Advanced Monte Carlo methods for the Hull–White model

by

Finnan Tewolde and Jiahui Zhang

Kandidatarbete i matematik / tillämpad matematik
Bachelor thesis in mathematics / applied mathematics

Date:
2018-05-31

Project name:
Advanced Monte Carlo methods for the Hull–White model

Author(s):
Finnan Tewolde and Jiahui Zhang

Version:
8th June 2018

Supervisor(s):
Anatoliy Malyarenko

Reviewer:
Christopher Engström

Examiner:
Linus Carlsson

Comprising:
15 ECTS credits
Acknowledgements

It has been a great honour to have Professor Anatoliy Malyarenko as our thesis instructor. We deeply thank him for his dedication and assistance to our thesis work. We truly appreciate him for his mentoring, leading and supporting of the whole time. We would also like to thank our Mälardalen University for providing us with this amazing bachelor program — Analytical Finance, for enriching our knowledge in mathematics and applied mathematics with specialisation in finance.
# Contents

Acknowledgements 1

1 Introduction 3
   1.1 The Formulation of the Task 3
   1.2 Review of Literature 7

2 Theoretical Considerations 9
   2.1 The Idea 9
   2.2 The Stratonovich Integral 12
   2.3 The Moment Matching Condition 15

3 Applications to the Hull–White Model 23
   3.1 A brief explanation of classical Monte Carlo methods 23
   3.2 The method of Lyons and Victoir for the Hull–White model 26

4 Conclusions 30

A Criteria for a Bachelor Thesis 31

Bibliography 32

Index 34
Chapter 1

Introduction

1.1 The Formulation of the Task

Financial engineering is a part of applied mathematics that studies market models. According to [6], a typical market model contains the following components.

- A probability space \((\Omega, \mathcal{F}, P)\) with a physical (historical) probability measure \(P\). It is usually supposed that \(\mathcal{F}\) contains all subsets of all sets of \(P\)-measure 0.

- The filtration \(\{\mathcal{F}_t: t \in [0, T]\}\), generated by one or more standard Brownian motions \(W_i(t), 1 \leq i \leq d, 0 \leq t \leq T\). The \(\sigma\)-field \(\mathcal{F}_t\) is the intersection of all \(\sigma\)-fields \(\mathcal{G}\) that contain all subsets of all sets of \(P\)-measure 0 and with respect to which the random variables \(\{W_i(s): 1 \leq i \leq d, 0 \leq s \leq t\}\) are measurable. Usually

- A \(\mathcal{F}_t\)-adapted stochastic process, \(S(t)\) — the price process, is measurable with respect to \(\mathcal{F}_t\) for every \(t \in [0, T]\);

- A non-negative \(\mathcal{F}_t\)-adapted stochastic process \(r(t)\) — the time \(t\) instantaneous spot rate.

- A positive \(\mathcal{F}_t\)-adapted stochastic process \(\sigma(t)\) satisfying \(\int_0^T E[\sigma^2(t)] dt < \infty\), the volatility.

- An \(\mathcal{F}_t\)-adapted stochastic process \(\delta(t)\) satisfying \(\int_0^T E[|\delta(t)|] dt < \infty\) — the continuous dividend rate.

Remark 1. For simplicity, in what follows we will put \(r(t) = \delta(t) = 0\). In other words, we suppose that the stock price \(S(t)\) is already discounted and pays no dividends.

Example 1 (The Hull–White model, [5]). The following model was proposed by Hull and White,

\[
\begin{align*}
    dX_t &= \mu X_t dt + Y_t X_t dW_t \\
    dY_t &= \nu Y_t dt + \xi Y_t dZ_t
\end{align*}
\]  

(1.1)
where $W$ and $Z$ are correlated Brownian motions with correlation coefficient $\rho$, that is,

$$dW_t = \sqrt{1 - \rho^2} \, d\tilde{W}_t + \rho \, dZ_t, \quad -1 \leq \rho \leq 1,$$

(1.2)

Here, $\tilde{W}_t$ and $Z_t$ are independent standard Brownian motions.

**Definition 1** (A contingent claim). [6]

A contingent claim is a random variable $X$ representing a pay-off at some future time $T$.

**Example 2** (A European call option). [6]

European call options is that the holder has the right to buy the underlying asset by maturity, and it can only be exercised at the maturity.

How to find a fair price of a contingent claim?

Answer: The fair price of a contingent claim can be determined by no-arbitrage pricing arguments, it is described as following [6].

Assume that the market contains $n + 1$ securities. Denote the time-$t$ price of security $i$ by $S_i(t)$, $0 \leq i \leq n$. The security $S_0(t)$ is risk-free, while the other securities are risky. We suppose that

$$S(t) = (S_0(t), \ldots, S_n(t))^\top$$

is a stochastic process. We call it the price process.

**Definition 2** (The portfolio process, [6]). Let $\theta_i(t)$ denote the amount of security $i$ at time $t$. Put the vector $\theta(t)$ to $\theta(t) = (\theta_0(t), \theta_1(t), \ldots, \theta_n(t))^\top$. The process $\{\theta(t): 0 \leq t \leq T\}$ is called a portfolio process.

**Definition 3** (The value process, [6]). The value process is

$$V(t) = \sum_{i=0}^{n} \theta_i(t)S_i(t).$$

Denote by $d_i(t)$ the dividend paid by the $i$th security at time $t$ and let

$$D_i(t) = \int_0^t d_i(s) \, ds$$

be the cumulative dividend.

**Definition 4** (The gain process, [6]). The gain process is

$$G_i(t) = S_i(t) + D_i(t), \quad 0 \leq i \leq n.$$
**Definition 6** (Replicating portfolio, [6]). A replicating portfolio for a contingent claim \( X \) is a self-financing portfolio \( \theta(t) \) such that \( V(T) = X \).

That is,
\[
X = V(0) + \sum_{i=0}^{n} \int_0^T \theta_i(s) \, dG_i(s).
\] (1.3)

**Definition 7** (Attainable contingent claim, [6]). A contingent claim is called attainable if a replicating portfolio exists.

**Definition 8** (Arbitrage opportunity, [6]). An arbitrage opportunity is the existence of some self-financing trading strategy
\[
\{ \theta(t) = (\theta_0(t), \theta_1(t), \ldots, t = 1, 2, \ldots, T) \}
\] (1.4)
such that (a) \( V(0) = 0 \), and (b) \( V(T) \geq 0 \) almost surely (abbreviation as a.s.) and \( V(T) > 0 \) with positive probability.

An arbitrage opportunity is a measure of risk-free profit making. We start with nothing and will not go into debt, however, there is a possibility of resulting in a positive amount of money. If there exists such an opportunity, it will affect the prices of securities as people will use it and cause non-equilibrium.

If there are no arbitrage opportunities in the market, then the initial price of a replicating portfolio, \( V(0) \), is the correct price of the contingent claim \( X \). Indeed, if not so, we can easily construct an arbitrage opportunity.

Note: \( \theta(t) \) might change due to the actual price process.

The left-hand side of (1.4) is the value of the portfolio at time \( t \), which represents the value of the portfolio before transactions, while the right-hand side one is the value of the portfolio at time \( t \) after the transactions.

**Definition 9** (Risk-neutral measure). A measure \( P^* \) is said to be risk–neutral if

1. \( P^* \) is equivalent to \( P \), i.e., \( P(A) > 0 \) if and only if \( P^*(A) > 0 \) for all \( A \in \mathcal{F} \), and

2. \( E^*[S^*_i(t)] = S^*_i(s), 0 \leq s \leq t \leq T \), where
\[
S^*_i(t) = \frac{S_i(t)}{S_0(t)}.
\]

In other words, all the price processes \( S_i(t) \), discounted with respect to the money-market account \( S_0(t) \), are martingales under \( P^* \).

**Theorem 1** (Fundamental theorem of financial engineering, part 1). There are no arbitrage opportunities in a securities market, if and only if there exists a risk neutral probability measure. In this case, the price for an attainable contingent claim \( X \) is given by
\[
V(0) = E^*\left[ \frac{X}{S_0(T)} \right]
\]
for every replicating trading strategy.
Example 3 (Risk-neutral measures in the Hull–White model). Consider the Hull–White model (1.1)–(1.2).

Definition 10. A stochastic process $\gamma(t)$ is called a market price of volatility risk if it is adapted to the filtration generated by the Wiener process $Z(t)$, for every $t > 0$ we have

$$P\left\{ \int_0^t \gamma^2(s) \, ds < \infty \right\} = 1,$$

and the stochastic process

$$\varepsilon_\gamma(t) = \frac{1}{2} \int_0^t \gamma^2(s) \, ds + \int_0^t \gamma(s) \, dZ(s)$$

satisfies the following condition

$$E[\exp(-\varepsilon_\gamma(t))] = 1, \quad t \geq 0.$$

Let $\zeta(t)$ be a stochastic process satisfying

$$P\left\{ \int_0^t \zeta^2(s) \, ds < \infty \right\} = 1,$$

and let $P_\zeta,\gamma_T$ be the measure on the $\sigma$-field $\mathfrak{F}_T$ defined by

$$P_\zeta,\gamma_T(A) = \int_A \exp(-\varepsilon_\zeta,\gamma(t)) \, dP(\omega), \quad A \in \mathfrak{F}_T. \quad (1.5)$$

Let $\mathfrak{F}_\omega$ be the minimal $\sigma$-field containing $\mathfrak{F}_T$ for all $t \geq 0$. There exists a unique measure $P^{\zeta,\gamma}$ on $\mathfrak{F}_\omega$ such that the restriction of $P^{\zeta,\gamma}$ to $\mathfrak{F}_T$ coincides with $P_\zeta,\gamma_T$ for all $t \geq 0$.

Theorem 2 ([3, Conclusion 2.19, Conclusion 2.34, Conclusion 2.37]). If the correlation coefficient $\rho$ of the Hull–White model is 0, then put $\zeta(t) = \mu$. For any market price of volatility risk $\gamma(t)$ the measure $P^{\mu,\gamma}_T$ defined by (1.5) is risk-neutral.

If $\rho < 0$, then the Hull–White model is risk-neutral under the physical probability measure $P$.

If $\rho > 0$, then put $\zeta(t) = 0$, $\gamma(t) = bY(t)$. The measure $P^{0,bY}_T$ is risk-neutral if and only if $b > \rho$.

In what follows we consider only the case of negative correlation, because this case reflects the so called leverage effect, that is, the volatility of the stock increases when the stock price decreases. Then, the Hull–White model (1.1)–(1.2) is already written under the risk-neutral measure.

Definition 11 (Complete market). A securities market is said to be complete, if every contingent claim is attainable: otherwise, the market is incomplete.

Theorem 3 (Fundamental theorem of financial engineering, part 2). Suppose that a securities market admits no arbitrage opportunities. Then, it is complete if and only if there exists a unique risk–neutral probability measure.
Generally, there are two general methods to find the price of a contingent claim: the Feynman–Kac theorem and the Monte Carlo method. In this thesis, we will use a modification of the Monte Carlo method initially proposed by [7]. The review of relevant literature follows.

1.2 Review of Literature

As we focus on Advanced Monte-Carlo Methods for Hull–White Model, we have skimmed through bunches of scientific papers concerning this topic and the databases we use mainly are GoogleScholar and MathSciNet. By entering keywords or synonyms we obtained corresponding contents.

The Hull–White model is introduced in [5], where they put forward a problem about one option-pricing. The authors claim that the pricing of a European call on an asset has a stochastic volatility. Nevertheless, this critical problem has remained unsettled. The option price is set in series form where the stochastic volatility does not depend on stock price. They generated the numerical solutions for the case in which the volatility has correlation with the stock price. They also have figured out that the Black–Scholes price frequently overvalues the prices of the options and that the degree of overpricing increases with the time to maturity [5].


[10] establishes another approximation approach named after Kusuoka, where they implement discrete random variables to non-commutative multi-factor models. Heath–Jarrow–Morton model is used to valuate the interest-rate derivatives, which is obtainable if the Kusuoka approximation is employed with the tree-based branching algorithm [10].

[12] introduces Fujiwara’s method, which is regarded as an extrapolation of order 6 of the Ninomiya-Victoir weak approximation scheme to achieve numerical approximation of SDEs solution. Comparing with other approach, this way is much simpler as it only processes half of number of steps [12].

Litterer and Lyons have mentioned Monte Carlo method, although they assume another method called particle as it is much more accurate for describing evolving measures. Since it has disadvantage in a higher order case, they develop an algorithm that can be used to give an application to the cubature on Wiener space method developed by Lyons and Victoir [8].

Bayer, Christian and Friz have concluded in their article [1] that cubature on Wiener space renders a powerful choice to Monte Carlo simulation for the integration of certain functionals on Wiener space. Rapid computation of European option prices is permitted in cubature in generic diffusion models [1].

By easing off the routine presumption of the cubature approach, article [2] proposes the theory of interest rate of productive numerical methods for generic Heath–Jarrow–Morton equations using high-order weak approximation schemes. Those schemes permit quasi-Monte Carlo implementations as the integration space has relatively low-dimensions.

So long as the order of quasi-Monte Carlo convergence is determined, we are confirmed that the complexity of the resulting algorithm is remarkably lower than that of the multilevel
Monte Carlo algorithms. To apply the approach to practice, the authors describe the setting of weighted function spaces, and numerical analysis for stochastic partial differential equations due to those spaces. At least, they offer an application where they efficiently calibrate a Heath–Jarrow–Morton equation to the caplet market [2].
Chapter 2

Theoretical Considerations

2.1 The Idea

Denote $W^*_0(t) = t$, and consider a general model

$$dX(t) = \sum_{i=0}^{d} \tilde{V}_i(X(t)) dW^*_i(t),$$

(2.1)

$$X(0) = x_0,$$

where $x_0 \in \mathbb{R}^N$, $\tilde{V}_i: \mathbb{R}^N \to \mathbb{R}^N$, and $W^*_1(t), \ldots, W^*_d(t)$ are independent standard Brownian motions under the risk-neutral measure $P^*$. In the Itô integral form, we obtain

$$X(t) = x_0 + \sum_{i=0}^{d} \int_0^t \tilde{V}_i(X(s)) \, dW^*_i(s).$$

(2.2)

Equation 2.1 is exactly the same as equation 2.2 since it is written in different forms.

Recall the Itô formula.

Theorem 4 (The Itô formula, [6]). Let $f(x)$ be a twice continuously differentiable function. The Itô formula in the multivariate setting can be written in the Itô integral form as,

$$f(X(t)) = f(x_0) + \sum_{i=0}^{d} \int_0^t (\tilde{V}_i f)(X(s)) \, dW^*_i(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t (\tilde{V}_i \tilde{V}_j f)(X(s)) \, dW^*_i(s),$$

where we denote

$$(\tilde{V}_i f)(x) = \sum_{j=1}^{N} \tilde{V}_i^j(x) \frac{\partial f}{\partial x_j}(x)$$

(2.3)

and where

$$(\tilde{V}_i \tilde{V}_j f)(x) = \tilde{V}_i(\tilde{V}_j f)(x) = \sum_{k=1}^{N} \tilde{V}_i^k(x) \frac{\partial}{\partial x_k} \left( \sum_{l=1}^{N} \tilde{V}_j^l(x) \frac{\partial f}{\partial x_l}(x) \right).$$
Proof. Integrate [6, Equation (13.47)] from 0 to \( t \).

Denote by \( C_0([0, T]; \mathbb{R}^d) \) the set of all continuous functions \( f: [0, T] \rightarrow \mathbb{R}^d \) with \( f(0) = 0 \). Denote

\[
\| f \| = \max_{t \in [0, T]} \| f(t) \|,
\]

where

\[
\| f(t) \| = \sqrt{f_1^2(t) + \cdots + f_d^2(t)}.
\]

It is possible to prove that the function \( \| f \| \) has the following properties.

1. \( \| f \| \geq 0 \). Moreover, \( \| f \| = 0 \) if and only if \( f = 0 \).

2. \( \| \lambda f \| = |\lambda| \cdot \| f \| \), \( \lambda \in \mathbb{R} \).

3. \( \| f + g \| \leq \| f \| + \| g \| \).

In other words, \( \| f \| \) is a norm on the space \( C_0([0, T]; \mathbb{R}^d) \).

We prove Property 3 when \( d = 1 \). We need to prove that

\[
\max_{t \in [0, T]} |f(t) + g(t)| \leq \max_{t \in [0, T]} |f(t)| + \max_{t \in [0, T]} |g(t)|.
\]

Indeed, let \( t_0 \in [0, T] \) be an arbitrary point where the continuous function \( |f(t) + g(t)| \) attains its maximal value. Then we have

\[
\max_{t \in [0, T]} |f(t) + g(t)| = |f(t_0) + g(t_0)| \leq |f(t_0)| + |g(t_0)| \leq \max_{t \in [0, T]} |f(t)| + \max_{t \in [0, T]} |g(t)|,
\]

where we use the inequality \( |x + y| \leq |x| + |y| \) with \( x = f(t_0) \) and \( y = g(t_0) \).

Definition 12. A set \( A \subseteq C_0([0, T]; \mathbb{R}^d) \) is called open if for any \( f \in A \) there is an \( \varepsilon > 0 \) such that the open ball

\[
B(f, \varepsilon) = \{ g \in C_0([0, T]; \mathbb{R}^d) : \| g - f \| < \varepsilon \}
\]

is a subset of \( A \).

Define \( \mathcal{F} \) as the minimal \( \sigma \)-field of the subsets of \( \Omega = C_0([0, T]; \mathbb{R}^d) \) that contains all open sets. The following result is proved in [14].

Theorem 5 ([14]). There is a unique probability measure \( P^* \) on the \( \sigma \)-field \( \mathcal{F} \) such that the stochastic process defined by

\[
W^*(t, \omega) = \omega(t), \quad \omega \in \Omega, \quad 0 \leq t \leq T
\]

is the \( d \)-dimensional standard Brownian motion.
Let $X(\omega)$ be the path of the stochastic process $X$ that corresponds to $\omega \in \Omega$. The price of a contingent claim $Y = Y(X(t, \omega))$ in our model is

$$P = E^*[Y] = \int_\Omega Y(X(\omega)) \, dP^*(\omega),$$

where the first equality follows from Theorem 1, and the second one is just the definition of the expected value. The idea is to replace the complicated space $\Omega$, the complicated probability measure $P^*$, and the complicated path $X(\omega)$ with a “simple” space $\Omega'$, a “simple” probability measure $Q$, and a “simple” path $X'(\omega)$ and write

$$P = \int_\Omega Y(X(\omega)) \, dP^*(\omega) \approx \int_{\Omega'} Y(X'(\omega')) \, dQ(\omega').$$

Kusuoka [7] proposed $\Omega'$ to be the so called space of functions of bounded variation. We propose even more simple space.

**Definition 13.** A function $g: [0, T] \to \mathbb{R}$ is called absolutely continuous if the Fundamental Theorem of calculus is true for $g$:

$$g(t) = g(0) + \int_0^t g'(s) \, ds, \quad t \in [0, T].$$

(2.4)

**Remark 2.** If the integral in (2.4) is a Riemann integral, then Definition 13 is too restrictive. For example, the function $g(t) = |t|$ is not absolutely continuous on the interval $[-1, 1]$, because $g'(0)$ does not exist. These complications disappear if the integral is understood as the Lebesgue integral. The details are outside the scope of this thesis.

Denote by $\Omega' = C_{0,AC}([0, T]; \mathbb{R}^{d+1})$ the set of functions $g = (g^0(s), \ldots, g^d(s))^\top$ with $g^i(0) = 0$ and $g^i$ being absolutely continuous, $0 \leq i \leq d$. Define

$$\|g^i\|(s) = \int_0^s |(g^i)'(u)| \, du.$$  

Denote $dg^i(s) = (g^i)'(s) \, ds$.

Let $\mathcal{F}'$ be the minimal $\sigma$-field of subsets of $\Omega'$ that contains all open balls. Let $Q$ be a probability measure on $\mathcal{F}'$. Along with the system (2.2), consider the following system

$$X(t) = x_0 + \sum_{i=0}^d \int_0^t V_i(X(s)) \, dg^i(s, \omega'),$$

(2.5)

where $g(s, \omega') = (g^0(s, \omega'), \ldots, g^d(s, \omega'))^\top$ is the stochastic process defined on the probability space $(\Omega', \mathcal{F}', Q)$ by

$$g(s, \omega') = \omega'(s).$$

In fact, this is a system of random ordinary differential equations. Indeed, by the Fundamental Theorem of calculus we obtain

$$\frac{dX(t)}{dt} = \sum_{i=0}^d V_i(X(t)) (g^i)'(t),$$

$$X(0) = x_0.$$  

(2.6)
Define the approximate price of the contingent claim $Y$ by

$$P \approx \int_{\Omega'} Y(\bar{X}(\omega')) \, dQ(\omega'),$$

where $\bar{X}$ is the solution of the system (2.6). The next task is to try to estimate the absolute error

$$\left| \int_{\Omega} Y(X(\omega)) \, dP^*(\omega) - \int_{\Omega'} Y(\bar{X}(\omega')) \, dQ(\omega') \right|$$

(2.7)

of our approximation.

Assume for simplicity that $Y$ is a contingent claim of European type, that is, $Y(X) = Y(X(T))$, the value of the claim depends only on the final value $X(T)$ of the process $X$. On the one hand, by the Itô formula, Theorem 4,

$$Y(X(T)) = Y(x_0) + \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(X(s)) \, dW_i^*(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^T (\tilde{V}_i \tilde{V}_j Y)(X(s)) \, dW_0^*(s),$$

where we assume that $Y$ is a twice continuously differentiable function. On the other hand, by the Fundamental Theorem of Calculus, we have

$$Y(\bar{X}(T)) = Y(x_0) + \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(\bar{X}(s)) \, ds.$$

The absolute error (2.7) becomes

$$\left| \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(X(s)) \, dW_i^*(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^T (\tilde{V}_i \tilde{V}_j Y)(X(s)) \, dW_0^*(s) - \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(\bar{X}(s)) \, ds \right|,$$

where the terms $Y(x_0)$ and $-Y(x_0)$ cancelled each other. It can be proved that the term

$$\left| \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(X(s)) \, dW_i^*(s) - \sum_{i=0}^{d} \int_0^T (\tilde{V}_i Y)(\bar{X}(s)) \, ds \right|$$

is small. We will give the exact formulation of this result later. The remaining term

$$\frac{1}{2} \sum_{i,j=1}^{d} \int_0^T (\tilde{V}_i \tilde{V}_j Y)(X(s)) \, dW_0^*(s)$$

is difficult to estimate. It would be good if it were equal to 0. Fortunately, it is possible, if we pass from the Itô integral to the so called Stratonovich integral.

### 2.2 The Stratonovich Integral

The most utilised alternative stochastic integral to the Itô integral is known as the Stratonovich integral. The Stratonovich integral is easy to manipulate as they are defined so that the chain rule of ordinary calculus is valid contrary to Itô integral. In other words, it obeys the ordinary calculus.
Definition 14. [6] A process $\psi(t)$ measurable with respect to the filtration $\{\mathcal{F}_t : t \in [0, T]\}$ is called simple if there exist a positive integer $n$, time moments $0 = t_0 < t_1 < \cdots < t_n = T$ and $\mathcal{F}_{t_i}$-measurable random variables $\xi_i$, $1 \leq i \leq n - 1$ such that

$$\psi(t) = \psi(0)1_{(0,t]}(t) + \sum_{i=1}^{n-1} \xi_i 1_{(t_i,t_{i+1}]}(t),$$

where $1_A(t)$ denotes the indicator function of the set $A$.

Definition 15. [6] Let $t \in (t_i,t_{i+1}]$, and let $\psi(t)$ be a simple process. The Itô integral is

$$I(t) = \sum_{k=0}^{i-1} \psi(t_k)[z(t_{k+1}) - z(t_k)] + \psi(t_i)[z(t) - z(t_i)],$$

while the Stratonovich integral is

$$J(t) = \sum_{k=0}^{i-1} \psi\left(\frac{t_k + t_{k+1}}{2}\right)[z(t_{k+1}) - z(t_k)] + \psi\left(\frac{t_i + t_{i+1}}{2}\right)[z(t) - z(t_i)].$$

Using tools from functional analysis, see, e.g., [11], both definitions may be extended to include non-simple processes. We denote the Itô integral by

$$I(t) = \int_0^t \psi(s) dW(s)$$

and the Stratonovich integral by

$$J(t) = \int_0^t \psi(s) \circ dW(s).$$

Consider a system of stochastic differential equations under the risk-neutral measure in Stratonovich integral form:

$$X(t) = x_0 + \sum_{i=0}^{d} \int_0^t V_i(X(s)) \circ dW_i^s(s). \quad (2.8)$$

Theorem 6. [11] Let $f(x)$ be a continuously differentiable function. The Itô formula in the multivariate setting can be written in the Stratonovich integral form as,

$$f(X(t)) = f(x_0) + \sum_{i=0}^{d} \int_0^t (V_i f)(X(s)) \circ dW_i^s(s).$$

Consider the system (2.2) and assume it has a unique solution. Do there exists a system in the Stratonovich integral form (2.8) that has the same solution?
Theorem 7. [11] Assume that the system (2.2) has a unique solution. The system (2.8) with
\[ V_i(x) = \tilde{V}_i(x), \quad 1 \leq i \leq d \]
and
\[ V^*_0(x) = \tilde{V}^*_0(x) - \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \tilde{V}_j^k(x) \frac{\partial \tilde{V}_j^k}{\partial x_k}(x) \] (2.9)
has the same unique solution.

Equation (2.9) is called the Stratonovich correction.

Example 4 (The Hull–White model). Consider the model of Example 1. We have \( N = d = 2 \).

Denote
\[ W^*_1(t) = \tilde{W}(t), \quad W^*_2(t) = Z(t). \]

Then the Itô integral form of the Hull–White model becomes
\[ X(t) = X(0) + \int_0^t \mu X(s) \, dW^*_0(s) + \int_0^t \rho X(s) Y(s) \, dW^*_1(s) + \int_0^t \rho X(s) Y(s) \, dW^*_2(s), \]
\[ Y(t) = Y(0) + \int_0^t \nu Y(s) \, dW^*_0(s) + \int_0^t \xi Y(s) \, dW^*_2(s). \]

It follows that
\[ \dot{V}_0(x) = \left( \begin{array}{c} \mu x_1 \\ \nu x_2 \end{array} \right), \quad \dot{V}_1(x) = \left( \begin{array}{c} \sqrt{1 - \rho^2} x_1 x_2 \\ 0 \end{array} \right), \quad \dot{V}_2(x) = \left( \begin{array}{c} \rho x_1 x_2 \\ \xi x_2 \end{array} \right), \]
where \( x = (X Y)^T \). We calculate the Stratonovich correction by Equation (2.9).

\[ \dot{V}_0^*(x) = \tilde{V}_0^*(x) - \frac{1}{2} \sum_{j=1}^2 \sum_{k=1}^2 \tilde{V}^k_j(x) \frac{\partial \tilde{V}^k_j}{\partial x_k}(x). \]

Put \( i = 1 \). Then
\[ \dot{V}_0^1(x) = \tilde{V}_0^1(x) - \frac{1}{2} \sum_{j=1}^2 \sum_{k=1}^2 \tilde{V}^k_j(x) \frac{\partial \tilde{V}^k_j}{\partial x_k}(x). \] (2.10)

The partial derivatives \( \frac{\partial \tilde{V}^k_j}{\partial x_k}(x) \) are
\[ \frac{\partial \tilde{V}^1_1}{\partial x_1}(x) = \sqrt{1 - \rho^2} x_2, \quad \frac{\partial \tilde{V}^1_1}{\partial x_2}(x) = \sqrt{1 - \rho^2} x_1, \]
\[ \frac{\partial \tilde{V}^1_2}{\partial x_1}(x) = \rho x_2, \quad \frac{\partial \tilde{V}^1_2}{\partial x_2}(x) = \rho x_1. \]

Substitute these values to Equation (2.10). We obtain
\[ V_0^1(x) = \mu x_1 - \frac{1}{2} \left( \sqrt{1 - \rho^2} x_1 x_2 \sqrt{1 - \rho^2} x_2 + 0 \cdot \sqrt{1 - \rho^2} x_1 + \rho x_1 x_2 \rho x_2 + \xi x_2 \rho x_1 \right) \]
\[ = \mu x_1 - \frac{1}{2} \left( (1 - \rho^2) x_1 x_2^2 + \rho^2 x_1 x_2^2 + \xi \rho x_1 x_2 \right) \]
\[ = \mu x_1 - \frac{1}{2} \left( x_1 x_2 + \xi \rho \right). \]
Similarly, put \( i = 2 \). Then

\[
V_0^2(x) = V_0^2(x) - \frac{1}{2} \sum_{j=1}^{2} \sum_{k=1}^{2} \tilde{V}_k^j(x) \frac{\partial \tilde{V}_j^2}{\partial x_k}(x).
\]  \hspace{1cm} (2.11)

The partial derivatives \( \frac{\partial \tilde{V}_j^2}{\partial x_k}(x) \) are

\[
\frac{\partial \tilde{V}_1^2}{\partial x_1}(x) = 0, \quad \frac{\partial \tilde{V}_1^2}{\partial x_2}(x) = 0,
\]

\[
\frac{\partial \tilde{V}_2^2}{\partial x_1}(x) = 0, \quad \frac{\partial \tilde{V}_2^2}{\partial x_2}(x) = \xi.
\]

Substitute these values to Equation (2.11). We obtain

\[
V_0^2(x) = v x_2 - \frac{1}{2} \xi x_2 \cdot \xi
\]

\[
= v x_2 - \frac{1}{2} \xi^2 x_2.
\]

Finally,

\[
V_0(x) = \left( \mu x_1 - \frac{1}{2} x_1 x_2 (x_2 + \xi \rho) \right)
\]

and the Stratonovich integral form of the Hull–White model becomes

\[
X(t) = X(0) + \int_0^t \left[ \mu - \frac{1}{2} (Y(s)(Y(s) + \xi \rho)) \right] X(s) \circ dW_0^x(s) + \int_0^t \sqrt{1 - \rho^2} X(s) Y(s) \circ dW_1^x(s)
\]

\[
+ \int_0^t \rho X(s) Y(s) \circ dW_2^x(s),
\]

\[
Y(t) = Y(0) + \int_0^t \left[ v - \frac{1}{2} \xi^2 \right] Y(s) \circ dW_0^x(s) + \int_0^t \xi Y(s) \circ dW_2^x(s).
\]

### 2.3 The Moment Matching Condition

In Theorem 6, assume that the functions \( V_i(x) \) are several times continuously differentiable. We may then apply Theorem 6 again and again to the term \( (V_i f)(X(s)) \). To formulate the result, we need a notation.

Denote

\[
\mathcal{I} = \{ \emptyset \} \cup \bigcup_{k=1}^{\infty} \{0, 1, \ldots, d\}^k,
\]

\[
\mathcal{A}_n = \{ \alpha \in \mathcal{I} : \|\alpha\| \leq n \},
\]

\[
\mathcal{A}_n = \{ \alpha = (\alpha_0, \ldots, \alpha_k) \in \mathcal{I} : \|\alpha\| > n, \| (\alpha_1, \ldots, \alpha_k) \| \leq n \},
\]

\[
\|\alpha\| = \begin{cases} 0, & \text{if } \alpha = \emptyset, \\ k + \sum_{j=1}^{k} \{ j : 1 \leq j \leq k, \alpha_i = 0 \}, & \text{otherwise}, \end{cases}
\]  \hspace{1cm} \text{(2.12)}

\[
I(t, \alpha, \circ dW^x) = \begin{cases} 1, & \text{if } \alpha = \emptyset, \\ \int_0^t \int_0^{t_1} \cdots \int_0^{t_k} \circ dW_{\alpha_1}(t_1) \cdots \circ dW_{\alpha_k}(t_k), & \text{otherwise}, \end{cases}
\]
where \( \sharp \) denotes the number of elements in the set, and where \( \{0, 1, \ldots, d\}^k \) denotes the \( k \)th Cartesian power of the set \( \{0, 1, \ldots, d\} \), that is,

\[ \{0, 1, \ldots, d\}^k = \{(\alpha_1, \ldots, \alpha_k) : 0 \leq \alpha_i \leq d, 1 \leq i \leq k\}. \]

**Example 5.** As in the Hull–White model, put \( d = 2 \). We calculate \( \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_0, \mathcal{A}_1 \). For each \( \alpha \) that belong to any of the four above sets, we calculate \( I(t, \alpha, \circ dW^*) \).

**Solution:**

\( \mathcal{A}_0 = \alpha_0 = \emptyset \), as when \( \| \alpha_0 \| = 0 \), \( \alpha_0 = \emptyset \).

\( \mathcal{A}_1 = \{\emptyset, 1, 2\} \),

As \( \| \alpha_0 \| = 0 \), \( \alpha_0 = \emptyset \),

\( \| \alpha_1 \| = k + \sharp \{j : 1 \leq j \leq k, \alpha_i = 0\} = 1 + 0 = 1, \alpha_1 = 1 \)

and \( \| \alpha_2 \| = k + \sharp \{j : 1 \leq j \leq k, \alpha_i = 0\} = 1 + 1 = 2, \alpha_2 = 2 \), they all satisfy the condition \( \| \alpha_0 \| \leq 0, \| \alpha_1 \| \leq 1, \| \alpha_2 \| \leq 2 \), and all of \( \alpha_0, \alpha_1, \alpha_2 \in \mathcal{A} \).

\( \mathcal{A}_0 = \{0, 1, 2\} \)

As \( \| \alpha \| = \|0\| = 0 \leq 0 \), \( \| \alpha \| = \|1\| = 1 \leq 1 \| \alpha \| = \|2\| = 2 \leq 2 \), and \( \alpha \in \mathcal{A} \).

\( \mathcal{A}_1 = \{0, (0, 1), (0, 2), (1, 1), (1, 2), (2, 1), (2, 2)\} \) \((0, 0)\) and \((2, 0)\) are excluded as \( \alpha_k \neq 0 \).

To calculate \( I(t, \alpha, \circ dW^*) \), use the last part of Equation (2.12). We have

\[ I(t, \emptyset, \circ dW^*) = 1. \]

When \( \alpha = 1 \), we have

\[ I(t, 1, \circ dW^*) = \int_0^t \circ dW^*_1(s). \]

Similarly

\[ I(t, 2, \circ dW^*) = \int_0^t \circ dW^*_2(s). \]

When \( \alpha = 0 \), we have

\[ I(t, 0, \circ dW^*) = \int_0^t \circ dW^*_0(s) = \int_0^t ds = t. \]

When \( \alpha = (0, 1) \), we have

\[ I(t, (0, 1), \circ dW^*) = \int_0^t \int_0^{t_2} \circ dW^*_0(t_1) \circ dW^*_1(t_2) = \int_0^t \int_0^{t_2} \circ dt_1 \circ dW^*_1(t_2) = \int_0^t t_2 \circ dW^*_1(t_2). \]

Similarly,

\[ I(t, (0, 2), \circ dW^*) = \int_0^t t_2 \circ dW^*_2(t_2). \]

When \( \alpha = (1, 1) \), we have

\[ I(t, (1, 1), \circ dW^*) = \int_0^t \int_0^{t_2} \circ dW^*_1(t_1) \circ dW^*_1(t_2). \]
Similarly,
\[ I(t, (1, 2), \circ dW^*) = \int_0^t \int_0^2 \circ dW_1^*(t_1) \circ dW_2^*(t_2), \]
\[ I(t, (2, 1), \circ dW^*) = \int_0^t \int_0^2 \circ dW_2^*(t_1) \circ dW_1^*(t_2), \]
\[ I(t, (2, 2), \circ dW^*) = \int_0^t \int_0^2 \circ dW_2^*(t_1) \circ dW_2^*(t_2). \]

**Theorem 8** (The stochastic Stratonovich–Taylor expansion). Assume that \(X(t)\) is the solution to (2.8) and the functions \(V_i(x)\) and \(f(x)\) are \(n+1\) times continuously differentiable. Then
\[
f(X(t)) = \sum_{\alpha \in A_n} (V_{\alpha_0} \cdots V_{\alpha_k}) f(x_0) I(t, \alpha, \circ dW^*) + \sum_{\alpha \in \tilde{A}_n} \int_0^t \int_0^{t_k} \cdots \int_0^{t_0} (V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0)) \circ dW_{\alpha_0}^*(t_0) \circ \cdots \circ dW_{\alpha_k}^*(t_k). \tag{2.13}
\]

**Proof.** We use mathematical induction.

**Induction base** Put \(n = 0\). The expansion (2.13) becomes
\[
f(X(t)) = f(x_0) + \sum_{\alpha_0 = 0}^d \int_0^t (V_{\alpha_0} f)(X(t_0)) \circ dW_{\alpha_0}^*(t_0), \tag{2.14}
\]
which is true by Theorem 6.

**Induction hypothesis** Assume that expansion (2.13) is correct.

**Induction step** We have to prove Equation (2.13) for \(n + 1\), that is,
\[
f(X(t)) = \sum_{\alpha \in \tilde{A}_{n+1}} (V_{\alpha_0} \cdots V_{\alpha_k}) f(x_0) I(t, \alpha, \circ dW^*) + \sum_{(i, \alpha_0, \ldots, \alpha_k) \in \tilde{A}_{n+1}} \int_0^t \int_0^{t_k} \cdots \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW_i^*(s) \circ dW_{\alpha_0}^*(t_0) \circ \cdots \circ dW_{\alpha_k}^*(t_k). \tag{2.15}
\]

In Equation (2.14), replace \(f\) with \((V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0))\). We obtain
\[
(V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0)) = (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) + \sum_{i=0}^d \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW_i^*(s). \tag{2.16}
\]
Substitute this equation to the second term in the right hand side of Equation (2.13). We obtain

\[
\begin{align*}
  f(X(t)) &= \sum_{\alpha \in \mathcal{A}_n} (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) I(t, \alpha, \circ dW^*) \\
  &+ \sum_{\alpha \in \mathcal{A}_n} (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) \int_0^t \int_0^{t_1} \cdots \int_0^{t_k} \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k) \\
  &+ \sum_{\alpha \in \mathcal{A}_n} \sum_{i=0}^d \int_0^t \int_0^{t_0} \cdots \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW^*_i(s) \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k) \\
  &= \int_0^t \int_0^{t_0} \cdots \int_0^{t_0} \left[ (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) + \sum_{i=0}^d \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW^*_i(s) \right] \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k).
\end{align*}
\] (2.17)

When \( \alpha \in \mathcal{A}_n \), there can be two cases.

**Case 1** \( \alpha = (\alpha_0, \ldots, \alpha_k) \in \mathcal{A}_{n+1} \). In this case, the corresponding part in the second term in the right hand side of Equation (2.17) goes to the first term in the right hand side of Equation (2.15). Similarly, the corresponding part in the third term in the right hand side of Equation (2.17) goes to the second term in the right hand side of Equation (2.15).

**Case 2** \( \alpha = (\alpha_0, \ldots, \alpha_k) \in \mathcal{A}_{n+2} \setminus \mathcal{A}_{n+1} \). The corresponding term is

\[
\begin{align*}
  (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) \int_0^t \int_0^{t_0} \cdots \int_0^{t_0} \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k) \\
  &+ \sum_{i=0}^d \int_0^t \int_0^{t_0} \cdots \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW^*_i(s) \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k) \\
  &= \int_0^t \int_0^{t_0} \cdots \int_0^{t_0} \left[ (V_{\alpha_0} \cdots V_{\alpha_k} f)(x_0) + \sum_{i=0}^d \int_0^{t_0} (V_i V_{\alpha_0} \cdots V_{\alpha_k} f)(X(s)) \circ dW^*_i(s) \right] \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k).
\end{align*}
\]

By Equation (2.16), the term becomes

\[
\int_0^t \int_0^{t_0} \cdots \int_0^{t_0} (V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0)) \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k),
\]

and this part goes to the second term in the right hand side of Equation (2.15).

\[\square\]

**Example 6.** We write down the stochastic Stratonovich–Taylor expansion for the Hull–White model for the values of \( n = 0 \).
For \( n = 0 \), we use Equations (2.14) and (2.3) as follows.

\[
\begin{align*}
 f(X(t)) &= f(x_0) + \sum_{a_0=0}^d \int_0^t (V_{a_0} f)(X(t_0)) \circ dW_{a_0}^*(t_0) \\
 &= f(x_0) + \sum_{a_0=0}^2 \int_0^t (V_{a_0} f)(X(t_0)) \circ dW_{a_0}^*(t_0) \\
 &= f(x_0) + \int_0^t (V_0 f)(X(t_0)) \circ dW_0^*(t_0) + \int_0^t (V_1 f)(X(t_0)) \circ dW_1^*(t_0) \\
 &\quad + \int_0^t (V_2 f)(X(t_0)) \circ dW_2^*(t_0) \\
 &= f(x_0) + \int_0^t f = \sum_{j=1}^2 V_j \frac{\partial f}{\partial x_j}(X(t_0)) \circ dW_0^*(t_0) + \int_0^t \sum_{j=1}^2 V_j \frac{\partial f}{\partial x_j}(X(t_0)) \circ dW_1^*(t_0) \\
 &\quad + \int_0^t \sum_{j=1}^2 V_j \frac{\partial f}{\partial x_j}(X(t_0)) \circ dW_2^*(t_0).
\end{align*}
\]

Substituting the values of \( V_i^I(x) \), we obtain

\[
\begin{align*}
 f(X(t),Y(t)) &= f(x_0) + \int_0^t \left[ \left( \mu X(t_0) - \frac{1}{2} X(t_0) Y(t_0) (Y(t_0) + \xi \rho) \right) \frac{\partial f}{\partial X}(X(t_0),Y(t_0)) + \frac{1}{2} \xi Y(t_0) \frac{\partial f}{\partial Y}(X(t_0),Y(t_0)) \right] dt_0 \\
 &\quad + \int_0^t \sqrt{1 - \rho^2 \xi^2 Y(t_0) \left( \frac{\partial f}{\partial X}(X(t_0),Y(t_0)) \right)^2} \circ dW_1^*(t_0) \\
 &\quad + \int_0^t \left[ \rho Y(t_0) \frac{\partial f}{\partial X}(X(t_0),Y(t_0)) + \xi Y(t_0) \frac{\partial f}{\partial Y}(X(t_0),Y(t_0)) \right] \circ dW_2^*(t_0).
\end{align*}
\]

Consider the system (2.5). Introduce the following notation.

\[
J(t, \alpha, dg) = \begin{cases} 
1, & \text{if } \alpha = \emptyset, \\
\int_0^t \int_0^{t_k} \cdots \int_0^{t_2} \circ dg^{\alpha_l}(t_1) \cdots dg^{\alpha_k}(t_k), & \text{otherwise.}
\end{cases}
\]

The Taylor expansion takes the following form.

**Theorem 9 (Taylor expansion).** Assume that \( \bar{X}(t) \) is the solution to (2.5) and the functions \( V_i(x) \) and \( f(x) \) are \( n + 1 \) times continuously differentiable. Then

\[
f(\bar{X}(t)) = \sum_{\alpha \in s_n^I} (V_{\alpha_1} \cdots V_{\alpha_k} f)(x_0) J(t, \alpha, dg)
\]

\[
+ \sum_{\alpha \in s_n^I} \int_0^t \int_0^{t_k} \cdots \int_0^{t_2} (V_{\alpha_0} \cdots V_{\alpha_k} f)(\bar{X}(t_0)) \circ dg^{\alpha_0}(t_0) \cdots dg^{\alpha_k}(t_k).
\]

**Proof.** In Proof of Theorem 8, add overline to \( X(t) \), replace \( I \) with \( J \), delete the circles \( o \), and replace each symbol \( W \) with the symbol \( g \). \( \square \)
Now we are able to formulate the idea of Kusuoka [7].

**Definition 16.** A probability measure $Q$ satisfies the moment matching condition of order $m$ if for all $\alpha$ with $\|\alpha\| \leq m$ we have

$$E^Q[f(1, \alpha, dg)] = E^*[f(1, \alpha, \circ dW^*)].$$

**Remark 3.** It’s time to explain the strange definition of $\|\alpha\|$. Why we calculate zeroes twice? In other words, the task at hand is to prove that;

$$I(t, \alpha, \circ dW^*) = t^\|\alpha\|/2 I(1, \alpha, \circ dW^*).$$

Let us prove it starting by definition, for $\alpha = \emptyset$;

$$I(t, \alpha, \circ dW^*) = t^\|\alpha\|/2 I(1, \alpha, \circ dW^*) \quad 1 = t^0 \cdot 1$$

Now, for $\alpha \neq \emptyset$ we need to make a change of variables, i.e., $t_k = t \cdot s_k, t_{k-1} = t \cdot s_{k-1}, \ldots, t_1 = t_1 \cdot s_1$. As a result;

$$dW^*_\alpha(t_i) = \begin{cases} t \cdot ds_i, & \text{if } \alpha_i = 0, \\ dW^*_\alpha(t_s_i) = \sqrt{t} dW^*_\alpha(s_i), & \alpha_i \neq 0. \end{cases}$$

Finally, the proof for the second case is as follows,

$$I(t, \alpha, \circ dW^*) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_2} \circ dW^*_\alpha(t_1) \circ \cdots \circ dW^*_\alpha(t_k)$$

$$= \int_0^1 \int_0^{s_1} \cdots \int_0^{s_2} t^\|\alpha\|/2 \circ dW^*_\alpha(s_1) \circ \cdots \circ dW^*_\alpha(s_k)$$

$$= t^\|\alpha\|/2 I(1, \alpha, \circ dW^*).$$

It is enough to formulate the moment matching condition for $t = 1$. Then it is automatically true for all positive $t$.

When we approximate $E^*[f(X(t))]$, the price of the contingent claim $f$, by the expected value $E^Q[f(\bar{X}(t))]$, then the absolute approximation error is

$$|E^*[f(X(t))] - E^Q[f(\bar{X}(t))]| = |E^*[R_1(t)] - E^Q[R_2(t)]|,$$

where we denote

$$R_1(t) = \sum_{\alpha \in \mathcal{M}_n} \int_0^t \int_0^{t_1} \cdots \int_0^{t_1} (V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0)) \circ dW^*_{\alpha_0}(t_0) \circ \cdots \circ dW^*_{\alpha_k}(t_k),$$

$$R_2(t) = \sum_{\alpha \in \mathcal{M}_n} \int_0^t \int_0^{t_1} \cdots \int_0^{t_1} (V_{\alpha_0} \cdots V_{\alpha_k} f)(X(t_0)) d^*_{\alpha_0}(t_0) \cdots d^*_{\alpha_k}(t_k).$$

and must be small. Indeed, we have the following result.
Theorem 10. Assume that a probability measure $Q$ satisfies the moment matching condition of order $m$ and the function $f$ has bounded derivatives of all orders. Then, there exists a constant $C$ that depends on $m$ and $f$ such that

$$|E^*[f(X(t))] - E^Q[f(X(t))]| \leq Ct^{(m+1)/2}.$$ 

Proof of this theorem is complicated and may be found in [7].

In the following examples, for simplicity, put $T = 1$.

Example 7. Define

$$g^i(t) = \begin{cases} 
    t, & \text{if } i = 0 \\
    W_i(1)t, & \text{otherwise.}
\end{cases}$$

In other words, let $A \subset \Omega'$ be the following set

$$A = \{ g(t) = (t, y_1t, \ldots, y_dt)^\top: (y_1, \ldots, y_d) \in \mathbb{R}^d \}.$$ 

It is easy to see that the derivatives $g'_i(t)$ are Riemann integrable. The measure $Q$ of a measurable subset $B \subset A$ is

$$Q(B) = P\{(t, W_i(1)t, \ldots, W_d(1)t)^\top \in B\}.$$ 

We prove that this measure satisfies the moment matching condition of order 3, following [13].

For each $0 < i < d$,

$$g^i(t) = W_i(1)t$$

where $g^i(t)$ defines a degree 3 formula. By Gaussian law of symmetry of $W_i(1)t$, if $\alpha$, if $\| \alpha \| = 1$ or 3 we obtain,

$$E^Q[J(1, \alpha, dg)] = E^*[I(1, \alpha, \circ dW^*)] = 0.$$ 

So it is adequate to just to check when $\| \alpha \| = 2$.

If $\alpha_1 \neq \alpha_2$, we then have,

$$E^Q[J(1, \alpha, dg)] = E^*[I(1, \alpha, \circ dW^*)] = 0.$$ 

If $\alpha_1 = \alpha_2 = i \geq 1$, we have by Itô’s formula,

$$J(1, \alpha, dg) = I(1, \alpha, \circ dW^*) = \frac{(W_i(1)t)^2}{2}.$$ 

Example 8. Put $d = 2$. Define

$$g_1(t) = (t, -t, -t)^\top,$$
$$g_2(t) = (t, -t, t)^\top,$$
$$g_3(t) = (t, t, -t)^\top,$$
$$g_4(t) = (t, t, t)^\top.$$ 

and put

$$Q(g_i) = \frac{1}{4}, \quad 1 \leq i \leq 4.$$ 

Lyons and Victoir [9] proved that this measure satisfies the moment matching condition of order 3. In their Table 3, they give 13 functions $g_i$ and 13 probabilities $Q(g_i)$ such that the measure $Q$ satisfies the moment matching condition of order 5.
Theorem 10 tell us that the approximation error is small if $t$ is small and $f$ has bounded derivatives of all orders. What to do if these conditions fail? Assume $t = T$, the time to maturity, is not small. Let $Q$ be a measure on the space $C_{0,AC}([0,1];\mathbb{R}^{d+1})$ satisfying the moment matching condition of order $m$. Let $g_1(t), \ldots, g_n(t)$ be independent elements of the above space chosen at random according to the measure $Q$. Construct a random function $g(t)$ by the following algorithm: if $(i-1)T/n \leq t < iT/n$, then the value of $g(t)$ is

$$g^j(t) = \begin{cases} \frac{1}{n} g_i^0(nt - (i-1)T), & \text{if } j = 0, \\ \frac{1}{\sqrt{n}} g_i^j(nt - (i-1)T), & \text{otherwise}. \end{cases}$$

Note that this function is not continuous at the points $iT/n$. Nevertheless, we have the following result.

**Theorem 11.** There exists a constant $C$ that depends on $m$ and $f$ such that

$$|E^*[f(X(T))] - E^Q[f(X(T))]| \leq \frac{C}{n^{(m-1)/2}}.$$

The case when $f$ does not have bounded derivatives of all orders is more complicated and will not be considered here.
Chapter 3

Applications to the Hull–White Model

In this Chapter, we will apply the theory of Chapter 2 to the Hull–White model. First note the difference between example 7 on the one hand, and example 8 on the other hand. In the latter example, the measure $Q$ is concentrated in finitely many points $g_i$. It follows that the expected value $E^Q$ of a random variable $X$ is given by

$$E^Q[X] = \int_{\Omega'} X(\omega) dQ(\omega) = \sum_{i=1}^{m} X(g_i) P\{g_i\},$$

that is, by a finite sum. Such an expectation is very easy to calculate. In example 7, the measure $Q$ is not concentrated in finitely many points, and we cannot replace the integral with a finite sum. In this case, we would use classical Monte Carlo methods, which will be briefly explained in Section 3.1.

For the modern Monte Carlo method, when the measure $Q$ is concentrated in finitely many points, we apply the method of Example 8 to the Hull–White model in Section 3.2.

3.1 A brief explanation of classical Monte Carlo methods

Monte Carlo method is using stochastic techniques which engages in obtaining the Monte Carlo estimators. It is relating a probability event with a set of outcomes and of which, the probability is regarded as its volume or measure. In our thesis, we do not use this classical method. However, we introduce it shortly as following.

Monte Carlo method was initially developed by Jon von Neumann, Stanislaw Ulam and Nicholas Metroplis in the 1940s during which they were researching on the Manhattan Project, the first nuclear weapons’ project in the laboratory of the Los Alamos National. The method’s name came after the homage to the known Monte Carlo Casino in Monaco, where Ulam’s uncle would always gamble.

Consider the integral of a function $f$ over the unit interval

$$\alpha = \int_{0}^{1} f(x) dx$$
For expectation $E[f(U)]$, where $U$ is uniformly distributed over the interval of 0 and 1. Assume there exists a mechanism of drawing points of $U_1$ to $U_n$, independently and uniformly over $[0,1]$.

Evaluate the function $f$ at $n$ of these random points of $U$. We obtain the Monte Carlo estimate

$$\hat{\alpha}_n = \frac{1}{n} \sum_{i=1}^{n} f(U_i)$$

Then the convergence of the estimate can be described as:

if $\int_0^1 |f(x)| < \infty$, by the strong law of large numbers,

$$P\{ \lim_{n \to \infty} \hat{\alpha}_n = \alpha \} = 1$$

Suppose that $\int_0^1 [f(x)]^2 dx < \infty$. Let

$$\sigma_f^2 = \int_0^1 [f(x) - \alpha]^2 dx$$

Central Limit Theorem tells that the distribution of the error, which is $\alpha_n - \alpha$, is approximately normal distribution with mean 0 and standard deviation $\frac{\sigma_f}{\sqrt{n}}$. We may estimate the unknown $\sigma_f$ using sample standard deviation:

$$s_f = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} [f(U_i) - \hat{\alpha}_n]^2}$$

In which the standard error is in proportion to $\frac{1}{\sqrt{n}}$. In order to cut it twice we are supposed to increase $n$ by a factor of 4.

**European call in the Black–Scholes model example:**

- $t$ current time
- $S(t)$ the price of the stock at time $t$
- $K$ the strike price of a European call option
- $T$ maturity
- $r$ annualised interest rate
- $\sigma$ annualised volatility
- $W^*(t)$ the Wiener process

The stochastic differential equation describing the evolution of the stock price is

$$\frac{dS(t)}{S(t)} = r dt + \sigma dW^*(t)$$
The value of its solution at maturity is

\[ S(T) = S(0) \exp \left( \left[ r - \frac{\sigma^2}{2} \right] T + \sigma W^*(T) \right). \]  

(3.1)

Indeed, put \( X(t) = f(t, S(t)) = \ln S(t) \). By the Itô lemma,

\[ dX(t) = \mu_X(t) \, dt + \sigma_X(t) \, dW^*(t), \]

where

\[ \mu_X(t) = \frac{\partial f}{\partial t}(S(t)) + \frac{\partial f}{\partial S}(S(t)) rS(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(S(t)) \sigma^2 \]

\[ = 0 + \frac{1}{S(t)} rS(t) + \frac{1}{2} \left( -\frac{1}{S^2(t)} \right) \sigma^2 S^2(t) \]

\[ = r - \sigma^2/2 \]

and

\[ \sigma_X(t) = \frac{\partial f}{\partial S}(S(t)) \sigma S(t) = \frac{1}{S(t)} \sigma S(t) = \sigma, \]

that is

\[ dX(t) = (r - \sigma^2/2) \, dt + \sigma \, dW^*(t). \]

Integrate both hand sides from 0 to \( T \). We obtain

\[ \int_0^T dX(t) = \int_0^T (r - \sigma^2/2) \, dt + \int_0^T \sigma \, dW^*(t). \]

All integrals may be easily calculated:

\[ X(T) - X(0) = (r - \sigma^2/2)T + W^*(T), \]

or

\[ \ln S(T) = \ln S(0) + (r - \sigma^2/2)T + W^*(T). \]

Calculating the exponent of both hand sides, we obtain (3.1).

The random variable \( W(T) \) is distributed as \( \sqrt{T} Z \), where \( Z \) is a standard normal random variable. The option’s price can be formed as following:

\[ C = E \left[ e^{-rT} \max\{S(T) - K, 0\} \right] \]

**Asian call in the Black–Scholes model:**

Set \( m \in \mathbb{N}^+ \), and set \( 0 = t_0 < t_1 < \ldots < t_m = T \) as a fixed set of dates. The average price of the stock is

\[ \bar{S} = \frac{1}{m} \sum_{j=1}^{m} S(t_j) \]

The pay-off of the Asian call option is \( \max\{\bar{S} - K, 0\} \). In order to get the simulation value of \( S(t_j) \), the property of Wiener process is introduced. \( W(t_{j+1}) - W(t_j) \) are the increments
that are independently normally distributed with mean 0 and variance $t_{j+1} - t_j$. The following equation 3.2 shows us the straightforward simulating of $S(t_{j+i})$ from $S(t_j)$:

$$S(t_{j+i}) = S(t_j) \exp \left( \left[ r - \frac{\sigma^2}{2} \right] (t_{j+1} - t_j) + \sigma \sqrt{t_{j+1} - t_j} Z_j \right)$$

Efficiency of estimators Let $C_i$ be IID (independent, identically distributed) random variables with $\mathbb{E}[C_i] = C$ and $\text{Var}[C_i] = \sigma^2 < \infty$. The estimator

$$\hat{C}_n = \frac{1}{n} \sum_{i=1}^{n} C_i,$$

As our estimator is unbiased. The sequence of random variables $\frac{\hat{C}_n - C}{\sigma \sqrt{n}}$ converges in distribution to the standard normal distribution by Central Limit Theorem. This implies that for any real $x$ we have

$$\lim_{n \to \infty} \mathbb{P} \left\{ \frac{\hat{C}_n - C}{\sigma \sqrt{n}} \leq x \right\} = \Phi(x).$$

Conclusion: The estimator with less variance is better.

Model discretization error Assume that we do not know the exact solution of the stochastic differential equation (1.2). Then one can divide the time interval $[0, T]$ into small intervals of length $h$ and approximate changes in $S$ by the following recursive formula:

$$S((j+1)h) = S(jh) + rS(jh)h + \sigma S(jh) \sqrt{h} Z_j, \quad 1 \leq j \leq n,$$

where $Z_j$ are independent standard normal random variables, and where $T = nh$. The distribution of the last value $S(nh)$ is not the same as that of $S(T)$ calculated by (1.3). Therefore, $\mathbb{E}[S(nh)] \neq \mathbb{E}[S(T)]$, i.e., the estimator is biased.

discretization error

Suppose that the discrete average for an Asian option is substituted by the continuous average

$$\tilde{S} = \frac{1}{T} \int_0^T S(u) \, du.$$

We approximate the continuous average by the discrete one and obtain a biased estimator.

The standard measure of the performance of an estimator $\hat{\alpha}$ of a quantity $\alpha$ is the mean square error:

$$\mathbb{E}[(\hat{\alpha} - \alpha)^2] = \mathbb{E}[(\mathbb{E}[\hat{\alpha}] - \alpha)^2] + \mathbb{E}[(\hat{\alpha} - \mathbb{E}[\hat{\alpha}])^2].$$

3.2 The method of Lyons and Victoir for the Hull–White model

Lyons and Victoir [9] found the points and coefficients of a cubature formula of degree 5. They are as follows. Let $\omega_1(t)$ and $\omega_2(t)$ be continuous piecewise-linear functions on $[0, 1]$
where the coefficients \( \theta_{ij}, \; 1 \leq i \leq 2, \; 1 \leq j \leq 4 \) are given in [9, Table 2]. The 13 paths \( g_k(t) : [0, 1] \to \mathbb{R}^3 \), and the 13 weights \( \lambda_k \), given by [9, Table 3], describe a cubature formula of degree 5. For example, \( g_2(t) = (t, \omega_1(t), \sqrt{3} \omega_2(t))^\top \) and \( \lambda_2 = \frac{1}{24} \). Similarly, all 13 paths are piecewise-linear, and their derivatives are piecewise constants. Proof of this fact uses sophisticated mathematical tools and lies outside the scope of this thesis. For future use, we write the function \( g_k(t) \) in the form

\[
g_k^j(t) = g_k^j \left( \frac{j-1}{4} \right) + \theta_{ijk} \left( t - \frac{j-1}{4} \right), \quad \frac{j-1}{4} < t \leq \frac{j}{4},
\]

where \( 0 \leq i \leq 2, \; 1 \leq j \leq 4, \; 1 \leq k \leq 13 \). Note that \( \theta_{0jk} = 1 \), because \( g_k^0(t) = t \). The coefficients \( \theta_{ijk} \) with \( i \neq 0 \) can be calculated by combining information from [9, Table 2, Table 3]. For example, \( \theta_{12} = 4 \theta_{1} \) and \( \theta_{22} = 4 \sqrt{3} \theta_{2} \).

Now let us use the method of Lyons and Victoir for the Hull–White model. In the system (2.5), we have \( d = 2, \; \omega^j(s) \) is given by (3.3). The above system takes the form

\[
X_k(t) = X_k \left( \frac{j-1}{4} \right) + \sum_{i=0}^{2} \int_{(j-1)/4}^{t} V_i(X_k(s)) \theta_{ijk} \, ds, \quad (j-1)/4 < t \leq j/4.
\]

The Fundamental Theorem of calculus gives

\[
\frac{dX_k(t)}{dt} = \sum_{i=0}^{2} \theta_{ijk} V_i(X_k(t)),
\]

\[
X_k \left( \frac{j-1}{4} \right) = x_{jk},
\]

where \( x_{jk} = (X(0), Y(0))^\top \). First, we solve this system for \( 0 < t \leq 1/4 \). Then, we calculate \( X_k(1/4) \) and use this value as the initial condition for the next system.

Using the values of \( V_i(x_1, x_2) \) calculated in Section 3.2

\[
V_0(x) = \left( \frac{\mu x_1 - \frac{1}{2} x_1 x_2 (x_2 + \xi \rho)}{v x_2 - \frac{1}{2} \xi^2 x_2} \right), \quad V_1(x) = \left( \sqrt{1 - \rho^2} x_1 x_2 \right), \quad V_2(x) = \left( \frac{\rho x_1 x_2}{\xi x_2} \right),
\]

we get

\[
\frac{dX^1(t)}{dt} = \mu X^1(t) - \frac{1}{2} X^1(t) X^2(t) (X^2(t) + \xi \rho) + \theta_{111} \sqrt{1 - \rho^2} X^1(t) X^2(t) + \theta_{211} \rho X^1(t) X^2(t),
\]

\[
\frac{dX^2(t)}{dt} = v X^2(t) - \frac{1}{2} \xi X^2(t) + \theta_{211} \xi X^2(t),
\]

\[
X^1(0) = X(0), \quad X^2(0) = Y(0),
\]

27
where we put \( j = k = 1 \) and omitted the lower index 1 in the expressions \( X^1_1(t) \) and \( X^2_1(t) \) for simplicity.

We solve the system of two equations. Observe that the equation in the second line is separable. The corresponding initial value problem may be written as

\[
\frac{dX^2(t)}{X^2(t)} = \left( \nu - \frac{1}{2} \xi^2 + \theta_{211} \xi \right) dt,
\]

\[X^2(0) = Y(0).\]

Integrating this system, we obtain

\[
\int_0^t \frac{dX^2(s)}{X^2(s)} = \int_0^t \left( \nu - \frac{1}{2} \xi^2 + \theta_{211} \xi \right) ds
\]

or

\[
\ln(X^2(t)) - \ln(X^2(0)) = \left( \nu - \frac{1}{2} \xi^2 + \theta_{211} \xi \right) t,
\]

By exponentiating both hand sides and using the initial condition, we have

\[
X^2(t) = Y(0) \exp \left[ \left( \nu - \frac{1}{2} \xi^2 + \theta_{211} \xi \right) t \right].
\]

Observe that when \( X^2(t) \) is known, the first equation of our system is also separable. The corresponding initial value problem may be written as

\[
\frac{dX^1(t)}{X^1(t)} = \left( \mu - \frac{1}{2} X^2(t)(X^2(t) + \xi \rho) + \theta_{111} \sqrt{1 - \rho^2 X^2(t)} + \theta_{211} \rho X^2(t) \right) dt,
\]

\[X^1(0) = X(0).\]

Integrate this system:

\[
\int_0^t \frac{dX^1(s)}{X^1(s)} = \int_0^t \left( \mu - \frac{1}{2} X^2(s)(X^2(s) + \xi \rho) + \theta_{111} \sqrt{1 - \rho^2 X^2(s)} + \theta_{211} \rho X^2(s) \right) ds.
\]

It is trivial to show that

\[
\int_0^t X^2(s) ds = Y(0) a^{-1}(e^{at} - 1),
\]

\[
\int_0^t (X^2(s))^2 ds = Y^2(0)(2a)^{-1}(e^{2at} - 1),
\]

where

\[ a = \nu - \frac{1}{2} \xi^2 + \theta_{211} \xi. \]
Our system becomes

$$\ln(X^1(t)) - \ln(X^1(0)) = \mu t - \frac{1}{2} Y^2(0)(2a)^{-1}(e^{2at} - 1)$$

$$+ \left( -\frac{1}{2} \xi \rho + \theta_{111} \sqrt{1 - \rho^2 + \theta_{211} \rho} \right) Y(0) a^{-1}(e^{at} - 1)$$

which gives

$$X^1(t) = X(0) \exp \left[ \mu t - \frac{1}{2} Y^2(0)(2a)^{-1}(e^{2at} - 1) \right.$$

$$\left. + \left( -\frac{1}{2} \xi \rho + \theta_{111} \sqrt{1 - \rho^2 + \theta_{211} \rho} \right) Y(0) a^{-1}(e^{at} - 1) \right].$$

Thus, the solution of the $k$th system on the interval $[0, 1/4]$ has the form

$$\bar{X}^1_k(t) = X(0) \exp \left[ \mu t - \frac{1}{2} Y^2(0)(2a_k)^{-1}(e^{2a_k t} - 1) \right.$$

$$\left. + \left( -\frac{1}{2} \xi \rho + \theta_{111} \sqrt{1 - \rho^2 + \theta_{211} \rho} \right) Y(0) a_k^{-1}(e^{a_k t} - 1) \right],$$

(3.4)

$$\bar{X}^2_k(t) = Y(0) \exp(a_k t),$$

where

$$a_k = v - \frac{1}{2} \xi ^2 + \theta_{211} \xi .$$

Substituting the value of $t = 1/4$, we obtain the updated initial conditions for the next system on the interval $(1/4, 1/2]$, where $j = 2$. The solution of this system is described by Equation (3.4), where $\theta_{11k}$ is replaced with $\theta_{12k}$, $\theta_{21k}$ is replaced with $\theta_{22k}$, and $X(0)$ and $Y(0)$ are replaced with the updated initial conditions. Proceeding in the same way, we finally obtain the values $X_k^1(1)$, $1 \leq k \leq 13$. The approximate price of the option with strike price $K$ is

$$C \approx \sum_{k=1}^{13} \lambda_k \max \{X_k^1(1) - K, 0\}.$$
Chapter 4

Conclusions

In this thesis, we considered a model method of pricing financial derivatives, which uses Cubature on Wiener space. The Wiener space is a space where all continuous functions taking values in a metric space on a given domain. The extension of this model with different maturities which are different than 1, and low maturities are outside our scope of the thesis. During the study of the thesis we have learnt the Hull White model, the Stratonovich Integral, the Moment Matching conditions, Cubature methods on Wiener space, and so on.
Appendix A

Criteria for a Bachelor Thesis

According to the objectives and criteria of Swedish National Agency for Higher Education to Bachelor thesis in mathematics, mathematical statistics, financial mathematics and actuarial science, the authors have gained understanding of the basis of the field of the study through introducing theoretical frameworks with the help of the book written by Kijima [6]. The main subject of the thesis work is the Hull White model by Cubature formula, and the methods that can be applied to this are divided into two parts, one is to use standard Monte-Carlo and the other is to apply the systems of differential equations. The information and the theorems and examples demonstrated in this thesis is gained by comprehensive reading and conclusions from scientific journals, books from the university library, and papers provided by Professor Anatoliy Malyarenko.

Defending the thesis is scheduled on May 31st 2018. All the explanation of understanding the study will be shown and questions may be asked by the examiner.


Index

arbitrage opportunity, 5
Attainable claim, 5

financial engineering, 3
gain process, 4
Hull–White model, 3
Itô formula, 9
Itô integral, 13

leverage effect, 6

market model, 3
moment matching condition, 20

norm, 10
portfolio process, 4
replicating portfolio, 5
risk-neutral measure, 5
self-financing portfolio, 4
simple process, 13
Stratonovich correction, 14
Stratonovich integral, 13
value process, 4