Electronic Thesis and Dissertation Repository

2-3-2021 10:00 AM

The Mean-Reverting 4/2 Stochastic Volatility Model: Properties And Financial Applications

Zhenxian Gong, *The University of Western Ontario*

Supervisor: Escobar-Anel, Marcos, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Statistics and Actuarial Sciences
© Zhenxian Gong 2021

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Applied Statistics Commons, Other Applied Mathematics Commons, Probability Commons, and the Statistical Models Commons

Recommended Citation

Gong, Zhenxian, "The Mean-Reverting 4/2 Stochastic Volatility Model: Properties And Financial Applications" (2021). *Electronic Thesis and Dissertation Repository*. 7686.
https://ir.lib.uwo.ca/etd/7686

Abstract

In this thesis, we target commodity and volatility index markets, and develop a novel stochastic volatility model that incorporates mean-reverting property and 4/2 stochastic volatility process. Commodities and volatility indexes have been shown to be mean-reverting. The 4/2 stochastic volatility process integrates two processes that have contrary behaviors. As a result, not only is the 4/2 stochastic volatility process able to reproduce "smile" and "skew", but also model the asset price time series very well even in the most extreme situation like financial crisis where the assets' prices become highly volatile. In the one-dimensional study, we derive a semi-closed form conditional characteristic function (c.f.) for our model and propose two feasible approximation approaches. Numerical study shows that the approximations are accurate in a large region of the parametric space. With the approximations, we are able to price options using c.f. based pricing algorithms which outperform Monte Carlo simulation in speed. We study two estimation methods for the 4/2 stochastic volatility process. We show numerically that the estimation methods produce consistent estimators. By applying our model to empirical data, especially volatility index data, we find evidence for an embedded 4/2 stochastic volatility process, which also can be seen by observing drastic spikes from data. In option pricing applications, we realize there can be 20% difference on option prices if the underlying model is specified by a 4/2 stochastic volatility process as opposed to a 1/2 stochastic volatility process. We further test our approximation approaches by comparing the option prices generated by Monte Carlo simulation and those obtained from Fast Fourier Transform using the approximated c.f., the error turns out to be negligible. We next consider a generalized multivariate model based on our one-dimensional mean-reverting 4/2 stochastic volatility model and principal component stochastic volatility framework to capture the behavior of multiple commodities or volatility indexes. The structure enables us to express the model in terms of a linear combination of independent one-dimensional mean-reverting 4/2 stochastic volatility processes. We find a quasi-closed form c.f. for the generalized model and analytic approximations of the c.f. under certain model assumptions. We propose a scaling factor to connect empirical variance series to theoretical variances and estimate the parameters with the methodology developed for our one-dimensional mean-reverting 4/2 stochastic volatility model. The effectiveness of our approximation approaches is supported by comparing the Value-at-Risk (VaR) values of a portfolio of two risky assets and a cash account using Monte Carlo simulations and the approximated distributions.

Keywords: Historical Volatility, Implied Volatility, VIX, Commodity, Multivariate Models, Mean-Reverting Property, Stochastic Volatility, Risk Management, Option Pricing

Summary for Lay Audience

Financial markets and instruments are continuously evolving, displaying new and more refined stylized facts. This requires regular reviews and empirical evaluations of advanced models. There is evidence in literature that supports stochastic volatility models over constant volatility models in capturing stylized facts such as "smile' and "skew" presented in implied volatility surfaces. In this thesis, we target commodity and volatility index markets, and develop a novel stochastic volatility model that incorporates mean-reverting property and 4/2 stochastic volatility process. Commodities and volatility indexes have been proved to be mean-reverting, which means their prices tend to revert to their long term mean over time; the 4/2 stochastic volatility process is able to reproduce "smile" and "skew", but also model the asset price time series very well even in situations like financial crisis where the assets' prices become unstable. In the study of single asset, we study theoretical properties of our model and propose two approximation approaches. With the approximations, we are able to price options using fast pricing algorithms instead of simulations. We are also the first to study two estimation methods for the 4/2 stochastic volatility process. Our estimation methods can produce accurate estimators when sample size is large enough. By applying our model to empirical data, especially volatility indexes data, we find evidence for an embedded 4/2 stochastic volatility process. In option pricing applications, we realize there can be 20% difference on option prices if the underlying model is incorrectly specified. We compare the option prices generated by simulation and those obtained from approximation approaches, the error turns out to be negligible. We next consider a multi-asset setting as an extension of our single asset study aiming to capture the behavior of multiple commodities or volatility indexes. The model structure enables us to express one asset as a linear combination of independent artificial "assets". We propose a scaling factor to connect empirical variance series to theoretical variances and estimate the parameters. As an application, we further construct a portfolio that consists of two risky assets and a cash account and calculate Value-at-Risk using simulations and the approximated distributions.

Acknowledgments

I would like to take this opportunity to express my greatest gratitude to the people who have helped me directly or indirectly in my journey to achieve a Doctor of Philosophy degree in statistics at The University of Western Ontario.

It was not an easy decision for me to make on whether or not to continue my academic career as a PhD student. My master supervisor at Queen's University, Dr. Bingshu Chen, strongly encouraged me to accept the challenge of doing research at PhD level. He convinced me that this is a unique experience that few people have the chance to have, and I would benefit from it in my future life no matter what I do after that. So I decided to take the challenge and even more, I decided to challenge myself to do a PhD in an area which I am not familiar with, but I am passionate about–Financial Modelling.

I still remember the day when I received an email from Prof. Marcos Escobar-Anel. I spent the Chinese New Year holiday with my family at the beginning of the winter term of my master program–something I would not do in winter terms when I was a student–because my sister had a type of tumor that was rare to see in clinical studies. One day morning, when I woke up I saw the email from Prof. Marcos Escobar-Anel expressing his interest in taking me as his PhD student. This email was definitely the best new year gift I had ever had and also a relief for me and my family after witnessing all the suffer my sister had been through. I took Prof. Marcos Escobar-Anel's offer knowing that it would be a difficult journey for me.

Prof. Marcos Escobar-Anel carefully designed my PhD career path to make sure that I would be successful. He guided me through all the challenging parts of my research patiently. Whenever I had new ideas, he always had an intuition on the consequences my ideas could lead to, which ensured I was on the right direction. His style of supervision helped me build a strong intuition. Other than research skills, he also offered invaluable advice on the life of a PhD sharing his own PhD experience. I would not make this far without Prof. Marcos Escobar-Anel's support. He is not just my PhD supervisor, but also my mentor and friend.

I also would like to thank Prof. Sebastian Ferrando from Ryerson University, Prof. Rogemar Mamon, Prof. Lars Stentoft and Prof. Xingfu Zou, who served on my PhD thesis committee and provided remarkable comments on this thesis. Without Ms. Miranda Fullerton's excellent work, my thesis defense would not have been so smooth. A big thanks to my friends Junhe Chen, Yiyang Chen, Boquan Cheng, Yuyang Cheng, Xing Gu, Wenjun Jiang, Yifan Li, Ang Li, Yuying Li, Yang Miao and Guqian Zhao who supported me through my PhD career. It is my honor to study and live with a group of top researchers and life lovers.

As always, my family backs me up for my decisions. They did not push me to live a routine life as most people would. They always understand me and see things more clearly than me. When I am lost, they always have a way to make me see things through. They have been through hard time for the past couple years, but still supporting me to continue my PhD career. I owe them more than just "Thank you". A special thanks to Shanshan Liu who always stands by my side.

Declaration

I hereby declare that this thesis proposal and its preliminary results incorporate materials that are direct results of my efforts.

The content of Chapter 3 is based on the paper *Marcos Escobar-Anel and Zhenxian Gong. The mean-reverting 4/2 stochastic volatility model: Properties and financial applications. Applied Stochastic Models in Business and Industry. 2020; 36: 836– 856*

The content of Chapter 4 is partially selected from the paper *Yuyang Cheng, Marcos Escobar-Anel and Zhenxian Gong. Generalized mean-reverting 4/2 Factor Model. Journal of Risk and Financial Management; 2019, 12.4: 159.*

The content of Chapter 5 is based on a paper in progress titled *Multivariate Mean-Reverting 4/2 Stochastic Volatility Model*.

With the exception of the guidance on formulating modelling frameworks from Dr. Marcos Escobar-Anel and contributions by Yuyang Cheng, I certify that this document is a product of my own work.

This research was conducted from June 2017 to present under the supervision of Dr. Marcos Escobar-Anel at the University of Western Ontario.

London, Ontario

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

In 1900, the publication of a PhD thesis "The theory of speculation" by Louis de Bachelier marked a milestone in quantitative finance. Bachelier's masterpiece is recognized as a breakthrough in finance and arguably considered to be the beginning of modern finance. Bachelier creatively introduced the concept of Brownian motion from physics to finance as an essential tool for the study of stochastic processes in his thesis. Despite the fact that Bachelier's study of Brownian motion in finance was revolutionary, the theory was largely overlooked for decades. In the middle of the twentieth century, the quantitative finance field was ignited by the rediscovery of Bachelier's work.

In 1951, the famous Japanese mathematician, Kiyoshi Ito, published his pioneering work "On stochastic differential equations" , which honored Ito as the founding father of stochastic calculus, placing one more layer of foundation in quantitative finance on top of Bachelier's theory. In Ito's work, he, in particular, demonstrated the rule that governs differentiating a time-dependent function of a stochastic process. The rule is known as Ito's lemma. Ito's lemma is so powerful that one can derive not only new stochastic differential equations and therefore new stochastic processes from existing ones, but also the differential equations describing the value of financial derivatives. Later in the 1970s, a series of game-changing works were carried out by Fischer Black, Myron Scholes, Robert Merton, which revolutionize the world of finance. In particular the Nobel prize winning Black-Scholes' model [21] helped establishing quantitative finance as a new branch in mathematics. This new branch applies mathematical and statistical techniques to solving problems in finance.

Since the 1970s financial markets have seen explosive growth in both volume and types of financial products. With the new techniques, novel products like exotic and multi-asset options can be engineered from traditional financial products using quantitative finance principles. As a result of increase in the complexity of financial markets, the time series of assets' prices in financial markets have developed new stylized facts. Existing quantitative finance methodologies are thus challenged by the complex problems emerging from the evolution of financial markets. In this thesis, we focus on one of these problems, which concerns modeling of instantaneous volatility using an advanced stochastic volatility framework for commodities and volatility indexes (VIX, VSTOXX etc.). We propose and study a Schwartz's one-factor mean-reverting model, as per [122] with a 4/2 stochastic volatility in chapter 3. Before we disclose

more details about our model, we first briefly present the motivation of our research—why a mean-reverting 4/2 stochastic volatility model is appealing and necessary.

The novelty of our research is that we are the first to consider a Schwartz one-factor model with a state-of-the-art stochastic volatility process—the 4/2 process. The 4/2 stochastic volatility model was first published in 2016 and originally used to model stocks. Let's pause here for a moment and spend some time reviewing some history to help readers understand why 4/2 stochastic volatility is necessary, and the strength this unique process possesses.

The renowned Black-Scholes model [21] assumes a constant volatility that drives the process of a stock price; the model leads to a closed-form and easy-to-compute option price formula thanks to its simple construction. On the other hand, the simple construction of Black-Scholes model also brings a number of limitations of which the most well-known is its inability in explaining volatility "smile". Volatility "smile" together with volatility "skew" are characteristics of the implied volatility surface. We would expect a flat implied volatility surface given different, say call option, strike prices and maturities if the Black-Scholes's model were compatible with empirical data. Reality shows that this is not true. The shape of the surface is somewhat U-shaped ("smile") or downward sloping ("skew") with respect to strike prices. To better explain the nature of the problem, Heston proposed a stochastic volatility framework [80] assuming a Cox-Ingersoll-Ross (CIR) process that drives the volatility of a stock in 1993. The introduction of stochastic volatility increases the complexity of the underlying model, nonetheless an analytic option pricing formula can still be derived. More importantly, Heston model is able to reproduce the volatility "smile" that is not possible in the case of Black-Scholes model. Undoubtedly, Heston's work supports the supremacy of stochastic volatility over constant volatility assumption.

The story does not end with Heston's model. Empirical data shows that Heston model would predict a mild trend in the price time series when "spikes" are actually observed in the stock price time series due to high volatility in the market. The reason for the converse behavior predicted by Heston's model is that Heston model needs a large volatility of volatility parameter to capture the "spikes", which in turn jeopardizes the Feller's condition for the variance process of Heston's model. This explains why calibration of Heston's model to market data suggests violation of Feller's condition. In 1997, Heston and Platen independently proposed the 3/2 stochastic volatility framework as an alternative to Heston's model, see [81] and [117]. 3/2 model considers an inverse CIR process—the power of the diffusion term in an inverse CIR process is 3/2—instead of the original CIR process for the stochastic volatility component of the model. Thus, Heston's model's shortcoming when facing "spikes" in stock price time series is solved by the 3/2 process, see [49] for a detailed study on the comparison between Heston's model and 3/2 model. In particular, the author mentions that calibration to empirical data yields a large (-99.0%) correlation parameter for the 3/2 model, which suggests the 3/2 model is less flexible in capturing smaller leverage than Heston's model; another difficulty faced by the 3/2 model is in the application of pricing options on realized variance, which involves first and second moments of the integrated variance. It is shown in [49] that under the 3/2 model, computation of the first moment is challenging and the second moment does not have closed-form representation. So it is clear that we cannot use only Heston's model or the 3/2 model for all the

situations, we need both. Recently, Grasselli remarkably combines Heston's model and the 3/2 model together introducing the "ultimate" 4/2 stochastic volatility [76]. The word "ultimate" here does not mean Grasselli's model is the final model in the world of stochastic volatility, but the state-of-the-art model extending the Heston's family because it inherits the advantages of both Heston and 3/2 models, and still maintains an analytical option pricing formula. Regrettably, it suffers a similar pitfall as 3/2 models does: a risk-neutral measure equivalent to the historical measure may not exist. A solution to this problem is to use the Benchmark approach for derivative pricing as suggested in [76]. The Benchmark approach does not require the existence of an equivalent risk-neutral measure, it is can be done using a portfolio of numéraire under historical probability measure, see [10] for details.

Grasselli's 4/2 model is tractable and flexible for equity modeling and equity derivative pricing. Nevertheless, the model is not appropriate for modeling commodity prices or volatility indexes as it lacks mean-reversion, which is considered the principal property evidenced in commodity prices and volatility indexes. We will review in the following context the progress in modeling commodities and volatility indexes.

Commodity is a popular and widely studied asset class in finance. Arguably, the most celebrated work in the area is a paper written by Schwartz [122] in which the author proposed three models for commodities: one-factor model, two-factor model, and three-factor model. In all of Schwartz's models, volatility is assumed constant; due to mounting evidence supporting stochastic volatility in commodities, a one-factor Schwartz model with stochastic volatility was proposed by Eydeland and Geman in 1998 [63]. More recently Benth [16] studied two advanced stochastic volatility models for the one-factor Schwartz model using Barndoff-Nielsen and Shephard [12]'s O-U process (BNS) and a square-root process for the variance. The latter was inspired on the Heston's model [80] from the equity modeling literature. The case for stochastic volatility is nowadays trivial to make. Empirically, CBOE publishes volatility indexes based on futures for commodity market, which are becoming popular for traders. The volatility indexes can be understood as implied volatilities of options on the underlying commodities. The randomness of these indexes is a clear indication that the constant volatility assumption is violated. Therefore, the appropriateness of stochastic volatility for commodities has been proved empirically and supported theoretically.

This allows to introduce another asset class—volatility indexes, like VIX and VSTOXX. This asset class is quite recent. Introduced by CBOE in 1993, VIX is the first volatility index that can be traded in the market. Current VIX is based on S&P 500, and estimates expected volatility by averaging the weighted prices of S&P 500 puts and calls over a wide range of strike prices. Hence, VIX is model-free and only depends on option prices. Derivatives on VIX did not receive enough attention before the financial crisis in 2008; now these derivatives are largely traded (about 2 million contracts exchanged daily) by investors everyday as a new way to hedge risk. Eurex introduced a similar volatility index for Euro STOXX 50 index–VSTOXX in 2005, and soon after Eurex launched VSTOXX futures contracts. Current data shows that the trading volume of VSTOXX derivatives is around 100,000 contracts daily. Other major markets have also set up their own volatility indexes, such as VIXC in Canada and VNKY for Nikkei 225 in Japan. The modeling methodology of a volatility index can be classified into

two categories: indirect modeling and direct modeling. The indirect approach (eg. [130] and [98]) basically models the volatility index in combination and as a byproduct of the underlying equity index. The direct method treats the volatility index as the underlying asset of interest independently of the underlying asset that inspires the volatility index. Fernandes et.al [66] studied the time series property of VIX by considering heterogeneous autoregressive (HAR) model. Whaley [128] noted that volatility tends to follow a mean-reverting process, which was confirmed by Goard and Mazur [74]. In their work, not only is the mean-reverting property of VIX confirmed, the volatility of VIX is also shown to be stochastic. In an early work done by Kaeck and Alexander [87], they studied one-factor affine and non-affine models as well as a two-factor stochastic volatility model, see [52] for a detailed treatment on affine models. Kaeck and Alexander [87] concluded that for one-factor models, non-affine structure captures the extreme behavior of VIX better than affine models. This finding coincides with Goard and Mazur [74] as the 3/2 model is non-affine. Although the non-affine model considered in [87] has diffusion term proportional to the VIX level or log VIX, they also emphasized that a two-factor stochastic volatility model is superior than one-factor models for the behavior of VIX. The historical data of VVIX undeniably proves the randomness of the volatility of VIX, therefore confirming the point that a two-factor model for VIX shall be better.

The following points outline the contributions of this thesis:

- We study a new stochastic process that incorporates mean-reversion and the state-of-the-art 4/2 stochastic volatility.

- We provide a quasi-closed-form representation for the conditional characteristic function (c.f.) that makes computations more efficient.

- An accurate closed-form approximation of the conditional c.f. is obtained for two relevant cases: mean-reverting 1/2 stochastic volatility, and the mean-reverting 4/2 model with no leverage.

- Two economically meaningful changes of measure are studied, conditions are provided for well-definiteness.

- Two estimation methods are studied and the consistency of the estimators is shown numerically.

- The appropriateness of the model is confirmed by fitting two financial asset classes: commodity prices and volatility indexes.

- Our model is compared to the nested Heston model, showing a significant impact of up to 20% on option prices.

- We generalize our mean-revering 4/2 stochastic volatility model by considering a multivariate mean-reverting 4/2 stochastic volatility factor model to capture either multivariate commodity behavior or multivariate volatility indexes.

- Our setting for the multivariate model reduces the dimension of the parametric space making parameters identifiable, permitting the use of popular estimation methods.

- The presence of independent common and intrinsic factors in the multivariate model, each with its own stochastic volatility, enables an elegant separable structure for characteristic functions (c.f.s) while capturing several stylized facts, such as: stochastic volatility, stochastic correlation among stocks, co-movements in the variances, multiple factors in the volatilities, and spillover effects.

- We derive a semi-closed form c.f. for the multivariate model.

- We propose analytic approximations of the semi-closed form c.f..

- A scaling factor inspired by findings in literature is proposed to match empirical variance series to theoretical variances in a multivariate setting; we then estimate the parameters following the same methodology developed for the one-dimensional mean-reverting 4/2 stochastic volatility model.

- We construct a portfolio of two risky assets and a cash account based on the multivariate model. Value-at-Risk (VaR) is computed using both simulation-based method and approximated distributions. The results support our approximation approaches in multi-dimensions.

In this chapter, we first give a conceptual review of volatilities in section 1.1 introducing three major volatility categories: historical volatility, implied volatility and integrated volatility. In section 1.2, we briefly introduce commodities and commodity markets as well as trading of commodities. In section 1.3, we provide some mathematical background related to the main findings of the thesis.

The rest of the thesis is organized as follows: chapter 2 offers a collection of models relevant to the thesis including existing models and some interesting new models that have either been studied recently or not been studied at all. Chapter 3 is a comprehensive treatment of our mean-reverting 4/2 stochastic volatility model—the main topic and major contribution of this thesis. We present our model and study its properties and financial applications. We also propose approximation approaches that result in closed-form approximation of the characteristic function. Chapter 4 is an extension of the mean-revering 4/2 stochastic volatility model to multi-dimensions taking into account two categories of factors that drive the diffusion of the model: common factors and intrinsic factors. The two categories of factors represent systematic risk and unsystematic risk respectively. We consider a convenient structure for the covariance matrix for the multivariate model in line with [57], [58], [59] and [60] by considering Principal Component Analysis (PCA). Thanks to principal component decomposition, we are able to boil down the multivariate structure of the model in terms of independent one dimensional mean-reverting 4/2 stochastic volatility processes that we have studied in chapter 3. We derive a semi-closed form c.f. for the multivariate model and changes of measure with conditions that guarantee they are well defined. In the application, we study the effects of 3/2 process and the factors on European call options and risk measures such as VaR and Expected Shortfall (ES). In chapter 5, we consider a sub model of the generalized 4/2 factor model in chapter 4, which consists of only common factors (no intrinsic factor, along the lines of PCA). In addition to the results obtained in chapter 4, we further apply our approximation approaches

developed in chapter 3 to tackle the semi-closed c.f. with analytic functions. Another important ingredient of chapter 5 is parameter estimation in a higher dimension under principal component framework. The estimation method we use is a generalization of the one described in chapter 3. In application part of chapter 5, we construct a portfolio with two risky assets and a cash account with flexible trading strategies. We then compute VaR's for the portfolio using two methods: Monte Carlo simulation and approximated density functions. The results support the appropriateness of our approximation approaches.

1.1 Volatility

In finance, volatility measures the variation of return for an underlying asset using standard deviation of the return. Therefore, volatility reflects the riskiness of the underlying asset. High volatility refers to large dispersion around the mean of the return and small volatility indicates that the return is relatively stable at a level. That is why volatility plays a key role in financial risk management. Depending on the objective, volatility can be subdivided into several categories. In this section, we focus on three major categories of volatility: historical volatility, implied volatility and integrated volatility.

1.1.1 Historical Volatility

Historical volatility, also interchangeable with realized volatility[1], measures the standard deviation of asset price over a period of time, normally 10 days to 30 days. Hence, historical volatility is also referred as statistical volatility.

Historical volatility indicates how volatile the price of an asset is, but it is not simply calculating the standard deviation of asset price over a period of time. In fact, the calculation concerns the return of the asset, which is the same as percentage price changes of the asset. Given daily prices, the returns are calculated for each day with respect to the price on the previous day. A standard deviation is derived from these returns. The standard deviation is then multiplied by the square root of the number of trading days in a calendar year (eg. $\sqrt{250}$) to yield the historical volatility because historical volatility is typically an annualized quantity. The probabilistic property of historical volatility has been studied in details in [13].

Historical volatility can also be thought of as percentage. For example, a stock with 10% historical volatility is considered to have very low volatility; while a stock with 80% historical volatility is considered very volatile. Therefore, historical volatility indicates the degree of riskiness for an asset. Derivatives are also affected by the level of riskiness of an asset, in particular, derivatives' premiums for risky assets are larger than they are for lesser risky assets.

1.1.2 Implied Volatility

Unlike historical volatility, implied volatility, as its name suggests, is a volatility that can not be observed or calculated directly, but it is "hidden" in the option prices. Sometimes the move-

---

[1]In the following sections, when we talk about realized volatility, we also mean historical volatility.

ment of option prices does not coincide with investors' expectation when the underlying asset prices make anticipated moves. To see this, we take bearish stock market as an example. In a bearish stock market, the value of stocks decreases. Normally, when the value of the underlying asset decreases, the prices of related options are also affected in the same way. In a bearish market, opposite to the trend in stocks, a surge of the options is often observed. Scenarios like this can be explained by demand and supply of options—demand for options increases in order to protect investment in stocks, thus higher implied volatility. In practice, implied volatility is used as a guidance for investors to perceive how the market is going to react to a series of events happening during the life of the option.

Option strike price, underlying asset price and maturity of an option are observable and considered key fundamentals for option pricing. Implied volatility can be extracted from option price given the above ingredients under concrete modeling assumptions. Implied volatility is believed to be more informative than historical volatility regarding the variations of underlying asset prices, because option prices reflect investors' perception of overall market performance in the future. Hence, implied volatility reflects investors' expectation of the market.

Not only does implied volatility better assess the magnitude of the future movement in the underlying asset price, it is also considered as an indicator of the risk associated with the underlying asset. When the underlying asset is believed to be risky, which results in potential large fluctuation in its price, investors will purchase the options on the asset by a large amount to hedge the risk, and the corresponding implied volatility increases due to surge in the options demand. On the other hand, volatility index provides a direct way to visualize and hedge against implied volatility. For a risky asset, we expect high transaction volume of futures and options on the volatility index for the asset for hedging purpose, which in turn drives up the implied volatility directly. Conversely, options on assets with relatively low risk attracts lower attention from investors, thus the corresponding implied volatility is low.

The implied volatility can be computed using either model-based or model-free methods. We briefly discuss both methodologies in the following sections.

1.1.2.1   Model-Based Implied Volatility

Model-based implied volatility is straight forward: the implied volatility is calculated from a given model. For instance, the model used to derive implied volatility could have both constant volatility and an analytic option price formula. Given an option price, the constant volatility that is traced back from the option price is the implied volatility. One popular model that is widely used for equity market is the Black-Scholes model. The pricing formula for a European call option has a closed form representation:

$$C(X(t), t) = X(t)e^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2),$$

$$d_1 = \frac{ln(\frac{X(t)}{K}) + (r - q + \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}}, d_2 = \frac{ln(\frac{X(t)}{K}) + (r - q - \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}},$$

where $X(t)$ is the stock price at time t, K is the strike price, r is the risk-neutral interest rate, q is the continuous dividend rate of the stock and T is the maturity of the option. For every call option price from the market, we can solve the equation for σ that corresponds to each call option price. The theory for the uniqueness of σ is that the Black-Scholes call price is a monotonically increasing function in σ and also continuous (positive vega), so the function is invertible. The σ value is regarded as implied volatility. The same can be done for European put option, and we expect the same value of σ for the same strike and maturity due to put-call parity for European options.

In this case, if we plot the implied volatilities against strike price, we will see either a U-shape curve with a minimum at-the-money or a downward sloping skew. The former is called "smile", and latter is known as "skew". Both "smile" and "skew" are commonly observed facts in option markets. These facts prove that the volatility is not constant as prescribed by Black-Scholes's model. Therefore, stochastic volatility models such as Heston model are developed to capture these facts.

1.1.2.2   Model-Free Implied Volatility

Model-free implied volatility is calculated without assuming any models for the underlying asset, in other words, it solely depends on the inputs (the characteristics of an option) from the market. Hence, it does not suffer from "the inconsistency of forecasting changes in volatility from a model based on constant volatility" [23].

Model-free implied volatility was first studied by Dupire [55], Neuberger [112], Demeterfi et.al (DDKZ) [42]. In the early work, deterministic volatility is considered. However, later research shows that deterministic volatility is quite restrictive, and thus non-deterministic volatility is preferred [54] [53] [26]. Britten-Jones and Neuberger [23] then proposed a reverse method that takes a complete set of option prices to extract as much information as possible about the underlying price process. Based on [23], the integrated squared volatility is expressed as the integration of call option prices:

$$\mathbb{E}\left[\int_0^T \left(\frac{dX(t)}{X(t)}\right)^2\right] = 2\int_0^\infty \frac{C(T, K) - (X(0) - K)^+}{K^2}dK, \tag{1.1}$$

where, $X(t)$ denotes the price of the underlying asset at time t. $C(T, K)$ denotes the European call option price at time T, and K is the strike price. In practice, since strike prices are finite

and discrete, a discrete approximation to the formula above is:

$$\mathbb{E}\left[\int_0^T \left(\frac{dX(t)}{X(t)}\right)^2\right] \approx 2\sum_{i=1}^n \frac{C(T, K_i) - (X(0) - K_i)^+}{K_i^2}\Delta K,$$

$$\Delta K = \frac{K_{max} - K_{min}}{n}, K_i = K_{min} + i\Delta K. \tag{1.2}$$

As a consequence of the approximation, the formula has two types of errors: truncation error and discretization error. The first type of error arises from the limited range of strikes prices $[K_{min}, K_{max}]$ instead of $(0, \infty)$; the second type derives from discretization of the integral. Jiang and Tian [85], however, discovered that the two types of errors can be ignored under certain conditions: if $K_{min} < X(0) - 2\sigma X(0)$ or $K_{max} > X(0) + 2\sigma X(0)$, then truncation error is negligible as the truncation errors diminish when truncation points K_{min} and K_{max} move away from initial price $X(0)$; if $\Delta K < 0.35\sigma X(0)$, discretization error can be ignored. In both cases, σ is the realized volatility of the underlying asset for the remaining maturity of the option. Jiang and Tian [85] also suggested cubic splines to fit the implied volatilities because not all strike prices in the range are accessible in the market.

According to Jiang and Tian [85], model-free implied volatility "aggregates information across options with different strike prices and should be informationally more efficient". They also asserted that Equation 1.1 is also satisfied when the underlying asset price process has jumps. More importantly, the model-free implied volatility is found to contain all the information of the Black-Scholes implied volatility and of past realized volatility, thus it is a more efficient forecast for future realized volatility [85]. We can see that model-free implied volatility is advantageous over model-based implied volatility except the two errors when using Formula 1.2 as noted by Jiang and Tian, which are truncation error and discretization error [85]. Hence, it beneficial for investors to hedge risk if a list of standardized model-free implied volatilities is available for different financial markets.

It is worth noting a useful reference on implied volatility: volatility indexes. In 1993, CBOE published the first volatility index—VIX—to reflect implied volatility based on S&P 100 index. Today VIX futures and options are largely traded by investors as a way to hedge risk. VIX was first developed based on Whaley's work [127] in 1993 as a model-based implied volatility based on Black-Scholes' model and prices of S&P 100 options. The calculation of VIX took into account only eight option contracts to determine the value. The 1993 approach is limited in that, first, it uses a limited number of American option contracts; second, it focuses on S&P 100 index that has a much narrower scope than S&P 500 index in terms of completeness of the stock market. In 2003, CBOE updated the definition and calculation of VIX based on a discrete approximation to DDKZ, which is in fact equivalent to Britten-Jones and Neuberger [23] theoretically [86]. The new methodology was developed in collaboration with Goldman Sachs. The methodology is based on S&P 500 and estimates expected volatility by averaging the weighted prices of S&P 500 puts and calls over a wide range of strike prices. So current VIX is model-free and only depends on option prices. Nowadays there is a broad range of volatility indexes available pretty much covering from stock markets to commodity markets.

1.1.3 Integrated Volatility

We talked about realized volatility and implied volatility above. Volatility family has two other members that are as important: instantaneous volatility and integrated volatility. Instantaneous volatility refers to the volatility of an asset over a short period of time, as this period approaches 0. We denote instantaneous volatility as $\sigma(t)$. Integrated volatility is then defined as $\sigma^*(t) = \int_0^t \sigma(s)ds$. An intuitive interpretation of the definition is that integrated volatility is the "total" or "true" volatility from time 0 to time t. Studying integrated volatility is crucial, especially for stochastic volatility model. The instantaneous volatility component, $\sigma(t)$, which is the object modeled by most stochastic volatility models, is the derivative with respect to (w.r.t) time of the integrated volatility. This is why it is important to understand the integrated volatility. In reality what brings challenges to practitioners is that exact calculation of integrated volatility is difficult since financial time series are discrete in time. The financial statistic that is closely related to integrated volatility is realized volatility. To see this, suppose we record n log asset prices $Y(t_i)$ for $t_i \in [t-1, t]$ and calculate the returns as $r_{t_i}^{(n)} = Y\left(t_i\right) - Y\left(t_i - \frac{1}{n}\right)$, $t_i = t-1 + \frac{i}{n}$ for $i = 1, \ldots, n$. By definition, the realized volatility over the period $[t-1, t]$ is given by $RV_t(n) = \sum_{i=1}^{n} r_{t-1+\frac{i}{n}}^{(n)2}$. Barndorff-Nielsen and Shepard [13] [14] [15] provide the asymptotic distribution for $RV_t(n)$ over the period $[t-1, t]$:

$$\sqrt{n}\left(RV_t(n) - \int_{t-1}^{t} \sigma^2(s)ds\right) \rightarrow N\left(0, 2 \int_{t-1}^{t} \sigma^4(s)ds\right).$$

Hence, realized volatility is a discrete time analog to integrated volatility as well as an estimator for integrated volatility, thus there exists estimation error. Meddahi [105] further explored the relationship between realized volatility and integrated volatility.

In calculus, to estimate an integral as accurately as possible, taking smaller increment is a common practice. In finance, the idea refers to using high frequency data to estimate integrated volatility because high frequency data is taken at very small time interval for example every 30 seconds, 5 minutes or 10 minutes throughout a trading day. There have been numerous publications on estimation of integrated volatility using high frequency data since 2000 when high frequency financial data first became available. Some of the articles outline an approximation of the integrated volatility using the sum of squared high frequency intraday returns, which is the realized volatility, see [5] and [13]. However, studies also found out that using high frequency data is not necessarily the best approach because of the microstructure effects in the market, see [6] and [8].

Microstructure or microstructure noise is a property that is hidden in high frequency data. The cause of microstructure consists of several sources, one of the best known source is the bid-ask spread. To see this intuitively, imagine our goal is to estimate intraday volatility. On an hourly price time series, a point means the price at that specific hour. If we narrow down the scope of time to minute, we will observe another time series of prices at every minute, which is more refined comparing to hourly data. In statistics, larger data size usually gives better results; however, calculating volatility using high frequency leads to unstable realized volatility and tends to overestimate the integrated volatility. In literature, one solution is to use data sampled

at lower frequency, for example 15 to 30 minutes for intraday volatility estimation. To solve the microstructure in high frequency data so that the data can be used for the original purpose, Zhang et.al provided a better solution in [131], which uses high frequency data to produce an unbiased and consistent estimator of integrated volatility.

1.1.4 Relationship Among The Volatilities

So far we have briefly introduced three types of volatilities that attract the most attention both academically and practically. In this section, we will talk about the relationship between them.

- **Realized Volatility vs. Integrated Volatility** In section 1.1.3, we talk about the relationship between realized volatility and integrated volatility while we introduce the concept of integrated volatility. Realized volatility is directly calculated from data. Depending on the frequency of the data being used to calculate realized volatility, the results will be interpreted as, for example, the intraday, daily or annual volatility. It reflects the magnitude of the price change in the given period of time. On the other hand, realized volatility can be treated as least squares estimator or maximum likelihood estimator in the context of parametric volatility. We have learned that realized volatility is also an estimator of the derivative with respect to time of the integrated volatility–the "true" volatility, which is important in stochastic volatility models. When high frequency data became available, estimation of integrated volatility using realized volatility based on high frequency data was common practice until microstructure effects were discovered. To reduce these effects, a solution against using high frequency data was proposed in [131]. Hence the connection between realized volatility and integrated volatility is tight.

- **Realized Volatility vs. Implied Volatility** Implied volatility reflects investors expectation on the future trend of market, hence it is tied to investors' perception of the risk in the market in the future. Realized volatility, on the other hand, more directly reflects current risk in the market. So a natural connection is that implied volatility can be used to predict future realized volatility. However, in early research, the results on the forecasting power of implied volatility are mixed. Christensen and Prabhala [35] carefully reviewed previous work and argued in their paper that implied volatility is capable of predicting future volatility, confirming that in an efficient option market, implied volatility should be an efficient forecast of future realized volatility. Other factors that may affect the forecasting power of implied volatility are moneyness of options and risk premiums, see [33]. With the introduction of VIX, we have a simpler and straight forward way to assess the relationship between realized volatility and implied volatility. Normally implied volatility is larger than realized volatility [77]. When compared with S&P 500 realized volatility, VIX tends to be larger than realized volatility resulting in overestimation of future volatility.

We have seen that realized volatility plays a crucial role in finance due to its connection to integrated volatility and implied volatility. There is a rich literature on the properties of realized volatility. It is also the center of the popular stochastic volatility models. Understanding realized volatility helps us understand the market and its risk.

1.2   Commodity

Nowadays financial markets consist of many different kinds of sub-markets. Among all the markets, fixed income market (eg. corporate bonds, government bonds) and stock market draw most attention from the investors due to their liquidity and volume. Commodity market has a history almost as long as human civilization. Although relatively small today in terms of market volume, it is closely related to our life and affects other major financial markets. If an abnormal movement is witnessed in commodity prices one can foresee crisis and impact in the economic cycle. For example, a surge of oil prices caused a global economic crisis in 1970s. Hence, commodity is a vital and widely studied asset classes in finance.

1.2.1   Definition

A basic economic definition of commodity is that a commodity is a physical good attributable to a natural resource that is tradable and supplied without substantial differentiation by the general public. The concept of commodities varies from industry to industry. For example, corn can be made into menthol and starch, these are considered commodities to, say, medical and food industries. Hence, it is not easy to classify all the commodities into certain groups based on their physical characteristics; however, the commodities used in the initial phase of production cycle are from two types of sources:

- **Soft Commodities:** Goods that are agricultural products or livestock. For example: rice, corn and wheat.

- **Hard Commodities:** Typically natural resources that are mined or extracted. For example: gold, iron ore and oil.

The products made from these commodities are also considered commodities to different industries. Note that being a physical good is not enough to be a commodity. One can notice another important characteristic of commodities described in definition is "without substantial differentiation by the general public" or "fungibility", which means the value added to the commodities due to differentiation based on non-competitive factors is minimum. In other words, the pricing of commodities highly depend on supply and demand and competition as well as logistics, so the production location and manufacturer of the commodities do not affect the pricing process as much as they do in a monopolistic market[2]. Then it is worth talking about a concept called commoditization.

Some goods are not commodities when they are first produced due to scarcity and high technology, so the marginal profits of these good are quite large. As technology evolves, entry level to the industry lowers and more competitors emerge, which brings down the marginal profits early players enjoyed and calls for competition. Eventually these good become commodities. Over-the-counter drugs and health products like fish oils are best examples of commoditization.

---

[2]We in no way mean to say that there is absolutely no differentiation among the same type of commodities. For a single type of commodity, say steel, the quality of iron ore and the technology a manufacturer uses and even the reputation of the manufacturer or preference of clients are factors that differentiate one steel manufacturer from other steel manufacturers. However, the steel price is a result of competition.

1.2.2 Financial Markets and Trading

So far we have talked about the definition of commodities. From raw materials to final products, there are many kinds of commodities. Among these commodities, hard and soft commodities draw most attention from investors as they directly determine the prices of all other commodities. In general, investing in commodities is time consuming and requires specialties, that is why for a long time trading these commodities has been a highly specialized activity. Only some specialized companies or specialized departments of general organizations get involved in the trading of commodities.

Today individual investors also have access to trading commodities thanks to the evolution of financial markets. Trading commodities has various ways, such as investing in the stocks of commodity companies and trading ETF's of the commodities, among which trading in physical (spot) markets and derivative markets are considered the most common and direct ways.

Physical markets, as the name suggests, is where commodities exchange hands physically, so the physical markets are also called exchanges. The exchanges normally only carry a few commodities with some specializing in only one kind of commodities like London Metal Exchange. Before the transaction happens, vendors and buyers mainly negotiate about the prices since commodities are standardized. Other details like means of delivery and financial transaction are also negotiated, but prices are the main concern to both parties. If an agreement on a certain price is reached, the transaction is carried out. Upon delivery, if the quality of the commodities is better than standard, the vendor asks for a higher price, otherwise the buyer asks for a discount [119]. Depending on the agreement, the quality of the commodities may be inspected upon delivery; under a more standard and strict agreement, such inspection is not necessary.

Derivative markets consist of derivatives whose values are derived from underlying commodities. Futures contracts, swaps, exchange-traded commodities and forward contracts are most traded commodity derivatives. The derivatives are not only used as a convenient way to invest in commodities, it is also an important way to hedge risk. As we introduced above, the prices of commodities depend largely on supply and demand. The demand for oil, for example, is inelastic, so buyers of oil purchase forwards/futures in case suppliers (eg. OPEC) decides to reduce supply or there is political crisis that may boost the oil price. On the other hand, suppliers like corn farmers use derivatives in the case of overproduction.

In the end, the prices of commodities are also affected by physical markets and derivative markets since the vendors and buyers from both markets complete the picture of supply and demand of certain commodities. Even tough most commodities see large fluctuations in price [119], unlike stocks, the prices of commodities tend to revert to their overall average over long time. This is a well known property for commodities, hence it is a key to select models for modeling of commodities.

Next, we briefly talk about modeling of commodities. We then review some popular models to model commodities in chapter 2.

For modeling of stocks, the most important and famous models are no doubt Black-Scholes model [21] and Heston model [80]. Arguably, the most celebrated work in the area is a paper written by Schwartz [122], in which the author proposed three models for commodities: one-factor model, two-factor model, and three-factor model. In all of Schwartz's models, volatility is assumed constant; not long thereafter, due to mounting evidence supporting stochastic volatility in commodities, a one-factor Schwartz model with stochastic volatility was proposed by Eydeland and Geman in 1998 [63]. More recently Benth [16] studied two advanced stochastic volatility models for the one-factor Schwartz model using Barndoff-Nielsen and Shephard's Non-Gaussian O-U process (BNS) [12] and a square-root process for the variance. The latter is also known as the Heston model [80] in the equity modeling literature. Empirically, CBOE publishes volatility indexes based on futures for commodity market. The volatility indexes, which are perceived as implied volatilities by investors, show strong volatile nature. The variations in the value of the indexes is an indicator that constant volatility assumption is violated. Therefore, stochastic volatility models are ideal candidates in modeling commodities.

1.3 Mathematical Background

In this section, we cover mathematical preliminaries that are related to our research. In section 1.3.1, we give brief introduction to probability spaces and stochastic processes; section 1.3.2 is about characteristic function; in section 1.3.3 we describe Ito's calculus, core in mathematical finance; section 1.3.4 is a introduction to Feynman-Kac theorem; in section 1.3.5, we present the concept of risk measures. We consult the following literature as references: [7], [17], [65], [88], [93], [101], [100], [104], [114], [125], [129] for the materials presented in this section.

1.3.1 Probability Spaces and Stochastic Process

Definition 1.3.1 (Probability Space) *If Ω is a given non-empty set, then a σ-algebra $\mathcal{F}$ on Ω is a family of subsets of Ω with the following properties:*

- *$\emptyset \in \mathcal{F}$,*

- *$A \in \mathcal{F} \implies A^C \in \mathcal{F}$, where $A^C = \Omega \setminus A$ is the complement of $A \in \Omega$,*

- *$A_1, A_2, \cdots \in \mathcal{F} \implies \mathbf{A} := \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.*

The pair $(\Omega, \mathcal{F})$ is called a measurable space. A probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \Omega \to [0, 1]$ such that

- *$\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$,*

- *If $A_1, A_2, \cdots \in \mathcal{F}$ are pairwise disjoint (i.e $A_i \cap A_j = \emptyset$ if $i \neq j$), then*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space. A set $A_0 \in \mathcal{F}$ with $\mathbb{P}(A_0) = 0$ is called a null set. $(\Omega, \mathcal{F}, \mathbb{P})$ is called a complete probability space if $\mathcal{F}$ contains all subsets of the null sets.

Definition 1.3.2 (Filtration) *A filtration $\mathbb{F}$ is a non-decreasing family of sub-sigma-algebras $(\mathcal{F}_t)_{t \geq 0}$ with $\mathcal{F}_t \subset \mathcal{F}$ and $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s < t < \infty$. The quadruple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a filtered probability space if*

- *$\mathcal{F}_0$ contains all subsets of the null sets of $\mathcal{F}$,*

- *$\mathbb{F}$ is right-continuous, i.e $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s$.*

Note that in the rest of the thesis, we may also use $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ to denote a filtered probability space.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a function $Y : \Omega \to \mathbb{R}^n$ is called $\mathcal{F} - measurable$ if

$$Y^{-1}(U) := \{\omega \in \Omega; Y(\omega) \in U\} \in \mathcal{F}$$

for all open sets $U \in \mathbb{R}^n$ (or, equivalently, for all Borel sets $U \subset \mathbb{R}^n$) [114].

A random variable is an $\mathcal{F} - measurable$ function $X : \Omega \to \mathbb{R}^n$. Every random variable induces a probability measure P_X on $\mathbb{R}^n$, defined by

$$P_X(B) = \mathbb{P}(X^{-1}(B)).$$

P_X is called the distribution of X [114].

Definition 1.3.3 (Random Vector and Distribution Function) *A random vector X is a real function $X : \Omega \to \mathbb{R}^d, d \in \mathbb{N}$, which is measurable with respect to its underlying σ-algebra $\mathcal{F}$. For $d = 1$, X is a random variable. The function F defined by $F(x) = \mathbb{P}(X \leq x)$ is called the distribution function of X.*

The definitions covered so far are basic definitions in probability theory, and foundational for stochastic processes, which is defined next.

Definition 1.3.4 (Stochastic Process) *A stochastic process is a family $\mathbf{X}(\mathbf{t}) = (X(t))_{t \geq 0}$ of random vectors $X(t)$ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The stochastic process $\mathbf{X}(\mathbf{t})$ is called*

- *adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if $X(t)$ is $\mathcal{F}_t$ measurable for all $t \geq 0$,*

- *measurable if the mapping $\mathbf{X}(\mathbf{t}) : [0, \infty) \times \Omega \to \mathbb{R}^d, d \in \mathbb{N}$ is $(\mathcal{B}([0, \infty)) \otimes \mathcal{F})$ measurable with $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$ denoting the product sigma-algebra of $\mathcal{B}([0, \infty))$ and $\mathcal{F}$, where $\mathcal{B}(\mathbf{A})$ denotes the Borel σ-algebra of $\mathbf{A}$,*

- *progressively measurable if the mapping $\mathbf{X}(\mathbf{t}) : [0, t] \times \Omega \to \mathbb{R}^d$ is $(\mathcal{B}([0, \infty)) \otimes \mathcal{F}_t)$ measurable for each $t \geq 0$.*

Note that for each t fixed, we have a random variable

$$\omega \to X(t, \omega), \omega \in \Omega.$$

When fixing ω, we have a function in t

$$t \to X(t, \omega),$$

which is called a path of X(t). If the paths are continuous, i.e t $\to$ X(t, ω) is a continuous function for all ω, X(t) is a continuous process.

Definition 1.3.5 (Quadratic Variation) *For real-valued stochastic process X_t defined on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with $t \geq 0$, the quadratic variation over the time interval $[0, t]$ is defined as:*

$$\langle X_t \rangle = \lim_{\Delta \to 0} \sum_{i=0}^{n-1} |X(t_{i+1}) - X(t_i)|^2$$

for $t_i \in [0, t], i = 0, 2, \ldots, n - 1$ and $\Delta = max(t_{i+1} - t_i)$

More generally, the variation of two different processes (let's call this covariation), say $X(t)$ and $Y(t)$, is defined as:

$$\langle X(t), Y(t) \rangle = \lim_{\Delta \to 0} \sum_{i=0}^{n-1} (X(t_{i+1}) - X(t_i))(Y(t_{i+1}) - Y(t_i))$$

Definition 1.3.6 (Brownian motion) *An adapted process W(t) for $t \geq 0$ is called a Brownian motion (or a Wiener process), if W(t) satisfies the following properties:*

- $W(0) = 0$,

- *W(t) has independent increments, i.e. W(t) − W(s) is independent of W(t') − W(s') where $0 \leq s' \leq t' \leq s \leq t < \infty$,*

- *W(t) has Gaussian increments, W(t + s) − W(t) $\sim$ N(0, s), for all s > 0.*

Definition 1.3.7 (n-dimensional Brownian Motion) $\mathbf{W_t}$ *given by* $\mathbf{W_t} = (W_1(t), \ldots, W_n(t))'$ *for $t \geq 0$ is called a d-dimensional Brownian motion (or n-dimensional Wiener process) if its components $W_j(t)$, for $j = 1, \ldots, n$ are independent Brownian motions.*

Definition 1.3.8 (Martingale) *Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ be a filtered probability space. A stochastic process $\mathbf{X(t)} = (X(t))_{t \geq 0}$ is called a martingale relative to $(\mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ if $\mathbf{X(t)}$ is adapted, $\mathbb{E}(\|\mathbf{X(t)}\|) < \infty$ for all $t \geq 0$, and*

$$\mathbb{E}(\mathbf{X(t)} \,|\mathcal{F}_s) = \mathbf{X(s)}, \forall \, 0 \leq s \leq t < \infty.$$

1.3.2   Distribution Functions and Characteristic Functions

In this section, we give an overview of distribution functions and characteristic function. In the following, we start with introduction of distribution functions, moment generating functions and characteristic functions in one dimension followed by an introduction to characteristic functions in higher dimensions. Then we introduce the class of analytic characteristic functions.

Definition 1.3.9 (Distribution Functions) *A point function F on a line is a distribution function if*

- *F is non-decreasing, that is $a < b \implies F(a) \leq F(b)$,*

- *F is right-continuous, that is $F(a) = \lim_{x \to a+} F(x)$,*

- *$F(-\infty) = 0$ and $F(\infty) < \infty$.*

F is a probability distribution function is it is a distribution function and $F(\infty) = 1$.

Definition 1.3.10 (Algebraic and Absolute Moments) *Let X be a random variable with probability distribution function F. The algebraic moment of oder k of $F(x)$, $x \in \mathbb{R}$, is then given by*

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k dF(x).$$

Similarly, the absolute moment of order k of $F(x)$ is defined by

$$\mathbb{E}(|X|^k) = \int_{-\infty}^{\infty} |x|^k dF(x).$$

Next theorem provides necessary and sufficient conditions for algebraic moment to exist.

Theorem 1.3.1 *The algebraic moment of order k of a distribution function $F(x)$ exists if and only if its absolute moment of order k exists. Suppose that the algebraic moment of order k of $F(x)$ exists then the moments $\mathbb{E}(X^n)$ and $\mathbb{E}(|X|^n)$ exist for all orders $n \leq k$.*

See [100] for proof. Now we give the definition for moment generating function and characteristic function.

Definition 1.3.11 (Moment Generating Function and Characteristic Function) *Let X be a random variable with probability distribution function F. The moment generating function of $F(x)$, $x \in \mathbb{R}$ (or of X) is defined for real u by*

$$\phi_X(u) = \int_{-\infty}^{\infty} e^{ux} dF(x).$$

The characteristic function of X is defined by

$$\psi_X(u) = \int_{-\infty}^{\infty} e^{iux} dF(x) = \int_{-\infty}^{\infty} \cos(ux) dF(x) + i \int_{-\infty}^{\infty} \sin(ux) dF(x).$$

$\psi_X(u)$ is also called the Fourier transform of $F(X)$.

Note that $\phi_X(u)$ may not be defined for every $u \in \mathbb{R}$, but $\psi_X(u)$ does not suffer from this problem as $|\psi_X(u)| \leq 1$. So $\phi_X(iu) = \psi_X(u)$ only if $\phi_X(u)$ is defined for every $u \in \mathbb{R}$.

Characteristic function is essential to option pricing because Carr and Madan's Fast Fourier Transform (FFT) approach [29] offers a fast yet robust way for option pricing. The following theorem is also an important theoretical result that supports the FFT option pricing approach.

Theorem 1.3.2 (Inversion Theorem) *Let X be a random variable with probability distribution function F and $\psi_X(u)$ be the characteristic function of the distribution of X. Then X has a bounded continuous density $f(x)$, $x \in \mathbb{R}$ given by*

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \psi_X(u) du.$$

See [65] for proof.

The characteristic function in higher dimensions is closely related to the characteristic function in one dimension.

Definition 1.3.12 (Characteristic Function in Higher Dimensions) *Let $\mathbf{X}$ be a vector of random variables $X_1, \ldots, X_n$ with probability distribution $F(\mathbf{X})$. The characteristic function of $\mathbf{X}$ is the function $\psi_{\mathbf{X}}(\mathbf{u})$ defined for a vector of real numbers $\mathbf{u}$:*

$$\psi_{\mathbf{X}}(\mathbf{u}) = \int_{-\infty}^{\infty} e^{i\mathbf{u}'\mathbf{x}} dF(\mathbf{x})$$

The multidimensional version of Theorem 1.3.2 is still valid.

Theorem 1.3.3 (Inversion Theorem in Higher Dimensions) *Let $\psi_{\mathbf{X}}(\mathbf{u})$ be the characteristic function of of $\mathbf{X}$ and suppose $\psi_{\mathbf{X}}(\mathbf{u}) \in \mathbf{L}^1$. Then $F(\mathbf{X})$ has a bounded continuous density $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ given by*

$$f(\mathbf{x}) = F'(\mathbf{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\mathbf{u}'\mathbf{x}} \psi_{\mathbf{X}}(\mathbf{u}) d\mathbf{u}.$$

Next, we introduce the class of analytic characteristic functions. Suppose now u is a complex number defined by $u = a + ib$ where $a, b \in \mathbb{R}$.

Definition 1.3.13 (Analytic Characteristic Function) *A characteristic function $\psi(u)$ is said to be an analytic characteristic function is there exists a function $e(u)$ of the complex variable u which is regular in a circle $|u| < c$ ($c > 0$) and a constant $\varepsilon > 0$ such that $e(a) = \psi(a)$ for $|a| < \varepsilon$.*

An informal alternative to express Definition 1.3.13 is by saying that an analytic characteristic function is a characteristic function that is equivalent to a holomorphic function in some neighborhood of the origin in the complex u-plane. A holomorphic function, according to [93], is a complex function defined on a region $D \subset \mathbb{C}$ and is complex differentiable at every point in D.

Theorem 1.3.4 *If a characteristic function $\psi(u)$ is regular in a neighbourhood of the origin, then it is also regular in a horizontal strip and can be represented by a Fourier integral. This strip is either the whole plane, or it has one or two horizontal lines, The purely imaginary points on the boundary of the strip of regularity (if this strip is not the whole plane) are singular points of $\psi(u)$.*

See [100] for a proof.

An example considered in [101] is the characteristic function given by:

$$\psi(u) = \left(1 - \frac{iu}{\lambda}\right)^{-1}.$$

We recognize that this is the characteristic function for an exponential random variable with parameter λ if $u \in \mathbb{R}$. As we defined $u = a + bi$, the characteristic function has a singularity when $u = -i\lambda$ and is regular near the origin in the strip $-\lambda < Im(u) < \infty$, where $Im(u)$ denotes the imaginary part of u.

Given a complex number $u = a + ib$, according to [125], the existence of the Fourier transform $Ft(u) = \int_{-\infty}^{\infty} e^{iux} f(x) dx$ implies certain restrictions on $f(x)$ at infinity. If $Ft(u)$ does not exist, partial integrals:

$$Ft^+(u) = \int_0^{\infty} e^{iux} f(x) dx$$

$$Ft^-(u) = \int_{-\infty}^0 e^{iux} f(x) dx$$

may exist for sufficiently large positive b and large negative b respectively. For the inversion we may have:

$$f(x) = \frac{1}{2\pi} \left(\int_{ib_1 - \infty}^{ib_1 + \infty} Ft^+(u) e^{-iux} du + \int_{ib_2 - \infty}^{ib_2 + \infty} Ft^-(u) e^{-iux} du \right)$$

where b_1 is a sufficiently large positive number and b_2 is a sufficiently large negative number.

1.3.3   Ito Process

In this section, we briefly introduce core components in Ito calculus, which are also essential to stochastic processes. Topics we cover in this section are Ito process, Ito's lemma and Ito isometry. Next we give the definition of Ito process.

Definition 1.3.14 (Ito Process) *Let $\mathbf{X_t} = (X_1(t), \dots, X_n(t))'$ be an n-dimensional stochastic process defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}), \mathbb{P})$. It is called an Itô process if it satisfies*

the following equation:

$$\mathbf{X_t} = \mathbf{X_0} + \int_0^t \boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{\Sigma}_s d\mathbf{W}_s \tag{1.3}$$

$$= \begin{pmatrix} X_1(0) + \int_0^t \mu_1(s)ds + \sum\limits_{j=1}^m \int_0^t \Sigma_{1,j}(s)dW_j(s) \\ \vdots \\ X_n(0) + \int_0^t \mu_n(s)ds + \sum\limits_{j=1}^m \int_0^t \Sigma_{m,j}(s)dW_j(s) \end{pmatrix}, \tag{1.4}$$

where:

- $\mathbf{X_0} = (X_1(0), \ldots, X_n(0))'$ *is an n-dimensional $\mathcal{F}_0$-measurable random variable,*

- $\boldsymbol{\mu_t} = (\mu_1(t), \ldots, \mu_n(t))'$ *is an n-dimensional stochastic process, where μ_i is adapted and $\int_0^t |\mu_i(s)|ds < \infty$, for all $i = 1, \ldots, n$,*

- $\boldsymbol{\Sigma_t} = \begin{pmatrix} \sigma_{11}(t) & \cdots & \sigma_{1m}(t) \\ \vdots & \cdots & \vdots \\ \sigma_{n1}(t) & \cdots & \sigma_{nm}(t) \end{pmatrix}$ *is an $n \times m$-dimensional stochastic process, where $\sigma_{ij}(t)$ is*

  adapted, $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$-measurable and $\int_0^t \mathbb{E}[(\sigma_{ij}(t))^2]ds < \infty$, for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$,

- $\mathbf{W_t} = (W_1(t), \ldots, W_m(t))'$ *is an m-dimensional Brownian motion.*

For convenience, we can rewrite Equation (1.4) in the following form:

$$d\mathbf{X_t} = \boldsymbol{\mu_t}dt + \boldsymbol{\Sigma_t}d\mathbf{W_t}.$$

We call this expression a stochastic differential equation (SDE) with drift $\boldsymbol{\mu_t}$ and diffusion term $\boldsymbol{\Sigma_t}$.

Generally speaking, an Ito process is the integral representation of a stochastic differential equation. Studying an Ito process leads to many statistical properties, for example moment generating function and characteristic function of $X(T)$ conditioning on $\mathcal{F}_t$. For Ito processes $X(t)$ and $Y(t)$ with SDE's in general forms:

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t)$$
$$dY(t) = \mu(Y(t), t)dt + \sigma(Y(t), s)dB(t)$$

$W(t)$ and $B(t)$ are two standard Wiener processes with correlation $-1 \leq \rho \leq 1$. The quadratic variation of $dX(t)$ is given by:

$$\langle dX_t \rangle = \sigma^2(X(t), t)dt.$$

The covariation of $dW(t)$ and $dB(t)$ is $< dW(t), dB(t) = \rho dt$, hence the covariation of $dX(t)$ and $dY(t)$ is

$$\langle dX(t), dY(t) \rangle = \rho\sigma(X(t), t)\sigma(Y(t), t)dt$$

The following theorem is the famous Ito's lemma, which produces the SDE of a transformed process.

Theorem 1.3.5 (Ito's Lemma) *Let $W(t)$ be Brownian motion and $X(t)$ be an Ito process satisfying following stochastic differential equation:*

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t).$$

If $f(X(t), t) \in C^2(\mathbb{R}, \mathbb{R})$, then $f(X(t), t)$ is also an Ito process with the stochastic differential equation:

$$
\begin{aligned}
d(f(X(t), t)) &= \frac{\partial f(X(t), t)}{\partial t}dt + \frac{\partial f(X(t), t)}{\partial X(t)}dX(t) + \frac{1}{2}\frac{\partial^2 f(X(t), t)}{\partial X^2(t)}\langle dX_t\rangle \\
&= \Big(\frac{\partial f(X(t), t)}{\partial t} + \frac{\partial f(X(t), t)}{\partial X(t)}\mu(X(t), t) + \frac{1}{2}\frac{\partial^2 f(X(t), t)}{\partial X^2(t)}\sigma^2(X(t), t)\Big)dt \\
&\quad + \frac{\partial f(X(t), t)}{\partial X(t)}\sigma(X(t), t)dW(t).
\end{aligned}
$$

Before we get into Ito isometry, we first briefly introduce another important concept called Ito integral. If $Y(t)$ is some bounded continuous stochastic process which is adapted to the same filtration $\mathcal{F}_t$ generated by $W(t)$, then $\int_t^T Y(s)dW(s)$ is an Ito integral. An Ito integral is a martingale. Ito isometry is a key fact about Ito integrals, which is given by:

$$\mathbb{E}\Big[\Big(\int_t^T Y(s)dW(s)\Big)^2\Big|\mathcal{F}_t\Big] = \mathbb{E}\Big[\int_t^T Y^2(s)ds\Big|\mathcal{F}_t\Big].$$

Next lemma and theorem lead to the theoretical foundation of change of measure for Itô processes.

Lemma 1.3.1 (Novikov's condition) *Let $\mathbf{W_t} = (W_1(t), \ldots, W_n(t))'$ be a n-dimensional Brownian motion and $\boldsymbol{\gamma_t} = (\gamma_1(t), \ldots, \gamma_n(t))'$ for $0 \le t \le T$ is a measurable, adapted, n-dimensional process with $\int_0^T \gamma_i(t)dt < \infty$ a.s., for all $i = 1, \ldots, n$. Assume that*

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \|\boldsymbol{\gamma_s}\|^2 ds\Big)\Big] < \infty.$$

Then process $L(t)$ for $0 \le t \le T$ is defined as follows:

$$L(t) := \exp\Big(-\int_0^t \boldsymbol{\gamma_s}d\mathbf{W_s} - \frac{1}{2}\int_0^t \|\boldsymbol{\gamma_s}\|^2 ds\Big)$$

is a continuous martingale.

Under Novikov's condition: $\int_0^t \|\boldsymbol{\gamma_s}\|^2 ds < \infty$ a.s., for all $t \in [0, T]$ so that the stochastic integral in $L(t)$ is well-defined. For a detailed explanation see [129].

Novikov's condition is a critical for change of measure practice in financial mathematics, which is given in following theorem.

Theorem 1.3.6 (Girsanov theorem) *Let processes* $\mathbf{W_t}$, $\boldsymbol{\gamma}_t$ *and* $L(t)$ *be as in Lemma 1.3.1. Define the equivalent probability measure* $\mathbb{Q}$ *on* $(\Omega, \mathcal{F}_T)$ *by:*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = L(T) \text{ i.e. } \mathbb{Q}(A) = \mathbb{E}[1_A \, L(T)] = \int_A L(T)d\mathbb{P}, \forall A \in \mathcal{F}_T.$$

Then, the process $\tilde{\mathbf{W}}_\mathbf{t} = (\tilde{W}_1(t), \dots, \tilde{W}_n(t))'$ *given by:*

$$\tilde{W}_i(t) = W_i(t) + \int_0^t \gamma_i(s)ds,$$

for $t \in [0, T]$ *and* $i = 1, \dots, n$ *is a* $\mathbb{Q}$-*Brownian motion. Process* $L(t)$ *is called the density or the change of measure from* $\mathbb{P}$ *to* $\mathbb{Q}$.

A typical application of Girsanov theorem is to change the probability measure (eg. historical measure) where the model is defined to another equivalent probability measure (eg. risk neutral measure). In chapter 2 we will see models defined under historical measure, denoted $\mathbb{P}$ as well as models defined under risk neutral measure, denoted $\mathbb{Q}$. We often interpret historical measure as real world. Models defined under this measure reflects risks undertaken by investors i.e investors bearing high risk expect high returns. Risk neutral measure, on the other hand, is a world where there is no risk or arbitrage. Everyone is in a fair game with one price. Next definition gives the mathematical aspect of risk neutral measure.

Definition 1.3.15 (Risk Neutral Measure) *A measure* $\mathbb{Q} \sim \mathbb{P}$ *such that the discounted price process at risk-free interest rate is a (local) martingale with respect to* $\mathbb{Q}$ *is called an equivalent (local) martingale measure.*

The martingale property guarantees that the price under risk neutral measure is unique, thus arbitrage is excluded.

1.3.4   Feynman-Kac Representation in One Dimension

Feynman-Kac formula provides stochastic representations for solutions of partial differential equations. There are formulas for one-simensional and multidimensional cases, we first state Feynman-Kac's representation in one dimension.

Theorem 1.3.7 (Feynman-Kac) *Define* $X(t)$ *on the probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ *with following SDE:*

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t).$$

with initial condition $X(t) = x$ *and* $u(X(t), t) \in C^2(\mathbb{R}, \mathbb{R}^+)$ *be a solution for the PDE:*

$$u_t + \mu(x, t)u_x + \frac{1}{2}\sigma^2(x, t)u_{xx} - V(x, t)u + f(x, t) = 0$$

where $u_t = \frac{\partial u(x,t)}{\partial t}, u_x = \frac{\partial u(x,t)}{\partial x}$ *and* $u_{xx} = \frac{\partial^2 u(x,t)}{\partial x^2}$ *for all* $t \in [0, T]$ *and* $x \in \mathbb{R}$ *with terminal condition*

$$u(X(T), T) = h(X(T))$$

then

$$u(x, t) = \mathbb{E}^{\mathbb{P}}\left[\int_t^T e^{-\int_t^s V(X(\tau), \tau)d\tau} f(X(s), s)ds + e^{-\int_t^T V(X(\tau), \tau)d\tau} h(X(T)) \Big| X(t) = x \right]$$

Next we give Feynman-Kac's representation in higher dimensions.

Theorem 1.3.8 (Feynman-Kac in Higher Dimensions) *Define a vector of n stochastic processes $\mathbf{X_t} = (X_1(t), \ldots, X_n(t))'$ and a vector of independent Brownian motions $\mathbf{W_t} = (W_1(t), \ldots, W_n(t))'$ on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}), \mathbb{P})$. The SDE for individual process $X_i(t)$ is given by*

$$dX_i(t) = \mu_i(\mathbf{X_t}, t)dt + \sum_{j=1}^{n} \sigma_{ij}(\mathbf{X_t}, t)dW_j(t).$$

with initial condition $\mathbf{X_t} = \mathbf{x}$ and $u(\mathbf{X_t}, t) \in C^2(\mathbb{R}^n, \mathbb{R}^+)$ be a solution for the PDE:

$$u_t + \sum_{i=1}^{n} \mu_i(\mathbf{x}, t)u_{x_i} + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{2}\sigma_{ij}^2(\mathbf{x}, t)u_{x_i x_i} + \sum_{i=1}^{n}\sum_{j \neq i}^{n} \sigma_{ii}(\mathbf{x}, t)\sigma_{ji}(\mathbf{x}, t)u_{x_i x_j} - V(\mathbf{x}, t)u + f(\mathbf{x}, t) = 0$$

where $u_t = \frac{\partial u(\mathbf{x},t)}{\partial t}, u_{x_i} = \frac{\partial u(\mathbf{x},t)}{\partial x_i}$, $u_{x_i x_i} = \frac{\partial^2 u(\mathbf{x},t)}{\partial x_i^2}$ and $u_{x_i x_j} = \frac{\partial^2 u(\mathbf{x},t)}{\partial x_i \partial x_j}$ for all $t \in [0, T]$ and $\mathbf{x} \in \mathbb{R}^n$ with terminal condition

$$u(\mathbf{X_T}, T) = h(\mathbf{X_T})$$

then

$$u(\mathbf{x}, t) = \mathbb{E}^{\mathbb{P}}\left[\int_t^T e^{-\int_t^s V(\mathbf{X_\tau}, \tau)d\tau} f(\mathbf{X_s}, s)ds + e^{-\int_t^T V(\mathbf{X_\tau}, \tau)d\tau} h(\mathbf{X_T}) \Big| \mathbf{X_t} = \mathbf{x} \right]$$

1.3.5 Risk Measures

In this section, we discuss risk measures that are used in the area of financial risk management. Risk management in general is "a discipline for living with the possibility that future events may cause adverse effects" [90]. Financial institutions like banks and insurance companies make business by taking on risks in the form of loans and insurance policies. Hence, risk management is the core of banks and insurance companies to keep them away from crisis or catastrophes. Financial institutions use various approaches to manage risk, part of this process involves measuring the risk.

A risk measure is a statistic used to determine the amount of assets to be kept in reserve. This amount is set by regulators for the risk taken by financial institutions so that there won't be dramatically change in the operations in the worst case scenario. In this section, we introduce properties that a good measure of risk should have by listing a set of properties (axioms) that a good risk measure should have. The risk measures that satisfy these axioms are called coherent risk measures. The characteristics of these risk measures are studied by Artzner et.al [7] in great detail. Here, we briefly introduce some widely used risk measures in practice.

1.3.5.1 The Axioms of Coherence

We first define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a time horizon Δ. Define $\mathbf{M}$ be the set of all finite random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, which represents losses over the time horizon Δ. Then financial risks are represented by a subset of $\mathbf{M}$ i.e $\mathbf{R} \subset \mathbf{M}$. It is often assumed that $\mathbf{R}$ is a convex cone i.e $R_1 \in \mathbf{R}, R_2 \in \mathbf{R} \implies R_1 + R_2 \in \mathbf{R}$ and $aR_1 \in \mathbf{R}$ for every $a > 0$. Risk measures

are real-valued functions $\rho : \mathbf{R} \to \mathbb{R}$. We follow the interpretation of $R \in \mathbf{R}$ as in [104] that R represents a loss rather than future value as in [7]. We also assume zero interest rate so that there is no discounting. The axioms are defined in the same way as in [104]. Let $\rho(R)$ denote the risk measure for random variable R (the amount of capital that should be added to a position given a loss R), a coherent risk measure should satisfy:

- **Axiom 1 (Translation Invariance):** For all $R \in \mathbf{R}$ and every $a \in \mathbb{R}$, we have $\rho(R + a) = \rho(R) + a$;

- **Axiom 2 (Subadditivity):** For all $R_1, R_2 \in \mathbf{R}$, we have $\rho(R_1 + R_2) \leq \rho(R_1) + \rho(R_2)$;

- **Axiom 3 (Positive Homogeneity):** For all $R \in \mathbf{R}$ and every $a \in \mathbb{R}$, we have $\rho(aR) = a\rho(R)$;

- **Axiom 4 (Monotonicity):** For all $R_1, R_2 \in \mathbf{R}$ such that $R_1 \leq R_2$, then $\rho(R_1) \leq \rho(R_2)$.

The definition of coherent risk measure is given by:

Definition 1.3.16 (Coherent Risk Measure) *A risk measure ρ whose domain includes the convex cone $\mathbf{M}$ is called coherent (on $\mathbf{M}$) if it satisfies Axioms 1-4.*

It is worth noting that a risk measure is not equivalent to a coherent risk measure. As Definition 1.3.16 suggests, a coherent risk measure is more formally and strictly defined whereas a risk measure is a general term. In next section, we discuss two popular risk measures used in practice.

1.3.5.2   VaR and Expected Shortfall

We start with the definition of VaR. Consider a portfolio of some risky assets over the time horizon Δ. Let R denote the size of the loss experienced by the portfolio. VaR is defined by

Definition 1.3.17 (Value-at-Risk (VaR)) *Given a confidence level $\alpha \in (0, 1)$, the VaR of our portfolio at the confidence level α is given by the smallest number r^* such that the probability that the loss R exceeds r^* is no larger than $(1 - \alpha)$:*

$$VaR_\alpha(R) = inf\{r^* \in \mathbb{R} : P(R > r^*) \leq 1 - \alpha\} = inf\{r^* \in \mathbb{R} : F_R(r^*) \geq \alpha\}$$

Another commonly used interpretation of α is that it is the probability that R does exceeds a certain amount r^*. No matter which version of the interpretations of α is used, it does not affect the fact that VaR is simply a quantile of the loss distribution in a probabilistic context. A formal definition of quantiles is given by the next definition.

Definition 1.3.18 (General Inverse and Quantile Function)

- *Given an increasing function $T : \mathbb{R} \to \mathbb{R}$, the generalized inverse of T is defined by $T^{-1}(t^*) = inf\{t \in \mathbb{R} : T(t) \geq t^*\}$, where we use the convention that the infimum of an empty set is ∞.*

- *Given a distribution function F, the generalized inverse F^{-1} is called the quantile function of F. For $\alpha \in (0, 1)$, the $\alpha-$th quantile of F is given by*

$$q_\alpha(F) = F^{-1}(\alpha) = inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$$

Next lemma is a direct result of Definition 1.3.18 and the right-continuity property of distribution functions.

Lemma 1.3.2 *A point $x_\alpha \in \mathbb{R}$ is the $\alpha-$th quantile of a distribution function F iff. the following two conditions are satisfied:*

- *$F(x_\alpha) \geq \alpha$;*

- *$F(x) < \alpha$ for all $x < x_\alpha$.*

In practice, normal distribution is usually assumed for the loss distribution and the computation of VaR is straightforward: if the loss $R \sim N(\mu, \sigma^2)$, then the VaR at level α is then:

$$VaR_\alpha(R) = \mu + \sigma\Phi^{-1}(\alpha),$$

where $\Phi^{-1}(\alpha)$ is the $\alpha-$th quantile of standard normal distribution. Similarly, if t distribution is assumed with the same mean and variance $\frac{v\sigma^2}{v-2}$ when degrees of freedom $v > 2$, we have

$$VaR_\alpha(R) = \mu + \sigma t_v^{-1}(\alpha),$$

where $t_v(x)$ is the standard t distribution function with degrees of freedom v. It can be seen that VaR is easy to calculate under normal or t distribution assumptions. VaR is also recognized as a key statistic in Basel Accords (Basel II and Basel III) for risk measurement and capital requirement calculation. However, VaR has been criticized for its weakness in loss aggregations. The critique originally outlined in [7] is that VaR violates subadditivity axiom of coherent risk measures. Thus, VaR is not a coherent risk measure.

As a illustration, consider a portfolio with individual loss distribution F_{R_1} and F_{R_2}, and the merged loss distribution $S = R_1 + R_2$ for the portfolio is F_S. It is not necessarily true that $q_\alpha(F_S) \leq q_\alpha(F_{R_1}) + q_\alpha(F_{R_2})$. This contradicts our intuition that diversification reduces risk; we can not use aggregated VaR's to set a bound for the overall risk either.

Another drawback of VaR comes from model risk. That is the models for modeling the losses are misspecified or the model assumptions are violated in practice. A typical example is that the symmetric distributions (normal or t distribution) are adopted for modeling losses, but in fact the distribution of losses is heavy-tailed, in which case symmetric distribution assumptions are not appropriate, see Figure 12.4 of [83].

To overcome the lack of subadditivity issue of VaR, another loss distribution-based coherent risk measure is introduced, the expected shortfall (ES). ES is closely related to VaR and preferred by risk managers in practice. The formal definition of ES is given as follows.

Definition 1.3.19 (Expected Shortfall) *For a loss R with $\mathbb{E}(|R|) < \infty$ and distribution function F_R, the expected shortfall (ES) at confidence level $\alpha \in (0, 1)$ is defined as*

$$ES_\alpha(R) = \frac{1}{1 - \alpha} \int_\alpha^1 q_u(F_R)du,$$

where $q_u(F_R)$ is the u−th quantile of F_R.

By Definition 1.3.17, therefore the relationship between ES and VaR is established by

$$ES_\alpha(R) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_u(R)du.$$

Again, depending on the interpretation of α, we also have following relationship:

$$ES_\alpha(R) = -\frac{1}{\alpha} \int_0^\alpha VaR_u(R)du,$$

which is also true. See [104] for proof of ES to be coherent risk measure. If the loss distributions are continuous, then a more intuitive expression of ES is given by the next lemma,

Lemma 1.3.3 *For an integrable loss R with continuous distribution function F_R and for $\alpha \in (0, 1)$, we have*

$$ES_\alpha(R) = \frac{\mathbb{E}(R; R \geq q_\alpha(R))}{1 - \alpha} = \mathbb{E}(R \,|R \geq VaR_\alpha)$$

where $\mathbb{E}(R; A)$ denotes $\mathbb{E}(RI_A)$ for a generic integrable random variable R and a generic set $A \in \mathcal{F}$.

See [104] for proof. So for losses with a continuous distribution, the ES_α is simply the conditional expectation of loss R given the loss already exceeds VaR_α or the expected tail loss. Sometimes, this is also called "Tail-Value-at-Risk" (TVaR). For completeness, a formula similar to Lemma 1.3.3 is also available for a discontinuous loss distribution function F_R, but is more complicated:

$$ES_\alpha(R) = \frac{1}{1 - \alpha}(\mathbb{E}(R; R \geq q_\alpha(R)) + q_\alpha(R)(1 - \alpha - P(R \geq q_\alpha(R)))).$$

See [1] for proof.

Next we give the ES for normal distribution with mean μ variance σ^2 and t distribution as examples. For normal distribution:

$$ES_\alpha(R) = \mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha}.$$

where $\phi(x)$ is the density function for standard normal distribution. For t distribution with the same mean and variance $\frac{\nu\sigma^2}{\nu-2}$ when degrees of freedom $\nu > 2$, we have

$$ES_\alpha(R) = \mu + \sigma\frac{g_\nu(t_\nu^{-1}(\alpha))}{1 - \alpha}\left(\frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1}\right),$$

where $g_\nu(x)$ is the density function of standard t distribution.

Chapter 2

Overview of Relevant Models

We dedicate this chapter to provide an overview of relevant models in the literature and interesting extensions which has not been studied in literature. Since mean-reverting models and Grasselli's 4/2 stochastic volatility model [76] are central to this thesis, we will spend most of the chapter covering mean-reverting models and the 4/2 stochastic volatility model as well as a few potential innovations based on the 4/2 type of stochastic volatility model. In section 1.2.2, we highlighted the most important models in the context of commodities, which is one of the oldest asset classes with the mean-reverting property. In finance, mean reversion property indicates that the asset prices will eventually revert to their long-run means. Graphically, the trend of the asset prices' series moves up and down with respect to their long-run means. As typical representatives from the mean-reverting family, commodities can be used as a guide to select models or to develop interesting models. Exchange rates and interest rates are yet other examples of well known financial asset classes that behave in a mean-reverting way. Most recently, volatility indexes (e.g. VIX) are yet another financial asset class that has been shown to be mean-reverting in the long run. In this perspective, popular models in finance like GBM or Heston model, that does not have mean-reverting property, are not suitable for modeling commodities, or volatility indexes.

The applications of mean-reverting models are quite versatile in finance. As mentioned above, other than commodities and volatility indexes, mean-reverting models are also used to model foreign exchange rates and interest rates therefore the whole world of fixed income products and currency modeling are influenced, but these last two asset classes won't be the focus of the thesis. The most basic yet famous mean-reverting model is the Vasicek model [126], which is better known as an application of Ornstein-Uhlenbeck (O-U) processes to interest rates. A problem or limitation of the Vasicek modelling is that the process can generate negative values, which is not realistic for positive asset classes. Later, Cox et.al [38] proposed the Cox-Ingersoll-Ross (CIR) model that guarantees positive values given Feller's condition on the parameters is satisfied, this way CIR model solves the negative values issue faced by Vasicek model [126]. Since then, the mean-revering model family has continued evolving and growing with more complex models, developed to better understand the market and to capture new stylized facts, e.g. stochastic volatility, spillover effects, leverage effect, and multivariate models.

In this chapter, we denote $\mathbb{P}$ as historical probability measure and $\mathbb{Q}$ as risk-neutral probability measure. In general, $\mathbb{P}$ measure is convenient for parameter estimation while $\mathbb{Q}$ measure is crucial for pricing practice where arbitrage should be excluded. Change of measure is required to perform certain task that is not suitable under the probability measure where the model is currently defined. For example, if the model is defined under $\mathbb{P}$ measure, and we would like to assess the model performance in option pricing, then we need to change the measure from $\mathbb{P}$ to $\mathbb{Q}$ for the model. We will see models defined under $\mathbb{P}$ measure as well as $\mathbb{Q}$. The models we will be reviewing cover continuous time mean-reverting models and stochastic volatility models that inspire our research. In particular, in section 2.1, we focus on one-factor models, i.e. with only the asset price as the stochastic factor. In section 2.2, we review models that has a stochastic volatility component, hence two factors. In section 2.3, we provide an overview of models in a multidimensional setting.

Discrete time stochastic volatility models such as ARCH and GARCH models are also essential in modelling volatility in econometrics. A continuous limit of GARCH models are also studied in e.g [111] [25] [50]. The continuous limit of GARCH(1,1) model found in [111] has a mean-reverting stochastic variance process that drives by an independent Brownian motion, which is different from the original GARCH(1,1) process in which case both the asset process and the stochastic variance process are driven by the same white noise. Even though the stochastic variance process has mean-reverting property, unlike the CIR process considered in Heston model [80], it does not preserve positivity. For higher order of GARCH(p, q) when $p > 1, q > 1$, it adds more complexity with more parameters to estimate. For these reasons, we do not consider GARCH family and its continuous time limit in this thesis.

2.1   One-Factor Mean-Reverting Models

In this section, we review five fundamental continuous-time mean-reverting models. These models have either constant volatility or constant elasticity of variance (CEV, local volatility), e.g. they do not use a separated correlated process as a driver for the instantaneous volatility. What all these models have in common is that the asset price is driven by a single stochastic process. We will review models with one more stochastic process that drives the stochastic volatility in section 2.2.

2.1.1   One-Factor Schwartz Model

Schwartz [122] proposed three models to study commodity prices, one of which (the one-factor) assumes the logarithm of the spot price of the commodity follows a mean-reverting process of O-U type. The first model is a simple one-factor O-U process driving the logarithm of commodity prices as the factor; the second model adds convenience yield of the commodity as the second factor besides the log of the commodity price; the third model includes stochastic interest rates as an additional factor. The one-factor Schwartz model can be also considered to model volatility indexes in that volatility indexes have the same mean-reverting property and positivity as commodity prices.

Assume the asset price follows the stochastic process on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ under historical measure $\mathbb{P}$:

$$\frac{dX(t)}{X(t)} = (L - \beta ln(X(t)))dt + \sigma dW^{\mathbb{P}}(t).$$

β, L and σ are positive constants. Let $Y(t) = ln(X(t))$ then we obtain the process for $Y(t)$ by Ito's lemma:

$$dY(t) = (L^* - \beta Y(t))dt + \sigma dW^{\mathbb{P}}(t),$$

$$L^* = L - \frac{\sigma^2}{2}.$$

We can see that $Y(t)$ is a standard O-U process with β measuring the mean-reverting speed and $\frac{L^*}{\beta}$ indicating the long-run average of the asset price. σ is the volatility parameter. Hence, $Y(t)$ is a normally distributed random variable with mean and variances [122]:

$$\mathbb{E}(Y(t)|Y(0)) = e^{-\beta t}Y(0) + (1 - e^{-\beta t})\frac{L^*}{\beta},$$

$$Var(Y(t)|Y(0)) = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}),$$

while the covariance between two times (autocovariance) is,

$$Cov(Y(t), Y(s)) = \frac{\sigma^2}{2\beta}\left(e^{-\beta|t-s|} - e^{-\beta(t+s)}\right).$$

It directly follows that $X(t)$ has lognormal distribution whose mean and variance can be obtained from those of $Y(t)$.

The models to be discussed in the following were originally developed for the term structures of interest rates and commodity prices. Since interest rates (non-negative) and volatility indexes share mean-reverting and non-negative properties, these models can also be applied in modeling volatility indexes.

2.1.2 Cox-Ingersoll-Ross Model

Cox et.al [38] developed a model ("CIR") that is defined by the following stochastic differential equation for $X(t)$ on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ under historical measure $\mathbb{P}$:

$$dX(t) = \kappa(\theta - X(t))dt + \sigma \sqrt{X(t)}dW^{\mathbb{P}}(t).$$

κ and θ are strictly positive constants, which determine the speed of mean reversion and the the long term mean of $X(t)$, respectively. $\sigma > 0$ is the volatility of the process. $W(t)$ is a standard Brownian motion. An important condition (Feller condition) of the model is $\kappa\theta \geq \frac{\sigma^2}{2}$ which ensures that the process remains positive. So the CIR model itself can be a good candidate to model commodities and volatility indexes due to its positivity and mean-reverting properties.

Cox et.al [38] also have shown that the conditional distribution of $X(t)$ given $X(0)$ is non-central χ^2, $\chi^2[2cX(0); 2q + 2, 2u]$ with degrees of freedom $2q + 2$ and parameter of noncentrality $2u$, where $c = \frac{2\kappa}{\sigma^2(1-e^{-\kappa t})}$, $u = cX(0)e^{-\kappa t}$ and $q = \frac{2\kappa\theta}{\sigma^2} - 1$. The conditional characteristic function for $X(t)$ given $X(0)$ can be obtained as follows, see [67]:

$$\psi(\tau) = \mathbb{E}(e^{i\tau X(t)}|X(0)) = \left(1 - \frac{\sigma^2}{2\kappa}(1 - e^{-\kappa t})i\tau\right)^{-\frac{2\kappa\theta}{\sigma^2}} exp\left\{\frac{e^{-\kappa t}i\tau X(0)}{1 - \frac{\sigma^2}{2\kappa}(1 - e^{-\kappa t})i\tau}\right\}.$$

The mean and variance can be calculated as:

$$\mathbb{E}(X(t)|X(0)) = X(0)e^{-\kappa t} + \theta(1 - e^{-\kappa t}),$$

$$Var(X(t)|X(0)) = X(0)(\frac{\sigma^2}{\kappa})(e^{-\kappa t} - e^{-2\kappa t}) + \theta(\frac{\sigma^2}{2\kappa})(1 - e^{-\kappa t})^2.$$

An autocovariance expression is not given explicitly in the original CIR paper [38]. It is not hard to derive the covariance formula for CIR process, see eg. [103]. For $0 \le s \le t$:

$$Cov(X(t), X(s)) = X(0)(\frac{\sigma^2}{\kappa})(e^{-\kappa t} - e^{-\kappa(t+s)}) + \theta(\frac{\sigma^2}{2\kappa})(e^{-\kappa(t-s)} - e^{-\kappa(t+s)})$$

CIR model is not only used to model asset price or interest rates, it is also used to capture the instantaneous variance in stochastic volatility models. For instance, the renowned Heston model [80] assumes the variance of the equity process follows a CIR.

2.1.3   The 3/2 Process

Heston model has been proved to be able to capture many stylized facts observed in equity option market and it is computationally tractable. However, in reality, Feller condition is not always satisfied when calibrating on data as seen by Da Fonseca and Grasselli [39]. Heston model is also not sensitive to high volatility-of-volatility because it assigns more weight to lower and vanishing volatility as it predicts that the skew in equity option market will flatten when the instantaneous volatility increases [76].

Heston [81] and Platen [117] independently developed a model for the instantaneous volatility as the inverse of a CIR process, which is known as the 3/2 model, for the purpose of eliminating the drawbacks of the original Heston model. Baldeaux and Platen [11] give a brief introduction on 3/2 process. Goard et.al [74] applied this 3/2 process to model VIX and for pricing VIX options.

Suppose $X(t)$ is a 3/2 process on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ under historical measure $\mathbb{P}$, the dynamics is given by:

$$dX(t) = \kappa X(t)(\theta - X(t))dt + \sigma X(t)^{\frac{3}{2}}dW^{\mathbb{P}}(t).$$

Let $\bar{X}(t) = \frac{1}{X(t)}$, we recover the "underlying" CIR process by Ito's lemma:

$$d\bar{X}(t) = (\kappa + \sigma^2 - \kappa\theta\bar{X}(t))dt - \sigma\sqrt{\bar{X}(t)}dW^{\mathbb{P}}(t).$$

If we rearrange the parameters in the $d\bar{X}(t)$ expression: $\tilde{\kappa} = \kappa\theta$, $\tilde{\theta} = \frac{\kappa+\sigma^2}{\tilde{\kappa}}$ and $\tilde{\sigma} = -\sigma$, then $\bar{X}(t)$ is a CIR process. So the conditional distribution of $\bar{X}(t)$ is non-central chi-square. We can find the mean and variance of $X(t)$ given its relationship to $\bar{X}(t)$. There are methods available to find inverse moments of non-central chi-square random variables analytically [115] or asymptotically [68]. Let $X \sim \chi^2(n, \lambda)$, a general formula for $\mathbb{E}(X^s)$, $s \in \mathbb{R}$, $s > -\frac{n}{2}$ is provided in [115] as:

$$\mathbb{E}(X^s) = \int_0^\infty x^s f_X(x) dx = 2^s \sum_{i=0}^\infty \omega_{i,\lambda} \frac{\Gamma(\frac{n}{2} + i + s)}{\Gamma(\frac{n}{2} + i)},$$

$\omega_{i,\lambda}$ is the probability mass function of a Poisson random variable with parameter $\frac{\lambda}{2}$.

Even though an expression to calculate the moments of 3/2 process is given above, there is no closed-form c.f. available for us to apply FFT option pricing method [29] directly. Thanks to the relationship between 3/2 process and CIR process, the best way to handle a 3/2 process is via CIR process as Goard et.al did in pricing options with 3/2 process [74] as the underlying process. Hence, we can infer the statistical properties of 3/2 process with the help of CIR process.

2.1.4   The Hull-White Model

Hull and White [84] proposed two models to be consistent with current term structure of interest rates and the current volatilities of all interest rates. Their work is an extension of Vasicek [126] model and CIR [38] model. Vasicek and CIR models on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ under historical measure $\mathbb{P}$ can be summarized in one expression as:

$$dX(t) = \kappa(\theta - X(t))dt + \sigma X(t)^a dW^{\mathbb{P}}(t).$$

When $a = 0$, it becomes a Vasicek model; when $a = 0.5$, it is a CIR model. Other parameters κ, θ and σ are positive constants controlling mean-reverting speed, long-run average and volatility. The disadvantages of these two models, as critiqued by Hull and White, "are that they involve several unobservable parameters and do not provide a perfect fit to the initial term structure of interest rates." [84] Hull and White further assume time-dependent parameters with an extra time-dependent drift term:

$$dX(t) = [b(t) + \kappa(t)(\theta - X(t))]dt + \sigma(t)X(t)^a dW^{\mathbb{P}}(t).$$

The processes in the two models ($a = 0$ and $a = 0.5$) can be deduced from the term structure of interest rates and the term structure of spot or forward interest-rate volatilities. The deterministic functions $b(t)$ and $\kappa(t)$ of the process can be determined analytically in the case of the extended Vasicek model ($a = 0$), and numerically in the case of the extended CIR model ($a = 0.5$). $\sigma(t)$ should be chosen to reflect the current and future volatilities of $X(t)$ [84].

The distribution for extended Vasicek model (a=0) is normal. Since the parameters are time-varying, the mean and variance depend on time. Let $\theta(t) = b(t) + \kappa(t)\theta$, then the above expression becomes:

$$dX(t) = (\theta(t) - \kappa(t)X(t))dt + \sigma(t)dW^{\mathbb{P}}(t).$$

Then,

$$X(t) = e^{-\int_0^t \kappa(s)ds}[X(0) + \int_0^t e^{\int_0^s \kappa(u)du}\theta(s)ds + \int_0^t e^{\int_0^s \kappa(u)du}\sigma(s)dW^{\mathbb{P}}(s)],$$

so, $X(t)$ has the following moments conditional on $X(0)$:

$$\mathbb{E}(X(t)|X(0)) = e^{-\int_0^t \kappa(s)ds}[X(0) + \int_0^t e^{\int_0^s \kappa(u)du}\theta(s)ds],$$

$$Var(X(t)|X(0)) = e^{-2\int_0^t \kappa(s)ds} \int_0^t e^{2\int_0^s \kappa(u)du}\sigma^2(s)ds,$$

We can also find the autocovariance between $X(t)$ and $X(s)$ for $s < t$ to be:

$$Cov(X(t), X(s)) = e^{-\int_0^t \kappa(u)du - \int_0^s \kappa(u)du} \int_0^s e^{2\int_0^u \kappa(v)dv}\sigma^2(u)du.$$

One important limitation of this modelling, for $a = 1/2$, is the lack of closed-form expressions for the conditional distribution, not only in terms of densities but also the lack of an analytical moment generating function.

2.1.5   The Black-Karasinski Model

Black-Karasinski model [20] is a generalization of Black-Derman-Toy ("BDT") model [19] in continuous time as well as a special case of Hull and White model. BDT model has a continuous time representation on the risk-neutral probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$:

$$dln(X(t)) = \left[b(t) + \frac{\sigma'(t)}{\sigma(t)}ln(X(t))\right]dt + \sigma(t)dW^{\mathbb{Q}}(t).$$

where $\sigma'(t)$ denotes first derivative.

Let $\phi(t) = -\frac{\sigma'(t)}{\sigma(t)}$, Black and Karasinski proposed their model as:

$$dln(X(t)) = \phi(t)(ln(\mu(t)) - ln(X(t)))dt + \sigma(t)dW^{\mathbb{Q}}(t).$$

$\mu(t)$ is defined as the "target interest rate" [20].

Black-Karasinski model has a similar form to the one-factor Schwartz model after log transformation, but it takes time-varying parameters instead of constants. We can obtain the Gaussian distribution for $ln(X(t))$ given $ln(X(0))$ the same way as we did for Hull-White extended Vasicek model:

$$\mathbb{E}(ln(X(t))|ln(X(0))) = e^{-\int_0^t \phi(s)ds}[ln(X(0)) + \int_0^t e^{\int_0^s \phi(u)du}\mu'(s)ds],$$

$$Var(ln(X(t))|ln(X(0))) = e^{-2\int_0^t \phi(s)ds} \int_0^t e^{2\int_0^s \phi(u)du}\sigma^2(s)ds$$

The autocovariance between $ln(X(t))$ and $ln(X(s))$ for $s < t$ is:

$$Cov(ln(X(t)), ln(X(s))) = e^{-\int_0^t \phi(u)du - \int_0^s \phi(u)du} \int_0^s e^{2\int_0^u \phi(v)dv}\sigma^2(u)du.$$

where $\mu'(t) = \phi(t)ln(\mu(t))$.

2.2 Two-Factor Mean-Reverting Models

In this section, we review models with two stochastic factors driving the commodity or volatility index $X(t)$. This is because usually a single process is not able to capture all the stylized facts of a financial instrument. For example, Heston model [80] is able to reproduce the "smile" and "skew" surface of implied volatility, which proves that the single factor Black-Scholes model is not appropriate. So far, we have covered some CEV models in Section 2.1. It is common to assume volatility as another stochastic factor that is driven by a different but correlated stochastic process due to leverage effect, which refers to the negative correlation between an asset return and its volatility. The models reviewed in this section are defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with historical probability measure $\mathbb{P}$.

2.2.1 One-Factor Schwartz Model with CIR Process for Volatility

The model is given by the following system of SDEs:

$$\begin{cases} \frac{dX(t)}{X(t)} & = (L - \beta ln(X(t)))dt + \sqrt{v(t)}dW^{\mathbb{P}}(t), \\ dv(t) & = \alpha(\theta - v(t))dt + \xi \sqrt{v(t)}dB^{\mathbb{P}}(t). \end{cases}$$

where α, θ and ξ are all positive constants with $\alpha\theta \geq \frac{\xi^2}{2}$ and $< W^{\mathbb{P}}(t), B^{\mathbb{P}}(t) > = \rho t$, $-1 < \rho \leq 1$. Benth [16] briefly discussed about the distribution of $Y(t) = ln(X(t))$ with stochastic volatility following a CIR process while assuming independence between $W^{\mathbb{P}}(t)$ and $B^{\mathbb{P}}(t)$. In this case, the characteristic function of $Y(t)$ is found to be:

$$\mathbb{E}[exp(i\tau Y(t))|Y(0)] = exp(i\tau(e^{-\beta t}Y(0) + (1 - e^{-\beta t})\frac{L}{\beta}))\mathbb{E}\left[\left(-\frac{1}{2}\tau^2 \int_0^t v(s)e^{-2\beta(t-s)}ds\right)\right]. \quad (2.1)$$

The expression involves an expectation term whose value is not possible to evaluate analytically. This is a similar problem as the one encountered in the Hull-White model with $a = 1/2$ where the parameters of the CIR process where time dependent. Also noted by Hikspoors and Jaimungal [82], the existence of an analytic solution for the forward price with mean-reverting spot dynamics with Heston model is "doubtful". Benth further analyzed the structure of Equation 2.1 and commented that it "is a Feynman-Kac solution of a partial differential equation...the coefficients in the solution will themselves be solutions of Riccatti equations which seem difficult to solve analytically" [16]. Hence, both [16] and [82] do not provide analytic solution to Equation 2.1. Considering the possibilities of future application to derivative pricing, we have to overcome those shortcomings by using this model. For example, one solution is to approximate $\mathbb{E}\left[\left(-\frac{1}{2}\tau^2 \int_0^t v(s)e^{-2\beta(t-s)}ds\right)\right]$ by a closed form expression. We will explore such approximation in chapter 3.

2.2.2 The Benth Model

Benth model is essentially a one-factor Schwartz model with the addition of stochastic volatility as a second factor. A Barndorff-Nielsen Shephard non-Gaussian O-U process [12] is deliberately chosen for the stochastic process so that the model has a simple structure and offers a

flexible framework for modeling time dependence structures. Analytic pricing of some derivatives is also possible, see [16].

Define the asset process:

$$\frac{dX(t)}{X(t)} = (L + \frac{v^2(t)}{2} - \beta ln(X(t)))dt + \sqrt{v(t)}dW^{\mathbb{P}}(t),$$

then the process for $Y(t) = ln(X(t))$ is:

$$dY(t) = (L - \beta Y(t))dt + \sqrt{v(t)}dW^{\mathbb{P}}(t).$$

Next, we define the process for $v(t)$. First, define following O-U processes:

$$dZ_j(t) = -\gamma_j Z_j(t)dt + dL_j^{\mathbb{P}}(t), j = 1, \ldots, n,$$

where $\gamma_j > 0$ is constant, $L_j^{\mathbb{P}}(t)$ is an increasing and driftless Lévy process with Lévy measure P_{L_j}. The process for $v(t)$ is defined as:

$$v(t) = \sum_{j=1}^{n} w_j Z_j(t),$$

with the conditions that $w_j > 0$ and $\sum_{j=1}^{n} w_j = 1$. The case of $n = 1$ would lead to a two-factor model, appropriate for this section. Any $n > 1$ would lead to multiple factors driving the variance and therefore a $n + 1$ factor model.

Recall the process we defined earlier for $Y(t)$, Benth showed that $Y(t) - \frac{L}{\beta}$ is a sum of mean-variance mixture models with limiting properties:

$$\lim_{t\to\infty} \mathbb{E}(Y(t)) = \frac{L}{\beta},$$

$$\lim_{t\to\infty} Var(Y(t)) = \lim_{t\to\infty} \mathbb{E}[v(t)]/2\beta.$$

These results match what we have for one-factor Schwartz models when constant parameters are assumed if $\sigma^2 =: \lim_{t\to\infty} \mathbb{E}(v(t))$. The autocovariance of $Y(t)$ is proved to be:

$$Cov(Y(t), Y(t + \tau)) = e^{-\beta\tau} \int_0^t \mathbb{E}[v(s)]e^{-2\beta(t-s)}ds,$$

where

$$\int_0^t \mathbb{E}[v(s)]e^{-2\beta(t-s)}ds = \sum_{j=1}^{n} w_j \left[Z_j(0)\gamma(t; 2\beta, \gamma_j) + \int_0^\infty l_j dP_{L_j} \int_0^t \gamma(u; 2\beta, \gamma_j)du, \right]$$

with

$$\gamma(s; a, b) = \begin{cases} \frac{1}{a-b} & (e^{-bs} - e^{-as}), a \neq b \\ & se^{-as}, a = b \end{cases}$$

2.3 4/2 Stochastic Volatility Models

In previous two section, we reviewed some important mean-reverting models in the literature. These mean-reverting models lay the foundation for our development of novel mean-reverting model in this thesis. Recently, a 4/2 stochastic volatility model has been studied, which expands the renowned Heston family, demonstrating interesting properties and versatility. As the other important component in our novel mean-reverting 4/2 stochastic volatility model, we introduce this new member in Heston family and an innovation based on a mean-reverting model in this section.

2.3.1 Grasselli's 4/2 Model

Grasselli [76] proposed what he calls a "4/2" model that combines Heston model and 3/2 model. This model is proposed to solve issues faced by Heston and by 3/2 separately.

As we pointed out in chapter 1, an issue with Heston model is that it is not capable of modeling extreme behaviors of stock prices. In these situations, Heston volatility process requires high volatility-of-volatility parameter to capture the extreme behaviors, thus Feller condition is likely violated. This explains why Heston model is not supported as evidenced in calibration to market data exercises. On the other hand, 3/2 process admits extreme paths, therefore it can capture the stylized facts such as steep "skew" due to abnormal movements in the stock prices' series. Despite the capability of modeling extreme behaviors in stock prices' series, 3/2 process is also found to be "wild" and suffer computational difficulties [49]. The 4/2 model incorporates both Heston and 3/2 models to resolve the issues Heston model and 3/2 model suffer. Inspired by Grasselli's 4/2 model, we extend this model to account for the mean-reverting property so that it can also be applied to other markets like commodity market and volatility index market. A detailed study of our mean-reverting 4/2 model will be performed in chapter 3.

In the original Heston framework, a geometric Brownian motion is assumed for the underlying, $X(t)$, with stochastic volatility as a CIR process. In the 4/2 model, an extra inverse of volatility term is added:

$$\begin{cases} \frac{dX(t)}{X(t)} & = rdt + (a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{Q}}(t), \\ dv(t) & = \alpha(\theta - v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{Q}}(t). \end{cases}$$

a and b are some constants. α, θ and ξ are defined as usual for a CIR process. The Feller condition is also required. $W^{\mathbb{Q}}(t)$ and $B^{\mathbb{Q}}(t)$ are defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, \mathbb{Q})$ with risk-neutral probability measure $\mathbb{Q}$ and $\langle W^{\mathbb{Q}}(t), B^{\mathbb{Q}}(t)\rangle = \rho t$. By defining under $\mathbb{Q}$ measure, it is straightforward to justify the existence of an equivalent $\mathbb{Q}$ measure by checking whether the discounted price process is a true martingale. If the model is defined under historical measure, it takes redundant steps for the same purpose. When $a = 0$, this is the 3/2 model while it becomes the Heston model when $b = 0$. When a and b are strictly positive, $(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})$ is uniformly bounded away from 0 and also greater than $\sqrt{ab}$. This is a new feature that is not presented in either Heston or 3/2 models. In fact, the discovery of this new feature is in line with the stylized fact that volatility processes should not be too close to zero, see e.g. [71].

Grasselli also discussed about a multifactor extension of 4/2 model taking $\mathbf{a} = (\mathbf{a_1}', \mathbf{0}_{1\times m})$, $\mathbf{b} = (\mathbf{0}_{1\times n}, \mathbf{b_2}')$, $\boldsymbol{v}(t) = (\boldsymbol{v_1}(t)', \boldsymbol{v_2}(t)')'$ with $\mathbf{a_1}, \boldsymbol{v_1}(t), \mathbf{0}_{1\times n} \in \mathbb{R}^n$ and $\mathbf{b_2}, \boldsymbol{v_2}(t), \mathbf{0}_{1\times m} \in \mathbb{R}^m$. The processes $\boldsymbol{v_1}(t)$ and $\boldsymbol{v_2}(t)$ are independent, but they are correlated to (vector) Brownian motion $\mathbf{W_t^Q}$ in the usual way. In the simplest case, when $m = n = 1$, the model has the following format:

$$\begin{cases} \frac{dX(t)}{X(t)} & = rdt + a\sqrt{v_1(t)}dW_1(t) + \frac{b}{\sqrt{v_2(t)}}dW_2^Q(t), \\ dv_1(t) & = \alpha_1(\theta_1 - v_1(t))dt + \xi_1\sqrt{v_1(t)}dB_1^Q(t), \\ dv_2(t) & = \alpha_2(\theta_2 - v_2(t))dt + \xi_2\sqrt{v_2(t)}dB_2^Q(t), \end{cases}$$

where $W_1^Q(t)$ and $W_2^Q(t)$ are independent, $B_1^Q(t)$ and $B_2^Q(t)$ are independent and $\langle W_k^Q(t), B_k^Q(t)\rangle = \rho_k t, k = 1, 2$. Since the CIR factor and 3/2 factor are independent, this model shares the same limitations as the usual affine specification, that is the positive factors must be driven by independent noises, see [76]. Also note that this setting with two extra factors does not lead to a closed-form characteristic function. Nonetheless, this setting offers more flexibility than a single factor driving the volatility, which receives a comprehensive treatment in [76].

In the single factor case, an analytical conditional characteristic function is derived using Lie symmetries for partial differential equations. The results is summarized in Proposition 3.1 in [76] and presented here for completeness.

$$\psi_{0,t}(\tau) = \mathbb{E}\Big(e^{\tau Y(t)}|Y(0)\Big) = exp\Big\{\tau Y(0) + \frac{\alpha^2\theta}{\xi^2}t + \tau\Big(r - ab - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi}\Big)t + \tau^2(1-\rho^2)abt\Big\}$$

$$\cdot\Big(\frac{\sqrt{A_\tau}}{\xi^2 sinh(\frac{\sqrt{A_\tau}}{2}t)}\Big)^{m_\tau + 1}v(0)^{\frac{1}{2} + \frac{m_\tau}{2} - \frac{\alpha\theta}{\xi^2} - \frac{\tau b\rho}{\xi}}\Big(K_\tau(t) - \frac{\tau a\rho}{\xi}\Big)^{-\left(\frac{1}{2} + \frac{m_\tau}{2} + \frac{\alpha\theta}{\xi^2} + \frac{\tau b\rho}{\xi}\right)}$$

$$\cdot exp\Big\{\frac{v(0)}{\xi^2}\Big(-\sqrt{A_\tau}coth\Big(\frac{\sqrt{A_\tau}t}{2}\Big) + \alpha - \tau a\rho\xi\Big)\Big\}\frac{\Gamma(\frac{1}{2} + \frac{m_\tau}{2} + \frac{\alpha\theta}{\xi^2} + \frac{\tau b\rho}{\xi})}{\Gamma(m_\tau + 1)}$$

$$\cdot {}_1F_1\Big(\frac{1}{2} + \frac{m_\tau}{2} + \frac{\alpha\theta}{\xi^2} + \frac{\tau b\rho}{\xi}, m_\tau + 1, \frac{A_\tau v(0)}{\xi^4 sinh^2(\frac{\sqrt{A_\tau}t}{2})(K_\tau(t) - \frac{\tau a\rho}{\xi})}\Big),$$

where,

$$A_\tau = \alpha^2 - 2\xi^2\Big(\tau\Big(\frac{a\rho\alpha}{\xi} - \frac{1}{2}a^2\Big) + \frac{1}{2}\tau^2(1-\rho^2)a^2\Big),$$

$$m_\tau = \frac{2}{\xi^2}\sqrt{\Big(\alpha\theta - \frac{\xi^2}{2}\Big)^2 - 2\xi^2\Big(\tau\Big(\frac{b\rho}{\xi}\Big(\frac{\xi^2}{2} - \alpha\theta\Big) - \frac{1}{2}b^2\Big) + \frac{1}{2}\tau^2(1-\rho^2)b^2\Big)},$$

$$K_\tau(t) = \frac{1}{\xi^2}\Big(\sqrt{A_\tau}coth\Big(\frac{\sqrt{A_\tau}t}{2}\Big) + \alpha\Big),$$

(2.2)

and $\Gamma(\cdot)$, ${}_1F_1(\cdot, \cdot, \cdot)$ denote the Gamma and hypergeometric confluent functions respectively. The characteristic function is well defined for all $t \geq 0$ when the complex number τ belongs to the strip $\mathcal{D}_{0,+\infty} = \mathcal{A}_{0,+\infty} + i\mathbb{R} \subset \mathbb{C}$, where $\mathcal{A}_{0,+\infty}$ is given by:

$$\mathcal{A}_{0,+\infty} = \{\tau \in \mathbb{R} : A_\tau \geq 0 \text{ and } f_i(\tau), (i = 1, \ldots, 4)\},$$

with

$$f_1(\tau) = \tau\left(\frac{a\rho\alpha}{\xi} - \frac{1}{2}a^2\right) + \frac{1}{2}\tau^2(1 - \rho^2)a^2 - \frac{\alpha^2}{2\xi^2} \le 0,$$

$$f_2(\tau) = \frac{1}{2} + \frac{m_\tau}{2} + \frac{\alpha\theta}{\xi^2} + \frac{\tau b\rho}{\xi} > 0,$$

$$f_3(\tau) = \left(\alpha\theta - \frac{\xi^2}{2}\right)^2 - 2\xi^2\left(\tau\left(\frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right) - \frac{1}{2}b^2\right) + \frac{1}{2}\tau^2(1 - \rho^2)b^2\right) \ge 0$$

$$f_4(\tau) = \sqrt{A_\tau} + \alpha - \tau a\rho\xi \ge 0.$$

When $t < \infty$, $\tau \in \mathcal{D}_{0,t} = \mathcal{A}_{0,t} + i\mathbb{R}$ with $\mathcal{A}_{0,t} = \{\tau \in \mathbb{R} : A_\tau \ge 0$ *and* $f_i(\tau), (i = 1, \ldots, 3)$ *and* $f_4(\tau) <$ $0\}$. Then there exists a maximal (explosion) time t^* given by:

$$t^* = \frac{1}{\sqrt{A_\tau}}log\left(1 - \frac{2\sqrt{A_\tau}}{\alpha - \tau a\rho\xi + \sqrt{A_\tau}}\right).$$

i.e when $f_1(\tau) - f_3(\tau)$ are satisfied but $f_4(\tau) < 0$, there is a t^* such that for $t < t^*$, $K_\tau(t) - \frac{\tau a\rho}{\xi}$ remains positive, therefore $\psi_{0,t}(\tau)$ is well defined. Grasselli also provides an exact simulation approach for the 4/2 model. Exact simulation approach does not suffer the problem that a discretization approach normally does, which introduces a bias. A drawback of exact simulation in the case of the 4/2 model is that it is time-consuming comparing to discretization. The goal of this exact simulation approach is to simulate the pair $(Y(t), v(t))$ for any t, hence producing the distribution of the pair at any time. If we know the distribution of $Y(t)$ conditional on $v(t)$ and the distribution of $v(t)$, then we can sample $(Y(t), v(t))$ exactly. Given $v(t)$ has non-central χ^2 distribution, the key step is to solve for the distribution of $Y(t)$ given $v(t)$. Proposition 4.1 of [76] gives this special moment generating function:

$$\mathbb{E}\left(e^{uY(t)}|v(t)\right) = exp\left\{uY(0) + u\left(r - ab - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi}\right)t + u^2(1 - \rho^2)abt\right\}$$

$$\cdot exp\left\{\frac{ua\rho}{\xi}(v(t) - v(0)) + \frac{ub\rho}{\xi}log\frac{v(t)}{v(0)}\right\}$$

$$\cdot \frac{\sqrt{A_u}sinh(\frac{\alpha t}{2})}{\alpha sinh(\frac{\sqrt{A_u}t}{2})}e^{\frac{v(0)+v(t)}{\xi^2}\left(\alpha coth(\frac{\alpha t}{2}) - \sqrt{A_u}coth(\frac{\sqrt{A_u}t}{2})\right)}$$

$$\cdot \frac{I_{\frac{2}{\xi^2}\sqrt{\left(\alpha\theta - \frac{\xi^2}{2}\right)^2 + 2\xi^2 B_u}}\left(\frac{2\sqrt{A_u v(t)v(0)}}{\xi^2 sinh\left(\frac{\sqrt{A_u}t}{2}\right)}\right)}{I_{\frac{2\alpha\theta}{\xi^2}-1}\left(\frac{2\alpha\sqrt{v(t)v(0)}}{\xi^2 sinh\left(\frac{\alpha t}{2}\right)}\right)},$$

with

$$A_u = \alpha^2 - 2\xi^2\left(u\left(\frac{a\rho\alpha}{\xi} - \frac{1}{2}a^2\right) + \frac{1}{2}u^2(1 - \rho^2)a^2\right),$$

$$B_u = u\left(\frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right) - \frac{1}{2}b^2\right) + \frac{1}{2}u^2(1 - \rho^2)b^2,$$

where $I_z(x)$ is the modified Bessel function of the first kind.

Grasselli puts a note on pricing using 4/2 model that a risk-neutral measure may not exist, which is also the case for the 3/2 model. Further, he shows that if the discounted stock price is a true martingale under (putative) risk-neutral measure, the process $v(t)$ has to satisfy Feller condition under both historical and risk-neutral measure. At the end, the process $v(t)$ is discovered to have different behavior at 0 under the two probability measures if:

$$2\alpha\theta + 2\rho\xi b < \xi^2 \le 2\alpha\theta.$$

Hence, risk-neutral pricing methodology is not applicable in this case, but one can still use the Benchmark approach, which only assumes the existence of a numéraire portfolio for both portfolio optimization and derivative pricing. Using the numéraire portfolio, derivative pricing can be done with historical probability measure, thus it does not require a risk-neutral measure to exist, see [10] for the theory of benchmark approaches to pricing.

2.3.2   Mean-Reverting 4/2 Stochastic Volatility Model with Time-Dependent Parameters

The model considered in this section is a variation of the mean-reverting 4/2 stochastic volatility model that will be studied in detail in chapter 3. The model we present here is a partial improvement as it allows for time-dependent parameters while imposing some structural conditions. Let us define the process on the interval $[0, T]$ on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$:

$$\begin{cases} \frac{dX(t)}{X(t)} & = [L + c(t)(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta(t)ln(X(t))]dt + g(t)\left(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)dW^{\mathbb{P}}(t), \\ dv(t) & = \alpha(\theta - v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{P}}(t) \\ \langle dW^{\mathbb{P}}(t), dB^{\mathbb{P}}(t)\rangle & = \rho dt, \end{cases}$$

where a, b, L are arbitrary constants; α, θ, ξ are positive constants satisfying Feller condition: $\alpha\theta \ge \frac{\xi^2}{2}$, and $-1 \le \rho \le 1$; $c(t) = \frac{1}{2}g^2(t) + c_0$, $g(t) = \exp\left(\int_t^T \beta(s)ds\right)$ and $\beta(t)$ are deterministic functions of time. Let $Y(t) = ln(X(t))$ then we have:

$$dY(t) = \left[L + c_0(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta(t)Y(t)\right]dt + g(t)\left(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)dW^{\mathbb{P}}(t)$$

Define $Z(t) = Y(t)\exp\left(-\int_t^T \beta(s)ds\right)$ then:

$$dZ(t) = \left[L + c_0\left(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)^2\right]dt + \left(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)dW^{\mathbb{P}}(t).$$

The SDE of $Z(t)$ process represents the SDE of the log stock price in [76], which means we are able to find a closed-form c.f. for $Z(T)$ conditional on $\mathcal{F}_t$, i.e $\psi_{t,T}(\tau) = \mathbb{E}\left(e^{\tau Z(t)}|\mathcal{F}_t\right)$, by directly using the results from [76] with L instead of r.

A natural question is what the possible candidates for $\beta(t)$ are. We propose two candidates for $\beta(s)$ in the examples below.

Example 1:

$$\begin{aligned}
\beta(t) &= \beta_0 > 0 \\
g(t) &= \exp\left((T-t)\beta_0\right)
\end{aligned}$$

Note in this example the speed of reversion $\beta(t)$ is constant β_0 but the volatility as controlled by $g(t)$ decreases with time since $g(0) = \exp(\beta_0 T)$ is greater than $g(T) = 1$, which seems practically counterintuitive because the volatility should increase as clock tics, in other words, we would expect the underlying asset price to be more volatile further into future due to uncertainty becoming more obvious.

Example 2:

$$\begin{aligned}
\beta(t) &= \beta_0 - \beta_1 t > 0 \text{ (with } 0 \leq \beta_1 \leq \frac{\beta_0}{T}) \\
g(t) &= \exp\left((T-t)\beta_0 - \left(\frac{T^2}{2} - \frac{t^2}{2}\right)\beta_1\right)
\end{aligned}$$

In this example, the mean reversion speed decreases with time. On the other hand, we have control on how "slowly" the variance decreases with time from $g(0) = \exp\left(\beta_0 T - \frac{T^2}{2}\beta_1\right)$ to $g(T) = 1$ compared to Example 1. Note that an increase in β_0 leads to drastic decrease in $g(t)$, hence a reduction of the volatility.

2.4 Multidimensional Models

In previous sections we review models for a single asset from the basic one-factor model with constant volatility to complex stochastic volatility models with volatility as another stochastic factor, and there is potential for more factors. In practice, when it comes to pricing multi-asset options, this requires modeling joint behavior of many underlying assets with likely as many extra factors. That is why we need to study multidimensional models.

In this section we first review some existing multidimensional models in sections 2.4.1-2.4.2. We list some advanced models incorporating the newly developed 4/2 stochastic volatility in section 2.4.3. The model introduced in section 2.4.3.1 has not been studied in details, but it can be a topic for future research.

Similar to section 2.2, we define the following models on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ under historical measure $\mathbb{P}$.

2.4.1 One-Factor Schwartz Model In Multi-dimension

Ware et.al [47] studied a portfolio with multiple commodities and the optimal allocation problem for such portfolio. A multivariate model based on one-factor Schwartz model [122] is

proposed in the aforementioned paper. Such model has all the ingredients to be entertained in
the context of multiple volatility indexes.

Suppose we observe prices for n assets. We define the model as follows:

$$\frac{dX_i(t)}{X_i(t)} = (L_i - \beta_i ln(X_i(t)))dt + \sum_j a_{ij} dW_j^{\mathbb{P}}(t), \ X_i(t) > 0, \ j = 1, \dots, n, \tag{2.3}$$

where $X_i(t)$ is the price for i-th asset; $\beta_i > 0$ is the mean-reversion speed; $\frac{L_i}{\beta_i}$ is the long term
average of each $X_i(t)$; $W_j^{\mathbb{P}}(t)$ is standard Brownian motion and independent for all j; a_{ij} is the
ij-th entry of volatility matrix $\mathbf{\Sigma}$.

Define $Y_i(t) = ln(X_i(t))$, then by Ito's lemma $Y_i(t)$ is a standard O-U process in the follow-
ing form:

$$dY_i(t) = \beta_i(L_i^* - Y_i(t))dt + \sum_j a_{ij} dW_j^{\mathbb{P}}(t),$$

$$L_i^* = L_i - \frac{\sum_j a_{ij}^2}{2\beta_i}.$$

So $\mathbf{Y(t)} = (Y_1(t), \dots, Y_n(t))$ is a normally distributed random vector with mean and covari-
ance [47]:

$$\mathbb{E}(Y_i(t)) = e^{-\beta_i t} Y_i(0) + (1 - e^{-\beta_i t})L_i^*,$$

$$Cov(Y_i(t), Y_j(t)) = \frac{\sum_k a_{ik} a_{jk}}{\beta_i + \beta_j}(1 - e^{-(\beta_i + \beta_j)t}).$$

It directly follows that $\mathbf{X(t)} = (X_1(t), \dots, X_n(t))$ is a multivariate lognormal random vector. The
mean and variance of $X_i(t)$ can be calculated using the mean and variance of $Y_i(t)$.

2.4.2   General Multivariate One-factor Schwartz Model

The one-factor Schwartz model can be generalized to multidimensions in a way that allows for
asset i to influence the expected returns of asset j. Let us take $Y_i(t)=ln(X_i(t))$ where $X_i(t)$ is
either a commodity or a volatility index, the model would be:

$$dY_i(t) = (L_i - \sum_j \beta_{ij} Y_j(t))dt + \sum_j a_{ij} dW_j^{\mathbb{P}}(t), \ j = 1, \dots, n. \tag{2.4}$$

The above models are present in Larsen's paper [95] in the context of optimal investment
strategies in the international economy. The model is used to model short interest rates in two
countries.
In practice, one could have used the model for commodities or volatility indexes. For example,
one could use the following model for VIX and VIXC jointly:

$$d(ln(VIX_t)) = (L_1 - \beta_{11} ln(VIX_t) - \beta_{12} ln(VIXC_t))dt + a_{11} dW_1^{\mathbb{P}}(t) + a_{12} dW_2^{\mathbb{P}}(t), \tag{2.5}$$

$$d(ln(VIXC_t)) = (L_2 - \beta_{21} ln(VIX_t) - \beta_{22} ln(VIXC_t))dt + a_{21} dW_1^{\mathbb{P}}(t) + a_{22} dW_2^{\mathbb{P}}(t). \tag{2.6}$$

Where L_1, L_2, β_{11}, β_{12}, β_{21} and β_{22} are all positive constants. a_{11}, a_{12}, a_{21} and a_{22} are entries affecting Σ. The correlation between the volatility indexes of two countries (US and Canada in the example) are incorporated in a_{ij} for $i, j = 1, 2$. An important property of this model is that the interdependence of the drift in volatility indexes across markets can be captured, namely the spillover effects are taken into account for. Generally speaking, spillover effects refer to the impact of one asset or underlying stock on the remaining assets/stocks in regards to trends (drift) or the long term average of a second asset. We will emphasize the concept of spillover effects again in chapter 5.

In such situation, depending on the relative effect that one index has on the other, we can choose the drift parameters accordingly. In the example, Canadian market is tied to US market, so VIX is more influential than VIXC. Hence, we can set β_{12} to 0, and (2.5) becomes a standard O-U process. On the other hand, if we are interested in studying VIX and VSTOXX (the volatility index for Euro STOXX 50) together, then β_{12} need not be 0 since US market and European market are comparable, VIX and VSTOXX shall impact one another. It is also assumed that $\sqrt{(\beta_{11} - \beta_{22})^2 + 4\beta_{12}\beta_{21}} \leq \beta_{11} + \beta_{22}$. This assumption is originally from [94] to ensure that the eigenvalues of $\begin{pmatrix} -\beta_{11} & -\beta_{12} \\ -\beta_{21} & -\beta_{22} \end{pmatrix}$ are all negative so that the term structure does not go to infinity. Later in Larsen's paper, she shows that zero-coupon bond prices can go to infinity, in finite time, without such assumption. In chapter 5, we also impose this assumption on our model.

Nielsen and Saá-Requejo [113] proposed similar models in 1993 in a Cox-Ingersoll-Ross framework. Larsen argued that "the reason for choosing a Gaussian setup instead of a CIR setup is the tractability of a Gaussian model." [95]. The argument again points to the limitations of finding closed-form expressions for moment generating function in the context of mean-reverting stochastic volatility models, obviously affecting larger dimensions.

2.4.3   General Multivariate Multifactor Models

In mathematical finance, factor models are used to explore the relationship between asset returns and other market/risk factors. For example, the capital asset pricing model (CAPM) is a one-factor model. The "one" here refers to the number of common factors excluding intrinsic factor; therefore this is equivalent to a two-factor model in the language of this thesis. The common factor in CAPM, called systematic factor, explains the dependence among stock returns. Later, Fama and French [64] extend this basic model to include three common factors: excess market return, stock size and book-to-market equity. These are discrete-time models targeting the empirical discrete time data available, the power of these factors in explaining stock returns are empirically studied via linear regression models in the literature.

The extension from discrete-time to continuous-time has been neglected in the literature. For stochastic processes, one is not only interested in factors that explain the daily behaviour of stocks, but also in terms of capturing the stylized facts reported on the assets under study. The factors in the context of continuous-time stochastic process can be observable (extension of CAPM) or unobservable (Factor Analysis), naturally they would be continuous-time stochastic

processes themselves and could appear in nonlinear ways.

In previous sections we reviewed one-factor Schwartz model, which has only the asset price as stochastic factor. Schwartz also proposes two-factor and three-factor models which include convenience yield and interest rate as, potentially unobservable/hidden, stochastic factors [122]. The well known Heston model is also a two-factor model with stock price and volatility as factors. Although Heston outperforms the classic Black-Scholes model in producing "volatility smile" and "volatility skew/smirk", it cannot "explain such largely independent fluctuations in its level and slope over time" [36]. Christoffersen et.al [36] first performed a principal component analysis on Black-Scholes' implied variance data to investigate the factors that determine the variation of the data. They found the first two of the four principal components explain more than 95% of the variation. Furthermore, the authors found that the first two principal components are highly correlated to volatility level and the slope of the smirk. A simple stylized fact about volatility level and the slope of the smirk is that they are largely independent. All the analysis point to a two-factor model which can capture the stylized fact. the authors proposed a multifactor stochastic covariance model based on Heston model trying to reproduce the shape and term structure of the "smirk". This is an example of multiple unobservable stochastic factors been used to capture stylized facts of a single asset.

Similarly, the authors assume a Heston-type multifactor stochastic covariance model with principal component analysis treatment for the purpose of modeling foreign exchange rates in [41]. Escobar and Moreno further studied a multifactor Heston model with jumps in the context of dynamic portfolio optimization [62]. Another direction of studying the structure of multifactor covariance models is by applying principal component analysis techniques on the stochastic covariance matrix, see for example [41] [60]. In [41], the authors consider a similar setting to [36]. The methodology undertaken in [60] directly goes into the heart of principal component–the authors decompose the covariance matrix using principal component techniques yielding constant eigenvectors and stochastic eigenvalues. Structurally the diffusions in [41] and [60] may seem the same; however, with the model in [60], we are able to rewrite the original multifactor stochastic covariance model in terms of independent one dimensional processes (the principal components) as the eigenvectors contains the correlation information among assets, which in fact reduces the dimension of the model in that theory developed for one dimensional model, in this case Heston, can be directly applied thanks to the independence property.

In this section, we review advanced stochastic factor models in the spirit of [60] by assuming advanced processes on the systematic risk factors, which accounts for the total common variance among the assets, and also advanced processes on the idiosyncratic factors. Generally, the systematic risk is the risk that affects the market and it is therefore non-diversifiable. Without the addition of the systematic factors, the variation of the model only comes from the assets themselves. We assume no interaction between the systematic factor and the existing idiosyncratic factors. See [32] for details.

We first provide the processes under the historical measure $\mathbb{P}$. Suppose that $\mathbf{X_t} = (X_1(t), ..., X_n(t))'$

is a vector of asset prices with the following $\mathbb{P}$-measure representation:

$$
\begin{cases}
\frac{dX_i(t)}{X_i(t)} &= \left\{ L_i + c_i \sum_{j=1}^{p} a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum_{j=1}^{p} \beta_{ij} \ln(X_j(t)) + \tilde{c}_i \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right)^2 \right\} dt \\
&\quad + \sum_{j=1}^{p} a_{ij} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t) + \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) d\widetilde{W}_i^{\mathbb{P}}(t) \\
dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \quad j = 1, ..., p \\
d\tilde{v}_i(t) &= \tilde{\alpha}_i(\tilde{\theta}_i - \tilde{v}_i(t))dt + \tilde{\xi}_i \sqrt{\tilde{v}_i(t)} d\tilde{B}_i^{\mathbb{P}}(t), \quad i = 1, ..., n
\end{cases}
$$

The quadratic variation structure is $\left\langle dB_j^{\mathbb{P}}(t), dW_j^{\mathbb{P}}(t) \right\rangle = \rho_j dt, \left\langle d\tilde{B}_i^{\mathbb{P}}(t), d\widetilde{W}_i^{\mathbb{P}}(t) \right\rangle = \tilde{\rho}_i dt$ and zero otherwise. In the language of factor analysis, a_{ij} is the ijth entry of the matrix of factor loadings (A) that captures the correlations among assets. The commonalities are represented by $V_j(t) = \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2$ (in matrix form, $\mathbf{\Lambda}_{nxp} = \mathbf{A} diag\left(\mathbf{V_t^{1/2}} \right)$) and the intrinsic residual variance is $\Psi = diag(\widetilde{\mathbf{V_t}})$, with $\widetilde{V}_j(t) = \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right)^2$. This leads to a factor decomposition of the quadratic variation of asset prices as follows:

$$
\mathbf{\Sigma_t} dt = (\mathbf{\Lambda\Lambda'} + \Psi) dt = \left(\mathbf{A} diag(\mathbf{V_t}) \mathbf{A'} + diag(\widetilde{\mathbf{V_t}}) \right) dt
$$

Note the correlation between assets is incorporated in a_{ij}'s, See [32] and chapter 4 for more details about this model. We include a spacial case in section 2.4.3.1 where the impact of intrinsic factor is ignored.

Whenever necessary, we assume $n = p$ and $A = (a_{ij})_{n \times p}$ to be an orthogonal matrix. In this setting, c_i and $\tilde{c}_i$ represent risk premiums of asset $X_i(t)$ associated with the common and intrinsic factors, respectively. $\beta = (\beta_{ij})_{n \times n}$ is an invertible matrix, which captures the spillover at the expected return level $X_i(t)$ on asset $X_j(t)$. In other words, it represents the impact from other assets on the long term average price of the current one.

Based on the quadratic variation relationship defined in this model, if we assume that $B_j^{\mathbb{P}}$, $B_j^{\mathbb{P}}(t)^{\perp}$, $\tilde{B}_i^{\mathbb{P}}(t)$, $\tilde{B}_i^{\mathbb{P}}(t)^{\perp}$ are independent Brownian motions with $-1 \le \rho_j \le 1$ and $-1 \le \tilde{\rho}_i \le 1$. Then,

$$
dW_j^{\mathbb{P}}(t) = \rho_j dB_j^{\mathbb{P}}(t) + \sqrt{1 - \rho_j^2} dB_j^{\mathbb{P}}(t)^{\perp}
$$

$$
d\widetilde{W}_i^{\mathbb{P}}(t) = \tilde{\rho}_i d\tilde{B}_i^{\mathbb{P}}(t) + \sqrt{1 - \tilde{\rho}_i^2} d\tilde{B}_i^{\mathbb{P}}(t)^{\perp}.
$$

v_j $j = 1, .., n$ and $\tilde{v}_i$ $i = 1, .., n$ follow standard CIR processes, hence α_j, θ_j, and ξ_j are positive constants satisfying $\alpha_j \theta_j \ge \frac{\xi_j^2}{2}$ (the Feller condition). Similarly, $\tilde{\alpha}_i$, $\tilde{\theta}_i$, and $\tilde{\xi}_i$ are positive constants satisfying $\tilde{\alpha}_i \tilde{\theta}_i \ge \frac{\tilde{\xi}_i^2}{2}$. Note that the Feller condition guarantees that the CIR processes remain positive and therefore the covariance matrix is definite positive.

The transformation $\mathbf{Y_t} = \ln(\mathbf{X_t})$ would create a multivariate Ornstein–Uhlenbeck process with

a 4/2 stochastic factor structure:

$$
\begin{aligned}
dY_i(t) &= \left\{ L_i + \left(c_i - \frac{1}{2}\right) \sum_{j=1}^{n} a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \sum_{j=1}^{n} \beta_{ij} Y_j(t) + \left(\tilde{c}_i - \frac{1}{2}\right)\left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}}\right)^2 \right\} dt \\
&\quad + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right) dW_j(t) + \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}}\right) d\widetilde{W}_i(t)
\end{aligned}
$$

To stay in line with [94] and [95], we enforce the same assumption that the eigenvalues of the matrix

$$
\begin{pmatrix}
-\beta_{11} & \cdots & -\beta_{1n} \\
\vdots & \ddots & \vdots \\
-\beta_{n1} & \cdots & -\beta_{nn}
\end{pmatrix}
$$

are all negative.

2.4.3.1 Special Case: No Intrinsic Factor

In this section, we introduce a model as a special case of our generalized mean-reverting 4/2 factor model. In this case, the only factors that impacts the asset are the common/idiosyncratic factors, i.e $\Psi = 0$. This can be interpreted as a Principal Component model.

Suppose $X_t = (X_1(t), \ldots, X_n(t))'$ is a vector of assets, the dynamics of log price $Y_i(t) = ln(X_i(t))$ is then:

$$
\begin{cases}
dY_i(t) &= \left[L_i + (c_i - \frac{1}{2}) \sum_{j=1}^{n} a_{ij}^2 (\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}})^2 - \sum_{j=1}^{n} \beta_{ij} Y_j(t) \right] dt + \sum_{j=1}^{n} a_{ij}(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}) dW_j^{\mathbb{P}}(t), \\
dv_j(t) &= \alpha_j(\theta_j - v_j(t)) dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \quad j = 1, \ldots, n.
\end{cases}
$$

$$(2.7)$$

For the covariance matrix Σ_t, the factor decomposition is therefore:

$$
\Sigma_t dt = \Lambda\Lambda' dt = A\, diag(V_t) A'\, dt
$$

For example, we are modeling two assets leads to:

$$
\begin{aligned}
\frac{dX_1(t)}{X_1(t)} &= \left[L_1 + a_{11}^2 c_1 (\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}})^2 + a_{12}^2 c_1 (\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}})^2 - \beta_{11} ln(X_1(t)) - \beta_{12} ln(X_2(t)) \right] dt \\
&\quad + a_{11}(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}}) dW_1^{\mathbb{P}}(t) + a_{12}(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}}) dW_2^{\mathbb{P}}(t), \\
\frac{dX_2(t)}{X_2(t)} &= \left[L_2 + a_{21}^2 c_2 (\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}})^2 + a_{22}^2 c_2 (\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}})^2 - \beta_{21} ln(X_1(t)) - \beta_{22} ln(X_2(t)) \right] dt \\
&\quad + a_{21}(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}}) dW_1^{\mathbb{P}}(t) + a_{22}(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}}) dW_2^{\mathbb{P}}(t),
\end{aligned}
$$

In this chapter, we have provided an overview on models that are related to this thesis. We focused mostly on mean-reverting models as they are the fundamental models to this thesis.

The majority of the models we covered are well-known. In cases like section 2.3.2, the model is an extension of Grasselli's 4/2 model, yet has not been studied. We entertain this extension because it is a natural generalization of the 4/2 model with the potential of a closed-form c.f.. The challenge is, as we pointed out, how to determine the time-dependent coefficients such that they are financially reasonable. The model discussed section 2.4.3 has been studied in a recent paper by Cheng et.al [32]. The model introduced in Section 2.4.3.1 will be studied in details in chapter 5.

Chapter 3

The Mean-Reverting 4/2 Stochastic Volatility Model

In this chapter we define and study a stochastic process that combines two important stylized facts of financial data: reversion to the mean, and a flexible generalized stochastic volatility process: the 4/2 process. Our work is motivated by the modeling of at least two financial asset classes: commodities and volatility indexes. Thus, we do not consider convenience yield in our model since this is not a common property shared by all mean-reverting assets. Our model targets a broad range of asset classes, not only commodities. We provide analytical expressions for the conditional characteristic functions and closed-form approximations to relevant particular cases such as a mean-reverting Heston stochastic volatility model. The results also cover feasible changes of measure with the final aim of pricing financial products. The empirical analysis and the estimation methodology confirm the need of such a model in several examples from the targeted asset classes. Applications to option pricing corroborate the substantial impact on the implied volatility surfaces of the new parameters.

The chapter is organized in as follows: the models (Type I and Type II) and its main properties and simulation are covered in section 3.1. We derive a semi-closed conditional characteristic functions and provide closed-form approximations to the characteristic function in section 3.2. We study feasible changes of measure in section 3.3. In section 3.4, we propose two estimation methods, their performance is shown numerically and they are applied to several financial time series. In section 3.5, we apply our model to the pricing of vanilla European call options two different markets: VIX and Gold ETF (GLD) using simulation and Fast Fourier Transform (FFT) approaches.

3.1   Definition and Properties

In this section, we first give the definition of the mean-reverting 4/2 stochastic volatility model as well as volatility process and volatility-of-volatility process. We also derive a representation for the leverage effect, which displays interesting property and coincides with empirical experience. Next, we outline the Euler discretization scheme for simulation of our model including a discussion of other available discretization schemes.

3.1.1 Model Specifications

We assume a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with historical probability measure $\mathbb{P}$. Our model has the following representation:

$$\begin{cases} \frac{dX(t)}{X(t)} & = [L + c(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta ln(X(t))]dt + (a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{P}}(t), \\ dv(t) & = \alpha(\theta - v(t))dt + \xi \sqrt{v(t)}dB^{\mathbb{P}}(t). \end{cases} \quad (3.1)$$

where $X(t)$ is represents the underlying asset price (commodity or volatility index); $W^{\mathbb{P}}(t)$ and $B^{\mathbb{P}}(t)$ are correlated Brownian motions with $< W^{\mathbb{P}}(t), B^{\mathbb{P}}(t) >= \rho t$; and a, b, L and c are arbitrary constants. The positive constants α, θ and ξ satisfy Feller condition (Condition I): $2\alpha\theta \geq \xi^2$. When $a = 0, b \neq 0$, our model becomes a 3/2 model while $a \neq 0, b = 0$ leads to the 1/2 (Heston) model. When a and b are strictly positive, $a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}$ is uniformly bounded away from 0 and is also greater than $\sqrt{ab}$.

Practitioners and researchers are also interested in the logarithm of asset price, especially in commodity market. In fact, Schwartz's models are developed based on what is known as exponential O-U process: the logarithm of asset price takes normal O-U process. Let $Y(t) = ln(X(t))$ then we have:

$$dY(t) = \left[L + (c - \frac{1}{2})(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta Y(t)\right]dt + (a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{P}}(t). \quad (3.2)$$

Some stylized facts of our model are presented next. First, the variance,

$$V^2(t) = a^2 v(t) + \frac{b^2}{v(t)} + 2ab \quad (3.3)$$

is a combination of a 1/2 and 3/2 processes, both driven by $v(t)$, therefore it is also mean-reverting, the volatility follows the SDE:

$$dV(t) = \left\{\left[\frac{a}{2\sqrt{v(t)}} - \frac{b}{2v^{\frac{3}{2}}(t)}\right]\alpha(\theta - v(t)) + \xi^2\left[\frac{3b}{8v^{\frac{3}{2}}(t)} - \frac{a}{8v^{\frac{1}{2}}(t)}\right]\right\}dt + \left[\frac{a\xi}{2} - \frac{b\xi}{2v(t)}\right]dB^{\mathbb{P}}(t).$$

Another byproduct of this model is that the variance of volatility, $(v')^2(t) = \frac{\xi^2}{4}\left[a - \frac{b}{v(t)}\right]^2$, is a fractional function of $v(t)$, which is also mean-reverting, the SDE of the volatility of volatility is given next:

$$dv'(t) = \frac{b\xi}{2v^2(t)}[\alpha\theta - \xi^2 - \alpha v(t)]dt + \frac{b\xi^2}{2v^{\frac{3}{2}}(t)}dB^{\mathbb{P}}(t) \quad (3.4)$$

The leverage usually refers to the negative correlation between the return of the asset and the change of asset volatility. In our case we obtain:

$$\rho_0 = sign\left(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)sign\left(\left[\frac{a\xi}{2} - \frac{b\xi}{2v(t)}\right]\right)\rho \quad (3.5)$$

Because ξ and $v(t)$ are positive, then Equation 3.5 is further simplified to:

$$\rho_0 = sign\left(a + \frac{b}{v(t)}\right) sign\left(a - \frac{b}{v(t)}\right)\rho \tag{3.6}$$

From Equation 3.5, we can see that the sign of leverage effect is determined by both sign of volatility and sign of volatility of volatility, which is not surprising. From Equation 3.6, it is clear that the sign of leverage effect changes over time whenever $v(t)$ crosses $|\frac{b}{a}|$, this is similar to Stein and Stein's model [123]. The changing sign of leverage effect suggests that the correlation between the return of the asset and the change of asset volatility is not always negative over time, sometimes higher volatility leads to better return on asset. This coincides with the fact that in the long run risky asset tends to generate higher payoff.

3.1.2   Model Simulation

There is rich literature on how to simulate Heston model, see [99] for a summary of discretization schemes of a CIR process. Andersen[4] proposed two different discretization schemes for volatility process in Heston model using moment-matching method. Other than discretization methods, exact simulation methods are also available. Broadie and Kaya[24] first developed an algorithm for exact simulation of CIR process. Inspired by Broadie and Kaya, Baldeaux[9] proposed an exact simulation method for 3/2 process and Grasselli[76] included an exact simulation method for his 4/2 model. Exact simulation method requires a closed-form expression for the characteristic function and does not work well when a full path of the volatility is required. Hence, we will use a discretization method for simulation.

It is widely agreed that traditional Euler-Maruyama scheme leads to negative values for CIR process. In practice, two fixes are considered to solve this problem: *absorption* and *reflection*. In our case, *absorption*-like methods can not be considered as they may drive volatility of underlying asset to infinity due to the 3/2 part. So we choose *reflection* method for variance process and Euler-Maruyama scheme for underlying asset process. The discretization of $v(t)$ and $Y(t)$ for simulations looks as follow:

$$\begin{aligned} v(t + \Delta t)^{'} &= |v(t)| + \alpha(\theta - |v(t)|)\Delta t + \xi\sqrt{|v(t)|}B\Delta t, \\ v(t + \Delta t) &= |v(t + \Delta t)^{'}|. \end{aligned} \tag{3.7}$$

$$\begin{aligned} Y(t + \Delta t) &= Y(t) + (L + (c - \frac{1}{2})(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta Y(t))\Delta t + \left(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right) \\ &\quad \cdot (\rho B\Delta t + \sqrt{1 - \rho^2}B^{\perp}\Delta t). \end{aligned} \tag{3.8}$$

where B and $B^{\perp}$ are independent standard normal variables. As we mentioned at the beginning of this section, exact simulation method generates better results than discretization; however, due to the need of a characteristic function in closed form, it is not always possible for any process to be simulated via an exact simulation method or does an exact simulation method outperform discretization when Monte Carlo simulation is required. In our case, we find a semi-closed form characteristic function for our model. Since there is an expectation term which is not solved analytically, Monte Carlo method is still required for simulation even we

reach an exact simulation scheme for our model. Hence, in our case an exact simulation method does not show absolute advantage over discretization and Monte Carlo. We apply discretization scheme to simulate our model and then use Monte Carlo for pricing purpose in section 4.

The study of convergence rate of the Euler scheme of our model itself can be an interesting research topic, detailed discussion of the convergence rate is beyond the scope of this chapter. However, we briefly touch on the known results of convergence rate for CIR process and Heston model and challenges in finding the convergence rate for our model.

Convergence rate has two categories: strong convergence and weak convergence. Strong convergence refers to how fast an approximation converges to the true solution of the SDE on average; weak converge concerns such convergence in the values some functions. A typical example of weak convergence is studying how fast the option prices from simulation using Euler scheme converges to the prices evaluated from a closed-form solution if one exists. There has been rich literature on the study of discretization schemes and their convergence rate for CIR process, see [99] and [89] for a scan of methods in this area. According to [99] and [89] and references therein, the strong convergence rate for certain CIR discretization schemes is $\frac{1}{2}$ with conditions on the process coefficients. In particular, the discretization we choose is shown to have weak convergence rate 1 and strong convergence $\frac{1}{2}$ in $L^p (p \geq 2)$ sense under a restrictive condition [46] [22]. In the case of Heston model, Lord et.al show that reflection scheme has a weak order of convergence slightly under $\frac{1}{2}$. For our model, both strong and weak convergence are difficult to show due to model complexity and theoretical limitations.

In the end, we would like to comment a little bit on [4] Truncated Gaussian (TG) and Quadratic Exponential (QE) method. These two methods also discretize the variance process, but in a way that is neither *absorption* nor *reflection* style. They are based on the idea of moment-matching. For full details of algorithms for these two methods, please see [4]. Both methods produce non-negative values and account for the density of CIR process, as commented by Andersen, QE is accurate enough in practice. In our simulation, QE is also considered, but this method should be used with caution. Since QE assigns a mass at origin, based on the choice of parameters, there is probability that simulated variance can hit zero, even though all positive values for variance are generated during our simulation process with parameters: $\alpha = 1.8, \theta = 0.04, \xi = 0.2, \Delta t = 0.001$ and the switching rule 1.5.

3.2 Characteristic Function

In this section we present a semi-closed form solution for the conditional characteristic function (c.f.). Define $Z(t) = e^{\beta t} Y(t)$. The dynamics of $Z(t)$ follows Ito's lemma:

$$dZ(t) = \left[Le^{\beta t} + e^{\beta t}(c - \frac{1}{2})(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 \right] dt + e^{\beta t}(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}) dW^{\mathbb{P}}(t). \qquad (3.9)$$

In integral form:

$$Z(T) = Z(t) + \frac{L}{\beta}(e^{\beta T} - e^{\beta t}) + \int_t^T e^{\beta s}(c - \frac{1}{2})(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})^2 ds + \int_t^T e^{\beta s}(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})dW^{\mathbb{P}}(s)$$

(3.10)

The conditional characteristic function of $Z(T)$ given $\mathcal{F}_t$ is given by next proposition.

Proposition 3.2.1 *Let $Z(T)$ follows Equation 3.10 then the c.f. has the following representation:*

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{ iuZ(t) + iu\left(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c - \frac{1}{2})\right)\frac{(e^{\beta T} - e^{\beta t})}{\beta} - u^2(1 - \rho^2)\frac{(e^{2\beta T} - e^{2\beta t})ab}{2\beta} \right\}$$

$$\times exp\left\{ -\frac{iua\rho e^{\beta t}}{\xi}v(t) - \frac{iub\rho e^{\beta t}}{\xi}ln(v(t)) \right\} \times \mathbb{E}\left[v(T)^{-A} exp\left\{ -Bv(T) - C\int_t^T e^{\beta s}ln(v(s))ds \right.\right.$$

$$\left.\left. - \int_t^T (De^{\beta s} + Ee^{2\beta s})\frac{1}{v(s)}ds - \int_t^T (Fe^{\beta s} + Ge^{2\beta s})v(s)ds \right\}\Big|\mathcal{F}_t \right].$$

where,

$$A = -\frac{iub\rho}{\xi}e^{\beta T}, B = -\frac{iua\rho}{\xi}e^{\beta T}, C = \frac{iub\rho\beta}{\xi}, D = -\frac{iub\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right) - iub^2(c - \frac{1}{2}),$$

$$E = \frac{1}{2}u^2(1 - \rho^2)b^2, F = \frac{iua\rho(\beta - \alpha)}{\xi} - iua^2(c - \frac{1}{2}), G = \frac{1}{2}u^2(1 - \rho^2)a^2.$$

(3.11)

See the proof in the Appendix.

The characteristic function involves an expected value with respect to only one Brownian motion ($B(t)$), i.e. the second Brownian $W(t)$ is eliminated. Therefore it constitute a simplification as oppose to the original expectation; however, the conditional expectation term in the expression can not be solved analytically due to its complex structure. In the next section, we propose approximation approaches to the conditional expectation term so that the semi-closed characteristic function can be approximated by an analytical expression.

3.2.1 Approximation to Characteristic Function

In this section, we propose two approaches and apply them to two scenarios in which the semi-closed characteristic function can be approximated by an analytic formula that is available in literature. Then we show the goodness of our approximation approaches by showing the density functions obtained via inversion of the true and approximated characteristic functions.

Two reasons lead to this semi-closed c.f. First, the expectation of integrated CIR process with time dependent factor is not easy to solve analytically [16]. Second reason, integrated log of CIR process with time-dependent factor has not been studied in the literature. Grasselli [72] offers a solution to his geometric 4/2 process, but there is no integrated log of CIR process

involved in the conditional expectation that needs to be solved.

Hence, we propose two approximation approaches to approximate the integrals $\int_t^T (De^{\beta s} + Ee^{2\beta s})\frac{1}{v(s)}ds$ and $\int_t^T (Fe^{\beta s} + Ge^{2\beta s})v(s)ds$ so that we can apply existing analytical expressions to approximate our c.f. We will show our approximations for cases $\rho = 0, b \neq 0$; $b = 0, \rho \neq 0$; and $b = 0, \rho = 0$.

We propose midpoint approximation approach and average approximation approach. The goal is two approximate the time-dependent factors by a constant. We will show numerically the performance of our approximation approaches.

- **Midpoint Approximation:** $g_1(t)$ and $g_2(t)$ are bounded functions on $[t, T]$, so the integrals are bounded by:

$$\min_{[t,T]}(g_1(s)) \int_t^T \frac{1}{v(s)}ds \leq \int_t^T g_1(s)\frac{1}{v(s)}ds \leq \max_{[t,T]}(g_1(s)) \int_t^T \frac{1}{v(s)}ds,$$

$$\min_{[t,T]}(g_2(s)) \int_t^T v(s)ds \leq \int_t^T g_2(s)v(s)ds \leq \max_{[t,T]}(g_2(s)) \int_t^T v(s)ds,$$

then we approximate the integrals by the midpoints:

$$\int_t^T g_1(s)\frac{1}{v(s)}ds \approx m \int_t^T \frac{1}{v(s)}ds,$$

$$\int_t^T g_2(s)v(s)ds \approx n \int_t^T v(s)ds,$$

where,

$$m = \frac{\min_{[t,T]}(g_1(s)) + \max_{[t,T]}(g_1(s))}{2}, n = \frac{\min_{[t,T]}(g_2(s)) + \max_{[t,T]}(g_2(s))}{2}$$

- **Average Approximation:** In last example we claim that $g_1(t)$ and $g_2(t)$ are bounded functions and deduce m and n as midpoints. Average is also a choice for approximation. We approximate our target integrals by averages as:

$$\int_t^T g_1(s)\frac{1}{v(s)}ds \approx \frac{1}{T-t} \int_t^T g_1(s)ds \int_t^T \frac{1}{v(s)}ds = m \int_t^T \frac{1}{v(s)}ds$$

$$\int_t^T g_2(s)v(s)ds \approx \frac{1}{T-t} \int_t^T g_2(s)ds \int_t^T v(s)ds = n \int_t^T v(s)ds$$

where,

$$m = \frac{1}{T-t}\left[\frac{D'}{\beta}(e^{\beta T} - e^{\beta t}) + \frac{E'}{2\beta}(e^{2\beta T} - e^{2\beta t})\right],$$

$$n = \frac{1}{T-t}\left[\frac{F'}{\beta}(e^{\beta T} - e^{\beta t}) + \frac{G'}{2\beta}(e^{2\beta T} - e^{2\beta t})\right].$$

3.2.1.1   No Correlation

The first scenario we will study is no correlation between $W(t)$ and $B(t)$. When $\rho = 0$, Formula 3.11 reduces to a simpler expression:

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{ iuZ(t) + iu\left(L + 2ab(c - \frac{1}{2})\right)\frac{(e^{\beta T} - e^{\beta t})}{\beta} - u^2\frac{(e^{2\beta T} - e^{2\beta t})ab}{2\beta}\right\}$$

$$\times \mathbb{E}\left[exp\left\{ -\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds - \int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds\right\}\Big|\mathcal{F}_t\right].$$

where,

$$D' = -iub^2(c - \frac{1}{2}), E' = \frac{1}{2}u^2b^2, F' = -iua^2(c - \frac{1}{2}), G' = \frac{1}{2}u^2a^2.$$

(3.12)

It is a shown in Grasselli [76] that the conditional expectation of the form $\mathbb{E}\left[exp\left\{ -m\int_t^T \frac{1}{v(s)}ds - n\int_t^T v(s)ds\right\}\Big|\mathcal{F}_t\right]$ has an analytical solution for constants $m > -\frac{(2\alpha\theta - \xi^2)^2}{8\xi^2}$ and $n \geq -\frac{\alpha^2}{2\xi^2}$. Hence, we aim to approximate the integrals $\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds$ and $\int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds$ as follows:

$$\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds \approx m\int_t^T \frac{1}{v(s)}ds,$$

(3.13)

$$\int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds \approx n\int_t^T v(s)ds,$$

(3.14)

such that c.f can be approximated by an analytical function. Let $g_1(s) = D'e^{\beta s} + E'e^{2\beta s}$ and $g_2(s) = F'e^{\beta s} + G'e^{2\beta s}$, we approximate $g_1(s)$ and $g_2(s)$ on interval $[t, T]$ by m and n for some $s_1, s_2 \in [t, T]$ respectively. The expectation term in the characteristic function can be approximated by a function described by following proposition:

Proposition 3.2.2

$$\mathbb{E}\left[exp\left\{ -\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds - \int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds\right\}\Big|\mathcal{F}_t\right]$$

$$\approx \left(\frac{\gamma(T, v(t))}{2}\right)^{k+1} v(t)^{-\frac{\alpha\theta}{\xi^2}} K(T)^{-\left(\frac{1}{2} + \frac{k}{2} + \frac{\alpha\theta}{\xi^2}\right)} e^{\frac{1}{\xi^2}\left(\theta(T-t) - \sqrt{H}v(t)coth\left(\frac{\sqrt{H}(T-t)}{2}\right) + \alpha v(t)\right)}\frac{\Gamma\left(\frac{1}{2} + \frac{k}{2} + \frac{\alpha\theta}{\xi^2}\right)}{\Gamma(k+1)}$$

$$\times {}_1F_1\left(\frac{1}{2} + \frac{k}{2} + \frac{\alpha\theta}{\xi^2}, k+1, \frac{\gamma(T, v(t))^2}{4K(T)}\right),$$

(3.15)

$$k = \frac{1}{\xi^2}\sqrt{(2\alpha\theta - \xi^2)^2 + 8m\xi^2}, H = \alpha^2 + 2n\xi^2, \gamma(T, v(t)) = \frac{2\sqrt{H v(t)}}{\xi^2 sinh\left(\frac{\sqrt{H}(T-t)}{2}\right)},$$

$$K(T) = \frac{1}{\xi^2}\left(\sqrt{H}v(t)coth\left(\frac{\sqrt{H}(T - t)}{2}\right) + \alpha\right).$$

3.2.1.2 Schwartz Model With Heston Stochastic Volatility

In this section we study the scenario where there is no 3/2 diffusion term, but $W(t)$ and $B(t)$ are correlated i.e $b = 0, \rho \neq 0$. We follow the approximation approaches introduced in Section 3.2.1.1. In this case, the characteristic function can be approximated by following function:

Proposition 3.2.3

$$
\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) \approx exp\left\{iuZ(t) + iu\left(L - \frac{a\rho\alpha\theta}{\xi}\right)\frac{(e^{\beta T} - e^{\beta t})}{\beta}\right\}
$$

$$
\times exp\left\{-\frac{iua\rho e^{\beta t}}{\xi}v(t)\right\} \times \left(\frac{(B\xi^2 + \alpha)(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}\right)^{-\frac{2\alpha\theta}{\xi^2}}
$$

$$
\times e^{v(t)\left(\frac{(B\alpha - 2n)(e^{\sqrt{A}(T-t)} - 1) - B\sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}{(B\xi^2 + \alpha)(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}\right)}
$$

where,

$$
B = -\frac{iua\rho}{\xi}e^{\beta T}, F = \frac{iua\rho(\beta - \alpha)}{\xi} - iua^2(c - \frac{1}{2}), G = \frac{1}{2}u^2(1 - \rho^2)a^2, A = \alpha^2 + 2n\xi^2,
$$

$$
\tag{3.16}
$$

Value of n depends on the approximation method being chosen.

3.2.1.3 Schwartz Model with Heston Stochastic Volatility with No Correlation

In this section, we explore a special case in which there is no 3/2 diffusion term or correlation. We will apply average approach to approximate the characteristic function. When $b = 0, \rho = 0$, we have:

$$
\begin{cases}
dY(t) &= [L + (c - \frac{1}{2})a^2 v(t) - \beta Y(t)]dt + a\sqrt{v(t)}dW^{\mathbb{P}}(t), \\
dv(t) &= \alpha(\theta - v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{P}}(t).
\end{cases}
\tag{3.17}
$$

Taking the transformation $Z(t) = Y(t)e^{\beta t}$ gives:

$$
\begin{cases}
dZ(t) &= \left[Le^{\beta t} + e^{\beta t}(c - \frac{1}{2})a^2 v(t)\right]dt + e^{\beta t}a\sqrt{v(t)}dW^{\mathbb{P}}(t) \\
dv(t) &= \alpha(\theta - v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{P}}(t).
\end{cases}
\tag{3.18}
$$

The conditional characteristic function in this case is:

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{iuZ(t) + iuL\frac{(e^{\beta T} - e^{\beta t})}{\beta}\right\} \times \mathbb{E}\left[exp\left\{-\int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds\right\}\Big|\mathcal{F}_t\right]$$

$$\approx exp\left\{iuZ(t) + iuL\frac{(e^{\beta T} - e^{\beta t})}{\beta}\right\} \times \left(\frac{\alpha(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}\right)^{-\frac{2\alpha\theta}{\xi^2}}$$

$$\times e^{v(t)\left(\frac{-2n(e^{\sqrt{A}(T-t)}-1)}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}\right)}$$

where,

$$F' = -iua^2(c - \frac{1}{2}), G' = \frac{1}{2}u^2a^2, A = \alpha^2 + 2n\xi^2, n = \frac{1}{T-t}\left[\frac{F'}{\beta}(e^{\beta T} - e^{\beta t}) + \frac{G'}{2\beta}(e^{2\beta T} - e^{2\beta t})\right]$$

Next exercise, we show that when $\beta \to 0$, $a = 1$ and $c = 0$, the characteristic function of $Z(T)$ actually coincides with the characteristic function of Heston model. The characteristic function of $Z(T)$ when $\beta \to 0$, $a = 1$, $c = 0$ and $\rho = 0$:

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{iuZ(t) + iuL\frac{(e^{\beta T} - e^{\beta t})}{\beta}\right\} \times \mathbb{E}\left[exp\left\{-\int_t^T (\frac{iu}{2}e^{\beta s} + \frac{u^2}{2}e^{2\beta s})v(s)ds\right\}\Big|\mathcal{F}_t\right]$$

As $\beta \to 0$,

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{iuZ(t) + iuL(T - t)\right\} \times \mathbb{E}\left[exp\left\{-\frac{u^2 + iu}{2}\int_t^T v(s)ds\right\}\Big|\mathcal{F}_t\right].$$

The conditional expectation can be solved by Pitman and Yor's [116]:

$$\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t) = exp\left\{iuZ(t) + iuL(T - t)\right\} \times \left(\frac{\alpha(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}\right)^{-\frac{2\alpha\theta}{\xi^2}}$$

$$\times e^{v(t)\left(\frac{-(u^2+iu)(e^{\sqrt{A}(T-t)}-1)}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}\right)},$$

where

$$A = \alpha^2 + \xi^2(u^2 + iu)$$

Recall the characteristic function of Heston's model:

$$\psi_{t,T}(u) = exp\left\{iuY(t) + iur(T - t) + C(T - t, u) + D(T - t, u)v(t)\right\}$$

where,

$$C(T - t, u) = \frac{\alpha\theta}{\xi^2}\left\{(\alpha + \sqrt{\alpha^2 + \xi^2(u^2 + iu)})(T - t) - 2ln\left(\frac{1 - ge^{-(T-t)\sqrt{\alpha^2+\xi^2(u^2+iu)}}}{1 - g}\right)\right\}$$

$$D(T - t, u) = \frac{\alpha + \sqrt{\alpha^2 + \xi^2(u^2 + iu)}}{\xi^2} \times \frac{1 - e^{-(T-t)\sqrt{\alpha^2+\xi^2(u^2+iu)}}}{1 - ge^{-(T-t)\sqrt{\alpha^2+\xi^2(u^2+iu)}}}$$

$$g = \frac{\alpha + \sqrt{\alpha^2 + \alpha\theta(u^2 + iu)}}{\alpha - \sqrt{\alpha^2 + \alpha\theta(u^2 + iu)}}$$

See proof in Appendix that

$$ln\left(\frac{\alpha(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)}{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}\right)^{-\frac{2\alpha\theta}{\xi^2}} = C(T-t, u)$$

$$\frac{-(u^2 + iu)(e^{\sqrt{A}(T-t)} - 1)}{\alpha(e^{\sqrt{A}(T-t)} - 1) + \sqrt{A}(e^{\sqrt{A}(T-t)} + 1)} = D(T-t, u)$$

Back to approximation, when $b = 0$, it is not hard to see that as $\beta \to 0$, $n \to F' + G'$. Under the conditions that $a = 1$ and $c = 0$, we have $n \to u^2 + iu$. Hence, as $\beta \to 0$, average approximation is able to recover the characteristic function of Heston's model when $a = 1$, $\rho = 0$ and $c = 0$. In the next section, we study the goodness of our approximation approaches to the characteristic function numerically by comparing densities via inversion of characteristic functions.

3.2.2 Characteristic Function Inversion Algorithm

- Step 1: Generate a vector $\mathbf{U} = [u_1, u_2, \ldots, u_{m-1}, u_m]$ as well as simulate N paths of $\mathbf{Y}(\mathbf{t}) = [Y(t), Y(t_1), \ldots, Y(t_{n-1}), Y(T)]$ and N paths of $\mathbf{V}(\mathbf{t}) = [v(t), v(t_1), \ldots, v(t_{n-1}), v(T)]$ using the simulation scheme (Equation 3.7 and Equation 3.8) introduced in Section 3.1.2. $\mathbf{Z}(\mathbf{t})$ is obtained by multiplying $\mathbf{Y}(\mathbf{t})$ by $e^{\beta\mathbf{T}}$ with $\mathbf{T} = t, t_1, \ldots, t_n - 1, T$. For approximations, no need to simulate, proceed to Step 3 directly.

- Step 2: For each element in $\mathbf{U}$, compute the integrals $\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds$ and $\int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds$ numerically using each path of $\mathbf{V}(\mathbf{t})$. We end up with two vectors of length N each. Taking average of the sum of these vectors gives estimates of $\mathbb{E}\left[exp\left\{-\int_t^T (D'e^{\beta s} + E'e^{2\beta s})\frac{1}{v(s)}ds - \int_t^T (F'e^{\beta s} + G'e^{2\beta s})v(s)ds\right\}\Big|\mathcal{F}_t\right]$.

- Step 3: Given characteristic function, density function is given by inverse Fourier transform of characteristic function:

$$f(z|\mathcal{F}_t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iuz}\mathbb{E}(e^{iuZ(T)}|\mathcal{F}_t)du.$$

  From Step 1, we have N simulated values of $Z(T)$. To speed up the computation, we draw (uniformly) a smaller sample of $Z(T)$ between $min(Z(T))$ and $max(Z(T))$. Then we integrate numerically over $\mathbf{U}$ for every value in the new sample of $Z(T)$ to obtain the density function.

Following the algorithm steps, we implement it in Matlab. Parameters used to generate the density plots are summarized in Table 3.1. Starting point of $Z(t)$ is $log(0.22)$. 10000 paths of length 251 are generated for $\mathbf{Z}(\mathbf{t})$. $\mathbf{U}$ is a vector consists of 500 points evenly distributed between -40 and 40. Figures 3.1 and 3.2 both show the densities as inverse Fourier transformation of the characteristic function. A histogram is also included as reference. From the figures we can see that as β gets larger, the goodness of approximation of both approaches decrease, but at different speed. When β is close to 0, as expected, the approximated densities coincide with

Table 3.1: Parameters for Density(t=0,T=1).

L	β	c	α	θ	ξ	a	b	ρ
log(0.22)	(0.0001,0.5,2)	(0,0.5)	17	0.0153	0.27	1	(0,0.1)	(0,-0.6)
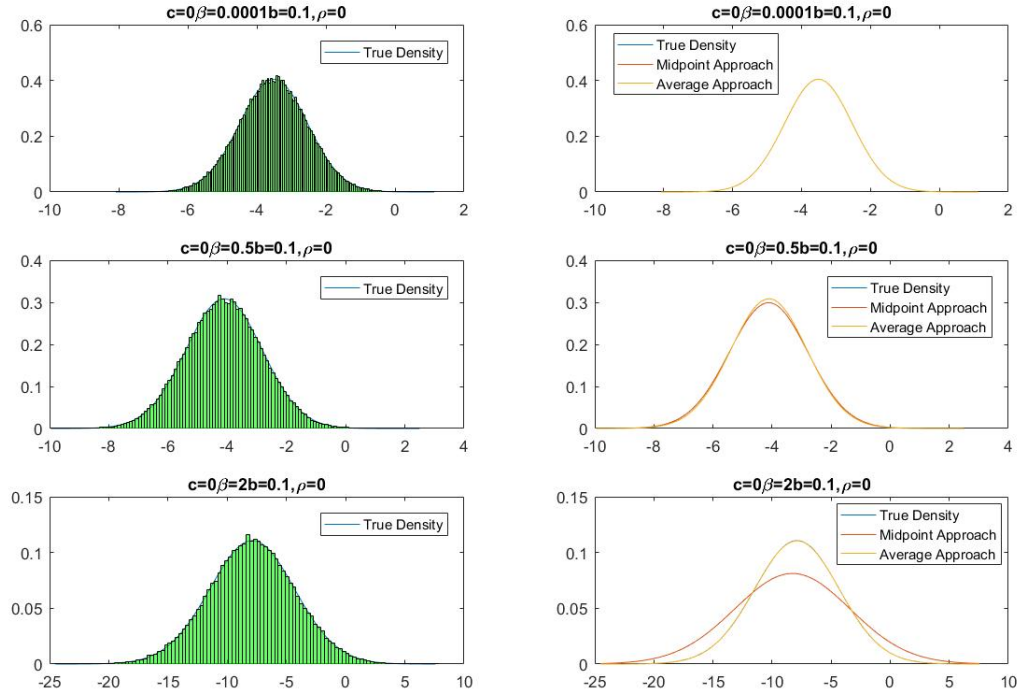
Figure 3.1: Density with Histogram

the true density. Midpoint approach fails to approximate the true density when β is 2. All of the cases, only average approach gives the density that stays close to the true. In particular, the true density greatly overlaps with the average approximation as shown in Figure 3.1. Overall, the difference between true density and average approximation is negligible, and the actual size of error will be assessed later.

3.3 Change of Measure

So far we have worked with the historical probability measure. However, this is not appropriate for option pricing according to the fundamental theorem of asset pricing. Therefore, in this section we study two families of feasible changes of measure for our model, these families come from specifications of the excess return.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{Q})$ denote a probability measure space where $\mathbb{Q}$ is equivalent to $\mathbb{P}$. Let
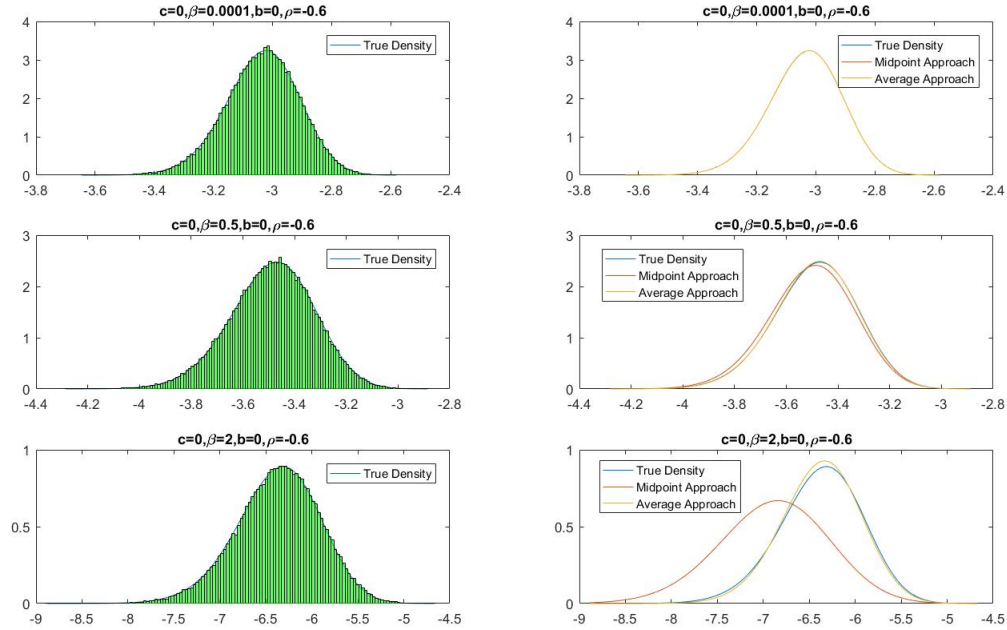
Figure 3.2: Density with Histogram

$W^{\mathbb{Q}}(t)$ be a Wiener process under $\mathbb{Q}$. Since our model consists of two processes (underlying asset and volatility), we will explore change of measures with sound economical interpretation and permitting for the processes to remain within the same family.

For convenience we split $W(t)$ into two Brownian motions: $B(t)$ and its orthogonal part $B(t)^\perp$: $W(t) = \rho B(t) + \sqrt{1 - \rho^2} B(t)^\perp$. Using Girsanov's theorem, we need to find a vector of processes $(\gamma(t), \gamma^\perp(t))$ that satisfy Novikov's condition $\mathbb{E}[exp(\frac{1}{2}(\int_0^T \gamma^2(s)ds + \int_0^T (\gamma^\perp)^2(s)ds))] < \infty$. In such case we have:

$$\begin{cases} dB^{\mathbb{Q}}(t) = dB(t) + \gamma(t)dt, \\ dB^{\mathbb{Q}}(t)^\perp = dB(t)^\perp + \gamma(t)^\perp dt. \end{cases}$$

The change of measure for $Y(t)$ would be $dW^{\mathbb{Q}}(t) = dW(t) + \gamma(t)'dt$, where $\gamma(t)' = \rho\gamma(t) + \sqrt{1 - \rho^2}\gamma(t)^\perp$. The processes $\gamma(t)'$ and $\gamma(t)$ can be seen as the market price of commodity (volatility index) risk and market price of volatility risk for commodity (volatility of volatility index) respectively. Let us now define $g(v(t)) = (a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})\gamma(t)'$ which represents the change in the drift of Y. The next proposition gives sufficient conditions for two feasible changes of measure that are compatible with the financial literature. In particular, Options 1 and 2 prescribes a change of measure where the excess return of $Y(t)$ ($g(v(t))$) is proportional to the variance of the underlying process (see [80]). On the other hand in Option 1 the excess return of $v(t)$ i.e $\gamma(t) = \lambda\sqrt{v(t)}$ is proportional to itself, while in Option 2 it is constant.

Proposition 3.3.1 *Let's assume a, b non-zero, the following two changes of measure are well-defined:*

- **Option 1**

$$
\begin{cases}
\gamma(t) = \lambda \sqrt{v(t)} \\
g(v(t)) = \Lambda(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2
\end{cases}
$$

where (λ, Λ) shall satisfy the following constraints: $I = \left\{\lambda > -\frac{\alpha}{\xi}\right\}$ and $II = \{\Lambda_3 < \Lambda < \Lambda_4\}$

with $\Lambda_3 = max\left[\frac{1}{a}\left(-\frac{\alpha}{\xi} + \rho\lambda\right), -\frac{1}{|b|}\left(\frac{\alpha\theta}{\xi} - \frac{\xi}{2}\right)\right]$ and $\Lambda_4 = min\left[\frac{1}{a}\left(\frac{\alpha}{\xi} + \rho\lambda\right), \frac{1}{|b|}\left(\frac{\alpha\theta}{\xi} - \frac{\xi}{2}\right)\right]$

Here the model under risk-neutral measure becomes:

$$
\begin{cases}
\frac{dX(t)}{X(t)} = [L + (c - \Lambda)(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta ln(X(t))]dt + (a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{Q}}(t) \\
dv(t) = (\alpha\theta - (\alpha + \lambda\xi)v(t))dt + \xi \sqrt{v(t)}dB^{\mathbb{Q}}(t)
\end{cases}
$$

- **Option 2**

$$
\begin{cases}
\gamma(t) = \frac{\lambda'}{\sqrt{v(t)}} \\
g(v(t)) = \Lambda'(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2
\end{cases}
$$

where (λ', Λ') shall satisfy the following constraints: $I = \left\{-\left(\frac{\alpha\theta}{\xi} - \frac{\xi}{2}\right) \le \lambda' \le 0\right\}$ and $II =$

$\{\Lambda_3 < \Lambda < \Lambda_4\}$ with $\Lambda_3 = max\left[-\frac{1}{b}\left(\frac{\alpha\theta}{\xi} - \frac{\xi}{2} - \rho\lambda'\right), -\frac{\alpha}{|a|\xi}\right]$ and $\Lambda_4 = min\left[\frac{1}{b}\left(\frac{\alpha\theta}{\xi} - \frac{\xi}{2} + \rho\lambda'\right), \frac{\alpha}{|a|\xi}\right]$

Here the model under risk-neutral measure is:

$$
\begin{cases}
\frac{dX(t)}{X(t)} = [L + (c - \Lambda')(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta ln(X(t))]dt + (a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{Q}}(t) \\
dv(t) = (\alpha\theta - \lambda'\xi - \alpha v(t))dt + \xi \sqrt{v(t)}dB^{\mathbb{Q}}(t)
\end{cases}
$$

The change of measure we studied is different from [76]. In the aforementioned paper, the stock price process is directly defined under the risk-neutral measure. The author checked whether the discounted stock price is a true martingale. However, our model is defined under historical measure and our change of measure is done by taking into account the market price of risk and market price of volatility risk.

3.4 Estimation

In our model, there are nine parameters to be determined: $L, c, \beta, a, b, \alpha, \theta$ and ξ plus correlation ρ. For estimation we divide them into two groups: drift group (L, c, β, ρ) and volatility group $(a, b, \alpha, \theta, \xi)$ with the later being our main objective. Volatility index (eg. VIX and VSTOXX) and commodities (eg. crude oil) are applications of our model, therefore we use daily observations of VVIX, V-VSTOXX (the volatility of VSTOXX), OVX (volatility of oil ETF, USO) and GVZ (volatility of gold ETF, GLD) to estimate parameters in volatility group respectively. We use volatility indexes data and commodity price data for estimation of drift group parameters. In section 3.4.1, we first estimate b, then propose two methods to estimate the remaining parameters in the volatility group. We present a method to estimate parameters in the drift group in section 3.4.2. In section 3.4.3 we show, via simulation, the consistency of the estimator for volatility group parameters using the methods proposed in 3.4.1. In section 3.4.4, we estimate parameters in volatility group and drift group from empirical data.

3.4.1   Estimation Method For Volatility Group

In our model, the observed volatility data is modeled by square root of Equation (3.3). Let $V(t)$ denote the observed volatility at time t with $a = 1$. By setting $a = 1$ we avoid a potential "identifiability" problem (i.e infinitely many solutions for the pair (a, b)). We can then interpret the volatility as "Heston plus a fraction of 3/2 model". This way, the celebrated Heston model becomes a benchmark model.

We first estimate b. From $V^2(t) = v(t) + \frac{b^2}{v(t)} + 2b$, one observes that the infimum value of $V^2(t)$ occurs when the process $v(t)$ hits b. Using the fact that CIR processes can attain any strictly positive barrier, with probability one (see eg. [3] and [108]), then we conclude $b = \min_{0 \leq t < \infty} \frac{V^2(t)}{4}$, i.e. $P(\min_{0 \leq t < \infty} \frac{V^2(t)}{4} = b) = 1$. Since the data is in discrete time, we propose the following estimator for b: $\hat{\mathbf{b}} = \frac{\min_{t_i \in \Omega_T} V^2(t_i)}{4}$, where $\Omega_T = \{t_i\}_{i=0}^n, t_0 = 0, t_n = T$. The consistency of this estimator is studied numerically in the coming sections.

Next we consider two methods that involve moments to estimate the parameters in the volatility process.

Method 1
The first method starts by solving Equation (3.3) as a quadratic function of $v(t)$, which results in two solutions. We use $V(t_i)$ and $\hat{\mathbf{b}}$ to calculate $\sqrt{\widehat{v^\pm}(t_i)} = \frac{V(t_i) \pm \sqrt{V^2(t_i) - 4\hat{\mathbf{b}}}}{2}$. Then we use MLE on $\widehat{v^+}(t_i)$ and $\widehat{v^-}(t_i)$ respectively to obtain two sets of estimates of α, θ and ξ: $\hat{\Theta}^+ = (\hat{\alpha}^+, \hat{\theta}^+, \hat{\xi}^+)$ and $\hat{\Theta}^- = (\hat{\alpha}^-, \hat{\theta}^-, \hat{\xi}^-)$. A selection criteria to determine which one of $\hat{\Theta}^+$ and $\hat{\Theta}^-$ is a better candidate to estimate the true parameters $\Theta = (\alpha, \theta, \xi)$ is based on the observations and moments of $V^2(t)$.
We now use [38], the authors proved that as $t \to \infty$, $v(t)$ is asymptotically gamma distributed with shape parameter $\frac{2\alpha\theta}{\xi^2}$ and scale parameter $\frac{\xi^2}{2\alpha}$, $\frac{1}{v(t)}$ is hence asymptotic inverse gamma distribution with corresponding paramters $\frac{2\alpha\theta}{\xi^2}$ and $\frac{2\alpha}{\xi^2}$. Since the moments of gamma distribution and inverse gamma distribution are known, our selection criteria proceeds with calculation of the following moments:

$$\mu_1' = \mathbb{E}\left(v(t) + \frac{\hat{\mathbf{b}}^2}{v(t)}\right) = \frac{2\alpha\hat{\mathbf{b}}^2}{2\alpha\theta - \xi^2} + \theta$$

$$\mu_2' = \mathbb{E}\left\{\left(v(t) + \frac{\hat{\mathbf{b}}^2}{v(t)}\right)^2\right\} = \frac{2\alpha\theta(2\alpha\theta + \xi^2)}{4\alpha^2} + \frac{4\alpha^2\hat{\mathbf{b}}^4}{(2\alpha\theta - \xi^2)(2\alpha\theta - 2\xi^2)} + 2\hat{\mathbf{b}}^2$$

$$\mu_3' = \mathbb{E}\left\{\left(v(t) + \frac{\hat{\mathbf{b}}^2}{v(t)}\right)^3\right\}$$

$$= \frac{2\alpha\theta(2\alpha\theta + \xi^2)(2\alpha\theta + 2\xi^2)}{8\alpha^3} + \frac{8\alpha^3\hat{\mathbf{b}}^6}{(2\alpha\theta - \xi^2)(2\alpha\theta - 2\xi^2)(2\alpha\theta - 3\xi^2)} + 3\hat{\mathbf{b}}^2\theta + \frac{6\alpha\hat{\mathbf{b}}^4}{2\alpha\theta - \xi^2},$$

where we have used the moments up to third degree from gamma distribution and inverse gamma distribution because we have three parameters to estimate. In order for the third mo-

ment of inverse gamma distribution to exist, we need a condition:

$$2\alpha\theta \geq 3\xi^2, \tag{3.19}$$

which we denote as Condition II. This condition is stronger than Condition I. Then we substitute $\hat{\mathbf{\Theta}}^+$ and $\hat{\mathbf{\Theta}}^-$ for $\mathbf{\Theta}$ in μ_1', μ_2' and μ_3' respectively, a candidate is determined by the set of estimates that minimizes the objective function

$$F = \left(\mu_1' - \frac{1}{n+1}\sum_{i=0}^{n}(V^2(t_i) - 2\hat{\mathbf{b}})\right)^2 + \left(\mu_2' - \frac{1}{n+1}\sum_{i=0}^{n}(V^2(t_i) - 2\hat{\mathbf{b}})^2\right)^2$$
$$+ \left(\mu_3' - \frac{1}{n+1}\sum_{i=0}^{n}(V^2(t_i) - 2\hat{\mathbf{b}})^3\right)^2.$$

Remark: Because we can obtain $\hat{\mathbf{\Theta}}^+$ and $\hat{\mathbf{\Theta}}^-$ from data directly, they may not satisfy Condition II. In this case, one can use as many moments as feasible. For example, if the estimates only satisfy Condition I, then the selection criteria only concerns μ_1'.

Method 2

This method can be viewed as an improvement over *Method 1*. Unlike *Method 1*, we use methods of moments by taking into account all three moments and imposing Condition II. We estimate the parameters by minimizing F to find $\hat{\alpha}, \hat{\theta}$ and $\hat{\xi}$ over the region $2\alpha\theta \geq 3\xi^2$. The optimization step is done numerically with initial values $\hat{\mathbf{\Theta}}^+$ and $\hat{\mathbf{\Theta}}^-$ leading to two sets of estimates. The set with smallest F value is our estimate for $\mathbf{\Theta}$.

The key of this method is that the series $v(t) + \frac{b^2}{v(t)}$ is weakly stationary and ergodic under Condition II. [78] has shown stationarity and ergodicity are requirements for method of moments to ensure consistency. It is known that CIR process is stationary and ergodic as shown in [38]. [2] examined conditions under which 3/2 process is stationary and ergodic, which requires mean-reverting speed to be positive. Note in our case, the mean-reverting speed of $\frac{1}{v(t)}$ is $\alpha\theta - \xi^2$. Hence, a condition for $\frac{1}{v(t)}$ to be stationary and ergodic is:

$$\alpha\theta \geq \xi^2 \tag{3.20}$$

denoted as Condition III. Condition III is stronger than Condition I but weaker than Condition II. By Proposition 2.1.7 and Corollary 2.1.8 of [120], our process $v(t) + \frac{b^2}{v(t)}$ satisfies weak stationary and ergodic hence we can ensure consistency via method of moments.

Remark: In both methods above we estimate b via $\hat{\mathbf{b}}$. We could have alternatively tried estimating b via method of moments at the expense of much stronger condition $2\alpha\theta \geq 4\xi^2$ (Condition IV).

3.4.2   Estimation Method For Drift Group

The method above provides us estimated volatility group parameters. Next, we use least squares method to estimate parameters in drift group.

We first start with discretization scheme of $Y(t)$ described in Equation (3.8) yielding:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i, i = 0, \ldots, n$$

where,

$$y_i = \frac{Y(t_{i+1}) - Y(t_i)}{\sqrt{v(t_i)} + \frac{b}{\sqrt{v(t_i)}}}, x_{1i} = \frac{1}{\sqrt{v(t_i)} + \frac{b}{\sqrt{v(t_i)}}}, x_{2i} = \sqrt{v(t_i)} + \frac{b}{\sqrt{v(t_i)}}, x_{3i} = \frac{Y(t_i)}{\sqrt{v(t_i)} + \frac{b}{\sqrt{v(t_i)}}},$$

$$\beta_1 = L\Delta t, \beta_2 = (c - \frac{1}{2})\Delta t, \beta_3 = -\beta\Delta t.$$

$\epsilon_i = W(t_{i+1}) - W(t_i)$ is the error term with $N(0, \Delta t)$ distribution. β_1, β_2 and β_3 can be estimated from data using least square method. Hence, the estimated drift group parameters follow as: $\hat{L} = \frac{\hat{\beta}_1}{\Delta t}, \hat{c} = \frac{\hat{\beta}_2}{\Delta t} + 0.5$ and $\hat{\beta} = -\frac{\hat{\beta}_3}{\Delta t}$.
To estimate ρ, we obtain a similar regression form for $v(t)$ process

$$v_i = \beta_4 x_{4i} + \beta_5 x_{5i} + \epsilon'_i, i = 0, \ldots, n$$

where,

$$v_i = \frac{v(t_{i+1}) - v(t_i)}{\xi \sqrt{v(t_i)}}, x_{4i} = \frac{1}{\sqrt{v(t_i)}}, x_{5i} = \sqrt{v(t_i)}, \beta_4 = \frac{\alpha\theta}{\xi}, \beta_5 = -\frac{\alpha}{\xi}.$$

$\epsilon'_i = B(t_{i+1}) - B(t_i)$ is the error term. ρ is estimated as the correlation between two residual terms: $\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ and $\mathbf{v} - \tilde{\mathbf{X}}\tilde{\boldsymbol{\beta}}$ where $\mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$, $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)'$, $\tilde{\mathbf{X}} = \{\mathbf{x_4}, \mathbf{x_5}\}$ and $\tilde{\boldsymbol{\beta}} = (\hat{\beta}_4, \hat{\beta}_5)'$.

3.4.3 Simulation Results

In this section, we first show that b can be recovered as sample size increases. To show this, we simulate time series of $\{V(t_i)\}_{i=0}^{i=n}$ with increment $\Delta t = \frac{1}{250}$ for different α, θ, ξ, b and n and calculate $\hat{\mathbf{b}}$. Then we report the absolute error of $|\hat{\mathbf{b}} - b|$ versus n in Table 3.2. From Table 3.2 we can see that the absolute error decreases as n gets larger. This is a numerical indication that $\hat{\mathbf{b}}$ is a consistent estimator for b. Second, we run 1000 simulations for each of sample sizes

Table 3.2: $|\hat{\mathbf{b}} - b|$ vs. n for different α, θ, ξ, b

(α, θ, ξ, b)	n=1000	n=10000	n=100000	n=1000000
(18, 0.02, 0.3, 0.1)	0.015	0.0064	0.0028	0.0013
(15, 0.5, 1.5, 0.05)	0.0106	0.007	0.0043	0.0035
(1, 0.15, 0.3, 0.005)	0.0055	0.0015	7.2e-11	7.4e-9

1000, 10000, 100000 and 1000000 with true parameters $\alpha = 18, \theta = 0.02, \xi = 0.3, b = 0.1$ and report the mean as well as standard errors (s.e) of the estimated parameters from *Method 1* and *Method 2* in Table 3.3 and Table 3.4, respectively. We can see from both tables that, as sample size gets larger, the estimated parameters get closer to the true value with smaller standard deviation. As expected *Method 2* has smaller bias for small sample sizes and converges faster to the true parameters. The results confirm that under Condition II our method provide good estimation of α, θ and ξ.

Table 3.3: Simulation Results for $\alpha = 18, \theta = 0.02, \xi = 0.3, b = 0.1$ With Method 1

Sample Size n	mean of $\hat{\mathbf{b}}$	s.e of $\hat{\mathbf{b}}$	mean of $\hat{\alpha}$	s.e of $\hat{\alpha}$	mean of $\hat{\theta}$	s.e of $\hat{\theta}$	mean of $\hat{\xi}$	s.e of $\hat{\xi}$
1000	0.1165	1.52e-4	22.4114	0.1146	0.3525	0.007608	1.408	0.02307
10000	0.1091	2.5e-5	18.7259	0.01313	0.0471	0.001046	0.4262	0.003072
100000	0.1051	4.74e-6	18.447	0.001166	0.0227	2.62e-6	0.3286	2.88e-5
1000000	0.1028	9.18e-7	18.081	1.03e-4	0.0215	5.02e-7	0.317	5.8e-6

3.4.4 Estimation With Empirical Data

In this section we will apply our estimation method to data. First, we estimate volatility group parameters. We use four datasets: VVIX, V-VSTOXX , OVX and GVZ. The estimated parameters with standard errors (s.e) obtained from parametric bootstrap are summarized in Table 3.5. Next we estimate drift group parameters following the method described in section 3.4.2 as well as correlation parameter ρ. Here we use volatility index data (eg. VVIX) as well as asset (index) data (VIX). Sample size of asset price data is tailored to meet the sample size of volatility index data as asset price data are larger. Results, estimated parameters and p-values are summarized in Table 3.7.

We can see from Table 3.5 that the results for VVIX and V-VSTOXX are comparable. Because $\hat{\mathbf{b}}$ is small for OVX and GVZ data, we could possibly ignore the 3/2 process component of $V(t)$ in which case our model becomes Schwartz one-factor model with stochastic volatility (Schwartz Heston model), a popular model used in modeling commodity prices. This shows that our estimation method is in line with empirical experience and our model is flexible that it can capture different financial instruments with mean-reverting property. To compare, we also include MLE estimators for OVX and GVZ data in Table 3.6. It can be seen that, the estimators using our estimation approach with small b are close to MLE estimators.

From Table 3.7 we can determine the final models for all three datasets based on p-values. Note for L, β and ρ, the hypothesis against 0 is tested; for c the hypothesis against 0.5 is tested. So the final model for USO only has volatility component; models for VIX, VVIX and GLD have non-zero drift component.

Table 3.4: Simulation Results for $\alpha = 18, \theta = 0.02, \xi = 0.3, b = 0.1$ With Method 2.

Sample Size n	mean of $\hat{\mathbf{b}}$	s.e of $\hat{\mathbf{b}}$	mean of $\hat{\alpha}$	s.e of $\hat{\alpha}$	mean of $\hat{\theta}$	s.e of $\hat{\theta}$	mean of $\hat{\xi}$	s.e of $\hat{\xi}$
1000	0.1165	1.52e-4	22.1666	0.1164	0.2433	0.007773	1.1591	0.02709
10000	0.1091	2.5e-5	18.9948	0.011237	0.0266	2.85e-4	0.3584	9.78e-4
100000	0.1051	4.74e-6	18.175	9.94e-4	0.0228	2.88e-6	0.3326	3.35e-5
1000000	0.1028	9.18e-7	18.2613	1.33e-4	0.0214	4.91e-7	0.3157	5.2e-6

Table 3.5: Estimated Volatility Group Parameters With Empirical Data

	Data: VVIX	Data: V-VSTOXX	Data: OVX	Data: GVZ
$\hat{\mathbf{b}}$	0.0892	0.067	0.0053	0.002
Mean of $\hat{\mathbf{b}}$ (s.e)	0.104 (7.54e-5)	0.0848 (7.99e-5)	0.0061 (1.91e-5)	0.0026 (7.81e-6)
$\hat{\alpha}$	17.00512	19.48345	3.6754	4.7937
Mean of $\hat{\alpha}$ (s.e)	18.691 (0.06836)	21.309 (0.05850)	4.09 (0.01806)	5.0976 (0.01977)
$\hat{\theta}$	0.015293	0.00961	0.1472	0.0395
Mean of $\hat{\theta}$ (s.e)	0.0273 (9.51e-4)	0.0172 (4.28e-5)	0.1444 (3.54e-4)	0.0382 (7.62e-5)
$\hat{\xi}$	0.27014	0.21462	0.6006	0.2873
Mean of $\hat{\xi}$ (s.e)	0.36 (0.001407)	0.315 (7.15e-4)	0.5946 (9.45e-4)	0.2876 (4.3e-4)

Table 3.6: MLE Estimates For OVX and GVZ Data

	Data: OVX	Data: GVZ
$\hat{\alpha}$	3.9042	5.028
Mean of $\hat{\alpha}$ (s.e)	4.2791 (0.01785)	5.4036 (0.01878)
$\hat{\theta}$	0.1530	0.0426
Mean of $\hat{\theta}$ (s.e)	0.1522 (4.02e-4)	0.0423 (9.33e-5)
$\hat{\xi}$	0.6849	0.4149
Mean of $\hat{\xi}$ (s.e)	0.6848 (1.8e-4)	0.4149 (1.07e-4)

Table 3.7: Estimated Drift Group Parameters

Data	$\hat{L}$	p-value	$\hat{c}$	p-value	$\hat{\beta}$	p-vale	$\hat{\rho}$	p-value
VIX&VVIX	-1.6411	0.44	-5.320	0.000145	3.7020	3.44e-5	-0.67	0
VSTOXX&VVSTOXX	-7.060	0.044252	-1.580	0.360	5.4841	0.000483	-0.49	0
USO&OVX	0.0870	0.678	-0.579	0.313	-0.0376	0.793	-0.35	0
GLD&GVZ	3.7672	0.015	0.572	0.97	0.78	0.014	-0.08	0

3.5   Pricing Financial Derivatives

In this section, we price vanilla European call options based on our model for different assets: VIX, USO Oil ETF and Gold ETF (GLD). We use the parameters estimated in section 3.4.4 and assume risk-free interest rate of 2%. Our purpose is to assess the effect of b (as a was set to 1) on the price. We investigate whether the presence of b makes a difference by looking at the absolute value of the relative change (ARC) in implied volatility (IV):$ARC = |\frac{IV_{b=0}-IV_{b=\hat{b}}}{IV_{b=\hat{b}}}|$ for $b = 0$ against $b = \hat{\mathbf{b}}$ (Figure 3.3 and Figure 3.4). We also report the maximum absolute value of the relative change in terms of option prices for completeness.

In section 3.5.1, we display how b affects the call option prices for VIX. In section 3.5.2, we study the effect on call option prices for USO. In section 3.5.3, we compare two popular option pricing methods for GLD: Fast Fourier Transform (FFT) developed by Carr and Madan [29] and Monte Carlo simulation in a special setting (Schwartz Heston model) where

the characteristic function can be approximated by an analytical solution. We list in section 3.3 the options for change of measure. For illustration purpose we only consider Option 1 in this section following [80].

3.5.1 Price VIX Call Options

Our model for VIX under risk-neutral measure is defined as:

$$\begin{cases} \frac{dX(t)}{X(t)} = [(c - \Lambda)(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2 - \beta ln(X(t))]dt + (\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})dW^{\mathbb{Q}}(t), \\ dv(t) = (\alpha\theta - (\alpha + \lambda\xi)v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{Q}}(t), \\ dW^{\mathbb{Q}}(t)dB^{\mathbb{Q}}(t) = \rho dt. \end{cases}$$

With the parameters obtained from section 3.4.1, we determine that $I = \{\lambda > -63\}$ and $II = \{-9.3 < \Lambda < \Lambda_4\}$ with $\Lambda_4 = min(63 - 0.67\lambda, 9.3)$. [37] find that market price of volatility risk is negative. [51] assume the price to be -0.1 in their simulation. Λ, in this case, represents the market price of volatility risk. Therefore, we price VIX options under two assumptions of Λ: $\Lambda = -0.1, 0$. We then pick $\lambda = -0.5, 0, 0.5$ so that $-0.1 \in II$ and $0 \in II$. Strike prices are 0.2, 0.21, 0.22, 0.23, 0.24, 0.25 with expiry dates 0.2, 0.4, 0.6, 0.8 and 1 year assuming current VIX at 0.22.

Implied volatilities are obtained by matching Black-Scholes option price formula with call option prices then solving for the volatility parameter. In our case, the dynamics of $Y(t)$ is an O-U process if we assume constant volatility:

$$dY(t) = (L + (c - 0.5 - \Lambda)\sigma^2 - \beta Y(t))dt + \sigma dW^{\mathbb{Q}}(t).$$

So we need to match call prices for VIX with a formula based on exponential O-U process. [44] derived call option price formula for $X(t)$ with strike price K and maturity T:

$$C(X(0), K) = e^{-rT}\left[X(0)^{\phi_T}exp\left\{\frac{\theta(\sigma)}{\beta}(1 - \phi_T) + \frac{1}{2}a_T^2\right\}N(d + a_T) - KN(d)\right],$$

where,

$$\phi_T = e^{-\beta T}, \theta(\sigma) = L + (c - 0.5 - \Lambda)\sigma^2, a_T = \frac{\sigma}{\sqrt{2\beta}}(1 - \phi_T^2)^{\frac{1}{2}},$$

$$d = \frac{1}{a_T}(\phi_T ln(X(0)) - ln(K) + \frac{\theta(\sigma)}{\beta}(1 - \phi_T)).$$

and r is risk-free interest rate. We can extract implied volatilities from given call option prices using this formula. When $b = 0$, we re-estimate volatility group parameters using MLE: $\hat{\alpha} = 19.345, \hat{\theta} = 0.8, \hat{\xi} = 1.617$ and drift group parameters remain the same except correlation $\hat{\rho} = 0.76$, then we calculate the implied volatilities. Figure 3.3 shows the surface of ARC. It is observed from Figure 3.3 that $\lambda = -0.5$ results in larger ARC, in particular there could be a difference of up to 4.5% in implied volatility and up to 20% in terms of call price due to a wrong choice of model. Another observation is λ has greater impact than Λ.

3.5.2 Price USO Call Options

From Table 3.7, the constrains on λ and Λ are: $I = \{\lambda > -3.21\}$ and $II = \{\Lambda_3 < \Lambda < 3.21 - 0.35\lambda\}$ with $\Lambda_3 = max(-3.21 - 0.35\lambda, -60.13)$. Kolos and Ronn [91] showed that market price of risk is positive for crude oil. In this case, Λ is the market price of risk and we assume $\Lambda = 0.05, 0.06$. Doran and Ronn [48] found that market price of volatility risk is negative for crude oil. Hence, we assume $\lambda - 0.1, -0.15, -0.2$. Note the values assumed for λ and Λ satisfy the constrains. Strike prices range from \$12 to \$15. Expiry date setup is same as above. We assume current value of USO is \$13.

Implied volatilities are calculated based on following call option price formula:

$$C(X(0), K) = X(0)e^{-(r+\Lambda\sigma^2-0.5\sigma^2)T}N(d_1) - Ke^{-rT}N(d_2),$$

where, $d_1 = \frac{ln\left(\frac{X(0)}{K}\right)+(1-\Lambda)\sigma^2 T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$. Next, we calculate implied volatilities for $b = 0.0053$ and $b = 0$, for the latter we re-estimate volatility group parameters using MLE: $\hat{\alpha} = 3.904$, $\hat{\theta} = 0.153$, $\hat{\xi} = 0.685$. Figure 3.4 shows the surface of ARC. The shape of the surface in Figure 3.4 do not vary as much as they do in Figure 3.3 for different λ values. Figure 3.4 shows a difference, of up to 5% in implied volatility and up to 10% in terms of call price due to a wrong choice of model (i.e. working with $b = 0$ instead of the true b).

3.5.3 Price GLD Options with Schwartz Heston Model

Since GVX data yields the smallest b that is negligible, we assume Schwartz Heston model for GLD&GVX data and price GLD options accordingly. We first perform pricing via Monte Carlo simulation. Then we apply the result given by Proposition 3.2.3 to perform FFT pricing method and compare the results with simulation. We use parameters from Table 3.6 and Table 3.7; we also assume $\Lambda = 0.25$ and $\lambda = 0.9$.

We price GLD options based on following model:

$$\begin{cases} \frac{dX(t)}{X(t)} = [L + (c - \Lambda)v(t) - \beta ln(X(t))]dt + \sqrt{v(t)}dW^{\mathbb{Q}}(t), \\ dv(t) = (\alpha\theta - (\alpha + \lambda\xi)v(t))dt + \xi\sqrt{v(t)}dB^{\mathbb{Q}}(t), \\ dW^{\mathbb{Q}}(t)dB^{\mathbb{Q}}(t) = \rho dt. \end{cases}$$

So far we have done option pricing via Monte Carlo simulation. Pricing via simulation is a fundamental approach that can be applied for all models; however, simulation scheme has to be carefully chosen so that the discretized process converges to the continuous-time process. In practice, option pricing using Fast Fourier Transform (FFT) is considered faster and more robust than simulation provided an analytical characteristic function is available. In this section, we compare two popular option pricing methods: FFT and Monte Carlo simulation under risk-neutral measure $\mathbb{Q}$.

A widely used FFT pricing formula was first introduced by Carr and Madan [29]. The formula expresses the call option price as a function that involves the characteristic function of log price $Y(t)$. Carr and Madan's method does not separate payoff from density, so it requires
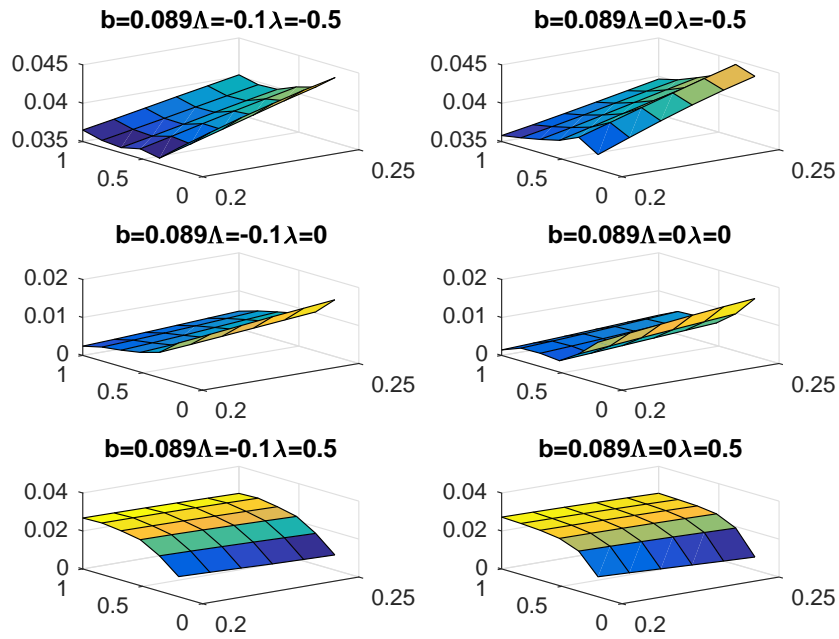
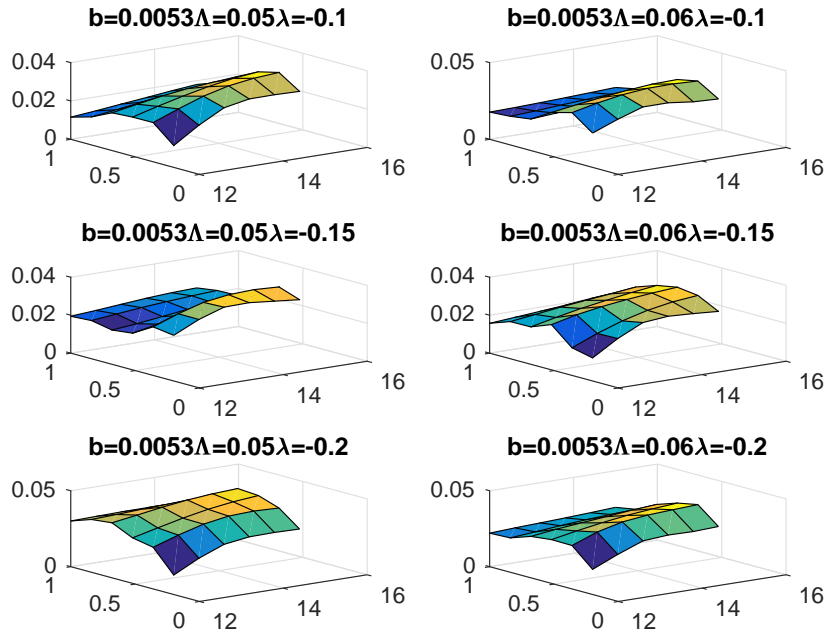Figure 3.3: Absolute Relative Change in Implied Volatility

Figure 3.4: Absolute Relative Change in Implied Volatility

an exponential damping factor to ensure the integrability of option payoff. Lewis [96] further proposed a generalized Fourier-transform approach to European options pricing. Both methods require the conditional characteristic function of the underlying process to be known explicitly or at least efficiently computed. Instead of introducing the damping factor, Lewis's method works with the region where the conditional characteristic function is defined. In this section, we follow Carr and Madan's framework but modify the formula a little bit since our characteristic function and its approximations are based on $Z(t)$ process instead of $Y(t)$.

Suppose a given option with maturity T and strike K at given time $t, t \geq 0$, whose payoff, P, is a function of $V(T)$, i.e $P = F(V(T))$ (for a European call option $P = F(V(T)) = (V(T) - K)^+$). Recall that $Z(t) = e^{\beta t} Y(t)$ and $Y(t) = ln(X(t))$, so $X(t) = exp(e^{-\beta t} Z(t))$, or we express $X(t)$ as a function of $Z(t)$, i.e $X(t) = g(Z(t))$. Then we have the relationship between the payoff of the VIX option and $Z(T)$:

$$P = F(g(Z(T))) = G(Z(T)).$$

The price, $C(t, T)$, of the option is the conditional expected value of P under risk-neutral measure:

$$C(t, T) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}(P|\mathcal{F}_t)$$
$$= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}(G(Z(T))|\mathcal{F}_t),$$

where r is the risk-neutral interest rate.

Let $f_T(z)$ be the risk-neutral density function of $Z(T)$ and the price at $t = 0$ be $C_T(k) = C(0, T)$, a function of strike price k with $k = g^{-1}(K) = e^{\beta T} ln(K)$. By definition the characteristic function of $Z(T)$ is the Fourier transform of $f_T(z)$:

$$\psi(u) = \int_{-\infty}^{\infty} e^{iuz} f_T(z) dz,$$

As we work on European call option on of VIX and $Z(T)$, hence the payoff function is:

$$P = G(Z(T)) = (exp(e^{-\beta T} Z(T)) - exp(e^{-\beta T} k))^+.$$

So we have the option price expressed in terms of z:

$$C_T(k) = e^{-rT} \int_k^{\infty} (exp(e^{-\beta T} z) - exp(e^{-\beta T} k)) f_T(z) dz. \tag{3.21}$$

An issue with 3.21 is that the integral is not square integrable as $k \to -\infty$. To solve the problem, Carr and Madan [29] modified the price by multiplying a damping factor e^{ck} to the price. Let $c_T(k)$ denote the modified price, so:

$$c_T(k) = e^{ck} C_T(k).$$

Now we apply Fourier transform to $c_T(k)$:

$$\psi'(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk,$$

for $v \in \mathbb{R}$. Then $C_T(k)$ is the inverse Fourier transform of $c_T(k)$:

$$C_T(k) = \frac{e^{-ck}}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \psi'(v) dv. \tag{3.22}$$

The price being real implies that $\psi'(v)$ is odd in its imaginary part and even in its real part [29]. We can also simplify $C_T(k)$ to be:

$$C_T(k) = \frac{e^{-ck}}{\pi} \int_{0}^{\infty} e^{-ivk} \psi'(v) dv. \tag{3.23}$$

If we can find an expression for $\psi'(v)$, we can plug it into 3.23 to find the price numerically using Fast Fourier Transform (FFT) algorithm. Therefore, we don't need to worry about the density function $f_T(z)$ in 3.21 as long as we have found the characteristic function for the underlying process.

From above we have:

$$\psi'(v) = e^{-rT} \int_{-\infty}^{\infty} e^{ivk} e^{ck} \int_{k}^{\infty} (exp(e^{-\beta T}z) - exp(e^{-\beta T}k)) f_T(z) dz dk$$

By Fubini's Theorem:

$$\psi'(v) = e^{-rT} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} (exp(e^{-\beta T}z) - exp(e^{-\beta T}k)) e^{(iv+c)k} dk \right) f_T(z) dz$$

$$= e^{-rT} \int_{-\infty}^{\infty} \left(\frac{e^{(e^{-\beta T}+iv+c)z}}{iv+c} - \frac{e^{(e^{-\beta T}+iv+c)z}}{e^{-\beta T}+iv+c} \right) f_T(z) dz$$

$$= e^{-(r+\beta)T} \int_{-\infty}^{\infty} \frac{e^{(e^{-\beta T}+iv+c)z}}{c^2 + e^{-\beta T}c - v^2 + (2c + e^{-\beta T})iv} f_T(z) dz$$

$$= \frac{e^{-(r+\beta)T} \psi(v - i(e^{-\beta T} + c))}{c^2 + e^{-\beta T}c - v^2 + (2c + e^{-\beta T})iv}$$

$\psi'(v)$ is now a function which involves the characteristic function of $Z(T)$. So the original price $C_T(k)$ can be written as:

$$C_T(k) = \frac{e^{-ck}}{\pi} \int_{0}^{\infty} e^{-ivk} \frac{e^{-(r+\beta)T} \psi(v - i(e^{-\beta T} + c))}{c^2 + e^{-\beta T}c - v^2 + (2c + e^{-\beta T})iv} dv. \tag{3.24}$$

Now we apply FFT to compute the price numerically. Assuming current GLD price is \$120, we set the strike prices \$118, \$119.4, \$120.8, \$122.2, \$123.6, \$125 with quarterly maturity throughout a year i.e $t = 0$ and $T \in [0.25, 0.5, 0.75, 1]$. We calculate the absolute relative price difference with respect to the prices from simulation and summarize our findings in Table 3.8. From the table, we can see that option prices using FFT are less that 1% different from prices from simulation in absolute value. The results, on the other hand, suggest that our approximation approach is able to capture the behavior of the characteristic function of Schwartz Heston model. Hence, we can use FFT method to price options, which is faster than simulation.

Table 3.8: Absolute Relative Price Difference: FFT vs. Simulation

Strike	T=0.25	T=0.5	T=0.75	T=1
118	0.0027	0.0015	0.0035	0.005
119.4	0.0032	0.0021	0.0036	0.0056
120.8	0.0036	0.0026	0.0038	0.0063
122.2	0.0042	0.0032	0.004	0.0071
123.6	0.005	0.0038	0.0042	0.0079
125	0.0059	0.0046	0.0045	0.0087

3.6 Conclusion

In this chapter we define and show properties of a new model displaying mean-reversion and stochastic volatility. We find a semi-analytic solution for the characteristic function. Even though simulation is still required to find this characteristic function, the cost of computation is greatly reduced from simulating two stochastic processes to one process (the underlying CIR process that drives the stochastic volatility). The benefit of the semi-closed characteristic function becomes more significant in higher dimensions. We further propose two approximation approaches to improve the computational efficiency of the characteristic function. The numerical results are promising showing our approximation approaches perform well in a wide parametric region, especially the average approach yields negligible approximation errors. We are also the first proposing a method to estimate the parameters of the 4/2 family of stochastic volatility models, the methods are simple and requires low computational power. From the tables in section 3.4.4, it can be seen that our estimation method, though biased for low sample, still provides consistent estimates for the true parameters. By adding a 3/2 component in the volatility, we generate more flexibility in our model than [16]. For volatility index markets (VIX and VSTOXX) and for the commodity USO market, the effect of this 3/2 addition is confirmed empirically and substantial in terms of impact on option pricing. Since our model incorporates two models ([122] and [76]), there could exist various applications to different markets (eg. interest rate and exchange rate). Table 3.8 again supports the appropriateness of our average approximation approach as the option prices from FFT method using approximated characteristic function is close to the option prices from Monte Carlo simulation, the absolute price difference is at 10^{-3} level. One exercise that is not demonstrated in this chapter is revisiting estimation with approximated characteristic function and methods that depend on closed-form characteristic functions. With semi-closed form characteristic functions, we cannot use estimation methods such as efficient method of moments (EMM) [70], generalized method of moments (GMM) [78] and quasi-maximum likelihood (QML) [110] [79], which rely on analytic density/transition density functions. We will mention about this again in chapter 5. The approximation approaches give us analytic functions, thus the EMM, GMM or QML can be used and their performance can be evaluated against our proposed estimation method. This is an interesting exercise for future research.

Chapter 4

Generalized Mean-Reverting 4/2 Factor Model

Continuous-time stochastic covariance models are crucial in capturing many stylized facts in financial data, from heteroscedasticity and fat tails to changing correlations and leverage effects. Early work in this field focused on discrete time models in the form of generalized autoregressive conditional heteroskedasticity (GARCH) models (see [56]). The best-known representatives in continuous time, are the stochastic Wishart family (see [75, 40]) and the Ornstein–Uhlenbeck (OU) family (see [109]) of models, as well as general linear-quadratic jump-diffusions (see [31]). These approaches are more realistic than the classical Black–Scholes lognormal model, but they quickly become intractable as dimensions increase in terms of the number of parameters and simulation paths, commonly known as the "curse of dimensionality". Recent papers (see [57, 41]) have presented models built from linear combination of tractable one-dimensional counterparts. These models involve fewer parameters than Wishart- or OU-type approaches, owing to a reduction in dimensionality while providing a closed-form solution to financial problems.

In this chapter, we introduce a multivariate mean-reverting stochastic volatility factor model that combines 1/2 (Heston-type, [80]) and 3/2 processes [117] for the modeling of volatility. Such underlying volatility processes were coined 4/2 by [76]. In section 4.1, we give the definitions of our model. Our model takes advantage of the factor structure in asset prices and allows for a mean-reverting structure on the assets thereby aiming at capturing either multivariate commodity behavior or multiple volatility indexes (see [73] for an alternative multivariate non-mean-reverting generalization based on a pairwise-structure applied to the exchange-rate market). In particular, our setting reduces the dimension of the parametric space which is a way of controlling the "curse of dimensionality" making parameters identifiable and popular estimation methods feasible. Secondly, the presence of independent common and intrinsic factors, each with its own stochastic volatility, enables an elegant separable structure for characteristic functions (c.f.s) and captures several stylized facts, such as: stochastic volatility, stochastic correlation among stocks (see [56]), co-movements in the variances (see [45]), multiple factors in the volatilities (see [36]) and stock correlations (see [40]). Thirdly, a factor representation is compatible with economical interpretations, where common factors are exogenous variables explaining financial markets, and intrinsic factors relate to companies' intrinsic risks. Lastly,

closed-form expressions are available for joint c.f.s; this is useful for derivative pricing and risk management calculations via Fourier transformations, and it makes c.f.-based estimations methods feasible (see [29] [69] [27]).

The rationale for a 4/2 volatility process rather than a 1/2 or 3/2 model is masterly presented in [76] for a one-dimensional structure. For instance, as observed by the author, the 1/2 process predicts that the implied volatility skew will flatten when the instantaneous volatility increases (crises), while the 3/2 model predicts steepening skews. The empirical violation of the Feller condition in the 1/2 model is also noted, which makes volatility paths stay closer to 0 for a longer period than empirically supported, while the 3/2 model admits extreme paths with spikes in instantaneous volatility. The two processes complement each other as they imply very different dynamics for the evolution of the implied volatility surface. It stands to reason that such a convenient underlying driver for multidimensional structures should be used to improve not only marginal volatility behavior, but also the dependence structure.

In section 4.2, we obtain an analytical representation for the c.f. of the vector of asset prices, which is in closed-form for non-mean-reverting nested cases. This type of c.f. is helpful for derivative pricing purposes. We also produce a second conditional c.f. that can be used for exact simulations of the non-mean reverting assets given the terminal volatilities, where the latter can be simulated exactly via chi-squares. We identify a set of conditions that not only produces well-defined changes of measure, but also avoids local martingales; hence, it can be used for risk-neutral pricing purposes.

In section 4.3, our results are applied numerically to parameters inspired by commodity prices[1]. There is a vast literature on commodity modeling (see, for instance, [122] [34], and more recently [121]). In our numerical study, we investigated the impact of the new parameters (b, the weight of 3/2 in the overall instantaneous volatility) on the shape of the implied volatility surface. This chapter is a result of collaboration with Prof. Marcos Escobar-Anel and Yuyang Cheng.

4.1   Model Description

Next, we define the model in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ where $\mathcal{F}_0$ contains all subsets of the $(\mathbb{P}-)$ null sets of $\mathcal{F}$ and $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous. We first provide the processes under the historical measure $\mathbb{P}$, then followed by the processes under a (conveniently chosen) risk-neutral measure $\mathbb{Q}$. Suppose that $\mathbf{X_t} = (X_1(t), ..., X_n(t))'$ is a vector of asset prices with the

---

[1]This can also be applied to volatility indexes, such as those reported by the Chicago Board Options Exchange (CBOE), which are clearly a mean-reverting asset class with stochastic volatility.

following $\mathbb{P}$-measure representation:

$$
\begin{aligned}
\frac{dX_i(t)}{X_i(t)} &= \left\{ L_i + c_i \sum_{j=1}^{p} a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum_{j=1}^{p} \beta_{ij} \ln(X_j(t)) + \tilde{c}_i \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right)^2 \right\} dt \\
&\quad + \sum_{j=1}^{p} a_{ij} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t) + \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) d\widetilde{W}_i^{\mathbb{P}}(t) \\
dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \quad j = 1, ..., p \\
d\tilde{v}_i(t) &= \tilde{\alpha}_i(\tilde{\theta}_i - \tilde{v}_i(t))dt + \tilde{\xi}_i \sqrt{\tilde{v}_i(t)} d\tilde{B}_i^{\mathbb{P}}(t), \quad i = 1, ..., n
\end{aligned}
$$

The quadratic variation structure is $\langle dB_j^{\mathbb{P}}(t), dW_j^{\mathbb{P}}(t) \rangle = \rho_j dt, \langle d\tilde{B}_i^{\mathbb{P}}(t), d\widetilde{W}_i^{\mathbb{P}}(t) \rangle = \tilde{\rho}_i dt$ and zero otherwise. In the language of factor analysis, a_{ij} is the ijth entry of the matrix of factor loadings (A) that captures the correlations among assets. The commonalities are represented by $V_j(t) = \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2$ (in matrix form, $\Lambda_{nxp} = A diag(V_t^{1/2})$) and the intrinsic residual variance is $\Psi = diag(\widetilde{V}_t)$, with $\widetilde{V}_j(t) = \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right)^2$. This leads to a factors decomposition of the quadratic variation of asset prices as follows:

$$
\Sigma_t dt = (\Lambda\Lambda' + \Psi) dt = \left(A diag(V_t)A' + diag(\widetilde{V}_t) \right) dt
$$

Whenever necessary, we assume $n = p$ and $A = (a_{ij})_{n \times p}$ to be an orthogonal matrix. In this setting, c_i and $\tilde{c}_i$ represent risk premiums of asset $X_i(t)$ associated with the common and intrinsic factors, respectively. $\mathbf{B} = (\beta_{ij})_{n \times n}$ is an invertible matrix, which captures the spillover at the expected return level $X_i(t)$ on asset $X_j(t)$. In other words, it represents the impact from other assets on the long term average price of the current one.

Based on the quadratic variation relationship defined in this model, if we assume that $B_j^{\mathbb{P}}$, $B_j^{\mathbb{P}}(t)^{\perp}$, $\tilde{B}_i^{\mathbb{P}}(t)$, $\tilde{B}_i^{\mathbb{P}}(t)^{\perp}$ are independent Brownian motions with $-1 \leq \rho_j \leq 1$ and $-1 \leq \tilde{\rho}_i \leq 1$. Then,

$$
dW_j^{\mathbb{P}}(t) = \rho_j dB_j^{\mathbb{P}}(t) + \sqrt{1 - \rho_j^2} dB_j^{\mathbb{P}}(t)^{\perp}
$$

$$
d\widetilde{W}_i^{\mathbb{P}}(t) = \tilde{\rho}_i d\tilde{B}_i^{\mathbb{P}}(t) + \sqrt{1 - \tilde{\rho}_i^2} d\tilde{B}_i^{\mathbb{P}}(t)^{\perp}.
$$

v_j $j = 1, .., n$ and $\tilde{v}_i$ $i = 1, .., n$ follow standard CIR processes, hence α_j, θ_j, and ξ_j are positive constants satisfying $\alpha_j \theta_j \geq \frac{\xi_j^2}{2}$ (the Feller condition). Similarly, $\tilde{\alpha}_i$, $\tilde{\theta}_i$, and $\tilde{\xi}_i$ are positive constants satisfying $\tilde{\alpha}_i \tilde{\theta}_i \geq \frac{\tilde{\xi}_i^2}{2}$. Note that the Feller condition in CIR model guarantees that the process remains positive.

The transformation $\mathbf{Y_t} = \ln(\mathbf{X_t})$ would create a multivariate Ornstein–Uhlenbeck process with a 4/2 stochastic factor structure:

$$
\begin{aligned}
dY_i(t) &= \left\{ L_i + \left(c_i - \frac{1}{2} \right) \sum_{j=1}^{n} a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum_{j=1}^{n} \beta_{ij} Y_j(t) + \left(\tilde{c}_i - \frac{1}{2} \right) \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right)^2 \right\} dt \\
&\quad + \sum_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j(t) + \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) d\widetilde{W}_i(t)
\end{aligned}
$$

To stay in line with [94] and [95], we enforce the same assumption that the eigenvalues of the matrix

$$\begin{pmatrix} -\beta_{11} & \cdots & -\beta_{1n} \\ \vdots & \ddots & \vdots \\ -\beta_{n1} & \cdots & -\beta_{nn} \end{pmatrix}$$

are all negative.

This general model includes a notable particular case, which is a direct generalization of [76] to a factor setting when $\beta_{ij} = 0$, $i, j = 1, .., n$; this case is studied in more detail and it is named FG, given its analytical flexibility.

4.1.1 Special Case: No Intrinsic Factor

In this section, we introduce a model as a special case of our generalized mean-reverting 4/2 factor model. In this case, the only factor that impacts the asset is the common factor of the asset i.e $\Psi = 0$.

Suppose $X_t = (X_1(t), \ldots, X_n(t))'$ is a vector of assets. The dynamics for each asset $X_i(t)$ under historical measure $\mathbb{P}$ is defined as:

$$\begin{cases} \dfrac{dX_i(t)}{X_i(t)} & = \left[L_i + c_i \sum\limits_{j=1}^{n} a_{ij}^2 \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum\limits_{j=1}^{n} \beta_{ij} ln(X_j(t)) \right] dt + \sum\limits_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t), \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \ j = 1, \ldots, n \end{cases}$$

$$(4.1)$$

The dynamics of log price $Y_i(t) = ln(X_i(t))$ is then:

$$\begin{cases} dY_i(t) & = \left[L_i + (c_i - \tfrac{1}{2}) \sum\limits_{j=1}^{n} a_{ij}^2 \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum\limits_{j=1}^{n} \beta_{ij} Y_j(t) \right] dt + \sum\limits_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t), \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \ j = 1, \ldots, n. \end{cases}$$

$$(4.2)$$

For the covariance matrix $\boldsymbol{\Sigma_t}$, the factor decomposition is therefore:

$$\boldsymbol{\Sigma_t} dt = \boldsymbol{\Lambda\Lambda'} dt = \mathbf{A} diag(\mathbf{V_t}) \mathbf{A}' dt$$

For example, we are modeling two assets with SDE 4.1:

$$\frac{dX_1(t)}{X_1(t)} = \left[L_1 + a_{11}^2 c_1 \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right)^2 + a_{12}^2 c_1 \left(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}} \right)^2 - \beta_{11} ln(X_1(t)) - \beta_{12} ln(X_2(t)) \right] dt$$

$$+ a_{11} \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right) dW_1^{\mathbb{P}}(t) + a_{12} \left(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}} \right) dW_2^{\mathbb{P}}(t),$$

$$\frac{dX_2(t)}{X_2(t)} = \left[L_2 + a_{21}^2 c_2 \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right)^2 + a_{22}^2 c_2 \left(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}} \right)^2 - \beta_{21} ln(X_1(t)) - \beta_{22} ln(X_2(t)) \right] dt$$

$$+ a_{21} \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right) dW_1^{\mathbb{P}}(t) + a_{22} \left(\sqrt{v_2(t)} + \frac{b_2}{\sqrt{v_2(t)}} \right) dW_2^{\mathbb{P}}(t),$$

4.2 Results

This section describes viable changes of measure and two key characteristic functions of the targeted multivariate process; one for pricing and the other for simulations. The proofs are presented in B.1.

4.2.1 Change of Measure

Here, we explore the topic of creating a risk-neutral measure $\mathbb{Q}$ for pricing purposes. As noted by [76] [118] and [10] among others, a risk-neutral measure may not be supported by data in the presence of a 3/2 model (e.g., $\frac{1}{\sqrt{v(t)}}$), as the parametric constraints needed for the discounted asset price process to be a $\mathbb{Q}$- martingale are violated with real data; hence, we can only produce a strict $\mathbb{Q}$-local martingale (i.e., $\mathbb{Q}$ would be absolute continuous but not equivalent to $\mathbb{P}$). In such situation, the standard risk-neutral pricing methodology would fail (biased prices), and we have to turn to the benchmark approach for pricing (see [10]).

The next proposition entertains the following changes of measure with constant λ_j, $\lambda_j^\perp$, $\widetilde{\lambda}_i$ and $\widetilde{\lambda}_i^\perp$ (see chapter 3 for other types of changes of measures) then identifies the parametric conditions needed for the existence of a valid risk-neutral measure $\mathbb{Q}$.

$$dB_j^Q(t) = \lambda_j \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dt + dB_j^P(t), \, d\tilde{B}_i^Q(t) = \widetilde{\lambda}_i \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) dt + d\tilde{B}_i^P(t)$$

$$dB_j^Q(t)^\perp = \lambda_j^\perp \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dt + dB_j^P(t)^\perp, \, d\tilde{B}_i^Q(t)^\perp = \widetilde{\lambda}_i^\perp \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) dt + d\tilde{B}_i^P(t)^\perp$$

Proposition 4.2.1 *The change of measure is well-defined for pricing purposes under the following four conditions:*

$$\xi_j^2 \leq 2\alpha_j\theta_j - 2\xi_j \max \left\{ |b_j\lambda_j|, |\lambda_j^\perp b_j|, |b_j a_{1j}\rho_j|, ..., |b_j a_{nj}\rho_j| \right\} \tag{4.3}$$

$$\widetilde{\xi}_i^2 \leq 2\widetilde{\alpha}_i\widetilde{\theta}_i - 2\widetilde{\xi}_i \max \left\{ |\tilde{b}_i\widetilde{\lambda}_i|, |\widetilde{\lambda}_i^\perp \tilde{b}_i|, |\tilde{b}_i\widetilde{\rho}_i| \right\} \tag{4.4}$$

$$\max \left\{ |\lambda_j|, |\lambda_j^\perp| \right\} < \frac{\alpha_j}{\xi_j} \tag{4.5}$$

$$\max \left\{ |\widetilde{\lambda}_i|, |\widetilde{\lambda}_i^\perp| \right\} < \frac{\widetilde{\alpha}_i}{\widetilde{\xi}_i} \tag{4.6}$$

Moreover, if $\beta_{ij} = 0$ for $i, j = 1, .., n$, then the following must also be satisfied:

$$L_i = r, \, c_i = \sum_{j=1}^n a_{ij} \left(\rho_j\lambda_j + \sqrt{1 - \rho_j^2}\lambda_j^\perp \right), \, \widetilde{c}_i = \widetilde{\rho}_i\widetilde{\lambda}_i + \sqrt{1 - \widetilde{\rho}_i^2}\widetilde{\lambda}_i^\perp \tag{4.7}$$

Proof is included in Appendix.

4.2.2 Characteristic Function

This section aims at obtaining an analytical representation for the c.f.. If $\mathbf{Z(t)} = e^{\mathbf{B}t}\mathbf{Y(t)}$ is defined such that $e^{\mathbf{B}t}$ is a matrix exponential, then $Z_i(t)$ is represented as:

$$
\begin{aligned}
dZ_i(t) = \sum_{j=1}^{n} \left(e^{\mathbf{B}t}\right)_{ij} \Bigg\{ & L_i + \left(c_j - \frac{1}{2}\right) \sum_{k=1}^{n} a_{jk}^2 \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right)^2 \\
& + \left(\tilde{c}_j - \frac{1}{2}\right) \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}}\right)^2 \Bigg\} dt \\
& + \sum_{j=1}^{n} \left(e^{\mathbf{B}t}\right)_{ij} \Bigg\{ \sum_{k=1}^{n} a_{jk} \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right) dW_k(t) + \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}}\right) d\widetilde{W}_j(t) \Bigg\}
\end{aligned}
\tag{4.8}
$$

For convenience, we use $\left(e^{\mathbf{B}t}\right)_{ij}$ as the ij component of the matrix $e^{\mathbf{B}t}$. Note that $Z_i(t)$ is no longer a mean-reverting process although it accounts for time dependent coefficients.

Next, we find the conditional c.f. for the increments of $\mathbf{Z(t)}$, defined as

$$
\Phi_{\mathbf{Z(t)},v(t)}(T,\omega) = E\left[\exp\{i\omega'(\mathbf{Z(T)} - \mathbf{Z(t)})\} \mid \mathbf{Z(t)} = z_t, v(t) = \mathbf{v_t}\right]
\tag{4.9}
$$

Under a risk neutral measure, this c.f. can be used for pricing some financial products, given the integrability conditions (a discussion of the generalized c.f. as per [76] is beyond the scope of this chapter.). For convenience, we formulate the information given at t on volatility as $v(t) = \left(\sqrt{V_1(t)}, .., \sqrt{V_n(t)}, \sqrt{\tilde{V}_1(t)}, .., \sqrt{\tilde{V}_n(t)}\right)$ where $\sqrt{V_k(t)} = \sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}$ and $\sqrt{\tilde{V}_k(t)} = \sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}}$, $j, k = 1, \ldots, n$.

Proposition 4.2.2 *Let* $(\mathbf{Z(t)})_{t \geq 0}$ *evolve according to the model in Equation* (4.8). *The c.f.* $\Phi_{Z(t),v(t)}$ *is then given as follows:*

$$
\begin{aligned}
\Phi_{\mathbf{Z(t)},v(t)}(T,\omega) &= E\left[\exp i\omega'(\mathbf{Z(T)} - \mathbf{Z(t)}) \mid \mathbf{Z(t)} = z_t, v(t) = \mathbf{v_t}\right] \\
&= \prod_{k=1}^{n} \Phi_{GG}\left(T, 1; L_k(\omega), h_k(\omega), g_k(\omega), \kappa_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, S^*_{k,t}\right) \\
&\quad \times \prod_{j=1}^{n} \Phi_{GG}\left(T, 1; 0, L_j(\omega), h_j(\omega), g_j(\omega), \widetilde{\kappa}_j, \widetilde{\theta}_j, \widetilde{\xi}_j, \widetilde{\rho}_j, \tilde{b}_j, \widetilde{c}_k, \widetilde{v}_{j,t}, S^{*j}_t\right)
\end{aligned}
$$

where Φ_{GG} *is a one-dimensional generalization of the c.f. from [76] provided in Lemma B.1.1.*

Proof is provided in Appendix. The c.f. above involves single expected values with respect to Brownian motion $B(t)$. In each term, Φ_{GG} (i.e., the second set of Brownian $W(t)$) is eliminated, hence this is a drastic simplification compared to the original $2n$ dimensional joint expectation.

A particular, fully solvable case is the FG model ($\beta_{ij} = 0$, $i, j = 1, .., n$).

Corollary 4.2.3 *Let $(\mathbf{Z}(t))_{t \geq 0}$ evolve according to the FG model ($\beta_{ij} = 0$, $i, j = 1, .., n$). The c.f. $\Phi_{\mathbf{Z}(t),v(t)}$ is subsequently presented as follows:*

$$\Phi_{\mathbf{Z}(t),v(t)}(T, \omega) = E\left[\exp i\omega'(\mathbf{Z(T)} - \mathbf{Z(t)}) \mid \mathbf{Z(t)} = z_t, v(t) = \mathbf{v_t}\right]$$

$$= \prod_{k=1}^{n} \Phi_G\left(T, 1; L_k(\omega), h_k(\omega), g_k(\omega), \kappa_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, S_{k,t}^*\right)$$

$$\times \prod_{j=1}^{n} \Phi_G\left(T, 1; 0, L_j(\omega), h_j(\omega), g_j(\omega), \widetilde{\kappa}_j, \widetilde{\theta}_j, \widetilde{\xi}_j, \widetilde{\rho}_j, \widetilde{b}_j, \widetilde{c}_k, \widetilde{v}_{j,t}, S_t^{*j}\right)$$

where Φ_G is the one-dimensional c.f provided by [76] in Proposition 3.1 and given in the Appendix for completeness.

See Appendix for proof. Next, we turn to the conditional c.f. of the increments of $Z(t)$ given the terminal value of the CIR processes. This is defined as follows:

$$\Phi_{\mathbf{Z}(t),v(T)}(\tau, \omega) = E\left[\exp\left[\omega'(\mathbf{Z(T)} - \mathbf{Z(t)})\right] \mid \mathbf{Z(t)} = z_t, v(T) = \mathbf{v_T}\right] \tag{4.10}$$

The above is useful when we need to work with the joint distribution of $(\mathbf{Z(T)}, v(T))$ given $(\mathbf{Z(t)}, v(t))$. For such cases, we can try to rely on a convenient simulation scheme combining the distribution of $\mathbf{Z(T)}$ given $(\mathbf{Z(t)}, v(T))$ (via Equation (4.10)) with that of $v(T)$ given $v(t)$, the latter is known to be non-centered chi-squared. In this way, we can avoid usual discretization algorithms such as the Euler–Maruyama or Milstein schemes, which are generally not suitable for the CIR process (due to failure of the Lipschitz condition at 0).

In this vein, when working with the non mean-reverting factor model ($\beta_{ij} = 0$, $i, j = 1, .., n$), we can easily adapt the procedures in [76] to provide an exact simulation scheme for the model given the vector of the independent CIR process at maturity T (i.e., $v(T)$). This requires only the c.f. provided next:

Corollary 4.2.4 *Let $(\mathbf{Z}(t))_{t \geq 0}$ evolve according to the FG model ($\beta_{ij} = 0$, $i, j = 1, .., n$). Then, the c.f. $\Phi_{\mathbf{Z}(t),v(T)}$ is then given as follows:*

$$\Phi_{\mathbf{Z}(t),v(T)}(T, \omega) = \prod_{j=1}^{n} \Phi_{G,T}\left(T, \phi; L, h_j, g_j, \kappa_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,T}, S_{j,t}^*\right)$$

$$\times \prod_{i=1}^{n} \Phi_{G,T}\left(T, 1; 0, h_i, g_i, \widetilde{\kappa}_i, \widetilde{\theta}_i, \widetilde{\xi}_i, \widetilde{\rho}_i, \widetilde{b}_i, \widetilde{c}_i, \widetilde{v}_{i,T}, S_t^{*i}\right)$$

where $\Phi_{G,T}$ is the one-dimensional c.f provided by [76] in Proposition 4.1 and given in the Appendix for completeness.

Proof of this result is provided in Appendix. Unsurprisingly, the previous result cannot be extended to the mean-reverting case, due to the absence of closed formulas for the object:

$$\mathbb{E}\left[\exp\left\{u\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds\right)\right\} \mid v(T)\right]$$

which is not solvable even when two of the three deterministic functions $B(s)$, $C(s)$ and $D(s)$ are zero.

4.3 Discussion: One Common Factor in Two Dimensions

We assume two assets, i.e., $X_1(t)$ and $X_2(t)$, with one common volatility component, and one intrinsic factor each. The asset prices thereby follow the system of SDE for $i = 1, 2$:

$$
\begin{aligned}
dY_i(t) &= (L_i - \beta_i Y_i(t)) \, dt \\
&+ \left((c_i - \frac{1}{2})[a_i^2(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}})^2] + (\tilde{c}_i - \frac{1}{2})(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}})^2 \right) dt \\
&+ a_i \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right) dW_1(t) + \left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}} \right) d\widetilde{W}_i(t)
\end{aligned}
$$

$$
\begin{aligned}
dv_1(t) &= \alpha_1(\theta_1 - v_1(t))dt + \xi_1 \sqrt{v_1(t)}dB_1(t) \\
d\tilde{v}_i(t) &= \tilde{\alpha}_i(\tilde{\theta}_i - \tilde{v}_i(t))dt + \tilde{\xi}_i \sqrt{\tilde{v}_i(t)}d\tilde{B}_i(t)
\end{aligned}
$$

with $\langle dB_j(t), dW_j(t) \rangle = \rho_j dt, \langle d\tilde{B}_i(t), d\widetilde{W}_i(t) \rangle = \tilde{\rho}_i dt$ for $j = 1; i = 1, 2$.

The following table (Table 4.1) gives a baseline parameter set for the one-factor, two-dimensional 4/2 factor model used in the subsequent sections. The choice of parameters in Scenario A was made by combining the seminal works of [122] (see Oil and Copper in Tables IV and V) and [80]. Scenario B combines [122] (see Oil and Copper, Tables IV and V) with [36]. In both cases, we assume a simple structure for the market price of risk $(c_1 = c_2 = \tilde{c}_1 = \tilde{c}_2 = 0)^2$.

The $\tilde{\theta}_i$, $i = 1, 2$ in the table are set to match the long term volatilities as estimated in [122], which are 0.334 (Oil, Table IV) and 0.233 (Copper, Table V):

$$
\mathbb{E}\left[a_1^2 \left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right)^2 + \left(\sqrt{\tilde{v}_1(t)} + \frac{\tilde{b}_1}{\sqrt{\tilde{v}_1(t)}} \right)^2 \right] \tag{4.11}
$$

$$
= a_1^2 \left(\frac{2\alpha_1 b_1^2}{2\alpha_1\theta_1 - \xi_1^2} + 2b_1 + \theta_1 \right) + \frac{2\tilde{\alpha}_1\tilde{b}_1^2}{2\tilde{\alpha}_1\tilde{\theta}_1 - \tilde{\xi}_1^2} + 2\tilde{b}_1 + \tilde{\theta}_1 = (0.334)^2 \tag{4.12}
$$

This explains the values of $\tilde{\theta}_i$ in the table. The present section considers two independent cases. First, we study the impact of the parameters $b_1, \tilde{b}_1$ and $\tilde{b}_2$ on implied volatility surfaces and on two risk measures for a portfolio of underlyings. We then assess the impact of the commonalities a_1 and a_2 on these same targets, i.e., implied volatilities and risk measures. To ensure that the cases lead to reasonable assets behavior, we report the expected return, variance of return for each asset, as well as the correlation between two assets and the leverage effects in Tables 4.2 and 4.3 under Scenarios A and B, respectively.

We simulated 500,000 paths with $dt = 0.1$ and considered the following scenarios for b: $b_1 = 0.008$, $\tilde{b}_1 = \tilde{b}_2 = 0$; $b_1 = 0, \tilde{b}_1 = \tilde{b}_2 = 0.008$; $b_1 = \tilde{b}_1 = \tilde{b}_2 = 0$ and $b_1 = \tilde{b}_1 = \tilde{b}_2 = 0.008$. Similarly, we considered the following scenarios for a: $a_1 = a_2 = 0$; $a_1 = 0.75, a_2 = 0$; $a_1 = 0$, $a_2 = 0.75$ and $a_1 = a_2 = 0.75$. Tables 4.4 and 4.5 present key statistics for the returns under Scenarios A and B, respectively.

---

[2]Variations on c will be studied in future research as part of a calibration exercise (see [106] for viable approaches and [61] for some pitfalls).

Table 4.1: Toy Parametric Values

Initial Values
$X_1(0) = 18, X_2(0) = 100$
$v_1(0) = \theta_1, \tilde{v}_1(0) = \tilde{\theta}_1, \tilde{v}_2(0) = \tilde{\theta}_2$
Commodity Drift, [122]
$\beta_{11} = 0.301, \beta_{12} = 0, \beta_{21} = 0, \beta_{22} = 0.369$
$L_1 = 3.09\beta_{11} = 0.93, L_2 = 4.85\beta_{22} = 1.79$
Commodity St. Volatility, [80, 122]. Scenario A
$\alpha_1 = \tilde{\alpha}_1 = \tilde{\alpha}_2 = 2$
$\theta_1 = 0.01, \tilde{\theta}_1 = 0.0753, \tilde{\theta}_2 = 0.0124$
$\xi_1 = \tilde{\xi}_1 = \tilde{\xi}_2 = 0.1$
$\rho_1 = \widetilde{\rho}_1 = \widetilde{\rho}_2 = -0.5$
Commodity St. Volatility, [36, 122]. Scenario B
$\alpha_1 = \tilde{\alpha}_1 = \tilde{\alpha}_2 = 0.2098$
$\theta_1 = 0.1633, \tilde{\theta}_1 = 0.0685, \tilde{\theta}_2 = 0.0689$
$\xi_1 = \tilde{\xi}_1 = \tilde{\xi}_2 = 0.1706$
$\rho_1 = \widetilde{\rho}_1 = \widetilde{\rho}_2 = -0.9$
New parameters.
$c_1 = c_2 = \tilde{c}_1 = \tilde{c}_2 = 0$
$a_1 = a_2 = 0.75$
$b_1 = \widetilde{b}_1 = \widetilde{b}_2 = 0.008$

Table 4.2: First four moments for scenarios on 3/2 component (b), Scenario A.

	$b_1 = 0.008, \tilde{b}_1 = \tilde{b}_2 = 0$	$b_1 = 0, \tilde{b}_1 = \tilde{b}_2 = 0.008$	$b_1 = \tilde{b}_1 = \tilde{b}_2 = 0$	$b_1 = \tilde{b}_1 = \tilde{b}_2 = 0.008$
$\mathbb{E}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0494	0.0489	0.0503	0.0480
$\mathbb{E}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0760	0.0754	0.0764	0.0750
$\mathbb{V}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0663	0.0680	0.0618	0.0729
$\mathbb{V}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0367	0.0390	0.0323	0.0445
Corr$(\ln X_1(T), \ln X_2(T))$	0.3194	0.0896	0.1060	0.2799
Corr$(\ln X_1(T), < \ln X_1(T) >)$	−0.4287	−0.4520	−0.4443	−0.4406
Corr$(\ln X_2(T), < \ln X_2(T) >)$	-0.4148	−0.4511	−0.4420	−0.4280

Table 4.3: First four moments for scenarios on 3/2 component (b), Scenario B.

	$b_1 = 0.008, \tilde{b}_1 = \tilde{b}_2 = 0$	$b_1 = 0, \tilde{b}_1 = \tilde{b}_2 = 0.008$	$b_1 = \tilde{b}_1 = \tilde{b}_2 = 0$	$b_1 = \tilde{b}_1 = \tilde{b}_2 = 0.008$
$\mathbb{E}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0514	0.0504	0.0527	0.0499
$\mathbb{E}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0774	0.0775	0.0787	0.0754
$\mathbb{V}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0360	0.0606	0.0247	0.1022
$\mathbb{V}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0359	0.1508	0.0248	0.0723
Corr$(\ln X_1(T), \ln X_2(T))$	0.7533	0.0099	0.4698	0.0156
Corr$(\ln X_1(T), < \ln X_1(T) >)$	−0.4509	−0.2031	−0.5560	−0.0273
Corr$(\ln X_2(T), < \ln X_2(T) >)$	−0.4496	−0.0398	−0.5517	−0.2444

4.3.1   Pricing Option

The section prices European call option on the asset X_1 based on our 4/2 generalized factor model. It explores the implied volatility surface in a three-dimensional plot with strike prices

Table 4.4: First four moments for scenarios on commonalities (a), Scenario A.

	$a_1 = a_2 = 0$	$a_1 = 0.75, a_2 = 0$	$a_1 = 0, a_2 = 0.75$	$a_1 = a_2 = 0.75$
$\mathbb{E}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0491	0.0492	0.0493	0.0492
$\mathbb{E}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0761	0.0757	0.0746	0.0759
$\mathbb{V}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0658	0.0727	0.0660	0.0732
$\mathbb{V}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0371	0.0370	0.0444	0.0447
Corr$(lnX_1(T), lnX_2(T))$	0.0000	-0.0011	0.0004	0.2841
Corr$(lnX_1(T), < lnX_1(T) >)$	-0.4523	-0.4400	-0.4529	-0.4430
Corr$(lnX_2(T), < lnX_2(T) >)$	-0.4507	-0.4521	-0.4285	-0.4279

Table 4.5: First four moments for scenarios on commonalities (a), Scenario B.

	$a_1 = a_2 = 0$	$a_1 = 0.75, a_2 = 0$	$a_1 = 0, a_2 = 0.75$	$a_1 = a_2 = 0.75$
$\mathbb{E}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0514	0.0496	0.0512	0.0500
$\mathbb{E}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0768	0.0772	0.0756	0.0752
$\mathbb{V}[\frac{X_1(T)-X_1(0)}{X_1(0)}]$	0.0420	0.0733	0.0420	0.0857
$\mathbb{V}[\frac{X_2(T)-X_2(0)}{X_2(0)}]$	0.0507	0.0564	0.0746	0.0719
Corr$(lnX_1(T), lnX_2(T))$	0.0004	-0.0000	-0.0005	0.0039
Corr$(lnX_1(T), < lnX_1(T) >)$	-0.1126	-0.0047	-0.3203	-0.2634
Corr$(lnX_2(T), < lnX_2(T) >)$	-0.1600	-0.2544	-0.0164	-0.0111

as the x-axis, time to maturity as the y-axis, and corresponding implied volatility as z-axis. We take the strike prices K to be 15, 16.4, 17.8, 19.2, 20.6, and 22 and the expiry dates T are 0.2, 0.36, 0.52, 0.68, 0.84 and 1.0. By choosing these strike prices, we account for the in-the-money, at-the-money, and out-of-the-money options, given the initial asset price 18. Subsequently, for each strike price and expiry date, we can obtain a simulated call price as follows

$$c(T, K) = e^{-rT}\mathbb{E}^Q[(X_1(T) - K)^+],$$

where $X_1(T)$ is approximated using the Euler method.

We extract the implied volatility by matching the Black–Scholes option price formula with simulated call prices and solve for the volatility parameter. Hence, we can treat the dynamics of $Y(t)$ as an O-U process such that:

$$dY(t) = (L_1 - 0.5\sigma^2 - \beta Y(t))dt + \sigma dW^*(t).$$

Next, we consider the two cases described above. The first one studies the impact of b, which represents the size of the 3/2 component on the covariance; and the second examines the impact of a, the size of the commonality.

In the case of b, we first extract the implied volatility surface by matching the standard BS formula for changes on b and $\tilde{b}$ respectively (see Figures 4.1 and 4.2).

(a) $\tilde{b}_1$=0, b_1 between (0, 0.008)

(b) $\tilde{b}_1$=0.008, b_1 between (0, 0.008)
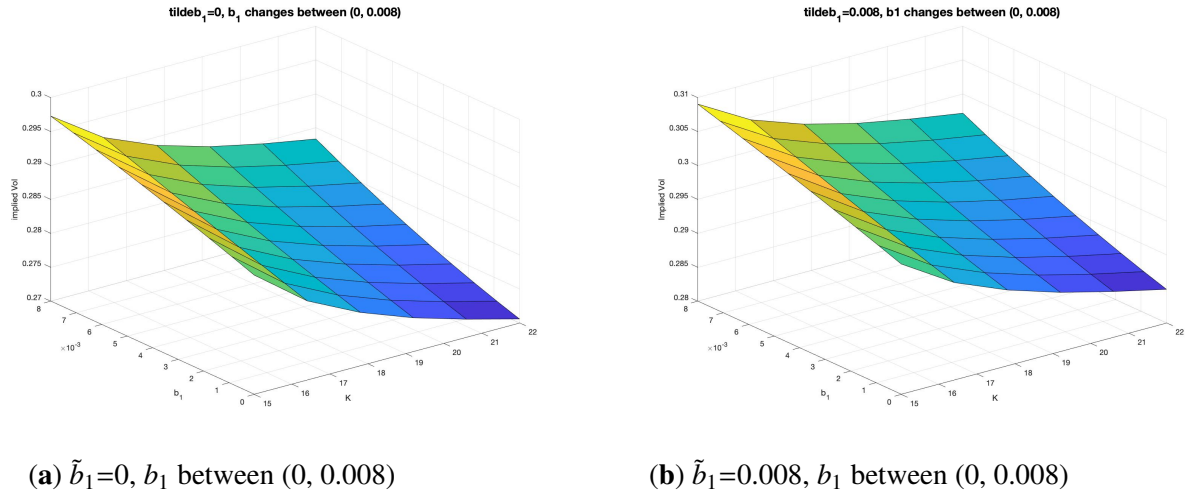
Figure 4.1: Impact of b_1 (common factor, 3/2 component) on implied volatility, Scenario A.

(a) $\tilde{b}_1$=0, b_1 between (0, 0.008)

(b) $\tilde{b}_1$=0.008, b_1 between (0, 0.008)
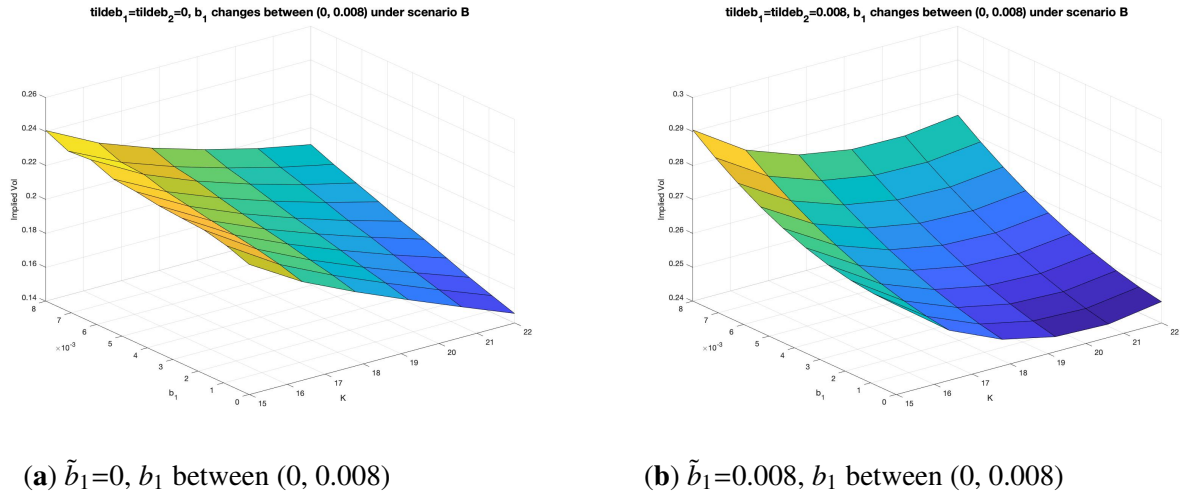
Figure 4.2: Impact of b_1 (common factor, 3/2 component) on implied volatility, Scenario B.

For Scenario **A**, Figure 4.1a,b illustrate that even small changes in (b_1) the common factor 3/2 component (from 0 to 0.008) can lead to a 7% difference in implied volatility (from 0.275 to 0.295, or 0.285 to 0.305). The joint effect of the common and intrinsic 3/2 components (b_1 and $\tilde{b}_1$) can be obtained by combining those two figures leading to a 11% change (from 0.275 to 0.305) in the presence of relatively small values of b.

For Scenario **B**, we observe that the impact of intrinsic factor on volatility surface is more significant than in Scenario **A** through a comparison of Figure 4.2a,b. The effect of b_1 on implied volatility increase by approximately 31% (0.145 to 0.19), as shown in Figure 4.2a, when only the common factor is present. In Figure 4.2b, we observe a volatility "smile" with the difference of approximately 12.2% (0.245 to 0.275). The joint effect of the common and intrinsic 3/2 components in this case is 100% (0.145 to 0.29).

Figures 4.1 and 4.2 jointly demonstrate that, given different underlying process for common and intrinsic factors, the impact of the 3/2 component can be crucial.

Next, we study a, the weight of the common factor (commonality). We again extract the implied volatility surface from matching the standard BS formula for changes on a.

Figure 4.3a,b displays the significant increase in implied volatility due to the commonality of the asset with the market (a_1). The change in implied volatility can increase up to 12.5% (from 0.28 to 0.315) in Scenario A and up to 30% (from 0.22 to 0.32) in Scenario B.
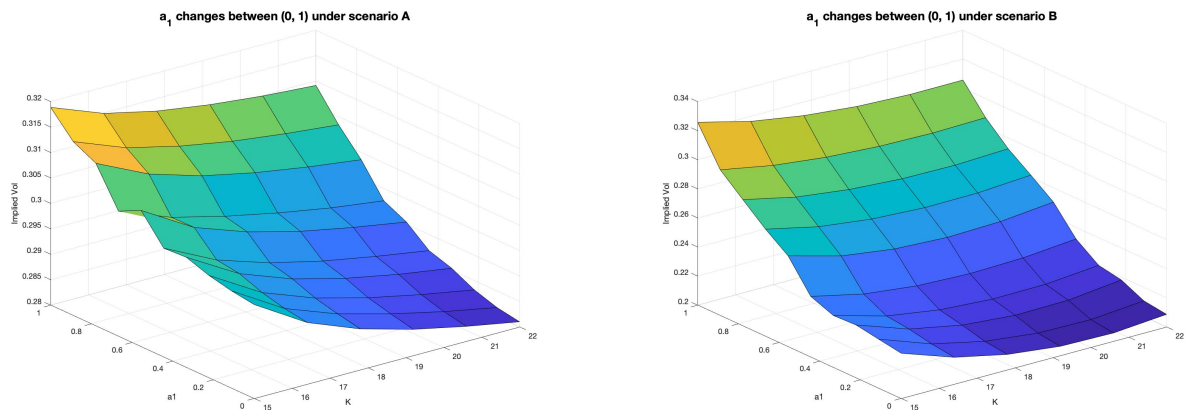
(**a**) a_1 between (0, 1). Scenario A

(**b**) a_1 between (0, 1). Scenario B

Figure 4.3: Impact of commonality (a_1) on implied volatility.

4.4 Conclusions

In previous chapter 3, we consider a single asset whose diffusion is affected by common factor or systematic/market risk, and we assume the factor follows a 4/2 porcess because 4/2 process is capable of capturing the stylized facts in asset price time series through out the entire business cycle. In this chapter, we propose and explore the generalized mean-reverting 4/2 factor model. The generalized model is a more realistic model than the one-dimensional model since in the real market there is more than the common factors that influence the asset prices. There are also intrinsic factors and spillover effects from other assets. The generalized model incorporates spillover effects, common factors and intrinsic factor. We assume 4/2 processes for both common factor and intrinsic factors not only because 4/2 processes are powerful capturing the stylized facts, they are also flexible by adjusting the "weights" of 3/2 components.

In this study, we first provide analytical expressions for key characteristic functions and conditions for well-defined changes of measures. Thanks to principal component decomposition, we are able to write the generalized model in terms of a linear combination of independent mean-reverting 4/2 stochastic volatility models. Thus the generalized model inherits the same theoretical challenges of the mean-reverting 4/2 stochastic volatility model, which is a closed-form solution of the characteristic function is not available. As we mentioned in chapter 3, even with the semi-closed form characteristic function, we can still reduce the computational cost. We are also able to derive an analytic characteristic function for a special case, which we call the generalized Grasselli's 4/2 model [76]. This opens a new window for studying multi-factor model in stock market. In numerical application section, we assess the impact of b, i.e. the 3/2 component of the model in the volatility process, and a, the commonalities in the absence and presence of stochastic volatility on the common factor. These impacts were measured with respect to implied volatility surfaces and two important risk measures. The results demonstrate that even small values of the 3/2 component (b) can lead to a 100% change in the implied volatility surface, which in turn shows the significance of the 3/2 component in the factors' processes. In next chapter, we will focus on a subclass of the generalized model and further evaluate the model performance.

Chapter 5

Multivariate Mean-Reverting 4/2 Stochastic Volatility Model

Chapter 4 introduced a generalized multivariate mean-reverting 4/2 factor model. Thanks to a principal component decomposition (P.C), we expressed the stochastic covariance matrix in terms of constant eigenvectors that drives the correlation and stochastic eigenvalues modeled by 4/2 stochastic processes. This means we can rewrite our multidimensional model in the form of linear combination of independent one-dimensional mean-reverting 4/2 processes. In chapter 4, we entertained interesting structures and studied properties of the model such as changes of measure and characteristics functions (c.f.s), we also applied the model to option pricing.

Now, in chapter 5, we extend two of the results of chapter 3 to the principal component setting of chapter 4. These are: 1) obtaining a closed-form approximation of the c.f.s; 2) estimation of the parameters of the model. Moreover we will expose our model to a setting of portfolio optimization for the fist time, to study its implications.

In particular, we remind the reader that the c.f.'s we found in chapter 4 have quasi-closed form expressions, which requires simulation of a lower dimensional process. This lack of analyticity is not surprising as we have already shown in chapter 3 the challenges of finding closed-form c.f. in one dimension, which can be traced back to the combination of mean-reverting and stochastic volatility features. As a result, in chapter 3, we were left with a quasi-closed form c.f., and some analytical expressions for very special cases of our model. An alternative solution was to find an analytic function that approximates the c.f. with acceptable accuracy within a region of the parametric space, i.e. for embedded models. In chapter 5, the two approximation approaches developed in chapter 3 are implemented in multidimension: midpoint approximation and average approximation. Here we need to carefully design the embedded models where the methodology works, which leads to three special cases. We illustrate their performance in higher dimensions, demonstrating high accuracy of the average approximation in capturing not only marginal densities but also joint behaviour.

This chapter also answers another question that is not addressed in chapter 4, namely estimation. We use the estimation approaches developed in chapter 3 to estimate the parameters

for special cases of P.C mean-reverting 4/2 model. For this exercise we use two pairs of bivariate time series capturing underlying and variance behaviour. Estimation of multi-dimensional processes is rare in the literature. Popular estimation methods, for example efficient method of moments (EMM) [70], generalized method of moments (GMM) [78] and quasi-maximum likelihood (QML) [110] [79], rely on analytic density/transition density functions of the underlying process, which is not the case for our model as our model has semi-closed form c.f.. Our work shows that many but not all the parameters are statistical significant, confirming and revealing stylized facts of commodity prices and volatility indexes like stochastic correlation and spillover effects.

In the applications part, we focus on the risk management of a portfolio that follows constant proportion strategies. In particular, we calculate a distribution based risk measure for this portfolio, i.e Value at Risk (VaR). As the quasi-closed form c.f. of chapter 4 does not bring efficiency to calculating distribution based risk measures due to simulation, we use the analytic approximations in chapter 5. We demonstrate that the error computing VaR with the approximation is negligible. The various examples presented in the context of portfolio allocation bring justification to the validity of the multidimensional approximation via c.f..

5.1   Model Definition

In this section, we start with the definition of the general model targeted in this chapter. In the general model setup, we introduce a model with spillover effects and only common factors that drive the volatility processes. Later we cover models with separable spillover effects and no spillover effects as special cases of the general model. These models are particular cases of those developed in chapter 4. As a reminder, spillover effects refer to the impact of one asset or underlying stock on the remaining assets/stocks in regards to trends (drift), in other words the impact of one asset on the long term average, "stationary price" of a second asset. In particular, the presence of (positive) spillover implies that large increases in the value of one stock lead to an increase in the mean reverting value of a second stock, the opposite for negative spillovers.

5.1.1   General Model Setup

Recall in section 4.1.1, we briefly introduce a special case of our generalized factor 4/2 model, which has no intrinsic factor, hence inspired in Principal Component Analysis. Suppose $X_t = (X_1(t), \ldots, X_n(t))'$ is a vector of assets. The dynamics for each asset $X_i(t)$ under historical measure $\mathbb{P}$ is defined as:

$$
\begin{cases}
\dfrac{dX_i(t)}{X_i(t)} & = \left[L_i + \sum\limits_{j=1}^{n} c_{ij} \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum\limits_{j=1}^{n} \beta_{ij} ln(X_j(t)) \right] dt + \sum\limits_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \dfrac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t), \\
dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \; j = 1, \ldots, n
\end{cases}
$$

$$(5.1)$$

where $W_i^{\mathbb{P}}(t)$ and $B_j^{\mathbb{P}}(t)$ are independent Brownian motions if $i \neq j$ and are correlated if $i = j$, i.e the quadratic variation $\langle W_i^{\mathbb{P}}(t), B_i^{\mathbb{P}}(t) \rangle = \rho_i t$, where ρ_i is constant. The parameters for each $v_j(t)$ process are positive and satisfy Feller condition: $\alpha_j \theta_j > \frac{\xi_j^2}{2}$. Moreover, we assume the mean-

reverting level of $v_j(t)$ decreases as j increases, i.e $0 < \theta_j < \theta_{j-1}$, for $j = 2, \ldots, n$. This last feature is intended to sort the eigenvalues in order of importance. b_j's measure the "weights" of 3/2 components.

This model is not like a traditional mean-reverting model as it takes into account the spillover effects in the drift, which appears in the form of $\beta_{ij}, i \neq j$. We mentioned above that the spillover effects show the impacts of one assets on others, which shall not be confused with correlations. The correlations are reflected in the price trend of both assets, capturing co-movements between assets. Spillover effects describe the impact on the mean-reverting level of one asset by others, i.e the shift in the long term mean due to movements of other assets. The concept of spillover effects can be understood as how much, for example, a demand curve of one good shifts according to the change of factors on other goods.

Writing Equation 5.1 in matrix form:

$$\mathbf{dX_t} = diag(\mathbf{X_t})[(\mathbf{L} + \mathbf{CV_t} - \mathbf{B} \cdot \mathbf{ln(X_t)})dt + \Sigma_t^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}]. \tag{5.2}$$

where $\mathbf{W_t} = (W_1^{\mathbb{P}}(t), \ldots, W_n^{\mathbb{P}}(t))'$ is a vector of independent standard Brownian motions; $\mathbf{B} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{pmatrix}$, $\mathbf{L} = (L_1, \ldots, L_n)'$, $\mathbf{C} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}$, $\mathbf{V_t} = \left(\left(\sqrt{v_1(t)} + \frac{b_1}{\sqrt{v_1(t)}} \right)^2, \ldots, \left(\sqrt{v_n(t)} + \frac{b_n}{\sqrt{v_n(t)}} \right)^2 \right)'$ and $\mathbf{ln(X_t)} = (ln(X_1(t)), \ldots, ln(X_n(t)))'$. We first assume that the eigenvalues of the matrix

$$\begin{pmatrix} -\beta_{11} & \cdots & -\beta_{1n} \\ \vdots & \ddots & \vdots \\ -\beta_{n1} & \cdots & -\beta_{nn} \end{pmatrix}$$

are all negative as per [94] and [95]. This is assumption will be used to justify estimation results in estimation section. Similarly to chapter 4, we have $\mathbf{B}$ captures the spillover effects; $\mathbf{C}$ contains risk premiums associated to the assets; long term average for the assets is determined by $\mathbb{E}[\mathbf{B}^{-1}(\mathbf{L} + \mathbf{CV_t})]$. We next assume a principal component decomposition on the instantaneous covariance matrix Σ_t:

$$\Sigma_t dt = \mathbf{A} diag(\mathbf{V_t})\mathbf{A}' dt,$$

where $\mathbf{A} = (a_{ij})_{n \times n}$ is an orthogonal matrix with constant entries, which captures the correlation among assets.

We craft the matrix $\mathbf{C}$ in such a way that allows for c.f. analytical approximations, this is $\mathbf{C} = \mathbf{A}\tilde{\mathbf{C}} + \frac{1}{2}(\mathbf{A} \circ \mathbf{A})$ where $\tilde{\mathbf{C}} = diag(c_1, \ldots, c_n)$ and $(\mathbf{A} \circ \mathbf{A})$ denotes Hadamard product of $\mathbf{A}$. The dynamics of log price $Y_i(t) = ln(X_i(t))$ is as follows:

$$\begin{cases} dY_i(t) & = \left[L_i + \sum_{j=1}^{n} a_{ij}c_j \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum_{j=1}^{n} \beta_{ij}Y_j(t) \right]dt + \sum_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)dW_j^{\mathbb{P}}(t), \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \quad j = 1, \ldots, n. \end{cases}$$

$$\tag{5.3}$$

Based on the applications, our model can be reduced to three sub cases for which we are able to approximate the c.f. with analytic functions.

- **$b_j = 0, \rho_j \neq 0$**: The advantage of 4/2 model is that it integrates two diffusion processes with different behavior, which enables the model to capture the stylized facts of the underlying asset under different financial situations. During the period when the financial markets enjoy stable business cycles, 3/2 diffusion process may be redundant in modeling the underlying asset as 1/2 diffusion is good enough to capture the stylized facts. Therefore, the dynamics for log price is defined as:

$$
\begin{cases}
dY_i(t) & = \left[L_i + \sum\limits_{j=1}^{n} a_{ij}c_j v_j(t) - \sum\limits_{j=1}^{n} \beta_{ij} Y_j(t) \right] dt + \sum_{j=1}^{n} a_{ij} \sqrt{v_j(t)} dW_j^{\mathbb{P}}(t), \\
dv_j(t) & = \alpha_j(\theta_j - v_j(t)) dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \ \ j = 1, \ldots, n, \\
\langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t) \rangle & = \rho_j dt
\end{cases}
$$

(5.4)

If $n = 1$, we have a mean-reverting Heston model, which is also considered in [16].

- **$b_j \neq 0, \rho_j = 0$**: This case applies to the assets whose price series demonstrates abnormal increase or decrease, but there is no leverage effect observed for the assets of interest. The term "leverage effect" was first defined and studied in [18]. It describes the negative correlation between asset's volatility and asset's return. In this situation, we know that the trend of asset's volatility provides no information on the trend of asset's return. The model for log price is therefore defined as:

$$
\begin{cases}
dY_i(t) & = \left[L_i + \sum\limits_{j=1}^{n} a_{ij}c_j \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 - \sum\limits_{j=1}^{n} \beta_{ij} Y_j(t) \right] dt \\
& \quad + \sum_{j=1}^{n} a_{ij} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j^{\mathbb{P}}(t), \\
dv_j(t) & = \alpha_j(\theta_j - v_j(t)) dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \ \ j = 1, \ldots, n, \\
\langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t) \rangle & = 0
\end{cases}
$$

(5.5)

- **$b_j = 0, \rho_j = 0$**: This case can be generated by either of previous two cases. It better applies to assets that exhibit mild behavior in their price series, at the same time no leverage effect is identified. Then the model for log price follows from either Equation 5.4 or Equation 5.5:

$$
\begin{cases}
dY_i(t) & = \left[L_i + \sum\limits_{j=1}^{n} a_{ij}c_j v_j(t) - \sum\limits_{j=1}^{n} \beta_{ij} Y_j(t) \right] dt + \sum_{j=1}^{n} a_{ij} \sqrt{v_j(t)} dW_j^{\mathbb{P}}(t), \\
dv_j(t) & = \alpha_j(\theta_j - v_j(t)) dt + \xi_j \sqrt{v_j(t)} dB_j^{\mathbb{P}}(t), \ \ j = 1, \ldots, n, \\
\langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t) \rangle & = 0
\end{cases}
$$

(5.6)

We will show how to approximate the c.f. with analytic functions for these three cases in section 5.2.

5.1.1.1 Separable Spillover effect

In this section, we assume a convenient structure in the spillover matrix **B** together with a slight variation of the methodology to obtain the c.f.. By doing so we obtain another solvable case.

In model setup described above, we defined our model using principal component stochastic volatility framework, we can further simplify our model by rewriting it in terms of n independent processes as we did in chapter 4. Recall that in chapter 4, we define the transformation $\mathbf{Z(t)} = e^{\mathbf{B}t}\mathbf{Y(t)}$. It is also possible that we directly work with log price processes $\mathbf{Y(t)}$ instead of going through with the transformation $\mathbf{Z_t}$, avoiding the matrix exponential.

We demonstrate this procedure by first writing Equation 5.3 in matrix form:

$$\mathbf{dY_t} = (\mathbf{L} + \mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{BY_t})dt + \mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \tag{5.7}$$

Multiplying both sides of Equation 5.7 by $\mathbf{A}^{-1}$, we get:

$$\mathbf{A^{-1}dY_t} = (\mathbf{A^{-1}L} + \mathbf{A^{-1}A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{A^{-1}BY_t})dt + diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \tag{5.8}$$

Suppose the matrix **B** can be written as follows:

$$\mathbf{B} = \mathbf{A}\tilde{\mathbf{B}}\mathbf{A}^{-1}$$

where $\tilde{\mathbf{B}} = diag(\tilde{\beta}_1, \ldots, \tilde{\beta}_n)$ is a diagonal matrix, i.e whose entries are eigenvalues of **B**. Using this result, Equation 5.8 can be written as

$$\mathbf{A^{-1}dY_t} = (\mathbf{A^{-1}L} + \tilde{\mathbf{C}}\mathbf{V_t} - \tilde{\mathbf{B}}\mathbf{A^{-1}Y_t})dt + diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

Applying a simple transformation $\mathbf{M_t} = \mathbf{A^{-1}Y_t}$, we end up with a new mean-reverting process with a diagonal matrix $\tilde{\mathbf{B}}$:

$$\mathbf{dM_t} = (\mathbf{A^{-1}L} + \tilde{\mathbf{C}}\mathbf{V_t} - \tilde{\mathbf{B}}\mathbf{M_t})dt + diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

It is easy to see that each element of $\mathbf{dM_t}$ is a mean-reverting 4/2 stochastic volatility process as per chapter 3 i.e:

$$\begin{cases} dM_j(t) & = \left[\tilde{L}_j + \tilde{c}_j(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}})^2 - \tilde{\beta}_j M_j(t)\right]dt + (\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}})dW_j^{\mathbb{P}}(t) \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t) \\ \langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t)\rangle & = \rho_j dt \end{cases}$$

$$\tag{5.9}$$

where $\tilde{L}_j = \sum_{i=1}^{n} L_i a_{ij}^*$ and a_{ij}^* are the entries of $\mathbf{A}^{-1}$. $\mathbf{dM_t}$ is also a vector of independent processes.

5.1.1.2 Model with no Spillover Effects

In this section we assume no spillover effects among the assets, i.e matrix **B** is diagonal, which means other assets movements do not affect the mean-reverting level of current asset. This

further simplified our model to:

$$
\begin{cases}
\frac{dX_i(t)}{X_i(t)} &= \left[L_i + \sum_{j=1}^{n} c_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \beta_i ln(X_i(t))\right]dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t), \\
dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \ \ j = 1,\ldots,n
\end{cases}
$$
(5.10)

The corresponding matrix representation has the same form:

$$
\mathbf{dX_t} = diag(\mathbf{X_t})[(\mathbf{L} + \mathbf{CV_t} - \mathbf{B} \cdot \mathbf{ln(X_t)})dt + \mathbf{\Sigma_t^{\frac{1}{2}}dW_t^{\mathbb{P}}}],
$$
(5.11)

with $\mathbf{B} = diag(\beta_1,\ldots,\beta_n)$. The dynamics of log price $Y_i(t) = ln(X_i(t))$ is then:

$$
\begin{cases}
dY_i(t) &= \left[L_i + \sum_{j=1}^{n} a_{ij}c_j\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 - \beta_i Y_i(t)\right]dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^{\mathbb{P}}(t), \\
dv_j(t) &= \alpha_j(\theta_j - v_j(t))dt + \xi_j \sqrt{v_j(t)}dB_j^{\mathbb{P}}(t), \ \ j = 1,\ldots,n.
\end{cases}
$$
(5.12)

5.1.2   Properties of The Variance Vector

We devote this subsection to explore the properties of the variance vector. This is important for understanding the instantaneous volatilities implied by our model. Recall empirical data is related to these volatilities, therefore one should ensure that these implied processes reflect stylized facts of the data they cater to.

Let σ_t^2 denote the variance vector. By definition, we have:

$$
\sigma_t^2 = diag(\mathbf{\Sigma_t}) = (\mathbf{A} \circ \mathbf{A})\mathbf{V_t}.
$$

As defined before, $\mathbf{V_t}$ is a vector of 4/2 processes (the sum of 1/2 and 3/2 processes). Therefore, σ_t^2 can be written in terms of linear combination of these two processes:

$$
\sigma_t^2 =
\begin{pmatrix}
\sum_{j=1} a_{1j}^2 (v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j) \\
\vdots \\
\sum_{j=1} a_{nj}^2 (v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j)
\end{pmatrix}
=
\begin{pmatrix}
\sum_{j=1} a_{1j}^2 v_j(t) \\
\vdots \\
\sum_{j=1} a_{nj}^2 v_j(t)
\end{pmatrix}
+
\begin{pmatrix}
\sum_{j=1} a_{1j}^2 (\frac{b_j^2}{v_j(t)}) \\
\vdots \\
\sum_{j=1} a_{nj}^2 (\frac{b_j^2}{v_j(t)})
\end{pmatrix}
+
\begin{pmatrix}
\sum_{j=1} a_{1j}^2 2b_j \\
\vdots \\
\sum_{j=1} a_{nj}^2 2b_j
\end{pmatrix}
$$

This model for the variance can be interpreted as factor model with n 4/2 factors. Due to the popularity of factor models to explain asset classes, it stands to reason that volatility indexes (these variances) can also be expressed in terms of factors, which could reflect intrinsic and systemic economical movements.

One can obtain the dynamics of σ_t^2 :

$$
\mathbf{d}\sigma_t^2 = \begin{pmatrix} \sum_{j=1} a_{1j}^2 dv_j(t) \\ \vdots \\ \sum_{j=1} a_{nj}^2 dv_j(t) \end{pmatrix} + \begin{pmatrix} \sum_{j=1} a_{1j}^2 (b_j^2 d\frac{1}{v_j(t)}) \\ \vdots \\ \sum_{j=1} a_{nj}^2 (b_j^2 d\frac{1}{v_j(t)}) \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{j=1} a_{1j}^2 \left[\left(\alpha_j(\theta_j - v_j(t)) + \frac{b_j^2}{v_j^2(t)}(\xi_j^2 - \alpha_j\theta_j + \alpha_j v_j(t)) \right) dt + \xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t)) dB_j^{\mathbb{P}}(t) \right] \\ \vdots \\ \sum_{j=1} a_{nj}^2 \left[\left(\alpha_j(\theta_j - v_j(t)) + \frac{b_j^2}{v_j^2(t)}(\xi_j^2 - \alpha_j\theta_j + \alpha_j v_j(t)) \right) dt + \xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t)) dB_j^{\mathbb{P}}(t) \right] \end{pmatrix}
$$

From the above SDE, we are able to obtain the variance and covariance of the vector $\mathbf{d}\sigma_t^2$ via quadratic variations. Note these can be interpreted as the volatility of variance and the correlation among variances respectively:

$$
\langle d\sigma_i^2(t), d\sigma_i^2(t) \rangle = \sum_{j=1}^n a_{ij}^4 \left[\xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t)) \right]^2 dt \tag{5.13}
$$

$$
\langle d\sigma_i^2(t), d\sigma_j^2(t) \rangle = \sum_{k=1}^n a_{ik}^2 a_{jk}^2 \left[\xi_j(\sqrt{v_j(t)} - b_j^2 v_j^{\frac{3}{2}}(t)) \right]^2 dt \tag{5.14}
$$

Equations 5.13 and 5.14 suggest that the instantaneous variance and covariance of σ_t^2 are stochastic in a local way, i.e. driven by the same Brownian as the underlying.

5.2 Characteristic Functions and Approximations.

In this section, we approximate the c.f.'s from Proposition 4.2.2 in line with the approximation approaches of chapter 3. Note that in chapter 3, we obtained analytical approximations of the c.f.'s for the special cases: $\rho = 0, b \neq 0$; $b = 0, \rho \neq 0$ using results from [76]. Solutions for a third case $b = 0, \rho = 0$ can be obtained from either one of previous two special cases. Similarly in this chapter, the two approximation approaches lead to analytical representations for special cases of the generalized mean-reverting 4/2 factor model. Taking advantage of the principal component structure of the model, the representations boil down to a multiplication of the approximations in sections 3.2.1.1 to 3.2.1.3.

In the following, we first show the c.f. for the general model and its sub models described in 5.1.1, namely the model with general spillover effects, the model with separable spillover effects and the model with no spillover effects. Then we present the principle used to approximate the c.f.'s in section 5.2.3, especially under the special scenarios defined by Equations 5.4-5.6, as well as the approximation results.

5.2.1   Characteristic Function for Model with Spillover Effects

Similarly to chapter 4, we first define $\mathbf{Z_t} = e^{\mathbf{B}t}\mathbf{Y_t}$ such that $e^{\mathbf{B}t}$ is a matrix exponential, then $Z_i(t)$ is represented as:

$$dZ_i(t) = \sum_{j=1}^{n} \left(e^{\mathbf{B}t}\right)_{ij} \left\{ L_j + \sum_{k=1}^{n} a_{ij}c_j \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}} \right)^2 \right\} dt + \sum_{j=1}^{n} \left(e^{\mathbf{B}t}\right)_{ij} \left\{ \sum_{k=1}^{n} a_{jk} \left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}} \right) dW_k(t) \right\}$$

(5.15)

For convenience, we use $\left(e^{\mathbf{B}t}\right)_{ij}$ as the ij-th component of the matrix $e^{\mathbf{B}t}$. Note that $Z_i(t)$ is no longer a mean-reverting process although it accounts for time dependent coefficients.

Recall we define $v(t) = \left(\sqrt{V_1(t)}, .., \sqrt{V_n(t)}, \sqrt{\tilde{V}_1(t)}, .., \sqrt{\tilde{V}_n(t)} \right)$ where $\sqrt{V_k(t)} = \sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}$ and $\sqrt{\tilde{V}_k(t)} = \sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}}$, $j, k = 1, \ldots, n$ to denote a vector of 4/2 volatility processes with v_k^v accounting for common factors and $\tilde{v}_j^v(t)$ accounting for intrinsic factors. Since we do not consider intrinsic factors in this chapter, $v(t) = (v_1^v(t), .., v_n^v(t))$. The conditional c.f. for the increments of $\mathbf{Z(t)}$ follows from Proposition 4.2.2:

Corollary 5.2.1 *Let $(\mathbf{Z(t)})_{t\geq 0}$ evolve according to the model in Equation* (5.15). *The c.f.* $\Phi_{\mathbf{Z(t),v(t)}}$ *is then given as follows:*

$$\Phi_{\mathbf{Z(t)},v_t}(T, \omega) = E\left[\exp i\omega'(\mathbf{Z(T)} - \mathbf{Z(t)}) \mid \mathbf{Z(t)} = z_t, v(t) = \mathbf{v_t}\right]$$

$$= \prod_{k=1}^{n} \Phi_{GG}\left(T, 1; L(\omega), h_k(\omega), g_k(\omega), \alpha_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, Z(t)\right)$$

where $h_k(\omega, t) = \sum_{j=1}^{n} a_{jk}c_k f_j(\omega, t)$, $L(\omega, t) = \sum_{j=1}^{n} \frac{L_j}{n} f_j(\omega, t)$, $g_k(\omega, t) = \sum_{j=1}^{n} a_{jk} f_j(\omega, t)$ and $f_j(\omega, t) = \sum_{m=1}^{n} i\omega_m \left(e^{\beta t}\right)_{mj}$. Φ_{GG} is a one-dimensional generalization of the c.f. from [76] *provided in Lemma B.1.1.*

The proof follows as a direct application of the proof of Proposition 4.2.2.

5.2.2   Characteristic Function for Model with Separable Spillover Effects

In order to derive the c.f. of $M_i(t)$, we do the transformation $S_j(t) = e^{\tilde{\beta}_j t} M_j(t)$, recognizing that the c.f. of $S_j(t)$ has been derived in chapter 3. Hence, the c.f. of $\mathbf{Y_t}$ is just a product of the corresponding c.f. of $S_j(t)$. The result is summarized in the following corollary.

Corollary 5.2.2 *Let $\Phi_{MR}\left(T, u; L, c, a, \beta, \alpha, \theta, \xi, \rho, b, v(t), Z_(t)\right)$ denote the characteristic function provided in Proposition 3.2.1, then the characteristic function of $\mathbf{Y_t}$ is given by the following equation:*

$$\mathbb{E}(e^{i\mathbf{u}'\mathbf{Y_T}}|\mathcal{F}_t) = \prod_{j=1}^{n} \Phi_{MR}\left(T, \mathbf{u}^*; \tilde{L}_j, \tilde{c}_j, 1, \tilde{\beta}_j, \alpha_j, \theta_j, \xi_j, \rho_j, b_j, v_j(t), S_j(t)\right)$$

(5.16)

where $\mathbf{u}^*$ is a new vector of real numbers with element $u_j^* = \sum_{i=1} u_i a_{ij} e^{-\tilde{\beta}_j t}$.

The proof is straightforward using the relationship $\mathbf{Y_t} = \mathbf{AM_t}$, we know that each individual process $Y_i(t)$ is a linear combination of $M_j(t)$, $j = 1, 2, \ldots, n$, and therefore $S_j(t)$, processes. The product $\mathbf{u'Y_T}$ can be further written in terms of $S_j(t)$:

$$\mathbf{u'Y_T} = \mathbf{u'AM_T} = \sum_{j=1}^{n}\sum_{i=1}^{n} u_i a_{ij} M_j(T) = \sum_{j=1}^{n}\sum_{i=1}^{n} u_i a_{ij} e^{-\tilde{\beta}_j t} S_j(T) = \sum_{j=1}^{n} u_j^* S_j(T).$$

The independence property of random variables leads to Equation 5.16.

5.2.3 Characteristic Function for Models with no Spillover Effects

The process $\mathbf{Z_t} = e^{\mathbf{B}t}\mathbf{Y_t}$ has a less complicated representation if B were diagonal, in this case the matrix exponential $e^{\mathbf{B}t}$ is a diagonal matrix with diagonal entries $(e^{\beta_1 t}, \ldots, e^{\beta_n t})$. In such case, $Z_i(t)$ can be represented as:

$$dZ_m(t) = e^{\beta_m t}\left\{ L_m + \sum_{j=1}^{n} a_{mj} c_j \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right)^2 \right\} dt + e^{\beta_m t}\left\{ \sum_{j=1}^{n} a_{mj} \left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}} \right) dW_j(t) \right\}$$

$$(5.17)$$

The characteristic function for the increment $Z(T) - Z(t)$ is provided in the next corollary:

Corollary 5.2.3 *Let $(Z(t))_{t\geq 0}$ evolve according to the model in Equation* (5.17). *The c.f.* $\Phi_{Z_i(t),\mathbf{v(t)}}$ *is then given as follows:*

$$\Phi_{Z(t),v_t}(T, \omega) = E\left[\exp i\omega'(Z(T) - Z(t)) \mid Z(t) = z_t, v(t) = v_\mathbf{t} \right]$$

$$= \prod_{j=1}^{n} \Phi_{GG}\left(T, 1; L(\omega), h_j(\omega), g_j(\omega), \alpha_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,t}, Z(t) \right)$$

where Φ_{GG} is provided in Lemma B.1.1. $L(\omega, t) = \sum_{j=1}^{n} \frac{i\omega_j e^{\beta_j t} L_j}{n}$, $h_j(\omega, t) = \sum_{m=1}^{n} i\omega_m e^{\beta_m t} a_{mj} c_j$ *and*

$g_j(\omega, t) = \sum_{m=1}^{n} i\omega_m e^{\beta_m t} a_{mj}$

Proof follows similarly to the proof of Proposition 4.2.2.

5.2.4 Approximation Principle and Results

From above, we have learned that the c.f. in both categories (the model with spillover effects and the model without spillover effects) can be written in terms of a product of the c.f.'s of n independent one dimensional processes thanks to principal component decomposition. These one dimensional processes defined in Equation 5.15 and Equation 5.17 are only different in the structure of the matrix exponential term i.e $e^{\mathbf{B}t}$, which is deterministic, and they resemble the same $Z(t)$ process we studied in chapter 3. Therefore, the principles to approximate $\Phi_{Z(t),v_t}(T, \omega)$ follow those we adopted in chapter 3. In other words, we only need to work out an approximation to the individual c.f. Φ_{GG}, and the approximation can be realized under three scenarios as we have shown in section 5.1.1: $b_j = 0, \rho_j \neq 0$; $b_j \neq 0, \rho_j = 0$ and $b_j = 0, \rho_j = 0$.

Recall that we give the formula for Φ_{GG} in Lemma B.1.1 as follows:

$$\Phi_{GG}\left(T, u; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_t, Z_t\right) = \exp\left\{iu \int_t^T A\left(s\right) ds\right\} v(t)^{iu\frac{b\rho}{\xi} g(t)} \exp\left\{-iu\rho \frac{g\left(t\right) v(t)}{\xi}\right\}$$

$$\times \mathbb{E}\left[v(T)^{iu\frac{b\rho}{\xi} g(T)} \exp\left\{iu\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds\right.\right.\right.$$

$$\left.\left.\left. +\rho\frac{g\left(T\right) v(T)}{\xi}\right)\right\} \mid \mathcal{F}_t\right].$$

As we pointed out in chapter 4, Φ_{GG} can not be solved in closed-form for a core reason that can be traced down to simpler 1/2 models, namely the lack of a representation for the moment generating function of an integrated CIR process with time dependent integrands. Therefore, we propose an analytic function that approximates Φ_{GG}. Recall that we studied two cases where the c.f. can be approximated by analytic functions in chapter 3; here Φ_{GG} has the same structure as the c.f. we studied in one dimension, so we can approximate Φ_{GG} in a similar fashion. Since the unsolvable conditional expectation brings all the complexity, we approximate the c.f. by approximating the conditional expectation. In this regard, following the same principle, we come up with two approximation approaches. Under certain conditions, as we will see in Corollary 5.2.4, the conditional expectation in Φ_{GG} is approximated in the following form by a conditional expectation whose analytic representation is known in the literature:

$$\mathbb{E}\left[v(T)^{\frac{b\rho}{\xi} g(T)} \exp\left\{\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds + \rho\frac{g\left(T\right) v(T)}{\xi}\right)\right\} \mid \mathcal{F}_t\right]$$

$$\approx \mathbb{E}\left[v(T)^{\frac{b\rho}{\xi} g(T)} \exp\left\{\rho\frac{g\left(T\right) v(T)}{\xi} - n\int_t^T v(s)ds - m\int_t^T \frac{1}{v(s)}ds\right\} \mid \mathcal{F}_t\right],$$

for some complex constants m and n:

$$m \approx -\int_t^T C(s)ds,$$

$$n \approx -\int_t^T B(s)ds,$$

and,

$$D(s) = 0,$$

for $s \in [t, T]$. Recall that in chapter 3, we obtain m and n via two approximation approaches for the integrals $\int_t^T C(s)ds$ and $\int_t^T B(s)ds$: midpoint and average. We briefly cover the results, see chapter 3 for details:

- **Midpoint:**

$$m = \frac{\min\limits_{[t,T]}(C(s)) + \max\limits_{[t,T]}(C(s))}{2}, n = \frac{\min\limits_{[t,T]}(B(s)) + \max\limits_{[t,T]}(B(s))}{2}.$$

- **Average:**

$$m = \frac{1}{T-t} \int_t^T C(s)ds, \, n = \frac{1}{T-t} \int_t^T B(s)ds.$$

The approximated conditional expectation is solvable as shown in [76]. We summarize the results in the following corollary for the general model with spillover effects. Since the model with separable spillover effects and the model without spillover effects are special cases of the general model, we can easily apply the results to the special cases. We leave some comments after next corollary on approximation procedures for the special cases.

Corollary 5.2.4 *Given deterministic functions $B_j(s)$ and $C_j(s)$ defined in Lemma B.1.1 and $g_j(s)$ defined in Corollary 5.2.1 for $s \in [t, T]$,*

$$\mathbb{E}\left[v_j(T)^{\frac{b_j\rho_j}{\xi_j}g_j(T)} \exp\left\{ \int_t^T B_j(s)v_j(s)ds + \int_t^T C_j(s)\frac{1}{v_j(s)}ds + \int_t^T D_j(s)\ln(v_j(s))ds \right.\right.$$
$$\left.\left. + \rho_j\frac{g_j(T)\,v_j(T)}{\xi_j}\right\} \mid \mathcal{F}_t\right]$$

can be approximated by analytic functions for constants m_j and n_j satisfying

$$n_j = \begin{cases} \frac{\min_{[t,T]}(B_j(s))+\max_{[t,T]}(B_j(s))}{2}, & \text{if Midpoint approach is considered} \\ \frac{1}{T-t}\int_t^T B_j(s)ds, & \text{if Average Approach is considered} \end{cases}$$

$$m_j = \begin{cases} \frac{\min_{[t,T]}(C_j(s))+\max_{[t,T]}(C_j(s))}{2}, & \text{if Midpoint approach is considered} \\ \frac{1}{T-t}\int_t^T C_j(s)ds, & \text{if Average Approach is considered} \end{cases}$$

under three scenarios:

- $\mathbf{b_j = 0, \rho_j \neq 0}$: *Given $b_j = 0, \rho_j \neq 0$, $C(s) = 0$ and $D(s) = 0$, $s \in [t, T]$. If $n_j \geq -\frac{\alpha_j^2}{2\xi_j^2}$, then the conditional expectation is approximated by:*

$$\mathbb{E}\left[\exp\left\{ \int_t^T B_j(s)v_j(s)ds + \rho_j\frac{g_j(T)\,v_j(T)}{\xi_j}\right\} \mid \mathcal{F}_t\right]$$
$$\approx \mathbb{E}\left[\exp\left\{ \rho_j\frac{g_j(T)\,v_j(T)}{\xi_j} - n_j\int_t^T v_j(s)ds\right\} \mid \mathcal{F}_t\right]$$
$$= \left(\frac{(B_j\xi_j^2 + \alpha)(e^{\sqrt{A_j}(T-t)} - 1) + \sqrt{A_j}(e^{\sqrt{A_j}(T-t)} + 1)}{2\sqrt{A_j}e^{\frac{\sqrt{A_j}+\alpha_j}{2}(T-t)}}\right)^{-\frac{2\alpha_j\theta_j}{\xi_j^2}} e^{v_j(t)\left(\frac{(B_j\alpha_j-2n_j)\left(e^{\sqrt{A_j}(T-t)}-1\right)-B_j\sqrt{A_j}\left(e^{\sqrt{A_j}(T-t)}+1\right)}{(B_j\xi_j^2+\alpha_j)\left(e^{\sqrt{A_j}(T-t)}-1\right)+\sqrt{A_j}\left(e^{\sqrt{A_j}(T-t)}+1\right)}\right)},$$

$$B_j = -\frac{\rho_jg_j(T)}{\xi_j}, A_j = \alpha_j^2 + 2n_j\xi_j^2,$$

$$(5.18)$$

- $\mathbf{b_j \neq 0, \rho_j = 0}$: *Given $b_j = 0, \rho_j \neq 0$ and $D(s) = 0$, $s \in [t, T]$. If $m_j > -\frac{(2\alpha_j\theta_j - \xi_j^2)^2}{8\xi_j^2}, n_j \geq -\frac{\alpha_j^2}{2\xi_j^2}$, then the conditional expectation is approximated by:*

$$
\mathbb{E}\left[\exp\left\{ \int_t^T B_j(s)v_j(s)ds + \int_t^T C_j(s)\frac{1}{v_j(s)}ds \right\} \,|\, \mathcal{F}_t \right]
$$

$$
\approx \mathbb{E}\left[\exp\left\{ -n_j \int_t^T v_j(s)ds - m_j \int_t^T \frac{1}{v_j(s)}ds \right\} \,|\, \mathcal{F}_t \right]
$$

$$
= \left(\frac{\gamma_j(T, v_j(t))}{2} \right)^{k_j+1} v_j(t)^{-\frac{\alpha_j\theta_j}{\xi_j^2}} K_j(T)^{-\left(\frac{1}{2}+\frac{k_j}{2}+\frac{\alpha_j\theta_j}{\xi_j^2}\right)} e^{\frac{1}{\xi_j^2}\left(\theta_j(T-t) - \sqrt{H_j}v_j(t)\coth\left(\frac{\sqrt{H_j}(T-t)}{2}\right) + \alpha_j v_j(t)\right)} \frac{\Gamma\left(\frac{1}{2} + \frac{k_j}{2} + \frac{\alpha_j\theta_j}{\xi_j^2}\right)}{\Gamma(k_j + 1)}
$$

$$
\times {}_1F_1\left(\frac{1}{2} + \frac{k_j}{2} + \frac{\alpha_j\theta_j}{\xi_j^2}, k_j + 1, \frac{\gamma_j(T, v_j(t))^2}{4K_j(T)} \right),
$$

$$
k_j = \frac{1}{\xi_j^2}\sqrt{(2\alpha_j\theta_j - \xi_j^2)^2 + 8m_j\xi_j^2}, \quad H_j = \alpha_j^2 + 2n_j\xi_j^2, \quad \gamma_j(T, v_j(t)) = \frac{2\sqrt{H_j v_j(t)}}{\xi_j^2 \sinh\left(\frac{\sqrt{H_j}(T-t)}{2}\right)},
$$

$$
K_j(T) = \frac{1}{\xi_j^2}\left(\sqrt{H_j}v_j(t)\coth\left(\frac{\sqrt{H_j}(T-t)}{2}\right) + \alpha_j \right).
$$

$$(5.19)$$

- $\mathbf{b_j = 0, \rho_j = 0}$: *Given $b_j = 0, \rho_j = 0$, $C(t) = 0$ and $D(t) = 0$, $s \in [t, T]$. If $n_j \geq -\frac{\alpha_j^2}{2\xi_j^2}$, then the conditional expectation is approximated by:*

$$
\mathbb{E}\left[\exp\left\{ \int_t^T B_j(s)v_j(s)ds \right\} \,|\, \mathcal{F}_t \right]
$$

$$
\approx \mathbb{E}\left[\exp\left\{ -n_j \int_t^T v_j(s)ds \right\} \,|\, \mathcal{F}_t \right]
$$

$$
= \left(\frac{\alpha_j(e^{\sqrt{A_j}(T-t)} - 1) + \sqrt{A_j}(e^{\sqrt{A_j}(T-t)} + 1)}{2\sqrt{A_j}e^{\frac{\sqrt{A_j}+\alpha_j}{2}(T-t)}} \right)^{-\frac{2\alpha_j\theta_j}{\xi_j^2}} \times e^{v_j(t)\left(\frac{-2n_j(e^{\sqrt{A_j}(T-t)}-1)}{\alpha_j(e^{\sqrt{A_j}(T-t)}-1)+\sqrt{A_j}(e^{\sqrt{A_j}(T-t)}+1)} \right)},
$$

$$
A_j = \alpha_j^2 + 2n_j\xi_j^2.
$$

$$(5.20)$$

Corollary 5.2.4 follows directly from Proposition 3.2.2 and Proposition 3.2.3. The approximation for the c.f. when there are no spillover effects follows the same procedure as presented in Corollary 5.2.4. For the case when the spillover effects are separable, note that after the decomposition, we end up with a series of independent mean-reverting 4/2 stochastic volatility processes as shown in Formula 5.9, for which we have studied complete approximation procedure in chapter 3. As a result, Proposition 3.2.2 and Proposition 3.2.3 can be directly applied to approximate the c.f.s for these processes.

5.3   Estimation

In this section, we consider the model with separable spillover effects as the underlying model for estimation. This way we fulfill the purpose of studying spillover effects among assets, and on the other hand we avoid the complexity of matrix exponential. Recall the model with separable spillover effects can be expressed as follows:

$$\mathbf{dY_t} = (\mathbf{L} + \mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t} - \mathbf{BY_t})dt + \mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

where $\mathbf{B}$ is constructed in a way such that it can be decomposed into a product of three matrices: $\mathbf{B} = \mathbf{A}\tilde{\mathbf{B}}\mathbf{A}^{-1}$. In this case, $dY_i(t)$ is a linear combination of independent processes $dM_j(t)$, i.e $d\mathbf{Y_t} = \mathbf{A}d\mathbf{M_t}$:

$$\begin{cases} dM_j(t) & = \left[\tilde{L}_j + \tilde{c}_j(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}})^2 - \tilde{\beta}_j M_j(t)\right]dt + (\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}})dW_j^{\mathbb{P}}(t) \\ dv_j(t) & = \alpha_j(\theta_j - v_j(t))dt + \xi_j\sqrt{v_j(t)}dB_j^{\mathbb{P}}(t) \\ \langle dW_j^{\mathbb{P}}(t), dB_j^{\mathbb{P}}(t)\rangle & = \rho_j dt \end{cases}$$

For simplicity, we focus on two dimensions hence studying pairs of assets with their respective volatility indexes. For example, VIX (VVIX) and VSTOXX (VVSTOXX); USO (OVX) and GLD (GVZ), or USO (OVX) and SLV (VXSLV). Then we follow the same estimation procedure as outlined in chapter 3 splitting the parameters into two groups: volatility group and drift group.

After a section on data description in 5.3.1, section 5.3.2 estimates the parameters in the volatility group. Under the model with separable spillover effects, we first need to estimate covariance matrix $(\widehat{\boldsymbol{\Sigma}})$ from asset data as a long term average of the stochastic covariance matrix $(\boldsymbol{\Sigma_t})$. This permits us to produce and estimate for the constant eigenvectors, denoted as $\hat{\mathbf{A}}$. With the estimated eigenvectors, we decompose our original asset processes into sum of independent mean-reverting 4/2 models. The volatility group then consists of parameters for the underlying CIR processes driving the principal components: $b_j, \alpha_j, \theta_j, \xi_j$, and the estimation method follows along the lines of chapter 3. Section 5.3.3 tackles the estimation of drift group parameters. This is done using least squares as in chapter 3.

Here we would like to remind the reader about volatility indexes and their meaning before we proceed further. As we discussed in chapter 1, volatility indexes are a kind of implied volatility that is model free and directly calculated from option prices from the market, thus they are a good measure of market's expectation about future volatility. See chapter 1 for a discussion on the relation among four broadly used volatility measures: realized volatility, implied volatility, instantaneous volatility and integrated volatility, and how tightly connected they are. In chapters 3 and 5, we use volatility indexes data as a convenient proxy for instantaneous volatility. In fact, instantaneous volatility is rather impossible to capture from empirical data, even with high frequency data, as it requires instantaneous periods rather than the available discrete periods. On the other hand, once a model is specified, volatility indexes can be used to represent instantaneous volatility with some multiplicative (scaling) adjustment or factor, see eg. [102], [130]

and references therein. These works fall under "indirect modelling" methods for volatility indexes and assume either Heston's model or Heston with jumps for the underlying asset process. In these Heston-related modelling methodologies, the relationship between the instantaneous volatility and the volatility index of interest, for example VIX, can be expressed in terms of a closed-form equation, where the difference between the two, under some conditions, lies on a multiplicative factor. In a recent paper, see [97], the author assumes Grasselli's 4/2 model with the addition of jumps as the underlying, obtaining a far more complicated relationship, nonetheless still with a multiplicative factor. All these works demonstrate that due to the short horizon of volatility indexes (21 day options), the multiplicative factor could be close to one in a large region of the parametric space, which implies that volatility indexes are almost equal to instantaneous volatilities regardless of the structural choice of underlying model.

Inspired by these pioneer works, we introduce scale parameters to adjust empirical volatility indexes data to estimate instantaneous volatilities. This is done in a way that the empirical means of the observed variance series ($\mathbf{V_t^{obs}}$) match the corresponding long term asset variances. These new scaling parameters can be estimated at an early stage of the estimation process and are methodologically independent of other parameters. The reason we did not introduce this scaling parameter earlier in chapter 3 is that in a one-dimensional case, their values would be absorbed by the mean-revering level of the underlying CIR process; in other words, the final estimation results already reflect the scaling operation. In this chapter, due to the multiple dimensions and the principal components decomposition, it turns out that the information of the scaling could be absorbed by either the eigenvectors (denoted A), the mean reverting level (MRL) of the principal components (i.e. θ) or both. This creates scenarios where the model is not compatible with data, e.g. given A, there is no reasonable θ matching data-implied MRLs. This can be fixed by acknowledging scaling parameters for every underlying CIR process. Hence, the estimation introduces the concept of "scaling" to accommodate the instantaneous covariances with the vector of observable volatility indexes.

5.3.1   Data Description

Before we officially engage in estimation procedures with empirical data, we devote this section to briefly describe the data we use and how we process the data for our purpose.

Generally, the model with separable spillover effects is good for multiple assets. As a demonstration, we consider the following pairs of assets: VIX and VSTOXX; USO and GLD; SLV and GLD. Recall from chapter 3 that our estimation method is a two-stage method due to the structure of our model. A two-stage method means that we need the asset data as well as its corresponding volatility index data, then we estimate a group of the parameters in the first stage with volatility index data and the other group of parameters next with asset data and volatility index data. Therefore, we also consider the related volatility indexes (VVIX and VVSTOXX, OVX and GVZ, VXSLV and GVZ) for the corresponding assets.

The sample size of the raw data is different across all the assets and volatility indexes. Hence, we have to further process the data to better suit our estimation purpose, in particular ensuring we take only the trading days when both assets and their volatilities can be observed. Fig-

ure 5.1 to Figure 5.3 show the pairs of asset data and their volatility indexed. Note that the volatility index data is quoted as annualized volatility multiplied by 100. When we use the volatility index for estimation, we transform the volatility index to daily volatility by dividing by $100 * \sqrt{250}$.
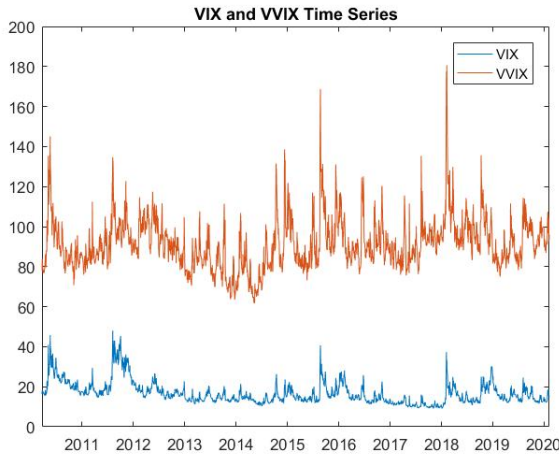
(a) Data: VIX and VVIX
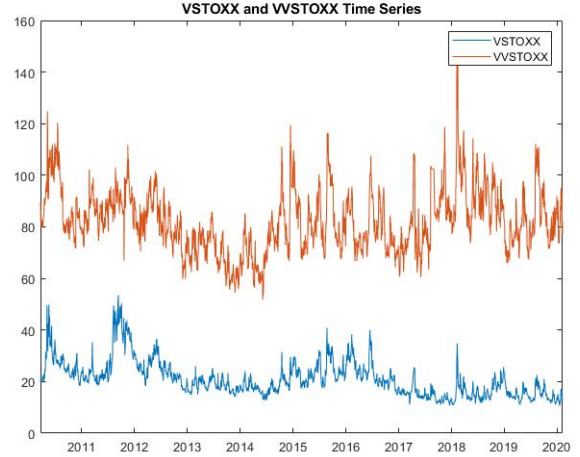
(b) Data: VSTOXX and VVSTOXX

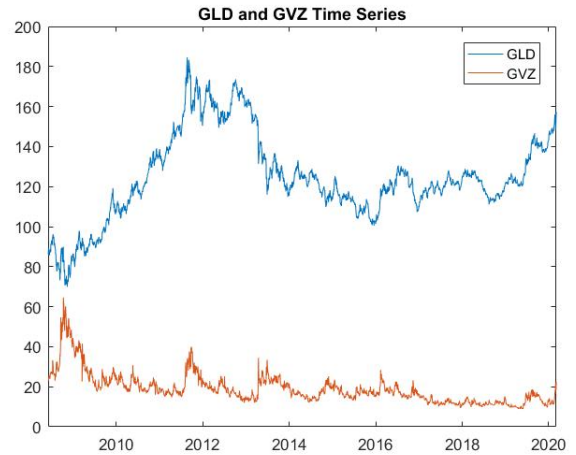Figure 5.1: Historical VIX (VVIX) and VSTOXX (VVSTOXX) Data
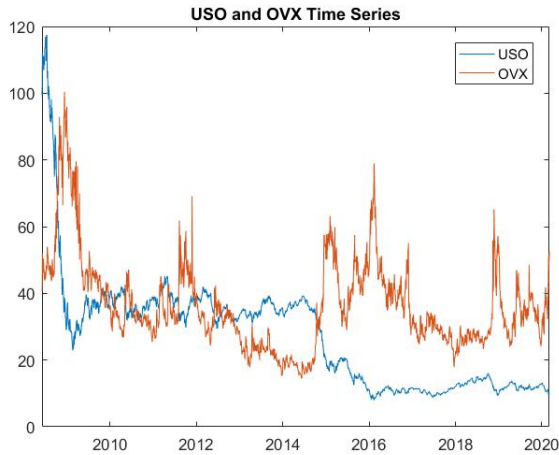
(a) Data: USO and OVX

(b) Data: GLD and GVZ

Figure 5.2: Historical USO (OVX) and GLD (GVZ) Data

5.3.2   Estimation of Volatility Group parameters.

We first take the chance to emphasize that there are two possible models that can be used for estimation of the "volatility group" parameters in this section: the model with separable

(**a**) Data: SLV and VXSLV
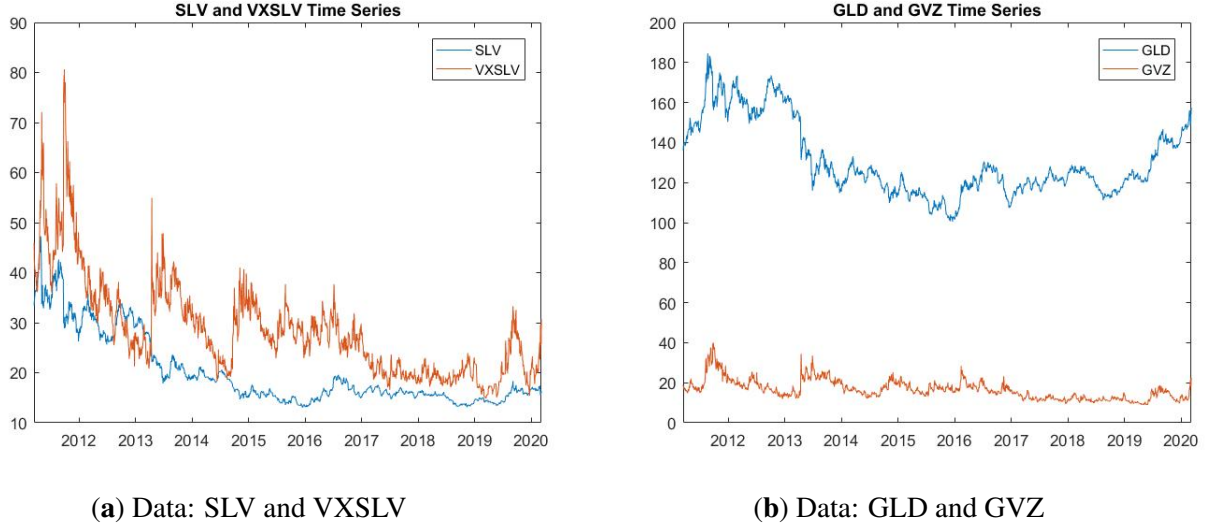

(**b**) Data: GLD and GVZ

Figure 5.3: Historical SLV (VXSLV) and GLD (GVZ) Data

spillover effects and without spillover effects. The former is the model we consider in this section. Both models are described next for completeness.

- **Model with Separable Spillover Effects**

  Recall that Equation 5.9 gives us the representation for each principal component that reflects on our mean-reverting 4/2 stochastic volatility model, this is $dlog(\mathbf{X_t}) = d\mathbf{Y_t} = \mathbf{A}d\mathbf{M_t}$ with $dM_i(t)$ defined as:

$$
\begin{cases}
dM_i(t) & = \left[\tilde{L}_i + \tilde{c}_i\left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}\right)^2 - \tilde{\beta}_i M_i(t)\right]dt + \left(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}\right)dW_i^{\mathbb{P}}(t) \\
dv_i(t) & = \alpha_i(\theta_i - v_i(t))dt + \xi_i\sqrt{v_i(t)}dB_i^{\mathbb{P}}(t) \\
\langle dW_i^{\mathbb{P}}(t), dB_i^{\mathbb{P}}(t)\rangle & = \rho_i dt
\end{cases}
$$

  The estimation procedure for this model setup can be summarized as we first transform the data using matrix $\mathbf{A}$ to produce $M_i(t)$ process following the relationship among $\mathbf{X_t}$, $\mathbf{Y_t}$ and $\mathbf{M_t}$; then we can use the estimation method developed in chapter 3 for each $M_i(t)$ process, finally we recover the parameters for each $X_i(t)$ process.

The model without spillover effects, defined by Equation 5.17, is simpler in setup because it does not involve spillover effects, thus the matrix $\mathbf{B}$ is a diagonal matrix with mean-reverting speed parameters for $\mathbf{X_t}$ processes on the diagonal. We consider the transformation $\mathbf{Z_t} = e^{\mathbf{B}t}\mathbf{Y_t} = e^{\mathbf{B}t}log(\mathbf{X_t})$. We can see from this relationship that in order to obtain $\mathbf{Z_t}$ we need the matrix $\mathbf{B}$ available, which is difficult to estimate at an early stage of the estimation process. Assuming $\mathbf{B}$ can be estimated at the beginning of the estimation process, we can proceed in the following procedures.

- **Model without Spillover Effects**

  We write Equation 5.17 in matrix form:

$$\mathbf{dZ_t} = e^{\mathbf{B}t}(\mathbf{L} + \mathbf{A}\tilde{\mathbf{C}}\mathbf{V_t})dt + e^{\mathbf{B}t}\mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \tag{5.21}$$

Multiplying both sides by $\mathbf{A}^{-1}$, we get:

$$\mathbf{A}^{-1}\mathbf{dZ_t} = e^{\mathbf{B}t}(\mathbf{A}^{-1}\mathbf{L} + \tilde{\mathbf{C}}\mathbf{V_t})dt + e^{\mathbf{B}t}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}. \tag{5.22}$$

Similarly, we take the transformation $\tilde{\mathbf{Z}}_\mathbf{t} = \mathbf{A}^{-1}\mathbf{Z_t}$, then we obtain dynamics for process $\tilde{\mathbf{Z}}_\mathbf{t}$ as follows:

$$\mathbf{d\tilde{Z}_t} = e^{\mathbf{B}t}(\mathbf{A}^{-1}\mathbf{L} + \tilde{\mathbf{C}}\mathbf{V_t})dt + e^{\mathbf{B}t}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}, \tag{5.23}$$

which indicates $\tilde{\mathbf{Z}}_\mathbf{t}$ is a vector of independent processes in the form:

$$\begin{cases} d\tilde{Z}_i(t) & = e^{\beta_i t}\left[\tilde{L}_i + \tilde{c}_i(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})^2\right]dt + e^{\beta_i t}(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})dW_i^{\mathbb{P}}(t) \\ dv_i(t) & = \alpha_i(\theta_i - v_i(t))dt + \xi_i\sqrt{v_i(t)}dB_i^{\mathbb{P}}(t) \\ \langle dW_i^{\mathbb{P}}(t), dB_i^{\mathbb{P}}(t)\rangle & = \rho_i dt \end{cases}$$

The key in this second model is to decompose our original process into a linear combination of independent mean-reverting 4/2 processes, so that we can apply the same estimation methodology developed in chapter 3 to each $\tilde{Z}_i(t)$ process. We can see that the matrix $\mathbf{B}$ is required for the decomposition, which represents a challenge and requires a modification to our chapter 3 estimation methodology. Recall that in chapter 3, our estimation considers log price. With no spillover effects, we have to work with $\mathbf{dZ_t}$ as it leads to linear combination of independent 4/2 processes to which we can apply our estimation methodology.

5.3.2.1   Estimation of Matrix A and the Scaling Parameters S.

In this section and the following section, we perform a preliminary analysis prior to estimating parameters in volatility group. Some metrics from this analysis play key roles in our estimation process. The empirical results are summarized in Table 5.1.

The first step is to estimate matrix $\mathbf{A}$ as it connects log asset prices $\mathbf{Y_t}$ and principal components $\mathbf{M_t}$. Recall that $\mathbf{A}$ is an orthogonal matrix comprising the eigenvectors of covariance matrix $\mathbf{\Sigma_t}$. Given daily data, we estimate $\mathbf{A}$ by first calculating the covariance matrix $\widehat{\mathbf{\Sigma}}$ and applying eigenvalue decomposition:

$$\widehat{\mathbf{\Sigma}} = \hat{\mathbf{A}}diag(\sigma^{(1)}, \ldots, \sigma^{(n)})\hat{\mathbf{A}}'.$$

$(\sigma^{(1)}, \ldots, \sigma^{(n)})$ is a vector of eigenvalues of $\widehat{\mathbf{\Sigma}}$ and $\hat{\mathbf{A}}$ is the estimate of matrix $\mathbf{A}$. In Table 5.1, we include the results for $\widehat{\mathbf{\Sigma}}$, $\hat{\mathbf{A}}$ and eigenvalues $(\sigma^{(1)}, \sigma^{(2)})$ from empirical data. Note that $\hat{\mathbf{A}}$ is not unique in that the signs of each element in the matrix can be manipulated such that the column vectors are still the eigenvectors for the corresponding eigenvalues, while $\hat{\mathbf{A}}$ preserve its orthogonality.

As we mentioned at the beginning of this section, volatility indexes are good proxy for instantaneous volatility, but it may require scaling adjustment. Let $\mathbf{V_t^{obs}} = (\mathbf{V_t^{(1)}}, \ldots, \mathbf{V_t^{(n)}})'$ denote the squared observed volatility indexes data for n assets respectively, we introduce a scale parameter s_i to bridge observed volatility indexes series $\mathbf{V_t^{(i)}}$'s to theoretical variances via the following relationship:

$$\mathbf{SV_t^{obs}} = (\mathbf{A} \circ \mathbf{A})\mathbf{V_t} \tag{5.24}$$

$$\mathbf{V_t} = (\mathbf{A} \circ \mathbf{A})^{-1}\mathbf{S}\mathbf{V_t^{obs}}. \tag{5.25}$$

where $\mathbf{S}$ is a diagonal matrix with diagonal vector $(s_1, \ldots, s_n)$. In theory, s_i is a function of $t, T, b_i, \alpha_i, \theta_i, \xi_i$, see [102], [130], some of which fall into volatility group and are to be estimated. Hence, it is crucial to come up with an estimate that does not depend on these parameters. We propose an estimate that matches the long run first empirical moment of both sides of Equation 5.24. The long term average of the left-hand side of Equation 5.24 can be directly calculated from squared volatility indexes data. The long term average of the right-hand side may not seem as straight forward as it is a stochastic process. In our definition $(\mathbf{A} \circ \mathbf{A})\mathbf{V_t}$ is the diagonal of covariance matrix and refers to instantaneous variance process. So in the long run, the expectation of the variance process should converge to the variance of the underlying asset. Let $\hat{\sigma}_i^2$ denote the empirical long term variance of asset i and $\hat{\mu}_i$ denote the long term average of the corresponding squared volatility index data. We estimate s_i as:

$$\hat{s}_i = \frac{\hat{\sigma}_i^2}{\hat{\mu}_i}. \tag{5.26}$$

Substituting Equation 5.26 and $\hat{\mathbf{A}}$ back to the right-hand side of Equation 5.25, then the long term average gives us:

$$\begin{pmatrix} \sigma^{(1)} \\ \vdots \\ \sigma^{(n)} \end{pmatrix} = \mathbb{E}\left[(\mathbf{A} \circ \mathbf{A})^{-1}\hat{\mathbf{S}}\mathbf{V_t^{obs}}\right],$$

which matches the left-hand side of Equation 5.25 in long term average as $\mathbf{V_t}$ is the eigenvalue of $\mathbf{\Sigma}_t$ and converges to $(\sigma^{(1)}, \ldots, \sigma^{(n)})$ in the long run. Table 5.1 also shows the results for $\hat{\mathbf{\Sigma}}$, $(\sigma^{(1)}, \sigma^{(2)})$ and $(\hat{s}_1, \hat{s}_2)$.

	VIX&VSTOXX	USO&GLD	SLV&GLD
$\widehat{\mathbf{\Sigma}}$	$\begin{pmatrix} 0.0062 & 0.0030 \\ 0.0030 & 0.0048 \end{pmatrix}$	$\begin{pmatrix} 4.779e-4 & 4.616e-5 \\ 4.616e-5 & 1.2547e-5 \end{pmatrix}$	$\begin{pmatrix} 0.0724 & 0.0331 \\ 0.0331 & 0.0237 \end{pmatrix}$
$(\hat{\mu}_1, \hat{\mu}_2)$	$(0.0034, 0.0028)$	$(6.0109e\text{-}4, 1.6689e\text{-}4)$	$(3.6258e\text{-}4, 1.1843e\text{-}4)$

(a): Covariance Matrix and Long Term Average of Squared Volatility Indexes

	VIX&VSTOXX	USO&GLD	SLV&GLD
$\hat{\mathbf{A}}$	$\begin{pmatrix} 0.7825 & -0.6226 \\ 0.6226 & 0.7825 \end{pmatrix}$	$\begin{pmatrix} 0.9918 & -0.1278 \\ 0.1278 & 0.9918 \end{pmatrix}$	$\begin{pmatrix} 0.8925 & -0.451 \\ 0.451 & 0.8925 \end{pmatrix}$
$\dot{\mathbf{A}}$	$\begin{pmatrix} 2.725 & -1.725 \\ -1.725 & 2.725 \end{pmatrix}$	$\begin{pmatrix} 1.0169 & -0.0169 \\ -0.0169 & 1.0169 \end{pmatrix}$	$\begin{pmatrix} 1.3429 & -0.3429 \\ -0.3429 & 1.3429 \end{pmatrix}$
$(\sigma^{(1)}, \sigma^{(2)})$	$(2.1372, 0.592)$	$(0.121, 0.0299)$	$(0.0891, 0.007)$
$(\hat{s}_1, \hat{s}_2)$	$(1.8074, 1.7229)$	$(0.795, 0.7518)$	$(0.799, 0.8)$

(b): Eigenvectors, Eigenvalues and Scaling Factors

Table 5.1: Empirical Results

5.3.2.2   Estimation of Volatility Group

Now we apply the estimation methods we developed in chapter 3. In chapter 3, we work with 1 dimensional process, and the volatility process $(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})$ is mapped to the observed volatility for the corresponding asset (eg. VVIX for VIX and VSTOXX for VSTOXX). In a multidimensional setting, the volatility process is driven by more than one diffusion, but the idea is the same.

Let $(\hat{\mathbf{A}} \circ \hat{\mathbf{A}})^{-1} = \dot{\mathbf{A}} = \{\dot{a}_{ij}\}_{i,j=1}^{n}$ and j-th eigenvalue of $\boldsymbol{\Sigma_t}$ be defined as $V_j(t) = v_j(t) + \frac{b_j^2}{v_j(t)} + 2b_j$, then $V_j(t)$ is represented by $\sum_{i=1}^{n} \dot{a}_{ji}\hat{s}_i \mathbf{V_t^{(i)}}$ according to Equation 5.25. Suppose $\mathbf{V_t^{(i)}} = (V^{(i)}(t_1), \ldots, V^{(i)}(t_n))$ is a series of squared volatility indexes for asset i observed on $\Omega_T = \{t_i\}_{i=0}^{m}, t_0 = 0, t_m = T$, then at time $0 \le t_k \le T$, we have $V_j(t_k) = \sum_{i=1}^{n} \dot{a}_{ji}\hat{s}_i V^{(i)}(t_k)$.

In theory, we expect the $\mathbf{V_t}$ series to be non-negative since it is related to the series of instantaneous variances $((\mathbf{A} \circ \mathbf{A})\mathbf{V_t})$ for the underlying assets. In practice, however, we observe inconsistencies in some cases. For example, as it can be seen in Figure 5.4a, $V_2(t)$ has quite a few negative values (labeled by "V2"), which are non negligible. Next we perform a preliminary analysis aiming at locating the root of the issue.

We first consider the limitations of our model to be the source of the inconsistency. Since we assume fixed constant orthogonal matrix $\mathbf{A}$, which contains information about correlation between assets, our model lacks the flexibility to capture varying correlation between assets. This could create a mismatch with the market's opinion on the actual correlation implied by a vector of volatility indexes. In other words, the correlation between assets captured by $\mathbf{A}$ might not be supported by their corresponding volatility indexes, otherwise we should observe positive $V_2(t)$.

In order to test this assumption, we estimate $\mathbf{A}$ at a higher frequency, for example, once a decade to once every year, thus we allow the matrix $\mathbf{A}$ to change in time aiming to capture the varying correlation between assets. A quick way to assess whether time-dependent $\mathbf{A}$ serves as a good solution, we picture the time series each year and check if we can still observe negative values in $V_2(t)$ series. We have 2450 observations for $V_1(t)$ and $V_2(t)$ respectively from 2011 to 2020. We break the two series into 10 observing sections with 245 observations in each section, roughly representing observations in a trading year. This way we have $\mathbf{A}$ changing every year instead of being fixed for 10 years. The result is shown in Figure 5.5. In Figure 5.5a, there are still negative values in $V_2(t)$ series. Rather, between 2018 and 2019, we even see negative $V_1(t)$. Figure 5.5b records the percentages of negative values in $V_2(t)$ in each observing section. We see a decreasing trend from section 1 to section 7, but it suddenly increases between section 7 to section 10.

Even though the preliminary analysis does not show exactly what causes the inconsistency, it on the other hand verifies that the cause is not necessarily due to deficiencies of our model in terms of matrix $\mathbf{A}$. Even more, by letting $\mathbf{A}$ change from year to year, this allows the scaling factors $(s1, s2)$ to vary as well. This means the model has plenty of flexibility to adapt to data.

The percentage of negative values, however, does not decrease with this flexibility. It stands to reason that for a very high frequency, i.e. time stochastic $\mathbf{A_t}$, $s1_t$, $s2_t$, all negative values in $V_2(t)$ shall disappear. This tells us that there might be a need for an extra stochastic driver in the covariance to fully capture the observable volatility indexes. We also suspect a disconnection between asset prices and their trade-able variances as a reason for this problem, which can be noticed more evidently at the multivariate level. Such disconnection could imply arbitrage opportunities (completing further the market), potential material for future research.

To solve the problem of inconsistency without modifying our model, we deal with the negative values as if they are missing values. So we replace the negative values by the weekly averages centered at those negative values. Figure 5.4b shows the series of $V_1(t)$ and $V_2(t)$ after this modification.

Figure 5.6 shows two series: $V_1(t)$ and $V_2(t)$ after transforming the original OVX and GVZ data. In this case, we do not observe the inconsistency as shown in Figure 5.4b, which means the data supports our model. The figure also illustrates the trend as expected with the first principal component generating largest variation ($V_1(t)$) in asset price compared to the second principal component ($V_2(t)$).

In Figure 5.7a, we also observe some inconsistency in silver ETF (SLV) and gold ETF (GLD) data between 2011 and 2012. Because the correlation between SLV and GLD is large, $V_2(t)$ series stays very close to 0, which implies that the two assets are probably driven by the same random factor. Since the negative values do not appear as often as in Figure 5.4a and are close to 0, we simply take absolute value of the negative values and show the modified series in Figure 5.7b.

Now we have prepared all the data for estimation. We apply the estimation method we have developed in chapter 3 to $V_1(t)$ and $V_2(t)$ to estimate $b_1, \alpha_1, \theta_1, \xi_1$ and $b_2, \alpha_2, \theta_2, \xi_2$. Note that, in all three scenarios, the minimum of $V_2(t)$ is approximately 0, which implies that b_2 is zero as seen from Figures 5.4b, 5.6 and 5.7. Therefore, it is sufficient to model $V_2(t)$ as a CIR (1/2) process instead of a 4/2 process. On the other hand, the "spikes" occurred frequently in $V_1(t)$ (labeled by "V1") shown by the figures are signals that $V_1(t)$ should be a 4/2 process given all three pairs of assets-volatility indexes data. Since we assume $V_2(t)$ follows a CIR process, we estimate $\alpha_2, \theta_2, \xi_2$ using maximum likelihood. Tables 5.2 to 5.4 show the estimated parameters and their standard errors (s.e's) with the chosen data sets for parameters in volatility group. The inference on the parameters (asymptotic mean and variance) is done via parametric bootstrap. In other words, we simulate the corresponding processes with the estimated parameters 1000 times and repeat the estimation procedure for each simulation. In the end, we obtain a pool of 1000 sets of estimates. The law of large numbers suggests that the means calculated from the pool of estimates are the asymptotic means for each estimator.

5.3.3   Estimation of Drift Group

Similarly, we use the least square approach from chapter 3 to estimate parameters in the drift group. Tables 5.5 to 5.7 show the results. Some parameters are assessed to be non significant
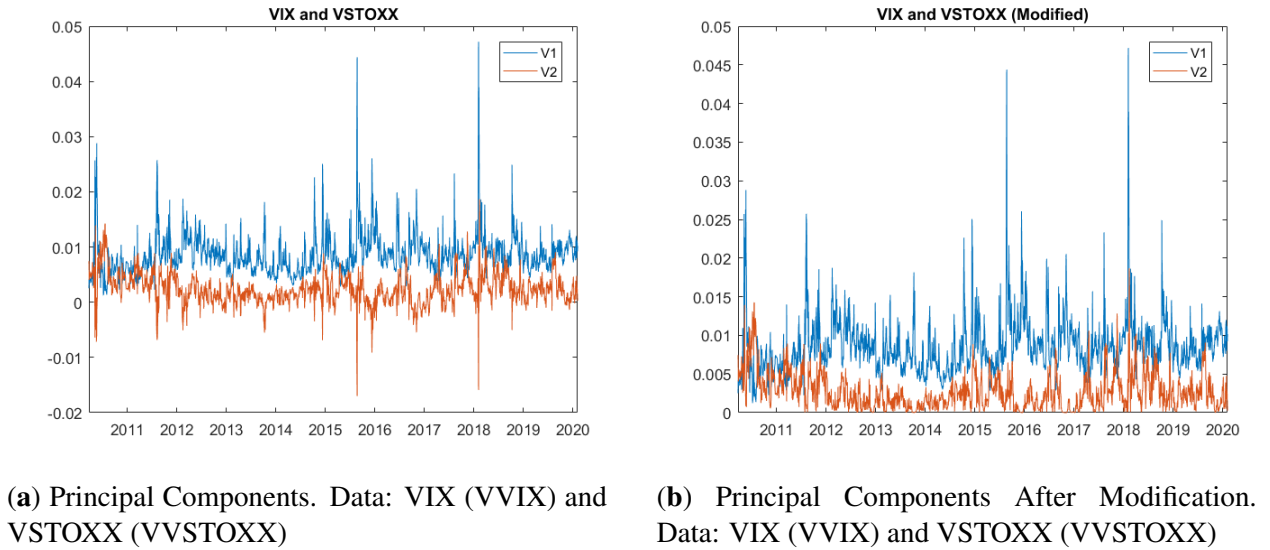
(a) Principal Components. Data: VIX (VVIX) and VSTOXX (VVSTOXX)

(b) Principal Components After Modification. Data: VIX (VVIX) and VSTOXX (VVSTOXX)

Figure 5.4: Principal Components With Volatility Indexes Data.

(a) Principal Components With Changing A. Data: VIX (VVIX) and VSTOXX (VVSTOXX)

(b) Percentage of Negative Values Each Section
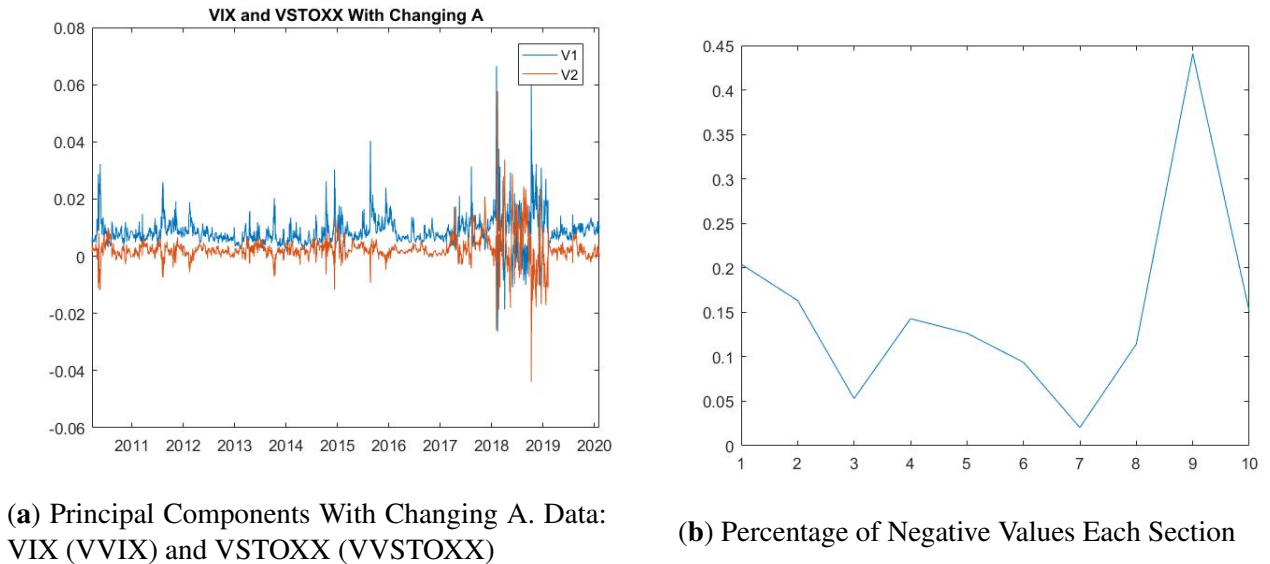
Figure 5.5: Principal Components With Volatility Indexes Data (Varying A).

based on the p-values. We decide to keep all the parameters because our sample sizes are not large enough to draw concrete conclusions on the significance of the parameters.

Note that the estimated parameters reported in the tables are for the parameters of the $M_1(t)$ and $M_2(t)$ processes. We can recover the estimates for original parameters using the relationship
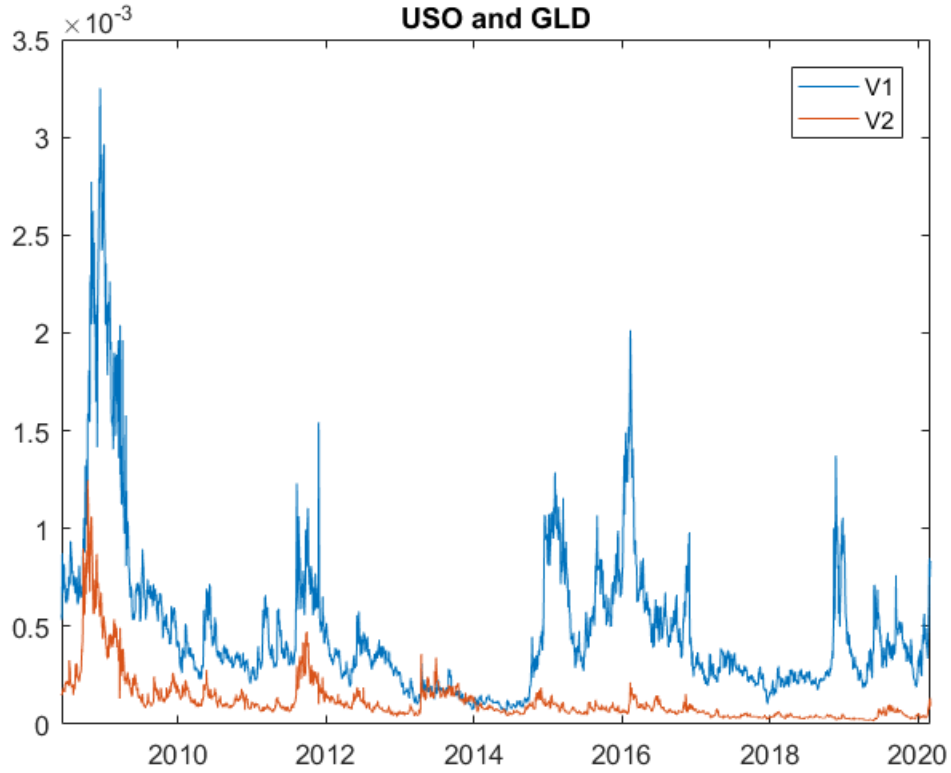
Figure 5.6: Principal Components. Data: USO (OVX) and GLD (GVZ)
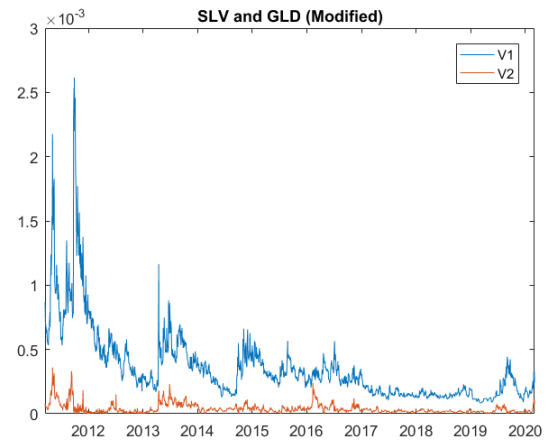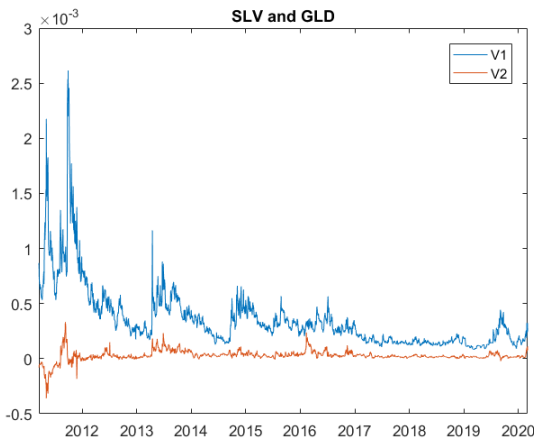
(**a**) Principal Components. Data: SLV (VXSLV) and GLD (GVZ)

(**b**) Principal Components After Modification. Data: SLV (VXSLV) and GLD (GVZ)

Figure 5.7: Principal Components With Volatility Indexes Data.

we defined earlier:

$$\hat{\mathbf{L}} = \hat{\mathbf{A}}\tilde{\mathbf{L}}$$

$$\hat{\mathbf{B}} = \hat{\mathbf{A}}\tilde{\mathbf{B}}\hat{\mathbf{A}}^{-1}$$

$$\hat{\mathbf{C}} = \hat{\mathbf{A}}\tilde{\mathbf{C}} + \frac{1}{2}(\hat{\mathbf{A}} \circ \hat{\mathbf{A}})$$

Table 5.2: Estimated Volatility Group Parameters With Empirical Data (VIX and VSTOXX)

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	3.11e-4	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	4.511e-4(1.1581e-6)	0(0)
$\hat{\alpha}$	42.1811	19.1624
Mean of $\hat{\alpha}$ (s.e)	41.6498(0.0695)	19.5827(0.0552)
$\hat{\theta}$	0.0079	0.0027
Mean of $\hat{\theta}$ (s.e)	0.0076(4.8578e-6)	0.0027(8.0796e-6)
$\hat{\xi}$	0.3436	0.3885
Mean of $\hat{\xi}$ (s.e)	0.3537(1.3937e-4)	0.3877(5.1103e-4)

Table 5.3: Estimated Volatility Group Parameters With Empirical Data (USO and GLD)

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	1.6776e-5	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	7.5578e-5(1.3892e-6)	0(0)
$\hat{\alpha}$	3.62	5.3597
Mean of $\hat{\alpha}$ (s.e)	5.9253(0.0465)	5.6994(0.0186)
$\hat{\theta}$	8.9803e-4	1.1859e-4
Mean of $\hat{\theta}$ (s.e)	4.6621e-4(7.4231e-6)	1.1779e-4(2.4962e-7)
$\hat{\xi}$	0.0271	0.0231
Mean of $\hat{\xi}$ (s.e)	0.02(2.4446e-4)	0.0231(5.3704e-6)

Table 5.4: Estimated Volatility Group Parameters With Empirical Data (SLV and GLD)

	Data: $V_1(t)$	Data: $V_2(t)$
$\hat{\mathbf{b}}$	2.0968e-5	~ 0
Mean of $\hat{\mathbf{b}}$ (s.e)	7.4123e-5(2.85e-7)	0(0)
$\hat{\alpha}$	5.178	24.3083
Mean of $\hat{\alpha}$ (s.e)	7.7687(0.0664)	24.8349(0.057)
$\hat{\theta}$	8.3026e-4	3.8644e-5
Mean of $\hat{\theta}$ (s.e)	4.6417e-4(7.5518e-6)	3.8753e-5(6.5274e-8)
$\hat{\xi}$	0.0307	0.0343
Mean of $\hat{\xi}$ (s.e)	0.024(3.0304e-4)	0.0344(1.1187e-5)

The estimates for original parameters are reported in Table 5.8. The diagonal entries in $\hat{\mathbf{B}}$ matrices give us the information on the mean-revering speed for all the assets. Note that for the pair SLV&GLD, one of the eigenvalues (-0.5401) of $\hat{\mathbf{B}}$ does not satisfy the assumption imposed on the eigenvalues of $\mathbf{B}$, which is a sign that the data does not support our model. The correlation coefficients are not included since they are not affected by the transformation.
It is worth noting that $\hat{\mathbf{L}}$ in Table 5.8 does not reflect the actual mean-reverting level. To see what the mean-reverting level is for each asset, we must go back to Equation 5.7 and rewrite it

Table 5.5: Estimated Drift Group Parameters (VIX and VSTOXX)

Data	$\hat{\tilde{L}}_i$	p-value	$\hat{\tilde{c}}_i$	p-value	$\hat{\tilde{\beta}}_i$	p-vale	$\hat{\rho}$	p-value
$M_1(t)\&V_1(t)$	25.6041	0	-114.6521	0.436	6.1077	0	0.5621	0
$M_2(t)\&V_2(t)$	9.4459	0	73.6448	0.734	15.6266	0	0.00419	0

Table 5.6: Estimated Drift Group Parameters (USO and GLD)

Data	$\hat{\tilde{L}}_i$	p-value	$\hat{\tilde{c}}_i$	p-value	$\hat{\tilde{\beta}}_i$	p-vale	$\hat{\rho}$	p-value
$M_1(t)\&V_1(t)$	0.8096	0.154	-416.2006	0.214	0.214	0.134	-0.3723	0
$M_2(t)\&V_2(t)$	2.6418	0.07	-646.7339	0.384	0.5701	0.079	-0.00294	0

Table 5.7: Estimated Drift Group Parameters (SLV and GLD)

Data	$\hat{\tilde{L}}_i$	p-value	$\hat{\tilde{c}}_i$	p-value	$\hat{\tilde{\beta}}_i$	p-vale	$\hat{\rho}$	p-value
$M_1(t)\&V_1(t)$	4.5459	0.009	842.936	0.153	1.0183	0.008	-0.2323	0
$M_2(t)\&V_2(t)$	-1.5092	0.23	-1787.0546	0.16	-0.5401	0.191	0.1228	0

Table 5.8: Estimated Original Drift Group Parameters

Data	$\hat{\mathbf{L}}$	$\hat{\mathbf{C}}$	$\hat{\mathbf{B}}$
VIX&VSTOXX	(25.9172,8.5495)	$\begin{pmatrix} -89.4091 & 46.0451 \\ -71.1886 & -57.3209 \end{pmatrix}$	$\begin{pmatrix} 9.7978 & -4.6378 \\ -4.6378 & 11.9365 \end{pmatrix}$
USO&GLD	(0.4653,2.7236)	$\begin{pmatrix} -412.2959 & -82.6444 \\ -53.1823 & 641.9225 \end{pmatrix}$	$\begin{pmatrix} 0.2198 & -0.0451 \\ -0.0451 & 0.5643 \end{pmatrix}$
SLV&GLD	(4.7379,0.7032)	$\begin{pmatrix} 752.7178 & -805.8599 \\ 380.2658 & 1595.3 \end{pmatrix}$	$\begin{pmatrix} 0.7013 & 0.6273 \\ 0.6273 & -0.2231 \end{pmatrix}$

in following format:

$$\mathbf{dY_t} = \mathbf{B}(\mathbf{B^{-1}L} + \mathbf{B^{-1}A\tilde{C}V_t} - \mathbf{Y_t})dt + \mathbf{A}diag(\mathbf{V_t})^{\frac{1}{2}}\mathbf{dW_t^{\mathbb{P}}}.$$

We can see that the mean-revering level is $\mathbf{B^{-1}L}$ plus a random component, which we define as $H(\mathbf{V_t}) = \mathbf{B^{-1}A\tilde{C}V_t}$. The long term mean indicated by the model is basically $\mathbf{B^{-1}L} + \mathbb{E}[H(\mathbf{V_t})]$. We report these estimates in Table 5.9 and compare with the averages calculated from empirical data.

Table 5.9: Estimated Original Drift Group Parameters

Data	$\mathbf{B^{-1}L}$	$\mathbb{E}[H(\mathbf{V_t})]$	Estimated Mean-Reverting Level	Empirical Averages
VIX&VSTOXX	(2.9042,3.0832)	(-0.0952,-0.0631)	(2.809,3.0219)	(2.6497,2.9139)
USO&GLD	(3.1598,5.0794)	(-0.961,-0.2547)	(2.244,4.8247)	(3.1342,4.8096)
SLV&GLD	(2.7242,4.5075)	(0.2215,0.2158)	(2.9457,4.7232)	(2.9468, 4.8645)

From Table 5.9, we can see that the estimated MRLs are closed to empirical log price averages except for USO case where the estimated mean is quite smaller than the empirical mean. This latest point might be due to the impact of the initial value on the stationary value of a 4/2 process. We can see that VIX&VSTOXX have the largest mean-reverting speed comparing to the other two commodity ETF pairs. This is not a surprise as evidenced by empirical data. Volatility indexes tend to return to mean faster due to economic cycle whilst commodities normally have longer time horizon to revert to mean level due to scarcity, demand and supply.

5.4 Risk Measures

In section 1.3.5, we provided some definitions and concepts related to risk measures and financial risk management. Risk measures determine the minimum amount of capital to be kept in reserve in worst case scenarios as a way of protecting financial institutions. There are many risk measures in the literature see, for example, [7] and [104], one of which is considered fundamental yet crucial: Value-at-Risk (VaR). VaR is a distribution-based risk measure, in other words, VaR calculation takes into account the distribution of the underlying (VaR is in fact a quantile). It is more robust to outliers than mean and variance.

In this section, we compute the VaR of a portfolio consisting of two assets, for example USO and GLD. To calculate the VaR, we first need to find the distribution of this portfolio, which might not be available due to, in particular, the correlation among the underlyings. In language of mathematical statistics, we need to find the joint distribution of USO and GLD in order to compute VaR. Theoretically, in general, it is difficult to find closed-form expressions for the joint distribution of two non-Gaussian stochastic processes. In fact, USO and GLD have very complex distribution functions under our multidimensional 4/2 model setting. Fortunately, using principal component decomposition, we can decompose a multidimensional structure into independent one dimensional pieces. What is more attractive is that we are able to express the joint distribution at any given date of USO and GLD in terms of two independent random variables, which simplifies our problem significantly and allow for the use of c.f.s to compute properties of the portfolio distribution.

5.4.1 Portfolio Setup

Suppose we have a portfolio Π consisting of two assets $X_1(t)$ and $X_2(t)$:

$$\Pi(t) = \phi_1(t)X_1(t) + \phi_2(t)X_2(t) + \phi_3(t)B(t), \tag{5.27}$$

where ϕ_1 and ϕ_2 represent the weights of $X_1(t)$ and $X_2(t)$ in the portfolio and $B(t)$ is a cash account with interest rate r. In a short period of time, we can also write the problem using the self-financing condition and relative portfolio weights π_1, π_2 and $(1 - \pi_1 - \pi_2)$ (i.e. proportion allocated to the assets and cash account, allocations) see [28]:

$$\frac{d\Pi(t)}{\Pi(t)} = \pi_1 \frac{dX_1(t)}{X_1(t)} + \pi_2 \frac{dX_2(t)}{X_2(t)} + (1 - \pi_1 - \pi_2)rdt. \tag{5.28}$$

For the purpose of this chapter we will only consider constant allocations, i.e. constant π, as this is the most popular investment strategy in the market, supported by [107]. From the process $\frac{dX_i(t)}{X_i(t)}$, we can easily obtain $dY_i(t)$ ($Y_i(t) = log(X_i(t))$) by using Ito's lemma. When comparing $\frac{dX_i(t)}{X_i(t)}$ and $dY_i(t)$, we observe that only the drift term is adjusted, diffusion terms stay the same. Assuming $(X_1(t), X_2(t))$ are modeled by Equation 5.2, the log prices $(Y_1(t), Y_2(t))$ then have the SDE specified in Equation 5.7 under the PCSV framework. Moreover, we can also write $\frac{dX_i(t)}{X_i(t)}$ in terms of $dY_i(t)$:

$$\frac{dX_i(t)}{X_i(t)} = dY_i(t) + \frac{1}{2} \sum_j a_{ij}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt.$$

Hence, we rewrite Equation 5.28 as:

$$\frac{d\Pi(t)}{\Pi(t)} = \pi_1 \left[dY_1(t) + \frac{1}{2} \sum_{j=1}^{2} a_{1j}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt \right] + \pi_2 \left[dY_2(t) + \frac{1}{2} \sum_{j=1}^{2} a_{2j}^2 \left(\sqrt{v_j(t)} + \frac{b_j}{v_j(t)} \right)^2 dt \right]$$

$$+ (1 - \pi_1 - \pi_2) r dt.$$

$$(5.29)$$

We have also demonstrated that $Y_1(t)$ and $Y_2(t)$ are linear combinations of two independent stochastic processes/random variables $M_1(t)$ and $M_2(t)$, hence we can write the dynamics of $Y_1(t)$ and $Y_2(t)$ in terms of $M_1(t)$ and $M_2(t)$ as follows:

$$dY_1(t) = a_{11} dM_1(t) + a_{12} dM_2(t), \qquad (5.30)$$

$$dY_2(t) = a_{21} dM_1(t) + a_{22} dM_2(t). \qquad (5.31)$$

We now substitute $dY_1(t)$ and $dY_2(t)$ in Equation 5.29 with Equations 5.30 and 5.31:

$$\frac{d\Pi(t)}{\Pi(t)} = (\pi_1 a_{11} + \pi_2 a_{21}) dM_1(t) + \frac{1}{2} (\pi_1 a_{11}^2 + \pi_2 a_{21}^2) \left(\sqrt{v_1(t)} + \frac{b_1}{v_1(t)} \right)^2 dt$$

$$+ (\pi_1 a_{12} + \pi_2 a_{22}) dM_2(t) + \frac{1}{2} (\pi_1 a_{12}^2 + \pi_2 a_{22}^2) \left(\sqrt{v_2(t)} + \frac{b_2}{v_2(t)} \right)^2 dt \qquad (5.32)$$

$$+ (1 - \pi_1 - \pi_2) r dt.$$

From Equation 5.32, we can conclude that $\frac{d\Pi(t)}{\Pi(t)}$ is also a linear combination of $dM_1(t)$ and $dM_2(t)$ with adjustment done to the drift terms, which does not affect the independent relationship between $dM_1(t)$ and $dM_2(t)$. We organize Equation 5.32 into the following expression:

$$\frac{d\Pi(t)}{\Pi(t)} = dM_1^*(t) + dM_2^*(t) \qquad (5.33)$$

where $dM_1^*(t)$ and $dM_2^*(t)$ are independent with

$$dM_i^*(t) = \left[L_i^* + c_i^* (\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})^2 - \beta_i^* M_i^*(t) \right] dt + a_i^* (\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}) dW_i^{\mathbb{P}}(t),$$

where,

$$L_1^* = \tilde{L}_1 a_1^* + (1 - \pi_1 - \pi_2)r, L_2^* = \tilde{L}_2 a_2^*$$

$$c_i^* = \tilde{c}_i a_i^* + \frac{1}{2} a_i^{**},$$

$$\beta_i^* = \tilde{\beta}_i a_i^*,$$

$$a_i^* = \sum_{j=1}^{2} \pi_j a_{ji},$$

$$a_i^{**} = \sum_{j=1}^{2} \pi_j a_{ji}^2.$$

Note we included the growth rate on the cash account into long term average of $M_1^*(t)$ for convenience. From a mathematical perspective, $M_1^*(t)$ is constructed based on $M_1(t)$, which is the first principal component that determines most variation among the assets. As a result, $M_1^*(t)$ affects the performance of $\Pi(t)$ more than $M_2^*(t)$ financially. For this reason, the effect of the growth rate on the portfolio can also be interpreted as if it impacts the long term average of $M_1^*(t)$. Now we apply Ito's lemma to Equation 5.33 to obtain the dynamics for $ln(\Pi(t))$:

$$dln(\Pi(t)) = d\tilde{M}_1^*(t) + d\tilde{M}_2^*(t) \tag{5.34}$$

where,

$$d\tilde{M}_i^*(t) = \left[L_i^* + (c_i^* - \frac{1}{2}(a_i^*)^2)(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})^2 - \beta_i^* \tilde{M}_i^*(t) \right] dt + a_i^*(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}}) dW_i^{\mathbb{P}}(t).$$

It is straightforward to find the characteristic function of $dln(\Pi(t))$ using results from chapter 4 and chapter 5.

5.4.2 The Density Function of The Portfolio $\Pi(t)$

From Equation 5.34, our portfolio now basically contains two new "assets" that are independent of each other. Thanks to this independence, we can derive the characteristic function as well as the density function of our portfolio. Since our goal is to calculate VaR, it is convenient to use density function and integrate numerically. In this section, we list two approaches to obtain such density function.

5.4.2.1 The Density Function via Convolution

One way to obtain the conditional density function for $ln(\Pi(T))$ is via convolution of two conditional density functions for $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$. In probability, if two random variables X and Y are independent with density functions $f_X(x)$ and $f_Y(y)$ respectively. Let $Z = X + Y$, then the density for Z, $f_Z(z)$, can be found via convolution, i.e:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx. \tag{5.35}$$

If X and Y have analytical density functions, then convolution method is straightforward. In our case, we obtain the conditional c.f. first. Theoretically, the Fourier inversion of the conditional c.f. gives the density function. Due to the structure of our original c.f. and the approximations, we need to invert the original c.f. and the approximated c.f.'s numerically for the corresponding density functions. For $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$, we can obtain their conditional corresponding density functions $f_1(m_1|\mathcal{F}_t)$ and $f_2(m_2|\mathcal{F}_t)$ by inverting the c.f. of $\tilde{Z}_1^*(t) = e^{\beta_1^* t}\tilde{M}_1^*(t)$ and $\tilde{Z}_2^*(t) = e^{\beta_2^* t}\tilde{M}_2^*(t)$:

$$f_1(m_1|\mathcal{F}_t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ium_1}\Phi_{MR}\left(T, u; L_1^*, c_1^*, a_1^*, \beta_1^*, \alpha_1, \theta_1, \xi_1, \rho_1, a_1^* b_1, v_1(t), e^{-\beta_1^* t}\tilde{Z}_1^*(t)\right) du.$$
(5.36)

$$f_2(m_2|\mathcal{F}_t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ium_2}\Phi_{MR}\left(T, u; L_2^*, c_2^*, a_2^*, \beta_2^*, \alpha_2, \theta_2, \xi_2, \rho_2, a_2^* b_2, v_2(t), e^{-\beta_2^* t}\tilde{Z}_2^*(t)\right) du.$$
(5.37)

We can now write Equation 5.40 as a convolution of Equations 5.36 and 5.37:

$$f(p|\mathcal{F}_t) = \int_{-\infty}^{\infty} f_1(m_1|\mathcal{F}_t)f_2(p - m_1|\mathcal{F}_t)dm_1.$$
(5.38)

A challenging part of this method is that we need to invert the semi-closed c.f.s first to obtain the density functions for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$ (artificial assets), which involves approximations. As we have well-developed approximation approaches for Φ_{MR}, we can apply the results to obtain analytic function as approximation of the c.f. for individual artificial asset and then find the density via Fourier inversion. Then we can use Equation 5.38 to obtain the density of the portfolio. Recall that our approximation approaches work well in a parametric region as seen in Figures 3.1 and 3.2 for three scenarios ($b_i = 0, \rho_i \neq 0; b_i \neq 0, \rho_i = 0; b_i = 0, \rho_i = 0$), the goodness of approximations depend on $L_i^*, c_i^*, a_i^*, \beta_i^*, \alpha_i, \theta_i, \xi_i, \rho_i, b_i$.

Also in the scenario where the 3/2 component is not negligible, we involve a hypergeometric function in the approximated function, which requires numerical computation methods, either we find the hypergeometric function numerically or avoid the hypergeometric function with partial simulation (only need to simulate $v_i(t)$ process to obtain the c.f.). These are the difficulties involved in finding the individual c.f. Despite these difficulties, our approximation approaches have been proved effective in chapter 3. With analytic approximations, we can find the individual c.f. efficiently.

5.4.2.2   Density Function via Fourier Inversion

Another way to obtain density function is to apply inverse Fourier transform to the characteristic function. Before we give the formula for the characteristic function, we consider the transformation $\tilde{Z}_i^*(t) = e^{\beta_i^* t}\tilde{M}_i^*(t)$. By Ito's lemma, we have:

$$d\tilde{Z}_i^*(t) = e^{\beta_i^* t}\left[L_i^* + (c_i^* - \frac{1}{2}(a_i^*)^2)(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})^2\right]dt + e^{\beta_i^* t}a_i^*(\sqrt{v_i(t)} + \frac{b_i}{\sqrt{v_i(t)}})dW_i^{\mathbb{P}}(t).$$

Next corollary illustrates how to derive the characteristic function:

Corollary 5.4.1 *Let* $\Phi_{MR}(T, u; L, c, a, \beta, \alpha, \theta, \xi, \rho, b, v(t), Z_t(t))$ *denote the characteristic function provided in Proposition 3.2.1, then the characteristic function of* $ln(\Pi(T))$ *is given by the following equation:*

$$\mathbb{E}(e^{iuln(\Pi(T))}|\mathcal{F}_t) = \prod_{i=1}^{2} \Phi_{MR}\left(T, u; L_i^*, c_i^*, a_i^*, \beta_i^*, \alpha_i, \theta_i, \xi_i, \rho_i, a_i^* b_i, v_i(t), \tilde{Z}_i^*(t)\right) \tag{5.39}$$

where $L_1^* = \tilde{L}_1 a_1^* + (1 - \pi_1 - \pi_2)r$, $L_2^* = \tilde{L}_2 a_2^*$, $c_i^* = \tilde{c}_i a_i^* + \frac{1}{2} a_i^{**}$, $\beta_i^* = \tilde{\beta}_i a_i^*$, $a_i^* = \sum_{j=1}^{2} \pi_j a_{ji}$, $a_i^{**} = \sum_{j=1}^{2} \pi_j a_{ji}^2$.

Let p denote all possible values in the domain of $ln(\Pi(T))$, the density function hence follows:

$$f(p|\mathcal{F}_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iup} \mathbb{E}(e^{iuln(\Pi(T))}|\mathcal{F}_t) du. \tag{5.40}$$

It is not hard to see that the c.f. of the portfolio does not have a closed-form representation since it is a product of semi-closed c.f.s (Φ_{MR}) from chapter 3. As a result, we only have a term (a product of conditional expectations in the case of portfolio c.f.) that requires approximations for fast computations. Hence, the Fourier inversion method still face both theoretical and numerical difficulties as we outlined at the end of previous section. The convolution method and the Fourier inversion method both require finding the c.f.s first, which may require either working with a hypergeometric function, a partial simulation (simulating $v_i(t)$ process) or an approximation.

5.4.2.3 Numerical Implementation of Selected Method

Theoretically, the Fourier inversion method and the convolution method yield the same density function for the portfolio. In a portfolio which only consists of two (artificial) risky assets, both methods are not complicated to implement. We decide to implement convolution method with partial simulation as it directly uses the well-developed algorithm from chapter 3. However, it would be more efficient to use the Fourier inversion method when the portfolio has a large pool of assets (e.g. over 100). To see this, note the convolution method involves, for $n, n \gg 2$ assets, the simulation of n processes with $n+1$ integrations, while the Fourier inversion method reduces the number of integrations to just one.

We summarize the numerical implementation to compute the conditional density function of the portfolio in following steps:

- Step 1: Simulate two CIR processes $v_1(t)$ and $v_2(t)$ and compute Φ_{MR} for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$.

- Step 2: Invert the c.f.s obtained in Step 1 to get $f_1(m_1|\mathcal{F}_t)$ and $f_2(m_2|\mathcal{F}_t)$.

- Step 3: Integrate numerically the product of the conditional density of $\tilde{M}_1^*(T)$ and $\tilde{M}_2^*(T)$ for the conditional density function of $ln(\Pi(T))$.

Even though we use partial simulation to get the density function of the portfolio, partial simulation is not time-consuming as there are efficient methods to simulate CIR processes, see for example [4]. Moreover, both of the convolution method and the direct Fourier inversion method require fewer simulations or no simulations with appropriate approximations, compared the the simulation of 4 processes required otherwise. In the worst case where the semi-closed c.f.s are involved, we would only need to simulate at most $n, n \geq 2$ processes ($v_i(t)$ for $i = 1, \ldots, n$ process), instead of an simulation of the $\tilde{M}_1^*(t)$ processes altogether ($2n$ processes). Most importantly, thanks to P.C principles, we would likely need m such volatility drivers to explain the stochastic covariance of n assets with $m << n$. This means a substantial reduction in computational complexity (partial, simulations, integrations or approximations).

To sum up, under a PCSV framework partial simulation is a good choice in terms of efficiency. The P.C principles reduces computational complexity as fewer diffusions may be required in explaining the variation of all assets. Our approximations further improve the efficiency for computing the c.f.s with analytic functions.

In the next application section, we entertain the convolution method from section 5.4.2.1 to compute the VaR at popular quantile $\alpha_q = 0.95$ ($VaR_{0.95}$).

5.4.3   VaR For A Portfolio of USO and GLD

We consider a pair of risky assets: USO and GLD. Based on the findings and results in section 5.3, we realize that our model is not supported by VIX and VSTOXX data as a number of negative principal component values are observed, which should not be possible. Even though a negligible number of negative principal component values are observed using SLV and GLD data, the estimation result does not agree with our assumption on the eigenvalues of **B**. These findings are indications that the data (VIX and VSTOXX; SLV and GLD) does not support our model, thus it is logically unreasonable to consider these data in the application as the portfolio which we will be constructing to calculate $VaR_{0.95}$ is based on our model described in section 5.3. In the following we study $VaR_{0.95}$ under two different investment strategies: equally weighted risky assets only ($\pi_1 = \pi_2 = 0.5$) and equally weighted assets ($\pi_1 = \pi_2 = \frac{1}{3}$). These have been proved to be robust and reliable strategies in the seminal work of DeMiguel et.al [43]. In Tables 5.11 and 5.13, we report the $VaR_{0.95}$ values and their standard errors (s.e). We consider a well-known asymptotic result for quantiles to calculate the standard errors for VaR_{α_q} as derived in [124]:

$$s.e(VaR_{\alpha_q}) = \sqrt{\frac{\alpha_q(1 - \alpha_q)}{nf(VaR_{\alpha_q})^2}},$$

where α_q is the quantile of the portfolio distribution, in this case it is 95%; n is the sample size and $f(VaR_{\alpha_q})$ is the probability distribution function (density function) evaluated at VaR_{α_q}.

Case 1: $\pi_1 = \pi_2 = 0.5$

In this case, we use the information from Tables 5.1b, 5.3 and 5.6 to obtain parameters that generate $\tilde{M}_2^*(T)$ or $\tilde{M}_1^*(T)$. Since only drift group parameters are transformed, we show the

new drift group parameters in the following table:

Table 5.10: Estimated Drift Group Parameters (USO and GLD)

Data	L_i^*	c_i^*	β_i^*	a_i^*	a_i^{**}	$\hat{\rho}$
$\tilde{M}_1^*(t)\&V_1(t)$	0.4532	0.5933	0.1198	0.5598	0.5	0
$\tilde{M}_2^*(t)\&V_2(t)$	1.1413	0.6567	0.2463	0.432	0.5	0

As in chapter 3, we show Figures 5.8 to 5.9 to justify that density functions from theory are in line with simulations. We compute and compare $VaR_{0.95}$ values from four sources: simulation

(a) Density and Histogram for $\tilde{M}_1^*(t)$

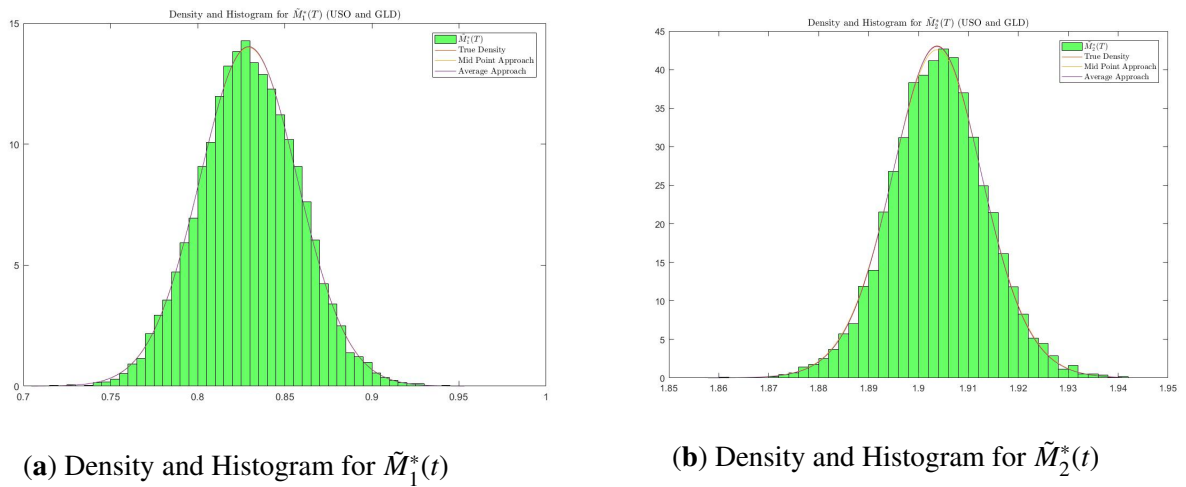(b) Density and Histogram for $\tilde{M}_2^*(t)$

Figure 5.8: Case 1: Density and Histogram for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$

of the portfolio (Simulation), density function without approximation (Density w/o Approximation) and approximated density function using midpoint approach (Approx. Density (M)) and average approach (Approx. Density (A)) respectively:

Table 5.11: $VaR_{0.05}$ for $ln(\Pi(t))$ From Four Sources

	Simulation	Density w/o Approximation	Approx. Density (M)	Approx. Density (A)
$VaR_{0.05}$	2.7833 (0.004006)	2.783 (0.00401)	2.7832 (0.003988)	2.783 (0.004008)

We use linear interpolation here to calculate the quantile if α_q falls in between of two critical levels calculated from histogram and density functions. Standard errors are reported in brackets.

Case 2: $\pi_1 = \pi_2 = \frac{1}{3}$

In this case, we consider the strategy "1/N" as supported by [43]. Similarly as in Case 1, we show the new drift group parameters in Table 5.12.
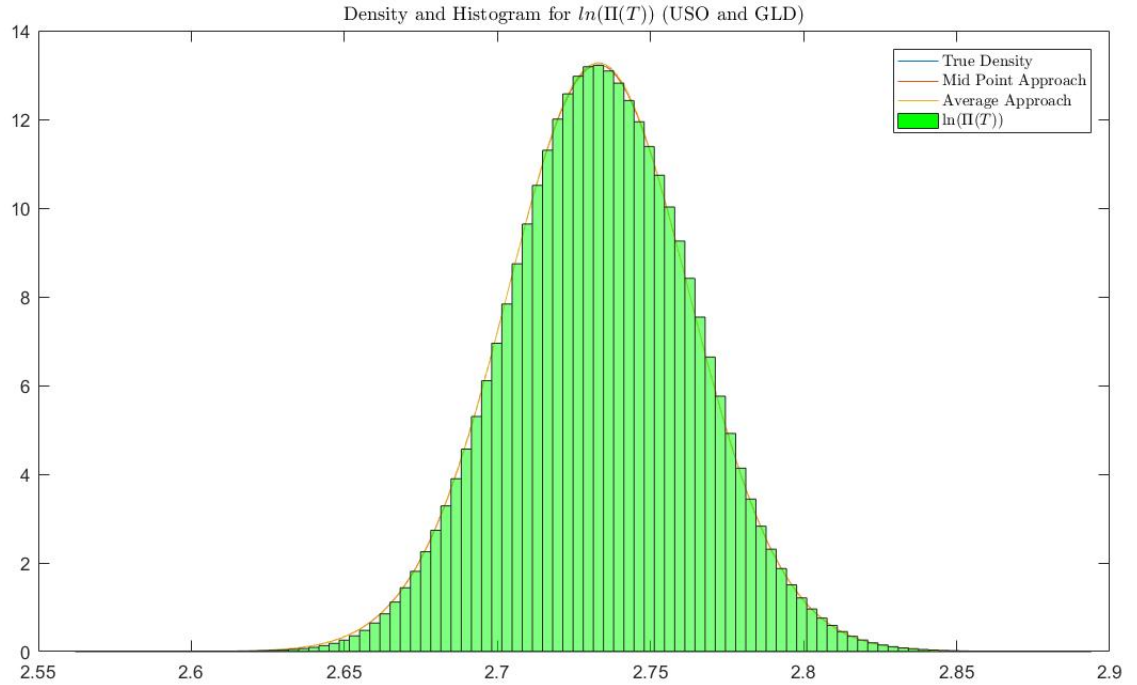
Figure 5.9: Case 1: Density and Histogram for $ln(\Pi(T))$

Table 5.12: Estimated Drift Group Parameters (USO and GLD)

| Data	L_i^*	c_i^*	β_i^*	a_i^*	a_i^{**}	$\hat{\rho}$
$\tilde{M}_1^*(t)\&V_1(t)$	0.3088	0.5970	0.0799	0.3732	0.3333	0
$\tilde{M}_2^*(t)\&V_2(t)$	0.7608	0.6252	0.1642	0.2880	0.3333	0

Similarly, Figures 5.10 to 5.11 demonstrate that density functions from theory are in line with simulations. Here we also consider four approaches: simulation of the portfolio (Simulation), density function without approximation (Density w/o Approximation) and approximated density function using midpoint approach (Approx. Density (M)) and average approach (Approx. Density (A)) respectively to compute and compare $VaR_{0.95}$ values. Linear interpolation is used

Table 5.13: $VaR_{0.05}$ for $ln(\Pi(t))$ From Four Sources

	Simulation	Density w/o Approximation	Approx. Density (M)	Approx. Density (A)
$VaR_{0.05}$	1.9801 (0.004454)	1.9802 (0.004294)	1.9803 (0.004282)	1.9802 (0.004293)

to calculate the quantile if α_q falls in between of two critical levels calculated from histogram and density functions. Standard errors are reported in brackets.

We can see from both cases that our approximation methods perform well given the estimated parameters. The $VaR_{0.95}$ results further prove the effectiveness of our approximation methods.
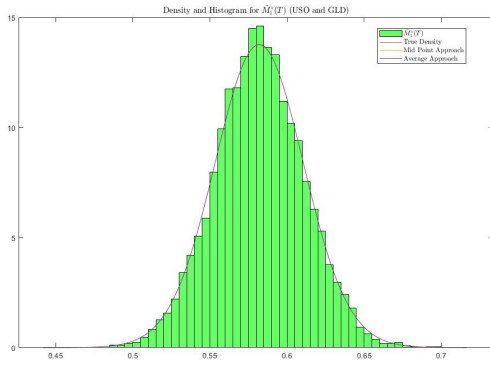
(a) Density and Histogram for $\tilde{M}_1^*(t)$
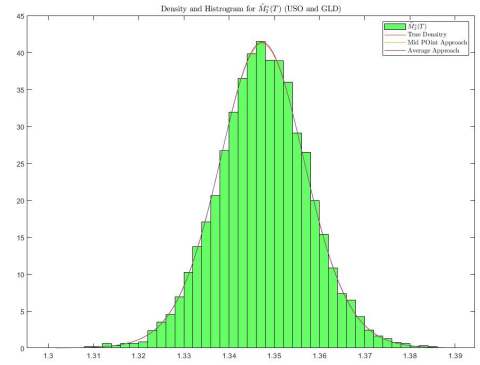
(b) Density and Histogram for $\tilde{M}_2^*(t)$

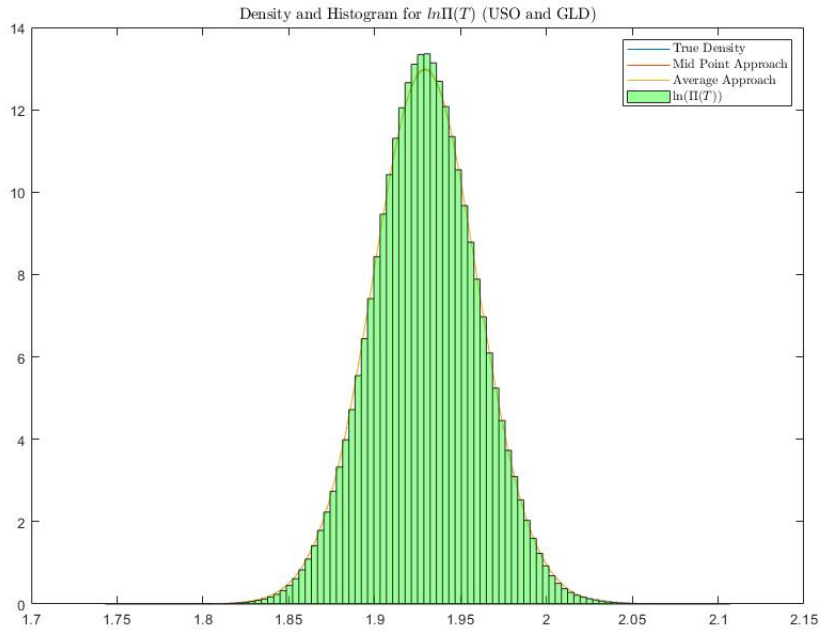Figure 5.10: Case 2: Density and Histogram for $\tilde{M}_1^*(t)$ and $\tilde{M}_2^*(t)$

Figure 5.11: Case 2: Density and Histogram for $ln(\Pi(T))$

Another finding from $VaR_{0.95}$ results is that the $VaR_{0.95}$ values produced by Case 2 investment strategy is smaller than those produced by Case 1 strategy. This is not surprise because in Case 1 we only invest in risky assets whereas in Case 2 we also invest in a risk-free cash account, hence we should have less exposure to risk in Case 2, and therefore smaller $VaR_{0.95}$ values.

5.5   Conclusion

In this chapter, we study the properties of a generalized mean-reverting 4/2 model in more details than chapter 4. We study c.f. approximations in the first two sections. In section 5.3, we applied our estimation methodology developed in chapter 3 to the generalized model. During the study, we confirm the role of the scaling parameters between instantaneous variance and VIX plays in a multidimensional setting. Another issue we find is the flaw of constant eigenvector matrix assumption imposed on principal component stochastic volatility framework. This can be seen from the estimation exercise on VIX (VVIX) and VSTOXX (VVSTOXX) where we find negative eigenvalues in the analysis of principal components.

In section 5.4, we further test our approximation methods in a risk management setting by computing one of the most popular risk measures VaR. Since VaR is a distribution based risk measure, it is ideal to assess the effectiveness of our approximation in a multidimensional setting. From Figures 5.8 to 5.11, we can see that the approximated density based on the average approach still out performs the one based on mid point in overall fitting. Using average approximation can greatly saves time in calculating distribution based risk measures. $VaR_{0.95}$ results in Tables 5.11 and 5.13 further prove that our theoretical and approximated density functions are consistent with simulation. Therefore, our approximation methods are proved to be effective in multidimension.

5.6   Summary and Future Research

In this thesis, we study a novel mean-reverting model with the state-of-the-art 4/2 stochastic volatility. We first study the theoretical properties of the model in one-dimensional setting in chapter 3. Even though the c.f. we find has quasi-closed form representation, it helps reduces computational cost for simulation purpose, especially in a multivariate model setting as we see in chapter 4 and 5, as we only need to simulate one process instead of the complete model that has two processes. To further improve computational efficiency, we propose two approximation approaches. Numerical results show that the approximation approaches have good performance in a certain parametric region. With approximated c.f., we can price options with FFT methods, completely avoiding Monte Carlo methods. Later in chapters 4 and 5, we extend the one-dimensional mean-reverting 4/2 stochastic volatility model to a multivariate case, also considering risk factors from different sources. We assume a convenient principal component stochastic volatility framework for the models studied in chapters 4 and 5. The benefit of this framework is that the original models can eventually be interpreted as linear combinations of independent processes after the decomposition of the stochastic covariance, which reduces the dimension of the model. However, the one of the drawback as we see in chapter 5 is that principal component stochastic volatility framework is not flexible with constant eigenvectors, thus it is not flexible to capture varying correlations.

Since our models are highly parametric, they embed interesting theoretical research topics

as well as numerical research topics. In chapter 3, the structure of the c.f. brings theoretical challenge, which is hard to solve given evidence from current literature, one can start with a simpler case by considering models in section 2.3.2 or a Grasselli's 4/2 model [76] with time-dependent coefficients. The parametric region is also worth exploring theoretically. For estimation, we currently claim $v(t) + \frac{b^2}{v(t)}$ satisfies weak stationary and ergodic properties, a rigorous proof is helpful to support this claim and also a useful fact about 4/2 process. Another interesting topic emerges from both theory and numerical application is regarding the choice of the stochastic covariance matrix in the context of chapter 5. As we have seen, the current choice or the decomposition of the stochastic covariance matrix is not consistent with some of the applications, which means the empirical data does not support our assumption. So a more robust principal component stochastic volatility framework is worth exploring due to the theoretical convenience of principal component decomposition. This way principal component stochastic volatility framework can be considered for a wide range of applications.

Bibliography

[1] Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

[2] Dong-Hyun Ahn and Bin Gao. A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies*, 12(4):721–762, 1999.

[3] Claudio Albanese and Alexey Kuznetsov. Transformations of markov processes and classification scheme for solvable driftless diffusions. *arXiv preprint arXiv:0710.1596*, 2007.

[4] Leif BG Andersen. Efficient simulation of the heston stochastic volatility model. 2007.

[5] Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Paul Labys. The distribution of realized exchange rate volatility. *Journal of the American statistical association*, 96(453):42–55, 2001.

[6] Elena Andreou and Eric Ghysels. Rolling-sample volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business & Economic Statistics*, 20(3):363–376, 2002.

[7] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[8] Xuezheng Bai. *Beyond Merton's utopia: effects of non-normality and dependence on the precision of variance estimaters using high-frequency financial data*. PhD thesis, University of Chicago, Graduate School of Business, 2000.

[9] Jan Baldeaux. Exact simulation of the 3/2 model. *International Journal of Theoretical and Applied Finance*, 15(05):1250032, 2012.

[10] Jan Baldeaux, Martino Grasselli, and Eckhard Platen. Pricing currency derivatives under the benchmark approach. *Journal of Banking & Finance*, 53:34–48, 2015.

[11] Jan Baldeaux and Eckhard Platen. *Functionals of multidimensional diffusions with applications to finance*, volume 5. Springer Science & Business Media, 2013.

[12] Ole E Barndorff-Nielsen and Neil Shephard. Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241, 2001. Wiley Online Library.

[13] Ole E Barndorff-Nielsen and Neil Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280, 2002.

[14] Ole E Barndorff-Nielsen and Neil Shephard. Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, 17(5):457–477, 2002.

[15] Ole E Barndorff-Nielsen, Neil Shephard, et al. Realized power variation and stochastic volatility models. *Bernoulli*, 9(2):243–265, 2003.

[16] Fred Espen Benth. The stochastic volatility model of barndorff-nielsen and shephard in commodity markets. *Mathematical Finance*, 21(4):595–625, 2011.

[17] N.H. Bingham and R. Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer Finance. Springer London, 2013.

[18] Fischer Black. Studies of stock market volatility changes. *1976 Proceedings of the American Statistical Association Bisiness and Economic Statistics Section*, 1976.

[19] Fischer Black, Emanuel Derman, and William Toy. A one-factor model of interest rates and its application to treasury bond options. *Financial analysts journal*, pages 33–39, 1990. JSTOR.

[20] Fischer Black and Piotr Karasinski. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, pages 52–59, 1991. JSTOR.

[21] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[22] Mireille Bossy and Awa Diop. *An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form— x—ˆ a, a in [1/2, 1*. PhD thesis, INRIA, 2007.

[23] Mark Britten-Jones and Anthony Neuberger. Option prices, implied price processes, and stochastic volatility. *The Journal of Finance*, 55(2):839–866, 2000.

[24] Mark Broadie and Özgür Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations research*, 54(2):217–231, 2006.

[25] Peter Brockwell, Erdenebaatar Chadraa, Alexander Lindner, et al. Continuous-time garch processes. *Annals of Applied Probability*, 16(2):790–826, 2006.

[26] Andrea Buraschi and Jens Jackwerth. Explaining option prices: Deterministic vs. stochastic models. Technical report, Working Paper, London Business School, 1998.

[27] Ruggero Caldana and Gianluca Fusai. A general closed-form spread option pricing formula. *Journal of Banking & Finance*, 37(12):4893–4906, 2013.

[28] John Y Campell, Andrew W Lo, and AC MacKinlay. The econometrics of financial markets. *New Jersey*, 1997.

[29] Peter Carr and Dilip Madan. Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73, 1999.

[30] Peter Carr and Jian Sun. A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, 10(2):87–150, 2007.

[31] Peng Cheng and Olivier Scaillet. Linear-quadratic jump-diffusion modeling. *Mathematical Finance*, 17(4):575–598, 2007.

[32] Yuyang Cheng, Marcos Escobar-Anel, and Zhenxian Gong. Generalized mean-reverting 4/2 factor model. *Journal of Risk and Financial Management*, 12(4):159, 2019.

[33] Mikhail Chernov. On the role of risk premia in volatility forecasting. *Journal of Business & Economic Statistics*, 25(4):411–426, 2007.

[34] Carl Chiarella, Boda Kang, Christina Sklibosios Nikitopoulos, and Thuy-Duong Tô. Humps in the volatility structure of the crude oil futures market: New evidence. *Energy Economics*, 40:989–1000, 2013.

[35] Bent J Christensen and Nagpurnanand R Prabhala. The relation between implied and realized volatility. *Journal of financial economics*, 50(2):125–150, 1998.

[36] Peter Christoffersen, Steven Heston, and Kris Jacobs. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 55(12):1914–1932, 2009.

[37] Joshua D Coval and Tyler Shumway. Expected option returns. *The journal of Finance*, 56(3):983–1009, 2001.

[38] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985. JSTOR.

[39] José Da Fonseca and Martino Grasselli. Riding on the smiles. *Quantitative Finance*, 11(11):1609–1632, 2011. Taylor & Francis.

[40] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180, 2007.

[41] Alvise De Col, Alessandro Gnoatto, and Martino Grasselli. Smiles all around: Fx joint calibration in a multi-heston model. *Journal of Banking & Finance*, 37(10):3799–3818, 2013.

[42] Kresimir Demeterfi, Emanuel Derman, Michael Kamal, and Joseph Zou. A guide to volatility and variance swaps. *The Journal of Derivatives*, 6(4):9–32, 1999.

[43] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953, 2009.

[44] Jérôme Detemple and Carlton Osakwe. The valuation of volatility options. *Review of Finance*, 4(1):21–50, 2000.

[45] Francis X Diebold and Marc Nerlove. The dynamics of exchange rate volatility: a multivariate latent factor arch model. *Journal of Applied econometrics*, 4(1):1–21, 1989.

[46] Awa Diop. *Sur la discrétisation et le comportement à petit bruit d'EDS unidimension-nelles dont les coefficients sont à dérivées singulières*. PhD thesis, 2003. Thèse de doctorat dirigée par Talay, Denis et Bossy, Mireille Mathématiques Nice 2003.

[47] Gordana Dmitrašinović-Vidović and Antony Ware. Optimal portfolios of mean-reverting instruments. *SIAM Journal on Financial Mathematics*, 2(1):748–767, 2011. SIAM.

[48] James S Doran and Ehud I Ronn. Computing the market price of volatility risk in the energy commodity markets. *Journal of Banking & Finance*, 32(12):2541–2552, 2008.

[49] Gabriel G Drimus. Options on realized variance by transform methods: a non-affine stochastic volatility model. *Quantitative Finance*, 12(11):1679–1694, 2012.

[50] Feike C Drost and Bas JM Werker. Closing the garch gap: Continuous time garch modeling. *Journal of Econometrics*, 74(1):31–57, 1996.

[51] Jefferson Duarte and Christopher S Jones. The price of market volatility risk. *Working Paper*, 2007.

[52] Darrell Duffie, Damir Filipović, Walter Schachermayer, et al. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

[53] Bernard Dumas, Jeff Fleming, and Robert E Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[54] Bruno Dupire. Pricing and hedging with smiles. *Mathematics of derivative securities*, 1(1):103–111, 1997.

[55] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[56] Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[57] Marcos Escobar. A stochastic volatility factor model of heston type. statistical properties and estimation. *Stochastics*, 90(2):172–199, 2018.

[58] Marcos Escobar, Hamidreza Arian, and Luis Seco. Creditgrades framework within stochastic covariance models. *Journal of Mathematical Finance*, 2(04):303, 2012. Scientific Research Publishing.

[59] Marcos Escobar, Sebastian Ferrando, and Alexey Rubtsov. Optimal investment under multi-factor stochastic volatility. *Quantitative Finance*, 17(2):241–260, 2017.

[60] Marcos Escobar, Barbara Götz, Luis Seco, and Rudi Zagst. Pricing a cdo on stochastically correlated underlyings. *Quantitative Finance*, 10(3):265–277, 2010.

[61] Marcos Escobar and Christoph Gschnaidtner. Parameters recovery via calibration in the heston model: A comprehensive review. *Wilmott*, 2016(86):60–81, 2016.

[62] Marcos Escobar-Anel and Harold A Moreno-Franco. Dynamic portfolio strategies under a fully correlated jump-diffusion process. *Annals of Finance*, 15(3):421–453, 2019.

[63] A. Eydeland and H. Geman. Pricing power derivatives. *Risk*, 1998.

[64] Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

[65] W. Feller. *An Introduction to Probability Theory And Its Applications, 2nd ED*. Number v. 2. Wiley India Pvt. Limited, 2008.

[66] Marcelo Fernandes, Marcelo C Medeiros, and Marcel Scharth. Modeling and predicting the cboe market volatility index. *Journal of Banking & Finance*, 40:1–10, 2014. Elsevier.

[67] Damir Filipović and Eberhard Mayerhofer. Affine diffusion processes: Theory and applications. *Radon Series Comp. Appl. Math*, 8:1–40, 2009.

[68] Teruo Fujioka. Asymptotic approximations of the inverse moment of the noncentral chi-squared variable. *Journal of the Japan Statistical Society*, 31(1):99–109, 2001. The Japan Statistical Society.

[69] Gianluca Fusai, Marina Marena, and Andrea Roncoroni. Analytical pricing of discretely monitored asian-style options: Theory and application to commodity markets. *Journal of Banking & Finance*, 32(10):2033–2045, 2018.

[70] A Ronald Gallant and George Tauchen. Which moments to match? *Econometric theory*, pages 657–681, 1996.

[71] Jim Gatheral. Consistent modeling of spx and vix options. In *Bachelier congress*, volume 37, pages 39–51, 2008.

[72] Paul Glasserman and Kyoung-Kuk Kim. Gamma expansion of the heston stochastic volatility model. *Finance and Stochastics*, 15(2):267–296, 2011.

[73] Alessandro Gnoatto, Martino Grasselli, and Eckhard Platen. A penny saved is a penny earned: Less expensive zero coupon bonds. *arXiv preprint arXiv:1608.04683*, 2018.

[74] Joanna Goard and Mathew Mazur. Stochastic volatility models and the pricing of vix options. *Mathematical Finance*, 23(3):439–458, 2013. Wiley Online Library.

[75] Christian Gouriéroux. Continuous time wishart process for stochastic risk. *Econometric Reviews*, 25(2-3):177–217, 2006.

[76] Martino Grasselli. The 4/2 stochastic volatility model: a unified approach for the heston and the 3/2 model. *Mathematical Finance*, 2016. Wiley Online Library.

[77] G.N. Gregoriou, C. Hoppe, and C.S. Wehn. *The Risk Modeling Evaluation Handbook: Rethinking Financial Risk Management Methodologies in the Global Capital Markets*. McGraw-Hill finance & investing. McGraw-Hill Education, 2010.

[78] Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

[79] Andrew Harvey, Esther Ruiz, and Neil Shephard. Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264, 1994.

[80] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993. Soc Financial Studies.

[81] Steven L Heston. A simple new formula for options with stochastic volatility. *Technical report*, 1997. Washington University at St. Louis.

[82] Samuel Hikspoors and Sebastian Jaimungal. Asymptotic pricing of commodity derivatives using stochastic volatility spot models. *Applied Mathematical Finance*, 15(5-6):449–477, 2008. Taylor & Francis.

[83] John Hull. *Risk management and financial institutions*, volume 743. John Wiley & Sons, 2015.

[84] John Hull and Alan White. Pricing interest-rate-derivative securities. *Review of financial studies*, 3(4):573–592, 1990. Soc Financial Studies.

[85] George J Jiang and Yisong S Tian. The model-free implied volatility and its information content. *The Review of Financial Studies*, 18(4):1305–1342, 2005.

[86] George J Jiang and Yisong S Tian. Extracting model-free volatility from option prices: An examination of the vix index. 2007.

[87] Andreas Kaeck and Carol Alexander. Continuous-time vix dynamics: On the role of stochastic volatility of volatility. *International Review of Financial Analysis*, 28:46–56, 2013.

[88] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 2014.

[89] Peter Kloeden and Andreas Neuenkirch. Convergence of numerical methods for stochastic differential equations in mathematical finance. In *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*, pages 49–80. World Scientific, 2013.

[90] H Felix Kloman. Risk management agonistes 1. *Risk Analysis*, 10(2):201–205, 1990.

[91] Sergey P Kolos and Ehud I Ronn. Estimating the commodity market price of risk for energy prices. *Energy Economics*, 30(2):621–641, 2008.

[92] Holger Kraft. Optimal portfolios and heston's stochastic volatility model: an explicit solution for power utility. *Quantitative Finance*, 5(3):303–313, 2005.

[93] Steven G Krantz and Steven George Krantz. *Handbook of complex variables*. Springer Science & Business Media, 1999.

[94] Terence C Langetieg. A multivariate model of the term structure. *The Journal of Finance*, 35(1):71–97, 1980.

[95] Linda Sandris Larsen. Optimal investment strategies in an international economy with stochastic interest rates. *International Review of Economics & Finance*, 19(1):145–165, 2010. Elsevier.

[96] Alan L Lewis. A simple option formula for general jump-diffusion and other exponential lévy processes. *Working paper, Envision Financial Systems and OptionCity.net.*

[97] Wei Lin, Shenghong Li, Xingguo Luo, and Shane Chern. Consistent pricing of vix and equity derivatives with the 4/2 stochastic volatility plus jumps model. *Journal of Mathematical Analysis and Applications*, 447(2):778–797, 2017.

[98] Yueh-Neng Lin and Chien-Hung Chang. Consistent modeling of s&p 500 and vix derivatives. *Journal of Economic Dynamics and Control*, 34(11):2302–2319, 2010.

[99] Roger Lord, Remmert Koekkoek, and Dick Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[100] E. Lukacs. *Characteristic Functions*. Griffin books of cognate interest. Hafner Publishing Company, 1970.

[101] Eugene Lukacs, Otto Szász, et al. On analytic characteristic functions. *Pacific J. Math*, 2(4):615–625, 1952.

[102] Xingguo Luo and Jin E Zhang. The term structure of vix. *Journal of Futures Markets*, 32(12):1092–1123, 2012.

[103] V. Mackevicius. *Stochastic Models of Financial Mathematics*. Elsevier Science, 2016.

[104] A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton Series in Finance. Princeton University Press, 2005.

[105] Nour Meddahi. A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17(5):479–508, 2002.

[106] Alexey Medvedev and Olivier Scaillet. Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility. *The review of Financial Studies*, 20(2):427–459, 2006.

[107] Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. In *Stochastic Optimization Models in Finance*, pages 621–661. Elsevier, 1975.

[108] Yuliya Mishura, Vladimir I Piterbarg, Kostiantyn Ralchenko, and Anton Yurchenko-Tytarenko. Stochastic representation and pathwise properties of fractional cox-ingersoll-ross process. *arXiv preprint arXiv:1708.02712*, 2017.

[109] Johannes Muhle-Karbe, Oliver Pfaffel, and Robert Stelzer. Option pricing in multivariate stochastic volatility models of ou type. *SIAM Journal on Financial Mathematics*, 3(1):66–94, 2012.

[110] Daniel B Nelson. *The time series behavior of stock market volatility and returns*. PhD thesis, Massachusetts Institute of Technology, 1988.

[111] Daniel B Nelson. Arch models as diffusion approximations. *Journal of econometrics*, 45(1-2):7–38, 1990.

[112] Anthony Neuberger. The log contract. *Journal of portfolio management*, 20:74–74, 1994.

[113] Lars Tyge Nielsen and Jesús Saá-Requejo. Exchange rate and term structure dynamics and the pricing of derivative securities. *Unpublished manuscript, INSEAD*, 1993.

[114] Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

[115] Marc S Paolella. *Intermediate probability: A computational approach*, chapter 10. John Wiley & Sons, 2007.

[116] Jim Pitman and Marc Yor. A decomposition of bessel bridges. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 59(4):425–457, 1982.

[117] E Platen. A non-linear stochastic volatility model. *Financial Mathematics Research Report No. FMRR005-97*, 1997. Center for Financial Mathematics, Australian National University, Canberra.

[118] Eckhard Platen and David Heath. *A benchmark approach to quantitative finance*. Springer Science & Business Media, 2006.

[119] J. Priolon. *Financial Markets for Commodities*. Wiley, 2019.

[120] Gennady Samorodnitsky. *Stochastic processes and long range dependence*, volume 26. Springer, 2016.

[121] Lorenz Schneider and Bertrand Tavin. From the samuelson volatility effect to a samuelson correlation effect: An analysis of crude oil calendar spread options. *Journal of Banking & Finance*, 95:185–202, 2018.

[122] Eduardo S Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997. Wiley Online Library.

[123] Elias M Stein and Jeremy C Stein. Stock price distributions with stochastic volatility: an analytic approach. *The review of financial studies*, 4(4):727–752, 1991.

[124] Alan Stuart, Steven Arnold, J Keith Ord, Anthony O'Hagan, and Jonathan Forster. *Kendall's advanced theory of statistics*. Wiley, 1994.

[125] Edward C Titchmarsh. Introduction to the theory of fourier integrals. 1948.

[126] Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188, 1977. Elsevier.

[127] Robert E. Whaley. Derivatives on market volatility. *The Journal of Derivatives*, 1(1):71–84, 1993.

[128] Robert E Whaley. Understanding the vix. *The Journal of Portfolio Management*, 35(3):98–105, 2009.

[129] Rudi Zagst. *Interest-rate management*. Springer Science & Business Media, 2013.

[130] Jin E Zhang and Yingzi Zhu. Vix futures. *Journal of Futures Markets*, 26(6):521–531, 2006.

[131] Lan Zhang, Per A Mykland, and Yacine Aït-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005.

Appendix A

Proofs for Theoretical Results in Chapter 3

Proof Proof of Proposition 3.2.1: Let $\tilde{v}(t) = e^{\beta t}v(t)$ and $\hat{v}(t) = e^{\beta t}ln(v(t))$, we have following:

$$d\tilde{v}(t) = \alpha\theta e^{\beta t}dt + (\beta - \alpha)e^{\beta t}v(t)dt + e^{\beta t}\xi\sqrt{v(t)}dB(t), \tag{A.1}$$

and

$$d\hat{v}(t) = \beta\hat{v}(t)dt + \frac{\partial\hat{v}(t)}{\partial v(t)}dv(t) + \frac{1}{2}\frac{\partial^2\hat{v}(t)}{\partial v(t)^2} < dv(t) >$$

$$= \beta\hat{v}(t)dt + \frac{e^{\beta t}\xi}{\sqrt{v(t)}}dB(t) + e^{\beta t}(\frac{\alpha\theta}{v(t)} - \alpha)dt - e^{\beta t}\frac{\xi^2}{2v(t)}dt. \tag{A.2}$$

From B.9 and B.10, we solve for $\int_t^T e^{\beta s}\sqrt{v(s)}dB(s)$ and $\int_t^T \frac{e^{\beta s}}{\sqrt{v(s)}}dB(s)$:

$$\int_t^T e^{\beta s}\sqrt{v(s)}dB(s) = \frac{e^{\beta T}v(T) - e^{\beta t}v(t)}{\xi} - \frac{\alpha\theta(e^{\beta T} - e^{\beta t})}{\beta\xi} - \frac{\beta - \alpha}{\xi}\int_t^T e^{\beta s}v(s)ds, \tag{A.3}$$

$$\int_t^T \frac{e^{\beta s}}{\sqrt{v(s)}}dB(s) = \frac{1}{\xi}ln\frac{v(T)^{e^{\beta T}}}{v(t)^{e^{\beta t}}} + \frac{\alpha(e^{\beta T} - e^{\beta t})}{\beta\xi} + \frac{1}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right)\int_t^T \frac{e^{\beta s}}{v(s)}ds - \frac{\beta}{\xi}\int_t^T e^{\beta s}ln(v(s))ds. \tag{A.4}$$

Split $W(t)$ into $B(t)$ and its orthogonal part $B(t)^\perp$:

$$Z(T) = Z(t) + \frac{L}{\beta}(e^{\beta T} - e^{\beta t}) + \int_t^T e^{\beta s}(c - \frac{1}{2})(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})^2 ds$$

$$+ \int_t^T e^{\beta s}(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})(\rho dB(s) + \sqrt{1 - \rho^2}dB(s)^\perp),$$

127

then substitute B.11 and B.12 to get rid of $dB(t)$. $Z(t)$ can be rewritten as:

$$
Z(T) = Z(t) + \left(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c - \frac{1}{2})\right)\frac{(e^{\beta T} - e^{\beta t})}{\beta} - \left[\frac{a\rho(\beta - \alpha)}{\xi} - a^2(c - \frac{1}{2})\right]
$$

$$
\times \int_t^T e^{\beta s}v(s)ds - \frac{b\rho\beta}{\xi}\int_t^T e^{\beta s}ln(v(s))ds + \frac{a\rho(e^{\beta T}v(T) - e^{\beta t}v(t))}{\xi} + \frac{b\rho}{\xi}ln\frac{v(T)^{e^{\beta T}}}{v(t)^{e^{\beta t}}}
$$

$$
+ \left[\frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right) + b^2(c - \frac{1}{2})\right]\int_t^T \frac{e^{\beta s}}{v(s)}ds + \sqrt{1 - \rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})dB(s)^\perp
$$

$$
(A.5)
$$

Let $(\mathcal{G}_t)_{t\geq 0}$ denote the filtration generated by $v(t), t \geq 0$. Using a property of iterated expectation, we can write the conditional moment generating function of $Z(t)$ as:

$$
\mathbb{E}(e^{uZ(T)}|\mathcal{F}_t) = \mathbb{E}\left[\mathbb{E}[e^{uZ(T)}|\mathcal{F}_t \bigcup \mathcal{G}_t]\Big|\mathcal{F}_t\right]
$$

$$
= exp\Big\{uZ(t) + u\Big(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c - \frac{1}{2})\Big)\frac{(e^{\beta T} - e^{\beta t})}{\beta}
$$

$$
- \frac{ua\rho}{\xi}e^{\beta t}v(t) - \frac{ub\rho e^{\beta t}}{\xi}ln(v(t))\Big\}
$$

$$
\times \mathbb{E}\Big[\mathbb{E}\Big[exp\Big\{-\frac{ub\rho\beta}{\xi}\int_t^T e^{\beta s}ln(v(s))ds + u\Big[\frac{b\rho}{\xi}\Big(\frac{\xi^2}{2} - \alpha\theta\Big) + b^2(c - \frac{1}{2})\Big]
$$

$$
\times \int_t^T \frac{e^{\beta s}}{v(s)}ds - u\Big[\frac{a\rho(\beta - \alpha)}{\xi} - a^2(c - \frac{1}{2})\Big]\int_t^T e^{\beta s}v(s)ds
$$

$$
+ u\sqrt{1 - \rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})dB(s)^\perp
$$

$$
+ \frac{ua\rho}{\xi}e^{\beta T}v(T) + \frac{ub\rho}{\xi}e^{\beta T}lnv(T)\Big\}\Big|\mathcal{F}_t \bigcup \mathcal{G}_t\Big]\Big|\mathcal{F}_t\Big]
$$

By independence of random variables, we have:

$$
\mathbb{E}\Big[\mathbb{E}[e^{uZ(T)}|\mathcal{F}_t \bigcup \mathcal{G}_t]\Big|\mathcal{F}_t\Big] = exp\Big\{uZ(t) + u\Big(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c - \frac{1}{2})\Big)\frac{(e^{\beta T} - e^{\beta t})}{\beta}
$$

$$
- \frac{ua\rho}{\xi}e^{\beta t}v(t) - \frac{ub\rho e^{\beta t}}{\xi}ln(v(t))\Big\}
$$

$$
\times \mathbb{E}\Big[exp\Big\{-\frac{ub\rho\beta}{\xi}\int_t^T e^{\beta s}ln(v(s))ds + u\Big[\frac{b\rho}{\xi}\Big(\frac{\xi^2}{2} - \alpha\theta\Big) + b^2(c - \frac{1}{2})\Big]
$$

$$
\times \int_t^T \frac{e^{\beta s}}{v(s)}ds - u\Big[\frac{a\rho(\beta - \alpha)}{\xi} - a^2(c - \frac{1}{2})\Big]\int_t^T e^{\beta s}v(s)ds
$$

$$
+ \frac{ua\rho}{\xi}e^{\beta T}v(T) + \frac{ub\rho}{\xi}e^{\beta T}lnv(T)\Big|\mathcal{F}_t\Big\}
$$

$$
\times \mathbb{E}\Big[exp\Big\{u\sqrt{1 - \rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}})dB(s)^\perp\Big\}\Big|\mathcal{F}_t \bigcup \mathcal{G}_t\Big]\Big|\mathcal{F}_t\Big]
$$

$$
(A.6)
$$

The inner double expectation of A.6 involves

$$\mathbb{E}\left[exp\left\{u\sqrt{1-\rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)}+\frac{b}{\sqrt{v(s)}})dB(s)^\perp\right\}\Big|\mathcal{F}_t\bigcup\mathcal{G}_t\right].$$

Conditioning on $\mathcal{F}_t\bigcup\mathcal{G}_t$ implies that $\sqrt{1-\rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)}+\frac{b}{\sqrt{v(s)}})dB(s)^\perp$ is a normal random variable with mean 0. By Ito's Isometry, the variance can be easily found to be $(1-\rho^2)\int_t^T e^{2\beta s}(a^2 v(s)+\frac{b^2}{v(s)}+2ab)ds$. So

$$\mathbb{E}[e^{u\sqrt{1-\rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)}+\frac{b}{\sqrt{v(s)}})dB(s)^\perp}|\mathcal{F}_t\bigcup\mathcal{G}_t] = e^{\frac{1}{2}u^2(1-\rho^2)\int_t^T e^{2\beta s}(a^2 v(s)+\frac{b^2}{v(s)}+2ab)ds}.$$

Consequently, the conditional moment generating function of $\sqrt{1-\rho^2}\int_0^t e^{\beta s}(a\sqrt{v(s)}+\frac{b}{\sqrt{v(s)}})dB(s)^\perp$ is found to be:

$$\mathbb{E}\left[\mathbb{E}\left[e^{u\sqrt{1-\rho^2}\int_t^T e^{\beta s}(a\sqrt{v(s)}+\frac{b}{\sqrt{v(s)}})dB(s)^\perp}\Big|\mathcal{F}_t\bigcup\mathcal{G}_t\right]\Big|\mathcal{F}_t\right] = \mathbb{E}[e^{\frac{1}{2}u^2(1-\rho^2)\int_t^T e^{2\beta s}(a^2 v(s)+\frac{b^2}{v(s)}+2ab)ds}|\mathcal{F}_t] \quad (A.7)$$

After reorganizing $\mathbb{E}(e^{uZ(T)}|\mathcal{F}_t)$, we have:

$$\begin{aligned}
\mathbb{E}(e^{uZ(T)}|\mathcal{F}_t) =&\, exp\Big\{uZ(t) + u\Big(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c-\frac{1}{2})\Big)\frac{(e^{\beta T}-e^{\beta t})}{\beta} + u^2(1-\rho^2) \\
&\times \frac{(e^{2\beta T}-e^{2\beta t})ab}{2\beta}\Big\} \times exp\Big\{-\frac{ua\rho}{\xi}e^{\beta t}v(t) - \frac{ub\rho e^{\beta t}}{\xi}ln(v(t))\Big\} \\
&\times \mathbb{E}\Big[exp\Big\{-\frac{ub\rho\beta}{\xi}\int_t^T e^{\beta s}ln(v(s))ds + u\Big[\frac{b\rho}{\xi}\Big(\frac{\xi^2}{2}-\alpha\theta\Big)+b^2(c-\frac{1}{2})\Big]\int_t^T \frac{e^{\beta s}}{v(s)}ds \quad (A.8) \\
&+ \frac{1}{2}u^2(1-\rho^2)a^2\int_t^T e^{2\beta s}v(s)ds - u\Big[\frac{a\rho(\beta-\alpha)}{\xi}-a^2(c-\frac{1}{2})\Big]\int_t^T e^{\beta s}v(s)ds \\
&+ \frac{1}{2}u^2(1-\rho^2)b^2\int_t^T \frac{e^{2\beta s}}{v(s)}ds + \frac{ua\rho}{\xi}e^{\beta T}v(T) + \frac{ub\rho}{\xi}e^{\beta T}lnv(T)\Big\}\Big|\mathcal{F}_t\Big].
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}(e^{uZ(t)}|\mathcal{F}_t) =&\, exp\Big\{uZ(t) + u\Big(L - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} + 2ab(c-\frac{1}{2})\Big)\frac{(e^{\beta T}-e^{\beta t})}{\beta} + u^2(1-\rho^2)\frac{(e^{2\beta T}-e^{2\beta t})ab}{2\beta}\Big\} \\
&\times exp\Big\{-\frac{ua\rho e^{\beta t}}{\xi}v(t) - \frac{ub\rho e^{\beta t}}{\xi}ln(v(t))\Big\} \times \mathbb{E}\Big[v(T)^{-A}exp\Big\{-Bv(T) - C\int_t^T e^{\beta s}ln(v(s))ds \\
&- \int_t^T (De^{\beta s}+Ee^{2\beta s})\frac{1}{v(s)}ds - \int_t^T (Fe^{\beta s}+Ge^{2\beta s})v(s)ds\Big\}\Big|\mathcal{F}_t\Big].
\end{aligned}$$

where,

$$\begin{aligned}
A &= -\frac{ub\rho}{\xi}e^{\beta T}, B = -\frac{ua\rho}{\xi}e^{\beta T}, C = \frac{ub\rho\beta}{\xi}, D = -\frac{ub\rho}{\xi}\Big(\frac{\xi^2}{2}-\alpha\theta\Big)-ub^2(c-\frac{1}{2}), \\
E &= -\frac{1}{2}u^2(1-\rho^2)b^2, F = \frac{ua\rho(\beta-\alpha)}{\xi}-ua^2(c-\frac{1}{2}), G = -\frac{1}{2}u^2(1-\rho^2)a^2.
\end{aligned}$$

$$(A.9)$$

∎

Proof Proof of the results in the example of Section 3.2.1.1: First:

$$
ln\left(\frac{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}\right)^{-\frac{2\alpha\theta}{\xi^2}}=\frac{2\alpha\theta}{\xi^2}ln\left(\frac{2\sqrt{A}e^{\frac{\sqrt{A}+\alpha}{2}(T-t)}}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}\right)
$$

$$
=\frac{2\alpha\theta}{\xi^2}\left\{\frac{\sqrt{A}+\alpha}{2}(T-t)+ln\left(\frac{2\sqrt{A}}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}\right)\right\}
$$

$$
=\frac{\alpha\theta}{\xi^2}\left\{(\sqrt{A}+\alpha)(T-t)-2ln\left(\frac{\sqrt{A}-\alpha+(\alpha+\sqrt{A})e^{\sqrt{A}(T-t)}}{2\sqrt{A}}\right)\right\}
$$

$$
=\frac{\alpha\theta}{\xi^2}\left\{(\sqrt{A}+\alpha)(T-t)-2ln\left(\frac{\alpha-\sqrt{A}-(\alpha+\sqrt{A})e^{\sqrt{A}(T-t)}}{-2\sqrt{A}}\right)\right\}
$$

$$
=\frac{\alpha\theta}{\xi^2}\left\{(\alpha+\sqrt{\alpha^2+\xi^2(u^2+iu)})(T-t)-2ln\left(\frac{1-ge^{-(T-t)\sqrt{\alpha^2+\xi^2(u^2+iu)}}}{1-g}\right)\right\}=C(T-t,u)
$$

Next:

$$
\frac{-(u^2+iu)(e^{\sqrt{A}(T-t)}-1)}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}=\frac{\frac{\alpha^2-A}{\xi^2}(e^{\sqrt{A}(T-t)}-1)}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}
$$

$$
=\frac{1}{\xi^2}\frac{(\alpha+\sqrt{A})(\alpha-\sqrt{A})(e^{\sqrt{A}(T-t)}-1)}{\alpha(e^{\sqrt{A}(T-t)}-1)+\sqrt{A}(e^{\sqrt{A}(T-t)}+1)}=\frac{\alpha+\sqrt{A}}{\xi^2}\times\frac{e^{\sqrt{A}(T-t)}-1}{\frac{\alpha}{(\alpha-\sqrt{A})}(e^{\sqrt{A}(T-t)}-1)+\frac{\sqrt{A}}{(\alpha-\sqrt{A})}(e^{\sqrt{A}(T-t)}+1)}
$$

$$
=\frac{\alpha+\sqrt{A}}{\xi^2}\times\frac{1-e^{\sqrt{A}(T-t)}}{1-ge^{\sqrt{A}(T-t)}}=D(T-t,u)
$$

∎

Proof Proof of Proposition 3.3.1:

    Option 1

    Now we define $\gamma(t)=\lambda\sqrt{v(t)}$ as in [80] for some constant λ such that $v(t)$ remains in the family of CIR processes under the risk-neutral measure. The process for volatility becomes:

$$
dv(t)=(\alpha\theta-(\alpha+\lambda\xi)v(t))dt+\xi\sqrt{v(t)}dB^{\mathbb{Q}}(t).
$$

Since $v(t)$ is unbounded from above, we need the conditions under which $\gamma(t)$ is well defined i.e $\mathbb{E}[exp(\frac{1}{2}\int_0^T\gamma^2(s)ds)]<\infty$. Proposition 5.1 of [92] provides a sufficient condition on λ that for $-\frac{\alpha}{\xi}\leq\lambda<\infty$, Novikov's condition is satisfied and hence,

$$
\mathbb{E}\left[exp\left(-\int_0^t\gamma(s)dB^{\mathbb{Q}}(s)-\frac{1}{2}\int_0^t\gamma^2(s)ds\right)\right]=1,
$$

which means $\gamma(t)=\lambda\sqrt{v(t)}$ is well defined for $-\frac{\alpha}{\xi}\leq\lambda<\infty$. When $\lambda\neq-\frac{\alpha}{\xi}$, Feller condition for volatility process after the change of measure is now $2(\alpha+\lambda\xi)\frac{\alpha\theta}{\alpha+\lambda\xi}=2\alpha\theta\geq\xi^2$, which is the same as we have under measure $\mathbb{P}$.

Next we need to find $\gamma(t)^{\perp}$. Under measure $\mathbb{Q}$, we have:

$$
\frac{dX(t)}{X(t)}=(L+c(a\sqrt{v(t)}+\frac{b}{\sqrt{v(t)}})^2-g(v(t))-\beta ln(X(t)))dt+(a\sqrt{v(t)}+\frac{b}{\sqrt{v(t)}})dW^{\mathbb{Q}}(t)
$$

where $g(v(t)) = (a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})\gamma(t)'$ and $\gamma(t)' = \rho\gamma(t) + \sqrt{1-\rho^2}\gamma(t)^\perp$. Since the drift term contains $(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2$, a natural choice of $g(v(t))$ is $g(v(t)) = \Lambda(a \sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2$. Substitute $\gamma(t)$ for $\lambda \sqrt{v(t)}$ and solve for $\gamma(t)^\perp$:

$$\gamma(t)^\perp = \frac{1}{\sqrt{1-\rho^2}}\left[(\Lambda a - \rho\lambda) \sqrt{v(t)} + \frac{\Lambda b}{\sqrt{v(t)}}\right].$$

$\gamma(t)^\perp$ is a linear combination of $\sqrt{v(t)}$ and $\frac{1}{\sqrt{v(t)}}$, we then write $\gamma(t)^\perp = \Lambda_1 \sqrt{v(t)} + \frac{\Lambda_2}{\sqrt{v(t)}}$,

$$\mathbb{E}\left[exp\left(\frac{1}{2}\int_0^T \gamma^2(s)^\perp ds\right)\right] = \mathbb{E}\left\{exp\left[\frac{1}{2}\int_0^T (\Lambda_1^2 v(t) + 2\Lambda_1\Lambda_2 + \frac{\Lambda_2^2}{v(t)})ds\right]\right\}$$
$$= e^{\Lambda_1\Lambda_2 T}\mathbb{E}\left[exp\left(\frac{\Lambda_1^2}{2}\int_0^T v(s)ds + \frac{\Lambda_2^2}{2}\int_0^T \frac{1}{v(s)}ds\right)\right].$$

Theorem A.1 of [76] gives a closed form solution for

$$\mathbb{E}\left[v(T)^{-q}e^{-lv(T)-m\int_0^T v(s)ds-n\int_0^T \frac{1}{v(s)}ds}\right].$$

With $q = 0$ and $l = 0$, the above expectation exists as long as $m, n \in \mathbb{R}$, $m > -\frac{\alpha^2}{2\xi^2}, n \geq -\frac{(2\alpha\theta-\xi^2)^2}{8\xi^2}$. In our case $m = -\frac{\Lambda_1^2}{2}$ and $n = -\frac{\Lambda_2^2}{2}$. From the range of m and n we have for $|\Lambda_1| < \frac{\alpha}{\xi}$ and $|\Lambda_2| \leq \frac{2\alpha\theta-\xi^2}{2\xi}$, $\mathbb{E}\left[exp\left(\frac{\Lambda_1^2}{2}\int_0^T v(s)ds + \frac{\Lambda_2^2}{2}\int_0^T \frac{1}{v(s)}ds\right)\right]$ is finite and thus defined. We combine the range of Λ_1 and Λ_2 to get the constrain on Λ: $\Lambda \in \{\Lambda_3 < \Lambda < \Lambda_4\}$ where $\Lambda_3 = max\left[\frac{1}{a}(-\frac{\alpha}{\xi}+\rho\lambda), -\frac{1}{|b|}(\frac{\alpha\theta}{\xi}-\frac{\xi}{2})\right]$ and $\Lambda_4 = min\left[\frac{1}{a}(\frac{\alpha}{\xi}+\rho\lambda), \frac{1}{|b|}(\frac{\alpha\theta}{\xi}-\frac{\xi}{2})\right]$ for a and b non-zero.

Option 2

We define $\gamma(t) = \frac{\lambda'}{\sqrt{v(t)}}$ for some constant λ'. This change will also keep $v(t)$ within CIR process family under risk-neutral measure by subtracting a constant to mean-reverting level:

$$dv(t) = (\alpha\theta - \lambda'\xi - \alpha v(t))dt + \xi \sqrt{v(t)}dB^\mathbb{Q}(t).$$

The corresponding Feller condition will be $2\alpha\frac{\alpha\theta-\lambda'\xi}{\alpha} = 2(\alpha\theta - \lambda'\xi) \geq \xi^2$ when $\lambda' \leq 0$. Let's check if Novikov's condition $\mathbb{E}[exp(\frac{\lambda'^2}{2}\int_0^T \frac{1}{v(t)}ds)] < \infty$ is satisfied. Note that $\frac{1}{v(t)} = v(t)'$ is a 3/2 process with

$$dv(t)' = v(t)'\left(\alpha - (\alpha\theta - \xi^2)v(t)'\right)dt - \xi v(t)'^{\frac{3}{2}}dB^\mathbb{P}(t).$$

[30] showed that the conditional Fourier-Laplace transformation of joint log return, $R(T)$, and integrated variance, $\mathbb{E}[e^{iuR(T)-s\int_t^T v(z)'dz}|\mathcal{F}_t]$ have a closed-form solution for stochastic volatility model with volatility process being 3/2 process. In our case, setting $u = 0$ and substituting $s = -\frac{\lambda'^2}{2}$, we get $\mathbb{E}\left[exp\left(\frac{\lambda'^2}{2}\int_0^T \frac{1}{v(s)}ds\right)\right]$ and the expectation is finite as long as $_1F_1\left(x; w; -\frac{2}{\xi^2 y(0,v(0))}\right)$ exists, in another word w is positive. The problem simplifies to solve the inequality $\left(\frac{1}{2}+\frac{\alpha\theta-\xi^2}{\xi^2}\right)^2 - \frac{\lambda'^2}{\xi^2} \geq 0$ for the range of λ': $|\lambda'| \leq \frac{\alpha\theta}{\xi} - \frac{\xi}{2}$. So when $-(\frac{\alpha\theta}{\xi}-\frac{\xi}{2}) \leq \lambda' \leq 0$, $\gamma(t) = \frac{\lambda'}{\sqrt{v(t)}}$ is well

defined.

Let $g(v(t)) = \Lambda'(a\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}})^2$. We repeat the procedures in **Option 1** for solving $\gamma(t)^{\perp}$:

$$\gamma(t)^{\perp} = \frac{1}{\sqrt{1-\rho^2}}\left[(\Lambda'a\sqrt{v(t)} + \frac{\Lambda'b - \rho\lambda'}{\sqrt{v(t)}}\right].$$

In this case, $\Lambda_1 = \Lambda'a$ and $\Lambda_2 = \Lambda'b - \rho\lambda'$, the constrain on Λ is $\Lambda \in \left\{\Lambda_3 < \Lambda' < \Lambda_4\right\}$ where $\Lambda_3 = max\left[-\frac{1}{b}(\frac{\alpha\theta}{\xi} - \frac{\xi}{2} - \rho\lambda'), -\frac{\alpha}{|a|\xi}\right]$ and $\Lambda_4 = min\left[\frac{1}{b}(\frac{\alpha\theta}{\xi} - \frac{\xi}{2} + \rho\lambda'), \frac{\alpha}{|a|\xi}\right]$ for a and b non-zero.

Data Descriptions:

VVIX Data Description: Data is obtained from *Cboe* website. Dataset consists of daily data starting from January 3, 2007 to March 13, 2018, 2817 entries in total excluding November 11, 2010 data that is not available.

V-VSTOXX Data Description: Euro STOXX 50 Volatility of volatility (V-VSTOXX) historical daily data is available on STOXX. The data is recorded from March 23, 2010 to March 13, 2018 with 15 missing data entries. The missing data entries are not evenly distributed throughout the time frame. Majority of the missing data are in 2017. Almost half month data of August 2017 are not available. The size of data is 2017 after excluding missing data.

Volatility Index of Crude Oil ETF Data Description: According to *Cboe* website, "the Cboe Crude Oil ETF Volatility index (USO) measures the market's expectation of 30-day volatility of crude oil prices by applying the VIX methodology to United States Oil Fund, LP options spanning a wide range of strike prices." We obtain daily data from May 10, 2007 to March 13, 2018 with 2730 entries in total.

Volatility Index of Gold ETF Data Description: Similar to USO, the Cboe Gold ETF Volatility index (GVZ) measures the market's expectation of 30-day volatility of gold prices by applying the VIX methodology Gold ETF options spanning a wide range of strike prices. We obtain daily data from June 3, 2008 to May 15, 2019 with 2757 entries in total.

Appendix B

Proofs and Helpful Results for Chapter 4

B.1 Proofs

Proof Proof of Proposition 4.2.1.

The first step is to ensure the change of measure is well-defined and for this we use Novikov's condition, i.e., generically

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \lambda^2 \left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)^2 ds\right)\right] = e^{\lambda^2 bT}\mathbb{E}\left[\exp\left(\frac{\lambda^2}{2}\int_0^T v(s)ds + \frac{\lambda^2 b^2}{2}\int_0^T \frac{1}{v(s)}ds\right)\right] < \infty.$$

From Grasselli, in order for this expectation to exist, we need two conditions:

$$-\frac{\lambda^2}{2} > -\frac{\alpha^2}{2\xi^2} \implies |\lambda| < \frac{\alpha}{\xi} \tag{B.1}$$

and

$$-\frac{\lambda^2 b^2}{2} \geq -\frac{(2\alpha\theta - \xi^2)^2}{8\xi^2} \implies |\lambda| \leq \frac{2\alpha\theta - \xi^2}{2|b|\xi} \implies \xi^2 \leq 2\alpha\theta - 2|\lambda||b|\xi \tag{B.2}$$

The latter condition in Equation (B.2) implies, in particular, that our volatility processes satisfy Feller's condition under $\mathbb{P}$ and $\mathbb{Q}$; in other words, it ensures all our CIR processes stay away from zero under both measures.

Applying Equation (B.2) to our setting leads to ($i, j = 1, .., n$):

$$\xi_j^2 \leq 2\alpha_j\theta_j - 2\xi_j \max\left\{\left|\lambda_j b_j\right|, \left|\lambda_j^\perp b_j\right|\right\} \tag{B.3}$$

$$\widetilde{\xi}_i^2 \leq 2\widetilde{\alpha}_i\widetilde{\theta}_i - 2\widetilde{\xi}_i \max\left\{\left|\widetilde{\lambda}_i\widetilde{b}_i\right|, \left|\widetilde{\lambda}_i^\perp\widetilde{b}_i\right|\right\} \tag{B.4}$$

Now, we apply Equation (B.1) producing two extra set of conditions ($i, j = 1, .., n$):

$$\max\left\{\left|\lambda_j\right|, \left|\lambda_j^\perp\right|\right\} < \frac{\alpha_j}{\xi_j} \tag{B.5}$$

$$\max\left\{\left|\widetilde{\lambda}_i\right|, \left|\widetilde{\lambda}_i^\perp\right|\right\} < \frac{\widetilde{\alpha}_i}{\widetilde{\xi}_i} \tag{B.6}$$

The second step applies to the case $\beta_{ij} = 0$ for $i, j = 1, .., n$ and it is to ensure the drift of the asset price equal the short rate:

$$L_i = r, \, c_i = \sum_{j=1}^{n} a_{ij}\left(\rho_j \lambda_j + \sqrt{1 - \rho_j^2}\lambda_j^\perp\right), \, \widetilde{c}_i = \widetilde{\rho}_i\widetilde{\lambda}_i + \sqrt{1 - \widetilde{\rho}_i^2}\widetilde{\lambda}_i^\perp$$

For the most general case ($\beta_{ij} \neq 0$ for some i or j), the second step should be adapted to any particular prescribed drift structure under the $\mathbb{Q}$-measure.

The third step is to ensure the drift-less asset price process is a true $\mathbb{Q}$-martingale and not just a local $\mathbb{Q}$-martingale:

$$\frac{dX_i(t)}{X_i(t)} = (.)\,dt + \sum_{j=1}^{n} a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j^Q(t) + \left(\sqrt{\widetilde{v}_i(t)} + \frac{\widetilde{b}_i}{\sqrt{\widetilde{v}_i(t)}}\right)d\widetilde{W}_i^Q(t)$$

Here, we test the martingale property using the Feller nonexplosion test for volatilities, hence considering the following $n^2 + n$ changes of Brownian motion for the volatility processes and checking the processes do not reach zero under the various measures:

$$dB_{ij}^Q(t) = a_{ij}\rho_j\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dt + dB_j^P(t), \, d\widetilde{B}_i^Q(t) = \widetilde{\rho}_i\left(\sqrt{\widetilde{v}_i(t)} + \frac{\widetilde{b}_i}{\sqrt{\widetilde{v}_i(t)}}\right)dt + d\widetilde{B}_i^P(t)$$

This leads to the following conditions:

$$\xi_j^2 \leq 2\alpha_j\theta_j - 2\left|a_{ij}\rho_j b_j\right|\xi_j, \, i, j = 1, ..., n \tag{B.7}$$

$$\widetilde{\xi}_i^2 \leq 2\widetilde{\alpha}_i\widetilde{\theta}_i - 2\left|\widetilde{\lambda}_i\widetilde{\rho}_i\widetilde{b}_i\right|\widetilde{\xi}_i, \, i = 1, ..., n \tag{B.8}$$

We can combine the first and third steps in Equations (B.3), (B.7), (B.4) and (4.4) into the final conditions.

Proof Proof of Proposition 4.2.2.
We start by defining new processes $dS_{ik,t}$ and dS_t^{ij} with $i, j, k = 1, 2, \ldots, n$ and $t \geq 0$.

$$dZ_i(t) = \sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \left[\left(\widetilde{c}_j - \frac{1}{2}\right)\left(\sqrt{\widetilde{v}_j(t)} + \frac{\widetilde{b}_j}{\sqrt{\widetilde{v}_j(t)}}\right) dt + \left(\sqrt{\widetilde{v}_j(t)} + \frac{\widetilde{b}_j}{\sqrt{\widetilde{v}_j(t)}}\right) d\widetilde{W}_j(t)\right]$$

$$+ \sum_{k=1}^{n} \left[\left(\sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \frac{L_j}{n}\right) dt + \left(\sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \left(c_k - \frac{1}{2}\right) a_{jk}^2\right)\left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right)^2 dt\right]$$

$$+ \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} a_{jk}\right)\left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right) dW_k(t)$$

$$= \sum_{j=1}^{n} dS_t^{ij} + \sum_{k=1}^{n} dS_{ik,t}$$

By the dependence structure implied by the model, it follows that all S are independent for a fix i, hence we can transform the characteristic function using the processes S as follows:

$$\Phi_{Z(t),v(t)}(T, \omega) = \prod_{k=1}^{n} E\left[\exp\{i\omega'\left(S_{\cdot k,T} - S_{\cdot k,t}\right)\} \mid S_t, v(t)\right] \prod_{j=1}^{n} E\left[\exp\{i\omega_i\left(S_T^{\cdot j} - S_t^{\cdot j}\right)\} \mid S_t, v(t)\right]$$

For each factor $j = 1, 2, \ldots, n$ we define $S_{k,t}^* = \omega' S_{\cdot k,t} = \sum_{i=1}^{n} \omega_i S_{ik,t}$; the dynamics of $S_{k,t}^*$ can be expressed as

$$dS_{k,t}^* = \omega' dS_{\cdot k,t}$$

$$= \left[L(\omega, t) + h_k(\omega, t)\left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right)^2\right] dt + g_k(\omega, t)\left(\sqrt{v_k(t)} + \frac{b_k}{\sqrt{v_k(t)}}\right) dW_{k,t}$$

where,

$$h_k(\omega, t) = \sum_{j=1}^{n} \left(c_k - \frac{1}{2}\right) a_{jk}^2 f_j(\omega, t), \; L(\omega, t) = \sum_{j=1}^{n} \frac{L_j}{n} f_j(\omega, t), \; g_k(\omega, t) = \sum_{j=1}^{n} a_{jk} f_j(\omega, t), \; f_j(\omega, t) = \sum_{i=1}^{n} \omega_i \left(e^{\beta t}\right)_{ij}.$$

These three functions are deterministic, linear combinations of $f_j(\omega, t)$.

Next, we find the characteristic function for the increments of $S_{k,t}^*$:

$$E\left[\exp\{i\phi\left(S_{k,T}^* - S_{k,t}^*\right)\} \mid S_{k,t}^*, v_k(t) = v_{k,t}\right] = \Phi_{GG}\left(T, \phi; L(\omega), h_k(\omega), g_k(\omega), \kappa_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, S_{k,t}^*\right).$$

The generic function Φ_{GG} is provided in Lemma B.1.1. It follows similarly for idiosyncratic

factors:

$$dS_t^{*j} = \omega' dS_t^{\cdot j}$$

$$= \left(\sum_{i=1}^{n} \omega_i \left(\sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \left(\tilde{c}_j - \frac{1}{2}\right) \right) \right) \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}} \right) dt$$

$$+ \left(\sum_{i=1}^{n} \omega_i \sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \right) \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}} \right) d\widetilde{W}_{j,t}$$

$$= h_j(\omega, t) \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}} \right)^2 dt + g_j(\omega, t) \left(\sqrt{\tilde{v}_j(t)} + \frac{\tilde{b}_j}{\sqrt{\tilde{v}_j(t)}} \right) d\widetilde{W}_{j,t}$$

where $h_j(\omega, t) = \sum_{i=1}^{n} \omega_i \left(\sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij} \left(\tilde{c}_j - \frac{1}{2}\right) \right)$ and $g_j(\omega, t) = \sum_{i=1}^{n} \omega_i \sum_{j=1}^{n} \left(e^{\beta t}\right)_{ij}$.

Combining all pieces together, we obtain:

$$\Phi_{Z(t),v(t)}(T, \omega) = \prod_{k=1}^{n} E\left[\exp\{i\omega'(S_{\cdot k,T} - S_{\cdot k,t})\} \mid S_t, v(t)\right] \prod_{j=1}^{n} E\left[\exp\{i\omega'(S_T^{\cdot j} - S_t^{\cdot j})\} \mid S_t, v(t)\right]$$

$$= \prod_{k=1}^{n} \Phi_{GG}\left(T, 1; L_k(\omega), h_k(\omega), g_k(\omega), \kappa_k, \theta_k, \xi_k, \rho_k, b_k, c_k, v_{k,t}, S_{k,t}^*\right)$$

$$\times \prod_{j=1}^{n} \Phi_{GG}\left(T, 1; 0, h_j(\omega), g_j(\omega), \widetilde{\kappa}_j, \widetilde{\theta}_j, \widetilde{\xi}_j, \widetilde{\rho}_j, \widetilde{b}_j, \widetilde{c}_j, \widetilde{v}_{j,t}, S_t^{*j}\right)$$

Lemma B.1.1 *Let the generic process be:*

$$dZ(t) = \left[L(t) + h(t) \left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right)^2 \right] dt + g(t) \left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right) dW_t$$

$$dv(t) = \alpha(\theta - v(t))dt + \xi \sqrt{v(t)}dB(t)$$

$$\langle dB(t), dW(t) \rangle = \rho dt$$

with $g(t)$ differentiable, then

$$\Phi_{GG}(T, \phi; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_t, Z_t) = \exp\left\{ u \int_t^T A(s)\,ds \right\} v(t)^{u\frac{b\rho}{\xi}g(t)} \exp\left\{ -u\rho \frac{g(t)v(t)}{\xi} \right\}$$

$$\times \mathbb{E}\left[v(T)^{u\frac{b\rho}{\xi}g(T)} \exp\left\{ u\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds + \rho \frac{g(T)v(T)}{\xi} \right) \right\} \mid \mathcal{F}_t \right]$$

where A, B, C and D are provided in the proof.

Proof Proof of Lemma B.1.1.
Let the generic process be:

$$dZ(t) = \left[L(t) + h(t) \left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right)^2 \right] dt + g(t) \left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right) dW_t$$

$$dv(t) = \alpha(\theta - v(t))dt + \xi \sqrt{v(t)}dB(t)$$

$$\langle dB(t), dW(t) \rangle = \rho dt$$

We want to find

$$\mathbb{E}\left[e^{uZ(T)}|\mathcal{F}_t\right] = e^{uZ(t)}\Phi_{GG}\left(T,\phi;L,h,g,\kappa,\theta,\xi,\rho,b,c,v_t,Z_t\right)$$

Letting $\tilde{v}(t) = g(t)v(t)$ and $\hat{v}(t) = g(t)\ln(v(t))$, we have following:

$$d\tilde{v}(t) = \alpha\theta g(t)\,dt + (g'(t) - \alpha g(t))v(t)dt + g(t)\xi\sqrt{v(t)}dB(t), \qquad (B.9)$$

and

$$\begin{aligned}
d\hat{v}(t) &= \frac{g'(t)}{g(t)}\hat{v}(t)dt + \frac{\partial\hat{v}(t)}{\partial\ln(v(t))}d\ln(v(t)) + \frac{1}{2}\frac{\partial^2\hat{v}(t)}{\partial\ln(v(t))^2} < d\ln(v(t)) > \qquad (B.10)\\
&= \frac{g'(t)}{g(t)}\hat{v}(t)dt + \frac{g(t)\xi}{\sqrt{v(t)}}dB(t) + g(t)\left(\frac{\alpha\theta}{v(t)} - \alpha\right)dt - g(t)\frac{\xi^2}{2v(t)}dt
\end{aligned}$$

From Equations (B.9) and (B.10), we solve for $\int_t^T g(s)\sqrt{v(s)}dB(s)$ and $\int_t^T \frac{g(s)}{\sqrt{v(s)}}dB(s)$:

$$\int_t^T g(s)\sqrt{v(s)}dB(s) = \frac{g(T)v(T) - g(t)v(t)}{\xi} - \frac{\alpha\theta}{\xi}\int_t^T g(s)\,ds - \frac{1}{\xi}\int_t^T (g'(s) - \alpha g(s))v(s)ds, \qquad (B.11)$$

$$\int_t^T \frac{g(s)}{\sqrt{v(s)}}dB(s) = \frac{1}{\xi}\ln\frac{v(T)^{g(T)}}{v(t)^{g(t)}} + \frac{\alpha}{\xi}\int_t^T g(s)\,ds + \frac{1}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right)\int_t^T \frac{g(s)}{v(s)}ds - \frac{1}{\xi}\int_t^T g'(s)\ln(v(s))ds. \quad (B.12)$$

Split $W(t)$ into $B(t)$ and its orthogonal part $B(t)^\perp$:

$$\begin{aligned}
Z(T) = Z(t) &+ \int_t^T L(s)\,ds + \int_t^T h(s)\left(\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}}\right)^2 ds\\
&+ \int_t^T g(s)\left(\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}}\right)(\rho dB(s) + \sqrt{1-\rho^2}dB(s)^\perp),
\end{aligned}$$

then substitute Equation (B.11) and (B.12) to eliminate $dB(t)$. $Z(t)$ can be rewritten now as:

$$\begin{aligned}
Z(T) = Z(t) &+ \int_t^T L(s)\,ds + \int_t^T h(s)\left(v(s) + 2b + \frac{b^2}{v(s)}\right)ds + \sqrt{1-\rho^2}\int_t^T g(s)\left(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}}\right)dB(s)^\perp\\
&+ \rho\frac{g(T)v(T) - g(t)v(t)}{\xi} - \frac{\alpha\theta\rho}{\xi}\int_t^T g(s)\,ds - \frac{\rho}{\xi}\int_t^T (g'(s) - \alpha g(s))v(s)ds\\
&+ \frac{b\rho}{\xi}\ln\frac{v(T)^{g(T)}}{v(t)^{g(t)}} + \frac{\alpha b\rho}{\xi}\int_t^T g(s)\,ds + \frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right)\int_t^T \frac{g(s)}{v(s)}ds - \frac{b\rho}{\xi}\int_t^T g'(s)\ln(v(s))ds
\end{aligned}$$

Grouping conveniently, we obtain:

$$\begin{aligned}
Z(T) = Z(t) &+ \int_t^T \overline{A}(s)\,ds + \int_t^T \overline{B}(s)v(s)ds + \int_t^T \overline{C}(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds\\
&+ \frac{b\rho}{\xi}\ln\frac{v(T)^{g(T)}}{v(t)^{g(t)}} + \rho\frac{g(T)v(T) - g(t)v(t)}{\xi} + \sqrt{1-\rho^2}\int_t^T g(s)\left(a\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}}\right)dB(s)^\perp
\end{aligned}$$

where

$$\begin{aligned}
\overline{A}(s) &= L(s) + 2bh(s) + \left(\frac{\alpha b\rho}{\xi} - \frac{\alpha\theta\rho}{\xi}\right)g(s)\\
\overline{B}(s) &= h(s) - \frac{\rho}{\xi}(g'(s) - \alpha g(s))\\
\overline{C}(s) &= b^2 h(s) + \frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta\right)g(s)\\
D(s) &= -\frac{b\rho}{\xi}g'(s)
\end{aligned}$$

Let $(\mathcal{G}_t)_{t\geq 0}$ denote the filtration generated by $v(t), t \geq 0$. Using iterated expectation and independence, we can write the conditional moment generating function of $Z(T)$ as:

$$\mathbb{E}\left[e^{uZ(T)}|\mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{E}\left[e^{uZ(T)}\mid\mathcal{F}_t\bigcup\mathcal{G}_t\right]\mid\mathcal{F}_t\right] = \exp\left\{u\left(Z(t) + \int_t^T A(s)\,ds\right)\right\}v(t)^{u\frac{b\rho}{\xi}g(t)}\exp\left\{-u\rho\frac{g(t)v(t)}{\xi}\right\}$$

$$\times\mathbb{E}[\exp\left\{u\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds + \frac{b\rho}{\xi}\ln v(T)^{g(T)} + \rho\frac{g(T)v(T)}{\xi}\right)\right\}$$

$$\times\mathbb{E}\left[\exp\left\{u\sqrt{1-\rho^2}\int_t^T g(s)\left(\sqrt{v(s)} + \frac{b}{\sqrt{v(s)}}\right)dB(s)^{\perp}\right\}\mid\mathcal{F}_t\bigcup\mathcal{G}_t\right]\mid\mathcal{F}_t]$$

The inner expectation, conditioned on $\mathcal{F}_t\bigcup\mathcal{G}_t$, leads to a normal random variable with mean 0 and variance (Ito's Isometry) $u^2(1-\rho^2)\int_t^T g^2(s)(v(s) + \frac{b^2}{v(s)} + 2b)ds$. Putting all together:

$$\mathbb{E}\left[e^{uZ(T)}|\mathcal{F}_t\right] = \exp\left\{u\left(Z(t) + \int_t^T A(s)\,ds\right)\right\}v(t)^{u\frac{b\rho}{\xi}g(t)}\exp\left\{-u\rho\frac{g(t)v(t)}{\xi}\right\}$$

$$\times\mathbb{E}\left[v(T)^{u\frac{b\rho}{\xi}g(T)}\exp\left\{u\left(\int_t^T B(s)v(s)ds + \int_t^T C(s)\frac{1}{v(s)}ds + \int_t^T D(s)\ln(v(s))ds + \rho\frac{g(T)v(T)}{\xi}\right)\right\}\mid\mathcal{F}_t\right]$$

where

$$A(s) = \overline{A}(s) + u(1-\rho^2)bg^2(s)$$
$$B(s) = \overline{B}(s) + \frac{1}{2}u^2(1-\rho^2)g^2(s)$$
$$C(s) = \overline{C}(s) + \frac{1}{2}u^2(1-\rho^2)b^2g^2(s)$$

Proof Proof of Corollary 4.2.3.

The proof starts similarly to Proposition 4.2.2. We start by defining new processes $dS_{ij,t}$ and dS_t^i with $i, j = 1, 2, \ldots, n$ and $t \geq 0$.

$$dZ_i(t) = \left[\left(\widetilde{c}_i - \frac{1}{2}\right)\left(\sqrt{\widetilde{v}_i(t)} + \frac{\widetilde{b}_i}{\sqrt{\widetilde{v}_i(t)}}\right)dt + \left(\sqrt{\widetilde{v}_i(t)} + \frac{\widetilde{b}_i}{\sqrt{\widetilde{v}_i(t)}}\right)d\widetilde{W}_i\right]$$

$$+ \sum_{j=1}^n\left[\frac{L_i}{n}dt + \left(c_j - \frac{1}{2}\right)a_{ij}^2\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2 dt + a_{ij}\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)dW_j\right]$$

$$= dS_t^i + \sum_{j=1}^n dS_{ij,t}$$

By the dependence structure implied by the model, it follows that all S are independent for a fix i, hence we can transform the characteristic function using the processes S as follows:

$$\Phi_{Z(t),v(t)}(T,\omega) = \prod_{j=1}^n E\left[\exp\{i\omega'\left(S_{\cdot j,T} - S_{\cdot j,t}\right)\} \mid S_t, v(t)\right]\prod_{i=1}^n E\left[\exp\{i\omega_i\left(S_T^i - S_t^i\right)\} \mid S_t, v(t)\right]$$

For each factor $j = 1, 2, \ldots, n$ we define $S_{j,t}^* = \omega' S_{\cdot j,t} = \sum_{i=1}^n \omega_i S_{ij,t}$, the dynamics of $S_{j,t}^*$ can be expressed as

$$dS_{j,t}^* = \omega'dS_{\cdot j,t}$$

$$= \left[L(\omega) + h_j(\omega)\left(\sqrt{v_j(t)} + \frac{b_j}{\sqrt{v_j(t)}}\right)^2\right]dt + g_j(\omega)\left(\sqrt{v_j(t)} + \frac{b_k}{\sqrt{v_j(t)}}\right)dW_{j,t}$$

where $h_j(\omega) = \sum_{i=1}^{n} \omega_i \left(c_i - \frac{1}{2}\right) a_{ij}^2$, $L(\omega) = \sum_{i=1}^{n} \omega_i \frac{L_i}{n}$ and $g_j(\omega) = \sum_{i=1}^{n} \omega_i a_{ij}$. Next, we find the characteristic function for the increments of $S_{j,t}^{*}$:

$$E\left[\exp\{i\phi\left(S_{j,T}^{*} - S_{j,t}^{*}\right)\} \mid S_{j,t}^{*}, v_j(t) = v_{j,t}\right] = \Phi_G\left(T, \phi; L, h_j, g_j, \kappa_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,t}, S_{j,t}^{*}\right).$$

The generic function Φ_G is provided in the Appendix B.2.

Similarly for the idiosyncratic factors dS_t^i with $i = 1, 2, \ldots, n$:

$$dS_t^{*i} = \omega_i dS_t^i = h_i(\omega)\left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}}\right)^2 dt + g_i(\omega)\left(\sqrt{\tilde{v}_i(t)} + \frac{\tilde{b}_i}{\sqrt{\tilde{v}_i(t)}}\right) d\widetilde{W}_{i,t}$$

where $h_i(\omega) = \omega_i\left(\tilde{c}_i - \frac{1}{2}\right)$ and $g_i(\omega) = \omega_i$.

Combining all pieces together, we obtain:

$$\Phi_{Z(t),v(t)}(T, \omega) = \prod_{j=1}^{n} E\left[\exp\{i\omega'\left(S_{\cdot j,T} - S_{\cdot j,t}\right)\} \mid S_t, v(t)\right] \prod_{i=1}^{n} E\left[\exp\{i\omega_i\left(S_T^i - S_t^i\right)\} \mid S_t, v(t)\right]$$

$$= \prod_{j=1}^{n} \Phi_G\left(T, \phi; L, h_j, g_j, \kappa_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,t}, S_{j,t}^{*}\right) \times \prod_{i=1}^{n} \Phi_G\left(T, 1; 0, h_i, g_i, \widetilde{\kappa}_i, \widetilde{\theta}_i, \widetilde{\xi}_i, \widetilde{\rho}_i, \widetilde{b}_i, \widetilde{c}_i, \widetilde{v}_{i,t}, S_t^{*i}\right)$$

Proof Proof of Corollary 4.2.4.

The proof uses Corollary 4.2.3, where we express the joint c.f. as the product of one dimensional c.f.s of 4/2 type.

$$\Phi_{Z(t),v(t)}(T, \omega) = \prod_{j=1}^{n} E\left[\exp\{i\omega'\left(S_{\cdot j,T} - S_{\cdot j,t}\right)\} \mid S_t, v(t)\right] \prod_{i=1}^{n} E\left[\exp\{i\omega_i\left(S_T^i - S_t^i\right)\} \mid S_t, v(t)\right]$$

$$= \prod_{j=1}^{n} \Phi_G\left(T, \phi; L, h_j, g_j, \kappa_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,t}, S_{j,t}^{*}\right) \times \prod_{i=1}^{n} \Phi_G\left(T, 1; 0, h_i, g_i, \widetilde{\kappa}_i, \widetilde{\theta}_i, \widetilde{\xi}_i, \widetilde{\rho}_i, \widetilde{b}_i, \widetilde{c}_i, \widetilde{v}_{i,t}, S_t^{*i}\right)$$

Hence, every one of these functions $(\Phi_G\left(T, \phi; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_t, Z_t\right) = \mathbb{E}\left[e^{uZ(T)} \mid \mathcal{F}_t\right])$ capture the c.f. of a process of the type:

$$dZ(t) = \left[L + h\left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right)^2\right] dt + g\left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}}\right) dW_t$$

$$dv(t) = \alpha(\theta - v(t))dt + \xi\sqrt{v(t)}dB(t)$$

$$\langle dB(t), dW(t)\rangle = \rho dt$$

It is not difficult to realize therefore that the c.f. given $v(T)$ can be similarly computed for every one of those processes, hence one can infer:

$$\Phi_{Z(t),v(T)}(T, \omega) = \prod_{j=1}^{n} \Phi_{G,T}\left(T, \phi; L, h_j, g_j, \kappa_j, \theta_j, \xi_j, \rho_j, b_j, c_j, v_{j,T}, S_{j,t}^{*}\right)$$

$$\times \prod_{i=1}^{n} \Phi_{G,T}\left(T, 1; 0, h_i, g_i, \widetilde{\kappa}_i, \widetilde{\theta}_i, \widetilde{\xi}_i, \widetilde{\rho}_i, \widetilde{b}_i, \widetilde{c}_i, \widetilde{v}_{i,T}, S_t^{*i}\right)$$

where $\Phi_{G,T}\left(T, \phi; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_T, Z_t\right) = \mathbb{E}\left[e^{uZ(T)} \mid \mathcal{F}_t \cup v(T)\right]$ is provided next in Appendix B.2.

B.2   Helpful Results

Given the 4/2 process, the following c.f. are used in this paper:

$$dZ(t) = \left[L + h\left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right)^2 \right] dt + g\left(\sqrt{v(t)} + \frac{b}{\sqrt{v(t)}} \right) dW_t$$

$$dv(t) = \alpha(\theta - v(t))dt + \xi \sqrt{v(t)}dB(t)$$

$$\langle dB(t), dW(t) \rangle = \rho dt$$

-

$$\Phi_G\left(T, u; L, h, g, \alpha, \theta, \xi, \rho, b, c, v_t, Z_t\right) = \mathbb{E}\left[e^{uZ(T)} | \mathcal{F}_t \right]$$

$$= exp\left\{ uZ(t) + \frac{\alpha^2\theta}{\xi^2}(T - t) + u\left(r + 2(h - \frac{1}{2})g^2 b - \frac{g\rho\alpha\theta}{\xi} + \frac{gb\rho\alpha}{\xi} \right)(T - t) + u^2(1 - \rho^2)g^2 b(T - t) \right\}$$

$$\times \left(\frac{\sqrt{A_u}}{\xi^2 \sinh(\frac{\sqrt{A_u}}{2}t)} \right)^{m_u + 1} v(t)^{\frac{1}{2} + \frac{m_u}{2} - \frac{\alpha\theta}{\xi^2} - \frac{ugb\rho}{\xi}} \left(K_u(T) - \frac{ug\rho}{\xi} \right)^{-\left(\frac{1}{2} + \frac{m_u}{2} + \frac{\alpha\theta}{\xi^2} + \frac{ugb\rho}{\xi} \right)}$$

$$\times exp\left\{ \frac{v(t)}{\xi^2}\left(-\sqrt{A_u} \coth\left(\frac{\sqrt{A_u}(T - t)}{2} \right) + \alpha - ug\rho\xi \right) \right\} \frac{\Gamma\left(\frac{1}{2} + \frac{m_u}{2} + \frac{\alpha\theta}{\xi^2} + \frac{ugb\rho}{\xi} \right)}{\Gamma(m_u + 1)}$$

$$\times {}_1F_1\left(\frac{1}{2} + \frac{m_u}{2} + \frac{\alpha\theta}{\xi^2} + \frac{ugb\rho}{\xi}, m_u + 1, \frac{A_u v(t)}{\xi^4(K_u(T) - \frac{ug\rho}{\xi})\sinh^2\left(\frac{\sqrt{A_u}(T-t)}{2} \right)} \right),$$

with

$$A_u = \alpha^2 - 2\xi^2\left(u\left(\frac{g\rho\alpha}{\xi} + (h - \frac{1}{2})g^2 \right) + \frac{1}{2}u^2(1 - \rho^2)g^2 \right),$$

$$m_u = \frac{2}{\xi^2}\sqrt{\left(\alpha\theta - \frac{\xi^2}{2} \right)^2 - 2\xi^2\left(u\left(\frac{gb\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta \right) + (h - \frac{1}{2})g^2 b^2 \right) + \frac{1}{2}u^2(1 - \rho^2)g^2 b^2 \right)},$$

$$K_u(T) = \frac{1}{\xi^2}\left(\sqrt{A_u} \coth\left(\frac{\sqrt{A_u}(T - t)}{2} \right) + \alpha \right)$$

-

$$\Phi_{G,T}\left(T, u; L, h, g, \kappa, \theta, \xi, \rho, b, c, v_t, Z_t\right) = \mathbb{E}\left[e^{uZ(T)} | \mathcal{F}_t \cup v(T) \right]$$

$$= exp\left\{ uZ(t) + u\left(r + 2(h - \frac{1}{2})g^2 b - \frac{a\rho\alpha\theta}{\xi} + \frac{b\rho\alpha}{\xi} \right)(T - t) + u^2(1 - \rho^2)g^2 b(T - t) \right\}$$

$$\times exp\left\{ \frac{ug\rho}{\xi}(v(T) - v(t)) + \frac{ugb\rho}{\xi}\log\frac{v(T)}{v(t)} \right\}$$

$$\times \frac{\sqrt{A_u}\sinh\left(\frac{\alpha(T-t)}{2} \right)}{\alpha \sinh\left(\frac{\sqrt{A_u}(T-t)}{2} \right)} exp\left(\frac{v(T) + v(t)}{\xi^2}\left(\alpha \coth\left(\frac{\alpha(T - t)}{2} \right) - \sqrt{A_u} \coth\left(\frac{\sqrt{A_u}(T - t)}{2} \right) \right) \right)$$

$$\times \frac{I_{\frac{2}{\xi^2}\sqrt{\left(\alpha\theta - \frac{\xi^2}{2} \right)^2 + 2\xi^2 B_u}}\left(\frac{2\sqrt{A_u v(T)v(t)}}{\xi^2 \sinh\left(\frac{\sqrt{A_u}(T-t)}{2} \right)} \right)}{I_{\frac{2\alpha\theta}{\xi^2} - 1}\left(\frac{2\alpha\sqrt{v(T)v(t)}}{\xi^2 \sinh\left(\frac{\alpha(T-t)}{2} \right)} \right)},$$

with

$$B_u = u\left(\frac{gb\rho}{\xi}\left(\frac{\xi^2}{2} - \alpha\theta \right) + (h - \frac{1}{2})g^2 b^2 \right) + \frac{1}{2}u^2(1 - \rho^2)g^2 b^2,$$

Curriculum Vitae

Name: Zhenxian Gong

Degrees: The University of Western Ontario
London, ON
2016-2021 Ph.D in Statistics

Queen's University
Kingston, ON
2015-2016 MSc in Statistics

University of Alberta
Edmonton, AB
2007-2011 BSc in Statistics

Honours and   Western Graduate Research Scholarship
Awards: 2016-2020

Related Work Teaching Assistant
Experience: The University of Western Ontario
2016 - 2021

Research Assistant
The University of Western Ontario
2017 - 2021

Publications:

Marcos Escobar-Anel and Zhenxian Gong. The mean-reverting 4/2 stochastic volatility model: Properties and financial applications. Applied Stochastic Models in Business and Industry. 2020; 36: 836– 856

Yuyang Cheng, Marcos Escobar-Anel and Zhenxian Gong. Generalized mean-reverting 4/2 Factor Model. Journal of Risk and Financial Management; 2019, 12.4: 159.