Mathematical Problems in the Theory of Incomplete Markets

by

David Saunders

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy, Graduate Department of Mathematics, University of Toronto

© by David Saunders (2001)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-63700-X
Mathematical Problems in the Theory of Incomplete Markets
Ph.D. 2001
David Saunders
Department of Mathematics
University of Toronto

Abstract

We study some mathematical problems that arise in studying the financial mathematics of incomplete markets. In the first chapter, we consider the pricing of derivative securities when markets are incomplete. In this case, the prices of all derivatives cannot be uniquely determined based on only the absence of arbitrage. We propose a method for determining the value of a derivative to a given investor based on the dual variables of the investor's utility maximization problem. We also derive results on the stability of these prices with respect to changes in the model inputs.

In the second chapter, we study some properties of the mean-reverting Ornstein-Uhlenbeck process that appears in many places in applied mathematics, and in particular in financial modelling. We study the estimation of the parameters of the model based on a constrained maximum likelihood problem (a hybrid of method of moments estimation and the method of maximum likelihood). We also derive the distribution of the range of the process over an interval.

The third and fourth chapters study the application of hidden Markov models to financial modelling. In particular, we study two dimensional stochastic differential equations driven by Brownian motion with one hidden variable. In the third chapter, we study the case where only the drift coefficient of the observed variable depends on the hidden variable, while in the fourth chapter we allow the diffusion coefficient
of the observed variable to depend on the hidden variable. In each chapter, we study the problem under the following headings: marginal distributions of the observed process, estimation of the path of the hidden process based on observations, and estimation of the parameters of the hidden process.
Acknowledgments

I would like to thank my thesis advisor, Prof. Luis Seco, for his guidance and patience, and many stimulating discussions. I would also like to thank the members of RiskLab, for their support and many interesting discussions. In particular, I would like to thank Dr. Gustavo Comezaña, Barindar Sandhu, and Dr. Stathis Tompaidis.

I am also grateful to Prof. I. Graham for support and helpful advice. I would also like to thank Prof. D. Dawson and Prof. J. Quastel for many helpful discussions on probability (in particular regarding the material from chapter 4).

During the third year of my graduate studies, I was a visiting student at the Centre for Financial Research at the University of Cambridge. I would like to thank the members of the Centre, and in particular Prof. Michael Dempster for his support and hospitality during this time.

While a graduate student, I have also had the benefit of conducting research on the mathematics of risk management in an industrial setting, as a member of the quantitative research group at Algorithmics Inc. I would like to thank the members of the research group, in particular Drs. O. Croissant, R. Dembo, J. Farvolden, I. Iscoe, A. Kreinin, A. Lakhany, H. Mausser and L. Merkolovitch for many interesting and fruitful discussions. I owe a particular debt of gratitude to Dr. Dan Rosen, for his advice and support throughout my studies.

I would like to thank the administrative staff of the University of Toronto department of mathematics, especially Ida Bulat, the Graduate Secretary for all her help during the last five years of my studies here.

I thank my family, and Ruth Basa for their love, support, and patience during the time I was working on my thesis.
Finally, I gratefully acknowledge the financial support of the National Sciences and Engineering Research Council of Canada, the Government of Ontario, Algorithmics Inc., and my supervisor through his research grants.
To my parents.


## Contents

1 Convex Duality and Derivatives Pricing 5
   1.1 Introduction ........................................ 5
   1.2 Motivation ........................................ 6
   1.3 Utility Maximization and Derivatives Pricing .......... 10
      1.3.1 The Probabilistic Setup ....................... 10
      1.3.2 The Market Model .......................... 10
      1.3.3 The Fundamental Theorem .................. 13
      1.3.4 Utility Maximization and the Absence of Arbitrage .... 14
      1.3.5 Pricing Derivative Securities ............ 18
   1.4 Utility Invariant Pricing of Derivative Securities .... 19
      1.4.1 Sensitivity Results ...................... 23
   1.5 Related Results .................................. 27

2 Mean-Reverting Processes in Finance 33
   2.1 Introduction ...................................... 33
   2.2 Estimation ....................................... 35
      2.2.1 Constrained Maximum Likelihood ............ 37
      2.2.2 Results .................................. 38
   2.3 The Distribution of the Range .................... 44
      2.3.1 Method for Computing the Range ............. 45
      2.3.2 Solution of the Partial Differential Equation .. 48

1
2.3.3 Double Lookback Options ............................................. 51

3 Stochastic Drift Hidden Markov Models .................................. 53
  3.1 Introduction .................................................................. 53
    3.1.1 The Model and Assumptions ..................................... 53
  3.2 The Marginal Distributions of Y ...................................... 55
  3.3 Filtering .................................................................. 60
  3.4 Parameter Estimation ................................................ 60
  3.5 Examples .................................................................. 70

4 Stochastic Volatility Hidden Markov Models .......................... 77
  4.1 Introduction ................................................................ 77
    4.1.1 The Model and Assumptions ..................................... 78
  4.2 Marginal Distributions of $Y_t$ ....................................... 79
  4.3 Filtering .................................................................. 82
    4.3.1 Discretization and Limit Theorems ............................ 84
    4.3.2 Results ................................................................ 85
    4.3.3 Nonzero Drift ...................................................... 94
  4.4 Parameter Estimation ................................................ 107

A Results from Optimization ................................................ 111
  A.1 Results from Optimization .......................................... 111
  A.2 Stability in Nonlinear Programming ............................... 113

B Results from Stochastic Analysis ......................................... 119
  B.1 Results from Stochastic Analysis ................................. 119
    B.1.1 Itô's Lemma ....................................................... 119
  B.2 Central Limit Theorems .............................................. 120
  B.3 Large Deviations for Martingale Difference Sequences .... 122
  B.4 Quadratic Variation .................................................. 122
  B.5 Ergodic Theorems ..................................................... 123
Chapter 1

Convex Duality and Derivatives Pricing

1.1 Introduction

In this chapter, we present some new results on the relationship between duality in convex optimization and the pricing of derivative securities in incomplete markets. The fact that there exists a connection between derivative pricing and optimization duality has been known for some time (see, e.g. [Duf96]). In the study of complete markets, the existence of a unique martingale measure for pricing all derivative securities (under the assumption of the absence of arbitrage) makes this connection less important. More recently, since the study of incomplete markets has come to the forefront of research in mathematical finance, optimization duality has been seen to play a vital role. For a sample of the more recent results, see [JCW01],[Kal98a], [Kal00], [KS98], and [KS99b].
1.2 Motivation

In [DR99] and [DRSO0], a simple, practical portfolio optimization model was considered, which we review here for the purpose of motivation. All technical terms will be made precise in the following sections. We work on a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $|\Omega| = M$ and $\mathbb{P}(\{\omega_i\}) = p_i > 0$ for $i = 1, \ldots, M$.

Assume that investors construct their portfolios at time 0 and hold them until time $T$, with no intermediate rebalancing. Also assume that there are $N$ securities to be traded in the market at prices $q_j$, $j = 1, \ldots, N$. Let the value of security $j$ if scenario $\omega_i$ prevails at the horizon be $D_{ij}$, and consider the matrix $D = \{D_{ij}\} \in \mathbb{R}^{M \times N}$. Then if an investor holds $x_j$ units of security $j$ the initial cost of her portfolio will be $q \cdot x$ and its value under scenario $\omega_i$ at time $T$ will be $(Dx)_i$.

We also assume the existence of a special security, referred to as the benchmark, with initial price $c$, which has value $\tau_i$ at time $T$ under scenario $\omega_i$. The investor is taken to measure risk by expected underperformance with respect to the benchmark (also referred to as regret), which is given by

$$R^-(x; \tau) = p \cdot (Dx - \tau)_-$$

$$= p \cdot ((\tau - Dx) \vee 0)$$

Benefit is measured in terms of expected overperformance of the benchmark (also referred to as reward):

$$R^+(x; \tau) = p \cdot (Dx - \tau)_+$$

$$= p \cdot ((Dx - \tau) \vee 0)$$

If we assume that the investor can spend no more than the price of the benchmark in constructing her portfolio, and is subject to trading restrictions that can be written as a set of linear constraints $Ax \leq b$ then her utility maximization problem becomes,
for a given level of risk aversion $\lambda \geq 1$

$$\max_{x, y^+, y^-} \ p \cdot y^+ - \lambda p \cdot y^- \quad (1.5)$$

$$Dx - y^+ + y^- = \tau \quad (\psi)$$

$$q \cdot x \leq c \quad (\alpha)$$

$$Ax \leq b \quad (\theta)$$

$$y^+, y^- \geq 0$$

which is a linear program. We observe for future reference that the dual of this linear program is

$$\min_{\phi, \alpha, \theta} \ \tau \cdot \psi + \alpha c + b \cdot \theta \quad (1.6)$$

$$D^T x + \alpha q + A^T \theta = 0 \quad (x)$$

$$p \leq -\psi \leq \lambda p \quad (y^+, y^-)$$

$$\alpha, \theta \geq 0$$

In a complete market, the absence of arbitrage is sufficient to determine the prices of all derivative securities and investor preferences are irrelevant. In the case where the market is incomplete (e.g. $M > N$), it is no longer possible to obtain unique prices for derivatives based solely on the absence of arbitrage (it is possible to derive bounds on derivatives prices using arbitrage arguments, but these bounds are often too loose to be of use in practice, see [KQ95],[SSC95]). It is therefore appropriate to ask what is the value of a particular derivative security for the investor represented by the utility maximization problem (1.5).

In [DRS00], it was proposed that the appropriate value of a derivative security to an investor should be the one at which she would be indifferent to trading in it. That is, it is the value at which, if it were introduced into the market, it would not affect the construction of her optimal portfolio, assuming that the prices of all other
securities remain constant. In, particular, this price is in some way the "worst" one for the investor. At a higher price she might sell the derivative, and at a lower price she might buy it, in both cases increasing her utility. However, when the derivative security is traded at this "utility invariant" price, the optimal portfolio (and therefore maximum utility) remains unchanged.

We now show how this "utility invariant" price can be calculated using the dual variables of the linear program (1.5). Assume that the market is opened for trading in a new security worth \( d_e \) under scenario \( \omega_e \) at the horizon. If the security is offered for trade at the price \( v \), then the utility maximization problem (1.5) becomes:

\[
\begin{align*}
\max_{x,z,y^+,y^-} & \quad p \cdot y^+ - \lambda p \cdot y^- \\
Dx + zd - y^+ + y^- &= \tau \\
q \cdot x + zu &\leq c \\
Ax &\leq b \\
y^+, y^- &\geq 0
\end{align*}
\]

where we have assumed that there are no restrictions on how investors can trade in the new security. The dual of the above problem is:

\[
\begin{align*}
\min_{\psi, \alpha, \theta} & \quad \tau \cdot \psi + \alpha c + b \cdot \theta \\
D^T \psi + \alpha q + A^T \theta &= 0 \\
d \cdot \psi + \alpha u &= 0 \\
p &\leq -\psi \leq \lambda p \\
\alpha, \theta &\geq 0
\end{align*}
\]

Suppose that (1.5) has an optimal solution. Then by the duality theorem of linear

\footnote{This assumption is not strictly correct from the viewpoint of economic equilibria, as the introduction of a new security to a market will likely change the prices of all traded securities, see [BW01].}
programming (see [Chv83]), the optimal solutions of the original utility maximization problem (1.5) and its dual (1.6) must satisfy

\[ p \cdot y^+ - \lambda p \cdot y^- = \tau \cdot \psi + \alpha c + b \theta \]  
\[ Dx - y^+ + y^- = \tau \]  
\[ q \cdot x \leq c \]  
\[ Ax \leq b \]  
\[ D^T x + \alpha q + A^T \theta = 0 \]  
\[ p \leq -\psi \leq \lambda p \]  
\[ y^+, y^-, \alpha, \theta \geq 0 \]  

while the optimal solutions to the problem (1.7) and its dual (1.8) where trading is allowed in the new instrument must satisfy

\[ p \cdot y^+ - \lambda p \cdot y^- = \tau \cdot \psi + \alpha c + b \theta \]  
\[ Dx + zd - y^+ + y^- = \tau \]  
\[ q \cdot x + vz \leq c \]  
\[ Ax \leq b \]  
\[ D^T x + \alpha q + A^T \theta = 0 \]  
\[ d \cdot \psi + \alpha v = 0 \]  
\[ p \leq -\psi \leq \lambda p \]  
\[ y^+, y^-, \alpha, \theta \geq 0 \]  

It is clear that the same portfolio \( x \) will satisfy both systems, with \( z = 0 \) in the second system (holding none of the new security is optimal) if \( \alpha > 0 \) and

\[ v = \frac{-\psi \cdot d}{\alpha} \]  

9
Therefore, the above equation gives an utility invariant price of the derivative security with payoff \( d \) for the investor whose utility maximization problem is (1.5). Notice that utility invariant prices of all derivative securities can be calculated after solving the utility maximization problem once, since they involve only simple manipulations with the security's payoff and the optimal dual variables of the problem (1.5).

1.3 Utility Maximization and Derivatives Pricing

In this section, we review some known results on the relationship between utility maximization and derivatives pricing in incomplete markets. The material for the exposition is drawn from [Duf96], [Kal98a], [Kal00], [MR98].

1.3.1 The Probabilistic Setup

Let \((\Omega, \mathcal{F})\) be a finite probability space. More specifically, \(\Omega\) is a finite set, with \(|\Omega| = M\), and \(\mathcal{F}\) is the \(\sigma\)-algebra consisting of all subsets of \(\Omega\). Let \(\mathbb{P}\) be a probability measure on \((\Omega, \mathcal{F})\). We denote the elements of \(\Omega\) by \(\omega_i, i = 1, \ldots, M\), and write \(p_i = \mathbb{P}(\{\omega_i\})\). For our probabilistic setup, we can define a simple filtration, with \(\mathcal{F}_0\) being the trivial \(\sigma\)-algebra, and \(\mathcal{F}_T = \mathcal{F}\), for some \(T > 0\). We denote the set of times \(\mathcal{T} = \{0, T\}\). Financially, this amounts to considering a problem with a single horizon, where investors construct their portfolios today and then hold them until the horizon date without rebalancing. We denote the filtration by \(\{\mathcal{F}_t\}\).

1.3.2 The Market Model

Definition 1. A security is a real-valued \(\{\mathcal{F}_t\}\) adapted stochastic process.

Definition 2. A derivative security is a \(\mathcal{F}_T\) measurable random variable.

Definition 3. A market is a pair \(\mathcal{M} = (S, \Phi)\) where \(S\) is an \(\mathbb{R}^N\) valued adapted stochastic process, and \(\Phi\) is a set of \(\mathbb{R}^N\)-valued predictable processes. In this case, we refer to \(N\) as the dimension of the market.
Processes $\varphi \in \Phi$ are referred to as trading strategies. If $(S, \Phi)$ is a market, then for any $\varphi \in \Phi$, since $S$ is adapted and $\varphi$ is predictable, the process $\varphi_t \cdot S_t, t \in \{0, T\}$ is adapted (a security).

**Remark 1.** Our simple filtration means that investors buy portfolios at time 0 and hold them until time T. Since predictable processes are constant under this filtration, we make no distinction between the vector $\varphi \in \mathbb{R}^N$ and the predictable $\mathbb{R}^N$ valued process $\varphi_0 = \varphi_T = \varphi$.

**Definition 4.** A security $X$ is an arbitrage if either of the following two conditions holds

1. $X_0 < 0$ and $X_T \geq 0$ a.s. $\mathbb{P}$

2. $X_0 \leq 0$, $X_T \geq 0$ a.s. and $\mathbb{P}(\{X_T > 0\}) > 0$

Thus, an arbitrage is an opportunity to earn a risk-free profit.

**Definition 5.** A market $\mathcal{M} = (S, \Phi)$ is said to be arbitrage-free if there does not exist $\varphi \in \Phi$ such that $\varphi \cdot S$ is an arbitrage.

**Definition 6.** A market $(S, \Phi)$ is said to be complete if for every derivative security $Y$, there exists $\varphi \in \Phi$ such that

$$\varphi \cdot S_T = Y$$

$\mathbb{P}$ almost surely.

That is, the payoff of any derivative security can be replicated exactly by a trading strategy $\varphi$.

We consider the problem of an investor seeking to maximize her utility over a prescribed horizon. This is formalized by the following.
Definition 7. An investor is a triple \((u, B, Z)\) where \(B \in \mathbb{R}\), \(Z\) is an \(\mathcal{F}_T\) measurable random variable and

\[ u : \Omega \times \mathbb{R} \rightarrow [-\infty, \infty) \]  

(1.26)

where \(x \rightarrow u(\omega, x)\) is a proper, concave function (in the sense of [Roc70]) for any \(\omega \in \Omega\). We refer to \(u\) as the investor's utility function, \(B\) as her budget, and \(X\) as her endowment.

Remark 2. In spite of the terminology, we do not assume that the investor's endowment must be positive. Furthermore, the utility functions above are not assumed to be increasing or differentiable. It is typical that \(u(\omega, x)\) will be independent of \(\omega\) (so that the utility is a function of the current level of wealth). The greater generality is required to deal with some hedging problems. For example, in section (1.2), the nondifferentiable utility functions

\[ u(\omega, x; \lambda) = x^+ - \lambda x^- \]

(1.27)

\[ = (x \vee 0) - \lambda \cdot (x \wedge 0) \]

(1.28)

were considered (this is equivalent to the utility function in section (1.2) if we assume that the investor is constrained to hold \(-1\) units of the benchmark security, or if the endowment is taken to be \(-\tau\)). Here \(\lambda \geq 1\) is a risk aversion parameter which measures the relative value of profit and loss to the investor. With \(|\Omega| < \infty\) and an affine set of trading constraints \(A\phi \leq b\) the investor's utility maximization problem
given endowment $0$ and budget $B$ becomes

$$\max_x \mathbb{E}[u(x)]$$

$$\varphi \cdot S_T = x$$

$$\varphi \cdot S_0 \leq B$$

$$A\varphi \leq b$$

which is a linear programming problem.

An important feature that causes incompleteness in many market models is the presence of trading constraints (classical models assume that investors can hold arbitrary amounts of any security; this is not a realistic, or insignificant, assumption).

**Definition 8.** A set of trading constraints is a convex set $K \subseteq \mathbb{R}^N$. The market induced by the set of trading constraints $K$ is the market $(S,\Phi_K)$ where $\Phi_K$ is the set of all trading strategies such that $(\varphi^1, \ldots, \varphi^N) (\omega) \in K$ for all $(\omega, t) \in \Omega \times T$.

**Remark 3.** Under the identification of simple processes with vectors in $\mathbb{R}^N$, the set $\Phi_K$ is identified with $K$.

**Assumption 1.** Throughout the remainder of this chapter, we will assume that the set $K$ can be expressed as all $\varphi \in \mathbb{R}^N$ such that

$$g_j(\varphi) \leq 0 \quad j = 1, \ldots, K_g$$

$$h_i(\varphi) = 0 \quad i = 1, \ldots, K_h$$

where $g_j : \mathbb{R}^N \rightarrow \mathbb{R}$ are convex functions for $j = 1, \ldots, K_g$ and $h_i$ are affine, $i = 1, \ldots, K_h$.

### 1.3.3 The Fundamental Theorem

**Definition 9.** Let $\mathcal{M} = (S, \mathbb{R}^N)$ be a market with dimension $N$. A vector $\psi \in \mathbb{R}_{++}^M$ is said to be a state-price vector for $\mathcal{M}$ if and only if $S^i_T \cdot \psi = S^i_0$ for $i = 1, \ldots, N$. 

The following is often called the fundamental theorem of asset pricing.

**Theorem 1.** Consider a market $\mathcal{M} = (S, \mathbb{R}^N)$.

- $\mathcal{M}$ is arbitrage free if and only if there exists a state-price vector (for $\mathcal{M}$).

- Suppose that $\mathcal{M}$ is arbitrage free. Then it is complete if and only if the state-price vector is unique (i.e. the set of all state-price vectors for $\mathcal{M}$ is a singleton).

**Proof.** The proof of this theorem is an application of the Separating Hyperplane Theorem from Convex Analysis. For a complete proof, see [Duf96] or [HP81].

### 1.3.4 Utility Maximization and the Absence of Arbitrage

Let $\mathcal{M} = (S, \mathbb{R}^N)$ be a market, and consider the investor $(u, Z, B)$ with $Z \in \mathbb{R}_+, B = 0$ and (with slight abuse of notation)

$$u(\omega, x) = u(x)$$ (1.35)

where $u : \mathbb{R} \to \mathbb{R}$ strictly increasing. Let $U : \mathbb{R}^M \to \mathbb{R}$ be defined by $U(c) = \sum_{i=1}^M p_i u(c_i)$. Interpreting $Z$ as an agent’s endowment and $U$ as her utility function leads to the following concave program.

$$\max_{c, \varphi} U(c) \quad (1.36)$$

$$Z(\omega_i) + \varphi_i \cdot S_T(\omega_i) = c_i \quad i = 1, \ldots, M$$

$$c_i \geq 0 \quad i = 1, \ldots, M$$

$$\varphi \cdot S_0 \leq 0$$

where the maximization is done over all $\varphi \in \Phi = \mathbb{R}^N$, and we have included a "no bankruptcy" constraint that terminal wealth must be positive under all possible scenarios ($c \geq 0$).
Assumption 2. We assume that there exists $\varphi \in \Phi$ such that $\varphi \cdot S_T(\omega_i)$ is positive for all $i$ and strictly positive for at least one $i$.

Proposition 1. Consider a market $\mathcal{M} = (S, \mathbb{R}^N)$. If the problem (1.36) has a solution, then there is no arbitrage. If there is no arbitrage and $U$ is continuous, the problem (1.36) has a solution.

Proof. See [Duf96].

When the utility function $U$ is differentiable, then there is an intimate connection between state price vectors and the gradient of $U$ at optimality. It is this relationship that we will exploit later when pricing derivative securities in incomplete markets.

Theorem 2. Let $\mathcal{M} = (S, \mathbb{R}^N)$ be a market. Suppose that $c^* \in \mathbb{R}^M_+$ solves (1.36), and that $\nabla U(c^*)$ exists and is strictly positive. Then there exists $\lambda \in \mathbb{R}$ such that $\lambda \nabla U(c^*)$ is a state-price vector.

Proof. See [Duf96].

Utility functions are often assumed to be concave, strictly increasing and differentiable (although we have already seen an example of a model where these assumptions don’t apply). In these circumstances, $\nabla U$ will be strictly positive and we have the following corollary.

Corollary 1. Let $\mathcal{M} = (S, \mathbb{R}^N)$ be a market. If $U$ is concave and differentiable, then $c^* \in \mathbb{R}^M_+$ is an optimal solution of (1.36) if and only if $\lambda \nabla U(c^*)$ is a state-price vector for some $\lambda > 0$.

Proof. See [Duf96].
Models with Trading Constraints and Counterexamples

Let \((S, \Phi)\) be a market and \((u, Z, B)\) an investor. Again let \(U : \mathbb{R}^M \to \mathbb{R}\) be defined as \(U(c) = \sum_{i=1}^{M} p_i u(c_i, \omega_i)\). The investor's utility maximization problem is then

\[
\max_{\varphi \in \Phi} U(c) \\
\varphi \cdot S_T(\omega_i) + Z(\omega_i) = c_i \\
\varphi \cdot S_0 \leq B
\]  

(1.37)

**Proposition 2.** Suppose that \(U\) is strictly increasing, \(\Phi\) is a convex cone, and that there exists an optimal solution to (1.37), then the market \((M, \Phi)\) is arbitrage-free.

**Proof.** Suppose to the contrary that \(\varphi^*, c^*\) is an optimal solution to (1.37) and that \(\tilde{\varphi} \in \Phi\) is such that \(\varphi \cdot S\) is an arbitrage. Then \(\varphi^* + \tilde{\varphi} \in \Phi\) (since \(\Phi\) is a convex cone) and

\[
(\varphi^* + \tilde{\varphi}) \cdot S_0 \leq B \\
\tilde{c} = (\varphi^* + \tilde{\varphi}) \cdot S_T > \varphi^* \cdot S_T = c^*
\]

Therefore \(U(\tilde{c}) > U(c^*)\), contradicting the assumption of optimality.

The assumption that \(\Phi\) is a convex cone allows, for example, the inclusion of constraints on short sales in the model.

**A Counterexample**

Now suppose that \(\Phi\) is not a cone. Then it is possible for there to be an optimal solution to (1.37) and yet for there still to be arbitrage. Consider the case where \(Z = 0\), and \(M = N = 1\). Suppose \(S_0 = B = 0\) and \(S_T = 1\), and observe that \(S\) so defined is an arbitrage. Let \(U(y) = y\), and \(\Phi = [0, 1]\). Then the utility maximization
problem simplifies to:

\[
\max_{\varphi \geq y} \quad \varphi = y \\
0 \leq \varphi \leq 1
\]

which clearly has optimal solution \( \varphi^* = 1 \). This simple example illustrates the more general rule that if there is an arbitrage in the portfolio selection model, then investors will hold as much of the arbitrage portfolio as they can (until they hit the boundary of \( \Phi \)).

Another Counterexample

We observe that, unlike for the model with positive endowment, the absence of arbitrage is not enough to guarantee the existence of an optimal solution to (1.37) in general. Consider a model with \( U(y) = 2y_1 + y_2 \), \( B = 0 \), \( M = N = 2 \), with

\[
S_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}
\]

\[
S_T(\omega_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\]

\[
S_T(\omega_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]
In this case, there is no arbitrage even if we take $\Phi = \mathbb{R}^M$, since $(0.5, 0.5)'$ is a state-price vector. The utility maximization problem reduces to:

$$\max_{\varphi, y} U(y) = 2y_1 + y_2$$

$$0.5\varphi_1 = y_1$$

$$0.5\varphi_2 = y_2$$

$$0.5\varphi_1 + 0.5\varphi_2 \leq 0$$

Clearly for this model $\varphi_1 = n, \varphi_2 = -n$ is feasible, and has objective value

$$U(y) = 2n - n = n$$

for all $n \in \mathbb{N}$ and therefore the problem is unbounded, despite the absence of arbitrage.

1.3.5 Pricing Derivative Securities

We now briefly consider the pricing of derivative securities.

**Definition 10.** Let $X$ be a derivative security, and let $(S, \Phi)$ be a market. We say that $\varphi \in \Phi$ replicates $X$ if

$$\varphi \cdot S_T = X$$

almost surely $\mathbb{P}$.

**Proposition 3.** Let $(S, \mathbb{R}^N)$ be a market with no arbitrage, and $X$ a derivative security. Suppose that both $\varphi_1 \in \Phi$ and $\varphi_2 \in \Phi$ replicate $X$. Then $\varphi_1 \cdot S_0 = \varphi_2 \cdot S_0$

**Proof.** Suppose $\varphi_1 \cdot S_0 > \varphi_2 \cdot S_0$. Then $\varphi_2 - \varphi_1$ is an arbitrage. 

Thus when markets are complete, we can determine the unique price of a derivative security, as the price of the corresponding replicating portfolio. For further details,
see [HK79], [HP81], and [MR98].

1.4 Utility Invariant Pricing of Derivative Securities

In this section we present a series of new results relating optimization duality to pricing in incomplete markets.

We begin by considering the basic utility maximization problem of the investor $(u, Z, B)$:

\[
\max_{\varphi \in \Phi} \sum_{k=1}^{M} p_k u(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) 
\]

\(\varphi \cdot S_0 \leq B\) \hspace{1cm} (1.40)

Here we assume that \(\Phi \subseteq \mathbb{R}^N\) is defined to be all \(\varphi \in \mathbb{R}^N\) such that:

\[g_j(\varphi) \leq 0 \hspace{1cm} i = 1, \ldots, K_g\]
\[h_i(\varphi) = 0 \hspace{1cm} i = 1, \ldots, K_h\]

Where \(g_j : \mathbb{R}^N \rightarrow \mathbb{R}\) are differentiable convex functions, and \(h_i : \mathbb{R}^N \rightarrow \mathbb{R}\) are affine. In this case, the problem (1.40) becomes a concave program, and we can use the theory of duality for concave programming to define a fair price for any derivative security.

Prices of derivative securities in incomplete markets cannot be uniquely determined by the absence of arbitrage. They will instead depend on investor preferences. We propose that a fair price of a derivative security for a given investor is one which will make the investor indifferent to trading the claim. That is, if adding a security \((q, X)\) with payoff equal to that of the claim and price \(q\) will not change the investor's
optimal portfolio, then we refer to \( q \) as an *utility invariant* price of \( X \). Related ideas for pricing in incomplete markets have been discussed in [Dav97] and [DRS00].

**Definition 11.** Consider the constrained utility maximization problem (1.40). Suppose \( \varphi^* \) is an optimal solution to (1.40), and let \( X \) be a derivative security. Then we say that \( q \) is an *utility invariant price* for \( X \) if \( \varphi^* \) and \( \varphi_n^* = 0 \) solve:

\[
\max_{\varphi^* \in \Phi, \varphi_n \in \mathbb{R}} \sum_{k=1}^{M} \sum_{k=1}^{M} p_k u(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) \tag{1.42}
\]

\[
\varphi \cdot S_0 + q \varphi_n \leq B
\]

**Theorem 3.** Let \((S, \Phi)\) be a market, \((u, Z, B)\) an investor with \( u(x, \omega) \) differentiable in \( x \) for each \( \omega \in \Omega \) and \( X \) a derivative security. Suppose that \( \varphi^* \) is an optimal solution to (1.40). Then there exists a utility invariant price \( q^* \) for the investor \((u, Z, B)\).

**Proof.** The Lagrangian for the problem (1.40) is:

\[
L(\varphi) = -\sum_{k=1}^{M} p_k u(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) + \alpha(\varphi \cdot S_0 - B) + \sum_{j=1}^{K_d} \lambda_j g_j(\varphi) + \sum_{i=1}^{K_h} \mu_i h_i(\varphi) \tag{1.43}
\]

By the necessity part of theorem 22 from appendix A \( \varphi^* \) satisfies

\[
-\sum_{k=1}^{M} \nabla_\varphi (u(Z(\omega_k) + \varphi^* S_T(\omega_k), \omega_k)) = \alpha S_0 + \sum_{j=1}^{K_d} \lambda_j \nabla_\varphi g_j(\varphi^*) + \sum_{i=1}^{K_h} \mu_i \nabla_\varphi h_i(\varphi^*) = 0 \tag{1.44}
\]
In the case of the extended problem (1.42) the Lagrangian becomes

\[
L(\varphi, \varphi_n) = -\sum_{k=1}^{M} p_k u(Z(\omega_k) + \varphi \cdot S_T(\omega_k) + \varphi_n X(\omega_k), \omega_k)
+ \alpha(\varphi \cdot S_0 + \varphi_n q - B) + \sum_{j=1}^{K_g} \lambda_j g_j(\varphi) + \sum_{i=1}^{K_h} \mu_i h_i(\varphi)
\]

(1.48)

By the sufficiency part of (theorem 22 from appendix A), we have that \((\varphi^*, \varphi_n^*)\) is an optimal solution to (1.42) if the following system is satisfied:

\[
-\sum_{k=1}^{M} p_k u(Z(\omega_k) + \varphi \cdot S_T(\omega_k) + \varphi_n^* X(\omega_k), \omega_k)
+ \alpha S_0 + \sum_{j=1}^{K_g} \lambda_j \nabla_\varphi g_j(\varphi^*) + \sum_{i=1}^{K_h} \mu_i \nabla_\varphi h_i(\varphi^*) = 0
\]

(1.49)

\[
-\frac{d}{d\varphi_n} \sum_{k=1}^{M} u(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) + \alpha q = 0
\]

(1.50)

\[
g_j(\varphi^*) \leq 0 \quad j = 1, \ldots, K_g
\]

(1.51)

\[
h_i(\varphi^*) = 0 \quad i = 1, \ldots, K_h
\]

(1.52)

\[
\lambda_j g_j(\varphi^*) = 0 \quad j = 1, \ldots, K_g
\]

(1.53)

for some \(\alpha \in \mathbb{R}_+, \lambda \in \mathbb{R}_{+}^{K_g}, \mu \in \mathbb{R}^{K_h}\).
But this system is satisfied by \( \varphi^* \) from (1.44) and \( \varphi_n^* = 0 \) if

\[
\alpha q^* = \frac{d}{d\varphi_n} \sum_{k=1}^{M} p_k u\left(Z(\omega_k) + \varphi^* \cdot S_T(\omega_k) + \varphi_n^* X(\omega_k)\right)
\]

(1.54)

\[
= \sum_{k=1}^{M} p_k u'(Z(\omega_k) + \varphi^* \cdot S_T(\omega_k))X(\omega_k)
\]

(1.55)

By definition, this means that \( q^* \) given above is an utility invariant price for the derivative security \( X \) and the investor \((u, Z, B)\).

\[\square\]

**Theorem 4.** Under the hypotheses of theorem 3, suppose that \( u(\cdot, \omega_k) \) is strictly concave, differentiable at \( Z(\omega_k) + \varphi \cdot S_T(\omega_k) \), \( k = 1, \ldots, M \) and \( g_j, j = 1, \ldots, K_g \) are differentiable and the set

\[
\{\nabla_{\varphi} g_j(\varphi^*), \nabla_{\varphi} h_i(\varphi^*), S_0|j \in I(\varphi^*), i = 1, \ldots, K_h, \}
\]

(1.56)

is linearly independent, where

\[
I(\varphi^*) = \{j|g_j(\varphi^*)\} = 0
\]

(1.57)

Also, suppose that the Lagrange multiplier \( \alpha \) from (1.50) is strictly positive. Then the utility invariant price is unique.

**Proof.** Strict concavity implies that if \( \varphi^*, \bar{\varphi} \) are two optimal solutions to (1.40) then
they must satisfy \( \varphi^* \cdot S_T(\omega_k) = \varphi \cdot S_T(\omega_k) \), \( k = 1, \ldots, M \). But then

\[
\frac{\partial}{\partial \varphi_j} \sum_{k=1}^{M} u(Z(\omega_k) + \varphi^* \cdot S_T(\omega_k), \omega_k) = \sum_{k=1}^{M} u'(Z(\omega_k) + \varphi^* \cdot S_T(\omega_k), \omega_k) \cdot S_T^j(\omega_k) \]

\[
= \sum_{k=1}^{M} u'(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) \cdot S_T^j(\omega_k) \]

\[
= \frac{\partial}{\partial \varphi_j} \sum_{k=1}^{M} u(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) \]

Thus the linear independence assumption implies that the Lagrange multipliers \( \lambda, \mu \) and in particular \( \alpha \) are unique. The uniqueness of the utility invariant price is therefore implied immediately by (1.54) and the fact that \( \alpha > 0 \).

Remark 4. If we define a new probability measure \( Q \) on \((\Omega, \mathcal{F})\) by

\[
\frac{dQ}{dp}(\omega_k) = \frac{u'(Z(\omega_k) + \varphi^* \cdot S_T(\omega_k), \omega_k)}{\sum_{i=1}^{M} p_i u'(Z(\omega_i) + \varphi^* \cdot S_T(\omega_i), \omega_i)}
\]

then the utility invariant price of \( X \) is given by

\[
q^* = kE_Q[X]
\]

where \( k = \alpha^{-1} \sum_{k=1}^{M} p_k u'(Z(\omega_k) + \varphi \cdot S_T(\omega_k), \omega_k) \)

1.4.1 Sensitivity Results

In this section, we consider the problem of determining the sensitivity of utility invariant prices with respect to perturbations of the model inputs. Specifically, we
consider the model:

\[
\max_{\varphi \in \Phi_\varepsilon} \sum_{k=1}^{M} p_k(\varepsilon) u(Z(\omega_k, \varepsilon) + \varphi \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon)
\]

(1.64)

\[\varphi \cdot S_0(\varepsilon) \leq B(\varepsilon)\]

(1.65)

where \(\varepsilon \in \mathbb{R}^{K_\varepsilon}\). We assume that the set of trading strategies \(\Phi_\varepsilon\) has the form:

\[
g_j(\varphi, \varepsilon) \leq 0 \quad j = 1, \ldots, K_g
\]

(1.66)

\[
h_i(\varphi, \varepsilon) = 0 \quad i = 1, \ldots, K_h
\]

(1.67)

**Assumption 3.** For each fixed \(\varepsilon \in \mathbb{R}^{K_\varepsilon}\)

1. \((u(\cdot, \cdot, \varepsilon), Z(\cdot, \varepsilon), B(\varepsilon))\) is an investor (i.e. for each \(\omega \in \Omega\), \(u(\cdot, \omega, \varepsilon)\) is a proper concave function, \(Z(\cdot, \varepsilon)\) is an \(\mathcal{F}_T\) measurable random variable, and \(B(\varepsilon) \in \mathbb{R}\)).

2. \(p_k(\varepsilon) > 0, S_0(\varepsilon) \in \mathbb{R}^n, \) and \(S_T(\cdot, \varepsilon)\) is \(\mathcal{F}_T\) measurable (so that \(S^i(\cdot, \varepsilon)\) is a security, \(i = 1, \ldots, n\)).

3. \(g_j(\cdot, \varepsilon) : \mathbb{R}^n \to \mathbb{R}\) is a convex function for \(j = 1, \ldots, K_g\) and \(h_i(\cdot, \varepsilon) : \mathbb{R}^n \to \mathbb{R}\), \(i = 1, \ldots, K_h\) are affine.

In order to derive sensitivity results for the problem (1.64) we also need the following smoothness assumptions on the coefficients. Let \(\varphi^*(\varepsilon_0)\), denote the optimal solution (if it exists) of (1.64) at \(\varepsilon = \varepsilon_0\).

**Assumption 4.**

1. \(u, g_j, j = 1, \ldots, K_g\) are twice continuously differentiable in \(\varphi\).

2. \(g_j, h_i, p_k, B, Z, S_0, S_T, \nabla_\varphi u, \nabla_\varphi g_j, \nabla_\varphi h_i\) are once continuously differentiable in \(\varepsilon\) in a neighbourhood of \((\varphi^*(0), 0), j = 1, \ldots, K_g, i = 1, \ldots, K_h\).
• Let \( I(\varphi, \varepsilon_0) = \{ j | g_j(\varphi, \varepsilon_0) = 0 \} \). The set of vectors
\[
\{ \nabla_\varphi g_j(\varphi^*(0), 0), \nabla_\varphi h_i(\varphi^*(0), 0), S_0(0) \} | j \in I(\varphi^*(0), 0), i = 1, \ldots, K_h \}
\] (1.68)
is linearly independent.

• Let
\[
U(\varphi, \varepsilon) = \sum_{k=1}^{M} p_k(\varepsilon) u(Z(\omega_k, \varepsilon)) + \varphi \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon)
\] (1.69)
then for all \( z \neq 0 \) such that
\[
\nabla_\varphi g_j(\varphi^*(0), 0) \cdot z \leq 0 \quad j \in I(\varphi^*(0), 0) \quad (1.70)
\]
\[
\nabla g_j(\varphi^*(0), 0) \cdot z = 0 \quad \forall j \text{ s.t. } \alpha_j^*(0) > 0 \quad (1.71)
\]
\[
\nabla_\varphi h_i(\varphi^*(0), 0) \cdot z = 0 \quad i = 1, \ldots, K_h \quad (1.72)
\]
we have that
\[
\langle z, H_\varphi L z \rangle > 0
\] (1.73)
(here \( H_\varphi L \) denotes the Hessian of \( L \) with respect to \( \varphi \)).

• Let \( \alpha^*(0), \lambda_j^*(0), \mu_i^*(0) \) be the optimal Lagrange multipliers of the problem (1.64) at \( \varepsilon = 0 \) (where we use the notation from the previous section, and the optimal Lagrange multipliers exist and are unique by the above assumptions and theorem 22). Then \( \alpha^*(0) > 0, \lambda_j^*(0) > 0, j \in I(\varphi^*(0), 0) \) (i.e. strict complementary slackness holds for problem (1.64) at \( \varepsilon = 0 \)).

**Theorem 5.** Let \( X \) be a derivative security and define
\[
q(\varepsilon) = \sum_{k=1}^{M} p_k(\varepsilon) \frac{\partial u}{\partial \varepsilon}(Z(\omega_k, \varepsilon)) + \varphi^*(\varepsilon) \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon)) \frac{X(\omega_k)}{\alpha(\varepsilon)}
\] (1.74)
where \( \alpha(\varepsilon_0) \) is the optimal Lagrange multiplier of the constraint (1.65) at \( \varepsilon = \varepsilon_0 \) (if it exists). Then \( q(\varepsilon) \) is continuously differentiable and

\[
\frac{\partial q}{\partial \varepsilon_i} = \frac{1}{\alpha(\varepsilon)} \left( \sum_{k=1}^{M} X(\omega_k) \left( \frac{\partial p_k}{\partial \varepsilon_i}(\varepsilon) \frac{\partial u}{\partial \omega_k}(\omega_k, \varepsilon) + \varphi^*(\varepsilon) \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon) \right.
\]

\[
+ p_k(\varepsilon) \frac{\partial^2 u}{\partial \omega_k^2}(\omega_k, \varepsilon) + \varphi^*(\varepsilon) \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon)
\]

\[
\times \left( \frac{\partial Z}{\partial \varepsilon_i}(\omega_k, \varepsilon) + \sum_{j=1}^{n} \left( \frac{\partial \varphi_j^*(\varepsilon)}{\partial \varepsilon_i} S_{T_j}(\omega_k, \varepsilon) + \varphi_j^*(\varepsilon) \frac{\partial S_{T_j}}{\partial \varepsilon_i}(\omega_k, \varepsilon) \right) \right)
\]

\[
+ p_k(\varepsilon) \frac{\partial u}{\partial \omega_k}(\omega_k, \varepsilon) + \varphi^*(\varepsilon) \cdot S_T(\omega_k, \varepsilon), \omega_k, \varepsilon) \right) \right)
\]

\[
= -\frac{1}{\alpha^2(\varepsilon)} \cdot \frac{\partial \alpha}{\partial \varepsilon_i} \left( \sum_{k=1}^{M} p_k(\varepsilon) \frac{\partial u}{\partial \omega_k}(\omega_k, \varepsilon) + \varphi^*(\varepsilon) \cdot S_T(\omega_k, \varepsilon) \right) X(\omega_k) \right) (1.75)
\]

for \( i = 1, \ldots, K_\varepsilon \)

Proof. The existence and uniqueness of optimal Lagrange multipliers, \( \alpha(0), \lambda_j(0)\mu_i(0), j = 1, \ldots, K_g, i = 1, \ldots, K_h \) follows from the linearly independence assumption and theorem 22. In fact, applying theorem 23, we have that there exists an \( \varepsilon^* > 0 \) such that the optimal solution \( \varphi^*(\varepsilon) \) exists and is unique for all \( \varepsilon \in [0, \varepsilon^*] \) and furthermore, there exist optimal Lagrange multipliers \( \alpha(\varepsilon), \lambda_1(\varepsilon), \ldots, \lambda_{K_g}(\varepsilon), \mu_1(\varepsilon), \ldots, \mu_{K_h}(\varepsilon) \) for each \( \varepsilon \in [0, \varepsilon^*] \) and that the map

\[
\varepsilon \rightarrow (\varphi^*(\varepsilon), \alpha(\varepsilon), \lambda_1(\varepsilon), \ldots, \lambda_p(\varepsilon), \mu_1(\varepsilon), \ldots, \mu_q(\varepsilon)) \quad (1.76)
\]

is continuously differentiable, with \( \alpha(\varepsilon) > 0 \) and \( \lambda_j(\varepsilon) > 0 \) for all \( j \in I(\varphi^*(0)) \) and all \( \varepsilon \in [0, \varepsilon^*] \). The expression for the derivative now follows by elementary calculus. \( \square \)
1.5 Related Results

It has recently come to our attention that the notion presented above and in [DRS00] under the name "utility invariant pricing" has been considered in a similar context by Kallsen in [Kal98a] (called there "utility neutral pricing"). We briefly summarize these results here, pointing out the similarities and differences between the two approaches.

Once again we consider a finite probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Now the set of times is finite: \(T = \{0, 1, \ldots, T\}\), and we consider a filtration \(\{\mathcal{F}_t\}_{t \in T}\). We observe that in this setting, investors are able to rebalance their portfolios at a discrete set of times. The definitions of a security, a trading strategy, a derivative security, and a market remain the same. We consider a fixed market \((S, \Phi)\) with dimension \(N + 1\).

**Definition 12.** A trading strategy \(\varphi \in \Phi\) is self-financing if:

\[
\varphi_t \cdot S_t = \varphi_{t+1} \cdot S_t
\]

for \(t = 1, \ldots, T\).

**Assumption 5.** We assume that \(S_t^1 > 0\) almost surely \(\forall t \in T\)

**Definition 13.** A set of trading constraints is a set of the form

\[
K' = \eta + K
\]

where \(\eta \in \mathbb{R}^N\) and \(K = \{\psi \in \mathbb{R}^N | g_j \cdot \psi \leq 0, j = 1, \ldots, p\ \text{and} \ g_j \cdot \psi = 0, j = p + 1, \ldots, q\}\), for some \(g_j \in \mathbb{R}^N, j = 1, \ldots, q\).

We note in particular that there are no restrictions on the holdings of \(S_t^1\), and that all constraint functions are affine.

27
Definition 14. If $K'$ is a set of trading constraints, then the market induced by $K'$ is the market $(S, \Phi_{K'})$ where $\Phi_{K'}$ is the set of all self-financing trading strategies such that $\varphi_t(\omega) \in K', \forall (\omega, t) \in \Omega \times \mathcal{T}$.

Definition 15. A consumption process $C$ is a real-valued adapted process.

Definition 16. An investor is a pair $(X, u)$ where $X$ is an adapted $\mathbb{R}$ valued stochastic process and

$$u : \Omega \times \{0, \ldots, T\} \times \mathbb{R} \to [-\infty, \infty]$$

(1.79)

where

- $(\omega, t) \to u(\omega, t, x)$ is a predictable $\mathbb{R}$ valued process for every $x \in \mathbb{R}$.

- $x \to u_t(\omega, x)$ is a proper, upper-semicontinuous concave function (in the sense of [Roc70]) for any $(\omega, t) \in \Omega \times \{0, \ldots, T\}$.

Given a market $(S, \Phi)$, an investor $(X, u)$, and a consumption process $C$, we define the corresponding set of discounted processes by

$$\hat{S} = \frac{S}{S_0}$$

(1.80)

$$\hat{u}_t(x) = u_t(S_t^1 x)$$

(1.81)

$$\hat{X}_t = \frac{X_0}{S_0} + \sum_{s=1}^{t} \frac{1}{S_s^1} \Delta X_s$$

(1.82)

$$\hat{C}_t = \frac{C_0}{S_0} + \sum_{s=1}^{t} \frac{1}{S_s^1} \Delta C_s$$

(1.83)

where for any stochastic process $Y$, we denote $\Delta Y_t = Y_t - Y_{t-1}$. For any self-financing
trading strategy \( \varphi \), we also consider the discounted value process

\[
\tilde{V}_t(X, C, \varphi) = \frac{1}{S_t} V_t(X, C, \varphi)
\]

\[
= \tilde{X}_t - \tilde{C}_t + \sum_{s=1}^{t} \varphi_s \cdot \Delta \tilde{S}_s \quad t = 0, \ldots, T
\]

**Definition 17.** Let \( \rho \) be a real-valued random variable and \( U \) an adapted process. Then we say that \( U \) is a \( \rho \)-**martingale** if

\[
\mathbb{E}[\rho(U_t - U_s) | \mathcal{F}_s] = 0
\]

for any \( s, t \in \{0, \ldots, T\} \) with \( s \leq t \). \( \rho \)-supermartingales and \( \rho \)-submartingales are defined analogously.

**Definition 18.** A real-valued random variable \( \rho \) is called a \( K \)-martingale density if \( \sum_{s=1}^{t} \varphi_s \cdot \Delta S_s \) is a \( \rho \)-martingale for any \( \varphi \) in the market induced by \( K \). (\( K \)-supermartingale and \( K \)-submartingale densities are defined analogously).

**Definition 19.** A probability measure \( \mathbb{Q} \sim \mathbb{P} \) is called a \( K \)-martingale measure if \( \frac{d\mathbb{Q}}{d\mathbb{P}} \) is a \( K \)-martingale density. (\( K \)-supermartingale and \( K \)-submartingale measures are defined analogously).

Consider the market \((S, \Phi_{K'})\) induced by the set \( K' \) and let \( \tilde{C} = \sum_{s=0}^{t} \) be a consumption process and \( \tilde{\varphi} \in \Phi_{K'} \) a trading strategy. Three notions of optimality are considered in [Kal98a]. We only discuss one, which is similar in spirit to the work presented in this thesis.

**Definition 20.** \((\tilde{C}, \tilde{\varphi})\) is \( \alpha \)-**optimal** if it maximizes

\[
\mathbb{E} \left[ \sum_{s=0}^{T} u_s(c_s) \right]
\]
over all consumption processes $C. = \sum_{s=0}^\infty c_s$ and all $\varphi \in \Phi_{K'}$ that satisfy

$$V_T(X, C, \varphi) = 0 \quad P \text{ a.s.} \quad (1.88)$$

**Definition 21.** Consider the market $(S, \Phi_{K'})$ and let $\check{C}. = \sum_{s=0}^\infty \hat{c}_s$ be a consumption process and $\check{\varphi} \in \Phi_{K'}$ a trading strategy. We say that $(\check{C}, \check{\varphi})$ is $\beta$-optimal if it maximizes

$$E \left[ \sum_{s=0}^T \hat{u}_s(\hat{c}_s) \right] \quad (1.89)$$

over all consumption processes $\check{C}. = \sum_{s=0}^\infty \hat{c}_s$ and all $\varphi \in \Phi_{K'}$ that satisfy

$$\check{C}_T = \check{X}_T + \sum_{s=1}^T \varphi_s \cdot \Delta \hat{S}_s \quad P \text{ a.s.} \quad (1.90)$$

The above utility maximization problems are dual to the following problem, in a sense that will be made precise in the following.

**Definition 22.** Let $\check{C}. = \sum_{s=0}^\infty \check{c}_s$ be a consumption process. Define $\check{X} = \check{X} - \eta \cdot (\check{S} - \check{S}_0)$ Let $\check{\rho}$ be a $K$-supermartingale density. We say that $(\check{\rho}, \check{C})$ is $\gamma$-optimal if $\check{\rho}$ minimizes

$$\sup \left\{ E \left[ \sum_{s=0}^T \hat{u}_s(\hat{c}_s) \right] \right\} \quad (1.91)$$

over all $\check{C} = \sum_{s=0}^\infty \check{c}_s$ consumption processes with

$$E[\rho(\check{C}_T - \check{X}_T)] \leq 0 \quad (1.92)$$

over all $K$-supermartingale densities $\check{\rho}$ and if, for $\check{\rho}$, the supremum is attained at $\check{C}$.

**Theorem 6.** Let $C$ be a consumption process and $\varphi$ a trading strategy in the market $(S, \Phi_{K'})$ induced by $K'$. Suppose that $E[\sum_{s=0}^T \hat{u}_s(\hat{c}_s)] \neq -\infty$, where $C. = \sum_{s=0}^\infty c_s$
and \( \hat{C} = \sum_{s=0}^{\gamma} \hat{c}_s \) then the following are equivalent:

- \((C, \varphi)\) is \(\alpha\)-optimal.

- \((\hat{C}, \varphi)\) is \(\beta\)-optimal.

- There exists a \(K'\)-supermartingale density \(\rho^*\) such that
  1. \((\rho^*, \hat{C})\) is \(\gamma\)-optimal.
  2. \(\hat{C}_T = \hat{X}_T + \sum_{s=1}^{T} \varphi_s \cdot \Delta \hat{S}_s\).

- There exists a \(K''\)-supermartingale density \(\rho\) such that
  1. \(\mathbb{E}[\rho|F_t] \in \partial \hat{u}_t(\hat{c}_t)\) a.s. for \(t = 0, \ldots, T\)
  2. \(\hat{C}_T = \hat{X}_T + \sum_{s=1}^{T} \varphi_s \cdot \Delta \hat{S}_s\).
  3. \(\sum_{s=1}^{T} (\varphi_s - \eta) \cdot \Delta \hat{S}_s\) is a \(\rho\)-martingale.

**Proof.** See [Kal98a].

**Assumption 6.** We assume that there are given \(r\) derivative securities denoted by \(U^k\) for \(k = N + 1, \ldots, N + r + 1\).

We denote \(\tilde{K'} = K' \times \mathbb{R}^r\). If \(S^{N+1}, \ldots, S^{N+r}\) are adapted processes, then \(\bar{\alpha}\)-optimality, etc. are defined analogously as above, and relative to the market induced by \(\tilde{K}'\). Identifying \((\varphi^1, \ldots, \varphi^{N+1}) \leftrightarrow (\varphi^0, \ldots, \varphi^{N+1}, 0, \ldots, 0)\) we can consider \(\Phi_{K'}\) as a subset of \(\tilde{\Phi}_{K'}\) (where there are no constraints on \(S^{N+1}, \ldots, S^{N+r+1}\)).

**Definition 23.** Adapted processes \(S^{N+1}, \ldots, S^{N+r}\) are called \(\alpha\)-neutral processes for the derivatives \(U^k\), \(k = N + 2, \ldots, N + r + 1\) if

1. \(S^k_T = U^k\) a.s. \(\mathbb{P}\) for \(k = N + 2, \ldots, N + r + 1\).

2. There exists an \(\alpha\)-optimal pair \((C, \varphi)\) which is \(\bar{\alpha}\)-optimal for the market \((S^1, \ldots, S^{N+r+1}), \tilde{\Phi}\).
Theorem 7. Suppose that $\dot{u}_T(\cdot)$ is strictly increasing on its domain and assume that there exists an $\alpha$-optimal pair $(C, \varphi)$ such that the expected utility is not $-\infty$. Then $\rho$ in theorem 6 equals $E[\rho] \frac{dP^*}{dP}$ for some $K$-supermartingale measure $P^* \sim P$. Also,

1. There exist $\alpha$-neutral processes $S^{N+2}, \ldots, S^{N+r+1}$ for the derivatives $U^k$, $k = N + 2, \ldots, N + r + 1$ such that $\dot{S}^{N+2}, \ldots, \dot{S}^{N+r+1}$ are $P^*$-martingales and the market $((S^1, \ldots, S^{N+r+1}), \Phi)$ induced by $\tilde{K}$ is arbitrage free.

2. If $\dot{u}_T(\cdot)$ is differentiable on its domain, then there exist (up to indistinguishability) no further $\alpha$-neutral processes.

Proof. See [Kal98a].
Chapter 2

Mean-Reverting Processes in Finance

2.1 Introduction

In this chapter, we study some properties of a stochastic process that arises frequently in the mathematics of finance. In particular, we are interested in the mean-reverting Ornstein-Uhlenbeck process, which is the solution of the stochastic differential equation

\[ dX_t = \alpha(\beta - X_t)dt + \sigma dW_t \quad X_0 = \xi \]  

(2.1)

where \( W_t \) is a standard Brownian motion, \( \xi \) is independent of \( W \) and \( \alpha, \beta, \sigma > 0 \). We study two particular problems for this process. The first is the estimation of its parameters based on discrete observations and knowledge of the stationary distribution. The second is the calculation of the distribution of the range of the process.

The following is a list of well-known properties of the Ornstein-Uhlenbeck process that we use in the following sections. We remark that a great deal is known about
this process and the list below is very far from exhaustive.

**Property 1.** The stochastic differential equation (2.1) has the explicit strong solution

\[ X_t = \beta + e^{-\alpha t}(X_0 - \beta) + \sigma \int_0^t e^{-\alpha(t-s)}dW_s \]  

(2.2)

**Property 2.** If \( \xi \) is Gaussian with mean \( E(\xi) \) and variance \( \text{var}(\xi) \), \( X \) is a Gaussian process, with mean \( M_t \) and covariance \( C(s, t) \) given by

\[ M_t = \beta + e^{-\alpha t}(E(\xi) - b) \]  

(2.3)

\[ C(s, t) = \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|} + e^{-\alpha(t+s)}(\text{var}(\xi) - \frac{\sigma^2}{2\alpha}) \]  

(2.4)

**Property 3.** The Ornstein-Uhlenbeck process \( X \) has a unique stationary distribution \( \mu \) where

\[ \mu(A) = \sqrt{\frac{\alpha}{\pi\sigma^2}} \int_A \exp \left( -\frac{\alpha(x - \beta)^2}{\sigma^2} \right) dx \]  

(2.5)

\( \forall A \in \mathcal{B}(\mathbb{R}) \) i.e. the stationary distribution is normal, with mean \( \beta \) and variance \( \frac{\sigma^2}{2\alpha} \).

**Property 4.** The transition densities of the process are known explicitly:

\[ p_t(x, y) = \sqrt{\frac{\alpha}{\sigma^2(1 - e^{-2\alpha t})}} \exp \left( -\frac{\alpha(y - \beta - \exp(-\alpha t)(x - \beta))^2}{\sigma^2(1 - \exp(-2\alpha t))} \right) \]  

(2.6)

**Property 5.** \( X \) is a regular diffusion (see [RW00b]), with speed measure \( m \) and scale function \( s \) given by

\[ m(dx) = \frac{2}{\sigma^2} \exp \left( \frac{-\alpha x^2}{\sigma^2} \right) dx \]  

(2.7)

\[ s(x) = \int_0^x \exp \left( \frac{\alpha y^2}{\sigma^2} \right) dy \]  

(2.8)
Property 6. X is ergodic. Furthermore, for any $\Delta > 0$, the process $X_{t\Delta}$ is ergodic. (This is true for any such discretization of an ergodic Markov process, so long as $\lim_{t\to\infty} \| p_t(x, y) - \mu \| = 0$ see [GCJL00]). For the Ornstein-Uhlenbeck process this can be proved directly, using properties 3 and 4, or by using property 5 and appealing to ([RW00b], theorem 54.5), as in [GCJL00].

2.2 Estimation

We review the case where the process can be observed continuously. In this case, (theorem 28 from appendix B) allows $\sigma^2$ to be calculated based on the observation of the sample path. We therefore consider the problem of estimating the drift coefficient, given $\sigma$ known. For simplicity, we set $\sigma = 1$.

Let $D^1$ be the Skorokhod space space of right continuous functions with left limits (see [JS87] chapter 6, or [EK86] chapter 3), and let $P_{a, \beta}$ be the measure on $D^1$ corresponding to particular values of $\alpha, \beta$, (see, for example [EK86]). By Girsanov's theorem, we have the likelihood ratio:

$$L(\alpha, \beta, T) = \left. \frac{dP_{a, \beta}}{dP_{0,0}} \right|_{\mathcal{F}_T} = \exp \left( \int_0^T \alpha (\beta - X_t) dX_t - \frac{1}{2} \int_0^T \alpha^2 (\beta - X_t)^2 dt \right)$$

(2.9)

The problem of maximizing this ratio becomes much easier if we perform the reparametrization $a = \alpha \beta, b = \alpha$, in which case the above model becomes:

$$dX_t = a - bX_t dt + dW_t$$

(2.10)

and we have, with a slight abuse of notation

$$L(a, b, T) = \left. \frac{dP_{a, b}}{dP_{0,0}} \right|_{\mathcal{F}_T} = \exp \left( \int_0^T (a - bX_t) dX_t - \frac{1}{2} \int_0^T (a - bX_t)^2 dt \right)$$

(2.11)
so that

\[ I_T(a,b) = \log L(a,b,T) \]

\[ = \int_0^T (a - bX_t)dX_t - \frac{1}{2} \int_0^T (a - bX_t)^2 dt \]

\[ = a(X_T - X_0) - b \int_0^T X_t dX_t - \frac{1}{2} \int_0^T (a^2 - 2abX_t + b^2X_t^2) dt \]

\[ = a(X_T - X_0) - b \int_0^T X_t dX_t - \frac{a^2T}{2} + ab \int_0^T X_t dt - \frac{b^2}{2} \int_0^T X_t^2 dt \]

and we have

\[ \frac{\partial I}{\partial a} = (X_T - X_0) - aT + b \int_0^T X_t dt \]

\[ \frac{\partial^2 I}{\partial a^2} = -T \]

\[ \frac{\partial I}{\partial b} = - \int_0^T X_t dX_t + a \int_0^T X_t dt - b \int_0^T X_t^2 dt \]

\[ \frac{\partial^2 I}{\partial b^2} = - \int_0^T X_t^2 dt \]

\[ \frac{\partial^2 I}{\partial a \partial b} = \int_0^T X_t dt \]

Any candidate for a maximizer of the log-likelihood will satisfy the necessary condition \( \frac{\partial I}{\partial a} = \frac{\partial I}{\partial b} = 0 \). Which yields estimates:

\[ \hat{b}_T = \frac{T \int_0^T X_t dX_t - (X_T - X_0) \int_0^T X_t dt}{\left( \int_0^T X_t dt \right)^2 - T \int_0^T X_t^2 dt} \]

\[ \hat{a}_T = \frac{X_T - X_0 - \hat{b} \int_0^T X_t dt}{T} \]

To see that these are in fact maximizers, we prove that the log-likelihood is concave. Let \( H \) be the Hessian of \( I \). Then to prove that \( I \) is concave, it is sufficient to prove
that \( G = -H \) is positive definite. But, from above:

\[
G = \begin{pmatrix}
-\frac{\partial^2 f}{\partial x^2} & -\frac{\partial f}{\partial x} \\
-\frac{\partial f}{\partial x} & -\frac{\partial^2 f}{\partial y^2}
\end{pmatrix} = \begin{pmatrix}
T & -\int_0^T X_t dt \\
-\int_0^T X_t dt & \int_0^T X_t^2 dt
\end{pmatrix}
\]  

(2.23)

To show that \( G \) is positive definite, it is sufficient that \( G_{11} > 0 \) (which is obvious) and \( |G| > 0 \) (see [AR86]). But

\[
|G| = T \int_0^T X_t^2 dt - \left( \int_0^T X_t dt \right)^2
\]  

(2.24)

which is positive by Jensen's inequality. Thus, the parameters of the OU process can be estimated via maximum likelihood. Asymptotic results about the distribution of the maximum likelihood estimators can be found in [Rao99b].

### 2.2.1 Constrained Maximum Likelihood

In this section, we investigate a hybrid approach towards estimating the parameters of the stochastic process \( X \). Since \( X \) is ergodic, by observing the sample path we should be able to estimate the parameters of the stationary distribution, which is normal with mean \( \beta \) and variance \( \frac{\sigma^2}{2a} \). Thus based on observations of \( X \) we know \( \beta \), and a function of \( \sigma \) and \( \alpha \). The process for \( X \) is still ergodic even when observed only at a discrete set of times \( k\Delta, k \in \mathbb{N} \) (property 6), so observation at discrete intervals will still yield consistent estimates of the mean and variance of the stationary distribution. Based on these observations, we want estimates of \( \sigma \) and \( \alpha \). We investigate the estimation of these parameters by maximizing the likelihood, given the information already known about the stationary distribution. This leads to a constrained maximum likelihood problem for the (discretely sampled) stochastic differential equation. This constrained likelihood problem only has a solution for certain sample paths, and the existence of a maximum likelihood estimate is related to the condition for the existence of (unconstrained) maximum likelihood estimates.
for the discretized Markov chain $X_{k\Delta}$. Constrained likelihood problems have been studied in [DW96], [Hey97], and [Sch96].

We assume that the process is observed discretely at equally spaced times $k\Delta$, $k = 1, \ldots, n$, that the process $X$ is started in its stationary distribution (i.e. $\xi \sim N(0, \frac{\sigma^2}{2\alpha})$).

By the properties of the Ornstein-Uhlenbeck process listed at the beginning of the chapter, we can determine the log-likelihood function for $X_{k\Delta}$ explicitly. We assume that we already know the parameters of the stationary distribution (which is $N(\beta, \frac{\sigma^2}{2\alpha})$). Letting $Y_{i\Delta} = X_{i\Delta} - \beta$ for ease of notation, we have the log-likelihood function:

\[
\ell(\alpha, \sigma) = -\frac{1}{2} \log(\pi \sigma^2 (1 - e^{-2\Delta \alpha}) \alpha^{-1}) - \left( \frac{\sum_{i=1}^{n} (Y_{i\Delta} - e^{-\alpha \Delta} Y_{i-1\Delta})^2 \alpha}{\sigma^2 (1 - e^{-2\alpha \Delta})} \right) \tag{2.25}
\]

We observe that the unconstrained maximum likelihood estimates for this problem are known and are:

\[
\hat{\alpha}_n = \frac{1}{\Delta} \log \left( \frac{\sum_{i=1}^{n} Y_{i-1\Delta} Y_{i\Delta}}{\sum_{i=1}^{n} Y_{i-1\Delta}^2} \right) \tag{2.26}
\]

\[
\hat{\sigma}_n = -\frac{2 \hat{\alpha}_n}{n (1 - e^{-2\Delta \alpha})} \sum_{i=1}^{n} (Y_{i\Delta} - Y_{i-1\Delta} e^{\Delta \alpha})^2 \tag{2.27}
\]

see [Rao99b]. Note that taking the logarithm in (2.26) is possible if and only if $\sum_{i=1}^{n} Y_{i-1\Delta} Y_{i\Delta} > 0$.

### 2.2.2 Results

Knowing the stationary distribution gives us the constraint $\frac{\sigma^2}{\alpha} = c$, with $c$ known. Using this constraint, we can reduce the log-likelihood to a function of only $\alpha$.

\[
\ell(\alpha) = -\frac{n}{2} \log(2\pi c (1 - e^{-2\Delta \alpha})) - \left( \frac{\sum_{i=1}^{n} (Y_{i\Delta} - e^{-\alpha \Delta} Y_{i-1\Delta})^2}{2c (1 - e^{-2\alpha \Delta})} \right) \tag{2.28}
\]
We want to know if this function has a maximum. We know a priori that the domain for the parameter $\alpha$ is $(0, \infty)$, so let us first examine the behaviour of the log-likelihood at the boundaries of this domain.

**Proposition 4.** Consider the function $l : (0, \infty) \to \mathbb{R}$ defined by

$$l(\alpha) = -\frac{n}{2} \log(2\pi c(1 - e^{-2\Delta})) - \left( \frac{\sum_{i=1}^{n}(Y_i - e^{-\alpha Y_{i-1}})^2}{2c(1 - e^{-2\Delta})} \right)$$  \hspace{1cm} (2.29)

and assume that the sequence $Y_i$ is not constant, then

$$\lim_{\alpha \to -\infty} l(\alpha) = \log \left( \left( \frac{1}{2\pi c} \right)^{\frac{n}{2}} \exp \left( -\sum_{i=1}^{n} \frac{Y_i^2}{2c} \right) \right)$$  \hspace{1cm} (2.30)

and

$$\lim_{\alpha \to 0^+} l(\alpha) = -\infty$$  \hspace{1cm} (2.31)

**Proof.** We begin by examining the behaviour as $\alpha \to \infty$. By considering each of the terms in $l$ separately that

$$\lim_{\alpha \to \infty} l(\alpha) = \lim_{\alpha \to \infty} -\frac{n}{2} \log(2\pi c(1 - e^{-2\Delta})) - \left( \frac{\sum_{i=1}^{n}(Y_i - e^{-\alpha Y_{i-1}})^2}{2c(1 - e^{-2\Delta})} \right)$$  \hspace{1cm} (2.32)

$$= -\frac{n}{2} \log(2\pi c) - \left( \frac{\sum_{i=1}^{n}(Y_i)^2}{2c} \right)$$  \hspace{1cm} (2.33)

So that as $\alpha \to \infty$, we have that the likelihood tends to:

$$\left( \frac{1}{2\pi c} \right)^{\frac{n}{2}} \exp \left( -\sum_{i=1}^{n} \frac{Y_i^2}{2c} \right)$$  \hspace{1cm} (2.34)

which is the likelihood of $n$ independent samples from the stationary distribution.
We now consider the case $\alpha \to 0$. We are interested in evaluating the limit:

$$\lim_{\alpha \to 0} l(\alpha) = -\frac{1}{2} \log(2\pi c) - \frac{1}{2} \lim_{\alpha \to 0} \left( n \log(1 - e^{-2\alpha}) - \frac{\sum_{i=1}^{n} (Y_i - e^{-\alpha}Y_{i-1})}{c(1 - e^{-2\alpha})^2} \right)$$

(2.35)

Letting $z = 1 - e^{-2\Delta}$, so that $e^{-\alpha} = \sqrt{1 - z}$ and $z \to 0^+$ as $\alpha \to 0^+$ we have

$$\lim_{\alpha \to 0} l(\alpha) = -\frac{1}{2} \log(2\pi c) - \frac{1}{2} \lim_{z \to 0} \left( \log(z) - \frac{\sum_{i=1}^{n} (Z_i - \sqrt{1 - z}Y_{i-1})^2}{cz} \right)$$

(2.36)

From which it is clear that the limit as $\alpha$ goes to 0 of the log-likelihood is $-\infty$. □

So we know that the likelihood tends to 0 as $\alpha$ tends to 0, and that it tends to a fixed constant (the likelihood of $n$ independent draws from the stationary distribution, evaluated at the data points) as $\alpha$ tends to $\infty$. Therefore there is a maximum likelihood estimator, so long as the function is not increasing towards its value at $\infty$. We now investigate the behaviour of $\frac{\partial l}{\partial \alpha}$ at the boundaries of the parameter domain. Recall that

$$l(\alpha) = -\frac{n}{2} \log(2\pi c) - \frac{n}{2} \log(1 - e^{-2\alpha}) - \frac{\sum_{i=1}^{n} (Y_i - e^{-\alpha}Y_{i-1})^2}{2c(1 - e^{-2\alpha})}$$

(2.37)

so differentiating yields:

$$\frac{\partial l}{\partial \alpha} = \frac{\Delta e^{-2\alpha}}{1 - e^{-2\alpha}} - \frac{\Delta e^{-\alpha} \sum_{i=1}^{n} Y_iY_{i-1}}{c(1 - e^{-2\alpha})}$$

$$+ \frac{\Delta e^{-2\alpha} \sum_{i=1}^{n} (Y_i - e^{-\alpha}Y_{i-1})^2}{2c(1 - e^{-2\alpha})^2}$$

$$= q(\alpha)$$

(2.38)

(2.39)

**Lemma 1.** Let $f \in C^1((0, \infty))$ and suppose that there exist $\epsilon, M > 0$ such that

- $f'(x) > 0 \quad x \in (0, \epsilon)$
- $f'(x) < 0 \quad x \in [M, \infty)$
then $f$ has an interior minimum on $(0, \infty)$, i.e. there exists $y \in (0, \infty)$ such that $f(y) \geq f(x) \forall x \in (0, \infty)$.

Proof. By the conditions on the derivative $f(\varepsilon) \geq f(x)$ for all $x \in (0, \varepsilon]$ and $f(M) \geq f(x)$ for all $x \in [M, \infty)$. Therefore the maximum of $f$ on $[\varepsilon, M]$ is also the maximum of $f$ on $(0, \infty)$ (the maximum is attained on $[\varepsilon, M]$ since $f$ is continuous on this compact set). \qed

**Theorem 8.** Consider the function $l'(\alpha) = q(\alpha) : (0, \infty) \to \mathbb{R}$ defined above. Then

$$\lim_{\alpha \to 0^+} q(\alpha) = \infty$$

and there exists $M > 0$ such that

$$q(\alpha) < 0$$

for all $\alpha \geq M$.

Proof. Let $z(\alpha) = 1 - e^{-2\alpha \Delta}$. Then $l(\alpha) = f(z(\alpha))$ where

$$f(z) = -\frac{n}{2} \log(2\pi c) - \frac{n}{2} \log(z) - \sum_{i=1}^{n} (Y_{i\Delta} - \sqrt{1 - 2Y_{(i-1)\Delta}})^2$$

Now, by the chain rule $l'(\alpha) = f'(z(\alpha))z'(\alpha)$. But

$$\frac{dz}{d\alpha} = 2\Delta e^{-2\alpha \Delta} > 0$$
So that the sign of \( l'(\alpha) \) is the same as that of \( f'(z(\alpha)) \). We note that as \( \alpha \downarrow 0, z \downarrow 0 \) and as \( \alpha \uparrow \infty, z \uparrow 1 \). Now

\[
\frac{df}{dz} = -\frac{n}{2z} - \frac{\sum_{i=1}^{n}(Y_{i\Delta} - \sqrt{1-z}Y_{i(-1)\Delta})Y_{(i-1)\Delta}}{2cz\sqrt{1-z}} + \frac{\sum_{i=1}^{n}(Y_{i\Delta} - \sqrt{1-z}Y_{i(-1)\Delta})^2}{2cz^2} \tag{2.44}
\]

\[
= \frac{1}{2z} \left( - n - \frac{\sum_{i=1}^{n}(Y_{i\Delta} - \sqrt{1-z}Y_{i(-1)\Delta})Y_{(i-1)\Delta}}{c\sqrt{1-z}} + \frac{\sum_{i=1}^{n}(Y_{i\Delta} - \sqrt{1-z}Y_{i(-1)\Delta})^2}{cz} \right) \tag{2.45}
\]

As \( z \downarrow 0 \), the third term in the above expression dominates and \( f'(z) \to \infty \) (in particular, it is eventually positive). As \( z \to 1 \), only the second term doesn’t have a finite limit. This term is:

\[
\frac{1}{2z} \left( - \frac{\sum_{i=1}^{n}(Y_{i\Delta} - \sqrt{1-z}Y_{i(-1)\Delta})Y_{(i-1)\Delta}}{2c\sqrt{1-z}} \right) \tag{2.46}
\]

Since \( 1/2z > 0 \), the sign is the same as that of

\[
- \frac{\sum_{i=1}^{n}Y_{i\Delta}Y_{(i-1)\Delta}}{2c\sqrt{1-z}} + \frac{\sum_{i=1}^{n}Y_{(i-1)\Delta}}{2c} \tag{2.47}
\]

which, for \( z \) close enough to 1, has the same sign as \( -\sum_{i=1}^{n}Y_{(i-1)\Delta}Y_{i\Delta} \) (which is negative by assumption). Thus, for large enough \( \alpha \), \( l'(\alpha) < 0 \).

**Corollary 2.** Suppose that the sequence \( Y_{i\Delta} \) is not constant and \( \sum_{i=1}^{n}Y_{i\Delta}Y_{(i-1)\Delta} > 0 \), then the constrained log-likelihood \( l(\alpha) \) has an interior maximum on \((0, \infty)\).

It is not always possible to find an interior maximum of the likelihood function, as is illustrated by the following theorem:

**Theorem 9.** Suppose that the log-likelihood function is as above, and suppose further that:
\[ \begin{align*} 
&\bullet \quad - \sum_{i=1}^{n} Y_{i\Delta} Y_{(i-1)\Delta} = r > 0 \\
&\bullet \quad \sum_{i=1}^{n} Y_{i\Delta}^2 + 2 \sum_{i=1}^{n} Y_{(i-1)\Delta} + r > cn \\

then the log-likelihood function is increasing on \((0, \infty)\) (and therefore has no interior maximum).
\end{align*} \]

**Proof.** In the proof of theorem 8, we proved that \(l'(\alpha)\) has the same sign as \(h(\alpha)\) where \(z(\alpha) = 1 - e^{-2\alpha\Delta}\), and

\[
\begin{align*}
  h(z) &= -n - \frac{\sum_{i=1}^{n} (Y_{i\Delta} - \sqrt{1 - z} Y_{(i-1)\Delta}) Y_{(i-1)\Delta}}{c\sqrt{1 - z}} + \frac{\sum_{i=1}^{n} Y_{i\Delta} - \sqrt{1 - z} Y_{(i-1)\Delta}}{cz} \\
  &= -n - \frac{\sum_{i=1}^{n} Y_{i\Delta} Y_{(i-1)\Delta}}{c\sqrt{1 - z}} + \frac{\sum_{i=1}^{n} Y_{i\Delta}^2}{c} + \frac{\sum_{i=1}^{n} (Y_{i\Delta} - \sqrt{1 - z} Y_{(i-1)\Delta})^2}{cz} \\
  &= -n + \frac{r}{c\sqrt{1 - z}} + \frac{\sum_{i=1}^{n} Y_{i\Delta}^2}{c} + \frac{\sum_{i=1}^{n} Y_{i\Delta}^2 + 2\sqrt{1 - z} r + \sum_{i=1}^{n} Y_{(i-1)\Delta}^2}{cz} \\
  &\geq -n + \frac{r}{c} + \frac{\sum_{i=1}^{n} Y_{i\Delta}^2}{c} + \frac{\sum_{i=1}^{n} (Y_{i\Delta}^2 + Y_{(i-1)\Delta}^2)}{c} \\
  &> 0
\end{align*}
\]

By the assumptions of the theorem. \(\square\)
2.3 The Distribution of the Range

In this section, we investigate the distribution of the range of an Ornstein-Uhlenbeck process. This problem has applications to the pricing of double lookback options. For related results see [BS96]. For example, let $W$ be a standard Brownian Motion on $\mathbb{R}$. The distribution of the range of $W$ conditional on $W$ starting at $x$ is

$$
\mathbb{P}_x \left( \left\{ \sup_{0 < s < t} W_s - \inf_{0 < s < t} W_s < y \right\} \right) = 8 \int_0^y \sum_{k=1}^{\infty} (-1)^{k-1} k^2 \phi \left( \frac{kt}{\sqrt{t}} \right) dr
$$

(2.55)

$$
= \sqrt{\frac{2}{\pi}} \int_0^y r^{-1} L \left( \frac{r}{2\sqrt{t}} \right) dr
$$

(2.56)

where $\phi$ is the density of the standard normal distribution and

$$
L(z) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2z^2)
$$

(2.57)

$$
= \frac{\sqrt{2\pi}}{z} \sum_{k=1}^{\infty} \exp \left( \frac{-(2k-1)^2\pi}{8z^2} \right)
$$

(2.58)

this is proved in [Fel51]. The proof of the formula for the density proceeds as follows.

First, one shows that the solution of the partial differential equation:

$$
\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial y^2}
$$

$$
u(t, y; a, b) = 0 \quad y = a \text{ or } y = b
$$

$$
u(0, y; a, b) = \delta_x
$$

(2.59)

gives the probability

$$
\mathbb{P}_x \left( \left\{ a < \inf_{0 < s < t} W_s \text{ and } \sup_{0 < s < t} W_s < b \text{ and } W_t \in A \right\} \right) = \int_A u(t, y; a, b) dy
$$

(2.60)

$$
= \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} \int_A \left( e^{-\left( y - z + 2k(b-a) \right)^2/2t} - e^{-\left( y + z - 2k(b-a) \right)^2/2t} \right)
$$

(2.61)
where the last line is obtained by solving (2.3) by the method of images. From this, it is shown that

\[ F(t; a, b) = \mathbb{P}_x \left( \left\{ a < \inf_{0<s<t} W_s \text{ and } \sup_{0<s<t} W_s < b \right\} \right) \]

\[ = \frac{1}{\sqrt{2\pi t}} \int_a^b \sum_{k=-\infty}^{\infty} \left( e^{-\frac{(y-x+2k(b-a))^2}{2t}} - e^{-\frac{(y+x-2a+2k(b-a))^2}{2t}} \right) dy \]  

(2.62)  

(2.63)

The corresponding probability density is given by the mixed partial derivative:

\[ f(t; a, b) = \frac{\partial^2}{\partial a \partial b} F(t; a, b) \]

(2.64)

and the density of the range is then given by

\[ \rho(t, r) = \int_0^r f(t; u, u - r) du \]

(2.65)

In the following, we consider the same problem for the Ornstein-Uhlenbeck process. The result relies on the fact that the partial differential equation considered above with $\frac{1}{2} \Delta$ is also solvable for the adjoint of the generator of the Ornstein-Uhlenbeck process. As asserted without proof in [Fel51], the argument for determining the distribution of the range holds for more general Markov processes. We demonstrate this in the following section. We then solve the relevant boundary problem by an eigenfunction expansion.

**2.3.1 Method for Computing the Range**

In this section, we provide a rigorous justification of the method for computing the range of a stochastic process described above for the case of a Brownian motion and due to Feller [Fel51].
Theorem 10. Let $X$ be the solution of the stochastic differential equation:

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$ (2.66)

where $b, \sigma : \mathbb{R} \to \mathbb{R}$ are Lipschitz functions, and $W$ is a standard Brownian motion.

Let

$$\mathbb{P}_x \left( \left\{ a < \inf_{0<s<t} X_s and \sup_{0<s<t} X_s < b and X_t \in A \right\} \right) = F(t, A; a, b)$$ (2.67)

Then $F(t, A; a, b)$ has a continuous density

$$F(t, A; a, b) = \int_A u(t, x, y; a, b)dy$$ (2.68)

and for all Borel sets $A \subseteq (a, b), u$ satisfies

$$\frac{\partial}{\partial t} \int_A u(t, x, y; a, b)dx = \int_A \frac{\partial^2}{\partial y^2} \left( \frac{\sigma^2(y)}{2} u(t, x, y; a, b) \right) - \frac{\partial}{\partial y} \left( b(y)u(t, x, y; a, b) \right)dy$$ (2.69)

and also $u(t, x, a; a, b) = u(t, x, b; a, b) = 0, \lim_{t \to 0^+} u(t, x, y; a, b) = \delta_x(y)$.

Furthermore, if we define

$$F(t; a, b) = \mathbb{P}_x \left\{ a < \inf_{0<s<t} X_s and \sup_{0<s<t} X_s < b \right\}$$ (2.70)

then

$$F(t; a, b) = \int_a^b u(t, x, y; a, b)dy$$ (2.71)

Proof. The first part is well known. See [IJ65] or [Fel54].
The second part is immediate since

\[
P_{\pi}\left(\left\{ a < \inf_{0<s<t} X_s \text{ and } \sup_{0<s<t} X_s < b \right\}\right) = P_{\pi}\left(\left\{ a < \inf_{0<s<t} X_s \text{ and } \sup_{0<s<t} X_s < b \text{ and } X_t \in (a, b) \right\}\right)
\]

(2.72)

(2.73)

(2.74)

\[
= \int_{a}^{b} u(t, y; a, b) dy
\]

Proposition 5. Let \( M_t = \sup_{0 \leq s \leq t} X_s \), \( m_t = \inf_{0 \leq s \leq t} X_s \) and suppose that the joint distribution of \((M_t, m_t)\) has continuous density \( \varphi(x, y) \). Let

\[
F(t; a, b) = P_{\pi}\left(\left\{ a < m_t \text{ and } M_t < b \right\}\right)
\]

(2.75)

then, defining

\[
f(t; a, b) = -\frac{\partial^2}{\partial a \partial b} F(t; a, b)
\]

(2.76)

and

\[
R(t) = M_t - m_t
\]

(2.77)

then

\[
P_{\pi}\left(\left\{ R(t) \in A \right\}\right) = \int_{A} \delta(t; r) dr
\]

(2.78)

for all \( A \in B(R) \), where

\[
\delta(t; r) = \int_{0}^{r} f(t; u, u - r) du
\]

(2.79)
Proof. First, we note that

\[ F(t; a, b) = \mathbb{P}_x \left( \{ a < m_t \text{ and } M_t < b \} \right) \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{a} \varphi(x, y) dxdy \]

so that by the fundamental theorem of calculus \( f(t, a, b) = \varphi(a, b) \). But then

\[ \mathbb{P}_x(\{ R_t \leq r \}) = \mathbb{P}_x(\{ M_t - X_t \leq r \}) \]

\[ = \int_{z-r}^{z} \int_{x}^{u} \varphi(u, v) dudv \]

\[ = \int_{0}^{r} \delta(t; u) du \]

\[ \Box \]

2.3.2 Solution of the Partial Differential Equation

In this section, we solve the partial differential equation (2.69) for the special case of the Ornstein-Uhlenbeck process. In this case, the general problem (2.69) becomes

\[ \frac{\partial u}{\partial t} = \frac{\sigma^2 \partial^2 u}{2 \, \partial x^2} + \alpha x \frac{\partial u}{\partial x} + \alpha u \]  

(2.85)

\[ u(t, y; a, b) = 0 \quad y = a \text{ or } y = b \]  

(2.86)

\[ u(0, y; a, b) = \delta_x \]  

(2.87)

(where we have assumed without loss of generality for computing the range, that \( \beta = 0 \)).

Eigenfunction Expansion

In [SH70], it is shown that the partial differential equation (2.85) can be solved by separation of variables. Assuming that the solution of (2.85) can be written as a
product $X(x)T(t)$ we obtain the following system of ordinary differential equations

\[
T'(t) + \lambda^2 T(t) = 0 \quad \text{(2.88)}
\]

\[
\frac{\sigma^2}{2} X''(x) + \alpha x X'(x) + (\alpha + \lambda^2) X(x) = 0 \quad \text{(2.89)}
\]

Letting

\[
X(x) = \exp \left( \frac{-\alpha x^2}{2\sigma^2} \right) Y(x)
\]

we obtain

\[
Y''(x) + \left( \frac{\alpha}{\sigma^2} + \frac{2\lambda^2}{\sigma^2} - \frac{\alpha^2}{\sigma^4} x^2 \right) = 0 \quad \text{(2.91)}
\]

so that by employing the change of variables

\[
x = s \sqrt{\frac{\sigma^2}{2\alpha}}
\]

we have

\[
\frac{d^2 Y(s)}{ds^2} + \left( \frac{1}{2} + \frac{\alpha^2}{\sigma^2} \right) Y(s) = 0 \quad \text{(2.93)}
\]

with $\vartheta = \frac{\lambda^2}{\alpha}$, and $s_0 = \sqrt{\frac{2\alpha}{\sigma^2}} x_0$ which is Weber’s equation (see [Olver74]).

We now proceed to solve (2.93), following [SH70]. Let

\[
y_0(\vartheta, s) = \exp \left( -\frac{s^2}{4} \right) M \left( -\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right)
\]

\[
y_0(\vartheta, s) = s \exp \left( -\frac{s^2}{4} \right) M \left( -\frac{1}{2}, \frac{1}{2}, \frac{3}{2} \right)
\]

where $M(a, b, z)$ is Kummer’s function [Olver74]. These functions obey the recurrence

49
The Wronskian of these solutions is \( W(y_e, y_0) = 1 \).

We begin with the case \( a \neq -b \). Then the eigenvalues of (2.85) satisfy

\[
y_e(\vartheta, s_a)y_0(\vartheta, s_b) + y_0(\vartheta, s_a)y_e(\vartheta, s_b) = 0 \tag{2.98}
\]

where \( s_a = a\sqrt{\frac{2a}{\vartheta}} \) and \( s_b = -b\sqrt{\frac{2a}{\vartheta}} \). The solution to (2.93) with \( Y_0 = s_0 \) and Dirichlet boundary conditions at \( s_a, s_b \) is thus given by

\[
\sum_{n=1}^{\infty} \frac{\exp(-\alpha \vartheta_n t) \exp(-\frac{1}{2} (s - s_0)^2) Z(\vartheta_n, s) Z(\vartheta_n, s_0)}{\int_{-s_n}^{s_n} Z^2(\vartheta_n, s) ds} \tag{2.99}
\]

where

\[
Z(\vartheta_n, s) = y_e(\vartheta_n, s) - y_0(\vartheta_n, s) \frac{y_e(\vartheta_n, s_a)}{y_0(\vartheta_n, s_a)} \tag{2.100}
\]

for the \( \theta_n \) which satisfy (2.98). In the case where \( a = -b \), the eigenvalues will satisfy

\[
y_0(\vartheta, s_a)y_e(\vartheta, s_a) = 0 \tag{2.101}
\]

and the solution is

\[
\sum_{j=1}^{\infty} \frac{\exp(-\alpha \vartheta_j t) \exp(-\frac{1}{2} (s^2 - s_0^2)) y_e(\vartheta_j, s)y_e(\vartheta_j, s_0)}{\int_{-s_a}^{s_a} y_e^2(\vartheta_j, s) ds} + \sum_{k=1}^{\infty} \frac{\exp(-\alpha \vartheta_k t) \exp(-\frac{1}{2} (s^2 - s_0^2)) y_0(\vartheta_k, s)y_0(\vartheta_k, s_0)}{\int_{-s_a}^{s_a} y_0^2(\vartheta_k, s) ds} \tag{2.102}
\]

50
where $\vartheta_j, \vartheta_k$ solve

\begin{align}
    y_e(\vartheta_j, s_a) &= 0 \quad (2.103) \\
    y_o(\vartheta_k, s_a) &= 0 \quad (2.104)
\end{align}

The distribution of the range of the Ornstein-Uhlenbeck process can now be calculated using the recipe from proposition 5.

### 2.3.3 Double Lookback Options

A model that has been proposed for the distribution of the prices of commodities and electricity is the exponential mean-reverting model where $S_t = \exp(X_t)$ and

\[ dX_t = \alpha(\beta - X_t)dt + \sigma dB_t \quad (2.105) \]

(see [Pil98]).

One form of double lookback option pays off the range of an asset over a particular time interval (see [HHR98]). In the case of the exponential mean-reverting model, the price of a double lookback on the return of the option can be calculated based on the distribution of the range of the Ornstein-Uhlenbeck process from the previous section. In particular, we define the log-return of the asset $S$ over the interval $[0, t]$ to be

\[ D_t = \log \left( \frac{S_t}{S_0} \right) = \log(S_t) - \log(S_0) = X_t - \log(S_0) \quad (2.106) \]

so that

\[ \max_{t \in [0,T]} D_t - \min_{t \in [0,T]} D_t = \max_{t \in [0,T]} X_t - \min_{t \in [0,T]} X_t \quad (2.107) \]

So that an option which pays off the maximum return minus the minimum return
on the commodity will simply pay the range of the Ornstein-Uhlenbeck process $X$ over the interval $[0, T]$. The risk-neutral valuation formula (see [MR98]) yields that (assuming that (2.105) is under the risk-neutral measure) the price of a double lookback option under a constant risk-free interest rate $r$ is given by

$$P = e^{-rT}E[R_T]$$\hspace{1cm} (2.108)

The eigenfunction expansion from the previous section can now be used to calculate a price for the option numerically.
Chapter 3

Stochastic Drift Hidden Markov Models

3.1 Introduction

In this chapter, we investigate the behaviour of a class of hidden Markov models with applications in mathematical finance to energy prices and interest rate modelling (see [Lan00], [Pi98]). In particular, we are interested in a two-dimensional stochastic differential equation driven by a Brownian motion, where only one variable is assumed to be observable. We study the case where the observable variable depends on the hidden variable only in its drift coefficient.

3.1.1 The Model and Assumptions

We consider the stochastic process that is defined as the solution to the two-dimensional stochastic differential equation:

\begin{align*}
  dY_t &= \rho(V_t - Y_t)dt + \sigma dB_t & Y_0 &= y \\
  dV_t &= b(V_t)dt + a(V_t)dW_t & V_0 &= \xi
\end{align*}
where $\rho, \sigma > 0$, $(B_t, W_t)_{t \geq 0}$ is a standard two-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t; 0 \leq t < \infty)$, and $\xi$ is independent of $(B_t, W_t)_{t \geq 0}$. Implicitly, we assume that the process $Y_t$ can be directly observed, while the process $V_t$ cannot. We shall consider the process (3.1) under the following headings: marginal distributions of $Y_t$, estimating $V_t$ based on observations of $Y_t$ (filtering), and parameter estimation. A particular special case, where we are often able to obtain explicit results is the following:

\[
\begin{align*}
    dY_t &= \rho(Y_t - Y_t)dt + \sigma dB_t \\
    dV_t &= \alpha(\beta - V_t)dt + \nu dW_t
\end{align*}
\]

where $\alpha, \beta, \nu > 0$, (i.e. when $V$ follows the Ornstein-Uhlenbeck process discussed in chapter 2).

We now proceed to make a series of assumptions on the coefficients of the model (3.1) and explain the purpose of each. Consider an interval $(l, r)$ with $-\infty < l < r < \infty$.

**Assumption 7.** $b, a \in C^1((l, r))$ and

\[
\exists K > 1, \ \forall u \in (l, r), \ \beta^2(u) + a^2(u) \leq K(1 + u^2)
\]

**Assumption 8.** For a fixed $v_0 \in (l, r)$, consider the function:

\[
    s(u) = \exp \left( -2 \int_{v_0}^{u} \frac{b(u)}{a^2(u)} du \right)
\]

defined for all $u \in \mathbb{R}$. We assume that:

- $\int_{l}^{r} s(u) du = \int_{l}^{r} s(u) du = \infty$
- $\bar{M} = \int_{l}^{r} \frac{du}{a^2(u)s(u)} < \infty$

Assumptions 7 and 8 guarantee the existence and uniqueness of a solution to (3.1).
Assumption 9. Denote

\[ \tilde{\pi}(u) = \frac{1}{M a^2(u) s(u)} 1_{(r,\infty)} \]  \quad (3.5)

We assume that \( \xi \) has distribution \( \tilde{\pi}(u) du \).

The above two assumptions ensure that the process for \( V \) is a positive recurrent, strictly stationary ergodic diffusion process. The stationary distribution of the process \( V \) is \( \tilde{\pi}(u) du \).

3.2 The Marginal Distributions of \( Y \)

In this section, we prove that the marginal distribution of \( Y_t \) is a continuous mean mixture of normal distributions.

Lemma 2. Let \( B \) be a standard real-valued Brownian motion on a filtered probability space \( (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}) \), and let \( G \subset \mathcal{F} \) be independent of \( B \). Let \( V \) be a continuous adapted process that is \( G \) measurable. Consider the stochastic process \( Y \) defined by

\[ Y_t = e^{-pt} y + e^{-pt} \int_0^t e^{ps} V_s ds + \sigma e^{-pt} \int_0^t e^{ps} dB_s \]  \quad (3.6)

Then for any \( t > 0 \)

\[ \mathbb{P}[Y_t | G](\omega) \sim N(e^{-pt} (y + \int_0^t e^{ps} V_s(\omega) ds), \frac{\sigma^2}{2p} (1 - e^{-2pt}) ) \]  \quad (3.7)

where equality in the above equation holds \( \mathbb{P} \) almost surely and "~" denotes equality in distribution.

Proof. The result is immediate since \( Y_t \) is the sum of the two independent random variables

\[ U_t = e^{-pt} y + e^{-pt} \int_0^t e^{ps} V_s ds \]  \quad (3.8)
Proposition 6. Let \((Y_t, V_t)\) be the solution of stochastic differential equation (3.1). Then, conditional on \(\mathcal{G}_t = \sigma(V_s; s \in [0, t])\),

\[
Y_t \sim N\left(e^{-\sigma t} \left(y + \int_0^t e^{\sigma s} V_s ds\right), \frac{\sigma^2}{2\rho}(1 - e^{-2\sigma t})\right)
\]  

Proof. Consider the continuous semimartingale \(X_t = (Y_t, V_t, t)\). Applying Itô’s lemma to the function \(f(X_t) = e^{\sigma t} Y_t\) yields

\[
e^{\sigma t} Y_t - Y_0 = \int_0^t e^{\sigma s} V_s ds + \int_0^t e^{\sigma s} dB_s
\]  

The result now follows immediately from the lemma.

Lemma 3. Let \(f \in C^1([0, t])\), and let \(V\) be as in model (3.2). Then \(H_t = \int_0^t f(s)V_s ds\) is normally distributed with mean

\[
\hat{\mu} = \beta \int_0^t f(s)V_s ds
\]  

and variance

\[
\hat{\sigma}^2 = 2 \int_0^t \int_0^u f(u)f(s)(e^{-\alpha(u-s)} + \beta^2)dsdu - \hat{\mu}^2
\]
Proof. We first prove that $H_t$ is normally distributed. We do this by showing that it is the $L^2(\Omega, \mathbb{P})$ limit of a sequence of normally distributed random variables (see [Oks95] theorem A.7). For any $n \in \mathbb{N}$ we have that $\frac{1}{n} \sum_{k=1}^{n} f(t_k) V_{t_k}$ is normal (as a linear combination of normal random variables). From the assumptions on $f$, there exist $M_1, M_2 > 0$ such that

$$\max_{0 \leq s \leq t} |f(s)| \leq M_1$$

$$\max_{0 \leq s \leq t} |f'(s)| \leq M_2$$

So that, letting $t_{n,k} = \frac{t_k}{n}$

\[
E \left[ \left( \int_{0}^{t} f(s) V_s ds - \frac{1}{n} \sum_{k=1}^{n} f(t_{n,k}) V_{t_{n,k}} \right)^2 \right] 
\]

\[
= E \left[ \left( \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} f(s) V_s ds - f(t_{n,k}) V_{t_{n,k}} \right)^2 \right] 
\]

\[
\leq E \left[ \sum_{k=1}^{n} \left( \int_{t_{n,k-1}}^{t_{n,k}} f(s) V_s ds - f(t_{n,k}) V_{t_{n,k}} \right)^2 \right] 
\]

\[
\leq E \left[ \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} (f(s) V_s - f(t_{n,k}) V_{t_{n,k}})^2 ds \right] 
\]

\[
= E \left[ \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} f(s) V_s - f(t_{n,k}) V_{t_{n,k}} + f(t_{n,k}) V_{t_{n,k}} - f(t_{n,k}) V_{t_{n,k}} \right]^2 ds 
\]

\[
\leq 2t \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} E[V_s^2 | f(s) - f(t_{n,k})|^2 + f^2(t_{n,k})(V_s - V_{t_{n,k}})^2] ds 
\]

\[
\leq 2tnM_2^3 \left( \frac{t}{n} \right)^3 \hat{m}_2 + 2n M_1^2 \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} E[(V_s - V_{t_{n,k}})^2] ds 
\]

\[
\leq 2tnM_2^3 \left( \frac{t}{n} \right)^3 \hat{m}_2 + 2n M_1^2 \sum_{k=1}^{n} \int_{t_{n,k-1}}^{t_{n,k}} K(s - t_{n,k}) ds 
\]

\[
= 2tnM_2^3 \left( \frac{t}{n} \right)^3 \hat{m}_2 + 2n M_1^2 \sum_{k=1}^{n} \frac{K \left( \frac{t}{n} \right)^2}{2} 
\]
which tends to 0 as $n \to \infty$. (3.24) holds for some $K \geq 0$ since for $\Delta \leq 1$

$$\mathbb{E}[(V_{t+\Delta} - V_t)^2] = \mathbb{E} \left[ \left( \alpha \int_t^{t+\Delta} (\beta - V_s) ds + \sigma \int_t^{t+\Delta} dW_s \right)^2 \right]$$

$$\leq 2\mathbb{E} \left[ \alpha^2 \left( \int_t^{t+\Delta} (\beta - V_s) ds \right)^2 + \sigma^2 (W_{t+\Delta} - W_t)^2 \right]$$

$$\leq 2\mathbb{E} \left[ \alpha^2 \Delta \int_t^{t+\Delta} (\beta - V_s)^2 ds \right] + 2\sigma^2 \Delta$$

$$= 2\alpha^2 \Delta^2 \text{var}(\xi) + 2\sigma^2 \Delta$$

Now, we derive the expressions for the moments of $H_t$. By Fubini's theorem

$$\mathbb{E} \left[ \int_0^t f(s) V_s ds \right] = \int_0^t f(s) \mathbb{E}[V_s] ds$$

$$= \beta \int_0^t f(s) ds$$

Let

$$U(t) = \left( \int_0^t f(s) V_s ds \right)^2$$

then $U(0) = 0$ and

$$U'(t) = 2f(t)V_t \int_0^t f(s) V_s ds$$

so that

$$U(t) = 2 \int_0^t f(u) V_u \int_0^u f(s) V_s ds du$$

$$= 2 \int_0^t \int_0^u f(u)f(s) V_u V_s ds du$$
and therefore

\[
\mathbb{E}\left[ \left( \int_0^t f(s) V_s \, ds \right)^2 \right] = \mathbb{E}[U(t)]
\]

(3.36)

\[
= 2 \int_0^t \int_0^u f(u) f(s) \mathbb{E}[V_u V_s] \, ds \, du
\]

(3.37)

\[
= 2 \int_0^t \int_0^u f(u) f(s) (e^{-\alpha(u-s)} + \beta^2) \, ds \, du
\]

(3.38)

The result now follows from property 2 of the Ornstein-Uhlenbeck process (see chapter 2). \qed

**Proposition 7.** In the model (3.2),

\[
Y_t \sim N \left( e^{-\rho t} y + \beta(1 - e^{-\rho t}), \eta^2 \right)
\]

(3.39)

where

\[
\eta^2 = \rho^2 e^{-2\rho t} c^2 - \beta^2 (1 - e^{-\rho t})^2
\]

(3.40)

and

\[
c^2 = \frac{\nu^2}{2\rho \alpha (\rho + \alpha)} (e^{2\rho t} - 1) - \frac{\nu^2 t}{\alpha (\rho + \alpha)} + 2 \left( \frac{\beta}{\rho} \right)^2 \left( e^{\rho t} \left( \frac{1}{2} e^{\rho t} - 1 \right) \right) + \frac{1}{2}
\]

(3.41)

if \( \rho = \alpha \) and

\[
c^2 = \frac{\nu^2}{2\rho \alpha (\rho + \alpha)} (e^{2\rho t} - 1) - \frac{\nu^2 t}{\alpha (\rho^2 - \alpha^2)} (e^{(\rho-\alpha)t} - 1) + 2 \left( \frac{\beta}{\rho} \right)^2 \left( e^{\rho t} \left( \frac{1}{2} e^{\rho t} - 1 \right) \right) + \frac{1}{2}
\]

(3.42)

otherwise.

59
Proof. By lemma 3,

$$\int_0^t e^{ps}V_s ds \sim N(\mu, \sigma^2)$$

(3.43)

with $\mu = \beta$ and

\begin{align*}
\sigma^2 &= 2 \int_0^t \int_0^u e^{\rho(u+s)} \frac{\sigma^2}{2\alpha} e^{-\alpha(u-s)} ds du \\
&= \eta^2
\end{align*}

(3.44)

(3.45)

by a long, but straightforward, calculation using only techniques from elementary calculus.

The assertion is now a consequence of the fact that

$$N(\zeta, \sigma_1^2) \wedge N(\mu, \sigma_2^2) \sim N(\mu, \sigma_1^2 + \sigma_2^2)$$

(3.46)

(see [PR82], p. 30). \qed

3.3 Filtering

The problem of estimating $V_t$ based on a path for $Y_t$ is the classical filtering problem for diffusion processes. We do not discuss it, but instead refer the reader to [Kal80], where the problem is treated in detail.

3.4 Parameter Estimation

In this section, we examine a method of parameter estimation for model (3.1) based on the fact that it is a hidden Markov model. In particular, we show the consistency and convergence of certain moment estimators based on observations of the process $Y$. The results in this section provide analogues of those developed for stochastic
volatility hidden Markov models in [GCJLO0]. Many of the proofs here are similar to those in this reference.

We begin with the formal definition of a hidden Markov model, which is based on [Ler92]. Let \((Z, \mathcal{B}(Z))\) and \((U, \mathcal{B}(U))\) be two Polish spaces equipped with their Borel \(\sigma\)-algebras.

**Definition 24.** A stochastic process \((Z_n, n \in \mathbb{N})\) with state-space \((Z, \mathcal{B}(Z))\) is a hidden Markov model if there exists a strictly stationary Markov chain \((U_n, n \in \mathbb{N})\) with state-space \((U, \mathcal{B}(U))\) such that

- For all \(n\), given \((U_1, U_2, \ldots, U_n)\) the \(Z_k, k \leq n\) are conditionally independent, and the conditional distribution of \(Z_k\) depends only on \(U_k\).

- The conditional distribution of \(Z_k\) given \(U_k = u\) does not depend on \(k\).

Sometimes, when discussing a hidden Markov model, we will refer to \(U\) as the hidden chain and \(Z\) as the observed chain. Here and below \(\alpha_R(\cdot)\) denote the \(\alpha\)-mixing coefficients of the stochastic process \(R\).

**Proposition 8.** The process \((Z_n, n \in \mathbb{N})\) is strictly stationary. If the hidden Markov chain \((U_n, n \in \mathbb{N})\) is ergodic, then \((Z_n, n \in \mathbb{N})\) is also ergodic. Moreover, if \((U_n, n \in \mathbb{N})\) is \(\alpha\)-mixing, then \((Z_n, n \in \mathbb{N})\) is also \(\alpha\)-mixing and

\[
\alpha_Z(n) \leq \alpha_U(n) \tag{3.47}
\]

**Proof.** See [GCJLO0]. \(\Box\)

We now turn our attention to (3.1). Given \(\Delta > 0\), \(I_k = [(k - 1)\Delta, k\Delta]\) we define

\[
Z_k = \tilde{V}_k + e^{-\rho \Delta} \int_{I_k} e^{\rho t} dW_s \quad (= Y_k - e^{-\rho \Delta} Y_{k-1}) \tag{3.48}
\]

\[
U_k = (\tilde{V}_k, V_{k\Delta}) \quad \tilde{V}_k = e^{-\rho \Delta} \int_{I_k} e^{\rho t} V_s ds \tag{3.49}
\]
Theorem 11. Under the assumptions listed at the beginning of the chapter, we have:

- \((U_n, n \in \mathbb{N})\) is a strictly stationary Markov chain.

- \((Z_n, n \in \mathbb{N})\) is a hidden Markov model with hidden chain \((U_n, n \in \mathbb{N})\).

Proof. Let \(G_t = \sigma(V_s, s \in [0, t])\), and let \(\varphi : \mathbb{R}^2 \to \mathbb{R}\) be a bounded Borel measurable function. By the Markov property for \(V_t\) we have that

\[
\mathbb{E}[\varphi(\tilde{V}_k, V_{k\Delta})|G_{(k-1)\Delta}] = \mathbb{E}[\varphi(\tilde{V}_k, V_{k\Delta})|V_{(k-1)\Delta}] \\
= \mathbb{E}\left[\varphi\left(e^{-\rho k\Delta} \int_{(k-1)\Delta}^{k\Delta} e^{\rho s} V_s ds, V_{k\Delta}\right)|V_{(k-1)\Delta}\right] \\
= \mathbb{E}\left[\varphi\left(e^{-\rho \Delta} \int_{(k-1)\Delta}^{k\Delta} e^{\rho(u-(k-1)\Delta)} V_{u+(k-1)\Delta} ds, V_{(k-1)\Delta}\right)|V_{(k-1)\Delta}\right] \\
= \mathbb{E}_{V_{(k-1)\Delta}} \left[\varphi\left(e^{-\rho \Delta} \int_{0}^{\Delta} e^{\rho s} V_s ds, V_{\Delta}\right)\right] \\
= \psi(V_{(k-1)\Delta})
\]

where

\[
\psi(v) = \mathbb{E}[\varphi(\tilde{V}_1, V_{\Delta})|V_0 = v]
\]

Thus \(U\) is Markovian since

\[
\mathbb{E}[\varphi(U_k)|G_{(k-1)\Delta}] = \mathbb{E}[\varphi(U_k)|\tilde{V}_{k-1} = \bar{v}, V_{(k-1)\Delta} = v] = \psi(v)
\]

For \(k \geq 1\) define \(X_k\) by \(X_k(s) = V_{(k-1)\Delta + s}\) for \(s \in [0, \Delta]\). Then \(X_k\) is a stochastic process taking values in \(C([0, \Delta], \mathbb{R})\) which we assume to be equipped with the Borel \(\sigma\)-algebra \(\mathbb{B}_C\) associated with the uniform topology.
The strict stationarity of $V_t$ implies that of $X_k$. But $U_k = T(X_k)$ where

$$T(x) = \left( e^{-\rho \Delta} \int_0^{\Delta} e^{\rho \Delta} x(s) ds, x(\Delta) \right)$$  \hspace{1cm} (3.58)

The continuity of $T$ then implies the strict stationarity of $U$.

To prove that $Z$ is a hidden Markov model with hidden chain $U$, we must now show that conditional on $U_1, \ldots, U_n$, the random variables $Z_1, \ldots, Z_n$ are independent, with the conditional distribution of $Z_k$ depending only on $U_k$, and furthermore that the conditional distribution of $Z_k$, given $U_k = u$ does not depend on $k$. But conditional on $G_{n \Delta}$ the random variables $Z_1, \ldots, Z_n$ are independent and $Z_k \sim N(\tilde{V}_k, \frac{\sigma^2}{2\rho}(1 - e^{-2\rho \Delta}))$. But then for any $\lambda \in \mathbb{R}^n$

$$E \left[ \exp \left( i \sum_{j=1}^{n} \lambda_j Z_j \right) \bigg| U_1, \ldots, U_n \right]$$

$$= E \left[ E \left[ \exp \left( i \sum_{j=1}^{n} \lambda_j Z_j \right) \bigg| G_{n \Delta} \right] \bigg| U_1, \ldots, U_n \right]$$

$$= E \left[ \exp \left( i \sum_{j=1}^{n} \lambda_j \tilde{V}_j - \frac{\sigma^2}{4\rho} \lambda_j^2 (1 - e^{-2\rho \Delta})^2 \right) \bigg| U_1, \ldots, U_n \right]$$

$$= \exp \left( i \sum_{j=1}^{n} \lambda_j \tilde{V}_j - \frac{\sigma^2}{4\rho} \lambda_j^2 (1 - e^{-2\rho \Delta})^2 \right)$$

So conditional on $U_1, \ldots, U_n$ we have that $Z_1, \ldots, Z_n$ are independent and $Z_k \sim N(\tilde{V}_k, \frac{\sigma^2}{2\rho}(1 - e^{-2\rho \Delta}))$ which, given that $\tilde{V}_k = \tilde{v}$ does not depend on $k$, which completes the proof.

\[\square\]

**Proposition 9.** The process $Z$ is $\alpha$-mixing with $\alpha_Z(n) \leq \alpha_V((n-1)\Delta)$.

**Proof.** See [GCJL00].

\[\square\]

Given a positive integer $d$ and a Borel measurable function $\varphi : \mathbb{R}^d \to \mathbb{R}$ define
\[ h_\varphi : \mathbb{R}^d \to \mathbb{R} \text{ by} \]
\[ h_\varphi(v_1, \ldots, v_d) = \mathbb{E}[\varphi(v_1 + \varepsilon_1, \ldots, v_d + \varepsilon_d)] \quad (3.63) \]

(when it exists), where here \( \varepsilon_1, \ldots, \varepsilon_d \) are independent and identically distributed with distribution \( N(0, \frac{\varepsilon^2}{2\rho}(1 - e^{-2\rho \Delta})) \).

**Theorem 12.** If \( \varphi \) is such that \( \mathbb{E}[|h_\varphi(\tilde{V}_1, \ldots, \tilde{V}_d)|] < \infty \) then

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n-d} \varphi(Z_{i+1}, \ldots, Z_{i+d}) = \mathbb{E}[h_\varphi(\tilde{V}_1, \ldots, \tilde{V}_d)] \quad (3.64) \]

almost surely.

**Proof.** By proposition 9 and theorem 11 we have that \( Z \) is a strictly stationary, \( \alpha \)-mixing Markov process, and is therefore ergodic (see [Bra86]). Therefore, by the ergodic theorem (see appendix B), all that remains to be shown is that \( \mathbb{E}[|\varphi(Z_1, \ldots, Z_d)|] < \infty \) and

\[ \mathbb{E}[\varphi(Z_1, \ldots, Z_d)] = \mathbb{E}[h_\varphi(\tilde{V}_1, \ldots, \tilde{V}_d)] \quad (3.65) \]

Again letting \( G_t = \sigma(V_s, s \in [0, t]) \) we have that conditional on \( G_{d\Delta} \) \( (Z_1, \ldots, Z_d)^T \sim N(\mu, \Sigma) \) where \( \mu = (\tilde{V}_1, \ldots, \tilde{V}_d)^T \) and \( \Sigma = \frac{\varepsilon^2}{2\rho}(1 - e^{-2\rho \Delta})I_d \), so that

\[ \mathbb{E}[\varphi(Z_1, \ldots, Z_d)] = \mathbb{E} \left[ \mathbb{E}[\varphi(Z_1, \ldots, Z_d) | G_{d\Delta}] \right] \quad (3.66) \]

\[ = \mathbb{E}[h_\varphi(\tilde{V}_1, \ldots, \tilde{V}_d)] \quad (3.67) \]

The proofs of the following results are identical to the corresponding ones for stochastic volatility models.

64
Theorem 13. Let \( \Phi_i = \varphi(Z_{i+1}, \ldots, Z_{i+d}) \) If there exists \( \delta > 0 \) such that

\[
\mathbb{E}[(\Phi_0)^{2+\delta}] < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \alpha_{\delta^2}^k (k\Delta)
\]

then

\[
\Sigma_\Delta(\varphi, d) = \text{var}(\Phi_0) + 2 \sum_{k=1}^{\infty} \text{cov}(\Phi_0, \Phi_k)
\]

is well-defined. If furthermore \( \Sigma_\Delta(\varphi, d) > 0 \), then

\[
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=0}^{n-d} \left( \Phi_k - \mathbb{E}[\varphi(V_1, \ldots, V_d)] \right) = N(0, \Sigma_\Delta(\varphi, d))
\]

where the convergence in the above equation is in distribution.

Proof. The proof is an application of the Ibragimov central limit theorem for strictly stationary \( \alpha \)-mixing sequences. See [GCJL00] for details.

There is also the following multi-dimensional corollary.

Corollary 3. Suppose that \( \Phi : \mathbb{R}^d \to \mathbb{R}^p \). Suppose that there exists \( \delta > 0 \) such that

\[
\sum_{k=1}^{\infty} \alpha_{\delta^2}^k (k\Delta) < \infty \quad \text{and} \quad \mathbb{E}[|\Phi_0^k|^{2+\delta}] < \infty \quad k = 1, \ldots, p
\]

Let

\[
\Sigma_\Delta(\Phi^k, \Phi^l; d) = \text{cov}(\Phi_0^k, \Phi_0^l) + \sum_{j=1}^{\infty} (\text{cov}(\Phi_0^k, \Phi_j^l) + \text{cov}(\Phi_0^l, \Phi_j^k))
\]

Then \( \Sigma_\Delta(\Phi^k, \Phi^l; d) \) is well-defined. Consider the matrix \( \Sigma_\Delta = (\Sigma_\Delta(\Phi^k, \Phi^l))_{1 \leq k, l \leq p} \).
If $\Sigma_\Delta$ is positive definite then

$$
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{j=0}^{n-d} \left( \Phi_0^j - \mathbb{E}[h_{\Phi_1}(\tilde{V}_1, \ldots, \tilde{V}_d)] \right) \to N(0, \Sigma_\Delta) \tag{3.73}
$$

where the convergence in the above equation is in distribution.

Proof. See [GCJL00]. \qed

The following is required in verifying the hypothesis of the above limit theorem.

**Proposition 10.** Suppose that there exists $K \geq 0$ such that

$$
|\varphi(z_1, \ldots, z_d)| \leq K \left(1 + \sum_{j=1}^{d} |z_j|^q\right) \tag{3.74}
$$

and suppose that there exists a positive integer $r$ such that $2r \geq (2+\delta)q$ and $\mathbb{E}[V_0^{2r}] < \infty$. Then $\mathbb{E}[|\Phi_0|^{2+\delta}] < \infty$.

Proof. For $p \leq 2r$ let $\tilde{m}_p = \mathbb{E}[V_0^p]$. By an application of Jensen’s inequality, there exists $K_1 \geq 0$ such that

$$
|\varphi(Z_1, \ldots, Z_d)|^{2+\delta} \leq K^{2+\delta} \tag{3.75}
$$

$$
\leq K_1 \left(1 + \sum_{j=1}^{d} |Z_j|^{(2+\delta)q}\right) \tag{3.76}
$$

so that we have

$$
\mathbb{E}[|\varphi(Z_1, \ldots, Z_d)|^{2+\delta}] \leq K_1 \left(1 + \sum_{j=1}^{d} \mathbb{E}[|Z_j|^{(2+\delta)q}]\right) \tag{3.77}
$$

$$
\leq K_1 \left(1 + \sum_{j=1}^{d} \mathbb{E}[|Z_j|^{2r}\mathbb{E}[G_j\Delta]}\right) \tag{3.78}
$$

$$
\leq K_1 \left(1 + \sum_{j=1}^{d} \mathbb{E}[|Z_j|^{2r}\mathbb{E}[G_j\Delta]}\right) \tag{3.79}
$$

66
\[
\begin{align*}
&= K_1 \left( 1 + \sum_{j=1}^{d} \mathbb{E} \left[ \sigma^{2r} \sum_{k=1}^{r} \frac{(2r-1)!}{(2k)!(r-k)!2^{r-k}\sigma^{2k}} \left( e^{-i\Delta} \int_{f_j} e^{iV_s} ds \right)^{2k} \right] \right) \quad (3.80) \\
&\leq K_1 \left( 1 + \sum_{j=1}^{d} \mathbb{E} \left[ \sigma^{2r} \sum_{k=1}^{r} \frac{(2r-1)!}{(2k)!(r-k)!2^{r-k}\sigma^{2k}} e^{-2kj\Delta} \Delta^{2k-1} \int_{f_j} e^{2ks} V_s^{2k} ds \right] \right) \quad (3.81) \\
&= K_1 \left( 1 + \sum_{j=1}^{d} \sigma^{2r} \sum_{k=1}^{r} \frac{(2r-1)!}{(2k)!(r-k)!2^{r-k}\sigma^{2k}} e^{-2kj\Delta} \Delta^{2k-1} \int_{f_j} e^{2ks} \mathbb{E}[V_s^{2k}] ds \right) \quad (3.82) \\
&= K_1 \left( 1 + \sum_{j=1}^{d} \sigma^{2r} \sum_{k=1}^{r} \frac{(2r-1)!}{(2k)!(r-k)!2^{r-k}\sigma^{2k}} \frac{\Delta^{2k-1} k^{2k}}{2k} \left( 1 - e^{-2k\Delta} \right) \right) \quad (3.83)
\end{align*}
\]

which is finite by assumption. \(\square\)

In practice, we are interested in using estimators of the form \(\varphi(z_1, z_2) = z_1^p z_2^q\) where \(p, q \geq 1\) are integers. The asymptotic variance from the central limit theorem can be calculated more explicitly in these cases. We employ the following notation. Suppose \(X \sim N(\mu, \sigma^2)\) then we define for integers \(r \geq 1\)

\[
g_r(\mu, \sigma) = \mathbb{E}[X^r] = \begin{cases} 
\sigma^{2r-1} \sum_{j=1}^{r} \frac{(2r-1)!}{(2j-1)!(2r-2j)!} \mu^{2j-1} & \text{if } p \text{ is odd, } p = 2r - 1; \\
\sigma^{2r} \sum_{j=1}^{r} \frac{(2r)!}{(2j)!(2r-2j)!} \mu^{2j} & \text{if } p \text{ is even, } p = 2r.
\end{cases}
\tag{3.84}
\]

(see, for example [PR82]).

**Theorem 14.** Let \(\theta_\Delta = \frac{\sigma^2}{2p}(1 - e^{-2\rho\Delta})\). Suppose that there exists \(\delta > 0\) such that

\[
\sum_{k=1}^{\infty} \alpha^2 \frac{\Delta^2}{2\pi^4} (k\Delta) < \infty \tag{3.85}
\]
If $p \geq 1$ is an integer and $\varphi(z) = z^p$, and if $\mathbb{E}[V_0^{2p}] < \infty$ then

$$
\Sigma_{\Delta}(z_1^p, z_2^p, 2) = \mathbb{E}[g_{2p}(\tilde{V}_1, \theta_{\Delta})] - \mathbb{E}[g_p(\tilde{V}_1, \theta_{\Delta})] 
+ 2 \sum_{k=1}^{\infty} \text{cov}(g_p(\tilde{V}_1, \theta_{\Delta}), g_p(\tilde{V}_k, \theta_{\Delta})) \tag{3.86}
$$

If $p, q \geq 1$ are integers and $\varphi(z_1, z_2) = z_1^p z_2^q$ and $\mathbb{E}[V_0^{2(p+q)}] < \infty$ then

$$
\Sigma_{\Delta}(z^p, 1) = \mathbb{E}[g_{2p}(\tilde{V}_1, \theta_{\Delta}) g_{2q}(\tilde{V}_2, \theta_{\Delta})] - \mathbb{E}[g_p(\tilde{V}_1, \theta_{\Delta}) g_q(\tilde{V}_2, \theta_{\Delta})]^2 
+ 2 \sum_{k=2}^{\infty} \text{cov}(g_p(\tilde{V}_1, \theta_{\Delta}) g_q(\tilde{V}_2, \theta_{\Delta}), g_p(\tilde{V}_{k+1}, \theta_{\Delta}) g_q(\tilde{V}_{k+2}, \theta_{\Delta})) 
+ \mathbb{E}[g_p(\tilde{V}_1, \theta_{\Delta}) g_{p+q}(\tilde{V}_2, \theta_{\Delta}) g_3(\tilde{V}_3, \theta_{\Delta})] 
- \mathbb{E}[g_p(\tilde{V}_1, \theta_{\Delta}) g_q(\tilde{V}_2, \theta_{\Delta})] \cdot \mathbb{E}[g_p(\tilde{V}_2, \theta_{\Delta}) g_q(\tilde{V}_3, \theta_{\Delta})] \tag{3.87}
$$

Proof. Recall from theorem (13) that

$$
\Sigma_{\Delta}(\varphi, d) = \text{var}(\Phi_0) + 2 \sum_{k=1}^{\infty} \text{cov}(\Phi_0, \Phi_k) \tag{3.88}
$$

For the first part, we consider $\varphi(z) = z^p$. Now

$$
\text{var}(\Phi_0) = \text{var}(Z_1^p) \tag{3.89}
$$

$$
= \mathbb{E}[Z_1^{2p}] - \mathbb{E}[Z_1^p]^2 \tag{3.90}
$$

but

$$
\mathbb{E}[Z_1^{2p}] = \mathbb{E}\left[\mathbb{E}[Z_1^{2p}|G_{\Delta}]\right] \tag{3.91}
$$

$$
= \mathbb{E}[g_{2p}(\tilde{V}_1, \theta_{\Delta})] \tag{3.92}
$$

68
Similarly,

\[ \mathbb{E}[Z_i^p] = \mathbb{E}[\mathbb{E}[Z_i^p | G_\Delta]] \]
\[ = \mathbb{E}[g_p(\bar{V}_1, \theta_\Delta)] \]  

(3.93)  

(3.94)

So that

\[ \text{var}(\Phi_0) = \mathbb{E}[g_{2^p}(\bar{V}_1, \theta_\Delta)] - \mathbb{E}[g_p(\bar{V}_1, \theta_\Delta)]^2. \]  

(3.95)

The covariance calculations are dealt with in a similar manner

\[
\text{cov}(\Phi_0, \Phi_k) = \mathbb{E}[Z_i^p Z_{k+1}^p] - \mathbb{E}[Z_i^p] \cdot \mathbb{E}[Z_{k+1}^p]
\]
\[
= \mathbb{E}[\mathbb{E}[Z_i^p Z_{k+1}^p | G_{(k+1)\Delta}]] - \mathbb{E}[\mathbb{E}[Z_i^p | G_\Delta]] \cdot \mathbb{E}[\mathbb{E}[Z_{k+1}^p | G_{(k+1)\Delta}]]
\]
\[
= \mathbb{E}[g_p(\bar{V}_1, \theta_\Delta) g_p(\bar{V}_k, \theta_\Delta)] - \mathbb{E}[g_p(\bar{V}_1, \theta_\Delta)] \cdot \mathbb{E}[g_p(\bar{V}_k, \theta_\Delta)]
\]
\[
= \text{cov}(g_p(\bar{V}_1, \theta_\Delta), g_p(\bar{V}_k, \theta_\Delta))
\]

(3.96)  

(3.97)  

(3.98)  

(3.99)

The finiteness of all the expectations is a consequence of Jensen’s inequality and the assumption on the finiteness of \( \mathbb{E}[V_0^{2^p}] \).

We now proceed to the second case, with \( \varphi(z_1, z_2) = z_1^{p} z_2^q \).

\[
\text{var}(\Phi_0) = \mathbb{E}[Z_i^{2^p} Z_2^q] - \mathbb{E}[Z_i^p Z_2^q]^2
\]
\[
= \mathbb{E}[\mathbb{E}[Z_i^{2^p} Z_2^q | G_{2\Delta}]] - \mathbb{E}[\mathbb{E}[Z_i^p Z_2^q | G_{2\Delta}]]^2
\]
\[
= \mathbb{E}[g_{2^p}(\bar{V}_1, \theta_\Delta) g_{2^q}(\bar{V}_2, \theta_\Delta)] - \mathbb{E}[g_p(\bar{V}_1, \theta_\Delta) g_q(\bar{V}_2, \theta_\Delta)]^2
\]

(3.100)  

(3.101)  

(3.102)

For the covariances, the case \( k = 1 \) needs to be handled separately from the rest. If \( k \geq 2 \)

\[
\text{cov}(\Phi_0, \Phi_k)
\]
\[
= \mathbb{E}[Z_i^p Z_2^q Z_{k+1}^p Z_{k+2}^q] - \mathbb{E}[Z_i^p Z_2^q] \cdot \mathbb{E}[Z_{k+1}^p Z_{k+2}^q]
\]

(3.103)  

(3.104)
The case when \( k = 1 \) is slightly different.

\[
\text{cov}(\Phi_0, \Phi_1) = \mathbb{E}[Z_1^p Z_2^p Z_3^p] - \mathbb{E}[Z_1^p Z_2^p] \cdot \mathbb{E}[Z_3^p]
\]

\[
= \mathbb{E}
\left[
\mathbb{E}[Z_1^p Z_2^p Z_3^p | G_2 \Delta]\right] - \mathbb{E}
\left[
\mathbb{E}[Z_1^p Z_2^p | G_2 \Delta]\right] \cdot \mathbb{E}
\left[
\mathbb{E}[Z_3^p | G_2 \Delta]\right]
\]

\[
= \mathbb{E}[g_p(\bar{V}_1, \theta \Delta) g_q(\bar{V}_2, \theta \Delta) g_q(\bar{V}_3, \theta \Delta)]
\]

\[
= \text{cov}(g_p(\bar{V}_1, \theta \Delta) g_q(\bar{V}_2, \theta \Delta), g_p(\bar{V}_{k+1}, \theta \Delta) g_q(\bar{V}_{k+2}, \theta \Delta))
\]

Again, the expectations are finite due to the assumptions on the finiteness of the moments of \( V_0 \) and by an application of Jensen's inequality.

\[\square\]

### 3.5 Examples

In this section, we show how the theory from the previous section can be used in order to estimate the parameters of the \( V \) process in the model (3.1). In particular, we consider three processes, each corresponding to different values of \( \lambda \) in the following model

\[
dY_t = \rho(V_t - Y_t)dt + \sigma dB_t \quad Y_0 = y
\]

\[
dV_t = b(V_t)dt + \nu V_t^\lambda dW_t \quad V_0 = \xi
\]

Specifically, we consider the cases \( \lambda = 0, \frac{1}{2}, 1 \). When \( \lambda = 0 \) the \( V \) process in (3.111) is the Ornstein-Uhlenbeck process studied in chapter 2. We shall refer to this as the
OU model. We refer to the case $\lambda = \frac{1}{2}$ as the CIR model (after [JCR85]), and the case $\lambda = 1$ as the GARCH mode (as in this case $V$ is the diffusion approximation of a GARCH process, see [Nel90]). The fact that OU process satisfies the assumptions at the beginning of the chapter when $\alpha > 0$ follows from the properties listed in chapter 2. The CIR process satisfies the assumptions from the beginning of the chapter if $\alpha > 0$ and $\alpha \beta \geq \frac{\nu^2}{2}$, and the GARCH process satisfies the assumptions if $\alpha \beta > 0$ and $\alpha > -\frac{\nu^2}{2}$ (see [GCJL00]).

Recall that if $\varphi : \mathbb{R}^d \to \mathbb{R}$, we denote (when it exists)

$$h_{\varphi}(v_1, \ldots, v_d) = \mathbb{E}[\varphi(v_1 + \varepsilon_1, \ldots, v_d + \varepsilon_d)]$$

(3.112)

where $\varepsilon_1, \ldots, \varepsilon_d$ are independent and identically distributed as $N(0, \frac{\nu^2}{2\rho}(1 - e^{-2\rho\Delta}))$.

**Lemma 4.** If $\varphi_1, \varphi_2 : \mathbb{R} \to \mathbb{R}$, and $\varphi_3 : \mathbb{R}^2 \to \mathbb{R}$ are defined by $\varphi_1(v) = v, \varphi_2(v) = v^2, \varphi_3(v_1, v_2) = v_1 v_2$ then

$$h_{\varphi_1}(v) = v$$

(3.113)

$$h_{\varphi_2}(v) = v^2 + \frac{\sigma^2}{2\rho}(1 - e^{-2\rho\Delta})$$

(3.114)

$$h_{\varphi_3}(v_1, v_2) = v_1 v_2$$

(3.115)

**Proof.** This follows from elementary calculations using the only the linearity of expectations and the definition of $h_{\varphi}$. □

We observe that by theorem 12, $\tilde{m}_1 = \frac{1}{n} \sum_{k=1}^n Z_k$, $\tilde{m}_2 = \frac{1}{n} \sum_{k=1}^n Z_k^2 - \frac{\sigma^2}{2\rho}(1 - e^{-2\rho\Delta})$, and $\tilde{m}_{1,2} = \frac{1}{n} \sum_{k=1}^{n-1} Z_k Z_{k+1}$ are consistent estimators of $\mathbb{E}[\tilde{V}_1], \mathbb{E}[\tilde{V}_1^2]$ and $\mathbb{E}[\tilde{V}_1 \tilde{V}_2]$ respectively.

**Lemma 5.** Consider the stochastic differential equation

$$dV_t = \alpha(\beta - V_t) dt + \nu V_t^\lambda dW_t$$

(3.116)
where $W_t$ is standard Brownian motion. Then if $t > s$

$$E[V_tV_s] = C_\lambda + K_\lambda e^{-\alpha(t-s)} \quad (3.117)$$

where

$$C_0 = C_{\frac{1}{2}} = C_1 = \beta^2 \quad (3.118)$$

and

$$K_0 = \frac{\nu^2}{2\alpha} \quad (3.119)$$
$$K_{\frac{1}{2}} = \frac{\beta\nu^2}{2\alpha} \quad (3.120)$$
$$K_1 = \frac{\nu^2\beta^2}{2\alpha} + \frac{\nu^2\beta^2}{2\alpha - \nu^2} \quad (3.121)$$

Proof. Applying Itô's lemma to the function $f(v, t) = e^{\alpha t} v$ and the continuous semimartingale $X_t = (V_t, t)$ we obtain

$$e^{\alpha t} V_t = V_0 + \alpha \beta \int_0^t e^{\alpha s} ds + \nu \int_0^t e^{\alpha s} V_s^\lambda dW_s \quad (3.122)$$

and so

$$V_t = \beta + e^{-\alpha t} (V_0 - \beta) + \nu e^{-\alpha t} \int_0^t e^{\alpha s} V_s^\lambda dW_s \quad (3.123)$$

Multiplying this expression out for $V_tV_s$ and using elementary properties of Brownian motion leads to

$$E[V_tV_s] = \beta^2 + e^{-\alpha(t-s)} \text{var}(V_0) + \nu^2 e^{-\alpha(t-s)} E\left[\left(\int_0^t e^{\alpha u} V_u^\lambda dW_u\right)^2\right] \quad (3.124)$$

$$= \beta^2 + e^{-\alpha(t-s)} \text{var}(V_0) + \nu^2 e^{-\alpha(t-s)} E\left[\int_0^t e^{2\alpha u} V_u^{2\lambda} du\right] \quad (3.125)$$

72
We now treat each case separately. For the OU case $\lambda = 0$ (3.125) becomes

$$E[V_tV_s] = \beta^2 + e^{-\alpha(t-s)}\text{var}(V_0) + \nu^2 e^{-\alpha(t-s)} \int_0^t e^{2\alpha u} du$$

(3.126)

$$= \beta^2 + e^{-\alpha(t+s)}\frac{\nu^2}{2\alpha} + \frac{\nu^2}{2\alpha} e^{-\alpha(t+s)}(e^{2\alpha s} - 1)$$

(3.127)

$$= \beta^2 e^{-\alpha(t-s)}\frac{\nu^2}{2\alpha}$$

(3.128)

(which we could also get from property 2 from chapter 2). For the case of the CIR process we get

$$E[V_tV_s] = \beta^2 + e^{-\alpha(t+s)}\text{var}(V_0) + \nu^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} E[V_u] du$$

(3.129)

$$= \beta^2 + e^{-\alpha(t+s)}\frac{\beta \nu^2}{2\alpha} + e^{-\alpha(t+s)}\frac{\beta \nu^2}{2\alpha}(e^{2\alpha s} - 1)$$

(3.130)

$$= \beta^2 + \frac{\beta \nu^2}{2\alpha} e^{-\alpha(t-s)}$$

(3.131)

where we have used the variance of the CIR process, which is derived in ([Shr96], pp. 172-173).

For the case $\lambda = 1$, we have

$$E[V_tV_s] = \beta^2 + e^{-\alpha(t+s)}\text{var}(V_0) + \nu^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} E[V_u^2] du$$

(3.132)

$$= \beta^2 + e^{-\alpha(t+s)}\text{var}(V_0) + \nu^2 e^{-\alpha(t+s)} \int_0^t e^{2\alpha u} (\text{var}(V_0) + \beta^2) du$$

(3.133)

$$= \beta^2 + e^{-\alpha(t+s)}\text{var}(V_0) + \frac{\nu^2}{2\alpha} e^{-\alpha(t+s)}(\text{var}(V_0) + \beta^2)(e^{2\alpha s} - 1)$$

(3.134)

$$= \beta^2 + \nu^2 \beta^2 \left( \frac{1}{2\alpha} + \frac{1}{2\alpha - \nu^2} \right)$$

(3.135)

where we have used that in this case

$$\text{var}(V_0) = \frac{\beta^2 \nu^2}{2\alpha - \nu^2}$$

(3.136)

(see [GCJL00], pp. 1072-1073).
Theorem 15. With $C_\lambda, K_\lambda$ as in the lemma, we have

$$E[\tilde{V}_1] = \frac{\beta}{\rho} (1 - e^{-\rho \Delta})$$

(3.137)

$$E[\tilde{V}_1 \tilde{V}_2] = \begin{cases} 
\frac{C_\lambda}{\rho} e^{-2\rho \Delta} (e^{\rho \Delta} - 1)^2 + \frac{K_\lambda}{\rho + \alpha} e^{-3\rho \Delta} (e^{(\rho + \alpha) \Delta} - 1) & \rho = \alpha \\
\frac{C_\lambda}{\rho} e^{-2\rho \Delta} (e^{\rho \Delta} - 1)^2 + \frac{K_\lambda}{\rho^2 - \alpha^2} e^{-2\rho \Delta} (e^{(\rho + \alpha) \Delta} - 1)(e^{(\rho - \alpha) \Delta} - 1) & \rho \neq \alpha 
\end{cases}$$

(3.138)

$$E[\tilde{V}_1^2] = 2e^{-\rho \Delta} \cdot \begin{cases} 
\frac{C_\lambda}{\rho^2} \left( \frac{1}{2} e^{2\rho \Delta} - e^{\rho \Delta} + \frac{1}{2} \right) + \frac{K_\lambda}{2\rho(\rho + \alpha)} (e^{2\rho \Delta} - 1) - \frac{K_\lambda}{\rho + \alpha} & \rho = \alpha \\
\frac{C_\lambda}{\rho^2} \left( \frac{1}{2} e^{2\rho \Delta} - e^{\rho \Delta} + \frac{1}{2} \right) + \frac{K_\lambda}{2\rho(\rho + \alpha)} (e^{2\rho \Delta} - 1) - \frac{K_\lambda}{\rho^2 - \alpha^2} (e^{(\rho - \alpha) \Delta} - 1) & \rho \neq \alpha 
\end{cases}$$

(3.139)

Proof. We begin with $E[\tilde{V}_1]$. Using Fubini’s theorem

$$E[\tilde{V}_1] = e^{-\rho \Delta} \int_0^\Delta e^s E[V_s] ds$$

(3.140)

$$= \frac{\beta}{\rho} e^{-\rho \Delta} (e^{\rho \Delta} - 1)$$

(3.141)

$$= \frac{\beta}{\rho} (1 - e^{-\rho \Delta})$$

(3.142)

Next we consider $E[\tilde{V}_1 \tilde{V}_2]$.

$$E[\tilde{V}_1 \tilde{V}_2] = E \left[ e^{-\rho \Delta} \int_0^\Delta e^s V_s ds \cdot e^{-2\rho \Delta} \int_\Delta^2 e^u V_u du \right]$$

(3.143)

$$= E \left[ e^{-3\rho \Delta} \int_\Delta^2 \int_0^\Delta e^{s+u} V_s V_u ds du \right]$$

(3.144)

$$= e^{-3\rho \Delta} \int_\Delta^2 \int_0^\Delta e^{s+u} E[V_s V_u] ds du$$

(3.145)

$$= e^{-3\rho \Delta} \int_\Delta^2 \int_0^\Delta e^{s+u}(C_\lambda + K_\lambda e^{-\alpha(u-s)}) ds du$$

(3.146)

The result for $E[\tilde{V}_1^2]$ now follows by a straightforward calculation using only techniques from elementary calculus. The case $E[\tilde{V}_1^2]$ is handled analogously. We
observe that

\[ \mathbb{E}[\tilde{Y}_t^2] = e^{-2\rho \Delta} \mathbb{E}[U(\Delta)] \]  

(3.147)

where

\[ U(t) = \left( \int_0^t e^{\rho s} V_s ds \right)^2 \]  

(3.148)

so that \( U(0) = 0 \) and by the chain rule (the integral in the above equation can be defined pathwise \( \mathbb{P} \) almost surely)

\[ U'(t) = 2e^{\rho t} V_t \int_0^t e^{\rho s} V_s ds \]  

(3.149)

so that

\[ U(t) = 2 \int_0^t \int_0^u e^{\rho u} V_u \int_0^u e^{\rho s} V_s ds du \]  

(3.150)

\[ = 2 \int_0^t \int_0^u e^{\rho(u+s)} V_u V_s ds du \]  

(3.151)

so that we have

\[ \mathbb{E}[U(\Delta)] = 2 \int_0^\Delta \int_0^u e^{\rho(u+s)} \mathbb{E}[V_u V_s] ds du \]  

(3.152)

\[ = 2 \int_0^\Delta \int_0^u e^{\rho(u+s)} (C_\lambda + K_\lambda e^{-\alpha(u-s)}) ds du \]  

(3.153)

The result now follows by an elementary calculation using only techniques from single-variable calculus.

The results of theorem 15 together with the estimators for moments of \( \tilde{V} \) can now be used to estimate the parameters of the process \( V \) (although, to obtain parameter values a nonlinear equation must be solved).
Chapter 4

Stochastic Volatility Hidden Markov Models

4.1 Introduction

In this chapter, we study hidden Markov stochastic volatility models that have been applied in finance. In order to facilitate comparison, we study the stochastic volatility models under the same headings as those in the previous chapter: marginal distributions, filtering, and parameter estimation. The results in this chapter were motivated by results on a similar model in [GCJL98],[GCJL99] and [GCJL00]. In particular, we employ similar assumptions on the diffusion processes concerned, and on the asymptotic framework for the sampling intervals. In order to facilitate comparisons, we have tried to keep our notation consistent with that in [GCJL98],[GCJL99] and [GCJL00].
4.1.1 The Model and Assumptions

We consider a two-dimensional diffusion process defined as the solution of the following system of stochastic differential equations:

\[ dY_t = V_t dB_t \quad Y_0 = 0 \]
\[ dV_t = b(V_t) dt + a(V_t) dW_t \quad V_0 = \xi \]

where \((B_t, W_t)_{t \geq 0}\) is a standard two-dimensional Brownian motion which is defined on \((\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t; 0 \leq t \leq \infty)\), and \(\xi\) is independent of \((B_t, W_t)_{t \geq 0}\). In financial applications, the process for \(Y_t\) could serve as a model for the value of the logarithm of the (suitably normalized) stock price. We now proceed to make a series of assumptions on the coefficients of the model (4.1) and explain the purpose of each.

**Assumption 10.** \(b, a \in C^1(\mathbb{R})\) and

\[ \exists K > 1, \quad \forall u \in \mathbb{R}, \quad b^2(u) + a^2(u) \leq K(1 + u^2) \]

**Assumption 11.** For a fixed \(v_0 \in \mathbb{R}\), consider the function:

\[ s(v) = \exp \left( -2 \int_{v_0}^v \frac{b(u)}{a^2(u)} du \right) \]

defined for all \(v \in \mathbb{R}\). We assume that:

- \( \int_{-\infty}^{\infty} s(u) du = \int_{-\infty}^{\infty} s(u) du = \infty \)
- \( \bar{M} = \int_{\mathbb{R}} \frac{du}{a^2(u)s(u)} < \infty \)

The first two assumptions guarantee the existence and uniqueness of a solution to (4.1). We also use the first assumption to estimate the small time behaviour of volatility fluctuations.
Assumption 12. Denote

\[ \pi(u) = \frac{1}{Ma^2(u)s(u)} \]  

We assume that \( \xi \) has distribution \( \pi(u)du \).

The above two assumptions ensure that the process for \( V \) is a positive recurrent, strictly stationary ergodic diffusion process. The stationary distribution of the process \( V \) is \( \pi(u)du \).

Assumption 13. We assume that \( \pi \) has moments of at least order eight. That is,

\[ \exists \alpha \geq 8, \int_{\mathbb{R}} |u|^\alpha \pi(u)du < \infty \]  

This assumption is needed to prove theorems 16, 17 and 18. For notational convenience, we denote:

\[ \bar{m}_p = \int_{\mathbb{R}} u^p \pi(u)du \]  

for \( 0 \leq p \leq \alpha \).

4.2 Marginal Distributions of \( Y_t \)

The following intuitive lemma is often used implicitly in the study of stochastic volatility models. The only reference for a proof that we know of is [Kal98b]. Since this is unpublished, we reproduce the proof here, but emphasize that it is due to [Kal98b].

Proposition 11. Let \( B \) be a standard real-valued Brownian motion on a filtered probability space \( (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}) \), and let \( \mathcal{G} \subset \mathcal{F} \) be independent of \( B \). Let \( V \) be a
continuous adapted process that is $G$ measurable. Then for any $T > 0$

$$P \left[ \int_0^T V_t dB_t | G \right] (\omega) \sim N \left( 0, \int_0^T V_t^2(\omega) dt \right)$$

(4.7)

where the above equation holds $P$ almost surely and $\sim$ denotes equality in distribution.

Proof. Suppose that $\tilde{V}$ is a simple predictable process, i.e. that $\tilde{V}$ has the form

$$\tilde{V} = \sum_{i=1}^{n} \alpha_i 1_{[s_i, s_{i+1}]}$$

(4.8)

where $0 \leq s_1 \leq \cdots \leq s_{n+1}$ are positive real numbers and $\alpha_1, \ldots, \alpha_n$ are $G$ measurable random variables. Then $\int_0^T \tilde{V}_t dB_t = \sum_{i=1}^{n} (B_{s_{i+1} \wedge T} - B_{s_i \wedge T}) \alpha_i$. Let $D^1$ denote the Skorokhod space of all functions $\tilde{\omega} : \mathbb{R}_+ \to \mathbb{R}$ which are right-continuous with left-hand limits (see [JS87] chapter 6, or [EK86] chapter 3), and let $P^{W}$ denote Wiener measure on $D^1$. Fixing $\omega \in \Omega$ and letting $g : D^1 \to \mathbb{R}$ by $g(\tilde{\omega}) = \sum_{i=1}^{n} (\tilde{\omega}(s_{i+1} \wedge T) - \tilde{\omega}(s_i \wedge T)) \alpha_i(\omega)$, then

$$P \left[ \int_0^T \tilde{V}_t dB_t | G \right] (\omega) \sim (P^{W})^g$$

(4.9)

$$\sim \mu_1 * \mu_2 * \cdots * \mu_n$$

(4.10)

$$\sim N \left( 0, \int_0^T \tilde{V}_t^2(\omega) dt \right)$$

(4.11)

where $(P^{W})^g(A) = P^{W}(g^{-1}(A))$, and

$$\mu_i \sim N(0, ((s_{i+1} \wedge T) - (s_i \wedge T))(\alpha_i(\omega))^2)$$

(4.12)

Now assume that $V$ is as in the assertion. We have that there exists a sequence $(V^k)_{k \in \mathbb{N}}$ of simple processes as defined above such that $V^k_t \to V_t$ uniformly on
[0, T] \ P \text{ almost surely and}

\int_0^T V_t^k dB_t \to \int_0^T V_t dB_t \quad \mathbb{P} \text{ a.s.} \quad (4.13)

as \ k \to \infty \ (\text{see [JS87], 1.4.44}). By the dominated convergence theorem

\begin{align*}
\lim_{k \to \infty} \int_\mathbb{R} f(x) \mathbb{P} \left[ \int_0^T V_t^k dW_t \mid \mathcal{G} \right] (dx) &= \lim_{k \to \infty} \mathbb{E} \left[ f \left( \int_0^T V_t^k dB_t \right) \mid \mathcal{G} \right] \\
&= \mathbb{E} \left[ f \left( \int_0^T V_t dB_t \right) \mid \mathcal{G} \right] \\
&= \int_\mathbb{R} f(x) \mathbb{P} \left[ \int_0^T V_t dW_t \mid \mathcal{G} \right] (dx)
\end{align*}

for any bounded continuous function \( f : \mathbb{R} \to \mathbb{R} \). Thus for almost all \( \omega \in \Omega \) the distribution of \( \mathbb{P}[\int_0^T V_t^k dt \mid \mathcal{G}] \) converges weakly to the distribution of \( \mathbb{P}[\int_0^T V_t dB_t \mid \mathcal{G}] \). But by the above, we have that

\( \mathbb{P} \left[ \int_0^T V_t^k dt \mid \mathcal{G} \right] \sim N \left( 0, \int_0^T (V_t^k(\omega))^2 dt \right) \quad (4.17) \)

So by the uniqueness of the weak limit, if we can show that

\( N \left( 0, \int_0^T (V_t^k(\omega))^2 dt \right) \to N \left( 0, \int_0^T (V_t(\omega))^2 dt \right) \quad (4.18) \)

weakly, then the assertion will be proved. The convergence in (4.18) is proved by considering characteristic functions. We have that

\( \int_0^T (V_t^k)^2 dt \to \int_0^T V_t^2 dt \quad \mathbb{P} \text{ a.s.} \quad (4.19) \)
Considering the characteristic functions of the measures in (4.17), and using (4.19) we have almost surely for any \( \lambda \in \mathbb{R} \)

\[
\lim_{k \to \infty} \exp \left( -\frac{\lambda^2}{2} \int_0^T (V_t^k)^2 dt \right) = \exp \left( -\frac{\lambda^2}{2} \int_0^T V_t^2 dt \right)
\]

(4.20)

which implies (4.18) by the continuity theorem (see [Bi179], theorem 26.3).

\[\square\]

Corollary 4. (Conditional Itô isometry). Under the hypotheses of the theorem

\[
\mathbb{E} \left[ \left( \int_0^T V_t dB_t \right)^2 \bigg\vert \mathcal{G} \right] (\omega) = \int_0^T V_t^2(\omega) dt
\]

\( \mathbb{P} \) almost surely.

**Proof.** This is an immediate consequence of the proposition, and ([Bi79], theorem 34.5).

\[\square\]

### 4.3 Filtering

In this section, we consider the problem of calculating the path for \( V_t \) based on observations of \( Y_t \). Theorem 28 on the quadratic variation (see appendix B) can be exploited in the current model to allow us to at least obtain the path of \( \int_0^t V_s^2 ds \).

We observe that this theorem implies that \( \int_0^t V_s^2 ds \) is measurable with respect to the \( \sigma \)-algebra generated by the observations \( \mathcal{H}_t = \{Y_s; 0 \leq s \leq t\} \) because of the following simple result.

**Lemma 6.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space and consider a complete \( \sigma \)-algebra \( \mathcal{G} \) with \( \mathcal{G} \subseteq \mathcal{F} \). Let \( X_n \) be a sequence of \( \mathcal{G} \) measurable random variables with \( X_n \) converging to \( X \) in probability. Then \( X \) is \( \mathcal{G} \) measurable.

**Proof.** There is a subsequence \( X_{n_k} \) which converges to \( X \) almost surely, thus implying the \( \mathcal{G} \) measurability of \( X \).

\[\square\]
This allows us (in the case of continuous observations), to obtain the path of $V^2$, due to the following result.

**Proposition 12.**

\[
\lim_{n \to \infty} \frac{1}{n} \int_{T - \frac{1}{n}}^{T} V_t^2 dt = V_T^2
\]  \hspace{1cm} (4.22)

almost surely.

**Proof.** The proof is identical to that of Wiener’s local ergodic theorem, (see [Kre85] p. 10-11, and appendix B). Let $\tilde{N}$ denote the complement of the set of points $(\omega, u) \in \Omega \times (0, \infty)$ with

\[
\lim_{n \to \infty} \frac{1}{n} \int_{u - \frac{1}{n}}^{u} V_t^2(\omega) dt = V_u^2(\omega)
\]  \hspace{1cm} (4.23)

Also let

\[
N_\omega = \{ u \in (0, \infty) : (\omega, s) \in \tilde{N} \} \hspace{1cm} (4.24)
\]

\[
N_u = \{ \omega \in \Omega : (\omega, u) \in \tilde{N} \} \hspace{1cm} (4.25)
\]

Since $V$ is almost surely continuous, for almost all $\omega \in \Omega$ we can apply the fundamental theorem of calculus to (4.23) to obtain $\lambda(N_\omega) = 0$ where $\lambda$ denotes Lebesgue measure.

It follows that $\mathbb{P} \otimes \lambda(\tilde{N}) = 0$ (by [Fol84], theorem 2.36) and by Fubini’s theorem, $\lambda$ almost all $u$ have the property that $\mathbb{P}(N_u) = 0$. But $N_T = \theta_{T-u}N_u$, and as $\theta_{T-u}$ is measure-preserving, we have $\mathbb{P}(N_T) = 0$. \qed

**Corollary 5.** $V_t^2$ is measurable with respect to the $\sigma$-algebra $\mathcal{H}_t$.

Therefore, it is always possible to estimate the path of the square of the volatility process. In many applications, the volatility process $V$ is assumed to be a diffusion

83
on the positive real line, where $x^2$ is a bijection, and we have that $V_t$ is also $\mathcal{H}_t$ measurable.

### 4.3.1 Discretization and Limit Theorems

**Notation**

We work in the following asymptotic framework. For each $n \in \mathbb{N}$, we partition the interval $[0, T_n]$ into $M_n$ subintervals of length $\Delta_n$. As $n \to \infty$, the number of observations $M_n$ tends to infinity, and the sampling interval $\Delta_n$, which is fixed for each $n$ tends to zero. We also assume that the length of the observation interval $T_n = M_n \Delta_n$ tends to infinity.

We employ the following notation:

\[
R_n = \frac{1}{\sqrt{\Delta_n T_n}} = \frac{1}{\Delta_n \sqrt{M_n}}
\]

\[
I_{n,k} = [(k-1)\Delta_n, k\Delta_n]
\]

\[
X_{n,k} = (Y(k\Delta_n) - Y((k-1)\Delta_n))^2 - \int_{(k-1)\Delta_n}^{k\Delta_n} V_t^2 dt
\]

\[
= \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt
\]

\[
Z_{n,k} = R_n X_{n,k}
\]

\[
V_T^\ast = \max_{t \in [0,T]} |V_t|
\]

\[
V_T^\dagger = \min_{t \in [0,T]} |V_t|
\]

\[
V_{n,k}^\ast = \max_{t \in I_{n,k}} |V_t|
\]

\[
V_{n,k}^\dagger = \min_{t \in I_{n,k}} |V_t|
\]

Finally, we denote by $\mathcal{F}_{n,k}$ the $\mathbb{P}$ completion of

\[
\sigma(Y(j\Delta_n), V_t; 0 \leq j \leq k, 0 \leq t < \infty)
\]
Notice that the $\sigma$-algebras $\mathcal{F}_{n,k}$ contain the information generated by the entire sample path of the volatility.

4.3.2 Results

Central Limit Theorem

Lemma 7. Let the processes $V$, $V^*$ and $V^\dagger$ be as in the previous section. Then

$$\lim_{T \to 0^+} \mathbb{E}[\int (V_T^\dagger)^4 - (V_T)^4] = 0$$

(4.36)

Proof. Consider the polynomials:

$$p(x) = 4K(x^3 + x^5) + 6K(x^2 + x^4)$$

(4.37)

$$q(x) = x^9 + Kx^6$$

(4.38)

where $K$ is as in assumption 10, and let

$$C_1 = \mathbb{E}[p(\mid \xi \mid)]$$

(4.39)

$$C_2 = \mathbb{E}[q(\mid \xi \mid)]$$

(4.40)

both of which are finite due to assumption 13. Let $T > 0$. By Itô’s lemma, we have that for any $s, t \in [0, T]$

$$|V_t^4 - V_s^4| = \left| \int_s^t 4V_u^3 b(V_u) + 6V_u^2 a^2(u)du \right|$$

$$+ \left| \int_s^t V_u^3 a(u) dW_u \right|$$

$$\leq \left| \int_s^t 4V_u^3 \| b(V_u) \| + V_u^2 a^2(u)du \right|$$

$$+ \left| \int_s^t V_u^3 a(u) dW_u \right|$$

(4.41)

(4.42)
where $M_t$ is the martingale:

$$M_t = \int_0^t V_u^a(u) dW_u$$  \hspace{1cm} (4.46)

and

$$M_T^* = \max_{u \in [0,T]} |M_u|$$  \hspace{1cm} (4.47)

Since the above inequality holds for all $t, s \in [0, T]$ we must have

$$(V_T^a)^4 - (V_T^d)^4 \leq \int_0^T p(|V_u|) du + 8M_T^*$$  \hspace{1cm} (4.48)

so

$$\mathbb{E}[(V_T^a)^4 - (V_T^d)^4] \leq \mathbb{E}\left[\int_0^T p(|V_u|) du\right] + 8\mathbb{E}[M_T^*]$$  \hspace{1cm} (4.49)

$$= C_1 T + 8\mathbb{E}[M_T^*]$$  \hspace{1cm} (4.50)

So it remains to show that $\mathbb{E}[M_T^*]$ tends to 0 as $T \to 0^+$. Using Doob’s inequality, the conditional Itô isometry (corollary 4) and Fubini’s theorem we have that there exists $C > 0$ such that:

$$\mathbb{E}[M_T^*] \leq \sqrt{\mathbb{E}[(M_T^*)^2]}$$  \hspace{1cm} (4.51)

$$\leq C \sqrt{\mathbb{E}[M_T^2]}$$  \hspace{1cm} (4.52)
\[ C^2 = A + C \]

This completes the proof.

The following is a corollary to (11).

**Proposition 13.** For almost all \( \omega \in \Omega \)

\[ \mathbb{P} \left[ \int_{I_{n,k}^\infty} V_t dB_t \bigg| \mathcal{F}_{n,k-1} \right] (\omega) \sim N \left( 0, \int_{I_{n,k}} V_t^2(\omega) dt \right) \]  

**Proof.** This follows immediately from proposition 11, the Markov property and the independence of increments for Brownian motion.

**Theorem 16.** Let \( Z_{n,k}, \mathcal{F}_{n,k} \) be as above and let

\[ S_n = \sum_{k=1}^{M_n} Z_{n,k} \]  

Then

\[ S_n \to Z \]  

where \( Z \) has the distribution \( N(0, 2\hat{m}_4) \).

**Proof.** The result is a consequence of the martingale central limit theorem (see appendix B). We have that \( \{ Z_{n,k}, \mathcal{F}_{n,k}, k = 1, \ldots, M_n, n \in \mathbb{N} \} \) is a martingale.
difference array since

\[
E[Z_{n,k} | \mathcal{F}_{n,k-1}] = R_n \left( E \left[ \left( \int_{l_{n,k}} \left( \int_{l_{n,k}} V_t dB_t \right)^2 dt \right) - \int_{l_{n,k}} V_t^2 dt | \mathcal{F}_{n,k-1} \right] \right) \quad (4.60)
\]

\[
= R_n \left( E \left[ \left( \int_{l_{n,k}} V_t dB_t \right)^2 | \mathcal{F}_{n,k-1} \right] - \int_{l_{n,k}} V_t^2 dt \right) \quad (4.61)
\]

\[
= R_n \left( E \left[ \int_{l_{n,k}} V_t^2 dt | \mathcal{F}_{n,k-1} \right] - \int_{l_{n,k}} V_t^2 dt \right) \quad (4.62)
\]

\[
= 0 \quad (4.63)
\]

Where we have used the conditional Itô isometry (corollary 4), and the fact that \( V_t \) is measurable with respect to \( \mathcal{F}_{n,k-1} \), for all \( t \in \mathbb{R} \) and for all \( n, k \).

To check the Lindeberg condition, it is sufficient to show that

\[
\lim_{n \to \infty} \sum_{k=1}^{M_n} E[Z_{n,k}^4 | \mathcal{F}_{n,k-1}] = 0 \quad (4.64)
\]

where the above limit is in probability (see appendix B).

Conditional upon \( \mathcal{F}_{n,k-1} \) we have by proposition 11

\[
\int_{l_{n,k}} V_t dB_t \sim U_{n,k} \quad (4.65)
\]

\( \mathbb{P} \) almost surely where \( \sim \) denotes equality in distribution and \( U_{n,k} \) is normally distributed with mean zero and variance

\[
\sigma_{n,k}^2 = \int_{l_{n,k}} V_t^2 dt \quad (4.66)
\]
Thus we have

\[
\mathbb{E}[Z_{n,k}^4 | \mathcal{F}_{n,k-1}] = R_n^4 \mathbb{E} \left[ \left( \left( \int_{I_{n,k}} V_i dB_i \right)^2 - \int_{I_{n,k}} V_i^2 dt \right)^4 \right]
\]

\[
= R_n^4 \mathbb{E}[(U_{n,k}^2 - \sigma_{n,k}^2)^4]
\]

\[
= R_n^4 (\mathbb{E}[U_{n,k}^6 - 4U_{n,k}^4 \sigma_{n,k}^2 + 6U_{n,k}^2 \sigma_{n,k}^4 - 4U_{n,k}^2 \sigma_{n,k}^6 + \sigma_{n,k}^8])
\]

\[
= R_n^4 \sigma_{n,k}^8 (105 - 60 + 18 - 4 + 1)
\]

\[
= 60R_n^4 \sigma_{n,k}^8
\]

Thus, verifying the Lindeberg condition is reduced to showing that:

\[
R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_i^2 dt \right)^4
\]

converges to zero in probability.

This follows since for any \( \varepsilon > 0 \)

\[
P \left( \left\{ \left( R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_i^2 dt \right)^4 \right) > \varepsilon \right\} \right) \leq \frac{R_n^4 \mathbb{E} \left[ \left( \int_{I_{n,k}} V_i^2 dt \right)^4 \right]}{\varepsilon}
\]

\[
\leq \frac{R_n^4 \sum_{k=1}^{M_n} \mathbb{E} \left[ \Delta_n^3 \int_{I_{n,k}} V_i^8 dt \right]}{\varepsilon}
\]

\[
= \frac{R_n^4 M_n \Delta_n^4 \tilde{m}_g}{\varepsilon}
\]

\[
= \frac{M_n \Delta_n^4 \tilde{m}_g}{M_n \Delta_n^4 \varepsilon}
\]

\[
= \frac{\tilde{m}_g}{\varepsilon M_n}
\]

which converges to 0 since \( M_n \to \infty \) as \( n \to \infty \). In the above, we have used Jensen’s inequality in (4.74) and Fubini’s theorem in (4.75).

In order to apply the martingale central limit theorem from appendix B, the only
The step remaining is to show that

\[ \sum_{k=1}^{M_n} \mathbb{E}[Z_{n,k}^2 | \mathcal{F}_{n,k}] \rightarrow \bar{m}_4 \]  \hspace{1cm} (4.78)

where the convergence is in probability (the measurability and nesting conditions in the theorem are obviously satisfied).

Using the same notation as in the previous calculation,

\[ \mathbb{E}[Z_{n,k}^2 | \mathcal{F}_{n,k}] = R_n^2 \mathbb{E}[(U_{n,k}^2 - \sigma_{n,k}^2)] \]  \hspace{1cm} (4.79)
\[ = 2R_n^2 \sigma_{n,k}^4 \]  \hspace{1cm} (4.80)
\[ = 2R_n^2 \left( \int_{I_{n,k}} V_t^2 \, dt \right)^2 \]  \hspace{1cm} (4.81)

So the proof will be complete if we are able to show that

\[ R_n^2 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 \, dt \right)^2 \rightarrow \bar{m}_4 \]  \hspace{1cm} (4.82)

in probability as \( n \rightarrow \infty \).

We have that

\[ \mathbb{P}\left( \left\{ \left| \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 \, dt \right)^2 - \bar{m}_4 \right| > \varepsilon \right\} \right) \]  \hspace{1cm} (4.83)
\[ \leq \mathbb{P}\left( \left\{ \left| \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 \, dt \right)^2 - \frac{1}{T_n} \int_0^{T_n} V_t^4 \, dt \right| > \frac{\varepsilon}{2} \right\} \right) \]
\[ + \mathbb{P}\left( \left\{ \left| \frac{1}{T_n} \int_0^{T_n} V_t^4 dt - \bar{m}_4 \right| > \frac{\varepsilon}{2} \right\} \right) \]  \hspace{1cm} (4.84)

The second term in the above expression converges to zero as \( n \rightarrow \infty \) by the ergodic theorem (see appendix B). After some manipulation, the first term can be controlled.
using the lemma.

\[
\mathbb{P}\left( \left\{ \left| R_n^2 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 dt \right)^2 - \frac{1}{T_n} \int_0^{T_n} V_t^4 dt \right| > \frac{\varepsilon}{2} \right\} \right)
\]

(4.85)

\[
= \mathbb{P}\left( \left\{ \left| R_n^2 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 dt \right)^2 - \frac{1}{T_n} \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^4 dt \right| > \frac{\varepsilon}{2} \right\} \right)
\]

(4.86)

\[
= \mathbb{P}\left( \left\{ \left| R_n^2 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 dt \right)^2 - \Delta_n \int_{I_{n,k}} V_t^4 dt \right| > \frac{\varepsilon}{2} \right\} \right)
\]

(4.87)

\[
\leq \mathbb{P}\left( \left\{ \left| R_n^2 \sum_{k=1}^{M_n} \Delta_n^2 \left( (V_{n,k}^*)^4 - (V_{n,k}^t)^4 \right) \right| > \frac{\varepsilon}{2} \right\} \right)
\]

(4.88)

\[
\leq \frac{2}{\varepsilon} \mathbb{E}\left[ R_n^2 \Delta_n^2 \sum_{k=1}^{M_n} \left( (V_{n,k}^*)^4 - (V_{n,k}^t)^4 \right) \right]
\]

(4.89)

\[
= \frac{2R_n^2 \Delta_n^2}{\varepsilon} \sum_{k=1}^{M_n} \mathbb{E}\left[ (V_{n,k}^*)^4 - (V_{n,k}^t)^4 \right]
\]

(4.90)

\[
= \frac{2}{\varepsilon M_n} \sum_{k=1}^{M_n} \mathbb{E}\left[ (V_{n,k}^*)^4 - (V_{n,k}^t)^4 \right]
\]

(4.91)

\[
= \frac{2}{\varepsilon} \mathbb{E}\left[ (V_{n,0}^*)^4 - (V_{n,0}^t)^4 \right]
\]

(4.92)

which tends to 0 as \( n \to \infty \) by the lemma. (4.91) follows from the strict stationarity of \( V \).

\[\square\]

**Corollary 6.** Suppose

\[
I_{n,k} = [(k-1)n^{-\alpha}, kn^{-\alpha}], k = 1, \ldots, n^\beta
\]

(4.93)

with \( \beta > \alpha > 0 \). So that \( T_n = n^{d-\alpha} \) and \( R_n = n^\gamma = n^{\alpha - \frac{\beta}{2}} \). Then

\[
n^\gamma \sum_{k=1}^{n^\beta} \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \to Z
\]

(4.94)

where the convergence is in distribution and \( Z \) has a normal distribution with mean 0 and variance \( 2\tilde{m}_4 \).
Notice that in the above corollary $\gamma$ be either positive or negative depending on the size of $\alpha$ and $\beta$.

**Large Deviations**

Given the recent results by Lesigne and Dolný on large deviations for martingale difference arrays (see [LV99], or appendix B), we can derive large deviation results for the limit theorem of the previous section. The criterion for applying the large deviation results for martingale difference sequences in [LV99] is the finiteness of moments of the random variables $Z_{n,k}$. In this section, we prove that the existence of the moments of the $Z_{n,k}$ defined in the previous section can be proven based on the existence of moments of the stationary distribution of the volatility process $V$.

**Lemma 8.** Let $p \geq 1$ and suppose that

$$\tilde{m}_{2p} = \int_{\mathbb{R}} u^{2p} \pi(u) du < \infty \quad (4.95)$$

Then

$$\mathbb{E}[Z_{n,k}^p] \leq \frac{a_p \tilde{m}_{2p}}{(M_n)^{\frac{p}{2}}} \quad (4.96)$$

where $a_p \in \mathbb{R}_+$ is the $p$th absolute moment about the mean of a $\chi^2(1)$ random variable. That is, if $U \sim \chi^2(1)$ then

$$a_p = \mathbb{E}[|U - 1|^p] \quad (4.97)$$

**Proof.** Using the fact that conditional on $\mathcal{F}_{n,k-1}$, $\int_{t_{n,k}} V_t dB_t$ is normally distributed with mean zero and variance $\int_{t_{n,k}} V_t^2 dt$ ($\mathbb{P}$ almost surely, by proposition 11) we have
that:

$$
E[|Z_{n,k}|^p] = E \left[ R_n^p \left( \int_{I_{n,k}} V_i dB_i \right)^2 - \int_{I_{n,k}} V_i^2 dt \right]^p
$$

(4.98)

$$
= E \left[ R_n^p E \left[ \left( \int_{I_{n,k}} V_i dB_i \right)^2 - \int_{I_{n,k}} V_i^2 dt \right]^p \right] |\mathcal{F}_{n,k-1}
$$

(4.99)

$$
= R_n^p a_p E \left[ \left( \int_{I_{n,k}} V_i^2 dt \right)^p \right]
$$

(4.100)

$$
\leq R_n^p a_p \Delta_n^{p-1} E \left[ \int_{I_{n,k}} V_i^{2p} dt \right]
$$

(4.101)

$$
= R_n^p a_p \Delta_n^{p-1} \bar{m}_{2p}
$$

(4.102)

$$
= \frac{a_p \bar{m}_{2p}}{(M_n)^{\frac{p}{2}}}
$$

(4.103)

\[ \square \]

**Theorem 17.** Let $p \geq 2, x > 0$ and suppose that

$$
\bar{m}_{2p} = \int_{\mathbb{R}} u^{2p} \hat{\pi}(u) du < \infty
$$

(4.104)

then

$$
\mathbb{P}(\{|S_n| > nx\}) \leq (18pq^{\frac{1}{2}})^p \frac{M^p}{x^p} \cdot \frac{1}{n^{\frac{p}{2}}}
$$

(4.105)

where

$$
M = \frac{a_p \bar{m}_{2p}}{(M_n)^{\frac{p}{2}}}
$$

(4.106)

and $q \in \mathbb{R}$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.

**Proof.** This follows directly from the result of the lemma together theorem 27 from appendix B. \[ \square \]
We observe that an exponential large deviation result using the corresponding result from [LV99] is not possible in this case, because of the following.

**Proposition 14.** $\mathbb{E}[\exp(|Z_{n,k}|)]$ is not finite.

**Proof.**

\begin{equation}
\mathbb{E}[\exp(|Z_{n,k}|)] = \mathbb{E}\left[ \exp \left( \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right) \right] \tag{4.107}
\end{equation}

\begin{equation}
= \mathbb{E}\left[ \mathbb{E}\left[ \exp \left( \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right) \bigg| \mathcal{F}_{n,k} \right] \right] \tag{4.108}
\end{equation}

Now, using the fact that given $\mathcal{F}_{n,k-1}$, $\int_{I_{n,k}} V_t dB_t$ is normally distributed with mean 0 and variance $\int_{I_{n,k}} V_t^2 dt$ ($\mathbb{P}$ almost surely, by proposition 11) we have that

\begin{equation}
\mathbb{E}\left[ \exp \left( \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right) \bigg| \mathcal{F}_{n,k} \right] < \infty \tag{4.109}
\end{equation}

if and only if

\begin{equation}
\int_{I_{n,k}} V_t^2 dt < \frac{1}{2} \tag{4.110}
\end{equation}

The result follows since

\begin{equation}
\mathbb{P}\left( \left\{ \int_{I_{n,k}} V_t^2 dt \geq \frac{1}{2} \right\} \right) > 0 \tag{4.111}
\end{equation}

\[ \square \]

### 4.3.3 Nonzero Drift

In this section, we prove an analogue of the result in the above section in the case where the model for the logarithm of the stock has a nonzero drift. Thus, we are
studying the solution of a two-dimensional stochastic differential equation

\[
dY_t = \varphi(V_t)dt + V_tdB_t \quad Y_0 = 0 \quad (4.112)
\]
\[
dV_t = b(V_t)dt + a(V_t)dW_t \quad V_0 = \xi \quad (4.113)
\]

where, as before, \((B_t, W_t)_{t \geq 0}\) is a standard two-dimensional Brownian motion defined on \((\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t; 0 \leq t \leq \infty)\), and \(\xi\) is independent of \((B_t, W_t)_{t \geq 0}\). In this case, our proof of the central limit theorem for the error in approximating the quadratic variation requires three additional assumptions.

**Assumption 14.** \(\varphi \in C^1(\mathbb{R})\) and

\[
\exists K_d > 1, \quad \forall u \in \mathbb{R}, \quad \varphi^2(u) \leq K_d(1 + u^2) \quad (4.114)
\]

**Assumption 15.** The stationary distribution of \(V\) has finite moments up to order at least 16

\[
\int_{\mathbb{R}} u^{16} \pi(u)du < \infty \quad (4.115)
\]

**Assumption 16.**

\[
\lim_{n \to \infty} \Delta_n \sqrt{M_n} = 0 \quad (4.116)
\]

We use the same notation as in the previous section except that now

\[
X_{n,k} = \left( Y(k\Delta_n) - Y((k-1)\Delta_n) \right)^2 - \int_{(k-1)\Delta_n}^{k\Delta_n} V_t^2 dt \quad (4.117)
\]

\[
X_{n,k} = \left( \int_{I_{n,k}} \varphi(V_t)dB_t + \int_{I_{n,k}} V_tdB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \quad (4.118)
\]

\[
Z_{n,k} = R_n X_{n,k} \quad (4.119)
\]
Theorem 18. Let \( Z_{n,k}, \mathcal{F}_{n,k} \) be as above and let

\[
S_n = \sum_{k=1}^{M_n} Z_{n,k}
\]

(4.120)

Then

\[
S_n \to Z
\]

(4.121)

where \( Z \) has the distribution \( N(0, 2\tilde{m}_4) \).

\textbf{Proof.} This is an application of the central limit theorem (theorem 25 from appendix B). We begin with the condition on the condition expectation.

\[
\mathbb{E}[X_{n,k} \mid \mathcal{F}_{n,k-1}] = \mathbb{E}
\left[
\left(\int_{I_{n,k}} \varphi(V_i)dB_t + \int_{I_{n,k}} V_i dB_t\right)^2
\right.
\]
\[
- \left. \int_{I_{n,k}} V_i^2 dt \bigg| \mathcal{F}_{n,k-1} \right]
\]  

(4.122)

\[
= \mathbb{E}
\left[
\left(\int_{I_{n,k}} \varphi(V_i)dt\right)^2 + 2 \int_{I_{n,k}} \varphi(V_i)dt \cdot \int_{I_{n,k}} V_i dB_t
\right]
\]
\[
+ \left(\int_{I_{n,k}} V_i dB_t\right)^2 - \int_{I_{n,k}} V_i^2 dt \bigg| \mathcal{F}_{n,k-1}
\]  

(4.123)

\[
= \left(\int_{I_{n,k}} \varphi(V_i)dt\right)^2 + 2 \int_{I_{n,k}} \varphi(V_i)dt \cdot \mathbb{E}
\left[
\int_{I_{n,k}} V_i dB_t \bigg| \mathcal{F}_{n,k-1}
\right]
\]
\[
+ \mathbb{E}
\left[
\left(\int_{I_{n,k}} V_i dB_t\right)^2 - \int_{I_{n,k}} V_i^2 dt \bigg| \mathcal{F}_{n,k-1}
\right]
\]  

(4.124)

\[
= \left(\int_{I_{n,k}} \varphi(V_i)dt\right)^2
\]

(4.125)

But this implies that \( \mathbb{E}[Z_{n,k} \mid \mathcal{F}_{n,k-1}] \) converges in probability to zero since for any
\[ \epsilon > 0 \]

\[
P \left( \left\{ \frac{R_n}{\varepsilon} \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 > \varepsilon \right\} \right) \leq P \left( \left\{ \Delta_n \frac{R_n}{\varepsilon} \sum_{k=1}^{M_n} \int_{I_{n,k}} \varphi^2(V_t)dt > \varepsilon \right\} \right) \]

(4.126)

\[
\leq \frac{\Delta_n R_n}{\varepsilon} \left[ \sum_{k=1}^{M_n} \int_{I_{n,k}} \varphi^2(V_t)dt \right] \]

(4.127)

\[
= \frac{\Delta_n R_n}{\varepsilon} \int_0^{T_n} \varphi^2(V_t)dt \]

(4.128)

\[
\leq \frac{\Delta_n R_n T_n}{\varepsilon} \cdot K_d (1 + \bar{m}_2) \]

(4.129)

\[
= \Delta_n \sqrt{M_n} K_d (1 + \bar{m}_2) \]

(4.130)

which tends to zero by assumption (16). Next, we consider the conditional variance of \( Z_{n,k} \) given \( F_{n,k-1} \). We have that

\[
E \left[ E \left[ \left( \int_{I_{n,k}} \varphi(V_t)dt + \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt - \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 \right] \bigg| F_{n,k-1} \right] \]

(4.131)

\[
= E \left[ E \left[ 2 \int_{I_{n,k}} \varphi(V_t)dt \cdot \int_{I_{n,k}} V_t dB_t + \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right]^2 \bigg| F_{n,k-1} \right] \]

(4.132)

\[
= E \left[ 4 \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 \left( \int_{I_{n,k}} V_t dB_t \right)^2 \right.

+ 4 \int_{I_{n,k}} \varphi(V_t)dt \int_{I_{n,k}} V_t dB_t \left( \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right)

+ \left( \left( \int_{I_{n,k}} V_t dB_t \right)^2 - \int_{I_{n,k}} V_t^2 dt \right)^2 \bigg| F_{n,k-1} \right] \]

(4.133)
So that the conditional variance of \( Z_{n,k} \) given \( \mathcal{F}_{n,k-1} \) is

\[
\text{Var}_{k-1}[Z_{n,k}] = 2R_n^2 \left( \int_{I_{n,k}} V_i^2 dt \right)^2 + 4 \int_{I_{n,k}} V_i^2 dt \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2
\]  

(4.137)

We now show that \( \sum_{k=1}^{M_n} \text{Var}_{k-1}[Z_{n,k}] \) converges in probability to \( 2\bar{m}_4 \).

\[
P \left( \left\{ \left| \frac{1}{M_n} \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 + 2 \int_{I_{n,k}} V_i^2 dt \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 \right| - \bar{m}_4 \right| > \varepsilon \right) \right)
\]  

(4.138)

\[
\leq P \left( \left\{ \left| 2R_n^2 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 \right| > \frac{\varepsilon}{2} \right\} \right)
\]  

(4.139)

\[
+ P \left( \left\{ \left| R_n^2 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 - \bar{m}_4 \right| > \frac{\varepsilon}{2} \right\} \right)
\]

The second term in the above expression tends to zero as in the proof of the limit theorem in the previous section. The first term is shown to converge to zero in
probability using the assumptions and Jensen's inequality.

\[
\mathbb{P}\left(\frac{2R_n^2}{\varepsilon} \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 > \frac{\varepsilon}{2}\right)
\leq \frac{4R_n^2}{\varepsilon} \mathbb{E}\left[ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^2 \right]
\leq \frac{4\Delta_n R_n^2}{\varepsilon} \mathbb{E}\left[ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \cdot \int_{I_{n,k}} \varphi^2(V_i) dt \right]
\leq \frac{4\Delta_n R_n^2}{\varepsilon} \mathbb{E}\left[ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \cdot \int_{I_{n,k}} K_d(1 + V_i^2) dt \right]
= \frac{4K_d}{\Delta_n M_n} \mathbb{E}\left[ \sum_{k=1}^{M_n} \left( \Delta_n + \int_{I_{n,k}} V_i^2 dt \right) \left( \int_{I_{n,k}} V_i^2 dt \right) \right]
= \frac{4K_d}{\Delta_n M_n} \mathbb{E}\left[ \Delta_n \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt + \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_i^2 dt \right)^2 \right]
\leq \frac{4K_d}{T_n} \mathbb{E}\left[ \Delta_n \int_0^{T_n} V_i^2 dt + \Delta_n \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \right]
= \frac{4K_d}{T_n} \mathbb{E}\left[ \int_0^{T_n} V_i^2 + V_i^4 dt \right]
= \frac{4K_d\Delta_n}{T_n} \cdot T_n (\bar{m}_2 + \bar{m}_4)
= 4K_d\Delta_n (\bar{m}_2 + \bar{m}_4)
\]

which tends to zero as \( n \to \infty \).

It remains to verify the Lindeberg condition. We again do this by proving that the sums of the conditional fourth moments of the \( Z_{n,k} \) converge to zero.

\[
\mathbb{E}[Z_{n,k}^4|\mathcal{F}_{n,k-1}] = R_n^4\mathbb{E}[X_{n,k}^4|\mathcal{F}_{n,k-1}] = \mathbb{E}[(a_{n,k} - b_{n,k})^4|\mathcal{F}_{n,k-1}]
\leq \mathbb{E}[a_{n,k}^4 - 4a_{n,k}^3 b_{n,k} + 6a_{n,k}^2 b_{n,k}^2 - 4a_{n,k} b_{n,k}^3 + b_{n,k}^4|\mathcal{F}_{n,k-1}]
\]

99
where

\[
a_{n,k} = \left( \int_{I_{n,k}} \varphi(V_i) dt + \int_{I_{n,k}} V_i dB_i \right)^2
\]  

(4.154)

\[
b_{n,k} = \int_{I_{n,k}} V_i^2 dt
\]  

(4.155)

The result will now follow if we can show that \( R_n^4 \sum_{k=1}^{M_n} c_{n,k} \) converges to zero in probability as \( n \to \infty \) where \( c_{n,k} \) can be any of the terms in (4.153). We will consider each of these in turn. In all that follows, we let \( \varepsilon > 0 \) be an arbitrary positive number.

We begin with \( \mathbb{E}[b_{n,k}^4] = (\int_{I_{n,k}} V_i^2 dt)^4 \).

\[
P\left( \left\{ R_n^4 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 > \varepsilon \right\} \right) \leq P\left( \left\{ R_n^4 \Delta_n^3 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^8 dt > \varepsilon \right\} \right)
\]

(4.156)

\[
= P\left( \left\{ R_n^4 \Delta_n^3 \int_{0}^{T_n} V_i^8 dt > \varepsilon \right\} \right)
\]

(4.157)

\[
\leq \frac{\Delta_n^3}{\varepsilon \Delta_n^3 M_n^2} \mathbb{E} \left[ \int_{0}^{T_n} V_i^8 dt \right]
\]

(4.158)

\[
= \frac{T_n \bar{m}_8}{\varepsilon T_n M_n}
\]

(4.159)

\[
= \frac{\bar{m}_8}{\varepsilon M_n}
\]

(4.160)

which converges to zero as \( n \to \infty \).

Next we consider \( \mathbb{E}[a_{n,k} b_{n,k}^3 | \mathcal{F}_{n,k-1}] \). Since

\[
R_n^4 \mathbb{E}[a_{n,k} b_{n,k}^3 | \mathcal{F}_{n,k-1}]
\]

(4.161)

\[
= R_n^4 \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_i) dt + \int_{I_{n,k}} V_i dB_i \right)^2 \cdot \left( \int_{I_{n,k}} V_i^2 dt \right)^3 | \mathcal{F}_{n,k-1} \right]
\]

(4.162)
\[ \leq R_n^4 \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 + \left( \int_{I_{n,k}} V_t dB_t \right)^2 \right] \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^3 \bigg| \mathcal{F}_{n,k-1} \right] \]
\[ = 2R_n^4 \left( \int_{I_{n,k}} V_t^2 dt \right)^3 \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 + \left( \int_{I_{n,k}} V_t dB_t \right)^2 \bigg| \mathcal{F}_{n,k-1} \right] \]  
(4.164)
\[ = 2R_n^4 \left( \int_{I_{n,k}} V_t^2 dt \right)^3 \cdot \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 + \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \right) \]  
(4.165)

we have that

\[ P \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 dt \right)^3 \cdot \left( \int_{I_{n,k}} \varphi(V_t)dt \right)^2 + \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \right) > \varepsilon \right\} \right) \]  
(4.166)
\[ \leq P \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} \varphi^2(V_t)dt \left( \int_{I_{n,k}} V_t^2 dt \right)^3 + \Delta_n^3 \int_{I_{n,k}} V_t^8 dt \right) > \varepsilon \right\} \right) \]  
(4.167)
\[ \leq P \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} K_d(1 + V_t^2)dt \right) \left( \int_{I_{n,k}} V_t^2 dt \right)^3 + R_n^4 \Delta_n^3 \int_0^{T_n} V_t^8 dt > \varepsilon \right\} \right) \]  
(4.168)

The second term in the above expression tends to zero in probability since

\[ P \left( \left\{ R_n^4 \Delta_n^3 \int_0^{T_n} V_t^8 dt > \varepsilon \right\} \right) \]  
(4.169)
\[ \leq \frac{R_n^4 \Delta_n^3}{\varepsilon \mathbb{E}} \left[ \int_0^{T_n} V_t^8 dt \right] \]  
(4.170)
\[ = \frac{m_8}{\varepsilon M_n} \]  
(4.171)

On the other hand the first term in (4.168) goes to zero in probability since

\[ P \left( \left\{ R_n^4 \Delta_n K_d \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 dt \right)^3 + \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \right) > \varepsilon \right\} \right) \]  
(4.172)
which converges to zero as $n \to \infty$. This concludes the proof of the $E[a_{n,k}b_{n,k}^3 | \mathcal{F}_{n,k-1}]$ term.

We now proceed to the term involving $E[a_{n,k}b_{n,k}^2 | \mathcal{F}_{n,k-1}]$, which is handled analogously.

\begin{align*}
R_n^4 E[a_{n,k}^2 b_{n,k}^2 | \mathcal{F}_{n,k}] &= R_n^4 E \left[ \left( \int_{I_{n,k}} \varphi(V_t) dt + \int_{I_{n,k}} V_t dB_t \right)^4 \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^2 | \mathcal{F}_{n,k-1} \right] \\
&\leq 8 R_n^4 \left[ \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^4 \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^2 \\
&\quad + \left( \int_{I_{n,k}} V_t dB_t \right)^4 \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^2 | \mathcal{F}_{n,k-1} \right] \\
&= 8 R_n^4 \left( \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^4 \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^2 + 3 \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \right)
\end{align*}

Now

\begin{align*}
P \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^4 \left( \int_{I_{n,k}} V_t^2 dt \right)^2 + 3 \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \right) > \varepsilon \right\} \right) &\leq P \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^4 \left( \int_{I_{n,k}} V_t^2 dt \right)^2 \right) + 24 R_n^4 \Delta_n^3 \int_0^T V_t^2 dt > \varepsilon \right\} \right)
\end{align*}
The second term in this expression converges to zero since

\[
P\left(\left\{ 24R_n^4 \Delta_n^3 \int_0^T V_t^8 dt > \varepsilon \right\} \right) \leq \frac{24R_n^4 \Delta_n^3}{\varepsilon} \mathbb{E} \left[ \int_0^T V_t^8 dt \right] \quad (4.183)
\]

\[
eq \frac{24 \tilde{m}_8}{\varepsilon M_n} \quad (4.184)
\]

on the other hand, the first term in the expression converges to zero due to the following argument:

\[
P\left(\left\{ 8 R_n^4 \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^4 \cdot \left( \int_{I_{n,k}} V_t^2 dt \right)^2 > \varepsilon \right\} \right) \quad (4.185)
\]

\[
\leq P\left(\left\{ 8 R_n^4 \Delta_n^4 \sum_{k=1}^{M_n} \int_{I_{n,k}} \varphi^4(V_t) dt \cdot \int_{I_{n,k}} V_t^4 dt > \varepsilon \right\} \right) \quad (4.186)
\]

\[
\leq P\left(\left\{ 8 K_d R_n^4 \Delta_n^4 \sum_{k=1}^{M_n} \int_{I_{n,k}} (1 + V_t^2)^2 dt \cdot \int_{I_{n,k}} V_t^4 dt > \varepsilon \right\} \right) \quad (4.187)
\]

\[
\leq P\left(\left\{ 8 K_d R_n^4 \Delta_n^4 \sum_{k=1}^{M_n} \int_{I_{n,k}} 1 + 2 V_t^2 + V_t^4 dt \cdot \int_{I_{n,k}} V_t^4 dt > \varepsilon \right\} \right) \quad (4.188)
\]

\[
= P\left(\left\{ 8 K_d R_n^4 \Delta_n^4 \sum_{k=1}^{M_n} \left( \Delta_n \int_{I_{n,k}} V_t^4 dt \right. \right.

\[
+ 2 \int_{I_{n,k}} V_t^2 dt \cdot \int_{I_{n,k}} V_t^4 dt + \left( \int_{I_{n,k}} V_t^4 dt \right)^2 \varepsilon \left. \right\} \right) \quad (4.189)
\]

\[
\leq P\left(\left\{ 8 K_d R_n^4 \Delta_n^5 \sum_{k=1}^{M_n} \left( \Delta_n \int_{I_{n,k}} V_t^4 + V_t^8 dt + 2 \int_{I_{n,k}} V_t^2 dt \cdot \int_{I_{n,k}} V_t^4 dt \right) > \varepsilon \right\} \right) \quad (4.190)
\]

The first part of (4.190) converges to zero in probability since

\[
P\left(\left\{ 8 K_d R_n^4 \Delta_n^5 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^4 + V_t^8 dt > \varepsilon \right\} \right) \quad (4.191)
\]

\[
= P\left(\left\{ 8 K_d R_n^4 \Delta_n^5 \int_0^T V_t^4 + V_t^8 dt > \varepsilon \right\} \right) \quad (4.192)
\]
Similarly, the second part of (4.190) converges to zero in probability since

\[
P\left( \left\{ 16K_d R_n \Delta_n^4 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_i^2 dt \cdot \int_{I_{n,k}} V_i^4 dt > \varepsilon \right\} \right) \leq \frac{16K_d R_n \Delta_n^4}{\varepsilon} \sum_{k=1}^{M_n} \mathbb{E} \left[ \left( \int_{I_{n,k}} V_i^2 dt \right)^2 \right] \cdot \mathbb{E} \left[ \left( \int_{I_{n,k}} V_i^4 dt \right)^2 \right]^{\frac{1}{2}} \
\leq \frac{16K_d R_n \Delta_n^4}{\varepsilon} \sum_{k=1}^{M_n} \mathbb{E} \left[ \Delta_n \int_{I_{n,k}} V_i^4 dt \right] \cdot \mathbb{E} \left[ \Delta_n \int_{I_{n,k}} V_i^3 dt \right]^{\frac{1}{2}} \
\leq \frac{16K_d R_n \Delta_n^4}{\varepsilon} \sum_{k=1}^{M_n} \sqrt{\Delta_n^2 \bar{m}_4 \cdot \Delta_n^2 \bar{m}_8} \
= \frac{16K_d \sqrt{\bar{m}_4 \bar{m}_8} \Delta_n^2}{\varepsilon M_n} 
\]  

(4.195)  

(4.196)  

(4.197)  

(4.198)  

(4.199)  

Thus the term involving \( \mathbb{E}[a_{n,k}^2 b_{n,k}^2 | \mathcal{F}_{n,k-1}] \) tends to zero. We now turn our attention to the term involving \( \mathbb{E}[a_{n,k}^2 b_{n,k}^2 | \mathcal{F}_{n,k-1}] \). The proof follows the same pattern as those above. We have that

\[
R_n^4 \mathbb{E}[a_{n,k}^3 b_{n,k}^3 | \mathcal{F}_{n,k-1}] \leq 32 R_n^4 \mathbb{E} \left[ \int_{I_{n,k}} V_i^2 dt \cdot \left( \int_{I_{n,k}} \varphi(V_i) dt + \int_{I_{n,k}} V_i dB_t \right)^6 | \mathcal{F}_{n,k-1} \right] \
\leq 32 R_n^4 \mathbb{E} \left[ \int_{I_{n,k}} V_i^2 dt \cdot \left( \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^6 + \left( \int_{I_{n,k}} V_i dB_t \right)^6 \right) | \mathcal{F}_{n,k-1} \right] \
= 32 R_n^4 \int_{I_{n,k}} V_i^2 dt \cdot \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_i) dt \right)^6 + \left( \int_{I_{n,k}} V_i dB_t \right)^6 | \mathcal{F}_{n,k-1} \right] 
\]  

(4.201)  

(4.202)  

(4.203)  

(4.204)
\[
= 32 R_n^4 \left( \int_{I_{n,k}} V_t^2 \, dt - \left( \int_{I_{n,k}} \varphi(V_t) \, dt \right)^6 + 15 \left( \int_{I_{n,k}} V_t^2 \, dt \right)^4 \right) \quad (4.205)
\]

We shall show that the sums of each of the terms in the above expression converge to zero in probability, which implies the result. For the second term we have

\[
\mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \left( \int_{I_{n,k}} V_t^2 \, dt \right)^4 > \varepsilon \right\} \right) \leq \mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^8 \, dt > \varepsilon \right\} \right) \quad (4.206)
\]

\[
\leq \mathbb{P} \left( \left\{ \int_{0}^{T_n} V_t^8 \, dt > \varepsilon \right\} \right) \quad (4.207)
\]

\[
\leq \frac{R_n^4 M_n^3}{\varepsilon} \left[ \int_{0}^{T_n} V_t^8 \, dt \right] \quad (4.208)
\]

\[
= \frac{m^8}{\varepsilon M_n} \quad (4.209)
\]

which converges to zero as \( n \to \infty \). For the first term in (4.205) we have

\[
\mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^2 \, dt \cdot \left( \int_{I_{n,k}} \varphi(V_t) \, dt \right)^6 > \varepsilon \right\} \right) \quad (4.210)
\]

\[
\leq \mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^2 \, dt \cdot \int_{I_{n,k}} \varphi^6(V_t) \, dt > \varepsilon \right\} \right) \quad (4.211)
\]

\[
\leq \mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^2 \, dt \cdot \int_{I_{n,k}} (1 + V_t^2)^3 \, dt > \varepsilon \right\} \right) \quad (4.212)
\]

\[
\leq \mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} (1 + V_t^2)^4 \, dt > \varepsilon \right\} \right) \quad (4.213)
\]

\[
\leq \mathbb{P} \left( \left\{ \sum_{k=1}^{M_n} \int_{I_{n,k}} (1 + V_t^2)^8 \, dt > \varepsilon \right\} \right) \quad (4.214)
\]

\[
\leq \frac{R_n^4 M_n^3}{\varepsilon} \left[ \int_{0}^{T_n} (1 + V_t^2)^8 \, dt \right] \quad (4.215)
\]

\[
= \frac{K^3 \Delta_n^3 \theta_1}{\varepsilon M_n} \quad (4.216)
\]
where

$$\theta_1 = \int_{\mathbb{R}} (1 + u^2)^8 \pi(u) du < \infty \quad (4.217)$$

by assumption 15.

Finally, we turn our attention to the term involving $\mathbb{E}[d_{n,k}^4 | \mathcal{F}_{n,k-1}]$. We have

$$R_n^4 \mathbb{E}[d_{n,k}^4 | \mathcal{F}_{n,k-1}] = R_n^4 \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_t) dt + \int_{I_{n,k}} V_t dB_t \right)^8 | \mathcal{F}_{n,k-1} \right] \quad (4.219)$$

$$\leq 128 R_n^4 \mathbb{E} \left[ \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^8 + \left( \int_{I_{n,k}} V_t dB_t \right)^8 | \mathcal{F}_{n,k-1} \right] \quad (4.220)$$

$$= 128 R_n^4 \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^8 + 105 \left( \int_{I_{n,k}} V_t^2 dt \right)^4 \quad (4.221)$$

Let $C_1 = 128$, $C_2 = 128 \cdot 105$, then

$$\mathbb{P} \left( \left\{ R_n^4 \sum_{k=1}^{M_n} \left( C_1 \left( \int_{I_{n,k}} \varphi(V_t) dt \right)^8 + C_2 \left( \int_{I_{n,k}} V_t^2 dt \right)^4 > \varepsilon \right) \right\} \right) \quad (4.222)$$

$$\leq \mathbb{P} \left( \left\{ R_n^4 \left( C_1 \Delta_n \sum_{k=1}^{M_n} \int_{I_{n,k}} \varphi(V_t) dt + C_2 \Delta_n^3 \sum_{k=1}^{M_n} \int_{I_{n,k}} V_t^2 dt \right) > \varepsilon \right\} \right) \quad (4.223)$$

$$= \mathbb{P} \left( \left\{ R_n^4 \left( C_1 \Delta_n \int_0^{T_n} \varphi(V_t) dt + C_2 \Delta_n^3 \int_0^{T_n} V_t^2 dt \right) > \varepsilon \right\} \right) \quad (4.224)$$

The second term in the above expression tends zero in probability since

$$\mathbb{P} \left( \left\{ R_n^4 \int_0^{T_n} V_t^2 dt > \varepsilon \right\} \right) \quad (4.225)$$

$$\leq \frac{R_n^4 \Delta_n^3 C_2}{\varepsilon} \mathbb{E} \left[ \int_0^{T_n} V_t^2 dt \right] \quad (4.226)$$

$$= \frac{C_2 \bar{m}_8}{\varepsilon M_n} \quad (4.227)$$

106
While the first term in (4.224) tends to zero since

\[
P \left( \left\{ C_1 R_n^4 \Delta_n^7 \int_0^{T_n} \varphi^8(V_t) \, dt > \varepsilon \right\} \right) \leq P \left( \left\{ C_1 K_4 R_n^4 \Delta_n^7 \int_0^{T_n} (1 + V_t^2)^4 \, dt > \varepsilon \right\} \right) \leq \frac{C_1 K_4 R_n^4 \Delta_n^7}{\varepsilon} \mathbb{E} \left[ \int_0^{T_n} (1 + V_t^2)^4 \, dt \right] = \frac{C_1 K_4 \theta_2}{\varepsilon M_n}
\]

where

\[
\theta_2 = \int_{\mathbb{R}} (1 + u^2)^4 \tilde{\pi}(u) \, du < \infty
\]

by assumption 15. This completes the proof.

4.4 Parameter Estimation

The following results are from [GCJL00]. We include them for the purpose of comparison with the results in chapter 3. In this section, we consider a model that is slightly different from the one in the previous sections. In particular

\[
dY_t = \sigma_t dB_t, \quad Y_0 = 0
\]

\[
dV_t = b(V_t) \, dt + a(V_t) \, dW_t, \quad V_0 = \eta
\]
here $\sigma_i^2 = V_i$, and $V_i$ is a positive diffusion process on the interval $(l, r)$. Let $\Delta > 0$, for $k = 1, 2, \ldots , \Delta = (k - 1)\Delta, k\Delta]$, and

$$Z_k = \frac{1}{\sqrt{\Delta}} \int_{I_k} \sigma_i dB_t$$

$$U_k = (\bar{V}_k, V_k)$$

$$V_k = \frac{1}{\Delta} \int_{I_k} V_t dt$$

Theorem 19. We have

- $(U_i, i \geq 1)$ is a strictly stationary Markov chain with state space $(l, r)^2$.
- $(Z_i, i \geq 1)$ is a hidden Markov model with hidden chain $(U_i, i \geq 1)$.

Proof. See [GCJL00].

Consider a Borel measurable function $\phi : \mathbb{R}^d \to \mathbb{R}$, and $h_\phi : \mathbb{R}^d \to \mathbb{R}$ given by (when the following expectation exists and is finite)

$$h_\phi(v_1, \ldots , v_d) = E[\phi(\epsilon_1\sqrt{v_1}, \ldots , \epsilon_d\sqrt{v_d})]$$

where $\epsilon_1, \ldots , \epsilon_d$ are i.i.d. standard normal random variables.

Theorem 20. If $\phi$ is such that $E[h_\phi(\bar{V}_1, \ldots , \bar{V}_d)] < \infty$ then

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n-d} \phi(Z_{i+1}, \ldots , Z_{i+d}) = E[h_\phi(\bar{V}_1, \ldots , \bar{V}_d)]$$

almost surely.

Proof. See [GCJL00].

Theorem 21. Let $\Phi = \phi(Z_{i+1}, \ldots , Z_{i+d})$ If there exists $\delta > 0$ such that

$$E[|\Phi_0|^{2+\delta}] < \infty$$

and

$$\sum_{k=1}^{\infty} \alpha_{\Phi}^{\frac{\delta}{\delta + 1}}(k\Delta)$$

108
\[ \Sigma_\Delta(\varphi, d) = \text{var}(\Phi_0) + 2 \sum_{k=1}^{\infty} \text{cov}(\Phi_0, \Phi_k) \]  
\[ (4.240) \]

is well-defined. If furthermore \( \Sigma_\Delta(\varphi, d) > 0 \), then
\[ \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=0}^{n-d} \left( \Phi_k - \mathbb{E}[h_{\varphi}(\bar{V}_1, \ldots, \bar{V}_d)] \right) = N(0, \Sigma_\Delta(\varphi, d)) \]
\[ (4.241) \]

where the convergence in the above equation is in distribution.

**Proof.** See [GCJL00]. \( \square \)

**Corollary 7.** Suppose that \( \Phi : \mathbb{R}^d \to \mathbb{R}^p \). Suppose that there exists \( \delta > 0 \) such that
\[ \sum_{k=1}^{\infty} \alpha_k^{2+\delta} (k \Delta) < \infty \quad \text{and} \quad \mathbb{E}[|\Phi_0^k|^{2+\delta}] < \infty \quad k = 1, \ldots, p \]
\[ (4.242) \]

Let
\[ \Sigma_\Delta(\Phi^k, \Phi^l; d) = \text{cov}(\Phi_0^k, \Phi_0^l) + \sum_{j=1}^{\infty} (\text{cov}(\Phi_0^k, \Phi_j^l) + \text{cov}(\Phi_j^k, \Phi_0^l)) \]
\[ (4.243) \]

Then \( \Sigma_\Delta(\Phi^k, \Phi^l; d) \) is well-defined. Consider the matrix \( \Sigma_\Delta = (\Sigma_\Delta(\Phi^k, \Phi^l))_{1 \leq k, l \leq p} \). If \( \Sigma_\Delta \) is positive definite then
\[ \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{j=0}^{n-d} \begin{pmatrix} \Phi_0^1 - \mathbb{E}[h_{\varphi}(\bar{V}_1, \ldots, \bar{V}_d)] \\ \vdots \\ \Phi_0^p - \mathbb{E}[h_{\varphi}(\bar{V}_1, \ldots, \bar{V}_d)] \end{pmatrix} \to N(0, \Sigma_\Delta) \]
\[ (4.244) \]

where the convergence in the above equation is in distribution.

**Proof.** See [GCJL00]. \( \square \)
Proposition 15. Assume that there exists $\delta > 0$ such that

$$
\sum_{k=1}^{\infty} \alpha_k \tilde{\nu}^2 (k\Delta) < \infty \tag{4.245}
$$

1. For $\varphi(z) = z^{2p}$, if $E[V_0^{2p(1+\frac{1}{2})}] < \infty$ then

$$
\Sigma_\Delta(z^{2p}, 1) = C_{2p}^2 \left( \text{var}(\tilde{V}_1^p) + 2 \sum_{k=1}^{\infty} \text{cov}(\tilde{V}_1^p, \tilde{V}_{k+1}) \right) + (C_{4p} - C_{2p}^2) E[\tilde{V}_1^{2p}] 
$$

(4.246)

2. For $\varphi(z_1, z_2) = z_1^{2p} z_2^{2q}$, if $q \leq p$ and $E[V_0^{p(1+\frac{1}{2})}] < \infty$

$$
\Sigma_\Delta(z_1^{2p}, z_2^{2q}, 2) = C_{2p}^2 C_{2q}^2 \left( \text{var}(\tilde{V}_1^p, \tilde{V}_2^q) + 2 \sum_{k=1}^{\infty} \text{cov}(\tilde{V}_1^p \tilde{V}_2^q, \tilde{V}_{k+1}^p \tilde{V}_{k+2}^q) \right) 
$$

$$
+ (C_{4p} C_{4q} - C_{2p}^2 C_{2q}^2) E[\tilde{V}_1^{2p} \tilde{V}_2^{2q}] 
$$

$$
+ 2C_{2p} C_{2q} (C_{2(p+q)} - C_{2p} C_{2q}) E[\tilde{V}_1^p \tilde{V}_2^q \tilde{V}_3^p] \tag{4.247}
$$

where in the above, we denote

$$
C_k = E[|\varepsilon_1|^k] \tag{4.248}
$$

Proof. See [GCJL00].

In particular, in [GCJL00] the authors consider models of the form

$$
dY_t = \sigma_t dB_t \quad Y_0 = 0 \tag{4.249}
$$

$$
dV_t = \alpha(\beta - V_t) dt + \alpha(V_t, \theta) dW_t \quad V_0 = \xi, \; V_t = \sigma_t^2 \tag{4.250}
$$

with $\alpha, \beta > 0$, and show how their results can be applied to estimate all the parameters, and calculate the asymptotic distribution of the estimation error.
Appendix A

Results from Optimization

A.1 Results from Optimization

In this section, we follow [Roc70] and [RW98].

Definition 25. A convex function \( f : \mathbb{R}^n \to \mathbb{R} \) is proper there exists \( x \in \mathbb{R}^n \) such that \( f(x) < \infty \).

Definition 26. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function, and \( x \in \mathbb{R}^n \). Then \( d \in \mathbb{R}^n \) is said to be a subgradient of \( f \) at \( x \) if

\[
f(z) \geq f(x) + \langle d, z - x \rangle \quad \forall z \in \mathbb{R}^n
\]  

(A.1)

The set of all subgradients of \( f \) at \( x \) is called the subdifferential of \( f \) at \( x \) and is denoted by \( \partial f(x) \).

We observe that if \( f \) is differentiable at \( x \) then \( \partial f(x) = \{\nabla f(x)\} \).
An "ordinary convex program" is an optimization problem of the form

\[
\begin{align*}
\min_{x \in C} f_0(x) & \\
f_i(x) & \leq 0 \quad i = 1, \ldots, r \quad (A.2) \\
f_i(x) & = 0 \quad i = r + 1, \ldots, m \quad (A.3)
\end{align*}
\]

where \( C \subseteq \mathbb{R}^n \) is a convex set, \( f_0 : \mathbb{R}^n \to \mathbb{R} \) is a proper convex function with \( \text{dom}(f_0) = C \), \( f_1, \ldots, f_r : \mathbb{R}^n \to \mathbb{R} \) are proper convex functions with

\[
\text{ri}(C) \subseteq \text{ri}(\text{dom}(f_i)) \quad C \subseteq \text{dom}(f_i) \quad (A.5)
\]

and \( f_i : \mathbb{R}^n \to \mathbb{R} \) is affine on \( C \) for \( i = r + 1, \ldots, m \). We allow the cases \( r = 0 \) (no inequality constraints) and \( r = m \) (no equality constraints).

**Definition 27.** A Kuhn-Tucker vector for the problem (A.2) is a vector \( u \in \mathbb{R}^m \) such that \( u_i \geq 0, i = 1, \ldots, r \) and the infimum of the proper convex function \( f_0 + \sum_{i=1}^{m} u_i f_i \) is finite and equal to the optimal value in (A.2).

**Definition 28.** The Lagrangian of the convex program (A.2) is the function \( L : \mathbb{R}^m \times \mathbb{R}^n \) defined by

\[
L(u, x) = \begin{cases} 
  f_0(x) + \sum_{i=1}^{m} u_i f_i(x) & u \in E_r, x \in C \\
  -\infty & u \notin E_r, x \in C \\
  \infty & x \notin C 
\end{cases} \quad (A.6)
\]

where

\[
E_r = \{ u \in \mathbb{R}^m | u_i \geq 0, i = 1, \ldots, r \} \quad (A.7)
\]

The variable \( u_i \) is referred to as the Lagrange multiplier associated with the \( i \)th constraint in (A.2). A point \( (u^*, x^*) \in \mathbb{R}^m \times \mathbb{R}^n \) is said to be a saddle-point for \( L \).
if
\[
L(u, x^*) \leq L(u^*, x^*) \leq L(u^*, x) \quad \forall u \in \mathbb{R}^m, x \in \mathbb{R}^n
\] (A.8)

**Theorem 22.** Consider the problem (A.2). Let \( u^* \in \mathbb{R}^m, x^* \in \mathbb{R}^n \). In order that \( u^* \) be a Kuhn-Tucker vector for (A.2) and \( x^* \) be an optimal solution to (A.2) it is necessary and sufficient that \((u^*, x^*)\) be a saddle-point of the Lagrangian \( L \).
Moreover, this conditions holds if and only if

- \( u^*_i \geq 0, f_i(x^*) \leq 0 \) and \( u^*_i f_i(x^*) = 0, i = 1, \ldots, r \).
- \( f_i(x^*) = 0 \) for \( i = r + 1, \ldots, m \)
- \( 0 \in \{ \partial f_0(x^*) + u^*_1 \partial f_1(x^*) + \cdots + u^*_m \partial f_m(x^*) \} \)

**Proof.** See ([Roc70], pp. 281-282). \qed

## A.2 Stability in Nonlinear Programming

In this section, we consider stability results for the optimal solutions and Lagrange multipliers of nonlinear programs. Our presentation is based on [Fia83]. The general form of nonlinear program that we consider is:

\[
\min_{x} f(x) \quad (A.9)
\]
\[
g_j(x) \leq 0 \quad j = 1, \ldots, p \quad (A.10)
\]
\[
h_i(x) = 0 \quad i = 1, \ldots, q \quad (A.11)
\]

**Definition 29.** A feasible point \( x^* \) of (A.9) is said to be a strict local minimizing point of (A.9) if there is a neighbourhood \( N \) of \( x^* \) such that there does not exist any feasible \( x \in N \) with \( x \neq x^* \) such that \( f(x) \leq f(x^*) \).
We consider the problem of determining a local solution of:

\[
\min_x f(x, \varepsilon) \tag{A.12}
\]

\[
g_j(x, \varepsilon) \leq 0 \quad j = 1, \ldots, p \tag{A.13}
\]

\[
h_i(x, \varepsilon) = 0 \quad i = 1, \ldots, q \tag{A.14}
\]

where \(x \in \mathbb{R}^n\) and \(\varepsilon \in \mathbb{R}^k\). The Lagrangian of the above problem is:

\[
L(x, \lambda, \mu, \varepsilon) = f(x, \varepsilon) + \sum_{j=1}^{p} \lambda_j g_j(x, \varepsilon) + \sum_{i=1}^{q} h_i(x, \varepsilon) \tag{A.15}
\]

**Proposition 16.** If \(f(\cdot, 0), g_j(\cdot, 0), h_i(\cdot, 0) \in C^2(N(x^*))\) where \(N(x^*)\) is some neighbourhood of \(x^*\), then \(x^*\) is a strict local minimum of the problem (A.12) with \(\varepsilon = 0\) if there exist \(\lambda^* \in \mathbb{R}^p\) and \(\mu^* \in \mathbb{R}^q\) such that the first order Kuhn-Tucker conditions hold:

\[
g_j(x^*, 0) \leq 0 \quad j = 1, \ldots, p \tag{A.16}
\]

\[
h_i(x^*, 0) = 0 \quad i = 1, \ldots, q \tag{A.17}
\]

\[
\lambda^*_j g_j(x^*, 0) = 0 \quad j = 1, \ldots, p \tag{A.18}
\]

\[
\lambda^*_j \geq 0 \quad j = 1, \ldots, p \tag{A.19}
\]

and

\[
\nabla_x L(x^*, \lambda^*, \mu^*, 0) = \nabla_x f(x^*, 0) + \sum_{j=1}^{p} \lambda^*_j \nabla_{zz} g_j(x^*, 0) + \sum_{i=1}^{q} \mu^*_i \nabla_{zz} h_i(x^*, 0) \tag{A.20}
\]

\[
= 0 \quad \tag{A.21}
\]

and further the following second-order constraint condition holds. For all \(z \neq 0\) such
we have that

\begin{align}
\nabla_z g_j(x^*, 0) \cdot z &\leq 0 \quad \forall j \text{ s.t. } g_j(x^*, 0) = 0 \\
\nabla_z g_j(x^*, 0) \cdot z &= 0 \quad \forall j \text{ s.t. } \lambda_j^* > 0 \\
\nabla_z h_i(x^*, 0) \cdot z &= 0 \quad i = 1, \ldots, q
\end{align}

where \( \nabla_x^2 L \) is the Hessian of \( L \).

**Proof.** See [Fia83].

**Theorem 23.** Let \( x^* \) be a feasible point of (A.9). Assume also that

- There exists a neighbourhood \( N \) of \( (x^*, 0) \) such that \( f, g_j, h_i \) are twice continuously differentiable with respect to \( x \) on \( N \) and \( \nabla_x f, \nabla_x g_j, \nabla_x h_i \) are continuously differentiable in \( \varepsilon \) in \( N \).

- The second order sufficient conditions from the proposition hold at \( x^* \), with associated vectors of Lagrange multipliers \( \lambda^*, \mu^* \).

- The vectors

\begin{align}
\nabla_z g_j(x^*, 0) &\quad j \in I(x^*) \\
\nabla_z h_i(x^*, 0) &\quad i = 1, \ldots, q
\end{align}

are linearly independent, where \( I(x^*) = \{ j \text{ s.t. } 1 \leq j \leq p, g_j(x^*, 0) = 0 \} \)

- \( \lambda_j > 0 \) when \( g_j(x^*, 0) = 0 \) for \( j = 1, \ldots, p \) (i.e. strict complementary slackness holds)

Then
1. $x^*$ is a local isolated minimizing point of (A.12) and the associated Lagrange multipliers $\lambda^*, \mu^*$ are unique.

2. There exists a neighbourhood $N_0$ of 0 and in $N_0$ a once continuously differentiable function

$$y(\varepsilon) = (x(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon))^T$$

(A.28)

satisfying the second order sufficient conditions for a local minimum of $P(\varepsilon)$ such that

$$y(0) = (x^*, \lambda^*, \mu^*)^T = y^*$$

(A.29)

and $x(\varepsilon)$ is a locally unique local minimum of $P(\varepsilon)$ with associated unique Lagrange multipliers $\lambda(\varepsilon), \mu(\varepsilon)$.

3. There exists a $\delta > 0$ such that for $\|\varepsilon\| < \delta$ the set of binding constraints $I(x(\varepsilon)) = I(x^*)$, and strict complementary slackness holds and the binding constraint gradients $\nabla x g_j(x(\varepsilon), \varepsilon), j \in I(x(\varepsilon))$ are linearly independent.

Proof. See [Fia83].

Corollary 8. If the conditions of the theorem hold, with the respective assumed orders of differentiability being $r - 1$ more than those assumed in the theorem, with $r \geq 1$ then

$$y(\varepsilon) = (x(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon))^T \in C^p(N_0)$$

(A.30)

for some neighbourhood $N_0$ of $\varepsilon = 0$. If the problem functions are analytic in $(x, \varepsilon)$ in a neighbourhood $N_a$ of $(x^*, 0)$ then $y(\varepsilon)$ is analytic in a neighbourhood of 0.

Proof. See [Fia83].

From the proof of the above theorem the following can also be seen.
Proposition 17. Using the notation from the theorem,

\[ M(\varepsilon)Y(\varepsilon) = N(\varepsilon) \]  \hspace{1cm} (A.31)

Here we employ the following notation:

\[
M(\varepsilon) = \begin{pmatrix}
H_zL & \nabla_z g_1 & \cdots & \nabla_z g_p & \nabla_z h_1 & \cdots & \nabla_z h_q \\
\lambda_1(\nabla_z g_1)^T & g_1 & 0 & 0 \\
\vdots & 0 & \ddots & 0 & 0 \\
\lambda_p(\nabla_z g_p)^T & 0 & \cdots & g_p \\
(\nabla_z h_1)^T & 0 & 0 \\
\vdots & 0 & 0 \\
(\nabla_z h_q)^T & 0 & 0
\end{pmatrix}
\]  \hspace{1cm} (A.32)

where all functions are evaluated at \( \varepsilon \) and \( H_zL \) is the Hessian of \( L \) with respect to \( x \):

\[
(H_zL)_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}(x(\varepsilon), \lambda(\varepsilon), \mu(\varepsilon), \varepsilon)
\]  \hspace{1cm} (A.33)

Also

\[
Y(\varepsilon) = \begin{pmatrix}
J_x \\
J_\lambda \\
J_\mu
\end{pmatrix}
\]  \hspace{1cm} (A.34)
where $J_z, J_\lambda$ and $J_\mu$ are the Jacobians of $x, \lambda$ and $\mu$ with respect to $\varepsilon$ and

$$
N(\varepsilon) = 
\begin{pmatrix}
-\partial_{xx} L \\
-\lambda_1 (\nabla_{x} g_1)^T \\
\vdots \\
-\lambda_p (\nabla_{x} g_p)^T \\
-(\nabla_{x} h_1)^T \\
\vdots \\
-(\nabla_{x} h_q)^T
\end{pmatrix}
$$

where

$$(\partial_{xx} L)_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}$$

There exists $\delta > 0$ such that $M(\varepsilon)$ is nonsingular for $\|\varepsilon\| < \delta$ so

$$Y(\varepsilon) = M(\varepsilon)^{-1} N(\varepsilon)$$

Also

$$\lim_{\varepsilon \to 0} Y(\varepsilon) = Y(0)$$

Proof. See [Fia83].

Corollary 9. Under the assumptions of the theorem

$$
\begin{pmatrix}
x(\varepsilon) \\
\lambda(\varepsilon) \\
\mu(\varepsilon)
\end{pmatrix} = 
\begin{pmatrix}
x^* \\
\lambda^* \\
\mu^*
\end{pmatrix} + M(0)^{-1} N(0) \varepsilon + o(\|\varepsilon\|)
$$
Appendix B

Results from Stochastic Analysis

B.1 Results from Stochastic Analysis

B.1.1 Itô’s Lemma

Definition 30. A continuous semimartingale is a stochastic process \( (X_t, \mathcal{F}_t)_{t \in \mathbb{R}^+} \)
which may be written as

\[
X = X_0 + M + A
\]  

(B.1)

where \( X_0 \) is \( \mathcal{F}_0 \) measurable, \( M \) is a continuous local martingale null at 0, and \( A \) is a continuous adapted finite variation process null at 0.

Theorem 24. (Itô’s Lemma). Let \( f \in C^2(\mathbb{R}^n, \mathbb{R}) \) and let \( X \) be a continuous semimartingale in \( \mathbb{R}^n \). Then

\[
f(X_t) - f(X_0) = \sum_{k=1}^{n} \int \frac{\partial f}{\partial x_k}(X_s) dX^k_s + \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_k \partial x_j}(X_s) d[X^k, X^j]_s \]  

(B.2)

Proof. See ([RW00b], pp 60-62). \( \square \)
B.2 Central Limit Theorems

Here we follow [HH80]. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space, and let \(\{X_{n,k} : k = 1, \ldots, k_n\}\) be an array of random variables. Let \(\{\mathcal{F}_{n,k} : 0 \leq k \leq k_n\}\) be any triangular array of sub \(\sigma\)-algebras of \(\mathcal{F}\) such that for each \(n\) and \(1 \leq k \leq k_n\), \(\mathcal{F}_{n,k-1} \subseteq \mathcal{F}_{n,k}\), and \(X_{n,k}\) is \(\mathcal{F}_{n,k}\) measurable for all \(n, k\).

**Theorem 25.** Suppose that

\[
\sum_{k=1}^{k_n} \mathbb{E}[X_{n,k}^2 I_{\{X_{n,k} > \varepsilon\}} | \mathcal{F}_{n,k-1}] \to 0 \quad \forall \varepsilon > 0 \tag{B.3}
\]

\[
\sum_{k=1}^{k_n} \mathbb{E}[X_{n,k} | \mathcal{F}_{n,k-1}] \to 0 \tag{B.4}
\]

\[
\sum_{k=1}^{k_n} \text{Var}[X_{n,k} | \mathcal{F}_{n,k-1}] \to 1 \tag{B.5}
\]

where all limits are in probability as \(n \to \infty\). Then

\[
\sum_{k=1}^{k_n} X_{n,k} = S_n \to N(0, 1) \tag{B.6}
\]

where the convergence is in distribution.

**Proof.** See [HH80]. \(\square\)

We observe that the "conditional Lindeberg condition" (B.3) is implied by any moment condition

\[
\sum_{k=1}^{k_n} \mathbb{E}[|X_{n,k}|^{2+\delta} | \mathcal{F}_{n,k-1}] \to 0 \tag{B.7}
\]

where the convergence is in probability and \(\delta > 0\). (This is because on \(\{X_{n,k} > \varepsilon\}\), we have that \(X^2 < e^{-\delta X^{2+\delta}}\).
Definition 31. If

\[ \mathbb{E}[X_{n,k} | \mathcal{F}_{n,k-1}] = 0 \]  \hspace{1cm} (B.8)

then we say that \( \{X_{n,k}, \mathcal{F}_{n,k}\} \) is a martingale difference array. In this case, we further denote:

\[ S_n = \sum_{t=1}^{k_n} X_{n,k} \]  \hspace{1cm} (B.9)

The following is one version of the martingale central limit theorem:

Theorem 26. Suppose that \( \{X_{n,k}, \mathcal{F}_{n,k}, 1 \leq t \leq k_n\} \) is a martingale difference array satisfying:

\begin{itemize}
  \item \( \sum_{t=1}^{k_n} \mathbb{E} \left[ X_{n,k}^2 \mathbb{1}(|X_{n,k}| > \varepsilon) \big| \mathcal{F}_{n,k-1} \right] \to 0 \) as \( n \to \infty \) for all \( \varepsilon > 0 \).
  \item \( \sum_{t=1}^{k_n} \mathbb{E}[X_{n,k}^2 | \mathcal{F}_{n,k-1}] \to 1 \) in probability.
\end{itemize}

Then

\[ S_n = \sum_{t=1}^{k_n} X_{n,k} \to N(0, 1) \]  \hspace{1cm} (B.10)

in distribution.

The first condition in the above theorem is often referred to as the Lindeberg condition and is implied by any result of the form

\[ \sum_{t=1}^{k_n} \mathbb{E}[X_{n,k}^{2p} | \mathcal{F}_{n,k-1}] \to 0 \]  \hspace{1cm} (B.11)

in probability for \( p > 1 \) (since on \( \{X_{n,k} > \varepsilon\} \), we have \( X^2 < \varepsilon^{-\delta} X^{2+\delta} \) for any \( \delta > 0 \)).
B.3 Large Deviations for Martingale Difference Sequences

The following is from [LV99]. Since it is unpublished, we include the proof.

**Theorem 27.** Let \((X_k)\) be a martingale difference sequence where \(X_k \in L^p, 2 \leq p < \infty, \) and \(\|X_k\|_p < M < \infty\) for all \(k\). Let \(x > 0\). Then if \(S_n = \sum_{k=1}^n X_k\)

\[
\mathbb{P}\{|S_n| > nx\} \leq (18pq^{1/p})^p \frac{M^p}{x^p} \cdot \frac{1}{n^q} \tag{B.12}
\]

where \(q\) is the real number such that \(\frac{1}{p} + \frac{1}{q} = 1\).

**Proof.** Let \(x > 0\). Using Burkholder's inequality ([HH80], theorem 2.10) and Jensen's inequality,

\[
\mathbb{E}\left[\left|\sum_{k=1}^n X_k\right|^p\right] \leq (18p\sqrt{q})^p \mathbb{E}\left[\sum_{k=1}^n |X_k|^q\right] \tag{B.13}
\]

\[
\leq n^{q/p - 1}(18p\sqrt{q})^p \mathbb{E}\left[\sum_{k=1}^n |X_k|^p\right] \tag{B.14}
\]

\[
\leq n^{q/p}(18p\sqrt{q})^p M^p \tag{B.15}
\]

The result now follows as

\[
\mathbb{P}(S_n > nx) \leq \frac{\mathbb{E}[|S_n|^p]}{x^p n^p} \leq \frac{(18p\sqrt{q})^p}{x^p} \cdot \frac{M^p}{n^q} \tag{B.16}
\]

\[\square\]

B.4 Quadratic Variation

The following is from [Mét82].

**Definition 32.** Let \(\mathcal{X}\) be a topological space. A stochastic process \((X, \{\mathcal{F}_t\}_{t \in \mathbb{R}_+})\) with state-space \(\mathcal{X}\) is said to be regular if \(X\) is adapted and all its paths have left
Theorem 28. (From [Mét82]). Let $Z$ be a real regular semimartingale and $[Z]$ be the following process:

$$[Z]_t = Z_t^2 - Z_0^2 - 2 \int_{(0,t)} Z_s dZ_s$$

(B.17)

- Then $[Z]$ is increasing adapted right-continuous and $[Z]_0 = 0$.

- Let us consider for every $n \in \mathbb{N}$ an increasing sequence $\Pi^n = \{0 = \tau^n_1 < \tau^n_2 < \ldots < \tau^n_k \ldots \}$ of stopping times with the properties

1. $\forall n \lim_{k \to \infty} \tau^n_k = +\infty$ a.s.
2. $\lim_{n \to \infty} \sup_k (\tau^n_{k+1} - \tau^n_k) = 0$ a.s.

We set

$$v_n(t, Z) = \sum_k (Z_{\tau^n_{k+1} \wedge t} - Z_{\tau^n_k \wedge t})^2$$

(B.18)

Then the sequence $(v_n(t, Z))_{n \geq 0}$ of random variables converges in probability for every fixed $t$ to the random variable $[Z]_t$. There exists a subsequence $(v_{n_k}(\cdot, Z))_{k \geq 0}$ of processes, the paths of which converge uniformly on any bounded interval to the paths of $[Z]$.

Proof. See ([Mét82], pp. 175-178).

\[ \square \]

B.5 Ergodic Theorems

Theorem 29. (Birkhoff’s Ergodic Theorem) If $(\Omega, \mathcal{F}, \mu)$ is a $\sigma$-finite measure space and $\{\tau_t, t \geq 0\}$ is a measure-preserving measurable semiflow, $f \in L_1(\mu)$, then

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(\tau_t \omega) dt = \int_{\Omega} f d\mu$$

(B.19)
holds $\mu$ a.e.

\textit{Proof.} See ([Kre85] pp. 7-10).

\textbf{Theorem 30.} (Wiener's Local Ergodic Theorem) If $(\Omega, \mathcal{F}, \mu)$ is a $\sigma$-finite measure space and $\{\tau_t, t \geq 0\}$ is a measure-preserving measurable semiflow, $f \in L_1(\mu)$, then

$$ \lim_{\varepsilon \to 0^+} \varepsilon^{-1} \int_0^\varepsilon f(\tau_\varepsilon \omega) d\alpha = f(\omega) $$

(B.20)

holds $\mu$ a.e.


\textbf{B.5.1 Markov Processes, Diffusions and Ergodic Theory}

\textbf{Definition 33.} A stochastic process $\{X_t : t \in \mathbb{R}\}$ is said to be strictly stationary if for all $n \in \mathbb{N}$, $t_1, \ldots, t_n \in \mathbb{R}$ and all $h > 0$

$$(X_{t_1}, \ldots, X_{t_n}) \sim (X_{t_1+h}, \ldots, X_{t_n+h})$$

(B.21)

where "\sim" denotes equality in distribution.

\textbf{Theorem 31.} Let $X$ be a regular diffusion on $\mathbb{R}$ in natural scale, with speed measure $m$. Suppose that $f, g : \mathbb{R} \to \mathbb{R}_+$ are measurable functions such that

$$ \int (f(x) + g(x)) m(dx) < \infty \quad \int g(x) m(dx) \neq 0 $$

(B.22)

then

$$ \lim_{t \to \infty} \frac{\int_0^t f(X_s) ds}{\int_0^t g(X_s) ds} = \frac{\int f(x) m(dx)}{\int g(x) m(dx)} $$

(B.23)

almost surely.
The following is based on [GCJL00]. Consider a probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) and a time-homogeneous Markov process \(X_t\) with state space \((\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))\). Assume that \(X\) has continuous sample paths and transition function \(P_t(x, dy)\). Assume that the process has a stationary distribution \(\pi\) and that \(X_0 \sim \pi\).

Suppose that \(\mathcal{F}_t = \sigma(X_s, s \leq t)\) and \(\mathcal{F} = \sigma(X_t, t \geq 0)\). For any function in \(L_1(\pi)\) define

\[
P_t f(x) = \int f(y) P_t(x, dy) = \mathbb{E}[f(X_t)|X_0 = x]
\]

Consider the space \(C(\mathbb{R}_+, \mathbb{R}^k)\) of continuous functions, with the topology of uniform convergence on compact subsets of \(\mathbb{R}^+\) and let \(\mathcal{B}_C\) be the corresponding Borel \(\sigma\)-algebra. Consider also the shift operator:

\[
\theta_t(x)(\cdot) = x(t + \cdot)
\]

Denote \(X(\omega)\) the sample path corresponding to \(\omega \in \Omega\) (so \(X\) is a random variable in the space \(C(\mathbb{R}_+, \mathbb{R}^k)\)). From strict stationarity for all \(B \in \mathcal{C}\):

\[
\mathbb{P}(X \in B) = \mathbb{P}(\theta_t(X) \in B)
\]

so we have that \(\theta_t\) is measure-preserving. Define the shift-invariant \(\sigma\)-algebra.

\[
\mathcal{J} = \{(X \in B)| B \in \mathcal{C}, \forall t, B = \theta_t^{-1}(B)\}
\]

A strictly stationary process \(X\) is said to be ergodic if \(\mathcal{J}\) is \(\mathbb{P}\) trivial (all sets in \(\mathcal{J}\) have probability either 1 or 0). If this is the case, then by Birkhoff's ergodic theorem when \(f \in L^1(\pi)\)

\[
\frac{1}{T} \int_0^T f(X_s)ds \to \mathbb{E}[f(X_0)|\mathcal{J}] = \int f d\pi
\]
where the convergence is almost everywhere, as $T \to \infty$. 
Bibliography


