Stochastic Correlation for Asset Pricing and Credit Derivative Products

Ashish Kumar

A thesis presented for the degree of
Doctor of Philosophy

Supervisor: László Márkus, Prof., PhD

Doctoral School of Mathematics
Director: István Faragó, Prof., D.Sc.

Doctoral Program of Applied Mathematics
Director: János Karátson, Prof., D.Sc.

Department of Probability Theory and Statistics
Eötvös Loránd University, Budapest, Hungary
November 2020
Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisor, Prof. László Márkus. His support and guidance were vital throughout my PhD studies. He introduced me to this exciting research topic and I appreciate his vast knowledge and skills in many areas. His expertise, encouragement, understanding, and patience helped me overcome many obstacles and finish this thesis. I am gratefully thankful to Stipendium Hungaricum Scholarship program for funding my studies. I would also like to acknowledge the Department of Probability Theory and Statistics, Eötvös Loránd University, for providing the resources and the atmosphere for conducting research. I sincerely thank Prof. Miklós Arató and Prof. László Márkus for making me a part of the additional research projects which helped to learn several statistical techniques and the precious experience of working in a team. I would also like to thank Prof. András Zempléni for his kind support throughout my doctoral studies, Prof. Tamás Móri and Prof. György Michaletzky for their help during the preparation of doctoral comprehensive exam.

I would like to express my gratitude to Prof. István Ágoston for his persistent help during the initial years of my doctoral studies. I would also like to thank Klaudia Szalay, the angel of the doctoral school, for all her patience, kindness and help during these years.

Besides, I am deeply grateful to Norbert Hári (Morgan Stanley and Citi) for letting me profit from his knowledge. His encouragement, suggestions and comments have been invaluable. Without his guidance and persistent help an important paper of this thesis would not have been possible. I hope the end of my doctoral program will become the lifelong friendship with Prof. Márkus and Norbert Hári. Lastly, I would like to thank my family and loved ones for their love, and for believing in me and supporting me.
List of publications


# Contents

1 Introduction ...................................................... 9
  1.1 Contributions .............................................. 10
  1.2 Motivation and Literature Review ....................... 12
  1.3 Future Research ........................................... 20

2 Preliminaries .................................................. 21
  2.1 Probability Theory .......................................... 21
  2.2 Stochastic Calculus ......................................... 22
  2.3 Copula Theory ............................................... 25
    2.3.1 Definitions and Basic Properties .................... 25
    2.3.2 Dependence Concepts ................................ 27
    2.3.3 Kendall’s function .................................... 27
    2.3.4 Tail Dependence ...................................... 28
  2.4 Derivative pricing and Counterparty Credit Risk ....... 30
    2.4.1 Derivative Pricing .................................... 30
    2.4.2 Counterparty Credit Risk ............................. 31
  2.5 Fractional Brownian Motion ................................. 34
    2.5.1 Properties of fBm ..................................... 34
  2.6 Fractals and fractal dimensions ............................ 35

3 Stochastic Correlation for Asset Pricing and Tail Dependence 39
  3.1 Comparison of Stochastic Correlation Models through quantiles of K-functions 41
    3.1.1 Stochastically correlated Wiener processes ........ 41
3.1.2 Models of Stochastic Correlation . . . . . . . . . . . . . . . . . . 43
3.1.3 Comparing the models . . . . . . . . . . . . . . . . . . . . . . . 46
3.2 Modeling Joint Behaviour of Asset Prices using Stochastic Correlation . . 50
3.2.1 Modeling a Single Stock Price; the Heston Model . . . . . . . . . 50
3.2.2 Stochastic Correlation Model: Goodness of Fit by Quantile Curves
of K-functions . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
3.2.3 Pseudo-algorithm of the methodology . . . . . . . . . . . . . . . 54
3.2.4 Data of Two Stock Prices; Fitting the Model . . . . . . . . . . . 55
3.2.5 Evaluating Model Goodness . . . . . . . . . . . . . . . . . . . . 56
3.2.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

4 Arbitrage-Free Pricing of CVA for Cross-Currency Swap with Wrong-Way Risk under Stochastic Correlation Modeling Framework 63
4.1 Credit Valuation Adjustment and Wrong Way Risk . . . . . . . . . . . 64
4.2 Exposure in Cross-Currency Swap . . . . . . . . . . . . . . . . . . . . 65
4.2.1 Model for Multi-Currency with Short-rate Interest Rates . . . . . 67
4.3 Modeling Stochastic Default Intensity . . . . . . . . . . . . . . . . . . 70
4.4 Stochastic Hybrid Model . . . . . . . . . . . . . . . . . . . . . . . . . 71
4.4.1 Wiggins’ Stochastic Volatility model, Discretization and Fitting . . 72
4.4.2 Interdependence Structure of the Stochastic Hybrid Model . . . . 74
4.5 Data Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
4.6 Monte-Carlo calculation of CVA . . . . . . . . . . . . . . . . . . . . . 76
4.6.1 Algorithm Implementation . . . . . . . . . . . . . . . . . . . . . . 77
4.7 Concluding remarks . . . . . . . . . . . . . . . . . . . . . . . . . . . . 83

5 Rough Correlation for Characterizing Herd Behaviour of Traders in Stock Markets 85
5.1 Brownian Motions with Rough Stochastic Correlation . . . . . . . . . 86
5.2 Wiggins’ Stochastic Volatility model, Discretization and Fitting . . . . 87
5.3 Modeling Rough Correlation . . . . . . . . . . . . . . . . . . . . . . . 88
Chapter 1

Introduction

The growth of the financial market in the last few decades has been tremendous. New products have been introduced by the Investment banks which are of value for existing and new customers. These customers as well as Investment banks will buy any substantial amounts of financial derivatives if the price and risk structure seem fair or appealing to them. Therefore, it is not only customary to develop innovative products but one also has to understand the dynamics of it. It may be necessary to adapt these dynamics to the market and find a suitable computational realization, which leads us to the field of quantitative and computational finance. The paramount task is to build a suitable model that can describe all the parts of the market which have an important influence on the assets and derivatives and to hedge its risk thereafter. Equally important is to find a balance between the complexity of the model and the speed of the implementation.

To understand the behavior of assets or derivatives, we must understand the relationship between their price movements. One way to measure the relationship between random variables is through their correlation, which can only measure linear dependence. But what if this relationship is not linear? This is the question which we will be addressing in this thesis and its application in asset pricing and credit derivative products.
1.1 Contributions

The main objective of this thesis is the study and development of continuous multi-dimensional processes applicable in asset pricing, portfolio management and credit risk and derivatives with emphasis on stochastic correlation. The results obtained are as follows:

1. **Comparison of Stochastic Correlation Models through quantiles of K-functions**

   In [Markus and Kumar(2019/1)], we compare five models of stochastic correlation on the basis of the generated association of Wiener processes. Associations are characterised by the copulas and their K-functions. A confidence domain for the randomly changing K-functions are build on the basis of simulated Wiener process pairs. The models are ordered by the magnitude of the domains or confidence bounds. Larger confidence domains or higher bounds represent higher correlation risk when the models are applied in mathematical finance.

2. **Modeling Joint Behaviour of Asset Prices using Stochastic Correlation**

   In Markus and Kumar(2019a), we use the quantile curves and tail dependence as goodness of fit (GoF) measure for model selection purpose for the daily frequency stock price data. At this point, we analyze and develop a criteria for a suitable model for the association or interdependence of the daily traded closing prices of two stocks. The association is generated by stochastic correlation, given by a stochastic differential equation (SDE) creating interdependent Wiener processes. These, in turn, may drive the SDEs in the stochastic volatility models for stock prices. To choose from possible stochastic correlation models two goodness of fit (GoF) procedures are proposed based on the copula of Wiener increments. One uses the confidence domain for the centered Kendall function, and the other relies on strong and weak tail dependence. The constant correlation and the corresponding Gaussian copula model is unanimously rejected by the methods, but all other are acceptable at 95% confidence level. The analysis also reveals that even for Wiener process, stochastic correlation is capable of creating tail-dependence unlike constant correlation which results in multivariate normal distributions and hence zero tail
dependence. Our study shows that the models with stochastic correlation are suitable to describe more dangerous situation in terms of correlation risk. With the proper modification the suggested model may serve well in multi-stock option pricing as well.

3. **Arbitrage Free Pricing of Credit Valuation Adjustment (CVA) under Wrong-Way Risk**

In Kumar, Markus and Hari(2019b), we revisit the problem of wrong way-risk and propose a stochastic correlation approach which is unique till date. A positive correlation between exposure and counterparty credit risk gives rise to the so called Wrong-Way Risk (WWR). Even after a decade of financial crisis when it played a crucial role, addressing WWR in a both sound and tractable way remains challenging. Academicians have proposed arbitrage-free set-ups through copula methods but those are computationally expensive and hard to use in practice. Resampling methods are proposed by the industry but they lack in mathematical foundations. In this thesis we bridge this gap between the approaches used by academicians and industry. To this end, we propose a stochastic correlation approach to asses WWR. The methods based on constant correlation to model the dependency between exposure and counterparty credit risk assume a linear dependency, thus fail to capture the tail dependence. Using a stochastic correlation we move further away from Gaussian copula and can capture the tail risk. This effect is reflected in the results where the impact of stochastic correlation on calculated CVA is substantial when compared to the case when a high constant correlation is assumed between exposure and credit.

4. **Rough Correlation for Characterizing Herd Behaviour of Traders in Stock Markets**

In Markus and Kumar(2019c) the main novelty is that we define the notion of rough stochastic correlation for Brownian motions, that has never been done before. We define the notion of rough stochastic correlation for Brownian motions. Next, we provide statistical evidence that minute-wise traded Microsoft and Apple stock
prices indeed have a rough correlation as to their association. By modeling those stock prices the estimations of the latent price driving Brownian motions and their temporally localized correlation as a stochastic process becomes available. Analyzing this stochastic correlation we conclude that it has rough paths. Moment scaling indicates fractal behavior and the fractal dimensions and Hurst exponent estimates point to rough paths. Modeling the rough correlation by suitably transforming a fractional Ornstein-Uhlenbeck process enables extensive simulations of stock prices and assessing model goodness of fit, compute tail dependence and the HIX. The HIX is found to be hardly variable in time; to the contrary tail dependence fluctuates heavily. As a result, a sudden coincident appearance of extreme log-return values is more likely than it is indicated by a steady HIX value.

1.2 Motivation and Literature Review

Stochastic modeling is an essential tool in different areas of research and application such as mathematical finance, medical science and biology etc. As soon as there is more than one factor to consider, the question arises on how to map the relationship between these factors. A common approach both in literature and industry is to use correlated stochastic processes where the magnitude of the correlation is measured by a single number the \( \rho \in [-1,1] \), the correlation coefficient. For example, in the case of two Brownian motions \( W_1 \) and \( W_2 \) correlated with \( \rho \), we can express the concept by the symbolic notation

\[
dW_1(t)dW_2(t) = \rho dt. \tag{1.1}
\]

In mathematical finance correlated Wiener processes appear for example in the Heston model [Heston (1993)]

\[
dS(t) = \mu S(t) dt + \sqrt{V(t)}S(t)dW(t) \tag{1.2}
\]
where $\mu > 0$, $\kappa > 0$, $\theta > 0$, $\sigma > 0$ are constant, $W$ and $Z$ are correlated Wiener processes. The first process describes the movement of an underlying asset $S$ and the second process describes the volatility $V$ behaviour. Another example of coupled stochastic process is the following generalization of Black-Scholes model for quantos

\[
dS_1(t) = \mu_1 S_1(t) dt + \sigma_1 S_1(t) dW(t),
\]

\[
dS_2(t) = \mu_2 S_2(t) dt + \sigma_2 S_2(t) dZ(t),
\]

with positive constants $\mu_1$, $\mu_2$, $\sigma_1$, $\sigma_2$. Here one of the stochastic differential equations is supposed to map the performance of a traded object (stock, index, for instance) in a currency $A$. The second stochastic differential equation (SDE) describes the exchange rate between currency $A$ and another currency $B$. Within both example one usually assumes the Wiener processes to be correlated

\[
dW(t).dZ(t) = \rho dt
\]

with a constant correlation factor $\rho \in [-1, 1]$. It is a well known fact that correlation coefficient $\rho$ can only capture linear dependence. However, from the market, we observe this interdependence is usually time-dependent and fluctuates randomly. Investors build and manage their portfolios comprising numerous assets taking into account the dependence structure between asset returns. In particular, the correlation between asset returns is a key ingredient for trading, portfolio optimization and risk management. It is very often considered as reflecting a structural correlation between fundamentals of asset returns and hence assumed not to vary a lot in time. While this may be true in the long run, but instantaneous, temporary or short-lived effects may have a variable influence on prices of shares of companies belonging otherwise to the
same sector, and having similar fundamentals. On the other hand, fire sales or, more generally, the sudden deleveraging of large financial portfolios may create unexpected spikes in volatility as well as in correlations of asset returns. The reason is that large investors build their portfolios with the use of indices and exchange-traded funds (ETFs) and assets traded on large national exchanges so their sells or buys affect those assets simultaneously [Cont and Wagalath (2016)]. This creates another key feature of co-movement of asset returns, i.e. they are more synchronized at large absolute values than at usual, frequent values. This phenomenon is called the tail dependence and by virtue of its existence, Pearson correlations, measuring linear dependence only, are not sufficient for describing the interdependence or association of these returns [McNeil et al. (2005)]. In our perception, the association should be modeled as an integral part of the joint dynamics of the time propagation of the returns. Therefore, although copulas (and for their part dynamic copulas [Cherubini et al. (2011)] as well) may represent the complete probabilistic association they, again, are not completely satisfactory for the description of the association of returns. In summary, intense trading may create an association of asset prices which is highly non-linear, time dependent and random, pointing to a stochastic process as a model for it. Similar to volatilities, there are two conceptually different ways to obtain inference on correlation. Estimation of the temporarily changing correlation is straightforward from historic asset price series e.g. by windowing techniques. We are using the term historic correlation instead of ”realized” one because in some studies the latter term is mixed with equicorrelation, having a restricted meaning of containing only one time dependent component multiplied by equal constants for all assets in question. In the spirit of implied volatility another approach draws on multi-asset or index based derivative prices quoted in the market creating implied correlation [Buss et al. (2017)]. Market data analysis [Driessen et al. (2006)] reveals that implied correlation, reflecting market expectations deviates from the realized correlation, creating a non-zero correlation risk premium [Buraschi et al. (2006)]. To the best of our knowledge, the earliest paper that apply stochastic correlation to the geometric Brownian motion (GBM) is [Hull et al (2005)]. They model the asset process with a GBM and then sample the correlation from a beta
distribution. However, their stochastic correlation is exogenously derived and does not
follow a time dependent stochastic process. Several other studies mention the term stochastic
correlation as [Buraschi et al. (2006)] or [Da Fonseca et al (2008)]. However, in these
papers correlation is stochastically sampled from a distribution or is inferred from a
historical price correlation matrix. The notion of stochastic correlation as a stochastic
process with proper temporal dynamics was first introduced by van [van Emmerich(2006)].
There, too, was the Jacobi process as well as the normalized tangent of an Ornstein–Uhlenbeck
process listed as examples for stochastic correlation. The Jacobi process as a suitable
candidate for stochastic correlation was extensively studied in [Ma (2009a)], in the context
of quanto derivative pricing. More recently [Carr(2017)] and [Teng et al. (2014)] consider
further transformations of Ornstein Uhlenbeck processes for modeling stochastic correlation.
[Itkin(2017)] uses the Lévy processes framework to describe stochastic correlation in the
FX market. Furthermore, there exist several approaches which generalize the constant
correlation to be stochastic, like e.g. the dynamical conditional correlation model of Engle
[Engle(2002)], local correlation models [Langnau(2009)] and the Wishart auto-regressive
process proposed by [Gourieroux et al. (2009)]. We suggest considering the stochastic
differential equation (SDE) description of asset prices together with a time-varying and
random association described by the quadratic covariation of the price SDE-driving Wiener
processes. The interdependence of these underlying Wiener processes may in turn generate
the association of prices observable in the market. The name of this approach was coined
as stochastic correlation in the pioneering work of Catherine van [van Emmerich(2006)]
and used in pricing several multi-asset derivatives like quantos [Ma (2009a)] and spreads
mentioned papers suggest either the Jacobi process or a suitably transformed Ornstein-Uhlenbeck
process as a model for stochastic correlations. Similar to volatilities, there are two conceptually
different ways to obtain inference on correlation. Estimation of the temporarily changing
correlation is straightforward from historic asset price series e.g. by windowing techniques.
In the spirit of implied volatility, another approach draws on multi-asset or index based
derivative prices quoted in the market creating *implied* correlation [Buss et al. (2017)].
In this thesis we confine ourselves to the historic approach only, noting that historic correlation should be used very carefully since generally, the correlation might be more unstable than volatility.

Credit rating models are used by financial institutions to determine the credit-worthiness of obligors. Most credit rating models determine the credit-worthiness of the obligors individually, neglecting possible correlations. In credit risk modeling, the correlation of the asset returns is an important component for the measurement of portfolio risk. Over the market (OTC) trades are the financial contracts between the two default-risky entities. Yet, prior to the crisis, financial institutions usually ignored the credit risk of high-quality rated counterparties but as recent history has shown this was a particularly dangerous assumption. To keep up with this sudden shift of market perception, dealers today make a number of adjustments when they book OTC trades. Credit valuation adjustment (CVA) corrects the price for the expected loss to the dealer in the event that the counterparty may default. CVA is defined as the difference between the risk-free portfolio and the true portfolio takes into account the possibility that a counterparty might default before or at the maturity of the contract. CVA was considered negligible before the credit crisis in 2007. Roughly two-thirds of the credit crisis risk losses were due to CVA losses and only one-third were due to actual defaults. The amount of attention paid to counterparty risk and CVA charges has consequently increased since then. A derivative dealer can have multiple counterparties with over millions of derivatives transactions in total. CVA is calculated for each counterparty based on the total net exposure to the counterparty. Calculating CVA is computationally expensive. An excellent discussion of the issues can be found in Gregory (2009). While calculating CVA, it is usually assumed that the counterparty’s probability of default and the dealer’s exposure are independent, in practice, this usually does not hold. A situation where there is a positive dependence between the two i.e., the counterparty default risk and its exposure move together, is referred to as “wrong-way risk” (WWR). There is no standard WWR approach widely adopted by the industry. There are a number of models in the literature which have tried to tackle the wrong way risk problem for example [Rosen and Saunders(2012)], [Ignacio Ruiz(2014)],
These approaches can be classified broadly into two groups. First family of models are those which use copula methods to model the joint probability distribution function driving exposures values and defaults, typically using a Gaussian copula with a constant correlation $\rho_{\text{Gauss}}$, and those who model the dependency between portfolio exposure and counterparty default events using an analytical approach linking portfolio value with default intensity. Both the families use a change of measure. Copula based models tend to be computationally expensive because the simulation of random paths from joint distributions is required. Models from the second family are faster to compute, but they require an analytical expression that links the value of the counterparty exposure with the default probability of the corresponding counterparty. Earlier works present statistical evidence that correlation should not, indeed, be assumed to be constant in time. Therefore, as correlation is not a constant quantity, the value of CVA may change with the time period. More often than not, using a linear correlation may fail to capture actual risk and thus can underestimate the value of CVA. In this thesis, we revisit the problem of CVA under WWR by proposing a new method to handle it in a sound but yet tractable way. We compute CVA under WWR by using a stochastic correlation approach. This way we move further away from Gaussian copula and thus can capture tail risk, which is not the case if a constant correlation is used. The models based on constant correlation have no tail dependence, thus, fail to model the wrong-way risk in extreme events. As it is said "demon lies in the tail" and Gaussian copula based models already have been criticized enough after the financial crisis. As far as authors know, no method to compute CVA by using a stochastic correlation has been proposed to date.

As we have seen so far that the interdependence/association of various financial instruments such as asset prices or credit values and the related derivatives, or credit events plays a basic role in investment strategies and risk management. Throughout this thesis, we will establish how using constant correlations the evolution of the dependence structure
which is often described by SDEs is either neglected, or this description is not compatibly
intertwined with the individual modeling. Understanding the fact that substantial amount
of financial instruments in market are short dated, we not only analyze the interdependence
for long time horizons financial instruments but also for the ones with the short time
horizons. For the short dated financial instruments we turn to High Frequency Trading
(HFT). HFT is a type of algorithmic financial trading that leverages high-frequency financial
data and electronic trading tools. Intraday, proportion of HFT may account for 0 - 100
percent of short-term trading volume. By regarding the short time ”localized” correlations
as a stochastic process, and relate them to the Brownian motions, the underlying driving
forces in the mentioned SDEs, a more coherent modeling framework can be built, and
empirically observed phenomena such as tail dependence can be better explained. This
non-linearly associated Brownian motions can be constructed by prescribing their quadratic
covariation as an integral of a diffusion or a rough process, called the stochastic correlation
and can in turn be described by another SDE generated from either a classical or a
fractional Brownian motion. To describe stochastic volatility a new modeling framework
has been introduced in the seminal paper [Gatheral et al.(2018)] by using fractional Brownian
Motions (fBm) as driving forces in stochastic differential equations. The peculiarity of
the approach is the small, typically of magnitude 0.1 or even lower, value of the Hurst
parameter of the fBm [M. Fukasawa (2017)]. These models are referred to as rough
volatility models, see [Bayer et al.(2016)], [Bayer et al.(2017)], [Bennedsen et al.(2017)],
[El Euch and Rosenbaum(2019)], [M. Fukasawa (2017)],
[Jacquier et al.(2017)] for more details and practical applications. Such small estimated
values for $H$ (between 0.05 and 0.2) have been found in [Mandelbrot and Ness (1968)]
when studying the volatility process of a great number of assets. In the light of the
mentioned literature we propose to model the stochastic correlation of Brownian motions,
with a suitable transformation of a rough, fractional Ornstein-Uhlenbeck (fOU) process
[T. Kaarakka and P. Salminen (2011)], generated from fBm as introduced in
[Cheridito et al.(2003)], with Hurst coefficient $H < 0.5$. Integration with respect to the
fractional Brownian motion (fBm) is necessary to specify the fOU process. It can be
introduced in several ways and the different resultant integrals are equal only for particular sets of integrands. We refer the reader to [Hu et al (2017)] for a review of stochastic calculus with fBm. An almost complete survey of the mentioned integrals is given in [Biagini et al.(2008)]. All these integrals coincide on particular deterministic functions, and for smooth integrands, they coincide with the limits of the Riemann-Stieltjes integral sums. In our case, since we only have constants as integrands, we will focus on these Riemann-Stieltjes integrals, referring to the path-wise approach. For further details on fBm and fOU processes, we refer the reader to [Biagini et al.(2008)], [Cheridito et al.(2003)], [L. Coutin (2007)]. We provide statistical evidence to justify the choice of a rough process instead of a diffusion in modeling stochastic correlation especially in high-frequency trading. As in the case of stochastic volatility [Gatheral et al.(2018)], [M. Fukasawa (2017)], moment scaling reveals the fractal character of the back-transformed stochastic correlation, that we describe by the fOU process. Although the various Hurst exponent estimations create diverse values, all of them are very well below the 0.5 bound, indicating rough paths. As a novel approach, the fractal dimension of the estimated stochastic correlation is considered and found that its high value is counter-intuitive to the choice of a traditional diffusion process. By using a rough correlation instead, a conforming match with this feature can be obtained. Once the Hurst exponent is set, the estimation of the fOU model parameters are straightforward as in [Biagini et al.(2008)]. We choose the Hurst exponent so that the model reconstructs well both the fractal dimension and the marginal distribution of the stochastic correlation process. The complex hierarchical model opens the way for extensive simulation of stochastic correlations and price processes for the goodness of fit assessment and further analysis. There are periods in the market when prices drop or grow together, by synchronized co-movements, and hence losses in one asset cannot be balanced by gains in others. This is particularly so in stressed or critical market situations. The reason behind such a phenomenon may be, that all investors or brokers may expect a similar impact on the market of an occurring piece of information and therefore, instinctively, start to act in their trade similar way. Such a spontaneous synchronized action, without centralized direction, is called herding behavior. To measure
its presence, the herding behavior index or HIX was created in [Dhaene et al.(2011)], (see also [Guillaume and Linders(2015)]) and used for measuring the level of possible diversification among stock prices. The HIX takes values between 0 and 1. Its high value signals a higher degree of herd behavior, which is positive dependence of market prices, allowing for less diversification in investments. It can naturally be expected that the co-movement of prices lead to the coincident appearance of extreme values. Since HIX measures the co-movement and tail dependence the coincidence of extreme values of stock prices a call for comparison occurs naturally, and the simulation of a high number of synthetic paths makes it possible. However, contrary to intuition we find that while the variability of the HIX index is very low, less than 10 percent around its mean value, tail dependence may eventually double. This means that relying only on the HIX index may result in miscalculations in the frequency of simultaneously occurring extreme situations.

1.3 Future Research

We are entering into an era which promises to be driven by Machine Learning and Artificial Intelligence. Techniques borrowed from Deep Learning are increasingly playing an important role in the calibration of the financial models. Stochastic correlation models are driven by a fractional Brownian motion can play a crucial role in modeling the interdependence between the different assets especially in the HFT. We plan to pursue this study of interdependence between different assets and credit derivatives using Deep Learning algorithms.
Chapter 2

Preliminaries

Stochastic analysis is a vast field of study that lies at the heart of quantitative finance or financial engineering. A comprehensive introduction to stochastic analysis is out of the scope of this thesis. In this chapter, we give a short and concise summary of the mathematical tools and its classical financial applications which are later used in this thesis. For in-depth study see [Øksendal(2000)] and [Karatzas and Shreve (1988)].

2.1 Probability Theory

The triplet \((\Omega, \mathcal{F}, \mathcal{P})\) is called a probability space where each of its components is given as

- \(\Omega\) - the sample space
- \(\mathcal{F}\) - \(\sigma\)-algebra of events which is a family of subsets of \(\Omega\) such that it contains the empty set and is closed under complements and countable unions
- \(\mathcal{P}\) - the probability measure

The triplet \((\Omega, \mathcal{F}, \mathcal{P})\) is called a complete probability space, if \(\mathcal{F}\) contains all the \(\mathcal{P}\)-null sets. A filtration is the set of information available up to a specific point in time. A filtered probability space \((\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F})\) is a probability space \((\Omega, \mathcal{F}, \mathcal{P})\) equipped with a filtration \(\mathcal{F}\). A filtration \(\mathcal{F}\) of sigma algebras is a family \((\mathcal{F}_t)_{t \geq 0}\) such that \(\mathcal{F}_t \subset \mathcal{F}\) and
\( \mathcal{F}_s \subset \mathcal{F}_t \) for all \( 0 \leq s < t < \infty \) and \( \mathbb{F} \) is right continuous, i.e. \( \mathcal{F}_t = \mathcal{F}_{t+} \). If \( \mathcal{F}_0 \) contains all subsets of the \( \mathcal{P} \)-null sets of \( \mathcal{F} \), we say that \((\Omega, \mathcal{F}, \mathcal{P}, \mathbb{F})\) is a complete filtered probability space.

Brownian motion is one of the most important stochastic processes and a key element in the quantitative finance. Let \((\Omega, \mathcal{F}, \mathcal{P}, \mathbb{F})\) be the underlying filtered probability space. \( \mathcal{F} \)-Brownian motion \( W = (W_t)_{t \geq 0} \) is a real stochastic process adapted to \( \mathbb{F} \) satisfying:

- \( W(0) = 0 \), a.s
- \( W \) has continuous paths, a.s
- For all \( 0 < s < t \), the random variable \( W_t - W_s \) is independent of \( \mathcal{F}_s \)
- For all \( 0 < s < t \), the law of \( W_t - W_s \) is \( \mathcal{N}(0, t-s) \)
- there exists a modification \( \tilde{W} \) that has continuous paths

Brownian motion is an example of a martingale. A stochastic process \( X = (X(t))_{t \geq 0} \) is a \( \mathcal{F} \)-martingale if it has a finite first moment and

\[
E[X_t | \mathcal{F}_s] = X_s \quad \mathcal{P} \text{-a.s} \quad \forall \quad 0 \leq s \leq t. \tag{2.1}
\]

### 2.2 Stochastic Calculus

The Wiener process (or Brownian motion) \( W(t, \omega) \) does not have bounded variation (neither in \( L^2 \) nor with probability 1). However, the quadratic variation is finite and equals the length of the considered time interval. As a result, \( \int_0^t f(s, \omega) dW(s) \) can not be defined as a Riemann-Stieltjes or Lebesgue-Stieltjes integral. Unless otherwise mentioned, we only consider Itô integrals with respect to the Wiener process. It is defined the standard way as limits of integrals of adapted simple processes. From simple processes it can first be extended to adapted processes in the \( L^2([0, T] \times \Omega) \) space, when the created integral process – the integral depending on the upper bound – is a genuine \( \mathbb{F} \)-martingale. The
integral can then be further extended to adapted processes with almost surely finite square (Lebesgue) integrals, when the integral process is local $F$-martingale only.

One of the most important consequences of the finite quadratic variation is that in a Taylor expansion the second order term still has a non-negligible contribution. The result is the celebrated Itô’s formula for the compound process, containing the second derivative of the outer function, as well.

**Theorem 2.2.1. (Itô’s Lemma)** Let $W = (W(t))_{t \geq 0}$ be a $m$-dimensional Brownian motion, $m \in \mathbb{N}$, and $X = (X(t))_{t \geq 0}$ be an Itô process with

\[
\begin{align*}
  dX(t) &= \mu(t)dt + \sigma(t)dW(t) = \mu(t)dt + \sum_{j=1}^{m} \sigma_j(t)dW_j(t). \\
  \text{Furthermee, let } G : \mathbb{R} \times [0, \infty) &\rightarrow \mathbb{R} \text{ be twice continuously differentiable in the first variable, with the derivatives denoted by } G_x \text{ and } G_{xx}, \text{ and once continuously differentiable in the second, with the derivative denoted by } G_t. \text{ Then we have for all } t \in [0, \infty) \\
  G(X(t), t) &= G(X(0), 0) + \int_0^t G_t(X(s), s)ds + \int_0^t G_x(X(s), s)dX(s) + \frac{1}{2} \int_0^t G_{xx}(X(s), s)d\langle X \rangle(s) \\
  \text{(2.3)}
\end{align*}
\]


Having these technical tools, stochastic differential equations (SDE) of the type

\[
  dX(t) = a(X(t))dt + \sigma(X(t))dW(t), \quad (2.4)
\]

where $W(t)$ is a Wiener process, $a(x), \sigma(x)$ are Borel-measurable real functions can be considered, under suitable conditions on $a$ and $\sigma$. The definitions can straightforwardly be extended to non-autonomous equations, as well.

Strong and weak solutions of the SDEs can be defined, strong solutions roughly meaning path-wise solutions whereas weak solutions are solutions – again roughly – in distribution. Well-known criteria, going back to Itô himself, for the existence and uniqueness (under
suitable initial condition) are the Lipschitz- and the linear growth conditions on the coefficient functions. These conditions ensure, e.g., the existence of the Ornstein-Uhlenbeck process satisfying in its mean reverting form the following SDE:

\[ dX(t) = \kappa [\vartheta - X(t)] dt + \sigma dW(t), \]

Unfortunately these conditions do not hold true in many other SDEs of mathematical finance. Skorokhod and independently Stroock and Varadhan established that a much weaker condition, namely bounded and continuous coefficients, is sufficient for the existence of the weak solution. Then Yamada and Watanabe’s concept of ”existing weak solution and path-wise uniqueness guarantees existence of strong solution” leads to the path-wise existence of square root diffusions. Among these processes is the solution of the famous Cox-Ingersoll-Ross (CIR) equation

\[ dX(t) = \kappa [\vartheta - X(t)] dt + \sigma \sqrt{X(t)} dW(t) \quad (2.5) \]

\[ \quad (2.6) \]

that we shall repeatedly use in the sequel. Note here that this process is non negative – therefore the square root is meaningful – and strictly positive under the Feller condition of \( 2\kappa \vartheta \geq \sigma^2 \). For a detailed and in depth introduction on the subject one may consult [Karatzas and Shreve (1988)].

The notion of stochastic correlation relies heavily on the following theorem, going back to Paul Lévy.

**Theorem 2.2.2. (Lévy’s Characterization of Brownian Motion )** If a stochastic process \( W(t) \) defined on probability space \((\Omega, \mathcal{F}, \mathcal{P})\) with filtration \( \mathcal{F} \in \mathbb{R}^+ \) satisfies the following conditions:

- \( \mathcal{P}(W(0) = 0) = 1 \)
- \( W(t) \) is a continuous martingale w.r.t. the filtration \( \mathcal{F} \in \mathbb{R}^+ \) under \( \mathcal{P} \)
- The quadratic variation \( \langle W(t) \rangle = t \) almost surely w.r.t. \( \mathcal{P} \)
then the stochastic process \( W(t) \) is a Brownian Motion.


2.3 Copula Theory

What are copulas? According to Nelsen (2006, p.12), copulas can be explained from two points of view: “From one point a view, copulas are functions that join or ‘couple’ multivariate distribution functions to their one dimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one dimensional margins are uniform on the interval \((0, 1)\).” Therefore, copulas are multivariate distribution functions that allow the decomposing of any n-dimensional joint distribution into its n marginal distributions and a copula function. In practice, a copula is often used to construct a joint distribution function by combining the marginal distributions and the dependence between the variables.

2.3.1 Definitions and Basic Properties

For simplicity, we only focus on bivariate copulas in this thesis. The extension to the multivariate case is theoretically straightforward but computationally expensive. Following Nelsen (2006), we begin with a definition.

Definition: A two dimensional copula is a function \( C : [0, 1]^2 \to [0, 1] \) with the following properties

1. For every \( u \in [0, 1] \)

\[
C(0, u) = C(u, 0) = 0 \tag{2.7}
\]

2. For every \( u \in [0, 1] \)
\( C(u,1) = u, C(1,u) = u. \) \hspace{1cm} (2.8)

3. For every \((u_1,u_2), (v_1,v_2) \in [0,1] \times [0,1]\) with \(u_1 \leq v_1\) and \(u_2 \leq v_2\):

\[
C(v_1,v_2) - C(v_1,u_2) - C(u_1,v_2) + C(u_1,u_2) \geq 0
\] \hspace{1cm} (2.9)

Property 1 is also referred to as the grounded property of a copula. It says that the joint probability of both outcomes is zero if the marginal probability of any outcome is zero. Property 3 is the two-dimensional analogue of a non-decreasing one-dimensional function. A function with this feature is therefore called 2-increasing.

**Sklar’s Theorem (1959)** Let \( H \) be a two-dimensional joint distribution function with marginal distributions \( F \) and \( G \). Then there exists a copula \( C \) such that for any \( x, y \):

\[
H(x,y) = C(F(x), G(y)),
\] \hspace{1cm} (2.10)

If \( F \) and \( G \) are continuous, then \( C \) is unique. Conversely, if \( C \) is a copula and \( F \) and \( G \) are distribution functions, then the function \( H \) is a joint distribution function with marginal distributions \( F \) and \( G \). The converse of Sklar’s theorem is very useful in modeling multivariate distributions in finance. It implies that if we combine two different marginal distributions with any copula, we will have defined a valid bivariate distribution. This provides great flexibility when modeling a portfolio as we can use different marginal distributions for each asset class and use a copula to link them together.

**Corollary**: Let \( H \) be a two-dimensional joint distribution function with continuous marginal distributions \( F, G \) and copula \( C \), satisfying Sklar’s theorem, then for \( u \in [0,1] \) and \( v \in [0,1] \):

\[
C(u,v) = H(F^{-1}(u), G^{-1}(v)),
\] \hspace{1cm} (2.11)

where \( F^{-1} \) and \( G^{-1} \) denote the inverses of \( F \) and \( G \). This corollary provides a motivation for calling a copula a dependence structure as the copula links the quantiles of the two
distributions rather than the original variables. Based on this corollary, one of the key properties of a copula is that the dependence structure is unaffected by a monotonically increasing transformation of the variables. For example, the same copula can be used for the joint distribution of \((X, Y)\) as well as for the joint distribution of \((\ln X, \ln Y)\). The above Theorem and Corollary extend naturally from the bivariate case to the multivariate case.

2.3.2 Dependence Concepts

Correlation is probably the single most important concept in modern portfolio theory ever since Markowitz (1952) proved the importance of diversification in portfolio allocation fifty years ago. The concept of correlation has become so popular that people now use the term “correlation” and “dependence” interchangeably. However, correlation is not an appropriate measure of dependence when returns are not normally distributed. Formally, two random variables \((X, Y)\) are dependent or associated if they don’t satisfy the condition of probabilistic independence, i.e. \(F(x, y) \neq F_1(x)F_2(y)\). A measure of dependence summarizes the dependence structure of two random variables in a single number. Dependence can be measured using several different concepts: linear correlation, concordance, and tail dependence.

2.3.3 Kendall’s function

An important tool to characterize a copula almost uniquely, up to the class of associative copulas, or up to a class of Bertino-copulas, – that is essentially up to Lipschitz continuity, is Kendall’s cumulative function. Given a random pair \((U, V)\) with distribution \(C\), define random variable \(Z = C(U, V)\). Then Kendall’s cumulative function or K-function is obtained through the probability integral transform as

\[
K(t) = P[C(U, V) \leq t]
\]  

(2.12)
Genest and Rivest (1993) introduced this function to choose among Archimedean Copulas. An estimation of K-function can be calculated as follows:

- for all \( i \in \{1, 2, \ldots, n\} \), compute \( Z_i \) as the proportion of observations in the lower quadrant, with upper corner \((X_i, Y_i)\), i.e.

\[
Z_i = \frac{1}{n - 1} \sum_{j \neq i} 1(X_j < X_i, Y_j < X_i)
\]  

(2.13)

- then compute the cumulative distribution function of \( Z_i \)’s.

2.3.4 Tail Dependence

Tail dependence measures the amount of dependence in the upper-right-quadrant tail or lower-left-quadrant tail of a bivariate distribution. Such a dependence measure relates to the conditional probability that one variable exceeds some value given that another exceeds some value. For continuous marginal distributions, tail dependence is a copula property; hence it is invariant under strictly increasing transformations.

Given two random variables, their tail dependence or asymptotic dependence measures association between their extreme values and depend only on their copula. Consider a random vector \( \mathbf{X} = (X, Y) \) with marginals \( F_X \) and \( F_Y \). The coefficient of lower tail dependence is then defined as the limit of the conditional probability that \( X \) is less than or equal to the quantile \( F_X^{-1}(u) \) provided that \( Y \) is less than or equal to \( F_Y^{-1}(u) \) as \( u \) approaches 0 i.e.

\[
\lambda_L = \lim_{u \to 0^+} P(X \leq F_X^{-1}(u) | Y \leq F_Y^{-1}(u))
\]  

(2.14)

In the context of financial risk management, the lower tail usually corresponds to losses, and hence is of utmost importance because the clustering of high-severity losses can have a devastating effect on the well-being of firms and is thus of pivotal importance in risk analysis. Gaussian copula based credit risk models like that of Li [Li (2000)] have
faced harsh criticism in the past, on the ground of having zero i.e. no tail dependence \[\text{[Shibuya (1960)]}\].

In bivariate random vectors with dependence structures given by the copula function \(C : [0, 1]^2 \to [0, 1]\) \[\text{[Joe (1990)]}\] suggested to measure the lower tail dependence as

\[
\lambda_L = \lambda(C) = \lim_{u \downarrow 0} \frac{C(u, u)}{u}. \tag{2.15}
\]

Non-zero values of \(\lambda_L\) suggest lower tail dependence in \(C\). In what follows we shall call this measure the strong tail dependence. The introduction of the distinctive adjective ”strong” is motivated by the fact that when the limit \(2.15\) is zero, it can be useful to rely on a more delicate index defined in \[\text{[Coles et al. (1999)]}\] and called the lower weak tail dependence \(\tilde{\lambda}_L \in [-1, 1]\) which is given as

\[
\tilde{\lambda}_L = \lim_{u \downarrow 0} \frac{2 \log(u)}{\log(C(u, u))} - 1. \tag{2.16}
\]

In \[\text{[Fisher and Klein(2007)]}\] one may find more details about the intuition behind this measure of extreme co-movements. Similarly as \(u \uparrow 1\) in \(2.15\) and \(2.16\), we may have the upper strong and weak tail dependences \(\lambda_U, \tilde{\lambda}_U\), but we only shall consider the lower dependences in this thesis so, we will not mention the upper ones any further.

The empirical counterpart of the strong lower tail dependence for a given sample size \(n\) is denoted as \(\hat{L}_n(u)\) and can be given as

\[
\hat{L}_n(u) = \frac{\sum_{i=1}^{n} 1_{\{X_i \leq \hat{F}_X^{-1}(u), Y_i \leq \hat{F}_Y^{-1}(u)\}}}{\sum_{i=1}^{n} 1_{\{X_i \leq \hat{F}_X^{-1}(u)\}}}, \tag{2.17}
\]

where \(\hat{F}\) is the empirical distribution function of the variable written in its index.

With the same notations the weak lower tail dependence is estimated as

\[
\hat{L}_n^w(u) = \frac{2 \log\left(\frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq \hat{F}_X^{-1}(u)\}}\right)}{\log\left(\frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq \hat{F}_X^{-1}(u), Y_i \leq \hat{F}_Y^{-1}(u)\}}\right)} - 1. \tag{2.18}
\]
2.4 Derivative pricing and Counterparty Credit Risk

2.4.1 Derivative Pricing

Quantitative finance aims at pricing the contracts with uncertain payoffs in the future. These contracts are known as contingent claims. In other words, if the underlying process is a vector \( S = (S_1, \ldots, S_n) \) of assets with generated information \( \mathcal{F}_t \) upto time \( t \), a contingent claim at maturity \( T \) is any \( \mathcal{F}_T \)-measurable random variable. Some of the popular contingent claims which are also a part of this thesis are as following:

1. **Forward Contract**: A forward contract is an agreement to buy or sell an underlying asset at a specific maturity date for a specific price. There are no upfront charges to enter a forward contract and the payoff of a sell contract is \( S_T - F_{0,T} \), where \( S \) is the spot price and \( F_{0,T} \) is the forward price seen at the beginning of the contract. By the risk neutral valuation, the value at \( t \) of a forward contract started at 0 is

\[
E_T^Q [\exp(-\int_t^T r(s) ds)(S_T - F_{0,T})]. \tag{2.19}
\]

2. **Forward Rate Agreement**: In a forward rate agreement, two parties agree that a certain interest rate applies to a specified principal during a period of time in the future.

3. **Swap**: Swaps are financial instruments that are used to exchange cash flows at fixed points in time. One type of swap is a cross currency swap in which the cash flows are exchanged in different currencies plus the notional of the contract is also exchanged at the maturity.

**Theorem 2.4.1.** (Risk-Neutral Valuation Formula) Assume that the price processes for the underlying assets is modeled by geometric Brownian motions, the market is complete and \( X \) is a given contingent claim. Let \( B_0 \) be a risk free asset with a price process given by \( dB_0(t) = B_0(t) r(t) dt, B_0(0) = 1 \), where \( r(t) \) is the risk free instantaneous interest rate.
The arbitrage price process of $X$ is given by the risk-neutral valuation formula

$$V_x(t) = B_0(t)E_{\mathcal{Q}}\left[\frac{X}{B(T)}|\mathcal{F}_t\right] = E_{\mathcal{Q}}[Xe^{-\int_t^T r(u)du}|\mathcal{F}_t],$$

where $\mathcal{Q}$ is the unique equivalent martingale measure.

Proof. See [Øksendal(2000)].

2.4.2 Counterparty Credit Risk

Over-the-counter (OTC) financial instruments are traded directly between two parties and not via an exchange or a so-called clearing house. The development and proliferation of the OTC derivatives market have arguably been one of the most important events in the last two decades. Being negotiated directly between counterparties, OTC derivatives can be tailored to the counterparties’ specific needs and thus offer unlimited possibilities for risk transferal.

In the wake of the financial crisis in 2007-2008, investors and financial institutions have been forced to rethink how they value and hedge contingent claims traded either in the OTC market or through central clearing house (CCPs). The credit valuation adjustment (CVA) which is defined as the difference between the risk-free portfolio and the true portfolio takes into account the possibility that a counterparty might default before or at the maturity of the contract. The credit valuation adjustment (CVA) corrects the price for the expected costs to the investor due to the possibility that the counterparty may default. Due to the large number of financial contracts that are traded over the counter, the importance of the credit quality of a counterparty is fundamental, and counterparty credit risk is introduced when evaluating a derivative contract. Market participants charge a risk premium when investing in default risky assets to account for the counterparty credit risk. As a result, the value of a contingent claim with a defaultable counterparty will be smaller than the value of the same claim with a non-defaultable counterparty.

Let $T \in \mathbb{R}_+$ be the maturity of the derivative contract. Considering the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{Q})$, we contemplate the valuation problem where only one of the two
parties can default. The problem reduces to computing the adjustment of the default-free price of the deal when entering a financial trade with a counterparty that has a positive probability of defaulting before the maturity of the trade. This adjustment is known as unilateral credit valuation (CVA, hereafter). CVA was considered negligible before the credit crisis in 2007. Roughly two-thirds of the credit crisis risk losses were due to CVA losses and only one-third were due to actual defaults. The amount of attention paid to counterparty risk and CVA charges has consequently increased since then. In the following, we see how to compute the value of this adjustment.

**General Pricing formula:** At the time of default $\tau < T$, the investor will calculate net present value (NPV) of the residual value of the deal until maturity. If the NPV is negative to the investor, the investor will pay in full the NPV to the defaulted counterparty. If the NPV is positive to the investor, the investor will face a loss and receive only a fraction of NPV, usually referred to as the recovered rate ($RR$). In the case when there is no default, i.e. $\tau > T$, the derivative trade is the same as the default-free case.

We denote by $\Pi(t,T)$ the sum of the discounted (at the risk-free rate) payoff at time $t<T$ happening over the time period $(t, T]$. Therefore, we have

$$NPV(\tau) = E_\tau[\Pi(\tau, T)].$$

(2.21)

The defaultable derivative price denoted as $\bar{V}(t)$ at time $\tau < T$ can be calculated as

$$\bar{V}(t) = E_\tau\{1_{\tau>T}\Pi(t,T) + 1_{\tau<T}[\Pi(t, \tau) + D(t, \tau)(RR(NPV(\tau))^+ - (-NPV(\tau))^+)]\},$$

(2.22)

$D(t, u)$ is the risk-free discount factor, given by

$$D(t, u) = \frac{B(t)}{B(u)},$$

(2.23)

where

$$dB(t) = r(t)B(t)dt$$

(2.24)
is the bank account driven by the risk-free instantaneous interest rate \( r \) and associated to the risk-neutral measure \( \mathcal{Q} \). The rate \( r \) is assumed to be \( (\mathcal{F}_t)_{t \in [0,T]} \) adapted.

**Proposition 1. (General unilateral counterparty risk pricing formula):** At time \( t < \tau \), the price of the derivative trade under counterparty credit risk is

\[
\bar{V}(t) = \mathbb{E}_t[\Pi(t)] - \text{LGD} \mathbb{E}_t[1_{\{t < \tau \leq T\}}D(t, \tau)(\text{NPV}(\tau))^+] + \text{CVA}(t),
\]

(2.25)

where \( \text{LGD} = 1 - \text{RR} \) is the loss given default, and the recovery rate \( \text{RR} \) is assumed to be deterministic.

We notice that the counterparty risk adjusted deal price consists of the default-free price and an adjustment term which is called credit valuation adjustment denoted by \( \text{CVA} \) given as

\[
\text{CVA}(t) = \text{LGD} \mathbb{E}_t[1_{\{t < \tau \leq T\}}D(t, \tau)(\text{NPV}(\tau))^+] \]

(2.26)

The \( \text{CVA} \) term takes the form of an option on NPV with zero strike, only when the event of default happens (\( \tau \leq T \)). Including counterparty credit risk into valuation makes the pricing model dependent, even when the original valuation is model independent (for instance an interest rate swap and cross-currency swap). If both parties in a transaction may default, the counterparty risk adjustment becomes bilateral credit valuation adjustment. However, in this thesis, we only care about unilateral \( \text{CVA} \). The full valuation formula in (2.25) depend on the joint distribution of the default times and the underlying asset.

Wrong way risk (WWR, hereafter) is the risk, the investor has when the underlying portfolio and the default of the counterparty are "correlated" in the worst possible way from the investor’s perspective.
2.5 Fractional Brownian Motion

2.5.1 Properties of fBm

Definition 2.5.1. (fractional Brownian motion)

A centered Gaussian process $W_H = \{ W^H_t, t \in \mathbb{R} \}$ is called a fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$ if it has the covariance function

$$C_H(t, s) = \mathbb{E}[W^H_t W^H_s] = \frac{1}{2} (s^{2H} + t^{2H} - |t - s|^{2H}) \quad s, t \in \mathbb{R}. \quad (2.27)$$

For a Gaussian process, its distribution is uniquely determined by its mean and covariance function. Therefore, the distribution of a fBm is uniquely specified by the above definition.

Definition 2.5.2. (homogeneous function)

A homogeneous function $f$ of variables $x$ and $y$ is a real-valued function that satisfies

$$f(tx, ty) = t^k f(x, y), \quad (2.28)$$

for some constant $k$ and all real numbers $t$. The constant $k$ is called the degree of homogeneity.

The covariance function fBm is then homogeneous of order $2H$. From this property it can be deduced that fBm is $H$ self-similar, that is, for $\alpha > 0$, $\{ W^H_{\alpha t}, t \in \mathbb{R} \}$ has the same distribution as $\{ \alpha^H W^H_t, t \in \mathbb{R} \}$. It can be derived from equation (2.27) that

$$\mathbb{E}[|W^H_t - W^H_s|^2] = |t - s|^{2H}, \quad s, t \in \mathbb{R}. \quad (2.29)$$

This implies that fBm has stationary increments. When $H = \frac{1}{2}$, we retrieve the standard Brownian motion $W^1_t$ with covariance function $\mathbb{E}[W^1_t W^1_s] = \min(|t|, |s|)$ for $t, s \in \mathbb{R}$ and independent increments. However, when $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, the increments of the fBm on disjoint intervals are not independent.

1. When $H \in (0, \frac{1}{2})$, $C_H(n) < 0$ and $\sum_{n=1}^{\infty} |C_H(n)| < \infty$ which implies that the process has short memory, and in contrast
2. When $H \in \left(\frac{1}{2}, 1\right)$, $C_H(n) > 0$ and $\sum_{n=1}^{\infty} |C_H(n)| = \infty$ indicating long memory.

Instead of the Wiener process it is possible to use fBm in stochastic differential equations as driving force. To this, the integral wrt. fBm has to be defined, that means technical difficulties, because fBm is not a semimartingale. Malliavin calculus or Skorokhod integral may be used to circumvent the issue. However, in certain cases the integral reduces to a Riemann-Stieltjes one. See, e.g., [Biagini et al.(2008)].

The dynamics of the mean reverting fractional Ornstein-Uhlenbeck process (as defined in [Cheridito et al.(2003)], which is different of [Kaarakka(2011)]) is given by a Langevin-type stochastic differential equation

\[
dX(t) = -\lambda (\theta - X(t))dt + \sigma dW^H(t)
\]

\[
X(0) = X_0 \in \mathbb{R}
\]

(2.30)

(2.31)

driven by a fBm $W^H = \{W^H(t), t \geq 0\}$ of Hurst parameter $H \in (0, 1)$; both the parameter $\lambda$ and the volatility $\sigma$ are positive real constants.

Assuming the Hurst exponent to be known, the drift and the volatility of the fOU process can be estimated as in [Hu et al (2017)]. The least squares estimation of the drift is $-\log\left(\frac{\int_0^T X(t)dX(t)}{\int_0^T X^2(t)dt}\right)$ where the approximating Riemann-Stieltjes sum can be used in the computation. For $\sigma$ the specific version, given for the constant case in Proposition 4.2. of [Hu et al (2017)] in terms of p-variations is applicable. Choosing $p = 2$ the estimate of $\sigma$ reduces to $n^H$ times the residual standard deviation.

2.6 Fractals and fractal dimensions

The word “fractal” originates in Mandelbrot’s works, e.g. [Mandelbrot 1977]. By this notion, he intended to describe the rough, broken and irregular character of certain objects he studied. It is characteristic of fractals that roughness is present at all scales. Fractals fill the Euclidean topological space they are embedded into more thickly than an object of
lower topological (integer) dimension. However, they still do not fill out completely any
open set of that Euclidean space. Hence they have a non-integer dimension, the so-called
fractal dimension.

A geometrically intuitive notion of dimension is an exponent that expresses the scaling
of an object’s bulk with its size (cf. [Theiler 1990]). It can also be interpreted as how
thickly an object fills part of the space of an integer topological dimension in which it is
embedded. Specifically, for time series the fractal dimension is a measure of the roughness
(or conversely, smoothness) of the paths; the rougher the path the higher the dimension
of the graph between 1 and 2. The earliest definition of a non-integer dimension is
Hausdorff’s, but since then various definitions have become known within this concept,
reflected also in the estimation methods. Quite a number of estimating methods are
thematically described in [Gneiting and Schlather 2004] and [Gneiting et al. 2010]. Below
we give a short overview of those we use in the present paper.

The Hausdorff Dimension
A theoretically rigorous definition of a non-integer dimension of a set was given in [Hausdorff(1919)].
The considered set is covered with balls of radius at most \( r \), the sum of the \( d \)-th power of
those radii is calculated, and its infimum is taken over all coverings of the type.

\[
\inf_{\text{coverings}} \left( \sum_{t_i < r} r_i^d \right)
\]

The limsup of this infimum as \( r \to 0 \) is either infinity or 0 and Hausdorff showed that the
transition happens at a unique value \( \dim_H \) as \( d \) grows. This transition value \( \dim_H \) is the
Hausdorff dimension of the set. For further details see e.g. [Theiler 1990].

\[p\text{-Variogram Based Estimators} \]
For a time dependent stochastic process \( \{X_t : t \in R^+\} \) with stationary increments, the
$(V_p(t)$ (semi)variogram function of order $p$ is

$$V_p(t) = \frac{1}{2} \mathbb{E}|X_u - X_{u+t}|^p$$

and its method of moments estimator at lag $t = \frac{1}{n}$ is

$$V_p\left(\frac{1}{n}\right) = \frac{1}{2(n-1)} \sum_{i=l}^n |X_i - X_{i+1}|^p. \quad (2.32)$$

When $p = 2$, we recover the usual variogram, the case $p = 1$ provides the madogram, while $p = \frac{1}{2}$ gives the rodogram. The FD estimator based on the $p$-th order variogram can be obtained from the regression fit of $\log(V_p(t)$ on $\log t$.

The Incremental estimator is created similarly using $p = 1$ and in (2.32) second order differences instead.

In case of $p = 1$ the relationship between the madogram and fractal dimension is particularly universal and can be explained through the Lipschitz-Hölder heuristics of [Mandelbrot 1977], page 304, see also [Carvalho and Caetano 2012]. A real function $f(x)$ is Hölder continuous in $x$ with exponent $0 < \beta < 1$, when

$$|f(x) - f(y)| < \text{const} \cdot |x - y|^{\beta} \quad \forall y : |x - y| < \epsilon$$

holds for its increments around $x$. [Kahane 1985] Chap. 10 Sec.7 relates Hölder continuity to the Hausdorff dimension. A special case of the result in one dimension, i.e. for real functions over $R$ is that a Hölder continuous function with the supremum of its Hölder exponents $\beta$ with $0 < \beta < 1$ has a fractal graph and its Hausdorff fractal dimension is $(2 - \beta)$. Such a function is irregular, and in general, it cannot be given by a formula. The trajectory of the fractional Brownian motion may serve as an example. Now, let’s take the case of the madogram. If by regressing $\log(V_1)(t)$ on $\log t$ we get the regression coefficient $\beta$, this just means a power like growth of the increments around $t$ with exponent $\beta$, i.e. Hölder continuity with maximal exponent $\beta$. Hence the fractal dimension of the observed path is $2 - \beta$. The relationship of FD with the variogram and the rodogram is less straightforward, see [Gneiting et al. 2010].
As is well-known [Adler and Taylor 2007], the paths of the Brownian motion are nowhere differentiable, are not Lipschitz continuous, however, they are Hölder continuous in all $t$ with any of the exponents $0 < \beta < \frac{1}{2}$, and their Hausdorff dimension is $\frac{3}{2}$. Under quite broad conditions $\frac{3}{2}$ is also the fractal dimension of the paths of solutions of SDEs. This makes it necessary to change the Brownian (Wiener process) driving force in the SDEs if a process with fractal dimension different from $\frac{3}{2}$ is to be modeled. The paths of the fractional Brownian motion of Hurst exponent $H$ are Hölder continuous in all $t$ with any $0 < \beta < H$, and their Hausdorff dimension is $2 - H$, showing, that the bound in Kahane’s theorem is sharp.

The mentioned estimators are available in the [R package `fractaldim`].
Chapter 3

Stochastic Correlation for Asset Pricing and Tail Dependence

We divide this chapter into two sections. In section 3.1, five models of stochastic correlation are compared on the basis of the generated associations of Wiener processes. Associations are characterized by the copulas and their K-functions. A confidence domain for the randomly changing K-functions are build on the basis of simulated Wiener process pairs. The models are ordered by the magnitude of the domains or confidence bounds. Larger confidence domains or higher bounds represent higher correlation risk when the models are applied in mathematical finance.

The comparison is made on the basis of simulations of 10000 Wiener process pairs. The empirical copulas, and the cdf-s of their probability integral transforms, the so-called Kendall functions, or K-functions for short, are then calculated. Goodness of fit (GoF) tests in terms of the difference of the empirical and hypothesized K-functions for copula models were first proposed in [Genest et al (2006)] and further studied in [Genest et al (2009)]. In our case however, we are confronted with a situation when there is no ”true” parametric copula, the association changes randomly. So, the ”best” from the given models is the one that possesses rich enough random variation to accommodate the observed sample reliably. Reliability may be best represented by a confidence domain, constructed by the corresponding ”quantile” of the simulated centered K-functions. However, the notion of
quantile is not defined – meaning there is no universally accepted approach to it – even for vector valued random variables let alone random processes, such as the K-function of stochastic correlations. One way to overcome this issue would be to use one functional, i.e. one of the statistics of the K-functions, e.g. the Cramér-von Mises type $S^K_n$, suggested in [Genest et al (2006)]. To substitute the parametric estimated K-function in that approach, one may choose the simulation average K-function, and compute the $S^K_n$ values in every simulation. Although no distributional results remain valid, the variances of those statistics may represent the variability of the stochastically changing association structures. Although we compute these values, here we advocate for a different, quantile curve based approach, which may be capable of making finer distinctions or at least allows for further analysis of the variable structures.

In section 3.2, the association or interdependence of two stock prices is analyzed and selection criteria for a suitable model developed. The association is generated by stochastic correlation, given by a stochastic differential equation (SDE) creating interdependent Wiener processes. These, in turn, drive the SDEs in the Heston model for stock prices. To choose from possible stochastic correlation models two goodness of fit procedures are proposed based on the copula of Wiener increments. One uses the confidence domain for the centered Kendall function, and the other relies on strong and weak tail dependence. The constant correlation model and two different stochastic correlation models, given by Jacobi and hyperbolic tangent transformation of Ornstein-Uhlenbeck (HtanOU) processes are compared by analyzing daily traded closing prices for Apple and Microsoft stocks. The constant correlation i.e. the Gaussian copula model is unanimously rejected by the methods, but all others are acceptable at 95% confidence level. The analysis also reveals that even for Wiener processes, stochastic correlation is capable of creating tail dependence, unlike constant correlation which results in multivariate normal distributions and hence zero tail dependence. Hence models with stochastic correlation are suitable to describe more dangerous situations in terms of correlation risk. Also, the model created tail dependencies, both strong [Joe (1990)] and weak [Coles et al. (1999)] are important
measures for model fit. As a practical realization of the procedure, some 3271 consecutive
daily traded closing prices of Apple and Microsoft stocks are analyzed. First, the Heston
model is fitted to them and then the association of resulted residuals – predicting the
generating Wiener processes – is modeled by constant, Jacobi and Hyperbolic tangent
transformed Ornstein-Uhlenbeck processes. The fit of these models is then compared by
the quantile curves of the corresponding K-functions and the created tail dependencies.
With suitable modifications, the suggested model may serve well in multi-stock option
pricing as well.

3.1 Comparison of Stochastic Correlation Models through
quantiles of K-functions

3.1.1 Stochastically correlated Wiener processes

Stochastic differential equations are used frequently to model data series such as interest
rate, asset price, exchange rate and so on. For diffusion processes described by SDEs,
the dependence between the series most often originates in correlated Brownian motions
driving the equations. Suppose we are given a finite time horizon \([0, T]\) and a filtered
probability space \((\Omega, \mathcal{F}, \mathcal{P}, F)\) with a filtration, satisfying the usual conditions. This
is e.g. the case when the filtration is the augmented filtration of a (multidimensional)
Brownian motion. Suppose, we are given two independent Brownian motions, \(W_0(t), W_1(t)\)
adapted to the filtration \(F\) (i.e. the filtration is rich enough). Considering two adapted
correlated Brownian motions \(W_1(t)\) and \(W_2(t)\) with \(corr(W_1(t), W_2(t)) = \rho\), their quadratic
covariation,

\[
[W_1(t), W_2(t)](t) = \int_0^t \rho ds = \rho \cdot t
\]  

(3.1)
equals the covariance and is symbolically written as

\[
dW_1(t) \, dW_2(t) = \rho dt.
\]  

(3.2)
Instead of a constant $-1 \leq \rho \leq 1$, let us given an adapted process $\rho(t)$ with $|\rho(t)| \leq 1$. Since $\rho(t)$ is bounded and the time horizon is finite $\rho(t) \in \mathcal{L}_2(\Omega \times [0, T])$. By using $\rho(t)$ we define the vector process $W(t) = (W_1(t), W_2(t))$ as

$$
dW(t) = d\begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ \rho(s), \sqrt{1 - \rho^2(t)} \end{bmatrix} d\begin{pmatrix} W_1(s) \\ W_0(s) \end{pmatrix}.
$$

(3.3)

Here the second coordinate $W_2(t)$ is clearly adapted to $F$, and as a sum of two stochastic integrals w.r.t. Brownian motions of $\mathcal{L}_2(\Omega \times [0, T])$ processes it is a continuous martingale. As a direct consequence of Itô’s formula the quadratic (co)variation of $W(t)$ is

$$[W, W](t) = \int_0^t \begin{bmatrix} 1, \rho(s) \\ \rho(s), 1 \end{bmatrix} dt,$$

(3.4)

and in particular, the quadratic variation of $W_2(t)$ is $t$. Being an adapted continuous martingale with quadratic variation $t$, by virtue of Lévy’s theorem [Øksendal(2000)] this implies that $W_2(t)$ is itself a Brownian motion. Further, the quadratic covariation of $W_1(t)$ and $W_2(t)$ is

$$[W_1, W_2](t) = \int_0^t \rho(s) ds.$$

(3.5)

To the analogy of the constant $\rho$ in (3.1), we shall call the process $\rho(t)$ the stochastic correlation of $W_1(t)$ and $W_2(t)$. The Pearson correlation is

$$corr(W_1(t), W_2(t)) = \frac{1}{t} E \int_0^t \rho(s) ds = \frac{1}{t} \int_0^t E \rho(s) ds.$$

In particular, $corr(W_1(t), W_2(t)) = R$ when $\rho(t)$ is a stationary ergodic process with mean value $R$. By the ergodic theorem $\frac{1}{t} \int_0^t \rho(s) ds \to R$ as $t \to \infty$ and hence $[W_1, W_2](t) \approx R \cdot t$ when $t$ is large, indicating that the long term behaviour of stochastically correlated Brownian motions with stationary and ergodic stochastic correlation becomes similar to Brownian motions with constant correlation. Therefore, such a model choice is interesting primarily in the finite (short) time horizon context.

42
### 3.1.2 Models of Stochastic Correlation

In the applications, the correlation fluctuates around a constant value and the boundaries -1 and +1 of the correlation process are non-attractive and unattainable. Therefore we require the stochastic correlation process to satisfy the following properties:

- it takes values within the interval $(-1, 1)$,
- varies around a constant mean value in the mean reversion sense, and
- the probability mass tends to zero at the boundaries $-1, +1$.

Having this in mind, we mention two basic approaches, known in the literature and capable to satisfy the above conditions, to model the stochastic correlation:

#### 1. Bounded diffusion.
A properly normalized diffusion process may serve as stochastic correlation. One of the popular choices for modeling stochastic correlation is a Jacobi process \[^{[van Emmerich(2006)]}\] \[^{[Ma (2009a)]}\]. For more details on the Jacobi process we refer the reader to \[^{[Gouriéroux and Valéry (2004)]}\]. If $\rho(t)$ follows a Jacobi process then the time depending evolution of $\rho(t)$ is given by the following stochastic differential equation

$$d\rho(t) = k(R - \rho(t))dt + \sigma_3 \sqrt{(R_+ - \rho(t))(\rho(t) - R_-)}dZ(t). \quad (3.6)$$

Here $k$ represents the mean-reverting parameter, $R$ the mean of the process and $\sigma_3$ the diffusion coefficient, while $Z(t)$ is again a Wiener process, independent of the previous ones.

This process is known to be geometrically ergodic. It takes values between $R_-$ and $R_+$. When using it as stochastic correlation, the process is parameterized so as to be bounded either between -1 and 1, or between 0 and 1 and be centered on an equilibrium value – usually obtained by inference in the applications. Specifically, the term under the square root reduces to either $1 - \rho^2$ or $\rho - \rho^2$, respectively. The marginal distribution of a stationary Jacobi process with $\rho - \rho^2$ under the square root (i.e. $R_- = 0, R_+ = 1$) is beta,
with parameters
\[
\frac{2kR}{\sigma_3^2}, \quad \frac{2k(1-R)}{\sigma_3^2}.
\] (3.7)

2. *Transformation of a mean-reverting diffusion.* The stochastic correlation process \( \rho(t) \) can be expressed in terms of transformation of an arbitrary diffusion

\[
\rho(t) = g(X(t)),
\] (3.8)

where \( X(t) \) is a diffusion process and \( g(x) \) is a real function with values in the interval \([-1,1]\). Some of the transformations used previously in the literature are:

1. Normalized inverse tangent \([\text{van Emmerich}(2006)]\),
2. Normal cdf, \([\text{Carr}(2017)]\),
3. Hyperbolic tangent \([\text{Teng, Ehrhardt and Günther}(2016)]\).

Let us also mention the recent work \([\text{Itkin}(2017)]\), which, instead of diffusions, uses a transformed Lévy process for stochastic correlation. This approach is not subject of this thesis.

Neither of the mentioned papers analyze the properties of the created interdependence structure on the basis of historically observed or simulated data but rather concentrate on pricing derivatives written on stochastically correlated assets.

In this section, we compare the diffusion process approaches, either bounded or transformed, on the basis of simulations. For the role of diffusion in the transformations we use a mean-reverting Ornstein-Uhlenbeck process \( X(t) \):

\[
dX(t) = -\alpha(\theta - X(t))dt + \sigma_3 dW(t)(t)
\] (3.9)

with the Wiener-process \( W(t) \), independent of any previous one. In the sequel, we refer
to the corresponding transformed models as TgOU, PhiOU, HtanOU, respectively.

To do the comparison properly, suitable parameters in the model of the stochastic correlation have to be chosen. It is secured so that the models are fitted to the same estimated realization of the stochastic correlation of two stock price processes.

As a starting point, we have a given, albeit estimated, sample from the stochastic correlation. Once the realization of the stochastic correlation is given, the fit of the Jacobi model can be done by using the method of moment parameter estimation as described in [Gouriéroux and Valéry (2004)]. In the other cases, the Ornstein-Uhlenbeck parameters are fitted by the method of least squares to the back-transformed (by inverse tangent, inverse normal cdf and inverse tangent hyperbolic transformations) realization of the stochastic correlation. The inverse normal cdf is computed by the built-in function of the statistical language R package. Note here that the model in [Carr(2017)] initially creates only positive correlations, but it can also be renormalized to obtain correlations in the range $[-1, 1]$.

To illustrate that the stochastic correlation models obtained above are indeed comparable we calculate the means and the variances of those processes. The stationary Jacobi and Ornstein-Uhlenbeck processes are well known to have beta – with parameters as given in (3.7) – and normal $N(\theta, \frac{\sigma^2}{\alpha})$ distributions, respectively. To compute the mean and variance in the Jacobi case is straightforward, unlike in the case of the transformed distributions. Instead of integrating out - what would only be numerically possible anyway - we compute the approximate mean and variance of the transformed stationary Ornstein-Uhlenbeck processes via simulations only. We generate 100000 normally distributed variables corresponding to the stationary distribution of the fitted Ornstein-Uhlenbeck process, take its transform and estimate the mean and the variance from this sample. For the illustrative purpose, this precision is quite sufficient. In table 3.1 the parameters of the Ornstein-Uhlenbeck processes fitted to the back-transformed observed stochastic correlation are given together with the estimated mean and variance of the transformed Ornstein-Uhlenbeck processes.
We see that the means and variances of the various stochastic correlations do not differ significantly, so the differences and the variability in the interdependence structure of the generated Wiener processes are due to the shape of the distributions and the interdependence structure of the stochastic correlation process. Furthermore, we also consider the constant correlation case as a reference. The constant correlation of the Wiener processes is chosen as the mean of the mentioned empirical stochastic correlation process, its value is 0.5642. This way we have 5 models to compare.

### 3.1.3 Comparing the models

We simulate 10000 - 10000 pairs of Wiener increments from the Jacobi, HtanOU, TgOU, PhiOU and constant stochastic correlation models. Since the constant correlation results in the Gaussian copula for the increments, we shall also refer to it as the Gaussian model or case. In each simulation, we compute the empirical copula of the Wiener increment pair.

We only consider empirical K-functions, so we shall drop the empirical adjective from here on. We compute the K-functions in every simulation. Every K-function in this thesis is calculated in the 401 division points, which are the endpoints of a uniform partition of the [0,1]. (The endpoints 0 and 1 are also called division points.) We experienced higher resolution as well but it did not give a different outcome. We then take what we call the "centered K-functions": the difference of each simulated K-function and the average of

<table>
<thead>
<tr>
<th>Statistics type</th>
<th>Models</th>
<th>Jacobi</th>
<th>HtanOU</th>
<th>TgOU</th>
<th>PhiOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td></td>
<td>0.7068</td>
<td>0.3352</td>
<td>0.1989</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td></td>
<td>0.03342</td>
<td>0.07001</td>
<td>0.07479</td>
<td></td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td></td>
<td>0.1078</td>
<td>0.2846</td>
<td>0.2274</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>0.5675</td>
<td>0.5662</td>
<td>0.5643</td>
<td>0.5645</td>
</tr>
<tr>
<td>variance</td>
<td></td>
<td>0.04968</td>
<td>0.05118</td>
<td>0.05152</td>
<td>0.05150</td>
</tr>
</tbody>
</table>

Table 3.1: Fitted OU parameters when applicable and means and variances of stochastic correlations in the considered models
all the empirical K-functions of the corresponding model. Now, in every division point, we compute modelwise the 2.5% and 97.5% quantiles of the 10000 centered K-function values, giving the 95% confidence bounds for every model. By doing this for every division point we obtain a closed curve and inside it a domain to every model. In what sense – if at all – can this domain be regarded as confidence domain is the subject of the coming discussion.

**Definition 3.1.1.** For a centered K-function of a given model select the division points where it is out (either upwards or downwards) of the confidence bounds of that model. We call this set the exceedance set.

**Definition 3.1.2.** For a centered K-function of a given model compute the number of points of the exceedance set. Dividing it by the overall number of division points (i.e. by 401 in the given instance) we obtain a statistic that we call the exceedance proportion of the given centered K-function.

Let us note first, that the confidence bounds were computed pointwise in every division point, from 10000 K-function values. Let us now consider only one K-function, its complete curve in the division points. Were the values of a centered K-function independent in every division point, only then would roughly 5% of them be out of the bounds obtained previously. But this is clearly not the case because of the smoothness of the curve (and therefore highly autocorrelated values) of centered K-functions. This means e.g. that once the curve is out of the bounds at one division point, it is more likely (as compared to the independent case) that it will be out in its neighboring division points as well. Therefore, the exceedance proportions, computed curvewise, cannot be expected to fluctuate nicely around the 5% level, they will exceed that level much too frequently. So, the 95% quantile of the exceedance proportion will be much higher than 5%. We give these values modelwise in the first row of table 3.2. The same can be observed in the second row where we give the percentage of the simulations when the centered K-function exceeds the bounds in more than 5% of the division points, i.e. the estimation of the probability that the exceedance proportion exceeds 5%. As the numbers vary across models, the curves obtained from the pointwise confidence bounds are unsuitable for comparing the
models.

<table>
<thead>
<tr>
<th>Exceedance proportion</th>
<th>Models</th>
<th>Jacobi</th>
<th>HtanOU</th>
<th>TgOU</th>
<th>PhiOU</th>
<th>Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% quantile</td>
<td>0.2597</td>
<td>0.2419</td>
<td>0.2394</td>
<td>0.1726</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability estimate</td>
<td>0.2200</td>
<td>0.2063</td>
<td>0.2389</td>
<td>0.3008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiplier for 95% CD</td>
<td>1.3112</td>
<td>1.3139</td>
<td>1.3281</td>
<td>1.3342</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S^K_n$</td>
<td>0.5041</td>
<td>0.5928</td>
<td>0.4816</td>
<td>0.4008</td>
<td>0.1807</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Goodness of fit statistics and their significance

So, we use the curves of pointwise confidence bounds only as reference curves and blow them with a constant multiplier so that exactly 5% of the centered K-functions will exceed the curves in 5% of the division points in every model. We give the multipliers in the 3rd row of table 3.2. These closed curves and the domains they bound may then serve for comparison of models, by representing the variability of the generated interdependence structures and copulas in terms of 95% “quantiles” or 95% confidence domains (CD) in a certain well defined, as above, sense. By choosing $p\%$ instead of 5% with variable $p$ we may get a set of curves characterizing each model. In the 4th row of table 3.2 we give the variances of the Cramér-von Mises type $S^K_n$ statistics [Genest et al (2006)]. According to this the order of variability of the created interdependences is HtanOU, Jacobi, TgOU, PhiOU and Gaussian. We compare the obtained curves in Figure 3.1.

![Comparison of 95% confidence domains](image)

Figure 3.1: (95% confidence domains for the various models)

As is clear from figure 3.1 the HtanOU model creates the largest confidence domain,
meaning that it has the greatest variability in this sense and hence it is the most flexible in accommodating an eventual sample from an observed or estimated stochastic correlation.

In order of domain magnitude the Jacobi model follows and it does not differ much to the TgOU model, the curves are hardly distinguishable in the figure. The domain of

![Graph showing confidence bounds near unity](image)

Figure 3.2: (logarithm of the 95% confidence bounds of models near the unity)

the PhiOU is a little smaller meaning the model creates less variable associations. The domain corresponding to the constant correlation i.e. the Gaussian copula is strikingly smaller, showing how little variability is inherent in this model as compared to the others. So the order of variability, judged by the magnitude of the confidence domains coincides with the one obtained by the Cramér-von Mises statistic. However, the situation changes when considering extreme dependence, which influences the K-function values taken in division points near the unity. In order to see the differences there, we take the logarithm of the bounds on the positive upper halfplane (avoiding endpoint 1 where the logarithm is infinity), as displayed in Figure 3.2. It is well visible that the order of models by the magnitude of the confidence bounds changes here. The largest now is the TgOU, while the HtanOU, Jacobi and PhiOU are indistinguishable, at least on the given level of the simulations. Again, the Gaussian remains uniquely the smallest. This results in a different model preference if an observed K-function goes out of the confidence bounds near unity. So, considering the confidence domains instead of just one function of the K-functions as the Cramér-von Mises statistic, may alter the model choice and may open a way to further analysis of the observed K-function.
Thus in this section, we compared 5 models of stochastic correlation. The models are compared according to their capability to produce more or less variable associations between the Wiener processes. Beyond the variability of usual GoF statistic, we proposed to build certain confidence domains and judge the models by the magnitude of the domains, or their bounds. The order of model preference may vary depending on which part of the domain of the K-function is taken into account. In financial applications, one concern may be the variability of the association of assets, increasing the correlation risk. An equally important concern may be the very strong association when diversification of risk is not possible. It turns out that our confidence domain approach may take into account the two different approaches. This way it can help investors of whom the more conservative ones may want to use the more variable models when inferring about the risks and pricing their financial products, while risk takers may use less variable models.

3.2 Modeling Joint Behaviour of Asset Prices using Stochastic Correlation

In this section, for the association or interdependence of two stock prices, a suitable model is developed. The association is generated by stochastic correlation, given by a stochastic differential equation (SDE) creating interdependent Wiener processes. To choose from possible stochastic correlation models two goodness of fit procedures are proposed based on the copula of Wiener increments. One uses the confidence domain for the centered Kendall function, and the other relies on strong and weak tail dependence.

3.2.1 Modeling a Single Stock Price; the Heston Model

In order to analyze the association of the generating Wiener processes for observed asset prices, a proper model is needed and its discretized version fitted to the price data. Time dependent and random nature of volatility has been noted for a long in quantitative analyses
of financial assets and to our days the Heston model \cite{Heston1993} became one of the industry standards to incorporate this property into the description. In the model the asset price follows a square-root diffusion while the spot volatility process follows a CIR process. The general form of the model is given as follows:

\begin{align}
    dS(t) &= \mu S(t)dt + \sqrt{V(t)}dW_s(t) \\
    dV(t) &= \kappa(\theta - V(t))dt + \sigma_v\sqrt{V(t)}dW_v(t)
\end{align}

(3.10) \quad (3.11)

where, $W_s(t)$ and $W_v(t)$ are Wiener processes with a constant correlation $\rho_H$. In our setup, when two stock prices are considered, the price driving Wiener processes i.e. the corresponding $W_s(t)$-s are supposed to be stochastically correlated creating the association of prices.

In order to estimate the parameters from a discrete sample and obtain residual processes predicting the Wiener generators, we consider the Euler-Maruyama scheme to obtain a discretized version of the model. First, by Itô’s lemma, we have

\begin{align}
    d\log S(t) &= (\mu - \frac{1}{2}V(t))dt + \sqrt{V(t)}dW_s(t), \\
    d\log S(t) &= (\mu - \frac{1}{2}V(t))dt + \sqrt{V(t)}dW_s(t)
\end{align}

(3.12)

Consider the log-returns within $\Delta$ period of time $Y(t) = \log S(t + \Delta) - \log S(t)$, then the approximation of (3.12) and (3.11) can be written in the discrete setup as

\begin{align}
    Y(t) &= (\mu - \frac{1}{2}V(t))\Delta + \sqrt{V(t)}\sqrt{\Delta}\varepsilon_s(t), \quad t = 0, \Delta, 2\Delta, \ldots
\end{align}

(3.13)

\begin{align}
    V(t + \Delta) - V(t) &= \kappa(\theta - V(t))\Delta + \sigma_v\sqrt{V(t)}\sqrt{\Delta}\varepsilon_v(t), \quad t = 0, \Delta, 2\Delta, \ldots
\end{align}

(3.14)

\begin{align}
    V(0) = v_0
\end{align}

(3.15)

where $\varepsilon_s(t)$ and $\varepsilon_v(t)$ are i.i.d $N(0, 1)$ sequences with correlation $\text{corr}(\varepsilon_s(t), \varepsilon_v(t)) = \rho_H$.

Equation (3.13) can be further rearranged to obtain an expression for $\varepsilon_s(t)$.
\[ \varepsilon_s(t) = \frac{Y(t) - \mu \Delta t + \frac{1}{2} V(t) \Delta t}{\sqrt{V(t-1) \Delta t}} \] (3.16)

Now, the aim is to estimate the parameters of the Heston model from \( n + 1 \) discrete, \( \Delta \)-period time-equidistant observations of the stock price \( S(t) \) in the time interval \([0, T]\), i.e. from the sample \( S(0), S(\Delta), S(2\Delta), \ldots, S(n\Delta) \) where \( T \) is given by \( n \cdot \Delta = T \). In this model, the parameters to be estimated are \( \mu \) (expected return of asset price), \( \kappa \) (mean reversion speed), \( \theta \) (long run mean of variance), \( \sigma_v \) (volatility of variance) and \( \rho_H \) (correlation between asset price and its variance process).

The estimation of the Heston model’s parameters is not straightforward, and various methods have been elaborated to obtain it. \([\text{Li et al. (2008)}]\) suggested a sophisticated MCMC algorithm to obtain the estimations, and more recently, \([\text{Cape et al. (2015)}]\) derived the necessary prior distributions for the effective implementation of the algorithm. It is mentioned in \([\text{Cape et al. (2015)}]\) that the variance process’s estimation is not perfect in the MCMC procedure. This is completely in line with our experience. Beyond that, in the application described in the subsequent subsection, the obtained residuals predicting the Wiener processes do not fit normal distributions; neither of the known normality tests accepts them. Hence we suggest a minor modification in the procedure. Originally the estimated volatility process is updated by a normally distributed random variable with 0 mean and a small variance in the MCMC algorithm. We suggest changing the distribution to a gamma one. Our motivation is that the volatility process satisfies a CIR equation known to have a stationary solution with gamma marginal distribution, and the latter has the same gamma distribution as its conjugate. By fine-tuning the applied gamma update parameters, a delicate balance can be achieved that brings the distribution of the residuals of both the price driving equation (let us call them price-residuals for short) and the variance equation equally close to normality. As a result of the Heston model fit, the price-residuals are obtained from the observed stock price series, opening the way to study the dependence structure of the price driving Wiener-processes.
3.2.2 Stochastic Correlation Model: Goodness of Fit by Quantile Curves of K-functions

Having the fitted models to the prices, the price-residuals as predictors of the price driving Wiener processes can be analyzed further. We stress here that as the $\Delta$ period Wiener increment pairs have the same two-dimensional distribution, the differenced price-residual pairs can be regarded as a sample from that distribution. In the case of two stocks, the first aim is to model their time-dependent stochastic correlation. To this end, "local" correlations of the two price-residual processes are generated by taking a backward-looking window around a time point $t$. Looking through it to the two residuals, the Pearson correlation is estimated the usual way, and the value is allocated to $t$. Running then the window in time through the price residuals, the time-dependent estimated stochastic correlation is obtained. Each of the mentioned stochastic correlation models: HtanOU, Jacobi, and constant is fitted to this estimated correlation process.

The choice of the window length is crucial in this procedure, and it can be chosen by using a "proxy" t-copula. First, we fit a t-copula to the differenced price-residuals. Then, having the estimated correlation process, we can simulate with that stochastically correlated Wiener process pairs, differencing it and fitting a t-copula again to these pairs. Although the t-copula fit is far from perfect, the window size, which preserves best the parameter of the t-copula, is likely to be the best choice.

Once we have an estimated stochastic correlation process, it is pretty straightforward to fit the mentioned models. The fitting of a Jacobi process can be carried out using the estimator described in [Gouriéroux and Valéry (2004)]. To obtain the HtanOU model parameters, the estimated stochastic correlation process is back-transformed first, and then the usual least squares estimator of the Ornstein-Uhlenbeck process is applied. The constant correlation model has only one parameter, the constant, which is chosen as the average of the estimated stochastic correlation process.

Assessment of the goodness-of-fit of the mentioned three stochastic correlation models is possible based on the copulas they induce and their Kendall functions or K-functions for short. Following [Markus and Kumar(2019/1)], the simulated quantile curves of the
centered K-functions are determined. The difference between the empirical K-function and the averaged simulated K-functions is tested against the curves. Remark here that throughout the paper, the empirical copula’s notion is used in the rigorous sense of [Ghoudi and Remillard], section 3.2 or more specifically, and explicitly section 3.4. Correspondingly the empirical K-function is the Kendall function of the empirical copula. We recall, in short, the procedure described in [Markus and Kumar(2019/1)] in detail.

In order to create the empirical quantile curves, 10000 - 10000 pairs of Wiener increments are simulated from the Jacobi, HtanOU, and constant stochastic correlation models. In each simulation, the Wiener increment pair’s empirical copula, and subsequently, their empirical K-function, is computed. These empirical K-functions are then centered model-wise by their average value, taken in the 10000 simulations from the same model. In every point where these centered K-functions are evaluated, the 2.5% and 97.5% quantiles of the 10000 values are computed, giving the pointwise 95% confidence bounds for every model. By doing so for every point, a closed curve and in it a domain are obtained for every model. These are only used as reference curves and domains, and a constant multiplier is determined to them so that exactly 5% of the centered K-functions go out from the domain. These curves and domains are considered as the confidence curves and domains of the corresponding models. As it has been said, the fit is good or statistically acceptable on a 95% level if the difference of the price-residuals induced empirical K-function and the simulated model average K-function remains within the domain.

### 3.2.3 Pseudo-algorithm of the methodology

This section gives a high-level overview of the methodology explained in sections [3.2.1] and [3.2.2].

1. From the given market data (i.e., closing prices of the two stocks), calculate the log-returns, and estimate the parameters of the discretized Heston model by MCMC, as explained in section [3.2.1] With the estimated parameters, price residuals of each stock can be estimated by using equation [3.16].

2. Next, test the normality and independence of the above price residuals processes
using any standard statistical tests (Anderson-Darling, Shapiro-Wilk, etc.). Having passed the tests, the price residual pairs can be regarded as a sample of Wiener increment pairs, as explained in section 3.2.2.

3. Estimate the 'local' correlations of two price residuals processes by using a sliding window technique. The reasoning behind the choice of an appropriate window size is explained in section 3.2.2. This procedure gives the estimated stochastic correlation process.

4. Fit the three models: HtanOU, Jacobi, and constant to the estimated stochastic correlation process, as explained in section 3.2.2.

5. Simulate the desired number of paths of the stochastic correlation process and assess each model’s fit based on the copulas and K-functions they induce, as explained in section 3.2.2. For the in-depth details of this step, follow the methodology presented in [Markus and Kumar(2019/1)].

6. Simulate the price driving Wiener process increment pairs associated according to the simulated stochastic correlations and compute the strong and weak tail dependence as defined in section 2.3.4. Compare it to the tail dependence of the price residual processes.

3.2.4 Data of Two Stock Prices; Fitting the Model

In the present study, we analyze the association of prices of two stocks, Microsoft (MSFT) and Apple (AAPL). The data are registered 12 years long between January 05, 2006 - January 04, 2019, and consist of 3271 daily close prices. First, we fit Heston stochastic volatility models to both MSFT and AAPL stock price data, using the MCMC algorithm mentioned in section 3.2.1, with the gamma-distributed update (shape = 5.6, rate = 236 for both Apple and Microsoft) for the variance process. From the fitted Heston models, two price-residuals $\varepsilon_{A,s}$ and $\varepsilon_{M,s}$ are obtained for the two stocks. Kolmogorov-Smirnov, Anderson-Darling and Shapiro-Wilk tests reject deviation from the normal distribution for all residuals. True though, while the serial dependence of the price-residuals is rejected,
e.g., by the powerful Brock-Dechert-Scheinkmann test and some other simpler independence tests (difference-sign, turning-point, or rank tests), the volatility-residuals show a minor autocorrelation for the first three lags. The obtained variance process nicely follows the empirical one obtained by a windowed estimation. To our experience, however, the use of normal updates suggested in [Cape et al. (2015)] for the volatility process for the AAPL and MSFT data seriously distorts the empirical variance processes. As a consequence, the generated residual processes of AAPL and MSFT are no longer normal and independent, although the estimated model parameters do not change substantially. Since we could not observe a similar phenomenon when instead of AAPL and MSFT, the procedure is applied to ”artificial” prices, simulated from the Heston model, we conclude that the need for gamma update may lay in the fact that the observed two stock data do not perfectly follow Heston models.

The two price-residuals are regarded as predictors of the price-governing Wiener processes. The pair of their increments \((\varepsilon_{A,s}(t + \Delta) - \varepsilon_{A,s}(t)), (\varepsilon_{M,s}(t + \Delta) - \varepsilon_{M,s})\) where \(\Delta\) is the discretisation time-length, can be regarded for every \(t = 1, 2, \ldots, 3271\) as a predicted sample from the model of the same two dimensional distribution, i.e. the joint distribution of the two Wiener process increments in \(\Delta\) time. Hence, this quasi-sample makes it possible to compute the empirical copula of the predicted Wiener process increment pair, its K-function and ultimately its strong and weak tail dependences.

### 3.2.5 Evaluating Model Goodness

The association of stock price processes is inherent from the stochastically correlated Wiener processes governing the SDEs. As a result, the association is random for both the prices and the Wiener processes. Therefore, the ”observed” snapshot of the association, represented by the price-residuals’ empirical copula, i.e., the predicted Wiener increments, is not a fixed model characteristic, but only one realization, a ”path”, from the randomly changing structure. Consequently, the question of model goodness of fit should not be raised so as ”which model creates a copula closest to the empirical”, but the way ”which model creates sufficiently rich variability of copulas so as to accommodate the observed
sample conveniently, reliably”.

We now have to check whether the obtained correlations, copulas, and tail dependence are coherent with those obtained from the models. For that purpose, the ∆ length increments of $3271 \cdot \Delta$ long stochastically correlated Wiener process pairs are simulated 10000-10000 times from the fitted HtanOU, Jacobi, and constant stochastic correlation models, respectively. Consider first the correlations, as presented in Table 3.3. Even though the price paths correlate by 0.9104, the log-returns do only by 0.456, due to eliminating the exponential growth. The correlations, averaged for simulated synthetic log-returns in both stochastic correlation models (third column), fit this value well. However, the constant correlation model shows some bias in this respect. The standard deviation of these correlations (4th column) are comparable for the TanhOU and Jacobi models, the first being slightly larger, while markedly lower for the constant correlation model showing less variability in the dependence of the created log-returns.

Table 3.3: Correlation of log returns of the stock prices and their simulated counterparts. The mean and standard deviation of the estimated stochastic correlation process of the price residuals and the average and standard deviation of the means and standard deviations of the 10000 simulated stochastic correlation processes.

<table>
<thead>
<tr>
<th>Model</th>
<th>Pearson Correlations</th>
<th>Stochastic Correlation Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed log-returns</td>
<td>Simulated 10000 log-returns</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>st.dev</td>
</tr>
<tr>
<td></td>
<td>Estimated from Price</td>
<td>means</td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>Std. Deviations</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>st.dev</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>st.dev</td>
</tr>
<tr>
<td></td>
<td>std.dev</td>
<td>std.dev</td>
</tr>
<tr>
<td>TanhOU</td>
<td>0.456</td>
<td>0.301</td>
</tr>
<tr>
<td></td>
<td>0.399</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>0.401</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>0.280</td>
<td>0.011</td>
</tr>
<tr>
<td>Jacobi</td>
<td>0.437</td>
<td>0.269</td>
</tr>
<tr>
<td></td>
<td>0.392</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>0.247</td>
<td>0.008</td>
</tr>
<tr>
<td>Constant</td>
<td>0.393</td>
<td>0.211</td>
</tr>
</tbody>
</table>

This is a consequence of the random association rather than any error or noise in the model and reflects the randomly changing environment where the prices are set. So, it is rather a strength than a weakness of the model.

The mean of the estimated (from the price residuals) stochastic correlation process is 0.386, and the means of the 10000 simulated stochastic correlation processes (7th column) coincide stunningly well with this value. The variability of these values is
very low in the simulations (8th column), showing their robustness in this respect. The standard deviations of the same processes are considerable, reflecting the random nature of the temporally localized association structure. The average standard deviations of the simulated stochastic correlation processes (9th column) are, again, close to the estimated one and pretty stable (10th column), although the Jacobi model shows somewhat less variability. In the case of the constant correlation, such considerations are meaningless, of course.

Summarizing, we can say those correlations are realistic and reflect well the observed situation while giving sufficient room to random changes, an essential source of the risk modeled.

Next, the empirical copula and the empirical K-function are computed from the 3270 Wiener increment pair in every simulation. We use these simulations to obtain the quantile curves for centered K-functions as in [Markus and Kumar(2019/1)] and the model induced lower tail dependence by using the expression in equation (2.17).

Figure 3.3: Confidence bounds of HtanOU, Jacobi and constant stochastic correlation models, and the model-centered K-functions of price residuals of Apple and Microsoft stocks

We create the 5% confidence domains model-wise for the centered K-functions, as described in section 3.2.2. Having the empirical K-function computed from the stocks by the described procedure, we center it corresponding to each model by the simulation averages, obtaining three centered empirical K-function curves, one to each considered model. All we have to check now is whether these curves are enclosed within the corresponding
model's confidence domains. This can be observed in Figure 1. (Note that the original image contains more than 12 million K-function points; hence it has to be processed, and the vertical white stripes result from image compression only.)

Clearly, the domains corresponding to the Jacobi and HtanOU models conveniently accommodate the curves, while in the constant correlation case, the curve goes out from the domain. This means that the fit is acceptable for the Jacobi and HtanOU models, whereas it has to be rejected for constant correlation.

Turning to tail dependence, we have 10000 simulations from each model, and so, 10000 copulas of Wiener process increments model-wise. With these, we use (2.17) and (2.18) to obtain an approximation of the strong and weak tail dependence, respectively. The expression in (2.17) and (2.18) are still functions of \( u \), and for tail dependence, we are interested in it only for the smallest \( u \)-s since theoretically, we would need the limit in 0. However, for the lowest \( u \) values, the estimation is based on just a few simulated values; hence it is unreliable. So, we need a balance between closeness to 0 and the reliability of estimation. Therefore, we drop the first 5 values to get rid of the bias they cause and average \( \hat{L}_n(u) \) and \( \hat{L}^w_n(u) \) for the 6 to 15 smallest \( u \)-values in order to capture the true behavior in the left tail. In order to characterize the models, we present in table 1 the mean and the standard deviation along with the 95% confidence bounds of these values characterizing the weak and strong tail dependence corresponding to realizations of the randomly changing association structure.

<table>
<thead>
<tr>
<th>Model type</th>
<th>Strong Tail dependence</th>
<th>Weak Tail dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empirical case</td>
<td>0.175</td>
<td>0.45</td>
</tr>
<tr>
<td>Mean</td>
<td>Standard deviation 95% quantile</td>
<td>Mean Standard deviation 95% quantile</td>
</tr>
<tr>
<td>Jacobi</td>
<td>0.159</td>
<td>0.056</td>
</tr>
<tr>
<td>HtanOU</td>
<td>0.147</td>
<td>0.055</td>
</tr>
<tr>
<td>Constant</td>
<td>0.072</td>
<td>0.049</td>
</tr>
</tbody>
</table>
As the table shows, the strong and weak mean tail dependence of both Jacobi and HtanOU models are close to the observed values, and the Jacobi model is proven to be the better in these terms. The 95% bounds are large enough to accommodate the observed values for both models. However, the constant correlation model fails to fit in this sense as well. The mean values for constant correlation reflect the known fact that the multivariate normal distribution has no tail dependence. Even the confidence intervals are too narrow to include the observed values. Therefore, based on both strong and weak tail dependence, the assumption of constant correlation has to be rejected on this level.

It is also worth noticing that the mean tail dependencies of the Jacobi and HtanOU models are also not included in the confidence interval of the constant correlation model. This indicates that at least in some sense, these latter models truly create tail dependence.

### 3.2.6 Conclusion

In this chapter we first analyzed 5 models of stochastic correlations. The models were compared according to their capability to produce more or less variable associations between the Wiener processes. Thereafter, we built up models to describe associations of two stock prices, based on stochastic correlation described by Jacobi and transformed Ornstein-Uhlenbeck processes. In financial applications, one concern may be the variability of the association of assets, increasing the correlation risk. An equally important concern may be the very strong association when diversification of risk is not possible. We developed two selection criteria based on the copula of the price driving Wiener process increments in the Heston model. One of them tests whether the centered K-function obtained from the price residuals of the corresponding model remains within the 95% confidence domain determined by simulated quantile curves. The other tests the strong and weak tail dependence of the mentioned empirical copula to the 95% confidence bounds, obtained again from simulations. We apply these criteria in the case of daily close prices of the Apple and Microsoft stocks. The three different models to compare or select from are defined by stochastic correlation types of constant correlation, Jacobi process, and hyperbolic tangent transformation of an Ornstein-Uhlenbeck (HtanOU) process. All
criteria reject the constant correlation and hence the Gaussian copula for the Wiener processes. Both Jacobi and HtanOU models can be accepted on the 95% levels by all criteria. As a consequence it can be said that while the Gaussian model is inappropriate for describing the interdependence, both the HtanOU and the Jacobi models can be suitable alternatives. In terms of K-functions the HtanOU model accommodates most easily the empirically observed interdependence. On the other hand, the mean tail dependence of the Jacobi model is the closest to the observed one, although the difference from the HtanOU is not large.

It is also important to see that stochastic correlations are capable of creating tail dependence even for Wiener processes, when the marginals are normally distributed. This is in sharp contrast to constant correlation which, in case of Wiener processes, results in a bivariate normal distribution of the increments, and as such, corresponds to the Gaussian copula having 0 i.e. no tail dependence [Shibuya (1960)].

We found that the price residuals are tail dependent and hence only tail dependent stochastic correlation may work for simultaneous modeling of the stock prices. This indicates that the source of stock price association is not only the similarity of the random environment in which they are set, i.e. the volatility, but also the internal price driving forces i.e. some fundamentals must act in a synchronized way. These effects increase the correlation risk, hence they must be accounted for e.g. in portfolio selection or other investment decisions.

A question for further research is how much tail dependence is inherited from the Wiener generators to the prices themselves. The association of the variance processes, their driving Wiener processes, and their effect on the price associations are also subject to further study.
Chapter 4

Arbitrage-Free Pricing of CVA for Cross-Currency Swap with Wrong-Way Risk under Stochastic Correlation Modeling Framework

In this chapter we revisit the problem of wrong way-risk and propose a stochastic correlation approach which is unique till date. A positive correlation between exposure and counterparty credit risk gives rise to the so called Wrong-Way Risk (WWR). Even after a decade of financial crisis, addressing WWR in a both sound and tractable way remains challenging. Academicians have proposed arbitrage-free set-ups through copula methods but those are computationally expensive and hard to use in practice. Resampling methods are proposed by the industry but they lack in mathematical foundations. The purpose of this chapter is to bridge this gap between the approaches used by academicians and industry. To this end, we propose a stochastic correlation approach to asses WWR. The methods based on constant correlation to model the dependency between exposure and counterparty credit risk assume a linear dependency, thus fail to capture the tail dependence. Using a stochastic correlation we move further away from Gaussian copula and can capture the tail risk. This effect is reflected in the results where the impact of stochastic correlation on
calculated CVA is substantial when compared to the case when a high constant correlation is assumed between exposure and credit.

4.1 Credit Valuation Adjustment and Wrong Way Risk

We denote the underlying filtered probability space as \((\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})\), where, \(\mathcal{F}_t\) is the filtration which contains the market information, including default monitoring up to time \(t\) and \(\mathcal{P}\) represents the physical measure. We consider two parties X and Y, where X can be an investment bank and Y which can be any financial institution entering into a cross-currency swap with a maturity of ten years. In addition to measure \(\mathcal{P}\) we also suppose two risk neutral pricing measures associated with domestic and foreign markets denoted \(\mathcal{Q}\) and \(\mathcal{Q}^*\), respectively. Here, we only consider unilateral CVA assuming that domestic investor X is default free, therefore, all the calculations are done from the point of view of X by using domestic risk neutral measure \(\mathcal{Q}\). This measure is associated with the risk-free money market account numéraire denoted \(B_t\), evolving according to the risk-free rate \(r_t\)

\[
  dB_t = r_t B_t dt, \quad B_0 = 1 \tag{4.1}
\]

Following this setup, CVA can be computed as the \(\mathcal{Q}\)-expectation of the loss given default due to counterparty’s default. The risk-neutral exposure faced by firm X is denoted as \(V_t^+\). The general formula for the CVA can be written as

\[
  CVA = \mathbb{E}^{\mathcal{Q}} \left[ (1 - RR) 1_{\tau \leq T} \frac{V_t^+}{B_\tau} \right] = (1 - RR) \mathbb{E}^{\mathcal{Q}} \left[ \mathbb{E}^{\mathcal{Q}} \left[ \sigma(\xi_u, 0 \leq u \leq t) \right] \right] \tag{4.2}
\]

where \(\mathbb{E}^{\mathcal{Q}}\) denotes the expectation under the risk neutral measure \(\mathcal{Q}\), \(\xi = (\xi_t)_{t \geq 0}\) is the default indicator process which is defined as \(\xi_t = 1_{\tau \leq t}\) and \(RR\) is the recovery rate which is assumed be constant at 25%. Second equality follows from the assumption of constant
recovery rate and the tower property. Equation (4.2) can be expressed in integral form as

\[ CVA = - (1 - RR) \int_0^T E^Q \left[ \frac{V^+_\tau}{B_\tau} \mid \tau = t \right] dS(t), \] (4.3)

where \( S(t) \) is the risk-neutral survival probability given as

\[ S(t) = \mathcal{Q}[\tau > t] = E^Q[1_{\tau > t}] \] (4.4)

Wrong-way risk (WWR) occurs when the exposure \( V^+_t \) tends to grow as the default probability of the counterparty becomes larger. In this case, one has to account for the dependency between exposure and credit. To capture this effect, exposure and credit need to be modeled jointly. This chapter aims to model the dependence between the exposure and probability by using stochastic correlation to capture this dependence. Estimating the CVA itself is model independent, however, the underlying derivative is model dependent, this leads to more complications in the calculation of CVA.

### 4.2 Exposure in Cross-Currency Swap

Cross-currency swap (CCS), a special case of interest rate swap is a derivative contract, agreed between two counterparties, which specifies the nature of the exchange of payments denominated in two different currencies. We take into account the case of floating vs floating cross currency swap which is seen as an exchange of two floating rate bonds one in the domestic currency with notional \( N^d \) and one in a foreign currency with notional \( N^f \). The payment of floating legs are indexed to the domestic and the foreign interest rates, for example LIBOR. Here, we assume that the domestic investor pays the floating legs in USD while the foreign investor pays in EUR, therefore, all the foreign leg cash flows need to be converted into the domestic currency by using the spot exchange rate prevailing in the market. When exchanging two currencies on the market for immediate delivery the spot exchange rate, \( \psi_0 \), is defined as
ψ_0 = \frac{\text{Number of units of domestic currency}}{\text{Number of units of foreign currency}} \quad (4.5)

The two notional amounts are related as \( N_d = \psi_0 N_f \). In depth details on valuation of CCS can be found in [Brigo, Morini, Pallavivini(2013)]. A prototypical floating vs floating CCS exchanges cash flows between two indexed legs, starting from a future time. Consider the time partition of the interval \([0, T]\) as \( \{0 = T_\alpha, ..., T_\beta = T\} \), where \( T \) is the maturity of the CCS contract. We assume that there are 360 trading days and cash flow occur quarterly. For simplicity, we also assume that the notional is 1 unit of domestic currency and both the floating-rate payments occur at the same dates and with the same year fractions. Thus, the value of the CCS at time \( t < T_\alpha \), from the payer’s point of view is

\[
\Pi_{\text{payer}}(t) = N_f \left[ \psi(t)(P^f(t, T_\alpha) - P^f(t, T_\beta)) + 1_{\{t = T\}} N^f \psi(T) - N^d \left[ P(t, T_\alpha) - P(t, T_\beta) \right] - 1_{\{t = T\}} N^d \right],
\]

where \( \psi(t) \) is the spot foreign exchange rate. \( N^f \psi(T) \) denotes the notional value received by party X at the maturity of CCS converted into domestic currency. We can look at the value of the swap from the side of the party Y, then the value is simply obtained by changing the sign of the cash flows

\[
\Pi_{\text{receiver}}(t) = -\Pi_{\text{payer}}(t).
\]

The detailed description on how to derive equation (4.6) can be found in Appendix A.

The exposure for the domestic (payer) investor can then be calculated by using the following relation:

\[
V^+(t) = \max(\Pi_{\text{payer}}(t) - C(t), 0)
\]

where \( \Pi_{\text{payer}}(t) \) is given in equation (4.6) and \( C \) is the collateral (if any) posted by firm Y.

In the aftermath of financial crisis of 2007-08, OIS rates are usually used for discounting purposes, however, in this study we do not take into account the collateralized portfolio.
(i.e., we assume firm Y posts no collateral), therefore we use LIBOR rates as proxies for risk-free rates and thus for discounting the cash flows to estimate the CVA at time 0. This reasoning can be further investigated in [Hull and White(2013)].

4.2.1 Model for Multi-Currency with Short-rate Interest Rates

Two markets need to be modeled simultaneously, therefore, we need to understand the standard concept of no-arbitrage. Two saving accounts for the role of numéraire are defined, one for the domestic market

\[ B^d(t) = e^{r^d t}, \]  

(4.9)

and one for the foreign market

\[ B^f(t) = e^{r^f t}. \]  

(4.10)

The foreign exchange rate process \( \psi(t) \) represents the domestic price at time \( t \) of one unit of the foreign currency. It is denominated in units of domestic currency per unit of the foreign currency. The foreign Exchange rate process is modeled by using the [Garman and Kohlhagen(1983)] model. This model is an extension of the Black-Scholes model in order to manage the two interest rates, one domestic and one foreign. The dynamics of \( \psi(t) \) is given by the following stochastic differential equation:

\[ d\psi(t) = \psi(t)(\mu_\psi dt + \sigma_\psi dW(t)), \]  

(4.11)

with constant drift \( \mu_\psi \) and volatility \( \sigma_\psi \) and \( W(t) \) is a standard Brownian motion under the physical measure \( \mathcal{P} \). In order to exclude arbitrage the drift of the foreign exchange rate process \( \psi(t) \) need to be adjusted and its dynamics is thus given under the domestic risk-neutral measure \( \mathcal{Q} \). The following result is an application of Girsanov’s theorem and can be found in [Garman and Kohlhagen(1983)].
Proposition 2. The dynamics of foreign exchange rate process $\psi(t)$ under the domestic martingale measure $\mathcal{Q}$ is described by

$$d\psi(t) = \psi(t)((r_d - r_f)dt + \sigma_\psi dW^\mathcal{Q}_\psi(t)),$$

(4.12)

where $W^\mathcal{Q}_\psi(t)$ follows a Brownian motion under $\mathcal{Q}$ and $r_d$ and $r_f$ are the prevailing domestic and foreign interest rates, respectively. Furthermore,

$$\psi(t) = \psi(0) \exp((r_d - r_f - \frac{1}{2}\sigma^2)dt + \sigma_\psi W^\mathcal{Q}_\psi(t)),$$

(4.13)

Now, we are ready for the analysis of the underlying interest rate processes $r_d(t)$ and $r_f(t)$. It is assumed that the interest rate dynamics are defined via short-rate processes, which under their spot measures, i.e., $\mathcal{Q}$-domestic and $\mathcal{Q}^*$-foreign are driven by Hull-White one factor (HW1F) model given as,

$$dr_d(t) = [\theta_d(t) - \beta_d r_d(t)]dt + \sigma_d dW^\mathcal{Q}_d(t)$$

(4.14)

$$dr_f(t) = [\theta_f(t) - \beta_f r_f(t)]dt + \sigma_f dW^\mathcal{Q}^*_f(t),$$

(4.15)

where $W^\mathcal{Q}_d(t)$ and $W^\mathcal{Q}^*_f(t)$ are the Brownian motion under $\mathcal{Q}$ and $\mathcal{Q}^*$ respectively. Parameters $\beta_d$, $\beta_f$ and $\sigma_d$, $\sigma_f$ are positive constants and $\theta_d$ and $\theta_f$ are deterministic functions chosen so as to exactly fit the term structure of interest rates being currently observed in the market.

The above processes, under the appropriate measures, are linear in their state variables, so that for a given maturity $T$ ($0 < t < T$) the zero-coupon bonds (ZCB) have the Affine structure,

$$P^d(t,T) = A^d(t,T) \exp(-B^d(t,T)r_d(t)),$$

(4.16)

$$P^f(t,T) = A^f(t,T) \exp(-B^f(t,T)r_f(t)),$$

(4.17)
where \( A^d(t, T), A^f(t, T), B^d(t, T), B^f(t, T) \) are analytically known quantities and can be found in [Brigo, Morini, Pallavicini(2013)].

The spot rates at time \( t \) are defined by
\[
\begin{align*}
   r^d(t) &\equiv f^M_d(t, t), \\
   r^f(t) &\equiv f^M_f(t, t),
\end{align*}
\]
where, \( f^M_d(t, t) \) and \( f^M_f(t, t) \) are the domestic and foreign instantaneous forward rates prevailing at time \( t \).

In general, under HW1F, the market instantaneous forward rate at time \( 0 \) for the maturity \( T \) is defined as
\[
f^M(0, T) = -\frac{\partial \ln P^M(0, T)}{\partial T},
\] (4.18)
where \( P^M(0, T) \) are the market implied ZCB prices prevailing at time \( 0 \) for the maturity \( T \), we must have

\[
\theta(t) = \frac{\partial f^M(0, T)}{\partial T} + \beta f^M(0, t) + \frac{\sigma^2}{2\beta^2} (1 - e^{-2\beta t}),
\] (4.19)
where \( \frac{\partial f^M}{\partial T} \) denotes partial derivative of \( f^M \) w.r.t its second derivative. This gives us the values of \( \theta_d \) and \( \theta_f \) in equations (4.14) or (4.15).

An alternative way to express HW1F model is by integrating equation (4.14) or (4.15), which gives us
\[
r(t) = r(s)e^{-\beta(t-s)} + \alpha(t) - \alpha(s)e^{-\beta(t-s)} + \sigma \int_s^t e^{-\beta(t-s)} dW^\mathbb{Q}_u,
\] (4.20)
where,
\[
\alpha(t) = f^M(0, t) + \frac{\sigma^2}{2\beta^2} (1 - e^{-\beta t})^2
\] (4.21)
Thus, conditional on \( \mathcal{F}_s \), \( r_t \) is normally distributed under \( \mathbb{Q} \) with mean and variance
\[
E[r(t)|\mathcal{F}_s] = r(s)e^{-\beta(t-s)} + \alpha(t) - \alpha(s)e^{-\beta(t-s)}
\] (4.22)
\[
Var[r(t)|\mathcal{F}_s] = \frac{\sigma^2}{2\beta} (1 - e^{-2\beta(t-s)})
\] (4.23)

69
This way we no longer need to estimate $\theta_d$ and $\theta_f$ in equations (4.14), (4.15) and the
domestic and foreign interest rates can be simulated by using equations (4.22) and (4.23)
which can reduce the computational complexity of the model.

### 4.3 Modeling Stochastic Default Intensity

Let $\tau$ be the stopping time that represents the default of the counterparty $Y$ and
$N(t) = 1_{\tau<t}$ be the right continuous increasing process adapted to the filtration $\mathcal{F}(t) = \sigma(N(s), s \leq t)$. Furthermore, $N(t)$ is a sub-martingale and $N(0) = 0$.

\[ E[N(t)|\mathcal{F}(s)] \geq N(s), \quad \forall s < t \tag{4.24} \]

Therefore, by Doob-Meyer decomposition, there exits a unique, increasing, predictable
process $\eta(t)$ with $\eta(0) = 0$ such that $M(t) = N(t) - \eta(t)$ is uniformly integrable martingale.

Assuming that $\eta(t)$ is defined as

\[ \eta(t) = \int_0^t \lambda(s)ds, \tag{4.25} \]

Since $M(t)$ is a martingale, therefore

\[ E[M(t+dt) - M(t)|\mathcal{F}(t)] = 0. \tag{4.26} \]

Thus,

\[ P(t < \tau < t + dt) = E[N(t+dt) - N(t)|\mathcal{F}(t)] = E[\int_t^{t+dt} \lambda(s)|\mathcal{F}(t)] \approx \lambda(t)dt \tag{4.27} \]

$\lambda(t)$ is called default intensity is modeled by Cox-Ingersoll-Ross (CIR) model. The risk
neutral dynamics of CIR model is given as

\[ d\lambda(t) = \kappa_\lambda (\theta_\lambda - \lambda)dt + \sigma_\lambda \sqrt{\lambda(t)}dW^{\mathbb{Q}}_\lambda(t), \quad \kappa_\lambda, \theta_\lambda, \sigma_\lambda > 0 \tag{4.28} \]
where $W^Q(\lambda)$ is a Wiener process under the risk neutral framework. The intensity is strictly positive if Feller condition $2\kappa_\lambda \theta_\lambda > \sigma^2_\lambda$ is satisfied. The relationship between default intensity $\lambda_t$ and survival probability $S(t)$ is given by

$$S(t) = \mathbb{E}(e^{-\int_0^t \lambda(u) \, du}) = \tilde{A}(s,t) e^{-\tilde{B}(s,t)\lambda_t}$$

(4.29)

where,

$$\tilde{A} = \left[ \frac{2h e^{(\kappa_\lambda+h)(t-s)/2}}{2h + (\kappa_\lambda + h)\{e^{h(t-s)} - 1\}} \right]^{2\kappa_\lambda \theta_\lambda / \sigma^2_\lambda}$$

(4.30)

and

$$\tilde{B}(s,t) = \frac{2\{e^{h(t-s)} - 1\}}{2h + (\kappa_\lambda + h)\{e^{h(t-s)} - 1\}}$$

(4.31)

$$h = \sqrt{\kappa^2_\lambda + 2\sigma^2_\lambda}$$

(4.32)

In many practical applications, the curve $S$ is given exogenously from market quotes (e.g., credit default swaps).

### 4.4 Stochastic Hybrid Model

The stochastic hybrid model (SHM), where we consider four processes $\psi$ the FX rate, $r_d$ the domestic interest rate, $r_f$ the foreign interest rate and $\lambda$ the default intensity, under the domestic risk neutral measure $Q$ satisfy the following stochastic differential equations:

$$d\psi(t) = \psi(t)((r_d - r_f)dt + \sigma_{\psi}dW^Q_{\psi}(t)), \quad (4.33)$$

$$dr_d(t) = [\theta_d(t) - \beta_d r_d(t)]dt + \sigma_{r_d}dW^Q_{r_d}(t) \quad (4.34)$$

$$dr_f(t) = [\theta_f(t) - \beta_f r_f(t)]dt + \sigma_{r_f}dW^Q_{r_f}(t) \quad (4.35)$$
\[ d\lambda(t) = \kappa \lambda(t) dt + \sigma \lambda(t) \sqrt{\lambda(t)} dW_\lambda^0(t), \] (4.36)

FX rate \( \psi \) plays an important role in estimating exposure given by equation (4.6) as the foreign payments need to be converted to domestic currency. FX rate is affected by domestic and foreign interest rates, \( r_d \) and \( r_f \), respectively. Therefore, if the temporal association is introduced between FX and default intensity, it’s inherited by \( r_d \) and \( r_f \). Two or more temporal structure between \( r_d \) and \( r_f \) other than the one inherited can be introduced, however, to keep transparency and let the dependence structure of the model readable we do not consider this approach.

Consider the stochastic differential equations (SDE-s) for FX rate (4.33) and default intensity (4.36). The two Brownian motions \( dW_\psi^0(t) \) and \( dW_\lambda^0(t) \) can be correlated by using the theory developed in section 3.1.1 as

\[ dW_\lambda^0(t) = \rho(t) dW_\psi^0(t)) + \sqrt{1 - \rho^2(t)} dW_0^0(t), \] (4.37)

where, the Wiener process \( W_0^0(t) \) is independent of \( W_\psi^0(t) \). Using Lévy theorem ([Øksendal(2000)]) we can prove that \( W_\lambda^0(t) \) is indeed a one dimensional Wiener process.

### 4.4.1 Wiggins’ Stochastic Volatility model, Discretization and Fitting

The underlying Brownian motions (BM)s driving the SDEs of the observed FX rate, default intensity, domestic and foreign interest rates are directly unobservable, latent processes. In order to analyze their association a suitable model is needed, its discretized version has to be fitted to the data and the underlying BMs have to be retrieved in a way that the characteristics of the BMs i.e.independence and stationarity of increments and normal distribution need to be preserved. Simple models like geometric Brownian motion are not suitable for this estimation. A large literature in the field of mathematical finance has acknowledged the temporally and stochastically changing nature of the volatility and
a number of models have been created to incorporate this fact. Among them, the Heston model became one of the widely adopted methodology by industry. However, in our study, when fitted to underlying data (see section 5.4 for the description of the data) — the residuals obtained did not sufficiently satisfy the characteristics of the BM. This phenomenon is well-known for practitioners and experienced by us in the current situation as well. After careful examination of a number of possibilities, in our case Wiggins’ model (Wiggins (1987)) turned out to be creating probably the best residuals in terms of both normality of the distribution and the independence and stationarity of the increments.

Let us consider the following continuous-time model, proposed first by Wiggins (1987), where the spot volatility process is allowed to follow a diffusion process.

\[
\frac{dS(t)}{S(t)} = \mu dt + \sigma(t)dW_i(t) \tag{4.38}
\]

\[
\log(\sigma(t)^2) = Y_i(t) \tag{4.39}
\]

\[
dY_i(t) = \alpha(\theta - Y(t))dt + dV(t) \tag{4.40}
\]

It is known that the Euler-Maruyama discretization of Wiggins’ model turns into Taylor’s stochastic volatility (SV) time series model with proper parametrization. The fitting of the latter via an MCMC algorithm is readily available, together with its full description, in R, in package "stochvol" [Kastner(2016)], and we use that to obtain the parameters and retrieve the residuals. The latter serve as estimations of the price driving Brownian motions, regarded as a quasi sample from its increments, and are the basis for recovering the stochastic correlation for the further analysis described in the next sections.

For more details on Wiggins model see [Wiggins(1987)].

The powerful BDS test (after W.A.Brock, W.Dechert and J. Scheinkman) just like some more simple ones reject dependence in the residuals. Kolmogorov-Smirnov, Anderson-Darling, and Shapiro-Wilk tests reject deviation from the normality of the residuals. It’s well-known fact that Wiener process has Gaussian and independent increments, so the generated (normal and independent) residuals can be regarded as a prediction of the Wiener increments. So the cumulation of these two residuals data series lead to two Wiener processes. The
"local" correlations of the two Wiener processes generated by cumulating the residuals is estimated by a sliding window technique.

### 4.4.2 Interdependence Structure of the Stochastic Hybrid Model

In the applications, the locally (in time) estimated correlation fluctuates around a constant value in the mean reversion sense and the boundaries -1 and +1 of the correlation process are non-attractive and unattainable. Therefore we require these properties from the model as well. The empirical stochastic correlation process $\rho(t)$ can be modeled as a transformation of a diffusion process. Teng et al \cite{Teng, Ehrhardt and Günther(2016)} proposed the hyperbolic tangent of a mean-reverting stochastic process $G_t$, like an Ornstein-Uhlenbeck (OU) process

\[
dG(t) = \theta(\mu - G(t))dt + \sigma dW(t),
\]

\[
G(0) = G_0 \in R
\]

For modeling stochastic correlation we map the values of OU into the [-1, 1] interval by using an hyperbolic tangent transformation $y(.)$ as $\rho(t) = y(G(t))$. In our analysis, we shall use $y(.)$ as the tangent hyperbolic function abbreviated in the literature as "tanh" and its inverse as "atanh". To obtain the parameters of the OU model the estimated stochastic correlation process is back-transformed first and then the usual least squares estimator of the Ornstein-Uhlenbeck process is applied.

We suppose that the FX $\psi$, domestic interest rate $r_d$ and foreign interest rate $r_f$ follow a correlated three dimensional Wiener process. The vector process $W(t)$ consisting of underlying Wiener processes of FX, domestic and foreign interest rates is given as $W(t) = [W_{\psi}^\mathcal{Q}(t), W_d^\mathcal{Q}(t), W_f^\mathcal{Q}(t)]^T$, which are assumed to be correlated by using a constant correlation and given as $\text{corr}(W_{\psi}^\mathcal{Q}(t), W_d^\mathcal{Q}(t)) = \rho_{12}$, $\text{corr}(W_{\psi}^\mathcal{Q}(t), W_f^\mathcal{Q}(t)) = \rho_{13}$ and $\text{corr}(W_d^\mathcal{Q}(t), W_f^\mathcal{Q}(t)) = \rho_{23}$. We denote this correlation matrix by $\mathcal{R}$
\[ R = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} \]

Since the matrix \( R \) is positive semi-definite, it can be decomposed by Cholesky factorization

\[ R = A A^T, \quad (4.43) \]

where \( A \) is a lower triangular matrix with zeros in the upper right corner.

Because the Cholesky matrix is triangular, the factors can be found by successive substitution. For the correlation matrix \( R \), the matrix \( A \) is given as

\[
\begin{bmatrix}
1 & 0 & 0 \\
\rho_{12} & \sqrt{1 - \rho_{12}^2} & 0 \\
\rho_{13} & \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1 - \rho_{12}^2}} & \frac{1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23}}{1 - \rho_{12}^2} \\
\end{bmatrix}
\]

Next, for the given vector process \( W(t) \) we set \( dW(t) = A \, d\zeta \), where, \( \zeta \) is a three-dimensional vector, which is composed of independent variables all with unit variances. Therefore, with this specification complete dependence structure of the SHM model can be given as

\[
dW^\varphi_{\psi}(t) = d\zeta_1 \\
dW^\varphi_{\bar{d}}(t) = \rho_{12} \, dW^\varphi_{\psi}(t) + \sqrt{1 - \rho_{12}^2} \, d\zeta_2 \\
dW^\varphi_f(t) = \rho_{13} \, dW^\varphi_{\psi}(t) + \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1 - \rho_{12}^2}} \, d\zeta_2 + \frac{1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23}}{1 - \rho_{12}^2} \, d\zeta_3 \\
dW^\varphi_{\lambda}(t) = \rho(t) \, dW^\varphi_{\psi}(t) + \sqrt{1 - \rho(t)^2} \, d\zeta_4, \quad (4.47)
\]
where (4.47) follows from (4.37) and \( \zeta_4 \) is independent of \( W^\varphi_{W} \) and \( \rho(t) \) is the stochastic correlation process. By Lévy characterization theorem it can be shown that \( W^\varphi_{W}(t) \), \( W^\varphi_{\varPhi}(t) \) and \( W^\varphi_{\varOmega}(t) \) are indeed Wiener processes along with \( W^\varphi_{\lambda}(t) \).

### 4.5 Data Description

To construct the domestic and foreign yield curves we consider the USD-LIBOR, USD-SWAP and EUR-LIBOR, EUR-SWAP data observed in the market on December 23, 2018. For the implied calibration of HW1F model we use the Swaption price in USD and EUR observed in market on December 23, 2018. Other data used for the calibration purposes are implied volatility of EUR/USD FX option, Credit default swap (CDS) data of the firm Y, USD and EUR-yield curves weakly data and EUR/USD FX rate. A detailed description of the data is given in the data description table given in Appendix A.

### 4.6 Monte-Carlo calculation of CVA

Here in this section, we give the numerical results for CVA calculation based on the methodology developed so far. When there is a dependency between the default intensity of the counterparty and the exposure level, there is no simple way to calculate CVA as no closed-form solution exists. We, therefore, use Monte Carlo simulations for the results. We calculate unilateral CVA from the perspective of firm X where we assume that recovery rate (RR) to be constant at 25\%. The credit rating of the firm Y is BBB- as per Standard & Poor’s rating agency. We assume that firm X pays the domestic leg in Dollar and receives the foreign leg from the firm Y in Euro. For simplicity, we assume that the notional amount is $1 and, furthermore, the domestic leg and foreign leg payments occur at the same dates and with same year fractions. We choose time steps \( t_i (0 \leq i \leq n) \) with \( t_0 = 0 \), \( t_n = T \) and \( t_0 < t_1 < t_2 < ... < t_n \) and set

\[
CVA = (1 - RR) \sum_{i=1}^{n} E^\varphi [DF(t_i) V^+(t_i) \{S(t_{i-1}) - S(t_i)\}]
\]  

\[ (4.48) \]
\[(1 - RR) \sum_{i=1}^{n} E^Q[DF(t_i)V^+(t_i)q(t_i)], \quad (4.49)\]

where, \(q(t_i)\) is the probability of default within \((t_{i-1}, t_i]\). Note that \(q(t_i)\) is the unconditional risk-neutral probability of default between time \(t_{i-1}\) and \(t_i\) (as seen from time 0). It is not the probability of the default conditional on no earlier default. These \(q(t_i)\)'s can be calculated from the credit spread data of the firm Y.

In order to give a clear understanding of the method for calculating CVA under SHM model we present the steps in an algorithm representation.

### 4.6.1 Algorithm Implementation

We divide our algorithm into two parts. The first part consists of methods used for calibrating the interest rate model, stochastic default intensity model and estimating of correlation matrix \(\hat{R}\). The second part includes simulations of the SHM model.

**Step 1 - Calibration of HW1F model**

The first step in the implementation process is the estimation of the parameters of the underlying stochastic interest rate model i.e., HW1F model using market implied price of financial instruments. The common method for calibration of HW1F model is by using Jamshidian’s trick. Using Jamshidian’s trick we can decompose the price of a swaption as a sum of Zero-Coupon bond options. For more details on Jamshidian’s trick please see [Jamshidian(1989)](#). The goal is to minimize the market-implied swaption price and the model-implied swaption price, and therefore, to minimize the objective function

\[
(\text{market implied swaption price} - \text{model implied swaption price})^2 \quad (4.50)
\]

As we have only one market implied price and two parameters including domestic and foreign markets, the mean reversion parameters \(\beta_d, \beta_f\) and the volatility parameters \(\sigma_d, \sigma_f\). To avoid the over-fitting, we fix the mean reversion parameters \(\beta_d, \beta_f\) at 1%. The
estimated values of the volatility parameters of the domestic and foreign markets are (4.554%, 3.525%), respectively.

Step 2 - Calibration of CIR model

The stochastic default intensity model explained in section 4 is calibrated by using the historical CDS data of the firm Y. Let $s_i$ be the credit spread for a maturity of $t_i$, then the default intensity $\lambda_i$ can be expressed as

$$\lambda_i = \frac{s_i}{1 - RR}, \quad (4.51)$$

The survival probability can be calculated by using equation (4.33) and therefore, default probabilities $q(t_i)'s$ follow

$$q(t_i) = S(t_{i-1}) - S(t_i). \quad (4.52)$$

The parameters of CIR model are estimated by well-known maximum likelihood estimation. The estimated values of the parameters $(\kappa, \theta, \sigma)$ are (0.2975, 0.3045, 0.1432), respectively.

Step 3 - Estimating correlation matrix $R$

Next, to fully specify the interdependence structure of the SHM model we need to estimate the correlation matrix $R$. We estimate $\rho_{12}, \rho_{13}$ and $\rho_{23}$ from the historical domestic and foreign interest rates estimated from the historical yield data and the historical Eur/USD FX rate data. The residuals are estimated by fitting Wiggins model and tested for normality and independence. As the increments of Wiener process are independent and normal, therefore, the estimated residuals can be regarded as increments of the Wiener processes. The estimated correlations for $(\rho_{12}, \rho_{13}, \rho_{23})$ are (0.1162656, 0.01965914, 0.1383345), respectively.
respectively. With the estimated correlations the correlation matrix $\mathbf{R}$ can be written as

$$
\begin{bmatrix}
1 & 0.1162656 & 0.01965914 \\
0.1162656 & 1 & 0.1383345 \\
0.01965914 & 0.1383345 & 1
\end{bmatrix}
$$

Now we ready for the simulation of the SHM model.

**Step 4 - Simulating SHM model and estimating exposure**

We simulate 10000 paths of the stochastic hybrid model by substituting the correlated Wiener processes from equations (4.44), (4.45), (4.46) and (4.47) into the equations (4.33), (4.34), (4.35) and (4.36), respectively. Domestic and foreign interest rates are simulated by the method of exact simulation using equations (4.22) and (4.23). For that cause we need to calculate $\alpha(t)$ given by equation (4.21). The instantaneous forward rate $f^M(0,t)$ is estimated by using equation (4.18). We first construct the domestic and foreign markets’ yield curves by using Nelsen-Siegel model. Short ends of both the yield curves are constructed by USD and EUR LIBOR rates while the longer ends are constructed by USD and EUR swap rates. For more details on the construction of yield curve using Nelsen-Siegel model we refer reader to [Nelson-Siegel(1987)]. The market implied discount factors $P^M(t,T)$ can be estimated from the constructed yield curves as follows

$$
P^M(t,T) = \exp(-R^M(t,T) \tau(t,T)),
$$

(4.53)

where, $R^M(t,T)$ is the continuously-compounded spot interest rate prevailing at time $t$ for the maturity $T$ in the market and $\tau(t,T)$ is the time difference. Once we have the simulated $r_d$ and $r_f$, zero coupon bond prices can be simulated by equations (4.16) and (4.17). The simulated $r_d$ and $r_f$ are used to simulate the FX spot rate $\psi(t)$, given by
Proposition (2). The survival and consequently default probabilities are calculated from the simulated default intensity as given by equation (4.29) and (4.52), respectively. As we have assumed that the maturity of cross currency swap is 10 years and both the domestic and foreign leg payments occur at the same dates quarterly, therefore the dimension of exposure matrix is $40 \times 10000$. Consequently, default probability, domestic and foreign interest rates and FX spot rates matrices also have the same dimensions. The net discounted exposure of the CCS as seen from the perspective of the firm X is shown in Figure 4.1.

Figure 4.1: Net discounted exposure of cross-currency swap for firm X at time 0.

**Step 5 - Calculation of CVA**

We calculate CVA using the SHM model under wrong-way risk path-wise by using equation (4.49). For each simulation $j$, for $j \in \{1, 2, ..., 10000\}$ we estimate CVA path-wise using the discretized equation (4.49). This gives us a CVA value for every simulated path. By averaging out all 10000 pathwise CVA values we get the estimated CVA to be charged from the counterparty. To draw a comparison between the CVA value calculated by using
SHM model and the usual market practiced methods, we assume a constant correlation between default intensity and FX spot rate too, keeping the correlation matrix \( \mathcal{R} \) same. The first column of the Table 4.1 corresponds to the correlation type used in SHM, keeping the correlation matrix \( \mathcal{R} \) fixed. Therefore, only the correlation between FX rate and default intensity is varied to calculate CVA with wrong-way risk. In the first case, this correlation is modeled as a stochastic process as explained in section 3.1.1. In the other five cases, this correlation is kept constant and decreased linearly starting from a perfect correlation of 1 as shown in first column of Table 4.1. Mean, standard deviation and 95% quantile is then calculated for the 10000 estimated CVA values by generating 10000 paths of the SHM under all six models. The mean of these values corresponds the estimated CVA to be charged from the counterparty Y while entering into CCS at time 0. The steps to be followed to implement this algorithm is given in Figure 4.2. The estimated CVA values under wrong-way risk using SHM and constant correlations are given in Table 4.1.

Table 4.1: Statistics table for 10000 CVA calculated pathwise

<table>
<thead>
<tr>
<th>Statistics type</th>
<th>CVA</th>
<th>Standard deviation</th>
<th>95% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastic hybrid model using ( \rho = 1 )</td>
<td>0.0358243</td>
<td>0.04002753</td>
<td>0.1199763</td>
</tr>
<tr>
<td>( \rho = 0.75 )</td>
<td>0.01858226</td>
<td>0.02209931</td>
<td>0.0612432</td>
</tr>
<tr>
<td>( \rho = 0.50 )</td>
<td>0.01790981</td>
<td>0.02042456</td>
<td>0.0581961</td>
</tr>
<tr>
<td>( \rho = 0.25 )</td>
<td>0.01699416</td>
<td>0.01905533</td>
<td>0.0558105</td>
</tr>
<tr>
<td>( \rho = 0 )</td>
<td>0.01662678</td>
<td>0.01821355</td>
<td>0.0525241</td>
</tr>
</tbody>
</table>

The first column shows the correlation type used between FX spot rate and default intensity in SHM, keeping the correlation matrix \( \mathcal{R} \) fixed. The second column of the
Step 1: Simulation of domestic, foreign interest rates and FX rates
- Estimate the parameters of the interest rate model HW1F using market implied Swaption prices
- Simulate the interest rate paths and zero-coupon bond prices by using closed form expression of HW1F
- Using change of measure technique simulate FX rate using Garman-Kohlhagen model

Step 2: Simulation of default intensities
- Estimate the default intensity vector by using the historical CDS market spread data
- Simulate the default intensities by estimating the parameters of CIR model by using MLE

Step 3: Estimate the stochastic correlation and decompose the correlation matrix
- Estimate the stochastic correlation process by using sliding window technique
- Model stochastic correlation as a proper transformation of a diffusion process
- Decompose the correlation matrix by using Cholesky decomposition

Step 4: Monte Carlo calculation of CVA
- Simulate 10000 paths of SHM model under stochastic correlation and various values of constant correlation from the interval [0,1]
- Calculate CVA path-wise using equation (49)
- Average the 10000 expected values obtained to get the actual value of CVA

Figure 4.2: Summary of the steps for the implementation of the model.

table contains the CVA values estimated by using Stochastic hybrid model under different correlation types, starting from stochastic correlation. It can be seen that the CVA increases linearly as the constant correlation between FX spot rate and default intensity increases from 0 to 1 as expected. However, this increase in CVA values is not substantial. CVA values estimated from a 0 correlation to a perfect correlation between FX spot rate and default intensity increased by around 12.66%. At the same time when CVA is estimated by using stochastic correlation in SHM is 0.0358243. The CVA estimated by using stochastic correlation in SHM almost doubles when compared to the perfect constant correlation case. To understand this impact first we need to look into the reasons why the increase in CVA values under the constant correlation model is not significant. We know that correlation captures only the linear dependencies between two random variables.
What if this interdependence is non-linear? More often than not, this interdependence when observed in the market is usually time-dependent and fluctuates randomly. So, while modeling wrong way risk using constant correlation it is assumed that all the dependencies within the model are linear and this may lead to so-called correlation risk. The fourth column of Table 4.1 contains the 95% quantile values for 10000 CVA values. It is evident from these values that even if the average of the 10000 simulated CVA values is 0.0358243, it can be even higher in some cases and reach up to 0.1199763 at 95% confidence level. However, the same value for the SHM with constant correlation is quite low when compared to SHM with stochastic correlation. This shows that using stochastic correlation the model allows more variability when compared to a constant correlation model. Another important factor why we see a higher CVA in the case of stochastic correlation framework is because of tail dependence. In general, tail dependence measures the association between the extreme values of two random variables. As it is well-known fact that the constant correlation model does not take into account tail dependence i.e., the tail dependence under the constant correlation model is zero. This is evident from the fact that Gaussian copula has tail dependence zero until the underlying correlation is perfect. Gaussian copula based credit risk models like that of [Li(1999)] have faced harsh criticism in the past, on the ground of having zero i.e. no tail dependence. Using the model with stochastic correlation we move further away from Gaussian copula and thus can be able to capture this tail risk because tail dependencies should be considered when assessing the diversification and risk of any portfolio.

4.7 Concluding remarks

In this chapter, we revisited the problem of pricing CVA with wrong-way risk for a cross-currency swap. Even after a decade of the financial crisis, addressing wrong-way risk still remains challenging. More often than not, a constant correlation model has been used to model wrong-way risk in one way or the other which can lead to underestimation
of WWR and thus the calculated value of CVA can be wrong leading to catastrophic situations like 2007-08. Here, we propose a new approach, a stochastic correlation modeling framework to calculate CVA with wrong-way risk. This helps us to move away from the Gaussian copula models, thus, enables our model to capture the tail risk. It is well-known fact that constant correlation models overlook the tail dependence i.e., the constant correlation models have zero tail dependence until and unless a perfect positive correlation is used which may not be practically true in the market. Our results show that the impact of stochastic correlation in modeling the wrong-way risk is significant because of the rich variability and heavier tails that our model produces when compared to the constant correlation model. This way the stochastic correlation model may be able to capture the extreme events which are not possible to model by using constant correlation models. Tail risk should not be ignored when pricing any financial asset or a derivative. Thus, for a more conservative and risk-averse investor, a stochastic correlation modeling approach can help her/him to hedge against the credit risk.
Chapter 5

Rough Correlation for Characterizing Herd Behaviour of Traders in Stock Markets

We denote the notion of rough stochastic correlation for Brownian motions as the integrand in the formula of their quadratic covariation. By modeling minute-wise traded Apple and Microsoft stock prices the estimations of the latent price driving Brownian motions and their temporally localized correlation as a stochastic process becomes available. Analyzing this stochastic correlation we give statistical evidence for the roughness of its paths. Moment scaling indicates fractal behavior and the fractal dimensions and Hurst exponent estimates point to rough paths. Modeling the rough correlation by suitably transforming a fractional Ornstein-Uhlenbeck process enables extensive simulations of stock prices and assessing model goodness of fit, compute tail dependence and the Herding Behavior Index (HIX). The HIX is found to be hardly variable in time; to the contrary tail dependence fluctuates heavily. As a result, a sudden coincident appearance of extreme log-return values are more likely than it is indicated by a steady HIX value.
5.1 Brownian Motions with Rough Stochastic Correlation

Stochastic differential equations are used frequently to model data series such as interest rate, asset price, exchange rate and so on. For diffusion processes described by SDEs, the dependence between the solutions is most often originates in correlated Brownian motions driving the equations. Suppose we are given a finite time horizon \([0, T]\) and a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with a filtration, satisfying the usual conditions. This is e.g. the case when the filtration is the augmented filtration of a (multidimensional) Brownian motion. Suppose, we are given two independent Brownian motions, \(W_0(t)\) and \(W_1(t)\) adapted to the filtration \(\mathcal{F}_t\) (i.e. the filtration is rich enough). Considering two adapted correlated Brownian motions \(W_1(t)\) and \(W_2(t)\) with \(\text{corr}(W_1(t), W_2(t)) = \rho\), their quadratic covariation,

\[
[W_1, W_2](t) = \int_0^t \rho \, ds = \rho \cdot t
\]  

(5.1)
equals the covariance and is symbolically written as \(dW_1(t)dW_2(t) = \rho \, dt\). Instead of a constant \(-1 \leq \rho \leq 1\), let us given an adapted stochastic process \(\rho(t)\) with \(|\rho(t)| \leq 1\).

Since \(\rho(t)\) is bounded and the time horizon is finite \(\rho(t) \in \mathcal{L}_2(\Omega \times [0, T])\). By using \(\rho(t)\) we define the vector process \(W(t) = \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix}\) by \(W(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\) and the differential

\[
dW(t) = d\begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho(t), \sqrt{1 - \rho^2(t)} \end{pmatrix} \begin{pmatrix} W_1(t) \\ W_0(t) \end{pmatrix}.
\]  

(5.2)

Here the second coordinate \(W_2(t)\) is clearly adapted to \(\mathcal{F}_t\), and as a sum of two stochastic integrals w.r.t. Brownian motions of \(\mathcal{L}_2(\Omega \times [0, T])\) processes it is a continuous martingale. As a direct consequence of Itô’s formula, the quadratic (co)variation of \(W(t)\) is

\[
[W, W](t) = \int_0^t \begin{bmatrix} 1, \rho(s) \\ \rho(s), 1 \end{bmatrix} \, ds,
\]  

(5.3)

and in particular, the quadratic variation of \(W_2(t)\) is \(t\). Being an adapted continuous martingale with quadratic variation \(t\), by virtue of Lévy’s theorem, (cf. e.g. theorem 8.1.6. in [Øksendal(2000)]) \(W_2(t)\) is itself a Brownian motion. However, the two dimensional
process \((W_1(t), W_0(t))\) is not a two dimensional Brownian motion, what is more, the two dimensional distribution of the vector is not Gaussian despite the normality of the marginals. Further, the quadratic covariation of \(W_1(t)\) and \(W_2(t)\) is

\[
[W_1, W_2](t) = \int_0^t \rho(s) \, ds.
\]  

(5.4)

To the analogy of the constant \(\rho\) in (5.1), the process \(\rho(t)\) is called the \textit{stochastic correlation} of \(W_1(t)\) and \(W_2(t)\). When in addition \(\rho(t)\) has rough paths we call it in short the \textit{rough correlation}. Note here that no semimartingale property was required of \(\rho(t)\) for what has been said above. Let us also remark that stochastic correlation is only defined here for Brownian motions. An analog of Goldstein’s theorem for the quadratic variation of SDEs seems to be easy to obtain, and that provides a way to extend the definition at least to diffusion processes. It is, however, out of the scope of this thesis since we intend to avoid technical discussions, that are not absolutely necessary.

### 5.2 Wiggins’ Stochastic Volatility model, Discretization and Fitting

The generating Brownian motions driving the SDEs of observed asset prices are directly unobservable, latent processes. Therefore, in order to analyze their association a suitable model is needed, its discretized version has to be fitted to the price data and estimated Brownian motions have to be retrieved. A reliable analysis requires that these estimations feature discretized Brownian motion characteristics: independence and stationarity of increments and normal distribution. Simple models like geometric Brownian motion do not live up to this expectation. The temporally and stochastically changing nature of volatility has long been noted in the literature of mathematical finance and a great number of models has been built to incorporate it to price description. Among them the Heston model became one of the industry standards in our days. However, when fitted to the minute-wise registered trading data – the subject of the present chapter – the residuals
obtained from this model did not sufficiently feature the mentioned Brownian motion characteristics. This phenomenon is well-known for practitioners and experienced by us in the current situation as well. After careful examination of a number of possibilities, in our case Wiggins’ model ([Wiggins(1987)]) turned out to be creating probably the best residuals in terms of both normality of the distribution and the independence and stationarity of the increments. Let us consider the following continuous time model, proposed first by Wiggins (1987), where the spot volatility process is allowed to follow a diffusion process.

\[
\frac{dS(t)}{S(t)} = \mu dt + \sigma(t)dW(t)
\]

(5.5)

\[
\log(\sigma(t)^2) = Y(t)
\]

(5.6)

\[
dY(t) = \alpha(\theta - Y(t))dt + dV(t)
\]

(5.7)

It is known that the Euler-Maruyama discretization of Wiggins’ model turns into Taylor’s stochastic volatility (SV) time series model with proper parametrization. The fitting of the latter via an MCMC algorithm is readily available, together with its full description, in R, in package ”stochvol” [Kastner(2016)], and we use that to obtain the parameters and retrieve the residuals. The latter serve as estimations of the price driving Brownian motions, regarded as a quasi sample from its increments, and are the basis for recovering the stochastic correlation for further analysis described in the next sections.

### 5.3 Modeling Rough Correlation

In the applications, the locally (in time) estimated correlation fluctuates around a constant value in the mean reversion sense and the boundaries -1 and +1 of the correlation process are non-attractive and unattainable. Therefore we require these properties from the model as well. Evidence for the roughness of paths, when found, influence the model choice for stochastic correlation. One natural candidate would be the fractional Jacobi process. [Dung(2014)] proved the existence and uniqueness theorems for fBm generated Jacobi process, however, the properties and the estimation theory of that process are not yet
elaborated, therefore, we omit its use in the present chapter. Instead, we shall consider the fOU process, and transform it according to the purpose of modeling rough correlation with the above-mentioned properties. Having this in mind, the dynamics of the mean-reverting fOU process is given by a Langevin-type stochastic differential equation (SDE) driven by a fBm \( W^H = \{W^H(t), t \geq 0\} \) of Hurst parameter \( H \in (0, 1) \), and both the parameter \( \lambda \) and the volatility \( \sigma \) are positive real constants.

Assuming the Hurst exponent to be known, the drift and the volatility of the fOU process can be estimated as in [Hu et al. (2017)]. According to formula (5.1) of that paper the least-squares estimation of the drift is \( -\log \left( \frac{\int_0^T X(t) dX(t)}{\int_0^T X^2(t) dt} \right) \) and we use the approximating Riemann-Stieltjes sum to compute this term. As for \( \sigma \) we use the specific version, given for the constant case in Proposition 4.2. of [Hu et al. (2017)] in terms of \( p \)-variations, and we choose \( p = 2 \). With that, the estimate of \( \sigma \) reduces to \( n^H \) times the residual standard deviation.

A crucial point and novelty of our analysis is to present statistical evidence that the correlation process is really rough and to estimate its most important characteristic, the Hurst exponent. When introducing rough processes into modeling stochastic volatilities [Gatheral et al. (2018)] rely on moment scaling, more precisely its linearity on the log-log scale, widely used primarily in the physics literature. Exposing theory to the test on real data [Fukasawa et al. (2019)] confirm statistically the roughness of stochastic volatility using 5-minute realized volatility data of stock indices. Beyond moment scaling they estimate the Hurst coefficients by using a specific Whittle type estimator that they show to be more accurate than the one obtained by moment scaling; the latter is either heavily biased or inaccurate in their case and - as we shall see later in this chapter - does not match to other estimators in our analysis as well.

A consequence of linear moment scaling – showed in both mentioned papers and also in the application part of the present one – is the fractal character of the process, more precisely: of its paths. A measurement of roughness for fractals is their fractal dimension, a notion dating back to [Hausdorff (1919)]. Several definitions and estimation procedures are known for fractal dimensions, a survey of them is given in [Gneiting et al. (2012)].
Hausdorff dimension is a theoretical quantity, it is intractable empirically. Boxcount, the traditionally most preferred computable counterpart, is notorious for its very slow convergence and hence unreliable. The authors in [Gneiting et al.(2012)] find the use of the madogram more suitable in a wide range of situations. We shall also rely on the madogram. Although the 'fractaldim' package of R readily gives the value of the other estimators, they – like the variogram, rodogram, ”incr“ and Genton’s robust estimator – either give the same results or – like the periodogram based estimators – produce meaningless values in the application. It is known ([Adler(1981)] and [Gneiting et al.(2012)]) that the path of a fBm as well as the fOU process has fractal dimension $2 - H$. So, it naturally lends itself to consider the fractal dimension of stochastic correlation, the measurement of roughness of the correlation paths, at the same time as an estimator of the Hurst exponent. This is done in the present work, and it turns out that the Hurst parameter obtained from the fractal dimension is in good agreement with some more conventional estimators, like the aggregated variance, R/S, and Peng’s method (or detrended fluctuation analysis). To that fit however, the log-log scale regressions have to be suitably truncated and that makes the general use of these latter, traditionally applied methods at least more circumstantial, if not dubious.

Next, a transformed fOU model is fitted to stochastic correlation. A crucial fitting criterion is that the Hurst exponent of the driving fBm in [2,30] has to be chosen so that the generated transformed fOU process fits to the stochastic correlation in terms of both fractal dimension and marginal distribution. Fractal dimension is a characteristic of a continuous sample, whereas we only can have discrete time observations. So, although theoretically, the fractal dimensions and thus the Hurst exponents of the fOU process and its driving fBm are the same, (see Chapter 8 in [Adler(1981)] and also [Gneiting and Schlather(2001)]) due to the continuous nature of the notion, it is only so in the limit if sampling frequency goes to infinity. While in the limit the local or short time smoothing effect of the drift parameter is negligible, it is not so for the finite sample. The effect of smoothing is analytically intractable, so we rely on simulations to take it into account. Besides the fractal dimension, the match of the empirical and the simulated distribution in terms of
mean, standard deviation, skewness, and estimated density is also a criterion for the fit.

For modeling stochastic correlation we map the fOU values into the [-1, 1] interval by using an appropriate transformation with a smooth, bounded, monotonous real function \( g(x) \) as \( \rho(t) = g(X(t)) \), preserving this way the rough path property. In our analysis we shall use for \( g(x) \) the tangent hyperbolic function abbreviated in the text as "tanh" and its inverse as "atanh". Using the tangent or the normal cdf for outer function in the transformation worsen either the fit of the fractal dimension or the distribution. A fully non-parametric estimation of the transforming function may be desirable, but we do not intend to develop it in this chapter.

5.4 Trading data of two stocks

In the present study we analyze the association of high-frequency closing prices of two stocks, Apple (AAPL) and Microsoft (MSFT). The data are registered between July 19, 2016 - August 03, 2016, and consists of 5040 minute-wise closing prices. The closing price of AAPL and MSFT with their log-returns are given in figure 5.1, respectively. Just as in [Fukasawa et al.(2019)] and in many other papers we model the price dynamics in the business time scale, meaning that the time in the model evolves only when a market of the asset is open.

Euler-Maruyama discretization of Wiggins’ stochastic volatility model (i.e. Taylor’s SV time series model) is fitted to the AAPL and MSFT log-return data by using the ‘svsample’ command of the ‘stochvol’ package in R. As a result of the MCMC algorithm the estimated model parameters are obtained as sample averages from the posterior distribution. The standardized residuals can also be extracted by in-built ‘resid’ method. The powerful BDS test ([Brock et al.(1996)]) just like the more simple difference-sign or turning point etc. tests reject dependence in the residuals. Kolmogorov-Smirnov, Anderson-Darling, and Shapiro-Wilk tests reject deviation from the normality of the residuals. Hence the obtained (normal and independent) residuals can be regarded as estimations of the increments of the driving Brownian motions. The volatilities can also be subtracted - the volatility plot is provided readily by the package, to get the object itself needs some further coding.
Figure 5.1: Minutewise closing prices and log-returns of stocks AAPL and MSFT in a two week period

in R. The volatilities are very highly correlated, by 0.965, that probably is specific for the time period, when the samples were taken. This high correlation is the reason why we use stochastic correlation modeling only for the price residuals. When simulating the synthetic stock prices we create volatilities according to this constant correlation.

The normal plots of the residuals for the two stocks, AAPL and MSFT are shown in figures 5.1. Supposing that the association of the two Brownian motions, driving the price equations of the two stocks, originates from stochastic correlation with stationary correlation process (constant correlation included as the extreme case), the pairs of residuals sample the same two dimensional distribution, opening the way to estimate the copula of the latter. If the two Brownian motions were simply Pearson correlated, the increments would create a Gaussian copula. But the plot of the copula of residuals and of normal correlated variables (figure 5.2) with identical Pearson correlation show clear deviation.

We now create estimations for the two price driving Brownian motions by cumulating the residuals and estimate their "local" correlations in non-overlapping windows of size 5 minutes. We call this estimation the (stochastic/rough) correlation of price residuals for short and denote it by \( \hat{\rho}(t_i) \), where \( t_i \) grows by 5 minutes. Allowing for overlapping windows would work as a smoother, inflating fractal dimension, and hence producing
misleading results. This way we receive an estimation for the stochastic correlation process of the price generating Brownian motions.

5.5 Evidences for roughness of the stochastic correlation

The next step is to provide statistical evidence to the presence of rough correlation. Similarly to the correlation of price residuals we can compute temporally localized correlations of the observed log-returns as well, and consider the inverse tangent hyperbolic of both. Although in the sense of the definition only the correlation of price residuals is an estimation of a stochastic correlation, the fractal dimension of all four processes can be computed. The madogram values ($\hat{M}$) are given in table 5.1 together with $2 - \hat{M}$ that - under some conditions - would be Hurst exponent estimate. As is observable from the table, the

<table>
<thead>
<tr>
<th>Empirical quantities</th>
<th>Estimated 5 min</th>
<th>atanh-transf.</th>
<th>Estimated 5 min</th>
<th>atanh-transf.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log-returns</td>
<td></td>
<td>residuals</td>
<td>residuals</td>
</tr>
<tr>
<td>Madogram ($\hat{M}$)</td>
<td>1.9560</td>
<td>1.9514</td>
<td>1.9568</td>
<td>1.9520</td>
</tr>
<tr>
<td>$H$ estimate ($2 - \hat{M}$)</td>
<td>0.0440</td>
<td>0.0486</td>
<td>0.0432</td>
<td>0.0480</td>
</tr>
</tbody>
</table>

Table 5.1: Madogram fractal dimension and Hurst exponent estimation of correlations and their transformations
Wiggins model does not change the fractal character of the estimated correlations, and the inverse tangent hyperbolic transformation causes only minor change in the madogram values. Other fractal dimension estimators could have also been given from the ‘fractaldim’ package of R, but most of them give similar results, so presenting them would only hamper transparency. The boxcount values are significantly different but that is because of the slow convergence, while frequency domain based estimators indifferently give a meaningless trivial value, namely 2.

The subject of all further analysis is the correlation of price residuals that is the estimated stochastic correlation process \( \hat{\rho}(t_i) \) of the Brownian motions and its back-transform by the atanh function \( \hat{X}(t_i) \). Just as the price residuals are regarded as a quasi-sample from the Brownian motions, the back-transformed correlation of price residuals can be thought of as a quasi sample from the fOU process mentioned in section 5.3 taken in the discrete time set \( t_i, i = 1, \ldots, 1008 \). With this quasi sample the moment scaling can be studied as is done in the case of the volatility process in [Gatheral et al.(2018)] and also in [Fukasawa et al.(2019)]. While computing the moments we change the level of aggregation in 2-power magnitude. This choice – although differs from [Gatheral et al.(2018)] and [Fukasawa et al.(2019)] – is quite customary in fractal literature (see e.g. [Tessier et al.(1993)], figures 11-13.), and results in more transparency. On the log-log scale the moments grow linearly with aggregation. This can be seen as statistical evidence for the fractal property. The coefficients of the log-log scale regression sit nicely on a line of slope \( \hat{H} = 0.0216 \) – as displayed in Figure 5.3 – and give another Hurst exponent estimate. The Hurst exponent estimates obtained from fractal dimension and moment scaling do not match, but the unreliability of moment scaling estimate was already noted in [Fukasawa et al.(2019)].

In order to obtain further inference on the Hurst coefficient of the back-transformed correlation of price residuals the ‘fArma’ package given in R is utilized and the R/S, aggregated variance, Peng, Higuchi, and Whittle estimators are computed, as mentioned in section 5.3. Their values are given in table 5.2.

The values by the aggregated variance method and the Whittle estimator match best
Table 5.2: Hurst exponent estimations of the back-transformed correlation of price residuals

<table>
<thead>
<tr>
<th>Hurst exp. estimations</th>
<th>R/S method</th>
<th>aggr. variance</th>
<th>Higuchi est.</th>
<th>Peng’s method</th>
<th>Whittle est.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{H} )</td>
<td>0.0537</td>
<td>0.0447</td>
<td>-0.001</td>
<td>0.0381</td>
<td>0.0466</td>
</tr>
</tbody>
</table>

to the one corresponding to fractal dimension, and the R/S and Peng’s methods produce only slightly different values; only Higuchi’s method fails to give meaningful result. Note here that the lack of distribution theory for those estimators does not permit considering the significance of the differences.

With the evidence that the atanh-transformed process of stochastic correlation of price residuals has fractal paths, its estimated fractal dimension is as high as nearly 1.95, and consistently with it various Hurst exponent estimations produce as low values as roughly 0.05 we can safely assert that this process has rough paths.

5.6 Simulation of synthetic stock prices

Next, we model the rough process of inverse tangent hyperbolic transform of the correlation of price residuals as a fOU process. The goodness of fit (GoF) of the fOU process is measured by the fit of the madogram fractal dimension and the marginal distribution of the process. Since no theory is available for the properties of fractal dimension estimation and we are not aimed at developing one either, we rely on simulation results only. Supposing
the Hurst exponent $H_B$ of the fOU driving fBm to be known, the estimation of the drift and volatility parameters is straightforward as is given in section 5.3. So, we need $H_B$.

We look for an approximate value $\hat{H}_B$ by a simple grid search, in 0.005-wise steps. In each step 1000 simulations are performed with the actual $\hat{H}_B$ value and the corresponding parameters creating 1000 tanh-transformed fOU processes. Their madogram fractal dimension, together with the mean, standard deviation and skewness of the marginal distribution are computed for assessing GoF and compared to the corresponding values obtained from $\hat{\rho}(t_i)$. On the basis of their fit, in the grid search $\hat{H}_B = 0.105$ can be selected as the estimation of the Hurst exponent. This in turn determines the values of the fitted fOU parameters exhibited in the first two rows of table 5.3. In the further rows the GoF statistics values and their standard deviations are presented for tanh-transformed fOU processes of lengths increasing roughly 10 times by rows, simulated with these same parameters.

The simulated tanh-transformed fOU process is of short memory, because $\hat{H}_B = 0.105 < 0.5$ so, the asymptotic normality of the mean, standard deviation and skewness holds. The madogram is a variogram type estimator, with absolute differences instead of the squares in sum, so for its values again asymptotic normality holds true under broad conditions, but it doesn’t hold for the fractal dimension computed from it. Remember, the latter has values between 1 and 2. The sixth row in table 5.3 displays the standard deviations of these GoF statistics used now as of simulated tanh-transformed fOU processes i.e. simulated rough correlations of the same length as the quasi sample. So, these may be regarded as an estimation of the standard deviations of the GoF statistics of the correlation of price residuals $\hat{\rho}(t_i)$ - from that latent process so far we only had one quasi observation. Together with the asymptotic normality, this means that at least the first three statistics presented in the fifth row are well within the 95% confidence bound, so the given Hurst exponent estimate can be accepted on that basis. Even if we increase the length of the simulated tanh-transformed fOU processes 10 and 100 times i.e. simulate that much longer correlations of price residuals, the GoF statistics and in particular the madogram fractal dimension shows remarkable stability. Remark here that the boxcount dimensions
Parameter estimation | $\hat{H}_B$ | $\hat{\lambda}$ | $\hat{\theta}$ | $\hat{\sigma}$
--- | --- | --- | --- | ---
Fitted values | 0.115 | 0.8907 | 0.7891 | 1.7322

<table>
<thead>
<tr>
<th>Process characteristics</th>
<th>mean</th>
<th>std. dev.</th>
<th>skewness</th>
<th>madogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}(t_i)$ correlation of price residuals</td>
<td>0.5305</td>
<td>0.4067</td>
<td>-1.2787</td>
<td>0.0480</td>
</tr>
<tr>
<td>Simulation; process length 1008</td>
<td>mean</td>
<td>0.5301</td>
<td>0.4045</td>
<td>-1.1517</td>
</tr>
<tr>
<td></td>
<td>std. dev.</td>
<td>0.0048</td>
<td>0.0089</td>
<td>0.0522</td>
</tr>
<tr>
<td>Simulation; process length 10000</td>
<td>mean</td>
<td>0.5330</td>
<td>0.4002</td>
<td>-1.1556</td>
</tr>
<tr>
<td></td>
<td>std. dev.</td>
<td>0.0014</td>
<td>0.0028</td>
<td>0.0149</td>
</tr>
<tr>
<td>Simulation; process length 100000</td>
<td>mean</td>
<td>0.5342</td>
<td>0.4002</td>
<td>-1.1574</td>
</tr>
<tr>
<td></td>
<td>std. dev.</td>
<td>0.0004</td>
<td>0.0009</td>
<td>0.0058</td>
</tr>
</tbody>
</table>

Table 5.3: Fitted parameters of the fOU model; Descriptive statistics and fractal dimension of the rough correlation of observed stock log-returns and of simulated ones

(not presented in the table) change significantly with length as evidence of its very slow convergence. The distribution of the simulated processes matches also well to that of $\hat{\rho}(t_i)$. It is illustrated in figure 5.4 where the quantile-quantile plot of $\hat{\rho}(t_i)$ and its simulated counterpart is displayed together with the estimated density function of $\hat{\rho}(t_i)$ and the simulated processes. Note here that the latter estimation is made by taking all simulated processes as one sample which is 100 times more data, hence a much smoother and more accurate function could be obtained. Further, a beta density function fitted by MLE as implemented in the fitdistr function in R is also displayed. Its imperfect fit implies that the distribution is not beta, an argument against using a Jacobi process in modeling. The good fit is remarkable, the higher peak of the density of simulated data may in part be resulted from the higher number of data.
5.7 Estimating the Herding Behavior Index and Tail Dependence for the Stocks

Consider two stocks with price or log-return processes \( \{X_i(t), 0 \leq t \leq T; i = 1, 2\} \). Suppose that their dynamics are described by SDEs. Consider the prices at a certain time instant \( t \) and for a more transparent formula use temporarily the notations \( X_1 = X_1(t) \) and \( X_2 = X_2(t) \). The HIX index defined in [Dhaene et al.(2012)] simplifies for this case to:

\[
HIX = \frac{\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2 \cdot corr[X_1,X_2] \cdot \sigma_{X_1} \cdot \sigma_{X_2}}{\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2 \cdot corr[F_{X_1}^{-1}(U), F_{X_2}^{-1}(U)] \cdot \sigma_{X_1} \cdot \sigma_{X_2}},
\]

(5.8)

where \( \sigma_{X_1}, \sigma_{X_2} \) are the variances, \( F_{X_1}(\cdot), F_{X_2}(\cdot) \) the distribution functions, and \( U \) is a uniform random variable in \([0,1]\). Note that with these choices \( corr(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U)) \) is the maximal correlation between \( X_1 \) and \( X_2 \) (see again [Dhaene et al.(2012)]).

We use the ‘somebm’ package of R to generate 10000 fractional Brownian motions of length 5040 – the same as the original data – and with them hyperbolic tangent transformed fOU processes for rough correlations. Using these latter we create 10000 rough-correlated Brownian motion pairs of the same length. These serve as driving forces in Wiggins’ models with the estimated parameters from the fit to AAPL and MSFT prices, and by that we simulate 10000 pairs of paths of similarly associated synthetic stock price log-returns. This way we can create at every time instant an empirical copula of log-returns. A single, non-time-dependent copula is also created from the original observations, when
observations in time are supposed to sample the same two dimensional distribution, i.e. based on the principle of ergodicity, that is taken for granted. In Figure 5.5 we compare this latter copula with the simulated one in an arbitrarily chosen time point – we display in the figure the copula of midday log-returns in the 14th trading day, but in any time point it looks almost absolutely the same – and with a Gaussian counterpart of the same Pearson correlation. For the sake of better comparison we only display 5040 points in all three plots – no more is available from the original observations.

A straightforward visual inspection concludes that the obtained copulas are different. Further, the first two plots indicate much stronger tail dependence than the third one, so both observations and simulations indicate strongly the presence of (non-zero) tail dependence. From the simulated paths in every time point we have 10000 samples from the copula of the a log-returns, hence we can calculate the tail dependence. Having 5040 time points, we could compute the same amount of tail dependence values, but it would be a tedious and time and computer memory consuming work with doubtful advantages. Given the stationary character of our model we do the calculation only for the last 500 trading times. We summarize the descriptive statistics of these tail dependence values in Table 5.4. To control for the effect of the rough correlation on tail dependence, we also simulate log-returns with constant Pearson correlations of the price driving Brownian motions. A source of interdependence in this case can be the very highly correlated (almost common) Brownian motions driving the stochastic volatility process. So, we
simulate both with such volatility process, and with independent volatility processes. The results are given in the last two rows of Table 5.4.

The model with rough correlation creates tail dependence with astounding match to the one computed from the observed log-returns by the principle of ergodicity, i.e. using temporal observations for estimating space characteristics. Stochastic volatilities themselves – although very highly correlated, almost being identical – are not capable of reaching that level of tail dependence. When correlation is constant in price driving and volatilities are independent the created tail dependence is similar to that of the finite sample from a Gaussian copula. The values present a rather convincing evidence for rough correlation being an essential component in creating tail dependence, equally strong to the observed one.

<table>
<thead>
<tr>
<th>Model type</th>
<th>Statistics</th>
<th>Lower Tail Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed log-returns</td>
<td>0.3267</td>
</tr>
<tr>
<td>Simulated log-returns</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Rough Correlation</td>
<td>0.2948</td>
<td>0.2975</td>
</tr>
<tr>
<td>Constant Correlation</td>
<td>Correlated volatility</td>
<td>0.1877</td>
</tr>
<tr>
<td></td>
<td>Independent volatility</td>
<td>0.0922</td>
</tr>
</tbody>
</table>

Table 5.4: Comparison of tail dependencies of simulated minute-wise log-returns by their mean, median, standard deviation and 90%, 95%, 99% quantiles with their counterparts from Gaussian copulas

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Process</th>
<th>Observed</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation in Observed</td>
<td>0.5896</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>HIX in Observed</td>
<td>0.7967</td>
<td>0.7645</td>
<td>0.7646</td>
</tr>
<tr>
<td>HIX in Rough Model</td>
<td>0.7632</td>
<td>0.7630</td>
<td>0.0046</td>
</tr>
<tr>
<td>HIX in Const. Corr. Model</td>
<td>0.5268</td>
<td>0.5270</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

Table 5.5: On the basis of 10000 simulated pairs of stock log-return processes the HIX and the correlations of minute-wise log-returns are compared using the rough stochastic correlation and a constant correlation model. The mean, median, standard deviation and 90%, 95%, 99% quantiles of the last 500 trading times are computed. The values of their counterparts computed by the ergodicity principle from observed stock log-returns are presented in rows 2-3.

For the observed log-return processes we calculate their correlation and the HIX by the ergodicity principle. The simulated synthetic stock log-return values make the estimation
of correlation and with that the HIX index possible at every trading time. We do that for the last 500 trading times and present its descriptive statistics in table 5.5.

The correlation of the observed two log-return processes slightly deviates from the temporal mean of the correlations computed in individual trading times from simulated log-return values. It is true for both the rough correlation and the constant correlation model. The reason may most probably be that temporally localized correlations are weaker and do not completely reflect the overall long term association. Indeed, we built up our model for temporally localized correlations, and their mean is 0.5308, which is reflected well by the model and the numbers in the last row of table 5.5. As a result of this deviance the HIX index computed from the observations slightly exceeds the mean, and even the 99% quantile of the simulated ones from both models.

What is the most interesting, however, is the very low variability in both the HIX and the correlations. This is a result of the stationarity of the model, built. Correlations are almost constant due to the correct modeling and since the distribution of the simulated values is also steady maximal correlations are almost unchangeable. As a result the HIX has very low variance. This also explains why there is practically no difference between the rough and the constant correlation model in their mean HIX values – even though variability is somewhat higher in the rough model but it is still very low. To the contrary of this, the tail dependence values show much greater variability.

Figure 5.6: HIX index against tail dependence for the last 500 trading times, computed from 10000 simulated log-return processes
Parallel to this, the HIX values displayed in figure 5.6 against the tail dependence values show very weak association. This circumstance calls for caution in risk management, in particular if the risk of coincident appearance of extreme values or events are of interest. Irrespective of whether this is a finite sample phenomenon or the result of the random fluctuation of local correlation, on this level of available information the steady HIX values may disguise an existing random fluctuation in the considered risk.

5.8 Conclusion

The notion of rough stochastic correlation of Brownian motions is introduced in the chapter, as the integrand in the expression of their quadratic covariation. Modeling Apple and Microsoft prices in frequent minute-wise trading by Wiggins’ stochastic volatility model, the latent Brownian motions driving the price dynamics are estimated. Analyzing their temporally localized correlations statistical evidence is found for the fractal property of those correlations. Further, the estimated fractal dimension of the path is very high – close to the enveloping space – and correspondingly, the estimated Hurst exponent is much lower than 0.5, it is close rather to 0.05. These are features consistent with the rough path property, in particular so on the given sample size level. Therefore, localized correlations are modeled with a rough process, a hyperbolic tangent transformed fOU process. It is fitted to the quasi sample obtained from the correlations of estimated Brownian motions. The model allows for extensive simulations of synthetic Apple and Microsoft stock prices. Fractal dimensions, Hurst exponents, distribution functions of local correlations are obtained from the simulations. These quantities are used as measures for the goodness of fit to approve the model choice, strengthening the claim on the roughness of the stochastic correlation, considered.

The rough correlation model creates tail dependence, similar to the one obtained from the observed data under hypothesized ergodicity. A fitted constant correlation model with highly correlated stochastic volatilities create far weaker tail dependence as observed, while the constant correlation with independent volatilities is not distinguishable from the no-tail-dependence model as opposed to observation. This indicates that the source of
stock price association is not only the similarity of the random environment in which they are set, i.e. the volatility, but internal price driving forces i.e. fundamentals also influence prices in a synchronized way and their effect significantly increase the probability of the occurrence of joint extreme events.

Relationship of tail dependence with the HIX index invented to describe co-movement of prices is also analyzed. Both HIX and tail-dependence are meant to describe the behavior of random variables in extreme situations, that typically exist in a stressed market. Our finding is that the HIX index has very similar values under the assumption of constant and rough correlation, and shows great stability in time. To the contrary of this, tail dependence in the rough correlation model has a greater mean value, almost double that of the constant correlation model. Moreover, tail dependence fluctuates more intensively in the rough model. While HIX judges the constant and rough correlation models equally risky, by considering fractal dimensions the model choice and with that, the risk of association in terms of tail dependence becomes quantitatively well accountable in real situations such as portfolio selection or management or other investment decisions.

Further research in the topic may concern model estimation issues, derivatives based inference and multiple stock price processes. In view of the hierarchical setup a Bayesian approach for model identification may be useful, while deep learning techniques may be considered in parallel with known rough volatility model estimations ([Bayer et al.(2019/2)]). Instead of historical data similar analysis on option prices would also be a matter of interest. Perhaps the most pressing question is how to model several stock prices instead of two. In our case the condition $|\rho(t)| \leq 1$ ensures the positive definiteness of the covariance matrix of the two log-returns at any time instant $t$. For multiple stocks, it is not so straightforward to find a proper substitute of this condition. For rough modeling fractional Wishart processes may be of use but their theory, in particular, the case $H < 0.5$ contains quite some peculiarity and is not so well developed in these days yet.
Chapter 6

Bibliography
Bibliography


[Dhaene et al.(2011)] Dhaene, J., Linders, D., Schoutens, W., Vyncke, D. The Herd Behavior Index: A New Measure for Systemic Risk in


[R package 'fractal'] https://CRAN.R-project.org/package=fractal
[R package ’fractaldim’] https://CRAN.R-project.org/package=fractaldim


Appendix A

Deriving the value of the CCS contract

Forward rates can be defined as the interest rates that can be locked today for an investment in a future time period. Forward rate can be understood through prototypical Forward Rate Agreement (FRA). Forward Rate Agreement (FRA) is an interest rate derivative contract between two parties who want to protect themselves against future movements in interest rates. FRA can be defined by three time instants: current time $t$, the expiry time $T_1$, and the maturity time $T_2$. This contract gives its holder an interest-rate payment for the period between $T_1$ and $T_2$. At the maturity $T_2$, a payment based on the fixed rate $K$ is exchanged against a floating payment based on the spot rate $L(T_1, T_2)$. In other words, this contract lets one to lock-in the interest rate between times $T_1$ and $T_2$ at a desired value $K$, for a contract with simply compounded rates. Therefore, the expected cash flows are discounted from time $T_2$ and $T_1$. At time $T_2$ one receives $\Delta(T_1, T_2)KN$ units of cash and pays the amount $\Delta(T_1, T_2)L(T_1, T_2)N$. Here $N$ denotes the contract’s nominal value and $\Delta(T_1, T_2)$ denotes the year fraction for the contract period $[T_1, T_2]$. The value of the FRA contract, at time $T_2$ can be expressed as

$$N \cdot \Delta(T_1, T_2) \cdot (K - L(T_1, T_2)) \quad \text{(A.1)}$$

Furthermore, $L(T_1, T_2)$ can also be expressed as

$$L(T_1, T_2) = \frac{1 - P(T_1, T_2)}{\Delta(T_1, T_2)P(T_1, T_2)} \quad \text{(A.2)}$$
Substituting it into equation 4.6

\[ N \left[ \Delta(T_1, T_2)K - \frac{1}{P(T_1, T_2)} + 1 \right] \]  

(A.3)

As per the no arbitrage theory, the implied forward rate between the time \( t \) and \( T_2 \) can be derived from two consecutive zero coupon bonds due to the equality (Filipovic, 2009)

\[ P(t, T_2) = P(t, T_1)P(T_1, T_2) \]  

(A.4)

The value of FRA in terms of simply compounded forward interest rate gives

\[ FRA(t, T_1, T_2, \Delta(T_1, T_2), N, K) = NP(t, T_2)\Delta(T_1, T_2)(K - F_s(t; T_1, T_2)), \]  

(A.5)

where \( F_s(t; T_1, T_2) \) is the simply compounded forward interest rate prevailing at time \( t \) for the expiry \( T_1 > t \) at maturity \( T_2 > T_1 \), defined as

\[ F_s(t; T_1, T_2) = \frac{P(t, T_1) - P(t, T_2)}{\Delta(t, T_2)P(t, T_2)} = \frac{1}{\Delta(T_1, T_2)} \left( \frac{P(t, T_1)}{P(t, T_2)} - 1 \right) \]  

(A.6)

Seeing CCS contract from X’s side, as a portfolio of FRA, every individual FRA can be evaluated by using equation (A5). The the value of the CCS for X denoted \( \pi_{\text{payer}}(t) \), is given as

\[ \pi_{\text{payer}}(t) = \sum_{i=\alpha+1}^{\beta} FRA(t, T_1, T_2, \Delta(T_{i-1}, T_i), N, F_d) \]  

(A.7)

\[ \pi_{\text{payer}}(t) = N \sum_{i=\alpha+1}^{\beta} \Delta_iP(t, T_i)(F_f(t; T_{i-1}, T_i) - F_d(t; T_{i-1}, T_i)) \]  

(A.8)

Using equation (A5) and (A6) in (A8) we get
\[
\pi_{\text{payer}}(t) = N \sum_{i=\alpha+1}^{\beta} \left[ P^f(t, T_{i-1}) - P^f(t, T_i) - (P^d(t, T_{i-1}) - P^d(t, T_i)) \right] \quad (A.9)
\]

The above sum can be decomposed into 2 sums

\[
N \sum_{i=\alpha+1}^{\beta} (P^f(t, T_{i-1}) - P^f(t, T_i)) + N \sum_{i=\alpha+1}^{\beta} (P^d(t, T_i) - P^d(t, T_{i-1})) \quad (A.10)
\]

The first sum can be simplified into

\[
N \sum_{i=\alpha+1}^{\beta} (P^f(t, T_{i-1}) - P^f(t, T_i)) = N(P^f(t, T_{\alpha}) - P^f(t, T_{\beta})) \quad (A.11)
\]

The second sum can be simplified into

\[
N \sum_{i=\alpha+1}^{\beta} (P^d(t, T_{i-1}) - P^d(t, T_i)) = N(P^d(t, T_{\beta}) - P^d(t, T_{\alpha})) \quad (A.12)
\]

However, the floating leg need to converted into domestic currency by multiplying it by the spot foreign exchange rate \(\psi_t\) and the notionals are also exchanged at the maturity date. Using the relation \(N^d = \psi_0 N^f\) and the equations (A11) and (A12) we get

\[
\pi_{\text{payer}}(t) = N^f [\psi(t)(P^f(t, T_{\alpha}) - P^f(t, T_{\beta})) + N^f \psi(T) - N^d[P(t, T_{\alpha}) - P(t, T_{\beta})] - N^d, \quad (A.13)
\]
<table>
<thead>
<tr>
<th>Source</th>
<th>Data Type</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://www.global-rates.com/">https://www.global-rates.com/</a></td>
<td>USD-LIBOR and EUR-LIBOR</td>
<td>1-day, 1-week, 3-months, 6-months and 1-year</td>
</tr>
<tr>
<td><a href="https://www.interestrateswapstoday.com/">https://www.interestrateswapstoday.com/</a></td>
<td>USD-SW AP and EUR-SW AP</td>
<td>January 05, 2014 - April 01, 2018</td>
</tr>
<tr>
<td><a href="https://www.feddata.com/">https://www.feddata.com/</a></td>
<td>USD YIELD</td>
<td>February 25, 2018 - February 03, 2019</td>
</tr>
<tr>
<td><a href="https://www.investing.com/">https://www.investing.com/</a></td>
<td>EUR/USD FX</td>
<td>January 05, 2014 - April 01, 2018</td>
</tr>
<tr>
<td><a href="https://www.riskmetricstoday.com/">https://www.riskmetricstoday.com/</a></td>
<td>IMPLIED VOL of EUR/USD FX OPTION</td>
<td>1-10 years</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>CDS</td>
<td>January 05, 2014 - April 01, 2018</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>EUROSWAP</td>
<td>February 03, 2019</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>USD/JPY FX</td>
<td>February 03, 2019</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>EUR/USD FX</td>
<td>February 03, 2019</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>IMPLIED VOL of EUR/USD FX OPTION</td>
<td>1-10 years</td>
</tr>
<tr>
<td>Bloomberg</td>
<td>CDS</td>
<td>January 05, 2014 - April 01, 2018</td>
</tr>
</tbody>
</table>