

A Square Root Process for Modelling Correlation



Dissertation

zur Erlangung des akademischen Grades eines
Doktor der Naturwissenschaften (Dr. rer. nat.)

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

Cathrin van Emmerich

Promotionsausschuß:

Gutachter/Prüfer: Prof. Dr. Michael Günther
Gutachter/Prüfer: Prof. Dr. Michael Nelles
Prüfer: Prof. Dr. Reinhard Michel
Prüfer: Prof. Dr. Wil Schilders

Diese Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20090117

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20090117>]

Acknowledgment

When I started my PhD in 2005, I wanted to learn more about mathematics and finance, gaining deeper knowledge and a better understanding. Now two years later I have more questions than ever before...

But I am tremendously grateful that I had the possibility of freedom and shelter at university to learn and teach, to develop my own answers and ask my own questions. I would like to express my gratefulness at this point knowing that I cannot mention everyone who should be mentioned.

Prof. Dr. Michael Günther gave me this opportunity and from the very beginning credit. Thank you very much for this.

Also I would like to thank all recent and former members of the Numerical Analysis working group for support, cooperation and company. Particularly I would like to mention Dr. Andreas Bartel. I could always rely on his interest, accuracy, and skills.

I had the pleasure to supervise some bright and motivated students at university. Thanks for making my work so enjoyable!

Besides I would like to express thank to Prof. Dr. Michael Nelles and Dr. Martin Uzik for letting me profit from their knowledge.

Thanks also goes to Prof. Dr. Wil Schilders, who gave me the possibility to spend some months at the TU Eindhoven.

Further I would like to thank Prof. Dr. Reinhard Michel, who awoke with his excellent lecture my interest in stochastics.

During my PhD I could spend a great time with the equity quant team of Bear Stearns in London headed by Oli Jonsson. It was very valuable to me. Thanks to the whole team, especially to Valer Zetocha.

Christian Kahl influenced my PhD years and the thesis itself very strongly and always to the best. Thank you so much.

Lastly I would like to thank my parents and Eike for being as they are and doing what they have done. I am very proud of you.

Contents

- Acknowledgment** **i**
- Contents** **ii**
- 1 Introduction** **1**
 - 1.1 Motivation 1
 - 1.2 Linear correlation coefficient 4
 - 1.3 Outline 6
- 2 Model** **8**
 - 2.1 Observed characteristics of correlation 8
 - 2.2 Stochastically correlated Brownian motions 10
 - 2.3 Bounded mean reversion model 12
- 3 Analytical Properties** **14**
 - 3.1 Boundary behaviour 14
 - 3.2 Stationary density 20
 - 3.3 Moments 26
 - 3.4 Summary 32
- 4 Maximum-Likelihood estimator** **33**
 - 4.1 Overview 33
 - 4.2 Estimating the integral of an Ornstein-Uhlenbeck process 36
 - 4.3 Fitting the stochastic correlation process 40
 - 4.4 Numerical tests 42
 - 4.5 Summary 47
- 5 Application** **48**
 - 5.1 Foreign exchange model 48
 - 5.2 Large homogeneous portfolio model (LHPM) 54
 - 5.3 Summary 61
- 6 Conclusion** **62**
- A Appendix** **65**
- Bibliography** **72**
- Index** **75**

List of Figures

2.1	Correlation between DJI and EURUSD, 1998 - 2005	9
2.2	Correlation between DJI and its volatility, 02 - 07	9
2.3	Average correlation in SX5E, 02 - 07	10
3.1	Restriction on κ for non-attractiveness	20
3.2	Stationary density for varying κ and θ	24
3.3	First three centralised moments of correlation	28
3.4	Recursive computation of moments of integrated correlation	30
3.5	First three centralised moments of realised correlation	31
4.1	Density of integrated correlation process vs. normal density	34
4.2	Density of integrated correlation process vs. normal density	35
4.3	Distribution of estimation results for mean	44
4.4	Distribution of estimation results for diffusion parameter.	45
4.5	Distribution of estimation results for speed of mean reversion	45
5.1	Density (3.21) fitted to correlation between DJI and EURUSD	51
5.2	Call price in FX model with fixed and stochastic correlation for different initial values	52
5.3	Prices for call in FX model with fixed and stochastic correlation for different level of correlation	53
5.4	Relative price difference and implied correlation in FX model (A)	53
5.5	Relative price difference and implied correlation in FX model (B)	54
5.6	Impact of number of paths on density in SLHPM	56
5.7	Densities in LHPM and SLHPM for varying speed of mean reversion	57
5.8	Densities in LHPM and SLHPM for varying level of correlation	58
5.9	Densities in LHPM and SLHPM for varying default probabilities	58
5.10	Tranche prices in LHPM and SLHPM (A)	59
5.11	Tranche prices in LHPM and SLHPM (B)	60
5.13	Implied correlation in SLHPM for varying default probability and mean	60
5.12	Implied correlation for LHPM and SLHPM	61
A.1	Percentage of paths preserving the boundaries.	68

List of Tables

1.1	Example of correlation swap	2
4.1	Numerical results for estimation of Ornstein-Uhlenbeck process (A)	43
4.2	Numerical results for estimation of Ornstein-Uhlenbeck process (B)	44
4.3	Numerical results for estimation of Ornstein-Uhlenbeck process (C)	46
4.4	Numerical results for estimation of correlation process	46

CHAPTER I

Introduction

Financial markets have been growing substantially over the last decades. This growth could be realised primarily by innovation. Investment banks introduce new products which are of value for existing and new customers. The motives for buying financial products differ depending on the counter-parties. In any case investment banks as well as their customers will only buy substantial amounts of financial derivatives if price and risk structure seem appealing or at least fair to them. Thus it is not only necessary to develop innovative products but one also has to understand the dynamics of it. Further it may be necessary to adapt these dynamics to the market and find a suitable computational realisation.

That is where quantitative finance or financial mathematics and computational finance come into play. The task is to set up a suitable model describing all parts of the market which have important influence on the derivative and to hedge its risk thereafter. This requires a recalibration of the model parameters on a day to day basis. Lastly it should be implemented to obtain prices and risk analysis. Naturally a balance between complexity of the model and speed of the implementation has to be found.

The motivation underlying this thesis is one particular rather new product: the **correlation swap**. We are going to suggest a model for correlation to price such a swap. But let us first have a short look on the structure of a correlation swap:

1.1 Motivation

The correlation swap is a swap on realised correlation. It is yet not publicly traded but the major investment banks trade correlation swaps over-the-counter. The correlation swap is a financial derivative on a basket of assets which pays off the correlation between the logarithmic changes in the assets.

Definition 1.1 (Correlation swap on m -dimensional basket) An unweighted correlation swap on $i = 1, \dots, m$ underlyings S_i is defined by an observation structure $\mathcal{T} = \{t_0, t_1, \dots, t_n\}$, start date, maturity and its payoff:

$$N \cdot (C - K)$$

where N denotes the notional and K the strike. C denotes the **averaged realised correlation** between the underlying assets:

$$(1.1) \quad C = \frac{2}{m \cdot (m-1)} \sum_{i=1}^m \sum_{j=i+1}^m \frac{\sum_{k=1}^n (X_i^k - \sum_{\ell=1}^n X_i^\ell) (X_j^k - \sum_{\ell=1}^n X_j^\ell)}{\sqrt{\sum_{k=1}^n (X_i^k - \sum_{\ell=1}^n X_i^\ell)^2 \sum_{k=1}^n (X_j^k - \sum_{\ell=1}^n X_j^\ell)^2}}.$$

Due to the implicitly assumed lognormal distribution of the assets, the realised correlation is calculated based on the logarithmic changes:

$$X_i^k = \log \left(\frac{S_i^k}{S_i^{k-1}} \right), \quad k = 1, \dots, n.$$

Often the payoff is simplified to

$$(1.2) \quad C_{\text{simple}} = \frac{2}{m \cdot (m-1)} \sum_{i=1}^m \sum_{j=i+1}^m \frac{\sum_{k=1}^n (X_i^k) (X_j^k)}{\sqrt{\sum_{k=1}^n (X_i^k)^2 \sum_{k=1}^n (X_j^k)^2}}.$$

Exemplarily a correlation swap could look like:

Example 1.2 (Correlation Swap) The contract is specified by

Underlying	EuroStoxx50: S_1, \dots, S_{50}
Start date	$t_0 = 02.01.2007$
maturity and pay off date	$t_n = 29.06.2007$
Observation dates	closing prices of days when all S_i are traded
Strike	$K = 0.4$
Payoff	$20000 \cdot (C_{\text{simple}} - K)$

Table 1.1: Example of correlation swap

In an analogous way it is possible to construct European and American style calls, puts etc on realised correlation.

Note that this product makes only sense if the underlying asset - correlation - is not constant and not known in advance. Thus the market assumes that correlation is stochastic. For pricing the correlation risk itself and correlation dependent derivatives, it is necessary to develop a suitable model.

We consider the correlation swap as the main motivation for the work presented in this thesis, as it shows a particularly clear distinction between correlation risk and other risk sources. Nonetheless there are other areas of application for stochastic correlation as well. For example in credit, the pricing of **Collateralised Debt Obligation** (CDO) strongly depends on the assumed interdependency structure of the relevant debtors. It is a type of asset-backed security and we define it as follows:

Definition 1.3 (Collateralised Debt Obligation) A CDO is a financial product derived from a basket of $i = 1, \dots, m$ defaultable assets S_i with notional N_i . The respective recovery rates are denoted by R_i . The payoff depends on tranches $[0, U_1], (U_1, U_2], (U_{n-1}, U_n]$ with $0 < U_1 \leq \dots \leq U_n \leq 1$. Over the period t_0 until maturity T the total loss is added, resulting in:

$$L = \sum_{i=1}^m (N_i - R_i) \cdot 1_{\{S_i, \text{defaulted}\}}$$

where $1_{\{S_i, \text{defaulted}\}}$ takes the value one if asset S_i defaulted and zero otherwise. The payoff for the j -th tranche is specified by

$$(L - U_{j-1})^+ - (L - U_j)^+.$$

The tranches are called equity, mezzanine and senior tranches in ascending order.

Note that one must consider time and definition of default, interest rate etc for a real CDO, which we neglect here. The risk structure of the single tranches depends on the correlation. It is particularly easy to see for the most-senior tranche. The tranche is only hit if a very high number of assets default. This will usually only happen if they are highly correlated. Some models like the large homogeneous portfolio model (LHPM) use constant correlation, but they fail in reproducing market quotes. Throughout the tranches one observes an implied correlation skew. Introducing stochastic correlation could lead to a better market fit.

So the CDO as well as the credit field in general could turn out to be useful application for stochastic correlation.

Correlation swap and CDO are products which are clearly correlation dependent. In addition there are already well known models which use the concept of correlated Brownian motions, for example stochastic volatility models such as the Heston model [Hes93] where the correlation describes the dependence between two Brownian motions where one of them drives the asset and the other one the asset's (stochastic) volatility:

$$(1.3) \quad dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t, \quad S_0 = s_0,$$

$$(1.4) \quad dV_t = \kappa(\theta - V_t) dt + \sigma \sqrt{V_t} dZ_t, \quad V_0 = v_0,$$

$$(1.5) \quad dW_t dZ_t = \rho dt.$$

Here $\mu, \kappa, \theta, \sigma, s_0$ and v_0 denote positive constants, $\rho \in [-1, 1]$ is also constant.

Another example is a basic foreign exchange model, abbreviated with FX model in the following:

$$(1.6) \quad dS_t^1 = \mu_1 S_t^1 dt + \sigma_1 S_t^1 dW_t, \quad S_0^1 = s_0^1,$$

$$(1.7) \quad dS_t^2 = \mu_2 S_t^2 dt + \sigma_2 S_t^2 dZ_t, \quad S_0^2 = s_0^2,$$

$$(1.8) \quad dW_t dZ_t = \rho dt,$$

with positive constants $\mu_1, \mu_2, \sigma_1, \sigma_2$ as well as positive initial values s_0^1 and s_0^2 . $\rho \in [-1, 1]$ denotes again a constant. Hereby the first geometric Brownian motion describes an asset in a foreign currency. Whilst the second prescribes the exchange rate between that asset's

currency and the domestic currency.

In the Heston model the correlation coefficient is important because it describes the shape of the volatility skew. In the FX model the correlation leads to a parameter restriction because of arbitrage arguments, see chapter 5.

The Brownian motions in both models are correlated with the (constant) correlation coefficient ρ :

$$(1.9) \quad dW_t dZ_t = \rho dt$$

which is a symbolic notation for

$$(1.10) \quad E[W_t Z_t] = \rho \cdot t.$$

As this thesis tries to extend the concept of the linear correlation coefficient we say a few more words about it.

1.2 Linear correlation coefficient

There are different concepts of dependence. We would like to mention the copula approach which is often used in credit modelling see for example Schönbucher [Sch03]. The **copula ansatz** is a very flexible framework. It features a clear distinction between marginal and joint distribution. Originally a static approach, recently one tries to extend the idea to the dynamic case via Lévy copulae (see for example Cont and Tankov [CT04] or Tappe [Tap07]). But if the risk described by a copula is supposed to be sold as a product, the mathematical problem arises how to measure that risk. From a market point of view it would be necessary that all participants agree on one copula. At least they should comprehend the risk concept and how the risk measure is calculated. Currently such an agreement cannot be observed in the market.

For the sake of completeness we mention a totally different concept of dependence, the long-range-dependence as it is for example represented in form of a **fractional Brownian motions**. But as we are rather interested in the dependence between assets and not in the dependence through time it does not seem appropriate for our purposes. For properties of fractional Brownian motion and theoretical framework see for example Duncan et al [DHPD00].

In case of the CDO various dependence approaches can be considered, but the correlation swap clearly indicates the use of the linear correlation coefficient, which is the idea we pursue in this thesis. It is closely connected to covariance. For example Gourieaux et al [GJS04] show a multivariate covariance model where the matrix is modelled as a **Wishart process**. We prefer to stick to the pure correlation, because we think that the modelling, calibrating and also the market itself for variance is well developed. Here we refer to the stochastic volatility models of Heston [Hes93] and the local volatility models such as the one presented by Dupire [Dup94]. There is extensive work on the numerical treatment, for example Kahl et al [KJ06] or Andersen [And06]. We would like to use these accomplishments by separating the modelling of variance and correlation strictly.

From a market point of view the linear correlation coefficient has several advantages compared to more complex dependence approaches like long-term memory processes and copulae. They are useful modelling approaches but it is difficult to create products which exclusively map their specific risk. Moreover it makes only sense to specify a product on a risk exposure which everyone can easily comprehend. In contrast to that the linear correlation coefficient is a measure of dependence which is known to everyone. Further it can be easily reproduced by every market participant. We give the definition for two random variables

Definition 1.4 (Linear correlation coefficient) The linear correlation coefficient ρ is defined for two random variables X and Y as

$$(1.11) \quad \rho := \frac{E[(X - E[X])(Y - E[Y])]}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

It is a meaningful measure of dependence between normal random variables.

In view of the definition above for two random variables it is between -1 and 1 . It is often quoted in form of the correlation matrix. For multivariate X_1, \dots, X_n it reads:

$$(1.12) \quad C = \begin{pmatrix} 1 & \rho_{1,2} & \dots & \rho_{1,n} \\ \rho_{1,2} & 1 & \ddots & \rho_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{1,n} & \rho_{2,n} & \dots & 1 \end{pmatrix}$$

where $\rho_{i,j} = \rho_{j,i} = \frac{E[(X_i - E[X_i])(X_j - E[X_j])]}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}}$ denotes the pairwise correlation for $1 \leq i, j \leq n$. Note that the correlation matrix must be symmetric and positive definite.

By definition the correlation coefficient ρ is a number. We consider two random variables X and Y . If realisations $\bar{X}_1, \dots, \bar{X}_n$ and $\bar{Y}_1, \dots, \bar{Y}_n$ are given it can be estimated pairwise as

$$(1.13) \quad \rho \approx \tilde{\rho} = \frac{\sum_{i=1}^n \left(\bar{X}_i - \sum_{j=1}^n \bar{X}_j \right) \left(\bar{Y}_i - \sum_{j=1}^n \bar{Y}_j \right)}{\sqrt{\sum_{i=1}^n \left(\bar{X}_i - \sum_{j=1}^n \bar{X}_j \right)^2 \sum_{i=1}^n \left(\bar{Y}_i - \sum_{j=1}^n \bar{Y}_j \right)^2}}.$$

Short calculations show that the estimator $\tilde{\rho}$ is unbiased.

If the realisations are observed at t_0, \dots, t_n , the quantity $\tilde{\rho} = \tilde{\rho}(t_0, \dots, t_n)$ is also denoted as **realised correlation** over t_0, \dots, t_n .

Related to this estimator is the idea of **rolling correlation**. Given $n + 1$ data points at t_0, \dots, t_n , the rolling correlation over intervals of length $m < n$ reads:

$$\tilde{\rho}(t_0, \dots, t_{m-1}), \tilde{\rho}(t_1, \dots, t_m), \dots, \tilde{\rho}(t_{n-m+1}, \dots, t_n).$$

In consideration of definition 1.4, rolling correlation should be constant - less numerical and estimation errors. As observed rolling correlation is not constant even for large m , we will suggest and analyse an extension of the concept of the linear correlation coefficient in this thesis:

1.3 Outline

In the beginning of chapter 2 we present some visualisations of correlation. Due to the lack of market data, we solely focus on the historically observed correlation. From this we will derive some properties which correlation should possess. Mainly we will see that correlation is indeed non-deterministic but its behaviour rather resembles a stochastic process.

As a next step we introduce a univariate stochastic process which can be used to model correlation. In analogy to interest rates and having in mind the structure of a correlation swap, the model describes **short correlation**.

Chapter 3 shows some analysis on the suggested stochastic correlation process. Particularly it shows that with regard to boundaries and density the process seems suitable for modelling correlation, but we also derive some necessary parameter restrictions. Additionally a system of ordinary differential equations is derived which describes all moments of correlation and integrated correlation.

We use this knowledge in chapter 4, in order to develop an estimator for the correlation process. The estimation procedure works with historical data, but notably it bases on the observed integrated correlation. One can argue that there is only information about that data available. The estimator is of Maximum-Likelihood type, where we need to use the similarity to Ornstein-Uhlenbeck processes.

The subsequent chapter 5 shows two examples. The first example deals with a call in a FX framework as presented above. We compare the constant correlation case to the model where we use the stochastic correlation process from chapter 2. The second example is taken from credit. We consider the LHPM for pricing CDOs and analyse the influence of stochastic correlation on implied correlation, density etc.

The thesis concludes with a short summary and an outlook showing what would be possible or necessary areas for further work from a practical and from a theoretical point of view respectively.

There is also a short appendix where we say a couple of words about the numerical treatment, which we particularly need for the examples. Further some of the computations needed in chapter 3 can be found here.

Assumptions

To ease reading, we want to introduce some of the assumptions we make throughout this thesis already at this point:

- Without loss of generality we assume $t = 0$ as the starting point of stochastic processes.
- We assume that all stochastic differential equations have constant initial value without loss of generality. We do so for the sake of simplicity, although one can argue that for some of the considered processes the initial value is not observable, and should therefore be featured with some noise. One possibility to include stochastic initial values into financial models is presented in [PE07].

- Throughout this thesis we assume that the appropriate filtrations \mathcal{F}_t underly the stochastic processes. Thus we denote the stochastic processes by Z_t , they generate the filtration: $\mathcal{F}_t = \mathcal{F}(Z_{(0 \leq s \leq t)})$.

Moreover we need some assumptions on the financial environment:

- Interest rate is constant.
- Market is frictionless and arbitrage free. In the last chapter we also postulate completeness.
- Assets are lognormally distributed with constant mean and variance.

Further we will use the following notation:

- Given a random variable X , we denote one realisation with \bar{X} .
- The estimator of a parameter p is described as \tilde{p} .
- The stepsize is usually denoted forward in time: $\Delta_i = t_{i+1} - t_i$.
- A process X_{t_i} at t_i is abbreviated with X_i .
- The process for stochastic correlation is given in X_t whereas the variable ρ describes constant correlation.
- \mathbb{Z}_- denotes the non-positive integers.

Additionally if there is no risk of misunderstanding, we will use the term correlation for denoting the linear correlation coefficient.

CHAPTER II

Model

In this chapter the concept of the linear correlation coefficient is extended from a single number to a stochastic process. Naturally the extension should still cover the case of constant correlation. Further it should fit the observed characteristics of correlation. We start with the latter. Then we will show how to stochastically correlate Brownian motions. Lastly we introduce a stochastic correlation model.

2.1 Observed characteristics of correlation

In the introduction it is motivated why it is necessary to extend the concept of the constant correlation coefficient to stochastic correlation. The main motivation is the existence of a correlation swap. Unfortunately this correlation swap is not exchange traded. Thus we do not have sufficient data for a market analysis. Instead of that we will look on the historic behaviour of correlation. It indicates that 'CORRELATIONS ARE EVEN MORE UNSTABLE THAN VOLATILITIES' as mentioned by Wilmott[Wil00a]. Figure 2.1 shows the estimated correlations between Dow Jones Composite (DJI) and Euro/US-Dollar exchange rate (EURUSD) on a daily basis¹. Considered is the period from 1998 till 2005. Hereby it is assumed that both processes follow a lognormal distribution as in (1.6) and (1.7). Shown is the rolling correlation over time frames of two weeks. The usual estimator as given in (1.13) is used.

Figure 2.1 (A) clearly shows that correlation is not constant over time. Moreover, correlation seems even to be nondeterministic. In figure 2.1 (B) we receive a first impression how the density of a stochastic process modelling correlation could look like. The distribution possesses one peak. From that peak it decreases while approaching the boundaries where it vanishes.

The next pictures in Fig. 2.2 are the analogue for the DJI and its volatility. It is assumed that the index S moves like a geometric Brownian motion. The volatility v is

¹Source of data: yahoo.com

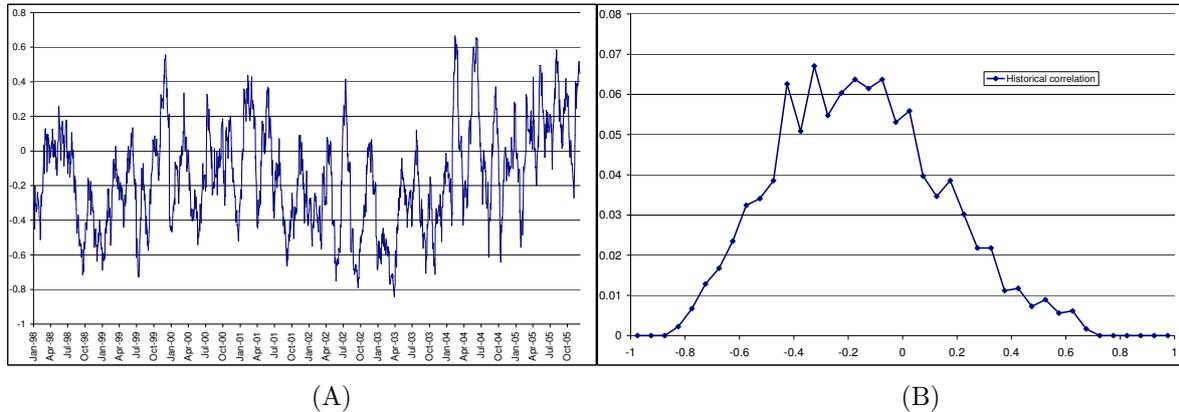


Figure 2.1: Correlation between Dow Jones and Euro/US-Dollar exchange rate from 1998 until 2005, process and density. (A): Estimated historical correlation over time. (B): Empirical density.

measured as standard deviation. For $n + 1$ data points it is calculated as

$$(2.1) \quad \tilde{v} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\bar{X}_i - \frac{1}{n} \sum_{j=1}^n \bar{X}_j \right)^2}.$$

Hereby X denotes the logarithmic returns in the index from t_{i-1} to t_i :

$$(2.2) \quad X_i = \ln \left(\frac{S_i}{S_{i-1}} \right)$$

Fig. 2.2 (A) shows the rolling correlation between the DJI and its historical volatility². The historical standard deviation is calculated over time intervals of 20 business days. As can be read from the legend, the rolling correlation refers to 20 and 60 business days.

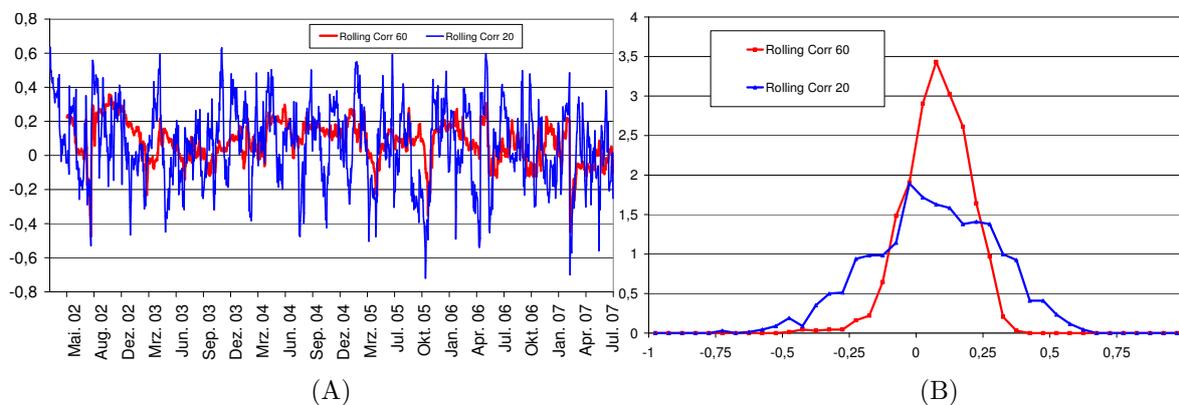


Figure 2.2: Correlation between Dow Jones Composite and its volatility from 2002 until 2007, process and density estimated over $n = 20$ and $n = 60$ business days. (A): Estimated historical correlation over time. (B): Empirical density.

²Source of data: yahoo.com

As in case of DJI and EURUSD correlation is not constant. Again one observes a peaked density, which vanishes in the boundary. The longer observation interval of 60–days leads hereby to a more stable path and therewith a more centered density compared to the 20–day rolling correlation.

Lastly the correlation between all stocks in the EuroStoxx50 (SX5E) is visualised, measured as the average of the pairwise correlations as in (1.1).

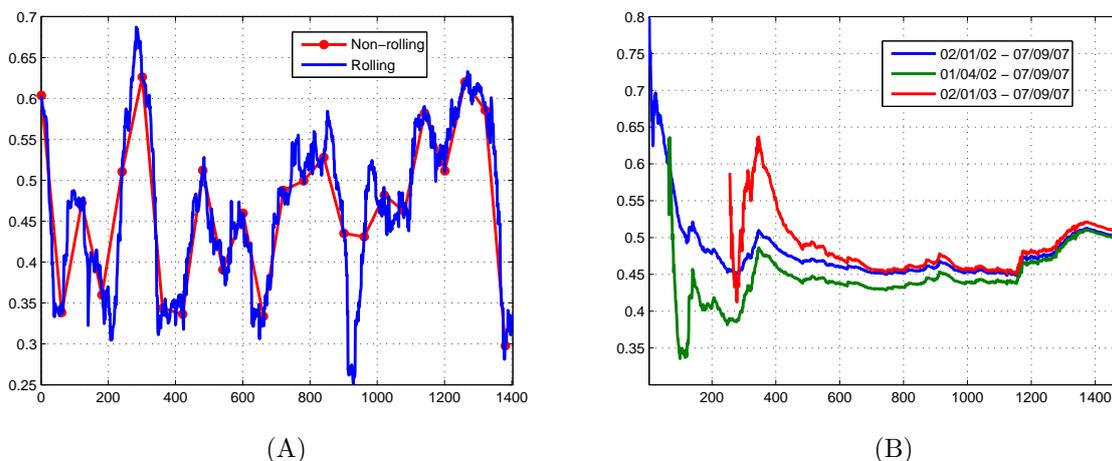


Figure 2.3: Average correlation in SX5E. (A): Rolling correlation versus correlation calculated for non-overlapping intervals from 02/01/02 until 07/09/07. (B): Realised correlation until 07/09/07 with different starting dates.

Fig. 2.3(A) shows the path of correlation over a time window of 60 days. Hereby the rolling correlation is shown (corresponding to overlapping time intervals) and the correlation estimated for distinct intervals of 60 days. Fig. 2.3(B) depicts the realised correlation for different initial dates. For each day they show the respective value of a correlation swap maturing at this day with according start date.

All graphs indicate that correlation is not deterministic. Therefore assuming constant correlation in a financial model neglects the correlation risk. In view of the figures above we sum up as follows:

- Correlation seems to vary stochastically around a mean.
- There is some reverting to that mean over time.
- Notably correlation does not reach the boundaries.

2.2 Stochastically correlated Brownian motions

Before introducing the model analysed within this thesis it is illustrated what we understand if we say stochastically correlated. Hereby we consider only ordinary Brownian

motions. An Itô-process X_t given by

$$(2.3) \quad dX_t = a(t, X_t)dt + b(t, X_t)dB_t, \quad X_0 \in \mathbb{R} \text{ constant}$$

with $X_t \in [-1, 1] \forall t$ can be used to correlate two Brownian motions W and Z . Hereby denotes B a Brownian motion. For the construction of Z let W and V be two independent Brownian motions, also independent of B . Set

$$(2.4) \quad Z_t = \int_0^t X_s dW_s + \int_0^t \sqrt{1 - X_s^2} dV_s, \quad Z_0 = 0.$$

For proving that Z (2.4) is a Brownian motion, we remark that $Z_0 = 0$ and show that $E[Z_t^2] = t$. Because of the Itô-isometry (see for example [Øks00]), we can calculate:

$$E[Z_t^2] = E \left[\left(\int_0^t X_s dW_s \right)^2 + \left(\int_0^t \sqrt{1 - X_s^2} dV_s \right)^2 + 2M_t \right] = t.$$

with $M_t := \int_0^t X_s dW_s \int_0^t \sqrt{1 - X_s^2} dV_s$. We used that $E[M_t] = 0$ which follows from $M_0 = 0$ and

$$dM_t = X_t dW_t \int_0^t \sqrt{1 - X_s^2} dV_s + \sqrt{1 - X_t^2} dV_t \int_0^t X_s dW_s + X_t \sqrt{1 - X_t^2} \underbrace{dW_t dV_t}_{=0}.$$

In a similar way, one verifies that $E[Z_t | \mathcal{F}_s] = Z_s, t \geq s$. By virtue of Lévy's Theorem (see again [Øks00]) it follows that Z is a Brownian motion. Thus we have two Brownian motions W and Z correlated by the stochastic process X . Especially it holds for W and Z that

$$E[Z_t \cdot W_t] = E \left[\int_0^t X_s ds \right]$$

which agrees for constant $X_s = \rho \forall s$ with

$$E[Z_t \cdot W_t] = \rho t.$$

This construction should be borne in mind throughout the remaining parts of this thesis. We have not mentioned the correlation between B and Z , B and W respectively. For simplicity we assume independence throughout this thesis, but clearly generalisation is possible and desirable. Observing market behaviour leads to the clear conclusion that correlation and market movement are not independent.

2.3 Bounded mean reversion model

The model we suggest and analyse within this thesis is a time-continuous bounded mean reverting model described by a stochastic differential equation. The rate of mean reversion is denoted by κ , the level of mean reversion by θ . Both as well as the diffusion parameter ω are constant. The process we suggest for describing stochastic correlation X_t reads therewith:

$$(2.5) \quad dX_t = \kappa(\theta - X_t) dt + \omega \sqrt{(r - X_t)(X_t - \ell)} dW_t$$

with $\ell < r$ and initial value $X_0 \in (\ell, r)$. Clearly the process is only well-defined on the interval $[\ell, r]$. The parameters κ and ω are assumed to be positive. The mean θ shall be in (ℓ, r) . The stochastic correlation process (2.5) is driven by a Brownian motion W_t . As already mentioned for simplicity this Brownian motion is assumed to be independent of the Brownian motions (or more generally the random processes) which are correlated by X_t .

If the process describes the correlation between two random variables, one could set $\ell = -1, r = 1$. If a higher number of identically distributed assets is considered (see also chapter 5) the upper boundary stays the same $r = 1$, but the lower boundary increases. Therefore if not stated otherwise we will set $r = 1$ in the following and the left boundary remains unspecified. As it describes the minimal value we change the notation from ℓ to $m \in [-1, 1)$ and therefore consider the stochastic process

$$(2.6) \quad dX_t = \kappa(\theta - X_t) dt + \omega \sqrt{(1 - X_t)(X_t - m)} dW_t .$$

In general the need for positive definiteness can make particular choices of the boundaries ℓ and r necessary.

For pricing a correlation swap on realised correlation as mentioned above the raw correlation is not explicitly needed. But there is a similarity to the approximation of an integral.

Let V and W denote Brownian motions stochastically correlated with X . Let the realised correlation observed at the market between 0 and t be denoted with C . One can approximate:

$$C \approx \frac{E[V_t W_t | \mathcal{F}_0]}{t} = \frac{1}{t} E \left[\int_0^t X_s ds \right]$$

For a single realisation holds:

$$\bar{C}_r \approx \frac{1}{t} \int_0^t \bar{X}_s ds \approx \frac{1}{t} \sum_{i=1}^n \bar{X}_i \cdot \Delta_i$$

with equidistant stepsizes $\Delta_i = \frac{t}{n}$. Thus in the following we will also consider the stochastic process Y for realised correlation. It is given as the time scaled integral over the

correlation process X :

$$(2.7) \quad Y_t := \frac{1}{t} \int_0^t X_s ds.$$

A possible interpretation of the relationship between realised correlation Y and X is that the process (2.6) describes the instantaneous or **short correlation** in analogy to short rate models in interest rate modelling.

CHAPTER III

Analytical Properties

In chapter 2 we suggested the mean reverting process (2.6) for modelling correlation. In accordance to the visualisations of historically observed correlation in Fig. 2.1, 2.2 and 2.3 and coinciding with intuition, the process possesses a mean reverting property. Naturally this is not sufficient for making (2.6) a suitable modelling approach. The most obvious requirement for such a process is that it preserves the boundaries, which are -1 and 1 in the univariate case. But it is as important that the model shares some of the observed characteristics of correlation. Therefore we will do some analysis in this chapter which shows that the idea of modelling correlation via (2.6) can be pursued.

We start by analysing the process (2.6) suggested in chapter 2 with regard to its behaviour at the boundaries m and 1 . One will observe that certain restrictions on the parameters are necessary. If these conditions are not satisfied, the process can reach the degenerated state of perfect positive or negative correlation.

For the appropriate parameter configurations, where the boundaries are not reached, the transition density for $t \rightarrow \infty$ is calculated in the second section of this chapter. This density is stationary. It will enable us to compare the model density with the one extracted from historical data. Additionally the stationary density can be used for calibration.

Further insight will be gained in the third section. All moments of correlation given by (2.6) and realised correlation (2.7) are calculated. This knowledge has various applications. One of them will be shown in chapter 4 where it is used to derive a historical parameter estimator. Further it can be used for fitting the model to historical observation or to market.

3.1 Boundary behaviour

The processes (2.5) and (2.6) respectively are only useful for modelling correlation if they are bounded. We show this property by proving that there are parameter configurations for which the boundaries are not reached. As the process is continuous, it therefore does not leave the interval given by the boundaries ℓ and r if $X_0 \in (\ell, r)$. Hereby we follow the notation of Karlin and Taylor[KT81] for classifying the boundaries $\ell = m$ and $r = 1$. We

start with a general bounded Itô diffusion x_t with $x_t \in [\ell, r]$ and omit time-dependence as the process under consideration does not feature time-dependent parameters:

$$(3.1) \quad dx_t = a(x_t)dt + b(x_t)dW_t, \quad x_0 \in \mathbb{R}.$$

We concentrate on the left boundary ℓ in the following. Later we will set $\ell = m$ as in (2.6). The analysis of the right one works analogously. Now we introduce some notations, thereby we assume $x \in (\ell, r)$,

$$s(v) = \exp\left(-\int_{v_0}^v \frac{2a(w)}{b^2(w)} dw\right), \quad v_0 \in (\ell, x), \quad v \in (v_0, r),$$

$$S(x) = \int_{x_0}^x s(v)dv, \quad x_0 \in (\ell, x),$$

$$S[c, d] = S(d) - S(c) = \int_c^d s(v)dv, \quad (c, d) \in (\ell, r),$$

$$S(\ell, x) = \lim_{a \rightarrow \ell} S[a, x].$$

We already indicate that x_0 and v_0 will be of no relevance in the following. We use that $dS(x) = s(x)dx$. S is called the scale measure whereas N is the speed measure:

$$n(x) = \frac{1}{b^2(x)s(x)},$$

$$N[c, d] = \int_c^d n(x)dx.$$

Analogously we can use the relation $dN(x) = n(x)dx$. Lastly we need the expression

$$\Sigma(\ell) = \lim_{a \rightarrow \ell} \int_a^x N[v, x]dS(v).$$

Using these notations we can classify the left boundary.

Definition 3.1 (Attractiveness and Attainability) The boundary ℓ is called attractive if and only if there is an $x \in (\ell, r)$ such that

$$(3.2) \quad S(\ell, x) < \infty.$$

Otherwise it is non-attractive. Furthermore ℓ is classified as an attainable boundary if and only if

$$(3.3) \quad \Sigma(\ell) < \infty.$$

Otherwise it is unattainable.

Attractiveness

We want to analyse if there are parameters configurations for which the boundaries of (2.6) are not attractive. We focus on the left boundary m . Note that the procedure can be used for the right boundary 1 in a very similar way. It is also straightforward to extend the results to (2.5).

First we are going to determine $S(m, x]$ for the mean-reverting process (2.6). We compute $s(x)$ as a first step:

$$\begin{aligned}
-\log(s(v)) &= \int_{v_0}^v \frac{2\kappa(\theta - z)}{\omega^2(1-z)(z-m)} dz \\
&= \left[\frac{2\kappa}{(m-1)\omega^2} ((\theta-1)\log(x-1) + (m-\theta)\log(x-m)) \right]_{v_0}^v \\
&= \log \left[(-1+v)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} \left((-1+v_0)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v_0)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} \right)^{-1} \right] \\
&= \log \left[(1-v)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} \left((1-v_0)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v_0)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} \right)^{-1} \right] \\
&= \log \left[(1-v)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} c_{v_0}^{-1} \right]
\end{aligned}$$

with $c_{v_0}^{-1} = (1-v_0)^{\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v_0)^{\frac{2\kappa(m-\theta)}{\omega^2(m-1)}}$. Consequently

$$s(v) = c_{v_0} (1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}}.$$

For the behaviour near the left boundary m we have to investigate

$$S(m, x] = \lim_{a \rightarrow m} c_0 \cdot \int_a^x (1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dv.$$

We note that $1-a \geq 1-v \geq 1-x > 0$ for $v \in [a, x] \subset (m, 1)$ and $-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)} \leq 0$.

Therefore the term $(1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}}$ is bounded for $v \in [a, x]$:

$$(3.4) \quad (1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} \geq (1-a)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} \geq (1-m)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}},$$

$$(3.5) \quad (1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} \leq (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}}.$$

Applying the first inequality (3.4) we can estimate as follows

$$\begin{aligned}
S(m, x] &\geq \lim_{a \rightarrow m} (1-a)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \int_a^x (-m+v)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dv \\
&= \lim_{a \rightarrow m} (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \left[\left(-\frac{2\kappa(m-\theta)}{\omega^2(m-1)} + 1 \right)^{-1} (-m+v)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \right]_a^x
\end{aligned}$$

for $\frac{2\kappa(m-\theta)}{(m-1)\omega^2} \neq 1$. The first factor converges to a constant. Therefore we only need to look at

$$(3.6) \quad (-m + a)^{1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2}}.$$

The expression $(-m + a)$ is positive and approaching 0 for $a \in (m, x)$, $a \rightarrow m$. Thus (3.6) converges if the exponent is positive and diverges if it is negative. If $\frac{2\kappa(m-\theta)}{(m-1)\omega^2} = 1$ one obtains

$$S(m, x) \geq \lim_{a \rightarrow \ell} (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} [\ln(-m+v)]_a^x \geq \infty.$$

In all we can write

$$S(m, x) \geq \begin{cases} \text{const} & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} > 0 \\ \infty & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} \leq 0. \end{cases}$$

The exponent $1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2}$ is not positive and the boundary m therewith non-attractive if

$$\frac{2\kappa(m-\theta)}{(m-1)\omega^2} \geq 1.$$

By applying the inequality (3.5) we show in which case the left boundary is attractive. For that we calculate as before:

$$\begin{aligned} S(m, x) &\leq c_{v_0} (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} \lim_{a \rightarrow m} \left[\left(-\frac{2\kappa(m-\theta)}{\omega^2(m-1)} + 1 \right)^{-1} (-m+v)^{1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \right]_a^x \\ &= \begin{cases} \text{const} & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} > 0 \\ \infty & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} < 0. \end{cases} \end{aligned}$$

again for $\frac{2\kappa(m-\theta)}{(m-1)\omega^2} \neq 1$. Including this case as before we get:

$$S(m, x) \leq \begin{cases} \text{const} & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} > 0 \\ \infty & \text{for } 1 - \frac{2\kappa(m-\theta)}{(m-1)\omega^2} \leq 0. \end{cases}$$

We can sum up, the left boundary m is attractive if

$$(3.7) \quad \frac{2\kappa(m-\theta)}{(m-1)\omega^2} < 1.$$

It is not attractive if

$$(3.8) \quad \frac{2\kappa(m-\theta)}{(m-1)\omega^2} \geq 1$$

which is the case suitable for modelling correlation.

Attainability

Now we are going to figure out if the left boundary is attainable. We remark that only an attractive boundary can be attainable, which follows directly from the definition of Σ . For that reason we assume $\frac{2\kappa(m-\theta)}{(m-1)\omega^2} < 1$ in the following. For notational simplicity, let c denote a suitable constant in the remaining part of this section. We consider

$$\begin{aligned}\Sigma(m) &= \lim_{a \rightarrow m} \int_a^x N[v, x] dS(v) = \lim_{a \rightarrow m} \int_a^x S[a, v] dN(v) \\ &= \lim_{a \rightarrow m} \int_a^x \int_a^v s(w) dw \frac{1}{b^2(v)s(v)} dv.\end{aligned}$$

First we calculate $\Sigma(m)$ for $\kappa = 0$:

$$\begin{aligned}\Sigma(m) &= \lim_{a \rightarrow m} \frac{1}{(m-1)\omega^2} \cdot [(a-1) \log(a-1) + (m-a) \log(a-m) \\ &\quad - (a-1) \log(x-1) - (m-a) \log(x-m)] \\ &= \frac{1}{(m-1)\omega^2} \lim_{a \rightarrow m} \left[(a-1) \log\left(\frac{a-1}{x-1}\right) + (m-a) \log\left(\frac{a-m}{x-m}\right) \right].\end{aligned}$$

It holds

$$\begin{aligned}\lim_{a \rightarrow m} (a-1) \log\left(\frac{a-1}{x-1}\right) &= \lim_{a \rightarrow m} (a-1) \log\left(\frac{1-a}{1-x}\right) < \infty, \\ \lim_{a \rightarrow m} (m-a) \log\left(\frac{a-m}{x-m}\right) &= \lim_{a \rightarrow m} (m-a) \log\left(\frac{m-a}{m-x}\right) = 0.\end{aligned}$$

Hence the left boundary m is attainable if $\kappa = 0$. Now we consider the case $\kappa > 0$. We can deduce that

$$\begin{aligned}\Sigma(m) &= \lim_{a \rightarrow -1} \int_a^x \left(\int_a^v c_{v_0} (1-w)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+w)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dw \right) \frac{1}{\beta^2(v)s(v)} dv \\ &\leq (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \lim_{a \rightarrow -1} \int_a^x \left(\int_a^v (-m+w)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dw \right) \frac{1}{\beta^2(v)s(v)} dv \\ &= (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \left(-\frac{2\kappa(m-\theta)}{\omega^2(m-1)} + 1 \right)^{-1} \lim_{a \rightarrow -1} \frac{1}{\omega^2 c_{v_0}} \\ &\quad \left[\int_a^x \left[(-m+w)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \right]_a^v \frac{1}{(1-v)(v-m)} \frac{1}{(1-v)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{-\frac{2\kappa(m-\theta)}{\omega^2(m-1)}}} dv \right] \\ &= \frac{1}{\omega^2} (1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \left(-\frac{2\kappa(m-\theta)}{\omega^2(m-1)} + 1 \right)^{-1} \lim_{a \rightarrow -1} \int_a^x \left((-m+v)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \right. \\ &\quad \left. - (-m+a)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \right) (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-1) (-m+v)^{-1+\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dv.\end{aligned}$$

We will use the abbreviation $K = \frac{1}{\omega^2}(1-x)^{-\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} c_{v_0} \left(-\frac{2\kappa(m-\theta)}{\omega^2(m-1)} + 1\right)^{-1}$ which is positive and does not depend on a . Therewith:

$$\begin{aligned}
\Sigma(m) &\leq -K \lim_{a \rightarrow -1} \int_a^x (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} \\
&\quad - (-m+a)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{-1+\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dv \\
&= K \lim_{a \rightarrow -1} (-m+a)^{1-\frac{2\kappa(m-\theta)}{(m-1)\omega^2}} \int_a^x (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (-m+v)^{-1+\frac{2\kappa(m-\theta)}{\omega^2(m-1)}} dv \\
&\quad - \int_a^x (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} dv \\
&\leq K \lim_{a \rightarrow -1} \int_a^x (1-v)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} dv - K (1-m)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (x-m) \\
&\leq K \lim_{a \rightarrow -1} (1-x)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (x-a) - K (1-m)^{-1+\frac{2\kappa(-1+\theta)}{\omega^2(m-1)}} (x-m).
\end{aligned}$$

Due to the attractiveness condition we already assumed $-1 + \frac{2\kappa(m-\theta)}{(m-1)\omega^2} < 0$. Therefore $\Sigma(m) < \infty$ holds. Thus if the left boundary m is attractive, it is also attainable. If condition (3.8) holds, the left boundary is neither attractive nor attainable.

Following the same line of thought one derives that the right boundary $r = 1$ is attractive and attainable if

$$(3.9) \quad \frac{\kappa}{\omega^2}(1-\theta) < 1.$$

Otherwise it is neither attractive nor attainable.

We sum up in a theorem:

Theorem 3.2 (Boundary condition) The stochastic process given by (2.6) possesses unattractive and unattainable boundaries m and 1 if and only if

$$(3.10) \quad \frac{2\kappa}{\omega^2} \geq \max\left(\frac{2}{1-\theta}, \frac{m-1}{m-\theta}\right).$$

Only if this condition is fulfilled, we will consider the stochastic process as a suitable model for correlation.

We will only consider (2.6) with (3.10) in the following.

Fig. 3.1 depicts the lowest possible values for the mean reversion parameter κ according to Thm. 3.2. The red lines in 3.1(A) show the lowest possible value for κ such that the boundaries $m = -0.1$ and 1 are not attractive (and therewith not attainable) resulting

from $\omega = 1$ for different mean values. The blue line shows the respective information for $\omega = 2$. The dashed lines describe the restriction which is necessary for the left boundary to be not attractive, the solid lines show the restrictions for the upper bound. Fig. 3.1(B) visualises the lowest possible value for κ such that the lower boundary is not attractive for different diffusion parameter and differing mean.

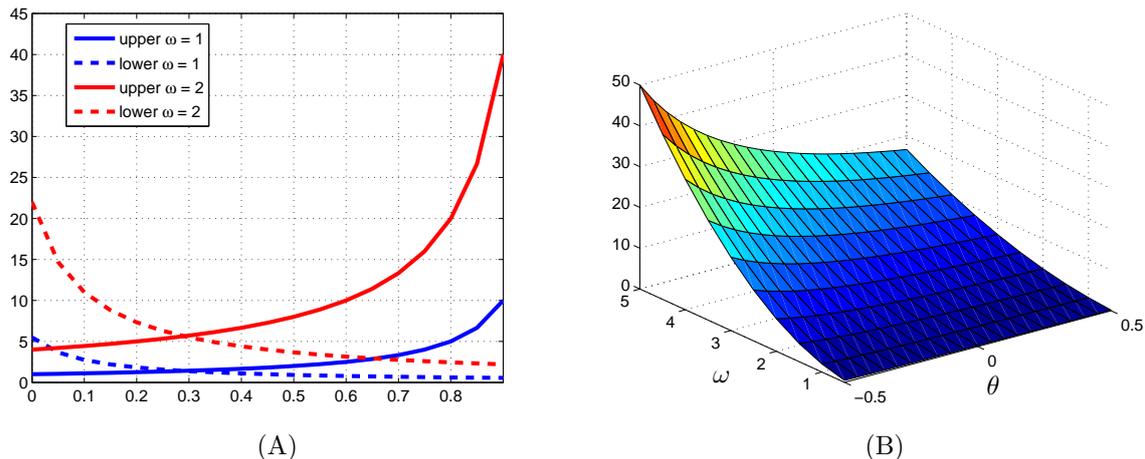


Figure 3.1: (A): Lowest κ such that (2.6) is not attractive for $m = -0.1, \omega = 1, 2, \theta \in [0, 0.9]$. (B): Lowest κ such that (2.6) is not attractive for $m = -1, \omega \in [0.5, 5], \theta \in [-0.5, 0.5]$.

One observes that a higher diffusion parameter forces higher level for the mean reversion. Fig. 3.1(A) shows that the restrictions are not symmetric. The restriction resulting from the upper boundary is more often active (for $\theta \gtrsim 0.3$). As already indicated by (A), (B) shows that increasing ω must be balanced by an over-proportional increasing of the rate of mean reversion.

In the following we will only deal with the case of unattainable boundaries. Thus we exclusively allow for parameter configurations which satisfy (3.10).

3.2 Stationary density

The transition density describes the distribution of a stochastic process in t starting at t_0 . In this section we will compute the transition density of (2.6) for $t \rightarrow \infty$, where it becomes stationary. By determining this density we hope to gain further insight into the behaviour of the stochastic correlation process. Particularly we will be able to compare it to the observed distributions of correlation as shown for example in Fig. 2.1.

We will use the **Fokker-Planck equation**, also known as Kolmogoroff forward equation. It is a partial differential equation of parabolic type. The transition density solves it. We restrict ourselves to non-time dependent coefficient functions. More precisely, assuming that the stochastic differential equation

$$(3.11) \quad dX_t = a(X_t)dt + b(X_t)dW_t, \quad X_0 = x_0,$$

possesses a transition density $p(t, y|x_0)$, then p satisfies the Fokker-Planck equation

$$(3.12) \quad \frac{\partial}{\partial t} p(t, x) + \frac{\partial}{\partial x} (a(x)p(t, x)) - \frac{1}{2} \frac{\partial^2}{\partial x^2} (b(x)^2 p(t, x)) = 0.$$

For example in case of an ordinary Brownian motion the Fokker-Planck equation reduces to a diffusion equation:

$$\frac{\partial}{\partial t} p(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (p(t, x)).$$

For a thorough discussion of the Fokker-Planck equation and its properties we refer to the book by Risken[Ris89]. We want to derive the transition density of the correlation process for $t \rightarrow \infty$. We demand the solution p to fulfill two structural conditions. Firstly p is required to be a density

$$(3.13) \quad \int_m^1 p(t, x) dx = 1.$$

Moreover we postulate that p preserves the expectation¹

$$(3.14) \quad \int_m^1 x \cdot p(t, x) dx \xrightarrow{t \rightarrow \infty} \theta.$$

For the mean reversion process (2.6) one can show that every two solutions of (3.12) for $t \rightarrow \infty$ are equal, see Risken[Ris89], section 6.1. As a consequence it suffices to show that a stationary solution exists to know that it is the unique solution. This is how we proceed in the following: We will derive a stationary solution which fulfills the stationary Fokker Planck equation

$$\frac{\partial}{\partial x} (a(x)p(x)) - \frac{1}{2} \frac{\partial^2}{\partial x^2} (b(x)^2 p(x)) = 0$$

as well as (3.13) and (3.14). Firstly we examine the simplified case $\theta = 0$, $\omega = 1$ and $m = -1$ and look for the stationary solution $p_0(x) = \lim_{t \rightarrow \infty} p_0(t, x)$. We consider the stationary Fokker-Planck equation for $a(x) = -\kappa x$, $b(x) = \sqrt{1 - x^2}$:

$$(3.15) \quad (1 - \kappa)p_0(x) + x(2 - \kappa)p_0'(x) - \frac{1}{2}(1 - x^2)p_0''(x) = 0.$$

We receive as a solution to (3.15)

$$(3.16) \quad p_0(x) = (1 - x^2)^{(\kappa-1)} \left(c + b \int_x^1 (1 - z^2)^\kappa dz \right)$$

¹Short calculations show that for the expectation of (2.6) holds $E[X_t] = \theta + (X_0 - \theta) \exp(-\kappa t)$, which converges to θ for $t \rightarrow \infty$. See also section 3.3.

with constants $b, c \in \mathbb{R}$. If $b = 0$, p_0 is symmetric around $x = 0$. Hence

$$\int_{-1}^1 x \cdot p_0(x) dx = 0 = \theta,$$

and (3.14) holds. Thus we choose $b = 0$. The still free variable c must be chosen such that condition (3.13) is fulfilled. The primitive of p can be described via the hyper-geometric function. See for example [AS72] and [WW72]. The classical standard hyper geometric series ${}_2F_1$ for $a, b, c, z \in \mathbb{R}$ is given by

$$(3.17) \quad {}_2F_1(a, b, c, y) = 1 + \frac{ab}{c}y + \frac{a(a+1)b(b+1)}{2!c(c+1)}y^2 + \dots$$

If c is a negative integer, one of the denominators becomes zero. Therefore we exclude that case. If a or b is a negative integer, the series terminates at the $(1-a)$ th or $(1-b)$ th term respectively and thus converges. For all other parameter configurations, one can prove that ${}_2F_1$ is absolute convergent if

$$(3.18) \quad a + b - c < 0.$$

It is possible to formulate many functions with the help of ${}_2F_1$, among others:

$$(1 - x^2)^{\kappa-1} = {}_2F_1(1 - \kappa, 1, 1, x^2).$$

Therewith one can easily integrate for $\kappa \in [1, \infty)$:

$$\int (1 - x^2)^{(\kappa-1)} dx = x \cdot {}_2F_1(1 - \kappa, \frac{1}{2}, \frac{3}{2}, x^2) > 0.$$

Thus we must choose c as

$$c = \frac{1}{2 \cdot {}_2F_1(1 - \kappa, \frac{1}{2}, \frac{3}{2}, 1)}.$$

With this choice

$$(3.19) \quad p_0(x) = c(1 - x^2)^{(\kappa-1)}$$

becomes a density.

As a next step we consider the mean-reverting process (2.6) with $\theta \neq 0$, but still $\omega = 1$, $m = -1$. The density p is solution of

$$(3.20) \quad (1 - \kappa)p(x) + x(2 - \kappa)p'(x) + \kappa\theta p'(x) - \frac{1}{2}(1 - x^2)p''(x) = 0.$$

Lengthy calculations show that

$$p(x) = \left(\frac{1-x}{1+x} \right)^{-\kappa\theta} (1-x^2)^{\kappa-1} \cdot \left[c + b \int_x^1 e^{(-\kappa(2\theta \tanh^{-1}(z) + \log(z^2-1)))} dz \right]$$

solves (3.20) with $b, c \in \mathbb{R}$. For $\theta = 0$ we want p to be equal to the above solution (3.19). This is only possible if $b = 0$. Thus the final solution is²

$$p(x) = c \left(\frac{1-x}{1+x} \right)^{-\kappa \cdot \theta} (1-x^2)^{\kappa-1},$$

with c such that $\int_{-1}^1 p(x) = 1$ holds. Note that p is only well-defined on $[-1, 1]$ if the parameter constraint (3.10) holds. We sum up the results in a theorem:

Theorem 3.3 (Stationary Density) The transition density p of

$$dX_t = \kappa(\theta - X_t) dt + \sqrt{1 - X_t^2} dW_t$$

for $t \rightarrow \infty$ under the constraints (3.13) and (3.14) reads

$$(3.21) \quad p(x) = c \left(\frac{1-x}{1+x} \right)^{-\kappa \cdot \theta} (1-x^2)^{\kappa-1},$$

with

$$(3.22) \quad c = \left(\frac{1}{\kappa + \theta \kappa} {}_2F_1(1, (\theta - 1)\kappa + 1; \theta \kappa + \kappa + 1; -1) + \frac{1}{\kappa - \theta \kappa} {}_2F_1(1, 1 - (\theta + 1)\kappa; -\theta \kappa + \kappa + 1; -1) \right)^{-1}.$$

Proof: It remains to calculate the constant c and to show that (3.21) with (3.22) preserves the expectation (3.14). The necessary computations can be found in the appendix A.2. \square

Figure 3.2 shows how the stationary densities may look like.

We can observe that

- p is concentrated on $[-1, 1]$.
- p is symmetric with respect to θ in the sense of $p^\theta(x) = p^{-\theta}(-x)$.
- If $\theta = 0$ the global maximum is attained at $x = 0$.
- The probability mass decreases when getting away from the mean. In particular, it vanishes in the boundary values.

These properties agree with intuition and what we demanded in the chapter 2. Especially they correspond to figure 2.1, 2.2 and 2.3.

²Note that this corresponds to a β -distribution.

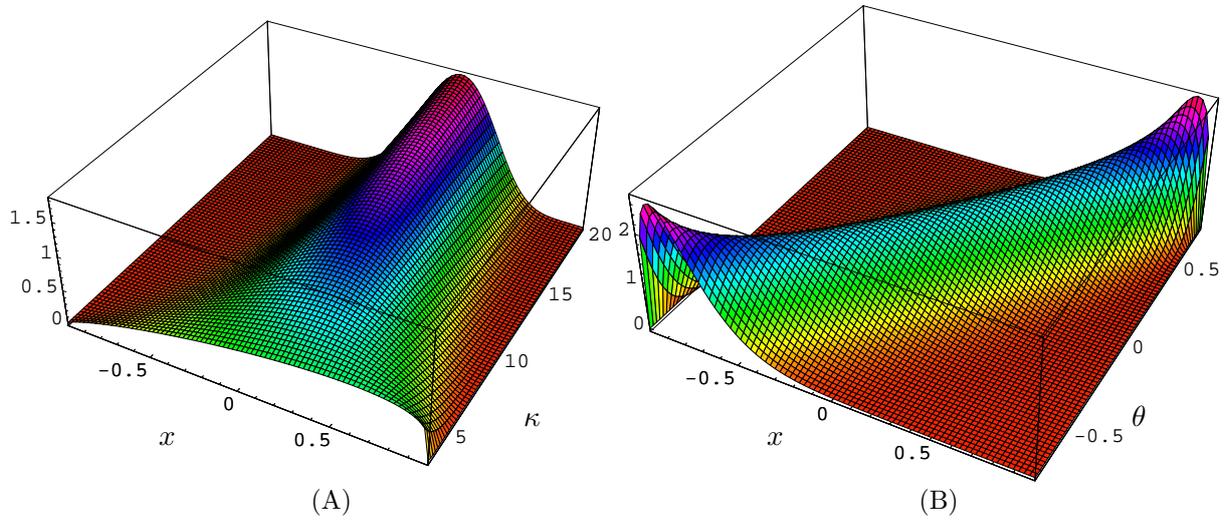


Figure 3.2: (A): Stationary density (3.21) of (2.6) with $m = -1$ and $\omega = 1$ for varying $\kappa \in [3, 20]$, $\theta = 0.2$, (B): Stationary density (3.21) of (2.6) with $m = -1$ and $\omega = 1$ for varying $\theta \in [-0.75, 0.75]$, $\kappa = 12$.

Remark 3.4 (Calibration via stationary density) The transition density can be used for calibrating the parameters. Hereby one fits the historically observed density to the stationary density (3.21). See for example Wilmott [Wil00b], chapter 46. An example from FX is shown in Fig. 5.1.

Remark 3.5 (Static stochastic correlation) A random variable distributed according to the density (3.21) can be used in static models as stochastic correlation. We show an example from credit in chapter 5.

One can easily extend the result to the more general case of (2.6). We do so in two steps. We start by allowing for arbitrary lower bound $m \in [-1, 1)$ in (2.6). The diffusion parameter is still fixed, $\omega = 1$. As throughout this chapter, we assume that (3.10) holds.

Corollary 3.6 The stationary density for

$$dX_t = \kappa(\theta - X_t) dt + \sqrt{(1 - X_t)(X_t - m)} dW_t$$

reads

$$(3.23) \quad p_m(x) = c \frac{\left(\frac{1-x}{x-m}\right)^{1-\frac{\kappa(m-2\theta+1)}{m-1}} ((1-x)(x-m))^\kappa}{(x-1)^2}$$

with

$$c = \left((1-m)^{2\kappa-1} \Gamma(\kappa-1) \Gamma\left(\frac{2(m-\theta)\kappa}{m-1}\right) \cdot {}_2F_1\left(\frac{2(m-\theta)\kappa}{m-1}, \frac{(m-2\theta+1)\kappa}{m-1} - 1; \frac{(3m-2\theta-1)\kappa}{m-1} - 1; 1\right) \right)^{-1} \cdot \Gamma\left(\frac{(3m-2\theta-1)\kappa}{m-1} - 1\right).$$

Proof: We transform

$$dX_t = \kappa(\theta - X_t) dt + \sqrt{(1-X_t)(X_t+1)} dW_t$$

with

$$x \rightarrow y = f(x) = \frac{(1-m)}{2}(x+1) + m \in [m, 1].$$

The inverse function is

$$y \in [m, 1] \rightarrow x = f^{-1}(y) = (y-m) \frac{2}{(1-m)} - 1 \in [-1, 1].$$

By applying Itô's formula we obtain the following stochastic differential equation for the Itô process $Y_t = f(X_t)$:

$$\begin{aligned} dY_t &= f'(X_t) dX_t = \frac{1-m}{2} \left(\kappa(\theta - X_t) dt + \sqrt{(1-X_t)(X_t+1)} dW_t \right) \\ &= \kappa \left((\theta+1) \frac{(1-m)}{2} + m - Y_t \right) dt + \sqrt{(1-Y_t)(Y_t-m)} dW_t \\ &= \kappa \left(\tilde{\theta} - Y_t \right) dt + \sqrt{(1-Y_t)(Y_t-m)} dW_t \end{aligned}$$

where $\tilde{\theta} = (\theta+1) \frac{1-m}{2} + m$. We know the stationary density for the process X_t from theorem 3.3:

$$p(x) = \left(\frac{1-x}{1+x} \right)^{-\kappa \tilde{\theta}} (1-x^2)^{\kappa-1},$$

less scaling by a constant. The density for the transformed process can be derived from that:

$$\begin{aligned} h(x) &= p(f^{-1}(x)) (f^{-1}(x))' \\ &= \frac{2^{2\kappa-3} (m-1)^3 \left(\frac{x-1}{m-x}\right)^{1-\kappa \tilde{\theta}} \left(\frac{(m-x)(x-1)}{(m-1)^2}\right)^\kappa}{(x-1)^2}. \end{aligned}$$

The scaling constant c is calculated in A.2, see (A.9). Therewith the original process X_t possesses the stationary density as given. \square

Next we state the stationary density for the correlation process with arbitrary $\omega > 0$, which corresponds to (2.6):

Corollary 3.7 The stationary density for (2.6)

$$(3.24) \quad dX_t = \kappa (\theta - X_t) dt + \omega \sqrt{(1 - X_t)(X_t - m)} dW_t$$

is given by

$$(3.25) \quad p_\omega(x) = c \frac{\left(\frac{1-x}{x-m}\right)^{1 - \frac{\kappa}{\omega^2} \frac{(m-2\theta+1)}{m-1}} \left((1-x)(x-m)\right)^{\frac{\kappa}{\omega^2}}}{(x-1)^2}$$

with

$$\begin{aligned} c = & \left((1-m)^{\frac{2\kappa}{\omega^2}-1} \Gamma\left(\frac{\kappa}{\omega^2}-1\right) \Gamma\left(\frac{2(m-\theta)\kappa}{(m-1)\omega^2}\right) \right. \\ & \cdot {}_2F_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2}-1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2}; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2}-1; 1\right) \Big)^{-1} \\ & \cdot \Gamma\left(\frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2}-1\right). \end{aligned}$$

Proof: The stationary Fokker-Planck equation for (2.6) reads.

$$\begin{aligned} 0 &= \frac{\partial}{\partial x} (\kappa (\theta - x) p_\omega(x)) - \frac{1}{2} \frac{\partial^2}{\partial x^2} (\omega^2 (1-x)(x-m) p_\omega(x)) \\ \iff 0 &= \frac{\partial}{\partial x} \left(\frac{\kappa}{\omega^2} (\theta - x) p_\omega(x) \right) - \frac{1}{2} \frac{\partial^2}{\partial x^2} ((1-x)(x-m) p_\omega(x)) \end{aligned}$$

Now one can use corollary 3.6 and obtains the result. The constant c is computed in the appendix , see (A.8). \square

3.3 Moments

In this section we will derive a system of ordinary differential equations describing the moments of the stochastic correlation process (2.6). As already indicated in chapter 2, the integral (2.7) over this stochastic process can be interpreted as an approximation to realised correlation, which is for example used in the correlation swap and of importance as well. Therefore we will also calculate its moments.

3.3.1 Moments of stochastic correlation

Dufresne [Duf01] derived a system of ordinary differential equations for the moments of the Cox-Ingersoll-Ross process [CIR85]

$$(3.26) \quad dZ_t = \alpha (\beta - Z_t) dt + \sigma \sqrt{Z_t} dW_t, \quad Z_0 = z_0$$

and its integral

$$(3.27) \quad \int_0^t Z_s ds .$$

This process is famous for modelling the short rate in an interest rate framework or for modelling volatilities. Its popularity is due to (among others) its analytical tractability, which allows for semi-analytical solutions. In contrast to (2.6) it is an affine linear diffusion such that the characteristic function is given as a solution of a system of Riccati equations. We refer to Duffie, Pan and Singleton [DPS00] for a detailed discussion. Here we apply the methodology of Dufresne to the correlation process (2.6).

In this section we use the abbreviation $m_i(t) = E[X_t^i]$. We can directly calculate that

$$m_0(t) = 1 \quad \text{and} \quad m_1(t) = \theta + (X_0 - \theta) \exp(-\kappa t)$$

holds³. Therefore let $i \geq 2$ in the following. By virtue of Itô's formula we can calculate the stochastic differential equations for X^i :

$$\begin{aligned} dX_t^i &= i \cdot X_t^{i-1} dX_t + \frac{i(i-1)}{2} X_t^{i-2} (dX_t)^2 \\ &= iX_t^{i-1} \left(\kappa(\theta - X_t) dt + \omega \sqrt{(1-X_t)(X_t-m)} dW_t \right) \\ &\quad + \frac{i(i-1)}{2} X_t^{i-2} \omega^2 (1-X_t)(X_t-m) dt \\ &= \left[-\frac{im(i-1)}{2} \omega^2 X_t^{i-2} + \left(i\kappa\theta + \frac{i(i-1)}{2} \omega^2 (1+m) \right) X_t^{i-1} \right. \\ &\quad \left. - \left(i\kappa + \frac{i(i-1)}{2} \omega^2 \right) X_t^i \right] dt + iX_t^i \omega \sqrt{(1-X_t)(X_t-m)} dW_t . \end{aligned}$$

Using the equivalent integral presentation and taking the expectation one obtains:

$$\begin{aligned} E[X_t^i] &= X_0^i - \frac{im(i-1)}{2} \omega^2 \int_0^t E[X_s^{i-2}] ds + \left(i\kappa\theta + \frac{i(i-1)}{2} \omega^2 (1+m) \right) \\ &\quad \cdot \int_0^t E[X_s^{i-1}] ds - \left(i\kappa + \frac{i(i-1)}{2} \omega^2 \right) \int_0^t E[X_s^i] ds . \end{aligned}$$

Differentiating with respect to t , we obtain the following system of ordinary differential equations describing the moments of X :

$$(3.28) \quad m'_i(t) = a_i m_{i-2}(t) + b_i m_{i-1}(t) + c_i m_i(t) , \quad \forall i \geq 2$$

$$(3.29) \quad m_1(t) = \theta + (X_0 - \theta) \exp(-\kappa t) ,$$

$$(3.30) \quad m_0(t) = 1 ,$$

³The first equation $m_0(t) = 1$ goes without saying. Standard calculation show the second: $X_t = X_0 + \kappa \int_0^t \theta - X_s ds + \omega \int_0^t \sqrt{(1-X_s)(X_s-m)} dW_s$ leads in expectation to $E[X_t] = X_0 + \kappa\theta t - \kappa \int_0^t E[X_s] ds$. The solution to this differential equation is $m_1(t)$.

with the auxiliary parameters

$$(3.31) \quad a_i = -\frac{im(i-1)}{2}\omega^2,$$

$$(3.32) \quad b_i = i\kappa\theta + \frac{i(i-1)}{2}\omega^2(1+m),$$

$$(3.33) \quad c_i = -\left(i\kappa + \frac{i(i-1)}{2}\omega^2\right).$$

Solving this system of differential equations exemplarily for the second moment leads to the following results:

$$\begin{aligned} E[X_t^2] &= \frac{1}{\omega^4 + 3\kappa\omega^2 + 2\kappa^2} e^{-t(\omega^2+2\kappa)} \\ &\cdot \left((\omega^4 + 3\kappa\omega^2 + 2\kappa^2) X_0^2 + (-1 + e^{t(\omega^2+\kappa)}) (\omega^2 + 2\kappa) ((m+1)\omega^2 + 2\theta\kappa) X_0 \right. \\ &\quad - (-1 + e^{t(\omega^2+2\kappa)}) m\omega^2 (\omega^2 + \kappa) \\ &\quad - (m+1)\theta\omega^2 \left((-1 + 2e^{t(\omega^2+\kappa)} - e^{t(\omega^2+2\kappa)}) \kappa - e^{t(\omega^2+\kappa)} (-1 + e^{t\kappa}) \omega^2 \right) \\ &\quad \left. - 2\theta^2\kappa \left((-1 + 2e^{t(\omega^2+\kappa)} - e^{t(\omega^2+2\kappa)}) \kappa - e^{t(\omega^2+\kappa)} (-1 + e^{t\kappa}) \omega^2 \right) \right). \end{aligned}$$

As the solutions for higher moments become even more complicated, their explicit presentations are omitted, but recursive calculation directly leads to the solutions.

Figure 3.3 shows the first three centralised moments for two different parameter configurations:

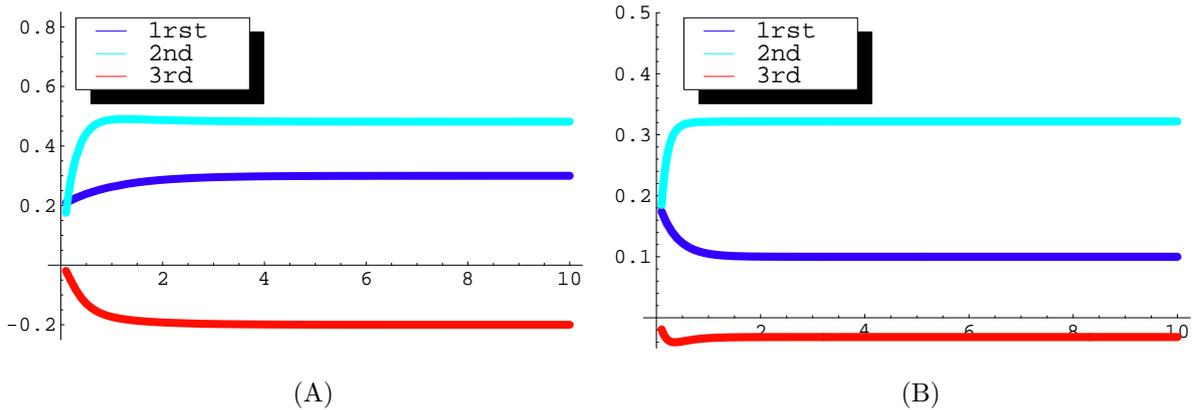


Figure 3.3: (A): First three centralised moments of correlation for $\kappa = 1, \theta = 0.3, \omega = 1.5, m = -1$ and $X_0 = 0.2$. (B): First three centralised moments of correlation for $\kappa = 3, \theta = 0.1, \omega = 1.7, m = -1$ and $X_0 = 0.2$.

As expected the first moment converges to the long term mean θ . The second moment increases up to $t \approx 1$ in (A) and up to $t \approx \frac{1}{3}$ in (B), slightly decreases and then stays stable afterwards. The third moment in (A) shows a decrease which becomes flatter for higher t . In (B) one observes a small hump. This behaviour agrees with the prospects about a mean reverting process with constant mean. After a certain period of time the system finds its equilibrium.

3.3.2 Moments of integrated stochastic correlation

Now we consider the moments $M_{(i,j)} := E [X_t^i Y_t^j]$ where

$$(3.34) \quad Y_t := \int_0^t X_s ds .$$

We already know various cases. If $i = j = 0$, one directly sees that $M_{(0,0)} = 0$. For $j = 0$ we can use the results of the previous section:

$$M_{(i,0)}(t) = m_i(t) .$$

As well one can directly observe that

$$M'_{(0,j)}(t) = jM_{(1,j-1)}$$

holds. Short calculation shows

$$M'_{(1,j)}(t) = jM_{(2,j-1)} + \kappa\theta M_{(0,j)} - \kappa M_{(1,j)} .$$

Therewith we can now state the stochastic differential equations for the remaining $M_{(i,j)}(t)$. Again we use Itô's formula:

$$\begin{aligned} d(X_t^i Y_t^j) &= X_t^i dY_t^j + Y_t^j dX_t^i \\ &= X_t^i j Y_t^{j-1} X_t dt + Y_t^j \left(i X_t^{i-1} dX_t + \frac{i(i-1)}{2} (dX_t)^2 \right) + \dots dW_t \\ &= j X_t^{i+1} Y_t^{j-1} dt + Y_t^j (a_i X_t^{i-2} + b_i X_t^{i-1} + c_i X_t^i) dt + \dots dW_t . \end{aligned}$$

As we are only interested in the expectation, the coefficient function of dW_t is of no significance and left out. Thus the raw (non-centralised) moments fulfill the following system of ordinary differential equations:

$$(3.35) \quad M'_{(i,j)}(t) = j \cdot M_{(i+1,j-1)}(t) + a_i M_{(i-2,j)}(t) + b_i M_{(i-1,j)}(t) + c_i M_{(i,j)}(t) ,$$

$$(3.36) \quad M_{(i,0)}(t) = m_i(t) ,$$

$$(3.37) \quad M'_{(0,j)}(t) = j M_{(1,j-1)} , j \geq 1$$

$$(3.38) \quad M'_{(1,j)}(t) = j M_{(2,j-1)} + \kappa\theta M_{(0,j)} - \kappa M_{(1,j)} , j \geq 1$$

$$(3.39) \quad M_{(i,j)}(0) = 0 , \quad \forall j > 0 .$$

Note that one recovers the moments of the integrated stochastic correlation by setting $i = 0$. This system can be solved recursively as illustrated in the following graph 3.4 for $i = 0, j = 3$:

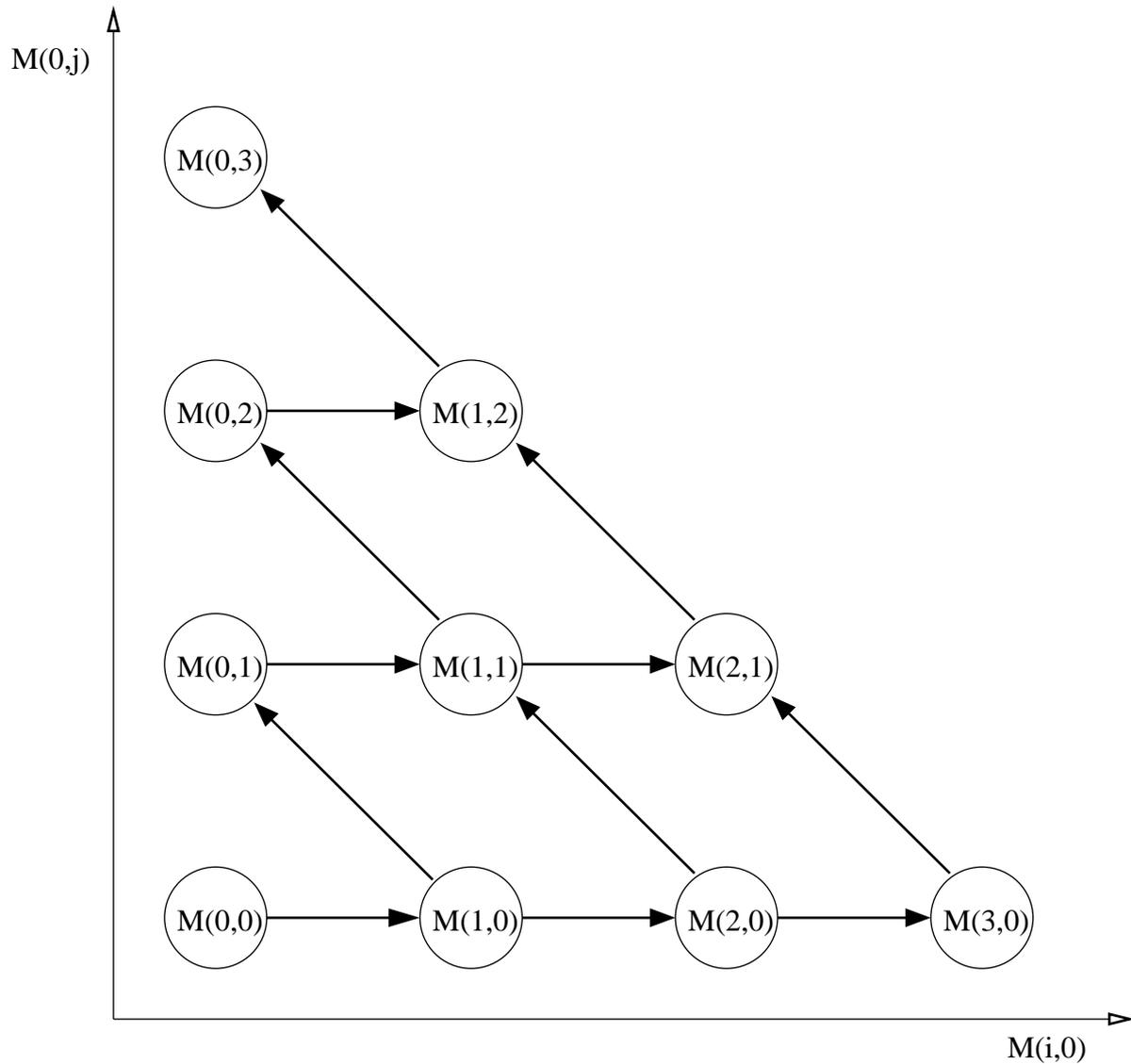


Figure 3.4: Recursive computation of moments of integrated correlation for $M_{(0,3)}$

As an example we present the solutions for the first non-centralised moment of integrated correlation

$$E \left[\int_0^t X_s ds \right] = \theta t + \frac{1}{\kappa} (X_0 - \theta) (1 - \exp(-\kappa t))$$

and the second:

$$\begin{aligned}
E \left[\left(\int_0^t X_s ds \right)^2 \right] &= \frac{1}{\kappa^2 (\omega^2 + \kappa)^2 (\omega^2 + 2\kappa)^2} e^{-t(\omega^2 + 3\kappa)} \\
&\cdot \left(2e^{t\kappa} ((X_0 - 1)(X_0 - m)\omega^4 + 2(X_0 - \theta)^2 \kappa^2 \right. \\
&+ (3X_0^2 - 2(\theta + 1)X_0 + \theta + m(-2X_0 + \theta + 1)) \kappa \omega^2) \kappa^2 \\
&- 2e^{t(\omega^2 + 2\kappa)} (\omega^2 + 2\kappa)^2 \cdot ((3\theta^2 - 2m\theta - 2X_0\theta - 2\theta \\
&+ m + mX_0 + X_0 + t(m - \theta + 1)(X_0 - \theta)\kappa) \omega^4 \\
&+ \kappa (X_0^2 - 2\theta X_0 + 2\theta^2 + m - m\theta - \theta + (m + 1)t(X_0 - \theta)\kappa) \omega^2 \\
&+ (X_0 - \theta)\kappa^2 (X_0 + \theta(t\kappa - 1))) \\
&+ e^{t(\omega^2 + 3\kappa)} (\omega^2 + \kappa)^2 ((t^2\theta^2\kappa^2 + 2t(m(\theta - 1) + (X_0 - 2\theta + 1)\theta)\kappa \\
&+ 2(3\theta^2 - 2(m + X_0 + 1)\theta + m + mX_0 + X_0)) \omega^4 \\
&+ 2\kappa(X_0^2 + (4t\kappa\theta - 6\theta + 2)X_0 + m(2X_0 - 5\theta + 2t(\theta - 1)\kappa + 3) \\
&\left. + \theta(2t\kappa + 2\theta(t\kappa(t\kappa - 3) + 4) - 5)) \omega^2 + 4\kappa^2 (X_0 + \theta(t\kappa - 1))^2 \right).
\end{aligned}$$

Again the solutions for higher moments become even more complicated, and their explicit presentations are omitted.

Figure 3.5 shows the first three centralised moments of realised correlation, which we obtain by scaling with time, for two different parameter configurations:

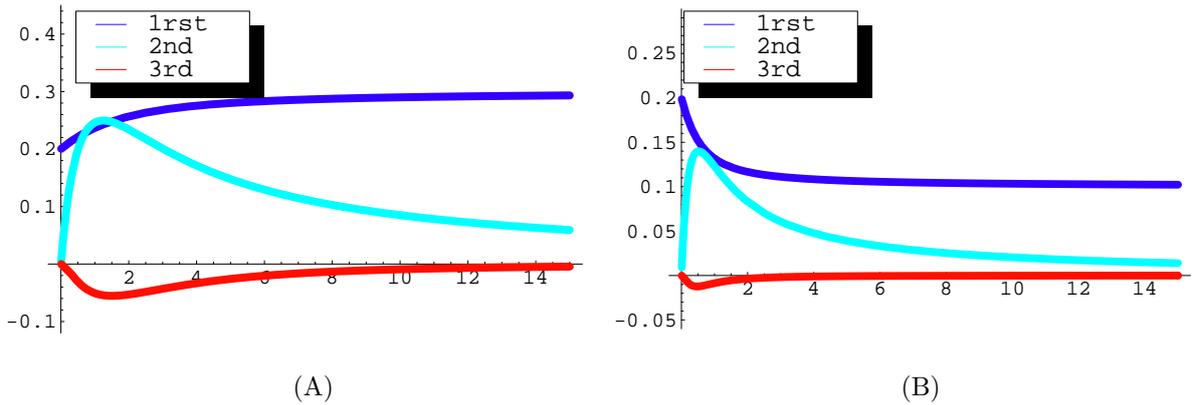


Figure 3.5: (A): First three centralised moments of realised correlation for $\kappa = 1, \theta = 0.3, \omega = 1.5, m = -1$ and $X_0 = 0.2$. (B): First three centralised moments of realised correlation for $\kappa = 3, \theta = 0.1, \omega = 1.7, m = -1$ and $X_0 = 0.2$.

Again the first moment starts at the initial value and converges to the long term mean θ . The development of second and third moment shows that uncertainty increases in the beginning. The underlying process diffuses from the initial value. This diffusing is stronger than the averaging effect of the integral. Uncertainty reaches a peak. Afterwards both moments decrease towards zero which is clear as the higher moments of the underlying process tend to a constant. The peak is reached around $t = \frac{1}{\kappa}$.

3.4 Summary

We could show that under the parameter restrictions in Thm. 3.2 the process given by (2.6) does not reach the boundaries. For these parameter configurations it was also possible to compute the transition density for $t \rightarrow \infty$. Further we could derive a system of ordinary differential equations for the correlation process as well as for the realised correlation. This knowledge will turn out to be important for the algorithm to calibrate the parameters presented in the following chapter.

CHAPTER IV

Maximum-Likelihood estimator

After having established in the previous chapter that (2.6) is suitable for modelling correlation for certain parameter configurations, there is still the problem of calibration. We noted that the transition density can be used for calibration. However this methodology only works under the assumption that correlation itself is observable.

In this chapter we do not need that assumption. Instead the presented algorithm bases on data which indeed can be observed: the realised correlation calculated from market data as in (1.13). In chapter 2 we already mentioned the similarities to the integrated correlation.

Here we derive a maximum-likelihood estimator which solely uses historical values of the time scaled integrated process (the realised correlation). As the density of the correlation integral is not known in closed form, the similarity to an Ornstein-Uhlenbeck process is used¹.

4.1 Overview

The aim is the historical calibration of the stochastic correlation process (2.6)

$$dX_t = \kappa(\theta - X_t)dt + \omega\sqrt{(1 - X_t)(X_t - m)}dW_t, \quad X_0 = \text{const}_1$$

to the observed realised correlation. Hereby we interpret the realised correlation as the average for a certain time period as in (2.7):

$$Y_t = \frac{1}{t} \int_0^t X_s ds.$$

The density of the process Y_t is not known analytically. However, since Y_t is an averaged quantity, we may hope that the approximation by the aid of a normal process leads to satisfactory results.

¹This chapter bases on cooperation with Valer Zetocha from Bear Stearns, London.

The Ornstein-Uhlenbeck process is a well known and analysed stochastic process with the same drift behaviour as the stochastic correlation process:

$$(4.1) \quad dx_t = a(b - x_t)dt + \sigma dV_t, \quad x_0 = \text{const}_2.$$

Moreover its integral is normally distributed for every t . For more information on the Ornstein-Uhlenbeck process the reader is referred to [UO30].

Here we want to make use of these similarities. At first we will compare normal densities to the densities of the integrated correlation process to get an impression how close the distribution of the stochastic correlation process is to the Ornstein-Uhlenbeck distribution for every t . Fig. 4.1 visualises the estimated densities of integrated stochastic correlation and densities from a normal distribution where standard deviation and mean are fitted to standard deviation and mean of the correlation density. The correlation density stems from a Monte Carlo simulation with the implicit Milstein scheme² and $M = 10000$ paths³.

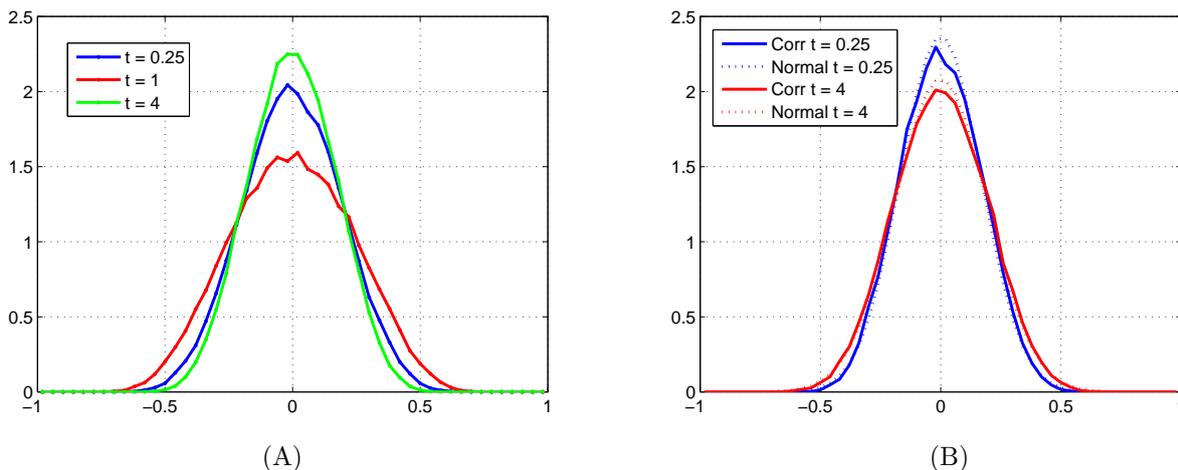


Figure 4.1: Simulated densities of $\frac{1}{t} \int X_s ds$ with $\kappa = 2, \theta = X_0 = 0, \omega = 0.5, m = -1$. (A): Varying time $t = 0.25, 1, 4$. (B): Simulated densities compared to fitted normal distributions for $t = \frac{1}{4}, t = 4$.

Fig. 4.1 (A) shows the density of realised correlation for different t . As already indicated by the behaviour of the second and third moment, observable in Fig. 3.5, for small t the density becomes wider and flatter. There is a point in time after which the density becomes more and more centered. Theoretically for $t \rightarrow \infty$ it converges to a Dirac measure.

Fig. 4.1(B) shows the fitted normal densities compared to the correlation densities for $t = \frac{1}{4}$ and $t = 4$. Both fits are rather close.

²For an explanation concerning the choice of the integration scheme see appendix A.1.

³The estimated densities in (A) and (B) stem from different Monte Carlo simulations. The difference between them is due to this comparatively low number of paths.

Fig. 4.2 gives the analogous information but for a correlation process with $m = -1$, which is not centered and therewith not symmetrically distributed in contrast to the normal distribution.

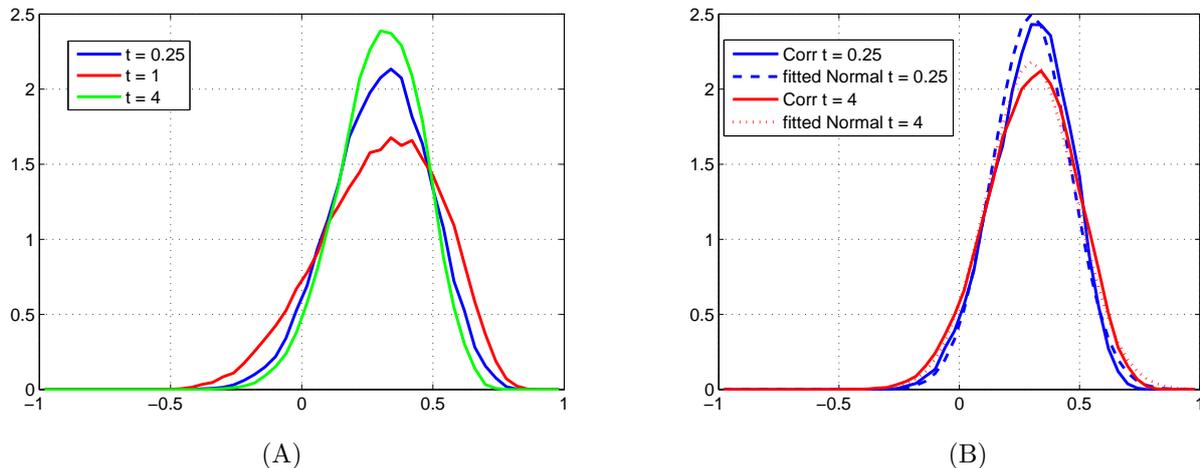


Figure 4.2: Simulated densities of realised correlation with $\kappa = 2, \theta = X_0 = 0.3, \omega = 0.5, m = -1$. (A): Varying time $t = 0.25, 1, 4$. (B): Simulated densities compared to fitted normal distributions for $t = 0.25, t = 4$.

As expected the fit worsens. But it is still rather good such that we can hope that the approximation via an Ornstein-Uhlenbeck process will work.

In the following we denote the average over the Ornstein-Uhlenbeck process (4.1) by

$$(4.2) \quad y_t = \frac{1}{t} \int_0^t x_s ds.$$

For this process the density can be calculated, see section 4.2. Here we would like to stress that the method of parameter calibration is constructed based on the values of the realised correlation (2.7). We assume that our given data only stems from the realisation of the integrals.

We will proceed as follows. First a maximum-likelihood estimator for the Ornstein-Uhlenbeck process is derived in section 4.2. No analytical features from the correlation process enter the procedure, such that it could also be used in a different context to purely calculate the parameters of (4.1). At the end of this section we will be able to estimate the parameters of the Ornstein-Uhlenbeck process. Afterwards the parameters of the correlation process (2.6) will be calibrated such that Ornstein-Uhlenbeck and correlation process are close.

4.2 Estimating the integral of an Ornstein-Uhlenbeck process

The method bases on $n + 1$ observations of the integral. They are denoted by $\bar{y}_0, \dots, \bar{y}_n$ for time points t_0, \dots, t_n . The length of the time steps is denoted by $\Delta_i = t_{i+1} - t_i$. Note that they do not need to be equidistant.

Following the idea of [Fra03], the likelihood function will eventually be maximised with respect to one variable, the rate of mean reversion. This simplifies and speeds up the optimisation considerably.

For the maximum likelihood estimator we need conditional expectation and conditional variance of the Ornstein-Uhlenbeck process (4.1).

The process x_t given in (4.1) possesses the following presentation

$$(4.3) \quad x_t = x_0 e^{-at} + b(1 - e^{-at}) + \sigma \int_0^t e^{a(s-t)} dV_s .$$

See for example [UO30]. Therewith we can compute the first two unconditional moments of the integrated process (4.2). For the expectation one can easily see:

$$(4.4) \quad E[y_t] = b + \frac{x_0 - b}{at} (1 - \exp(-at)) .$$

Of course this agrees with the results from chapter 3 where the moments of the correlation process are calculated since both processes only differ in the diffusion term. More tedious but straightforward calculations show that the variance is given by

$$(4.5) \quad Var[y_t] = \frac{1}{t^2} \frac{\sigma^2}{a^2} \left(t - \frac{3}{2a} + \frac{2 \exp(-at)}{a} - \frac{\exp(-2at)}{2a} \right) .$$

Therewith we can write up conditional expectation and variance of y_i when the integral y_{i-1} at t_{i-1} is given. For the first moment holds

$$\begin{aligned} E[y_i | \bar{y}_{i-1}] &= E \left[\frac{1}{t_i} \int_0^{t_{i-1}} x_s ds + \frac{1}{t_i} \Delta_{i-1} \frac{1}{\Delta_{i-1}} \int_{t_{i-1}}^{t_i} x_s ds \middle| \bar{y}_{i-1} \right] \\ &= \frac{t_{i-1}}{t_i} \bar{y}_{i-1} + \frac{1}{t_i} \Delta_{i-1} E[y_{\Delta_{i-1}}] \\ &= \frac{t_{i-1}}{t_i} \bar{y}_{i-1} + \frac{1}{t_i} \Delta_{i-1} b \left(1 - \frac{1}{a \Delta_{i-1}} (1 - \exp(-a \Delta_{i-1})) \right) \\ &\quad + \frac{1}{t_i} x_{i-1} \frac{1}{a} (1 - \exp(-a \Delta_{i-1})) , \end{aligned}$$

where y_{Δ_i} is defined as $y_{\Delta_i} = \frac{1}{\Delta_i} \int_{t_{i-1}}^{t_i} x_s ds$. The value x_{i-1} of the underlying process at t_{i-1} is not known since we can only observe the realisation of the integral at discrete

observation points. Thus we need a suitable approximation. We suggest the following for $i \geq 1$:

$$(4.6) \quad \begin{aligned} t_i y_i - t_{i-1} y_{i-1} &\approx \Delta_{i-1} (q x_i + (1-q)x_{i-1}) \\ \Leftrightarrow x_i &\approx \frac{1}{q} \left(\frac{1}{\Delta_{i-1}} (t_i y_i - t_{i-1} y_{i-1}) - (1-q)x_{i-1} \right) =: \tilde{x}_i, \end{aligned}$$

with $q \in [\frac{1}{2}, 1]$ and initial value $x_0 = \bar{y}_0$. This discretisation coincides with the so called θ -method. Here we substituted the traditional weighting variable θ with q , to avoid potential confusing with the level of mean reversion in the correlation process. For information on the θ -scheme see for example [GJ03]. For $q = 1$ it is fully implicit and numerically stable. It corresponds to an approximation through Riemann sums. If $q = \frac{1}{2}$ one obtains the trapezoidal rule

$$(4.7) \quad x_i \approx \frac{2}{\Delta_{i-1}} (t_{i-1} y_i - t_{i-1} y_{i-1}) - x_{i-1} \text{ with } x_0 = y_0.$$

It is advantageous concerning consistency, but at the border of numerical stability⁴. Unfortunately it often leads to oscillation in the likelihood function (4.10) below. The trapezoidal rule does not damp but preserve oscillation. Here we have stochastic input values, thus there is always noise or oscillation in every step. The trapezoidal rule preserves and sums up all those oscillations. See for example Hairer and Wanner [HW02], chapter 15. With a suitable choice of q is possible to avoid these oscillations and (4.6) turns out to be an appropriate approximation.

It may be possible to use other discretisation schemes as well, but one should use them with care because of the possible oscillation problem.

After having approximated the conditional expectation, we state the conditional variance:

$$(4.8) \quad Var[y_i | \bar{y}_{i-1}] = \frac{1}{t_i^2} \frac{\sigma^2}{a^2} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right),$$

which tends to zero for large t . Knowing that y_t is normally distributed with expectation μ and variance σ^2 , its density reads

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right).$$

We want to determine a, b, σ such that the probability for the independent realisations $\bar{y}_0, \dots, \bar{y}_n$ is maximised. The likelihood function $P(a, b, \sigma) = P(a, b, \sigma; \bar{y}_0, \dots, \bar{y}_n)$ for the realisations is given by

$$(4.9) \quad P(a, b, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi Var[y_i | \bar{y}_{i-1}]}} \exp\left(-\frac{1}{2} \frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{Var[y_i | \bar{y}_{i-1}]}\right).$$

Taking the logarithm of $P(a, b, \sigma)$ leads to

$$(4.10) \quad g(a, b, \sigma) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n \ln(Var[y_i | \bar{y}_{i-1}]) - \frac{1}{2} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{Var[y_i | \bar{y}_{i-1}]} \right)$$

⁴For $q \in [0, \frac{1}{2})$ specific constraints are necessary to guarantee stability.

Lemma 4.1 The function

$$g(a, b, \sigma) : \mathcal{L} := (0, \infty) \times (-\infty, \infty) \times (0, \infty) \longrightarrow \mathbb{R}$$

defined in (4.10) does not attain its global maximum at the boundary $\partial\mathcal{L}$.

Proof: We use $A := \frac{1}{2} \sum_{i=1}^n \ln(\text{Var}[y_i|\bar{y}_{i-1}])$ and $B := \frac{1}{2} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i|\bar{y}_{i-1}])^2}{\text{Var}[y_i|\bar{y}_{i-1}]} \right)$ as abbreviations in (4.10). Without loss of generality we assume that $\exists \bar{y}_{i_0}$ such that $\bar{y}_{i_0} \neq E[y_i|\bar{y}_{i-1}]$. We begin with the diffusion parameter σ . One directly observes

$$\text{Var}[y_i|\bar{y}_{i-1}] \longrightarrow \begin{cases} 0 & \text{for } \sigma \longrightarrow 0 \\ \infty & \text{for } \sigma \longrightarrow \infty. \end{cases}$$

For $\sigma \rightarrow \infty$, it holds that $A \rightarrow \infty$ and $B \rightarrow 0$. Thus

$$g(a, b, \sigma) \longrightarrow -\infty \quad \text{for } \sigma \rightarrow \infty.$$

For $\sigma \rightarrow 0$, it holds $A \rightarrow -\infty$ and $B \rightarrow \infty$. As $\frac{\ln(\sigma^2)}{\sigma^2} \rightarrow 0$, A is dominated by B and thus also in this case

$$g(a, b, \sigma) \longrightarrow -\infty \quad \text{for } \sigma \rightarrow 0.$$

holds. For the rate of mean reversion we note that

$$E[y_i|\bar{y}_{i-1}] \longrightarrow \begin{cases} \frac{t_{i-1}}{t_i} \bar{y}_{i-1} + \frac{\Delta_{i-1}}{t_i} \theta & \text{for } a \longrightarrow \infty \\ \frac{t_{i-1}}{t_i} \bar{y}_{i-1} + \frac{\Delta_{i-1}}{t_i} x_{i-1} & \text{for } a \longrightarrow 0. \end{cases}$$

The influence of a on the variance (which is non-negative by construction) is similar to the influence of σ , but the other way round:

$$\text{Var}[y_i|\bar{y}_{i-1}] \longrightarrow \begin{cases} \infty & \text{for } a \longrightarrow 0 \\ 0 & \text{for } a \longrightarrow \infty. \end{cases}$$

Therefore analogous argumentation as above leads to

$$g(a, b, \sigma) \longrightarrow -\infty \quad \text{for } a \rightarrow \{0, \infty\}.$$

If $a, \sigma \rightarrow \infty$ or $a, \sigma \rightarrow 0$ the limit behaviour is dominated by a see (4.8).

Concerning the dependence on b one only has to consider the expectation which depends linearly on the mean factor: $A \rightarrow \infty$. Thus

$$g(a, b, \sigma) \longrightarrow -\infty \quad \text{for } b \rightarrow \{-\infty, \infty\}.$$

If $b, \sigma \rightarrow \infty$, B converges to a constant. As $A \rightarrow \infty$:

$$g(a, b, \sigma) \longrightarrow -\infty$$

holds in this case. The same result is obtained if $b \rightarrow -\infty, \sigma \rightarrow 0$. Again A and B diverge, but the limit behaviour is dominated by B .

The results for the remaining combinations can be shown as easily. \square

The likelihood expression shall be maximised. In view of the previous lemma 4.1, we can calculate the derivatives and set them to zero:

$$\begin{aligned} 0 \stackrel{!}{=} g(a, b, \sigma)_b &= \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i|\bar{y}_{i-1}])}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \cdot E[y_i|\bar{y}_{i-1}]_b \\ &= \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i|\bar{y}_{i-1}])}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \cdot \frac{1}{t_i} \Delta_{i-1} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right). \end{aligned}$$

We want to derive an expression for b by rearranging this equation:

$$\begin{aligned} &\sum_{i=1}^n \left(\frac{\frac{1}{t_i} \Delta_{i-1} b \left(1 - \frac{(1 - \exp(-a\Delta_{i-1}))}{a\Delta_{i-1}} \right)}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \cdot \frac{\Delta_{i-1}}{t_i} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right) \\ &= \sum_{i=1}^n \left(\frac{\bar{y}_i - \frac{t_{i-1}}{t_i} \bar{y}_{i-1} - \frac{\bar{x}_{t_{i-1}}(1 - \exp(-a\Delta_{i-1}))}{t_i a}}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \cdot \frac{\Delta_{i-1}}{t_i} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right), \end{aligned}$$

which results in

$$\begin{aligned} b &= \frac{\sum_{i=1}^n \left(\frac{\bar{y}_i - \frac{t_{i-1}}{t_i} \bar{y}_{i-1} - \frac{\bar{x}_{t_{i-1}}(1 - \exp(-a\Delta_{i-1}))}{t_i a}}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \frac{\Delta_{i-1}}{t_i} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right)}{\sum_{i=1}^n \left(\frac{\frac{1}{t_i} \Delta_{i-1} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right)}{\text{Var}[y_i|\bar{y}_{i-1}]} \right) \frac{\Delta_{i-1}}{t_i} \left(1 - \frac{1}{a\Delta_{i-1}} (1 - \exp(-a\Delta_{i-1})) \right)} \\ (4.11) \quad &= \frac{\sum_{i=1}^n \left(\frac{\bar{y}_i - \frac{t_{i-1}}{t_i} \bar{y}_{i-1} - \frac{\bar{x}_{t_{i-1}}(1 - \exp(-a\Delta_{i-1}))}{t_i a}}{\frac{1}{t_i} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right)} \right) \left(\Delta_{i-1} - \frac{1}{a} (1 - e^{-a\Delta_{i-1}}) \right)}{\sum_{i=1}^n \left(\frac{\left(\Delta_{i-1} - \frac{1}{a} (1 - \exp(-a\Delta_{i-1})) \right)^2}{\left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{\kappa} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right)} \right)}. \end{aligned}$$

Note that this expression does not depend on the parameter σ , but only on the speed of mean reversion.

Analogously we compute the derivative with respect to the diffusion parameter σ . We need the derivative of the variance (4.8):

$$\text{Var}[y_i|\bar{y}_{i-1}]_\sigma = \frac{2}{t_i^2} \frac{\sigma}{a^2} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right).$$

Therewith we can calculate $g(a, b, \sigma)_\sigma$:

$$g(a, b, \sigma)_\sigma = -\frac{1}{2} \sum_{i=1}^n \frac{\text{Var}[y_i|\bar{y}_{i-1}]_\sigma}{\text{Var}[y_i|\bar{y}_{i-1}]} + \frac{1}{2} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i|\bar{y}_{i-1}])^2}{\text{Var}[y_i|\bar{y}_{i-1}]^2} \right) \cdot \text{Var}[y_i|\bar{y}_{i-1}]_\sigma.$$

In view of lemma 4.1 it must hold for the maximum value that $g(a, b, \sigma)_\sigma = 0$. We

obtain

$$\begin{aligned}
\sum_{i=1}^n \frac{2}{\sigma} &= \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{\text{Var}[y_i | \bar{y}_{i-1}]} \right) \cdot \frac{2}{\sigma} \\
\iff n &= \frac{1}{\sigma^2} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{\frac{1}{t_i^2} \frac{1}{a^2} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right)} \right) \\
\implies \sigma &= \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{\frac{1}{t_i^2} \frac{1}{a^2} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right)} \right)}.
\end{aligned}$$

To sum up we obtained an expression for σ which depends on the mean reversion factor a and the level of mean reversion b :

$$(4.12) \quad \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{(\bar{y}_i - E[y_i | \bar{y}_{i-1}])^2}{\frac{1}{t_i^2} \frac{1}{a^2} \left(\Delta_{i-1} - \frac{3}{2a} + \frac{2 \exp(-a\Delta_{i-1})}{a} - \frac{\exp(-2a\Delta_{i-1})}{2a} \right)} \right)}.$$

As b can be expressed in terms of a , independent on σ , see (4.11), we can conclude that we can also express σ as a function in the mean reversion factor. We denote those functions given by (4.11) and (4.12) by

$$(4.13) \quad b = h_b(a) \quad \text{and} \quad \sigma = h_\sigma(a).$$

By aid of these functions the maximisation of the likelihood function (4.10) reduces to a one-dimensional optimisation problem in a single variable a , the speed of mean reversion:

$$(4.14) \quad g(a) = g(a, h_b(a), h_\sigma(a)) \rightarrow \max$$

This problem can be solved with standard optimisation methods, see for example [NM02].

Now we are able to calibrate an Ornstein-Uhlenbeck process if historical data for the integral are given. In the next section we want to use that procedure for the calibration of the stochastic process (2.6) given in the introduction.

4.3 Fitting the stochastic correlation process

We use the observation that the integral y_t over the Ornstein-Uhlenbeck process (4.1)

$$dx_t = a(b - x_t)dt + \sigma dV_t$$

and the integral Y_t over the correlation process

$$dX_t = \kappa(\theta - X_t)dt + \omega \sqrt{(1 - X_t)(X_t - m)} dW_t,$$

are similarly distributed as indicated by Fig. 4.1 and 4.2. For the first moment of y_t holds

$$E[y_t] = b + \frac{x_0 - b}{at} (1 - \exp(-at)),$$

which is equal in expectation to the integrated correlation process (2.6)

$$E[z_t] = \theta + \frac{X_0 - \theta}{\kappa t} (1 - \exp(-\kappa t)).$$

In section 4.2, we computed that the variance of y_t is given as

$$\text{Var}[y_t] = \frac{\sigma^2}{a^2} \left(t - \frac{3}{2a} + \frac{2 \exp(-at)}{a} - \frac{\exp(-2at)}{2a} \right)$$

with the obvious parameter identification. One can also calculate the second moment of Y_t , see chapter 3:

$$\begin{aligned} E[(Y_t)^2] &= \frac{1}{\kappa^2 t^2 (\omega^2 + \kappa)^2 (\omega^2 + 2\kappa)^2} e^{-t(\omega^2 + 3\kappa)} \\ &\cdot \left(2e^{t\kappa} ((X_0 - 1)(X_0 - m)\omega^4 + 2(X_0 - \theta)^2 \kappa^2 \right. \\ &\quad + (3X_0^2 - 2(\theta + 1)X_0 + \theta + m(-2X_0 + \theta + 1)) \kappa \omega^2) \kappa^2 \\ &\quad - 2e^{t(\omega^2 + 2\kappa)} (\omega^2 + 2\kappa)^2 \cdot ((3\theta^2 - 2m\theta - 2X_0\theta - 2\theta \\ &\quad + m + mX_0 + X_0 + t(m - \theta + 1)(X_0 - \theta)\kappa)\omega^4 \\ &\quad + \kappa (X_0^2 - 2\theta X_0 + 2\theta^2 + m - m\theta - \theta + (m + 1)t(X_0 - \theta)\kappa) \omega^2 \\ &\quad + (X_0 - \theta)\kappa^2 (X_0 + \theta(t\kappa - 1))) \\ &\quad + e^{t(\omega^2 + 3\kappa)} (\omega^2 + \kappa)^2 ((t^2\theta^2\kappa^2 + 2t(m(\theta - 1) + (X_0 - 2\theta + 1)\theta)\kappa \\ &\quad + 2(3\theta^2 - 2(m + X_0 + 1)\theta + m + mX_0 + X_0))\omega^4 \\ &\quad + 2\kappa(X_0^2 + (4t\kappa\theta - 6\theta + 2)X_0 + m(2X_0 - 5\theta + 2t(\theta - 1)\kappa + 3) \\ &\quad \left. + \theta(2t\kappa + 2\theta(t\kappa(t\kappa - 3) + 4) - 5))\omega^2 + 4\kappa^2(X_0 + \theta(t\kappa - 1))^2 \right). \end{aligned}$$

Therewith we obtain

$$(\tilde{\kappa}, \tilde{\theta}, \tilde{\omega}) \in \mathcal{L}_2 := (0, \infty) \times (m, 1) \times (0, \infty).$$

as solution of the optimisation problem:

$$(4.15) \quad (\tilde{\kappa}, \tilde{\theta}, \tilde{\omega}) = \{ \kappa, \theta, \omega \in \mathcal{L}_2 \mid \|F\| \text{ minimal} \}.$$

The function F to minimise can be chosen according to different criteria which are of particular importance in the given application. Using the abbreviation $K[\theta, \kappa, \omega] =$

$E[(X_t - E[X_t])^3]$ for the third centralised moment we suggest⁵

$$F = \alpha_1 \cdot |E_{OU} - E_{Corr}[\kappa, \theta]| + \alpha_2 \cdot |Var_{OU} - Var_{Corr}[\kappa, \theta, \sigma]| \\ + \alpha_3 \cdot |K[\theta, \kappa, \omega]|$$

to fit the first three centralised moments. The third moment of the centralised Ornstein-Uhlenbeck is zero because of its normal distribution. The weighting factors $\alpha_i > 0, i = 1, 2, 3$ should hereby sum up to 1. In this notation the focus is on one (arbitrary) point in time, but naturally different horizons should be considered to make the solution unique. The actual choice of the term structure t as well as the number clearly influences the quality of the estimation.

The concept can be extended to the lower bound m of the correlation process, by including m in the optimisation problem (4.15), i. e. E_{Corr} , Var_{Corr} and K would then depend on κ, θ, ω and m .

4.4 Numerical tests

This section is subdivided into two parts. First we show numerical results for the calibration of the Ornstein-Uhlenbeck process. Afterwards the estimator for the correlation process is tested.

4.4.1 Numerical test for calibration of the Ornstein-Uhlenbeck process

For the numerical tests, we simulate the integral (4.2) over the stochastic process (4.1). We do not use a discretisation scheme but use the known density, i. e. we directly draw from its normal distribution. The presented results base on $M = 500$ paths, simulated from $t = 0$ until $t = 15$ with time discretisation $\Delta_t = \frac{1}{200}$, corresponding to 15 years with 200 observation dates each year. Every Δ_t is further discretised with $H = 4$ equidistant discretisation steps⁶. We consider four time periods:

- $\mathcal{T}_1 = \{t_0, \dots, t_{200}\}$: 1 year;
- $\mathcal{T}_2 = \{t_0, t_1, \dots, t_{1000}\}$: 5 years;
- $\mathcal{T}_3 = \{t_0, t_1, \dots, t_{2000}\}$: 10 years;
- $\mathcal{T}_4 = \{t_0, t_1, \dots, t_{3000}\}$: 15 years.

⁵Here we use intuitive abbreviations: E_{OU} and Var_{OU} denote the quantities for expectation and variance resulting from fitting the realisations to an integrated Ornstein-Uhlenbeck process. $E_{Corr}[\cdot, \cdot, \cdot]$ and $Var_{Corr}[\cdot, \cdot, \cdot]$ denote expectation and variance of realised correlation as functions of κ, θ and ω .

⁶We restricted the numerical tests to only $M = 500$ paths because of computation time. An increase is surely sensible. It would make the results more reliable. Obviously the distributions in Fig. 4.3-4.5 would become smoother.

We computed an analytical expression (4.11) for the mean which depends on a . Although theoretically justified this expression sometimes leads to numerical problems. In particular it slows down the computation. This problem is enforced because \tilde{b} enters the Maximum-Likelihood procedure twice, as the estimator of the level of mean reversion and as an input parameter for the diffusion parameter σ . Therefore from a practical point of view it would be nice, to make the computation of \tilde{b} more stable and faster.

Usually the estimation of the level of mean reversion does not pose a real problem. It can usually be fixed rather closely from historical observation or be fitted to market expectation. Therefore we include in the numerical tests as an alternative the estimation of b as the mean over the realisations:

$$(4.16) \quad \hat{h}_b = \frac{1}{n} \sum_{i=1}^n \bar{y}_i .$$

Note that this estimator also influences the estimation of a and σ .

For (4.11) and (4.16) we present the calibration results for the parameters a , b and σ . The input parameters for the Monte Carlo simulation are chosen as

$$(4.17) \quad a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4.$$

Table 4.1 shows the results when b is estimated via the Maximum-Likelihood expression (4.11). For every simulated path the parameters are estimated. The table shows the percentage of paths leading to estimation results in the respective integral. For a the intervals are $[3.1, 3.9]$, $[2.7, 4.3]$ and $[2.3, 4.7]$, corresponding to deviations of ± 0.4 , ± 0.8 and ± 1.2 from the input parameter $a = 3.5$. In case of b they are chosen as $[0.35, 0.45]$, $[0.3, 0.5]$ and $[0.25, 0.55]$. The intervals for σ are set to $[0.79, 0.81]$, $[0.78, 0.82]$ and $[0.77, 0.83]$.

	a			b			σ		
	± 0.4	± 0.8	± 1.2	± 0.05	± 0.1	± 0.15	± 0.01	± 0.02	± 0.03
\mathcal{T}_1	3%	7%	12%	20%	37%	52%	21%	37%	53%
\mathcal{T}_2	16%	33%	46%	34%	62%	80%	39%	67%	85%
\mathcal{T}_3	26%	50%	66%	43%	75%	94%	47%	78%	95%
\mathcal{T}_4	31%	60%	81%	54%	86%	98%	49%	86%	99%

Table 4.1: Percentage of paths leading to correct parameter estimation for the Ornstein-Uhlenbeck process with $a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4$ and $q = 0.8$. Level of mean reversion calculated via Maximum-Likelihood estimator.

As expected the estimators improve, the longer the observation period is. But one notices that the estimation of the rate of mean reversion does not lead to precise estimation results. To this respect the estimator can only give an idea about the magnitude of a . In contrast to that, the results for b are rather good, for σ they are even highly accurate. Thus the maximum-likelihood estimator for level of mean reversion and diffusion can compensate for some of the weaknesses in the estimation of the speed of mean reversion.

As the latter is an input parameter it does influence the quality of the estimation for b and σ . This is clearly observable, when we compare Fig. 4.3 and Fig. 4.5.

In table 4.2 the corresponding results are presented when the level of mean reversion is estimated as the average of the integral realisations:

	a			b			σ		
	± 0.4	± 0.8	± 1.2	± 0.05	± 0.1	± 0.15	± 0.01	± 0.02	± 0.03
\mathcal{T}_1	6%	12%	18%	28%	54%	74%	21%	37%	52%
\mathcal{T}_2	19%	35%	49%	31%	60%	78%	39%	67%	85%
\mathcal{T}_3	26%	49%	67%	40%	71%	90%	46%	77%	94%
\mathcal{T}_4	36%	61%	80%	45%	78%	94%	49%	85%	98%

Table 4.2: Percentage of paths leading to correct parameter estimation for the Ornstein-Uhlenbeck process with $a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4$ and $q = 0.8$. Level of mean reversion calculated as average.

The results are similar, but one observes that averaging via (4.16) leads to slightly better results, especially for shorter time frames. As the Maximum-Likelihood estimator for the mean does not seem to be clearly superior to the average, we are going to use the average for the remaining tests, as it is computationally faster.

Note that both estimators for b seem to be unbiased as we can see from Fig. 4.3.

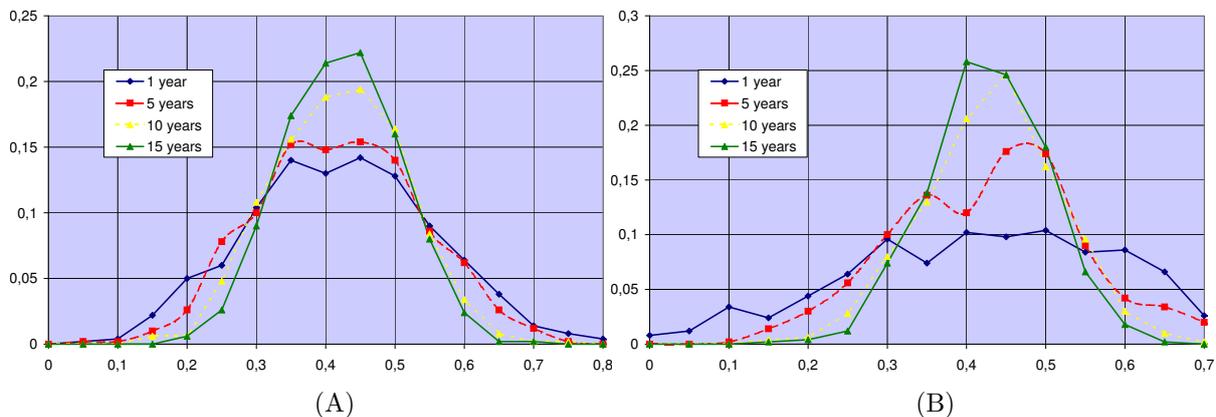


Figure 4.3: Distribution of estimation results for mean with $a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4, q = 0.8$. (A): Mean as average (4.16). (B): Mean via Maximum-Likelihood estimator (4.11).

Also the estimation of σ seems to be unbiased, see table 4.4. One can observe that both graphs are very similar. Thus we can record that the estimation of σ seems to depend only weakly on the estimation of b .

In contrast to the unbiased calibration of b and σ there is a clear bias in the estimation of the speed of mean reversion. The maximum-likelihood estimator tends to overestimate a . The bias decreases for increasing number of observations. Here one also sees that

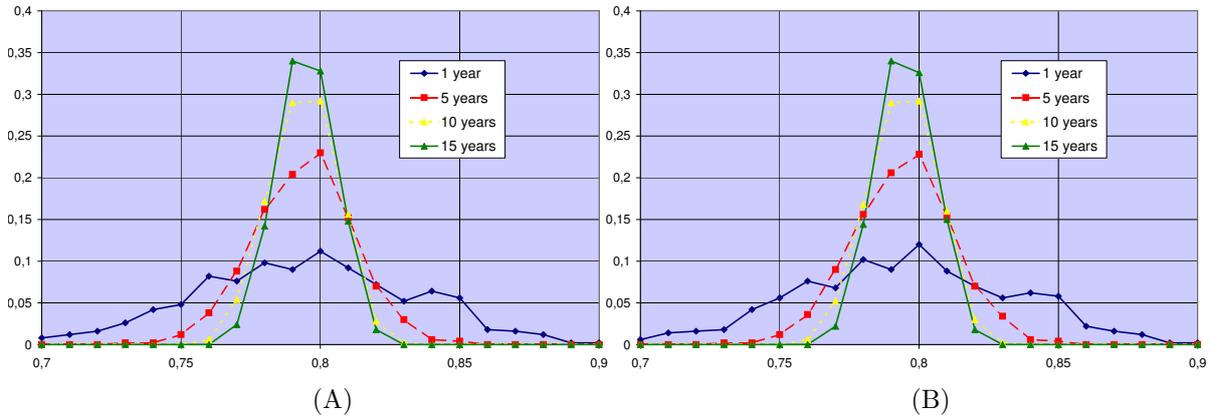


Figure 4.4: Distribution of estimation results for σ with $a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4, q = 0.8$. (A): Mean as average (4.16). (B): Mean via Maximum-Likelihood estimator (4.11).

$M = 500$ paths are not sufficient for making a precise statement on the distribution of the results.

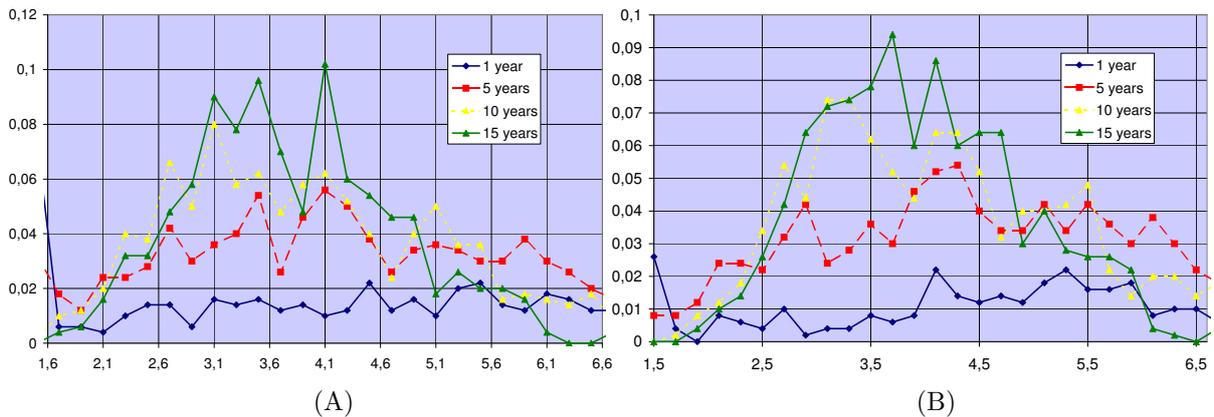


Figure 4.5: Distribution of estimation results for a with $a = 3.5, b = 0.4, \sigma = 0.8, x_0 = 0.4, q = 0.8$. (A): Mean as average (4.16). (B): Mean via Maximum-Likelihood estimator (4.11).

In Fig. 4.3-4.5 one observes a steepening in the distribution of the estimates through time. This steepening is rather slow in case of the mean-reversion factor, resulting in unprecise estimation results. But it gives hope that increasing the amount of input data improves the estimation results. This could either be done by prolonging the observation period. Although it is questionable if there are many fields of application where more than 15 years of data are available and meaningful. Refining the observation stepsize could turn out to be successful.

For comparison we show the results for a second parameter configuration:

$$(4.18) \quad a = 2, b = 0, \sigma = 0.6, x_0 = 0 .$$

Table 4.3 shows the estimation results. Hereby we estimated the level of mean reversion as the average:

	a			b			σ		
	± 0.4	± 0.8	± 1.2	± 0.1	± 0.15	± 0.2	± 0.01	± 0.02	± 0.03
\mathcal{T}_1	7%	12%	18%	33%	60%	78%	26%	50%	69%
\mathcal{T}_2	23%	39%	53%	34%	60%	79%	55%	84%	97%
\mathcal{T}_3	36%	60%	77%	38%	66%	85%	71%	96%	100%
\mathcal{T}_4	44%	72%	87%	44%	73%	91%	70%	97%	100%

Table 4.3: Percentage of paths leading to correct parameter estimation for the Ornstein-Uhlenbeck process with $a = 2, b = 0, \sigma = 0.6, x_0 = 0, q = 0.8$. Level of mean reversion calculated as average.

The parameter configuration given as in (4.17) seems to be easier to calibrate with the Maximum-Likelihood method. Particularly the results for the rate of mean reversion have improved.

4.4.2 Numerical results for calibration of the correlation process

Instead of simulating the Ornstein-Uhlenbeck integral as before, we now simulate the correlation process (2.6) and then build the integral accordingly. We consider the same time horizons as before: 1 year, 5, 10 and 15 years. The parameters are chosen as

$$\kappa = 3.5, \theta = 0.4, \omega = 0.8, m = -0.5, x_0 = 0.4.$$

The simulation uses the implicit Milstein scheme⁷. Each $\Delta_t = \frac{1}{200}$ is again discretised with $H = 4$ equidistant discretisation steps. None of the paths exceeded the boundaries. We only use the simplified way of calculating the mean θ . For the approximation (4.6) we used $q = 0.71$. Note that the intervals for the diffusion parameter have been adapted. As in case of the Ornstein-Uhlenbeck process the Maximum-Likelihood method can only

	κ			θ			ω		
	± 0.4	± 0.8	± 1.2	± 0.05	± 0.1	± 0.15	± 0.02	± 0.04	± 0.06
\mathcal{T}_1	10%	17%	23%	40%	66%	85%	14%	27%	39%
\mathcal{T}_2	19%	31%	42%	48%	80%	94%	25%	47%	59%
\mathcal{T}_3	25%	44%	61%	59%	91%	99%	36%	59%	70%
\mathcal{T}_4	32%	60%	73%	67%	95%	100%	41%	66%	75%

Table 4.4: Percentage of paths leading to correct parameter estimation for correlation process with $\kappa = 3.5, \theta = 0.4, \omega = 0.8, m = -0.5$. $H = 4$ and $q = 0.71$. Mean is estimated as average.

give a rough estimate for the mean-reversion κ . This was to be expected. The level of mean reversion is estimated as well as before. The results for ω are worse, than before but still rather accurate.

⁷See appendix for an explanation of the choice of this integration scheme.

4.5 Summary

In this chapter a Maximum-Likelihood estimator for the parameters of an Ornstein-Uhlenbeck process (4.1) was developed, which employs data from the integrated Ornstein-Uhlenbeck process, only. It was possible to reduce the optimisation of the likelihood function to an optimisation problem in one variable. The relationship to the remaining parameters is given analytically. Applying this estimation procedure to the integrated correlation process, it became possible to calibrate the correlation process (2.6).

The estimation of κ and ω as the stochasticity parameters often poses problems. We observe this phenomenon here as well in case of the speed of mean-reversion. It is fair to say that the results are not very accurate. On the other hand the diffusion factor could be estimated with rather high precision. This holds particularly for the Ornstein-Uhlenbeck process, but also for the correlation process.

Application

The theoretical results presented in the previous chapters are of no practical use, as long as we cannot incorporate stochastic correlation into modelling/pricing. In section 2.2 we demonstrated how two Brownian motions can be stochastically correlated.

The main motivation given in chapter 1 was the correlation swap. If we assume the easiest case of identically correlated (geometric) Brownian motions representing the assets, the valuation of the correlation swap on a basket of these assets works straightforward via simulating of the integrated correlation process.

But the concept of stochastic correlation can be used in a broader sense. We present two examples which show two structurally different areas of application.

In section 5.1 two geometric Brownian motions are correlated via the full stochastic process (2.6). The two geometric Brownian motions hereby build a foreign exchange model as presented in chapter 2.

The second example presented in section 5.2 comes from credit modelling where it can be used for pricing collateralised debt obligations. The model is static, thus we only include correlation as a random variable which is drawn from the stationary density (3.25).

5.1 Foreign exchange model

A quanto is a cross-currency option. It is a combination of a regular option (European, American, Asian etc.) which pays off in currency A and a currency option which hedges against the exchange risk between A and another currency B. The quanto pays off in B. Thus the payoff is defined with respect to an underlying noted in currency A. Then this payoff is converted to currency B.

As an example we can think of a call on the Dow Jones whose payoff is paid in Euro

$$(Dow_T - Strike)^+ \cdot exchange - rate_0 .$$

For pricing a quanto option a model is needed which takes into account asset, exchange rate and the correlation between them. An easy example of such a model is the following foreign exchange (FX) model. We assume a Black-Scholes world with lognormal

distribution for the underlying asset

$$dS_t = \mu_S S_t dt + \sigma_S S_t dW_t^S, \quad S_0 = s_0$$

and lognormally distributed exchange rate

$$dR_t = \mu_R R_t dt + \sigma_R R_t dW_t^R, \quad R_0 = r_0.$$

For simplification we will price a call in this framework. If the correlation ρ between W^S and W^R is constant,

$$dW_t^S dW_t^R = \rho dt,$$

an analytical solution for pricing a call exists. It turns out that pricing a call in the FX model reduces to pricing a call in the original Black-Scholes setting with the adjusted constant dividend yield

$$r_f - r_h + \rho \sigma_S \sigma_R.$$

See for example Wilmott[Wil00a], chapter 11. Hereby r_f denotes the risk-free interest rate in the currency in which S is traded (foreign currency). The risk-free interest rate of the other currency is r_h (home).

The remaining question is how to incorporate stochastic correlation X_t in this framework. We assume the model to be arbitrage free and complete. Thus the expected return of one unit of home currency, exchanged, risk-free invested in the foreign country and re-exchanged must equal the risk-free return on one unit:

$$(5.1) \quad \frac{1}{R_0} \exp(r_f T) E[R_T] = \exp(r_h T).$$

The exchange rate R_t follows a geometric Brownian motion and thus $E[R_T] = R_0 \exp(\mu_R T)$. Insertion in (5.1) yields the restriction

$$(5.2) \quad \mu_R = r_h - r_f.$$

In addition the arbitrage argument leads to a second restriction. Analogous argumentation as above makes clear that

$$(5.3) \quad \frac{1}{R_0} \frac{1}{S_0} E[S_T R_T] = \exp(r_h T).$$

must hold. Hereby the left side describes the re-exchanged expectation of an investment of one home currency unit into the underlying S .

For calculating $E[S(T)R(T)]$ we need to compute

$$\begin{aligned} d(S_t R_t) &= S_t dR_t + R_t dS_t + dS_t dR_t \\ &= S_t R_t [(\mu_S + \mu_R)dt + \sigma_S dW_t^S + \sigma_R dW_t^R + \rho_t \sigma_S \sigma_R dt]. \end{aligned}$$

Itô's formula implies

$$d(\ln(x_t)) = \frac{1}{x_t} dx_t - \frac{1}{2} \frac{1}{x_t^2} (dx_t)^2.$$

Application to $S_t R_t$ leads to

$$\int_0^T \frac{d(S_t R_t)}{S_t R_t} = (\mu_S + \mu_R)T + \sigma_S W_T^S + \sigma_R W_T^R - \frac{1}{2} ((\sigma_S^2 + \sigma_R^2) T) .$$

A further application of Itô's formula leads to

$$(5.4) \quad E[S_T R_T] = S_0 R_0 E \left[\exp \left((\mu_S + \mu_R)T + \sigma_S \sigma_R \int_0^T X_t dt \right) \right] .$$

Thus with the choice

$$(5.5) \quad \mu_S = r_h - \mu_R - \sigma_S \sigma_R \frac{1}{T} \int_0^T X_t dt ,$$

the no-arbitrage condition (5.3) is fulfilled. Note that the choice

$$\mu_S = r_h - \mu_R - \frac{1}{T} \ln E \left[\exp \left(\sigma_S \sigma_R \int_0^T X_t dt \right) \right] ,$$

would also be consistent with (5.3), but in that case the stochasticity would have been lost.

The idea is to use a conditional Monte Carlo approach [Gla03] to price the FX option with stochastic correlation. Therefore let us denote the Black-Scholes formula with continuous dividend yield with

$$(5.6) \quad \text{Black}(S_0, K, r, D, \sigma, T) = P(0, T) (S_0 e^{(r-D)T} \mathcal{N}(d_1) - K \mathcal{N}(d_2))$$

where \mathcal{N} denotes the cumulative distribution function of the normal density, $P(0, T)$ is today's value of a zero coupon bond expiring at T , r the risk free interest rate and D the continuous dividend yield. Interpreting (5.5) as return minus continuous dividend

$$(5.7) \quad D(X_t) = r_h - (r_h - r_f + \sigma_S \sigma_R \frac{1}{T} \int_0^T X_t dt) ,$$

the fair value of the FX option is given by

$$(5.8) \quad FX = E \left[E \left[\text{Black}(S_0, K, r_h, D(X_t), \sigma, T) \mid \mathcal{F}(X_{\{0 \leq s \leq t\}}) \right] \right]$$

where \mathcal{F} is the filtration given the full information about the correlation process X_t . Note that the conditional Monte Carlo formula (5.8) is the same approach as the one used by Hull and White [HW87] to compute the option price in a stochastic volatility framework. Thus we can approximate the integral and obtain a constant expression for μ_S for every path. This allows to simply use the Black-Scholes price for stocks with continuous dividend payments (5.6) and get a solution for every path. Afterwards we compute the fair price of the quanto as the mean over all Black-Scholes prices. The formula (5.8) further allows to make a qualitative statement about the impact of stochastic correlation.

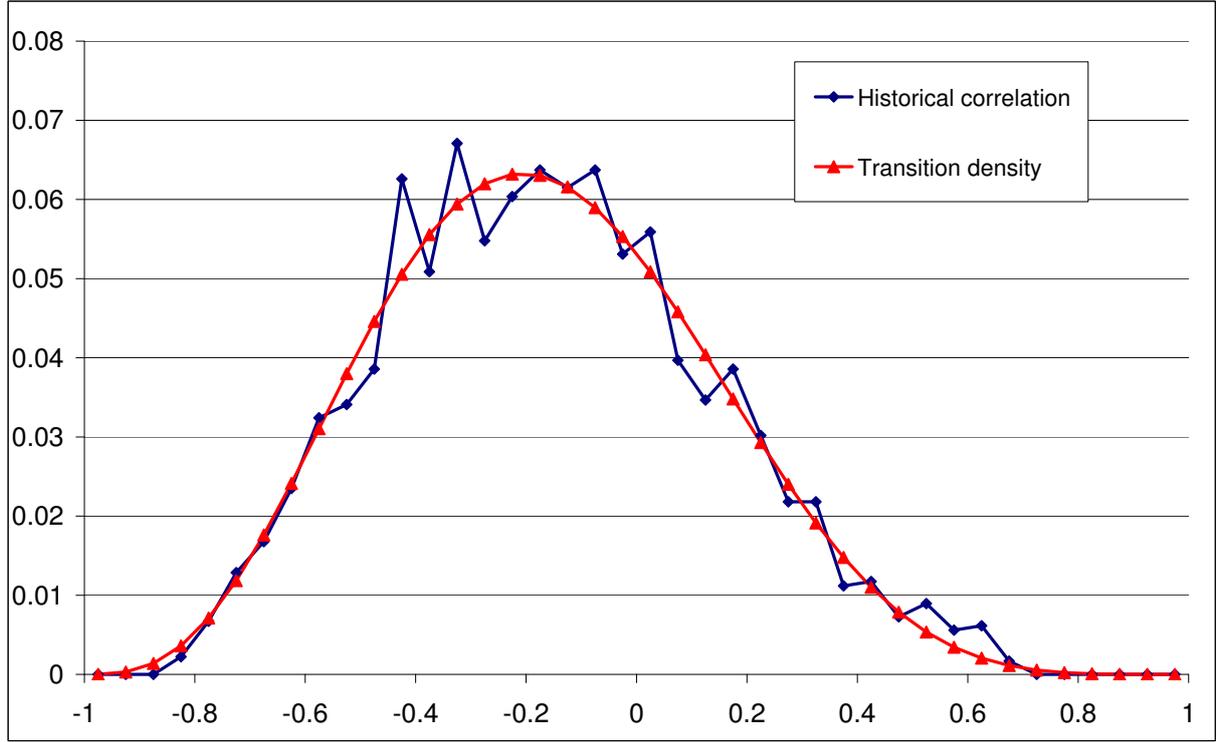


Figure 5.1: Correlation between Dow Jones and Euro/US-Dollar exchange rate; empirical distribution vs. density (3.21) computed with $\kappa = 10.6, \theta = -0.1, \omega = 1, m = -1$.

Lemma 5.1 Let us assume that $K = S_0$ then

$$(5.9) \quad E [\text{Black}(D(X_t)) | \mathcal{F}(X_{\{0 \leq s \leq t\}})] \geq \text{Black}(E [E [D(X_t) | \mathcal{F}(X_{\{0 \leq s \leq t\}})])],$$

where Black denotes the Black-Scholes formula (5.6). We omitted some of the parameters.

Proof: Formula (5.9) directly follows from the fact, that the Black-Scholes formula with continuous dividend yield is convex with respect to the dividend and applying Jensen's inequality for conditional expectations twice¹. In order to prove the convexity we use the formula by Wilmott [Wil00a, p. 127]

$$(5.10) \quad \frac{\partial}{\partial D} \text{Black}(S_0, S_0, r, D, \sigma, T) = -TS_0 e^{-DT} \mathcal{N}(d_1),$$

for the first derivative and a subsequent differentiation yields

$$(5.11) \quad \begin{aligned} \frac{\partial^2}{\partial D^2} \text{Black}(S_0, S_0, r, D, \sigma, T) &= T^2 S_0 e^{-DT} \mathcal{N}(d_1) - TS_0 e^{-DT} \frac{\partial}{\partial D} \mathcal{N}(d_1) \\ &= T^2 S_0 e^{-DT} \mathcal{N}(d_1) + TS_0 e^{-DT} \frac{\sqrt{T}}{\sigma} \phi(d_1) > 0 \end{aligned}$$

where ϕ is the density of the normal distribution. Note that both terms in (5.11) are positive. \square

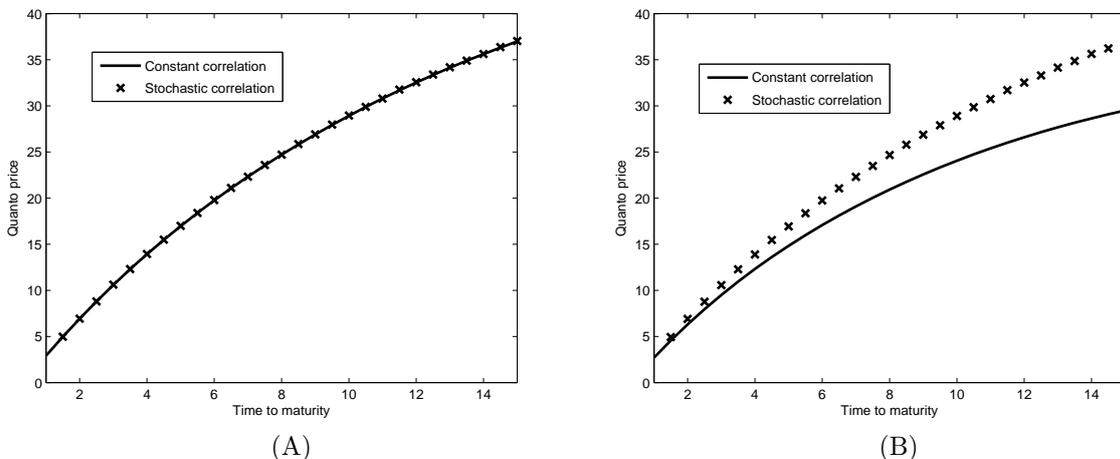


Figure 5.2: Call price with fixed and stochastic correlation with different initial values. Parameters: $\rho = -0.1$ and $\kappa = 10.6$, $\theta = -0.1$, $\omega = 1$ and $m = -1$. (A): $X_0 = -0.1$, (B): $X_0 = 0$.

Next we put theory into practice and test if stochastic correlation leads to different prices for calls compared to the constant correlation case. We consider the process (2.6) with $m = -1$ for stochastic correlation

$$dX_t = \kappa \cdot (\theta - X_t)dt + \omega \cdot \sqrt{1 - X_t^2}dW_t, \quad X_0 = x_0.$$

As above W denotes a Brownian motion. If not mentioned otherwise $\theta = -0.1$, $\kappa = 10.6$, $\omega = 1$. The parameter choice results from a least-square fitting of (2.6) to the historical data from figure 2.1. This fit is surprisingly good, see figure 5.1.

Underlying and exchange rate are supposed to follow lognormal distributions with $\sigma_S = 0.2$ and $\sigma_R = 0.4$, respectively. The risk free interest rates amounts to $r_h = 0.05$ and $r_f = 0.03$. The underlying starts in 100, the starting value of the exchange rate is 1. The strike amounts to 120. For numerical integration the implied Milstein scheme is used. The number of paths amounts to 10000, and the step size is set to 0.001. In none of the cases the boundaries $\ell = -1$ and $r = 1$ have been exceeded thus we consider the Milstein scheme as an appropriate choice. This effect perfectly agrees with the result that the boundaries are not attractive for this parameter configuration.

Figures 5.2 and 5.3 show prices computed for constant correlation $\rho = \theta$ using the analytic formula (continuous line). The crosses show the prices determined by simulation using our correlation process (2.6).

In figure 5.2 (A) there is no significant difference between the prices calculated with constant correlation and the prices calculated with (5.5) where X_t follows (2.6). The reason is that we do not use the whole path of correlation but eventually only the distribution of the integral at T . One can easily verify that with $X_0 = \theta$ the Milstein scheme generates random variables at T with mean θ . As the mean-reverting factor κ is relatively high this distribution is strongly concentrated around the mean. In contrast to that, in figure 5.2 (B) we can observe a difference: The prices for stochastic correlation are higher as for

¹See Øksendal [Øks00] for Jensen's inequality for conditional expectations.

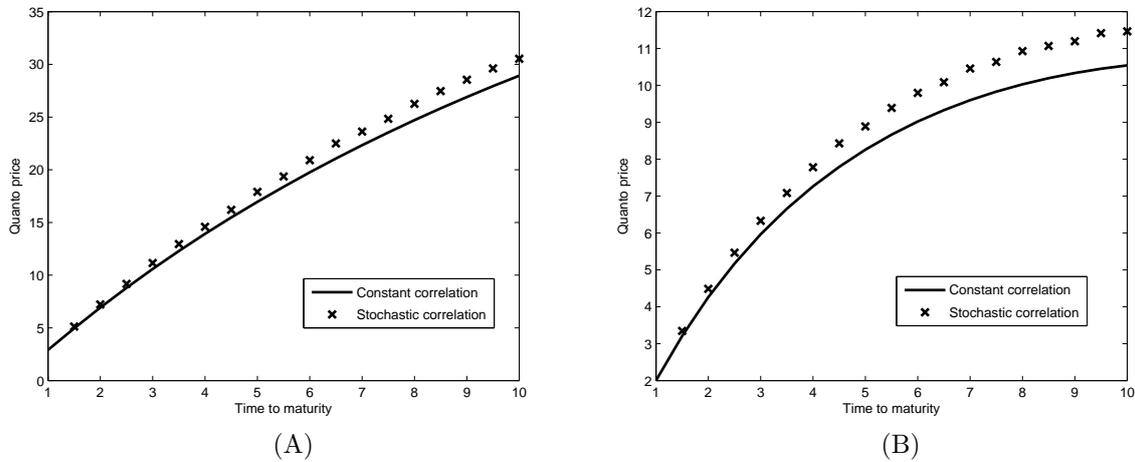


Figure 5.3: Prices for call with fixed and stochastic correlation for different level of correlation. (A): $X_0 = -0.1, \theta = -0.1, \kappa = 1.06, \omega = 1$ and $\rho = -0.1$, (B): $X_0 = 0.4, \theta = 0.4, \kappa = 1.06, \omega = 1$ and $\rho = 0.4$.

constant correlation. This is due to the fact that the expectation of correlation changes over time. Consequently the additional freedom we win by using stochastic processes for modelling correlation does have an influence on pricing.

In figure 5.3 (A) initial value and mean are equal but here the mean reverting factor κ is smaller compared to figure 5.2. We observe that stochastic correlation leads to higher prices than constant correlation although the expectation is equal. Thus pricing with constant correlation means neglecting correlation risk. The effect is even stronger if θ is higher as seen in figure 5.3 (B).

The first figures made clear that stochastic correlation leads to higher prices, and thus changes the model structurally. Fig. 5.4 gives even more insight how the model is changed.

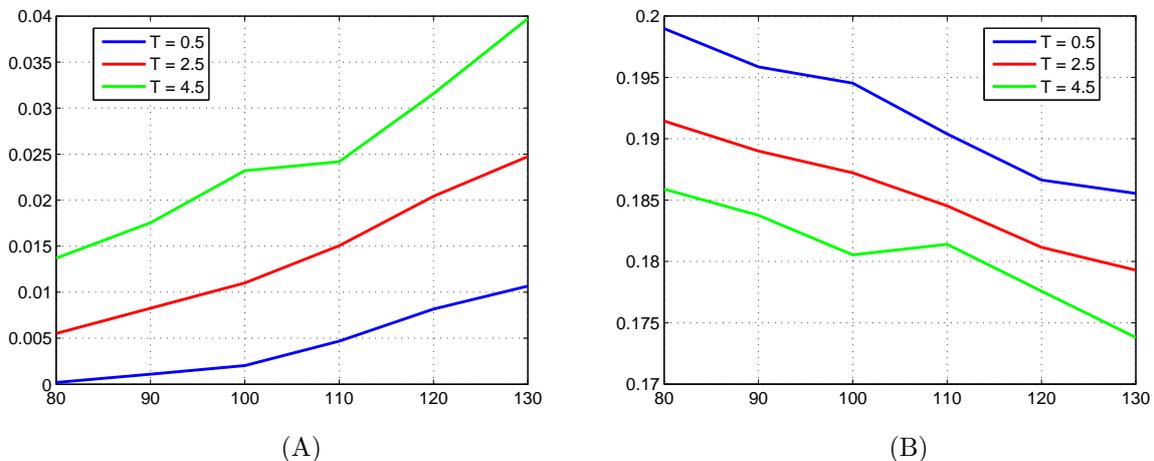


Figure 5.4: FX model with $X_0 = -0.1 = \theta = 0.2, \kappa = 10.6, \omega = 1, m = -1, M = 30000$. (A) Relative price difference. (B): Implied correlation.

Relative price difference between non-stochastic and stochastic correlation as well as implied correlation are presented. Implied correlation denotes the constant correlation we must insert in the FX model to obtain the price calculated with stochastic correlation where all other parameters remain the same. One observes a clear skew in the implied correlation. The skew is even more pronounced if the speed of mean reversion is decreased as in figure 5.5 where κ is chosen as 1.06. This behaviour is expected as the higher the speed of mean reversion is the closer stochastic correlation comes to constant correlation.

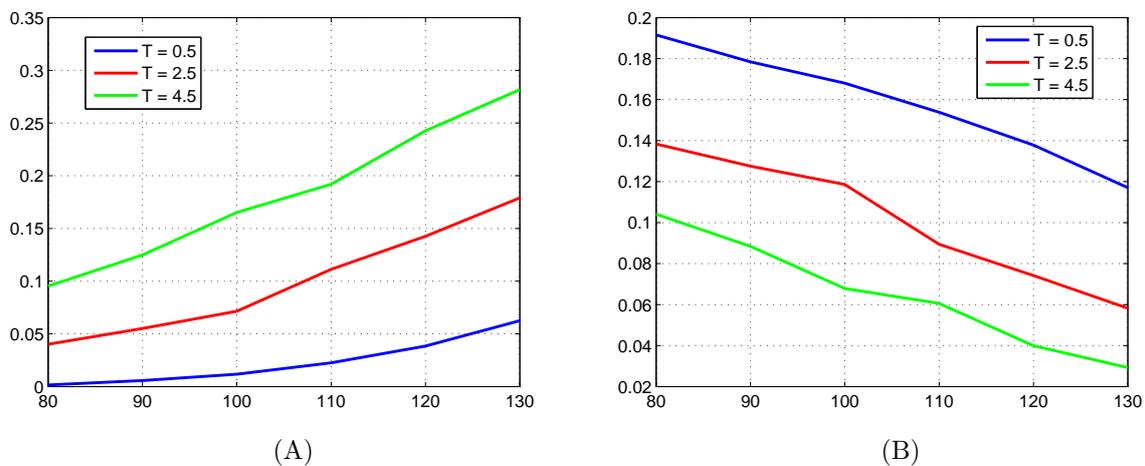


Figure 5.5: FX model with $X_0 = -0.1, \theta = 0.2, \kappa = 1.06, \omega = 1, m = -1, M = 30000$. (A) Relative price difference. (B): Implied correlation.

5.2 Large homogeneous portfolio model (LHPM)

The LHPM is used for pricing derivatives on a basket of defaultable assets. Hereby it is assumed that they are identically correlated. It is a static model. Therefore we do not model stochastic correlation by a process, but we will draw random numbers from the transition density (3.25). We will analyse how the incorporation of stochasticity in the correlation changes the characteristics of the LHPM. We will do this analysis with regard to prices, density and implied correlation.

5.2.1 Model

We start with a short description of the large homogeneous portfolio model (LHPM). For more details see Schönbucher [Sch03]. The LHPM can be used to model the default of a credit. It is based on N assets A_i which are constructed as

$$(5.12) \quad A_i = \rho Z + \sqrt{1 - \rho^2} Y_i .$$

Z, Y_1, \dots, Y_N are independent standard normal random variables. The associated correlation matrix

$$\begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \ddots & \rho \\ \vdots & \ddots & \ddots & \vdots \\ \rho & \dots & \rho & 1 \end{pmatrix}$$

is positive definite if $\rho \in \left(-\frac{1}{N-1}, 1\right]$. Therefore we assume $\rho \in [0, 1]$ in this section. By this construction the single assets are correlated with ρ . The model is static, the focus is on one single point in time. The default is triggered when the asset's value falls below the barrier K . This barrier is identical for every asset. The default probability is assumed to be

$$(5.13) \quad q = \Phi(K)$$

where Φ denotes the cumulative standard normal distribution.

We are interested in the loss fraction of the portfolio which is given as

$$(5.14) \quad L = \frac{1}{N} \sum_{i=1}^N 1_{\{A_i < K\}}.$$

This construction is the idea underlying the LHPM. In order to achieve further analytical tractability, in the LHPM an infinite number of assets is considered. Thus one takes the limit $N \rightarrow \infty$ and therewith the loss fraction (5.14) becomes continuous. Then density and cumulative distribution function (cdf) are known in closed form. The density is given as

$$(5.15) \quad p(x) = \sqrt{\frac{1-\rho^2}{\rho^2}} \exp\left(\frac{1}{2}(\Phi^{-1}(x))^2 - \frac{1}{2\rho^2}(\Phi^{-1}(p) - \sqrt{1-\rho^2}\Phi^{-1}(x))^2\right).$$

The cdf is as follows

$$(5.16) \quad P(x) = \Phi\left(\frac{1}{\rho}\left(\sqrt{1-\rho^2}\Phi^{-1}(x) - \Phi^{-1}(p)\right)\right).$$

In the following the LHPM is equipped with stochastic correlation. Referring for more details to previous chapters, we assume correlation to follow the stochastic process (2.6) with $m = 0$ and $\omega = 1$:

$$dX_t = \kappa(\theta - X_t)dt + \sqrt{X_t(1-X_t)}dZ_t, \quad X_0 = x_0$$

with $x_0, \theta \in (0, 1)$ and $\kappa > 0$. Regarding Thm. 3.2 the boundaries are neither attainable nor attractive if

$$(5.17) \quad \kappa \geq \max\left(\frac{1}{\theta}, \frac{1}{1+\theta}\right).$$

As the LHPM is time independent, we do not need to consider the whole process over time. Therefore we will only include the randomness from the transition density for $t \rightarrow \infty$ which becomes stationary and reads

$$(5.18) \quad f(x) = c \cdot \left(-2 + \frac{2}{x}\right)^{-\kappa\theta} (x(1-x))^{-1+\kappa}$$

with $c \in \mathbb{R}^+$ such that $\int_0^1 f(x)dx = 1$ see corollary 3.7. Thus in the following we will analyse the stochastic LHPM (SLHPM). With abuse of notation the random variable describing the correlation is denoted with X :

$$(5.19) \quad A_i = XZ + \sqrt{1 - X^2}Y_i.$$

where $X \sim P_f$ and P_f is the probability distribution implied by the stationary density f in (5.18). The probability that the loss fraction L is below a barrier $B \in [0, 1]$ reads

$$P(L \leq B) = \int_0^1 \int_0^B p(x, X) f(X) dx dX.$$

5.2.2 Results

We want to calculate densities, CDO prices and implied correlations in SLHPM and compare them to LHPM. As there is no closed-form solution if the correlation is stochastic, X is simulated with the help of numerical algorithms.

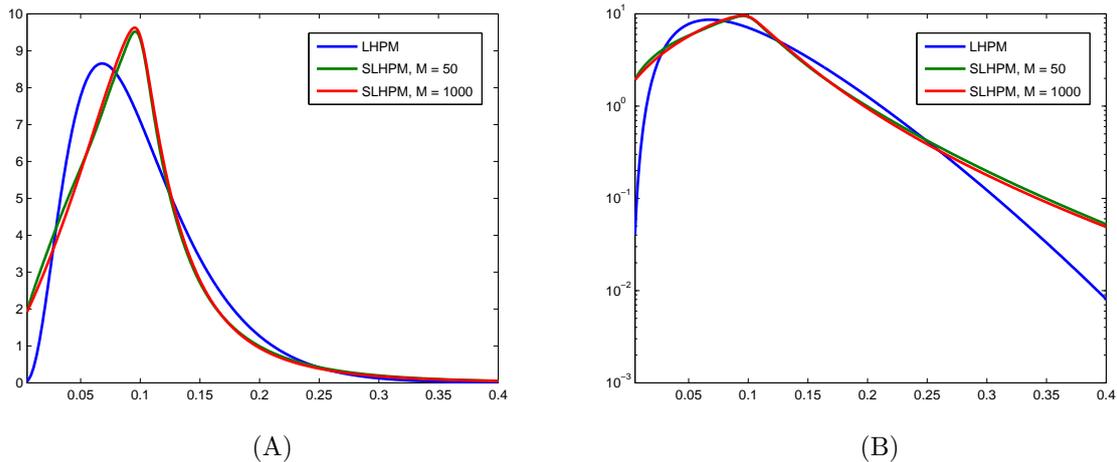


Figure 5.6: Densities of LHPM and SLHPM. (A): Different number of simulations of (5.15) in SLHPM with $\theta = 0.3, \kappa = 5, \omega = 1, q = 0.1$ compared to LHPM with $\rho = 0.3$. (B): Logarithmic coordinates on y-axes.

5.2.2.1 Density

We start by computing the densities. Hereby the SLHPM density is an estimated density. M random numbers for the correlation are drawn. For each of those the density is calculated. The mean over those M densities is considered as the SLHPM density. For making the two considered models comparable the mean parameter θ in the SLHPM is generally identical with the constant correlation ρ in the LHPM: $\rho = \theta$.

Fig. 5.6 result from the parameter configuration $\theta = 0.3, \kappa = 5$ and $q = 0.1$. One observes that the maximum is shifted to the right in case of SLHPM. Further both tails are fatter for the stochastic model which becomes visible in Fig. 5.6 (B). This is a rather interesting feature as many models based on normals like the LHPM, have the problem, that they underestimate the probability of rare events. Stochastic correlation as introduced here seems to lessen this problem.

The dependence on κ is illustrated in the next graph 5.7. One observes that for the considered values of κ , the density becomes stable for high mean reversion factor, see Fig. 5.7.

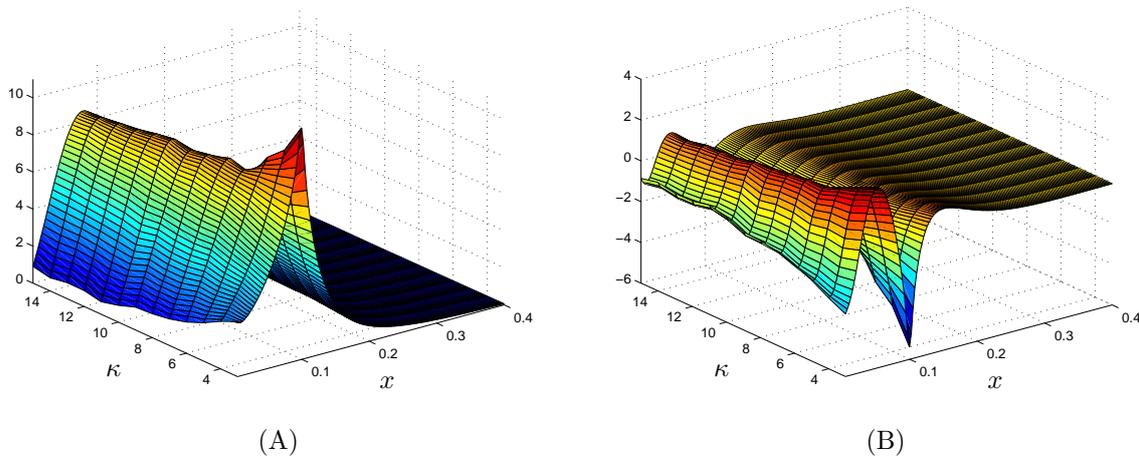


Figure 5.7: (A): Density (5.15) for fixed $\theta = 0.3, \omega = 1, q = 0.1$ and varying $\kappa \in [3, 15]$ in SLHPM. $M = 500$. (B): Difference of densities to LHPM density with $\rho = 0.3$ measured as LHPM - SLHPM

In Fig. 5.7 (B) the difference between the densities for increasing κ in SLHPM and the LHPM density is visualised. Although high mean reversion in (2.6) leads to a centralisation around the mean, the difference does not vanish for high mean reversion. Neither does it vanish for large argument values because of the fat tails. The difference at $x = 0.4$ is between -0.055 and -0.015 . Thus both models are structurally different.

The next graph 5.8 shows the LHPM densities and their stochastic counterparts for different mean θ and ρ respectively.

One observes that the differences between LHPM and SLHPM induced densities are relatively large for low correlation. But the higher correlation ρ in LHPM and θ in

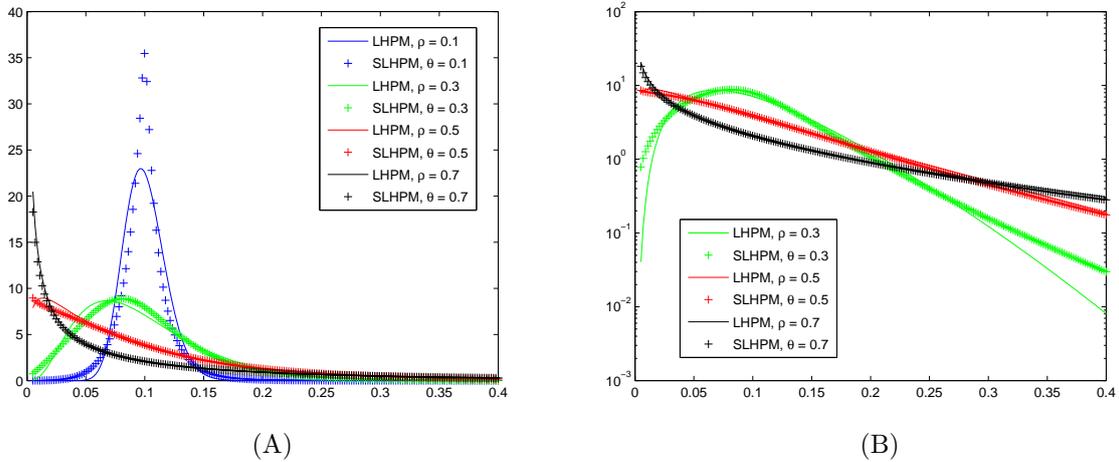


Figure 5.8: (A): Density (5.15) for fixed $\kappa = 12$, $\omega = 1$, $q = 0.1$ and varying θ in SLHPM and varying ρ in LHPM, respectively. $M = 500$. (B): Logarithmic scale.

SLHPM, the smaller this difference seems to become. This agrees with intuition as this corresponds to an approach to the degenerated case of perfect correlation. Particularly for $\rho = 0.5$ one observes that the density for constant correlation is more strongly peaked. This can be explained as there is one source of risk less.

The comparison of varying default probabilities q is presented in the subsequent graph. The stochastic densities still feature fat tails, see Fig 5.9:

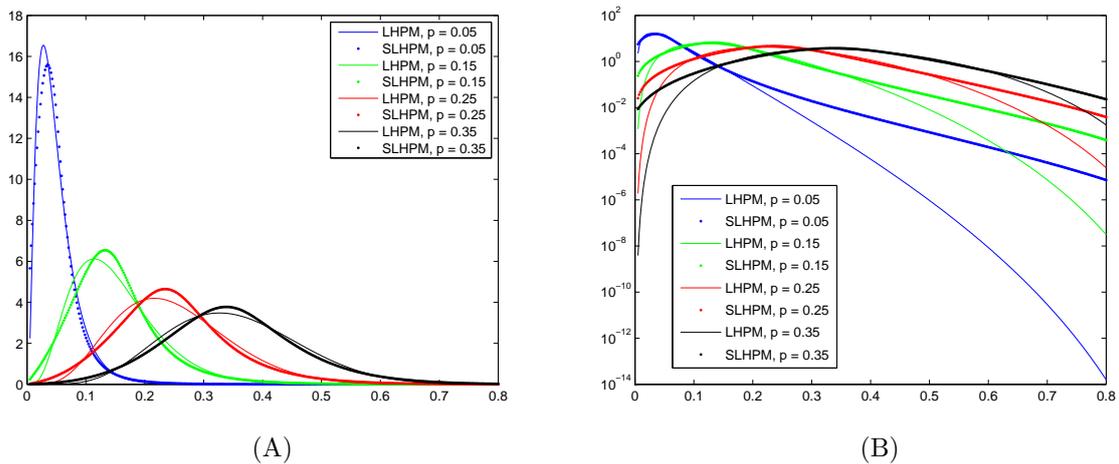


Figure 5.9: (A): Density (5.15) for $\theta = 0.3$, $\kappa = 12$, $\omega = 1$ compared to LHPM density with constant $\rho = 0.3$. Default probability q varies from. $M = 500$. (B): Logarithmic scale

Furthermore the maximum is always shifted towards higher loss fraction in case of the SLHPM which corresponds to a higher risk level. For small default probabilities the biggest differences show up in the right tail corresponding to the catastrophic events. For large q the differences at the very left tail appear bigger.

5.2.2.2 Price

As a next step the influence of randomness on the prices of a CDO, as defined in 1.3, is analysed. We consider an artificial CDO with detachment points $u_1 = 0.02, u_2 = 0.04$ and so on.

Fig. 5.10 shows the prices of tranches of a CDO with notional one. Thus we actually look at the loss fraction itself. The tranches $(u_{i+1}, u_i]$ and the pure equity tranches $[0, u_i]$ are considered. The interest rate is set to zero.

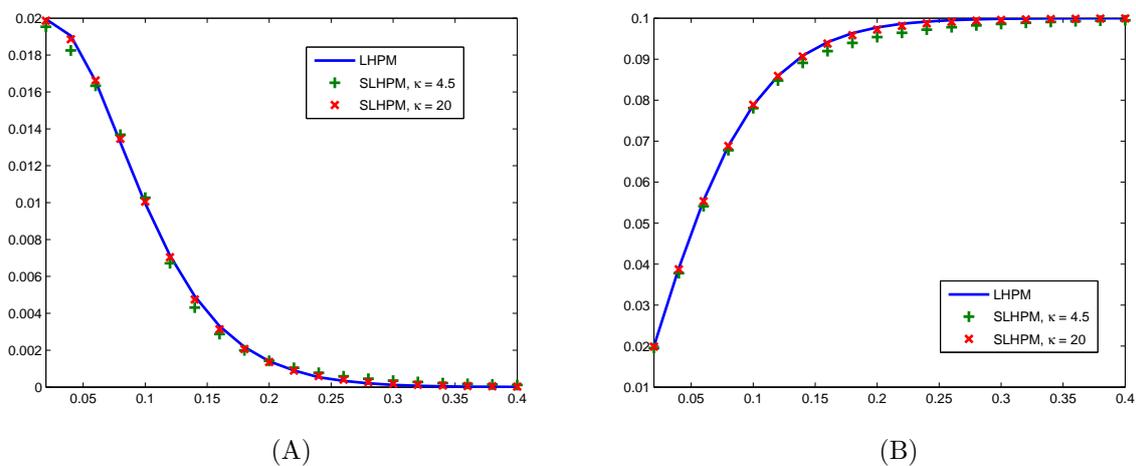


Figure 5.10: (A): Tranche prices for LHPM with $\rho = 0.3, q = 0.1$, and for SLHPM with $\theta = 0.3, \kappa = 4.5, 20, \omega = 1, q = 0.1, M = 1000$. (B): Equity tranches.

One observes that the prices from LHPM and SLHPM are very close to each other. The higher the mean reversion, the closer the prices are. But one also sees that the simulated prices show a lack of smoothness. For further analysis therefore the number of paths should be increased.

Fig. 5.11 shows the same comparison with θ and ρ being lowered. Again one notices that generally the prices of LHPM and SLHPM are rather close.

5.2.2.3 Implied Correlation

For getting more insight we calculate the implied correlations. Therefore we first simulate the prices with correlation according to the density (5.18). Afterwards we compute the constant correlation which has to be inserted into the LHPM to get this price. This input correlation is considered as implied correlation throughout this section. The calculation bases on the equity tranche prices. Hereby we follow the concept of base correlation, introduced by Ahluwalia et al in [RAW04]. The advantage of equity tranches is that its prices are monotone in correlation. This is in contrast to the prices of mezzanine tranches. Therewith the implied correlation is unique.

First Fig.5.12 depicts the implied correlation. It uses the same parameter configuration as used for the prices presented in Fig. 5.10 and 5.11.

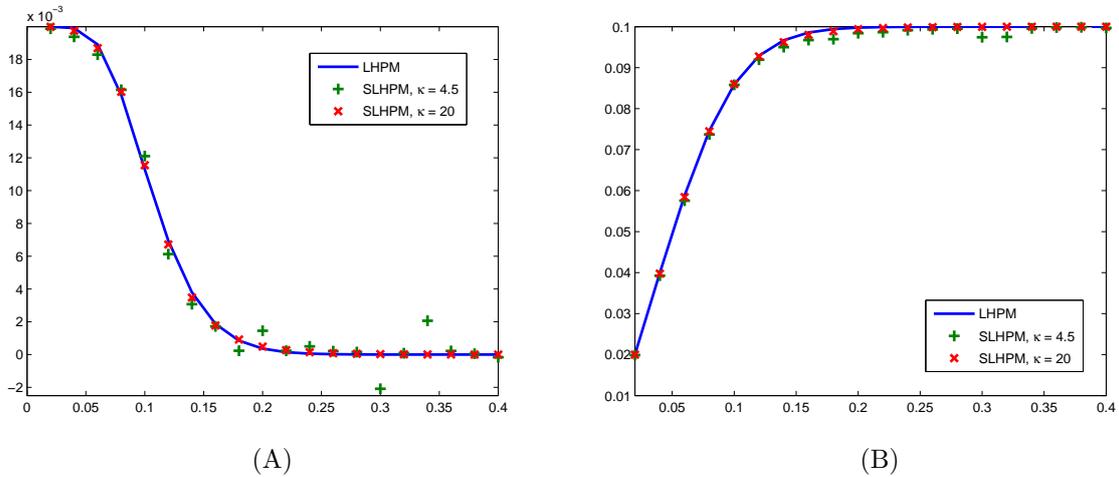


Figure 5.11: (A): Tranche prices for LHPM with $\rho = 0.2$, $q = 0.1$, and for SLHPM with $\theta = 0.2, \kappa = 4.5, 20, \omega = 1, q = 0.1$. $M = 1000$. (B): Equity tranches.

Although there is no obvious difference in the tranche price, the implied correlations do differ. There is a clear smile effect in the SLHPM. The lowest implied correlation is measured hereby for the tranche with detachment point 0.1 which also is the default probability.

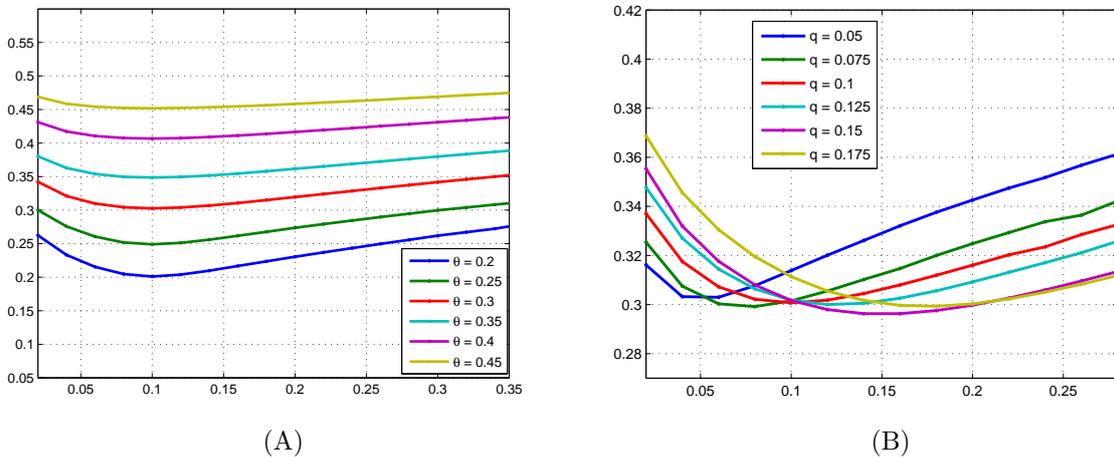


Figure 5.13: (A) Implied correlation in SLHPM for varying mean θ with $\kappa = 20, \omega = 1, q = 0.1$. $M = 500$. (B) Implied correlation in SLHPM for varying default probability q with $\theta = 0.3, \kappa = 20, \omega = 1$. $M = 500$.

Fig. 5.13 (A) visualises the changes in the smile according to the level of input correlation changes. Although the smile is clearly visible for all levels of correlation, it flattens with increasing θ . This goes in line with the result that LHPM and SLHPM become closer for increasing correlation, see also Fig. 5.8.

Fig. 5.13 (B) shows the implied correlation smile for different default probabilities q . The minimum of implied correlation - the smile - moves along with default probability. Fur-

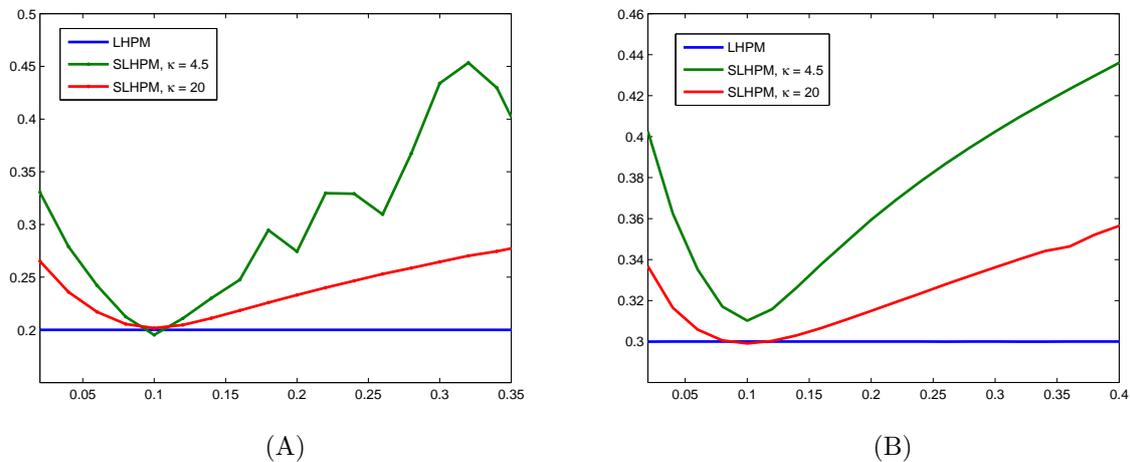


Figure 5.12: (A): Implied correlation for LHPM with $\rho = 0.3$, $q = 0.1$, and for SLHPM with $\theta = 0.2$, $\kappa = 4.5, 20$. $q = 0.1$. $M = 1000$. (B): Implied correlation for LHPM with $\rho = 0.3$ and for SLHPM with $\theta = 0.2$, $\kappa = 4.5, 20$. $q = 0.1$. $M = 1000$.

ther one can observe that lower default probabilities lead to steeper smiles. In the case of small probability of default, the smile effect rather resembles a skew. A skew is often observed in market prices, see again for example [RAW04]. Therewith we can hope, that the suggested model can successfully map market dynamics.

5.3 Summary

We presented two different ways to include stochastic correlation, first as a stochastic process and secondly as a random variable. As expected in both cases the incorporation led to structural changes in the models. In case of the FX model, we observed an increase in call prices, reflecting the new source of risk. Consequently neglecting stochasticity of correlation in the FX model counts for an underpricing of the risk associated with this derivative.

In the second example we saw that stochastic correlation between the defaultable assets led to significantly fatter tails in the distribution. This effect shows also up in the implied correlation, where we observed a non-symmetrical smile or a skew respectively. This is of particular interest, as there is an observed implied correlation skew in the market.

Conclusion

This thesis developed a model which can be used for pricing correlation swaps. We suggested a stochastic process for modelling correlation. Equation (2.6) defines a continuous and bounded mean reversion process driven by a Brownian motion. Chapter 3 was devoted to the analysis of this process, the following chapter to its calibration. In chapter 5 we included stochastic correlation into two models used within quantitative finance and illustrated its impact.

We can conclude that the analysis led to satisfactory results. There are parameter configurations where the process reaches the state of perfect correlation. This is undesirable because it does not agree with market observation. On the other hand, in the case of perfect correlation, we would clearly use a different modelling approach. By a specific parameter restriction given in Thm. 3.2, it is guaranteed that the process stays within the interval $(m, 1)$.

In the subsequent section we computed the stationary density by the virtue of the Fokker-Planck equation. In contrast to usual differential equations, we chose structural conditions instead of boundary conditions. To be specific, we demanded that it is a density and that it preserves the already known expectation. The knowledge of this density has many advantages. It allows for a comparison to reality and therewith calibration. It can be used by itself as a static model for stochastic correlation which features an intuitive explanation because it is coming from a bounded mean reversion model.

Afterwards we set up a system describing all moments of correlation. This knowledge can also be used for calibration, as we can now compare the market term structures of expectation and variance (possibly more moments) to the ones given by the market. But in our opinion this result is also promising from another point of view. By knowing the moments we basically know everything about the process. Further work will show if it is possible to set up the characteristic function, and therewith to obtain semi-analytical solutions.

The analysis on (2.6) showed that the process is suitable for modelling correlation. For practical usage it is necessary to calibrate the model. We already hinted at some ways of calibration. But as we assume that correlation is not constant, one can argue that the observed correlation estimated via (1.13) is not a single estimation of correlation

at a single point in time, but one (approximation to a) realisation over the respective time period. Therefore we developed an estimator which uses the historical data from the integral. Unfortunately integration over a process leads to loss of information through the averaging property of the integral. The developed estimation procedure suffers from that as well. The numerical tests show a lack of accuracy in estimating κ . On the other hand ω could be calibrated rather precisely. The estimation of the level of mean reverting is satisfactory in general. We would like to stress, that the numerical tests are not exhaustive. The fact that two out of three parameters are usually estimated quite good, gives hope, that maybe finer observation intervals or different weighting can improve the results.

Surely the topic of calibration should still be further researched. Especially the task of how to measure the dependence between correlation and asset.

In chapter 5 two examples were presented. First a call in an FX model was valued. Foreign exchange rate and asset are hereby stochastically correlated through time by (2.6). The call prices tend to be higher if correlation is taken into account. Moreover there is a skew in the implied correlation.

In the second example the correlation in the static LHPM is chosen as a random variable according to the density (3.25). The stochasticity led to fatter tails compared to the constant correlation case. There is a clear smile or skew respectively in the implied correlation. As there is a correlation skew in the market, the result is promising. Further analysis must show if the stochastic LHPM meets the market dynamics.

As there is no analytical solution for (2.6), it was necessary to simulate the process (2.6). The appendix explains why we chose the implicit Milstein scheme. Among the compared integration schemes it gives the highest probability of preserving the analytical boundaries. As all compared schemes stem from stochastic Taylor expansion, they cannot be expected to preserve both boundaries. Thus it remains the question of a suitable integration schemes which preserves upper and lower boundary.

It would be an interesting area of further research. As we know all moments, one could try an approach similar to the one by Andersen [And06], where the integration scheme is supposed to preserve the first two moments. One could extend the idea to the requirement that the boundaries may not be exceeded. Other approaches are also possible, see for example Alfonsi [Alf05] or Malham and Wiese [MW07].

We already mentioned various topics for further work. But there is still more to do. In the last stage of this work, Freddy Delbaen made me aware of joint work with Shirakawa [DS02], where they used (2.5) for modelling interest rates. They derived some interesting analytical results. For example they could compute the price of a zerobond. This gives hope, that correlation products can be solved analytically or in some sort of semi-analytical way.

Although the motivation in this work is mainly from equity, the effects showing up in the LHPM already indicate that stochastic correlation as presented here, may be successfully applied in credit modelling. This field of application should be further investigated.

There are only few comparisons with market data in this thesis. Especially the comparison with and fitting to market prices must be made, before the model can be used in

praxis. All in all more market analysis is of interest.

Deeper market analysis will possibly show that the concept of short correlation is not rich enough to cover the dynamics observed in the market. This is particularly likely if the market of correlations swaps develops further. A possible approach is to model the forward rates of correlation. This could be done in analogy to the Heath, Jarrow, and Merton framework [HJM92] for interest rates. Bühler [Büh06] extended the idea to modelling variance.

We restricted ourselves to the case of two correlated underlyings and the case of identically correlated underlyings. In the first case the positive definiteness is given if the process is in $[-1, 1]$. In the case of identically correlated random variables, the positive definiteness can be reached by adjusting the lower boundary from -1 to $-\frac{1}{N-1}$ where N denotes the number of random variables.

Related to this topic is the question of how to measure and include the correlation between the assets and the Brownian motion driving the correlation itself. A still unsolved problem in correlation modelling is how to model the whole matrix such that it stays positive definite. In the introduction we mentioned the Wishart process. Here the covariance matrix is described by a stochastic differential equation. There we also mentioned why we prefer to model variance and correlation separately. Similarly to the covariance matrix it is possible to model the correlation matrix C as follows [BG07]: One does not model C itself, but one splits C into Cholesky factor L and scaling matrix D : $C = DLL^T D$. Hereby D is defined via L such that $c_{i,i} = 1$, $c_{i,j} \leq 1$, $i, j = 1, \dots, n$. Now L can be modelled arbitrarily. This approach is feasible. But from a more general point of view, it would be nice, if the dimension of the correlation matrix would not matter when we model and calibrate the stochastic process for the correlation $c_{i,j}(t)$ between two assets.

Although there are many open questions which further research must investigate, we think that stochastic correlation is an important source of risk which should not be neglected. Based on the results of this thesis one can conclude that the suggested model (2.6) is suitable for modelling correlation. It is intuitive, fits observed characteristics of correlation, can be calibrated and numerically handled.

Appendix

Within the main part of this thesis we left aside the numerical treatment of the correlation process (2.6). Although at some points numerical solving of the stochastic differential equation in a Monte Carlo simulation was necessary. We devote section A.1 of the appendix to the task of numerical discretisation of (2.6), because the process has one feature which distinguishes it from the vast majority of stochastic processes considered in quantitative finance. It possesses a lower and an upper bound. So far there are mainly results on preserving one boundary, which is motivated from volatility and interest rate modelling, see for example Kahl and Schurz [KS06].

Further we postponed some of the computations within section 3.2. In A.2 we will compute the scaling constants necessary to make the solutions of the Fokker-Planck to density. Additionally we will show that constraint (3.14) holds, i. e. that the scaled solutions preserve the analytically known expectation θ .

A.1 Numerical treatment

Throughout this work it is necessary to simulate the stochastic process (2.6) as so far its analytical solution is not known. For calculating prices of correlation dependent options it may therefore be necessary to use Monte Carlo simulation techniques. Within this thesis we use the implicit Milstein scheme for discretising the stochastic differential equation (2.6). We reason this decision with a comparison between four integration schemes. Let

$$(A.1) \quad dx_t = a(t, x_t)dt + b(t, x_t)dW_t$$

be an Itô process driven by a Brownian motion W_t . For this stochastic differential equation and a time grid $t_0 < t_1 < \dots < t_n$ the integration schemes under consideration read:

- explicit Euler-Maruyama

$$(A.2) \quad x_{t_{i+1}} = x_{t_i} + a(t_i, x_{t_i})(t_{i+1} - t_i) + b(t_i, x_{t_i})(W_{t_{i+1}} - W_{t_i})$$

- drift-implicit Euler-Maruyama,

$$(A.3) \quad x_{t_{i+1}} = x_{t_i} + a(t_{i+1}, x_{t_{i+1}})(t_{i+1} - t_i) + b(t_i, x_{t_i})(W_{t_{i+1}} - W_{t_i})$$

- explicit Milstein

$$(A.4) \quad \begin{aligned} x_{t_{i+1}} &= x_{t_i} + a(t_i, x_{t_i})(t_{i+1} - t_i) + b(t_i, x_{t_i})(W_{t_{i+1}} - W_{t_i}) \\ &\quad + \frac{1}{2}b(t_i, x_{t_i})\frac{\partial}{\partial x}b(t_i, x_{t_i})\left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)\right) \end{aligned}$$

and

- drift-implicit Milstein

$$(A.5) \quad \begin{aligned} x_{t_{i+1}} &= x_{t_i} + a(t_{i+1}, x_{t_{i+1}})(t_{i+1} - t_i) + b(t_i, x_{t_i})(W_{t_{i+1}} - W_{t_i}) \\ &\quad + \frac{1}{2}b(t_i, x_{t_i})\frac{\partial}{\partial x}b(t_i, x_{t_i})\left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)\right). \end{aligned}$$

Naturally for the implicit schemes a rearrangement has to be undertaken. We will use the abbreviations $\Delta_n = t_{n+1} - t_n$ and $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ in the following. For more information on stochastic discretisation schemes like derivation and definition of convergence order see for example Kloeden and Platen[KP99].

We chose the implicit Milstein scheme because of the following reasons. Both Euler schemes possess a weak convergence order of 1 and a strong convergence order of 0.5. The Milstein schemes feature a higher strong order of convergence. It is 1 if the coefficient function are continuously differentiable.

This does not hold for the correlation process, since although the diffusion function $\sqrt{(1-x)(x-m)}$ is continuously differentiable in the inner interval $(m, 1)$, it is not in the boundaries. Nonetheless we hope that the strong order of convergence of 1 is practically valid if the boundaries are not attainable. Thus from the point of view of convergence the Milstein schemes should be preferred.

But the correlation process (2.6) inhibits a second problem, a discretisation scheme has to deal with: the treatment of the boundaries. Here we can state the following:

Corollary A.1 (Boundary preservation) Let the realisation \bar{X}_n at t_n be given and within the boundaries. Denote the probability that the explicit Euler scheme exceeds the upper boundary in the next step by

$$P_{EE}^1 := \mathcal{P}(X_{n+1} > 1 | \bar{X}_n).$$

Analogously the probability for violating the lower boundary shall be denoted by

$$P_{EE}^m := \mathcal{P}(X_{n+1} < m | \bar{X}_n).$$

Using the analogous notations for the other three discretisation schemes, it holds in every point $\bar{X}_n \in (m, 1)$:

$$\begin{aligned} P_{EE}^m &\geq P_{IE}^m & \text{and} & & P_{EE}^1 &\geq P_{IE}^m, \\ P_{EM}^m &\geq P_{IM}^m & \text{and} & & P_{EM}^1 &\geq P_{IM}^m \end{aligned}$$

Proof: First we look at the probability $\mathcal{P}(X_{n+1} < 1 | \bar{X}_n)$ at the upper boundary $r = 1$ and compare explicit and implicit Euler scheme.

$$\begin{aligned} \mathcal{P}(X_{n+1} < 1 | \bar{X}_n) &= \mathcal{P}\left(\Delta W_n < \frac{1 - \bar{X}_n - \kappa(\theta - \bar{X}_n)\Delta}{\omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}}\right) \\ &= \Phi\left(\frac{1 - \bar{X}_n - \kappa(\theta - \bar{X}_n)\Delta}{\omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}}\right). \end{aligned}$$

On the other hand the probability for the implicit scheme is

$$\begin{aligned} \mathcal{P}(X_{n+1} < 1 | \bar{X}_n) &= \mathcal{P}\left(\Delta W_n < \frac{1 + \kappa\Delta_n - \bar{X}_n - \kappa\theta\Delta_n}{\omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}}\right) \\ &= \Phi\left(\frac{1 + \kappa\Delta_n - \bar{X}_n - \kappa\theta\Delta_n}{\omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}}\right). \end{aligned}$$

By presumption $\bar{X}_n \in (m, 1)$, therefore $1 - \bar{X}_n - \kappa(\theta - \bar{X}_n)\Delta_n < 1 + \kappa\Delta_n - \bar{X}_n - \kappa\theta\Delta_n$ and thus:

$$P_{EE}^1 \geq P_{IE}^1.$$

The calculations work similar for the lower boundary, where we have to use that $\bar{X}_n > m$.

Next we have a look at the Milstein schemes. First we consider the explicit one:

$$\begin{aligned} \mathcal{P}(X_{n+1} < 1 | \bar{X}_n) &= \mathcal{P}\left(\bar{X}_n + \kappa(\theta - \bar{X}_n)\Delta_n + \omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}\Delta W_n \right. \\ &\quad \left. + \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n)((\Delta W_n)^2 - \Delta) < 1\right) \\ &= \mathcal{P}(\alpha + \beta\Delta W_n + \gamma(\Delta W_n)^2 < 0), \end{aligned}$$

where

$$\begin{aligned} \alpha &= -1 + \bar{X}_n + \kappa(\theta - \bar{X}_n)\Delta_n - \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n)\Delta_n \\ \beta &= \omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)} \\ \gamma &= \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n). \end{aligned}$$

One can derive the probabilities for the implicit scheme in analogous manner:

$$\begin{aligned} \mathcal{P}(X_{n+1} < 1 | \bar{X}_n) &= \mathcal{P}\left(-1 - \kappa\Delta_n + \bar{X}_n + \kappa\theta\Delta_n + \omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)}\Delta W_n \right. \\ &\quad \left. + \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n)((\Delta W_n)^2 - \Delta_n) < 0\right), \end{aligned}$$

with

$$\begin{aligned} \alpha_I &= -1 - \kappa\Delta_n + \bar{X}_n + \kappa\theta\Delta_n - \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n)\Delta_n = \alpha + \kappa\Delta_n\bar{X}_n - \kappa\Delta_n \\ &= \alpha + \kappa\Delta_n(\bar{X}_n - 1) < \alpha \\ \beta_I &= \omega\sqrt{(1 - \bar{X}_n)(\bar{X}_n - m)} = \beta \\ \gamma_I &= \frac{1}{4}\omega^2(1 + m - 2\bar{X}_n) = \gamma. \end{aligned}$$

Therewith:

$$\begin{aligned}
& \mathcal{P}(\alpha + \kappa\Delta_n(\bar{X}_n - 1) + \beta\Delta_n W_n + \gamma(\Delta W_n)^2 < 0) \\
&= \mathcal{P}(\alpha + \beta\Delta W_n + \gamma(\Delta W_n)^2 < 0) + \mathcal{P}(0 < \alpha + \beta\Delta W_n + \gamma(\Delta W_n)^2 < -\kappa\Delta_n(\bar{X}_n - 1)) \\
&> \mathcal{P}(\alpha + \beta\Delta_n W_n + \gamma(\Delta_n W_n)^2 < 0) .
\end{aligned}$$

□

Thus the implicit schemes perform better in preserving the boundaries. Fig. A.1 visualises the theoretical result:

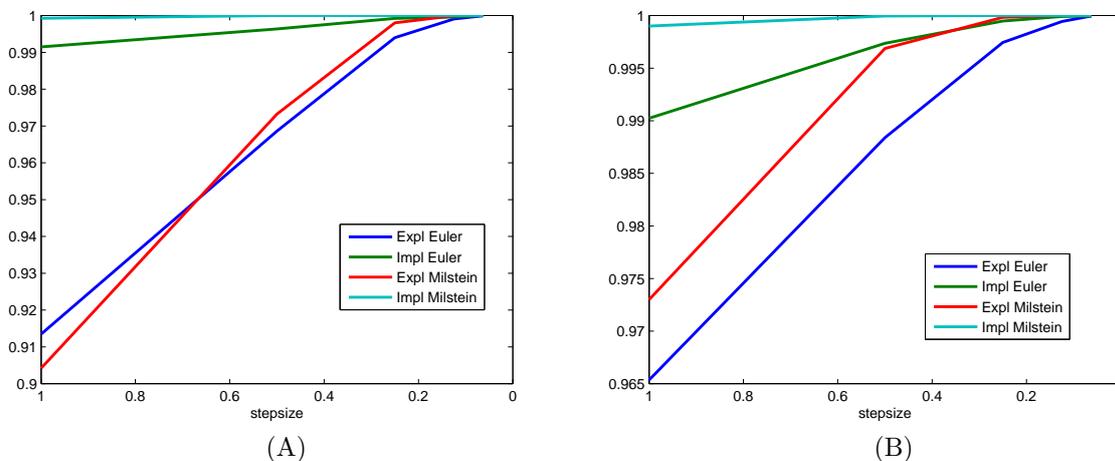


Figure A.1: Percentage of paths exceeding the boundaries, $m = -1$, $\sigma = 1$, $T = 5$, $M = 250000$. (A) $\theta = X_0 = 0.4$, $\kappa = 1.7$. (B) $\theta = X_0 = 0.0$, $\kappa = 1.1$.

Further it indicates that the second order scheme is better suited than the first order one in the explicit and the implicit case, respectively. Therefore in chapters 4 and 5, we numerically solved the stochastic differential equation (2.6) by virtue of the implicit Milstein scheme¹.

A.2 Calculating the integration constants for the stationary densities

We will only show the calculations for Cor. 3.7, as this is the most general case. The constants in Thm. 3.3 and Cor. 3.6 can be derived from that.

The function p given by (3.25)

$$p(x) = \frac{\left(\frac{1-x}{x-m}\right)^{1-\frac{\kappa}{\omega^2}(m-2\theta+1)} \left((1-x)(x-m)\right)^{\frac{\kappa}{\omega^2}}}{(x-1)^2}$$

¹As noted there the implicit Milstein schemes worked successfully in these applications. The boundaries were not exceeded.

fulfills the Fokker-Planck equation (3.24) of (2.6) for $t \rightarrow \infty$. Also every $p_C(x) = C \cdot p(x)$ with $C \in \mathbb{R}$ fulfills this differential equation. Now we want to calculate a constant C such that (3.13)

$$\int_m^1 p_C(x) dx = 1.$$

We will show that therewith also (3.14)

$$\int_m^1 x \cdot p_C(x) dx = \theta$$

holds. Mathematica[Wol05] computes the integral from m to 1 over p as

$$\begin{aligned} \int_m^1 p(x) dx &= (1-m)^{\frac{2\kappa}{\omega^2}-1} \Gamma\left(\frac{\kappa}{\omega^2} - 1\right) \Gamma\left(\frac{2(m-\theta)\kappa}{(m-1)\omega^2}\right) \\ &\cdot {}_2\tilde{F}_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2}; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} - 1; 1\right) \end{aligned}$$

where

$$(A.6) \quad {}_2\tilde{F}_1(\alpha, \beta; \gamma; z) := \frac{{}_2F_1(\alpha, \beta; \gamma; z)}{\Gamma(\gamma)}$$

denotes the **regularised hypergeometric function**². The definition of the hypergeometric function ${}_2F_1$ is given in (3.17). See also [AS72] and [WW72]. Thus we set C such that

$$\begin{aligned} \frac{1}{C} &= (1-m)^{\frac{2\kappa}{\omega^2}-1} \Gamma\left(\frac{\kappa}{\omega^2} - 1\right) \Gamma\left(\frac{2(m-\theta)\kappa}{(m-1)\omega^2}\right) \\ &\cdot {}_2\tilde{F}_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2}; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} - 1; 1\right). \end{aligned}$$

Note that by construction the right hand side is not zero. Thus C is well-defined. It remains to show that $p_\omega(x) := C \cdot p(x)$ satisfies condition (3.14). We calculate the expectation with Mathematica[Wol05]:

$$\int_m^1 x p_\omega(x) dx = m + \frac{2(\theta-m)\kappa {}_2\tilde{F}_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2} + 1; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2}; 1\right)}{\omega^2 {}_2\tilde{F}_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2}; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} - 1; 1\right)}.$$

According to [WW72] the hypergeometric function ${}_2F_1(a, b, c, 1)$ can be expressed in terms of the gamma function as follows:

$$(A.7) \quad {}_2F_1(\alpha, \beta, \gamma, 1) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}$$

²The gamma function is defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ for $x \notin \mathbb{Z}_-$.

as long as γ is not a negative integer, see [WW72], chapter XIV. As Mathematica cannot simplify this term any further, we do the necessary calculations ourselves. We introduce the following abbreviations:

$$\begin{aligned} a &= \frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \\ b &= \frac{2(m-\theta)\kappa}{(m-1)\omega^2} + 1, \\ c &= \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2}. \end{aligned}$$

Therewith we can write:

$$\int_m^1 xp_\omega(x)dx = m + \frac{2(\theta-m)\kappa {}_2F_1(a, b; c; 1)}{\omega^2 {}_2F_1(a, b-1; c-1; 1)} \cdot \frac{\Gamma(c-1)}{\Gamma(c)}.$$

Due to $c = \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} > 0$, neither c nor $c-1$ are negative integers. Using (A.7) and $\Gamma(z+1) = z\Gamma(z)$, we can thus calculate:

$$\begin{aligned} \int_m^1 xp_\omega(x)dx &= m + \frac{2(\theta-m)\kappa}{\omega^2} \frac{\Gamma(c)\Gamma(c-a-b)\Gamma(c-1-a)\Gamma(c-1-(b-1))}{\Gamma(c-a)\Gamma(c-b)\Gamma(c-1)\Gamma(c-1-a-(b-1))} \cdot \frac{\Gamma(c-1)}{\Gamma(c)} \\ &= m + \frac{2(\theta-m)\kappa}{\omega^2} \frac{\Gamma(c-1-a)}{\Gamma(c-a)} \\ &= m + \frac{2(\theta-m)\kappa}{\omega^2} \frac{\Gamma(c-1-a)}{(c-a-1)\Gamma(c-a-1)} \\ &= m + \frac{2(\theta-m)\kappa}{\omega^2} \frac{1}{(c-a-1)}. \end{aligned}$$

With

$$c-a-1 = \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} - \left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1 \right) - 1 = \frac{2\kappa}{\omega^2}$$

we can conclude that

$$\int_m^1 p_\omega(x)dx = m + \frac{2(\theta-m)\kappa}{\omega^2} \frac{\omega^2}{2\kappa} = \theta.$$

Thus in Cor. 3.7 we must choose

$$\begin{aligned} C &= \left((1-m)^{\frac{2\kappa}{\omega^2}-1} \Gamma\left(\frac{\kappa}{\omega^2} - 1\right) \Gamma\left(\frac{2(m-\theta)\kappa}{(m-1)\omega^2}\right) \right. \\ (A.8) \quad &\cdot \left. {}_2\tilde{F}_1\left(\frac{(m-2\theta+1)\kappa}{(m-1)\omega^2} - 1, \frac{2(m-\theta)\kappa}{(m-1)\omega^2}; \frac{(3m-2\theta-1)\kappa}{(m-1)\omega^2} - 1; 1\right) \right)^{-1}. \end{aligned}$$

Therefore the constant in Cor. 3.6 must be

$$(A.9) \quad C = \left((1-m)^{2\kappa-1} \Gamma(\kappa-1) \Gamma\left(\frac{2(m-\theta)\kappa}{m-1}\right) \cdot {}_2\tilde{F}_1\left(\frac{2(m-\theta)\kappa}{m-1}, \frac{(m-2\theta+1)\kappa}{m-1} - 1; \frac{(3m-2\theta-1)\kappa}{m-1} - 1; 1\right) \right)^{-1}.$$

With some simplification, one can write

$$(A.10) \quad C = \left(\frac{1}{\kappa + \theta\kappa} {}_2F_1(1, (\theta-1)\kappa + 1; \theta\kappa + \kappa + 1; -1) + \frac{1}{\kappa - \theta\kappa} {}_2F_1(1, 1 - (\theta+1)\kappa; -\theta\kappa + \kappa + 1; -1) \right)^{-1}$$

for the constant in Thm. 3.3. Thus all solutions fulfill the Fokker-Planck equation and the structural conditions (3.13) and (3.14).

Bibliography

- [Alf05] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) process. *Monte Carlo Methods and Applications*, 11:355–384, 2005.
- [And06] L. Andersen. Efficient Simulation of the Heston Stochastic Volatility Model. *SSRN*, 2006. <http://ssrn.com/abstract=946405>.
- [AS72] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Wiley, New York, 1972.
- [BG07] A. Bartel and B. Garbe. Private communication. University of Wuppertal, 2007.
- [Büh06] H. Bühler. Volatility Markets: Consistent Modeling, Hedging and Practical Implementation. *Dissertation, TU Berlin*, 2006.
- [CIR85] J. Cox, J. Ingersoll, and S. Ross. A Theory of the Term Structures of Interest Rates. *Econometrica*, 3:229–263, 1985.
- [CT04] R. Cont and P. Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall/CRC, 2004.
- [DHPD00] T. E. Duncan, Y. Hu, and B. Pasic-Duncan. Stochastic calculus for fractional Brownian motion. *SIAM J. of Control and Optimization*, 41:582–612, 2000.
- [DPS00] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.
- [DS02] F. Delbaen and H. Shirakawa. An Interest Rate Model with Upper and Lower Bounds. *Financial Engineering and the Japanese Markets*, 9(3):191–209, 2002.
- [Duf01] D. Dufresne. The integrated square-root process. <http://www.economics.unimelb.edu.au/actwww/no90.pdf>, 2001.
- [Dup94] B. Dupire. Pricing with a Smile. *Risk*, 7(1):18–20, 1994.
- [Emm06] C. van Emmerich. Modelling Correlation as a Stochastic Process. Technical report, University of Wuppertal, 2006. <http://www-num.math.uni-wuppertal.de/preprints.html>.

- [Emm07] C. van Emmerich. Stochastic Mean Reversion in Large Portfolio Model. Technical report, University of Wuppertal, 2007. <http://www-num.math.uni-wuppertal.de/preprints.html>.
- [Fra03] J. C. G. Franco. Maximum likelihood estimation of mean reverting processes. *Research paper*, 2003. http://investmentscience.com/Content/howtoArticles/MLE_for_OR_mean_reverting.pdf.
- [GJ03] M. Günther and A. Jüngel. *Finanzderivate mit MATLAB*. Vieweg, Wiesbaden, 2003.
- [GJS04] C. Gourierieux, J. Jasiak, and R. Sufana. The Wishart Autoregressive Process of Multivariate Stochastic Volatility. *Research paper*, 2004.
- [Gla03] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.
- [Hes93] S. Heston. A closed-form solution for options with stochastic volatility with application to bond and currency options. *Rev. Financ. Studies*, 6:327–343, 1993.
- [HJM92] D. Heath, R. Jarrow, and A. Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for contingent claims valuation. *Econometrica*, 60:77–105, 1992.
- [HW87] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *J. Finance*, 42:281–300, 1987.
- [HW02] E. Hairer and G. Warner. *Solving Ordinary Differential Equations I*. Springer, Berlin, 2002.
- [KJ06] C. Kahl and P. Jäckel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Journal of Quantitative Finance*, 6(6):513–536, December 2006. <http://www.math.uni-wuppertal.de/~kahl/publications.html>.
- [KP99] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1999.
- [KS06] C. Kahl and H. Schurz. Balanced Milstein Methods for Ordinary SDEs. *Monte Carlo Methods and Applications*, 12(2):143–170, April 2006. <http://www.math.uni-wuppertal.de/~kahl/publications.html>.
- [KT81] S. Karlin and H. M. Taylor. *A second course in stochastic processes*. Academic Press, New York, 1981.
- [MW07] S. J. A. Malham and A. Wiese. Stochastic Lie Group Integrators. *SIAM J. Sci. Comput.*, 2007. accepted.
- [NM02] K. Neuman and M. Morlock. *Operations Research*. Carl Hanser Verlag, 2002.

- [NS04] M. Nelles and C. Senft. *Basel II und der Einsatz von Kreditderivaten*, volume 6. Die Bank, 2004.
- [Øks00] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, Berlin, 2000.
- [PE07] R. Pulch and C. van Emmerich. Polynomial Chaos for Simulating Stochastic Volatilities. Technical report, University of Wuppertal, 2007. <http://www-num.math.uni-wuppertal.de/preprints.html>.
- [RAW04] L. McGinty R. Ahluwalia, E. Beinstein and M. Watts. Introducing Base Correlation. *Research paper*, 2004.
- [Ris89] H. Risken. *The Fokker-Planck equation*. Springer-Verlag, Berlin, 1989.
- [Sch03] H. Schurz. General theorems for numerical approximation of stochastic processes on the Hilbert space H_2 . *Electr. Trans. Numer. Anal.*, 16:50–69, 2003.
- [Tap07] K. Tappe. Launching a canon into multivariate Lévy processes. Technical report, University of Wuppertal, 2007. <http://www-num.math.uni-wuppertal.de/preprints.html>.
- [UO30] G. E. Uhlenbeck and L. S. Ornstein. On the theory of Brownian motion. *Phys. Rev.*, 36:823–841, 1930.
- [Wil00a] P. Wilmott. *Quantitative Finance*, volume 1. Wiley, 2000.
- [Wil00b] P. Wilmott. *Quantitative Finance*, volume 2. Wiley, 2000.
- [Wol05] Wolfram Research, Inc. *Mathematica*. Wolfram Research, Inc., Champaign Illinois, Version 5.2 edition, 2005.
- [WW72] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge at the University Press, 1972.

Index

- θ -method, 37
- attainability, 15, 18
- attractiveness, 15, 16
- averaged realised correlation, 2
- base correlation, 59
- boundary
 - behaviour, 14
 - condition, 19
- collateralised debt obligation, 2, 3
- conditional Monte Carlo, 50
 - stochastic correlation, 50
- copula, 4
- correlation
 - realised, 2, 5, 12
 - rolling, 5
 - short, 6, 13
 - swap, 1
- correlation skew, 3
 - FX model, 54
 - large homogeneous portfolio model, 59
- Euler-Maruyama scheme, 65
 - drift implicit, 66
- Fokker-Planck equation, 20
 - stationary, 21
- foreign exchange model, 3, 48
- fractional Brownian Motion, 4
- gamma function, 69
- Heston model, 3
- hyper geometric series, 22
- integrated
 - Ornstein-Uhlenbeck process, 35
 - stochastic correlation, 29
- Jensen's inequality, 52
- Kolmogoroff forward equation, 20
- large homogeneous portfolio model, 3, 54
- likelihood function, 37
- linear correlation coefficient, 4, 5, 8
- Milstein scheme, 66
 - drift implicit, 34, 66
- moments
 - integrated stochastic correlation, 29
 - stochastic correlation, 27
- Ornstein-Uhlenbeck process, 34
- quanto, 48
 - adjustment, 50
- realised correlation, 12
- regularised hyper geometric series, 69
- rolling correlation, 8
- scale measure, 15
- short correlation, 6, 13
- speed measure, 15
- stationary density, 23
 - solution, 26
- stochastic correlation model SDE, 12
- stochastically correlated Brownian motions, 10
- transition density condition, 21
- Wishart process, 4, 64