2\sigma^2(|\nu|+1)}(L^{-|\nu|}U^{2|\nu|} - L^{-|\nu|}U^{-2} - L^{|\nu|} - L^{-2-|\nu|}U^{2|\nu|} + L^{-2-|\nu|} + L^{|\nu|}U^{-2})\right\} \\
& /[(U/L)^{|\nu|} - (U/L)^{-|\nu|}]^2\,.
\end{aligned}$$

After simplification, we obtain

$$(\mathrm{V}) = \frac{U^{-2} - L^{-2} + U^{-2\nu}(L^{-2}-1) + L^{-2\nu}(1-U^{-2})}{2\sigma^2(1-\nu)(U^{-2\nu} - L^{-2\nu})}.$$

Finally, for the cost of adjustment term $\frac{K_L(L)}{E(H(L,U);R_H=L)} + \frac{K_U(U)}{E(H(L,U);R_H=U)}$, it can be expressed by $-\frac{K_L(L)}{\mathscr{C}_1} - \frac{K_U(U)}{\mathscr{C}_2}$ from the equations (.1.18) and (.1.19).