Asymptotic expansions and distribution properties for diffusion processes

She, Qihao

2017

She, Q. (2017). Asymptotic expansions and distribution properties for diffusion processes.
Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69543>

<https://doi.org/10.32657/10356/69543>

NANYANG TECHNOLOGICAL
UNIVERSITY

DOCTORAL THESIS

**Asymptotic expansions and
distribution properties for
diffusion processes**

Author:

QiHao SHE

Supervisor:

Associate Professor

Nicolas PRIVAULT

Co-Supervisor:

Professor Emmanuel

GOBET

School of Physical and Mathematical Sciences

*A thesis submitted to the Nanyang Technological University in
partial fulfilment of the requirement
for the degree of Doctor of Philosophy*

2016

Declaration of Authorship

I, QiHao SHE, declare that this thesis titled, “Asymptotic expansions and distribution properties for diffusion processes” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“Il est remarquable qu’une science qui a commencé par la considération des jeux de hasard devrait être devenue l’objet le plus important de la connaissance humaine.”

Pierre-Simon Laplace, DES Probabilités de Théorie Analytique, 1812

NANYANG TECHNOLOGICAL UNIVERSITY

Abstract

Division of Mathematical Sciences

School of Physical and Mathematical Sciences

Doctor of Philosophy

Asymptotic expansions and distribution properties for diffusion processes

by QiHao SHE

For this thesis, we derive new applications and theoretical results for some multidimensional mean-reverting stochastic processes via series expansion.

In Chapter 2, we use Malliavin calculus and moment cumulant formula to specify the conditions on the process $(v_s)_{s \in [0, T]}$, under which, we are able to obtain the following conditional Gaussian identities:

$$\mathbb{E} \left[\exp \left(i \int_0^T v_s dW_s \right) \middle| \int_0^T |v_s|^2 ds \right] = \exp \left(-\frac{1}{2} \int_0^T |v_s|^2 ds \right),$$

for Brownian stochastic integrals. In particular, if we have the matrix condition $A^\top A^2 = 0$, then from a characterization of Yor, we have the above mentioned formula for the quadratic Brownian integral $\int_0^s AW_t dW_t$.

In Chapter 3, we derive conditional Edgeworth-type expansions using result from Chapter 2. We then consider multiple stochastic integrals, where we obtain the Stein approximation bounds using the derived conditional Edgeworth-type expansions.

In Chapter 4, we derive a closed-form analytical approximations formula in terms of series expansion for the option prices, implied volatility and delta under the 2-Hypergeometric stochastic volatility model with correlated Brownian motions. Most notably, the approximation formula we derive for the implied volatility is capable of

recovering the well-known smiles and skew phenomenon on implied volatility surfaces, depending on the correlation.

Lastly, in Chapter 5, we let X be a multidimensional ergodic Ornstein-Uhlenbeck process, Y be a multidimensional stochastic process such that its stochastic differential equation is written as a drift-perturbation of X and μ^Y be the stationary distribution of Y . We derive a first and second order expansion of μ^Y in terms of X and their respective error estimates. Thereafter, we turn these approximations into a simulation scheme to sample μ^Y approximately.

Acknowledgements

This thesis would not have been possible if not for the kindness and help from these people.

Firstly, I would like to thank my supervisor Associate Professor Nicolas Privault, while we may not agree on everything, he always patiently explain the reasons to me. His ability in Mathematics have also never failed to amaze me. Having me as a PhD student is never easy on him, so I am thankful that he has never stopped guiding me.

I would also like to thank my co-supervisor, Professor Emmanuel Gobet from Ecole Polytechnique. Thank you for hosting me in Paris and showing me the rope of doing mathematical research and writings.

My PhD experience would never be wholesome if not for the wonderful colleague I have, especially Ian and Bunyamin.

The endless discussion with Ian allows me to think deeper and consider wider perspectives. I would also like to thank him for being incredibly helpful.

Having Bunyamin as my fellow PhD mate is refreshing, I have never seen someone complete their PhD with such finesse.

Thank you Professor Yue and Associate Professor Yeow for the help in securing the fee exemption.

I would also like to thank SPMS for the wonderful experience and of course the secretaries like Si Yun, Yean Chin and Shi Qin.

Lastly, of course I would like to take the chance to thank the reviewers and examiners for taking the time to review my thesis.

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
Contents	xi
List of Figures	xiii
List of Tables	xv
1 Introduction and Main Results	1
2 Conditionally Gaussian Stochastic Integrals	5
2.1 Introduction	5
2.2 Preliminaries	7
2.3 Conditional Characteristic Functions	14
2.4 Vanishing of $A^\top A^2$	17
2.5 Anti-Symmetric $A^\top A^2$	18
2.6 Commutation with Orthogonal matrices	20
2.7 Anti-symmetric orthogonal A	21
3 Conditional Stein approximation for Itô and Skorohod integrals	23
3.1 Introduction	23
3.2 Conditional Edgeworth type expansions	26
3.3 Conditional Stein approximation	29
3.4 Double stochastic integrals	31
4 Option pricing and implied volatilities in a 2-Hypergeometric stochastic volatility model	37
4.1 Introduction	37
4.2 Preliminaries	39
4.3 Stochastic Volatility	42
4.4 Deterministic Volatility	43
4.5 First Order Expansion	44
4.6 Second Order Expansion	46
4.7 Implied Volatility	51
4.8 Delta	53

4.9	Second Order Expansion Proof	54
5	Perturbation of Ornstein-Uhlenbeck Stationary Distributions: expansion and simulation	61
5.1	Introduction	61
5.2	Preliminary results	65
5.3	Main approximation results	71
5.3.1	Strategy	72
5.3.2	First and second order expansions	74
5.3.3	Proof of Theorem 5.1	77
5.3.4	Proof of Theorem 5.2	83
5.4	Application to simulation algorithm, numerical results	90
5.4.1	Generalized acceptance/rejection algorithm . .	90
5.4.2	Simulation algorithm for 1st and 2nd order approximation	93
5.4.3	Numerical experiments	99
A	Vector Calculus	105
A.1	Kronecker Product	105
A.2	vec operator and its relation to the Kronecker product .	106
A.3	vec and its relation to the Trace operator	108
A.4	Differential and Derivative	109
A.5	Derivatives of Vector Functions	110
A.6	Derivatives of Matrix Functions	111
A.7	Derivatives of Trace Operator	112
A.8	Divergence	112
	Bibliography	115

List of Figures

4.1	Option price f plotted against v with $\rho = -0.5$	50
4.2	Implied volatility σ^{imp} plotted against the moneyness K/x for $\rho = 0$	52
4.3	Implied volatility σ^{imp} plotted against the moneyness K/x for $\rho = -0.8$	53
5.1	Density plot for the stationary distribution of $d\tilde{Y}_t = (-A\tilde{Y}_t + \beta(\tilde{Y}_t))dt + dW_t$, where $a_{11} = 7$, $a_{22} = 7.1$ and $b = 0.1$	100
5.2	Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 1Y_te^{-Y_t^2})dt + dW_t$	101
5.3	Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 1.6Y_te^{-Y_t^2})dt + dW_t$	102
5.4	Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 1.7Y_te^{-Y_t^2})dt + dW_t$	103
5.5	Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 5Y_te^{-Y_t^2})dt + dW_t$	104

List of Tables

4.1	Values of $f_0, f_0 + \epsilon \tilde{f}_1, f_0 + \epsilon \tilde{f}_1 + \epsilon^2 \tilde{f}_2$ compared to Monte Carlo estimates.	50
-----	---	----

For/Dedicated to/To my...Repository

Chapter 1

Introduction and Main Results

In this manuscript, new applications and theoretical results for multidimensional stochastic processes are derived. Moreover, we develop a new analytical approximation formula for pricing options. In particular, the multidimensional stochastic processes which we are considering are those that are mean-reverting. In general, we begin by adopting a suitable expansion scheme. Next, we use the tools of Malliavin calculus to obtain the respective terms in the expansions.

In chapter 2, we use the moment cumulant formula which is developed in [Pri13] and [Pri15a] and Malliavin calculus to give the precise conditions under which we can derive the various variations of the following equation:

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_t dW_t \right) \middle| W_s \right] = \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |AW_t|^2 dt \right) \middle| W_s \right],$$

where A is a $d \times d$ anti-symmetric matrix and $0 \leq s \leq T$. For the stochastic integrals $\int_0^T v_s dW_s$ of a process $(v_s)_{s \in [0, T]}$ that is $(\mathcal{F}_s)$ -adapted. We will also make precise conditions for which the stochastic integral $\int_0^T v_s dW_s$ has a central normal distribution with variance $\int_0^T |v_s|^2 ds$ whenever $\int_0^T |v_s|^2 ds$ is given. In particular, we show that if we assume Yor's condition $A^\top A^2 = 0$ and let $(v_s)_{s \in [0, T]} = (AW_s)_{s \in [0, T]}$, then our identities holds. Moreover, we prove that the Yor's condition can be weakened to $A^\top A^2$ anti-symmetric if $A^\top A$ is proportional to a projection.

In Chapter 3, we use result obtain in Chapter 2 to derive a conditional Edgeworth-type expansions for Skorohod and Itô integrals with respect to the Brownian motion. Thereafter, we derive conditional Stein approximation bounds for multiple stochastic integrals.

In chapter 4, we develop an approximate pricing formula for the European option under the Hypergeometric stochastic volatility model introduced in [DM16] to address the shortcomings associated with pricing option using the Heston stochastic volatility model. We assume that the asset price S_t and volatility V_t at time t is governed by the following dynamics:

$$\begin{aligned} dS_t &= S_t e^{V_t} dW_t^1, \\ dV_t &= \left(a - \frac{c}{2} e^{\alpha V_t} \right) dt + \eta dW_t^2, \end{aligned}$$

under the α -Hypergeometric model, where $a \in \mathbb{R}$, $c > 0$, $\alpha > 0$, $\eta > 0$, W_t^1 and W_t^2 are correlated Brownian motions which satisfies $\langle W^1, W^2 \rangle_t = \rho t$. Since explicit solutions for stochastic volatility model are not usually possible, we use an approximation proposed in [Han+13] to obtain an approximation via a series expansion that is dependable on the underlying stochastic volatility. Consequently, we address a shortcoming in the approximation method detailed in [Fou+11]. Our approximations obtain an implied volatility that recovers the effect of smile and skew. Furthermore, we also obtain an approximation for the Delta of the option.

In chapter 5, we consider a perturbed multidimensional Ornstein-Uhlenbeck process:

$$dY_t = \left[-AY_t + \beta(Y_t) \right] dt + \Sigma dW_t, \quad Y_0 \text{ independent of } W,$$

where A is $d \times d$ -matrix, Σ is $d \times q$ -matrix and $\beta : \mathbb{R}^d \mapsto \mathbb{R}^d$ is a measurable function. We show that by making the following assumptions:

(H-i) $-A$ is a Hurwitz matrix.

Choose $\lambda_{0,A} \in (0, \min_{\lambda \in \lambda(A)} \operatorname{Re}(\lambda))$: for such $\lambda_{0,A}$, define

$$A_0 := A - \lambda_{0,A} \operatorname{Id};$$

note that $-A_0$ is a Hurwitz matrix. Denote by $c_A > 0$ the constant such that $\|\exp(-As)\| \leq c_A \exp(-\lambda_{0,A}s)$, $\forall s \geq 0$ (see Lemma 5.1).

(H-ii) $A\Sigma\Sigma^\top$ is symmetric.

(H-iii) $\Sigma\Sigma^\top$ is non-singular.

(H-iv) β is a $d \times 1$ vector and all its derivatives up till third order are bounded.

(H-v) h is a continuous bounded function and its derivatives up till third order exist and are bounded.

(H-vi) $\lambda_{0,A} > c_A \|\beta^{(1)}\|_\infty$, where $\lambda_{0,A}, c_A > 0$ are the constants from **(H-i)**. We set

$$c_{(5.16)} := \frac{c_A}{\lambda_{0,A} - c_A \|\beta^{(1)}\|_\infty}. \quad (1.1)$$

We can obtain an approximate stationary distribution for Y which is hardly tractable by using an instrumental process X with the following dynamics:

$$dX_t = -AX_t dt + \Sigma dW_t, \quad X_0 \text{ independent of } W.$$

Moreover, our approximation allowed us to develop an efficient algorithm to simulate from the stationary distribution of Y . Error estimates for our approximation are also established.

Chapter 2

Conditionally Gaussian Stochastic Integrals¹

Abstract

In this chapter, we start by giving an overview of Stochastic calculus and Malliavin calculus. Thereafter, we use Malliavin calculus and the moment cumulant formula of [Pri13] and [Pri15a] to specify the conditions for the process $(v_s)_{s \in [0, T]}$ which allow us to obtain conditional Gaussian identities of the following form

$$\mathbb{E} \left[\exp \left(i \int_0^T v_s dW_s \right) \middle| \int_0^T |v_s|^2 ds \right] = \exp \left(-\frac{1}{2} \int_0^T |v_s|^2 ds \right),$$

for Brownian stochastic integrals. More specifically, from a characterization of Yor (see [Yor79]), we have under the matrix condition $A^\top A^2 = 0$, the said formula for the quadratic Brownian integral $\int_0^s AW_t dW_t$.

2.1 Introduction

Suppose that $(W_s)_{s \in [0, T]}$ is a d -dimensional standard Brownian motion and denotes the filtration generated by $(\mathcal{F}_s)_{s \in [0, T]}$. In this chapter, we use methods from Malliavin calculus to find the conditions such that the stochastic integral $\int_0^T v_s dW_s$, where v_s is a stochastic process that is adapted to $\mathcal{F}_t$, has a normal distribution when conditioned on $\int_0^T |v_s|^2 ds$. The results demonstrate that using cumulant formula developed in [Pri15a], we can generalize the results related to the

¹This chapter is based on works published in [PS15].

problem of stochastic area studied in [Yor80] and [DEM16], where they took $v_s = AW_s$ and an anti-symmetric matrix A .

The following identity is proved in Theorem 2.1 of [DEM16]

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_t dW_t \right) \middle| W_s \right] = \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |AW_t|^2 dt \right) \middle| W_s \right], \quad (2.1)$$

where A is a $d \times d$ anti-symmetric matrix and $0 \leq s \leq T$. This is an extension to $d \geq 2$ for the formula of [Yor80], used in the computation of the characteristics function of Lévy stochastic area when $d = 2$.

The proof relies on a result of Yor (see [Yor79]) which prove that the filtration $(\mathcal{F}_s^m)_{s \in [0, T]}$ of $s \mapsto \int_0^s AW_t dW_t$ when $A^\top A^2 = 0$ is generated by m independent Brownian motions, where m is the number of distinct nonzero eigenvalues of $A^\top A$.

Using moment cumulant formula of [Pri13] and [Pri15a] and Malliavin calculus, we formulate the condition under which we are able to derive various variations of the Equation (2.1) for the stochastic integral $\int_0^T v_s dW_s$ of a process $(v_s)_{s \in [0, T]}$ that is $(\mathcal{F}_s)$ -adapted (see Theorems 2.4 and 2.5). Moreover, we derive the condition for which given $\int_0^T |v_s|^2 ds$, the stochastic integral $\int_0^T v_s dW_s$ has a central normal distribution with variance $\int_0^T |v_s|^2 ds$ (see Theorem 2.5).

In particular, assuming Yor's condition $A^\top A^2 = 0$, our identities holds when $(v_s)_{s \in [0, T]} = (AW_s)_{s \in [0, T]}$ (see Corollary 2.1). Furthermore, if $A^\top A$ is proportional to a projection we can weaken Yor's condition to $A^\top A^2$ anti-symmetric (see Corollary 2.2).

In [Oco93], we know that an Ocone martingale is a martingale which is Gaussian distributed with mean 0 and variance its quadratic variation when conditioned under its quadratic variation. Thus, martingales that satisfied our sufficient conditions are called Ocone martingales and are studied extensively in [Oco93] and [VY00]. In [Oco93], Ocone gave sufficient condition for a martingale to be Ocone martingale. In particular, he proved that if $M_T = \int_0^T v_s dW_s$ satisfied the H-invariance condition which requires that M_T have the same law as the process $\int_0^r v_s dW_s - \int_r^T v_s dW_s$, for all $r \in [0, T]$, $T \in \mathbb{R}_+$, then M_T is an Ocone martingale. Whereas, the sufficient condition we have is that

$\langle v, (Dv)^k v \rangle = 0$, for all $k \geq 1$. Hence, it seems like we have obtained a sufficient condition that is different from that found in [Oco93].

Furthermore, in the matrix case, an adaptation of the proof found in [Yor79], provided another sufficient condition that is disconnected with that of Ocone and vanishing of $\langle v, (Dv)^k v \rangle = 0$, for all $k \geq 1$. Therefore, we see that Ocone's condition is sufficient but not necessary and we have in this chapter derived new sufficient conditions for a martingale to be an Ocone martingale.

In Section 2.2, we provide an overview of the results of Malliavin calculus which is relevant in our derivation of the generalized results. After the section of preliminaries, we present our results under different conditions in Sections 2.3, 2.4, 2.5, 2.6 and 2.7.

2.2 Preliminaries

The following sections will make use of the results from Stochastic Calculus and Malliavin Calculus to derive an identity for the conditional expectation of the exponential of the Gaussian stochastic integrals. Thus, in this section, we provide an introduction to some of the results of stochastic calculus and Malliavin calculus. The aim is to highlight the definitions and results that are used in this thesis. For a detailed discussion of stochastic calculus and Malliavin calculus, please refer to [Üst10] and [Nua06].

Adapted Process, Martingales, Quadratic Variation, Wiener Process and Isonormal Gaussian Process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\mathcal{F}_t$ be a filtration for the probability space. We will start by recalling some definitions.

Definition 2.1 (Adapted Process). *A stochastic Process $(Y_t)_{t \geq 0}$ is an adapted process if for every $t \geq 0$, Y_t is $\mathcal{F}_t$ -measurable.*

Definition 2.2 (Martingales). *For any stochastic process, $(Y_s)_{s \in \mathbb{R}_+}$, if we have $\mathbb{E}[|Y_s|] < \infty$ and $\forall \tau, s > 0$, $\mathbb{E}[Y_{s+\tau} | \mathcal{F}_s] = Y_s$, then it is a martingale with respect to $\mathcal{F}_t$.*

Definition 2.3 (Quadratic Variation). *The quadratic variation of a Stochastic process Y_t , $[Y, Y]_t$ is:*

$$[Y, Y]_T = \lim_{m \rightarrow +\infty} \sum_{j=0}^{m-1} (Y_{\tau_{j+1}} - Y_{\tau_j})^2$$

where $(\tau_j)_{j \geq 0}$ is the partition of $[0, T]$.

Definition 2.4 (Standard Wiener Process). *A standard Wiener process, $(W_s)_{s \in \mathbb{R}^+}$, is a stochastic process that satisfies the following 4 conditions:*

1. $W_0 = 0$,
2. For any $0 \leq \tau < s$, the increments, $W_{s-\tau} = W_s - W_\tau$ follows a central normal distribution with variance $s - \tau$,
3. For $\tau > 0$, $\sigma(W_t : 0 \leq t \leq s)$ is independent of $W_{s+\tau} - W_s$,
4. With probability 1, it has continuous trajectories.

Therefore, a standard d -dimensional Brownian motion,

$\tilde{W}_s = (W_s^1, \dots, W_s^d)$ is one where W_s^j are mutually independent and $\forall j, W_s^j$ is a standard Brownian motion.

Theorem 2.1. *If W_t is a standard Brownian motion, then $[W, W]_T = T$ with probability 1.*

Isonormal Gaussian Process and Multiple Stochastic Integral

In this section, we provide the definitions of Isonormal Gaussian process and Multiple Stochastic Integral. We also discuss some results related to the stochastic integral, namely Itô isometry and Itô Lemma. Note that we restrict the discussion to the stochastic integral associated with the Brownian Motion in this thesis.

Definition 2.5 (Isonormal Gaussian Process). *Let H be a real and separable Hilbert space, if $W = \{W(h), h \in H\}$ is a Gaussian random variable for all $h \in H$ with mean 0 and $\mathbb{E}[W(h)W(f)] = \langle h, f \rangle_H$, for all $h, f \in H$, where $\langle h, g \rangle_H$ is the inner product of H , then we said that W is a Isonormal Gaussian Process or Gaussian process on H .*

Remark 2.1. We are only concerned with the Isonormal Gaussian Process associated with the Brownian motion, i.e.

$$W(h) = \sum_{i=1}^d \int_{\mathbb{R}^+} h_i(t) dW_t^i,$$

where $(W_t)_{t \geq 0}$ is a d -dimensional Brownian motion and $h \in H$, $H = L^2(\mathbb{R}^+; \mathbb{R}^d)$.

Let $g \in L^2([a, b]^d)$, where $0 \leq a < b < +\infty$ and Ξ_d be the vector space formed by the following set of elementary functions:

$$g(s_1, \dots, s_d) = \sum_{j_1, \dots, j_d=0}^{m-1} g_{j_1 \dots j_d} \mathbb{1}_{[\tau_{j_1}, \tau_{j_1+1}) \times \dots \times [\tau_{j_d}, \tau_{j_d+1})}(s_1, \dots, s_d),$$

where $a = \tau_0 < \tau_1 < \dots < \tau_m = b$, such that $g_{j_1 \dots j_d} = 0$ if $j_i = j_k$ for some $i \neq k$.

Theorem 2.2. Let $g \in L^2([a, b]^d)$ and $h \in L^2([a, b]^n)$, $d, n \geq 1$. If we define

$$I_d(g) = \sum_{j_1, \dots, j_d=0}^{m-1} g_{j_1 \dots j_d} W_{\tau_{j_1+1}-\tau_{j_1}} \dots W_{\tau_{j_d+1}-\tau_{j_d}}, \quad \text{for } g \in \Xi_d$$

and

$$\tilde{g}(s_1, \dots, s_d) = \frac{\sum_{\sigma \in P_d} g(s_{\sigma_1}, \dots, s_{\sigma_d})}{d!}$$

where P_d is the set of all the permutations of $\{1, \dots, d\}$, then for $g \in \Xi_d$ and $h \in \Xi_n$, we have that

1. $I_d(g)$ is linear,
2. $\mathbb{E}[I_d(g)] = 0$,
3. $I_d(\tilde{g}) = I_d(g)$,
- 4.

$$\mathbb{E}[I_d(g) I_n(h)] = \begin{cases} 0, & \text{for } d \neq n, \\ d! \langle \tilde{g}, \tilde{h} \rangle_{L^2([a, b]^d)}, & \text{for } d=n. \end{cases}$$

Since Ξ_d is dense in $L^2([a, b]^d)$, the result for Theorem (2.2) can be extended to the case of $g \in L^2([a, b]^d)$ and we have the Wiener-Itô integral

$$I_d(g) = \int_{[a,b]^d} g(s_1, \dots, s_d) dW_{s_1} \dots dW_{s_d}.$$

We see that (4) of Theorem (2.2) implies Itô isometry and we use it to derive the Itô Lemma or Itô formula, which is a very important result of stochastic calculus.

Proposition 2.1 (Itô Lemma or Itô formula, Proposition 1.1.5, p. 19 of [Nua06]). *Let $G : \mathbb{R} \rightarrow \mathbb{R}$ be a twice continuously differentiable function and $Y_T = Y_0 + \int_0^T \mu(t) dt + \int_0^T \sigma(t) dW_t$, where μ_t and σ_t are measurable and adapted processes such that $\int_0^s \sigma_t^2 dt < +\infty$ and $\int_0^s |\mu_t| dt < +\infty$ almost surely for every $s \in T$, then*

$$\begin{aligned} G(Y_T) &= G(Y_0) + \int_0^T G^{(1)}(Y_t) \mu_t dt + \int_0^T G^{(1)}(Y_t) \sigma_t dW_t \\ &\quad + \frac{1}{2} \int_0^T G^{(2)}(Y_t) \sigma_t^2 dt. \end{aligned}$$

Derivative Operator or Annihilation Operator

Let $G : \Omega \rightarrow \mathbb{R}$, be a square integrable random variable. The aim of this subsection is to provide the definition of the Derivative operator or Annihilation operator DG , where the derivative is taken with respect to $\omega \in \Omega$ and some properties related to it.

Definition 2.6 (Derivative Operator or Annihilation Operator). *Let $\mathcal{S}$ be the set of smooth random variables, where*

$$\mathcal{S} := \{G = g(W(h_1), \dots, W(h_d)), g \in \mathcal{C}_p^\infty(\mathbb{R}^d), h_j \in H, d \geq 1\}$$

where $\mathcal{C}_p^\infty(\mathbb{R}^d)$ is the set of polynomial growth functions g , from $\mathbb{R}^d$ to $\mathbb{R}$, such that all its derivatives exist and have polynomial growth. So for $G \in \mathcal{S}$, the derivative of DG is

$$DG = \sum_{j=1}^d \partial_j g(W(h_1), \dots, W(h_d)) h_j.$$

The integration by parts formula plays an important role in the discussion of Malliavin calculus, and it is very important in our discussion, so we provide the integration by parts formula in the following lemma.

Lemma 2.1 (Integration by parts, Lemma 1.2.1, p. 25 of [Nua06]). *If $G \in \mathcal{S}$ and $h \in H$, then*

$$\mathbb{E} [\langle DG, h \rangle_H] = \mathbb{E} [GW(h)].$$

As a consequence of the fact that $DGF = GDF + FDG$, we have the following lemma.

Lemma 2.2 (Lemma 1.2.2, p.26 of [Nua06]). *If $G, F \in \mathcal{S}$ and $g \in H$, then*

$$\mathbb{E} [GFW(g)] = \mathbb{E} [F \langle DG, g \rangle_H] + \mathbb{E} [G \langle DF, g \rangle_H].$$

Since the derivative operator D is closable from $L^p(\Omega)$ to $L^p(\Omega; H)$ for $p \geq 1$, we have the following definition for the domain of the derivative operator.

Definition 2.7. *Let $\mathbb{D}^{p,k}$ be the domain of the derivative operator D , then $\mathbb{D}^{p,k}$ is the completion of the set of smooth random variables $\mathcal{S}$ under the following norm:*

$$\|G\|_{p,k}^p = \mathbb{E} [|G|^p] + \sum_{i=1}^k \mathbb{E} [\|D^i G\|_{H^{\otimes i}}^p].$$

Now, we can give the decomposition of the derivative in the Wiener Chaos.

Theorem 2.3. *G is a square integrable variable in $\mathbb{D}^{2,1}$ if and only if*

$$D_s G = \sum_{d=1}^{\infty} d I_{d-1} (g_d(*, s))$$

where $g_d \in L^2([a, b]^d)$ is symmetric, $G = \sum_{d=0}^{\infty} I_d(g_d)$ and

$$\sum_{d=1}^{\infty} dd! \|g_d\|_{L^2([a,b]^d)}^2 < +\infty.$$

The Divergence Operator or Creation Operator

The definition and properties of the Divergence operator or Creation operator which is the adjoint of the Derivative operator will be consider in this subsection.

Definition 2.8 (The Divergence Operator, Definition 1.3.1, p. 36 [Nua06]). *Let the Divergence operator, δ be the adjoint of the Derivative operator D , then δ is an unbounded operator on $L^2(\Omega; H)$ taking values in $L^2(\Omega)$ such that:*

1. *The domain of δ , $\text{Dom } \delta$ is the set of square integrable H -valued random variables in $L^2(\Omega; H)$ satisfying the inequality:*

$$|\mathbb{E}[\langle DG, u \rangle_H]| \leq c_u \|G\|_{L^2(\Omega)},$$

for all $G \in \mathbb{D}^{2,1}$ and $u \in L^2(\Omega; H)$, where c_u is a constant that depends on u .

2. *Let $u \in \text{Dom } \delta$, then $\delta(u) \in L^2(\Omega)$ satisfying*

$$\mathbb{E}[\langle DG, u \rangle_H] = \mathbb{E}[G\delta(u)]$$

for every $G \in \mathbb{D}^{2,1}$.

We will now introduce a proposition that let us factor out a scalar random variable in a divergence.

Proposition 2.2 (Proposition 1.3.3, p. 39 of [Nua06]). *Suppose that $G \in \mathbb{D}^{2,1}$, $u \in \text{Dom } \delta$ such that $G u \in L^2(\Omega; H)$ and $G\delta(u) - \langle DG, u \rangle_H \in L^2(\Omega)$, then we have $G u \in \text{Dom } \delta$ and*

$$\delta(Gu) = G\delta(u) - \langle DG, u \rangle_H.$$

The Skorohod Integral

In this subsection, we take $H = L^2([a, b])$ and define the Skorohod Integral to be $\delta(u)$, since the divergence operator can be viewed as a stochastic integral when $H = L^2([a, b])$. In this case, we have

$$\delta(u) = \int_a^b u_s dW_s.$$

and $\text{Dom } \delta \subset L^2([a, b] \times \Omega)$. This also means that for any $u(s) \in \text{Dom } \delta$, it is square integrable and so we have

$$u(s) = \sum_{k=0}^{+\infty} I_k(g_k(*, s)),$$

where $g_k \in L^2([a, b]^{k+1})$, for every $k \geq 1$ is a symmetric function for the first k variables. As with other subsections, we will discuss some properties of the skorohod integral. We will also present a Wiener chaos expansion representation of the skorohod integral. We start by providing the Wiener chaos decomposition of the skorohod integral, δ .

Proposition 2.3 (Wiener Chaos Decomposition of δ , Proposition 1.3.7, p. 41 of [Nua06]). *Suppose that $u(s) = \sum_{k=0}^{+\infty} I_k(g_k(*, s)) \in L^2([a, b] \times \Omega)$, then $u \in \text{Dom } \delta$ if and only if*

$$\delta(u) = \sum_{k=0}^{+\infty} I_{k+1}(g_k) = \sum_{k=0}^{+\infty} I_{k+1}(\hat{g}_k),$$

converges in $L^2(\Omega)$. Here, $\hat{g}_k$ is the symmetrization of g_k defined as

$$\begin{aligned} \hat{g}_k(s_1, \dots, s_k, s) = & \frac{1}{k+1} (g_k(s_1, \dots, s_k, s) \\ & + \sum_{j=1}^k g_k(s_1, \dots, s_{j-1}, s, s_{j+1}, \dots, s_k, s_j)) \end{aligned}$$

We end this preliminaries section by introducing a proposition that enable us to compute the derivative of $\delta(u)$.

Proposition 2.4 (Derivative of $\delta(u)$, Proposition 1.3.8, p. 43 of [Nua06]). *Assume that $u \in \mathbb{D}^{2,1}(L^2([a, b]))$ and for every $r, s \in [a, b]$, $D_r u(s) \in \text{Dom } \delta$. In addition, we also assume that a version of $\delta(D_r u(s))$ exists, such that $\delta(D_r u(s)) \in L^2([a, b] \times \Omega)$. Then we have $\delta(u) \in \mathbb{D}^{2,1}$ and for all $r \in [a, b]$,*

$$D_r(\delta(u)) = u(r) + \delta(D_r u).$$

2.3 Conditional Characteristic Functions

We take $H = L^2([0, T]; \mathbb{R}^d)$ where $T > 0$. If on the d -dimensional Wiener space, we denote the Malliavin gradient with a domain of $\mathbb{D}^{2,1}$ by D (see §1.2 of [Nua06] for definitions) and suppose that v is in the domain $\mathbb{D}^{k,1}(H)$ of D in $L^2(\Omega; H)$, then from [Pri15a] we have

$$(Dv)^k v_s := \int_0^T \dots \int_0^T (D_{s_k} v_s)^* (D_{s_{k-1}} v_{s_k})^* \dots (D_{s_1} v_{s_2})^* v_{s_1} ds_1 \dots ds_k,$$

where $s \in [0, T]$, $k \geq 1$ and $(Dv)^*$ is the adjoint of Dv such that $\langle (Dv)^* u, h \rangle = \langle u, (Dv) h \rangle$, $u, h \in H$.

Theorem 2.4. Suppose that the process $v \in \bigcap_{k \geq 1} \mathbb{D}^{k,1}(H)$ is $(\mathcal{F}_s)$ -adapted and satisfies the following condition:

$$\left\langle v_s, (Dv)^k v_s \right\rangle_{\mathbb{R}^d} = 0, \quad s \in [0, T], \quad k \geq 1,$$

then

$$\mathbb{E} \left[\exp \left(i \int_0^T v_s dW_s \right) \middle| (|v_s|)_{s \in [0, T]} \right] = \exp \left(-\frac{1}{2} \int_0^T |v_s|^2 ds \right), \quad (2.2)$$

if $\frac{1}{2} \int_0^T |v_s|^2 ds$ is exponentially integrable.

Proof. From either Definition 1 of [Pri13] or Definition 4.1 of [Pri15a], we have for every $G \in \mathbb{D}^{2,1}$ and $k \geq 1$,

$$\Gamma_k^v G := \mathbb{1}_{k \geq 2} G \left\langle v, (Dv)^{k-2} v \right\rangle_H + \left\langle DG, (Dv)^{k-1} v \right\rangle_H. \quad (2.3)$$

Furthermore, for every $v \in \mathbb{D}^{2,1}(H)$ such that for all $i_1 + \dots + i_k \leq m$, $k = 1, \dots, m$, $\Gamma_{i_1}^v \dots \Gamma_{i_k}^v \mathbb{1}$ has finite expectation, we have by either Theorem 1 of [Pri13] or Proposition 4.3 of [Pri15a], the following identity for $G \in \mathbb{D}^{2,1}$:

$$\mathbb{E} \left[G \left(\int_0^T v_s dW_s \right)^m \right] = m! \sum_{j=1}^m \sum_{\substack{i_1 + \dots + i_j = m \\ i_1 \geq 1, \dots, i_j \geq 1}} \frac{\mathbb{E} \left[\Gamma_{i_1}^v \dots \Gamma_{i_j}^v G \right]}{i_1 (i_1 + i_2) \dots (i_1 + \dots + i_j)}, \quad (2.4)$$

Next, since $\left\langle v_s, (Dv)^k v_s \right\rangle_{\mathbb{R}^d} = 0$, for $s \in [0, T]$ and $k \geq 1$, we have for all $g \in \mathcal{C}_b^1(\mathbb{R})$ and $k \geq 1$,

$$\begin{aligned}
& \Gamma_k^v g \left(\int_a^b |v_s|^2 ds \right) \\
& \stackrel{(2.3)}{=} \mathbb{1}_{k=2} \left(\int_0^T |v_s|^2 ds \right) g \left(\int_a^b |v_s|^2 ds \right) \\
& \quad + g^{(1)} \left(\int_a^b |v_s|^2 ds \right) \int_0^T \left\langle D_s \int_a^b |v_t|^2 dt, (Dv)^{k-1} v_s \right\rangle_{\mathbb{R}^d} ds \\
& = \mathbb{1}_{k=2} \left(\int_0^T |v_s|^2 ds \right) g \left(\int_a^b |v_s|^2 ds \right) \\
& \quad + 2g^{(1)} \left(\int_a^b |v_s|^2 ds \right) \int_a^b \left\langle v_s, (Dv)^k v_s \right\rangle_{\mathbb{R}^d} ds \\
& = \mathbb{1}_{k=2} \left(\int_0^T |v_s|^2 ds \right) g \left(\int_a^b |v_s|^2 ds \right), \quad 0 \leq a \leq b \leq T.
\end{aligned}$$

On iteration, we obtain

$$\Gamma_{i_1}^v \dots \Gamma_{i_j}^v G = \mathbb{1}_{i_1=\dots=i_j=2} \left(\int_0^T |v_s|^2 ds \right)^j G, \quad i_1, \dots, i_j \geq 1, \quad j \geq 1, \quad (2.5)$$

for all random variable G that have the form

$$G = g \left(\int_{a_1}^{b_1} |v_s|^2 ds, \dots, \int_{a_n}^{b_n} |v_s|^2 ds \right), \quad 0 \leq a_l \leq b_l \leq T, \quad l = 1, \dots, n,$$

where $g \in \mathcal{C}_b^1(\mathbb{R}^n)$. Now, from Equations (2.4) and (2.5), we find

$$\begin{aligned}
& \mathbb{E} \left[\exp \left(i \int_0^T v_s dW_s \right) G \right] \\
& = \sum_{m=0}^{+\infty} \frac{(-1)^m}{2m!} \mathbb{E} \left[\left(\int_0^T v_s dW_s \right)^{2m} G \right] \\
& \quad + \sum_{m=0}^{+\infty} \frac{(i)^{2m+1}}{(2m+1)!} \mathbb{E} \left[\left(\int_0^T v_s dW_s \right)^{2m+1} G \right] \\
& \stackrel{(2.4),(2.5)}{=} \sum_{m=0}^{+\infty} \frac{(-1)^m}{m! 2^m} \mathbb{E} \left[\left(\int_0^T |v_s|^2 ds \right)^m G \right] \\
& = \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |v_s|^2 ds \right) G \right] \quad (2.6)
\end{aligned}$$

Taking conditional expectation of Equation (2.6) on $(|v_s|)_{s \in [0, T]}$ we proved Equation (2.2). $\square$

A similar argument for the proof of Theorem 2.4 can be used to obtain the following theorem.

Theorem 2.5. *Suppose that $v \in \bigcap_{k \geq 1} \mathbb{D}^{k,1}(H)$ is $(\mathcal{F}_s)$ -adapted and satisfies the following condition:*

$$\left\langle v, (Dv)^k v \right\rangle_H = 0, \quad k \geq 1,$$

then

$$\mathbb{E} \left[\exp \left(i \int_0^T v_s dW_s \right) \middle| \int_0^T |v_s|^2 ds \right] = \exp \left(-\frac{1}{2} \int_0^T |v_s|^2 ds \right),$$

if $\frac{1}{2} \int_0^T |v_s|^2 ds$ is exponentially integrable.

Proof. The proof is similar to Theorem 2.4 except that we take $a = a_1 = \dots = a_n = 0$, $b = b_1 = \dots = b_n = T$ and take conditional expectation of Equation (2.6) given $\int_0^T |v_s|^2 ds$ instead. $\square$

We note that if $v_s = R_s h$, $s \in [0, T]$, where R is a random adapted (or quasi-nilpotent) isometry of H , then we find that $\int_0^T |v_s|^2 ds = \int_0^T |h(s)|^2 ds$ is deterministic, and thus

$$\left\langle v, (Dv)^k v \right\rangle_H = \frac{1}{2} \left\langle D \langle v, v \rangle_H, (Dv)^{k-1} v \right\rangle_H = 0, \quad k \geq 1,$$

which by Theorem 2.5 shows that $\int_0^T (R_s h) dW_s$ has a central normal distribution with variance $\int_0^T |h(s)|^2 ds$, as in Theorem 2.1-(b) of [ÜZ95].

Furthermore, when $\int_0^T |v_s|^2 ds$ is random, Theorems 2.4 and 2.5 still holds, for example, when $(v_s)_{s \in [0, T]}$ is of the form $v_s = f(W_s)$, $s \in [0, T]$, where $f \in \mathcal{C}_b^1(\mathbb{R}^d; \mathbb{R}^d)$ satisfies the condition

$$\left\langle f(y), ((Df(y))^*)^k f(y) \right\rangle = 0, \quad (2.7)$$

for every $y \in \mathbb{R}^d$, $k \geq 1$.

In the next section, we check that when f is a linear mapping having the form $f(y) = Ay$, for every $y \in \mathbb{R}^d$. We have concrete examples from [Yor79], that satisfies the condition (2.7).

2.4 Vanishing of $A^\top A^2$

Let $(v_s)_{s \in [0, T]} := (AW_s)_{s \in [0, T]}$. We recall that $D_t W_s = \mathbb{1}_{[0, s]}(t)$, thus under Yor's condition $A^\top A^2 = 0$ (see [Yor79]) and Proposition A.7 of Appendix A, we have

$$\begin{aligned}
 & \left\langle v_s, (Dv)^k v_s \right\rangle_{\mathbb{R}^d} \\
 &= \int_0^T \dots \int_0^T \left\langle (D_{s_k} v_s)^* (D_{s_{k-1}} v_{s_k})^* \dots (D_{s_1} v_{s_2})^* v_{s_1}, v_s \right\rangle_{\mathbb{R}^d} ds_1 \dots ds_k \\
 &= \int_0^s \int_0^{s_k} \dots \int_0^{s_2} \left\langle (A^\top)^k AW_s, AW_s \right\rangle ds_1 \dots ds_k \\
 &= \int_0^s \int_0^{s_k} \dots \int_0^{s_2} \left\langle W_s, A^\top A^2 (A)^{k-1} W_s \right\rangle ds_1 \dots ds_k \\
 &= 0, \quad s \in [0, T], \quad k \geq 1.
 \end{aligned} \tag{2.8}$$

Applying Theorem 2.4, we yield the next corollary containing the Yor's condition, which includes the particular case of 2-nilpotent matrices.

Corollary 2.1. *Suppose that we have $A^\top A^2 = 0$, then*

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| (|AW_s|)_{s \in [0, T]} \right] = \exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right) \tag{2.9}$$

and

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| \int_0^T |AW_s|^2 ds \right] = \exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right). \tag{2.10}$$

Proof. From Equation (2.8), we have $\left\langle v_s, (Dv)^k v_s \right\rangle_{\mathbb{R}^d} = 0$ whenever $A^\top A^2 = 0$. Hence, we apply Theorem 2.4 to obtain Equation (2.9) and Theorem 2.5 to obtain (2.10). $\square$

Note that we have from Corollary 2 of [Yor79] that the filtration generated by $(|AW_s|)_{s \in [0, T]}$ coincides with the filtration $(\mathcal{F}_s^m)_{s \in [0, T]}$ generated by m independent Brownian motions, where m is the number of distinct nonzero eigenvalues of $A^\top A$.

2.5 Anti-Symmetric $A^\top A^2$

In this section, we see that the condition $A^\top A^2 = 0$ can be relaxed to the condition that $A^\top A^2$ is anti-symmetric by using stochastic calculus if we have $A^\top A$ is proportional to a projection, i.e. $A^\top A$ has only one distinct nonzero eigenvalue. We begin with the following lemma, which is a variation of Corollary 2 of [Yor79].

Lemma 2.3. *If we assume that $A^\top A$ has a unique nonzero eigenvalue λ and $A^\top A^2$ is anti-symmetric. Then the following processes:*

$$X_s^1 := \frac{1}{\sqrt{\lambda}} \int_0^s \frac{AW_t}{|AW_t|} dAW_t, \quad \text{and} \quad X_s^2 := \int_0^s \frac{AW_t}{|AW_t|} dW_t, \quad s \in [0, T],$$

(2.11)

are 2 independent Brownian motions.

Proof. We know that $A^\top A$ is symmetric so we can rewrite it as $A^\top A = R^\top DR$, where R is orthogonal and D is diagonal. Moreover, we recall that the standard Brownian motion is invariant under orthogonal transformation, thus, $(RW_s)_{s \in [0, T]}$ is also a standard Brownian motion. This means that we can let $A^\top A = (\lambda_i \mathbb{1}_{1 \leq i=r \leq m})_{1 \leq i, r \leq d}$ with $\lambda_j > 0$ when $1 \leq j \leq m$. Clearly $(X_s^2)_{s \in [0, T]}$ is a standard Brownian motion, since

$$d \langle X^2, X^2 \rangle_s = \frac{\langle AW_s, AW_s \rangle}{|AW_s|^2} ds = ds.$$

We also have $(X_s^1)_{s \in [0, T]}$ is a Brownian motion, as we have $dX_s^1 = \frac{\lambda^{-1/2}}{|AW_s|} \sum_{j=1}^m \lambda_j W_s^j dW_s^j$ and so

$$d \langle X^1, X^1 \rangle_s = \frac{(\lambda_1 W_s^1)^2 + \dots + (\lambda_m W_s^m)^2}{\lambda (\lambda_1 (W_s^1)^2 + \dots + \lambda_m (W_s^m)^2)} ds = ds.$$

The last equality comes from the fact that we only have a unique nonzero eigenvalue. Now, in addition, we have

$$\begin{aligned} d\langle X^1, X^2 \rangle_s &= \frac{\langle A^\top AW_s, AW_s \rangle}{\sqrt{\lambda} |AW_s|^2} \\ &= \frac{\langle W_s, A^\top A^2 W_s \rangle}{\sqrt{\lambda} |AW_s|^2} ds \\ &= 0, \end{aligned}$$

as $A^\top A^2 = 0$. Therefore, $(X_s^1)_{s \in [0, T]}$ and $(X_s^2)_{s \in [0, T]}$ are 2 independent Brownian motions. $\square$

The following corollary replaces the vanishing hypothesis of Corollary 2.1 with $A^\top A^2$ being anti-symmetric

Corollary 2.2. *Suppose $A^\top A^2$ is anti-symmetric and assume that $A^\top A$ has only a unique nonzero eigenvalue λ . Then*

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| (|AW_s|)_{s \in [0, T]} \right] = \exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right). \quad (2.12)$$

Proof. We recall that by Corollary 2 of [Yor79], $(|AW_s|)_{s \in [0, T]}$ generate a filtration that coincides with the filtration $(\mathcal{F}_s^1)_{s \in [0, T]}$ generated by $(X_s^1)_{s \in [0, T]}$. Now, if we let $S_s := |AW_s|^2$, then by Itô's formula, we have

$$\begin{aligned} S_s &= 2 \int_0^s AW_t dW_t + \text{Tr}(A^\top A) s \\ &\stackrel{(2.11)}{=} 2 \int_0^s \sqrt{\lambda S_t} dX_t^1 + m\lambda s, \quad s \in [0, T]. \end{aligned}$$

Hence, we see that $(|AW_s|)_{s \in [0, T]}$ is $(\mathcal{F}_s^1)_{s \in [0, T]}$ -adapted. Moreover, by Lemma 2.3, we have that it is independent of $(X_s^2)_{s \in [0, T]}$. Therefore,

$$\int_0^T AW_s dW_s = \int_0^T |AW_s| dX_s^2$$

has a central normal distribution with variance $\int_0^T |AW_s|^2 ds$ given $\mathcal{F}_T^1$. Now, by Itô isometry we obtain Equation (2.12). $\square$

The next corollary is the direct consequence of either Corollaries 2.1 or 2.2.

Corollary 2.3. *If we assume either the assumptions of Corollaries 2.1 or 2.2 i.e. $A^\top A^2 = 0$ or $A^\top A^2$ is anti-symmetric and $A^\top A$ has a unique nonzero eigenvalue, then we have*

$$\begin{aligned} & \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| |AW_s| \right] \\ &= \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right) \middle| |AW_s| \right], \end{aligned} \quad (2.13)$$

where $s \in [0, T]$.

Proof. Under the assumptions of Corollaries 2.1 or 2.2, we obtain Equation (2.13) by taking the conditional expectation of Equations (2.9) and (2.12) respectively given $|AW_s|$, where $s \in [0, T]$. $\square$

2.6 Commutation with Orthogonal matrices

If in addition to the assumptions of Corollaries 2.1 or 2.2, we assume that A commutes with orthogonal matrices, then our identities are still valid given AW_s , where $s \in [0, T]$.

Corollary 2.4. *Under the assumptions of Corollaries 2.1 or 2.2, i.e. $A^\top A^2 = 0$ or $A^\top A^2$ is anti-symmetric and $A^\top A$ has only a unique nonzero eigenvalue and if in addition A commutes with the orthogonal matrices, then*

$$\begin{aligned} & \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| AW_s \right] \\ &= \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right) \middle| AW_s \right], \end{aligned} \quad (2.14)$$

where $s \in [0, T]$.

Proof. Since Brownian motion is invariant under orthogonal transformation, for all $d \times d$ orthogonal matrix R , we check that we have

$$\begin{aligned} & \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| AW_s = y \right] \\ &= \mathbb{E} \left[\exp \left(i \int_0^T AR^\top W_s dR^\top W_s \right) \middle| AR^\top W_s = y \right] \\ &= \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| AW_s = Ry \right], \end{aligned}$$

for all $y \in \mathbb{R}^d$. This implies that

$$\begin{aligned} & \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| |AW_s| \right] \\ &= \mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| AW_s \right]. \end{aligned}$$

Applying similar argument for the right hand side of Equation (2.14), we conclude our proof by utilizing Equation (2.13). $\square$

2.7 Anti-symmetric orthogonal A

In the next corollary we assume the matrix A is anti-symmetric and orthogonal, where we notice that under those assumptions, $A^\top A^2$ is anti-symmetric. Hence, we have the following corollary.

Corollary 2.5. *If A is anti-symmetric and orthogonal, then we have*

$$\mathbb{E} \left[\exp \left(i \int_0^T AW_s dW_s \right) \middle| W_s \right] = \mathbb{E} \left[\exp \left(-\frac{1}{2} \int_0^T |AW_s|^2 ds \right) \middle| W_s \right], \quad (2.15)$$

where $s \in [0, T]$.

Proof. Under the assumptions that A is anti-symmetric and orthogonal, we have

$$\begin{aligned} (A^\top A^2)^\top &= A^\top A^\top A \stackrel{(\text{A orthogonal})}{=} A^\top \stackrel{(\text{A anti-symmetric})}{=} -A \\ &\stackrel{(\text{A orthogonal})}{=} -A^\top A^2. \end{aligned}$$

Therefore $A^\top A^2$ is anti-symmetric. Moreover, $A^\top A = \text{Id}_{\mathbb{R}^d}$ has only one unique nonzero eigenvalue $\lambda = 1$ and we have the invariant of Brownian motion under orthogonal transformation, hence we prove Equation (2.15) from (2.14). $\square$

Corollary 2.5 holds in particular when $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, in which case we have $A^\top A = \text{Id}_{\mathbb{R}^2}$ has only one unique nonzero eigenvalue $\lambda = 1$ and $A^\top A^2 = A$ is anti-symmetric. Therefore, we recover the result of [Yor80] which is used to show that Equation (2.15) holds when A is anti-symmetric but not necessarily orthogonal in Theorem 2.1 of [DEM16].

We also see that using different methods, we are able to derive sufficient conditions different from that found in [Oco93]. Thus, the sufficient conditions for A are sufficient but might be different from the necessary conditions.

Chapter 3

Conditional Stein approximation for Itô and Skorohod Integral¹

Abstract

In this Chapter, we use results from Chapter 2 to derive conditional Edgeworth-type expansions for Skorohod and Itô integrals with respect to Brownian motion, based on cumulant operators defined by the Malliavin calculus. As a consequence we obtain conditional Stein approximation bounds for multiple stochastic integrals.

3.1 Introduction

Let $(W_t)_{t \in [0, T]}$ be a d -dimensional Brownian motion generating the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ on the Wiener space Ω . It has been shown in Theorem 2.1 of [DEM16], that given W_t , the stochastic integral $\int_0^T AW_s dW_s$ is Gaussian $\mathcal{N}\left(0, \int_0^T |AW_s|^2 ds\right)$ -distributed given $\int_0^T |AW_s|^2 ds$ when the $d \times d$ matrix A is skew-symmetrix, as an extension of results of [Yor80] in the case of Lévy's stochastic area with $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Related results have been recently obtained in [PS15] (see Chapter 2) for the conditional distribution of $\int_0^T AW_s dW_s$ given $\int_0^T |AW_s|^2 ds$ when A is a 2-nilpotent $d \times d$ matrix, in connection with results of [Yor79] which shows that the filtration $(\mathcal{F}_t^k)_{t \in [0, T]}$ of $t \mapsto \int_0^t AW_s dW_s$ is generated by k independent Brownian motions, where k is the number of distinct eigenvalues of $A^\top A$.

¹This chapter is based on works with Associate Professor Nicolas Privault, NTU, Singapore.

More generally, this type of result has been extended in [PS15] (see Chapter 2) to stochastic integrals of the form $\int_0^T u_t dW_t$ where $(u_t)_{t \in [0, T]}$ is an $(\mathcal{F}_t)$ -adapted process, under conditions formulated in terms of the Malliavin calculus, based on the cumulant-moment formula of [Pri13] and [Pri15a]. Namely, sufficient conditions have been given for $\int_0^T u_t dW_t$ to be Gaussian $\mathcal{N}\left(0, \int_0^T |u_t|^2 dt\right)$ -distributed when we have $\int_0^T |u_t|^2 dt$, cf. Theorem 2 therein.

In this chapter we provide conditional estimates on the distance between the law of $\int_0^T u_t dW_t$ and the Gaussian $\mathcal{N}\left(0, \int_0^T |u_t|^2 dt\right)$ distribution given $\int_0^T |u_t|^2 dt$ using the Stein method on the Wiener space. For this, we rely on a conditional Edgeworth type expansions for random variables represented as the Itô or Skorohod integral $F = \delta(u)$ of a process u on the Wiener space, extending results of [Pri15b] to the conditional setting. This approach relies on properties of the Skorohod integral operator δ , which coincides with the Itô stochastic integral with respect to d -dimensional Brownian motion on the square-integrable adapted processes.

Letting $H = L^2(\mathbb{R}_+; \mathbb{R}^d)$, we consider the standard Sobolev spaces of real-valued, resp. H -valued, functionals $\mathbb{D}^{p,k}$, resp. $\mathbb{D}^{p,k}(H)$, $p, k \geq 1$, for the Malliavin gradient D on the Wiener space, cf. [Nua06] for definitions. Recall that the Skorohod operator δ is the adjoint of the gradient D through the duality relation

$$\mathbb{E}[F\delta(v)] = \mathbb{E}[\langle DF, v \rangle_H], \quad F \in u \in \mathbb{D}^{2,1}, \quad v \in \mathbb{D}^{2,1}(H), \quad (3.1)$$

and we have the commutation relation

$$D_t \delta(u) = u(t) + \delta(D_t u), \quad t \in \mathbb{R}_+,$$

provided that $u \in \mathbb{D}^{2,1}(H)$ and $D_t u \in \mathbb{D}^{2,1}(H)$, dt -a.e., cf. Proposition 1.3.2 of [Nua06]. The duality relation (3.1) shows that we have

$$\begin{aligned} \mathbb{E}[F\delta(u) f(\delta(u))] &= \mathbb{E}[F\langle u, u \rangle f'(\delta(u))] + \mathbb{E}[\langle u, DF \rangle f(\delta(u))] \\ &\quad + \mathbb{E}[F f'(\delta(u)) \langle u, \delta(Du) \rangle], \end{aligned}$$

for Lipschitz f , provided that $u \in \mathbb{D}^{2,1}(H) \cap \mathbb{D}^{2,2}(H)$. Applying this relation to $F = g(\langle u, u \rangle)$ where $g : (0, \infty) \rightarrow (0, \infty)$ is a measurable

function, under the condition $\langle u, (Du)u \rangle = 0$ we have

$$\begin{aligned} \mathbb{E}[F\delta(u)f(\delta(u))] &= \mathbb{E}[F\langle u, u \rangle f'(\delta(u))] + \mathbb{E}[Ff'(\delta(u))\langle u, \delta(Du) \rangle] \\ &\quad + \mathbb{E}[\langle u, D\langle u, u \rangle \rangle f(\delta(u))g'(\langle u, u \rangle)] \\ &= \mathbb{E}[F\langle u, u \rangle f'(\delta(u))] + \mathbb{E}[F\langle u, \delta(Du) \rangle f'(\delta(u))] \\ &\quad + 2\mathbb{E}[\langle u, (Du)u \rangle f(\delta(u))g'(\langle u, u \rangle)] \\ &= \mathbb{E}[F\langle u, u \rangle f'(\delta(u))] + \mathbb{E}[F\langle u, \delta(Du) \rangle f'(\delta(u))], \end{aligned}$$

which yields

$$\mathbb{E}_{|u|}[\delta(u)f(\delta(u)) - \langle u, u \rangle f'(\delta(u))] = \mathbb{E}_{|u|}[\langle u, \delta(Du) \rangle f'(\delta(u))], \quad (3.2)$$

for $u \in \mathbb{D}^{2,1}(H)$, $F \in \mathbb{D}^{2,1}$ and $f \in \mathcal{C}_b^1(\mathbb{R})$, where

$$\mathbb{E}_{|u|}[F] := \mathbb{E}[F | \langle u, u \rangle]$$

denote the conditional expectation given $\langle u, u \rangle$.

Let now $\mathcal{N}(0, g(\|u\|))$ denote a centered Gaussian random variable with variance $g(\|u\|)$, where $g : (0, \infty) \rightarrow (0, \infty)$ is a measurable function. Applying the above relation (3.2) to the solution f_x of the Stein equation

$$\mathbf{1}_{(-\infty, x]}(z) - \Phi_{g(\|u\|)}(x) = g(\|u\|)f'_x(z) - zf_x(z), \quad z \in \mathbb{R},$$

satisfying the bounds $\|f_x\|_\infty \leq \sqrt{2\pi}/4$ and $\|f'_x\|_\infty \leq 1/\sqrt{g(\|u\|)}$, cf. Lemma 2.2-(v) of [CGS11], yields the conditional expansion

$$\begin{aligned} P(\delta(u) \leq x | \|u\|) &= \Phi_{g(\|u\|)}(x) + \mathbb{E}_{|u|}[(g(\|u\|) - \langle u, u \rangle)f'_x(\delta(u))] \\ &\quad - \mathbb{E}[\langle u, \delta(Du) \rangle f'_x(\delta(u))], \end{aligned}$$

$x \in \mathbb{R}$, around the Gaussian CDF $\Phi_{g(\|u\|)}(x)$, with $u \in \mathbb{D}^{2,1}(H)$.

In Section 3.2 we expand this approach to Edgeworth type expansion of all orders, based on a family of cumulant operators that are associated to the process u .

In Section 3.3, we derive conditional Stein approximation bounds for the distance between $\delta(u)$ and the Gaussian distribution with variance $g(\|u\|)$.

Section 3.4 treats the case of double stochastic integrals, which includes the quadratic functional $\int_0^T AW_s dW_s$ as a particular case.

3.2 Conditional Edgeworth type expansions

Given $u \in \mathbb{D}^{2,1}(H)$ and $k \geq 1$, we define the operator composition $(Du)^k$ in the sense of matrix powers with continuous indices, namely, $(Du)^k$ denotes the random operator on H almost surely defined by

$$(Du)^k h_s = \int_0^{+\infty} \cdots \int_0^{+\infty} (D_{t_k} u_s D_{t_{k-1}} u_{t_k} \cdots D_{t_1} u_{t_2}) h_{t_1} dt_1 \cdots dt_k,$$

where $s \in \mathbb{R}_+$ and $h \in H$, cf. e.g. § 7 of [Pri15a], [Pri13] and [Pri12] for details. The adjoint D^*u of Du on H satisfies

$$\langle (Du)v, h \rangle = \langle v, (D^*u)h \rangle, \quad h, v \in H,$$

with $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_H$, and is given by

$$(D^*u)v_s = \int_0^{+\infty} (D_s u_t) v_t dt, \quad s \in \mathbb{R}_+, \quad v \in L^2(W; H).$$

The next proposition reformulates Proposition 2.1 of [Pri15b] in a conditional setting given the random value of $\langle u, u \rangle$, whereas the case where $\langle u, u \rangle$ is deterministic (random isometries) has been treated in [Pri15b] based on the relation $\langle u, (Du)^k u \rangle = \langle (Du)^{k-1} u, D\langle u, u \rangle \rangle / 2$, $u \in \mathbb{D}^{k+2,1}(H)$, $k \geq 1$.

Proposition 3.1. *Let $n \geq 1$ and assume that $u \in \mathbb{D}^{k,2}(H)$ for all $k = 1, \dots, n+2$ and $\langle u, (Du)^k u \rangle = 0$ for $k = 1, \dots, n$. Then for all $f \in \mathcal{C}_b^{n+1}(\mathbb{R})$ we have*

$$\begin{aligned} \mathbb{E}_{|u|}[\delta(u) f(\delta(u))] &= \mathbb{E}_{|u|}[(\langle u, u \rangle + \text{trace}(Du)^2) f'(\delta(u))] \\ &\quad + \sum_{k=2}^n \mathbb{E}_{|u|}[\langle D^*u, D((Du)^{k-1}u) \rangle f^{(k)}(\delta(u))] \\ &\quad + \mathbb{E}_{|u|}[\langle (Du)^n u, \delta(Du) \rangle f^{(n+1)}(\delta(u))]. \end{aligned} \quad (3.3)$$

Proof. From Proposition 2.1 of [Pri15b] we have

$$\begin{aligned} \mathbb{E}[F\delta(u)f(\delta(u))] &= \sum_{k=0}^n \mathbb{E}[f^{(k)}(\delta(u))\Gamma_{k+1}^u F] \\ &\quad + \mathbb{E}[F\langle(Du)^n u, \delta(Du)\rangle f^{(n+1)}(\delta(u))], \end{aligned} \quad (3.4)$$

where

$$\Gamma_k^u : \mathbb{D}^{2,1} \longrightarrow L^2(\Omega), \quad k \geq 1,$$

is defined for $u \in \mathbb{D}^{k,2}(H)$, by $\Gamma_1^u F := \langle u, DF \rangle$ and

$$\Gamma_k^u F := F\langle(Du)^{k-2} u, u\rangle + F\langle D^* u, D((Du)^{k-2} u)\rangle + \langle(Du)^{k-1} u, DF\rangle,$$

where $k \geq 2$. Next, from §2 of [PS15], we have for F of the form $F = g\left(\int_0^T |u_t|^2 dt\right)$, $g \in C_b^1(\mathbb{R})$ and $k \geq 1$:

$$\begin{aligned} \Gamma_k^u F &= \mathbf{1}_{\{k=2\}} \langle u, u \rangle g\left(\int_0^T |u_t|^2 dt\right) \\ &\quad + g'\left(\int_0^T |u_t|^2 dt\right) \int_0^T \langle D_t \int_0^T |u_s|^2 ds, (Du)^{k-1} u_t \rangle_{\mathbb{R}^d} dt \\ &\quad + g\left(\int_0^T |u_t|^2 dt\right) \langle D^* u, D((Du)^{k-2} u)\rangle \\ &= \mathbf{1}_{\{k=2\}} \langle u, u \rangle F + 2g'\left(\int_0^T |u_t|^2 dt\right) \int_0^T \langle u_s, (Du)^k u_s \rangle_{\mathbb{R}^d} dt \\ &\quad + \mathbf{1}_{\{k \geq 2\}} F \langle D^* u, D((Du)^{k-2} u)\rangle \\ &= \mathbf{1}_{\{k=2\}} \langle u, u \rangle F + \mathbf{1}_{\{k \geq 2\}} \langle D^* u, D((Du)^{k-2} u)\rangle F. \end{aligned} \quad (3.5)$$

Hence from (3.4) we find

$$\begin{aligned} \mathbb{E}[F\delta(u)f(\delta(u))] &= \mathbb{E}\left[F(\langle u, u \rangle + \text{trace}(Du)^2) f'(\delta(u))\right] \\ &\quad + \sum_{k=2}^n \mathbb{E}\left[F\langle D^* u, D((Du)^{k-1} u)\rangle f^{(k)}(\delta(u))\right] \\ &\quad + \mathbb{E}\left[F\langle(Du)^n u, \delta(Du)\rangle f^{(n+1)}(\delta(u))\right], \end{aligned}$$

where we used the relation

$$\Gamma_2^u \mathbf{1} = \langle u, u \rangle + \langle D^* u, Du \rangle_{H \otimes H} = \langle u, u \rangle + \text{trace} (Du)^2. \quad (3.6)$$

□

Proposition 3.1 also covers the following particular settings.

- (i) Quasi-nilpotent processes. Given $n \geq 0$, when $\text{trace} (Du)^k = 0$ for all $k = 2, \dots, n+1$, under the conditions of Proposition 3.1 we have

$$\begin{aligned} \mathbb{E}_{|u|}[\delta(u) f(\delta(u))] &= \langle u, u \rangle \mathbb{E}_{|u|}[f'(\delta(u))] \\ &\quad + \mathbb{E}_{|u|}[f^{(n+1)}(\delta(u)) \langle (Du)^n u, \delta(Du) \rangle], \end{aligned}$$

which recovers the conditional Gaussian integration by parts formula

$$\mathbb{E}_{|u|}[\delta(u) f(\delta(u))] = \langle u, u \rangle \mathbb{E}_{|u|}[f'(\delta(u))],$$

showing that $\delta(u) \simeq \mathcal{N}\left(0, \int_0^T |u_t|^2 dt\right)$ given $\int_0^T |u_t|^2 dt$. This setting includes the particular cases where $(u_t)_{t \in \mathbb{R}_+}$ is an $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ -adapted process, cf. e.g. Lemma 3.5 of [Pri12] and references therein, in which case $\delta(u)$ coincides with the Itô integral of u , cf. Proposition 1.3.11 of [Nua06].

- (ii) Multiple stochastic integrals. Taking $u_t := I_{m-1}(f_m(*, t))$, where $m \geq 1$ and f_m is a symmetric square-integrable function on $\mathbb{R}_+^m$, we have $\delta(u) = I_m(f_m)$ and

$$\delta(D_t u) = (m-1) I_{m-1}(f_m(*, t)) = (m-1) u_t, \quad t \in \mathbb{R}_+.$$

Hence, under the conditions of Proposition 3.1 applied to $u_t = I_{m-1}(f_m(*, t))$, we get

$$\begin{aligned} \mathbb{E}_{|u|}[I_m(f_m) f(I_m(f_m))] &= \langle u, u \rangle \mathbb{E}_{|u|}[f'(I_m(f_m))] \\ &\quad + \sum_{k=1}^n \mathbb{E}_{|u|}\left[\langle D^* u, D((Du)^{k-1} u) \rangle f^{(k)}(I_m(f_m))\right]. \end{aligned}$$

Remark 3.1. By replacing F in Equation (3.5) of the proof of Proposition 3.1 with F of the form

$$F = g \left(\int_{a_i}^{b_i} |u_t|^2 dt, \dots, \int_{a_d}^{b_d} |u_t|^2 dt \right), \quad 0 \leq a_i \leq b_i \leq T, \quad i = 1, \dots, d,$$

where $g \in \mathcal{C}_b^1(\mathbb{R}^d)$, we can rewrite (3.3) by conditioning with respect to $(u_t)_{t \in [0, T]}$, under the stronger condition $\langle u_t, (Du)^k u_t \rangle_{\mathbb{R}^d}$, $t \in [0, T]$, $k = 1, \dots, n + 2$.

3.3 Conditional Stein approximation

From now on we work with $d = 1$ and a one-dimensional Brownian motion $(W_t)_{t \in \mathbb{R}_+}$. Given $h : \mathbb{R} \rightarrow \mathbb{R}$ an absolutely continuous function with bounded derivative, the functional equation

$$h(z) - \mathbb{E}_{|u|}[h(\mathcal{N}_{g(\|u\|)})] = g(\|u\|) f'(z) - z f(z), \quad z \in \mathbb{R}, \quad (3.7)$$

has a solution $f_h \in \mathcal{C}_b^1(\mathbb{R})$ which is twice differentiable and satisfies the bounds

$$\|f'_h\|_\infty \leq \|h'\|_\infty / \sqrt{g(\|u\|)} \quad \text{and} \quad \|f''_h\|_\infty \leq 2 \|h'\|_\infty / g(\|u\|), \quad x \in \mathbb{R},$$

cf. Lemma 1.2-(v) of [NP09]. Let

$$d_{|u|}(F, G) = \sup_{h \in \mathcal{L}} |\mathbb{E}_{|u|}[h(F)] - \mathbb{E}_{|u|}[h(G)]|$$

denote the conditional Wasserstein distance between the conditional laws of F and G , where $\mathcal{L}$ denotes the class of 1-Lipschitz functions.

Proposition 3.2. Let $u \in \bigcap_{k=1}^3 \mathbb{D}^{k,2}(H)$, such that $\langle u, (Du)u \rangle = 0$. We have

$$\begin{aligned} d_{|u|}(\delta(u), \mathcal{N}_{g(\|u\|)}) &\leq \frac{1}{\sqrt{g(\|u\|)}} |g(\|u\|) - \text{Var}_{|u|}[\delta(u)]| \\ &\quad + \frac{2}{g(\|u\|)} \mathbb{E}_{|u|}[|\langle (Du)u, \delta(Du) \rangle|]. \end{aligned} \quad (3.8)$$

Proof. Applying Proposition 3.1 with $n = 1$ shows that

$$\begin{aligned}\mathbb{E}_{|u|}[\delta(u) f(\delta(u))] &= \mathbb{E}_{|u|}[(\langle u, u \rangle + \text{trace}(Du)^2) f'(\delta(u))] \\ &\quad + \mathbb{E}_{|u|}[\langle (Du)u, \delta(Du) \rangle f''(\delta(u))],\end{aligned}$$

hence for any continuous function $h : \rightarrow [0, 1]$, denoting by f_h the solution to (3.7) we have

$$\begin{aligned}&|\mathbb{E}_{|u|}[h(\delta(u))] - \mathbb{E}_{|u|}[h(\mathcal{N}_{g(\|u\|)})]| \\ &= |\mathbb{E}_{|u|}[\delta(u) f_h(\delta(u)) - g(\|u\|) f'_h(\delta(u))]| \\ &= |\mathbb{E}_{|u|}[(\langle u, u \rangle - g(\|u\|)) f'_h(\delta(u)) + \text{trace}(Du)^2 f'_h(\delta(u)) \\ &\quad + \langle (Du)u, \delta(Du) \rangle f''_h(\delta(u))]| \\ &\leq \frac{\|h'\|_\infty}{\sqrt{g(\|u\|)}} |\langle u, u \rangle - g(\|u\|) + \mathbb{E}_{|u|}[\text{trace}(Du)^2]| \\ &\quad + 2 \frac{\|h'\|_\infty}{g(\|u\|)} \mathbb{E}_{|u|}[\langle (Du)u, \delta(Du) \rangle],\end{aligned}$$

which yields (3.8) by (3.6) and the conditional Skorohod isometry

$$\text{Var}_{|u|}[\delta(u)] = \mathbb{E}_{|u|}[\delta(u)^2] = \langle u, u \rangle + \mathbb{E}_{|u|}[\text{trace}(Du)^2], \quad (3.9)$$

that follows from (3.2) with $f(x) = x$. $\square$

When $(u_t)_{t \in \mathbb{R}_+}$ is a quasi-nilpotent processes, and in particular when $(u_t)_{t \in \mathbb{R}_+}$ is $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ -adapted process, we obtain the following corollary.

Corollary 3.1. Let $u \in \bigcap_{k=1}^3 \mathbb{D}^{k,2}(H)$, such that $\langle u, (Du)u \rangle = 0$ and $\text{trace}(Du)^2 = 0$. We have

$$\begin{aligned}&d_{|u|}(\delta(u), \mathcal{N}_{g(\|u\|)}) \\ &\leq \frac{1}{\sqrt{g(\|u\|)}} |\langle u, u \rangle - g(\|u\|)| + \frac{2}{g(\|u\|)} \mathbb{E}_{|u|}[\langle (Du)u, \delta(Du) \rangle].\end{aligned}$$

Proof. This follows from Proposition 3.2 and the isometry (3.9). $\square$

Note that from Section 3.3, we have

$$d_{|u|}(\delta(u), \mathcal{N}_{g(\|u\|)}) = (g(\|u\|))^{-1/2} |\langle u, u \rangle - g(\|u\|)|,$$

if, in addition to $\text{trace } (Du)^2 = 0$, the condition $\langle u, (Du)^k u \rangle = 0$ holds for all $k \geq 1$.

Regarding multiple stochastic integrals, we have the following corollary.

Corollary 3.2. *Given $f_m \in L^2(\mathbb{R}_+^m)$ a symmetric function in $m \geq 1$ variables, let $u_t := I_{m-1}(f_m(*, t))$, $t \in \mathbb{R}_+$, and assume that $\langle u, (Du)u \rangle = 0$. Then we have*

$$d_{|u|}(I_m(f_m), \mathcal{N}_{g(\|u\|)}) \leq \frac{1}{\sqrt{g(\|u\|)}} |g(\|u\|) - \text{Var}_{|u|}[I_m(f_m)]|.$$

Proof. We have $\langle (Du)u, \delta(Du) \rangle = (m-1) \langle (Du)u, u \rangle = 0$, hence the conclusion follows from Proposition 3.2. $\square$

3.4 Double stochastic integrals

In this section we take $(u_t)_{t \in \mathbb{R}_+}$ of the form $u_t = I_1(f_2(*, t))$, $t \in \mathbb{R}_+$, where the function $f_2 \in L^2(\mathbb{R}_+^2)$ may not be symmetric in its two variables. In this case we have $\delta(u) = I_2(\tilde{f}_2)$, where $\tilde{f}_2$ is the symmetrization of f_2 , with $D_s u_t = f_2(s, t)$ and $\delta(D_s u) = I_1(f_2(s, *))$, $s, t \in \mathbb{R}_+$.

Proposition 3.3. *The condition $\langle u, (Du)u \rangle = 0$ is equivalent to $(Du)^2 = 0$, i.e. to the vanishing of the contraction*

$$(f_2 \otimes_1 f_2)(t_1, t_2) := \int_0^{+\infty} f_2(t_1, s) f_2(s, t_2) ds = 0, \quad t_1, t_2 \in \mathbb{R}_+, \quad (3.10)$$

and in this case, we have

$$d_{|u|}(I_2(\tilde{f}_2), \mathcal{N}_{g(\|u\|)}) \leq \frac{1}{\sqrt{g(\|u\|)}} |g(\|u\|) - \langle u, u \rangle|.$$

In particular, $I_2(\tilde{f}_2)$ is Gaussian $\mathcal{N}_{\|u\|^2}$ -distributed given $\langle u, u \rangle$.

Proof. By Itô calculus we have

$$\begin{aligned}\langle u, (Du)u \rangle &= \int_0^{+\infty} \int_0^{+\infty} I_1(f_2(*, s)) f_2(s, t) I_1(f_2(*, t)) ds dt \\ &= \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} f_2(r, s) f_2(s, t) f_2(r, t) dr ds dt \\ &\quad + \int_0^{+\infty} \int_0^{+\infty} f_2(s, t) I_2(f_2(*, s) \otimes f_2(*, t)) ds dt\end{aligned}$$

hence the condition $\langle u, (Du)u \rangle$ is equivalent to the vanishing

$$\int_0^{+\infty} \int_0^{+\infty} f_2(t_1, s) f_2(s, t) f_2(t_2, t) ds dt = 0, \quad t_1, t_2 \in \mathbb{R}_+, \quad (3.11)$$

i.e.

$$\begin{aligned}&\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} h(t_1) f_2(t_1, s) f_2(s, t) f_2(t_2, t) g(t_2) ds dt dt_1 dt_2 \\ &= 0,\end{aligned}$$

and taking

$$g(t_2) := \int_0^{+\infty} f_2(r, t_2) h(r) dr,$$

we get

$$\begin{aligned}&\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \left(\int_0^{+\infty} f_2(t_1, s) f_2(s, t) h(t_1) dt_1 \right) \\ &\quad \left(\int_0^{+\infty} f_2(r, t_2) f_2(t_2, t) h(r) dr \right) ds dt dt_2 \\ &= \int_0^{+\infty} \left(\int_0^{+\infty} \int_0^{+\infty} f_2(r, s) f_2(s, t) h(r) dr ds \right)^2 dt = 0,\end{aligned}$$

hence

$$\int_0^{+\infty} f_2(r, s) f_2(s, t) ds = 0, \quad r, t \in \mathbb{R}_+.$$

Consequently we have $\text{trace } (f_2)^2 = 0$, hence by (3.9) we obtain that $\text{Var}_{|u|} \left[I_2 \left(\tilde{f}_2 \right) \right] = \langle u, u \rangle$. Moreover, we note that

$$\begin{aligned}
& \langle (Du) u, \delta (Du) \rangle \\
&= \int_0^{+\infty} \int_0^{+\infty} u_s D_s u_t \delta (D_t u) \, ds dt \\
&= \int_0^{+\infty} \int_0^{+\infty} I_1 (f_2 (*, s)) f_2 (s, t) I_1 (f_2 (t, *)) \, ds dt \\
&= \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} f_2 (r, s) f_2 (s, t) f_2 (t, r) \, dr ds dt \\
&+ \frac{1}{2} I_2 \left(\int_0^{+\infty} \int_0^{+\infty} f_2 (s, t) (f_2 (*, s) f_2 (t, \cdot) + f_2 (\cdot, s) f_2 (t, *)) \, ds dt \right) \\
&= 0,
\end{aligned}$$

and we conclude by Proposition 3.2. $\square$

Examples of functions satisfying (3.10) can be obtained by choosing $(e_k)_{k \geq 1}$ an orthonormal system in $L^2([0, T])$ and letting

$$f_2(s, t) := \sum_{i,j=1}^d a_{i,j} e_i(s) e_j(t), \quad s, t \in [0, T],$$

where $A = (a_{i,j})_{1 \leq i,j \leq d}$ is a 2-nilpotent $d \times d$ matrix. Note that the vanishing (3.10) of the contraction $f_2 \otimes_2 f_2$ implies $f_2 = 0$ if the function f_2 is symmetric in its two variables, similarly $A = 0$ if A is symmetric and 2-nilpotent.

Next, given $A = (a_{i,j})_{1 \leq i,j \leq d}$ an $d \times d$ matrix, we note that the quadratic functional $\int_0^T AB_s dB_s$ of the d -dimensional Brownian motion

$$B_t := (W_t, W_{T+t} - W_T, \dots, W_{(d-1)T+t} - W_{(d-1)T}), \quad t \in [0, T],$$

can be represented as

$$I_2(\tilde{f}_2) = \int_0^T AB_s dB_s,$$

where f_2 is the function

$$f_2(s, t) := \sum_{i,j=1}^d a_{i,j} \mathbf{1}_{[(j-1)T, (j-1)T+t-(i-1)T]}(s) \mathbf{1}_{[(i-1)T, iT]}(t), \quad s, t \in [0, dT],$$

as we have

$$\begin{aligned} I_2(\tilde{f}_2) &= \sum_{i,j=1}^d a_{i,j} \int_0^T \int_0^{t+(i-1)T} \mathbf{1}_{[(j-1)T, (j-i)T+t+(i-1)T]}(s) dW_s dB_t^i \\ &\quad + \sum_{i,j=1}^d a_{i,j} \int_{(j-1)T}^{nT} \int_{(i-1)T}^{t \wedge (iT)} \mathbf{1}_{[(j-1)T, (j-i)T+s]}(t) dW_s dW_t \\ &= \sum_{1 \leq j \leq i \leq n} a_{i,j} \int_0^T \int_{(j-1)T}^{t+(j-1)T} dW_s dB_t^i \\ &\quad + \sum_{1 \leq i < j \leq n} a_{i,j} \int_{(j-1)T}^{jT} \int_{(i-1)T}^t dW_s dW_t \\ &= \sum_{1 \leq j \leq i \leq n} a_{i,j} \int_0^T \int_0^t dB_s^j dB_t^i + \sum_{1 \leq i < j \leq n} a_{i,j} \int_0^T \int_0^t dB_s^j dB_t^i \\ &= \int_0^T AB_s dB_s, \end{aligned}$$

with

$$\begin{aligned} &\int_0^{nT} (I_1(f_2(*, t)))^2 dt \\ &= \sum_{i=1}^d \int_{(i-1)T}^{iT} \left(\sum_{j=1}^d a_{i,j} (W_{(j-1)T+t-(i-1)T} - W_{(j-1)T}) \right)^2 dt \\ &= \int_0^T |AB_t|^2 dt. \end{aligned}$$

Note here that, the process $(u_t)_{t \in [0, T]} = (I_1(f_2(*, t)))_{t \in [0, T]}$ is not $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted.

As a consequence of Proposition 3.3 we have the following.

Corollary 3.3. *The condition $\langle u, (Du)u \rangle = 0$ is satisfied if and only if the $d \times d$ matrix A is 2-nilpotent, and in this case we have*

$$d_{|u|} \left(\int_0^T AB_s dB_s, \mathcal{N}_{g(\|u\|)} \right) \leq \frac{1}{\sqrt{g(\|u\|)}} \left| g(\|u\|) - \int_0^T |AB_t|^2 dt \right|.$$

Proof. We have

$$\begin{aligned}
& \int_0^{nT} \int_0^{nT} f_2(t_1, s) f_2(s, t) f_2(t_2, t) \, ds dt \\
&= \sum_{i,j=1}^d \sum_{k,l=1}^d \sum_{p,q=1}^d \int_0^{nT} \int_0^{nT} \mathbf{1}_{[(j-1)T, (j-1)T+t-(i-1)T]}(t_2) \\
&\quad \mathbf{1}_{[(q-1)T, (q-1)T+s-(p-1)T]}(t_1) a_{i,j} a_{k,l} a_{p,q} \mathbf{1}_{[(i-1)T, iT]}(t) \mathbf{1}_{[(k-1)T, kT]}(t) \\
&\quad \mathbf{1}_{[(l-1)T, (l-1)T+t-(k-1)T]}(s) \mathbf{1}_{[(p-1)T, pT]}(s) \, ds dt \\
&= \sum_{i,j=1}^d \sum_{l=1}^d \sum_{q=1}^d \int_{(l-1)T}^{iT} \int_{(i-1)T}^{iT} \mathbf{1}_{[(j-1)T, (j-1)T+t-(i-1)T]}(t_2) \\
&\quad \mathbf{1}_{[(q-1)T, (q-1)T+s-(l-1)T]}(t_1) a_{i,j} a_{i,l} a_{l,q} \mathbf{1}_{[(l-1)T, (l-1)T+t-(i-1)T]}(s) \, dt ds \\
&= \sum_{j,q=1}^d \int_0^T \int_0^T \mathbf{1}_{[(j-1)T, (j-1)T+t]}(t_2) \mathbf{1}_{[(q-1)T, (q-1)T+s]}(t_1) \mathbf{1}_{[0,t]}(s) \, ds dt \\
&\quad \sum_{i,l=1}^d a_{i,j} a_{i,l} a_{l,q} \\
&= \sum_{j,q=1}^d (A^\top A^2)_{j,q} \\
&\quad \int_0^T \int_0^T \mathbf{1}_{[(j-1)T, (j-1)T+t]}(t_2) \mathbf{1}_{[(q-1)T, (q-1)T+s]}(t_1) \mathbf{1}_{[0,t]}(s) \, ds dt,
\end{aligned}$$

$t_1, t_2 \in [0, T]$, hence from (3.11), $\langle u, (Du) u \rangle = 0$ implies $A^\top A^2 = 0$. This in turn will imply that $A^2 = 0$ by the relation $\langle A^2 x, A^2 x \rangle_{\mathbb{R}^d} = \langle Ax, A^\top A^2 x \rangle_{\mathbb{R}^d}$. $\square$

In particular, taking $g(x) = x^2$, Corollary 3.3 shows that, if we are given $\int_0^T |AB_t|^2 dt$, then the quadratic functional $\int_0^T AB_s dB_s$ is Gaussian with variance $\int_0^T |AB_t|^2 dt$ when $A^2 = 0$, which recovers Corollary 3 in [PS15] under the condition of [Yor79].

Chapter 4

Option pricing and implied volatilities in a 2-Hypergeometric stochastic volatility model¹

Abstract

In the preliminaries, we cover some definitions and provide the foundations of option pricing. We also provide motivations for considering the 2-Hypergeometric stochastic volatility model. Thereafter, the closed-form analytical approximations formula in terms of series expansion is derived for the option prices, implied volatilities and delta for the 2-Hypergeometric stochastic volatility model with correlated Brownian motions. Moreover, we show that as in [Han+13], our series expansion formula for implied volatilities is able to recover the long observed volatility smiles and skews phenomenon on implied volatility surfaces, depending on the correlation.

4.1 Introduction

It has long been observed that the volatility of a derivatives cannot be a constant but varies with the time to maturity and strike, as such for a more accurate modeling of the asset prices in financial markets, stochastic volatility models are introduced. The most well-known of which is the Heston model introduced in [Hes93]. The Heston model, though effective, has one major shortcoming as its stochastic volatility is not strictly positive unless we have the Feller condition. However, as

¹This chapter is based on works published in [PS16].

we see in §1.3.3 of [Jan+11] and §6.5.2 of [Hen09], Feller condition is rarely satisfied in reality. Thus, in [DM16], the α -Hypergeometric stochastic volatility model is introduced to ensure that the stochastic volatility is strictly positive. For an asset, the asset price S_t and volatility V_t at time t is governed by the following dynamics:

$$\begin{aligned} dS_t &= S_t e^{V_t} dW_t^1, \\ dV_t &= \left(a - \frac{c}{2} e^{\alpha V_t} \right) dt + \eta dW_t^2, \end{aligned} \quad (4.1)$$

under the α -Hypergeometric model, where $a \in \mathbb{R}$, $c > 0$, $\alpha > 0$, $\eta > 0$, W_t^1 and W_t^2 are correlated Brownian motions which satisfies $\langle W^1, W^2 \rangle_t = \rho t$. In this model, we take the risk free rate r to be 0 and the price of volatility risk can be set by the value of c .

In general, we do not have explicit solutions for stochastic volatility models and even if a model admits an explicit solution like the Heston model, it is usually difficult to evaluate the solution and so this motivated the development of approximate solution in terms of series expansions. In [Fou+11], a method is introduced for the Heston model to obtain the series expansion of European option prices. The method however failed to reproduce the skew and smile effect in model calibration as the first and second terms in the expansion is independent of the stochastic volatility which for Heston model is an important quantity.

In view of this, an approximation that is more accurate is introduced in [Han+13] to price European option under the Heston model via a series expansion that is dependable on the underlying stochastic volatility, thus allowing the authors to not only avoid the secular effect and terminal layer problems posed by the third term in the expansion of [Fou+11], but also to recover the effect of volatility skew and smile. See also [Kim02] under stochastic interest rates.

In this chapter, we extend the method used in [Han+13] and [ZHG13] to derive a closed-form analytical approximate solution based on approximations of the 2-Hypergeometric model of [DM16]. Our series expansion provides an approximate solution that depends on the stochastic volatility and allows us to derive an implied volatility estimate that recover the well observed phenomena of smile and skew.

Furthermore, we provide a series expansion to estimate Delta and

conduct a numerical study to show that our approximate solutions agree with the results of Monte Carlo simulation even if only the first order term is used.

Well known models and methods are discussed in Section 4.2 together with definitions of some financial jargon before detailing our approximation approach in Section 4.3. After which, we present the derivation of our zeroth order term of our expansion in Section 4.4, the first order term in Section 4.5 and the second order term in Section 4.6. Using the second order approximation formula we derived, we proceed to obtain the expansion to approximate the implied volatility and delta in Sections 4.7 and 4.8 respectively. Numerical studies on the effectiveness of our approximation scheme is found in Sections 4.6 and 4.7. As the computation for the second order expansion formula is extensive, we have separately placed it in Section 4.9.

4.2 Preliminaries

The systematic pricing of European call and put was introduced in the seminal paper by Black and Scholes. In [BS73], the now famous Black-Scholes formula was introduced under certain assumptions to price exotic securities. In this subsection, we will discuss about the shortcomings of the Black and Scholes model, which motivated the development of the stochastic volatility model. A short overview of the stochastic volatility model will also be presented in this section together with the some definitions that will be used in the later sections.

Implied Volatility

We observe that all the parameters in the Black-Scholes Model can be observed with the exception of the volatility σ . Thus, it is a common practice to take the implied volatility of the stock to be the σ .

Definition 4.1 (Implied Volatility). *The Black-Scholes implied volatility σ^{imp} is the value at which the Black-Scholes formula recovers the current market price of the option.*

$$V_c^{\text{BS}}(x, K, \sigma^{\text{imp}}, T) = V_c^m,$$

where V_c^m is the current market price of the call option and $V_c^{\text{BS}}(x, K, \sigma^{\text{imp}}, T)$ is the call option price obtain by the Black-Scholes Model with $\sigma = \sigma^{\text{imp}}$.

Shortcomings of Black-Scholes Model

It is clear that the Black-Scholes Model has its advantages, most notably the fact that all the parameters with the exception of σ the volatility is observable. However, the Black-Scholes Model also has its fair share of shortcomings and one of the more important ones is also related to the volatility parameter. We see that in the Black-Scholes Model, the volatility parameter σ is assumed to be a constant, but as we see in [JR96], the volatility is not a constant and displays the smile and skew phenomenon. Hence, it make sense to model the volatility in terms of random process.

Stochastic Volatility Model

In Black-Scholes Model, we assume that the volatility is constant. However we see in [JR96] that the implied volatility varies with the time to maturity and the strike price, more precisely, we observe that the implied volatility surface displays the phenomenon of volatility skew and smile. Thus, to address this shortcoming of the Black-Scholes Model, Stochastic Volatility Model is proposed.

Stochastic Volatility Model allows us to use a more accurate way to model derivatives as we let the volatility of the derivatives follow a random process. One notable stochastic volatility model is the Heston model.

The Heston model introduced in [Hes93] has the following dynamics:

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t dW_t^s, \\ dv_t &= \kappa(\theta - v_t) dt + \xi \sqrt{v_t} dW_t^v, \end{aligned}$$

where r is the rate of return of the stock, κ is the rate at which v_t reverts to θ , θ is the long run average price of volatility and ξ is the volatility of the volatility. Furthermore, W_t^s and W_t^v are two Brownian motions satisfying $d\langle W^s, W^v \rangle_t = \rho dt$.

In the seminal paper, a closed form formula is obtained for the Heston Model. However, the closed form solution is two integrals of complex function which can only be evaluated numerically with considerable difficulty (see [Lee04] and [LK07]). Therefore, numerous approximate solutions are proposed, such as [Fou+11] and [Han+13].

The Heston Model though effective is not without its flaws. One of which is that the stochastic volatility can be 0 in finite time unless the Feller condition is satisfied. But we see in §1.3.3 of [Jan+11] and §6.5.2 of [Hen09] that the Feller condition is seldom satisfied in reality.

As a result, in the following section, we will consider the 2-Hypergeometric model introduced in [DM16] which has the nice property whereby the volatility is strictly positive.

Greeks

Before we dive into the discussion of the 2-Hypergeometric model, it is important to discuss some important quantities that provide information on the sensitivity of the derivatives price to a change its underlying parameters. These quantities are called the Greeks.

One example of the first order Greeks is called Delta and it has the following definition:

Definition 4.2 (Delta). *The Delta, Δ , of a derivatives measures the change in the price of a derivatives with respect to the change in the price of the underlying asset, i.e.*

$$\Delta = \frac{\partial V}{\partial S},$$

where V is the value of the derivative and S is the price of the underlying asset.

4.3 Stochastic Volatility

We begin by considering the following general class of stochastic volatility model whereby the asset price and volatility are governed by:

$$\begin{aligned} dS_t^\epsilon &= S_t^\epsilon p(V_t^\epsilon, t) dW_t^1, \\ dV_t^\epsilon &= u(V_t^\epsilon, t) dt + \epsilon h(V_t^\epsilon, t) dW_t^2 \end{aligned}$$

where we have $\epsilon > 0$. Note that small time asymptotics can be derived using small volatility coefficients by Brownian rescaling (see §2.1.8 of [DM16]). Now, if we have a vanilla option with payoff function $g(S_T^\epsilon)$ then under the absence of arbitrage, the price of the vanilla option have the form

$$f(S_t^\epsilon, V_t^\epsilon, t) := \mathbb{E}[g(S_T^\epsilon) | \mathcal{F}_t],$$

where $(\mathcal{F}_t)_{t \in [0, T]}$ is the filtration generated by $(W_t^1, W_t^2)_{t \in [0, T]}$. Furthermore, we have that the function $f(x, v, t)$ solves the following pricing partial differential equation:

$$\begin{aligned} \frac{\partial f}{\partial t} + u(v, t) \frac{\partial f}{\partial v} + \frac{x^2}{2} p^2(v, t) \frac{\partial f}{\partial x^2} \\ + \epsilon \rho x p(v, t) h(v, t) \frac{\partial^2 f}{\partial x \partial v} + \frac{\epsilon^2}{2} h^2(v, t) \frac{\partial^2 f}{\partial v^2} = 0, \end{aligned} \quad (4.2)$$

with the terminal condition $f(x, v, T) = g(x)$ (see Equation (2.17) in [Fou+11]). Next, we expand $f(x, v, t)$ as

$$f(x, v, t) = f_0(x, v, t) + \epsilon f_1(x, v, t) + o(\epsilon). \quad (4.3)$$

Substituting the expansion (4.3) into (4.2), we obtain a system of equations:

$$\frac{\partial f_n}{\partial t} + \mathcal{L}_0 f_n + \mathcal{L}_1 f_{n-1} + \mathcal{L}_2 f_{n-2} = 0, \quad n \in \mathbb{N},$$

with $f_0(x, v, T) = g(x)$, $f_n(x, v, T) = 0$, $n \geq 1$ and $f_n(x, v, t) = 0$, for all t , when $n \leq -1$. The operators $\mathcal{L}_0$, $\mathcal{L}_1$ and $\mathcal{L}_2$ are

$$\begin{aligned} \mathcal{L}_0 &= u(v, t) \frac{\partial}{\partial v} + \frac{x^2}{2} p^2(v, t) \frac{\partial^2}{\partial x^2}, \quad \mathcal{L}_1 = \rho x p(v, t) h(v, t) \frac{\partial^2}{\partial x \partial v} \\ \text{and } \mathcal{L}_2 &= \frac{1}{2} h^2(v, t) \frac{\partial^2}{\partial v^2}. \end{aligned} \quad (4.4)$$

4.4 Deterministic Volatility

Since we have $\frac{\partial f_0}{\partial t} + \mathcal{L}_0 f_0 = 0$, when $n = 0$. Therefore, $(S_t^0)_{t \in [0, T]}$ and $(V_t^0)_{t \in [0, T]}$ are given by

$$\begin{aligned} dS_t^0 &= S_t^0 p(V_t^0, t) dW_t^1, \\ dV_t^0 &= u(V_t^0, t) dt. \end{aligned}$$

Moreover, the vanilla call option price $f_0(S_t^0, V_t^0, t) := \mathbb{E}[g(S_T^0) | \mathcal{F}_t]$, where $g(S_T^0) = (S_T^0 - K)^+$ when the strike is K and the maturity date is T , can be computed using the Black-Scholes formula to be

$$\begin{aligned} f_0(S_t^0, V_t^0, t) &= \mathbb{E} \left[(S_T^0 - K)^+ \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[\left(S_t^0 \exp \left(Z \gamma(V_t^0, t) - \frac{1}{2} \gamma^2(V_t^0, t) \right) - K \right)^+ \middle| \mathcal{F}_t \right], \end{aligned}$$

with Z being a standard normal random variable which is independent of $\mathcal{F}_t$ and $\gamma^2(V_t^0, t) := \int_t^T p^2(V_u^0, u) du$, $t \in [0, T]$.

We remark that when $\eta = 0$ in the α -Hypergeometric model (4.1), the close form solution of the integral $\int_t^T e^{\alpha V_u^0} du$ is

$$\begin{aligned} \int_t^T e^{\alpha V_u^0} du &= \frac{2}{\alpha c} \log \left(1 + \frac{\alpha c}{2} e^{\alpha V_t^0} \int_0^{T-t} e^{\alpha s} ds \right) \\ &= \frac{2}{\alpha c} \log \left(1 + \frac{\alpha c}{2} e^{\alpha V_t^0} \frac{e^{\alpha a(T-t)} - 1}{\alpha a} \right) \end{aligned} \quad (4.5)$$

(see §2.1.1 of [DM16]). This leads us to the following proposition.

Proposition 4.1. *If $\eta = 0$ in the 2-Hypergeometric stochastic volatility model (4.1), then the European call price with strike K and time to maturity T*

$$f_0(S_t^0, V_t^0, t) = \mathbb{E} \left[(S_T^0 - K)^+ \mid \mathcal{F}_t \right]$$

with the terminal condition $f_0(x, v, T) = (x - K)^+$ is

$$f_0(x, v, t) = x\Phi(d_+(x, v, t)) - K\Phi(d_-(x, v, t)), \quad (4.6)$$

where Φ is the standard normal cumulative distribution function,

$$d_{\pm}(x, v, t) = \frac{1}{\gamma(v, t)} \left(\log\left(\frac{x}{K}\right) \pm \frac{\gamma^2(v, t)}{2} \right) \quad (4.7)$$

and

$$\gamma^2(v, t) = \frac{1}{c} \log \left(1 + ce^{2v} \frac{e^{2a(T-t)} - 1}{2a} \right). \quad (4.8)$$

Proof. We have $\alpha = 2$, $p(V_t^0, t) = e^{V_t^0}$, thus

$$\begin{aligned} \gamma^2(V_t^0, t) &= \int_t^T e^{2V_u^0} du \\ &\stackrel{(4.5)}{=} \frac{1}{c} \log \left(1 + ce^{2v} \frac{e^{2a(T-t)} - 1}{2a} \right). \end{aligned}$$

We conclude the proof by using Black-Scholes formula to obtain Equation (4.6). $\square$

The price of the European put option with strike K and maturity date T can be obtained by a call-put parity argument to be

$$f_0(x, v, t) = -x\Phi(-d_+(x, v, t)) + K\Phi(-d_-(x, v, t)), \quad t \in [0, T].$$

We note here that for the remainder of this chapter, we work in the 2-Hypergeometric model where $\alpha = 2$.

4.5 First Order Expansion

In this section, small values of the volatility of volatility are considered by replacing η in 2-Hypergeometric model (4.1) with $\epsilon\eta e^{V_t^\epsilon}$, $0 < \epsilon \ll 1$,

i.e. we have

$$\begin{aligned} dS_t^\epsilon &= S_t^\epsilon e^{V_t^\epsilon} dW_t^1, \\ dV_t^\epsilon &= \left(a - \frac{c}{2} e^{2V_t^\epsilon}\right) dt + \epsilon \eta e^{V_t^\epsilon} dW_t^2. \end{aligned} \quad (4.9)$$

Next, from (4.4), the operators $\mathcal{L}_0$, $\mathcal{L}_1$ and $\mathcal{L}_2$ can be written as

$$\begin{aligned} \mathcal{L}_0 &= \left(a - \frac{c}{2} e^{2v}\right) \frac{\partial}{\partial v} + \frac{x^2}{2} e^{2v} \frac{\partial^2}{\partial x^2}, \quad \mathcal{L}_1 = \eta \rho x e^{2v} \frac{\partial^2}{\partial x \partial v}, \\ \text{and } \mathcal{L}_2 &= \frac{\eta^2}{2} e^{2v} \frac{\partial^2}{\partial v^2}. \end{aligned}$$

Now, when $n = 1$, we have

$$\frac{\partial f_1}{\partial t} + \mathcal{L}_0 f_1 + \mathcal{L}_1 f_0 = 0, \quad \text{with } f_1(x, v, T) = 0.$$

Before we proceed to the next proposition, we remark that our approximation $(S_t^\epsilon, V_t^\epsilon)_{t \in [0, T]}$ does not lie within the class of 2-Hypergeometric stochastic volatility models.

Proposition 4.2. *The partial differential equation $\frac{\partial f_1}{\partial t} + \mathcal{L}_0 f_1 + \mathcal{L}_1 f_0 = 0$ with the terminal condition $f_1(x, v, T) = 0$ has a solution given by*

$$\begin{aligned} f_1(x, v, t) &= -\rho K \frac{\eta}{c} d_-(x, v, t) \phi(d_-(x, v, t)) \\ &\quad \frac{e^{-c\gamma^2(v, t)} + c\gamma^2(v, t) - 1}{c\gamma^2(v, t)}, \quad t \in [0, T], \end{aligned}$$

where $\phi(x)$ is the probability density function of a standard normal distribution. $d_-(x, v, t)$ and $\gamma^2(v, t)$ are as given in (4.7) and (4.8) respectively.

Proof. Since $\phi(x)$ is the probability density function of a standard normal distribution, we obtain the relation

$$\begin{aligned} \phi(d_+(x, v, t)) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} (d_+(x, v, t))^2\right) \\ &= \frac{K}{x} \phi(d_-(t, x, v)). \end{aligned}$$

So using the Feynman-Kac formula with locally Lipschitz coefficients as in Theorem 1 of [HS00], we have

$$\begin{aligned}
 f_1(S_t^0, V_t^0, t) &= \int_t^T \mathbb{E} [\mathcal{L}_1 f_0(S_r^0, V_r^0, r) \mid \mathcal{F}_t] dr \\
 &= -\eta \rho K \int_t^T \frac{e^{2V_r^0}}{\gamma(V_r^0, r)} \frac{\partial \gamma}{\partial v}(V_r^0, r) \\
 &\quad \mathbb{E} [d_-(S_r^0, V_r^0, r) \phi(d_-(S_r^0, V_r^0, r)) \mid \mathcal{F}_t] dr \\
 &= -\eta \rho K \frac{d_-(S_t^0, V_t^0, t)}{\gamma^2(V_t^0, t)} \phi(d_-(S_t^0, V_t^0, t)) \\
 &\quad \int_t^T e^{2V_r^0} \gamma(V_r^0, r) \frac{\partial \gamma}{\partial v}(V_r^0, r) dr,
 \end{aligned}$$

by a standard computation based on the normal distribution since

$$\begin{aligned}
 &d_-(S_r^0, V_r^0, r) \\
 &\sim \mathcal{N} \left(\frac{1}{\gamma(V_r^0, r)} \left(\log \left(\frac{S_t^0}{K} \right) - \frac{\gamma^2(V_t^0, t)}{2} \right), \frac{\gamma^2(V_t^0, t)}{\gamma^2(V_r^0, r)} - 1 \right),
 \end{aligned}$$

given $\mathcal{F}_t$ and $r \in [t, T]$. Finally, from Equation (4.8) we obtain

$$\begin{aligned}
 &\int_t^T e^{2V_r^0} \gamma(V_r^0, r) \frac{\partial \gamma}{\partial v}(V_r^0, r) dr \\
 &= \frac{1}{c} \int_t^T e^{2V_r^0} \left(1 - e^{-c\gamma^2(V_r^0, r)} \right) dr \\
 &= \frac{1}{c^2} \left(e^{-c\gamma^2(V_t^0, t)} + c\gamma^2(V_t^0, t) - 1 \right). \tag{4.10}
 \end{aligned}$$

□

4.6 Second Order Expansion

In this section, we replace η in 2-Hypergeometric model (4.1) with $\epsilon \eta e^{V_t} \gamma^4(V_t, t)$ so that only even powers of $\gamma(v, r)$ are involved when extending the computation of Equation (4.10) in order to compute the second order correction term $\tilde{f}_2(x, v, t)$. As such, the operators $\mathcal{L}_1$ and $\mathcal{L}_2$ are replaced by

$$\tilde{\mathcal{L}}_1 = \eta \rho x e^{2v} \gamma^4(v, t) \frac{\partial^2}{\partial x \partial v}, \quad \tilde{\mathcal{L}}_2 = \frac{\eta^2}{2} e^{2v} \gamma^8(v, t) \frac{\partial^2}{\partial v^2},$$

respectively. We assume that we can expand $f(x, v, t)$ into the form

$$f(x, v, t) = f_0(x, v, t) + \epsilon \tilde{f}_1(x, v, t) + \epsilon^2 \tilde{f}_2(x, v, t) + o(\epsilon^2), \quad (4.11)$$

where $f_0(x, v, T) = (x - K)^+$, $\tilde{f}_1(x, v, T) = \tilde{f}_2(x, v, T) = 0$ and

$$\begin{aligned} \frac{\partial f_0}{\partial t} + \mathcal{L}_0 f_0 &= 0, & \frac{\partial \tilde{f}_1}{\partial t} + \mathcal{L}_0 \tilde{f}_1 + \tilde{\mathcal{L}}_1 f_0 &= 0, \\ \text{and } \frac{\partial \tilde{f}_2}{\partial t} + \mathcal{L}_0 \tilde{f}_2 + \tilde{\mathcal{L}}_1 \tilde{f}_1 + \tilde{\mathcal{L}}_2 f_0 &= 0. \end{aligned}$$

Proposition 4.3. *We can compute $\tilde{f}_1(x, v, t)$ and $\tilde{f}_2(x, v, t)$ of expression (4.11) to be*

$$\begin{aligned} \tilde{f}_1(x, v, t) &= -\eta \rho K \frac{d_-(x, v, t)}{c^4 \gamma^2(v, t)} \phi(d_-(x, v, t)) \\ &\quad \left(e^{-c\gamma^2(v, t)} (c^2 \gamma^4(v, t) + 2c\gamma^2(v, t) + 2) - 2 + \frac{c^3}{3} \gamma^6(v, t) \right), \end{aligned} \quad (4.12)$$

$$\begin{aligned} \tilde{f}_2(x, v, t) &= \frac{\eta^2}{c} K \phi(d_-(x, v, t)) \left(\frac{A_3(v, t)}{\gamma(v, t)} + d_-(x, v, t) B_3(v, t) \right. \\ &\quad \left. + \frac{(d_-(x, v, t))^2}{\gamma(v, t)} B_3(v, t) \right) + \eta^2 \rho^2 \frac{K}{c} \phi(d_-(x, v, t)) (C_3(v, t) \\ &\quad + \frac{2D(v, t)}{3c^7 \gamma^4(v, t)} \left(\frac{(d_-(x, v, t))^4}{3\gamma(v, t)} - d_-(x, v, t) + \frac{(d_-(x, v, t))^3}{3} \right) \\ &\quad \left. + \frac{(d_-(x, v, t))^2}{\gamma^5(v, t)} E_3(v, t) \right), \quad t \in [0, T], \end{aligned} \quad (4.13)$$

where the functions A_i, B_i, C_i, D, E_i for $i = 1, 2, 3$, are

$$\begin{aligned}
A_1(v, t) &= \frac{\gamma^8(v, t)}{2c} + \frac{5\gamma^6(v, t)}{4c^2} + \frac{2\gamma^4(v, t)}{c^3} + \frac{9\gamma^2(v, t)}{4c^4} + \frac{3}{2c^5} \\
&\quad + \frac{3c^{-6}}{8\gamma^2(v, t)}, \\
A_2(v, t) &= -\frac{\gamma^8(v, t)}{c} - \frac{5\gamma^6(v, t)}{c^2} - \frac{16\gamma^4(v, t)}{c^3} - \frac{24\gamma^2(v, t)}{c^4} - \frac{48}{c^5} \\
&\quad - \frac{24c^{-6}}{\gamma^2(v, t)}, \\
A_3(v, t) &= -\frac{\gamma^8(v, t)}{10c} + e^{-2c\gamma^2(v, t)} A_1(v, t) + \frac{93}{4c^5} + \frac{189c^{-6}}{8\gamma^2(v, t)} \\
&\quad + A_2(v, t) e^{-c\gamma^2(v, t)}, \\
B_1(v, t) &= -\frac{\gamma^6(v, t)}{4c^2} - \frac{\gamma^4(v, t)}{2c^3} - \frac{3\gamma^2(v, t)}{4c^4} - \frac{3}{4c^5} - \frac{3c^{-6}}{8\gamma^2(v, t)}, \\
B_2(v, t) &= \frac{\gamma^6(v, t)}{c^2} + \frac{4\gamma^4(v, t)}{c^3} + \frac{12\gamma^2(v, t)}{c^4} + \frac{24}{c^5} + \frac{24c^{-6}}{\gamma^2(v, t)}, \\
B_3(v, t) &= \frac{\gamma^8(v, t)}{10c} + e^{-2c\gamma^2(v, t)} B_1(v, t) - \frac{189c^{-6}}{8\gamma^2(v, t)} \\
&\quad + B_2(v, t) e^{-c\gamma^2(v, t)}, \\
C_1(v, t) &= -\frac{\gamma^5(v, t)}{c^2} + \frac{\gamma^3(v, t)}{2c^3} + \frac{3\gamma^2(v, t)}{c^4} + \frac{4\gamma(v, t)}{c^4} + \frac{9}{2c^5} \\
&\quad + \frac{21c^{-5}}{2\gamma(v, t)} + \frac{9c^{-6}}{2\gamma^2(v, t)} + \frac{9c^{-6}}{\gamma^3(v, t)} + \frac{9c^{-7}}{4\gamma^4(v, t)} + \frac{15c^{-7}}{4\gamma^5(v, t)}, \\
C_2(v, t) &= -\frac{3\gamma^5(v, t)}{c^2} - \frac{9\gamma^3(v, t)}{c^3} - \frac{6\gamma^2(v, t)}{c^4} - \frac{64\gamma(v, t)}{c^4} - \frac{36}{c^5} \\
&\quad - \frac{120c^{-5}}{\gamma(v, t)} - \frac{36c^{-6}}{\gamma^2(v, t)} + \frac{24c^{-6}}{\gamma^3(v, t)} - \frac{36c^{-7}}{\gamma^4(v, t)} + \frac{24c^{-7}}{\gamma^5(v, t)}, \\
C_3(v, t) &= -\frac{7\gamma^7(v, t)}{30c} - \frac{2\gamma(v, t)}{c^4} + C_1(v, t) e^{-2c\gamma^2(v, t)} \\
&\quad + C_2(v, t) e^{-c\gamma^2(v, t)} + \frac{189c^{-6}}{2\gamma^3(v, t)} + \frac{135c^{-7}}{4\gamma^4(v, t)} \\
&\quad - \frac{111c^{-7}}{4\gamma^5(v, t)}, \\
D(v, t) &= e^{-2c\gamma^2(v, t)} \left(e^{c\gamma^2(v, t)} (c^3\gamma^6(v, t) - 3) \right. \\
&\quad \left. + 3c\gamma^2(v, t) \left(\frac{c}{2}\gamma^2(v, t) + 1 \right) + 3 \right)^2,
\end{aligned}$$

$$\begin{aligned}
E_1(v, t) &= \frac{\gamma^{10}(v, t)}{c^2} + \frac{\gamma^8(v, t)}{c^3} - \frac{15\gamma^6(v, t)}{c^4} - \frac{27\gamma^4(v, t)}{c^5} \\
&\quad - \frac{51\gamma^2(v, t)}{2c^6} - \frac{12}{c^7}, \\
E_2(v, t) &= \frac{2\gamma^{10}(v, t)}{c^2} + \frac{2\gamma^8(v, t)}{c^3} + \frac{17\gamma^6(v, t)}{c^4} + \frac{216\gamma^4(v, t)}{c^5} \\
&\quad + \frac{24\gamma^2(v, t)}{c^6} + \frac{24}{c^7}, \\
E_3(v, t) &= \frac{\gamma^{12}(v, t)}{15c} + \frac{4\gamma^6(v, t)}{c^4} - \frac{189\gamma^2(v, t)}{2c^6} - \frac{492}{c^7} \\
&\quad + E_1(v, t) e^{-2c\gamma^2(v, t)} + e^{-c\gamma^2(v, t)} E_2(v, t).
\end{aligned}$$

Proof. The expression $\tilde{f}_1$ is computed using a similar argument as the proof of Proposition 4.2. Similarly, since we have $\frac{\partial \tilde{f}_2}{\partial t} + \mathcal{L}_0 \tilde{f}_2 + \tilde{\mathcal{L}}_1 \tilde{f}_1 + \tilde{\mathcal{L}}_2 f_0 = 0$, where $\tilde{f}_2(x, v, T) = 0$, $\tilde{f}_2$ can be obtained by the same arguments of using Feynman-Kac to have

$$\tilde{f}_2(S_t, V_t, t) = \int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_1 \tilde{f}_1(S_r, V_r, r) + \tilde{\mathcal{L}}_2 f_0(S_r, V_r, r) \mid \mathcal{F}_t \right] dr.$$

For a more detail expositions of the computations, please refer to Section 4.9, as the calculations involved are significantly longer than the proof of Proposition 4.2. $\square$

Notice that the corresponding European put option price can be obtained using the standard call-put parity argument, where we see that only the function $f_0(x, v, t)$ is modified. The higher orders term like $f_1(t, x, v)$, $\tilde{f}_1(x, v, t)$ and $\tilde{f}_2(x, v, t)$ remain unchanged.

In Figure 4.1 we plot the option price against the stochastic variance v with $\rho = -0.5$, $x = K$, $T = 0.1$, $t = 0$, $a = c/2 = 1$, $\eta = 2$ and $\epsilon = 0.01$. The Monte Carlo curve is generated with 300,000 samples on 30,000 time steps.

FIGURE 4.1: Option price f plotted against v with $\rho = -0.5$.

Figure 4.1 shows that our asymptotic solutions are in agreement with the Monte Carlo solution. We see that when v is in the interval $[0, 4]$ our analytical approximate solutions are accurate even if only f_0 is used. For larger values of v , the Monte Carlo estimate hovers around the analytical approximate solutions. Hence, for larger v , a smoother Monte Carlo graph can be obtained by further increasing the number of samples and time steps.

Next, in Table 4.1, we show the approximated values obtained from f_0 , $f_0 + \epsilon \tilde{f}_1$, $f_0 + \epsilon \tilde{f}_1 + \epsilon^2 \tilde{f}_2$, with the same parameters as above, except now, we generate the corresponding Monte Carlo estimates with 1,000,000 samples on 100,000 time steps. While the Monte Carlo took us a long time, the evaluation of the approximations is instantaneous.

Note that the larger number of time steps is due to the instabilities in the solutions of stochastic differential equations with non-Lipschitz (here exponential) coefficients such as those in (4.9).

v	f_0	$f_0 + \epsilon \tilde{f}_1$	$f_0 + \epsilon \tilde{f}_1 + \epsilon^2 \tilde{f}_2$	Monte Carlo
5.5	0.7239	0.7060	0.6998	0.6827
5.6	0.7289	0.7103	0.7036	0.6975
5.7	0.7338	0.7144	0.7072	0.7045
5.8	0.7386	0.7185	0.7107	0.7089
5.9	0.7433	0.7224	0.7140	0.7078
6.0	0.7478	0.7262	0.7172	0.7097

TABLE 4.1: Values of f_0 , $f_0 + \epsilon \tilde{f}_1$, $f_0 + \epsilon \tilde{f}_1 + \epsilon^2 \tilde{f}_2$ compared to Monte Carlo estimates.

We notice that our approximate solutions up till second correction terms gives values closer to the Monte Carlo estimates.

4.7 Implied Volatility

An estimation of the implied volatility, σ^{imp} is provided in this section. By definition, σ^{imp} is determined by the following equation:

$$f^{\text{BS}}(x, K, \sigma^{\text{imp}}, t, T) = f(x, v, t),$$

where $f^{\text{BS}}(x, K, \sigma^{\text{imp}}, t, T)$ is the classical Black-Scholes function (see [FG11] in multi-factor model).

Theorem 4.1. *Assume that we can expand as in Equation (4.11), then the implied volatility σ^{imp} admits the following series expansion*

$$\sigma^{\text{imp}}(x, v, t) = \sigma_0(x, v, t) + \epsilon \sigma_1(x, v, t) + \epsilon^2 \sigma_2(x, v, t) + o(\epsilon^2),$$

with

$$\sigma_0(x, v, t) := \frac{\gamma(v, t)}{\sqrt{T-t}}, \quad \sigma_1(x, v, t) := \frac{\tilde{f}_1(x, v, t)}{K\sqrt{T-t}\phi(d_-(x, v, t))},$$

and

$$\begin{aligned} \sigma_2(x, v, t) := & \frac{\tilde{f}_2(x, v, t)}{K\sqrt{T-t}\phi(d_-(x, v, t))} \\ & - d_+(x, v, t) d_-(x, v, t) \frac{\sigma_1^2(x, v, t)}{2\sigma_0(x, v, t)}. \end{aligned}$$

Proof. We determine the implied volatility σ^{imp} by equating

$$\begin{aligned} f^{\text{BS}}(x, K, \sigma^{\text{imp}}, t, T) &= f(x, v, t) \\ &= f_0(x, v, t) + \epsilon \tilde{f}_1(x, v, t) + \epsilon^2 \tilde{f}_2(x, v, t) + o(\epsilon^2), \end{aligned}$$

where f^{BS} is the Black-Scholes function with implied volatility σ^{imp} . Next, we expand the implied volatility σ^{imp} in ϵ into the following power series

$$\sigma^{\text{imp}}(x, v, t) = \sigma_0(x, v, t) + \epsilon \sigma_1(x, v, t) + \epsilon^2 \sigma_2(x, v, t) + o(\epsilon^2)$$

and used Taylor expansion in terms of σ on $f^{\text{BS}}(x, K, \sigma^{\text{imp}}, t, T)$ to obtain

$$\begin{aligned} & f^{\text{BS}}(x, K, \sigma^{\text{imp}}, t, T) \\ &= f^{\text{BS}}(x, K, \sigma_0(x, v, t), t, T) \\ &+ (\epsilon \sigma_1(x, v, t) + \epsilon^2 \sigma_2(x, v, t)) \frac{\partial f^{\text{BS}}}{\partial \sigma}(x, K, \sigma_0(x, v, t), t, T) \\ &+ \frac{1}{2} \epsilon^2 \sigma_1^2(x, v, t) \frac{\partial^2 f^{\text{BS}}}{\partial \sigma^2}(x, K, \sigma_0(x, v, t), t, T) + \dots \end{aligned}$$

By identification of coefficients, we obtain the first three terms of the implied volatility expansion. $\square$

With the parameters $a = \frac{c}{2} = v = 1$, $\eta = 3.5$, $\epsilon = 0.1$, $t = 0$ and $T = 1$, we plot the estimation of implied volatility σ^{imp} against the ratio $\frac{K}{x}$ (moneyness) in Figures (4.2) and (4.3).

FIGURE 4.2: Implied volatility σ^{imp} plotted against the moneyness K/x for $\rho = 0$.

FIGURE 4.3: Implied volatility σ^{imp} plotted against the moneyness K/x for $\rho = -0.8$.

We see that our estimate $\sigma_0(x, v, t) + \epsilon \sigma_1(x, v, t) + \epsilon^2 \sigma_2(x, v, t)$ for the implied volatility σ^{imp} exhibits the well observed smile and skew phenomena (Figures (4.2) and (4.3) respectively). Furthermore, the figures also show that higher order correction terms such as $\tilde{f}_1$ and $\tilde{f}_2$ is necessary for improved calibration. Estimates of implied volatility using Monte Carlo are not available due to the instabilities observed in Figure (4.1) for the numerical solution of SDEs such as (4.9).

4.8 Delta

In view of Proposition 4.3, we can obtain an approximation for the Delta, Δ , of the European option.

Proposition 4.4. *The Delta, Δ , of a European option can be approximated by using the following series expansion*

$$\Delta(x, v, t) = \frac{\partial f_0(x, v, t)}{\partial x} + \epsilon \frac{\partial \tilde{f}_1(x, v, t)}{\partial x} + \epsilon^2 \frac{\partial \tilde{f}_2(x, v, t)}{\partial x} + o(\epsilon^2), \quad (4.14)$$

where

$$\begin{aligned}
\frac{\partial f_0(x, v, t)}{\partial x} &= \begin{cases} \Phi(d_+(x, v, t)) + \frac{f_0(x, v, t)}{x\gamma(v, t)}, & \text{for call} \\ -\Phi(-d_+(x, v, t)) - \frac{f_0(x, v, t)}{x\gamma(v, t)}, & \text{for put} \end{cases} \\
\frac{\partial \tilde{f}_1(x, v, t)}{\partial x} &= \frac{\tilde{f}_1(x, v, t) (1 - (d_-(x, v, t))^2)}{x\gamma(v, t) d_-(x, v, t)}, \text{ for call and put} \\
\frac{\partial \tilde{f}_2(x, v, t)}{\partial x} &= -\frac{d_-(x, v, t) \tilde{f}_2(x, v, t)}{x\gamma(v, t)} + \frac{\eta^2}{c} K \phi(d_-(x, v, t)) \\
&\quad \left(\frac{A_3(v, t)}{\gamma(v, t)} + \frac{B_3(v, t)}{x\gamma(v, t)} + \frac{2d_-(x, v, t)}{x\gamma^2(v, t)} B_3(v, t) \right) \\
&\quad + \eta^2 \rho^2 \frac{K}{c} \phi(d_-(x, v, t)) \left(C_3(v, t) + \frac{2D(v, t)}{3c^7 \gamma^4(v, t)} \right. \\
&\quad \left. \left(\frac{4(d_-(x, v, t))^3}{3x\gamma^2(v, t)} - \frac{1}{x\gamma(v, t)} + \frac{(d_-(x, v, t))^2}{x\gamma(v, t)} \right) \right. \\
&\quad \left. + \frac{2d_-(x, v, t)}{x\gamma^6(v, t)} E_3(v, t) \right), \text{ for call and put.}
\end{aligned}$$

Proof. Equation (4.14) follows directly from the definition of Delta. $\square$

4.9 Second Order Expansion Proof

For the proof of the second order expansion, we split it into several steps.

Step 1) Preliminaries results. The following results are important for the proof of $\tilde{f}_1(x, v, t)$ and $\tilde{f}_2(x, v, t)$

Proposition 4.5. *If*

$$\begin{aligned}
&d_-(S_r^0, V_r^0, r) \\
&\sim \mathcal{N} \left(\frac{1}{\gamma(V_r^0, r)} \left(\log \left(\frac{S_t^0}{K} \right) - \frac{\gamma^2(V_t^0, t)}{2} \right), \frac{\gamma^2(V_t^0, t)}{\gamma^2(V_r^0, r)} - 1 \right),
\end{aligned}$$

then we have the following identities:

$$\mathbb{E} [\phi (d_- (S_r^0, V_r^0, r)) \mid \mathcal{F}_t] = \frac{\phi (d_- (S_t^0, V_t^0, t))}{\gamma (V_t^0, t)} \gamma (V_r^0, r), \quad (4.15)$$

$$\begin{aligned} & \mathbb{E} [d_- (S_r^0, V_r^0, r) \phi (d_- (S_r^0, V_r^0, r)) \mid \mathcal{F}_t] \\ &= \frac{d_- (S_t^0, V_t^0, t) \phi (d_- (S_t^0, V_t^0, t))}{\gamma^2 (V_t^0, t)} \gamma^2 (V_r^0, r), \end{aligned} \quad (4.16)$$

$$\begin{aligned} & \mathbb{E} [(d_- (S_r^0, V_r^0, r))^2 \phi (d_- (S_r^0, V_r^0, r)) \mid \mathcal{F}_t] \\ &= \frac{\phi (d_- (S_t^0, V_t^0, t))}{\gamma (V_t^0, t)} \gamma (V_r^0, r) - \frac{\phi (d_- (S_t^0, V_t^0, t))}{\gamma^3 (V_t^0, t)} \gamma^3 (V_r^0, r) \\ & \quad + \frac{(d_- (S_t^0, V_t^0, t))^2 \phi (d_- (S_t^0, V_t^0, t))}{\gamma^3 (V_t^0, t)} \gamma^3 (V_r^0, r), \end{aligned} \quad (4.17)$$

$$\begin{aligned} & \mathbb{E} [(d_- (S_r^0, V_r^0, r))^3 \phi (d_- (S_r^0, V_r^0, r)) \mid \mathcal{F}_t] \\ &= \frac{(d_- (S_t^0, V_t^0, t))^3 \phi (d_- (S_t^0, V_t^0, t))}{\gamma^4 (V_t^0, t)} \gamma^4 (V_r^0, r) \\ & \quad - \frac{3d_- (S_t^0, V_t^0, t) \phi (d_- (S_t^0, V_t^0, t))}{\gamma^4 (V_t^0, t)} \gamma^4 (V_r^0, r) \\ & \quad + \frac{3d_- (S_t^0, V_t^0, t) \phi (d_- (S_t^0, V_t^0, t))}{\gamma^2 (V_t^0, t)} \gamma^2 (V_r^0, r) \end{aligned} \quad (4.18)$$

and

$$\begin{aligned} & \mathbb{E} [(d_- (S_r^0, V_r^0, r))^4 \phi (d_- (S_r^0, V_r^0, r)) \mid \mathcal{F}_t] \\ &= \frac{(d_- (S_t^0, V_t^0, t))^4 \phi (d_- (S_t^0, V_t^0, t))}{\gamma^5 (V_t^0, t)} \gamma^5 (V_r^0, r) \\ & \quad - \frac{6(d_- (S_t^0, V_t^0, t))^2 \phi (d_- (S_t^0, V_t^0, t))}{\gamma^5 (V_t^0, t)} \gamma^5 (V_r^0, r) \\ & \quad + \frac{3\phi (d_- (S_t^0, V_t^0, t))}{\gamma^5 (V_t^0, t)} \gamma^5 (V_r^0, r) \\ & \quad + \frac{6(d_- (S_t^0, V_t^0, t))^2 \phi (d_- (S_t^0, V_t^0, t))}{\gamma^3 (V_t^0, t)} \gamma^3 (V_r^0, r) \\ & \quad - \frac{6\phi (d_- (S_t^0, V_t^0, t))}{\gamma^3 (V_t^0, t)} \gamma^3 (V_r^0, r) \\ & \quad + \frac{3\phi (d_- (S_t^0, V_t^0, t))}{\gamma (V_t^0, t)} \gamma (V_r^0, r). \end{aligned} \quad (4.19)$$

Proof. Equation (4.15) can be proved by using a standard computation based on normal distribution. Equations (4.16), (4.17), (4.18) and (4.19) are also proved by using a standard computation based on normal distribution with the use of moments formula. $\square$

In the following propositions, we evaluate the numerous integrals that we encounter in our proof.

Proposition 4.6. *If $\gamma^2(V_t^0, t)$ is defined as in (4.8), then we have the followings:*

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^4(V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr \\ &= \frac{\gamma^6(V_t^0, t)}{3} - \frac{e^{-c\gamma^2(V_t^0, t)} (-c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 2) - 2) + 2}{c^3}, \end{aligned} \quad (4.20)$$

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^8(V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr \\ &= \frac{\gamma^{10}(V_t^0, t)}{5} - \frac{e^{-c\gamma^2(V_t^0, t)}}{c^5} \left(-c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) \right. \\ & \quad \left. (c\gamma^2(V_t^0, t) + 4) + 12) + 24) \right) - \frac{24}{c^5} + \frac{24e^{-c\gamma^2(V_t^0, t)}}{c^5} \end{aligned} \quad (4.21)$$

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^{10}(V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr \\ &= \frac{\gamma^{12}(V_t^0, t)}{6} - \frac{e^{-c\gamma^2(V_t^0, t)}}{c^6} \left(-c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) \right. \\ & \quad \left. (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 5) + 20) + 60) + 120) - 120 \right) - \frac{120}{c^6} \end{aligned} \quad (4.22)$$

Proof. Equations (4.20), (4.21) and (4.22) are proved by the recursive use of integration by parts formula. $\square$

Proposition 4.7. *Similarly, we have*

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^4 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right)^2 dr \\ &= \frac{\gamma^6(V_t^0, t)}{3} - \frac{e^{-2c\gamma^2(V_t^0, t)}}{4c^3} (2c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 1) + 1) \\ & \quad + \frac{2e^{-c\gamma^2(V_t^0, t)}}{c^3} (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 2) + 2) - \frac{15}{4c^3}, \end{aligned} \quad (4.23)$$

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right)^2 dr \\ &= \frac{\gamma^{10}(V_t^0, t)}{5} + \frac{e^{-2c\gamma^2(V_t^0, t)}}{4c^5} (-2c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) \\ & \quad (c\gamma^2(V_t^0, t) + 2) + 3) + 3) - 3) - \frac{2e^{-c\gamma^2(V_t^0, t)}}{c^5} (-c\gamma^2(V_t^0, t) \\ & \quad (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 4) + 12) + 24) - 24) \\ & \quad - \frac{48}{c^5} + \frac{3}{4c^5}. \end{aligned} \quad (4.24)$$

Proof. The proof of Equations (4.23) and (4.24) is similar to the proof of Proposition 4.6 □

Proposition 4.8. *We also have*

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^6 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)}\right) dr \\ &= \frac{e^{-2c\gamma^2(V_t^0, t)}}{8c^4} (2c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (2c\gamma^2(V_t^0, t) + 3) + 3) + 3) \\ & \quad + \frac{e^{-c\gamma^2(V_t^0, t)}}{c^4} (-c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 3) + 6) - 6) \\ & \quad + \frac{6}{c^4} - \frac{3}{8c^4}, \end{aligned} \quad (4.25)$$

$$\begin{aligned} & \int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)}\right) dr \\ &= \frac{e^{-2c\gamma^2(V_t^0, t)}}{4c^5} (2c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) + 2) \\ & \quad + 3) + 3) + 3) + \frac{e^{-c\gamma^2(V_t^0, t)}}{c^5} (-c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) (c\gamma^2(V_t^0, t) \\ & \quad (c\gamma^2(V_t^0, t) + 4) + 12) + 24) - 24) - \frac{3}{4c^5} + \frac{24}{c^5}. \end{aligned} \quad (4.26)$$

Proof. Equations (4.25) and (4.26) are also proved by the recursive use of the integration by parts formula. $\square$

Step 2) Proof of $\tilde{f}_1(x, v, t)$. Using Propositions 4.5 and 4.6, we have

$$\begin{aligned}
 \tilde{f}_1(S_t^0, V_t^0, t) &= \int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_1 f_0(S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr \\
 &= -\eta \rho K \int_t^T e^{2V_r^0} \gamma^3(V_r^0, r) \frac{\partial \gamma}{\partial v}(V_r^0, r) \\
 &\quad \mathbb{E} [d_-(S_r^0, V_r^0, r) \phi(d_-(S_r^0, V_r^0, r)) \mid \mathcal{F}_t] dr \\
 &\stackrel{(4.16)}{=} -\eta \rho K \frac{d_-(S_t^0, V_t^0, t)}{\gamma^2(V_t^0, t)} \phi(d_-(S_t^0, V_t^0, t)) \\
 &\quad \int_t^T e^{2V_r^0} \gamma^5(V_r^0, r) \frac{\partial \gamma}{\partial v}(V_r^0, r) dr
 \end{aligned}$$

Next, from Equations (4.8) and using (4.20), we proved (4.12). $\square$

In the remaining of the proof, we use Propositions 4.5 to obtain the components of $\int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_1 \tilde{f}_1(S_r^0, V_r^0, r) + \tilde{\mathcal{L}}_2 f_0(S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr$ and use Propositions 4.6, 4.7 and 4.8 to prove Equation (4.13).

Step 3) Components of $\int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_1 \tilde{f}_1(S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr$. The components of $\int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_1 \tilde{f}_1(S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr$ consist of the followings:

The term involving $\int_t^T e^{2V_r^0} \gamma^8(V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr$:

$$\begin{aligned}
 M_1(S_t^0, V_t^0, t) &= \eta^2 \rho^2 K \frac{2\phi(d_-(S_t^0, V_t^0, t))}{c^2 \gamma^3(V_t^0, t)} \left((d_-(S_t^0, V_t^0, t))^2 - 1 \right) \\
 &\quad \int_t^T e^{2V_r^0} \gamma^8(V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr.
 \end{aligned}$$

The term involving $\int_t^T e^{2V_r^0} \gamma^{10} (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr$:

$$\begin{aligned}
& M_2(S_t^0, V_t^0, t) \\
&= \eta^2 \rho^2 K \frac{\phi(d_-(S_t^0, V_t^0, t))}{3c^2 \gamma^4(V_t^0, t)} \left((d_-(S_t^0, V_t^0, t))^3 - 3d_-(S_t^0, V_t^0, t) \right. \\
&\quad \left. + \frac{(d_-(S_t^0, V_t^0, t))^4 - 6(d_-(S_t^0, V_t^0, t))^2 + 3}{\gamma(V_t^0, t)} \right) \\
&\quad \int_t^T e^{2V_r^0} \gamma^{10} (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right) dr.
\end{aligned}$$

The term involving $\int_t^T e^{2V_r^0} \gamma^4 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right)^2 dr$:

$$\begin{aligned}
& M_3(S_t^0, V_t^0, t) \\
&= -\eta^2 \rho^2 K \frac{2\phi(d_-(S_t^0, V_t^0, t))}{c^5 \gamma^4(V_t^0, t)} \left((d_-(S_t^0, V_t^0, t))^3 - 3d_-(S_t^0, V_t^0, t) \right. \\
&\quad \left. + \frac{(d_-(S_t^0, V_t^0, t))^4 - 6(d_-(S_t^0, V_t^0, t))^2 + 3}{\gamma(V_t^0, t)} \right) \\
&\quad \int_t^T e^{2V_r^0} \gamma^4 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)}\right)^2 dr.
\end{aligned}$$

The term involving $\int_t^T e^{2V_r^0} \gamma^6 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)}\right) dr$:

$$\begin{aligned}
& M_4(S_t^0, V_t^0, t) \\
&= \eta^2 \rho^2 K \frac{2\phi(d_-(S_t^0, V_t^0, t))}{c^4 \gamma^4(V_t^0, t)} \left((d_-(S_t^0, V_t^0, t))^3 - 3d_-(S_t^0, V_t^0, t) + 3 \right. \\
&\quad \left. + \frac{(d_-(S_t^0, V_t^0, t))^4 - 6(d_-(S_t^0, V_t^0, t))^2}{\gamma(V_t^0, t)} \right) \\
&\quad \int_t^T e^{2V_r^0} \gamma^6 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)}\right) dr.
\end{aligned}$$

The term involving $\int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)} \right) dr$:

$$\begin{aligned} M_5 (S_t^0, V_t^0, t) &= \eta^2 \rho^2 K \frac{\phi(d_-(S_t^0, V_t^0, t))}{c^2 \gamma^3(V_t^0, t)} \left(\frac{(d_-(S_t^0, V_t^0, t))^3 - 3d_-(S_t^0, V_t^0, t)}{c\gamma(V_t^0, t)} \right. \\ &\quad \left. + \frac{(d_-(S_t^0, V_t^0, t))^4 - 6(d_-(S_t^0, V_t^0, t))^2 + 3}{c\gamma^2(V_t^0, t)} - 2d_-(S_t^0, V_t^0, t) + 2 \right) \\ &\quad \int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)} \right) dr. \end{aligned}$$

Step 4) Components of $\int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_2 f_0 (S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr$. The components of $\int_t^T \mathbb{E} \left[\tilde{\mathcal{L}}_2 f_0 (S_r^0, V_r^0, r) \mid \mathcal{F}_t \right] dr$ consist of the followings:

The term involving $\int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)} \right)^2 dr$:

$$\begin{aligned} M_6 (S_t^0, V_t^0, t) &= \eta^2 K \frac{\phi(d_-(S_t^0, V_t^0, t))}{2c^2 \gamma^2(V_t^0, t)} \left(d_-(S_t^0, V_t^0, t) + \frac{(d_-(S_t^0, V_t^0, t))^2 - 1}{\gamma(V_t^0, t)} \right) \\ &\quad \int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(1 - e^{-c\gamma^2(V_r^0, r)} \right)^2 dr. \end{aligned}$$

The term involving $\int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)} \right) dr$:

$$\begin{aligned} M_7 (S_t^0, V_t^0, t) &= \eta^2 K \frac{\phi(d_-(S_t^0, V_t^0, t))}{c\gamma(V_t^0, t)} \\ &\quad \int_t^T e^{2V_r^0} \gamma^8 (V_r^0, r) \left(e^{-c\gamma^2(V_r^0, r)} - e^{-2c\gamma^2(V_r^0, r)} \right)^2 dr. \end{aligned}$$

Step 5) Proof of $\tilde{f}_2(x, v, t)$. Finally, the proof of Equation (4.13) is completed by invoking Proposition 4.6, 4.7 and 4.8 to evaluate the integrals and sum up all the components. $\square$

Chapter 5

Perturbation of Ornstein-Uhlenbeck Stationary Distributions: expansion and simulation¹

Abstract

We consider a multidimensional stochastic differential equation Y written as a drift-perturbation of an ergodic Ornstein-Uhlenbeck process X . Under the condition of time-reversibility of X , we derive a first and second order expansion of the stationary distribution μ^Y of Y in terms of X . Error estimates are established. These approximations are then turned into a simulation scheme for sampling approximately according to μ^Y . Numerical experiments support the theoretical error estimates.

5.1 Introduction

The problem. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ be a filtered probability space supporting a q -dimensional Brownian motion W , with $q \geq 1$ and the usual assumptions on the filtration. Consider the $\mathbb{R}^d$ -valued solution $(Y_t)_{t \geq 0}$ of the stochastic differential equation

$$dY_t = \left[-AY_t + \beta(Y_t) \right] dt + \Sigma dW_t, \quad Y_0 \text{ independent of } W, \quad (5.1)$$

¹This chapter is based on a joint work with Professor Emmanuel Gobet from École Polytechnique.

where A is $d \times d$ -matrix, Σ is $d \times q$ -matrix and $\beta : \mathbb{R}^d \mapsto \mathbb{R}^d$ is a measurable function, quantities which satisfy these assumptions will be specified later. We assume that Y has a unique invariant measure μ^Y , and we are concerned with the approximation and simulation of μ^Y using an instrumental process X and its invariant distribution μ^X (assuming so far that it exists). Namely, X is the solution of the linear SDE (generalized Ornstein-Uhlenbeck equation)

$$dX_t = -AX_t dt + \Sigma dW_t, \quad X_0 \text{ independent of } W. \quad (5.2)$$

While μ^Y is hardly tractable and not easily simulable, μ^X is rather explicit (Gaussian distribution) and can be advantageously used to approximate μ^Y : our work provides a numerical scheme to achieve this purpose. Our main results (Theorems 5.1 and 5.2) state that

$$\mu^Y(dx) := (1 + c_k(x)) \mu^X(dx) \quad (5.3)$$

with some explicit correction terms c_k that depend of the approximation order k . Our analysis is supported by theoretical error estimates and numerical experiments. Essentially we consider that Y is a perturbation of X , and our analysis may be justified by β small in a suitable sense.

Background results. Our strategy is based on stochastic approximation techniques, in the spirit of the proxy expansion of [GM14]. This is quite different from the literature on small noise expansion [FS86], [FW12], or small time asymptotics [Wat87], [KP10], or even multiscale asymptotics of [Fou+11]. To the best of our knowledge, this is the first time that the problem of approximating invariant distribution is tackled, which gives rise to new challenges not encountered in previous works.

In this chapter, we will only derive a limited number of correction terms in (5.3) (first and second order approximation), but in principle any order could be achieved. The main term is the stationary distribution of the OU process X . In practice, the first and second order correction terms provide an excellent accuracy. The final second order approximation takes the form

$$\mu^Y(dx) \approx [1 + \mathbb{E}(g(X_{\tau_1}^x, X_{\tau_2}^x, X_{\tau_1+\tau_2}^x, \tau_1, \tau_2))] \mu^X(dx)$$

where g is explicit, X^x stands for the OU process starting from x and (τ_1, τ_2) are independent exponentially-distributed random variables. The above representation is very convenient to sample according to μ^Y , as we shall see later.

Approximation strategy. It consists of defining a suitable interpolation between the distributions of X and Y , and performing expansion along this interpolation. Actually, this is done for the marginal of the processes at time $T \rightarrow +\infty$. The crucial part lies in the derivation of explicit correction terms. As opposed to strong approximation techniques like in [Wat87], [FW12], [KP10], we have to control in the weak sense X_T and Y_T as $T \rightarrow +\infty$, since strong estimates explode in large time. To achieve this weak convergence analysis, we use in an essential manner the reversibility of the process X to pull back the computations in large time to small time, and derive in this way explicit representations. Furthermore, when taking $T \rightarrow +\infty$, we show that the expansion errors can be well controlled.

Outline of the chapter. We give the notations used throughout the paper. Section 5.2 gathers preliminary properties on X and Y . In particular, we provide necessary and sufficient conditions under which X is a reversible process; these results are interesting on their own, besides they will be crucially used in the subsequent expansion analysis. In Section 5.3, we state our main approximation results and analyse the errors. In Section 5.4, we investigate how to turn the expansion results into simulation algorithms of μ^Y and we perform some numerical experiments.

Notations and basic definitions. The following is used frequently throughout this paper.

Vector. $|x|$ is the Euclidean Norm of a vector $x \in \mathbb{R}^d$. Its sup-norm is denoted by $|x|_\infty$.

Its i th coordinate is x_i , or $x_{i,t}$ if $x = x_t$ depends on time.

Matrix. If $A = (A_{i,j})$ is a square matrix, $A^\top$ denote its transpose, $\text{Tr}(A)$ its trace and $\det(A)$ its determinant. $\lambda(A)$ denotes the spectrum of A .

Id is the $d \times d$ -identity matrix.

For a square matrix A , $\lambda_{\max,A}$ and $\lambda_{\min,A}$ stand for the maximum and minimum eigenvalues of A .

Definition 5.1 (p.135 of [Kha02]). A square matrix H is called a Hurwitz matrix if its eigenvalues have negative real parts, i.e. its spectrum $\lambda(H)$ is included in $\{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) < 0\}$.

$\|A\|$ is the matrix norm of A subordinated to the Euclidean norm.

For a vector (resp. matrix) valued function β , $|\beta|_\infty$ (resp. $\|\beta\|_\infty$) stands for $\sup_{t,x} |\beta(t,x)|$ (resp. $\sup_{t,x} \|\beta(t,x)\|$).

$\operatorname{vec}(A)$ is the vectorizing operator applied to A , it stacks the columns of A into a vector.

The Kronecker Product $A \otimes B$ of matrix $A = (A_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ and $B = (B_{ij})_{1 \leq i \leq k, 1 \leq j \leq l}$ is defined as

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix},$$

where $A \otimes B$ is a $mk \times nl$ matrix.

We also recall that if matrices M_1 , M_2 and M_3 are conformable for matrix multiplication then

$$\operatorname{vec}(M_1 M_2 M_3) = (M_3^\top \otimes M_1) \operatorname{vec}(M_2). \quad (5.4)$$

If M_1 and M_2 are two matrices of the same size, then

$$\operatorname{Tr}(M_1^\top M_2) = (\operatorname{vec}(M_1))^\top \operatorname{vec}(M_2).$$

Functions. For a vector (resp. matrix) valued function φ , $|\varphi|_\infty$ (resp. $\|\varphi\|_\infty$) stands for $\sup_{t,x} |\varphi(t,x)|$ (resp. $\sup_{t,x} \|\varphi(t,x)\|$).

For a smooth function $\varphi(w)$, $\partial_{w_i} \varphi(w)$ stands for the partial derivative of φ with respect to w_i .

If $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^q$, then $D\varphi(x)$ is defined as the $q \times d$ -matrix

$$D\varphi(x) := [\partial_{x_1}\varphi(x), \dots, \partial_{x_d}\varphi(x)].$$

For a smooth drift function $\beta : \mathbb{R}^d \mapsto \mathbb{R}^d$, we set $\beta^{(1)}(x) := D\beta(x)$, $\beta^{(2)}(x) := D \operatorname{vec}(\beta^{(1)})(x)$ and $\beta^{(3)}(x) := D \operatorname{vec}(\beta^{(2)})(x)$.

For a smooth test function $h : \mathbb{R}^d \mapsto \mathbb{R}$, we set $h^{(1)}(x) := Dh(x)$ and $h^{(2)}(x) := (\partial_{x_i, x_j} h(x))_{i,j}$.

The divergence of $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is $\operatorname{div}(\varphi) := \operatorname{Tr}(D\varphi)$.

Generic constants. We shall denote as c , all constants which depends only on universal constants and the dimension d .

5.2 Preliminary results

Exponential of a Hurwitz matrix

We will repeatedly use the large time behavior of e^{Ht} as $t \rightarrow +\infty$, for a given square matrix H . Consider its Jordan matrix canonical decomposition (Chapter 1, Section 7 of [FF07]), $H = T_H J_H T_H^{-1}$, for which T_H comprises the generalized eigenvectors of H and J_H is block diagonal, i.e.

$$J_H = \begin{pmatrix} J_{k_{1,H}}(\lambda_{1,H}) & 0 & 0 & \dots & 0 \\ 0 & J_{k_{2,H}}(\lambda_{2,H}) & 0 & \dots & 0 \\ 0 & 0 & J_{k_{3,H}}(\lambda_{3,H}) & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \dots & J_{k_{m,H}}(\lambda_{m,H}) \end{pmatrix}$$

and each block has the form

$$J_{k_{i,H}}(\lambda_{i,H}) = \begin{pmatrix} \lambda_{i,H} & 1 & 0 & \dots & 0 \\ 0 & \lambda_{i,H} & 1 & \dots & 0 \\ 0 & 0 & \lambda_{i,H} & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda_{i,H} \end{pmatrix} \quad (k_{i,H} \times k_{i,H})$$

where $k_{i,H}$ is the multiplicity of $\lambda_{i,H}$.

Then, we can write $\exp(Ht) = T_H \exp(J_H t) T_H^{-1}$. Since J_H is block diagonal, so is $\exp(J_H t)$; in particular, for each submatrix of the form $J_{k_{i,H}}(\lambda_{i,H})$ on the block diagonal of J_H , $\exp(J_H t)$ will contain the diagonal block

$$\begin{aligned} & \exp(J_{k_{i,H}}(\lambda_{i,H})t) \\ &= \exp(\lambda_{i,H}t) \begin{pmatrix} 1 & t & \frac{1}{2}t^2 & \dots & \frac{1}{(k_{i,H}-1)!}t^{k_{i,H}-1} \\ 0 & 1 & t & \dots & \frac{1}{(k_{i,H}-2)!}t^{k_{i,H}-2} \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & t \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} (k_{i,H} \times k_{i,H}). \end{aligned}$$

Then, it becomes clear that if H is a Hurwitz matrix, e^{Ht} converges to 0, with some precise exponential rate. We have proved the following result.

Lemma 5.1. *If H is a Hurwitz matrix, then there exists $c_H > 0$ and $c_{\text{vec}(H)}$ such that*

$$\|e^{Ht}\| \leq c_H e^{-\lambda_H t} \quad \text{and} \quad |\text{vec}(e^{Ht})| \leq c_{\text{vec}(H)} e^{-\lambda_H t} \quad \forall t \geq 0.$$

for any $\lambda_H \in (0, -\max_{\lambda \in \lambda(H)} \text{Re}(\lambda))$.

The generalized Ornstein-Uhlenbeck process

We collect in this paragraph the main results on X . It will serve as a basis for our work. The two following statements are standard, see for instance Section 5.6 of [KS91].

Proposition 5.1. *Assume that the random variable X_0 is independent of W and is square integrable. Then, there exists a unique square integrable solution to (5.2), represented as*

$$X_t = e^{-At} \left(X_0 + \int_0^t e^{Ar} \Sigma dW_r \right). \quad (5.5)$$

Its mean $m_t := \mathbb{E}[X_t]$ and covariance $V_{t,s} := \mathbb{E}[X_t X_s^\top]$ are given by

$$m_t := e^{-At} \mathbb{E}[X_0], \quad (5.6)$$

$$V_{t,s} := e^{-At} \left(\mathbb{E}[X_0 X_0^\top] + \int_0^{t \wedge s} e^{Ar} \Sigma \Sigma^\top e^{A^\top r} dr \right) e^{-A^\top s} \quad (5.7)$$

for any $s \geq 0$ and $t \geq 0$. Additionally, for any $t \geq s \geq 0$, the distribution of X_t conditioned on X_s is Gaussian

$$\mathcal{N} \left(e^{-A(t-s)} X_s, \int_0^{t-s} e^{-Ar} \Sigma \Sigma^\top e^{-A^\top r} dr \right). \quad (5.8)$$

Proof. (5.5) is a direct consequence of the Itô formula. Identities (5.6)-(5.7)-(5.8) are straightforward. $\square$

Proposition 5.2. Assume that $-A$ is a Hurwitz matrix. Then, X has a unique stationary distribution μ^X which is Gaussian with mean 0 and covariance

$$V_\infty := \int_0^{+\infty} e^{-Ar} \Sigma \Sigma^\top e^{-A^\top r} dr. \quad (5.9)$$

Under the stationary distribution, we have

$$V_{t,s} = e^{-A(t-s)} V_\infty \quad (5.10)$$

for any $t \geq s \geq 0$.

Proof. Owing to Lemma 5.1, both matrices e^{-At} and $e^{-A^\top t}$ converge to 0 exponentially fast as $t \rightarrow +\infty$. Then, the stochastic integral $e^{-At} \int_0^t e^{Ar} \Sigma dW_r$, as $t \rightarrow +\infty$, is Gaussian, centered, with covariance $\int_0^t e^{-As} \Sigma \Sigma^\top e^{-A^\top s} ds \rightarrow V_\infty$: thus, the related Wiener stochastic integral converges weakly to μ^X . Since $e^{-At} X_0$ converges almost surely to 0 and in view of (5.5), we have proved the weak convergence of X to μ^X . If X_0 has the Gaussian distribution μ^X , then X is a Gaussian process, with mean 0 (see (5.6)): the covariance matrix of X_t is (see (5.7))

$$V_{t,t} = e^{-At} \left(V_\infty + \int_0^t e^{Ar} \Sigma \Sigma^\top e^{A^\top r} dr \right) e^{-A^\top t} = V_\infty. \quad (5.11)$$

From the above and (5.7), we easily deduce (5.10). $\square$

In our subsequent approximation approach (Section 5.3), we make use of the property of time-reversibility of X , i.e. for any non-negative time t and s , under the stationary distribution, the processes $(X_r)_{s \leq r \leq t}$ and $(X_{s+t-r})_{s \leq r \leq t}$ have the same distribution. In the case of discrete-time Gaussian linear processes, this time-reversibility question is investigated in [TZ05]. Here in our continuous-time framework, we establish a rather explicit equivalence criterion with a specific proof. To the best of our knowledge, this statement is new.

Proposition 5.3. *Assume that the assumptions of Proposition 5.2 hold and that X is considered under the stationary distribution μ^X . Then X is reversible if and only if $A\Sigma\Sigma^\top$ is a symmetric matrix, i.e.*

$$A\Sigma\Sigma^\top = \Sigma\Sigma^\top A^\top. \quad (5.12)$$

Furthermore, we have in this case

$$V_\infty = \frac{1}{2}A^{-1}\Sigma\Sigma^\top. \quad (5.13)$$

Proof. Note that (5.13) directly follows from (5.9) and (5.12).

Because of the Gaussian properties of X , time-reversibility is equivalent to $V_{t,s} = V_{s,t}$ for any $t \geq s \geq 0$. Observe that $V_{s,t} = V_{t,s}^\top$.

Proof of $\Leftarrow$: we first prove by induction that $A^k\Sigma\Sigma^\top = \Sigma\Sigma^\top (A^\top)^k$ for any $k \geq 1$. This is true for $k = 1$ because of (5.12); assuming the property for k , then $A^{k+1}\Sigma\Sigma^\top = A(A^k\Sigma\Sigma^\top) = A(\Sigma\Sigma^\top (A^\top)^k) = \Sigma\Sigma^\top A^\top (A^\top)^k = \Sigma\Sigma^\top (A^\top)^{k+1}$, therefore the announced property.

Second, by writing the matrix exponential as a series, it readily follows that $e^{-At}\Sigma\Sigma^\top = \Sigma\Sigma^\top e^{-A^\top t}$ for any $t \geq 0$. We finally deduce

$$\begin{aligned} V_{t,s} &\stackrel{(5.10)}{=} e^{-A(t-s)}V_\infty \stackrel{(5.9)}{=} \int_0^{+\infty} e^{-Ar}e^{-A(t-s)}\Sigma\Sigma^\top e^{-A^\top r}dr \\ &\stackrel{(5.12)}{=} \int_0^{+\infty} e^{-Ar}\Sigma\Sigma^\top e^{-A^\top(t-s)}e^{-A^\top r}dr \\ &= V_\infty e^{-A^\top(t-s)} = V_{s,t}, \end{aligned}$$

therefore the time-reversibility of X .

Proof of $\Rightarrow$: assume $V_{t,s} = V_{s,t}$ for any $t \geq s \geq 0$. For such t and s , write $V_{t,s}$ in a form similar to (5.11), i.e.

$$V_{t,s} = e^{-At} \left(V_\infty + \int_0^s e^{Ar}\Sigma\Sigma^\top e^{A^\top r}dr \right) e^{-A^\top s}.$$

A straightforward differentiation w.r.t. s and t ($t > s$) gives

$$\begin{aligned} \partial_t V_{t,s} &= -AV_{t,s}, & \partial_s V_{t,s} &= -V_{t,s}A^\top + e^{A(s-t)}\Sigma\Sigma^\top, \\ \partial_{s,t}^2 V_{t,s} &= AV_{t,s}A^\top - Ae^{A(s-t)}\Sigma\Sigma^\top. \end{aligned} \quad (5.14)$$

Owing to $V_{s,t} = V_{t,s}^\top$, we deduce $\partial_{s,t}^2 V_{s,t} = (AV_{t,s}A^\top)^\top - (Ae^{A(s-t)}\Sigma\Sigma^\top)^\top$. Now, invoking the time-reversibility and identifying the previous equality with (5.14), we obtain $Ae^{-Ar}\Sigma\Sigma^\top = \Sigma\Sigma^\top e^{-A^\top r}A^\top$ for any $r = t - s > 0$. By taking the limit as $r \downarrow 0$, we obtain the advertised equality (5.12). $\square$

Examples 5.1. *Let us consider the assumptions of Proposition 5.2.*

- i) *If A is symmetric, and A and $\Sigma\Sigma^\top$ commute, then reversibility holds.*
- ii) *In the case $\Sigma\Sigma^\top = c \text{Id}$ for some $c > 0$, reversibility is equivalent to symmetry of A .*
- iii) *The symmetry of A is not sufficient for time-reversibility: for instance, take $A = \text{diag}(1, 2)$ and $\Sigma\Sigma^\top = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ for $\rho \neq 0$. Indeed $A\Sigma\Sigma^\top = \begin{pmatrix} 1 & \rho \\ 2\rho & 2 \end{pmatrix}$ is not symmetric, therefore X will not be time-reversible.*
- iv) *There are many situations of reversibility for which A is not symmetric: take $\Sigma\Sigma^\top$ as before and set $A = \begin{pmatrix} 1 & \rho(\lambda - 1) \\ 0 & \lambda \end{pmatrix}$ for some $\lambda \in (0, +\infty) \setminus \{1\}$. A is not symmetric but $A\Sigma\Sigma^\top$ is, and hence reversibility holds.*

Extension to non-centered Ornstein-Uhlenbeck process. For the sake of completeness, we briefly expose the method of handling the case of a constant term in the drift coefficient of (5.1) and (5.2), by using a simple translation of the variables. Namely, consider

$$dX_t^a = (-AX_t^a + a)dt + \Sigma dW_t, \quad X_0^a \text{ independent of } W,$$

with $a \in \mathbb{R}^d$, and similarly for Y . Consider the assumptions of Proposition 5.2 and set $m = a \int_0^\infty e^{-Ar} dr$. A similar analysis yields that X^a has a stationary Gaussian distribution $\mathcal{N}(m, V_\infty)$ (see Equations (6.10)-(6.11) p.355 of [KS91]): moreover, we have the relation $Am = a$ (see Equation (6.12) p.355 of [KS91]). Now, if we put $X_t = X_t^a - m$, we obtain $dX_t = (-AX_t + (a - Am))dt + \Sigma dW_t = -AX_t dt + \Sigma dW_t$. It means that all subsequent results can be translated from the case $a = 0$ to the case $a \neq 0$, by shifting the space variable from m . In the error estimates of Theorems 5.1 and 5.2, we would have to replace β by

$\beta - a$. Heuristically, for a given β , the parameter a should be chosen such that m is approximately equal to the mean of μ^Y . These heuristics are not captured in our subsequent quantitative analysis, however this centering property presumably brings further accuracy.

The perturbed Ornstein-Uhlenbeck process

Recall the process of interest Y defined by (5.1). We impose that β is a Lipschitz function, ensuring the existence and uniqueness of a strong solution (with square integrability) for any initial condition Y_0 independent of W and square integrable (see Theorem 2.9 p.289 of [KS91]). Reinforcing the conditions lead to ergodic properties for Y , as stated in the following standard result.

Proposition 5.4. *Assume that $-A$ is a Hurwitz matrix, that $\Sigma\Sigma^\top$ is non-singular and that β is a bounded Lipschitz function. Then, Y has a unique stationary distribution μ^Y , which has a density with respect to the Lebesgue measure; thus we denote $\mu^Y(dy) = \mu^Y(y)dy$.*

Furthermore, for any continuous bounded function $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ and any $x \in \mathbb{R}^d$, we have

$$\lim_{t \rightarrow +\infty} \mathbb{E}[\varphi(Y_t) | Y_0 = x] = \int_{\mathbb{R}^d} \varphi(y) \mu^Y(dy). \quad (5.15)$$

Proof. To prove the existence and uniqueness of a stationary distribution, we apply Theorem 4.1 and Corollary 4.4 of [Kha12] by checking the validity of Assumptions (B1)-(B2), p. 107 of [Kha12]. The related assumption (B1) is satisfied since $\Sigma\Sigma^\top$ is non-degenerate. To check (B2) (the mean recurrence time is locally uniform w.r.t. the starting point), we apply Theorem 3.9 of [Kha12] with the Lyapunov function $V(x) = x \cdot P^\# x$ where $P^\#$ is a symmetric non-negative matrix to be specified hereafter. Denoting by L^Y the infinitesimal generator associated to Y , we only have to check that $L^Y V(x) \leq -1$ for $|x|$ sufficiently large. Simple computations give

$$\begin{aligned}
L^Y V(x) &= 2x^\top P^\#(-Ax + \beta(x)) + \text{Tr}(P^\# \Sigma \Sigma^\top) \\
&= -x^\top (P^\# A + A^\top P^\#)x + 2x \cdot P^\# \beta(x) + \text{Tr}(P^\# \Sigma \Sigma^\top) \\
&\leq -x^\top (P^\# A + A^\top P^\#)x + \frac{1}{2}|x|^2 + 2|P^\# \beta|_\infty^2 + \text{Tr}(P^\# \Sigma \Sigma^\top).
\end{aligned}$$

Now, for a given symmetric positive-definite matrix Q , we consider solutions P to the Lyapunov equation

$$PA + A^\top P = Q.$$

By Theorem 4.6 of [Kha02] and since $-A$ is Hurwitz, there is a unique solution P_Q in the class of symmetric positive-definite matrices. Denote by $P^\#$, the solution associated to $Q = \text{Id}$. This choice ensures that the related Lyapunov function V satisfies $L^Y V(x) \leq -\frac{1}{2}|x|^2 + 2|P^\# \beta|_\infty^2 + \text{Tr}(P^\# \Sigma \Sigma^\top)$, therefore the requested behavior for $|x|$ large.

The existence of a density with respect to the Lebesgue measure is a direct consequence of Lemma 4.16 of [Kha12].

Finally, Theorem 4.3 of [Kha12] gives the convergence (5.15) and we are done. $\square$

5.3 Main approximation results

We start by listing the working assumptions we need, in order to ensure the validity of the approximation.

(H-i) $-A$ is a Hurwitz matrix.

Choose $\lambda_{0,A} \in (0, \min_{\lambda \in \lambda(A)} \text{Re}(\lambda))$: for such $\lambda_{0,A}$, define

$$A_0 := A - \lambda_{0,A} \text{Id};$$

note that $-A_0$ is a Hurwitz matrix. Denote by $c_A > 0$ the constant such that $\|\exp(-As)\| \leq c_A \exp(-\lambda_{0,A}s)$, $\forall s \geq 0$ (see Lemma 5.1).

(H-ii) $A \Sigma \Sigma^\top$ is symmetric.

(H-iii) $\Sigma \Sigma^\top$ is non-singular.

(H-iv) β is a $d \times 1$ vector and all its derivatives up till third order are bounded.

(H-v) h is a continuous bounded function and its derivatives up till third order exist and are bounded. (Note here that h is a test function used in our derivation.)

(H-vi) $\lambda_{0,A} > c_A \|\beta^{(1)}\|_\infty$, where $\lambda_{0,A}, c_A > 0$ are the constants from **(H-i)**. We set

$$c_{(5.16)} := \frac{c_A}{\lambda_{0,A} - c_A \|\beta^{(1)}\|_\infty}. \quad (5.16)$$

5.3.1 Strategy

For the sake of pedagogy, we outline the main arguments for the first-order approximation only. The full justification is given in Subsection 5.3.3. We aim at approximating

$$\int_{\mathbb{R}^d} h(y) \mu^Y(y) dy - \int_{\mathbb{R}^d} h(x) \mu^X(x) dx$$

for some test functions h in a certain class (assumption **(H-v)**). To study the approximation of μ^Y by μ^X , we use the solution denoted by $\tilde{Y}$ starting from μ^X (independent of W):

$$d\tilde{Y}_t = \left(-A\tilde{Y}_t + \beta(\tilde{Y}_t) \right) dt + \Sigma dW_t, \quad \tilde{Y}_0 \stackrel{d}{=} \mathcal{N}(0, V_\infty) = \mu^X.$$

By the sake of symmetry, we also denote by $\tilde{X}$ the process solution without the above β -term, starting from μ^X .

Finite-time approximation. Under the standing assumption **(H-i)**, recall that V_∞ is properly defined in (5.9). For a given $T > 0$, consider furthermore that

$$\begin{aligned} \mathbb{E}_x \left[h(\tilde{Y}_T) \right] &:= \mathbb{E} \left[h(\tilde{Y}_T) \mid \tilde{Y}_0 = x \right], \\ \mathbb{E}_{\mu^X} \left[h(\tilde{Y}_T) \right] &:= \int_{\mathbb{R}^d} \mathbb{E}_x \left[h(\tilde{Y}_T) \right] \mu^X(x) dx. \end{aligned}$$

First, we get the following lemma.

Lemma 5.2. *Assume (H-i)-(H-ii)-(H-iii)-(H-iv). We have for any continuous bounded function $h : \mathbb{R}^d \rightarrow \mathbb{R}$,*

$$\lim_{T \rightarrow +\infty} \mathbb{E}_{\mu^X} \left[h(\tilde{Y}_T) \right] = \int_{\mathbb{R}^d} h(y) \mu^Y(y) dy, \quad (5.17)$$

$$\lim_{T \rightarrow +\infty} \mathbb{E}_{\mu^X} \left[h(\tilde{X}_T) \right] = \int_{\mathbb{R}^d} h(x) \mu^X(x) dx. \quad (5.18)$$

Parameterization. Then, we introduce the d -dimensional parameterized process

$$d\tilde{X}_t^\epsilon = \left(-A\tilde{X}_t^\epsilon + \epsilon\beta(\tilde{X}_t^\epsilon) \right) dt + \Sigma dW_t, \quad \tilde{X}_0^\epsilon \stackrel{d}{=} \mu^X \quad (5.19)$$

for $\epsilon \in [0, 1]$ and we define

$$\tilde{X}_{i,t}^\epsilon = \frac{\partial^i \tilde{X}_t^\epsilon}{\partial \epsilon^i}, \quad \tilde{X}_{i,t} = \frac{\partial^i \tilde{X}_t^\epsilon}{\partial \epsilon^i} \Big|_{\epsilon=0}.$$

These derivatives up to the second order are well defined by an application of [Kun97]. A simple computation shows that $\tilde{X}_{0,t} = \tilde{X}_t$ and

$$d\tilde{X}_{1,t} = \left(-A\tilde{X}_{1,t} + \beta(\tilde{X}_t) \right) dt, \quad \tilde{X}_{1,0} = 0. \quad (5.20)$$

With this d -dimensional parameterization, we can derive a decomposition of $\mathbb{E}_{\mu^X} \left[h(\tilde{Y}_T) \right]$.

Decomposition. For any continuous bounded function h , having bounded derivatives, we expand $\mathbb{E}_{\mu^X} \left[h(\tilde{Y}_T) \right]$ to the first order. Applying the Taylor expansion to $\mathbb{E}_{\mu^X} \left[h(\tilde{X}_T^\epsilon) \right]$ for $\epsilon = 1$ around $\epsilon = 0$, we have

$$\mathbb{E}_{\mu^X} \left[h(\tilde{Y}_T) \right] = \mathbb{E}_{\mu^X} \left[h(\tilde{X}_T) \right] + C_1(T) + \text{Error}_1(T) \quad (5.21)$$

where

$$\begin{aligned} C_1(T) &:= \mathbb{E}_{\mu^X} \left[h^{(1)}(\tilde{X}_T) \tilde{X}_{1,T} \right], \\ \text{Error}_1(T) &:= \int_0^1 (1-u) \mathbb{E}_{\mu^X} \left[(\tilde{X}_{1,T}^u)^\top h^{(2)}(\tilde{X}_{0,T}^u) (\tilde{X}_{1,T}^u) \right. \\ &\quad \left. + h^{(1)}(\tilde{X}_{0,T}^u) \tilde{X}_{2,T}^u \right] du. \end{aligned} \quad (5.22)$$

Taking limits as $T \rightarrow +\infty$. Owing to Lemma 5.2, the first two terms of (5.21) converge respectively to

$$\int_{\mathbb{R}^d} h(y) \mu^Y(y) dy \quad \text{and} \quad \int_{\mathbb{R}^d} h(y) \mu^X(y) dy.$$

Besides, we will also show that the first correction term converges (see Equation (5.39)):

$$\lim_{T \rightarrow +\infty} C_1(T) = \int_{\mathbb{R}^d} h(x) c_1(x) \mu^X(x) dx$$

where $c_1(x)$ is given in Theorem 5.1. Here, we use *Time-Reversibility of X* to obtain Equation (5.36), which allows us to derive the above limit. Moreover, we also prove that the error term, $\text{Error}_1(T)$, is uniformly bounded (when $T \rightarrow +\infty$) and controlled like Estimate (5.26) by assuming Assumption **(H-vi)**. Therefore, this shows the approximate stationary distribution of μ^Y to be

$$\mu^{Y,1}(dx) = (1 + c_1(x)) \mu^X(dx),$$

with an error bound for smooth test functions h .

5.3.2 First and second order expansions

We now state the main results obtained by using the above approximation strategy.

Theorem 5.1 (First-order approximation). *If we assume **(H-i)** to **(H-vi)**, then we have*

$$\int_{\mathbb{R}^d} h(y) \mu^Y(y) dy = \int_{\mathbb{R}^d} h(x) (1 + c_1(x)) \mu^X(x) dx + \text{Error}_1, \quad (5.23)$$

where

$$c_1(x) = \mathbb{E}_x \left[\frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta \left(\tilde{X}_{\tau_1} \right)}{\lambda_{0,A}} - \frac{\text{Tr} \left(e^{-A_0 \tau_1} \beta^{(1)} \left(\tilde{X}_{\tau_1} \right) e^{-A \tau_1} \right)}{\lambda_{0,A}} \right] \quad (5.24)$$

and $\tau_1 \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ is independent of the Brownian motion W . Additionally, we have

$$|c_1(x)| \leq \frac{2c_A}{\lambda_{0,A}} |x| \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty + \frac{dc_A^2 |\beta^{(1)}|_\infty}{2\lambda_{0,A}}, \quad (5.25)$$

$$\begin{aligned} |\text{Error}_1| &\leq c \left(c_{(5.16)}^2 |\beta|_\infty^2 \|h^{(2)}\|_\infty \right. \\ &\quad \left. + (c_{(5.16)}^2 \|\beta^{(1)}\|_\infty |\beta|_\infty + c_{(5.16)}^3 \|\beta^{(2)}\|_\infty |\beta|_\infty^2) |h^{(1)}|_\infty \right). \end{aligned} \quad (5.26)$$

As a consequence, the stationary distribution of Y is approximated by

$$\mu^{Y,1}(\mathrm{d}x) := (1 + c_1(x)) \mu^X(\mathrm{d}x). \quad (5.27)$$

We refer to (5.23) as a first-order approximation since the error term is quadratic in β and its derivatives. For the purpose of simulation in Section 5.4, we shall compute an explicit upper bound on c_1 as we do in (5.25).

Observe that $\mu^{Y,1}(x)\mathrm{d}x$ is a finite measure on $\mathbb{R}^d$ (owing to (5.25)), its mass equals 1 (take $h = 1$ in (5.23)) but it is not necessarily a density (even for small β). This is discussed later in Section 5.4.

Theorem 5.2 (Second-order approximation). *If we assume (H-i) to (H-vi), then we have*

$$\begin{aligned} \int_{\mathbb{R}^d} h(y) \mu^Y(y) \mathrm{d}y &= \int_{\mathbb{R}^d} h(x) (1 + c_1(x) + c_{21}(x) + c_{22}(x)) \mu^X(x) \mathrm{d}x \\ &\quad + \text{Error}_2, \end{aligned} \quad (5.28)$$

where

$$c_{21}(x) = \mathbb{E}_x \left[\frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_1 + \tau_2})}{\lambda_{0,A}^2} \right. \\ \left. - \frac{1}{\lambda_{0,A}^2} \text{Tr} \left(\left(\left(\beta^\top(\tilde{X}_{\tau_1 + \tau_2}) e^{-A_0^\top \tau_2} \right) \otimes e^{-A_0 \tau_1} \right) \beta^{(2)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right. \right. \\ \left. \left. + e^{-A \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_2} \beta^{(1)}(\tilde{X}_{\tau_1 + \tau_2}) e^{-A_0(\tau_1 + \tau_2)} \right) \right]$$

and

$$c_{22}(x) = \mathbb{E}_x \left[\left(-2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0(\tau_1 + \tau_2)} \beta(\tilde{X}_{\tau_2}) \right. \right. \\ + \text{Tr} \left(\left(\left(\beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top(\tau_1 + \tau_2)} \right) \otimes e^{-A \tau_1} \right) \beta^{(2)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right) \\ + \text{Tr} \left(e^{-A \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0(\tau_1 + \tau_2)} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A \tau_2} \right) \\ + 4\beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top \tau_2} A^\top (\Sigma \Sigma^\top)^{-1} x x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \\ - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A \tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A_0(\tau_1 + \tau_2)} \beta(\tilde{X}_{\tau_1}) \\ - 2\beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top \tau_2} A^\top (\Sigma \Sigma^\top)^{-1} e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \\ - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_2}) \text{Tr} \left(e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right) \\ - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \text{Tr} \left(e^{-A_0 \tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A \tau_2} \right) \\ + \left(\text{vec} \left(e^{-2A^\top \tau_2} \right) \right)^\top \beta^{(2)}(\tilde{X}_{\tau_2}) e^{-A_0(\tau_1 + \tau_2)} \beta(\tilde{X}_{\tau_1}) \\ \left. + \text{Tr} \left(e^{-A_0 \tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A \tau_2} \right) \text{Tr} \left(e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right) \right] \frac{\mathbb{1}_{\tau_1 < \tau_2}}{\lambda_{0,A}^2}.$$

with $\tau_2 \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ independent of τ_1 and the Brownian motion W . Additionally, we have

$$|c_{21}(x)| \leq \frac{2c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty \|\beta^{(1)}\|_\infty |x|}{\lambda_{0,A}^2} + \frac{dc_A^3 |\beta|_\infty \|\beta^{(2)}\|_\infty}{2\lambda_{0,A}^2} \\ + \frac{dc_A^3 \|\beta^{(1)}\|_\infty^2}{4\lambda_{0,A}^2}$$

and

$$\begin{aligned}
& |c_{22}(x)| \\
& \leq \frac{c_A \|\beta^{(2)}\|_\infty |\beta|_\infty}{4\lambda_{0,A}^2} \left(dc_A^2 + \frac{c_{\text{vec}}(-2A^\top)}{3} \right) + \frac{dc_A^3 \|\beta^{(1)}\|_\infty^2}{8\lambda_{0,A}^2} (dc_A + 1) \\
& \quad + \frac{c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \|\beta^{(1)}\|_\infty |\beta|_\infty |x|}{\lambda_{0,A}^2} (1 + dc_A) \\
& \quad + \frac{c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty^2}{\lambda_{0,A}^2} \left(2|x|^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| + 1 \right).
\end{aligned}$$

Moreover, the Error term, bounded by

$$\begin{aligned}
& |\text{Error}_2| \\
& \leq c \left(|h^{(3)}|_\infty c_{(5.16)}^3 |\beta|_\infty^3 \right. \\
& \quad + \|h^{(2)}\|_\infty c_{(5.16)}^3 |\beta|_\infty^2 \left(\|\beta^{(1)}\|_\infty + c_{(5.16)} \|\beta^{(2)}\|_\infty |\beta|_\infty \right) \\
& \quad + |h^{(1)}|_\infty c_{(5.16)}^3 |\beta|_\infty \left(\left(\|\beta^{(1)}\|_\infty + c_{(5.16)} \|\beta^{(2)}\|_\infty |\beta|_\infty \right)^2 \right. \\
& \quad \left. \left. + |\beta|_\infty \left(\|\beta^{(2)}\|_\infty + c_{(5.16)} \|\beta^{(3)}\|_\infty |\beta|_\infty \right) \right) \right). \tag{5.29}
\end{aligned}$$

Consequently, the stationary distribution of Y is approximated by

$$\mu^{Y,2}(\mathrm{d}x) := (1 + c_1(x) + c_{21}(x) + c_{22}(x)) \mu^X(\mathrm{d}x). \tag{5.30}$$

Since the error term is cubic in β and its derivatives, (5.28) is a second-order approximation. Here again, $\mu^{Y,2}(x) \mathrm{d}x$ is a finite measure in $\mathbb{R}^d$ with mass equal to 1 but not necessarily a probability density.

5.3.3 Proof of Theorem 5.1

The proof is split into several steps.

Step 1) Preliminary estimates. The following result is instrumental in our analysis.

Proposition 5.5. *Let Z be the $\mathbb{R}^d$ -valued solution to*

$$dZ_s = ((-A + \alpha_s) Z_s + \gamma_s) \mathrm{d}s, \quad Z_0 = 0,$$

where α and γ are two bounded measurable functions (resp. a $d \times d$ matrix and a vector in $\mathbb{R}^d$) and A is a $d \times d$ matrix function such that $\|e^{-As}\| \leq c_A e^{-\lambda_{0,A}s}$ ($c_A > 0$, $\lambda_{0,A} > 0$). Then, if $\lambda_{0,A} > c_A \|\alpha\|_\infty$, we have

$$\sup_{T \geq 0} |Z_T| \leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A} - c_A \|\alpha\|_\infty}. \quad (5.31)$$

Moreover, if we assume that $B(t) := -A + \alpha_t$ commutes with $B(s)$ for every s and t , then Z can be written as

$$Z_T = \int_0^T \exp\left(\int_s^T B(u) du\right) \gamma_s ds. \quad (5.32)$$

Proof. The second statement is easily derived by writing (under the commutativity assumption on $(B(s) : s \geq 0)$)

$$e^{-\int_0^t B(u) du} Z_t = \int_0^t e^{-\int_0^s B(u) du} \gamma_s ds,$$

then simplifying (still under the commutativity assumption).

The representation (5.32) can be an intermediate step in the derivation of (5.31). However, in our analysis, we need (5.31) without the restrictive commutativity assumption. We employ another approach using Gronwall inequality. Use (5.32) with the commutative family $B(s) = -A$ (by adjusting γ to $\gamma + \alpha Z$), hence

$$Z_t = \int_0^t e^{-A(t-s)} \gamma_s ds + \int_0^t e^{-A(t-s)} \alpha_s Z_s ds.$$

Then take the norm and apply Minkowski inequality: it follows that

$$\begin{aligned} |Z_t| &\leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A}} + \int_0^t c_A \|\alpha\|_\infty e^{-\lambda_{0,A}(t-s)} |Z_s| ds, \\ e^{\lambda_{0,A}t} |Z_t| &\leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A}} e^{\lambda_{0,A}t} + \int_0^t c_A \|\alpha\|_\infty e^{\lambda_{0,A}s} |Z_s| ds. \end{aligned}$$

The Gronwall inequality yields

$$\begin{aligned}
e^{\lambda_{0,A}t} |Z_t| &\leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A}} e^{\lambda_{0,A}t} \\
&\quad + \int_0^t \frac{c_A^2 |\gamma|_\infty \|\alpha\|_\infty}{\lambda_{0,A}} e^{\lambda_{0,A}s} \exp(c_A \|\alpha\|_\infty (t-s)) ds, \\
|Z_t| &\leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A}} \\
&\quad + \frac{c_A^2 |\gamma|_\infty \|\alpha\|_\infty}{\lambda_{0,A}} \int_0^t \exp(-(\lambda_{0,A} - c_A \|\alpha\|_\infty)(t-s)) ds \\
&\leq \frac{c_A |\gamma|_\infty}{\lambda_{0,A} - c_A \|\alpha\|_\infty}.
\end{aligned}$$

□

The previous proposition allows us to directly control uniformly in time the process $\tilde{X}_{1,\cdot}^\epsilon$, solution to

$$d\tilde{X}_{1,t}^\epsilon = \left(-A\tilde{X}_{1,t}^\epsilon + \beta \left(\tilde{X}_{0,t}^\epsilon \right) + \epsilon \beta^{(1)} \left(\tilde{X}_{0,t}^\epsilon \right) \tilde{X}_{1,t}^\epsilon \right) dt, \quad \tilde{X}_{1,0}^\epsilon = 0.$$

Lemma 5.3. Assume **(H-i)**-**(H-iv)**-**(H-vi)**. We have, for any $\varepsilon \in [0, 1]$,

$$\sup_{T \geq 0} \left| \tilde{X}_{1,T}^\epsilon \right| \leq c_{(5.16)} |\beta|_\infty.$$

We proceed similarly for $\tilde{X}_{2,\cdot}^\epsilon$, solution to

$$\begin{aligned}
d\tilde{X}_{2,t}^\epsilon &= \left(-A\tilde{X}_{2,t}^\epsilon + 2\beta^{(1)} \left(\tilde{X}_{0,t}^\epsilon \right) \tilde{X}_{1,t}^\epsilon + \epsilon \beta^{(1)} \left(\tilde{X}_{0,t}^\epsilon \right) \tilde{X}_{2,t}^\epsilon \right. \\
&\quad \left. + \epsilon \partial_\epsilon \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\epsilon \right) \right) \tilde{X}_{1,t}^\epsilon \right) dt, \quad \tilde{X}_{2,0}^\epsilon = 0. \quad (5.33)
\end{aligned}$$

Lemma 5.4. Assume **(H-i)**-**(H-iv)**-**(H-vi)**. We have, for any $\varepsilon \in [0, 1]$,

$$\sup_{T \geq 0} \left| \tilde{X}_{2,T}^\epsilon \right| \leq c \left(c_{(5.16)}^2 \|\beta^{(1)}\|_\infty |\beta|_\infty + c_{(5.16)}^3 \|\beta^{(2)}\|_\infty |\beta|_\infty^2 \right).$$

Proof. Observe that the matrix $\partial_\epsilon \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\epsilon \right) \right)$ has a norm bounded by $c \|\beta^{(2)}\|_\infty \sup_{T \geq 0} \left| \tilde{X}_{1,T}^\epsilon \right|$. Therefore, Proposition 5.5 combined with Lemma 5.3 yields

$$\sup_{T \geq 0} \left| \tilde{X}_{2,T}^\epsilon \right| \leq c_{(5.16)} \left(2 \|\beta^{(1)}\|_\infty c_{(5.16)} |\beta|_\infty + c \|\beta^{(2)}\|_\infty (c_{(5.16)} |\beta|_\infty)^2 \right).$$

□

Additionally, using Lemma 5.1, we can easily prove the following.

Lemma 5.5. *Assume (H-i)-(H-iv). We have*

$$\left| \text{Tr} \left(e^{-A_0 r} \mathbb{E}_x \left[\beta^{(1)} \left(\tilde{X}_r \right) \right] e^{-A r} \right) \right| \leq d c_A^2 e^{-\lambda_{0,A} r} \|\beta^{(1)}\|_\infty.$$

Proof of Lemma 5.2. We prove a slightly more general statement regarding $\tilde{X}_T^\epsilon$ for any given $\epsilon \in [0, 1]$. Denote by μ^{X^ϵ} , its stationary distribution, which exists owing to Proposition 5.4. If $\tilde{X}_T^\epsilon$ follows the dynamics of Equation (5.19), then we have

$$\begin{aligned} \lim_{T \rightarrow +\infty} \mathbb{E}_{\mu^X} \left[h \left(\tilde{X}_T^\epsilon \right) \right] &= \lim_{T \rightarrow +\infty} \int_{\mathbb{R}^d} \mathbb{E}_x \left[h \left(\tilde{X}_T^\epsilon \right) \right] \mu^X(x) dx \\ &= \int_{\mathbb{R}^d} \mu^X(x) dx \int_{\mathbb{R}^d} h(y) \mu^{X^\epsilon}(y) dy \end{aligned}$$

(by the dominated convergence theorem and Proposition 5.4)

$$= \int_{\mathbb{R}^d} h(y) \mu^{X^\epsilon}(y) dy.$$

Substituting $\epsilon = 1$ and $\epsilon = 0$, we obtain the convergences (5.17) and (5.18) respectively. □

Step 2) Derivation of the first correction term. We start by a crucial lemma which strongly relies on the time-reversibility of X .

Lemma 5.6. *Let $C_1(T)$ be the first correction term of the finite-time expansion (5.21). Then*

$$C_1(T) = \int_{\mathbb{R}^d} h^{(1)}(x) \mathbb{E}_x \left[e^{-A_0 \tau_1} \beta \left(\tilde{X}_{\tau_1} \right) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right] \mu^X(x) dx, \quad (5.34)$$

where $\tau_1 \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ is independent of the Brownian motion W .

Proof. We will first derive the explicit form of $\tilde{X}_{1,T}$. Starting from (5.20) for the equation of $\tilde{X}_{1,\cdot}$, then applying Proposition 5.5 and finally

taking advantage of the Hurwitz property of A (e.g. **(H-i)**), we get

$$\tilde{X}_{1,T} = \int_0^T e^{-A(T-s)} \beta(\tilde{X}_s) ds = \int_0^T e^{-A_0 r} \beta(\tilde{X}_{T-r}) \frac{1}{\lambda_{0,A}} \lambda_{0,A} e^{-\lambda_{0,A} r} dr. \quad (5.35)$$

By the definition of $C_1(T)$, we have

$$\begin{aligned} C_1(T) &= \mathbb{E}_{\mu^X} \left[h^{(1)}(\tilde{X}_T) e^{-A_0 \tau_1} \beta(\tilde{X}_{T-\tau_1}) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right] \\ &= \mathbb{E}_{\mu^X} \left[h^{(1)}(\tilde{X}_0) e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right], \end{aligned} \quad (5.36)$$

where the last equality comes from the time-reversibility of $\tilde{X}$ (valid under assumption **(H-i)**-**(H-ii)**, see Proposition 5.3), and the independence between τ_1 and the Brownian motion W (and thus between τ_1 and $\tilde{X}$). $\square$

Completion of the derivation of the first order correction term. If $\varphi(x) : \mathbb{R}^d \rightarrow \mathbb{R}^d$, then recall that $\operatorname{div}(\varphi(x)) = \operatorname{Tr}(D\varphi(x))$. Using (5.13), the identity

$$\frac{D\mu^X(x)}{\mu^X(x)} = -2x^\top (\Sigma \Sigma^\top)^{-1} A \quad (5.37)$$

and (5.34), we obtain after an integration by parts

$$\begin{aligned} C_1(T) &= \int_{\mathbb{R}^d} -h(x) \operatorname{div} \left(\mathbb{E}_x \left[e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right] \mu^X(x) \right) dx \\ &= \int_{\mathbb{R}^d} -h(x) \left((D\mu^X(x)) \mathbb{E}_x \left[e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right] \right. \\ &\quad \left. + \mathbb{E}_x \left[\operatorname{div} \left(e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \right) \frac{\mathbb{1}_{\tau_1 < T}}{\lambda_{0,A}} \right] \mu^X(x) \right) dx \\ &\stackrel{(5.37)}{=} \int_{\mathbb{R}^d} h(x) \mathbb{E}_x \left[\frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1})}{\lambda_{0,A}} \mathbb{1}_{\tau_1 < T} \right. \\ &\quad \left. - \frac{\operatorname{Tr} \left(e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right)}{\lambda_{0,A}} \mathbb{1}_{\tau_1 < T} \right] \mu^X(x) dx \\ &:= \int_{\mathbb{R}^d} h(x) c_1(T, x) \mu^X(x) dx. \end{aligned} \quad (5.38)$$

By Lemmas 5.1 and 5.5, we see that $c_1(T, x)$ is bounded uniformly in T . We can easily apply Lebesgue dominated convergence theorem to show that

$$\begin{aligned} c_1(x) &:= \lim_{T \rightarrow +\infty} c_1(T, x) \\ &= \mathbb{E}_x \left[\frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1})}{\lambda_{0,A}} - \frac{\text{Tr} \left(e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_1} \right)}{\lambda_{0,A}} \right] \end{aligned}$$

which is exactly the announced formula for (5.24).

Step 3) Bounds on $c_1(T, x)$ and $c_1(x)$. From Equation (5.38), the above formula for c_1 and from Lemma 5.5, we have

$$\begin{aligned} &\sup_{T \geq 0} |c_1(T, x)| \vee |c_1(x)| \\ &\leq \int_0^{+\infty} \left(2c_A |x| \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| e^{-\lambda_{0,A} r} |\beta|_\infty + d c_A^2 |\beta^{(1)}|_\infty e^{-2\lambda_{0,A} r} \right) dr \\ &= \frac{2c_A}{\lambda_{0,A}} |x| \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty + \frac{d c_A^2 |\beta^{(1)}|_\infty}{2\lambda_{0,A}}. \end{aligned}$$

The bounds on $c_1(T, x)$ and $c_1(x)$ are now proved. They are sufficient to pass to the limit in (5.38) and obtain

$$\lim_{T \rightarrow +\infty} C_1(T) = \int_{\mathbb{R}^d} h(x) c_1(x) \mu^X(x) dx. \quad (5.39)$$

Step 4) Error estimates. In view of the representation (5.22) of the error at fixed time, we easily deduce, in view of Lemmas 5.3 and 5.4,

$$\begin{aligned} \sup_{T \geq 0} |\text{Error}_1(T)| &\leq c \left((c_{(5.16)} |\beta|_\infty)^2 \|h^{(2)}\|_\infty \right. \\ &\quad \left. + (c_{(5.16)}^2 \|\beta^{(1)}\|_\infty |\beta|_\infty + c_{(5.16)}^3 \|\beta^{(2)}\|_\infty |\beta|_\infty^2) \|h^{(1)}\|_\infty \right). \end{aligned}$$

Step 5) Conclusion. Take the limit of Equation (5.21) as $T \rightarrow +\infty$, in combination with (5.39) and Lemma 5.2: it shows that $\text{Error}_1(T)$ has a

limit as $T \rightarrow +\infty$, which we denote by Error_1 , and this limit is bounded by $\sup_{T \geq 0} \text{Error}_1(T)$. The proof of Theorem 5.1 is now complete. $\square$

5.3.4 Proof of Theorem 5.2

As before, the proof is split into multiple steps.

Step 1) Preliminary results. The results presented in this step are necessary for our analysis. First observe that under **(H-iv)**, $\beta^{(2)} = D \text{vec}(\beta^{(1)})$ and $\beta^{(3)} = D \text{vec}(\beta^{(2)})$ are well defined and bounded.

Proposition 5.6. *The following derivatives are computed w.r.t. x , the starting point of the OU process $\tilde{X}$:*

$$\begin{aligned} & D \left(e^{-As} \beta^{(1)} \left(\tilde{X}_s \right) e^{-At} \beta \left(\tilde{X}_{s+t} \right) \right) \\ &= \left(\left(\beta^\top \left(\tilde{X}_{s+t} \right) e^{-A^\top t} \right) \otimes e^{-As} \right) \beta^{(2)} \left(\tilde{X}_s \right) e^{-As} \\ &+ e^{-As} \beta^{(1)} \left(\tilde{X}_s \right) e^{-At} \beta^{(1)} \left(\tilde{X}_{s+t} \right) e^{-A(s+t)}, \end{aligned} \quad (5.40)$$

$$\begin{aligned} & D \left(2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-Aq} \beta \left(\tilde{X}_q \right) \right) \\ &= 2\beta^\top \left(\tilde{X}_q \right) e^{-A^\top q} A^\top (\Sigma \Sigma^\top)^{-1} \\ &+ 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-Aq} \beta^{(1)} \left(\tilde{X}_q \right) e^{-Aq}, \end{aligned} \quad (5.41)$$

and

$$D \text{Tr} \left(\left(e^{-2A^\top q} \right)^\top \beta^{(1)} \left(\tilde{X}_q \right) \right) = \left(\text{vec} \left(e^{-2A^\top q} \right) \right)^\top \beta^{(2)} \left(\tilde{X}_q \right) e^{-Aq}. \quad (5.42)$$

Proof. Note that $\partial_x \tilde{X}_{\tau_1} = e^{-A\tau_1}$ under the $\mathbb{P}_x$ -measure, see (5.5). To prove Equation (5.40), we just need to recall from Appendix A, Proposition A.3, the fact that $\text{vec} (M_1 M_2 M_3) = (M_3^\top \otimes M_1) \text{vec} (M_2)$. We prove Equation (5.41) by a direct matrix differentiation. Lastly, to prove Equation (5.42), we recall from Appendix A, Proposition A.4 that $\text{Tr} (M_1^\top M_2) = (\text{vec} (M_1))^\top \text{vec} (M_2)$. $\square$

Proposition 5.7. *The following equality holds:*

$$\int_0^{+\infty} \int_0^v e^{-m\lambda_{0,A}r} e^{-n\lambda_{0,A}v} dr dv = \frac{1}{n(m+n)\lambda_{0,A}^2}.$$

Proof. We have

$$\begin{aligned} \int_0^{+\infty} \int_0^v e^{-m\lambda_{0,A}r} e^{-n\lambda_{0,A}v} dr dv &= \int_0^{+\infty} \frac{e^{-n\lambda_{0,A}v} - e^{-(m+n)\lambda_{0,A}v}}{m\lambda_{0,A}} dv \\ &= \frac{1}{m\lambda_{0,A}} \left(\frac{1}{n\lambda_{0,A}} - \frac{1}{(m+n)\lambda_{0,A}} \right) \\ &= \frac{1}{n(m+n)\lambda_{0,A}^2}. \end{aligned}$$

□

Lemma 5.7. *Assume (H-i)-(H-iv)-(H-vi). We have, for any $\varepsilon \in [0, 1]$,*

$$\begin{aligned} \sup_{T \geq 0} |\tilde{X}_{3,T}^\varepsilon| &\leq c c_{(5.16)}^3 |\beta|_\infty \left((\|\beta^{(1)}\|_\infty + c_{(5.16)} \|\beta^{(2)}\|_\infty |\beta|_\infty)^2 \right. \\ &\quad \left. + |\beta|_\infty (\|\beta^{(2)}\|_\infty + c_{(5.16)} \|\beta^{(3)}\|_\infty |\beta|_\infty) \right). \end{aligned}$$

Proof. From Equation (5.33), we obtain

$$\begin{aligned} d\tilde{X}_{3,t}^\varepsilon &= \left(-A\tilde{X}_{3,t}^\varepsilon + \varepsilon\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \tilde{X}_{3,t}^\varepsilon + 3\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \tilde{X}_{2,t}^\varepsilon \right. \\ &\quad \left. + 3\partial_\varepsilon \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \right) \tilde{X}_{1,t}^\varepsilon + 2\varepsilon\partial_\varepsilon \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \right) \tilde{X}_{2,t}^\varepsilon \right. \\ &\quad \left. + \varepsilon\partial_\varepsilon^2 \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \right) \tilde{X}_{1,t}^\varepsilon \right) dt. \end{aligned}$$

Therefore, from Proposition 5.5, Lemma 5.3 and 5.4, the norm of $\partial_\varepsilon \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \right)$ is bounded by $c \|\beta^{(2)}\|_\infty \sup_{T \geq 0} |\tilde{X}_{1,T}^\varepsilon|$. Moreover, we observe that the norm of $\partial_\varepsilon^2 \left(\beta^{(1)} \left(\tilde{X}_{0,t}^\varepsilon \right) \right)$ can be easily proved to be bounded by $c \left(\|\beta^{(3)}\|_\infty \left(\sup_{T \geq 0} |\tilde{X}_{1,T}^\varepsilon| \right)^2 + \|\beta^{(2)}\|_\infty \sup_{T \geq 0} |\tilde{X}_{2,T}^\varepsilon| \right)$. Therefore, we obtain easily the announced result by factorization. □

Step 2) Derivation of the second correction term. The decomposition of $\mathbb{E}_{\mu^X} \left[h \left(\tilde{Y}_T \right) \right]$ to the second order correction term is

$$\begin{aligned} \mathbb{E}_{\mu^X} \left[h \left(\tilde{Y}_T \right) \right] &= \mathbb{E}_{\mu^X} \left[h \left(\tilde{X}_{0,T} \right) \right] + C_1(T) + C_{21}(T) + C_{22}(T) \\ &\quad + \text{Error}_2(T) \end{aligned} \quad (5.43)$$

where

$$\begin{aligned} C_{21}(T) &:= \frac{1}{2} \mathbb{E}_{\mu^X} \left[h^{(1)} \left(\tilde{X}_{0,T} \right) \tilde{X}_{2,T} \right], \\ C_{22}(T) &:= \frac{1}{2} \mathbb{E}_{\mu^X} \left[\left(\tilde{X}_{1,T} \right)^\top h^{(2)} \left(\tilde{X}_{0,T} \right) \tilde{X}_{1,T} \right] \end{aligned}$$

and $\text{Error}_2(T)$ as shown in Equation (5.48).

Now, we introduce 2 important lemmas which rely on the time-reversibility of X .

Lemma 5.8. $C_{21}(T)$ is the first part of the second correction term of our finite expression (5.43) and it can be represented as

$$\begin{aligned} C_{21}(T) &= \int_{\mathbb{R}^d} h^{(1)}(x) \\ &\quad \mathbb{E}_x \left[e^{-A_0 \tau_1} \beta^{(1)} \left(\tilde{X}_{\tau_1} \right) e^{-A_0 \tau_2} \beta \left(\tilde{X}_{\tau_1 + \tau_2} \right) \frac{\mathbb{1}_{\tau_1 + \tau_2 < T}}{\lambda_{0,A}^2} \right] \mu^X(x) dx, \end{aligned}$$

with τ_1 and τ_2 as in Theorems 5.1 and 5.2.

Proof. We take $\epsilon = 0$ for Equation (5.33) before using Proposition 5.5 and invoking Equation (5.35), we have

$$\begin{aligned} X_{2,T} &= \int_0^T \int_0^q 2e^{-A(T-q)} \beta^{(1)} \left(\tilde{X}_q \right) e^{-Au} \beta \left(\tilde{X}_{q-u} \right) dudq \\ &\stackrel{(r=T-q)}{=} \int_0^T \int_0^{T-r} 2e^{-A_0 r} \beta^{(1)} \left(\tilde{X}_{T-r} \right) e^{-A_0 u} \beta \left(\tilde{X}_{T-r-u} \right) \\ &\quad \frac{\lambda_{0,A} e^{-\lambda_{0,A} r} \lambda_{0,A} e^{-\lambda_{0,A} u}}{\lambda_{0,A}^2} dudr. \end{aligned}$$

From the time-reversibility of the OU process $\tilde{X}$ and the independent of τ_1 and τ_2 with Brownian Motion W and thus with $\tilde{X}$. We have

$$\begin{aligned}
 C_{21}(T) &= \mathbb{E}_{\mu^X} \left[\frac{h^{(1)}(\tilde{X}_T) e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{T-\tau_1}) e^{-A_0 \tau_2} \beta(\tilde{X}_{T-\tau_1-\tau_2})}{\lambda_{0,A}^2} \mathbb{1}_{\tau_1+\tau_2 < T} \right] \\
 &= \int_{\mathbb{R}^d} h^{(1)}(x) \\
 &\quad \mathbb{E}_x \left[\frac{e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_1+\tau_2})}{\lambda_{0,A}^2} \mathbb{1}_{\tau_1+\tau_2 < T} \right] \mu^X(x) dx.
 \end{aligned}$$

□

Lemma 5.9. $C_{22}(T)$ is the second part of the second order correction term of our finite-time expansion (5.43), it has the representation

$$\begin{aligned}
 C_{22}(T) &:= \int_{\mathbb{R}^d} \mathbb{E}_x \left[\frac{\beta^\top(\tilde{X}_{\tau_1}) e^{-A_0^\top \tau_1} h^{(2)}(x) e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_2})}{\lambda_{0,A}^2} \mathbb{1}_{\tau_1 < \tau_2 < T} \right] \mu^X(x) dx,
 \end{aligned}$$

with τ_1 and τ_2 as in Theorems 5.1 and 5.2.

Proof. We merely proceed as in the proof of Lemma 5.8. Start from the definition of $C_{22}(T)$, writing the quadratic term in $\tilde{X}_{1,T}$ as a double time-integral owing to (5.35), use the symmetry in the time-integral; it readily follows that

$$\begin{aligned}
 C_{22}(T) &= \mathbb{E}_{\mu^X} \left[\frac{\beta^\top(\tilde{X}_{T-\tau_1}) e^{-A_0^\top \tau_1} h^{(2)}(\tilde{X}_T) e^{-A_0 \tau_2} \beta(\tilde{X}_{T-\tau_2})}{\lambda_{0,A}^2} \mathbb{1}_{\tau_1 < \tau_2 < T} \right].
 \end{aligned}$$

We conclude by the time-reversibility property. □

Completion of the derivation of the second order correction term. From Lemma 5.8, an integration by parts and Proposition 5.6 with Equations

(5.37) and (5.4) gives

$$\begin{aligned}
& C_{21}(T) \\
&= \int_{\mathbb{R}^d} h(x) \mathbb{E}_x \left[\left(\frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_1 + \tau_2})}{\lambda_{0,A}^2} \right. \right. \\
&\quad - \text{Tr} \left(\frac{\left(\left(\beta^\top(\tilde{X}_{\tau_1 + \tau_2}) e^{-A_0^\top \tau_2} \right) \otimes e^{-A_0 \tau_1} \right) \beta^{(2)}(\tilde{X}_{\tau_1}) e^{-A \tau_1}}{\lambda_{0,A}^2} \right) \\
&\quad \left. \left. - \text{Tr} \left(\frac{e^{-A \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A \tau_2} \beta^{(1)}(\tilde{X}_{\tau_1 + \tau_2}) e^{-A_0(\tau_1 + \tau_2)}}{\lambda_{0,A}^2} \right) \right) \right] \\
&\quad \mathbb{1}_{\tau_1 + \tau_2 < T} \mu^X(x) dx \\
&:= \int_{\mathbb{R}^d} h(x) c_{21}(T, x) \mu^X(x) dx. \tag{5.44}
\end{aligned}$$

Next, recall that $\text{div}(uv^\top) = (Du)v + \text{div}(v)u$, if u and v are column vector functions. So, from Lemma 5.9, an integration by parts gives

$$\begin{aligned}
& C_{22}(T) \\
&= - \int_{\mathbb{R}^d} h^{(1)}(x) \\
&\quad \mathbb{E}_x \left[\text{div} \left(e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top \tau_2} \mu^X(x) \right) \frac{\mathbb{1}_{\tau_1 < \tau_2 < T}}{\lambda_{0,A}^2} \right] dx \\
&= - \int_{\mathbb{R}^d} h^{(1)}(x) \mathbb{E}_x \left[\left(e^{-A \tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0(\tau_1 + \tau_2)} \beta(\tilde{X}_{\tau_2}) \right. \right. \\
&\quad - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 \tau_2} \beta(\tilde{X}_{\tau_2}) e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \\
&\quad \left. \left. + \text{Tr} \left(e^{-A \tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A_0 \tau_2} e^{-A_0 \tau_1} \beta(\tilde{X}_{\tau_1}) \right) \right) \frac{\mathbb{1}_{\tau_1 < \tau_2 < T}}{\lambda_{0,A}^2} \right] \mu^X(x) dx.
\end{aligned}$$

Applying another integration by parts gives

$$\begin{aligned}
& C_{22}(T) \\
&= \int_{\mathbb{R}^d} h(x) \mathbb{E}_x \left[\left(-2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A\tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0(\tau_1+\tau_2)} \beta(\tilde{X}_{\tau_2}) \right. \right. \\
&\quad + \text{Tr} \left(\left(\left(\beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top(\tau_1+\tau_2)} \right) \otimes e^{-A\tau_1} \right) \beta^{(2)}(\tilde{X}_{\tau_1}) e^{-A\tau_1} \right) \\
&\quad + \text{Tr} \left(e^{-A\tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A_0(\tau_1+\tau_2)} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A\tau_2} \right) \\
&\quad + 4x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0\tau_2} \beta(\tilde{X}_{\tau_2}) x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0\tau_1} \beta(\tilde{X}_{\tau_1}) \\
&\quad - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A\tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A_0(\tau_1+\tau_2)} \beta(\tilde{X}_{\tau_1}) \\
&\quad - 2\beta^\top(\tilde{X}_{\tau_2}) e^{-A_0^\top\tau_2} A^\top (\Sigma \Sigma^\top)^{-1} e^{-A_0\tau_1} \beta(\tilde{X}_{\tau_1}) \\
&\quad - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0\tau_2} \beta(\tilde{X}_{\tau_2}) \text{Tr} \left(e^{-A_0\tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A\tau_1} \right) \\
&\quad - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0\tau_1} \beta(\tilde{X}_{\tau_1}) \text{Tr} \left(e^{-A_0\tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A\tau_2} \right) \\
&\quad + \left(\text{vec} \left(e^{-2A^\top\tau_2} \right) \right)^\top \beta^{(2)}(\tilde{X}_{\tau_2}) e^{-A_0(\tau_1+\tau_2)} \beta(\tilde{X}_{\tau_1}) \\
&\quad \left. + \text{Tr} \left(e^{-A_0\tau_2} \beta^{(1)}(\tilde{X}_{\tau_2}) e^{-A\tau_2} \right) \text{Tr} \left(e^{-A_0\tau_1} \beta^{(1)}(\tilde{X}_{\tau_1}) e^{-A\tau_1} \right) \right] \\
&\quad \frac{\mathbb{1}_{\tau_1 < \tau_2 < T}}{\lambda_{0,A}^2} \mu^X(x) dx \\
&:= \int_{\mathbb{R}^d} h(x) c_{22}(T, x) \mu^X(x) dx. \tag{5.45}
\end{aligned}$$

So, just like for $c_1(x)$, we can justify that $c_{21}(T, x) \rightarrow c_{21}(x)$ and $c_{22}(T, x) \rightarrow c_{22}(x)$, as $T \rightarrow +\infty$. Hence, we also have

$$\lim_{T \rightarrow +\infty} C_{21}(T) = \int_{\mathbb{R}^d} h(x) c_{21}(x) \mu^X(x) dx, \tag{5.46}$$

and

$$\lim_{T \rightarrow +\infty} C_{22}(T) = \int_{\mathbb{R}^d} h(x) c_{22}(x) \mu^X(x) dx. \tag{5.47}$$

The details are left for the readers.

Step 3) Bounds on $c_{21}(T, x)$ and $c_{21}(x)$. From (5.44) and Proposition 5.7, we have

$$\begin{aligned}
& \sup_{T \geq 0} |c_{21}(T, x)| \vee |c_{21}(x)| \\
& \leq \int_{[0, +\infty]^2} 2c_A^2 |x| \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| e^{-\lambda_{0,A}(r+v)} |\beta|_\infty \|\beta^{(1)}\|_\infty \\
& \quad + dc_A^3 e^{-\lambda_{0,A}(2r+v)} |\beta|_\infty \|\beta^{(2)}\|_\infty + dc_A^3 e^{-2\lambda_{0,A}(r+v)} \|\beta^{(1)}\|_\infty^2 dv dr \\
& = \frac{2c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty \|\beta^{(1)}\|_\infty |x|}{\lambda_{0,A}^2} + \frac{dc_A^3 |\beta|_\infty \|\beta^{(2)}\|_\infty}{2\lambda_{0,A}^2} \\
& \quad + \frac{dc_A^3 \|\beta^{(1)}\|_\infty^2}{4\lambda_{0,A}^2}.
\end{aligned}$$

The bounds on $c_{21}(T, x)$ and $c_{21}(x)$ are proved.

Step 4) Bounds on $c_{22}(T, x)$ and $c_{22}(x)$. From (5.45) and Proposition 5.7, we have

$$\begin{aligned}
& \sup_{T \geq 0} |c_{22}(T, x)| \vee |c_{22}(x)| \\
& \leq \int_0^{+\infty} \int_0^v c_A c_{\text{vec}}(-2A^\top) e^{-\lambda_{0,A}r - 3\lambda_{0,A}v} dr dv \|\beta^{(2)}\|_\infty |\beta|_\infty \\
& \quad + dc_A^3 (dc_A + 1) \int_0^{+\infty} \int_0^v e^{-2\lambda_{0,A}(r+v)} dr dv \|\beta^{(1)}\|_\infty^2 \\
& \quad + \int_0^{+\infty} \int_0^v dc_A^3 e^{-3\lambda_{0,A}r - \lambda_{0,A}v} dr dv \|\beta^{(2)}\|_\infty |\beta|_\infty \\
& \quad + 2c_A^2 |x| \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \|\beta^{(1)}\|_\infty |\beta|_\infty \\
& \quad \quad (dc_A + 1) \int_0^{+\infty} \int_0^v e^{-\lambda_{0,A}r - 2\lambda_{0,A}v} + e^{-2\lambda_{0,A}r - \lambda_{0,A}v} dr dv \\
& \quad + \left(4|x|^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| + 2 \right) \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty^2 \\
& \quad \quad \int_0^{+\infty} \int_0^v c_A^2 e^{-\lambda_{0,A}(r+v)} dr dv \\
& = \frac{c_A \|\beta^{(2)}\|_\infty |\beta|_\infty}{4\lambda_{0,A}^2} \left(dc_A^2 + \frac{c_{\text{vec}}(-2A^\top)}{3} \right) + \frac{dc_A^3 \|\beta^{(1)}\|_\infty^2}{8\lambda_{0,A}^2} (dc_A + 1) \\
& \quad + \frac{c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \|\beta^{(1)}\|_\infty |\beta|_\infty |x|}{\lambda_{0,A}^2} (1 + dc_A) \\
& \quad + \frac{c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty^2}{\lambda_{0,A}^2} \left(2|x|^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| + 1 \right).
\end{aligned}$$

The bound on $c_{22}(T, x)$ and $c_{22}(x)$ are proved.

Step 5) Error estimates. By a Taylor expansion of $\epsilon \mapsto \mathbb{E}_{\mu^x} [h(\tilde{X}_T^\epsilon)]$, we directly obtain that the error up to the second order correction term is given as

$$\begin{aligned} \text{Error}_2(T) = \int_0^1 \frac{(1-u)^2}{2} \mathbb{E}_{\mu^x} \left[\left(\left(\tilde{X}_{1,T}^u \right)^\top \otimes \left(\tilde{X}_{1,T}^u \right)^\top \right) h^{(3)} \left(\tilde{X}_T^u \right) \tilde{X}_{1,T}^u \right. \\ \left. + 3 \left(\tilde{X}_{2,T}^u \right)^\top h^{(2)} \left(\tilde{X}_{0,T}^u \right) \tilde{X}_{1,T}^u + h^{(1)} \left(\tilde{X}_{0,T}^u \right) \tilde{X}_{3,T}^u \right] du. \end{aligned} \quad (5.48)$$

Thus, in view of the above representation of the error at a fixed time T , we deduce easily the inequality (5.29) from Lemmas 5.3, 5.4 and 5.7.

Step 6) Conclusion. Taking the limit of Equation (5.43) as $T \rightarrow +\infty$, in combination with Equations (5.39), (5.46), (5.47) and Lemma 5.2, we show that $\text{Error}_2(T)$ has a limit as $T \rightarrow +\infty$ which is denoted by Error_2 and the limit is bounded by $\sup_{T \geq 0} \text{Error}_2(T)$. The proof of Theorem 5.2 is thus completed. $\square$

5.4 Application to simulation algorithm, numerical results

In the sequel, we assume the previous hypotheses of Theorems 5.1 and 5.2 are in force.

5.4.1 Generalized acceptance/rejection algorithm

We give a simulation scheme, in a form that suits well the decomposition (5.27) and (5.30). The aim is to sample according to a probability distribution μ^+ built from a slightly non-positive unit-mass measure μ ; moreover, we control the total variation distance between the two measures, in terms of the negative part of μ .

Theorem 5.3. Let μ be a measure on $\mathbb{R}^d$ with mass 1 but not necessarily positive, that can be written as:

$$\mu(\mathrm{d}x) = \mathbb{E} [\xi(Y^x, x)] \mu_0(\mathrm{d}x),$$

where $\mu_0(\mathrm{d}x)$ is a probability measure, the distribution of Y^x depends on x and there exists a function $H : \mathbb{R}^d \mapsto \mathbb{R}^+$ such that

$$|(\xi(y, x))^+| \leq H(x), \quad \forall x, y,$$

with $\int_{\mathbb{R}^d} H(x) \mu_0(\mathrm{d}x) < +\infty$.

Moreover, assume that $0 < \int_{\mathbb{R}^d} \mathbb{E} [(\xi(Y^x, x))^+] \mu_0(\mathrm{d}x) < +\infty$ and set

$$\mu^+(\mathrm{d}x) := \frac{\mathbb{E} [(\xi(Y^x, x))^+] \mu_0(\mathrm{d}x)}{\int_{\mathbb{R}^d} \mathbb{E} [(\xi(Y^x, x))^+] \mu_0(\mathrm{d}x)}.$$

Then Algorithm 1 produces a random variate with distribution μ^+ , with an acceptance rate as

$$p := \frac{\int_{\mathbb{R}^d} \mathbb{E} [(\xi(Y^x, x))^+] \mu_0(\mathrm{d}x)}{\int_{\mathbb{R}^d} H(z) \mu_0(\mathrm{d}z)}$$

and an error in total variation

$$|\mu - \mu^+|_{\mathrm{TV}} \leq 2 \int_{\mathbb{R}^d} \mathbb{E} [(\xi(Y^x, x))^-] \mu_0(\mathrm{d}x). \quad (5.49)$$

<pre> 1 repeat 2 Generate X with distribution $\frac{H(x)\mu_0(\mathrm{d}x)}{\int H(x)\mu_0(\mathrm{d}x)}$ (on $\mathbb{R}^d$) 3 Generate Y with distribution Y^x for $x = X$ 4 Generate a uniform $[0, 1]$ random variate U 5 until $UH(X) \leq (\xi(Y, X))^+$; 6 Return X </pre>

Algorithm 1: Generalized acceptance-rejection method

Proof. Denoting by $g_x(dy)$, the distribution of Y^x for any Borel set $\mathcal{B}$, we have

$$\begin{aligned} \mathbb{P}(X(\text{returned}) \in \mathcal{B}) &= \frac{1}{p} \mathbb{P}(X \in \mathcal{B}, UH(X) \leq (\xi(Y, X))^+) \\ &= \frac{1}{p} \int_{\mathcal{B}} \int \frac{(\xi(y, z))^+}{H(z)} \frac{H(z)\mu_0(dz)}{\int H(x)\mu_0(dx)} g_z(dy) \\ &= \frac{1}{p \int H(x)\mu_0(dx)} \int_{\mathcal{B}} \mathbb{E}[(\xi(Y^z, z))^+] \mu_0(dz) \\ &= \frac{\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^z, z))^+] \mu_0(dx)}{p \int H(x)\mu_0(dx)} \int_{\mathcal{B}} \mu^+(dz). \end{aligned}$$

Therefore, by setting $\mathcal{B} = \mathbb{R}^d$ it is clear that the acceptance probability is p , as advertised, and that the variate produced by the algorithm has distribution μ^+ . When $\xi(Y^x, x)$ is not non-negative, μ and μ^+ may be different. Then the total variation distance between μ^+ and μ is

$$\begin{aligned} &\sup_{h: |h|_{\infty} \leq 1} \left| \int_{\mathbb{R}^d} h(x) \mu(dx) - \int_{\mathbb{R}^d} h(x) \mu^+(dx) \right| \\ &= \sup_{h: |h|_{\infty} \leq 1} \left| \int_{\mathbb{R}^d} h(x) \left(\frac{\mathbb{E}[\xi(Y^x, x)] \int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx)}{\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx)} \right. \right. \\ &\quad \left. \left. - \frac{\mathbb{E}[(\xi(Y^x, x))^+]}{\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx)} \right) \mu_0(dx) \right| \end{aligned}$$

(since $\xi(Y^x, x) = (\xi(Y^x, x))^+ - (\xi(Y^x, x))^-$)

$$\begin{aligned} &= \sup_{h: |h|_{\infty} \leq 1} \left| \int_{\mathbb{R}^d} h(x) \left(\frac{\mathbb{E}[(\xi(Y^x, x))^+] (\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx) - 1)}{\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx)} \right. \right. \\ &\quad \left. \left. - \mathbb{E}[(\xi(Y^x, x))^-] \mu_0(dx) \right) \right| \end{aligned}$$

(furthermore we have $1 = \int_{\mathbb{R}^d} \mathbb{E}[\xi(Y^x, x)] \mu_0(dx)$)

$$\begin{aligned} &= \sup_{h: |h|_{\infty} \leq 1} \left| \int_{\mathbb{R}^d} h(x) \left(\frac{\mathbb{E}[(\xi(Y^x, x))^+] (\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^-] \mu_0(dx))}{\int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^+] \mu_0(dx)} \right. \right. \\ &\quad \left. \left. - \mathbb{E}[(\xi(Y^x, x))^-] \mu_0(dx) \right) \right| \\ &\leq 2 \int_{\mathbb{R}^d} \mathbb{E}[(\xi(Y^x, x))^-] \mu_0(dx). \end{aligned}$$

□

5.4.2 Simulation algorithm for 1st and 2nd order approximation

The algorithm for the multidimensional case using only first and second order correction term is presented here.

First order simulation scheme. We have from Theorem 5.1,

$$\mu^{Y,1}(\mathrm{d}x) = \mathbb{E}_x \left[\xi_1 \left(\tau_1, \tilde{X}_{\tau_1}, x \right) \right] \mu^X(\mathrm{d}x),$$

with

$$\xi_1(r, y, x) = 1 + \frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 r} \beta(y)}{\lambda_{0,A}} - \frac{\mathrm{Tr} (e^{-A_0 r} \beta^{(1)}(y) e^{-Ar})}{\lambda_{0,A}}.$$

Furthermore, if we let

$$p_{11} = 1 + \frac{dc_A^2 |\beta^{(1)}|_\infty}{\lambda_{0,A}}$$

$$p_{12} = 2^{1+0.5d} c_A \left| x \exp \left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right) \right|_\infty \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \frac{|\beta|_\infty}{\lambda_{0,A}},$$

then

$$\tilde{H}_1(x) = p_{11} + p_{12} \left(\frac{1}{2} \right)^{\frac{d}{2}} \exp \left(\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right)$$

and

$$\frac{\tilde{H}_1(x) \mu^X(\mathrm{d}x)}{\int \tilde{H}_1(x) \mu^X(\mathrm{d}x)} \stackrel{d}{=} \frac{p_{11}}{p_{11} + p_{12}} \mu^X(\mathrm{d}x) + \frac{p_{12}}{p_{11} + p_{12}} \mathcal{N}(0, A^{-1} \Sigma \Sigma^\top).$$

In this form, we observe that the distribution to simulate in Algorithm 1-Line 2 is simply a mixture of two Gaussian distributions. As a consequence, Algorithm 2 produces a random variate with distribution

$$(\mu^{Y,1})^+(\mathrm{d}x) = \frac{\mathbb{E}[(\xi_1(\tau_1, \tilde{X}_{\tau_1}, x))^+] \mu^X(\mathrm{d}x)}{\int_{\mathbb{R}^d} \mathbb{E}[(\xi_1(\tau_1, \tilde{X}_{\tau_1}, x))^+] \mu^X(\mathrm{d}x)} \text{ with an acceptance rate } p = \frac{\int_{\mathbb{R}^d} \mathbb{E}[(\xi_1(\tau_1, \tilde{X}_{\tau_1}, x))^+] \mu^X(\mathrm{d}x)}{p_{11} + p_{12}}.$$

Recall that when $\beta \rightarrow 0$, ξ_1 converges to 1: thus, for small β , the acceptance rate p is close to 1. Starting from (5.49) and invoking Gaussian tail estimates, we can derive a sharp estimate on the error in TV between $\mu^{Y,1}$ and $(\mu^{Y,1})^+$.

So, if $\|\beta^{(1)}\|_\infty < \frac{\lambda_{0,A}}{dc_A^2}$, then by a standard Gaussian computation, we obtain

$$\begin{aligned}
 & \left| \mu^{Y,1} - (\mu^{Y,1})^+ \right|_{\text{TV}} \\
 & \leq 2 \int_{\mathbb{R}^d} \mathbb{E} \left[\left(\xi_1 \left(\tau_1, \tilde{X}_{\tau_1}, x \right) \right)^- \right] \mu^X(dx) \\
 & \leq 2(p_{11} + p_{12} - 1) \int_{|x| \geq m_{1,1}} 2^{-0.5d} \exp \left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} Ax \right) \mu^X(dx) \\
 & \leq m_{1,2} e^{-\frac{m_{1,3}}{|\beta|_\infty^2}},
 \end{aligned}$$

where $m_{1,1} = \frac{\lambda_{0,A} - dc_A^2 \|\beta^{(1)}\|_\infty}{2c_A \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| |\beta|_\infty}$ and $m_{1,2}, m_{1,3} > 0$, are some constants. Thus, we prove the following proposition:

Proposition 5.8. *If $\|\beta^{(1)}\|_\infty < \frac{\lambda_{0,A}}{dc_A^2}$, then the total variation is controlled as*

$$\left| \mu^{Y,1} - (\mu^{Y,1})^+ \right|_{\text{TV}} \leq m_{1,2} e^{-\frac{m_{1,3}}{|\beta|_\infty^2}},$$

where the positive constants $m_{1,2}$ and $m_{1,3}$ depend on the model parameters and are locally uniform when $|\beta|_\infty$ and $\|\beta^{(1)}\|_\infty$ tend to 0.

This shows that $\mu^{Y,1}$ and $(\mu^{Y,1})^+$ are exponentially close to each other as $|\beta|_\infty$ and $|\beta^{(1)}|_\infty$ are small.

```

1 repeat
2   Generate  $K \stackrel{d}{=} \text{Bern}\left(\frac{p_{11}}{p_{11}+p_{12}}\right)$ ; /* sampling of the
   Gaussian mixture */
3   if  $K = 1$  then
4     Generate  $X \stackrel{d}{=} \mu^X$ 
5   else
6     Generate  $X \stackrel{d}{=} \mathcal{N}(0, A^{-1}\Sigma\Sigma^\top)$ 
7   Generate  $R \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ ; /* sampling of  $\tau_1$  */
8   Generate  $Y \stackrel{d}{=} \mathcal{N}\left(e^{-AR}X, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top R})\right)$ ;
   /* sampling of  $\tilde{X}_{\tau_1}$  using (5.8) */
9   Generate a uniform  $[0, 1]$  random variate  $U$ 
10 until  $U\tilde{H}_1(X) \leq (\xi_1(R, Y, X))^+$ ;
11 Return  $X$ 

```

Algorithm 2: sampling of μ^Y at the first-order accuracy

Second order simulation scheme. From Theorem 5.2, we have

$$\mu^{Y,2}(\mathrm{d}x) = \mathbb{E}_x \left[\xi_2 \left(\tau_1, \tau_2, \tilde{X}_{\tau_1}, \tilde{X}_{\tau_2}, \tilde{X}_{\tau_1+\tau_2}, x \right) \right] \mu^X(\mathrm{d}x).$$

with

$$\begin{aligned} & \xi_2(r, v, y, w, z, x) \\ &= 1 + \frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 r} \beta(y)}{\lambda_{0,A}} - \frac{\mathrm{Tr}(e^{-A_0 r} \beta^{(1)}(y) e^{-Ar})}{\lambda_{0,A}} \\ &+ \frac{2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 r} \beta^{(1)}(y) e^{-A_0 v} \beta(z)}{\lambda_{0,A}^2} \\ &- \frac{1}{\lambda_{0,A}^2} \mathrm{Tr} \left(\left(\left(\beta^\top(z) e^{-A_0^\top v} \right) \otimes e^{-A_0 r} \right) \beta^{(2)}(y) e^{-Ar} + e^{-Ar} \beta^{(1)}(y) e^{-Av} \right. \\ &\quad \left. \beta^{(1)}(z) e^{-A_0(r+v)} \right) \\ &+ \frac{\mathbb{1}_{r < v}}{\lambda_{0,A}^2} \left(-2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-Ar} \beta^{(1)}(y) e^{-A_0(r+v)} \beta(w) \right. \\ &\quad \left. + \mathrm{Tr} \left(\left(\left(\beta^\top(w) e^{-A_0^\top(r+v)} \right) \otimes e^{-Ar} \right) \beta^{(2)}(y) e^{-Ar} \right. \right. \\ &\quad \left. \left. + e^{-Ar} \beta^{(1)}(y) e^{-A_0(r+v)} \beta^{(1)}(w) e^{-Av} \right) \right. \\ &\quad \left. + 4\beta^\top(w) e^{-A_0^\top v} A^\top (\Sigma \Sigma^\top)^{-1} x x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 r} \beta(y) \right. \\ &\quad \left. - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-Av} \beta^{(1)}(w) e^{-A_0(r+v)} \beta(y) \right. \\ &\quad \left. - 2\beta^\top(w) e^{-A_0^\top v} A^\top (\Sigma \Sigma^\top)^{-1} e^{-A_0 r} \beta(y) \right. \\ &\quad \left. - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 v} \beta(w) \mathrm{Tr}(e^{-A_0 r} \beta^{(1)}(y) e^{-Ar}) \right. \\ &\quad \left. - 2x^\top (\Sigma \Sigma^\top)^{-1} A e^{-A_0 r} \beta(y) \mathrm{Tr}(e^{-A_0 v} \beta^{(1)}(w) e^{-Av}) \right. \\ &\quad \left. + \left(\mathrm{vec}(e^{-2A^\top v}) \right)^\top \beta^{(2)}(w) e^{-A_0(r+v)} \beta(y) \right. \\ &\quad \left. + \mathrm{Tr}(e^{-A_0 v} \beta^{(1)}(w) e^{-Av}) \mathrm{Tr}(e^{-A_0 r} \beta^{(1)}(y) e^{-Ar}) \right). \end{aligned}$$

Furthermore, if we let

$$\begin{aligned} & p_{21} \\ &= 1 + \frac{dc_A^2 \|\beta^{(1)}\|_\infty}{\lambda_{0,A}} + \frac{dc_A^3 (2 + dc_A) \|\beta^{(1)}\|_\infty^2}{\lambda_{0,A}^2} \\ &\quad + \frac{c_A (2dc_A^2 + c_{\mathrm{vec}(-2A^\top)}) \|\beta\|_\infty \|\beta^{(2)}\|_\infty}{\lambda_{0,A}^2} + \frac{2c_A^2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \|\beta\|_\infty^2}{\lambda_{0,A}^2} \end{aligned}$$

and

$$\begin{aligned}
& p_{22} \\
&= 2^{1+0.5d} \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \left| x \exp \left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right) \right|_\infty \\
&\quad \left(\frac{c_A \|\beta\|_\infty}{\lambda_{0,A}} + \frac{3c_A^2 \|\beta\|_\infty \|\beta^{(1)}\|_\infty}{\lambda_{0,A}^2} + \frac{2dc_A^3 \|\beta\|_\infty \|\beta^{(1)}\|_\infty}{\lambda_{0,A}^2} \right) + 2^{2+0.5d} c_A^2 \\
&\quad \frac{\left\| x x^\top \exp \left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right) \right\|_\infty \left\| (\Sigma \Sigma^\top)^{-1} \right\|^2 \|A\|^2 \|\beta\|_\infty^2}{\lambda_{0,A}^2}
\end{aligned}$$

then

$$\tilde{H}_2(x) = p_{21} + p_{22} \left(\frac{1}{2} \right)^{\frac{d}{2}} \exp \left(\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right)$$

and

$$\frac{\tilde{H}_2(x) \mu^X(dx)}{\int \tilde{H}_2(x) \mu^X(dx)} \stackrel{d}{=} \frac{p_{21}}{p_{21} + p_{22}} \mu^X(dx) + \frac{p_{22}}{p_{21} + p_{22}} \mathcal{N}(0, A^{-1} \Sigma \Sigma^\top).$$

Similar to the first-order simulation scheme, Algorithm 3 produces a random variate with distribution $(\mu^{Y,2})^+(dx)$ (defined in Theorem 5.3) with an acceptance rate of close to 1 as β is small and the distance in TV between $(\mu^{Y,2})^+$ and $\mu^{Y,2}$ is exponentially small as $\beta \rightarrow 0$.

Hence, if $p_{21} \leq 2$, which holds true when $\|\beta\|_\infty$, $\|\beta^{(1)}\|_\infty$ and $\|\beta^{(2)}\|_\infty$ are small enough. Then, we have

$$\begin{aligned}
& \left| \mu^{Y,2} - (\mu^{Y,2})^+ \right|_{\text{TV}} \\
&\leq 2 \int_{\mathbb{R}^d} \mathbb{E}_x \left[\left(\xi_2 \left(\tau_1, \tau_2, \tilde{X}_{\tau_1}, \tilde{X}_{\tau_2}, \tilde{X}_{\tau_1+\tau_2}, x \right) \right)^- \right] \mu^X(dx) \\
&\leq 2(p_{21} + p_{22} - 1) \int_{|x| \geq m_{2,1}} 2^{-0.5d} \exp \left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x \right) \mu^X(dx) \\
&\leq m_{2,4} e^{-\frac{m_{2,5}}{\|\beta\|_\infty^2}},
\end{aligned}$$

where

$$m_{2,1} = \frac{-m_{2,3} + \sqrt{m_{2,3}^2 + 4m_{2,2}(2 - p_{21})}}{2m_{2,2}},$$

$$m_{2,2} = \frac{4c_A^2 \left\| x x^\top \exp\left(-\frac{1}{2} x^\top (\Sigma \Sigma^\top)^{-1} A x\right) \right\|_\infty \left\| (\Sigma \Sigma^\top)^{-1} \right\|^2 \|A\|^2 |\beta|_\infty^2}{\lambda_{0,A}^2},$$

$$m_{2,3} = 2 \left\| (\Sigma \Sigma^\top)^{-1} \right\| \|A\| \left(\frac{c_A |\beta|_\infty}{\lambda_{0,A}} + \frac{3c_A^2 |\beta|_\infty \|\beta^{(1)}\|_\infty}{\lambda_{0,A}^2} + \frac{2dc_A^3 |\beta|_\infty \|\beta^{(1)}\|_\infty}{\lambda_{0,A}^2} \right)$$

and $m_{2,4}, m_{2,5} > 0$, are some constants. Hence we have

Proposition 5.9. *Suppose that $p_{21} \leq 2$, which holds when $|\beta|_\infty, \|\beta^{(1)}\|_\infty$ and $\|\beta^{(2)}\|_\infty$ are small enough. Then we have the total variation error to be*

$$\left| \mu^{Y,2} - (\mu^{Y,2})^+ \right|_{\text{TV}} \leq m_{2,4} e^{-\frac{m_{2,5}}{|\beta|_\infty^2}}.$$

The positive constants $m_{2,4}$ and $m_{2,5}$ depend on the model parameters and are locally uniform when $|\beta|_\infty, \|\beta^{(1)}\|_\infty$ and $\|\beta^{(2)}\|_\infty$ tend to 0.

We notice again that as $|\beta|_\infty, \|\beta^{(1)}\|_\infty$ and $\|\beta^{(2)}\|_\infty$ tend to 0, we have exponential convergence of $\left| \mu^{Y,2} - (\mu^{Y,2})^+ \right|_{\text{TV}} \rightarrow 0$. Hence, our approximation of $\mu^{Y,2}$ with $(\mu^{Y,2})^+$ is very accurate for small $|\beta|_\infty$,

$\|\beta^{(1)}\|_\infty$ and $\|\beta^{(2)}\|_\infty$.

```

1 repeat
2   Generate  $K \stackrel{d}{=} \text{Bern}\left(\frac{p_{21}}{p_{21}+p_{22}}\right)$ ; /* sampling of the
   Gaussian mixture */
3   if  $K = 1$  then
4     Generate  $X \stackrel{d}{=} \mu^X$ 
5   else
6     Generate  $X \stackrel{d}{=} \mathcal{N}(0, A^{-1}\Sigma\Sigma^\top)$ 
7   Generate  $R \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ ; /* sampling of  $\tau_1$  */
8   Generate  $V \stackrel{d}{=} \text{Exp}(\lambda_{0,A})$ ; /* sampling of  $\tau_2$  */
9   if  $R \leq V$  then
10    Generate  $Y \stackrel{d}{=} \mathcal{N}\left(e^{-AR}X, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top R})\right)$ 
11    Generate
12     $W \stackrel{d}{=} \mathcal{N}\left(e^{-A(V-R)}Y, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top(V-R)})\right)$ 
13    Generate  $Z \stackrel{d}{=} \mathcal{N}\left(e^{-AR}W, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top R})\right)$ 
14  else
15    Generate  $W \stackrel{d}{=} \mathcal{N}\left(e^{-AV}X, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top V})\right)$ 
16    Generate
17     $Y \stackrel{d}{=} \mathcal{N}\left(e^{-A(R-V)}W, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top(R-V)})\right)$ 
18    Generate  $Z \stackrel{d}{=} \mathcal{N}\left(e^{-AV}Y, \frac{A^{-1}\Sigma\Sigma^\top}{2}(\text{Id} - e^{-2A^\top V})\right)$ 
19  Generate a uniform  $[0, 1]$  random variate  $U$ 
20 until  $U\tilde{H}_2(X) \leq (\xi_2(R, V, Y, W, Z, X))^+$ ;
21 Return  $X$ 

```

Algorithm 3: sampling of μ^Y at the second-order accuracy

5.4.3 Numerical experiments

Two-dimensional example. In this subsection, we first conduct numerical study for $d = 2$, $\Sigma = \text{Id}$, with A and β given by

$$A = A^\top = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} \quad \text{and} \quad \beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} b \sin(x_1) \\ b \sin(x_2) \end{pmatrix},$$

where $a_{11}, a_{22}, b \in \mathbb{R}$. Using contour plots, we compare our approximate distribution generated with 4 millions sample points with the exact

distribution for $a_{11} = 7$, $a_{22} = 7.1$ and $b = 0.1$. The density of the exact distribution² of Y up to a constant can be solved to be $\text{Cst} \times \exp(-7x_1^2 - 7.1x_2^2 - 2b \cos(x_1) - 2b \cos(x_2))$.

FIGURE 5.1: Density plot for the stationary distribution of $d\tilde{Y}_t = \left(-A\tilde{Y}_t + \beta(\tilde{Y}_t)\right)dt + dW_t$, where $a_{11} = 7$, $a_{22} = 7.1$ and $b = 0.1$.

We see from Figure 5.1 that our approximation is accurate in approximating the exact distribution and that the second order approximation is better than the first order approximation. Furthermore, we also note that because of the Gaussian exponential decay, the errors are larger in the center of the distributions than in the tails.

One-dimensional example. Next, we conduct numerical studies for $d = 1$, $\Sigma = 1$, $A = 1.7$ and $\beta(x) = bxe^{-x^2}$, for $b = 1, 1.6, 1.7$ and 5 with 4 millions sample points. The density of the exact distribution for

²Indeed the drift is of the form $-\frac{1}{2}DV(x)$ so that the stationary density is $\text{Cst} \times e^{-V(x)}$.

this model is $Cst \times \exp(-1.7x^2 - be^{-x^2})$. Taking different values for b (thus changing $\|\beta^{(1)}\|_\infty$) for a fixed A serves to investigate the extent that the condition **(H-vi)** is important for the algorithm accuracy.

FIGURE 5.2: Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 1Y_t e^{-Y_t^2})dt + dW_t$.

In the case of $b = 1$, we notice again from Figure 5.2 that our approximation is very accurate and that the second order approximation is superior compare to the first order.

FIGURE 5.3: Density plot for the stationary distribution
of $dY_t = (-1.7Y_t + 1.6Y_t e^{-Y_t^2}) dt + dW_t$.

Here, we compare the results for $b = 1$ and $b = 1.6$ (respectively on Figures 5.2 and 5.3), to see that our approximation is more accurate when the ratio of $|\beta|_\infty$ against A is smaller. This is in agreement with our theoretical error estimates.

FIGURE 5.4: Density plot for the stationary distribution of $dY_t = (-1.7Y_t + 1.7Y_t e^{-Y_t^2}) dt + dW_t$.

FIGURE 5.5: Density plot for the stationary distribution
of $dY_t = (-1.7Y_t + 5Y_t e^{-Y_t^2}) dt + dW_t$.

In Figure 5.4, where $b = 1.7$, we see that although $\|\beta^{(1)}\|_\infty$ is not smaller than A , our approximation can still work. However, we see in Figure 5.5 ($b = 5$) that, when $\|\beta^{(1)}\|_\infty \gg A$, our approximation is not accurate anymore, as it only gives a rough approximation of the solution. In this case, both first and second order schemes are inaccurate; nevertheless, the second order approximation is seemingly worse, this reinforces the role of the assumptions in this analysis.

Appendix A

Vector Calculus

We are all aware that while vector calculus is well understood, matrix calculus remains a difficult problem. However, in [MN99], the authors prove that it is possible to treat the matrix calculus problem as a vector calculus problem. Therefore, in this appendix, we present the fundamental idea behind [MN99], which allows us to utilize the established tools of vector calculus to solve our problems in matrix calculus.

The central idea in [MN99] is to use a vectorizing operator to convert a matrix to a vector and instead of working directly with the derivative operator, we work with differentials, which allow a vector function to remain as a vector. Hence, this allows us to apply the well understood tools of vector calculus to handle our problem.

A.1 Kronecker Product

In this section, we define the Kronecker Product, which is critical in making sense of the matrix calculus through its relation with the vectorizing operator.

Definition A.1 (Kronecker Product). *The Kronecker Product $A \otimes B$ of matrix $A = (A_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ and $B = (B_{ij})_{1 \leq i \leq k, 1 \leq j \leq l}$ is defined as*

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}, \quad (\text{A.1})$$

where $A \otimes B$ is a $mk \times nl$ matrix.

Proposition A.1. *If A, B, C and D are $m \times n, k \times l, n \times q$ and $l \times p$ matrices respectively, then*

$$AC \otimes BD = (A \otimes B)(B \otimes D) \quad (\text{A.2})$$

and $AC \otimes BD$ is a $mk \times qp$ matrix.

Proof.

$$\begin{aligned} AC \otimes BD &\stackrel{(\text{A.1})}{=} \begin{bmatrix} \sum_{j=1}^n A_{1j}C_{j1}BD & \dots & \sum_{j=1}^n A_{1j}C_{jq}BD \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^n A_{mj}C_{j1}BD & \dots & \sum_{j=1}^n A_{mj}C_{jq}BD \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B & \dots & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{m1}B & \dots & A_{mn}B \end{bmatrix} \begin{bmatrix} C_{11}D & \dots & C_{1q}D \\ \vdots & \ddots & \vdots \\ C_{n1}D & \dots & C_{nq}D \end{bmatrix} \\ &\stackrel{(\text{A.1})}{=} (A \otimes B)(C \otimes D). \end{aligned}$$

□

A.2 vec operator and its relation to the Kronecker product

We provide the definition of the vec operator and prove some properties that relate the vec operator to the Kronecker product.

Definition A.2 (vec operator). *The vectorizing operator vec converts a matrix to a column vector by stacking the columns of the matrix, i.e. If*

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix}, \quad (m \times n)$$

then

$$\text{vec}(A) = \begin{bmatrix} A_{11} \\ \vdots \\ A_{m1} \\ A_{12} \\ \vdots \\ A_{m2} \\ \vdots \\ A_{1n} \\ \vdots \\ A_{mn} \end{bmatrix}, \quad (mn \times 1). \quad (\text{A.3})$$

Next, we prove 2 propositions that relate the vec operator to the Kronecker product.

Proposition A.2. *If A and B is a $m \times 1$ and $n \times 1$ column vector respectively then we have*

$$\text{vec}(AB^\top) = B \otimes A. \quad (\text{A.4})$$

Proof. We have

$$\begin{aligned} \text{vec}(AB^\top) &\stackrel{(\text{A.3})}{=} \begin{bmatrix} B_1 A \\ \vdots \\ B_n A \end{bmatrix} \\ &\stackrel{(\text{A.1})}{=} B \otimes A. \end{aligned}$$

□

Proposition A.3. *If A , B and C are $m \times n$, $n \times k$ and $k \times l$ matrices respectively, then*

$$\text{vec}(ABC) = (C^\top \otimes A) \text{vec}(B).$$

Proof. If we represent the i -th column of matrix B by $B_{.i}$ and let the i -th column of a $k \times k$ identity matrix be e_i , then

$$\begin{aligned}
 \text{vec}(ABC) &= \text{vec}\left(A \sum_{i=1}^k (B_{.i} e_i^\top) C\right) \\
 &= \sum_{i=1}^k \text{vec}\left((AB_{.i}) (e_i^\top C)\right) \\
 &\stackrel{\text{(A.4)}}{=} \sum_{i=1}^k ((C^\top e_i) \otimes (AB_{.i})) \\
 &\stackrel{\text{(A.2)}}{=} (C^\top \otimes A) \sum_{i=1}^k (e_i \otimes B_{.i}) \\
 &\stackrel{\text{(A.2)}}{=} (C^\top \otimes A) \text{vec}\left(\sum_{i=1}^k (B_{.i} e_i^\top)\right) \\
 &= (C^\top \otimes A) \text{vec}(B).
 \end{aligned}$$

□

A.3 vec and its relation to the Trace operator

Definition A.3 (Trace Operator). *The trace Tr of a $m \times m$ -square matrix A is defined as the sum of the main diagonal elements of matrix A , i.e.*

$$\text{Tr}(A) = \sum_{i=1}^m A_{ii}.$$

We note that the above definition is not effective in allowing us to derive the derivative of a trace on the product of matrices. So, in the following proposition, we introduce an alternative representation of trace, one which makes use of the vec operator.

Proposition A.4. *If A and B are both $m \times n$ matrices, then we can write*

$$\text{Tr}(A^\top B) = (\text{vec}(A))^\top \text{vec}(B).$$

Proof. We have

$$\begin{aligned}
 \text{Tr} (A^\top B) &= \sum_{k=1}^m \sum_{j=1}^n (A_{kj}^\top B_{jk}) \\
 &= \sum_{k=1}^m \sum_{j=1}^n (A_{jk} B_{jk}) \\
 &\stackrel{\text{(A.3)}}{=} (\text{vec} (A))^\top \text{vec} (B).
 \end{aligned}$$

□

A.4 Differential and Derivative

In order to use the tools of vector calculus, we vectorize any matrix and work solely with differentials in our analysis of matrix calculus.

Definition A.4 (Differential). *Let $F(X)$ be a matrix function and X be a matrix. The differential $d \text{vec} (F(X))$ is*

$$d (\text{vec} (F(X))) := D (\text{vec} (F(X))) d (\text{vec} (X)),$$

where $D \text{vec} (F(X))$ is the first derivative of $\text{vec} (F(X))$.

The differential defined above has the following fundamental properties.

Proposition A.5 (Fundamental properties of differentials). *Let $A(X)$ and $B(X)$ be matrix functions of similar dimensions. If C is a constant matrix and E is a constant matrix such that E and $A(X)$ are conformable, then*

1. $dC = O$, where O is a null matrix with the same dimensions as C .
2. $d (\text{Tr} (A(X))) = \text{Tr} (d (A(X)))$.
3. $d (EA(X)) = E d (A(X))$.
4. $d (A(X) + B(X)) = d (A(X)) + d (B(X))$.
5. $d (A(X) B(X)) = d (A(X)) B(X) + A(X) d (B(X))$.
6. $d (A(X) \otimes B(X)) = d (A(X)) \otimes B(X) + A(X) \otimes d (B(X))$.

Next, we see that the Chain Rule also holds in multivariate settings.

Proposition A.6 (Multivariate Chain Rule). *Let $f(x)$, $g(y)$ and $h(x)$ be some vector functions, with x being a vector and y a vector with the same dimension as $h(x)$. Then we have when $f(x) = g(h(x))$,*

$$D(f(x)) = (D(g(h(x))))(D(h(x))).$$

In the following section, we compute the derivatives of various functions, by working with differentials.

A.5 Derivatives of Vector Functions

Proposition A.7. *If $f(x) = Ax$, where A is a $m \times n$ matrix and x is $n \times 1$ column vector, then*

$$D(f(x)) = A.$$

Proof. We have

$$d(f(x)) = d(Ax) = Adx.$$

So,

$$D(f(x)) = A.$$

□

Proposition A.8. *If $f(x) = x^\top Ax$, where A is a $n \times n$ matrix and x is $n \times 1$ column vector, then*

$$D(f(x)) = x^\top (A + A^\top).$$

Proof. Since

$$\begin{aligned} d(f(x)) &= d(x^\top Ax) \\ &= (dx)^\top Ax + x^\top Adx \\ &= x^\top A^\top dx + x^\top Adx \end{aligned}$$

then

$$D(f(x)) = x^\top A^\top dx + x^\top A = x^\top (A + A^\top).$$

□

Proposition A.9. Let $f(x) = Ax$

A.6 Derivatives of Matrix Functions

Proposition A.10. We have $f(x) = A(x)x$, where $A(X)$ is a matrix functions and x is a column vector. Then

$$D(f(x)) = A(x) + (x^\top \otimes \text{Id}) D(\text{vec}(A(x))).$$

Proof. We write

$$\begin{aligned} d(f(x)) &= d \text{vec}(f(x)) \\ &= \text{vec}(d(A(x))x + A(x)dx) \\ &= (x^\top \otimes \text{Id}) d(\text{vec}(A(x))) + A(x)dx \\ &= (x^\top \otimes \text{Id}) D(\text{vec}(A(x))) dx + A(x)dx \end{aligned}$$

Thus, we have

$$D(f(x)) = A(x) + (x^\top \otimes \text{Id}) D(\text{vec}(A(x))).$$

□

Proposition A.11. Let $f(x) = AB(x)Cx$, where A and C are constant matrices, $B(x)$ is a matrix function and x is a column vector. Then we have

$$D(f(x)) = (x^\top C^\top \otimes A) D(\text{vec}(B(x))) + AB(x)C.$$

Proof. We have

$$\begin{aligned} d(f(x)) &= \text{vec}(d(AB(x)Cx)) \\ &= \text{vec}(A(dB(x))Cx + AB(x)Cdx) \\ &= (x^\top C^\top \otimes A) d(\text{vec}(B(x))) + AB(x)Cdx \\ &= (x^\top C^\top \otimes A) D(\text{vec}(B(x))) dx + AB(x)Cdx. \end{aligned}$$

Hence

$$D(f(x)) = (x^\top C^\top \otimes A) D(\text{vec}(B(x))) + AB(x)C.$$

□

A.7 Derivatives of Trace Operator

Proposition A.12. *Suppose that $f(X) = \text{Tr}(AXB)$, where A , B and X are matrices, then we have*

$$D(f(X)) = (\text{vec}(A^\top B^\top))^\top.$$

Proof. Recall that Trace is invariant under cyclic permutations, thus

$$\begin{aligned} d(f(X)) &= d(\text{Tr}(AXB)) \\ &= \text{Tr}(d(BAX)) \\ &= \text{Tr}(BAdX) \\ &= (\text{vec}(A^\top B^\top))^\top d(\text{vec}(X)). \end{aligned}$$

and so we prove that

$$D(f(X)) = (\text{vec}(A^\top B^\top))^\top.$$

□

A.8 Divergence

We provide the definition of the divergence of a matrix and derive some properties of divergence with the results of matrix calculus.

Definition A.5 (Divergence of a Matrix Function). *The divergence of a $n \times n$ matrix function $A(x)$, where x is a $n \times 1$ column vector is*

$$\text{div}(A(x)) = \begin{bmatrix} \sum_{j=1}^n \frac{\partial A_{1j}(x)}{\partial x_j} \\ \vdots \\ \sum_{j=1}^n \frac{\partial A_{nj}(x)}{\partial x_j} \end{bmatrix}.$$

Proposition A.13 (Product rule for divergence). *If $f(x)$ is a scalar function, $v(x)$ is a $n \times 1$ vector function and x is a $n \times 1$ column vector, then*

$$\operatorname{div}(f(x) v(x)) = D(f(x)) v(x) + f(x) \operatorname{div}(v(x)).$$

Proof. Recall for every vector function $g(x)$, where x , $\operatorname{div}(g(x)) = \operatorname{Tr}(D(g(x)))$. Therefore,

$$\begin{aligned} \operatorname{div}(f(x) v(x)) &= \operatorname{Tr}(D(f(x) v(x))) \\ &= D(f(x)) v(x) + \operatorname{Tr}(f(x) D(v(x))) \\ &= D(f(x)) v(x) + f(x) \operatorname{Tr}(D(v(x))) \\ &= D(f(x)) v(x) + f(x) \operatorname{div}(v(x)). \end{aligned}$$

□

Proposition A.14. *If $u(x)$ and $v(x)$ are $n \times 1$ vector functions, where x is a $n \times 1$ column vector, then we have*

$$\operatorname{div}\left(u(x) (v(x))^\top\right) = D(u(x)) v(x) + \operatorname{div}(v(x)) u(x).$$

Proof. The proof is completed by writing the vectors in its components and applying the definition of divergence. □

Bibliography

- [BS73] Fischer Black and Myron Scholes. “The pricing of options and corporate liabilities”. In: *J. Polit. Econ.* 81.3 (1973), pp. 637–654.
- [CGS11] Louis H. Y. Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein’s method*. Probability and its Applications (New York). Springer, Heidelberg, 2011, pp. xii+405. ISBN: 978-3-642-15006-7.
- [DEM16] Bruce K. Driver, Nathaniel Eldredge, and Tai Melcher. “Hypoelliptic heat kernels on infinite-dimensional Heisenberg groups”. In: *Trans. Amer. Math. Soc.* 368.2 (2016), pp. 989–1022.
- [DM16] José Da Fonseca and Claude Martini. “The α -hypergeometric stochastic volatility model”. In: *Stochastic Process. Appl.* 126.5 (2016), pp. 1472–1502.
- [FF07] Dmitry K. Faddeev and Vera N. Faddeeva. *Computational methods of linear algebra*. WH Freeman and company, 2007.
- [FG11] José da Fonseca and Martino Grasselli. “Riding on the smiles”. In: *Quant. Finance* 11.11 (2011), pp. 1609–1632.
- [Fou+11] Jean-Pierre Fouque et al. *Multiscale stochastic volatility for equity, interest rate, and credit derivatives*. Cambridge University Press, Cambridge, 2011, pp. xiv+441.
- [FS86] Wendell H. Fleming and Panagiotis E. Souganidis. “Asymptotic series and the method of vanishing viscosity”. In: *Indiana Univ. Math. J.* 35.2 (1986), pp. 425–447.
- [FW12] Mark I. Freidlin and Alexander D. Wentzell. *Random perturbations of dynamical systems*. Third. Vol. 260. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1979 Russian original by Joseph Szücs. Springer, Heidelberg, 2012, pp. xxviii+458.

- [GM14] Emmanuel Gobet and Mohammed Miri. “Weak approximation of averaged diffusion processes”. In: *Stochastic Process. Appl.* 124.1 (2014), pp. 475–504.
- [Han+13] Jiguang Han et al. “Option prices under stochastic volatility”. In: *Appl. Math. Lett.* 26.1 (2013), pp. 1–4.
- [Hen09] Pierre Henry-Labordère. *Analysis, geometry, and modeling in finance*. Chapman & Hall/CRC Financial Mathematics Series. Advanced methods in option pricing. CRC Press, Boca Raton, FL, 2009, pp. xiv+383.
- [Hes93] Steven L. Heston. “A closed-form solution for options with stochastic volatility with applications to bond and currency options”. In: *Review of financial studies* 6.2 (1993), pp. 327–343.
- [HS00] David Heath and Martin Schweizer. “Martingales versus PDEs in finance: an equivalence result with examples”. In: *J. Appl. Probab.* 37.4 (2000), pp. 947–957.
- [Jan+11] Agnieszka Janek et al. “FX smile in the Heston model”. In: *Statistical tools for finance and insurance*. Springer, Heidelberg, 2011, pp. 133–162.
- [JR96] Jens Carsten Jackwerth and Mark Rubinstein. “Recovering probability distributions from option prices”. In: *The Journal of Finance* 51.5 (1996), pp. 1611–1631.
- [Kha02] Hassan K. Khalil. *Nonlinear systems*. third. Prentice Hall, 2002.
- [Kha12] Rafail Khasminskii. *Stochastic stability of differential equations*. second. Vol. 66. Stochastic Modelling and Applied Probability. With contributions by G. N. Milstein and M. B. Nevelson. Springer, Heidelberg, 2012, pp. xviii+339.
- [Kim02] Yong-Jin Kim. “Option pricing under stochastic interest rates: an empirical investigation”. In: *Asia-Pacific Financial Markets* 9.1 (2002), pp. 23–44.
- [KP10] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. 4th corrected printing. Applications of Mathematics 23. Berlin: Springer, 2010.
- [KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Second. Vol. 113. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991, pp. xxiv+470.

- [Kun97] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*. Vol. 24. Cambridge Studies in Advanced Mathematics. Reprint of the 1990 original. Cambridge University Press, Cambridge, 1997, pp. xiv+346.
- [Lee04] Roger W. Lee. “Option pricing by transform methods: extensions, unification and error control”. In: *Journal of Computational Finance* 7.3 (2004), pp. 51–86.
- [LK07] Roger Lord and Christian Kahl. “Optimal Fourier inversion in semi-analytical option pricing”. In: (2007).
- [MN99] Jan R. Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and economics*. 1999.
- [NP09] Ivan Nourdin and Giovanni Peccati. “Stein’s method on Wiener chaos”. In: *Probab. Theory Related Fields* 145.1-2 (2009), pp. 75–118.
- [Nua06] David Nualart. *The Malliavin calculus and related topics*. Second. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006, pp. xiv+382.
- [Oco93] Daniel L. Ocone. “A symmetry characterization of conditionally independent increment martingales”. In: *Barcelona Seminar on Stochastic Analysis (St. Feliu de Guíxols, 1991)*. Vol. 32. Progr. Probab. Birkhäuser, Basel, 1993, pp. 147–167.
- [Pri12] Nicolas Privault. “Laplace transform identities and measure-preserving transformations on the Lie-Wiener-Poisson spaces”. In: *J. Funct. Anal.* 263.10 (2012), pp. 2993–3023.
- [Pri13] Nicolas Privault. “Cumulant operators and moments of the Itô and Skorohod integrals”. In: *C. R. Math. Acad. Sci. Paris* 351.9-10 (2013), pp. 397–400.
- [Pri15a] Nicolas Privault. “Cumulant operators for Lie-Wiener-Itô-Poisson stochastic integrals”. In: *J. Theoret. Probab.* 28.1 (2015), pp. 269–298.
- [Pri15b] Nicolas Privault. “Stein approximation for Itô and Skorohod integrals by Edgeworth type expansions”. In: *Electron. Commun. Probab.* 20 (2015), No. 35, 10.
- [PS15] Nicolas Privault and Qihao She. “Conditionally Gaussian stochastic integrals”. In: *C. R. Math. Acad. Sci. Paris* 353.12 (2015), pp. 1153–1158.

- [PS16] Nicolas Privault and Qihao She. “Option pricing and implied volatilities in a 2-hypergeometric stochastic volatility model”. In: *Appl. Math. Lett.* 53 (2016), pp. 77–84.
- [TZ05] Howell Tong and Zhiqiang Zhang. “On time-reversibility of multivariate linear processes”. In: *Statist. Sinica* 15.2 (2005), pp. 495–504.
- [Üst10] Ali S. Üstünel. “Analysis on Wiener space and applications”. In: *arXiv preprint arXiv:1003.1649* (2010).
- [ÜZ95] Ali S. Üstünel and M. Zakai. “Random rotations of the Wiener path”. In: *Probab. Theory Related Fields* 103.3 (1995), pp. 409–429.
- [VY00] L. Vostrikova and M. Yor. “Some invariance properties (of the laws) of Ocone’s martingales”. In: *Séminaire de Probabilités, XXXIV*. Vol. 1729. Lecture Notes in Math. Springer, Berlin, 2000, pp. 417–431.
- [Wat87] Shinzo Watanabe. “Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels”. In: *Ann. Probab.* 15.1 (1987), pp. 1–39.
- [Yor79] Marc Yor. “Les filtrations de certaines martingales du mouvement brownien dans $\mathbf{R}^n$ ”. In: *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*. Vol. 721. Lecture Notes in Math. Springer, Berlin, 1979, pp. 427–440.
- [Yor80] Marc Yor. “Remarques sur une formule de Paul Lévy”. In: *Seminar on Probability, XIV (Paris, 1978/1979) (French)*. Vol. 784. Lecture Notes in Math. Springer, Berlin, 1980, pp. 343–346.
- [ZHG13] Qiang Zhang, Jiguang Han, and Ming Gao. “Option price with stochastic volatility for both fast and slow mean-reverting regimes”. In: *C. R. Math. Acad. Sci. Paris* 351.9-10 (2013), pp. 411–414.