Modelling of Credit Risk and Correlation Risk: Time-Dependent and Stochastic Correlation Models

Dissertation

zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

dem Fachbreich C - Mathematik und Naturwissenschaften- der Bergischen Universität Wuppertal vorgelegt von

Dipl. Math. Long Teng

Promotionsausschuß

Gutachter/Prüfer:	Prof. Dr. Michael Günther
Gutachter/Prüfer:	Prof. Dr. Matthias Ehrhardt
Gutachter/Prüfer:	Prof. Dr. Carlos Vázquez Cendón
Prüfer:	Prof. Dr. Hanno Gottschalk
Prüfer:	Prof. Dr. Bruno Lang
Prüfer:	Dr. Jörg Kienitz

August, 2015

Die Dissertation kann wie folgt zitiert werden:
urn:nbn:de:hbz:468-20151124-114230-6
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20151124-114230-6]

Danksagung/Acknowledgements

Danksagung:

Die vorliegende Dissertation wurde während meiner Tätigkeit am Lehrstuhl für Angewandte Mathematik and Numerische Analysis der Begischen Universität Wuppertal verfasst. Daher möchte ich mich zuerst und ganz besonders beiProf. Dr. Michael Günther and Prof. Dr. Matthias Ehrhardt für ihre Betreuung und die Arbeitsgelegenheit am Lehrstuhl für Angewandte Mathematik in Wuppertal bedanken. Insbesondere möchte ich mich herzlich für ihren unermündlicher Einsatz and Vertrauen, sowie die Möglichkeit, meine Forschungsschwerpunkte frei zu gestalten, bedanken. Sie haben mir ein optimales Forschungsumfeld zur Verfügung gestellt und es war mir eine große Freude mit ihnen zusammenarbeiten.

Ferner möchte ich mich bei Prof. Carlos Vázquez Cendón dafür bedanken, diese Dissertation zu begutachten. Mein Dank richtet sich auch an Prof. Dr. Hanno Gottschalk and Dr. Jörg Kienitz, die eingewilligt haben, Mitglieder der Promotionskommission zu werden. Besonders danke ich hier Dr. Jörg Kienitz dafür, dass er immer Zeit für meine Fragen finden konnte. Ich habe von seiner großen Erfahrung in der Finanzindustrie stark profitiert.

An dieser Stelle ergeht auch herzlicher Dank an Kees De Graaf und Xueran Wu für das Korrekturlesen dieser Dissertation. Ganz besonders danke ich Xueran Wu nicht nur für intensive und hilfreiche Diskussionen, sondern auch für ihre Liebe und ihre Unterstützung.

Ich danke meinen derzeitigen und ehemaligen Kollegen in der Arbeitsgruppe “Angewandte Mathematik and Numerische Analysis” für mathematische Diskussionen und auch für das angenehme Arbeitsumfeld. Ich möchte genau einige Namen erwähnen, Dr. Andreas Bartel aus Wuppertal, Dr. Christian Kahl aus Commerzbank AG (London), Dr. Cathrin van Emmerich aus RWE Supply & Trading (Essen) and Patrick Deuß aus parcIT GmbH (Köln).

Mein herzlicher Dank richtet sich auch an meine Familie, die mir alle ihre Liebe und Unterstützung gegeben haben und geben.

Acknowledgement:

The present dissertation was written during my employment at the Chair of Angewandte Mathematik und Numerische Analysis of the Bergische Universität Wuppertal. Hence foremost and special thanks goes to my supervisors Prof. Dr. Michael Günther and Prof. Dr. Matthias Ehrhardt for their supervision and the opportunity to work at the Chair of Applied Mathematics in Wuppertal. I want to thank them particularly for the indefatigable commitment, the confidence in me and the freedom to choose my research focus. They provided an optimal research environment for me and it was my pleasure to work with them.

I would like to thank Prof. Carlos Vázquez Cendón for reviewing this thesis. Thanks also goes to Prof. Dr. Hanno Gottschalk and Dr. Jörg Kienitz who agreed to be members of my doctoral committee. I particularly thank Dr. Jörg Kienitz who always found the time to answer my questions, I have strongly benefited from his extensive experience in the financial industry.

I would also like to thank Kees De Graaf and Xueran Wu for proof reading this thesis. In particular I want to thank Xueran Wu not only for all the intensive and helpful discussions but also for giving me love and support.

I thank my colleagues in the working group "Applied Mathematics and Numerical Analysis", present and former, for mathematical discussions and also for being kind, I had an enjoyable working environment. Ich would like to mention just a few names, Dr. Andreas Bartel from Wuppertal, Dr. Christian Kahl from Commerzbank AG (London), Dr. Cathrin van Emmerich from RWE Supply & Trading (Essen) and Patrick Deuß from parcIT GmbH (Köln).

Finally, I would like to thank my family for giving me all their love and support.

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

CVA	credit valuation adjustment
WWR	wrong-way risk
CCR	counterparty credit risk
BVA	bilateral credit Valuation adjustment
CDS	credit default swap
OTC	over the counter
R_{EC}	recovery rate on a unit amount
L_{GD}	loss given default
$\mathbb{1}_{\{\cdots\}}$	indicator function
$B(t,T)$	zero bond price at time t for the maturity T
$D(t,T)$	stochastic discount factor at time t for the maturity T
PFE	potential future exposure
EFE	expected positive exposure
UCVA	unilateral credit valuation adjustment
DVA	debit valuation adjustment
UDVA	unilateral debit valuation adjustment
$\mathbb{Q}$	risk-neutral measure, equivalent martingale measure
$\mathbb{E}$	expectation under $\mathbb{Q}$
τ	default time
N_t	time homogeneous Poisson process
$\hat{N}_t$	time inhomogeneous Poisson process
λ	hazard rate, (constant) intensity
$\Gamma(t)$	hazard function, cumulated intensity
$\lambda(t)$	deterministic time-varying intensity
ξ	standard exponential random variable
$\Lambda(t)$	hazard process, cumulated intensity

λ_t	stochastic intensity
P	physical/objective/real-world measure
E	expectation under P
i.i.d.	independent and identically distributed
$\mathcal{L}_x$	loss given default of the counterparty x
$\mathcal{P}$	premium rate
$\mathcal{R}_x$	recovery rate of the counterparty x
$\Pi(t,T)$	payoff of a CDS with a default-free counterparty
$\mathbb{R}$	real numbers
$\mathbb{R}^+$	positive real numbers
$P_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)$	price of a CDS to the CDS seller at time t
x^+	positive part of x
$\overline{\Pi}(t,T)$	payoff of a CDS with defaultable counterparty
$\overline{P}_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)$	price of a CDS at time t under bilateral CCR
$\text{BVA}(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}})$	BVA of a CDS at time t
CIR	Cox-Ingersoll-Ross
CDF	cumulative distribution function
FFT	fast Fourier transform
FRFT	fractional fast Fourier transform
$\widehat{\Pi}(t,T)$	CDS payoff with simultaneously defaultable counterparties
$\widehat{P}_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)$	CDS price at t under bilateral CCR (simultaneous default)
$\widehat{\text{BVA}}(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}})$	BVA of a CDS with simultaneously defaultable counterparties
$\bar{\rho}_t$	dynamic (time-dependent correlation function)
SCP	stochastic correlation process
ρ_t	correlation process
$(\cdot)^+$	$\max(0,\cdot)$
Φ	standard normal distribution function

$\approx$	equivalent measures
W_t	brownian motion
$:=$	defined as
$E[X]$	expectation of X
$V[X]$	variance of X
σ_X	standard deviation of X
$\text{cov}(X, Y)$	linear covariance between X and Y
$\approx$	approximately

List of Publications

11. L. Teng, M. Ehrhardt and M. Günther, *On the Heston Model with Stochastic Correlation*, Preprint 15/22, University of Wuppertal, **submitted** for publication, April 2015.

10. L. Teng, M. Ehrhardt and M. Günther, *Numerical Simulation of the Heston Model with Stochastic Correlation*, Preprint 15/01, University of Wuppertal, **submitted** for publication, January 2015.

9. L. Teng, C. van Emmerich, M. Ehrhardt and M. Günther, *A Versatile Approach for Stochastic Correlation using Hyperbolic Functions*, Int. J. Comput. Math. **DOI:10.1080/00207160.2014.1002779**, January 2015.

8. L. Teng, C. van Emmerich, M. Ehrhardt and M. Günther, *Modelling Stochastic Correlation with modified Ornstein-Uhlenbeck process*, Proceedings that will be **published** by Springer Heidelberg as part of the ECMI book subseries of Mathematics in Industry, ISSN: **1612-3956**, 2015.

7.. L. Teng, M. Ehrhardt and M. Günther, *Option Pricing with Dynamically Correlated Stochastic Interest Rate*, Acta Mathematica Universitatis Comenianae, **85 (2)**, ISSN: **0862-9544**, 2014.

6. L. Teng, M. Ehrhardt and M. Günther, *The Pricing of Quanto Options under Dynamic Correlation (extended version)*, J. Comput. Appl. Math., **275C**, DOI: **10.1016/j.cam.2014.07.017**, p.304–310, 2014.

5. L. Teng, M. Ehrhardt and M. Günther, *The Pricing of Quanto Options under Dynamic Correlation (short version)*, Proceedings of the 14th International Conference on Mathematical Methods in Science and Engineering, ISBN: **978-84-616-9216-3**, 2014.

4. L. Teng, M. Ehrhardt and M. Günther, *The Dynamic Correlation Model and its Application to the Heston Model*, Preprint 14/09, University of Wuppertal, **submitted** for publication, April 2014.

3. L. Teng, M. Ehrhardt and M. Günther, *Modelling Stochastic Correlation*, Preprint 14/03, University of Wuppertal, **submitted** for publication, February 2014.

2. L. Teng, M. Ehrhardt and M. Günther, *Bilateral Counterparty Risk Valuation of CDS Contracts with Simultaneous Defaults*, Int. J. Theoret. Appl. Fin. **16 (7) 1350040**, 2013.

1. L. Teng, M. Ehrhardt and M. Günther, *Numerical Evaluation of Complex Logarithms in the Cox-Ingersoll-Ross Model*, Int. J. Comput. Math. **90 (5)**, p.1083–1095, 2013.

Chapter 1

Introduction

The degree of a relationship between the changes of two or more financial quantities in time can be measured by correlations, which play a key role in investing, trading, risk management and regulation. Generally, correlation risk refers to unexpected losses due to unpredictable changes of the correlation between financial variables. Contrast to other types of risk, correlation had often been disregarded until the global financial crisis between 2007 and 2009. Since correlation risk is an influential trigger of this financial crisis, it has thus become the focus of attention in finance, see Section 2.3.

Hedging correlation risk is more difficult than hedging other financial risks for two reasons indicated in [84]: (1) Hedging correlation risk involves two or more financial variables, since the correlation is measured between at least two financial variables. (2) There is principally no underlying instrument traded in the market as a hedge by buying or selling. Nevertheless, in order to hedge correlation risk one firstly has to realistically model the financial correlation. The one which is simplest and has been widely used is the *Pearson correlation coefficient* (see B.1), although it has several limitations for Finance. For more information on the limitations of Pearson correlation coefficient we refer the interested reader to [83, 84, 105]. For the case of modelling financial quantities as random variables, the Pearson correlation coefficients of these random variables are

used to represent the dependences among financial quantities within a time period. We usually correlate Brownian motions (BMs) of stochastic differential equations (SDEs) with a deterministic parameter that drive the financial quantities, in order to measure how the financial quantities move jointly in time.

Mathematically, constant correlated BMs imply that the corresponding stochastic processes are jointly covariance-stationary. Particularly, the instantaneous covariance of the jointly covariance-stationary stochastic processes is deterministic over time, in other words, they have stable correlation parameters. However, market observations indicate that the financial quantities are correlated in a strongly and highly non-linear way. Financial correlations behave even stochastically and unpredictably, thus, the financial quantities in the real market are not likely to be jointly stationary. Obviously, if one decides nevertheless to ignore this problem and insists on using the correlated BMs by a constant, this may lead to correlation risk. A possible solution for this problem is to find an appropriate non-linear function (time-dependent) or stochastic process to model financial correlations, in terms of which we can construct dynamically or stochastically correlated BMs. Employing the stochastic processes with dynamically or stochastically correlated BMs for financial quantities will certainly be more realistic.

This thesis is divided into three parts. Chapter 2 is devoted to briefly introduce some basic notions of counterparty credit risk (CCR), CVA and to study reduced form (intensity) models to model *default time* and the ways how to impose the default dependence between counterparties. This part focuses on understanding the impact of financial correlation to credit risk modelling by computing credit valuation adjustment (CVA) and analyzing wrong-way risk (WWR).

In Chapter 3 we investigate computing bilateral credit valuation adjustment (BVA) on Credit Default Swap (CDS) which is one of the most common credit derivatives. We investigate a computational problem of using the BVA-formula provided in [15] (abridged version [16]) and show how to address this problem by employing tailored

numerical methods. Another important contribution of this chapter is a new BVA-formula which allows simultaneous defaults among counterparties, where the default intensity is modelled by applying a Markov copula model.

From the first part we realize that the correlation between counterparties in a CDS contract plays a key role on their default risk management. Generally, the degree of relationship between financial products and financial instituations always plays an essential role on, e.g., pricing and hedging. However, intuitively, a time-dependent model or a stochastic model could better replicate the phenomena in the real world. Indeed, market observations clearly indicate that financial quantities are correlated in a strongly nonlinear way, correlation could even behave stochastically and unpredictably. This motivates us to finish the next two parts for the present thesis: modelling and application of local time-dependent and stochastic correlation.

In the second part of this thesis we provide a time-dependent correlation function and its applications for pricing financial derivatives and related financial products. In Chapter 4, we propose an appropriate and reasonable time-dependent correlation function and present the concept of dynamically (time-dependent) correlated Brownian motions (BMs) and its construction. As examples, we employ this new time-dependent correlation function to price European options and Quanto options. We analyze the effect (improvement) by using a time-dependent correlation instead of a constant correlation.

In the third part of this thesis we investigate how to model correlation as a stochastic process and its applications in finance. In Chapter 5, a general stochastic correlation model is provided, several stochastic correlation processes are discussed and analyzed. By applying stochastic correlation to price Quanto options we quantify the correlation risk caused by using a wrong (constant) correlation. Furthermore, we incorporate stochastic correlation into the Heston model and find, that the Heston model extended by introducing stochastic correlation provide a better fit to the skew and smile in the volatility

surface that is visible in the market than the pure Heston model and the double Heston model.

In Chapter 6, we summarise our findings developed and discussed throughout this thesis and mention an outlook about further research opportunities in the direction of modelling and application of stochastic correlation. Appendix A supplements some fundamental conditions and theorems, e.g., the *usual conditions which are assumed to hold throughout this thesis,* measure change based on Radon-Nikodym and Girsanov theorem for pricing purpose. Appendix B is devoted to the chosen basic knowledge. The proofs of propositions and theorems are provided in Appendix C.

Part I

Impact of Correlation to Credit Risk Modelling

This part is devoted to briefly introduce basic notions of counterparty credit risk (CCR), credit value adjustment (CVA) and to study default dependence between counterparties in a credit default swap (CDS) contract. To compute the highly accurate bilateral CVA (BVA) on CDS, we employ tailored numerical methods to obtain the cumulative distribution function (CDF) of the integrated Cox-Ingersoll-Ross (CIR) process, which is demanded to to compute the survival probabilities of the counterparties. Furthermore, we develop a new formula which allows simultaneous defaults among counterparties, the simultaneous default risk can thus be regarded.

Chapter 2

Counterparty Risk Valuation

The aim of this chapter is to briefly introduce basic notions of counterparty credit risk (CCR), credit valuation adjustment (CVA) and modelling the default time using reduced form (intensity) models as a preparation for Chapter 3. The goal is to emphasize the impact of correlation to credit risk modelling by computing CVA and analyzing wrong-way risk (WWR), so we will not discuss about the other crucial issues of credit risk modelling like collateral, re-hypothecation, netting, mitigating and so on. The interested reader is referred to [20] and [55].

2.1 Counterparty Credit Risk (CCR)

For evaluating financial contracts that are traded over the counter (OTC), one has to consider CCR, whereby the transactions are not backed by the guarantee of a clearing-house or an exchange. Therefore, each counterparty is exposed to the default risk of the other party. As its name implies, a *default risk* refers to the possibility that a counterparty in a financial contract will be unable or unwilling to make the required payments to meet the obligations stated in the contract. If this happens, instead of the agreed payments, just a fraction of the value at the instant of default will be paid. This leads to the concept of *recovery rate* (R_{EC}) which represents the percentage of the outstanding

claim recovered when a default event occurs. An accompanied variable to the recovery is the *loss given default (L_{GD})* which is equal to $1 - R_{EC}$ on a unit amount. The resulting amount of potential loss due to the defaulting counterparty is considered as counterparty credit exposure (hereafter known as exposure). It is worth to mention a characteristic of the exposure: At a general level, if a company holds a financial instrument which has a positive value on a defaulted counterparty, this can be only seen as a claim on this defaulted counterparty. However, for a negative value of a financial instrument the company is still obliged to honour his agreed payments. This is to say that the company will incur a loss if it is owed money and its counterparty defaults, whilst in the case of the company in debt it cannot profit from the default of its counterparty and has to admit its liability. For a detailed description on CCR we refer to [55, 98].

To know how the recovery rules work we study the *cash flow* of *a corporate coupon bond*, which is a simple defaultable contract. We consider a corporate coupon bond with face value 1 issued by counterparty "C", which matures at time $T = T_n$ and promises to pay coupons c_i at times $T_1 < T_2 < \cdots < T_n$. Let us assume, in case of the default event by the counterparty before or at the T, only the recovery payments with R_{EC} will be made at the maturity date. The cash flow of the bond is thus

$$\sum_{i=1}^{n} c_i \mathbb{1}_{\{\tau > T_i\}} + \mathbb{1}_{\{\tau > T\}} + R_{EC} \mathbb{1}_{\{\tau \leq T\}}, \tag{2.1}$$

where τ denotes the default time of the bond, and the discounted payoff at time t of this bond is given by

$$\sum_{i=1}^{n} c_i B(t, T_i) \mathbb{1}_{\{\tau > T_i\}} + B(t, T) \mathbb{1}_{\{\tau > T\}} + R_{EC} B(t, T) \mathbb{1}_{\{\tau \leq T\}}. \tag{2.2}$$

More detailed information about the corporate coupon bond can be found in [8, 19, 60].

Credit rating

The default probability of a counterparty depends on his credit quality which can be

measured by an important indicator, the *credit rating*. Credit ratings are attributed by major rating agencies such as Moody's Investors Service, Standard & Poor's Corporation or Fitch IBCA. Besides, some financial institutions assign their own *internal ratings*. To reflect the possibility of default, the financial institutions could either use the standard credit rating (standardized approach) or use the internal ratings (internal-ratings-based approach). An incentive for financial institutions to use internal ratings is that using the standardized approach can lead to higher capital requirements than using the internal ratings, because the standardized approach employe conservative measures of capital requirements based on simple calculations. For more information on credit rating we refer the interested reader to [7, 38, 81, 98].

Bilateral CCR

In the earlier years, the counterparty holding a higher rating, such as triple-A entities and global investment banks, could be seen as default-free for risk assessment. However, we have witnessed in recent years the increasing default events of these "default-free" institutions, e.g.: On March 16, 2008, the Federal Reserve Bank of New York assisted JP Morgan Chase to purchase Bear Stearns for just $2 a share which represented a shocking loss as Bear Stearns's stock had been traded at $93 only a month before; On September 14, 2008, Lehman Brothers announced it filed for bankruptcy protection; On the same day as Lehman Brothers's bankruptcy, Merrill Lynch agreed to be acquired by the Bank of America. Thus, it is no longer realistic to regard any financial institutions as default-free, no matter how prestigious or important it is. Considering a *bilateral* CCR has become the standard, where "bilateral" means that each party takes CCR with respect to the other party into account. For the arbitrage-free valuation of bilateral CCR for several financial derivatives we refer the readers to [15, 16, 17, 21].

2.2 Credit Valuation Adjustment (CVA)

Exposure can be computed by simulating many different scenarios of the price of the transaction with the given counterparty at different points in future time. One chooses statistics to characterise the generated price distributions, e.g., 99% quantile, called *Potential Future Exposure (PFE)* or the mean of the positive part of the price distributions, called *Expected Positive Exposure (EPE)*. For more details on exposure, we refer to [28].

Another important index for CCR is the cost of its hedging, *Credit Valuation Adjustment (CVA)* or sometimes called *Counterparty Valuation Adjustment*, or *Counterparty Value Adjustment* which is defined as the difference between the risk-free value of a derivative and its fair value when the counterparty default possibility is taken into account. This means that CVA is an adjustment to be subtracted from the default risk-free price in order to account for the counterparty default risk. Obviously, everyone would prefer to make a trade with a default risk-free party rather than with a risky one, so the risk-free prices need to be decreased by subtracting a positive CVA. In other words, we charge the default-risky one a supplementary amount besides the default-free cost of the contract, CVA is thus the price of CCR.

Furthermore, if an investor thinks himself as being default-free and computes CVA only by considering the default risk of his counterparty, this is called *unilateral CVA (UCVA)*. From the point of view of the counterparty this adjustment is also called *Debit Valuation Adjustment (DVA)*, more precisely *unilateral DVA (UDVA)* as the investor has been considered as default-free. In this case we see that UCVA (investor) = UDVA (counterparty) which will be subtracted by the investor and added by his counterparty, and UDVA (investor) = UCVA (counterparty) are equal to zero because the investor is default-free. We refer to [20] for more information about DVA.

Bilateral CVA

In the past, the investor with high credit quality had been often regarded as default-free such that we only needed to deal with UCVA. However, due to the witnessed de-

fault events by highly superior financial institutions, e.g., Lehman Brother's bankruptcy, counterparties (like corporate client) do not accept any financial institutions (such as the bank) as default free, i.e. the both parties will not agree on the price. To achieve the agreement on a price, the default probability of both parties must be considered. An new concept *bilateral CVA (BVA)* is arising, where the default risk of each party should be considered, no matter what kind of credit quality. Indeed, bilateral risk had been mentioned in the credit risk by Basel II, "Unlike a firms exposure to credit risk through loan, where the exposure to credit risk is unilateral and only the lending bank faces the risk of loss, the counterparty credit risk creates a bilateral risk of loss: the market value of the transaction can be positive or negative to either counterparty to the transaction." After the global financial crisis (2007-2009), the Basel III requirements "Basel III: A global regulatory framework for more resilient banks and banking systems" had been published which has a large portion of changes related to CCR and CVA and focuses on improving CCR management. For the regulatory aspects of CCR and CVA in more detail we refer to [22, 55].

The realistic calculation of CVA has thus become an urgent task. Calculating UVA is relatively straightforward, one simply needs to discount the cash flows and add any default payment taking account of the possible default events. However, in case of BVA, the calculation is much more difficult due to the bilateral nature which means all the cash flows and payments must be considered in both directions.

2.3 Impact of Correlation: Wrong-Way risk

To emphasize the impact of financial correlations to credit risk modelling one must keep wrong-way risk in mind. As a feature of CCR *Wrong-way risk* is defined by the International Swaps and Derivatives Association (ISDA) as the risk that occurs when exposure to a counterparty is adversely correlated with the credit quality of that counterparty.

Since a high credit quality relates to a small default probability, wrong-way risk exists when exposure to a counterparty is positively correlated with the default probability of that counterparty. An example could be trading an oil swap with a oil producer where a bank pays floating oil (spot price) and receives a fixed price. We may imagine a negative correlation between the default of the oil producer and the price of oil, since lower prices of oil will put business of the oil producer less profitable. When the correlation has a large negative value, the oil price decrease would worsen the credit quality of the oil producer, which causes that he will have a increased default probability. However, a decrease in spot oil price will increase in the value of the oil swap to the bank. While the oil price is decreasing, there will be a higher default probability from the oil producer due to the correlation, the bank's exposure is thus increasing. The bank faces a large loss if the oil producer defaults now. For more information we refer to [34, 67, 94]

Correlation risk

As we all know, the financial volatility plays an important role, e.g., in pricing and hedging. However, another equally important factor for the financial market, correlation had been disregarded until the global financial crisis between 2007 and 2009. Afterwards, the Basel committee has recognized the importance of correlation risk: Since CVA is an integral part of the Basel III as mentioned above, especially, its associated wrong-way risk arises from the correlation between the credit quality of a counterparty and the exposure. This is to say that correlation between counterparties plays a key role for wrong-way risk and must be considered for computing CVA. We should recognize that financial correlation is a critical factor in managing CCR. Moreover, correlation risk is also a critical part of the other financial risks, as market risk, systemic risk and so on. About correlation risk there exist many work, see e.g., [24, 25, 35, 59, 68, 88].

In connection with wrong-way risk, correlation risk is a form of risk, there is a strong dependence between financial value and default event which will increases the loss, e.g., wrong-way risk increases with correlation between counterparties. That is why we must

be able to handle this effect by the correlation for computing CVA, this is also one main part of the Basel III accords. In Chapter 3, as an example we study how to compute BVA for a CDS which is one of the most important credit derivatives. The effect of correlation on CVA will be analyzed, the problem of computation of CVA will be pointed out and its solution will be provided.

In a more general way, *correlation risk* refers to the risk of a financial loss due to change in correlations between financial variables. This also means that the financial loss can also arise using a constant correlation due to the fact that the actual correlations between two financial variables are unstable and likely to change over a small time interval. To hedge the correlation risk one has to model correlations properly which is very important for risk assessment. In Chapter 4 and 5 we focus on modelling correlation as a time-dependent functions, as suitable stochastic processes, and also provide their applications for pricing different financial derivatives.

Simultaneous default

We have seen that a higher wrong-way risk comes down to a larger negative correlation. One may ask whether it is possible that the correlation is so extremely high that counterparties default simultaneously? Indeed, a simultaneous default can happen in the real financial market. Mathematically we define simultaneous defaults among the counterparties as that the default times of them are exactly the same. However, in the real world we can already say that they default simultaneously if they filed for bankruptcy protection on one day or within a few days, e.g., the collapses of Lehman Brothers and Merrill Lynch were just within two days (September 13-14, 2008). Another example noticed in [4], 24 railway firms defaulted simultaneously on the same day, June 21, 1970. In reality, it is possible that the defaults among the counterparties do not occur simultaneously, but if one's default has triggered a jump in the default probability of the other one (e.g., if they are very highly correlated), which might end up defaulting only within a short time period. For example, the protection seller's default could trigger a jump in the default

probability of the reference credit so that the protection buyer has to suffer a loss. In the sequel we refer to this later case also as a simultaneous default. The investigation of computing BVA by allowing simultaneous defaults among counterparties for CDS is provided in Section 3.4.

2.4 Modelling Default Time: Reduced-form (Intensity) models

As mentioned before we will concentrate on computing BVA for CDS contracts in the next chapter. The most important and also the first task is to model default probability and default time, especially the first default time among the parties in a CDS contract. To model the default time there are two methodologies: the structural approach and the reduced-form approach.

The Structural approach

The structural models are based on the work by Merton [85] in which credit events are triggered by movements of the obligor's value relative to some barrier which can be deterministic or stochastic. The default time can thus be considered as the first instant where the firm value hits such barrier. In this work we will not present the structural approach , the interested reader is referred to the comprehensive list of references to this approach, e.g., [10, 51, 53, 13, 90, 32, 71, 76, 72] and a nice book [8].

The reduced-form (intensity) model

In this section, we focus on reduced-form (intensity) models which describe the default by means of exogenous jump process, the Poisson process (see Appendix B). In contrast to the structural approach, the value of obligor's (firm's) assests is not modelled at all in the reduced-form models. The default is not triggered by market observables but an exogenous component which is independent of all the default-free market information. Applying the intensity model is devoted to model the random default time and evaluate

conditional expectations under $\mathbb{Q}$ of functionals of the default time and corresponding cash flows. What attracted the practitioners most to this approach is the easy calibration to the market data (credit spreads). For this approach there exist a lot of work, see e.g., [14, 39, 40, 62, 69, 74, 80, 91, 92, 97] and the nice books [8] and [19].

So far we have already a rudimentary grasp of the reduced-form models in which the default time τ is the first jump of a Poisson process. Furthermore, to include WWR we should find a way how to impose the dependencies between default events, more precisely, between default times.

2.4.1 Time homogeneous Poisson Processes

First we consider the time-homogeneous Poisson process:

Definition 2.4.1 *A process N_t is called a (time-homogeneous) Poisson process with intensity $\lambda > 0$ if the following conditions are satisfied [8]:*

- $N_0 = 0$,
- *the increment $N_t - N_s$ is independent of $\mathcal{F}_s$ for $0 \leq s < t$,*
- *the increment $N_t - N_s$ is Poisson distributed with mean $\lambda(t-s)$; especially, for any $k = 0, 1, \ldots$ we have*

$$P(N_t - N_s = k|\mathcal{F}_s) = P(N_t - N_s = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}. \tag{2.3}$$

Lemma 2.4.1 *Useful properties of Poisson processes: Let N_t be a time homogeneous Poisson process with intensity $\lambda > 0$. Then [19]*

- $P(N_t = 0) = e^{-\lambda t}$,
- $\lim_{t\to 0} P(N_t \geq 2)/t = 0$,
- $\lim_{t\to 0} P(N_t = 1)/t = \lambda$.

In the sequel, we employ the fundamental properties of jumps of Poisson processes to model default time. We begin with *average arrival rate* and *variance per unit of time*. From (2.3) it is direct that N_t is a Poisson random variable with parameter λt for a fixed t :

$$\mathbb{E}[N_t] = \mathbb{V}[N_t] = \lambda t, \tag{2.4}$$

where the expectation and variance are under $\mathbb{Q}$ which is a risk-neutral measure and should be used for pricing purpose. More detailed information about $\mathbb{Q}$ can be found in Appendix A. By rewriting (2.4) we have

$$\lambda = \frac{\mathbb{E}[N_t]}{t} = \frac{\mathbb{V}[N_t]}{t}, \tag{2.5}$$

which can be interpreted as an average arrival rate or variance per unit of time and is called as *hazard rate* . For the case of a time homogeneous Poisson process, the hazard rate λ is assumed to be constant for all t.

We denote the jump times of the process N_t by $\tau^1, \tau^2, \cdots$, and from (2.3) we easily see

$$\mathbb{Q}(N_t = 0) = \mathbb{Q}(\tau^1 > t) = \mathrm{e}^{-\lambda t}, \tag{2.6}$$

which is the probability of no jumps before or at t and is called *survival probability*. As indicated in [19], (2.6) has the same structure as a discount factor, with the default intensity playing the role of the interest rate. One more intuitive interpretation: We are interested in the distribution of the jump times. The times between one jump and the next one, $\tau^1, \tau^2 - \tau^1, \tau^3 - \tau^2, \cdots$ are i.i.d. as an exponential random variable with mean $1/\lambda$. Therefore, $\lambda\tau^1$ is a standard exponential random variable ($\sim$ Exp(1)), we thus obtain

$$\mathbb{Q}(\tau^1 > t) = \mathbb{Q}(\lambda\tau^1 > \lambda t) = \mathbb{Q}(X > \lambda t) = \mathrm{e}^{-\lambda t}, \tag{2.7}$$

where $X \sim \mathrm{Exp}(1)$.

It is easy too see that the probability of defaulting between the time instances s and t is thus given by

$$\mathbb{Q}(s < \tau^1 \leq t) = \mathbb{Q}(\tau^1 > s) - \mathbb{Q}(\tau^1 > t) = \mathrm{e}^{-\lambda s} - \mathrm{e}^{-\lambda t}, \tag{2.8}$$

which is approximately equal to $\lambda(t-s)$ for t close to s. More generally, if we define the default time τ as the first default time, namely $\tau := \tau^1$, for an arbitrarily small dt we find:

$$\begin{aligned}\mathbb{Q}\left(\tau \in [t, t+dt) | \tau \geq t\right) &= \frac{\mathbb{Q}\left(\tau \in [t, t+dt) \cap \tau \geq t\right)}{\mathbb{Q}(\tau > t)} = \frac{\mathbb{Q}\left(\tau \in [t, t+dt)\right)}{\mathbb{Q}(\tau > t)} \\ &= \frac{\mathbb{Q}(\tau > t) - \mathbb{Q}(\tau > t+dt)}{\mathbb{Q}(\tau > t)} = \frac{\mathrm{e}^{-\lambda t} - \mathrm{e}^{-\lambda(t+dt)}}{\mathrm{e}^{-\lambda t}} \\ &\approx \lambda\, dt, \end{aligned} \tag{2.9}$$

which can be interpreted as in [19]: The probability that a company defaults in (arbitrarily small) time period of dt years given that it has not defaulted so far is $\lambda\, dt$.

2.4.2 Time inhomogeneous Poisson Processes

In contrast to an assumed constant intensity for all t we consider in this section a *deterministic time-varying* intensity $\lambda(t)$, which is assumed to be a positive and piecewise (right-) continuous function. We first define

$$\Gamma(t) := \int_0^t \lambda(s)\, ds, \tag{2.10}$$

the *cumulated intensity, cumulated hazard rate,* or *Hazard function.*

From the cumulated hazard rate, we can derive the survival probability structure for a deterministic time-varying intensity [19]. If M_t is a standard Poisson process, we can

define a time-inhomogeneous Poisson process $\hat{N}_t$ with intensity $\hat{\lambda}$ as

$$\hat{N}_t = M_{\Gamma(t)}, \tag{2.11}$$

which indicates that a time inhomogeneous Poisson process is a time-changed standard Poisson process. Obviously, the increments of $\hat{N}_t$ are no longer i.i.d. due to the time distortion, but they are still independent and increasing by jumps of size 1. Furthermore, we can consider the event "$\hat{N}$ has a first jump at τ" to be equivalent with the event "M has the first jump at $\Gamma(\tau)$". From Section 2.4.1 we have seen that the first jump time for the standard Poisson process M_t is a standard exponential random variable, say ξ

$$\Gamma(\tau) := \xi \sim \text{Exp}(1), \tag{2.12}$$

and by inverting (2.12) we obtain

$$\tau = \Gamma^{-1}(\xi). \tag{2.13}$$

Furthermore, we can easily calculate the survival probability as

$$\mathbb{Q}(\tau > t) = 1 - \mathbb{Q}(\tau \leq t) = 1 - (1 - \mathrm{e}^{-\Gamma(t)}) = \mathrm{e}^{-\int_0^t \lambda(s)\, ds}. \tag{2.14}$$

When we compare this expression to $B(0,t) = \mathrm{e}^{-\int_0^t r(s)\, ds}$ with the deterministic time-varying short rate $r(t)$ in an interest rate world we realize again that one can translate many ideas of the interest rate into credit modelling, see [19]. As in the case of using a time homogeneous Poisson process we calculate the probability of defaulting between s and t :

$$\begin{aligned}\mathbb{Q}\left(s < \tau \leq t\right) &= \mathbb{Q}\left(\Gamma(s) < \Gamma(\tau) \leq \Gamma(t)\right) = \mathbb{Q}\left(\Gamma(s) < \xi \leq \Gamma(t)\right) \\ &= \mathbb{Q}\left(\xi > \Gamma(s)\right) - \mathbb{Q}\left(\xi > \Gamma(t)\right) = \mathrm{e}^{-\Gamma(s)} - \mathrm{e}^{-\Gamma(t)}\end{aligned}$$

$$= \mathrm{e}^{-\int_0^s \lambda(u)\,du} - \mathrm{e}^{-\int_0^t \lambda(u)\,du} \approx \int_s^t \lambda(u)\,du, \tag{2.15}$$

where the approximation is satisfied for small exponents. Conditional on survival at time t, as for (2.9), we calculate the probability of default in arbitrarily small dt

$$\mathbb{Q}\left(\tau \in [t, t+dt) | \tau \geq t\right) \approx \lambda(t)\,dt. \tag{2.16}$$

Using the time-varying intensity we have the following limitations [19]: 1. ξ is imposed externally so that it is independent of all default free market quantities; 2. Although the time-varying intensity allows us to consider a possible term structure of credit spreads, however, this formulation does not take into account credit spread volatility. For those reasons we turn to a stochastic intensity in the next section.

2.4.3 Stochastic Default Intensity

Generally, an important feature of the default intensity is randomness. To include the randomness, one can let λ_t to be a stochastic intensity process, which is $\mathcal{F}_t$-adapted and right continuous positive. The *cumulated intensity* or *hazard process* is thus defined by

$$\Lambda(t) = \int_0^t \lambda_s\,ds. \tag{2.17}$$

It is worth mentioning: $\mathcal{F}_t$ represents all (default-free) observable market information up to time t, this is to say that λ_t is known from 0 to t at time t; credit spread volatility can be introduced due to the stochasticity of λ_t. The generalized Poisson process with stochastic intensity is called *Cox process* and also known as *doubly stochastic Poisson process* . Furthermore, conditional on λ_t all facts for the case with $\lambda(t)$ will still hold for the stochastic intensity λ_t. Especially, the default time can be defined by inverting the

cumulated stochastic intensity as

$$\tau := \Lambda^{-1}(\xi). \tag{2.18}$$

We see that ξ and λ_t are both random variables, but ξ is independent of λ_t, more precisely, of by λ_t generated filtration $\mathcal{F}_t^{\lambda} = \sigma\left(\{\lambda_s : s \leq t\}\right)$.

For the survival probability we calculate:

$$\begin{aligned}
\mathbb{Q}\left(\tau > t\right) &= \mathbb{Q}\left(\Lambda(\tau) > \Lambda(t)\right) = \mathbb{Q}\left(\xi > \int_0^t \lambda_u \, du\right) \\
&= \mathbb{E}\left[\mathbb{Q}\left(\xi > \int_0^t \lambda_u \, du | \mathcal{F}_t^{\lambda}\right)\right] \\
&= \mathbb{E}\left[\mathrm{e}^{-\int_0^t \lambda_u \, du}\right].
\end{aligned} \tag{2.19}$$

Again, we can compare with the zero bond price under a stochastic interest rate

$$D(0,t) = \mathbb{E}\left[\mathrm{e}^{-\int_0^t r_u \, du}\right] \tag{2.20}$$

and recognize that the Cox process allows to use the technologies and methodologies for stochastic short rate to model default. The default probability between s and t can be calculated as

$$\begin{aligned}
\mathbb{Q}\left(s < \tau \leq t | \, \mathcal{F}_s\right) &= \mathbb{Q}\left(\Gamma(s) < \Gamma(\tau) \leq \Gamma(t) | \mathcal{F}_s\right) \\
&= \mathbb{Q}\left(\xi > \Gamma(s) | \mathcal{F}_s\right) - \mathbb{Q}\left(\xi > \Gamma(t) | \mathcal{F}_s\right) \\
&= \mathrm{e}^{-\int_0^s \lambda_u \, du} - \mathrm{e}^{-(\int_0^s \lambda_u \, du + \int_s^t \lambda_u \, du)} \\
&\approx \int_s^t \lambda_u \, du,
\end{aligned} \tag{2.21}$$

where the approximation is accurate for small exponents. Analogously, the default probability in dt years, given no default event so far and given default-free market information,

can be immediately approximated by

$$\mathbb{Q}\left(\tau \in [t, t+dt) | \tau \geq t, \mathcal{F}_t\right) \approx \lambda_t \, dt. \tag{2.22}$$

For more information about survival probability we refer to [19, 75], for instant.

2.4.4 Impose Default Correlation

As explained before, in order to consider WWR we should introduce a dependence between counterparties. For this aim one way is to impose a dependence between default times of counterparties which will be discussed in this section. We consider the case of using stochastic intensity. Let us assume that we have two counterparties "1" and "2" with the following default times (cf. (2.18)):

$$\tau_1 := \Lambda_1^{-1}(\xi_1), \quad \tau_2 := \Lambda_2^{-1}(\xi_2). \tag{2.23}$$

Intuitively, we can impose the dependence either on the stochastic intensities $\lambda_{1,t}$ and $\lambda_{2,t}$ using two independent ξ_1 and ξ_2, or on the exponential distributed random variables ξ_1 and ξ_2 assuming two independent stochastic intensities. Especially, one can even put dependence together in both ways which will be complicated. For a review of these strategies we refer to [19].

We assume that the stochastic intensities are given by

$$d\lambda_{1,t} = a(t, \lambda_{1,t}) \, dt + b(t, \lambda_{1,t}) \, dW_{1,t}, \tag{2.24}$$

$$d\lambda_{2,t} = a(t, \lambda_{2,t}) \, dt + b(t, \lambda_{2,t}) \, dW_{2,t}. \tag{2.25}$$

The dependence can thus be induced by the correlated BMs

$$dW_{1,t} dW_{2,t} = \rho_{12} \, dt. \tag{2.26}$$

On the other hand, we take only two independent BMs $W_{1,t}$ and $W_{2,t}$, but put a dependence between ξ_1 and ξ_2 : By the fact that

$$U_i = 1 - \mathrm{e}^{-\xi_i}, \quad i = 1, 2 \tag{2.27}$$

are uniform random variables on $[0, 1]$, we can employ a copula function with a correlation matrix to indirectly introduce the dependence between ξ_1 and ξ_2,

$$C_R(u_1, u_2) = \mathbb{Q}\left(U_1 < u_1, U_2 < u_2\right), \tag{2.28}$$

with the correlation matrix

$$R = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{21} & 1 \end{pmatrix}. \tag{2.29}$$

A large amount of literature about copulas and its application to the intensity models is available, we refer to [19, 83, 89, 99]. More precisely, given the correlation matrix (2.29) we generate the uniform distributed random numbers using (2.28), then use them by (2.27) to get the correlated standard exponential distributed random numbers. As mentioned, we can also select both ways, namely taking correlated BMs and correlated exponential random variables, however, this will then be turn out to be complicated. Furthermore, the mechanism for imposing correlations on top can also be generalized to higher dimensions, i.e. for more counterparties.

Chapter 3

Application to Credit Default Swaps (CDS)

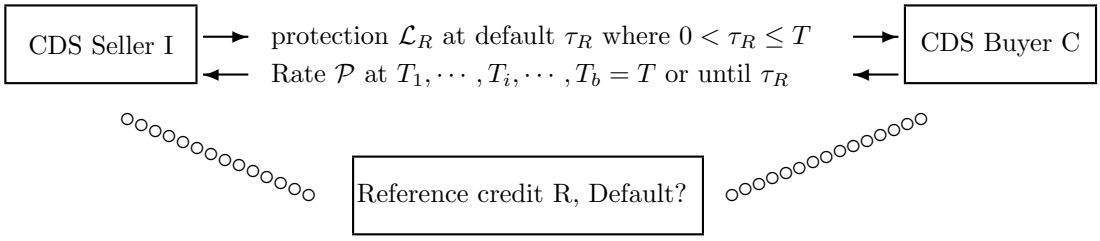
In this chapter we investigate computing BVA on CDS which is one of the most common credit derivatives. We point out a computational problem of using the BVA-formula provided in [15] (abridged version [16]) and show how to address this problem by employing particular numerical methods. Another important contribution of this chapter is a new BVA-formula which allows simultaneous defaults among counterparties, where the default intensity is modelled by applying a Markov copula model.

3.1 Bilateral Credit Valuation Adjustment on a CDS

In this section, we briefly introduce CDS and review the general arbitrage-free valuation framework for bilateral counterparty risk adjustments in [15] and [16] and its application to a CDS.

3.1.1 Credit Default Swap

A CDS is a financial swap agreement between two counterparties where the CDS buyer is protected against the loss in the default event by the reference credit as he will be compensated by the CDS seller. In turn, as long as there is no default event, the buyer pays a rate at certain times until maturity to the seller. More precisely, if we label by "C", "I" and "R" the counterparty (as CDS buyer), the investor (as CDS seller) and the underlying reference entity. With a CDS contract "C" and "I" will be in agreement as follows:



If "R" defaults at time $\tau_R \in (0, T)$, the investor pays the counterparty a certain cash amount according to L_{GD} of the underlying reference entity denoted by $\mathcal{L}_R$ which is called the *protection leg*. Conversely, the counterparty needs to pay the investor a premium rate denoted by $\mathcal{P}$ at certain times, say $(T_0 = 0), T_1, \cdots, T_i, \cdots, T_b = T$, this is the *premium leg*. Following this principle we can formally write the discounted payoff of a CDS with a default-free counterparty at time t from the perspective of the investor as the protection seller, given by the following definition.

Definition 3.1.1 *We define the discounted payoff of a CDS with a default-free counterparty at time t as [19]:*

$$\begin{aligned}\Pi(t,T) := {} & D(t,\tau_R)(\tau_R - T_{\gamma(\tau_R)-1})\mathcal{P}\mathbb{1}_{\{T_0<\tau_R<T_b\}} \\ & + \sum_{i=1}^{b} D(t,T_i)\alpha_i\mathcal{P}\mathbb{1}_{\{\tau_R \geq T_i\}} - D(t,\tau_R)\mathcal{L}_R\mathbb{1}_{\{T_0<\tau_R\leq T_b\}},\end{aligned} \tag{3.1}$$

where $t \in [T_{\gamma(t)-1}, T_{\gamma(t)})$, *i.e.* $T_{\gamma(t)}$ *is the first date among the* $T_i's$ *that follows* t, *and where* α_i *is the year fraction between* T_{i-1} *and* T_i. *We assume* $\mathcal{L}_R$ *to be deterministic and* $\mathcal{L}_R = 1 - \mathcal{R}_R$, *where* $\mathcal{R}_R$ *is also assumed to be deterministic and the notional is set to one.*

As we have seen in the introduction of CDS above, a single protection payment is made exactly at the default time $\tau_R \in (0, T)$, this CDS is called a running CDS. We remark that the time and the type for protection payment can be agreed upon in different ways. This leads to different definitions of a CDS, e.g., for the case of postponing the protection payment to the first time T_i following τ_R one defines a postponed payoff running CDS. In particular, in an upfront CDS there is an upfront payment made by the protection buyer in addition to the premium rate to match the present value of protection payment. In this work we consider only the running CDS (always called CDS for short) and the values or cash flows always from the perspective of the investor as the protection seller.

General Set-Up

We consider a probability space $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{Q})$ and define the enlarged filtration $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t$, $\forall t \in \mathbb{R}^+$ to model the whole information in the market, where the right-continuous and complete sub-filtration $\mathcal{F}_t$ represents all the observable market quantities and

$$\mathcal{H}_t = \sigma(\{\tau_R \leq u\} \vee \{\tau_C \leq u\} \vee \{\tau_I \leq u\} : \ u \leq t) \tag{3.2}$$

denotes the right-continuous filtration generated by the default events of three names under contract. In particular, the random times as non-negative random variables τ_j, $j = C, I, R$ are $\mathcal{G}_t$ stopping times (see B.5) for $t \in \mathbb{R}^+$, the stopped filtrations are thus given by

$$\mathcal{G}_{\tau_j} = \sigma\left(\mathcal{G}_t \cup \{t \leq \tau_j\}, \ t \geq 0\right). \tag{3.3}$$

This introduced setup will be used for analyzing CCR on CDS throughout this thesis.

Counterparty risk-free CDS price

Definition 3.1.2 *We denote by P_t^{CDS} the price of a counterparty risk-free CDS contract at time t and maturing at time T which is given by*

$$P_t^{CDS}(\mathcal{P}, \mathcal{L}_R) := \mathbb{E}\left\{\Pi(t,T)|\mathcal{G}_t\right\}, \quad t \in [0,T], \tag{3.4}$$

where $\Pi(t,T)$ is defined in (3.1).

Proposition 3.1.1 *Calculating the expectation* (3.4) *for $t = 0$ we obtain the CDS price "today" as*

$$\begin{aligned} P_0^{CDS}(\mathcal{P}, \mathcal{L}_R) = \mathcal{P}\Bigg[&- \int_0^{T_b} D(0,t)(t - T_{\gamma(t)-1})\, d\mathbb{Q}(\tau_R > t) \\ &+ \sum_{i=1}^{b} \alpha_i D(0,T_i)\mathbb{Q}(\tau_R > T_i)\Bigg] + \mathcal{L}_R\Bigg[\int_0^{T_b} D(0,t)\, d\mathbb{Q}(\tau_R > t)\Bigg]. \end{aligned} \tag{3.5}$$

For the detailed calculation we refer the reader to [19]. Similarly, we can straightforwardly update (3.5) to the price evaluated at time t, with $0 = T_0 < t < T_b = T$ and conditioning on the available information in the market at t, which is stated by the following proposition.

Proposition 3.1.2

$$\begin{aligned} P_t^{CDS}(\mathcal{P}, \mathcal{L}_R) = \mathbb{1}_{\{\tau_R > t\}} \Bigg\{ \mathcal{P}\Bigg[&- \int_t^{T_b} D(t,u)(u - T_{\gamma(u)-1})\, d\mathbb{Q}(\tau_R > u|\mathcal{G}_t) \\ &+ \sum_{i=\gamma(t)}^{b} \alpha_i D(t,T_i)\mathbb{Q}(\tau_R > T_i|\mathcal{G}_t)\Bigg] \\ &+ \mathcal{L}_R\Bigg[\int_t^{T_b} D(t,u)\, d\mathbb{Q}(\tau_R > u|\mathcal{G}_t)\Bigg]\Bigg\}. \end{aligned} \tag{3.6}$$

3.1.2 Arbitrage-free Valuation of bilateral CCR for a CDS

In the last section, we have seen the payoff formula of CDS under the assumption that the investor and the counterparty are both default-free. Indeed, in order to consider bilateral CCR in a CDS contract we have to take the default risk of all the parties into account. Following [15], in this section we evaluate bilateral CCR in a CDS contract in this section.

Let us define a new stopping time

$$\tau_{\text{IC}} := \min\{\tau_I, \tau_C\}, \tag{3.7}$$

which is the first default time between investor and his counterparty. According to τ_{IC} one can distinguish the following three cases for a traded CDS contract between the investor and his counterparty:

- $\tau_{\text{IC}} > T_b = T$: there is no default event of the investor and his counterparty during the life of the contract.

- $\tau_{\text{IC}} = \tau_C$: the counterparty defaults firstly. For this case, the residual value of CDS at τ_{IC} (or τ_C) until T, namely $P^{\text{CDS}}_{\tau_{\text{IC}}}(\mathcal{P}, \mathcal{L}_R)$ (see (3.6)) will play a key role. If $P^{\text{CDS}}_{\tau_{\text{IC}}}(\mathcal{P}, \mathcal{L}_R)$ is positive for the defaulted counterparty, he will then be paid completely by the investor. Contrarily, if the residual value is negative for the defaulted counterparty, then he can only pay a recovery fraction $\mathcal{R}_C$ of the exchanged $P^{\text{CDS}}_{\tau_{\text{IC}}}(\mathcal{P}, \mathcal{L}_R)$ to the investor. We can formally write this as

$$\mathcal{R}_C(P^{\text{CDS}}_{\tau_C}(\mathcal{P}, \mathcal{L}_R))^+ - (-P^{\text{CDS}}_{\tau_C}(\mathcal{P}, \mathcal{L}_R))^+. \tag{3.8}$$

- $\tau_{\text{IC}} = \tau_I$: the investor has a first default. Similar to the latter case, if $P^{\text{CDS}}_{\tau_I}$ is positive for the defaulted investor, the value will be completely paid by his counterparty. If it is negative value for the defaulted investor, his counterparty

can only receive a recovery fraction $\mathcal{R}_I$ of that amount. Formally,

$$(P^{\text{CDS}}_{\tau_I}(\mathcal{P}, \mathcal{L}_R))^+ - \mathcal{R}_I(-P^{\text{CDS}}_{\tau_I}(\mathcal{P}, \mathcal{L}_R))^+. \tag{3.9}$$

To write all terms in one formula we define the following events ordering the default times which should be mutually exclusive and exhaustive,

$$\begin{aligned} &\mathfrak{A} = \{\tau_I \le \tau_C \le T\}, \quad \mathfrak{B} = \{\tau_I \le T \le \tau_C\}, \quad \mathfrak{C} = \{\tau_C \le \tau_I \le T\}, \\ &\mathfrak{D} = \{\tau_C \le T \le \tau_I\}, \quad \mathfrak{E} = \{T \le \tau_I \le \tau_C\}, \quad \mathfrak{F} = \{T \le \tau_C \le \tau_I\}. \end{aligned} \tag{3.10}$$

Obviously, in the events $\mathfrak{E}$ or $\mathfrak{F}$ which means no defaults between the counterparties, the discounted value of the contract at time t is exactly the same to (3.1). For this we can write

$$\mathbb{1}_{\mathfrak{E}\cup\mathfrak{F}}\,\Pi(t, T). \tag{3.11}$$

If the counterparty defaults first, namely the events $\mathfrak{C}$ or $\mathfrak{D}$, we need to consider the value before the default $\Pi(t, \tau_C)$ and discount the residual value given by (3.8), formally given by

$$\mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}}\left\{\Pi(t, \tau_C) + D(t, \tau_C)[\mathcal{R}_C(P^{\text{CDS}}_{\tau_C}(\mathcal{P}, \mathcal{L}_R))^+ - (-P^{\text{CDS}}_{\tau_C}(\mathcal{P}, \mathcal{L}_R))^+]\right\}. \tag{3.12}$$

Similar in the case where the investor has a default for which we can obtain

$$\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\left\{\Pi(t, \tau_I) + D(t, \tau_I)[(P^{\text{CDS}}_{\tau_I}(\mathcal{P}, \mathcal{L}_R))^+ - \mathcal{R}_I(-P^{\text{CDS}}_{\tau_I}(\mathcal{P}, \mathcal{L}_R))^+]\right\}. \tag{3.13}$$

All together this leads to the following definition:

Definition 3.1.3 *We denote the discounted payoff of a CDS with a defaultable counterparty at time t and maturing at time T by $\overline{\Pi}(t,T)$:*

$$\begin{aligned}\overline{\Pi}(t,T) &= \mathbb{1}_{\mathfrak{C}\cup\mathfrak{F}}\,\Pi(t,T)\\ &+ \mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}}\left\{\Pi(t,\tau_C) + D(t,\tau_C)[\mathcal{R}_C(P^{CDS}_{\tau_C}(\mathcal{P},\mathcal{L}_R))^+ - (-P^{CDS}_{\tau_C}(\mathcal{P},\mathcal{L}_R))^+]\right\}\\ &+ \mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\left\{\Pi(t,\tau_I) + D(t,\tau_I)[(P^{CDS}_{\tau_I}(\mathcal{P},\mathcal{L}_R))^+ - \mathcal{R}_I(-P^{CDS}_{\tau_I}(\mathcal{P},\mathcal{L}_R))^+]\right\},\end{aligned} \tag{3.14}$$

where $\Pi(t,T)$ is defined in (3.1).

Obviously, the expectation of (3.14) under $\mathbb{Q}$ is the price of the CDS contract traded by defaultable counterparties.

Definition 3.1.4 *We denote by $\overline{P}^{CDS}_t$ the price of a counterparty defaultable CDS contract at time t and maturing at time T which is given by*

$$\overline{P}^{CDS}_t(\mathcal{P},\mathcal{L}_R) := \mathbb{E}\left\{\overline{\Pi}(t,T)|\mathcal{G}_t\right\}, \quad t\in[0,T], \tag{3.15}$$

where $\overline{\Pi}(t,T)$ is defined in (3.14).

Using the finding (general bilateral counterparty risk pricing formula) in [15], the price $\overline{P}^{\mathrm{CDS}}_t$ can be represented by P^{CDS}_t with additional terms so that we can obtain BVA formula for the CDS.

Proposition 3.1.3 *The price of a counterparty defaultable CDS contract at time t and maturing at time T is given by*

$$\begin{aligned}\mathbb{E}\left[\overline{\Pi}(t,T)|\mathcal{G}_t\right] =& \mathbb{E}\left[\Pi(t,T)|\mathcal{G}_t\right]\\ &+ \mathbb{E}\left[\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\cdot\mathcal{L}_I\cdot D(t,\tau_I)\cdot(-P^{CDS}_{\tau_I}(\mathcal{P},\mathcal{L}_R))^+|\mathcal{G}_t\right]\\ &- \mathbb{E}\left[\mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}}\cdot\mathcal{L}_C\cdot D(t,\tau_C)\cdot(P^{CDS}_{\tau_C}(\mathcal{P},\mathcal{L}_R))^+|\mathcal{G}_t\right].\end{aligned} \tag{3.16}$$

The detailed proof is given in [15]. Observing (3.16) one can find that the value of a CDS contract under defaultable counterparties is the value of the identical CDS contract

under default-free counterparties plus a long position in a put-option (with zero strike) on the residual value of CDS at the possible default time τ_I and plus a short position in a call-option (with zero strike as well) on the residual value of CDS at the possible default time τ_C. Furthermore, as introduced before, CVA is defined as the difference between the value of a default-free credit derivative and the fair value of this derivative when the counterparty default possibility is taken into account. In our case, we have considered both default possibilities of the investor and his counterparty in a CDS contract. The BVA expression can thus be directly given by the following proposition.

Proposition 3.1.4 *The bilateral Credit Valuation Adjustment for a CDS contract at time t and maturing at time T is given by*

$$\begin{aligned} BVA(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}}) := \mathcal{L}_C \cdot \mathbb{E}\left[\mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}} \cdot D(t,\tau_C) \cdot (P^{CDS}_{\tau_C}(\mathcal{P},\mathcal{L}_R))^+ | \mathcal{G}_t\right] \\ - \mathcal{L}_I \cdot \mathbb{E}\left[\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}} \cdot D(t,\tau_I) \cdot (-P^{CDS}_{\tau_I}(\mathcal{P},\mathcal{L}_R))^+ | \mathcal{G}_t\right]. \end{aligned} \tag{3.17}$$

Remark 3.1.1 *The value of $BVA(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}})$ might be positive or negative depending on whether the counterparty is more or less likely to default than the investor and certainly also on the correlation among all the counterparties.*

Remark 3.1.2 *The above discussion shows that the BVA equals the sum of the value of a long position in a zero-strike call option on the residual price of CDS and the value of a short position in a zero-strike put option on the residual price of CDS. The option only gives a contribution, if the corresponding party defaults earlier.*

Remark 3.1.3 *From* (3.17) *we realize that BVA (as seen from the investor) = CVA Counterparty − CVA Investor. As explained in Section 2.2 we know CVA Counterparty = DVA Investor, so BVA to Investor = DVA Investor − CVA Investor.*

3.2 Computing of Bilateral Credit Valuation Adjustment on a CDS

In order to compute (3.17) we use a stochastic intensity model introduced in Section 2.4. Following [14, 15, 18] we assume that the intensities in a CDS contract are given by

$$\lambda_{j,t} = y_{j,t} + \psi_j(t, \beta_j), \quad t \geq 0,\ j \in \{C, I, R\}, \tag{3.18}$$

where ψ_j is a deterministic function and each y_j follows a *Cox-Ingersoll-Ross (CIR) process [33]* given by

$$dy_{j,t} = \kappa_j(\mu_j - y_{j,t})\, dt + \sigma_j \sqrt{y_{j,t}}\, dW_{j,t}. \tag{3.19}$$

We remark: 1) The variable $\beta_j(y_{0,j}, \kappa_j, \mu_j, \sigma_j)$ of the function ψ_j is a vector whose components are all deterministic constants and β_j should be positive as well. 2) It is suggested to relax the Feller condition $2\kappa_j\mu_j > \sigma_j$ so that the CDS implied volatility will be not limited, see, e.g., [14] and [15]. For this the argument is that the assumed positivity of ψ_j does not allow λ_j to attain a zero value. 3) We assume that BMs under the risk neutral measure are independent for $j \in \{C, I, R\}$, as we will impose the correlation through a copula function.

We define the integrated quantities which will be used in the remainder of this section as follows:

$$\Lambda_j(t) = \int_0^t \lambda_{j,s}\, ds, \quad Y_j(t) = \int_0^t y_{j,s}\, ds, \quad \Psi_j(t) = \int_0^t \psi_{j,s}\, ds. \tag{3.20}$$

Then, the default times can be defined as

$$\tau_j := \Lambda_j^{-1}(\xi_j), \quad j \in \{C, I, R\}, \tag{3.21}$$

where ξ_C, ξ_I and ξ_R are standard exponential random variables with associated uniforms

$$U_j := 1 - \mathrm{e}^{-\xi_j}, \tag{3.22}$$

they are correlated through a trivariate copula function

$$C_{\mathcal{R}}(u_C, u_I, u_R) = \mathbb{Q}\left(U_C < u_C, U_I < u_I, U_R < u_R\right) \tag{3.23}$$

with the correlation matrix

$$\mathcal{R} = \begin{pmatrix} 1 & \rho^{CI} & \rho^{CR} \\ \rho^{IC} & 1 & \rho^{IR} \\ \rho^{RC} & \rho^{RI} & 1 \end{pmatrix}. \tag{3.24}$$

We can choose different kinds of copula functions, e.g., Gaussian copula, t-copula and so on. Besides, we notice that a trivariate copula implies a bivariate marginal copula. Before showing an example for computing (3.17) we firstly introduce how to imply survival probabilities from the market CDS curve, which is an essential part in a reduced form model. Secondly, we see how to calculate the conditional survival probabilities in (3.17) and point out the computational challenges.

3.2.1 Implied Survival Probabilities from the Market CDS Curve

In practice, the reduced form models have been most commonly used to imply the survival probabilities from market spreads. The idea is: we assume a time inhomogeneous Poisson process, with time varying intensity $\lambda(t)$ and hazard function $\Lambda(t) = \int_0^t \lambda(u)\, du$ which we have introduced in Section 2.4.2. One can take the CDS spread for T_b years to be premium rate $\mathcal{P}$ and solve $P_0^{\mathrm{CDS}}(\mathcal{P}, \mathcal{L}_R) = 0$ in (3.5) for the implied survival prob-

abilities over the payment dates $T_1, T_2, \cdots, T_b$ iteratively. This process of iteratively searching for the survival probabilities is known as *bootstrapping procedure* CDS curve . We have to know that the bootstrapping is model independent, say we just imply the market survival probabilities $\mathbb{Q}(\tau > t)_{\text{market}}$ from the market CDS curve. In Section 21.3.5. and 22.3 in [19], Brigo and Mercurio have described this bootstrapping in detail. Besides, there is also a matlab routine `cdsbootstrap` available in the financial instruments toolbox. Therefore, we do not repeat it and skip to an example for the illustrative purpose. We consider the market spread quotes for Lehman Brothers on May 1, 2008 which are presented in Table 3.1 below. For the calibration we may take the

Maturity	1y	2y	3y	4y	5y
Spread	203	188.5	166.75	152.25	145

Table 3.1: Market spread quotes in basis points for Lehman Brothers on May 1, 2008.

time varying $\lambda(t)$ to be piecewise constant or linear. We calculate the discount factors using the interest rates on the corresponding dates as spread quotes. One can choose stochastic discount factor or deterministic discount factor, because the interest rates will be anyway independent with the default time due to the deterministic intensity $\lambda(t)$. We set the deterministic recovery rate to be 40% and display the calibrated piecewise constant intensity and survival probability in Figure 3.1, the piecewise linear intensity and survival probability in Figure 3.2.

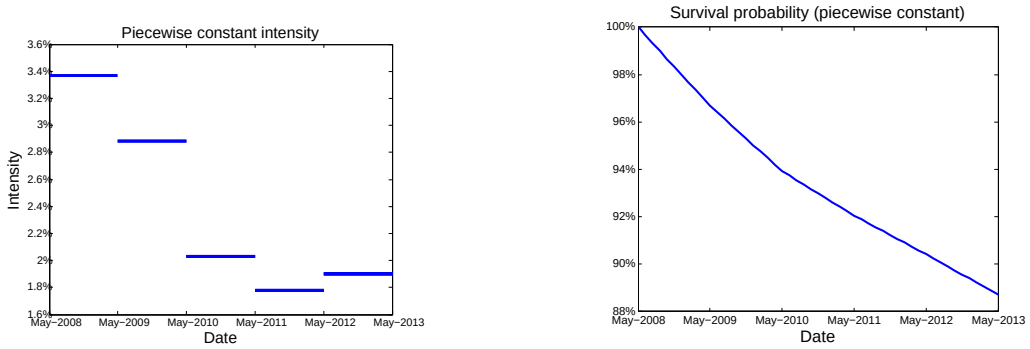


Figure 3.1: Piecewise constant intensity $\lambda(t)$ calibrated on spreads in 3.1 and corresponding survival probability

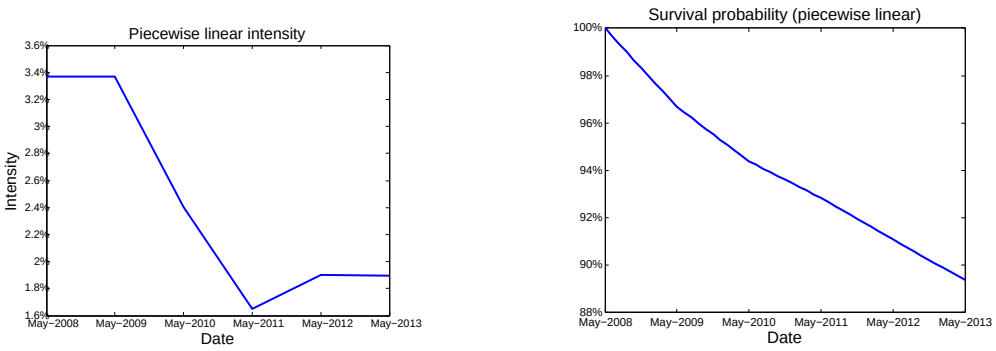


Figure 3.2: Piecewise linear intensity $\lambda(t)$ calibrated on spreads in Table 3.1 and corresponding survival probability

A comparison between using piecewise constant and linear intensity is provided in [19].

3.2.2 Survival Probability and its Computational Challenges

From (3.17) and (3.6) we observe that the only terms we need to know for using a Monte-Carlo evaluation to compute BVA are

$$\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\tau_R > t|\mathcal{G}_{\tau_I}) \tag{3.25}$$

and

$$\mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}}\mathbb{1}_{\tau_R>\tau_C}\mathbb{Q}(\tau_R > t|\mathcal{G}_{\tau_C}). \tag{3.26}$$

The formulas in a closed form for survival probabilities (3.25) and (3.26) have been found by Brigo and Capponi [15, 26], we adopt these formulas to a simpler one. We state our results in the following propositions. We define firstly

$$\overline{U}_{j,k} := 1 - \mathrm{e}^{-\Lambda_j(\tau_k)}, \quad j,k \in \{C,I,R\} \tag{3.27}$$

and denote the cumulative distribution function of the integrated CIR process $Y_j(t)$ by $F_{Y_j(t)}$. We thus have

Proposition 3.2.1

$$\begin{aligned}\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\tau_R>t|\mathcal{G}_{\tau_I}) = \mathbb{1}_{\tau_I\le T}\mathbb{1}_{\tau_I\le\tau_C}\big(\mathbb{1}_{\{t<\tau_I<\tau_R\}}+\\ \mathbb{1}_{\tau_I\le t}\mathbb{1}_{\tau_R>\tau_I}\int_{\overline{U}_{R,I}}^{1} F_{Y_R(t)}(-\log(1-u_R)-\Psi_R(t;\,\beta_R))dC_{R|I}(u_R;\,U_I)\big),\end{aligned} \tag{3.28}$$

with

$$C_{R|I}(u_R;\,U_I) = \frac{\frac{\partial C_{I,\,R}(u_I,u_R)}{\partial u_I}|_{u_I=U_I} - \frac{\partial C_{I,\,R}(u_I,\overline{U}_{R,\,I})}{\partial u_I}|_{u_I=U_I}}{1-\frac{\partial C_{I,\,R}(u_I,\overline{U}_{R,I})}{\partial u_I}|_{u_I=U_I}}. \tag{3.29}$$

The proof of the proposition can be found in Appendix C. And similarly,

Proposition 3.2.2

$$\begin{aligned}\mathbb{1}_{\mathfrak{C}\cup\mathfrak{D}}\mathbb{1}_{\tau_R>\tau_C}\mathbb{Q}(\tau_R>t|\mathcal{G}_{\tau_C}) = \mathbb{1}_{\tau_C\le T}\mathbb{1}_{\tau_C\le\tau_I}\big(\mathbb{1}_{\{t<\tau_C<\tau_R\}}+\\ \mathbb{1}_{\tau_C\le t}\mathbb{1}_{\tau_R>\tau_C}\int_{\overline{U}_{R,C}}^{1} F_{Y_R(t)}(-\log(1-u_R)-\Psi_R(t;\,\beta_R))dC_{R|C}(u_R;\,U_C)\big),\end{aligned} \tag{3.30}$$

with

$$C_{R|C}(u_R;\,U_C) = \frac{\frac{\partial C_{R,\,C}(u_R,u_C)}{\partial u_C}|_{u_C=U_C} - \frac{\partial C_{R,\,C}(\overline{U}_{R,C},u_C)}{\partial u_C}|_{u_C=U_C}}{1-\frac{\partial C_{R,\,C}(\overline{U}_{R,C},u_C)}{\partial u_C}|_{u_C=U_C}}. \tag{3.31}$$

We remark that the partial derivatives in (3.29) and (3.31) hold a formula in closed-form for some copula functions, e.g., Gaussian copula. Otherwise, one also can approximate the partial derivative numerically

$$\frac{\partial C(u_1,u_2)}{\partial u_2} = \lim_{\Delta\to 0^+}\frac{C(u_1,u_2+\Delta)-C(u_1,u_2)}{\Delta}. \tag{3.32}$$

Intuitively, the CDF of the integrated CIR process can be transformed by the approach of Carr and Madan [27] for numerically determining the option values using the Fast Fourier Transform (FFT) from the corresponding (analytically known) characteristic function. Since the characteristic function is given, an analytic expression for the Fourier transformed probability density can be developed and then solved numerically using FFT techniques. Unfortunately, we must tolerate a restriction between the grid

size for infinitesimal summands and the output grid size when applying the FFT. This is not convenient for application, because the generated CDF will have somewhat a restriction on the grids as well.

Chourdakis [30] adapted this methodology proposing the fractional Fast Fourier Transform (FRFT) instead of the FFT for the purpose of removing the grid sizes restriction. However, numerical tests for several CDFs showed that this advantage of the FRFT did not outweigh the speed of the FFT in our application. This was also mentioned for pricing options in [47]. Besides, jointly with the FRFT the so called control parameter (dampening parameter) is also introduced as in the FFT to resolve the problem of the divergence of the integrand at zero. For using the FFT to determine numerically the option values, Lord and Kahl [78] have explained how to choose the optimal dampening parameter. The choice of this parameter is essential and strongly depends on the model parameters. However, for the case of using the FRFT to determine numerically the CDF of the integrated CIR process, it is still unclear how to select an optimal dampening parameter. On the other hand, an integration of the characteristic function over the infinite domain is numerically instable due to cancellation effects and the fast growth of the characteristic function.

For this reason, we propose a new approach (see Section 3.3) to deal with such instability problems mentioned above. This new strategy allows to construct a very robust routine to determine numerically a highly accurate CDF of the integrated CIR process for almost any choices of parameters, so that highly accurate survival probabilities can be evaluated.

3.2.3 An example of Computing BVA with Wrong-Way risk

This section is dedicated to the example of evaluating BVA defined in (3.17) in a CDS contract. In particular, we present numerical results for some different default correlations. By analyzing the numerical results we stress the role of default correlations on

the WWR.

The Monte-Carlo approach

A Monte-Carlo approach is used to evaluate BVA, we introduce briefly the steps in implementing the numerical algorithm below:

- Given initial parameters we simulate CIR processes defined in (3.19) for the three names, namely $j \in \{C, I, R\}$. The parameters can be calibrated using market spread quotes. The simulation can be done according to the fact that the distribution of y_t given y_s as defined (3.19), for some $s < t$ follows, up to a scale factor, a noncentral χ^2-distribution. Alternatively, we can simulate sample paths numerically, see e.g., [2, 46, 82, 87].

- From the break-even spreads (generated using initial parameters) or market spread quotes we imply the market survival probabilities $\mathbb{Q}_{\text{market}}(\tau_j > t)$ for $j \in \{C, I, R\}$. It has been introduced in Section 3.2.1 how to imply market survival probabilities from the market spread quotes. By imposing

$$\mathbb{Q}_{\text{market}}(\tau_j > t) = \mathbb{Q}_{\text{model}}(\tau_j > t) \tag{3.33}$$

we obtain the quantities $\Psi_j(t)$ defined in (3.20) as

$$\Psi_j(t) = \ln\left(\frac{\mathbb{Q}_{\text{model}}(\tau_j > t)}{\mathbb{Q}_{\text{market}}(\tau_j > t)}\right), \tag{3.34}$$

where $\mathbb{Q}_{\text{model}}(\tau_j > t) = \mathbb{E}[\mathrm{e}^{-Y_j(t)}] = D^{\text{CIR}}(0, t, \beta_j)$, and $D^{\text{CIR}}(0, t, \beta_j)$ is the price at time 0 of a zero coupon bond maturing at time t under a stochastic interest rate given by the corresponding CIR processes initialized by $\beta_j(y_{0,j}, \kappa_j, \mu_j, \sigma_j)$ and known analytically as

$$A(t, T) \cdot \mathrm{e}^{-B(t,T)y_t}, \tag{3.35}$$

where

$$A(t,T) := \left[\frac{2h\exp\{(\kappa+h)(T-t)/2\}}{2h+(\kappa+h)(\exp\{(T-h)h\}-1)}\right]^{\frac{2\kappa\mu}{\sigma^2}},$$
$$B(t,T) := \frac{2(\exp\{(T-t)h\}-1)}{2h+(\kappa+h)(\exp\{(T-h)h\}-1)},$$
$$h := \sqrt{\kappa^2+2\sigma^2},$$

for more information we refer the reader to, e.g., [19].

- We generate then the default times τ_j using equations (3.21)–(3.23), in our example we choose a Gaussian copula. Depending on the default situation with respect to the generated default times, we need to compute one survival probability of (3.28) and (3.30), or none of them when no defaults occurs. A robust routine to determine numerically a highly accurate CDF of the integrated CIR process, which is needed for those survival probabilities, has been provided in Section 3.3.

- Finally, one can compute the BVA (3.17) by numerical integration.

Numerical result

We consider a five years CDS contract traded by an investor with a counterparty on a reference name, the investor and the counterparty are both subject to the default risk. We assume that the investor has low credit risk, the reference name has high credit risk and the counterparty has middle credit risk. This is a common scenario in the real market. We state the values of the parameter vectors $\beta_j(y_{0,j},\kappa_j,\mu_j,\sigma_j)$ in Table 3.2. The parties agree that the loss given defaults of each credit risk level are set to

	Credit risk level	$y(0)$	κ	μ	Volatility	σ
β_C	middle	0.01	0.7	0.02	middle	0.02
β_I	low	0.0001	1.1	0.001	low	0.02
β_R	high	0.03	0.6	0.05	high	0.5

Table 3.2: The credit risk levels of each parties on the CDS contract

$\mathcal{L}_C = 0.65$, $\mathcal{L}_I = 0.6$, $\mathcal{L}_R = 0.7$, respectively. And they will make the payments every three months, which indicates that the year fraction is given by $\alpha = \frac{1}{4}$.

Using these values in Table 3.2 one can generate the corresponding break-even spreads by solving the equation (3.5) for $\mathcal{P}$. We report the break-even spreads generated using the parameter in Table 3.2 for $T \in \{1, \cdots 10\}$ years in Table 3.3, which will be used to imply the market survival probabilities for our example. Since the aim of this example

Maturity	β_C	β_I	β_R
1y	90	3	239
2y	102	4	252
3y	111	5	258
4y	116	6	262
5y	120	6	264
6y	123	6	266
7y	125	6	267
8y	127	6	268
9y	128	6	269
10y	129	6	269

Table 3.3: The break-even spreads in basis points generated using the parameters of the CIR processes in Table 3.2.

is to show the impact of correlation on BVA, thus, we compute the BVAs by varying the correlation between the counterparty and the reference credit ρ^{CR}, together with varying the volatility of the counterparty σ_C. We denote Monte-Carlo values of the BVA for the CDS payer and seller respectively with BVA_p and BVA_s and take the five-year CDS spread of the reference credit as the premium rate $\mathcal{P}$ which is 264 (basis points) in Table 3.3. Finally, we report our results in Table 3.4. We consider firstly the case of negative correlations: Almost all the $\text{BVA}_p s$ approach zero for the reason, at the reference credit's default, the counterparty has probably no default due to the negative relationship with the reference credit, thus no adjustments will be required by the investor; furthermore, the investor as a protection seller requires a certain adjustment, since the default probability of the counterparty increases when the default risk of the reference entity (the investor as seller holds an option which is in the money) reduces. At the counterparty's

default, the investor can only get a value which can be recovered by the counterparty. Then, it is easy to understand that BVA_s increases for a lower negative value of ρ^{CR}.

On the contrary, for high and positive values of ρ^{CR} : It is obvious that the investor as a CDS seller requires almost no adjustment. However, as a CDS purchaser he needs much more adjustment, because the default probability of the counterparty increases with the rising of the reference entity default risk due to the high positive correlation between them (the investor as purchaser holds an option which is in the money). At the counterparty's default, he can only receives a value which can be recovered by the counterparty. This is to say, one should expect that BVA_p increases with ρ^{CR}. However, we look at Table 3.4, this phenomenon can be only observed when ρ^{CR} increases from 0 to 0.9. For $\rho^{CR} = 0.99$, which is a relatively large value, BVA_p decreases contrarily instead. Especially, we look at the first column where $\sigma_C = 0.02$, BVA_p even approaches zero. This means that we observe a strong decreasing WWR with low volatility and high correlation.

$(\rho^{CI}, \rho^{CR}, \rho^{IR})$	σ_R	0.02	0.1	0.2	0.3	0.4	0.5
$(0, -0.99, 0)$	BVA_p	0.0 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)
	BVA_s	33.7 (0.9)	33.6 (0.9)	33.8 (0.9)	33.0 (0.9)	35.4 (0.9)	34.3 (0.9)
$(0, -0.9, 0)$	BVA_p	0.0 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.2 (0.0)	0.1 (0.0)
	BVA_s	33.5 (0.8)	33.6 (0.8)	33.8 (0.8)	32.7 (0.8)	33.8 (0.8)	31.7 (0.8)
$(0, -0.5, 0)$	BVA_p	0.0 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)
	BVA_s	29.4 (0.8)	28.3 (0.8)	29.1 (0.8)	27.3 (0.7)	25.8 (0.7)	25.2 (0.7)
$(0, -0.2, 0)$	BVA_p	0.0 (0.0)	0.1 (0.0)	0.1 (0.0)	0.1 (0.0)	0.2 (0.0)	0.2 (0.0)
	BVA_s	11.7 (0.3)	12.9 (0.4)	12.9 (0.4)	12.3 (0.4)	11.5 (0.3)	12.1 (0.3)
$(0, 0, 0)$	BVA_p	6.5 (0.2)	4.6 (0.2)	4.0 (0.1)	3.8 (0.1)	4.0 (0.1)	2.6 (0.1)
	BVA_s	0.4 (0.0)	1.4 (0.1)	2.1 (0.1)	2.3 (0.1)	2.0 (0.1)	2.7 (0.1)
$(0, 0.2, 0)$	BVA_p	27.0 (0.8)	22.6 (0.8)	21.0 (0.7)	19.6 (0.7)	17.8 (0.6)	17.2 (0.6)
	BVA_s	0.3 (0.0)	0.7 (0.1)	0.8 (0.1)	1.0 (0.1)	1.1 (0.1)	1.1 (0.1)
$(0, 0.5, 0)$	BVA_p	64.4 (2.1)	57.4 (2.0)	47.9 (1.7)	43.7 (1.6)	42.0 (1.5)	33.5 (1.3)
	BVA_s	0.2 (0.0)	0.4 (0.0)	0.4 (0.0)	0.4 (0.0)	0.4 (0.0)	0.4 (0.0)
$(0, 0.9, 0)$	BVA_p	82.3 (3.6)	65.6 (3.0)	54.3 (2.5)	52.6 (2.3)	53.7 (2.3)	52.0 (2.5)
	BVA_s	0.3 (0.0)	0.2 (0.0)	0.2 (0.0)	0.2 (0.0)	0.2 (0.0)	0.1 (0.0)
$(0, 0.99, 0)$	BVA_p	1.2 (0.0)	2.1 (0.4)	5.8 (0.5)	11.0 (0.7)	19.0 (1.0)	30.1 (1.4)
	BVA_s	0.3 (0.0)	0.2 (0.0)	0.1 (0.0)	0.2 (0.0)	0.2 (0.0)	0.1 (0.0)

Table 3.4: BVAs in basis points for the scenario parameterized using the parameters in Table 3.2 while varying the values of ρ^{CR} and σ_R. The numbers in round brackets are the Monte-Carlo standard errors.

The reason for this decreasing WWR is explained as follows: when the counterparty and the reference credit are correlated with $\rho^{CR} = 0.99$, then the exponential triggers ξ_C and ξ_R (see (3.22)) will be almost identical. At the same time σ_C and σ_R have a small value 0.02, both default intensity processes are not so random, such that the inequality $\lambda_R > \lambda_C$ can hold almost all the time. Thus, by a first approximation $\tau_R \approx \frac{\xi_R}{\lambda_R} < \tau_C \approx \frac{\xi_C}{\lambda_C}$, we find the reference credit will default almost always before the counterparty, therefore, no adjustments take place.

Now we can imagine, if we increase the value of σ_R such that the default intensity process λ_R is more random. This should result in adjustments even for $\rho^{CR} = 0.99$. As expected, we look at the last row in Table 3.4 and observe that BVA_p increases with the value of σ_R. From this example, we see that correlation plays a key role in credit risk modelling.

3.3 The Cumulative Distribution Function of the integrated CIR Process

As introduced in Section 3.2, the CDF of the integrated CIR process must be exactly known for computing the survival probability (3.28) and (3.30) which is one of the substantial parts for computing the BVA in a CDS contract. In this section, we adapt the approach by Kahl and Jäckel [63] for option pricing to evaluate the CDF of the integrated CIR process. This will allow us to construct a very robust routine to determine numerically a highly accurate CDF of the integrated CIR process for almost any choices of parameters.

3.3.1 The CDF of the integrated CIR

The characteristic function of the integrated CIR process (see (3.20)) under the risk-neutral probability measure is defined as

$$\phi_{Y_t}(u) = \mathbb{E}[\mathrm{e}^{iuY_t}] \tag{3.36}$$

and can be rewritten as [43, 73]

$$\phi_{Y_t}(u) = \mathrm{e}^{A(t,u)+B(t,u)y_0}, \tag{3.37}$$

with

$$A(t,u) := \frac{2\kappa\mu}{\sigma^2}\left(\ln(2) + \ln\left(\frac{\frac{b(u)}{\kappa-b(u)}\mathrm{e}^{\frac{(\kappa+b(u))t}{2}}}{a(u)\mathrm{e}^{b(u)t}-1}\right)\right), \tag{3.38}$$

$$B(t,u) := \frac{2ui}{\kappa-b(u)}\left(\frac{\mathrm{e}^{b(u)t}-1}{a(u)\mathrm{e}^{b(u)t}-1}\right), \tag{3.39}$$

where i denotes the imaginary unit and with the auxiliary functions

$$a(u) := \frac{\kappa+b(u)}{\kappa-b(u)}, \qquad b(u) := \sqrt[+]{\kappa^2-2\sigma^2ui}. \tag{3.40}$$

Here, $\sqrt[+]{\ }$ denotes the branch of the square root with positive real part.

Using an inverse Fourier transform we obtain the probability density function of the integrated CIR process as

$$f(\tilde{y}_t) := \int_0^\infty \frac{\mathrm{Re}[\mathrm{e}^{-iu\tilde{y}_t}\phi_{Y_t}(u)]}{\pi}\,du. \tag{3.41}$$

Many authors (e.g., Bakshi and Madan [5]) determined the corresponding CDF numerically by

$$F_{Y_t}(\tilde{y}_t) := \frac{1}{2} - \frac{1}{\pi}\int_0^\infty g(u)\,du, \tag{3.42}$$

where the function g is defined as

$$g(u) := \mathrm{Re}\left[\frac{\mathrm{e}^{-iu\tilde{y}_t}\phi_{Y_t}(u)}{iu}\right]. \tag{3.43}$$

Apparently, the fact that the integrand (3.43) diverges at $u = 0$ leads to a cumbersome numerical integration. We further observe that the numerical integration is made even more complicated by the fact that this integrand (3.43) can be highly oscillatory depending on the choice of parameters. Besides, the fast growth of the characteristic function (3.37) is hard to handle in general, because it depends strongly on the model parameters. Therefore, a simple quadrature or a naive numerical integration is not appropriate for the integration in (3.42). We show in the sequel how to apply the adaptive Gauss-Lobatto quadrature [52].

However, in order to use the adaptive Gauss-Lobatto quadrature we need to solve two problems. First, this Gauss-Lobatto algorithm is designed only to operate on finite intervals. Secondly, another problem is the complication in the calculation of the embedded complex logarithms in equation (3.38) when the function g is evaluated repeatedly in this quadrature scheme. In the remainder of this section, we show in Section 3.3.2 how to solve the first problem and turn in Section 3.3.3 to the second problem.

3.3.2 The Transformation to a Finite Interval

In this section we show how to transform the original unbounded domain of integration $[0, \infty)$ in (3.42) to the finite interval $[0, 1]$ for applying later the Gauss-Lobatto quadrature. This transformation relies on the asymptotic behaviour of the integrand for $u \to \infty$, see Proposition 3.3.1. This transformation strategy leads to an improved stability of the adaptive quadrature scheme, cf. [63]. Besides, this modified integration scheme is significantly more efficient since less quadrature points for the evaluation are needed.

Proposition 3.3.1 *For the CIR model parameters κ, μ, σ, $t > 0$ we obtain the following asymptotics:*

$$\lim_{u\to\infty} \frac{b(u)}{\sqrt{u}} = \sqrt{2}\sigma e^{\frac{7\pi}{4}i}, \tag{3.44}$$

$$\lim_{u\to\infty} a(u) = -1, \tag{3.45}$$

$$\lim_{u\to\infty} \frac{A(u)}{\sqrt{u}} = -\frac{\sqrt{2}\kappa\mu t}{\sigma} e^{\frac{7\pi}{4}i}, \tag{3.46}$$

$$\lim_{u\to\infty} \frac{B(u)}{\sqrt{u}} = \frac{\sqrt{2}i}{\sigma} e^{-\frac{7\pi}{4}i}. \tag{3.47}$$

The proof can be found in the Appendix C.

Proposition 3.3.2 *For the CIR model parameters κ, μ, σ, $t > 0$ we obtain the asymptotics for the function $g = g(u)$ defined in* (3.43)*:*

$$\lim_{u\to\infty} g(u) \approx \exp^{-\sqrt{u}A_\infty} \cdot \mathrm{Re}\left(\frac{e^{-iu\tilde{y}_t + i\sqrt{u}t_\infty}}{iu}\right) = e^{-\sqrt{u}A_\infty} \cdot \frac{\sin(\sqrt{u}t_\infty - u\tilde{y}_t)}{u}, \tag{3.48}$$

with

$$A_\infty = t_\infty = \frac{\kappa\mu t + y_0}{\sigma}. \tag{3.49}$$

The proof follows immediately from Proposition 3.3.1.

Obviously A_∞ is positive and from the equation (3.48) we can see that the integrand g defined in (3.43) has at least exponential asymptotic decay for $u \to \infty$. Hence, we simply transform the integration interval in (3.42) as

$$\int_0^\infty g(u)\, du = -\frac{2}{A_\infty^2} \int_0^1 \frac{\ln x}{x} g(u(x))\, dx, \tag{3.50}$$

with the change of variable

$$u(x) := \left(\frac{\ln x}{A_\infty}\right)^2. \tag{3.51}$$

Note that in earlier work of our research [104] $u(x) := -\frac{\ln x}{A_\infty}$ was used in the transformation. However, we find applying (3.51) can provide more stable and accurate results for extreme parameters.

Up to now we have achieved the desired transformation to a finite integration interval. The second issue is the choice of the branch of the multivalued complex logarithm embedded in $A(t,u)$ in (3.38) for calculations based on the inverse Fourier transform (3.42) of the function g. However, the restriction on the choice of the principal branch leads to a discontinuous function (3.43), which would lead to incorrect results. In the next section we will show how to guarantee the continuity of the function g in (3.50).

3.3.3 Numerical Evaluation of Complex Logarithms

First, let us recall that the function $g(u)$ has discontinuities if we simply select the principal branch of the complex logarithm in $A(t,u)$. In Figure 3.3 we first present the imaginary part of the function $A(t,u)$ as defined in (3.38), the discontinuities appear very clearly.

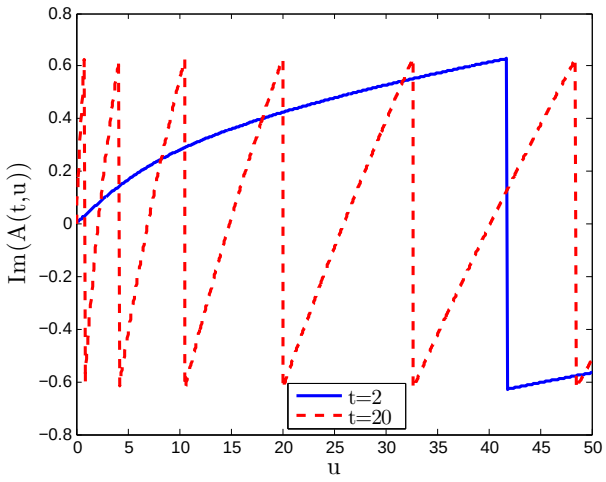


Figure 3.3: The function $A(t,u)$ defined in (3.38) with $\kappa = 0.5$, $\mu = 0.05$, $\sigma = 0.5$, $y_0 = 0.03$, implementation using the principal branch, blue curve: $t = 2$, red dashed curve: $t = 20$.

In particular, it is even worse for $t = 20$ and this observation explains why a simple approach for the integration in (3.42) must fail, since the integrands strongly depend on the chosen parameters. Figure 3.4 shows the function $g(u)$ which is implemented using its respective function $A(t, u)$, see Figure 3.3. For $t = 20$, the discontinuous peaks of the

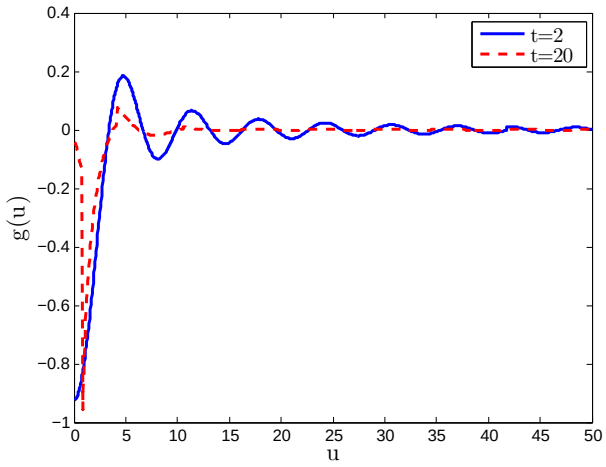


Figure 3.4: The function $g(u)$ defined in (3.43) with $\kappa = 0.5$, $\mu = 0.05$, $\sigma = 0.5$, $\tilde{y}_t = 1$, $y_0 = 0.03$, implementation using the principal branch, blue curve: $t = 2$, red dashed curve: $t = 20$.

function $g(u)$ are apparent, from its plot we can deduce that the integration of $g(u)$ will also be very cumbersome. Besides, we can also observe a discontinuity of the function g for $t = 2$ when the variable u equals that value in the interval $[40, 45]$ as shown in Figure 3.6.

In order to avoid the discontinuity of the function g we use the approach of Kahl and Jäckel [63] that was originally designed for the Heston model. To do so, we rewrite the characteristic function $\phi_{Y_t}(u)$ of the integrated CIR process defined in (3.37) as

$$\phi_{Y_t}(u) = 2^{\alpha} C(t, u)^{\alpha} \mathrm{e}^{B(t,u)y_0}, \tag{3.52}$$

where

$$\alpha := \frac{2\kappa\mu}{\sigma^2}, \tag{3.53}$$

$$c(u) := \frac{b(u)}{\kappa - b(u)}, \tag{3.54}$$

$$C(t,u) := \frac{c(u)\mathrm{e}^{\frac{(\kappa+b(u))t}{2}}}{a(u)\mathrm{e}^{b(u)t} - 1}. \tag{3.55}$$

Note that $A(t,u)$ in (3.38) can thus be written as $\alpha \ln 2 + \alpha \ln C(t,u)$, and $B(t,u)$ in (3.52) is already defined in (3.39).

From (3.52) we realize that we just shifted the problem from the complex logarithm to the evaluation of $C(t,u)^\alpha$, i.e. the evaluation of a complex logarithm is not necessary any more. Now it is easy to see that the function $C(t,u)^\alpha$ is the exact part of the function g as defined in (3.43) where the jump arises, because its argument $\arg(C)$ must have a discontinuity for any branch we selected. In other words, the branch switching of the complex logarithm is in fact not the main problem that gives rise to the jumps of the function g. For further details we refer the interested reader to [63].

In the literature different authors [70, 96, 100] proposed the straight forward idea to bookmark the number of jumps of $C(t,u)$ between two neighbouring quadrature points. However, in our case this may lead to a rather complicated routine since we prefer to use an adaptive quadrature scheme.

Rotation count correction

In the following, we describe a relative simple procedure, originally proposed by Kahl and Jäckel [63] for the Heston model, to guarantee the continuity of $C(t,u)$ by ensuring that the argument of $C(t,u)$ is continuous, such that the discontinuity of the integrand $g(u)$ in (3.42) is avoided. First, we introduce the polar and the rectangular representation for the complex valued coefficients $a(u)$ and $b(u)$ defined in (3.40):

$$a = r_a \mathrm{e}^{i\theta_a}, \tag{3.56}$$

$$b = p_b + iq_b. \tag{3.57}$$

Then the denominator of $C(t,u)$ in (3.55) can be written as

$$ae^{bt} - 1 = r_a e^{i\theta_a + p_b t + iq_b t} - 1 \tag{3.58}$$

$$= r^* e^{i(\chi^* + 2\pi m)}, \tag{3.59}$$

where

$$m := \text{int}\left[\frac{\theta_a + q_b t + \pi}{2\pi}\right], \tag{3.60}$$

$$\chi^* := \arg(ae^{bt} - 1), \tag{3.61}$$

$$r^* := |ae^{bt} - 1|. \tag{3.62}$$

Note that in (3.60) int$[\cdot]$ denotes the Gauss's integer brackets.

Restricting the argument $\theta_a \in [-\pi, \pi)$ means that we cut the complex plane along the negative real axis. When the function $C(t,u)$ in (3.55) crosses the negative real axis by varying u, the sign of the argument of $C(t,u)$ changes from $-\pi$ to π and therefore the argument of $C(t,u)^\alpha$ changes from $-\pi\alpha$ to $\pi\alpha$. Then the function jumps if α is not an integer, since

$$e^{i\pi} = e^{-i\pi} \Rightarrow \begin{cases} e^{i\alpha\pi} = e^{-i\alpha\pi} & \text{if } \alpha \in \mathbb{Z} \\ e^{i\alpha\pi} \neq e^{-i\alpha\pi} & \text{else.} \end{cases} \tag{3.63}$$

In general, we can assume χ^* and θ_a to be on the same argument interval $[-\pi, \pi)$, because the subtraction of 1 from ae^{bt} is simply a shift parallel to the real axis and cannot possibly move the complex number across the negative real axis as long as ae^{bt} never crosses the real axis in $[0,1]$. This essential property is guaranteed due to the following Proposition 3.3.3.

Proposition 3.3.3 *The absolute value of the function ae^{bt} is strictly greater than 1.*

The proof is given in the Appendix C.

Now we perform the same calculation with the numerator of $C(t,u)$. First we introduce the polar representation for $c(u)$ defined in (3.54):

$$c = r_c \mathrm{e}^{i\theta_c}. \tag{3.64}$$

Using the representation of $b(u)$ in (3.57) we have

$$c\mathrm{e}^{\frac{(\kappa+b)t}{2}} = r_c \mathrm{e}^{i\theta_c + \frac{t}{2}(k+p_b+iq_b)} \tag{3.65}$$

$$= r^{**}\mathrm{e}^{i(\chi^{**}+2\pi n)}, \tag{3.66}$$

with

$$n := \mathrm{int}\left[\frac{\theta_c + \frac{t}{2}q_b + \pi}{2\pi}\right], \tag{3.67}$$

$$\chi^{**} := \arg(c\mathrm{e}^{\frac{(\kappa+b)t}{2}}), \tag{3.68}$$

$$r^{**} := |c\mathrm{e}^{\frac{(\kappa+b)t}{2}}|. \tag{3.69}$$

This situation seems to be more intuitive; both χ^{**} and θ_c can be assumed to be in the same argument interval $[-\pi, \pi)$.

So far we obtained the following representation of $C(t,u)$ by combining the results above:

$$C(t,u) = \frac{c(u)\mathrm{e}^{\frac{(\kappa+b(u))t}{2}}}{a(u)\mathrm{e}^{b(u)t} - 1} = \frac{r^{**}}{r^*}\mathrm{e}^{i(\chi^{**}-\chi^*+2\pi(n-m))}, \tag{3.70}$$

and the rotation count correction

$$\ln C(t,u) = \ln\left(\frac{r^{**}}{r^*}\right) + i(\chi^{**} - \chi^* + 2\pi(n-m)). \tag{3.71}$$

By comparing the results with and without the rotation count correction (3.71) in Figure 3.5 we observe that the jump discontinuities can be removed.

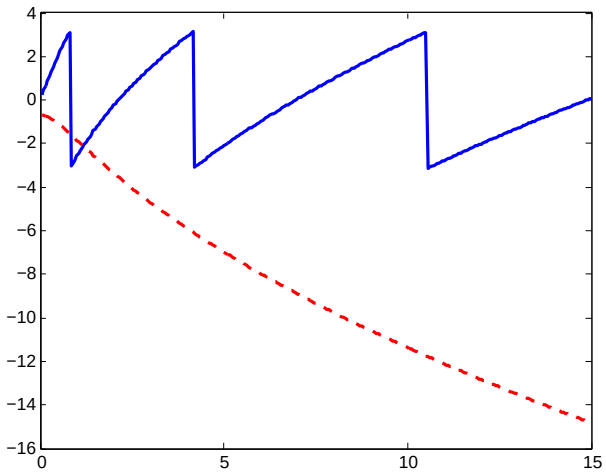


Figure 3.5: The blue solid curve is the argument of $C(t,u)$ as defined in (3.55) by using just the principal branch of $C(t,u)$ and the red dashed one is also the argument of $C(t,u)$ but with the rotation count correction (3.71) for $\kappa = 0.5$, $\mu = 0.05$, $\sigma = 0.5$, $y_0 = 0.03$, $t = 20$.

Now, we will consider the function $g(u)$ from Figure 3.4 and apply the rotation count correction. First we look at the $g(u)$ which is smoother (for $t = 2$), having only a discontinuity in the interval $[40, 45]$ as we have mentioned before, see Figure 3.6.

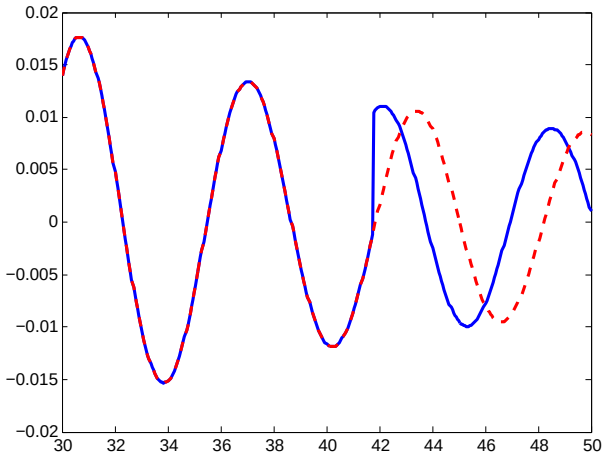


Figure 3.6: The blue solid curve is exactly a zoomed region of $g(u)$ (blue) shown in Figure 3.4 for $u \in [30, 50]$, the red dashed curve is obtained with the rotation count correction (3.71).

In Figure 3.7 we compare the functions $g(u)$ with and without the rotation count correction which are initialized with a high level of the CIR model parameter for a high

credit risk, in this case the function has stronger discontinuities. Let us note that high levels refer to the situations when the maturity of the CIR process is large, here $t = 20$. For the reason why the rotation count correction works we refer to [77].

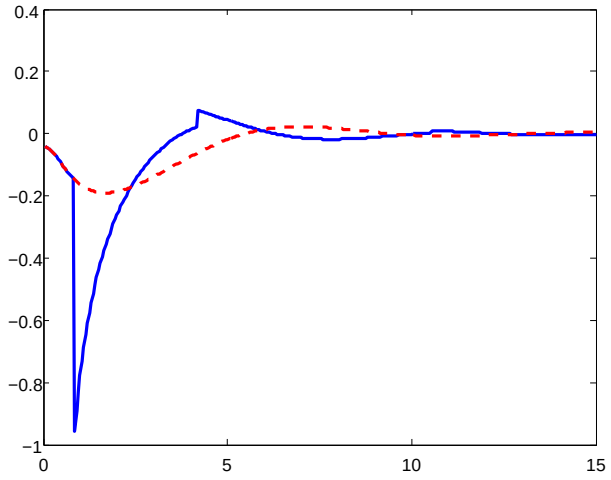


Figure 3.7: The blue solid curve is exactly the function $g(u)$ (red) shown in Figure 3.4 for $u \in [0, 15]$, the red dashed curve is obtained with the rotation count correction (3.71).

3.3.4 The Quadrature on the Finite Interval

We rewrite the CDF (3.42) using the transform (3.50) as

$$F_{Y_t}(\tilde{y}_t) = \int_0^1 \tilde{f}(x)\, dx, \tag{3.72}$$

where

$$\tilde{f}(x) := \frac{1}{2} + \frac{2 \ln x}{x \cdot \pi \cdot A_\infty^2} \cdot g\left(\frac{\ln x}{A_\infty}\right)^2. \tag{3.73}$$

This means that for the implementation using the adaptive Gauss-Lobatto quadrature we additionally need the limits of $\tilde{f}(x)$ at the boundaries 0 and 1 of the integral. For $x \to 0$ we observe that (3.48) and (3.51) imply

$$\lim_{x \to 0} \tilde{f}(x) = \frac{1}{2}. \tag{3.74}$$

In the following proposition we state the limit of the function $g(u)$, defined in (3.43), if u tends to zero.

Proposition 3.3.4 *For the function g defined by* (3.43) *we obtain that*

$$g(u) = \mathrm{Re}\left[\frac{e^{-iu\tilde{y}_t}\phi_{Y_t}(u)}{iu}\right] \tag{3.75}$$

has the following limit at zero:

$$\lim_{u\to 0} g(u) = -\tilde{y}_t + \mathrm{Im}(A(t,0)') + \mathrm{Im}(B(t,0)') \cdot y_0, \tag{3.76}$$

where

$$\mathrm{Im}(A(t,0)') = \frac{\mu\kappa e^{-\kappa t} + \mu\kappa(t\kappa - 1)}{\kappa^2}, \tag{3.77}$$

and

$$\mathrm{Im}(B(t,0)') = \frac{1 - e^{-\kappa t}}{\kappa}. \tag{3.78}$$

The proof can be found in Appendix C. The existence of the limit of the function $g(u)$ for u tending to zero implies directly

$$\lim_{x\to 1} \tilde{f}(x) = \frac{1}{2}. \tag{3.79}$$

Numerical results

Now we can implement the required Fourier inversion in (3.42) as a Gauss-Lobatto integration over the finite interval $[0, 1]$ instead of the infinite interval $[0, \infty)$ using the transformation (3.49)–(3.51). We see the first example in Figure 3.8. The stability of the Gauss-Lobatto integration over the finite interval $[0, 1]$ grants that even extreme probabilities can be also computed, like we choose a very long dated maturity $t = 30$ and see the corresponding CDF in Figure 3.8. It is worth mentioning that the parameters used in Figure 3.8 do not satisfy the Feller condition $2\kappa\mu > \sigma^2$. Therefore, we point out

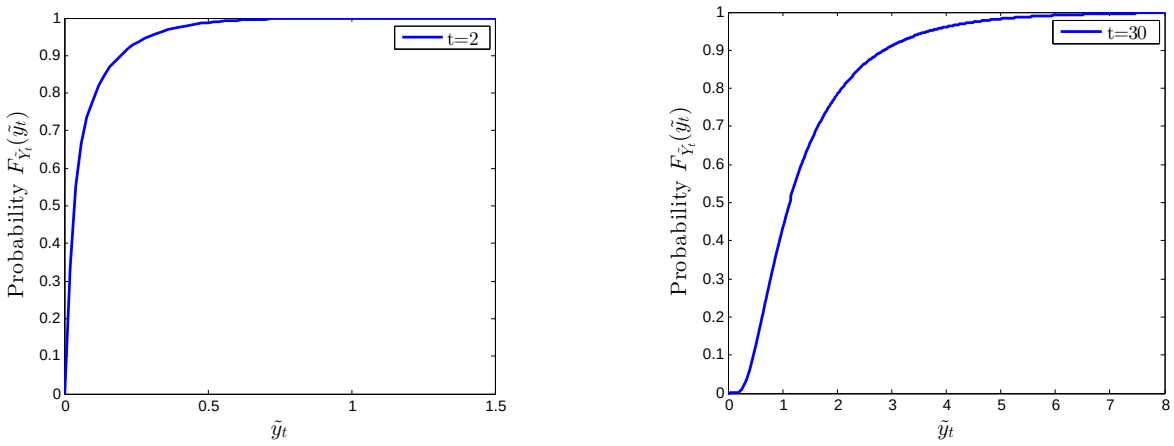
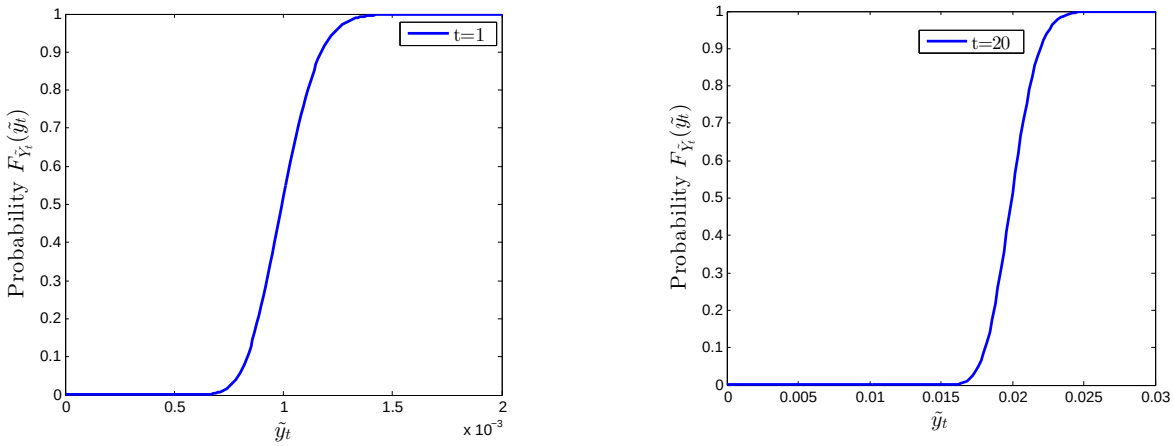


Figure 3.8: The CDF $F_{Y_t}(\tilde{y}_t)$ of the integrated CIR process Y_t with $\kappa = 0.5$, $\mu = 0.05$, $\sigma = 0.5$, $y_0 = 0.03$ computed with the adaptive Gauss-Lobatto scheme for a prescribed accuracy 10^{-12}.

that the accuracy of our numerical results is not limited by the Feller condition which guarantees the positivity of the value of the CIR process.

Besides, we show the CDFs of the integrated CIR processes which are computed for a lower value of parameters in Figure 3.9.



(a) (b)

Figure 3.9: The CDF $F_{Y_t}(\tilde{y}_t)$ of the integrated CIR process Y_t with $\kappa = 0.9$, $\mu = 0.001$, $\sigma = 0.01$, $y_0 = 0.001$ computed by the adaptive Gauss-Lobatto scheme for a prescribed accuracy 10^{-16}.

The method works surprisingly well for almost any model parameters. Only for some very extreme unrealistic parameters the numerical evaluation of the CDF cannot be real highly accurate, because the adaptive Gauss-Lobatto scheme may work for such extreme

parameters only with a low accuracy. In all practical relevant cases, the calibrated model parameters from real market data will not be that extreme.

3.4 Simultaneous Defaults risk (CDS)

As already briefly introduced in Section 2.3, we investigate to compute BVA by allowing simultaneous defaults among counterparties for a CDS contract. This extension is motivated with the indeed simultaneous default events[1] in real financial market: e.g., the collapses of Lehman Brothers and Merrill Lynch were just within two days (September 13-14, 2008); Another example noticed in [4], 24 railway firms defaulted simultaneously on the same day, June 21, 1970.

In reality, it is possible that the defaults among the counterparties do not occur simultaneously, but if one's default has triggered a jump in the default probability of the other one (e.g., if they are highly correlated), which might end up defaulting only within a short time period. For example, the protection seller's default could trigger a jump in the default probability of the reference credit so that the protection buyer has to suffer a loss. In the sequel we will refer to this later case also as a simultaneous default.

In particular, the effect of the simultaneous default between the CDS seller and the reference entity on the BVA is different to the corresponding effect, when the CDS buyer and the reference entity default simultaneously. If the protection seller and the reference entity default simultaneously, the buyer must suffer a loss, since the seller cannot honour the contractual obligations any more. However, when the protection buyer and the reference entity default simultaneously, the protection buyer can still be paid by the seller as the case, if only the reference entity defaults. This is to say, the simultaneous defaults between the CDS buyer and the reference entity have no substantial effects on the BVA, unlike the simultaneous defaults between the CDS seller and the reference

[1] Mathematically we define simultaneous defaults among the counterparties as that the default times of them are exactly the same. However, in the real world we can already say that they default simultaneously if they filed for bankruptcy protection on one day or within a few days.

entity. Thus, the possibility of the simultaneous defaults should be considered to price the counterparty risk.

3.4.1 Counterparty Risk of a CDS Contract with Simultaneous Defaults

We remind that the introduction to CDS has been provided in Section 3.1.1, where a general set-up including e.g., the definition of filtration can be found. Section 3.1.2 is devoted to arbitrage-free valuation framework for the bilateral counterparty default risk in a CDS contract, however, the simultaneous defaults among the counterparties are not considered.

We define the following events ordering the default times[2] of three names in the CDS contract between valuation t and maturity T [103]:

$$\begin{aligned}
&\mathfrak{A} = \{t < \tau = \tau_R \leq T\}, \quad \mathfrak{B} = \{t < \tau = \tau_C \leq T\}, \quad \mathfrak{C} = \{t < \tau = \tau_I \leq T\}, \\
&\mathfrak{D} = \{t < \tau = \tau_C = \tau_I \leq T\}, \quad \mathfrak{E} = \{t < \tau_{IC} = \tau_R \leq T\}.
\end{aligned} \tag{3.80}$$

Definition 3.4.1 *The discounted payoff of a counterparty-risky (simultaneous default included) CDS contract at time t can be written as:*

$$\begin{aligned}
\widehat{\Pi}(t,T) := {} & \mathbb{1}_{\mathfrak{A}} D(t,\tau)(-\mathcal{L}_R) \\
& + \mathbb{1}_{\mathfrak{B}} \left[D(t,\tau) \left(\mathcal{R}_C (P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}} \mathcal{L}_R)^+ - (P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}} \mathcal{L}_R)^- \right) \right] \\
& + \mathbb{1}_{\mathfrak{C}} \left[D(t,\tau) \left((P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}} \mathcal{L}_R)^+ - \mathcal{R}_I (P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}} \mathcal{L}_R)^- \right) \right] \\
& + \mathbb{1}_{\mathfrak{D}} \left[D(t,\tau) \left(-(P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}} \mathcal{L}_R) \right) \right] + \mathbb{1}_{\mathfrak{E}} \left[D(t,\tau) \mathcal{L}_R \right] \\
& + D(t,\tau)(\tau - T_{\gamma(\tau)-1}) \mathcal{P} \mathbb{1}_{\{T_a < \tau < T_b\}} + \sum_{i=a+1}^{b} D(t,T_i) \alpha_i \mathcal{P} \mathbb{1}_{\{\tau \geq T_i\}},
\end{aligned} \tag{3.81}$$

[2]See Section 3.1.2 for the definition of default times.

where the term $\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R$ *represents the exposure in case when the reference entity simultaneously defaults with any other counterparties.* $P_\tau^{CDS}(\mathcal{P}, \mathcal{L}_R)$ *is the price of a counterparty risk-free CDS contract at time* τ *defined in* (3.6).

- $\mathfrak{B}$: When the counterparty defaults, at the default time τ, the value of the CDS until maturity $P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R$ is computed. If this value is negative, the investor closes out the position by paying the defaulting counterparty this price. If the value is positive, the investor closes out the position and only receives a fraction $\mathcal{R}_C$ of this value from his counterparty. Therefore, in this case, we can define the close-out payment as

$$\mathcal{R}_C(P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - (P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-. \tag{3.82}$$

- $\mathfrak{C}$: In case of an investor default, if the value of CDS until maturity $P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R$ is positive, the counterparty closes out the position by paying this price in full. If this value is negative, the counterparty only receives a fraction $\mathcal{R}_I$ of this value to close out the position. Hence, the close-out payment is defined as

$$(P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - \mathcal{R}_I(P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-. \tag{3.83}$$

- $\mathfrak{D}$: If the investor and the counterparty default simultaneously, then compute the value of CDS like in case $\mathfrak{B}$ and $\mathfrak{C}$, that is $P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R$, and if it is negative, the counterparty receives a fraction $\mathcal{R}_I$ of this value; however, if it is positive, the investor receives a fraction $\mathcal{R}_C$ of this value. Together, we set the close-out payment for this case as

$$-(P_\tau^{\text{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R). \tag{3.84}$$

- $\mathfrak{E}$: If the investor or the counterparty default simultaneously with the reference entity, the investor receives a fraction $\mathcal{R}_C$ of the remaining recovery amount, $(-\mathcal{L}_R)^+$, when the counterparty defaults. Similarly, if the investor defaults, the counterparty

receives a portion $\mathcal{R}_I$ of the remaining recovery amount, $(-\mathcal{L}_R)^-$. The close-out payment for a joint default including the reference entity has the form

$$\mathcal{L}_R. \tag{3.85}$$

Definition 3.4.2 *We denote by $\widehat{P}_t^{CDS}(\mathcal{P}, \mathcal{L}_R)$ the price of a counterparty-risky CDS contract maturing at time T, i.e.*

$$\widehat{P}_t^{CDS}(\mathcal{P}, \mathcal{L}_R) = \mathbb{E}\{\widehat{\Pi}(t,T)|\mathcal{G}_t\}, \quad t \in [0,T]. \tag{3.86}$$

Proposition 3.4.1 *At the valuation time t, the BVA on a CDS contract maturing at time T defined as $\widehat{BVA}(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}}) = P_t^{CDS}(\mathcal{P},\mathcal{L}_R) - \widehat{P}_t^{CDS}(\mathcal{P},\mathcal{L}_R)$, can be written as*

$$\begin{aligned}\widehat{BVA}(t,T,\mathcal{P},\mathcal{L}_{\{C,I,R\}}) = \mathbb{E}\left\{\mathbb{1}_{\mathfrak{B}} \cdot \mathcal{L}_C \cdot D(t,\tau) \cdot \left(P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R\right)^+ |\mathcal{G}_t\right\} \\ - \mathbb{E}\left\{\mathbb{1}_{\mathfrak{C}} \cdot \mathcal{L}_I \cdot D(t,\tau) \cdot \left(P_\tau^{CDS} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R\right)^- |\mathcal{G}_t\right\},\end{aligned} \tag{3.87}$$

for every $t \in [0,T]$.

The proof can be found in the Appendix C.

Remark 3.4.1 *Similar to* (3.17)*, the BVA considering simultaneously defaultable counterparties also equals the sum of the values of a long position in a zero-strike call option on the residual price of CDS and the value of a short position in a zero-strike put option on the residual price of the CDS.*

Remark 3.4.2 *The formula* (3.87) *has the great advantage of being symmetric. This property means that the BVA from the point of view of the the counterparty is exactly opposite for the investor $(-BVA_t)$, this is to say that the parties agree on the value of the BVA. Besides, from (3.87) we can conclude that the value of this BVA can be negative and positive, the sign depends on which party is more risky to default.*

3.4.2 The Multivariate Markov Default Model

We propose an underlying stochastic model following [3, 9], based on which the BVA of a CDS contract considering simultaneously defaultable counterparties can be evaluated. We define a Markov Copula model of multivariate default times with factor processes $y = (y_C, y_I, y_R)$ and the corresponding default indicator processes $H = (H_C, H_I, H_R)$ for a CDS contract which have the following key features:

(i) The pair (y, H) is Markov in its natural filtration.

(ii) Each pair (y_j, H_j) is a Markov process.

(iii) At every instant, either each name on CDS contracts defaults individually or simultaneously with other names.

Remark 3.4.3 *The property (i) allows us to address in a dynamic and theoretically consistent way the issues of pricing and hedging credit derivatives. Property (ii) grants a quick valuation of single-name CDS contracts and independent calibration of each pair* (y_j, H_j)*, whereas (iii) will allow us to account for a dependence between defaults of each name.*

Instead of (3.18) we define henceforth the default intensities

$$\lambda_{j,t} = y_{j,t} + a_j, \quad t \geq 0,\ j = C, I, R, \tag{3.88}$$

where each a_j is a constant and each y_j is again a CIR process given in (3.19).

Remark 3.4.4 *We assumed that the processes* W_j *in* (3.19) *are independent of each other. Under this assumption the specification as defined in* (3.88) *has Markov consistency, i.e. the intensities of surviving names would not be affected by past defaults and the model dependence between defaults is only represented by the possibility of common jumps.*

3.4.3 The Model Specification and Simulation

We define a certain number of groups $M_l \subseteq \{C, I, R\} := M_{CDS}$, and set $M_l = l$ for $l \in \{\{C, I\}, \{C, R\}, \{I, R\}, \{C, I, R\}\} := \mathbb{L}$. However, λ_{M_l} can not solely be interpreted as the intensity of all parties in l defaulting simultaneously. For example, the reference credit R will also default with $M_l = \{C, R\}$ as long as he is still alive, if the investor I is already defaulted. Then for the default intensity we have the following:

- the counterparty C defaults with intensity $\lambda_{\{C\}} + \lambda_{\{C,I\}} + \lambda_{\{C,R\}} + \lambda_{\{C,I,R\}}$ as long as he is still alive,
- the investor I defaults with intensity $\lambda_{\{I\}} + \lambda_{\{C,I\}} + \lambda_{\{I,R\}} + \lambda_{\{C,I,R\}}$ as long as he is still alive,
- the reference credit R defaults with intensity $\lambda_{\{R\}} + \lambda_{\{C,R\}} + \lambda_{\{I,R\}} + \lambda_{\{C,I,R\}}$ as long as he is still alive,
- the counterparty C and the reference credit R default together with intensity $\lambda_{\{C,R\}} + \lambda_{\{C,I,R\}}$ as long as they are still alive,
- the counterparty C and the investor I default together with intensity $\lambda_{\{C,I\}} + \lambda_{\{C,I,R\}}$ as long as they are still alive,
- the investor I and the reference credit R default together with intensity $\lambda_{\{I,R\}} + \lambda_{\{C,I,R\}}$ as long as they are still alive,
- the counterparty C , the investor I and the reference credit R default together with intensity $\lambda_{\{C,I,R\}}$ as long as they are still alive.

Using this specification we first set the non-negative bounded intensity functions $\tilde{a}_j(t)$ as

$$\tilde{a}_{j,t} = \sum_{\{l \in \mathbb{L}; j \in l\}} \lambda_{l,t}, \tag{3.89}$$

but the calibration scheme will be computationally costly. It is thus useful to devise parsimonious model parameterizations. For instance, we use constant joint default intensities, setting $\lambda_l(t) = \lambda_l$ and thus for a_j in (3.88) we have

$$a_j = \sum_{\{l \in \mathbb{L}; j \in l\}} \lambda_l, \quad j \in M_{CDS}. \tag{3.90}$$

The default intensities for every $j \in M_{CDS}$ as defined in (3.88) can be written as

$$\lambda_{j,t} = y_{j,t} + \sum_{\{l \in \mathbb{L};\ j \in l\}} \lambda_l, \quad t \geq 0. \tag{3.91}$$

Analogous to the definition in (3.20) we define the following integrated quantities which will be used in the remainder of this section

$$\Lambda_j(t_1, t_2) := \int_{t_1}^{t_2} \lambda_{j,s} ds, \quad Y_j(t_1, t_2) := \int_{t_1}^{t_2} y_{j,s} ds, \quad \Lambda_l(t_1, t_2) := \int_{t_1}^{t_2} \lambda_{l,s} ds,$$

and

$$\Lambda_j(t) := \int_0^t \lambda_{j,s} ds, \quad Y_j(t) := \int_0^t y_{j,s} ds, \quad \Lambda_l(t) := \int_0^t \lambda_{l,s} ds,$$

where $j \in M_{CDS}$ and $l \in \mathbb{L}$.

It is obvious from (3.87) that we need the following conditional survival probabilities to compute the counterparty risk adjustment as defined in (3.6)

$$\mathbb{Q}(\tau_R > t | \mathcal{G}_{\tau_C}), \tag{3.92}$$

and

$$\mathbb{Q}(\tau_R > t | \mathcal{G}_{\tau_I}). \tag{3.93}$$

These two survival probabilities can be easily calculated by the following propositions.

Proposition 3.4.2

$$\begin{aligned}\mathbb{Q}(\tau_R > t|\mathcal{G}_{\tau_C}) &= \mathbb{E}\left\{\exp\left(-\Lambda_R(\tau_R,t)\right)|\mathcal{G}_{\tau_C}\right\} \\ &= \mathbb{E}\Big\{\exp(-Y_R(\tau_R,t) - \sum_{\{l\in\mathbb{L};\ R\in l\}}\Lambda_l(\tau_R,t))|\mathcal{G}_{\tau_C}\Big\}. \end{aligned} \tag{3.94}$$

Proposition 3.4.3

$$\begin{aligned}\mathbb{Q}(\tau_R > t|\mathcal{G}_{\tau_I}) &= \mathbb{E}\{\exp\left(-\Lambda_R(\tau_R,t)\right)|\mathcal{G}_{\tau_I}\} \\ &= \mathbb{E}\Big\{\exp(-Y_R(\tau_R,t) - \sum_{\{l\in\mathbb{L};\ R\in l\}}\Lambda_l(\tau_R,t))|\mathcal{G}_{\tau_I}\Big\}. \end{aligned} \tag{3.95}$$

The two propositions follow directly from the Markov probabilities, cf. [9].

Model simulation

As described in [9] the above model allows for a common shock model such that the simulation of a random time τ is computationally easy. However, the default time of each counterparty indicates not only its own default, but also the common defaulting among them. Therefore, we need to update the definition of default times. Given the previously simulated trajectories of the CIR processes y_j for $j \in M_{CDS}$, one essentially needs to simulate IID (Independent and identically) exponential random variables $\xi_{\hat{j}}$, for $\hat{j} \in \mathbb{L} \cup M_{CDS}$. Then one computes, for every $l \in \mathbb{L}$,

$$\hat{\tau}_l := \inf\{t > 0;\ \Lambda_l(t) \geq \xi_l\} \tag{3.96}$$

and for every $j \in M_{CDS}$,

$$\hat{\tau}_j := \inf\{t > 0;\ Y_j(t) \geq \xi_j\}. \tag{3.97}$$

Next, we set for every $j \in M_{CDS}$,

$$\tau_j = \hat{\tau}_j \wedge \Big(\bigwedge_{\{l\in\mathbb{L};\ j\in l\}} \hat{\tau}_l\Big). \tag{3.98}$$

3.4.4 Computing of BVA considering Simultaneous Defaults

Before computing BVA considering simultaneous defaults we firstly have to determine the constant a_j for $j = C, I, R$ in the model (3.90). If we assume that the processes y_j, $j \in M_{CDS}$ are always non-negative, then due to the definition (3.88) the constant a_j defined in (3.90) must be chosen as

$$a_j = \sum_{\{l \in \mathbb{L}; j \in l\}} \lambda_l \leq \lambda_j, \quad \forall j \in M_{CDS}. \tag{3.99}$$

For instance, we can set as in [3], for every $l \in \mathbb{L}$,

$$\lambda_l = \mathcal{C}_l \inf_{j \in l} \lambda_j \tag{3.100}$$

for some non-negative model dependence parameters $\mathcal{C}_l$ such that $\sum_{l \in \mathbb{L}} \mathcal{C}_l \leq 1$. The value of $\mathcal{C}_l$ determines the possibility of simultaneous defaults between the parties in the group l; a larger value refers to a higher possibility of simultaneous defaults. Conversely, if we e.g., set $\mathcal{C}_{\{C,R\}} = 0$, then the simultaneous defaults between the counterparty and the reference credit is not possible. In other words, the parameters $\mathcal{C}_l$ represent the dependence between defaults of the parties in the group l.

It has been shown in Section 3.2.1 that the market implied intensity (hazard rate) for name j with λ_j^* can be bootstrapped from the individual CDS quotes. We remark that the bootstrapping procedure is model independent. Now we can calibrate the constant a_j for every $j \in M_{CDS}$ by choosing appropriate model dependence parameters $\mathcal{C}_l$ and setting

$$a_j = \sum_{\{l \in \mathbb{L}; j \in l\}} \lambda_l \leq \lambda_j^*. \tag{3.101}$$

From the CDS quotes of the higher risk the bootstrapped intensity λ_j^* is larger; thus λ_l, $j \in l$ is larger due to (3.100). Besides, for the same reason as for the intensity λ_j^*, a larger dependence parameter $\mathcal{C}_l$ constructs the larger λ_l, $j \in l$. Hence, with the same

exponentially distributed trigger variable the simultaneous default time of the group l is smaller (earlier) through (3.96), if $\mathcal{C}_l$ is larger and consequently the possibility of the simultaneous defaults between parties in the group l is higher.

The Monte-Carlo approach

Now we can compute the BVA on a CDS contract. We perform the following steps based on Monte-Carlo simulations:

1. Produce default times τ_C, τ_I and τ_R using (3.96), (3.97) and (3.98).

2. In case of $\mathfrak{B}$ (see (3.80)), i.e. the counterparty defaults first, we need to compute the term inside the first expectation value which has positive sign. First, we check at the default time of counterparty whether the reference credit also defaults. For the case of a simultaneous default we just need the loss given default $\mathcal{L}_R$. Otherwise we compute $P^{\text{CDS}}_{\tau_C}$ given in (3.6), the survival possibility in $P^{\text{CDS}}_{\tau_C}$ can be computed by (3.94).

3. In the event of $\mathfrak{C}$ (see (3.80)) we need the term inside the second expectation value with negative sign, the computation is similar to the last step.

4. Finally, we produce the BVA by discounting and averaging.

3.4.5 An example with Wrong-Way risk

In this section we perform a numerical evaluation of the BVA as defined in (3.87) based on Monte-Carlo simulations. We study again a five years CDS contract on a reference entity traded by an investor and a counterparty, where both the investor and the counterparty are defaultable. We assume the payment dates to be every three months $\alpha = 0.25$ and the LGD of the three names are taken from a market provider and are fixed to 60%. Furthermore, we assume deterministic interest rates such that the default time τ and the discount factor are independent.

We set the three names having different levels of credit risk which are specified by the collections of the parameters in Table 3.5. As already introduced in Section 3.2.3,

Credit Risk Level	κ	μ	σ	y_0
Low	0.9	0.001	0.01	0.001
Medium	0.8	0.02	0.1	0.01
High	0.5	0.05	0.3	0.04

Table 3.5: Collection of parameters for initializing the CIR processes.

using the parameters in Table 3.5 we can compute break-even spreads which will be used as the market quotes. We show in Table 3.6 the premium rate $\mathcal{P}$ in basis points using the assumed deterministic LGD ($LGD_{\text{low}} = 0.6$, $LGD_{\text{medium}} = 0.65$ and $LGD_{\text{high}} = 0.7$) and collections of the parameters in Table 3.5.

Maturity	Low Risk	Medium Risk	High Risk
1y	6	85	293
2y	6	97	298
3y	6	105	301
4y	6	110	302
5y	6	113	302
6y	6	115	303

Table 3.6: Break-even spreads in basis points generated using the collections of the parameters of the CIR processes in Table 3.5.

We assume the following scenarios:

- Scenario 1. The investor has low credit risk, the reference entity has high credit risk and the counterparty has medium credit risk. This situation is the most common in the real market.

- Scenario 2. The investor has low credit risk, the reference entity has medium credit risk and the counterparty has high credit risk. We are facing a risky counterparty in this case.

- Scenario 3. The investor has high credit risk, the reference entity has medium credit risk and the counterparty has low credit risk. The investor is most risky itself.

- Scenario 4. Both investor and counterparty have medium credit risk, while the reference entity has high credit risk (Risky Reference I).

- Scenario 5. Both investor and counterparty have low credit risk, while the reference entity has high credit risk (Risky Reference II).

As an example, we assume the parameters $\mathcal{C}_l, l \neq \{C, R\}$ to be the same and equal with 0.01. and compute the BVAs for each following scenario by varying the parameter $\mathcal{C}_{\{C,R\}}$. We report our results in Table 3.7. Table 3.7 clearly shows the effect of the

Scenario	Base Scenario	Risky Counterparty	Risky Investor	Risky Ref I	Risky Ref II
$\mathcal{C}_{\{C,R\}}$	BVA_p BVA_s	BVA_p BVA_s	BVA_p BVA_s	BVA_p BVA_s	BVA_p BVA_s
0.01	7.0 (0.1) 4.1 (0.0)	6.5 (0.1) 0.7 (0.0)	-0.6 (0.0) -6.3 (0.1)	2.9 (0.1) -2.9 (0.1)	0.2 (0.0) -0.1 (0.0)
0.03	12.5 (0.2) 4.0 (0.0)	12.5 (0.2) 0.5 (0.0)	-0.4 (0.0) -6.3 (0.1)	8.3 (0.1) -3.1 (0.1)	0.4 (0.0) -0.1 (0.0)
0.05	18.4 (0.2) 3.9 (0.0)	19.1 (0.2) 0.3 (0.0)	-0.0 (0.1) -6.3 (0.1)	14.1 (0.2) -3.1 (0.1)	0.8 (0.1) -0.1 (0.0)
0.1	32.6 (0.3) 3.6 (0.0)	35.1 (0.3) 0.0 (0.0)	0.8 (0.1) -6.5 (0.1)	28.0 (0.2) -3.6 (0.1)	1.5 (0.1) -0.1 (0.0)
0.15	46.0 (0.4) 3.2 (0.0)	50.8 (0.4) -0.1 (0.0)	1.6 (0.1) -6.5 (0.1)	41.5 (0.4) -4.2 (0.1)	2.3 (0.1) -0.1 (0.0)
0.2	59.7 (0.4) 2.9 (0.0)	66.4 (0.4) -0.3 (0.0)	2.2 (0.1) -6.4 (0.1)	54.8 (0.4) -4.6 (0.1)	2.8 (0.1) -0.1 (0.0)
0.25	74.4 (0.5) 2.6 (0.0)	83.1 (0.5) -0.3 (0.0)	3.3 (0.1) -6.5 (0.1)	69.3 (0.5) -5.1 (0.1)	4.0 (0.1) -0.1 (0.0)

Table 3.7: The values of the BVA in basis points for the different scenarios, the number between parentheses represents the Monte-Carlo standard error.

wrong-way risk. For example, if one looks at the second column, one notices that as the possibility of the simultaneous defaults between counterparty and reference credit gets larger, the BVA_p increases significantly due to the reasons: (1) the counterparty is the riskiest name. (2) The higher represented positive correlation makes the spread of the reference entity larger at the counterparty default, thus the option on the residual price

of CDS for the investor as payer will be in the money and worth more. However, at the counterparty default the investor only gets a fraction of it proportional to the recovery value of the counterparty. (3) At the simultaneous default of the counterparty and the reference credit, that option must be deep into the money, but the payer investor only gets a fraction of it proportional to the recovery value of the counterparty, more BVA_p takes place.

The adjustments BVA_p in Scenario 1 (Base Scenario) and Scenario 2 (Risky Counterparty) are similar, since the possibility of simultaneous defaults between the counterparty and the reference entity are the same if one has medium credit risk and the other one has high risk. The adjustments BVA_p in Scenario 2 are a little bit larger than the corresponding adjustments in Scenario 1, because the counterparty in Scenario 2 is riskier. If one looks at the adjustments BVA_s in Scenario 1 and Scenario 2, at the simultaneous default between the counterparty and the reference entity, the option for the investor as the receiver will be out of the money, thus slight adjustments are required. In particular, as the dependence parameter $\mathcal{C}_{\{C,R\}}$ is larger, less adjustments take place, but the changes are very small.

In Scenario 3, the values of the adjustments BVA_p have only small changes. For the small dependence parameter $\mathcal{C}_{\{C,R\}}$ the adjustment is negative, because the investor is riskier. However, as the dependence parameter $\mathcal{C}_{\{C,R\}}$ gets larger, this is to say that the possibility of the simultaneous defaults between the counterparty and the reference is increasing, the investor as payer even requires the adjustments although he is risky, see the last three rows at the third column.

An interesting pattern emerges from the fourth column (Risky Ref I). Contrary to earlier works, e.g., [15, 16], by looking at BVA_p at the fourth column, one finds that as the possibility of the simultaneous defaults between counterparty and the reference entity is increasing, the BVA_p increases significantly. The reason is that the counterparty has medium credit risk while the reference entity has high risk, thus they have higher

possibility for larger $\mathcal{C}_{\{C,R\}}$ to default simultaneously, then the investor needs adjustments to hedge this risk. However, if the counterparty becomes safer while the reference entity is still riskier, then the possibility of the simultaneous defaults between the counterparty and the reference entity will be lower, thus less adjustments will take place as reported in the last column (Risky Ref II).

3.5 Summary

Either from the correlation matrix $\mathcal{R}$ in Section 3.2 or the dependence factor $\mathcal{C}_l$ in Section 3.4 we realize that the correlation between counterparties in a CDS contract plays a key role on their default risk management. As already mentioned in Section 2.3, to introduce the default correlation among the counterparties one can use a copula function, the correlated BMs in the SDE system, or both approaches jointly.

Generally, the degree of relationship between financial products and financial instituations always plays an essential role on, e.g., pricing and hedging, and thus must be considered. The most common one is the correlated BMs in a SDE system by a deterministic constant. However, intuitively, a time-dependent model or a stochastic model could better replicate the phenomena in the real world. Indeed, market observations clearly indicate that financial quantities are correlated in a strongly nonlinear way, correlation could even behave stochastically and unpredictably. This motivates us to finish the next two parts for the present thesis: modelling and application of local time-dependent and stochastic correlation.

Part II

Modelling and Applications of Local Time-dependent Correlation

In this part, we provide an appropriate and reasonable time-dependent correlation function and present the concept of dynamically (time-dependent) correlated Brownian motions (BMs) and its construction. As example, we apply this new time-dependent correlation function to price European options and Quanto options. We analyze the improvement by using a time-dependent correlation instead of a constant correlation.

Chapter 4

Time-dependent Correlation Model and its Application

We introduce an appropriate and reasonable time-dependent correlation function. The concept of dynamically (time-dependent) correlated BMs and its construction are presented. For the applicability of time-dependent correlation, for examples, we price Quanto options under time-dependent correlation and extend the Heston model [58] by incorparating time-dependent correlation.

4.1 The Dynamic Correlation Function

The key issue of modelling correlation as a time-dependent function is to ensure that the boundaries -1 and 1 of the correlation function are not attractive and unattainable for any time. Besides, the correlation function must converge for increasing time. Actually, it is demanding to find such a correlation function which satisfies these two properties. In this section, we build up a reasonable and appropriate time-dependent correlation function, so that one can reasonably choose additional parameters to increase the fitting quality on the one hand but also add an economic concept on the other hand. Thus many

problems of finance and economics can be treated under dynamic correlation which is much more realistic than using a constant correlation to model real world phenomena.

4.1.1 Build-up Model

A correlation function must satisfy the correlation properties: It provides only the values in the interval $(-1, 1)$ for any time; it converges to a value for increasing time. We find the following simple idea: We denote the dynamic correlation with $\bar{\rho}$ and propose simply using

$$\bar{\rho}_t := E\left[\tanh(X_t)\right], \quad t > 0, \tag{4.1}$$

for the *dynamic correlation function*, where X_t is any mean-reverting process with positive and negative values. For a fixed parameter of X_t, the correlation function $\bar{\rho}_t : [0, t] \to (-1, 1)$ depends only on t. We observe that the dynamic correlation model (4.1) satisfies the desired properties: First, it is obvious that $\bar{\rho}_t$ takes values only in $(-1, 1)$ for all t. Besides, it converges to a value for increasing time due to the mean reversion of the used process X_t.

The time-dependent correlation function based on OU process

X_t in (4.1) could be any mean-reverting process which alows for positive and negative outcomes. As an example, let X_t be the *Ornstein-Uhlenbeck (OU) process* [113]

$$dX_t = \kappa(\mu - X_t)dt + \sigma dW_t, \quad t \geq 0. \tag{4.2}$$

We are interested in computing $E[\rho_t]$ as a function of given parameters in (4.2). We compute $\bar{\rho}_t = E[\tanh(X_t)]$ as

$$\bar{\rho}_t = E[\tanh(X_t)] = E\left[1 - \mathrm{e}^{-X_t} \cdot \frac{2}{\mathrm{e}^{-X_t} + \mathrm{e}^{X_t}}\right] = 1 - E\left[\mathrm{e}^{-X_t} \cdot \frac{1}{\cosh(X_t)}\right]. \tag{4.3}$$

We set $g(X_t) = 1/\cosh(X_t)$, and apply the results by Chen and Joslin [29], the expectation in (4.3) can be found in closed-form expression (up to an integral) as

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{g}(u)\cdot E[\mathrm{e}^{-X_t}\mathrm{e}^{iuX_t}]\,du, \tag{4.4}$$

where $i=\sqrt{-1}$ denotes the imaginary unit and $\hat{g}$ is the Fourier transform of g, in this case is known analytically by $\hat{g}(u) = \pi/\cosh(\frac{\pi u}{2})$. Denoting $CF(t,u|X_0,\kappa,\mu,\sigma)$ as the characteristic function of X_t, the expectation in (4.4) can be presented by $CF(t,i+u|X_0,\kappa,\mu,\sigma)$. Thus, we obtain the closed-form expression for $\bar{\rho}_t$:

$$\bar{\rho}_t = 1-\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\cosh(\frac{\pi u}{2})}\cdot CF(t,i+u|X_0,\kappa,\mu,\sigma)du. \tag{4.5}$$

The next step is to calculate the expression of $CF(t,i+u|X_0,\kappa,\mu,\sigma)$. X_t is the OU process and its characteristic function $CF(t,u|X_0,\kappa,\mu,\sigma)$ can be obtained analytically, e.g., using the framework of the affine process, see [36]. Then, we only need to substitute $u+i$ for u in the characteristic function of X_t to calculate $CF(t,i+u|X_0,\kappa,\mu,\sigma)$ which is given by

$$CF(t,i+u|X_0,\kappa,\mu,\sigma) = \mathrm{e}^{-A(t)-\frac{B(t)}{2}+iu(A(t)+B(t))+u^2\frac{B(t)}{2}}, \tag{4.6}$$

with

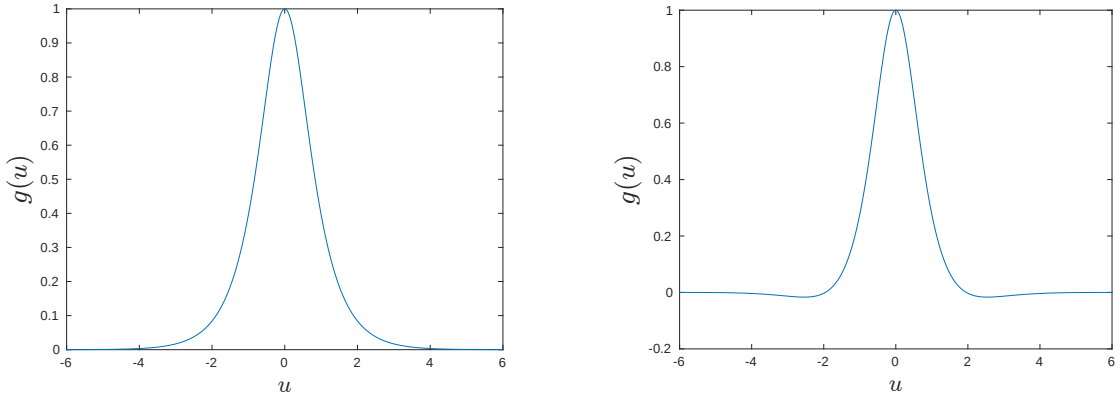
$$A(t) = \mathrm{e}^{-\kappa t}X_0+\mu(1-\mathrm{e}^{-\kappa t}),\quad B(t) = -\frac{\sigma^2}{2\kappa}(1-\mathrm{e}^{-2\kappa t}). \tag{4.7}$$

Finally, the dynamic correlation function $\bar{\rho}_t$ can be computed by

$$\bar{\rho}_t = 1-\frac{\mathrm{e}^{-A(t)-\frac{B(t)}{2}}}{2}\int_{-\infty}^{\infty}\underbrace{\frac{1}{\cosh(\frac{\pi u}{2})}\cdot\mathrm{e}^{iu(A(t)+B(t))+u^2\frac{B(t)}{2}}}_{:=g(u)}du, \tag{4.8}$$

where $A(t)$ and $B(t)$ are defined in (4.7). In fact, X_0 in $A(t)$ is equal to $\operatorname{artanh}(\bar{\rho}_0)$.

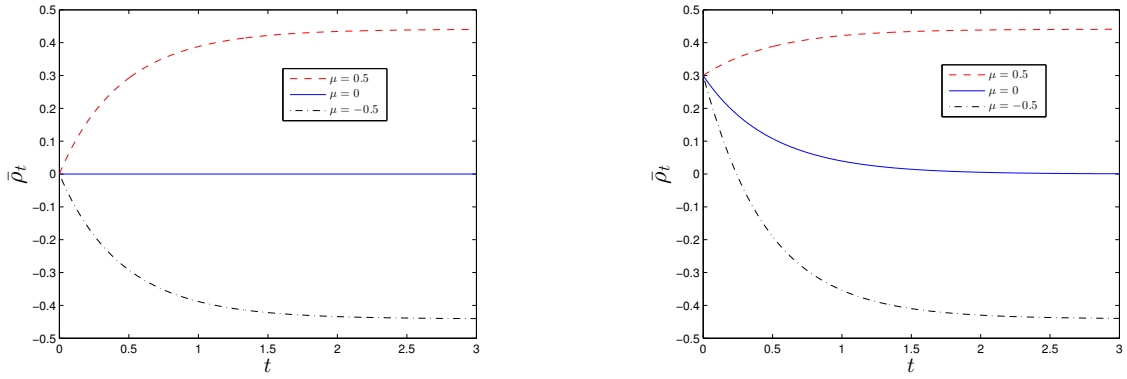
We observe that the integrand $g(u)$ is a symmetric function about $u = 0$ and vanishes (approaches zero) for a sufficiently large absolute value of u, see Figure 4.1. For these two reasons, the numerical integration in (4.8) is computationally fast.



(a) $t = 0.1$ (b) $t = 10$

Figure 4.1: $g(u)$ under $\rho_0 = 0.3, \kappa_\rho = 2, \mu_\rho = -0.8, \sigma_\rho = 0.1$.

To illustrate the role of each parameter in (4.8), we plot $\bar{\rho}_t$ for a couple of parameters. First in Figure 4.2, we let $\kappa = 2$ and $\sigma = 0.5$ and display $\bar{\rho}_t$ with different values of μ, which is set to be 0.5, 0 and -0.5, respectively. Obviously, μ determines the long



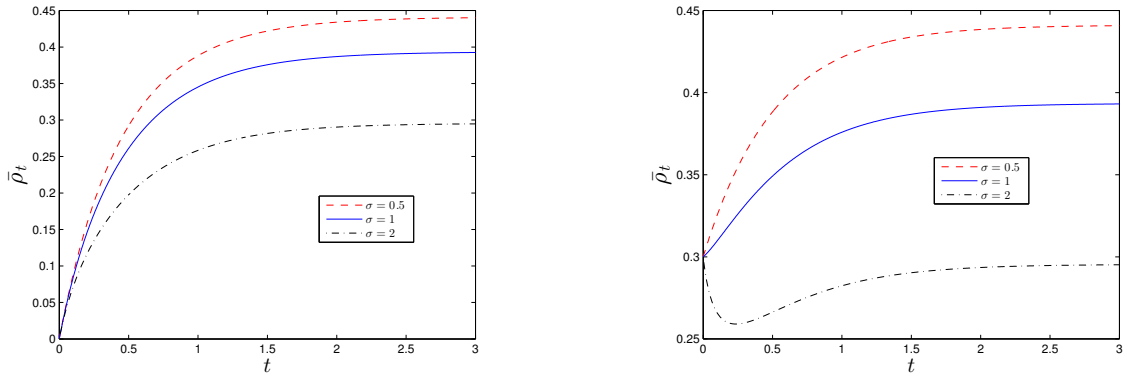
(a) $\bar{\rho}_0 = 0$ (b) $\bar{\rho}_0 = 0.3$

Figure 4.2: Dynamic correlation $\bar{\rho}_t$ for varying μ ($\kappa = 2$ and $\sigma = 0.5$).

term mean of $\bar{\rho}_t$. However, μ is not the exact limiting value. Considering Figure 4.2a where the initial value of the correlation function is 0, we see that $\bar{\rho}_t$ is increasing to a value around $\mu = 0.5$ and decreasing to a value around $\mu = -0.5$ as t goes on, when

$\mu = 0.5$ and -0.5, respectively. Besides, for $\mu = \bar{\rho}_0 = 0$ we observe that the correlation function $\bar{\rho}_t$ always yields 0 which is the same as the constant correlation $\rho = 0$. Now, we set $\bar{\rho}_0 = 0.3$ and keep the values of all other parameters to be unchanged, then display the curves of $\bar{\rho}_t$ in Figure 4.2b.

Next, we fix $\kappa = 2$ and $\mu = 0.5$ and show $\bar{\rho}_t$ for the varying $\sigma = 0.5,\ 1$ and 2 in Figure 4.3. Obviously, σ shows the magnitude of variation from the value around $\mu = 0.5$. In

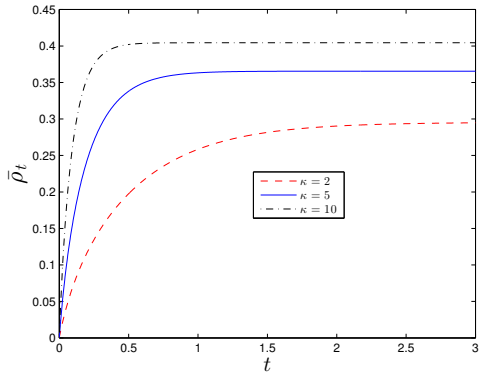


(a) $\bar{\rho}_0 = 0$ (b) $\bar{\rho}_0 = 0.3$

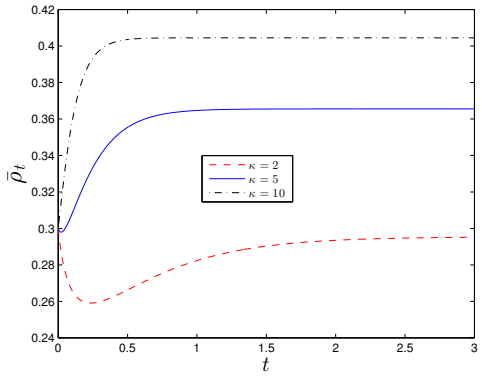
Figure 4.3: Dynamic correlation $\bar{\rho}_t$ for varying σ ($\kappa = 2$ and $\mu = 0.5$).

Figure 4.3a we see, the larger the value of σ is, the stronger the deviations of $\bar{\rho}_t$ is from the value around $\mu = 0.5$. More interesting is that $\bar{\rho}_t$ first decreases until $t \approx 0.25$, then increases and converges to a value, see Figure 4.3b where $\bar{\rho}_0 = 0.3$ and $\sigma = 2$.

Again, in order to illustrate the role of κ, we set $\mu = 0.5,\ \sigma = 2$ and vary the value of κ, see Figure 4.4. From Figure 4.4a it is easy to observe that κ represents the speed of $\bar{\rho}_t$ tending to its limit. Especially, as we have seen in Figure 4.3b, the curve is more unstable for $\kappa = 2$ and $\sigma = 2$ in Figure 4.4b. However, if σ remains constant while the value of κ is increasing, we can see that curves of $\bar{\rho}_t$ become more stable and tend straightly to its limit. If one incorporates the dynamic correlation function (4.8) to a financial model, the parameters $\bar{\rho}_0,\ \kappa,\ \mu,$ and σ could be estimated by fitting the model to market data.



(a) $\bar{\rho}_0 = 0$



(b) $\bar{\rho}_0 = 0.3$

Figure 4.4: Dynamic correlation $\bar{\rho}_t$ for varying κ ($\mu = 0.5$ and $\sigma = 2$).

4.1.2 Dynamically correlated BMs and its Construction

At a time $t > 0$, the correlation coefficient of two Brownian motions (BMs) $W_{t,1}$ and $W_{t,2}$ is defined as (see B.1)

$$\rho_t^{1,2} = \frac{E\left[W_{t,1}W_{t,2}\right]}{t}. \tag{4.9}$$

If we assume that $\rho_t^{1,2}$ is constant, $\rho_t^{1,2} = \rho^{1,2}$ for all $t > 0$, then $W_{t,1}$ and $W_{t,2}$ are correlated with the constant $\rho^{1,2}$. In the following, we show how to define the dynamically correlated Brownian motions, where the correlation is not the same for each time instant. Let $(\Delta_n)_{n\in\mathbb{N}} := \{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$ be a partition of $[0, t]$ with the mesh $\|(\Delta_n)\| := \max_{1\le i\le n}(t_i - t_{i-1})$, we calculate

$$\begin{aligned}
E\left[W_{t,1}W_{t,2}\right] &= E\left[\left((W_{t_n,1} - W_{t_{n-1},1}) + (W_{t_{n-1},1} - W_{t_{n-2},1}) + \cdots + (W_{t_1,1} - W_{t_0,1})\right)\right. \\
&\quad \left.\cdot\left((W_{t_n,2} - W_{t_{n-1},2}) + (W_{t_{n-1},2} - W_{t_{n-2},2}) + \cdots + (W_{t_1,2} - W_{t_0,2})\right)\right] \\
&= E\left[\sum_{i=1}^{n}(W_{t_i,1} - W_{t_{i-1},1})(W_{t_i,2} - W_{t_{i-1},2})\right] \\
&= E\left[\sum_{i=1}^{n}\rho_{t_i-t_{i-1}}^{1,2}(t_i - t_{i-1})\right] \overset{\substack{\|\Delta_n\|\to 0\\ n\to\infty}}{=} E\left[\int_0^t \rho_s^{1,2}ds\right].
\end{aligned} \tag{4.10}$$

Therefore, we give the definition of dynamically correlated BMs.

Definition 4.1.1 *Two Brownian motions $W_{t,1}$ and $W_{t,2}$ are called dynamically correlated with correlation function $\bar{\rho}_t$, if they satisfy*

$$E\left[W_{t,1}W_{t,2}\right] = \int_0^t \bar{\rho}_s ds, \tag{4.11}$$

where $\bar{\rho}_t : \mathbb{R}^+ \to (-1, 1)$. The average correlation of $W_{t,1}$ and $W_{t,2}$, ρ_{Av}, is given by $\rho_{Av} := \frac{1}{t}\int_0^t \bar{\rho}_s ds$.

Construction

The construction of dynamically correlated BMs can be easily done. We consider first the two-dimensional case. Let $\bar{\rho}_t$ be a correlation function. For two independent BMs $W_{t,1}$ and $W_{t,3}$ we define

$$W_{t,2} = \int_0^t \bar{\rho}_s dW_{s,1} + \int_0^t \sqrt{1-\bar{\rho}_s^2}\, dW_{s,3}, \tag{4.12}$$

with the symbolic expression

$$dW_{t,2} = \bar{\rho}_t dW_{t,1} + \sqrt{1-\bar{\rho}_t^2}\, dW_{t,3}. \tag{4.13}$$

It can be easily verified that $W_{t,2}$ is a BM and correlated with $W_{t,1}$ dynamically by ρ_t. Besides, the covariance matrix and the average correlation matrix of $\mathbb{W}_t = (W_{t,1}, W_{t,2})$ can be determined, given by

$$\begin{pmatrix} t & \int_0^t \bar{\rho}_s ds \\ \int_0^t \bar{\rho}_s ds & t \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & \frac{1}{t}\int_0^t \bar{\rho}_s ds \\ \frac{1}{t}\int_0^t \bar{\rho}_s ds & 1 \end{pmatrix}$$

respectively.

The construction above can be also generalized to n-dimensions. We denote a standard n-dimensional Brownian motion by $\mathbb{Z}_t = (Z_{1,t}, ..., Z_{n,t})$ and the matrix of dynamic correlations $\mathcal{R}_t = (\bar{\rho}_t^{i,j})_{1<i,j<n}$ which has the Cholesky decomposition for each

time t, $\mathcal{R}_t = \mathbb{A}_t\mathbb{A}_t^\top$ with $\mathbb{A}_t = (a_t^{i,j})_{1<i,j<n}$. We define a new n-dimensional process $\mathbb{W}_t = (W_{1,t}, ..., W_{n,t})$ by

$$W_{i,t} = \sum_{j=1}^{n} a_t^{ij}\, dZ_{j,t}, \quad i = 1, \cdots, n. \tag{4.14}$$

We can easily verify that $\mathbb{W}_t$ satisfies the following properties:

- $\mathbb{W}_0 = \mathbf{0}$ and the paths are continuous with probability 1.
- The increments $\mathbb{W}_{t_1} - \mathbb{W}_{t_0}$ and $\mathbb{W}_{t_2} - \mathbb{W}_{t_1}$ are independent for $0 \le t_0 < t_1 < t_2 < t$.
- For $0 \le s < t$, the increment $\mathbb{W}_t - \mathbb{W}_s$ is multivariate normally distributed with mean zero and covariance matrix $\Sigma : \mathbb{W}_t - \mathbb{W}_s \sim N(0, \Sigma)$ with

$$\Sigma = \begin{pmatrix} t-s & \int_s^t \bar{\rho}_u^{1,2} du & \cdots & \int_s^t \bar{\rho}_u^{1,n} du \\ \int_s^t \bar{\rho}_u^{2,1} du & t-s & \cdots & \int_s^t \bar{\rho}_u^{2,n} du \\ \vdots & \vdots & \ddots & \vdots \\ \int_s^t \bar{\rho}_u^{n,1} du & \int_s^t \bar{\rho}_u^{n,2} du & \cdots & t-s \end{pmatrix}.$$

We call the process $(\mathbb{W}_t)_{t\ge 0}$ *a n-dimensional dynamically correlated Brownian motion,* with the correlation matrix $\mathcal{R}_t$.

4.2 The Applications to Quanto Options

The Quanto option is a cash-settled, cross-currency derivative in which the underlying asset has a payoff in one currency, but the payoff is converted to another currency in which the option is settled. Thus, the correlation between assets and currency exchange rate must be considered. Instead of assuming a constant correlation, we develop a strategy for pricing the Quanto option under time-dependent correlations in a closed formula. An example of calibration to real market data is provided. We study the effect

of dynamic correlation on the option pricing and hedging. The numerical results show that the prices of Quanto option under dynamic correlation can be better fitted to the market prices than simply using a constant correlation.

4.2.1 Quanto Options under Dynamic Correlation

We derive the pricing formula of Quanto options with incorporated dynamic correlation. We define H as the exchange rate between domestic and foreign currency and S is the level of an index traded in the foreign countries. We assume that they satisfy

$$\begin{cases} dS_t &= \mu_S S_t\, dt + \sigma_S S_t\, dW_t^S \\ dH_t &= \mu_H H_t\, dt + \sigma_H H_t\, dW_t^H, \end{cases} \tag{4.15}$$

where W_t^S and W_t^H are correlated dynamically with the correlation function $\bar{\rho}_t$ defined in (4.8).

Following the methodologies in [116, 108] we construct a portfolio consisting of the quanto in question, hedged with foreign currency and the asset S :

$$\Pi = V(H, S, \bar{\rho}_t, t) - \Delta_H H - \Delta_S HS. \tag{4.16}$$

We remark that every term in this equation values in domestic currency. Δ_H is the number of foreign currency we hold short, so $-\Delta_H H$ is the value in domestic currency of that foreign currency. It is similar to understand $-\Delta_S HS$.

The change in the value of the portfolio due to the change in the value of its components and the interest rate of foreign currency (r_f) can be obtained with the aid of the

Itô lemma as

$$\begin{aligned} d\Pi = \Big(&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \bar{\rho}_t \sigma_H \sigma_S HS \frac{\partial^2 V}{\partial S \partial H} + \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 V}{\partial S^2} \\ &- \bar{\rho}_t \sigma_H \sigma_S \Delta_S HS - r_f \Delta_H H\Big) dt + \Big(\frac{\partial V}{\partial H} - \Delta_H - \Delta_S S\Big) dH \\ &+ \Big(\frac{\partial V}{\partial S} - \Delta_S H\Big) dS. \end{aligned} \tag{4.17}$$

We now choose

$$\Delta_H = \frac{\partial V}{\partial H} - \frac{S}{H}\frac{\partial V}{\partial S} \text{ and } \Delta_S = \frac{1}{H}\frac{\partial V}{\partial S} \tag{4.18}$$

to hedge the risk in the portfolio. Thus, the return on this risk-free portfolio must be equal to the domestic currency risk-free rate (r_d), which yields

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \bar{\rho}_t \sigma_H \sigma_S HS \frac{\partial^2 V}{\partial S \partial H} + \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 V}{\partial S^2} \\ &\quad + H\frac{\partial V}{\partial H}(r_d - r_f) + S\frac{\partial V}{\partial S}(r_f - \bar{\rho}_t \sigma_H \sigma_S) - r_d V = 0. \end{aligned} \tag{4.19}$$

To fully specify a particular quanto we consider a Quanto Put-option with the payoff at maturity

$$W(S_T, T) = H_0 \max(K - S_T, 0), \tag{4.20}$$

where H_0 is the exchange rate at the time zero (today). This means, it is agreed upon at the inception of the contract that the exchange rate at time-zero will be used at maturity. So there is no currency risk appears. By substituting (4.20) into (4.19) we obtain

$$\frac{\partial W}{\partial t} + \frac{1}{2}\sigma_S^2 S^2 \frac{\partial^2 W}{\partial S^2} + S\frac{\partial W}{\partial S}(r_f - \bar{\rho}_t \sigma_H \sigma_S) - r_d W = 0, \tag{4.21}$$

which is just the simple one factor Black-Scholes (BS) equation [11] with a time-dependent dividend yield of

$$D(\bar{\rho}_t) = r_d - r_f + \sigma_H \sigma_S \frac{1}{t}\int_0^t \bar{\rho}_s ds. \tag{4.22}$$

Finally, the price of a Quanto Put-Option in the extended BS model incorporating time-dependent dividend yield can be derived as

$$P = H_0\Big(K\exp^{-r_d T}\mathcal{N}(-d_2) - S_0\exp^{(r_f-r_d)T-\sigma_H\sigma_S\int_0^T\bar{\rho}_t dt}\mathcal{N}(-d_1)\Big), \tag{4.23}$$

with

$$d_1 = \frac{\log(\frac{S_0}{K}) + (r_f - \sigma_H\sigma_S\int_0^T\bar{\rho}_t dt + \frac{\sigma_S^2}{2})/T}{\sigma_S\sqrt{T}}, \qquad d_2 = d_1 - \sigma_S\sqrt{T}, \tag{4.24}$$

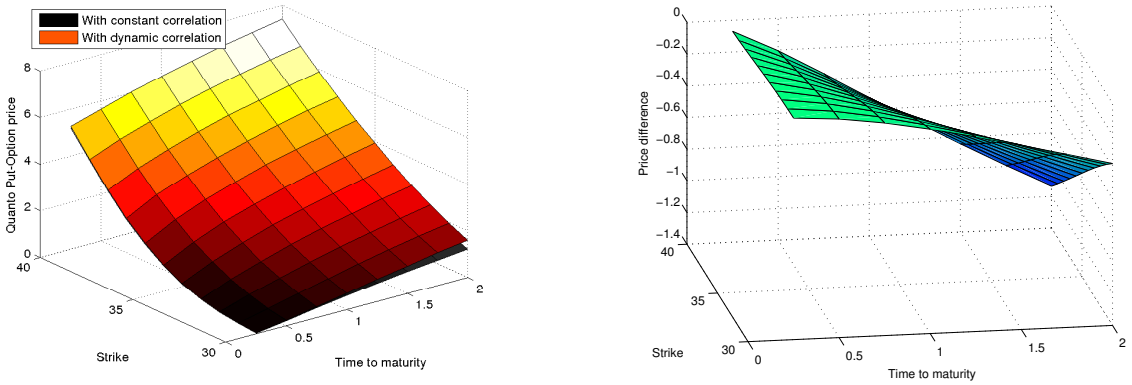
where the correlation function $\bar{\rho}_t$ is defined in (4.8). The price of a Quanto Call-Option can be derived easily from the put-call parity.

4.2.2 Dynamic Correlation vs. Constant Correlation

As an example, think of investing a Put-option on the Deutsche Bank stock traded in Euro (foreign currency) and converted to USD (domestic currency) at maturity. We assume that $S_0 = 36, H_0 = 1.3, r_d = 0.05, r_f = 0.03, \sigma_H = 0.3$ and $\sigma_S = 0.2$. For the dynamic correlation function we set $\bar{\rho}_0 = 0, \kappa = 2, \mu = 0.25, \sigma = 0.5$ and the value of the constant correlation to be 0.2.

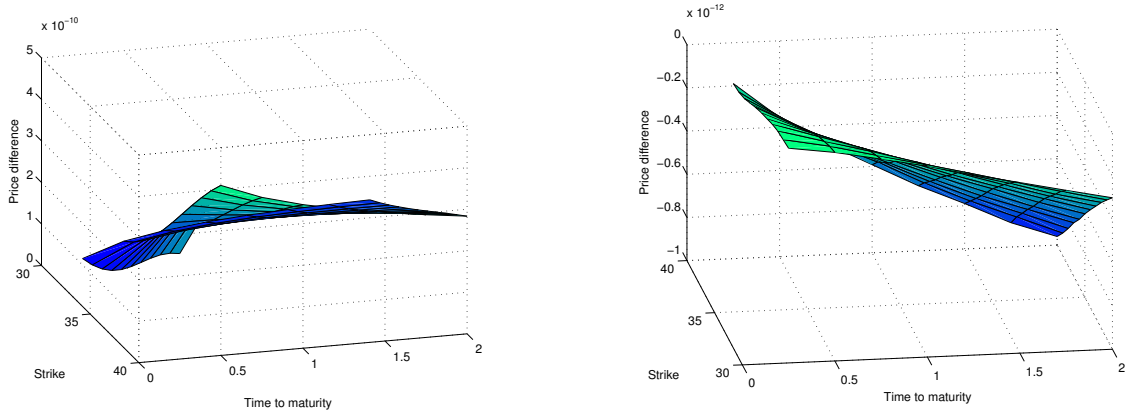
In Figure 4.5a we display the prices using constant and dynamic correlation for different strikes and maturities. We see that prices under the dynamic correlation are higher than the price using the constant correlation. To clarify the difference between them we show the difference in Figure 4.5b.

We now keep all the value of parameters to be the same except for setting $\mu = 0$. From (4.8) we see that the dynamic correlation function takes value around zero, which is the value of the constant correlation. This means that the price differences must be less than the last case. To see this, we plot the price differences for this case in Figure 4.6a and compare it to Figure 4.5b. Furthermore, we can set $\kappa = 8$ so that the dynamic correlation function reaches its limit (around zero) rapidly. For this case, the prices with



(a) (b)

Figure 4.5: Comparison of prices between using constant and dynamic correlation with $\kappa = 2$, $\mu = 0.25$, $\sigma = 0.5$ and $\bar{\rho}_0 = 0$ (correlation process parameters) and $\rho = 0$ (constant correlation).



(a) $\kappa = 2,\ \sigma = 0.5$ (b) $\kappa = 8,\ \sigma = 0.1$

Figure 4.6: Differences of prices between using constant and dynamic correlation with $\mu = 0$, and $\bar{\rho}_0 = 0$ (correlation process parameters) and $\rho = 0$ (constant Correlation).

and without dynamic correlation must be closer to each other, see the price differences in Figure 4.6b.

Effect on hedging

In the following, we discuss the effect of dynamic correlation on the hedging strategy. We consider the delta hedging as an example. By using a dynamic correlation, the delta is given by

$$\Delta_d = \Phi(d_1) - 1, \tag{4.25}$$

where Φ is standard normal distribution function and d_1 is defined in (4.24). Similarly, the delta by using a constant correlation is given by

$$\Delta_c = \Phi(d_1) - 1 \tag{4.26}$$

where d_1 is defined in (4.24) by setting $\bar{\rho}_t = \rho$. We take the same values for all BS parameters as in Figure 4.5 and set $\rho = \bar{\rho}_0 = 0, \kappa = 2, \mu = 0.6$ and $\sigma = 0.5$. Then, we compare the delta of a Quanto Put-option ($T = 1$) for different spot prices under the dynamic correlation to the corresponding delta with constant correlation in Figure 4.7. We observe that the delta values using the constant correlation are larger than the delta

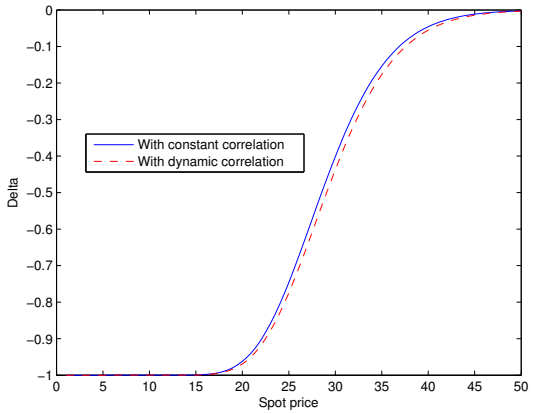


Figure 4.7: Comparison of the delta hedging with and without dynamic correlation.

values under the dynamic correlation.

4.2.3 Calibration to the Market Data

Here we illustrate the existing advantage of using dynamic correlation for the calibration to the market data. We take the Quanto puts on Deutsche Bank on July 30, 2013. The spot price is $S = 35.9$ Euro, the strike K_j ranges in [32, 33, 34, 35.9, 37, 38]. In the United States, if one invests these puts, the Euro-USD exchange rate is needed to convert the payoff in USD, which is $H_0 = 1.35$ on July 30, 2013. Furthermore, both

σ_S	σ_R	ρ	RMSE
0.32	0.20	-0.04	27×10^{-4}

Table 4.1: Estimated model parameters using constant correlation.

σ_S	σ_R	ρ_0	κ	μ	σ	RMSE
0.34	0.42	-0.57	2.07	0.49	0.3	9.3×10^{-4}

Table 4.2: Estimated model parameters using dynamic correlation.

interest rates r_f and r_d are 0.05 and the contract is considered for different maturities, $T_i \in [30, 90, 180, 240]$ days.

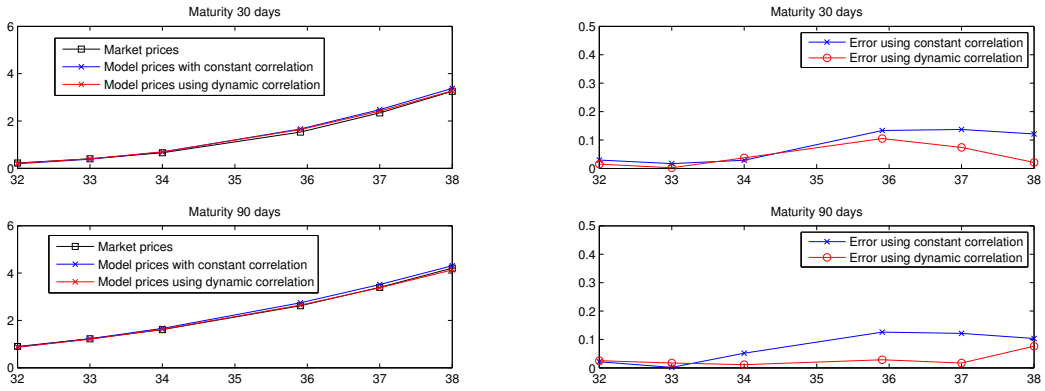
For each strike and maturity we denote the market price with $P^{Mkt}(\tau_i, K_j)$ and the corresponding model price with $P^{Mod}(\tau_i, K_j)$. We obtain the model parameters by minimizing, e.g., the relative mean square error (RMSE)

$$\frac{1}{N} \sum_{i,j} w_{ij} \frac{(P^{Mkt}(\tau_i, K_j) - P^{Mod}(\tau_i, K_j))^2}{P^{Mkt}(\tau_i, K_j)}, \tag{4.27}$$

where N is the number of prices and w_{ij} is an optional weight. Several numerical methods can be employed for this optimization, we choose a quasi-Newton approximation (e.g., matlab routine `fmincon`) in this example.

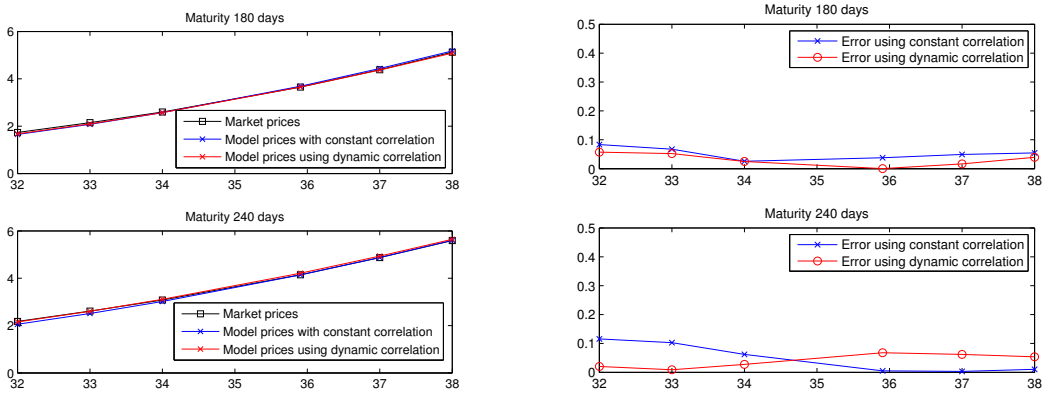
We estimate the parameters of the model using the constant and dynamic correlation, and report the estimated parameters and the errors in Table 4.1 and 4.2. We observe that the RMSE of the case using the constant correlation is almost three times larger than the RMSE of the case using the dynamic correlation. Furthermore, we present the plots of the market prices, the model prices with constant and dynamic correlation in Figure 4.8a and 4.9a. And we display the corresponding error, which is defined as $|P^{Mkt}(\tau_i, K_j) - P^{Mod}(\tau_i, K_j)|$, in 4.8b and 4.9b.

Either from the Table 4.1 and 4.2 or from the Figure 4.8 and 4.9, we directly conclude that the model under the dynamic correlation can be better fitted to the market prices. Furthermore, we can also expect better results using dynamic correlation for Quanto Caplets/Floorlets due to their pricing formula similar to (4.23). Especially, for pricing



(a) (b)

Figure 4.8: Comparison of market prices to model prices using constant and dynamic correlation for $T = 30$ and 60 days.



(a) (b)

Figure 4.9: Comparison of market prices to model prices using constant and dynamic correlation for $T = 180$ and 240 days.

Quanto caps/floor the advantages of using dynamic correlation are more obvious, since the instantaneous correlation for each time step is considered.

4.3 The Applications to the Heston Model

The Heston model [58] is one of the most widely used affine stochastic volatility models for equity prices. However, as mentioned before, in many situations the pure Heston model has a limitation on presenting a volatility smile, especially for a short maturity.

For this problem, several time-dependent Heston models have been proposed for a good fitting to implied volatilities, e.g., [6, 42, 86, 107], for a detailed explanation for those extensions we refer readers to [95]. Due to the fact that correlation mainly affects the slope of implied volatility smile, if the correlation is modelled with a time-dependent dynamic function, more realistic skews or smiles will be provided in the implied volatility surface by reasonably choosing additional parameters. In this section, we incorporate our time-dependent correlation function (4.8) into the Heston model.

4.3.1 Incorporating Dynamic Correlations

Heston's stochastic volatility model is specified as

$$dS_t = \mu_S S_t dt + \sqrt{\nu_t}\, S_t\, dW_t^S, \tag{4.28}$$

$$d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)dt + \sigma_\nu\sqrt{\nu_t}\, dW_t^\nu, \tag{4.29}$$

where (4.28) is assumed dynamics of the price of the spot asset, (4.29) is the volatility (variance) and W_t^S and W_t^ν are correlated with a constant $\rho_{S\nu}$. To incorporate the time-dependent correlation, we assume that dS_t and $d\nu_t$ are correlated by the time-dependent correlation function $\bar{\rho}_t$ instead of the constant correlation $\rho_{S\nu}$. The extended Heston model with a dynamic correlation $\bar{\rho}_t$ is specified as

$$dS_t = \mu_S S_t dt + \sqrt{\nu_t}\, S_t\, dW_t^1, \tag{4.30}$$

$$d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)dt + \sigma_\nu\sqrt{\nu_t}(\bar{\rho}_t\, dW_t^1 + \sqrt{1-\bar{\rho}_t^2}\, dW_t^2), \tag{4.31}$$

where W_t^1 and W_t^2 are independent. Applying Itô's lemma and no-arbitrage arguments yields [58]

$$\begin{aligned}\frac{1}{2}\nu^2 S^2\frac{\partial^2 U}{\partial S^2}+\bar{\rho}_t\sigma_\nu\nu S\frac{\partial^2 U}{\partial S\partial\nu}+\frac{1}{2}\sigma_\nu^2\nu\frac{\partial^2 U}{\partial\nu^2}+rS\frac{\partial U}{\partial S}\\ +[\kappa_\nu(\mu_\nu-\nu)-\tilde{\lambda}(S,\nu,\bar{\rho},t)\nu]\frac{\partial U}{\partial\nu}-rU+\frac{\partial U}{\partial t}=0,\end{aligned} \tag{4.32}$$

where $\bar{\rho}_t$ is defined in (4.8) but with the parameters $\bar{\rho}_0,\ \kappa_\rho,\ \mu_\rho,$ and ν_ρ. It is worth mentioning that the market price of volatility risk also depends on the dynamic correlation, which could be written as $\tilde{\lambda}(S,\nu,\bar{\rho}_t,t)$. This means, the price of correlation risk embedding in the price of volatility risk has been considered.

We consider e.g., a European call option with strike K and maturity T in the Heston model

$$C(S,\nu,t,\bar{\rho}_t)=SP_1-KP(t,T)P_2,\quad \tau=T-t, \tag{4.33}$$

where $P(t,T)$ is the discount factor and both probabilities P_1,P_2 must satisfy the PDE (4.32) as well as their characteristic functions, $f_1(S,\nu,\bar{\rho}_t,\phi,t)$ and $f_2(S,\nu,\bar{\rho}_t,\phi,t)$

$$f_j(S,\nu,\bar{\rho}_t,\phi,t)=\mathrm{e}^{C_j(\tau,\phi)+D_j(\tau,\phi)\nu+i\phi\ln S},\quad j=1,2. \tag{4.34}$$

By substituting this functional form (4.34) into the PDE (4.32) we can obtain the following ordinary differential equations (ODEs) for the unknown functions C and D :

$$-\frac{1}{2}\phi^2+\bar{\rho}_t\sigma_\nu\phi iD_j+\frac{1}{2}\sigma_\nu^2D_j^2+u_j\phi i-b_jD_j+\frac{\partial D_j}{\partial t}=0, \tag{4.35}$$

$$r\phi i+\kappa_\nu\mu_\nu D_j+\frac{\partial C_j}{\partial t}=0, \tag{4.36}$$

with the initial conditions $C_j(0,\phi)=D_j(0,\phi)=0$, and

$$u_1=0.5,\quad u_2=-0.5,\quad b_1=\kappa_\nu+\lambda-\bar{\rho}_t\sigma_\nu\quad\text{and}\quad b_2=\kappa_\nu+\lambda, \tag{4.37}$$

where

$$\bar{\rho}_t = 1 - \frac{\mathrm{e}^{-A(t)-\frac{B(t)}{2}}}{2} \int_{-\infty}^{\infty} \frac{1}{\cosh(\frac{\pi u}{2})} \cdot \mathrm{e}^{iu(A(t)+B(t))+u^2\frac{B(t)}{2}}\, du, \tag{4.38}$$

with $A(t) = \mathrm{e}^{-\kappa_\rho t}\mathrm{artanh}(\bar{\rho}_0) + \mu_\rho(1-\mathrm{e}^{-\kappa_\rho t}),\ B(t) = -\frac{\sigma_\rho^2}{2\kappa_\rho}(1-\mathrm{e}^{-2\kappa_\rho t})$.

Obviously, (4.35)-(4.36) cannot be solved analytically. Therefore, we need to find an efficient way to compute the option price numerically. For this we use an explicit Runge-Kutta method, the matlab routine `ode45`, to obtain C and D in (4.35)-(4.36) and thus also the characteristic functions (4.34). Finally, we employ the COS method [47] to obtain the option price $C(S, \nu, t, \bar{\rho})$ in (4.33). Thanks to the COS method, although we solved that ODE system numerically, the time for obtaining European option prices is less than 0.1 seconds such that a calibration can be performed quickly.

4.3.2 Calibration of the Heston Model under Dynamic Correlation

We calibrate the Heston model extended by our time-dependent correlation function to the real market data (Nikk300 index Call-options on July 16, 2012) and compare these to the pure Heston model [58] and the time-dependent Heston model [86].

We consider a set of N maturities $T_i,\ i = 1, \ldots, N$ and a set of M strikes $K_j,\ j = 1, \ldots, M$. Then for each combination of maturity and strike we have a market price $V^M(T_i, K_j) = V_{ij}^M$ and a corresponding model price $V(T_i, K_j; \Theta) = V_{ij}^\Theta$ generated by using (4.33). We choose the RMSE for the loss function $\frac{1}{M\times N}\sum_{i,j}\frac{(V_{ij}^M - V_{ij}^\Theta)^2}{V_{ij}^M}$, which can be minimized to obtain the parameter estimates

$$\hat{\Theta} = \arg\min \frac{1}{M \times N} \sum_{i,j} \frac{(V_{ij}^M - V_{ij}^\Theta)^2}{V_{ij}^M}. \tag{4.39}$$

For the optimization we restrict $\bar{\rho}_0$ to the interval $(-1, 1)$ but not the value of μ_ρ. Since it is not the direct limit of the correlation function but the mean reversion of the Ornstein-

Uhlenbeck process, thus, it could take any value in $\mathbb{R}$. Our experiments showed, that it is sufficient and appropriate to restrict μ_ρ to the interval $[-4, 4]$.

We state our estimated parameters and the estimation error for the pure Heston model (abbr. PH), the Heston model under our time-dependent correlations (CH), the time-dependent Heston model by Mikhailov and Nögel [86] (MN) in Table 4.3, 4.4 and 4.5, respectively.

The pure Heston model					
$\hat{\nu}_0$	$\hat{\kappa}_\nu$	$\hat{\mu}_\nu$	$\hat{\sigma}_\nu$	$\hat{\rho}$	Estimation Error
0.029	4.746	0.053	1.108	-0.355	1.10×10^{-3}

Table 4.3: The estimated parameters for the pure Heston model using Call-options on the Nikk300 index on July 16, 2012.

The extended Heston model by using our time-dependent correlation function								
$\hat{\nu}_0$	$\hat{\kappa}_\nu$	$\hat{\mu}_\nu$	$\hat{\sigma}_\nu$	$\hat{\bar{\rho}}_0$	$\hat{\kappa}_\rho$	$\hat{\mu}_\rho$	$\hat{\sigma}_\rho$	Estimation Error
0.027	5.542	0.055	1.224	-0.165	5.333	-0.752	0.434	2.38×10^{-4}

Table 4.4: The estimated parameters for the Heston model under time-dependent correlations using Call-options on the Nikk300 index on July 16, 2012.

The time-dependent Heston model by Mikhailov and Nögel						
Maturity	$\hat{\nu}_0$	$\hat{\kappa}_\nu$	$\hat{\mu}_\nu$	$\hat{\sigma}_\nu$	$\hat{\rho}$	Estimation Error
1/12	0.025	2.749	0.095	1.172	-0.201	1.78×10^{-4}
1/4	0.012	2.936	0.076	0.524	-0.411	2.45×10^{-5}
1/2	0.011	2.890	0.058	0.592	-0.430	1.14×10^{-5}
1	0.001	2.911	0.051	0.558	-0.389	4.28×10^{-6}

Table 4.5: The estimated parameters for the time-dependent Heston model by Mikhailov and Nögel using Call-options on the Nikk300 index on July 16, 2012.

We see that the estimation error using the CH model is significantly less than the error using the PH model and almost the same as the error (sum of errors for each maturity) under the MN model. To illustrate more clearly, for each maturity we compare the implied volatilities for all the models to the market volatilities in Figure 4.10. We can observe that the implied volatilities for the CH model are much closer to the market volatilities than the implied volatilities for the PH model, especially has the better

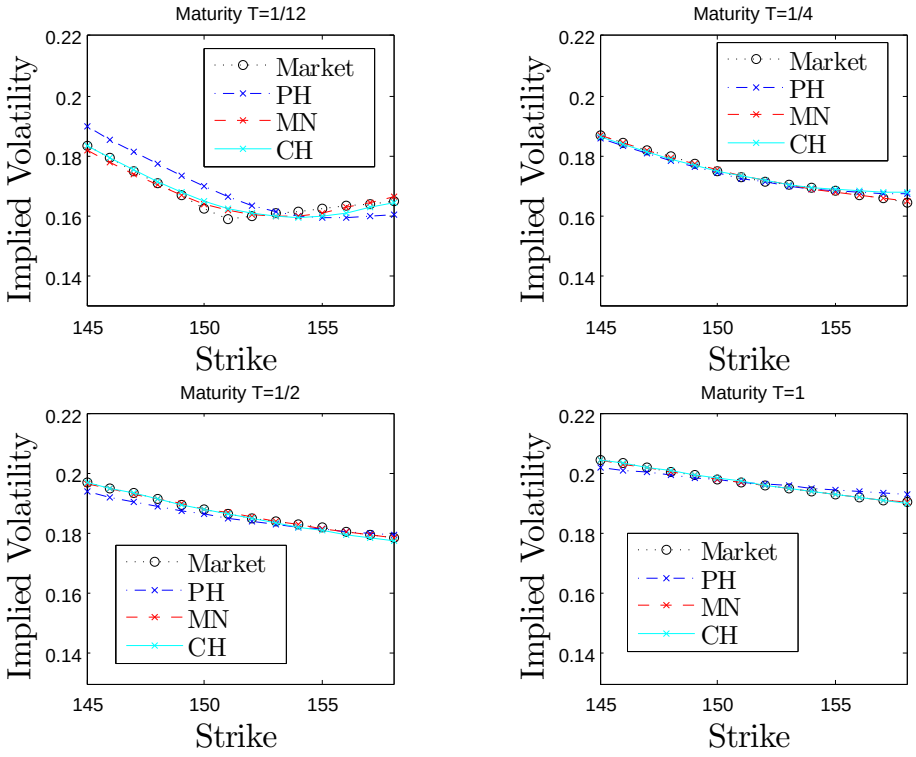


Figure 4.10: The comparison of implied volatilities for all the models to the market volatilities of the Call-options on the Nikk300 index on July 16, 2012, where the spot price is 150.9.

volatility smile for the short maturity $T = 1/12$. Comparing to the MN model, the implied volatilities for our model are almost the same. However, our CH model has an economic interpretation, namely the correlation is nonlinear and time-dependent as market requires. We conclude that the Heston model extended by incorporating our time-dependent correlations can provide a better volatility smile than the pure Heston model. Another nice issue is that the time-dependent correlation function can be easily and directly introduced into the financial models. The application of the time-dependent correlation model in pricing option with stochastic interest rate can be found in [106].

Part III

Modelling and Applications of Stochastic Correlation

As randomness features more generally, like moving from time-dependent interest rate to stochastic interest rate, from time-dependent volatility to stochastic volatility, in this part we investigate how to model correlation as a stochastic process and its applications in finance. We provide a general stochastic correlation model and discuss several stochastic correlation processes. We apply stochastic correlation to price Quanto options and quantify the correlation risk caused by using a wrong (constant) correlation. Furthermore, we incorporate stochastic correlation into the Heston model and find, that the Heston model extended by introducing stochastic correlation provide a better fit to the skew and smile in the volatility surface that is visible in the market.

Chapter 5

Stochastic Correlation Models and its Application

In Chapter 4 we have investigated a time-dependent correlation model and its application to some financial models. We find, instead of using a deterministic constant correlation, incorporating time-dependent correlations into a financial model can improve its performance. However, only a time-dependent correlation model might be not enough to model real phenomena in financial world due to the uncertainty associated with the future development of relationship between, e.g., financial parties and products. This is to say that applying stochasticity for correlation should better replicate correlation properties in reality. Like moving from time-dependent interest rate to stochastic interest rate, from time-dependent volatility to stochastic volatility, we turn in this chapter to stochastic correlation modelling and its application.

5.1 Stochastic Correlation Models

Firstly, we mention two concepts of stochastic correlation models which are widely applied in finance. The first one is dynamic conditional correlation by Engle [44, 45]. The second one is based on the Wishart process introduced by [23] and extended by [54],

and see [25, 49] for its application to the Heston model. In financial markets, the first problem of using a correlation concept is the *observability*. Unlike price, exchange rate and so on, the correlation cannot be observed directly in the market and can only be measured in the context of a model. The easiest estimator of the correlation is the sample correlation coefficient which has been given in (B.2). Furthermore, using (B.3) one can calculate a *historical rolling correlation* .

5.1.1 Historical Correlation

To find financial correlation properties in reality, we make an example of historical correlations between *S&P 500 index* and *Euro/US-Dollar exchange rate* on a daily basis. We take daily log-return series of S&P 500 and Euro/US-Dollar exchange rate and calculate the 15-day, 30-day and 60-day historical correlations using (B.3) which are displayed in Figure 5.1.

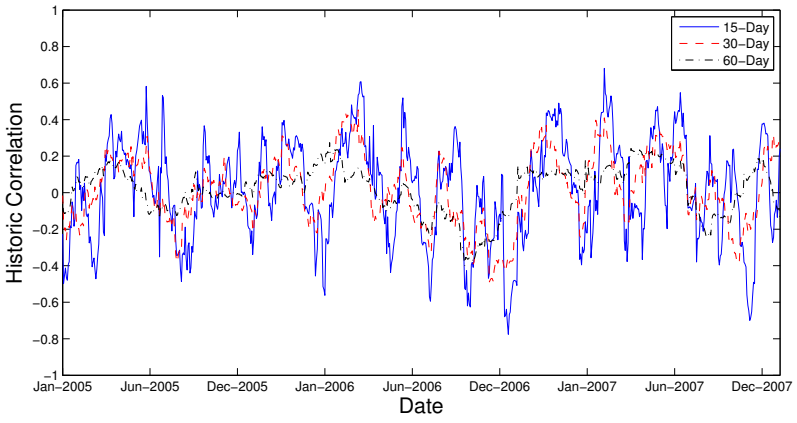


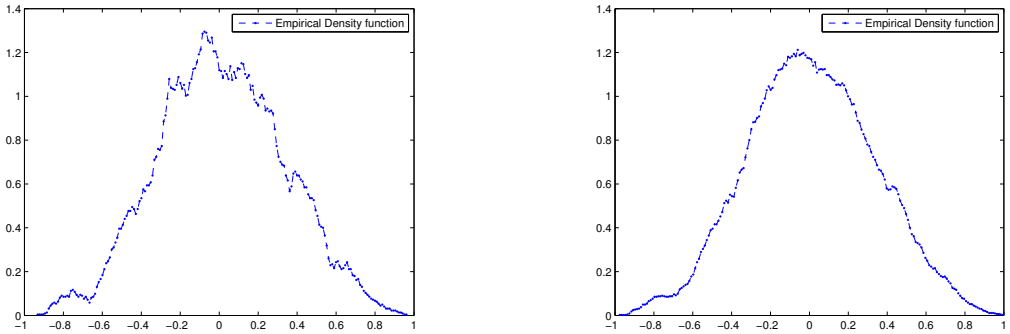
Figure 5.1: Historical Correlation between S&P 500 and Euro/US-Dollar exchange rate (Source of data: `www.yahoo.com`).

We observe that the longer a time window is, the less volatile a historical correlation is. In Figure 5.1, the 15-day historical correlation is more variable than the 30-day historical correlation which is again more variable than the 60-day correlation. With a longer averaging period a *long-term correlation* is calculated. If we choose the time window as

10 or 15 days, the estimated correlation for each time t using (B.2), could be seen as a *short-term correlation* of the current market phenomena whose immediate past returns are used for the estimation. It is worthwhile noting that the events, especially, some extreme events in a time window will affect the estimated correlation in the following time windows, even has a delayed effect on the long-term correlation.

If one assumes that the phenomena in the past could be a reflection of the future, one would like to use the historical correlation as a forecast for the future. It could be a better way for correlation forecasting, if one describes the correlation using a *mean-reverting stochastic process*. Besides, modelling correlation as a stochastic process, not only the variation of the short-term correlation can be reflected, also the attributes of long-term correlation is determined by the long-term parameter values, like long-term mean value and mean reversion speed. For more detailed information about historical correlation we refer to [1].

To see more properties, which a mean-reverting stochastic process should have to be a stochastic correlation process (SCP), we plot its empirical density functions in Figure 5.2 using different bandwidths. We refer to [12] for details about the estimation of density function from historical data.



(a) bandwidth 1/40 (b) bandwidth 1/30

Figure 5.2: Empirical Density function of the historical correlation between S&P 500 and Euro/US-Dollar exchange rate.

From the illustration of historical correlation above it seems like a good idea to model correlation as a stochastic process which should satisfy the following properties [109, 114]:

(i) only takes values in the interval $(-1, 1)$,

(ii) varies around a mean value,

(iii) the probability mass tends to zero at the boundaries -1, $+1$.

We remark that the first two properties are similar to those properties for a time-dependent correlation function introduced in Section 4.1.1.

5.1.2 A General Stochastic Correlation Model

For the motivations and the properties (i)–(iii) in Section 5.1.1, we propose the hyperbolic tangent function of a mean-reverting stochastic process X_t, like the OU process [113] or other square root diffusion processes (with positive and negative values)

$$dX_t = a(t, X_t)dt + b(t, X_t)\, dW_t, \quad t \geq 0,\ X_0 = x_0, \tag{5.1}$$

to model the correlations as

$$\rho_t = \tanh(X_t), \quad \rho_0 = \tanh(x_0) \in (-1, 1). \tag{5.2}$$

Note that we used the expectation of $\tanh(X_t)$ to model correlation as a time-dependent function in Section 4.1.1. Obviously, the properties (i)–(iii) are fulfilled due to the range of values of the hyperbolic tangent and mean reversion of the process. Besides, the function tanh is symmetric and measurable. Although the function tanh can not really attain -1 and 1, which respectively presents perfect negative and perfect positive dependence, one can still use it for modelling correlations, because the correlation equal to -1 or 1 is indeed an extreme event which happens very rarely in the real market, see, e.g., Figure 4.2. Besides, the function tanh tends to the boundaries -1 and 1 at infinity.

Applying *Itô's Lemma* [61] with (5.2)

$$d\rho_t = d\tanh(X_t) = \frac{\partial \tanh(X_t)}{\partial t}\,dt + \frac{\partial \tanh(X_t)}{\partial x}\,dX_t + \frac{1}{2}\frac{\partial^2 \tanh(X_t)}{\partial x^2}(dX_t)^2\,, \tag{5.3}$$

we obtain the *stochastic correlation process (SCP)*

$$d\rho_t = (1-\rho_t^2)\left((\tilde{a}-\rho_t\tilde{b}^2)dt + \tilde{b}dW_t\right), \quad t \geq 0, \tag{5.4}$$

where $\rho_0 \in (-1,1)$, $\tilde{a} = a(t, \operatorname{artanh}(\rho_t))$ and $\tilde{b} = b(t, \operatorname{artanh}(\rho_t))$. From (5.4) we see that there is a suitable number of free parameters to calibrate the model to market data. Besides, it is obvious, in this approach any mean-reverting process (with positive and negative values) can be considered. The free parameters are hidden in the functions a and b, which depend on the chosen underlying mean-reverting process.

Although we could intuitively observe that the function $\tanh(x)$ is eminently suitable for correlation modelling, one can still ask whether other functions having values inside the interval $(-1,1)$, like trigonometric functions or $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$, $x \in \mathbb{R}$ can also be applied for this purpose? In theory, such functions could be used for the SCP model above. However, the trigonometric function is a periodic function, the arising complex number will complicate further calculations. For the function $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$, its Itô's formula for the process (5.1) is given by

$$\begin{aligned} d\rho_t = d\frac{2}{\pi}\arctan\Big(\frac{\pi}{2}X_t\Big) = &\left(\frac{\tilde{a}}{\big(1+\tan^2(\frac{\rho_t\pi}{2})\big)} - \frac{\pi\tilde{b}^2\tan(\frac{\rho_t\pi}{2})}{2\big(1+\tan^2(\frac{\rho_t\pi}{2})\big)^2}\right)dt \\ &+ \frac{\tilde{b}}{\big(1+\tan^2(\frac{\rho_t\pi}{2})\big)}\,dW_t\,, \end{aligned} \tag{5.5}$$

which is rather complicate such that the further computation will turn out to be tedious. Nevertheless, we will additionally consider the function $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$ which is, like $\tanh(x)$ close to the identity in the neighbourhood of $x = 0$, see Figure 5.3. However, compared with $\tanh(x)$, the function $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$ grows much slower up to 1 and

down to -1, the estimation of the correlation will thus be worsened, similar to the estimation for the heavy tailed distributions.

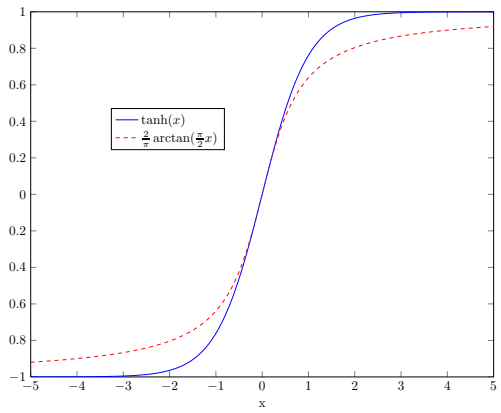


Figure 5.3: Comparison of $\tanh(x)$ and $\frac{2}{\pi}\arctan(\frac{\pi}{2}x)$: the later is less steep having larger tails.

Calibration

We can estimate the free parameters of (5.4) using the density function. If we choose for (5.1) a process which has the known density function, the density function of (5.2) thus can be derived and used for the calibration purpose, e.g., for the case of using OU process, see Section 5.1.3. Otherwise, we need to determine the transition density with the aid of the *Fokker-Planck equation* [93].

Only for simplicity, we rewrite (5.4) with the redefined parameters $\hat{a}$ and $\hat{b}$ as

$$d\rho_t = \underbrace{(1-\rho_t^2)(\tilde{a}-\rho_t\tilde{b}^2)}_{:=\hat{a}(t,\rho_t)}\,dt + \underbrace{(1-\rho_t^2)\tilde{b}}_{:=\hat{b}(t,\rho_t)}\,dW_t, \quad t \geq 0, \tag{5.6}$$

where $\rho_0 \in (-1,1)$. We assume that it possesses a transition density $p(t,\tilde{\rho}|\rho_0)$ which satisfies the *Fokker-Planck equation*

$$\frac{\partial}{\partial t}p(t,\tilde{\rho}) + \frac{\partial}{\partial\tilde{\rho}}(\hat{a}(t,\tilde{\rho})p(t,\tilde{\rho})) - \frac{1}{2}\frac{\partial^2}{\partial\tilde{\rho}^2}(\hat{b}(t,\tilde{\rho})^2 p(t,\tilde{\rho})) = 0. \tag{5.7}$$

For the calibration purpose we consider the stationary density (for $t \to \infty$)

$$p(\tilde{\rho}) := \lim_{t\to\infty} p(t, \tilde{\rho}|\rho_0). \tag{5.8}$$

With the above construction (5.4) is also a mean-reverting process, thus one can show that every two solutions of (5.7) are the same for $t \to \infty$, this is to say that a unique stationary solution $p(\tilde{\rho})$ exists, cf. [93]. Besides, the following two standard conditions for a density function should be fulfilled by $p(\tilde{\rho})$,

$$\int_{-1}^{1} p(\tilde{\rho}) d\tilde{\rho} = 1, \tag{5.9}$$

$$\int_{-1}^{1} \tilde{\rho} \cdot p(\tilde{\rho}) d\tilde{\rho} \xrightarrow[t\to\infty]{} \text{mean value}. \tag{5.10}$$

Up to now, we have just shown our structural idea of SCP. Several exact examples by choosing different mean-reverting processes as underlying process for (5.1) with the detailed stochastic calculus will be presented in the next sections, their calibration as well.

5.1.3 Variant I: Stochastic Correlation with an OU Process

We specify our SCP model using the OU process. For the basis process (5.1) we choose the OU process

$$dX_t = \kappa(\mu - X_t)dt + \sigma dW_t, \tag{5.11}$$

where $\kappa, \sigma > 0$ and $X_0, \mu \in \mathbb{R}$.

Proposition 5.1.1 *Applying Itô's Lemma with* $\rho_t = \tanh(X_t)$,

$$d\rho_t = \frac{\partial \tanh(X_t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 \tanh(X_t)}{\partial x^2} \sigma^2 dt \tag{5.12}$$

gives the SCP as

$$d\rho_t = (1-\rho_t^2)\left(\kappa(\mu - artanh(\rho_t)) - \rho_t\sigma^2\right)dt + (1-\rho_t^2)\sigma dW_t, \tag{5.13}$$

where $t \geq 0,\ \rho_0 \in (-1,1),\ \kappa, \sigma > 0$ *and* $\mu \in \mathbb{R}$.

The Proof can be found in Appendix C.

Density function and calibration

As mentioned in Section 5.1.2, we do not really need the transition density of (5.13) in this case, since the OU process $X_t \sim \mathcal{N}(x_0 \mathrm{e}^{-\kappa t} + \mu(1-\mathrm{e}^{-\kappa t}), \frac{\sigma^2}{2\kappa}(1-\mathrm{e}^{-2\kappa t}))$ is normal distributed, if the initial value x_0 is given. As $t \to \infty$, then $X_t \sim \mathcal{N}(\mu, \frac{\sigma^2}{2\kappa})$. Therefore, one can derive the density function for (5.13) as $t \to \infty$ as

$$f(\tilde{\rho}) = \frac{1}{1-\tilde{\rho}^2} \cdot \frac{\sqrt{\kappa}}{\sigma\sqrt{\pi}} \cdot \mathrm{e}^{-\frac{\kappa(\mathrm{artanh}(\tilde{\rho})-\mu)^2}{\sigma^2}}, \tag{5.14}$$

which can be used to cablibrate the model.

In the following, we still derive the transition density of (5.13) to show how this approach works. Besides, we want to compare the transition density of (5.13) to (5.14). As pointed out in Section 5.1.2, we assume that (5.13) possesses a transition density $p(t,\tilde{\rho}|\rho_0)$ which satisfies the following Fokker-Planck equation

$$\frac{\partial}{\partial t}p(t,\tilde{\rho}) + \frac{\partial}{\partial \tilde{\rho}}(\hat{a}(t,\tilde{\rho})p(t,\tilde{\rho})) - \frac{1}{2}\frac{\partial^2}{\partial \tilde{\rho}^2}(\hat{b}(t,\tilde{\rho})^2 p(t,\tilde{\rho})) = 0 \tag{5.15}$$

with

$$\hat{a}(t,\tilde{\rho}) = (1-\tilde{\rho}^2)\left(\kappa(\mu - \mathrm{artanh}(\tilde{\rho})) - \tilde{\rho}\sigma^2\right), \tag{5.16}$$

$$\hat{b}(t,\tilde{\rho}) = (1-\tilde{\rho}^2)\sigma. \tag{5.17}$$

For $t \to \infty$, the stationary density $p(\tilde{\rho})$ can be obtained by solving

$$\frac{\partial}{\partial\tilde{\rho}}((1-\tilde{\rho}^2)\left(\kappa(\mu-\operatorname{artanh}(\tilde{\rho}))-\tilde{\rho}\sigma^2\right)p(\tilde{\rho})) = \frac{1}{2}\frac{\partial^2}{\partial\tilde{\rho}^2}(((1-\tilde{\rho}^2)\sigma)^2 p(\tilde{\rho})) \tag{5.18}$$

with

$$p(\tilde{\rho}) = \frac{\left(m + n\ \operatorname{erf}\left(\frac{\sqrt{-\kappa}(\operatorname{artanh}(\tilde{\rho})-\mu)}{\sigma}\right)\right)\mathrm{e}^{-\frac{\kappa\operatorname{artanh}(\tilde{\rho})}{\sigma^2}(\operatorname{artanh}(\tilde{\rho})-2\mu)}}{\tilde{\rho}^2-1} \tag{5.19}$$

and constants $m,\ n \in \mathbb{R}$.

Now we try to simplify (5.19). Firstly we can easily observe, n must be zero, so that the condition (5.10) can be satisfied by (5.19). We can check this straightly by setting $\mu = 0$. Thus, (5.19) can be further written as

$$p(\tilde{\rho}) = \frac{m}{\tilde{\rho}^2-1}\cdot\mathrm{e}^{-\frac{\kappa\operatorname{artanh}(\tilde{\rho})}{\sigma^2}(\operatorname{artanh}(\tilde{\rho})-2\mu)}. \tag{5.20}$$

In theory we can compute m by solving the condition (5.9) with (5.19), but the integration of (5.19) will be tedious. However, due to the uniqueness of the asymptotic distribution, m can be specified by identifying (5.14) and (5.20) as

$$m = -\frac{\sqrt{\kappa}}{\sigma\sqrt{\pi}}\mathrm{e}^{\frac{-\mu^2\kappa}{\sigma^2}}. \tag{5.21}$$

By substituting (5.21) for m in (5.20) we can obtain the transition density function which is the same to (5.14).

As mentioned before, (5.14) can be used to estimate the parameters of (5.13). However, in this case of using the OU process, we can even calibrate the model considering each time step. The OU process $X_t \sim \mathcal{N}(x_0\mathrm{e}^{-\kappa t} + \mu(1-\mathrm{e}^{-\kappa t}), \frac{\sigma^2}{2\kappa}(1-\mathrm{e}^{-2\kappa t}))$ has the (conditional) probability density

$$f_x(x_{s+\Delta t}|x_s,\kappa,\mu,\sigma) = \sqrt{\frac{\kappa}{\pi\sigma^2(1-\mathrm{e}^{-2\kappa\Delta t})}}\cdot\mathrm{e}^{\frac{-\kappa(x_{s+\Delta t}-x_s\mathrm{e}^{-\kappa\Delta t}-\mu(1-\mathrm{e}^{-\kappa\Delta t}))^2}{\sigma^2(1-\mathrm{e}^{-2\kappa\Delta t})}}, \tag{5.22}$$

from which we derive the density of $\rho_t = \tanh(X_t)$ directly as

$$f_\rho(\tilde{\rho}_{s+\Delta t}|\tilde{\rho}_s, \kappa, \mu, \sigma) = \sqrt{\frac{a}{b}} \cdot \frac{1}{1-\tilde{\rho}_{s+\Delta t}^2} \cdot \mathrm{e}^{\frac{-\kappa(\operatorname{artanh}(\tilde{\rho}_{s+\Delta t}) - \operatorname{artanh}(\tilde{\rho}_s)\mathrm{e}^{-\kappa\Delta t} - \mu c)^2}{\sigma^2 b}} \tag{5.23}$$

with

$$a = \frac{\kappa}{\pi\sigma^2}, \quad b = (1-\mathrm{e}^{-2\kappa\Delta t}) \quad \text{and} \quad c = (1-\mathrm{e}^{-\kappa\Delta t}). \tag{5.24}$$

Therefore, we prefer to employ the *maximum-likelihood estimation* for the historical correlation, see [19, 50]. We use θ to denote the collection of the parameters κ, μ and σ, for the $n+1$ given observed correlations $(\tilde{\rho}_0, \tilde{\rho}_1, \cdots, \tilde{\rho}_t)$. We derive its *log-likelihood* function

$$\begin{aligned}\mathcal{L}(\theta|\tilde{\rho}_0, \tilde{\rho}_1, \cdots, \tilde{\rho}_t) = &\sum_{i=1}^{n} \log\left(\sqrt{\frac{\kappa}{\pi\sigma^2(1-\mathrm{e}^{-2\kappa(t_i-t_{i-1})})}}\right) + \sum_{i=1}^{n} \log\left(\frac{1}{1-\tilde{\rho}_{t_i}^2}\right) \\ &+ \sum_{i=1}^{n} \frac{-\kappa(\operatorname{artanh}(\tilde{\rho}_{t_i}) - \operatorname{artanh}(\tilde{\rho}_{t_{i-1}})\mathrm{e}^{-\kappa(t_{i+1}-t_i)} - \mu(1-\mathrm{e}^{-\kappa(t_{i+1}-t_i)}))^2}{\sigma^2(1-\mathrm{e}^{-2\kappa(t_{i+1}-t_i)})}.\end{aligned} \tag{5.25}$$

Furthermore, the parameter estimators $\hat{\theta} = (\hat{\kappa}, \hat{\mu}, \hat{\sigma})$ can be obtained by maximizing (5.25). This can be done for example by various numerical optimization methods. Besides, we remark that the derivatives of (5.25) with respect to μ and σ can be found analytically and only tedious with respect to κ. The expressions for $\hat{\mu}$ and $\hat{\sigma}$ can thus be obtained by solving

$$\frac{\partial\mathcal{L}(\theta|\tilde{\rho}_0, \tilde{\rho}_1, \cdots, \tilde{\rho}_t)}{\partial\sigma} = 0 \quad \text{and} \quad \frac{\partial\mathcal{L}(\theta|\tilde{\rho}_0, \tilde{\rho}_1, \cdots, \tilde{\rho}_t)}{\partial\mu} = 0. \tag{5.26}$$

We write the results as

$$\hat{\mu} = \sum_{i=1}^{n} \frac{\operatorname{artanh}(\tilde{\rho}_{t_i}) - \operatorname{artanh}(\tilde{\rho}_{t_{i-1}})\mathrm{e}^{-\kappa(t_i-t_{i-1})}}{1+\mathrm{e}^{-\kappa(t_i-t_{i-1})}} \Big/ \sum_{i=1}^{n} \frac{1-\mathrm{e}^{-\kappa(t_i-t_{i-1})}}{1+\mathrm{e}^{-\kappa(t_i-t_{i-1})}} \tag{5.27}$$

and

$$\hat{\sigma} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{2\kappa(\operatorname{artanh}(\tilde{\rho}_{t_i}) - \operatorname{artanh}(\tilde{\rho}_{t_{i-1}})\mathrm{e}^{-\kappa(t_i - t_{i-1})} - \mu(1 - \mathrm{e}^{-\kappa(t_i - t_{i-1})}))^2}{1 - \mathrm{e}^{-2\kappa(t_i - t_{i-1})}} \right)^{\frac{1}{2}}. \quad (5.28)$$

We see that $\hat{\mu}$ has the expression only with respect to κ, as well as $\hat{\sigma}$ by substituting μ in (5.28) by (5.27). Hence, we substitute $\hat{\mu}$ and $\hat{\sigma}$ in (5.25) to gain the *log-likelihood* function only with the parameter κ, which can be computed by maximizing this function. Finally, we can substitute the value of $\hat{\kappa}$ back to (5.27) and (5.28) to get values of $\hat{\mu}$ and $\hat{\sigma}$.

As an example we estimated the SCP parameters using the historical correlation in Figure 5.2b. Then, we compared (5.14) using the estimated parameters and the empirical density function of historical correlation, see Figure 5.4. We remark that no additional

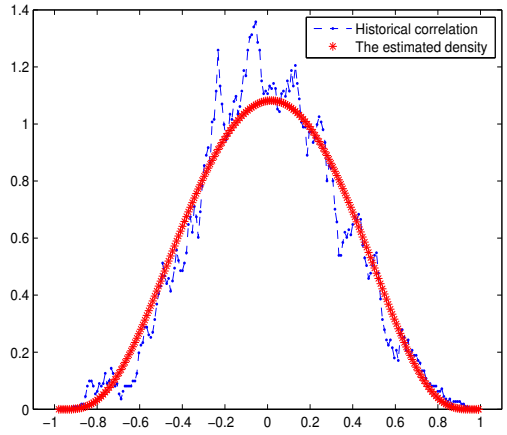


Figure 5.4: The estimated parameters: $\hat{\kappa} = 32.11$, $\hat{\mu} = 0.012$ and $\hat{\sigma} = 2.96$

parameter restrictions appear using this SCP, which simplifies the calibration procedure.

5.1.4 Variant II: Stochastic Correlation with a Modified OU Process

If one want to restrict the mean value μ in (5.11) to be only in (-1,1), it is reasonable to modify first the OU process (5.11) as [110]

$$dX_t = \kappa\big(\mu - \tanh(X_t)\big)\,dt + \sigma\,dW_t, \tag{5.29}$$

where κ, $\sigma > 0$ and X_0, $\mu \in (-1, 1)$.

Proposition 5.1.2 *Applying Itô's Lemma with* $\rho_t = \tanh(X_t)$ *yields the SCP*

$$d\rho_t = (1-\rho_t^2)\big(\kappa(\mu-\rho_t) - \sigma^2\rho_t\big)\,dt + (1-\rho_t^2)\sigma\,dW_t, \tag{5.30}$$

where $t \geq 0$, $\rho_0 \in (-1, 1)$, κ, $\sigma > 0$ *and* $\mu \in (-1, 1)$.

The proof is similar to Proposition 5.1.1, we leave this to the readers.

Next, we modify the notation and rewrite (5.30) as

$$\kappa^* = \kappa + \sigma^2, \quad \mu^* = \frac{\kappa\mu}{\kappa+\sigma^2}, \quad \sigma^* = \sigma, \tag{5.31}$$

$$\frac{d\rho_t}{1-\rho_t^2} = \kappa^*(\mu^* - \rho_t)dt + \sigma^*\,dW_t, \tag{5.32}$$

where $t \geq 0$, $\rho_0 \in (-1, 1)$, κ^*, $\sigma^* > 0$ and $\mu^* \in (-1, 1)$. Then, we generalize the correlation process (5.32) directly with the arbitrary parameter coefficients $\kappa > 0$, $\mu \in (-1, 1)$ and $\sigma > 0$, like

$$\frac{d\rho_t}{1-\rho_t^2} = \kappa(\mu - \rho_t)dt + \sigma\,dW_t. \tag{5.33}$$

Proposition 5.1.3 *The value of* (5.33) *is bounded in the interval* $(-1, 1)$, *if the condition*

$$\kappa > \frac{\sigma^2}{1 \pm \mu} \tag{5.34}$$

is fulfilled, and its stationary density function $f(\tilde{\rho})$ exists and can be analytically calculated as

$$f(\tilde{\rho}) = \frac{(1+\tilde{\rho})^{a_\rho+b_\rho}(1-\tilde{\rho})^{a_\rho-b_\rho}}{M}\,, \tag{5.35}$$

with

$$\begin{aligned} M :=& \frac{\Gamma(1+a_\rho-b_\rho)F(1,-a_\rho-b_\rho,2+a_\rho-b_\rho,-1)}{\Gamma(2+a_\rho-b_\rho)} \\ &+ \frac{\Gamma(1+a_\rho+b_\rho)F(1,-a_\rho+b_\rho,2+a_\rho+b_\rho,-1)}{\Gamma(2+a_\rho+b_\rho)}, \end{aligned} \tag{5.36}$$

where $a_\rho = \frac{\kappa-2\sigma^2}{\sigma^2}$, $b_\rho = \frac{\kappa\mu}{\sigma^2}$, F is the hypergeometric function (see B.6) and Γ is Gamma function.

The proof is given in Appendix C.

Transition density function

The transition density function $f(\tilde{\rho})$ in (5.35) can be used for calibration purposes. To further illustrate it we display in Figures 5.5, 5.6 and 5.7 the behavior of $f(\tilde{\rho})$ for different values of each parameter used in [105]. In Figure 5.5, we let $\kappa = 2$ and $\mu = 0$ and display $f(\tilde{\rho})$ with different values of σ, which is equal to 0.3, 0.4 and 0.5, respectively. Obviously, σ shows the magnitude of variation from the mean value $\mu = 0$. Next, we fix $\kappa = 2$ and $\sigma = 0.3$, the behavior of $f(\tilde{\rho})$ with different mean values $\mu = -0.5$, $\mu = 0$ and $\mu = 0.5$ can be found in Figure 5.6. However, whilst $\mu = -0.5$ and $\mu = 0.5$ we can observe that the peak of the corresponding $f(\tilde{\rho})$ does not locate exactly at the points $\tilde{\rho} = -0.5$ and $\tilde{\rho} = 0.5$, respectively. The reason is that, the value of κ, which is mean reversion rate, is not large enough. In order to illustrate the role of κ, we set $\mu = -0.5$, $\sigma = 0.5$ and vary the value of κ, see Figure 5.7. For $\kappa = 3$, the peak of the transition density function is far away from the mean value -0.5. However, in contrast the peak reaches the point $\tilde{\rho} = -0.5$ when $\kappa = 12$.

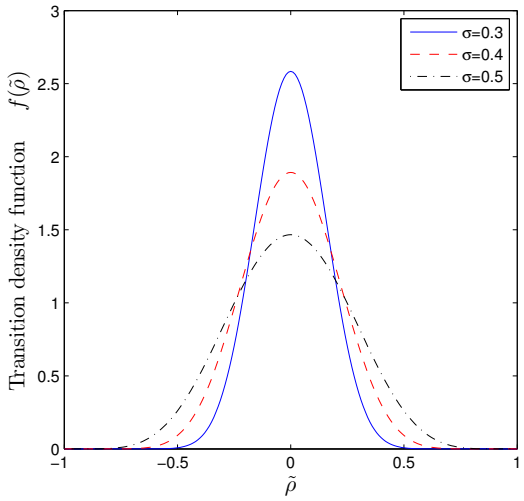


Figure 5.5: Comparison of $f(\tilde{\rho})$ for different values of σ ($\kappa = 2$ and $\mu = 0$).

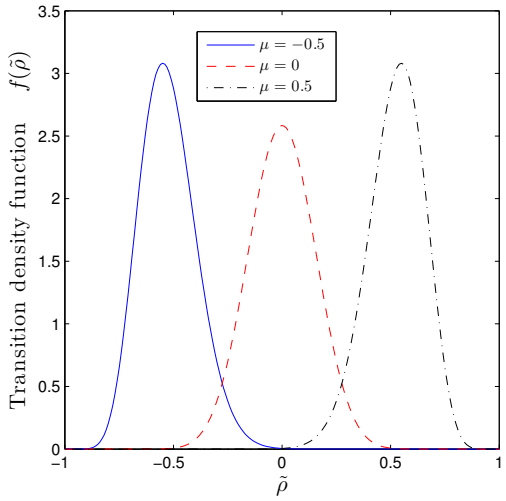


Figure 5.6: Comparison of $f(\tilde{\rho})$ for different values of μ ($\kappa = 2$ and $\sigma = 0.3$).

5.1.5 Variant III: Stochastic Correlation with a Bounded Jacobi Process

Ma [79] and van Emmerich [114] proposed to use the bounded Jacobi process to model stochastic correlation. Indeed, the bounded Jacobi process is also in the class of our general stochastic models, which means the bounded Jacobi process can be obtained by transforming a special mean-reverting process.

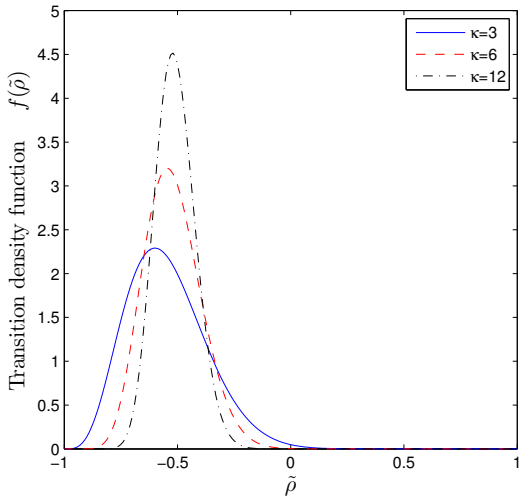


Figure 5.7: Comparison of $f(\tilde{\rho})$ for different values of κ ($\mu = -0.5$ and $\sigma = 0.5$).

We define the following mean-reverting process

$$dX_t = \frac{\kappa\big(\mu - \tanh(X_t)\big)}{1 - \tanh^2(X_t)}dt + \frac{\sigma}{\sqrt{1 - \tanh^2(X_t)}}dW_t, \quad t \geq 0,\ X_0 = x_0, \tag{5.37}$$

where κ and σ are positive, $\mu \in (-1, 1)$. Then, we transform (5.37) with $\rho_t = \tanh(X_t)$, with the aid of Itô's Lemma we obtain

$$d\rho_t = \big[(\kappa(\mu - \rho_t)) - \sigma^2\rho_t\big]\, dt + \sigma\sqrt{1 - \rho_t^2}\, dW_t. \tag{5.38}$$

The calculation is straightforward but tedious. Now, if we define

$$\kappa^* = \kappa + \sigma^2, \quad \mu^* = \frac{\kappa\mu}{\kappa + \sigma^2}, \quad \sigma^* = \sigma, \tag{5.39}$$

the correlation process (5.38) can be rewritten as

$$d\rho_t = \kappa^*(\mu^* - \rho_t)dt + \sigma^*\sqrt{1 - \rho_t^2}\, dW_t, \tag{5.40}$$

which is exactly the bounded Jacobi process used by van Emmerich in [114]. Due to the transformation with the function tanh, the correlations provided by (5.40), with coefficients (5.39), are obviously located in the interval $(-1, 1)$.

However, if we redefine (5.40) with arbitrary parameters κ, μ and σ as

$$d\rho_t = \kappa(\mu - \rho_t)dt + \sigma\sqrt{1-\rho_t^2}\, dW_t, \tag{5.41}$$

the following condition

$$\kappa \geq \frac{\sigma}{1 \pm \mu} \tag{5.42}$$

must be satisfied to ensure that the boundaries -1 and 1 are unattainable. For the proof and discussion about calibration we refer to [114].

5.1.6 Stochastically correlated BMs and its Construction

Following the way to define the dynamically correlated BMs, from (4.10) we give the following definition.

Definition 5.1.1 *Two Brownian motions $W_{t,1}$ and $W_{t,2}$ are called stochastically correlated with the stochastic process ρ_t, if they satisfy*

$$E\left[W_{t,1}W_{t,2}\right] = E\left[\int_0^t \rho_s ds\right], \tag{5.43}$$

where $\rho_t : \Omega \times \mathbb{R}^+ \to (-1, 1)$ is a SCP. The average correlation of $W_{t,1}$ and $W_{t,2}$, ρ_{Av}, is given by $\rho_{Av} := \frac{1}{t}E[\int_0^t \rho_s ds]$.

The construction of stochastically correlated BMs is also similar to the case of dynamically correlated BMs presented in 4.1.2. We consider first the two-dimensional case and let ρ_t be a stochastic correlation process. For two independent BMs $W_{t,1}$ and $W_{t,3}$ we define

$$W_{t,2} = \int_0^t \rho_s dW_{s,1} + \int_0^t \sqrt{1-\rho_s^2}\, dW_{s,3}, \tag{5.44}$$

with the symbolic expression

$$dW_{t,2} = \rho_t dW_{t,1} + \sqrt{1-\rho_t^2}\, dW_{t,3}. \tag{5.45}$$

It can be easily verified that $W_{t,2}$ is a BM and correlated with $W_{t,1}$ stochastically by ρ_t. Besides, the covariance matrix and the average correlation matrix of $\mathbb{W}_t = (W_{t,1}, W_{t,2})$ can be determined, given by

$$\begin{pmatrix} t & E\left[\int_0^t \rho_s ds\right] \\ E\left[\int_0^t \rho_s ds\right] & t \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & \frac{1}{t}E\left[\int_0^t \rho_s ds\right] \\ \frac{1}{t}E\left[\int_0^t \rho_s ds\right] & 1 \end{pmatrix}$$

respectively. The construction can be also straightforwardly generalized to n-dimensions.

5.2 Pricing Quanto Options with Stochastic Correlation

In Section 4.2 we have investigated the Quanto pricing under time-dependent correlations. In this section we utilize the extended Black-Scholes formula by using our stochastic correlation model to evaluate the fair price of the quanto options.

5.2.1 The Formula of Quanto Pricing

As an example we think of a Put-Option on the S&P 500 with payoff in USD

$$(\underbrace{\text{Strike}}_{:=K} - \underbrace{\text{S\&P500}_T}_{:=S_T})^+, \tag{5.46}$$

then the payoff of a currency-protected Quanto Put-option in Euro can be written as

$$\underbrace{\text{exchangeRate}_0}_{:=H_0} \cdot (\text{Strike} - \text{S\&P500}_T)^+, \tag{5.47}$$

where H_0 is the Euro/USD (number of Euro per USD) exchange rate agreed upon at the inception of the contract.

We recall that S_t and R_t are modelled by

$$\begin{cases} dS_t = \mu_S S_t dt + \sigma_S S_t dW_t^S \\ dH_t = \mu_H H_t dt + \sigma_H H_t dW_t^H \quad . \end{cases} \tag{5.48}$$

Instead of using a time-dependent correlation, the W_t^S and W_t^R are stochastically correlated by the SCP introduced in Section 5.1.3, namely,

$$d\rho_t = (1-\rho_t^2)\left(\kappa(\mu - \operatorname{artanh}(\rho_t)) - \rho_t\sigma^2\right)dt + (1-\rho_t^2)\sigma dW_t. \tag{5.49}$$

Furthermore, we assume nonzero relationships between the SCP and the price, the exchange rate process, say

$$dW_t dW_t^S = \rho_1 dt \quad \text{and} \quad dW_t dW_t^H = \rho_2 dt. \tag{5.50}$$

In fact, we are trying to incorporate the SCP (5.49) in the model (5.48) exogenously. For this reason we could assume that $\rho_1 = 0$ and $\rho_2 = 0$.

We denote the risk-free interest rate of Euro and US-Dollar by r_e and r_d, respectively. To incorporate the stochasticity of the correlation exogenously in the BS model, we consider the following strategy to obtain the no-arbitrage price: First we think that the expected return of one unit of US-Dollar, exchanged to Euro, risk-free invested in the Euro countries and re-exchanged to US-Dollar must be equal to the risk-free return on

one unit of US-Dollar in US-Dollar countries, which reads

$$\frac{\exp(r_e T)H_0}{\mathbb{E}[H_T]} = \exp(r_d T). \tag{5.51}$$

The exchange rate H_t follows a geometric Brownian motion and thus $\mathbb{E}[H_T] = H_0 \exp(\mu_H T)$, which can be substituted into (5.51) to get

$$\mu_H = r_e - r_d. \tag{5.52}$$

Besides, the expected value of an investment of one Euro unit into the underlying with price S must be equal to risk-free return on one unit of US-Dollar in US-Dollar countries, which gives

$$\frac{1}{H_0}\frac{1}{S_0}\mathbb{E}[S_T H_T] = \exp(r_d T). \tag{5.53}$$

For calculating $\mathbb{E}[S_T H_T]$, we apply firstly Itô's lemma to the function $u(t, S_t, H_t) = \ln(S_t H_t)$

$$du(t, S_t, H_t) = d\ln(S_t H_t) = (\mu_S + \mu_R - \frac{1}{2}(\sigma_S^2 + \sigma_H^2))dt + \sigma_S dW_t^S + \sigma_H dW_t^H. \tag{5.54}$$

Furthermore, the last equation can be rewritten as

$$\ln(S_T H_T) - \ln(S_0 H_0) = (\mu_S + \mu_H - \frac{1}{2}(\sigma_S^2 + \sigma_H^2))T + \sigma_S W_T^S + \sigma_H W_T^H \tag{5.55}$$

which implies

$$\mathbb{E}[S_T H_T] = S_0 H_0 \exp\left((\mu_S + \mu_H - \frac{1}{2}(\sigma_S^2 + \sigma_H^2))T\right)\mathbb{E}[\exp(\sigma_S W_T^S + \sigma_H W_T^H)]. \tag{5.56}$$

Now, we set $dX_S = \sigma_S dW_t^S$ and $dX_H = \sigma_H dW_t^H$. A further application of Itô's lemma to the function $f(t, X_S, X_H) = \exp(X_S + X_H)$ leads to

$$\mathbb{E}[\exp(\sigma_S W_T^S + \sigma_H W_T^H)] = \exp(\frac{T}{2}(\sigma_S^2 + \sigma_H^2))\mathbb{E}\left[\exp(\sigma_S\sigma_H \int_0^T \rho_t dt)\right]. \tag{5.57}$$

We substitute the last equation into (5.56) to obtain

$$\mathbb{E}[S_T H_T] = S_0 H_0 \exp((\mu_S + \mu_H)T)\mathbb{E}\left[\exp(\sigma_S\sigma_H \int_0^T \rho_t dt)\right]. \tag{5.58}$$

Thus, we can choose

$$\mu_S = r_d - \mu_H - \frac{1}{T}\ln\mathbb{E}\left[\exp(\sigma_S\sigma_H \int_0^T \rho_t dt)\right], \tag{5.59}$$

such that the no-arbitrage condition (5.53) can be fulfilled.

Remark 5.2.1 *The expectation in (5.59) should be considered under a (unique) risk-neutral measure. If we think the correlation between assets can also be traded directly, e.g., with correlation swaps, so that correlation risk could completely be hedged. In such complete market, there exists exactly one risk-neutral measure for the expectation in (5.59). Otherwise, one need to choose a particular pricing measure in the case of uniqueness, for example we can pick the one which is closest to the physical probability measure in terms of relative entropy.*

In the following, we assume that the market is complete for trading correlation. This is to say that the correlation risk can be hedged by choosing μ_S with (5.59). Now, in the framework of BS model, we interpret (5.59) as a return minus the continuous dividend, the dividend can thus be obtained as

$$D(\rho_t) := \mu_H + \frac{1}{T}\ln\mathbb{E}\left[\exp(\sigma_S\sigma_H \int_0^T \rho_t dt)\right]. \tag{5.60}$$

Together with (5.52) we have

$$D(\rho_t) = r_e - r_d + \frac{1}{T} \ln \mathbb{E}\left[\exp(\sigma_S \sigma_H \int_0^T \rho_t dt)\right]. \tag{5.61}$$

The integral of the stochastic correlation ρ_t can be computed numerically using, e.g., the Milstein scheme. see, e.g., [66].

Finally, we adopt the price formular (4.23) of a Quanto Put-Option to the one under stochastic correlation, which is denoted by $\mathbb{P}_{\text{Quanto}}$:

$$\mathbb{P}_{\text{Quanto}}(S_0, K, r_e, r_d, D(\rho_t), T) = R_0 \left(K \mathrm{e}^{-r_e T} \mathcal{N}(-d_2) - S_0 \mathrm{e}^{-D(\rho_t)T} \mathcal{N}(-d_1)\right) \tag{5.62}$$

with

$$d_1 = \frac{\log(\frac{S_0}{K}) + ((r_e - D(\rho_t)) + \frac{\sigma_S^2}{2})/T}{\sigma_S \sqrt{T}}, \qquad d_2 = d_1 - \sigma_S \sqrt{T}. \tag{5.63}$$

The price of a Quanto Call-Option can be easily derived from the put-call parity. If we apply for example the Monte-Carlo approach to approximate the expectation under risk-neutral measure in (5.61), the price $\mathbb{P}_{\text{Quanto}}$ can thus be directly computed.

5.2.2 The Effect of Stochastic Correlation on Hedging

Based on the formula on (5.62) we derive the Delta as an example to discuss the effect of stochastic correlation on the delta hedging strategy. For using a constant correlation, the delta is given by

$$\Delta_c = \Phi(d_1) - 1, \tag{5.64}$$

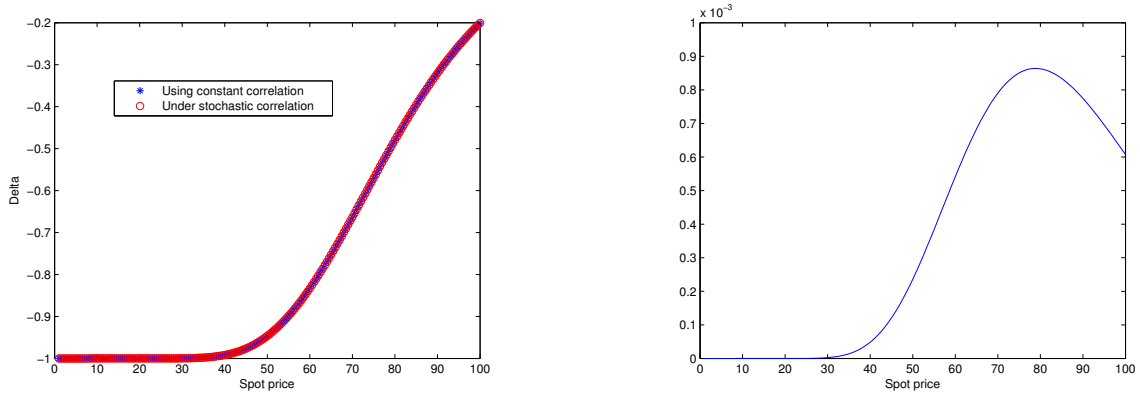
where Φ is the standard normal distribution function and d_1 is defined by

$$d_1 = \frac{\log(\frac{S_0}{K}) + ((r_d - D) + \frac{\sigma_S^2}{2})/T}{\sigma_S \sqrt{T}}, \tag{5.65}$$

where $D = r_e - r_d + \rho\sigma_S\sigma_H$ with a constant correlation ρ. Similarly, we have the Delta under a stochastic correlation as

$$\Delta_s = \Phi(d_1) - 1, \tag{5.66}$$

where d_1 is (5.63). One can use a Monte-Carlo approach to approximate (5.61) and then compute $D(\rho_t)$ in (5.63). Using the same parameters in Figure 5.9 and taking the maturity $T = 2$ we display Δ_c and Δ_s in Figure 5.8a. Since the value of μ is



(a) (b)

Figure 5.8: BS parameters: $K = 80$, $S_0 = 100$, $H_0 = 1$, $r_d = 0.03$, $r_e = 0.05$, $\sigma_S = 0.2$, $\sigma_H = 0.4$, Correlation process parameters: $\kappa = 32.11, \mu = 0.012, \sigma = 2.96$ and $\rho_0 = 0.025$.

close to the constant correlation, the difference between Δ_c and Δ_s is not apparent. Therefore, in Figure 5.8b, we plot the difference $\Delta_s - \Delta_c$ in Figure 5.8a and observe that delta values under the stochastic correlation are larger than the delta values using the constant correlation.

5.2.3 Numerical Results of an Example

In Figure 5.9, 5.10 and 5.11, we show our numerical results for pricing a quanto Put-Option and analyze the results centering around correlation risk.

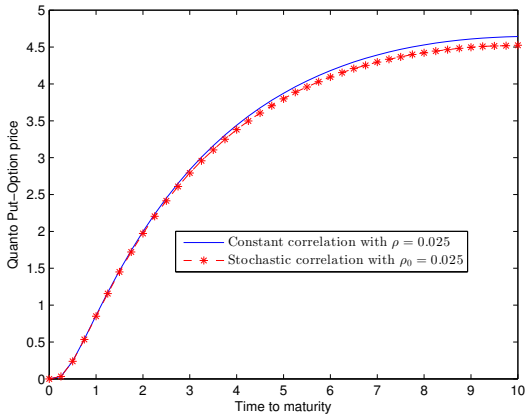


Figure 5.9: BS parameters: $K = 80,\ S_0 = 100,\ H_0 = 1,\ r_d = 0.03,\ r_e = 0.05,\ \sigma_S = 0.2,\ \sigma_H = 0.4$, Correlation process parameters: $\kappa = 31.11, \mu = 0.012, \sigma = 2.96$ and $\rho_0 = 0.025$.

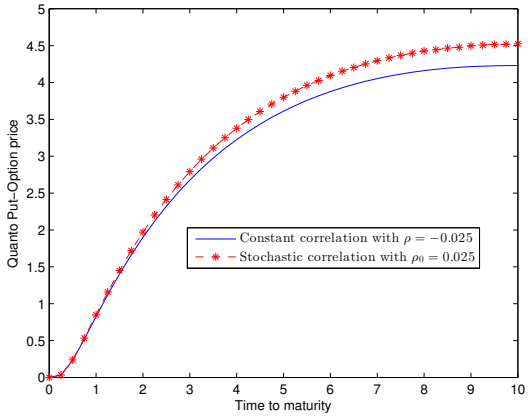


Figure 5.10: BS parameters: $K = 80,\ S_0 = 100,\ H_0 = 1,\ r_d = 0.03,\ r_e = 0.05,\ \sigma_S = 0.2,\ \sigma_H = 0.4$, Correlation process parameters: $\kappa = 32.11, \mu = 0.012, \sigma = 2.96$ and $\rho_0 = 0.025$.

First In Figure 5.9, we set the parameter for the BS model and use the estimated parameter for the SCP model (see Figure 5.4). Besides, we use the whole historical data (Jan 2003 - Mar 2013) of S&P 500 and Euro/US-Dollar exchange rate but only to estimate the constant correlation which is 0.025. At the same time, we can let the SCP ρ_t start from this value, $\rho_0 = 0.025$. It is clearly to see, the prices of Put-Option with the constant correlation are higher than the prices using the stochastic correlations. The

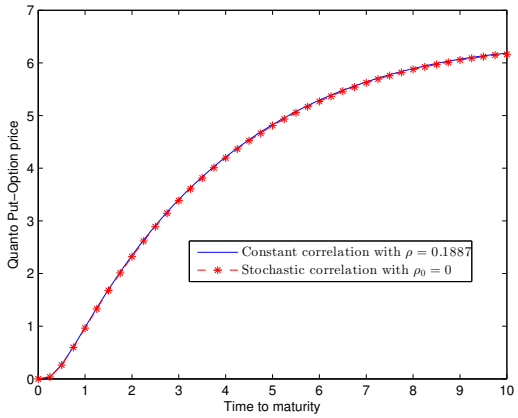


Figure 5.11: BS parameters: $K = 80,\ S_0 = 100,\ H_0 = 1,\ r_d = 0.03,\ r_e = 0.05,\ \sigma_S = 0.2,\ \sigma_H = 0.4$, Correlation process parameters: $\kappa = 10, \mu = 0.2, \sigma = 1$ and $\rho_0 = 0 \Rightarrow$ Mean value of the SCP model : 0.1887.

difference is even getting larger with the increasing maturity. We interpret this difference as the correlation risk using the wrong (constant) correlation.

In Figure 5.10, we change the constant correlation to -0.025 and keep the other parameter the same as in Figure 5.9. We observe that the prices with the constant correlation could be also lower than the prices applying the stochastic correlation. The sign of the price of correlation risk in this case is opposite of that sign in Figure 5.9.

It is very interesting to see the results in Figure 5.11, the prices using the constant correlation and stochastic correlation are very close, even almost the same for the longer time. We explain the reason for this phenomena as follows. In this case the parameters of BS model remain unchanged as the last two examples, but we set here $\kappa = 10, \mu = 0.2$ and $\sigma = 1$. Thus, we can compute numerically the mean value of (5.49) for these assumed parameters which is 0.1887. We then price the Quanto Put-Option with constant correlation $\rho = 0.1887$ and using the stochastic correlation with $\rho_0 = 0$. Because the value of the mean-reverting factor κ is large and the value of σ is small, such that the stochastic correlation process concentrates strongly on its mean value, this is why there is no obvious difference between the prices using the constant correlation and stochastic correlation in this special case. Analyzing the numerical results we conclude that the

correlation risk caused by a using wrong (constant) correlation can be priced through applying our new SCP model, which can not always be neglected.

5.3 Stochastic Correlation in the Heston Model

In Section 4.3 we have investigated the extension of the Heston model by time-dependent correlations. Our results have shown that performance of the Heston model regarding the calibration to real market data can be improved only by allowing an appropriate time-dependent correlations. Therefore, we believe that a significant improvement could be expected by imposing a SCP into the Heston model. This is the motivation for working on the Heston model in this section.

5.3.1 Stochastic Correlation in the Heston Model

Let us recall that the Heston's stochastic volatility model [58] under the risk-neutral measure is specified as

$$\begin{cases} dS_t = rS_t\,dt + \sqrt{\nu_t}S_t\,dW_t^S, & S_0 > 0, \\ d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)\,dt + \sigma_\nu\sqrt{\nu_t}\,dW_t^\nu, & \nu_0 > 0, \end{cases} \tag{5.67}$$

where BMs W_t^S and W_t^ν are correlated with a constant $\rho_{S\nu}$. Under the log-transform for the asset, i.e. $x_t = \log(S_t)$, the model is represented by

$$\begin{cases} dx_t = (r - \frac{1}{2}\nu_t)\,dt + \sqrt{\nu_t}\,dW_t^x, & x_0 = \log(S_0), \\ d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)\,dt + \sigma_\nu\sqrt{\nu_t}\,dW_t^\nu, & \nu_0 > 0, \end{cases} \tag{5.68}$$

which is in the class of affine diffusion processes (AD), see B.7. The discounted CF has been found by Heston [58]. We extend the model by imposing stochastic correlation

between the asset price and the volatility given by an appropriate SDE system:

$$\begin{cases} dx_t = (r - \frac{1}{2}\nu_t)\,dt + \sqrt{\nu_t}\,dW_t^x, & x_0 = \log(S_0), \\ d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)dt + \sigma_\nu\sqrt{\nu_t}\,dW_t^\nu, & \nu_0 > 0, \\ d\rho_t = a(t,\rho_t)dt + b(t,\rho_t)dW_t^\rho, & \rho_0 \in [-1,1], \end{cases} \tag{5.69}$$

where

$$dW_t^x dW_t^\nu = \rho_t\,dt, \quad dW_t^x dW_t^\rho = \rho_{x\rho}\,dt, \quad dW_t^\nu dW_t^\rho = \rho_{\nu\rho}\,dt, \tag{5.70}$$

i.e. the log price process and the volatility process are set to be correlated stochastically, driven by the correlation process ρ_t which is by itself correlated with the log-price process by $\rho_{x\rho}$ and with the volatility by $\rho_{\nu\rho}$, respectively.

Affinity

To conveniently check the affinity, we reformulate the SDE system (5.69) with respect to the independent BMs: We first rearrange the SDE system (5.69) as

$$\begin{cases} d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)dt + \sigma_\nu\sqrt{\nu_t}\,dW_t^\nu, \\ d\rho_t = a(t,\rho_t)dt + b(t,\rho_t)dW_t^\rho, \\ dx_t = (r - \frac{1}{2}\nu_t)\,dt + \sqrt{\nu_t}\,dW_t^x, \end{cases} \tag{5.71}$$

which has a family of correlation matrices

$$\mathcal{C}_t = \begin{pmatrix} 1 & \rho_{\nu\rho} & \rho_t \\ \rho_{\rho\nu} & 1 & \rho_{\rho x} \\ \rho_t & \rho_{x\rho} & 1 \end{pmatrix}, \quad t \geq 0, \tag{5.72}$$

which is symmetric, namely $\rho_{\nu\rho} = \rho_{\rho\nu}$ and $\rho_{x\rho} = \rho_{\rho x}$. To simplify the notation we set $\rho_1 := \rho_{\nu\rho}(\rho_{\rho\nu})$ and $\rho_2 := \rho_{x\rho}(\rho_{\rho x})$. Obviously, (5.72) is positive semi-definite if

$$1 - \rho_1^2 - \rho_2^2 + 2\rho_1\rho_2\rho_t - \rho_t^2 \geq 0 \tag{5.73}$$

which implies

$$\rho_1\rho_2 - \sqrt{(1-\rho_1^2)(1-\rho_2^2)} \leq \rho_t \leq \rho_1\rho_2 + \sqrt{(1-\rho_1^2)(1-\rho_2^2)}. \tag{5.74}$$

Under the condition (5.74), one can perform a Cholesky decomposition $\mathcal{C}_t = \mathcal{L}_t\mathcal{L}_t^\top$, where $\mathcal{L}_t$ is a family of lower triangular matrices given by

$$\mathcal{L}_t = \begin{pmatrix} 1 & 0 & 0 \\ \rho_1 & \sqrt{1-\rho_1^2} & 0 \\ \rho_t & \frac{\rho_2-\rho_1\rho_t}{\sqrt{1-\rho_1^2}} & \sqrt{1-\rho_t^2-\left(\frac{\rho_2-\rho_1\rho_t}{\sqrt{1-\rho_1^2}}\right)^2} \end{pmatrix}, \quad t \geq 0, \tag{5.75}$$

which can be employed to reformulate the SDE system (5.71) with respect to the independent BMs $\tilde{W}_t^\nu, \tilde{W}_t^\rho$ and $\tilde{W}_t^x$ as:

$$\begin{cases} d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)\,dt + \sigma_\nu\sqrt{\nu_t}\,d\tilde{W}_t^\nu, \\ d\rho_t = a(t,\rho_t)\,dt + \rho_1 b(t,\rho_t)\,d\tilde{W}_t^\nu + \sqrt{1-\rho_1^2}b(t,\rho_t)\,d\tilde{W}_t^\rho, \\ dx_t = (r - \frac{1}{2}\nu_t)\,dt + \rho_t\sqrt{\nu_t}\,d\tilde{W}_t^\nu + \frac{\rho_2-\rho_1\rho_t}{\sqrt{1-\rho_1^2}}\sqrt{\nu_t}\,d\tilde{W}_t^\rho \\ \qquad\qquad + \sqrt{1-\rho_t^2-\left(\frac{\rho_2-\rho_1\rho_t}{\sqrt{1-\rho_1^2}}\right)^2}\sqrt{\nu_t}\,d\tilde{W}_t^x. \end{cases} \tag{5.76}$$

The family of symmetric instantaneous covariance matrices for $\mathbf{X_t} := [\nu_t, \rho_t, x_t]^\top$ reads

$$\sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^\top = \begin{pmatrix} \nu_t\sigma_\nu^2 & \rho_1\sigma_\nu\sqrt{\nu_t}b(t,\rho_t) & \sigma_\nu\nu_t\rho_t \\ * & b^2(t,\rho_t) & \rho_2 b(t,\rho_t)\sqrt{\nu_t} \\ * & * & \nu_t \end{pmatrix}, \quad t \geq 0. \tag{5.77}$$

Since our main aim is to impose a stochastic correlation between the asset process and the stochastic volatility process, we first assume $\rho_1 = 0$ so that the latter SDE

system becomes

$$\begin{cases} d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)dt + \sigma_\nu\sqrt{\nu_t}d\tilde{W}_t^\nu, \\ d\rho_t = a(t,\rho_t)dt + b(t,\rho_t)d\tilde{W}_t^\rho, \\ dx_t = \left(r - \frac{1}{2}\nu_t\right)dt + \rho_t\sqrt{\nu_t}d\tilde{W}_t^\nu + \rho_2\sqrt{\nu_t}d\tilde{W}_t^\rho + \sqrt{1-\rho_t^2-\rho_2^2}\sqrt{\nu_t}d\tilde{W}_t^x, \end{cases} \tag{5.78}$$

and the family of symmetric instantaneous covariance matrices reads

$$\sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^\top = \begin{pmatrix} \nu_t\sigma_\nu^2 & 0 & \sigma_\nu\nu_t\rho_t \\ * & b^2(t,\rho_t) & \rho_2 b(t,\rho_t)\sqrt{\nu_t} \\ * & * & \nu_t \end{pmatrix}, \quad t \geq 0. \tag{5.79}$$

We define the discounted characteristic function $\phi(\mathbf{u}, \mathbf{X}_t, t, T) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-r(T-t)+i\mathbf{u}^\top\mathbf{X}_T}|\mathcal{F}_t\right]$, whose Kolmogorov's backward equation is given by

$$\begin{aligned} \frac{\partial\phi}{\partial t} + (r - \frac{1}{2}\nu)\frac{\partial\phi}{\partial x} + \kappa_\nu(\mu_\nu - \nu)\frac{\partial\phi}{\partial\nu} + a(t,\rho_t)\frac{\partial\phi}{\partial\rho} + \frac{1}{2}\nu\frac{\partial^2\phi}{\partial x^2} + \frac{1}{2}\nu\sigma_\nu^2\frac{\partial^2\phi}{\partial\nu^2} \\ + \frac{1}{2}b^2(t,\rho)\frac{\partial^2\phi}{\partial\rho^2} + \sigma_\nu\nu_t\rho_t\frac{\partial^2\phi}{\partial\nu\partial x} + \rho_2 b(t,\rho_t)\sqrt{\nu_t}\frac{\partial^2\phi}{\partial\rho\partial x} - r\phi = 0 \end{aligned} \tag{5.80}$$

subject to the terminal condition $\phi(\mathbf{u}, \mathbf{X}_T, T, T) = \mathrm{e}^{iux_t}$. Obviously, the system (5.78) is not in an affine form. We can use appropriate approximations in order to generate an affine form. We first consider $\sigma_\nu\nu_t\rho_t$: Assuming independence between ρ_t and ν_t we can straightforwardly take the following approximation

$$\sigma_\nu\nu_t\rho_t \approx \mathbb{E}\left[\sigma_\nu\nu_t\rho_t\right] = \sigma_\nu\mathbb{E}\left[\nu_t\right]\mathbb{E}\left[\rho_t\right]. \tag{5.81}$$

A better approximation could be

$$\sigma_\nu\nu_t\rho_t \approx \sigma_\nu\mathbb{E}\left[\nu_t\right]\rho_t, \tag{5.82}$$

which is justified by the assumption $\rho_1 = 0$, because the stochasticity of the correlation process is kept. We discuss the affinity of the terms including $a(t, \rho_t)$ and $b(t, \rho_t)$ in the next section, as it will depend on the chosen stochastic correlation process.

5.3.2 Incorporating the OU Process into the Heston Model

If one uses the OU process to model stochastic correlation, the major drawback of using an OU process for stochastic correlation is that the process is not bounded. This is to say the generated correlations can be out of the correlation interval $(-1, 1)$. This specially occurs for a small value of κ_ρ and a large value of σ_ρ. However, due to its analytical tractability, one would like to use it for modelling correlation; e.g., Düllmann et al. [41] estimated asset correlations from stock prices or default rates by assuming that correlation follows an OU process. In this section, we check the feasibility of using an OU process [113] to be a SCP.

We recall that the OU process is given by

$$d\rho_t = \kappa_\rho(\mu_\rho - \rho_t)\, dt + \sigma_\rho\, d\tilde{W}_t^\rho. \tag{5.83}$$

Therefore, the functions $a(t, \rho_t)$ and $b(t, \rho_t)$ defined in (5.78) and (5.79) are known as $\kappa_\rho(\mu_\rho - \rho_t)$ and σ_ρ, respectively.

We employ it for modelling stochastic correlations while we limit the mean value μ_ρ to be in $(-1, 1)$ and choose a relative large value of κ_ρ, a small value of σ_ρ. We name this extended Heston model as "HO" model. In the HO model, the remaining non-affine term is only $\rho_2\sigma_\rho\sqrt{\nu_t}$, see (5.79). For its approximation we use the following result [56]:

$$\rho_2\sigma_\rho\sqrt{\nu_t} \approx \rho_2\sigma_\rho\mathbb{E}\left[\sqrt{\nu_t}\right], \tag{5.84}$$

where $\mathbb{E}\left[\sqrt{\nu_t}\right]$ is given in the next proposition.

Proposition 5.3.1 $\mathbb{E}\left[\sqrt{\nu_t}\right]$ *can be approximated by*

$$\mathbb{E}\left[\sqrt{\nu_t}\right] \approx m + ne^{-lt}, \tag{5.85}$$

where

$$m := \sqrt{\mu_\nu - \frac{\sigma_\nu^2}{8\kappa_\nu}}, \quad n := \sqrt{\nu_0} - m, \quad l := -\log\left(n^{-1}\left(\hat{d} - m\right)\right), \tag{5.86}$$

$$\hat{d} := \sqrt{\left(\nu_0 e^{-\kappa_\nu} - \frac{\sigma_\nu^2(1-e^{-\kappa_\nu})}{4\kappa_\nu}\right) + \mu_\nu(1-e^{-\kappa_\nu}) + \frac{\sigma_\nu^2\mu_\nu(1-e^{-\kappa_\nu})^2}{8\kappa_\nu\mu_\nu + 8\kappa_\nu e^{-\kappa_\nu}(\nu_0 - \mu_\nu)}}. \tag{5.87}$$

The detailed derivation and the test of quality of the approximation can be found in [56].

Characteristic Function

We start to derive the CF for the HO model, according to [37]. We first assume that the discounted CF for the HO model is of the following form:

$$\phi_{HO}\left(\mathbf{u}, \mathbf{X_t}, \tau\right) = \mathrm{e}^{-r\tau + A(u,\tau) + B(u,\tau)x_t + C(u,\tau)\rho_t + D(u,\tau)\nu_t}, \tag{5.88}$$

with final conditions $A(u,0) = 0$, $B(u,0) = iu$, $C(u,0) = 0$, $D(u,0) = 0$ and $\tau := T - t$. By substituting (5.88) into (5.80) we obtain the ODEs related to the HO model given in the following lemma.

Lemma 5.3.1 *The functions in* (5.88) $A(u,\tau), B(u,\tau), C(u,\tau)$ *and* $D(u,\tau)$ *for the HO model satisfy the following ODE system:*

$$B'(u,\tau) = 0, \quad B(u,0) = iu, \tag{5.89}$$

$$C'(u,\tau) = \sigma_\nu \mathbb{E}[\nu_t] B(u,\tau) D(u,\tau) - \kappa_\rho C(u,\tau), \quad C(u,0) = 0, \tag{5.90}$$

$$D'(u,\tau) = \frac{1}{2}B^2(u,\tau) + \frac{1}{2}\sigma_\nu^2 D(u,\tau) - \frac{1}{2}B(u,\tau) - \kappa_\nu D(u,\tau), \quad D(u,0) = 0, \tag{5.91}$$

$$A'(u,\tau) = (B(u,\tau) - 1)r + \kappa_\nu\mu_\nu D(u,\tau) + \kappa_\rho\mu_\rho C(u,\tau) \tag{5.92}$$

$$+\frac{1}{2}\sigma_\rho^2 C^2(u,\tau)+\sigma_\rho\rho_2\mathbb{E}[\sqrt{\nu_t}]B(u,\tau)C(u,\tau),\quad A(u,0)=0.$$

Obviously, the discounted CF can be obtained as long as the closed-form solution of the latter ODE system is available.

Lemma 5.3.2 *The solution of the ODE system in Lemma 5.3.1 is given by*

$$B(u,\tau)=iu, \tag{5.93}$$

$$D(u,\tau)=\frac{\kappa_\nu-D_1}{\sigma_\nu^2}\cdot\frac{1-e^{-D_1\tau}}{1-D_2e^{-D_1\tau}}, \tag{5.94}$$

$$A(u,\tau)=H_1(u,\tau)+\alpha H_2(u,\tau)+\beta H_3(u,\tau)+\frac{\sigma_\rho^2}{2}H_4(u,\tau), \tag{5.95}$$

$$\begin{aligned}C(u,\tau)=&\frac{C_1(\mu_\nu-\nu_0)}{\kappa_\nu+\kappa_\rho-l_1}e^{(\kappa_\nu-l_1)\tau-\kappa_\nu T}+\frac{C_1(\nu_0-\mu_\nu)}{\kappa_\nu+\kappa_\rho}e^{\kappa_\nu(\tau-T)}\\&+\frac{C_1\mu_\nu}{\kappa_\rho}-\frac{C_1\mu_\nu}{\kappa_\rho-l_1}e^{-l_1}+C_1C_2e^{-\kappa_\rho\tau},\end{aligned} \tag{5.96}$$

where m, n, and l defined in (5.86) - (5.87) *and*

$$D_1=\sqrt{\kappa_\nu^2+\sigma_\nu^2(u^2+iu)},\quad D_2=\frac{\kappa_\nu-D_1}{\kappa_\nu+D_1},\quad C_1=iu\frac{\kappa_\nu-D_1}{\sigma_\nu^2}, \tag{5.97}$$

$$l_1=-\ln\left(\frac{e^{-D_1}-D_2e^{-D_1}}{1-D_2e^{-D_1}}\right),\quad \alpha=\kappa_\rho\mu_\rho+m\sigma_\rho\rho_2ui,\quad \beta=n\sigma_\rho\rho_2ui, \tag{5.98}$$

$$C_2=\frac{\mu_\nu-\nu_0}{\kappa_\nu+\kappa_\rho-l_1}e^{-\kappa_\nu T}+\frac{\nu_0-\mu_\nu}{\kappa_\nu+\kappa_\rho}e^{-\kappa_\nu T}-\frac{\mu_\nu}{\kappa_\rho}+\frac{1}{\kappa_\rho-l_1}, \tag{5.99}$$

$$H_1(u,\tau)=(iu-1)r\tau+\frac{\kappa_\nu\mu_\nu}{\sigma_\nu^2}\left((\kappa_\nu-D_1)\tau-2\ln\left(\frac{1-D_2e^{-D_1\tau}}{1-D_2}\right)\right), \tag{5.100}$$

$$\begin{aligned}H_2(u,\tau)=&\frac{C_1(\mu_\nu-\nu_0)e^{\kappa_\nu(\tau-T)-l_1\tau}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu-l_1)}+\frac{C_1(\nu_0-\mu_\nu)e^{\kappa_\nu(\tau-T)}}{\kappa_\nu(\kappa_\nu+\kappa_\rho)}+\frac{\mu_\nu\tau C_1}{\kappa_\rho}\\&+\frac{\mu_\nu C_1e^{-l_1\tau}}{(\kappa_\rho-l_1)l_1}-\frac{C_1C_2e^{-\kappa_\rho\tau}}{\kappa_\rho}+H_{2c},\end{aligned} \tag{5.101}$$

$$H_3(u,\tau) = \frac{C_1(\mu_\nu-\nu_0)e^{\tau(\kappa_\nu+l-l_1)-T(\kappa_\nu+l)}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+l-l_1)} + \frac{C_1(\nu_0-\mu_\nu)e^{(\tau-T)(l+\kappa_\nu)}}{(l+\kappa_\nu)(\kappa_\nu+\kappa_\rho)} + \frac{\mu_\nu C_1 e^{l(\tau-T)}}{\kappa_\rho l} - \frac{\mu_\nu C_1 e^{\tau(l-l_1)-lT}}{(\kappa_\rho-l_1)(l-l_1)} + \frac{C_1C_2e^{\tau(l-\kappa_\rho)-lT}}{l-\kappa_\rho} + H_{3c}, \tag{5.102}$$

$$H_{2c} = \frac{C_1(\nu_0-\mu_\nu)e^{-\kappa_\nu T}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu-l_1)} - \frac{C_1(\nu_0-\mu_\nu)e^{-\kappa_\nu T}}{\kappa_\nu(\kappa_\nu+\kappa_\rho)} - \frac{\mu_\nu C_1}{(\kappa_\rho-l_1)l_1} + \frac{C_1C_2}{\kappa_\rho}, \tag{5.103}$$

$$H_{3c} = \frac{C_1(\mu_\nu-\nu_0)e^{-T(\kappa_\nu+l)}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+l-l_1)} + \frac{C_1(\nu_0-\mu_\nu)e^{-T(l+\kappa_\nu)}}{(l+\kappa_\nu)(\kappa_\nu+\kappa_\rho)} + \frac{\mu_\nu C_1 e^{-lT}}{\kappa_\rho l} - \frac{\mu_\nu C_1 e^{-lT}}{(\kappa_\rho-l_1)(l-l_1)} + \frac{C_1C_2e^{-lT}}{l-\kappa_\rho}, \tag{5.104}$$

$$\begin{aligned} H_4(u,\tau) = {} & H_{4c_1}e^{2\kappa_\nu(\tau-T)} + H_{4c_4}e^{-\tau l_1} + H_{4c_5}e^{(-l_1-\kappa_\rho)\tau} + H_{4c_9}e^{\tau(\kappa_\nu-\kappa_\rho)-\kappa_\nu T} \\ & + H_{4c_2}e^{2\tau(\kappa_\nu-l_1)-2\kappa_\nu T} + H_{4c_3}e^{\tau(2\kappa_\nu-l_1)-2\kappa_\nu T} + H_{4c_{11}}e^{\tau(\kappa_\nu-\kappa_\rho-l_1)-\kappa_\nu T} \\ & + H_{4c_{12}}e^{\tau(\kappa_\nu-l_1)-\kappa_\nu T} + H_{4c_{13}}e^{\tau(\kappa_\nu-2l_1)-\kappa_\nu T} + H_{4c_6}e^{\tau(-2\kappa_\rho\tau} + H_{4c_7}e^{-\kappa_\rho\tau} \\ & + H_{4c_8}e^{-2l_1\tau} + H_{4c_{10}}e^{\tau(\kappa_\nu-l_1)-\kappa_\nu T} + H_{4c_{14}}e^{\kappa_\rho(\tau-T)} + \frac{C_1^2\mu_\nu^3\tau}{\kappa_\rho^2} + H_{4c}, \end{aligned} \tag{5.105}$$

$$\begin{aligned} H_{4c} = {} & (H_{4c_1}+H_{4c_2}+H_{4c_3})e^{-2\kappa_\nu T} + H_{4c_4} + H_{4c_5} + H_{4c_6} + H_{4c_7} + H_{4c_8} \\ & + (H_{4c_9}+H_{4c_{10}}+H_{4c_{11}}+H_{4c_{12}}+H_{4c_{13}}+H_{4c_{14}})e^{-\kappa_\nu T}, \end{aligned} \tag{5.106}$$

with

$$H_{4c_1} := \frac{C_1^2(\nu_0-\mu_\nu)^2}{2\kappa_\nu(\kappa_\nu+\kappa_\rho)^2}, \quad H_{4c_2} := \frac{C_1^2(\nu_0-\mu_\nu)^2}{2(2\kappa_\nu+\kappa_\rho-l_1)^2(\kappa_\nu-l_1)}, \tag{5.107}$$

$$H_{4c_3} := \frac{2C_1^2(\nu_0-\mu_\nu)^2}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+\kappa_\rho)(l_1-2\kappa_\nu)}, \quad H_{4c_4} := \frac{2C_1^2\mu_\nu^2}{\kappa_\nu l_1(\kappa_\rho-l_1)}, \tag{5.108}$$

$$H_{4c_5} := \frac{2\mu_\nu C_1^2C_2}{\kappa_\rho^2-l_1^2}, \quad H_{4c_6} := -\frac{1}{2}\frac{C_1^2C_2^2}{\kappa_\rho}, \quad H_{4c_7} := -\frac{2\mu_\nu C_1^2C_2}{\kappa_\rho^2}, \tag{5.109}$$

$$H_{4c_8} := -\frac{1}{2}\frac{\mu_\nu^2C_1^2}{l_1(\kappa_\rho-l_1)^2}, \quad H_{4c_9} := \frac{2(\nu_0-\mu_\nu)C_1^2C_2}{(\kappa_\nu+\kappa_\rho)(\kappa_\nu-\kappa_\rho)}, \tag{5.110}$$

$$H_{4c_{10}} := \frac{2C_1^2(\mu_\nu^2-\nu_0\mu_\nu)}{(\kappa_\nu+\kappa_\rho)(\kappa_\nu-l_1)(\kappa_\rho-l_1)}, \; H_{4c_{11}} := \frac{2(\mu_\nu-\nu_0)C_1^2C_2}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu-\kappa_\rho-l_1)}, \tag{5.111}$$

$$H_{4c_{12}} := \frac{2C_1^2(\mu_\nu^2-\nu_0\mu_\nu)}{\kappa_\rho(\kappa_\nu-l_1)(\kappa_\nu+\kappa_\rho-l_1)}, \; H_{4c_{13}} := \frac{2C_1^2(\nu_0\mu_\nu-\mu_\nu^2)}{(\kappa_\rho-l_1)(\kappa_\nu-2l)(\kappa_\nu+\kappa_\rho-l_1)}. \tag{5.112}$$

$$H_{4c_{14}} := \frac{2C_1^2(\nu_0\mu_\nu - \mu_\nu^2)}{\kappa_\nu\kappa_\rho(\kappa_\nu + \kappa_\rho)^2}, \tag{5.113}$$

The proof can be found in Appendix C.

5.3.3 Incorporating the Bounded Jacobi Process into the Heston Model

In this section we investigate how to incorporate the variant of our general stochastic correlation model, the bounded Jacobi process (see Section 5.1.4), into the Heston model. Let us recall that

$$d\rho_t = \kappa_\rho(\mu_\rho - \rho_t)\,dt + \sigma_\rho\sqrt{1-\rho_t^2}\,d\tilde{W}_t^\rho, \tag{5.114}$$

where the functions $a(t,\rho_t)$ and $b(t,\rho_t)$ defined in (5.78) and (5.79) are $\kappa_\rho(\mu_\rho - \rho_t)$ and $\sigma_\rho\sqrt{1-\rho_t^2}$, respectively. We call this extended Heston model as "HJ" model. Similar to the HO model, from (5.79) we observe that the non-affine terms in the HJ model are $b^2(t,\rho_t)$ and $\rho_2 b(t,\rho_t)\sqrt{\nu_t}$, as

$$b^2(t,\rho_t) = \sigma_\rho^2(1-\rho_t^2), \tag{5.115}$$

$$\rho_2 b(t,\rho_t)\sqrt{\nu_t} = \rho_2\sigma_\rho\sqrt{1-\rho_t^2}\sqrt{\nu_t}. \tag{5.116}$$

Approximation to affinity

We attempt to find appropriate approximations for (5.115) and (5.116) which are affine. We consider first (5.115) which could be approximated by

$$\sigma_\rho^2(1-\mathbb{E}[\rho_t^2]), \tag{5.117}$$

where $\mathbb{E}[\rho_t^2]$) is given by [115]

$$\begin{aligned}\mathbb{E}[\rho_t^2] =& \frac{1}{\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2} \mathrm{e}^{-t(\sigma_\rho^3+2\kappa_\rho)} \Big((\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2)\rho_0^2 \\ &+ 2\mu_\rho\kappa_\rho\rho_0(\sigma_\rho^2 + 2\kappa_\rho)(\mathrm{e}^{t(\sigma_\rho^2+\kappa_\rho)} - 1) + \sigma_\rho^2(\sigma_\rho^2 + \kappa_\rho)(\mathrm{e}^{t(\sigma_\rho^2+2\kappa_\rho)} - 1) \\ &- 2\mu_\rho^2\kappa_\rho\big(\kappa_\rho(2\mathrm{e}^{t(\sigma_\rho^2+\kappa_\rho)} - \mathrm{e}^{t(\sigma_\rho^2+2\kappa_\rho)} - 1) - \sigma_\rho^2\mathrm{e}^{t(\sigma_\rho^2+\kappa_\rho)}(\mathrm{e}^{t\kappa_\rho} - 1)\big)\Big).\end{aligned} \tag{5.118}$$

We see that the latter equation is rather complicated and not convenient for further calculation. Therefore, we introduce the following approximation.

Proposition 5.3.2 $\mathbb{E}[\rho_t^2]$ *can be approximated by*

$$f_2(t) := \mathbb{E}[\rho_t^2] \approx e^{-m_2 t} + b_2 e^{-n_2 t} + a_2, \tag{5.119}$$

where

$$a_2 = \frac{(\sigma_\rho^2 + \kappa_\rho)(\sigma_\rho^2 + 2\kappa_\rho\mu_\rho^2)}{\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2}, \quad b_2 = \rho_0^2 - a_2 - 1, \tag{5.120}$$

$$m_2 = -2\log\left(\gamma_1 - b_2 e^{-\frac{n_2}{2}}\right), \quad n_2 = -2\log\left(\frac{b_2\gamma_1 - \sqrt{b_1^2\gamma_1^2 - \gamma_2\gamma_3}}{\gamma_2}\right), \tag{5.121}$$

with

$$\gamma_1 := f_2(0.5) - a_2, \quad \gamma_2 := b_2 + b_1^2, \quad \gamma_3 := \gamma_1^2 + a_2 - f_2(1). \tag{5.122}$$

The proof and the test of quality of the approximation can be found in Appendix C.

Next, we investigate the approximation for the other non-affine term (5.116). Due to $\rho_1 = 0$ we propose to approximate (5.116) using its expectation

$$\rho_2\sigma_\rho\sqrt{\nu_t} \approx \rho_2\sigma_\rho\mathbb{E}[\sqrt{1-\rho_t^2}]\mathbb{E}[\sqrt{\nu_t}]. \tag{5.123}$$

$\mathbb{E}[\sqrt{\nu_t}]$ is already known, see Prop.5.3.1. The remaining task is to find a formula for $\mathbb{E}[\sqrt{1-\rho_t^2}]$, for this we apply the delta method which has also been used to find the approximation for $\mathbb{E}[\sqrt{\nu_t}]$ in [56]. Say that $\psi(X)$ is sufficiently smooth where the first

two moments of X exist, then with the aid of a Taylor expansion we have

$$\psi(X) \approx \psi(\mathbb{E}[X]) + (X - \mathbb{E}[X])\frac{\partial \psi}{\partial X}\mathbb{E}[X], \tag{5.124}$$

such that the variance of $\psi(X)$ can be given by

$$\mathbb{V}[\psi(X)] \approx \mathbb{V}\left[\psi(\mathbb{E}[X]) + (X - \mathbb{E}[X])\frac{\partial \psi}{\partial X}\mathbb{E}[X]\right] = \left(\frac{\partial \psi}{\partial X}\mathbb{E}[X]\right)^2 \mathbb{V}[X]. \tag{5.125}$$

On one hand, setting $\psi(\rho_t) = \sqrt{1-\rho_t^2}$ we obtain

$$\mathbb{V}\left[\sqrt{1-\rho_t^2}\right] = \frac{\mathbb{E}[\rho_t]^2}{1-\mathbb{E}[\rho_t]^2}\mathbb{V}[\rho_t]. \tag{5.126}$$

On the other hand, from the definition of the variance we also get

$$\mathbb{V}\left[\sqrt{1-\rho_t^2}\right] = \mathbb{E}[1-\rho_t^2] - \mathbb{E}\left[\sqrt{1-\rho_t^2}\right]^2. \tag{5.127}$$

Directly from the last two equations we obtain finally

$$\mathbb{E}\left[\sqrt{1-\rho_t^2}\right] = \sqrt{\mathbb{E}[1-\rho_t^2] - \frac{\mathbb{E}[\rho_t]^2}{1-\mathbb{E}[\rho_t]^2}\mathbb{V}[\rho_t]} = \sqrt{1 - \frac{\mathbb{E}[\rho_t^2] - \mathbb{E}[\rho_t]^4}{1-\mathbb{E}[\rho_t]^2}}, \tag{5.128}$$

where $\mathbb{E}[\rho_t^2]$ is given in (5.118) and its approximation in (5.119). Besides, we know $\mathbb{E}[\rho_t] = \mu_\rho + (\rho_0 - \mu_\rho)\mathrm{e}^{-\kappa_\rho t}$ for the correlation process ρ_t defined in (5.114). In the same way as above we try to find a suitable approximation for (5.128) which has a more convenient form.

Proposition 5.3.3 $\mathbb{E}[\sqrt{1-\rho_t^2}]$ *can be approximated by*

$$f_3(t) := \mathbb{E}[\sqrt{1-\rho_t^2}] \approx e^{-m_3 t} + b_3 e^{-n_3 t} + a_3, \tag{5.129}$$

where

$$a_3 = \sqrt{1 - \frac{(\sigma_\rho^2 + \kappa_\rho)(\sigma_\rho^2 + 2\kappa_\rho\mu_\rho^2) - \mu_\rho^4(\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2)}{(1-\mu_\rho^2)(\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2)}}, \tag{5.130}$$

$$b_3 = \sqrt{1-\rho_0^2} - a_3 - 1, \tag{5.131}$$

$$m_3 = -2\log\left(\eta_1 - b_3 e^{-\frac{n_3}{2}}\right),\ n_3 = -2\log\left(\frac{b_3\eta_1 - \sqrt{b_3^2\eta_1^2 - \eta_2\eta_3}}{\eta_2}\right), \tag{5.132}$$

with

$$\eta_1 := f_3(0.5) - a_3, \quad \eta_2 := b_3 + b_3^2, \quad \eta_3 := \eta_1^2 + a_3 - f_3(1). \tag{5.133}$$

We show the proof and measure the quality of the approximation in Appendix C.

Characteristic function

Again, we assume the discounted CF for the HJ model with the form:

$$\phi_{HJ}\left(\mathbf{u}, \mathbf{X_t}, \tau\right) = \mathrm{e}^{-r\tau + \tilde{A}(u,\tau) + \tilde{B}(u,\tau)x_t + \tilde{C}(u,\tau)\rho_t + \tilde{D}(u,\tau)\nu_t} \tag{5.134}$$

with final conditions $\tilde{A}(u,0) = 0$, $\tilde{B}(u,0) = iu$, $\tilde{C}(u,0) = 0$, $\tilde{D}(u,0) = 0$ and $\tau := T - t$. By substituting (5.134) into (5.80) we can obtain a similar ODE system as in Lemma 5.3.1.

Lemma 5.3.3 *The functions in* (5.134) $\tilde{A}(u,\tau), \tilde{B}(u,\tau), \tilde{C}(u,\tau)$ *and* $\tilde{D}(u,\tau)$ *for the HJ model satisfy the following ODE system:*

$$\tilde{B}'(u,\tau) = 0, \quad \tilde{B}(u,0) = iu, \tag{5.135}$$

$$\tilde{C}'(u,\tau) = \sigma_\nu \mathbb{E}[\nu_t]\tilde{B}(u,\tau)\tilde{D}(u,\tau) - \kappa_\rho\tilde{C}(u,\tau), \quad \tilde{C}(u,0) = 0, \tag{5.136}$$

$$\tilde{D}'(u,\tau) = \frac{1}{2}\tilde{B}^2(u,\tau) + \frac{1}{2}\sigma_\nu^2\tilde{D}(u,\tau) - \frac{1}{2}\tilde{B}(u,\tau) - \kappa_\nu\tilde{D}(u,\tau), \quad \tilde{D}(u,0) = 0, \tag{5.137}$$

$$\begin{aligned}\tilde{A}'(u,\tau) = {} & (\tilde{B}(u,\tau) - 1)r + \kappa_\nu\mu_\nu\tilde{D}(u,\tau) + \frac{1}{2}\sigma_\rho^2\mathbb{E}[1-\rho_t^2]\tilde{C}^2(u,\tau) \\ & + \kappa_\rho\mu_\rho\tilde{C}(u,\tau) + \sigma_\rho\rho_2\mathbb{E}[\sqrt{\nu_t}]\mathbb{E}[\sqrt{1-\rho_t^2}]\tilde{B}(u,\tau)\tilde{C}(u,\tau), \quad A(u,0) = 0.\end{aligned} \tag{5.138}$$

We observe that there is only a difference between the ODEs in Lemma 5.3.1 and 5.3.3 in $A(u,\tau)$ because of the distinct correlation processes used. This also means that the solutions of $\tilde{B}(u,\tau)$, $\tilde{C}(u,\tau)$ and $\tilde{D}(u,\tau)$ coincide with $B(u,\tau)$, $C(u,\tau)$ and $D(u,\tau)$ in the HO model. Therefore we only need to calculate (5.138) to gain the discounted CF for the HJ model (5.134). We state our result in the following lemma.

Lemma 5.3.4 *The solutions of* $\tilde{B}(u,\tau)$, $\tilde{C}(u,\tau)$, $\tilde{D}(u,\tau)$ *are respectively equal to* (5.93), (5.96), (5.94), *and*

$$\begin{aligned} A(u,\tau) &= \tilde{H}_1(u,\tau) + (\kappa_\rho\mu_\rho + a_3 m\zeta)\tilde{H}_2(u,\tau) + a_3 n\zeta\tilde{H}_3(u,\tau,l) \\ &\quad + b_3 m\zeta\tilde{H}_3(u,\tau,n_3) + m\zeta\tilde{H}_3(u,\tau,m_3) + b_3 n\zeta\tilde{H}_3(u,\tau,(l+n_3)) \\ &\quad + n\zeta\tilde{H}_3(u,\tau,(l+m_3)) + \frac{\sigma_\rho^2}{2}(1-a_2)\tilde{H}_4(u,\tau) - \frac{\sigma_\rho^2}{2}\tilde{H}(u,\tau,m_2) \\ &\quad - \frac{b_2\sigma_\rho^2}{2}\tilde{H}(u,\tau,n_2), \end{aligned} \tag{5.139}$$

where $\zeta = \sigma_\rho\rho_2 ui$, $\tilde{H}_1(u,\tau) = H_1(u,\tau)$ (5.100), $\tilde{H}_2(u,\tau) = H_2(u,\tau)$ (5.101), $\tilde{H}_4(u,\tau) = H_4(u,\tau)$ (5.105),

$$\begin{aligned} \tilde{H}_3(u,\tau,y) &= \frac{C_1(\mu_\nu-\nu_0)e^{\tau(\kappa_\nu+y-l_1)-T(\kappa_\nu+y)}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+y-l_1)} + \frac{C_1(\nu_0-\mu_\nu)e^{(\tau-T)(y+\kappa_\nu)}}{(y+\kappa_\nu)(\kappa_\nu+\kappa_\rho)} \\ &\quad + \frac{\mu_\nu C_1 e^{y(\tau-T)}}{\kappa_\rho y} - \frac{\mu_\nu C_1 e^{\tau(y-l_1)-yT}}{(\kappa_\rho-l_1)(y-l_1)} + \frac{C_1C_2 e^{\tau(y-\kappa_\rho)-yT}}{y-\kappa_\rho} + H_{3c}, \end{aligned} \tag{5.140}$$

$$\begin{aligned} \tilde{H}_{3c} &= \frac{C_1(\mu_\nu-\nu_0)e^{-T(\kappa_\nu+y)}}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+y-l_1)} + \frac{C_1(\nu_0-\mu_\nu)e^{-T(y+\kappa_\nu)}}{(y+\kappa_\nu)(\kappa_\nu+\kappa_\rho)} + \frac{\mu_\nu C_1 e^{-yT}}{\kappa_\rho y} \\ &\quad - \frac{\mu_\nu C_1 e^{-yT}}{(\kappa_\rho-l_1)(y-l_1)} + \frac{C_1C_2 e^{-yT}}{y-\kappa_\rho}, \end{aligned} \tag{5.141}$$

$$
\begin{aligned}
\tilde{H}(u,\tau,y) =& I_1 e^{(y+2\kappa_\nu-l_1)\tau-(y+2\kappa_\nu)T} + I_2 e^{(y+\kappa_\nu-l_1)\tau-(y+\kappa_\nu)T} \\
&+ I_3 e^{(y+\kappa_\nu-2l_1)\tau-(y+\kappa_\nu)T} + I_4 e^{(y+\kappa_\nu)(\tau-T)} \\
&+ I_5 e^{(y+\kappa_\nu-l_1)\tau-(y+\kappa_\nu)T} + I_6 e^{(y-2l_1)\tau-yT} \\
&+ I_7 e^{(y+2\kappa_\nu-2l_1)\tau-(y+2\kappa_\nu)T} + I_8 e^{(y+2\kappa_\nu)(\tau-T)} \\
&+ I_9 e^{(\kappa_\nu-\kappa_\rho+y-l_1)\tau-(y+\kappa_\nu)T} + I_{10} e^{(\kappa_\nu-\kappa_\rho+y)\tau-(y+\kappa_\nu)T} \\
&+ I_{11} e^{(y-2\kappa_\rho)\tau-yT} + I_{12} e^{(y-\kappa_\rho)\tau-yT} + I_{13} e^{y(\tau-T)} \\
&+ I_{14} e^{(y-l_1)\tau-yT} + I_{15} e^{(y-l_1-\kappa_\rho)\tau-yT} + \tilde{H}_c,
\end{aligned}
\tag{5.142}
$$

$$
\begin{aligned}
\tilde{H}_c = (I_1 + I_7 + I_8) e^{-(y+2\kappa_\nu)T} + (I_2 + I_3 + I_4 + I_5 + I_9 + I_{10}) e^{-(y+\kappa_\nu)T} \\
+ (I_6 + I_{11} + I_{12} + I_{13} + I_{14} + I_{15}) e^{-yT},
\end{aligned}
\tag{5.143}
$$

with

$$
\begin{aligned}
I_1 &= \frac{-2C_1^2(\nu_0-\mu_\nu)^2}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu+\kappa_\rho)(y+2\kappa_\nu-l_1)}, \quad I_2 = \frac{2C_1^2(\mu_\nu^2-\mu_\nu\nu_0)}{\kappa_\rho(\kappa_\nu+\kappa_\rho-l_1)(y+\kappa_\nu-l_1)}, \\
I_3 &= \frac{2C_1^2(\mu_\nu\nu_0-\mu_\nu^2)}{(\kappa_\rho-l_1)(\kappa_\nu+\kappa_\rho-l_1)(y+\kappa_\nu-2l_1)}, \quad I_4 = \frac{2C_1^2(\mu_\nu\nu_0-\mu_\nu^2)}{\kappa_\rho(\kappa_\rho+\kappa_\nu)(\kappa_\nu+y)}, \\
I_5 &= \frac{2C_1^2(\mu_\nu^2-\mu_\nu\nu_0)}{(\kappa_\rho-l_1)(\kappa_\nu+\kappa_\rho)(y+\kappa_\nu-l_1)}, \quad I_6 = \frac{C_1^2\mu_\nu^2}{(\kappa_\rho-l_1)^2(y-2l_1)}, \\
I_7 &= \frac{C_1^2(\nu_0-\mu_\nu)^2}{(\kappa_\nu+\kappa_\rho-l_1)^2(y+2\kappa_\nu-2l_1)}, \quad I_8 = \frac{C_1^2(\nu_0+\mu_\nu)^2}{(\kappa_\rho+\kappa_\nu)^2(2\kappa_\nu+y)}, \\
I_9 &= \frac{2C_1^2C_2(\mu_\nu-\nu_0)}{(\kappa_\nu+\kappa_\rho-l_1)(\kappa_\nu-\kappa_\rho+y-l_1)}, \quad I_{10} = \frac{2C_1^2C_2(\nu_0-\mu_\nu)}{(\kappa_\nu+\kappa_\rho)(\kappa_\nu-\kappa_\rho+y)}, \\
I_{11} &= \frac{C_1^2C_2^2}{y-2\kappa_\rho}, \quad I_{12} = \frac{2C_1^2C_2\mu_\nu}{y\kappa_\rho-\kappa_\rho^2}, \quad I_{13} = \frac{C_1^2\mu_\nu^2}{y\kappa_\rho^2}, \\
I_{14} &= \frac{-2C_1^2\mu_\nu^2}{\kappa_\rho(y\kappa_\rho-l_1\kappa_\rho-yl_1+l_1^2)}, \quad I_{15} = \frac{-2C_1^2C_2\mu_\nu}{(\kappa_\rho-l_1)(y-\kappa_\rho-l_1)}.
\end{aligned}
$$

C_1, l_1 and C_2 are respectively located in (5.97), (5.98) *and* (5.99).

The proof can be found in Appendix C.

5.3.4 Simulation of the Heston Model under Stochastic Correlation

We have now obtained the explicit CFs for both models. One can thus do fast pricing by inverting the CFs directly using numerical integration routines, e.g., see [27, 30] for Fourier methods and [47] for COS method. We also refer the readers to [65] for a detailed explanation for the both methods. However, in order to justify the proposed approximations of non-affine terms we want to compare the implied volatilities for the extended Heston model (HO and HJ) to the volatilities implied by performing a Monte-Carlo simulation as the benchmark. In this section, we study how to simulate the paths in these models.

Quadratic-exponential scheme

Basically, we will adopt the approach by Andersen [2] to simulate the HO and HJ model. Firstly, to discretize the stochastic variance ν_t we just employ the quadratic-exponential (QE) scheme. The concrete degrees of freedom and non-centrality parameter for the variance process in (5.71) are given by the following proposition.

Proposition 5.3.4 *For the variance process $d\nu_t$ in* (5.71) *we specify*

$$d = \frac{4\kappa_\nu \mu_\nu}{\sigma_\nu^2}, \qquad \lambda(t,T) = \frac{4\kappa_\nu e^{-\kappa_\nu (T-t)}}{\sigma_\nu^2 (1 - e^{-\kappa_\nu (T-t)})} \quad \textit{for } t < T. \tag{5.144}$$

Conditional on a value ν_t, ν_T is distributed as $\frac{e^{-\kappa_\nu (T-t)}}{\lambda(t,T)}$ times a non-central chi-squared distribution (see B.8) with d degrees of freedom and non-centrality parameter $\nu_t \lambda(t,T)$; this means

$$P(\nu_T < x | \nu_t) = F_{\chi^2}\left(\frac{x\lambda(t,T)}{e^{-\kappa_\nu (T-t)}}; d, \nu_t \lambda(t,T) \right). \tag{5.145}$$

From Proposition 5.3.4 we know that the non-centrality parameter is proportional to ν_t, which means that large or small values of the parameter correspond to large or small values of ν_t. Let $\hat{\nu}_t$ denote the discrete-time approximation to ν_t, for sufficiently large

realized values of $\hat{\nu}_t$, cf. [2]. Then one can approximate the non-central chi-square random variable by the power function

$$\hat{\nu}_{t+\Delta} = \bar{\alpha}(\bar{\beta} + Z_\nu)^2, \tag{5.146}$$

where Z_ν is a standard Gaussian random variable, $\bar{\alpha}$ and $\bar{\beta}$ are certain constants which can be determined by the moment-matching using the parameters of $d\nu_t$ in (5.71). We only state the formulas for $\bar{\alpha}$ and $\bar{\beta}$ in the following Proposition 5.3.5; the detailed calculations can be found in [2].

Proposition 5.3.5 *The mean and the variance of the variance process in* (5.71) *read*

$$\begin{aligned} m &= E[\nu_{t+\Delta}|\nu_t] = \mu_\nu + (\nu_t - \mu_\nu)e^{-\kappa_\nu(T-t)}, \\ s^2 &= \frac{\nu_t\sigma_\nu^2 e^{-\kappa_\nu(T-t)}}{\kappa_\nu}\left(1 - e^{-\kappa_\nu(T-t)}\right) + \frac{\mu_\nu\sigma_\nu^2}{2\kappa_\nu}\left(1 - e^{-\kappa_\nu(T-t)}\right)^2. \end{aligned}$$

If we set $\psi := \frac{s^2}{m^2}$ *and choose*

$$\bar{\beta}^2 = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1} - 1} \geq 0 \tag{5.147}$$

and

$$\bar{\alpha} = \frac{m}{1 + \bar{\beta}^2}, \tag{5.148}$$

then (5.146) *has a mean equal to* m *and a variance equal to* s^2. *Note that* $\psi \leq 2$.

However, the approximation (5.146) will not work properly for small values of $\hat{\nu}_t$, cf. [2]. For small values of $\hat{\nu}_t$, Andersen [2] suggested to use instead an approximated density for $\hat{\nu}_{t+\Delta}$ of the form:

$$P(\hat{\nu}_{t+\Delta} \in [x, x + dx]) \approx (p\delta(0) + q(1-p)\mathrm{e}^{-qx})dx, \quad x \geq 0, \tag{5.149}$$

where δ is a Dirac delta function, and $p \in [0,1]$ and $q > 0$ are constants which can be determined by the moment-matching. Next, we integrate (5.149)

$$\Psi(x) = P(\hat{\nu}_{t+\Delta} < x) = p + (1-p)(1 - \mathrm{e}^{-qx}), \quad x \geq 0 \tag{5.150}$$

and by inverting we obtain

$$\Psi^{-1}(u) = \Psi^{-1}(u; p, q) = \begin{cases} 0, & 0 \leq u \leq p \\ q^{-1} \ln(\frac{1-p}{1-u}), & p < u \leq 1 \end{cases}. \tag{5.151}$$

Thus, the sampling scheme for small values of $\hat{\nu}_t$ reads

$$\hat{\nu}_{t+\Delta} = \Psi^{-1}(U_\nu; p, q), \tag{5.152}$$

where U_ν is a uniform random variable. Again, we state the formulas for p and q in the following proposition; for the detailed calculations we refer to [2].

Proposition 5.3.6 *Let m, s^2 and ψ be defined as in Proposition 5.3.5. For $\psi \geq 1$ we choose*

$$p = \frac{\psi - 1}{\psi + 1} \in [0, 1) \tag{5.153}$$

and

$$q = \frac{1-p}{m} = \frac{2}{m(\psi+1)} > 0, \tag{5.154}$$

such that (5.152) *has a mean equal to m and a variance equal to s^2.*

We only need to select an arbitrary level $\psi_c \in [1, 2]$ and choose either (5.146) or (5.152) according to $\psi \leq \psi_c$ or $\psi > \psi_c$ to do the sampling for the variance process in (5.71).

Discretization for the log-price process

Next, we discuss the discretization for the log-price process. We remark: As indicated by Andersen [2], a straight discretization of dx_t in (5.71) may lead to the problem of

"leaking correlation". Suppose that we use an Euler scheme for simulating dx_t in (5.71)

$$\hat{x}_{t+\Delta} = \hat{x}_t + (r - \frac{1}{2}\hat{\nu}_t)\Delta + \sqrt{\hat{\nu}_t} Z_x \sqrt{\Delta}. \tag{5.155}$$

We know that the true correlation between $\hat{x}_{t+\Delta}$ and $\hat{\nu}_{t+\Delta}$ is always close to ρ_t given by $d\nu_t$ in (5.71). However, $\hat{\nu}_{t+\Delta}$ and Z_ν in (5.146) have a strong nonlinear relationship which will imply that the effective correlation between $\hat{\nu}_{t+\Delta}$ and $\hat{x}_{t+\Delta}$ will be closer to zero than ρ_t for the cases where $P(\bar{\beta} + Z_v < 0)$ is significant. The problem of "leaking correlation" can be tackled by using dx_t in (5.76) or in (5.78).

In the sequel, we try to follow the methodology suggested by Andersen [2] to discretize the log-price process in (5.78): The integral form of the stochastic variance in (5.71) reads

$$\nu_{t+\Delta} = \nu_t + \int_t^{t+\Delta} \kappa_\nu(\mu_\nu - \nu_u)du + \sigma_\nu \int_t^{t+\Delta} \sqrt{\nu_u} dW_u^\nu, \tag{5.156}$$

which can be rearranged as

$$\int_t^{t+\Delta} \sqrt{\nu_u} dW_u^\nu = \sigma_\nu^{-1} \left(\nu_{t+\Delta} - \nu_t - \kappa_\nu \mu_\nu \Delta + \kappa_\nu \int_t^{t+\Delta} \nu_u du \right). \tag{5.157}$$

Using the Cholesky decomposition, dx_t in (5.71) can be reformulated as

$$x_{t+\Delta} = x_t + r\Delta - \frac{1}{2}\int_t^{t+\Delta} \nu_u du + \rho_{x\nu} \int_t^{t+\Delta} \sqrt{\nu_u} dW_u^\nu + \sqrt{1-\rho_{x\nu}^2} \int_t^{t+\Delta} \sqrt{\nu_u} dW_u^x. \tag{5.158}$$

Now we substitute (5.157) into (5.158) to obtain

$$\begin{aligned} x_{t+\Delta} = x_t + r\Delta &+ \frac{\rho_{x\nu}}{\sigma_\nu} \left(\nu_{t+\Delta} - \nu_t - \kappa_\nu \mu_\nu \Delta \right) \\ &+ \left(\frac{\kappa_\nu \rho_{x\nu}}{\sigma_\nu} - \frac{1}{2} \right) \int_t^{t+\Delta} \nu_u du + \sqrt{1-\rho_{x\nu}^2} \int_t^{t+\Delta} \sqrt{\nu_u} dW_u^x. \end{aligned} \tag{5.159}$$

Andersen [2] proposed to discretize the integral in (5.159) using the approximation

$$\int_t^{t+\Delta} \nu_u du \approx \Delta \left(\gamma_1 \nu_t + \gamma_2 \nu_{t+\Delta}\right), \tag{5.160}$$

where γ_1 and γ_2 are given constants. For example, one could set $\gamma_1 = \gamma_2 = \frac{1}{2}$ for a trapezoidal quadrature. Thus, the other Itô integral can be approximated by

$$\int_t^{t+\Delta} \sqrt{\nu_u} dW_u^x \approx \sqrt{\Delta}\sqrt{\gamma_1 \nu_t + \gamma_2 \nu_{t+\Delta}} Z_\nu, \tag{5.161}$$

where Z_ν is a standard Gaussian random variable. However, using the approximations and discretization above, (5.159) will not be a martingale, while it must be under the risk-neutral measure. For this problem, on the one hand one can reduce the size of Δ, on the other hand the "martingale correction" proposed by Andersen [2] can be employed.

Now we come back to consider the asset process in (5.78) in the extended Heston model. In principle, one can also apply the discretization described above for dx_t in (5.78). However, due to the incorporated stochastic correlation and its more complicated structure, the approximation would not be satisfying as in the original Heston model. Besides, the corresponding martingale correction for dx_t in (5.78) will be tedious. In light of this problem, we suggest to use directly the Euler or Milstein scheme . The reason is, firstly, in dx_t in (5.78) all the correlation terms have been considered, such terms will be kept by using the Euler ([42, 82]) or Milstein ([87]) scheme. Secondly, the discretized process (5.162) by Euler or Milstein scheme will be a martingale. For a detailed description on the analysis and application of Euler and Milstein scheme we refer to [57, 101].

The discretization of dx_t in (5.78) by applying the Euler scheme reads

$$\begin{aligned} \hat{x}_{t+\Delta} = \hat{x}_t + (r - \frac{1}{2}\hat{\nu}_t)\Delta + \rho_2\sqrt{\Delta} Z_{\hat{\rho}}\sqrt{\hat{\nu}_t} \\ + \hat{\rho}_t\sqrt{\Delta} Z_{\hat{\nu}}\sqrt{\hat{\nu}_t} + \sqrt{1-\rho_2^2-\hat{\rho}_t^2}\sqrt{\Delta} Z_{\hat{x}}\sqrt{\hat{\nu}_t}, \end{aligned} \tag{5.162}$$

where $Z_{\hat{\rho}}, Z_{\hat{\nu}}$ and $Z_{\hat{x}}$ are independent standard Gaussian random variables. The discretization of dx_t in (5.78) by applying the Milstein scheme will be the same as (5.162), since all the derivatives included in the coefficients of the double integral terms (with respect to BMs) of the Milstein scheme are equal to zero.

Discretization for SCPs

We consider now the suitable discretization schemes for some SCPs. The OU process (see (5.83)) can be simulated in several ways, e.g., using its exact solution

$$\rho_{t+\Delta} = \rho_t \mathrm{e}^{-\kappa_\rho \Delta} + \mu_\rho (1 - \mathrm{e}^{-\kappa_\rho \Delta}) + \sigma_\rho \sqrt{\frac{1 - \mathrm{e}^{-2\kappa_\rho \Delta}}{2\kappa_\rho}} Z_\rho. \tag{5.163}$$

As mentioned before, the major drawback of using an OU process for stochastic correlation is that the process is not bounded. Hence, for the simulation we limit the mean value μ_ρ to be in $(-1, 1)$ and choose a relative large value of κ_ρ, a small value of σ_ρ.

To discretize other SCPs with unknown exact solution, one can either directly use the Euler or the Milstein scheme. For the discretization of (5.114) we remark that the implicit Milstein schemes should be preferred for a better convergence and preserving the boundaries, cf. [115]. The implicit Milstein scheme for (5.114) reads

$$\rho_{t+\Delta} = \rho_t + a\left(t + \Delta, \rho_{t+\Delta}\right) \Delta + b(t, \rho_t) \sqrt{\Delta} Z_\rho + \frac{1}{2} b(t, \rho_t) \frac{\partial}{\partial \rho} b(t, \rho_t) \Delta \left(Z_\rho^2 - 1\right). \tag{5.164}$$

A comparison of the numerical methods AM vs. EM

To verify the approaches introduced above to simulate the Heston model extended by stochastic correlation, we conduct the following example: considering a European Call-option whose numerically approximated price is

$$\hat{C} = \mathbb{E}\left[(\hat{S}_T - K)^+\right] = \mathbb{E}\left[(\mathrm{e}^{\hat{x}_T} - K)^+\right], \tag{5.165}$$

which can be approximated by a Monte-Carlo method

$$\hat{C} \approx \frac{1}{M} \sum_{i=1}^{M} \left(\mathrm{e}^{\hat{x}_T^i} - K \right)^+. \tag{5.166}$$

For $\hat{x}_T^i$ in the latter equation, on the one hand we simply discretize (5.158) by using the Euler or Milstein scheme, denoted "EM". On the other hand we choose the approach by Andersen [2], namely using (5.158)-(5.161) with a martingale correction, denoted by "AM". We denote the exact option price in the original Heston model with C obtained by computing the (semi-)analytical pricing formula in [58] and define the error of a discretization scheme as

$$\epsilon = |C - \hat{C}|, \tag{5.167}$$

which will be dependent on Δ. For all the numerical experiments we assume $S = 120, r = 1\%$ and three different levels of the strike $K = [80,\ 120,\ 160]$.

To initialize the variance process, we choose $\nu_0 = \mu_\nu = 0.04, \kappa_\nu = 0.6, \sigma_\nu = 1$, which do not obey the Feller condition $2\kappa_\nu\mu_\nu > \sigma_\nu^2$, set the constant correlation $\rho^{x\nu}$ to be -0.8, maturity $T = 10$ years. Let $\gamma_1 = \gamma_2$ in (5.160) to be 0.5. We use $M = 10^6$ for the Monte-Carlo method and report the errors in Table 5.1 by varying the values of Δ from 1/32 year to 1 year. We find that the discretization scheme AM has an advantage over EM when the Feller condition is not satisfied for the variance process. The advantage is considerable for the out-of-money options with $K = 160$. Next, we initialize the variance process using $\nu_0 = \mu_\nu = 0.04,\ \kappa_\nu = 2.6$ and $\sigma_\nu = 0.2$ which obviously fulfill the Feller condition, the other parameters are kept to be same as before. By analyzing the relative errors displayed in Table 5.2, we realize that there is no obvious difference between using AM and EM. From the results in Table 5.1 and 5.2 we conclude that the discretization of the dx_t process (with correlation terms) using the Euler or Milstein scheme is sufficiently accurate. In particular, it can be simply applied to the case of using stochastic correlation. In a word, for the extended Heston by imposing a stochastic correlation,

Δ	AM	EM
	$K = 80$	
1	0.136 (0.035)	0.614 (0.061)
1/2	0.057 (0.035)	0.414 (0.049)
1/4	0.022 (0.036)	0.313 (0.043)
1/8	0.022 (0.036)	0.236 (0.039)
1/16	0.009 (0.036)	0.163 (0.038)
1/32	0.008 (0.035)	0.072 (0.037)
	$K = 120$	
1	0.290 (0.024)	0.385 (0.056)
1/2	0.163 (0.025)	0.159 (0.043)
1/4	0.055 (0.026)	0.050 (0.035)
1/8	0.012 (0.026)	0.049 (0.031)
1/16	0.019 (0.026)	0.063 (0.029)
1/32	0.015 (0.026)	0.045 (0.027)
	$K = 160$	
1	0.780 (0.014)	5.465 (0.049)
1/2	0.310 (0.014)	4.260 (0.034)
1/4	0.067 (0.015)	2.938 (0.026)
1/8	0.004 (0.015)	1.876 (0.021)
1/16	0.027 (0.015)	1.140 (0.018)
1/32	0.007 (0.014)	0.662 (0.017)

Table 5.1: A comparison of the relative errors using AM and EM when the parameters of variance process do not fulfill the Feller condition, numbers in parentheses are standard deviations.

we suggest to use the QE scheme for ν_t and the Euler (or Milstein) scheme for x_t which contains stochastic correlation ρ_t using a Cholesky decomposition.

The role of stochastic correlation

We start to analyze the effect of imposing stochastic correlation on the implied volatilities. To illustrate clearly the role of using a correlation process in implied volatility, namely to see how the values of parameters of the correlation process will drive the implied volatilities, we display in Figure 5.12 the changes of the implied volatilities by varying each parameter of the correlation process. For this experiment, we prefer to use the SCP introduced in Section 5.1.3, namely transformed OU process for stochastic

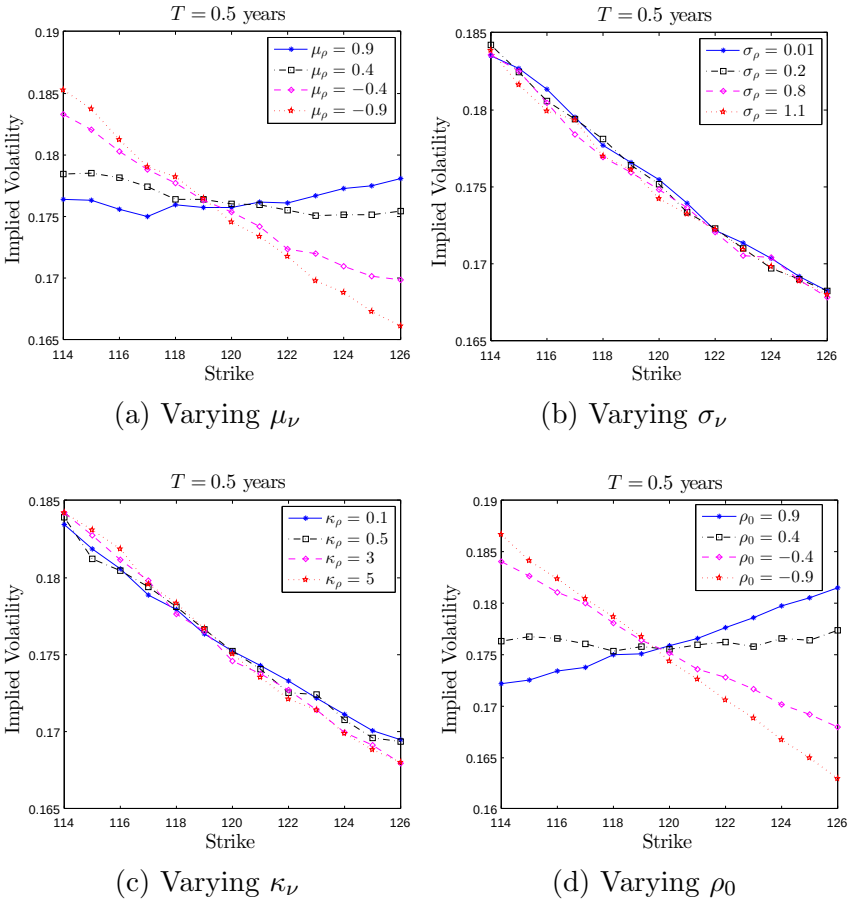
Δ	AM	EM
	$K = 80$	
1	0.031 (0.073)	0.047 (0.072)
1/2	0.042 (0.072)	0.117 (0.072)
1/4	0.065 (0.072)	0.011 (0.072)
1/8	0.044 (0.072)	0.044 (0.072)
1/16	0.021 (0.072)	0.057 (0.072)
1/32	0.088 (0.072)	0.023 (0.072)
	$K = 120$	
1	0.110 (0.063)	0.039 (0.062)
1/2	0.226 (0.062)	0.067 (0.062)
1/4	0.106 (0.062)	0.059 (0.062)
1/8	0.036 (0.063)	0.010 (0.062)
1/16	0.021 (0.063)	0.025 (0.063)
1/32	0.040 (0.063)	0.037 (0.063)
	$K = 160$	
1	0.281 (0.053)	0.178 (0.051)
1/2	0.143 (0.052)	0.038 (0.052)
1/4	0.121 (0.052)	0.075 (0.052)
1/8	0.038 (0.052)	0.027 (0.052)
1/16	0.028 (0.052)	0.032 (0.052)
1/32	0.028 (0.052)	0.020 (0.052)

Table 5.2: A comparison of the relative errors using AM and EM when the parameters of variance process do not fulfill the Feller condition, numbers in parentheses are standard deviations.

correlation. We will use the same parameters as above except for the one who is varying, and choose $T = 0.5$ year. The examples of using other SCPs can be found in [111].

5.3.5 Approximation Error

In this section, we conduct some numerical experiments to justify the proposed approximations of non-affine terms. We compare the implied volatilities for the extended Heston model (HO and HJ) to the volatilities implied by performing a Monte-Carlo simulation as the benchmark. We define the approximation error as the absolute difference between them. For a Monte-Carlo simulation of the extended Heston with stochastic correlation we use the method introduced in Section 5.3.4. For the Monte-Carlo simulation using



(a) Varying μ_ν (b) Varying σ_ν

(c) Varying κ_ν (d) Varying ρ_0

Figure 5.12: Comparison of implied volatilities for varying each parameter of stochastic correlation processes separately.

the OU process, in order to ensure that the generated correlations lie in the interval $(-1, 1)$, as mentioned before, we choose values of μ_ρ and ρ_0 from $(-1, 1)$ and a large value of κ_ρ, a small value of σ_ρ. For using the bounded Jacobi process we only need to take care of the condition (5.42).

We consider a Call-option ($S_0 = 100$) for the maturity of 5 years and present our results in Table 5.3 and 5.4, where $20T$ steps and 10^5 paths are used for the Monte-Carlo simulation; the implied volatilities and errors are expressed in percentage. We consider first Table 5.3 where σ_ρ is set to be 0.1. From the values of the error we see that the approximations in both models give highly accurate results. Besides, we observe that the values of implied volatilities are the same for the HO and HJ model; there is no significant difference by varying ρ_2. This observation can be explained as follows: The

Model		HO			HJ		
ρ_2	Strike	MC Imp. vol.	Approx	Err.	MC Imp. vol.	Approx.	Err.
-0.4	40	19.45 (0.16)	19.12	0.33	19.38 (0.16)	19.12	0.26
	80	17.26 (0.20)	17.46	0.20	17.22 (0.20)	17.46	0.24
	100	16.58 (0.23)	16.83	0.25	16.61 (0.23)	16.83	0.22
	120	16.28 (0.25)	16.26	0.02	16.27 (0.25)	16.26	0.01
	160	15.08 (0.30)	15.17	0.09	15.33 (0.30)	15.17	0.16
0	40	19.18 (0.16)	19.13	0.05	19.38 (0.16)	19.13	0.25
	80	17.32 (0.20)	17.46	0.14	17.27 (0.20)	17.46	0.19
	100	16.65 (0.23)	16.83	0.18	16.71 (0.23)	16.83	0.12
	120	16.16 (0.25)	16.26	0.10	16.06 (0.25)	16.26	0.20
	160	15.23 (0.30)	15.17	0.06	15.22 (0.30)	15.17	0.05
0.4	40	19.45 (0.16)	19.14	0.31	19.31 (0.16)	19.14	0.17
	80	17.30 (0.20)	17.46	0.16	17.25 (0.20)	17.46	0.20
	100	16.59 (0.23)	16.82	0.24	16.59 (0.23)	16.82	0.24
	120	16.22 (0.25)	16.25	0.03	16.10 (0.25)	16.25	0.15
	160	15.57 (0.30)	15.18	0.39	15.52 (0.30)	15.18	0.35

Table 5.3: The other parameters are assumed as: $\nu_0 = 0.02$, $\kappa_\nu = 2.1$, $\mu_\nu = 0.03$, $\sigma_\nu = 0.2$, $\rho_0 = -0.4$, $\kappa_\rho = 3.4$, $\mu_\rho = -0.6, \sigma_\rho = 0.1$, the numbers in round brackets represent the standard deviations.

OU process and the bounded Jacobi process are both mean-reverting processes; more exactly, they have the same structure for the drift. If the value of σ_ρ is so small that the random part in the correlation process will play a minor role, one obtains thus the same implied volatilities for using the OU and the bounded Jacobi process. Similarly, a small value σ_ρ of the correlation process leads to a rather small effect of ρ_2. In Table 5.4, we increase the value of σ_ρ to be 0.18, the mentioned differences between using the HO and HJ model, or for varying ρ_2 can be seen. The error values in this table showed again that the approximations give a rather accurate result. For the Monte-Carlo simulation we remark: While choosing the parameters one needs to pay attention to keep the values inside of the square root to be positive, see (5.78).

Model		HO			HJ		
ρ_2	Strike	MC Imp. vol.	Approx	Err.	MC Imp. vol.	Approx.	Err.
-0.4	40	19.24 (0.16)	19.51	0.27	19.27 (0.16)	19.02	0.25
	80	17.38 (0.20)	17.39	0.01	17.37 (0.20)	17.42	0.05
	100	16.82 (0.23)	16.86	0.04	16.75 (0.23)	16.84	0.08
	120	16.05 (0.25)	16.27	0.22	16.18 (0.25)	16.31	0.13
	160	15.31 (0.30)	15.35	0.04	15.16 (0.30)	15.35	0.19
0	40	19.29 (0.16)	19.72	0.43	19.25 (0.16)	19.03	0.22
	80	17.34 (0.20)	17.38	0.04	17.24 (0.20)	17.42	0.18
	100	16.70 (0.22)	16.86	0.16	16.71 (0.22)	16.83	0.12
	120	16.26 (0.25)	16.25	0.01	16.14 (0.25)	16.30	0.16
	160	15.22 (0.30)	15.37	0.15	15.41 (0.30)	15.36	0.05
0.4	40	19.36 (0.16)	20.00	0.64	19.33 (0.16)	19.04	0.29
	80	17.35 (0.20)	17.37	0.02	17.31 (0.20)	17.42	0.11
	100	16.61 (0.23)	16.86	0.25	16.79 (0.23)	16.82	0.03
	120	16.36 (0.25)	16.22	0.14	16.07 (0.25)	16.30	0.22
	160	15.63 (0.30)	15.39	0.24	15.46 (0.30)	15.36	0.10

Table 5.4: The other parameters are assumed as: $\nu_0 = 0.02$, $\kappa_\nu = 2.1$, $\mu_\nu = 0.03$, $\sigma_\nu = 0.2$, $\rho_0 = -0.4$, $\kappa_\rho = 3.5$, $\mu_\rho = -0.55, \sigma_\rho = 0.18$, the numbers in round brackets represent the standard deviations.

5.3.6 Calibration to Market Data

In order to recognize the performance of our models in a calibration setting, we compare the calibration using the Heston model extended with a stochastic correlation to the calibrations using the pure Heston model and the double Heston model. For the market data, we choose Put-options on the Nikk300 index on December 31, 2012, which is used in [112] and representative for the skew and patterns observed. Since our aim is to compare our models to the pure Heston model [58] and the double Heston model [31], we thus just use the standard optimization methods: We fit the prices computed by the different models to the market observed prices for several maturities T_i and strikes K_j; one can obtain the parameter estimates by minimizing, e.g., the mean square error (MSE)

$$\frac{1}{N}\sum_{i,j} w_{ij}(P^{Mkt}(T_i, K_j) - P^{Mod}(T_i, K_j))^2, \tag{5.168}$$

with the market price $P^{Mkt}(T_i, K_j)$ and the corresponding model price $P^{Mod}(T_i, K_j)$; w_{ij} is an optional weight.

We report our results in Table 5.5, where ν_0^k, κ_ν^k, μ_ν^k, σ_ν^k, σ_ν^k are the parameters for the two stochastic volatilities in the double Heston model, $k = 1, 2$. We see that the

Pure Heston	$\hat{\nu}_0$ 0.05	$\hat{\kappa}_\nu$ 4.13	$\hat{\mu}_\nu$ 0.05	$\hat{\sigma}_\nu$ 0.39	$\hat{\rho}$ -0.47			MSE 2.3×10^{-2}			
Double Heston	$\hat{\nu}_0^1$ 0.05	$\hat{\kappa}_\nu^1$ 6.36	$\hat{\mu}_\nu^1$ 0.02	$\hat{\sigma}_\nu^1$ 0.49	$\hat{\rho}^1$ -0.23	$\hat{\nu}_0^2$ 0.01	$\hat{\kappa}_\nu^2$ 5.69	$\hat{\mu}_\nu^2$ 0.03	$\hat{\sigma}_\nu^2$ 0.62	$\hat{\rho}^2$ -0.44	MSE 13.0×10^{-3}
Heston OU	$\hat{\nu}_0$ 0.06	$\hat{\kappa}_\nu$ 3.32	$\hat{\mu}_\nu$ 0.07	$\hat{\sigma}_\nu$ 2.02	$\hat{\rho}_0$ -0.01	$\hat{\kappa}_\rho$ 2.12	$\hat{\mu}_\rho$ -0.31	$\hat{\sigma}_\rho$ 0.33	$\hat{\rho}_{x\rho}$ -0.88	MSE 6.7×10^{-3}	
Heston Jacobi	$\hat{\nu}_0$ 0.05	$\hat{\kappa}_\nu$ 0.75	$\hat{\mu}_\nu$ 0.07	$\hat{\sigma}_\nu$ 0.50	$\hat{\rho}_0$ -0	$\hat{\kappa}_\rho$ 2.72	$\hat{\mu}_\rho$ -0.17	$\hat{\sigma}_\rho$ 0.02	$\hat{\rho}_{x\rho}$ -0.91	MSE 5.4×10^{-3}	

Table 5.5: Estimated model parameters for the Nikk300 index on December 31, 2012.

MSE values for the Heston model with stochastic correlation are smaller than the pure Heston model and the double Heston model.

To illustrate more clearly, we define the error as the absolute value of the difference between the implied market volatilities and the model implied volatilities, namely

$$Error := |Vol^{Mkt}(T_i, K_j) - Vol^{Mod}(T_i, K_j)|. \tag{5.169}$$

Then, we compare the errors for these models in Figure 5.13 for relatively short maturities $T = 30$, 90, 180, 360 days and in Figure 5.14 for relative long maturities $T = 2$, 3, 4, 5 years. We observe for all maturities, that the Heston model extended by incorporating the stochastic correlation (in the both cases HO and HJ) can be better fitted to real market data not only than the pure Heston model but also than the double Heston model, although the extended Heston model with stochastic correlation has one parameter less than the double Heston model. This proves that introducing a stochastic correlation can significantly improve the the calibration. About how each parameter of the stochastic correlation process effect the implied volatilities, see Section 5.3.4.

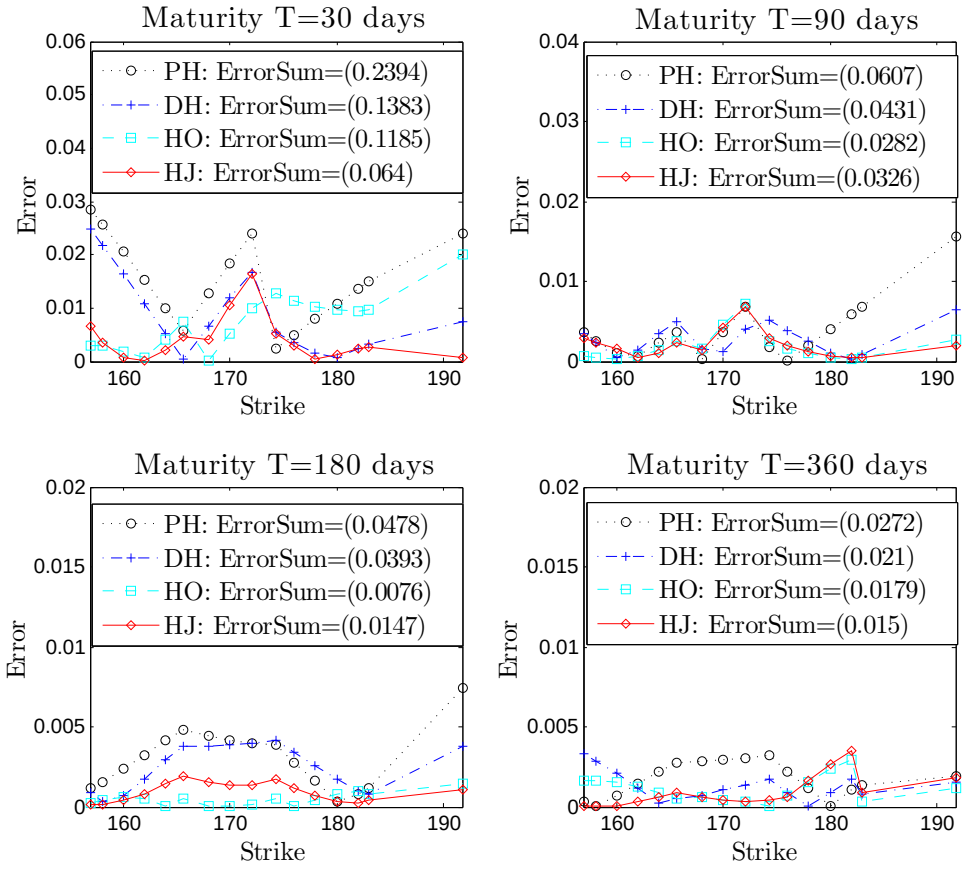


Figure 5.13: Using the Nikk300 index on December 31, 2012 where spot price is 174.3, and for some short maturities $T = 30, 90, 180, 360$ days, the errors (defined as the absolute value of the difference between the implied market volatilities and the implied volatilities for the models) are compared for the pure Heston model ('PH'), the double Heston model ('DH'), the HO model and the HJ model. ErrorSum denotes the sum of errors for each maturity with different strikes.

For an further example we take the Call-options on the Nikk300 index on May 10, 2012, where spot price is 159.7. We consider several maturities $T = 30, 90, 180, 360$ days and strike which ranges from 154 to 167. Instead of statistical error (MSE) we report the error defined in (5.169) and estimates for all the models in Table 5.6. By comparing the

Pure Heston	$\hat{\nu}_0$ 0.039	$\hat{\kappa}_\nu$ 6.984	$\hat{\mu}_\nu$ 0.051	$\hat{\sigma}_\nu$ 1.641	$\hat{\rho}$ -0.313			Error 4.16×10^{-4}			
Double Heston	$\hat{\nu}_0^1$ 0.001	$\hat{\kappa}_\nu^1$ 3.520	$\hat{\mu}_\nu^1$ 0.050	$\hat{\sigma}_\nu^1$ 0.585	$\hat{\rho}^1$ -0.791	$\hat{\nu}_0^2$ 0.033	$\hat{\kappa}_\nu^2$ 2.128	$\hat{\mu}_\nu^2$ 0.001	$\hat{\sigma}_\nu^2$ 1.163	$\hat{\rho}^2$ -0.068	Error 1.21×10^{-4}
Heston OU	$\hat{\nu}_0$ 0.029	$\hat{\kappa}_\nu$ 3.996	$\hat{\mu}_\nu$ 0.046	$\hat{\sigma}_\nu$ 0.591	$\hat{\rho}_0$ 0.052	$\hat{\kappa}_\rho$ 4.036	$\hat{\mu}_\rho$ -0.919	$\hat{\sigma}_\rho$ 0.034	$\hat{\rho}_{x\rho}$ 0.532	Error 1.00×10^{-4}	
Heston Jacobi	$\hat{\nu}_0$ 0.043	$\hat{\kappa}_\nu$ 5.489	$\hat{\mu}_\nu$ 0.049	$\hat{\sigma}_\nu$ 1.461	$\hat{\rho}_0$ 0.259	$\hat{\kappa}_\rho$ 5.487	$\hat{\mu}_\rho$ -0.404	$\hat{\sigma}_\rho$ 0.077	$\hat{\rho}_{x\rho}$ 0.166	Error 1.14×10^{-4}	

Table 5.6: Estimated model parameters for the Nikk300 index on May 10, 2012.

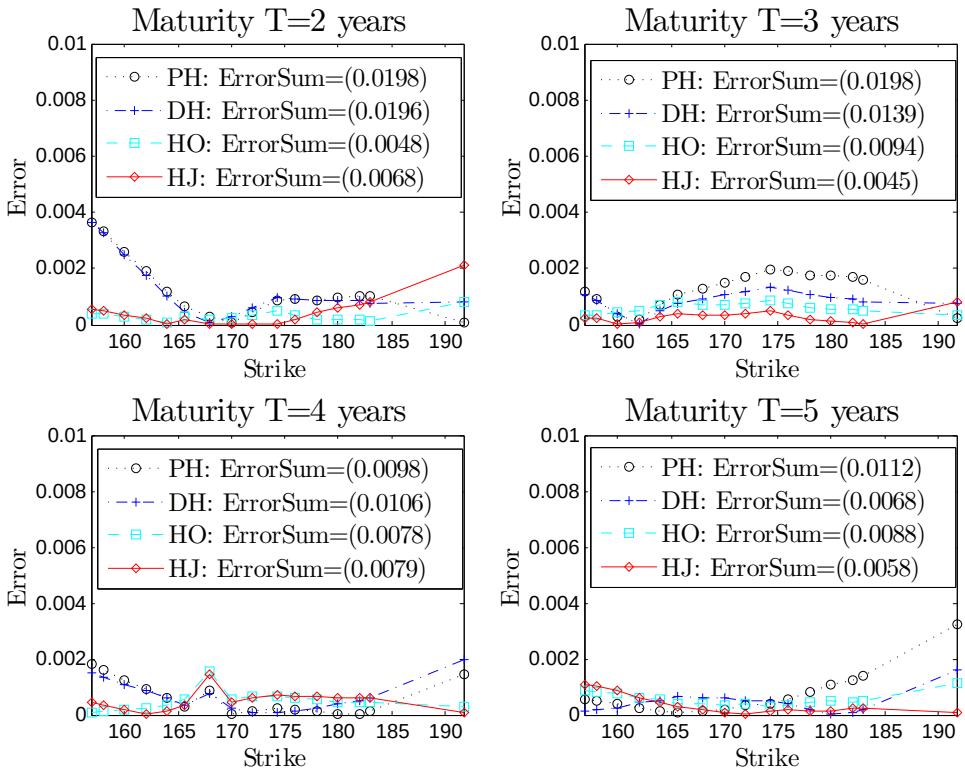


Figure 5.14: Using the Nikk300 index on December 31, 2012 where spot price is 174.3, and for some long maturities $T = 2, 3, 4, 5$ years, the errors (defined as absolute value of the difference between the implied market volatilities and the implied volatilities for the models) are compared for the pure Heston model ('PH'), the double Heston model ('DH'), the HO model and the HJ model. ErrorSum denotes the sum of errors for each maturity with different strikes.

error values in Table (5.169) we again conclude that our models provide a more realistic volatility smile than the pure Heston model and the double Heston model.

Furthermore, we display the implied volatilities for all the models in Figure 5.15 and compare them to the market volatilities. Obviously, either the HO model or the HJ model provides a better fit to the market volatilities, especially, the volatility smile as market requires for the short maturity $T = 30$ days.

The experiments on the calibration to the real market data has shown that introducing a stochastic correlation can not only improve significantly the performance of the pure Heston model, but also be better than the double Heston model. The great importance of modelling financial correlation as a stochastic process has thus been validated.

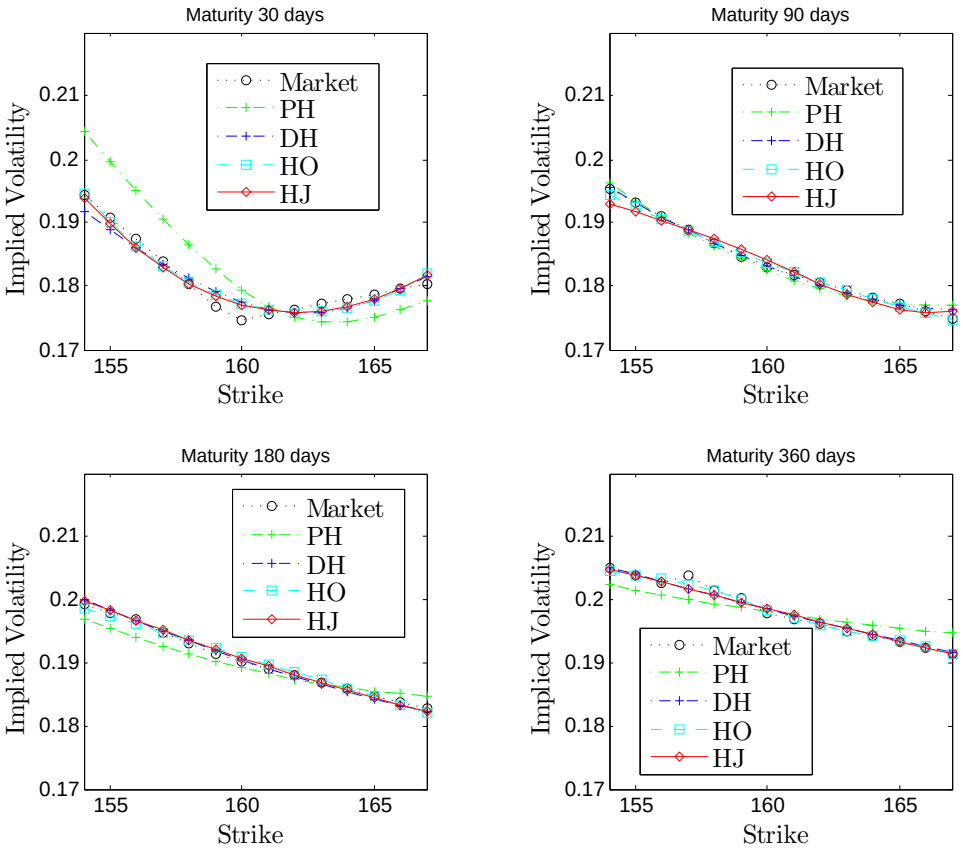


Figure 5.15: Using the Nikk300 index on 10 May, 2012 where spot price is 159.7, and for the maturities $T = 30, 90, 180, 360$ days, the implied volatilities are compared for the pure Heston model ('PH'), the double Heston model ('DH'), the HO model and the HJ model.

Chapter 6

Conclusion

In this thesis, Counterparty Credit risk with special regard to the correlation among counterparties, modelling and application of time-dependent correlation, modelling and application of stochastic correlation were introduced and analysed.

The main contributions of this thesis are fivefold. Firstly, from the investigation of computing Bilateral Value Adjustment on a Credit Default Swap contract we presented a problem of computing the cumulative distribution function of the integrated Cox-Ingersoll-Ross process, which has been a benchmark in finance for many years. This cumulative distribution function plays an important role not only in modelling of Credit risk but also on many other issues in finance. We developed a new strategy that allows to construct a very robust routine to numerically determine a highly accurate cumulative distribution function of the integrated Cox-Ingersoll-Ross process.

Secondly, motivated with the indeed simultaneous default events in the real financial market, e.g., the collapses of Lehman Brothers and Merrill Lynch were just within two days (September 13-14, 2008) and 24 railways firms defaulted simultaneously on the same day, June 21, 1970, a new formula for pricing Bilateral Value Adjustment on a Credit Default Swap contract is developed. Applying our new formula, we were able not only to fully capture the Wrong-Way risk on a Credit Default Swap contract, but also to confirm

the role of considering simultaneous defaults on the valuation of the Counterparty risk in Credit Default Swap contracts. Another important finding was that the effects of the simultaneous defaults on the Bilateral Value Adjustment are not identical for the contracting party as the Credit Default Swap seller and the Credit Default Swap payer.

Thirdly, we proposed a new time-dependent correlation function which firstly satisfies the correlation properties and secondly can be easily incorporated into financial models instead of using a constant correlation. The benefit of using our time-dependent correlation model is that additional parameters can be chosen to increase the fitting quality to the real market data. Compared to the way using time-dependent parameters our model has an economic meaning, namely the correlation between, e.g., financial quantities and parties is not constant but time-varying as observed in the market. To confirm our statement, as examples, we applied this time-dependent correlation to price Quanto options and the Heston model.

Fourthly, due to the uncertainty associated with the future development of relationship between, e.g., financial parties and products, modelling correlation stochastically should better replicate the correlation in reality. We proposed a general stochastic correlation model from which we could define many variants of stochastic correlation processes. To the best of our knowledge, almost all stochastic correlation processes proposed by other authors are within the class of our general stochastic correlation models.

The last important topic analysed in the framework of this thesis was the application of stochastic correlation process to price Quanto options and the Heston model. For pricing Quanto options we quantified the correlation risk caused by using a wrong (constant) correlation. Introducing a stochastic correlation into the Heston model can provide a better skew and smile in the volatility surface, not only than the pure Heston model but also than some other extensions of the Heston for the same aim, e.g., the double Heston model.

This thesis contains a wide range of topics related to the impact of financial correlation, its modellings and applications in finance. Of course, compared to other important financial quantities, e.g., stochastic volatility processes and stochastic interest rate processes, the modelling and application of financial correlations using a stochastic process is still at an early stage. Therefore, there remain many open problems that need to be solved for a wide application of stochastic correlation in finance. In particular, for credit risk management, it is an urgent task to establish a general framework for considering stochastically dependent default intensities among counterparties.

Appendix A

Preliminaries

Suppose a probability space $(\Omega, \mathcal{F}, P)$

Definition A.1 *A filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ is said to satisfy the usual conditions if the following conditions hold [64]:*

- *$\mathcal{F}_0$ (so that all $\mathcal{F}_t$) contains all P-negligible events in $\mathcal{F}$;*
- *$\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous for all $t \geq 0$.*

Definition A.2 *A self-financing trading strategy is called an arbitrage opportunity of its value process V satisfies [48]*

$$V_0 \leq 0, \quad V_T \geq 0 \text{ } P\text{-almost surely and } P(V_T) > 0.$$

Definition A.3 *Let $\tilde{P}$ a probability measure on $(\Omega, \mathcal{F})$, $\tilde{P}$ is said to be absolutely continuous with respect to P on $\mathcal{F}$, and we write $\tilde{P} \ll P$, if for all $A \in \mathcal{F}$ [48],*

$$P(A) = 0 \;\Rightarrow\; \tilde{P}(A) = 0.$$

If both $\tilde{P} \ll P$ and $P \ll \tilde{P}$ hold, $\tilde{P}$ and Q are said to be equivalent by $\tilde{P} \approx P$.

Theorem A.1 *$\tilde{P}$ is absolutely continuous with respect to P on $\mathcal{F}$ if and only if there exists an $\mathcal{F}$-measurable function $\psi \geq 0$ so that [48]*

$$\int F \, d\tilde{P} = \int F\psi \, dP \quad \textit{for all } \mathcal{F}\textit{-measurable functions } F \geq 0,$$

we say that ψ is the density or Radon-Nikodym derivative of $\tilde{P}$ with respect to P and write

$$\psi := \frac{d\,\tilde{P}}{d\,P},$$

which is uniquely determined.

Theorem A.2 *We denote the set of risk-neutral measures which are equivalent to P by*

$$\mathcal{Q} := \{\mathbb{Q} \mid \mathbb{Q} \textit{ is a risk-neutral measure with } \mathbb{Q} \approx P\}.$$

A market model is arbitrage-free if and only if $\mathcal{Q} \neq \emptyset$ and there exists a $\mathbb{Q} \in \mathcal{Q}$ which has a bounded density $d\,\mathbb{Q}/d\,P$. The proof can be found in [48].

Definition A.4 *A market model is complete if every derivative security can be hedged.*

Theorem A.3 *Consider a market model that has a risk-neutral probability measure, say a arbitrage-free market model. The model is complete if and only if the risk-neutral probability measure is unique, i.e. if $|\mathcal{Q}| = 1$.*

Theorem A.4 *Girsanov theorem in one dimension: Let $W_t,\ 0 \leq t \leq T$, be a Brownian motion (BM) on $(\Omega, \mathcal{F}, P)$, and $\mathcal{F}_t$ is the filtration for this BM. Let $\Theta_t,\ 0 \leq t \leq T$, be an adapted process. Define*

$$Z_t = e^{-\int_0^t \Theta_s \, dW_s - \frac{1}{2}\int_0^t \Theta_s^2 \, ds},$$
$$\tilde{W}_t = W_t + \int_0^t \Theta_s \, ds$$

and assume that

$$\mathbb{E}\left[\int_0^T \Theta_s^2 Z_s^2 \, ds\right] < \infty.$$

Set $Z = Z_T$. *Then* $\mathbb{E}(Z) = 1$ *and under the probability measure* $\mathbb{Q}$ *given by*

$$d\mathbb{Q} = e^{-\int_0^T \Theta_s \, dW_s - \frac{1}{2}\int_0^T \Theta_s^2 \, ds} \, dP,$$

the $\tilde{W}_t$, $0 \leq t \leq T$, *is a BM. See [102] for the proof and Girsanov theorem in multidimensional case.*

An Example (Stock price under $\mathbb{Q}$):

The stock price can be modelled by a geometric BM:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t. \tag{A.1}$$

We define $\tilde{W}_t = W_t + \frac{\mu - r}{\sigma} t$, *namely choosing* $\Theta_t = \frac{\mu - r}{\sigma}$ *as constant. Theorem A.4 says that* $\tilde{W}$ *is a BM under* $\mathbb{Q}$, *by substituting* $dW_t = d\tilde{W}_t - \frac{\mu - r}{\sigma} dt$ *into* (A.1) *we obtain*

$$dS_t = r S_t \, dt + \sigma S_t \, d\tilde{W}_t, \tag{A.2}$$

where r *is the risk-free interest rate and* $\frac{\mu - r}{\sigma}$ *is called market price of risk. By applying Itô's lemma to the discounted price* $e^{-rt} S_t$ *we obtain* $d\left(e^{-rt} S_t\right) = \sigma e^{-rt} S_t d\tilde{W}_t$ *which shows that the discounted price is obviously a martingale. The parameter* μ, σ, r *are constant in this example, they are also allowed to be adapted processes. For more detailed information see [43, 48, 102].*

Appendix B

Basic Definitions

Definition B.1 *The Pearson correlation coefficient is defined for two random variables X and Y as*

$$\rho_{X,Y} := \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E\left[(X-\mu_X)(Y-\mu_Y)\right]}{\sigma_X \sigma_Y}. \tag{B.1}$$

Definition B.2 *For the given realisations $\tilde{X}_1, \ldots, \tilde{X}_n$ and $\tilde{Y}_1, \ldots, \tilde{Y}_n$ of X and Y, the realised correlation can be estimated pairwise as*

$$\rho_{X,Y} \approx \tilde{\rho}_{X,Y} := \frac{\sum_{i=1}^{n}\left(\tilde{X}_i - \sum_{j=1}^{n}\tilde{X}_j\right)\left(\tilde{Y}_i - \sum_{j=1}^{n}\tilde{Y}_j\right)}{\sqrt{\sum_{i=1}^{n}\left(\tilde{X}_i - \sum_{j=1}^{n}\tilde{X}_j\right)^2 \sum_{i=1}^{n}\left(\tilde{Y}_i - \sum_{j=1}^{n}\tilde{Y}_j\right)^2}}. \tag{B.2}$$

Definition B.3 *Linked to the realised correlation, given $n+1$ data points at $t_0, \cdots, t_n$, the rolling correlation over a time window of length $m < n$ reads*

$$\tilde{\rho}(t_0, \ldots, t_{m-1}),\ \tilde{\rho}(t_1, \cdots, t_m),\ \ldots,\ \tilde{\rho}(t_{n-m+1}, \ldots, t_n). \tag{B.3}$$

Definition B.4 *A zero-coupon bond with one unit of currency principal and maturity T is a contract that guarantees its holder the payment of one unit of currency at T, with no intermediate coupons. The price of the bond at time $t < T$ is denoted by $B(t,T)$ and $B(T,T) = 1$.*

Definition B.5 *We consider a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\{\mathcal{F}_t\}$. A random variable τ is a stopping time of the filtration, if the event $\{\tau \leq t\} \in \mathcal{F}_t$, for all $t \geq 0$.*

Definition B.6 *A hypergeometric function F is defined as*

$$F(a, b, c, x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \frac{(a)_k (b)_k}{(c)_k}, \quad |x| < 1, \tag{B.4}$$

where $(\cdot)_k$ denotes the Pochhammer symbol,

$$(a)_k = a(a+1)(a+2)\cdots(a+k-1), \quad (a)_0 = 1. \tag{B.5}$$

Definition B.7 *Suppose a system of SDEs given by*

$$d\mathbf{X}_t = \mu(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)d\mathbf{W}_t, \tag{B.6}$$

which is said to be of the affine form [36, 37], if

$$\mu(\mathbf{X}_t) = a_0 + a_1 \mathbf{X}_t, \quad (a_0, a_1) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}, \tag{B.7}$$

$$\left(\sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^\top\right)_{i,j} = (b_0)_{i,j} + (b_1)_{i,j}^\top \mathbf{X}_t, \quad (b_0, b_1) \in \mathbb{R}^n \times \mathbb{R}^{n \times n \times n}, \tag{B.8}$$

for $i, j = 1, ..., n$. Then, the charateristic function under $\mathbb{Q}$ takes the form

$$\phi\left(\mathbf{u}, \mathbf{X}_t, t, T\right) = \mathbb{E}^{\mathbb{Q}}\left[e^{i\mathbf{u}^\top \mathbf{X}_T} | \mathcal{F}_t{}^1\right] = e^{A(\mathbf{u},\tau) + B(\mathbf{u},\tau)\mathbf{X}_t}. \tag{B.9}$$

By setting $\tau := T - t$, *the coefficients* $A(\mathbf{u}, \tau)$ *and* $B(\mathbf{u}, \tau)$ *in* (B.9) *must satisfy the following complex-valued ordinary differential equations:*

$$\frac{d}{d\tau}B(\mathbf{u},\tau) = a_1^\top B(\mathbf{u},\tau) + \frac{1}{2}B^\top(\mathbf{u},\tau)b_1 B(\mathbf{u},\tau), \tag{B.10}$$

$$\frac{d}{d\tau}A(\mathbf{u},\tau) = a_0 B(\mathbf{u},\tau) + \frac{1}{2}B^\top(\mathbf{u},\tau)b_0 B(\mathbf{u},\tau), \tag{B.11}$$

with boundary conditions $A(\mathbf{u}, 0) = 0$ *and* $B(\mathbf{u}, 0) = i\mathbf{u}$.

Definition B.8 *The cumulative distribution function for the non-central chi-squared distribution,* $F_{\chi^2}(x; d, \lambda)$, *with* d *degrees of freedom and non-centrality parameter* λ *is defined as*

$$F_{\chi^2}(x; d, \lambda) = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!}\frac{\gamma(\frac{d}{2}+j, \frac{x}{2})}{\Gamma(\frac{d}{2}+j)}, \tag{B.12}$$

where $\gamma(d, x) = \int_0^x y^{d-1}e^{-y}dy$ *is the lower incomplete Gamma function.*

[1]Assume that $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$ satisfies the usual conditions, and $\mathbf{X}$ is assumed to be Markov relative to $(\mathcal{F}_t)$.

Appendix C

Proofs

Proofs of propositions in Section 3.2

The proof of Proposition 3.2.1

Proof: We calculate firstly

$$\begin{aligned}\mathbb{1}_{\mathfrak{A}\cup\mathfrak{B}}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\tau_R>t|\mathcal{G}_{\tau_I}) &= \mathbb{1}_{\tau_I\leq T}\mathbb{1}_{\tau_I\leq\tau_C}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\tau_R>t|\mathcal{G}_{\tau_I})\\ &= \mathbb{1}_{\tau_I\leq T}\mathbb{1}_{\tau_I\leq\tau_C}\left(\mathbb{1}_{t<\tau_I<\tau_R}+\underbrace{\mathbb{1}_{\tau_I\leq t}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\tau_R>t|\mathcal{G}_{\tau_I})}_{:=\mathcal{M}}\right),\end{aligned}\tag{C.1}$$

where $\mathcal{M}$ can be calculated as:

$$\mathbb{1}_{\tau_I\leq t}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(\Lambda_R(t)<\xi_R|\mathcal{G}_{\tau_I}) = \mathbb{1}_{\tau_I\leq t}\mathbb{1}_{\tau_R>\tau_I}\mathbb{Q}(U_R<1-\mathrm{e}^{-Y_R(t)-\Psi_R(t;\,\beta_R)}|\mathcal{G}_{\tau_I}).\tag{C.2}$$

Condition on ξ_R (or u_R) we may rewrite the equation (C.2) as

$$\begin{aligned}\mathcal{M} &= \mathbb{1}_{\tau_I\leq t}\mathbb{1}_{\tau_R>\tau_I}\mathbb{E}\left[\mathbb{Q}(u_R<1-\mathrm{e}^{-Y_R(t)-\Psi_R(t;\,\beta_R)}|\mathcal{G}_{\tau_I},u_R)|\mathcal{G}_{\tau_I},\{\tau_R>\tau_I\}\right]\\ &= \mathbb{1}_{\tau_I\leq t}\mathbb{1}_{\tau_R>\tau_I}\mathbb{E}\left[\mathbb{Q}(Y_R(t)<-\log(1-u_R)-\Psi_R(t;\,\beta_R)|\mathcal{G}_{\tau_I},u_R)|\mathcal{G}_{\tau_I},\{\tau_R>\tau_I\}\right]\\ &= \mathbb{E}\left[F_{Y_R(t)}\left(-\log(1-u_R)-\Psi_R(t;\,\beta_R)\right)|\mathcal{G}_{\tau_I},\{\tau_R>\tau_I\}\right]\end{aligned}$$

$$= \int_0^1 F_{Y_R(t)} \left(-\log(1-u_R) - \Psi_R(t;\, \beta_R) \right) d\mathbb{Q} \left(U_R < u_R | \mathcal{G}_{\tau_I}, \{\tau_R > \tau_I\} \right). \tag{C.3}$$

Next, we compute the conditional distribution as follows: Denote

$$C_{R|I}(u_R; U_I) := \mathbb{Q} \left(U_R < u_R | \mathcal{G}_{\tau_I}, \{\tau_R > \tau_I\} \right), \tag{C.4}$$

which may be rewritten as

$$\begin{aligned} C_{R|I}(u_R; U_I) &= \mathbb{Q} \left(U_R < u_R | U_I, \{U_R > \overline{U}_{R,I}\} \right) \\ &= \frac{\mathbb{Q}(U_R < u_R | U_I) - \mathbb{Q}(U_R < \overline{U}_{R,I} | U_I)}{1 - \mathbb{Q}(U_R < \overline{U}_{R,I} | U_I)} \\ &= \frac{\frac{\partial C_{I,\,R}(u_I, u_R)}{\partial u_I}|_{u_I = U_I} - \frac{\partial C_{I,\,R}(u_I, \overline{U}_{R,\,I})}{\partial u_I}|_{u_I = U_I}}{1 - \frac{\partial C_{I,\,R}(u_I, \overline{U}_{R,I})}{\partial u_I}|_{u_I = U_I}}. \end{aligned}$$

The proof can be completed by substituting (C.4) into (C.3), and further (C.3) into (C.1). □

Proofs of propositions in Section 3.3

The proof of Proposition 3.3.1

Proof:

$$\lim_{u\to\infty} \frac{b(u)}{\sqrt{u}} = \lim_{u\to\infty} \sqrt{\frac{\kappa^2}{u} - 2i\sigma^2} = \sqrt{2}\sigma\sqrt{-i} = \sqrt{2}\sigma \mathrm{e}^{\frac{7\pi}{4}i},$$

Regarding this result we have straightforward $\lim_{u\to\infty} a(u) = -1$.

Now we prove the equation (3.46)

$$\lim_{u\to\infty} \frac{A(u)}{\sqrt{u}} = \lim_{u\to\infty} \frac{2\kappa\mu}{\sigma^2} \left(\ln(2) + \frac{t}{2}(\kappa - b(u)) + \ln\left(\frac{-\mathrm{e}^{tb(u)}}{a(u)\mathrm{e}^{tb(u)} - 1} \right) \right) / \sqrt{u}$$

$$= \lim_{u\to\infty} \frac{2\kappa\mu}{\sigma^2}\left(-\frac{t}{2}\frac{b(u)}{\sqrt{u}}\right)$$
$$= -\frac{\sqrt{2}\kappa\mu t}{\sigma}\mathrm{e}^{\frac{7\pi}{4}i}.$$

Finally, we show the equation (3.47)

$$\lim_{u\to\infty}\frac{B(u)}{\sqrt{u}} = \lim_{u\to\infty}\frac{2ui}{\kappa - b(u)}\left(\frac{1}{a(u)}\left(1+\frac{1-a(u)}{a(u)\mathrm{e}^{tb(u)}-1}\right)\right)/\sqrt{u}$$
$$= \lim_{u\to\infty}\frac{-2\sqrt{u}i}{\kappa - b(u)}$$
$$= \lim_{u\to\infty}\frac{2i}{\frac{b(u)}{\sqrt{u}}} = \frac{\sqrt{2}i}{\sigma}\mathrm{e}^{-\frac{7\pi}{4}i}.$$

□

The proof of Proposition 3.3.3

Proof: Since the function $b(u)$ and the parameter t are non-negative, we only need to prove

$$|a(u)| = \left|\frac{\kappa + b(u)}{\kappa - b(u)}\right| > 1.$$

This is to say that we have to show

$$|\kappa + b(u)|^2 > |\kappa - b(u)|^2. \tag{C.5}$$

We split $b(u)$ into a real and imaginary part as

$$b(u) = b_r + ib_i, \quad b_r, b_i \in \mathbb{R},$$

then the left hand side of (C.5) satisfies

$$0 < |\kappa + b(u)|^2 = |\kappa + b_r + ib_i|^2 = (\kappa + b_r)^2 + b_i^2 = \kappa^2 + 2\kappa b_r + |b(u)|^2,$$

and analogously the right hand side of (C.5) fullfills

$$0 < |\kappa - b(u)|^2 = \kappa^2 - 2\kappa b_r + |b(u)|^2.$$

The fact that $\kappa > 0$ and $b_r > 0$ completes the proof. □

The proof of Proposition 3.3.4

Proof: Combining the function $g(u)$ with the function $\phi_{Y_t}(u)$ as given in (3.37) we have

$$\lim_{u\to 0} g(u) = \lim_{u\to 0} \mathrm{Im}\left[\mathrm{e}^{-iu\tilde{y}_t}\frac{\mathrm{e}^{A(t,u)+B(t,u)y_0}}{u}\right].$$

First we consider the limits of $a(u)$ and $b(u)$ as defined in (3.40)

$$\lim_{u\to 0} b(u) = |\kappa| = \kappa, \qquad \lim_{u\to 0} a(u) = \infty, \tag{C.6}$$

where the last step in the limit of $b(u)$ follows the fact that the CIR model parameter κ is always positive. From (C.6) we can directly deduce

$$\lim_{u\to 0} A(t,u) = \lim_{u\to 0} B(t,u) = 0, \tag{C.7}$$

and thus

$$\lim_{u\to 0} \mathrm{e}^{-iu\tilde{y}_t + A(t,u) + B(t,u)y_0} = 1. \tag{C.8}$$

Now we split the exponent in the last equation into a real and imaginary part as

$$H(u) + iJ(u) := -iu\tilde{y}_t + A(t,u) + B(t,u)y_0, \tag{C.9}$$

with functions $H(u)$ and $J(u) : \mathbb{R} \to \mathbb{R}$. Furthermore, from (C.8) we also know that

$$\lim_{u\to 0} H(u) = 0, \qquad \lim_{u\to 0} J(u) = 0.$$

Now we can calculate the value of $g(u)$ at zero as follows

$$\begin{aligned} g(0) &= \lim_{u\to 0} \mathrm{Im}\left[\frac{\mathrm{e}^{H(u)+iJ(u)}}{u}\right] \\ &= \lim_{u\to 0} \mathrm{Im}\left[\mathrm{e}^{H(u)\frac{\cos(J(u))+i\sin(J(u))}{u}}\right] \\ &= \lim_{u\to 0} \mathrm{e}^{H(u)\frac{\sin(J(u))}{u}} \\ &= \lim_{u\to 0} \frac{\sin(J(u))}{u} \\ &\overset{\text{l'Hospital}}{=} \lim_{u\to 0} J'(u)\frac{\cos(J(u))}{1} \\ &= \lim_{u\to 0} J'(u). \end{aligned}$$

Using the equation (C.9) we obtain

$$g(0) = \lim_{u\to 0} J'(u) = J'(0) = -\tilde{y}_t + \mathrm{Im}(A'(t,u)) + \mathrm{Im}(B'(t,u)y_0).$$

The computation of $A'(t,u)$ and $B'(t,u)$ is straightforward but tedious. We obtain finally

$$\begin{aligned} \mathrm{Im}(A(t,0)') &= \frac{\mu\kappa\mathrm{e}^{-\kappa t} + \mu\kappa(t\kappa - 1)}{\kappa^2}, \\ \mathrm{Im}(B(t,0)') &= \frac{1-\mathrm{e}^{-\kappa t}}{\kappa}. \end{aligned}$$

□

The proof of Proposition 3.4.1

Proof: We have

$$\begin{aligned}\widehat{P}_t^{\mathrm{CDS}}(\mathcal{P},\mathcal{L}_R) \overset{(3.81)(3.86)}{=} \mathbb{E}\Bigg\{ & \mathbb{1}_{\mathfrak{A}} D(t,\tau)(-\mathcal{L}_R) \\ & + \mathbb{1}_{\mathfrak{B}}\left[D(t,\tau)\left(\mathcal{R}_C(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - (P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\right)\right] \\ & + \mathbb{1}_{\mathfrak{C}}\left[D(t,\tau)\left((P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - \mathcal{R}_I(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\right)\right] \\ & + \mathbb{1}_{\mathfrak{D}}\left[D(t,\tau)\left(-(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)\right)\right] \\ & + \mathbb{1}_{\mathfrak{E}}\left[D(t,\tau)\left(\mathcal{L}_R\right)\right] \\ & + \underbrace{D(t,\tau)(\tau - T_{\gamma(\tau)-1})\mathcal{P}\mathbb{1}_{\{T_a<\tau<T_b\}} + \sum_{i=a+1}^{b} D(t,T_i)\alpha_i\mathcal{P}\mathbb{1}_{\{\tau\geq T_i\}}}_{:=\mathcal{M}_1}\Bigg|\mathcal{G}_t\Bigg\}\end{aligned} \tag{C.10}$$

We first consider the expression regarding the event $\mathfrak{B}$ inside the above conditional expectation

$$\mathbb{1}_{\mathfrak{B}}\left[D(t,\tau)\left(\mathcal{R}_C(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - (P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\right)\right].$$

Since

$$\begin{aligned}\mathcal{R}_C(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ - (P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^- =& (\mathcal{R}_C - 1)(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+ \\ & + (P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R),\end{aligned}$$

this expression equals

$$\mathbb{1}_{\mathfrak{B}}\left[-D(t,\tau)(1-\mathcal{R}_C)(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+\right] + \underbrace{\mathbb{1}_{\mathfrak{B}} D(t,\tau)(P_\tau^{\mathrm{CDS}} - \mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)}_{=:\mathcal{M}_2}.$$

Similarly, the expression conditional on the event $\mathfrak{C}$ in (C.10) can be rewritten as

$$\mathbb{1}_{\mathfrak{C}}\left[D(t,\tau)(1-\mathcal{R}_I)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\right]+\underbrace{\mathbb{1}_{\mathfrak{C}}D(t,\tau)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)}_{=:\mathcal{M}_3}.$$

It is obvious that $\mathbb{1}_{\{\tau=\tau_R\}}P_\tau^{\text{CDS}}=0$, and therefore we observe $\mathcal{M}_2,\mathcal{M}_3$ and the last two expressions respectively regarding the event $\mathfrak{D}$ and $\mathfrak{E}$ in (C.10) together as follows,

$$D(t,\tau)\mathbb{1}_{\{\tau=\tau_R\}}(-\mathcal{L}_R)\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}-\mathbb{1}_{\mathfrak{E}}\right)+D(t,\tau)\mathbb{1}_{\{\tau\neq\tau_R\}}P_\tau^{\text{CDS}}\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}\right).$$

Recalling that $\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}-\mathbb{1}_{\mathfrak{E}}=0$, we can rewrite the terms inside $(C.10)$ as

$$\begin{aligned}\widehat{\Pi}(t,T)=\;&\mathbb{1}_{\mathfrak{B}}\left[-D(t,\tau)(1-\mathcal{R}_C)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+\right]\\&+\mathbb{1}_{\mathfrak{C}}\left[D(t,\tau)(1-\mathcal{R}_I)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\right]\\&+\mathbb{1}_{\mathfrak{A}}\left(-D(t,\tau)\mathcal{L}_R+\mathcal{M}_1\right)+\mathbb{1}_{\{\tau>T\}}\mathcal{M}_1\\&+\mathbb{1}_{\{\tau\neq\tau_R\}}\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}\right)\left(D(t,\tau)P_\tau^{\text{CDS}}+\mathcal{M}_1\right).\end{aligned}$$

Next, by comparing $\mathcal{M}_1$ with (3.1) we get

$$\begin{aligned}\widehat{P}_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)=\;&\mathbb{E}\left\{-\mathbb{1}_{\mathfrak{B}}D(t,\tau)(1-\mathcal{R}_C)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+\big|\,\mathcal{G}_t\right\}\\&+\mathbb{E}\left\{\mathbb{1}_{\mathfrak{C}}D(t,\tau)(1-\mathcal{R}_I)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\big|\,\mathcal{G}_t\right\}\\&+\mathbb{E}\left\{\left(\mathbb{1}_{\mathfrak{A}}+\mathbb{1}_{\{\tau>T\}}\right)\Pi(t,T)\big|\,\mathcal{G}_t\right\}\\&+\mathbb{E}\left\{\mathbb{1}_{\{\tau\neq\tau_R\}}\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}\right)\left(D(t,\tau)\mathbb{E}\{\Pi(\tau,T)|\mathcal{G}_\tau\}+\Pi(t,\tau)\right)\big|\,\mathcal{G}_t\right\}.\end{aligned}\tag{C.11}$$

Using $\mathbb{E}\{\mathbb{E}\{\cdot|\mathcal{G}_\tau\}|\mathcal{G}_t\}=\mathbb{E}\{\cdot|\mathcal{G}_t\}$ for $t<\tau$, the last expression in (C.11) equals

$$\mathbb{E}\left\{\mathbb{1}_{\{\tau\neq\tau_R\}}\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}\right)\Pi(\tau,T)\big|\,\mathcal{G}_t\right\}.$$

We see that

$$\mathbb{1}_{\{\tau\leq t\}}+\mathbb{1}_{\{\tau>T\}}+\mathbb{1}_{\mathfrak{A}}+\mathbb{1}_{\{\tau\neq\tau_R\}}\left(\mathbb{1}_{\mathfrak{B}}+\mathbb{1}_{\mathfrak{C}}-\mathbb{1}_{\mathfrak{D}}\right)=1$$

and the events in the terms of the sum are exclusive. We get finally

$$\begin{aligned}\widehat{P}_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)=&P_t^{\text{CDS}}(\mathcal{P},\mathcal{L}_R)-\mathbb{E}\left\{\mathbb{1}_{\mathfrak{B}}D(t,\tau)(1-\mathcal{R}_C)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^+\Big|\,\mathcal{G}_t\right\}\\&+\mathbb{E}\left\{\mathbb{1}_{\mathfrak{C}}D(t,\tau)(1-\mathcal{R}_I)(P_\tau^{\text{CDS}}-\mathbb{1}_{\{\tau=\tau_R\}}\mathcal{L}_R)^-\Big|\,\mathcal{G}_t\right\}\end{aligned}\tag{C.12}$$

□

Proofs of propositions in Section 5.1

The proof of Proposition 5.1.1

Proof: We calculate (5.12) as

$$\begin{aligned}&\operatorname{sech}^2(X_t)\kappa(\mu-X_t)dt-\operatorname{sech}^3(X_t)\sinh(X_t)\sigma^2dt+\operatorname{sech}^2(X_t)\sigma dW_t\\=\;&\operatorname{sech}^2(X_t)\kappa(\mu-X_t)dt-\operatorname{sech}^2(X_t)\frac{\sinh(X_t)}{\cosh(X_t)}\sigma dt+\operatorname{sech}^2(X_t)\sigma dW_t\\=\;&(1-\rho_t^2)\kappa(\mu-X_t)dt-(1-\rho_t^2)\rho_t\sigma^2dt+(1-\rho_t^2)\sigma dW_t.\end{aligned}$$

□

The proof of Proposition 5.1.3

Proof: Following the methodology described in Section 5.1.2 based on the Fokker-Planck equation, the stationary density function $f(\tilde{\rho})$ of the SCP (5.33) obviously satisfies

$$\frac{\partial}{\partial\tilde{\rho}}\Big((1-\tilde{\rho}^2)\big(\kappa(\mu-\tilde{\rho})\big)f(\tilde{\rho})\Big)=\frac{1}{2}\frac{\partial^2}{\partial\tilde{\rho}^2}\big((1-\tilde{\rho}^2)\sigma\big)^2f(\tilde{\rho}).\tag{C.13}$$

By solving the elliptic equation (C.13) we obtain the stationary density $f(\tilde{\rho})$ as

$$\begin{aligned} f(\tilde{\rho}) =& \frac{m}{2^{\frac{\kappa}{\sigma}}}(1+\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}+\frac{\kappa\mu}{\sigma^2}}(1-\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}-\frac{\kappa\mu}{\sigma^2}} \\ &+ \frac{n}{\tilde{\rho}^2-1}\left(\frac{1}{2}\right)^{\frac{2\sigma^2-\kappa}{\sigma^2}} F\left(1, \frac{2\sigma^2-2\kappa}{\sigma^2}, \frac{(-\mu-1)\kappa+2\sigma^2}{\sigma^2}, \frac{\tilde{\rho}}{2}+\frac{1}{2}\right) \end{aligned} \tag{C.14}$$

with the constants m, $n \in \mathbb{R}$ and the hypergeometric function F (see B.6). Next we need to fix the constants m and n in (C.14) to obtain the stationary density. Due to the mean reversion the stationary density $f(\tilde{\rho})$ must satisfy

$$\int_{-1}^{1} \tilde{\rho} f(\tilde{\rho})\, d\tilde{\rho} = \mu.$$

If we choose $\mu = 0$, we observe that the first term in (C.14) becomes

$$\frac{m}{2^{\frac{\kappa}{\sigma^2}}}(1+\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}}(1-\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}}, \tag{C.15}$$

which is obviously symmetric around $\tilde{\rho} = 0$, i.e. the condition (C.15) will be fulfilled for $n = 0$. In the sequel we assume that $n \equiv 0$ for all general $\mu \in (-1, 1)$ such that the transition density function (C.14) can be rewritten as

$$f(\tilde{\rho}) = \frac{m}{2^{\frac{\kappa}{\sigma^2}}}(1+\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}+\frac{\kappa\mu}{\sigma^2}}(1-\tilde{\rho})^{\frac{\kappa-2\sigma^2}{\sigma^2}-\frac{\kappa\mu}{\sigma^2}}. \tag{C.16}$$

To determine the value of m we can employ the basic property of a density function

$$\int_{-1}^{1} f(\tilde{\rho})\, d\tilde{\rho} = 1. \tag{C.17}$$

The constant m in (C.16) must be chosen such that the normalization condition (C.17) is always fulfilled. We set

$$a_\rho = \frac{\kappa - 2\sigma^2}{\sigma^2}, \quad b_\rho = \frac{\kappa\mu}{\sigma^2}, \tag{C.18}$$

and substitute it into (C.16) to obtain

$$f(\tilde{\rho}) = \frac{m}{2^{\frac{\kappa}{\sigma^2}}}(1+\tilde{\rho})^{a_\rho+b_\rho}(1-\tilde{\rho})^{a_\rho-b_\rho}. \tag{C.19}$$

As long as

$$a_\rho \pm b_\rho > -1, \tag{C.20}$$

the integral

$$\int_{-1}^{1}(1+\tilde{\rho})^{a_\rho+b_\rho}(1-\tilde{\rho})^{a_\rho-b_\rho}\,d\tilde{\rho}$$

has the solution

$$\begin{aligned} M :=& \frac{\Gamma(1+a_\rho-b_\rho)F(1,-a_\rho-b_\rho,2+a_\rho-b_\rho,-1)}{\Gamma(2+a_\rho-b_\rho)} \\ &+ \frac{\Gamma(1+a_\rho+b_\rho)F(1,-a_\rho+b_\rho,2+a_\rho+b_\rho,-1)}{\Gamma(2+a_\rho+b_\rho)}, \end{aligned} \tag{C.21}$$

with the hypergeometric function F defined in (B.6) and the Gamma function Γ.

We check the condition (C.20) as follows:

$$a+b>-1 \Leftarrow \tfrac{\kappa-2\sigma^2}{\sigma^2}+\tfrac{\kappa\mu}{\sigma^2}>-1 \Leftarrow \kappa(1+\mu)>\sigma^2 \Leftarrow \kappa>\tfrac{\sigma^2}{1+\mu},$$

$$a-b>-1 \Leftarrow \tfrac{\kappa-2\sigma^2}{\sigma^2}-\tfrac{\kappa\mu}{\sigma^2}>-1 \Leftarrow \kappa(1-\mu)>\sigma^2 \Leftarrow \kappa>\tfrac{\sigma^2}{1-\mu}.$$

This is to say that the condition (C.20) always holds as long as

$$\kappa > \frac{\sigma^2}{1\pm\mu}. \tag{C.22}$$

Under the condition (C.22), the constant m can be determined as

$$m = \frac{2^{\frac{\kappa}{\sigma^2}}}{M}. \tag{C.23}$$

Finally, we obtain the transition density function in a closed form as

$$f(\tilde{\rho}) = \frac{(1+\tilde{\rho})^{a+b}(1-\tilde{\rho})^{a-b}}{M}, \tag{C.24}$$

with a_ρ, b_ρ defined in (C.18) and M in (C.21). □

Proofs and Approximations in Section 5.3

The proof of Lemma 5.3.2

Proof: Recall the ODE system in Lemma 5.3.1

$$B'(u,\tau) = 0, \quad B(u,0) = iu, \tag{C.25}$$

$$C'(u,\tau) = \sigma_\nu \mathbb{E}[\nu_t] B(u,\tau) D(u,\tau) - \kappa_\rho C(u,\tau), \quad C(u,0) = 0, \tag{C.26}$$

$$D'(u,\tau) = \frac{1}{2}B^2(u,\tau) + \frac{1}{2}\sigma_\nu^2 D(u,\tau) - \frac{1}{2}B(u,\tau) - \kappa_\nu D(u,\tau), \quad D(u,0) = 0, \tag{C.27}$$

$$\begin{aligned} A'(u,\tau) = {} & (B(u,\tau)-1)r + \kappa_\nu \mu_\nu D(u,\tau) + \kappa_\rho \mu_\rho C(u,\tau) \\ & + \frac{1}{2}\sigma_\rho^2 C^2(u,\tau) + \sigma_\rho \rho_2 \mathbb{E}[\sqrt{\nu_t}] B(u,\tau) C(u,\tau), \quad A(u,0) = 0. \end{aligned} \tag{C.28}$$

Straightforwardly, due to the final condition $B(u,0) = iu$ we obtain $B(u,\tau) = iu$. We consider first the following Riccati-type equation:

$$\begin{aligned} \frac{\partial D(u,\tau)}{\partial \tau} &= \frac{1}{2}B^2(u,\tau) + \frac{1}{2}\sigma_\nu^2 D(u,\tau) - \frac{1}{2}B(u,\tau) - \kappa_\nu D(u,\tau), \quad D(u,0) = 0, \\ H_1(u,\tau) &= (iu-1)r\tau + \kappa_\nu \mu_\nu \int_0^\tau D(u,s)\, ds, \quad H_1(u,0) = 0, \end{aligned}$$

which has the same form as those in [58] so that we can gain the solution given by

$$D(u,\tau) = \frac{\kappa_\nu - D_1}{\sigma_\nu^2} \cdot \frac{1 - \mathrm{e}^{-D_1\tau}}{1 - D_2 \mathrm{e}^{-D_1\tau}}, \tag{C.29}$$

$$H_1(u,\tau) = (iu-1)r\tau + \frac{\kappa_\nu \mu_\nu}{\sigma_\nu^2}\left((\kappa_\nu - D_1)\tau - 2\ln\left(\frac{1 - D_2\mathrm{e}^{-D_1\tau}}{1 - D_2}\right)\right), \tag{C.30}$$

where $D_1 = \sqrt{\kappa_\nu^2 + \sigma_\nu^2(u^2 + iu)}$ and $D_2 = \frac{\kappa_\nu - D_1}{\kappa_\nu + D_1}$.

We turn to (C.26) where

$$\mathbb{E}[\nu_t] = (\nu_0 - \mu_\nu)\mathrm{e}^{-\kappa_\nu(T-\tau)} + \mu_\nu. \tag{C.31}$$

To find its analytical solution we use the approximation

$$1 - \mathrm{e}^{-l_1\tau} \approx \frac{1 - \mathrm{e}^{-D_1\tau}}{1 - D_2\mathrm{e}^{-D_1\tau}}, \tag{C.32}$$

where l_1 is defined in (C.47). The detailed information and the measure of the quality of this approximation can be found in Figure C.1. We can thus rewrite (C.29) as

$$D(u,\tau) = \frac{\kappa_\nu - D_1}{\sigma_\nu^2} \cdot (1 - \mathrm{e}^{-l_1\tau}), \tag{C.33}$$

and set

$$C_1 := iu\frac{\kappa_\nu - D_1}{\sigma_\nu^2}. \tag{C.34}$$

Sequentially, (C.26) can be rewritten as

$$C'(u,\tau) = \sigma_\nu C_1\big((\nu_0 - \mu_\nu)\mathrm{e}^{-\kappa_\nu(T-\tau)} + \mu_\nu\big) \cdot (1 - \mathrm{e}^{-l_1\tau}) - \kappa_\rho C(u,\tau),\ C(u,0) = 0, \tag{C.35}$$

which has an analytical solution, although its calculation is a bit tedious but straightforward. We obtain

$$\begin{aligned} C(u,\tau) = \frac{C_1(\mu_\nu - \nu_0)}{\kappa_\nu + \kappa_\rho - l_1}\mathrm{e}^{(\kappa_\nu - l_1)\tau - \kappa_\nu T} + \frac{C_1(\nu_0 - \mu_\nu)}{\kappa_\nu + \kappa_\rho}\mathrm{e}^{\kappa_\nu(\tau - T)} + \frac{C_1\mu_\nu}{\kappa_\rho} \\ - \frac{C_1\mu_\nu}{\kappa_\rho - l_1}\mathrm{e}^{-l_1} + C_1C_2\mathrm{e}^{-\kappa_\rho\tau}, \end{aligned} \tag{C.36}$$

where l_1 is defined in (C.47), C_1 is defined in (C.34) and C_2 is given by

$$C_2 := \frac{\mu_\nu - \nu_0}{\kappa_\nu + \kappa_\rho - l_1}\mathrm{e}^{-\kappa_\nu T} + \frac{\nu_0 - \mu_\nu}{\kappa_\nu + \kappa_\rho}\mathrm{e}^{-\kappa_\nu T} - \frac{\mu_\nu}{\kappa_\rho} + \frac{1}{\kappa_\rho - l_1}. \tag{C.37}$$

Finally, we rewrite (C.28) with approximations as

$$A(u,\tau) = H_1(u,\tau) + \underbrace{(\kappa_\rho\mu_\rho + m\sigma_\rho\rho_2 ui)}_{:=\alpha} H_2(u,\tau) + \underbrace{n\sigma_\rho\rho_2 ui}_{:=\beta} H_3(u,\tau) + \frac{\sigma_\rho^2}{2}H_4(u,\tau), \tag{C.38}$$

for solving which we only need to calculate the following integrals

$$H_2(u,\tau) = \int_0^\tau C(u,s)\,ds, \quad H_2(u,0) = 0, \tag{C.39}$$

$$H_3(u,\tau) = \int_0^\tau \mathrm{e}^{-(T-\tau)l}C(u,s)\,ds, \quad H_3(u,0) = 0, \tag{C.40}$$

$$H_4(u,\tau) = \int_0^\tau C^2(u,s)\,ds, \quad H_4(u,0) = 0, \tag{C.41}$$

where the constants m, n, and l are defined in (5.86) - (5.87). The calculation of the integrals above is straightforward but rather tedious. □

The proof of Lemma 5.3.4

Proof: As indicated before, the solutions of $\tilde{B}(u,\tau)$, $\tilde{C}(u,\tau)$ and $\tilde{D}(u,\tau)$ are the same as B, C and D in the HO model. We consider now only

$$\begin{aligned} A'(u,\tau) =& (\tilde{B}(u,\tau) - 1)r + \kappa_\nu\mu_\nu\tilde{D}(u,\tau) + \kappa_\rho\mu_\rho\tilde{C}(u,\tau) + \frac{1}{2}\sigma_\rho^2\mathbb{E}[1-\rho_t^2]\tilde{C}^2(u,\tau) \\ &+ \sigma_\rho\rho_2\mathbb{E}[\sqrt{\nu_t}]\mathbb{E}[\sqrt{1-\rho_t^2}]\tilde{B}(u,\tau)\tilde{C}(u,\tau), \quad \tilde{A}(u,0) = 0. \end{aligned} \tag{C.42}$$

By substituting the approximations of $\mathbb{E}[\rho_t^2]$, $\mathbb{E}[\sqrt{1-\rho_t^2}]$ and $\mathbb{E}[\sqrt{\nu_t}]$ into (C.42) we obtain

$$\begin{aligned}\tilde{A}'(u,\tau) =&(\tilde{B}(u,\tau)-1)r+\kappa_\nu\mu_\nu\tilde{D}(u,\tau)+\kappa_\rho\mu_\rho\tilde{C}(u,\tau)\\ &+\sigma_\rho\rho_2(m+n\mathrm{e}^{-l(T-\tau)})(\mathrm{e}^{-m_3(T-\tau)}+b_3\mathrm{e}^{-n_3(T-\tau)}+a_3)\tilde{B}(u,\tau)\tilde{C}(u,\tau)\\ &+\frac{1}{2}\sigma_\rho^2\tilde{C}^2(u,\tau)(1-\mathrm{e}^{-m_2(T-\tau)}-b_2\mathrm{e}^{-n_2(T-\tau)}-a_2),\quad \tilde{A}(u,0)=0,\end{aligned}\tag{C.43}$$

which can be reformulated as

$$\begin{aligned}\tilde{A}(u,\tau) =&\tilde{H}_1(u,\tau)+(\kappa_\rho\mu_\rho+a_3m\sigma_\rho\rho_2ui)\tilde{H}_2(u,\tau)+a_3n\sigma_\rho\rho_2ui\tilde{H}_3(u,\tau)\\ &+\frac{\sigma_\rho^2}{2}(1-a_2)\tilde{H}_4(u,\tau)+b_3m\sigma_\rho\rho_2ui\tilde{H}_5(u,\tau)+m\sigma_\rho\rho_2ui\tilde{H}_6(u,\tau)\\ &+b_3n\sigma_\rho\rho_2ui\tilde{H}_7(u,\tau)+n\sigma_\rho\rho_2ui\tilde{H}_8(u,\tau)-\frac{\sigma_\rho^2}{2}\tilde{H}_9(u,\tau)-\frac{b_2\sigma_\rho^2}{2}\tilde{H}_{10}(u,\tau)\end{aligned}$$

with the following integrals

$$\begin{aligned}&\tilde{H}_1(u,\tau)=(iu-1)r\tau+\kappa_\nu\mu_\nu\int_0^\tau\tilde{D}(u,s)\,ds,\ \tilde{H}_2(u,\tau)=\int_0^\tau\tilde{C}(u,s)\,ds,\\ &\tilde{H}_3(u,\tau)=\int_0^\tau\mathrm{e}^{-(T-\tau)l}\tilde{C}(u,s)\,ds,\ \tilde{H}_4(u,\tau)=\int_0^\tau\tilde{C}^2(u,s)\,ds,\\ &\tilde{H}_5(u,\tau)=\int_0^\tau\mathrm{e}^{-(T-\tau)n_3}\tilde{C}(u,s)\,ds,\ \tilde{H}_6(u,\tau)=\int_0^\tau\mathrm{e}^{-(T-\tau)m_3}\tilde{C}(u,s)\,ds,\\ &\tilde{H}_7(u,\tau)=\int_0^\tau\mathrm{e}^{-(T-\tau)(l+n_3)}\tilde{C}(u,s)\,ds,\ \tilde{H}_8(u,\tau)=\int_0^\tau\mathrm{e}^{-(T-\tau)(l+m_3)}\tilde{C}(u,s)\,ds,\\ &\tilde{H}_9(u,\tau)=\int_0^\tau\mathrm{e}^{-m_2(T-\tau)}\tilde{C}^2(u,s)\,ds,\ \tilde{H}_{10}(u,\tau)=\int_0^\tau\mathrm{e}^{-n_2(T-\tau)}\tilde{C}^2(u,s)\,ds,\\ &\tilde{H}_i(u,0)=0\text{ for }i=1\cdots10.\end{aligned}$$

It is easy to see that $\tilde{H}_1$, $\tilde{H}_2$, $\tilde{H}_3$ and $\tilde{H}_4$ are respectively equal to H_1, H_2, H_3 and H_4 which have been given before. Besides, the solutions of $\tilde{H}_5$, $\tilde{H}_6$, $\tilde{H}_7$, $\tilde{H}_8$ can be directly obtained by adopting the solution of $\tilde{H}_3$, as they have only different constant coefficients in the exponential function. For simplicity of notation, we let this coefficient to be a variable of $\tilde{H}_3$, namely $\tilde{H}_3(u,\tau,l)$. The solutions of $\tilde{H}_5$, $\tilde{H}_6$, $\tilde{H}_7$ and $\tilde{H}_8$ can thus be

immediately given by $\tilde{H}_3(u,\tau,n_3)$, $\tilde{H}_3(u,\tau,m_3)$, $\tilde{H}_3(u,\tau,(l+n_3))$ and $\tilde{H}_3(u,\tau,(l+m_3))$, respectively. Now, only the integral in the following form

$$\tilde{H}(u,\tau,y) = \int_0^\tau \mathrm{e}^{-y(T-\tau)} C^2(u,s)\, ds, \quad \tilde{H}(u,0) = 0$$

need to be calculated. The calculation is straightforward, however, rather tedious. It is obvious that $\tilde{H}(u,\tau,m_2) = \tilde{H}_9(u,\tau)$ and $\tilde{H}(u,\tau,n_2) = \tilde{H}_{10}(u,\tau)$. Finally, by defining $\zeta := \sigma_\rho \rho_2 u i$, $A(u,\tau)$ can be rewritten as

$$\begin{aligned} A(u,\tau) =& \tilde{H}_1(u,\tau) + (\kappa_\rho \mu_\rho + a_3 m\zeta)\tilde{H}_2(u,\tau) + a_3 n\zeta \tilde{H}_3(u,\tau,l) + b_3 m\zeta \tilde{H}_3(u,\tau,n_3) \\ &+ m\zeta \tilde{H}_3(u,\tau,m_3) + b_3 n\zeta \tilde{H}_3(u,\tau,(l+n_3)) + n\zeta \tilde{H}_3(u,\tau,(l+m_3)) \\ &+ \frac{\sigma_\rho^2}{2}(1-a_2)\tilde{H}_4(u,\tau) - \frac{\sigma_\rho^2}{2}\tilde{H}(u,\tau,m_2) - \frac{b_2\sigma_\rho^2}{2}\tilde{H}(u,\tau,n_2). \end{aligned}$$

□

Approximation I

We match $f_1(\tau) := \frac{1-\mathrm{e}^{-D_1\tau}}{1-D_2\mathrm{e}^{-D_1\tau}} \approx m_1 + n_1\mathrm{e}^{-l_1\tau} := \tilde{f}_1(\tau)$ for $\tau \to 0$, $\tau \to \infty$, $\tau \to 1$:

$$\lim_{\tau\to 0} f_1(\tau) = 0 = m_1 + n_1 = \lim_{\tau\to 0} \tilde{f}_1(\tau), \tag{C.44}$$

$$\lim_{\tau\to \infty} f_1(\tau) = 1 = m_1 = \lim_{\tau\to \infty} \tilde{f}_1(\tau), \tag{C.45}$$

$$\lim_{\tau\to 1} f_1(\tau) = \frac{1-\mathrm{e}^{-D_1}}{1-D_2\mathrm{e}^{-D_1}} = 1 - \mathrm{e}^{-l_1} = \lim_{\tau\to 1} \tilde{f}_1(\tau), \tag{C.46}$$

which give

$$m_1 = 1, \quad n_1 = -1, \quad l_1 = -\ln\left(\frac{\mathrm{e}^{-D_1} - D_2\mathrm{e}^{-D_1}}{1-D_2\mathrm{e}^{-D_1}}\right). \tag{C.47}$$

In order to measure the quality of this approximation we compare $f_1(\tau)$ to $\tilde{f}_1(\tau)$ for different randomly chosen parameters in Figure C.1.

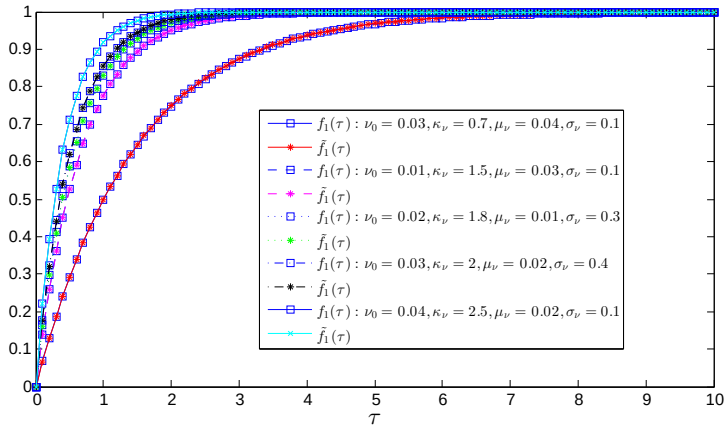


Figure C.1: The quality of the approximation $\tilde{f}_1(\tau)$ versus the original $f_1(\tau)$ for randomly chosen parameters.

Approximation II

We match $f_2(t) := \mathbb{E}[\rho_t^2] \approx \mathrm{e}^{-m_2 t} + b_2 \mathrm{e}^{-n_2 t} + a_2 := \tilde{f}_2(t)$ for $t \to 0$, $t \to \frac{1}{2}$, $t \to 1$, $t \to \infty$ as follows:

$$\lim_{t\to\infty} f_2(t) = \frac{(\sigma_\rho^2 + \kappa_\rho)(\sigma_\rho^2 + 2\kappa_\rho\mu_\rho^2)}{\sigma_\rho^4 + 3\kappa_\rho\sigma_\rho^2 + 2\kappa_\rho^2} = a_2 = \lim_{t\to\infty} \tilde{f}_2(t), \tag{C.48}$$

$$\lim_{t\to 0} f_2(t) = \rho_0^2 = 1 + b_2 + a_2 = \lim_{t\to 0} \tilde{f}_2(t), \tag{C.49}$$

$$\lim_{t\to\frac{1}{2}} f_2(t) = f_2(0.5) = \mathrm{e}^{-\frac{m_2}{2}} + b_2 \mathrm{e}^{-\frac{n_2}{2}} + a_2 = \lim_{t\to\frac{1}{2}} \tilde{f}_2(t), \tag{C.50}$$

$$\lim_{t\to 1} f_2(t) = f_2(1) = \mathrm{e}^{-m_2} + b_2 \mathrm{e}^{-n_2} + a_2 = \lim_{t\to 1} \tilde{f}_2(t). \tag{C.51}$$

From (C.48) and (C.49) one obtains directly $a_2 = \frac{(\sigma_\rho^2+\kappa_\rho)(\sigma_\rho^2+2\kappa_\rho\mu_\rho^2)}{\sigma_\rho^4+3\kappa_\rho\sigma_\rho^2+2\kappa_\rho^2}$ and $b_2 = \rho_0^2 - a_2 - 1$. Then one needs to solve the system of equations (C.50) and (C.51) to find m_2 and n_2 which has been given in (5.121). Like in the last section, we compare $f_2(\tau)$ to $\tilde{f}_2(\tau)$ for different randomly chosen parameters to measure the quality of the proposed approximation.

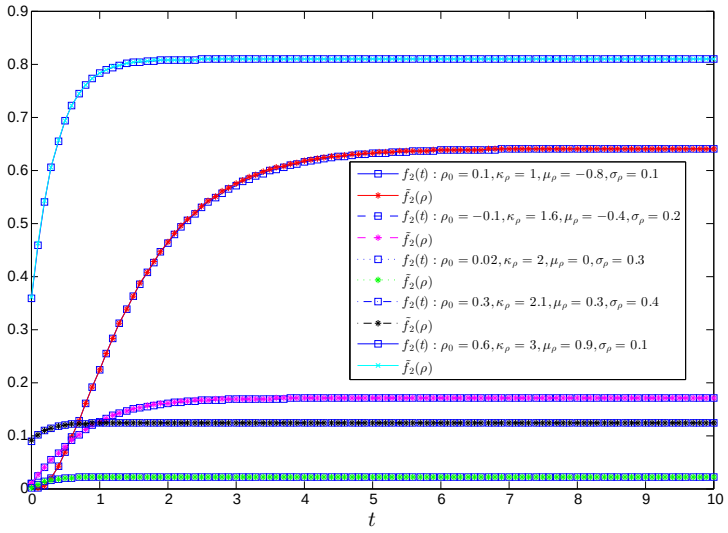


Figure C.2: The quality of the approximation $\tilde{f}_2(t)$ versus the original $f_2(t)$ for randomly chosen parameters.

Approximation III

We match $f_3(t) := \mathbb{E}[\sqrt{1-\rho_t^2}] \approx \mathrm{e}^{-m_3 t} + b_3\mathrm{e}^{-n_3 t} + a_3 := \tilde{f}_3(t)$ for $t \to 0$, $t \to \frac{1}{2}$, $t \to 1$, $t \to \infty$ as follows:

$$\lim_{t\to\infty} f_3(t) = \sqrt{1 - \frac{a_2 - \mu_\rho^4}{1-\mu_\rho^2}} = a_3 = \lim_{t\to\infty} \tilde{f}_2(t), \tag{C.52}$$

$$\lim_{t\to 0} f_3(t) = \sqrt{1-\rho_0^2} = 1 + b_3 + a_3 = \lim_{t\to 0} \tilde{f}_3(t), \tag{C.53}$$

$$\lim_{t\to\frac{1}{2}} f_3(t) = f_3(0.5) = \mathrm{e}^{-\frac{m_3}{2}} + b_3\mathrm{e}^{-\frac{n_3}{3}} + a_3 = \lim_{t\to\frac{1}{2}} \tilde{f}_3(t), \tag{C.54}$$

$$\lim_{t\to 1} f_3(t) = f_3(1) = \mathrm{e}^{-m_3} + b_3\mathrm{e}^{-n_3} + a_3 = \lim_{t\to 1} \tilde{f}_3(t). \tag{C.55}$$

From (C.52) and (C.53) one obtains directly

$$a_3 = \sqrt{1 - \frac{(\sigma_\rho^2+\kappa_\rho)(\sigma_\rho^2+2\kappa_\rho\mu_\rho^2) - \mu_\rho^4(\sigma_\rho^4+3\kappa_\rho\sigma_\rho^2+2\kappa_\rho^2)}{(1-\mu_\rho^2)(\sigma_\rho^4+3\kappa_\rho\sigma_\rho^2+2\kappa_\rho^2)}}$$

and $b_3 = \sqrt{1-\rho_0^2} - a_3 - 1$. Further, we solve the system of equations (C.54) and (C.55) to find m_3 and n_3 which has been given in (5.132). The comparison of $f_3(\tau)$ with $\tilde{f}_3(\tau)$ and

the measure of quality of the approximation for different randomly chosen parameters is exhibited in Figure C.3.

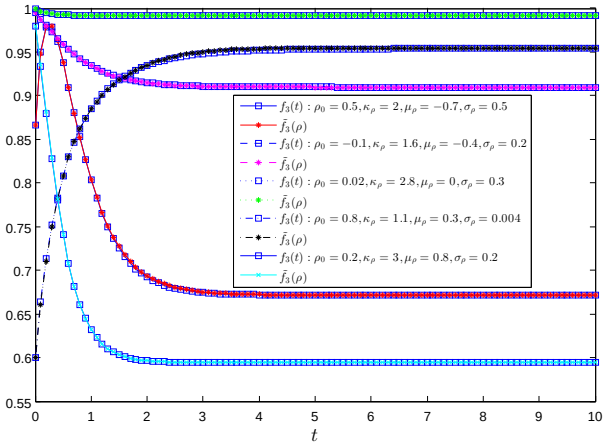


Figure C.3: The quality of the approximation $\tilde{f}_3(t)$ versus the original $f_3(t)$ for randomly chosen parameters.

Bibliography

[1] C. Alexander. *Market Models: A Guide to Financial Data Analysis.* John Wiley & Sons, Ltd., Chichester, West Sussex, UK, 2003.

[2] L. Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Finance,* **11**, pages 1–42, 2008.

[3] S. Assefa, T.R. Bielecki, S. Crépey, and M. Jeanblanc. CVA computation for counterparty risk assessment in credit portfolios. In *Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings, and Liquidity*, Bloomberg, 2011.

[4] S. Azizpour and K. Giesecke. Self-exciting corporate defaults: contagion vs. frailty. Working paper, Aug. 2008.

[5] G. Bakshi and D. Madan. Spanning and derivative-security valuation. *J. Fin. Econ.,* **55**, pages 205–238, 2000.

[6] E. Benhamou, E. Gobet, and M. Miri. Time Dependent Heston Model. *SIAM J. Fin. Math.,* **1(1)**, pages 289–325, 2010.

[7] T. R. Bielecki, D. Brigo, and F. Patras. *Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings, and Liquidity.* Wiley, 2011.

[8] T. R. Bielecki and M. Rutkowski. *Credit risk: modelling, valuation and hedging.* Springer, 2002.

[9] T.R. Bielecki, A. Cousin, S. Crépey, and A. Herbertsson. Pricing and Hedging Portfolio Credit Derivatives in a Bottom-up Model with Simultaneous Defaults. Working paper, Evry University, 2011.

[10] F. Black and J. C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *J. Finance,* **31**, pages 351–367, 1976.

[11] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *J. Polit. Econ.,* **81(3)**, pages 637–654, May-Jun 1973.

[12] A. W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis.* New York: Oxford University Press Inc., 1997.

[13] M. J. Brennen and E. S. Schwartz. Convertible bonds: valuation and optimal strategies for call and conversion. *J. Finance,* **32**, pages 1699–1715, 1977.

[14] D. Brigo and A. Alfonsi. Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model. *Financ. Stoch.,* **9(1)**, 2005.

[15] D. Brigo and A. Capponi. Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps. Working paper, November 2009.

[16] D. Brigo and A. Capponi. Bilateral counterparty risk with application to CDSs. *RISK,* **23(3)**, pages 85–90, 2010.

[17] D. Brigo, A. Capponi, and A. Pallavicini. Arbitrage-free bilateral counterparty risk valuation under collateralization and application to Credit Default Swaps. To appear in *Math. Financ.*, 2012.

[18] D. Brigo and N. El-Bachir. An exact formula for default swaptions' pricing in the SSRJD stochastic intensity model. *Math. Financ.,* **20(3)**, pages 365–382, 2010.

[19] D. Brigo and F. Mercurio. *Interest rate models-theory and practice*. Springer, 2006.

[20] D. Brigo, M. Morini, and A. Pallavicini. *Counterparty credit risk, collateral and funding*. Wiley, 2013.

[21] D. Brigo, A. Pallavicini, and V. Papatheodorou. Arbitrage-free valuation of bilateral counterparty risk for interest-rate products: Impact of volatilities and correlations. *Int. J. Theoret. Appl. Fin.,* **14 (6)**, pages 773–802, 2011.

[22] D. Brigo, A. Pallavicini, and R. Torresetti. *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models*. Wiley, 2010.

[23] M. F. Bru. Wishart Processes. *J. Theor. Probab.,* **4 (4)**, pages 725–743, 1991.

[24] A. Buraschi, A. Cieslak, and F. Trojani. Correlation Risk and the Term Structure of Interest Rates. Working paper, Tanaka Business School, Imperial College London and University of St. Gallen, June 2007.

[25] A. Buraschi, P. Porchia, and F. Trojani. Correlation Risk and Optimal Portfolio Choice. *J. Finance,* **65**, pages 393–420, February 2010.

[26] A. Capponi. *Credit risk and nonlinear filtering: computational aspects and empirical evidence*. Dissertation, California Institute of Technology, 2009.

[27] P. Carr and D. Madan. Option Valuation Using the Fast Fourier Transform. *J. Comput. Finance,* **2**, pages 61–73, 1999.

[28] G. Cesari, J. Aquilina, N. Charpillon, Z. Filipović, G. Lee, and L. Manda. *Modelling, pricing, and hedging counterparty credit exposure*. Springer, 2009.

[29] H. Chen and S. Joslin. Generalized Transform Analysis of Affine Processes and Applications in Finance. *Rev. Fin. Stud.,* **25(7)**, pages 2225–2256, 2012.

[30] K. Chourdakis. Option pricing using the fractional FFT. *J. Comput. Finance,* **8**, pages 1–18, 2005.

[31] P. Christoffersen, S. Heston, and K. Jacobs. The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. *Manage. Sci.,* **55**, pages 1914–1932, 2009.

[32] I. A. Cooper and A. S. Mello. The default risk of swaps. *J. Finance,* **46**, pages 597–620, 1983.

[33] J. Cox, J. Ingersoll, and S. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica,* **53**, pages 385–408, 1985.

[34] S. Crépey and T. R. Bielecki. *Counterparty Risk and Funding: A Tale of Two Puzzles*. CRC Press, 2014.

[35] J. Driessen, P. J. Maenhout, and G. Vilkov. The Price of Correlation Risk: Evidence from Equity Options. *J. Finance,* **64(3)** , pages 1377–1406, June 2009.

[36] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.,* **13(3)**, pages 984–1053, 2003.

[37] D Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica,* **68**, pages 1343–1376, 2000.

[38] D. Duffie and S. Schaefer. *Credit risk: pricing, measurement, and management*. Princeton University Press, 2003.

[39] D. Duffie, M. Schroder, and C. Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Ann. Appl. Probab.,* **6**, pages 1075–1090, 1996.

[40] D. Duffie and K. J. Singleton. Modeling term structure of defaultable bonds. *Rev. Financ. Stud.,* **12**, pages 687–719, 1996.

[41] K. Düllmann, J. Küll, and M. Kunish. Estimating asset correlations from stock prices or default rates - which method is superior? Working paper, Deutsche Bundesbank, 2008.

[42] A. Elices. Affine Concatenation. *Wilmott Journal,* **1(3)**, pages 155–162, 2009.

[43] R. J. Elliott and P. E. Kopp. *Mathematics of Financial Markets*. Springer Science + Business Media Inc., 2005.

[44] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica,* **50(4)**, pages 987–1008, 1982.

[45] R. F. Engle. Dynamic conditional correlation: a simple class of multivariate GARCH. *J. Business Econ. Stat.,* **20(3)**, pages 339–350, 2002.

[46] L. Euler. Institutiones Calculi Integralis. *St. Petersburg,* **1**, 1768, *reprinted in Opera Omnia,* **11**, 1913.

[47] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM J. Sci. Comput.,* **31**, pages 826–848, 2008.

[48] H. Föllmer and S. Alexander. *Stochastic finance: an introduction in discrete time.* De Gruyter, 2011.

[49] J. D. Fonseca, M. Grasselli, and F. Lelpo. Estimating the Wishart Affine Stochastic Correlation Model Using the Empirical Characteristic Function. Working paper, Auckland University of Technology, University of Padova and Université Paris I Panthéon-Sorbonne, 2008.

[50] J. C. G. Franco. Maximum likelihood estimation of mean reverting processes. Working paper, 2003.

[51] D. Galai and R. W. Masulis. The option pricing model and the risk factor of stock. *J. Finan. Econom.,* **3**, pages 53–81, 1976.

[52] W. Gander and W. Gautschi. Adaptive Quadrature – Revisited. *BIT,* **40**, pages 84–101, 2000.

[53] R. Geske. The valuation of corporate liabilities as compound options. *J. Finan. Quant. Anal.,* **12**, pages 541–552, 1977.

[54] C. Gourieroux and R. Sufana. Derivative Pricing With Wishart Multivariate Stochastic Volatility. *J. Bus. Econ. Statist.,* **28(3)**, pages 438–451, 2010.

[55] J. Gregory. *Counterparty credit risk and credit value adjustment.* Wiley, 2012.

[56] L. A. Grzelak and C. W. Oosterlee. On the Heston model with Stochastic Interest Rate. *SIAM J. Fin. Math.,* **2**, pages 255–286, 2011.

[57] M. Günther and A. Jüngel. *Finanzderivate mit Matlab: Mathematische Modellierung und numerische Simulation.* Springer, 2010.

[58] S. L. Heston. A Closed-Form Solution for options with Stochastic Volatility with Applications to Bond and Currency Options. *Rev. Fin. Stud.,* **6(2)**, pages 327–343, 1993.

[59] T. S. Ho, R. C. Stapleton, and M. G. Subrahmanyam. Correlation risk, cross-market derivative products and portfolio performance. *Europ. Finan. Manage.,* **1(2)**, pages 105–124, July 1995.

[60] J. C. Hull. *Options, Futures and Other Derivatives.* Pearson, 2011.

[61] K. Itô. Stochastic Integral. *Proc. Imperial Acad. Tokyo,* **20**, pages 519–524, 1944.

[62] R. A. Jarrow and D. B. Madan. Pricing derivatives on financial securities subject to credit risk. *Math. Finance,* **5**, pages 311–336, 1995.

[63] C. Kahl and P Jäckel. Not-so-complex Logarithms in the Heston Model. *Wilmott Magazine*, pages 94–103, Sep. 2005.

[64] I Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Springer, 1998.

[65] J. Kienitz and D. Wetterau. *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*. Wiley, 2012.

[66] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1992.

[67] R. W. Kolb and J. A. Overdahl. *Financial Derivatives: Pricing and Risk Management*. Wiley, 2010.

[68] C. N. V. Krishnan, R. Petkova, and P. Ritchken. Correlation Risk. Working paper, Weatherhead School of Management, Case Western Reserve University and Texas A&M University, July 2007.

[69] D. Lando. On Cox processes and credit-risky securities. *Rev. Derivatives Res.,* **2**, pages 99–120, 1998.

[70] R. Lee. Option Pricing by Transform Methods: Extensions, Unification, and Error Control. *J. Comput. Finance,* **7**, pages 51–86, 2005.

[71] H. E. Leland. Corporate debt value, bond covenants, and optimal capital structure. *J. Finance,* **49**, pages 1213–1252, 1994.

[72] H. E. Leland and K. Toft. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *J. Finance,* **51**, pages 987–1019, 1996.

[73] E. Lindström, H. Madsen, and J. N. Nielsen. *Statistics for Finance*. CRC Press, 2015.

[74] R. Litterman and T. Iben. Corporate bond valuation and the term structure of credit spreads. *J. Portfolio Management,* **17(3)**, pages 52–64, 1991.

[75] J. London. *Modeling Derivatives Applications in Matlab, C++ and Excel*. Pearson Financial Times, 2006.

[76] F. A. Longstaff. How much can marketability affect security values. *J. Finance,* **50**, pages 1767–1774, 1995.

[77] R. Lord and C. Kahl. Why the rotation count algorithm works. Working paper, Erasmus University Rotterdam and University of Wuppertal, July 2006.

[78] R. Lord and C. Kahl. Optimal Fourier inversion in semi-analytical option pricing. *J. Comput. Finance,* **10(4)**, pages 1–30, 2007.

[79] J. Ma. Pricing Foreign Equity Options with Stochastic Correlation and Volatility. *Ann. Econ. Fin.,* **10(2)**, pages 303–327, 2009.

[80] D. B. Madan and H. Unal. Pricing the risk of default. *Rev. Derivatives Res.,* **2**, pages 121–160, 1998.

[81] M. Martin, M. Reitz, and C. Wehn. *Kreditderivate und Kreditrisikomodelle: Eine mathematische Einführung.* Springer, 2014.

[82] G. Maruyama. Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo,* **4**, pages 48–90, 1955.

[83] A. J. McNeil, R. Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, 2005.

[84] G. Meissner. *Correlation risk modeling and management.* Wiley, 2014.

[85] R. C. Merton. On the pricing of corporate debt: the risk structure of interest rate. *J. Finance,* **29(2)**, pages 449–470, 1974.

[86] S. Mikhailov and U. Nögel. Heston's Stochastic Volatility Model: Implementation, Calibration and some Extensions. *Wilmott Magazine*, pages 74–79, 2003.

[87] G. N. Milstein. Approximate integration of stochastic differential equations. *Theory Probab. Appl.,* **19**, pages 557–562, 1974.

[88] P. Mueller, A. Stathopoulos, and A. Vedolin. International Correlation Risk. Working paper, London School of Economics and USC Marshall School of Business, April 2012.

[89] R. B. Nelsen. *An introduction to copulas.* Springer, 2006.

[90] C. Pitts and M. Selby. The pricing of corporate debt: A further note. *J. Finance,* **38**, pages 1311–1313, 1983.

[91] G Pye. Gauging the default premium. *Finan. Analysts J.,* **30(1)**, pages 49–52, 1974.

[92] K. Ramaswamy and S. M. Sundaresan. The valuation of floating-rate instruments, theory and evidence. *J. Finan. Econom.,* **17**, pages 251–272, 1986.

[93] H. Risken. *The Fokker-Planck equation.* Springer, 1989.

[94] A. Roncoroni, G. Fusai, and M. Cummins. *Handbook of Multi-Commodity Markets and Products: Structuring, Trading and Risk Management.* Wiley, 2015.

[95] F. D. Rouah. *The Heston Model and its Extensions in Matlab and C.* John Wiley & Sons, Inc., Hoboken, NJ, USA, 2013.

[96] R. Schöbel and J. Zhu. Stochastic Volatility With an Ornstein Uhlenbeck Process: An Extension. *Europ. Fin. Rev.,* **3**, pages 23–46, 1999.

[97] P. J. Schönbucher. Term structure modelling of defaultable bonds. *Rev. Derivatives Res.,* **2**, pages 161–192, 1998.

[98] P. J. Schönbucher. *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley, 2003.

[99] P. J. Schönbucher and D. Schubert. Copula-Dependent Default Risk in Intensity Models. Technical report, University of Bonn, December 2001.

[100] A. Sepp. Pricing European-Style Options under Jump Diffusion Processes with Stochastic Volatility: Application of Fourier Transform. Working paper, Institute of Mathematical Statistics, Faculty of Mathematics and Computer Science, University of Tartu, Sep. 2003.

[101] R. U. Seydel. *Tools for Computational Finance*. Springer, 2009.

[102] S. E. Shreve. *Stochastic calculus for finance II: continuous-time models*. Springer, 2004.

[103] L Teng, M. Ehrhardt, and M. Günther. Bilateral Counterparty Risk Valuation of CDS Contracts with Simultaneous Defaults. *Int. J. Theoret. Appl. Fin.,* **16 (7)**, page 1350040, 2013.

[104] L Teng, M. Ehrhardt, and M. Günther. Numerical evaluation of complex logarithms in the Cox-Ingersoll-Ross model. *Int. J. Comput. Math.,* **90 (5)**, pages 1083–1095, 2013.

[105] L Teng, M. Ehrhardt, and M. Günther. Modelling Stochastic Correlation. Working paper, University of Wuppertal, Preprint 14/03, February 2014.

[106] L Teng, M. Ehrhardt, and M. Günther. Option Pricing with Dynamically Correlated Stochastic Interest Rate. *Acta Mathematica Universitatis Comenianae,* **85 (2)**, 2014.

[107] L Teng, M. Ehrhardt, and M. Günther. The Dynamic Correlation Model and its Application to the Heston Model. Working paper, University of Wuppertal, Preprint 14/09, April 2014.

[108] L Teng, M. Ehrhardt, and M. Günther. The Pricing of Quanto Options under dynamic Correlation (extended version). *J. Comput. Appl. Math.,* **275C**, *DOI: 10.1016/j.cam.2014.07.017*, pages 304–310, 2014.

[109] L Teng, M. Ehrhardt, and M. Günther. A Versatile Approach for Stochastic Correlation using Hyperbolic Functions. *Int. J. Comput. Math., DOI:10.1080/00207160.2014.1002779*, 2015.

[110] L Teng, M. Ehrhardt, and M. Günther. Modelling Stochastic Correlation with modified Ornstein-Uhlenbeck process. In *ECMI book subseries of Mathematics in Industry, ISSN: 1612-3956*. Springer Heidelberg, 2015.

[111] L Teng, M. Ehrhardt, and M. Günther. Numerical Simulation of the Heston model with Stochastic Correlation. Working paper, University of Wuppertal, Preprint 15/01, January 2015.

[112] L Teng, M. Ehrhardt, and M. Günther. On the Heston model with Stochastic Correlation. Working paper, University of Wuppertal, Preprint 15/22, April 2015.

[113] G. E. Uhlenbeck and L. S. Ornstein. Gauging the default premium. *Phys. Rev.,* **36**, pages 823–841, 1930.

[114] C. Van Emmerich. Modelling Correlation as a Stochastic Process. Working paper, University of Wuppertal, Preprint 06/03, Jun. 2006.

[115] C. Van Emmerich. *A Square Root Process for Modelling Correlation*. PhD thesis, University of Wuppertal, 2007.

[116] P. Wilmott. *Paul Wilmott on Quantitative Finance*. John Wiley & Sons, Ltd, Second Edition, 2006.

Index