UNIVERSITY OF
GOTHENBURG

Simulation and Statistical Methods for Stochastic Differential Equations

Master's Thesis in Mathematical Statistics

CHRISTIAN KÄLLGREN

Department of Mathematical Sciences
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2019

Master's thesis 2019

Simulation and Statistical Methods for Stochastic Differential Equations

CHRISTIAN KÄLLGREN

UNIVERSITY OF
GOTHENBURG

Department of Mathematical Sciences
Division of Mathematical Statistics
University of Gothenburg
Gothenburg, Sweden 2019

Simulation and Statistical Methods for Stochastic Differential Equations
CHRISTIAN KÄLLGREN

© CHRISTIAN KÄLLGREN, 2019.

Supervisor: Patrik Albin, Department of Mathematical Sciences
Examiner: Johan Tykesson, Department of Mathematical Sciences

Master's Thesis 2019
Department of Mathematical Sciences
Division of Mathematical Statistics
University of Gothenburg
SE-412 96 Gothenburg

Simulation and Statistical Methods for Stochastic Differential Equations
CHRISTIAN KÄLLGREN
Department of Mathematical Sciences
University of Gothenburg

Abstract

We look at numerical methods for simulation of stochastic differential equations exhibiting volatility induced stationarity. This is a property of the process which means that the stationary behaviour is mostly imposed by how volatile the process is. The property creates issues in simulation and hence also in statistical methods. The methods considered for simulations are the fully implicit Euler scheme and time-changed simulation. We look at statistical methods for estimation of parameters. The specific statistical methods we investigate is the likelihood ratio, which gives expressions for the drift parameters for CKLS and least squares estimation, which is used together with quadratic variation to estimate parameters in different models.

Acknowledgements

I want thank my supervisor Patrik Albin for his guidance during this project.

I also want to thank my parents Anders and Gun-Britt and my brother William for their continued support.

Christian Källgren, Gothenburg, June 2019

Contents

List of Figures **xi**

1 Introduction **1**
1.1 Background and problem statement 1
1.2 Method 2
1.3 Outline 2

2 Theory **3**
2.1 Preliminaries 3
2.2 Volatility induced stationarity 4
2.3 Numerical simulation schemes 5
2.3.1 Issues with the Euler scheme 6
2.3.2 Implicit Euler scheme 7
2.3.3 Change-of-time simulation 7
2.4 Transformations 9
2.5 Stochastic differential equations 9
2.5.1 CKLS 9
2.5.2 Hyperbolic diffusion 10
2.5.3 A family of heavy tailed SDE 11

3 Methods **15**
3.1 Simulation 15
3.2 Likelihood ratio method 15
3.3 Least squares estimation 18

4 Results **19**
4.1 Simulation results 19
4.2 Statistical results 19
4.2.1 CKLS 19
4.2.2 Hyperbolic diffusion 21
4.2.3 A family of heavy tailed SDE 22

5 Conclusion **29**
5.1 Simulation 29
5.2 Statistical methods 29

Bibliography **31**

List of Figures

2.1 Simulated path of the CKLS, using original representation (2.7) . . . 11
2.2 Simulated path of the CKLS, transformed representation (2.8) 11
2.3 Sample path of (2.10) using the parameter set $(\alpha, \beta, \delta, \mu, \sigma) = (5, -4, 1, 7, 0.05)$ the scheme from Section 2.3.3 with $T = 500$. 12
2.4 Simulated path of the original representation (2.11) 13
2.5 Simulated path of the transformed representation (2.12) 13

3.1 Brownian motion where the steps are decreased from 2^{12} down to 2^3 . 16

4.1 Plot of the convergence results for the CKLS model using the parameter set $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$. 20
4.2 Plot of the convergence results for the transformed *CKLS* model using the parameter set $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$. 21
4.3 Plot of the convergence results for the family of heavy tailed SDE using $\alpha = -10$ 22
4.4 Plot of the convergence results for transformed SDE, (2.12), using $\alpha = -10$ 23
4.5 Plot of the convergence results for the hyperbolic diffusion using $(\alpha, \beta, \delta, \mu, \sigma) = (7, -6, 0.05, 7, 1)$ 24
4.6 Drift parameters estimated by (3.3) and (3.4) with $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$. 25
4.7 Drift parameters estimated by (3.3) and (3.4) with $(\alpha, \beta, \sigma, \gamma) = (1, -1, 1, 10)$. 25
4.8 Drift parameters estimated by (3.3) and (3.4) with $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 50)$. 26
4.9 Drift parameters estimated by (3.3) and (3.4) with $(\alpha, \beta, \sigma, \gamma) = (1, -1, 1, 50)$. 26
4.10 Parameters estimated for the hyperbolic diffusion, Section 2.5.2 using the parameter set $(\alpha, \beta, \delta, \mu, \sigma) = (7, -6, 1, 7, 0.05)$ with $h = 2^{18}$ and interval length $T = 50$. 27
4.11 The parameter α estimated for the transformed SDE Section 2.11 using step length $h = 2^{18}$ and interval length $T = 10$. 27

1
Introduction

In this chapter we give a short background and problem statement together with an outline of the thesis.

1.1 Background and problem statement

Simulation of stochastic differential equations, SDE, has been a topic of research for decades. Stochastic differential equations exhibiting the property called *volatility induced stationarity*, vis, have been a topic of research for a long time, however, generally, the concept has not been explicitly used in the academic literature. In [3] an early definition for the concept in the context of CKLS was introduced, a generalized CIR model. This was in 1997. Later [1] extended this definition and made a more generalized version. Stochastic differential equations that exhibit this property are generally difficult to simulate, certain traditional methods that are normally used, are not enough. They become unstable as they are not able to handle their wild behaviour of. In particular [1] produced results regarding this problem. Where, specifically the Euler scheme, a traditional method, fails with a positive probability. Hence one has to look for different schemes to solve this issue. One option is the fully implicit Euler scheme, Section 2.3.2, which can handle the high volatility, but instead the issue is to solve a (non-)linear equation. We also present a different method for simulations that works well for SDEs without drift. This method is based on the result of the Theorem 2.1. This method is very efficient numerically, since we only have to generate normal random numbers in each iteration.

Since traditional numerical methods have difficulties with stability, statistical methods regarding parameter estimation are also unstable. If the simulated path exhibits numerical issues then this will transfer to the parameter estimation. We investigate how the certain methods behave after using the fully implicit Euler scheme and also the second method mentioned above. We do not look at fitting any model to real data. We are only interested in the estimation behaviour with certain parameter configurations.

The SDEs we are interested in are presented briefly in Section 2.5.1, 2.5.2 and 2.5.3. The first model mentioned shortly above is CKLS. The second SDE is a so called hyperbolic diffusion, which has no drift. The last SDE has just one parameter, but behaves very wild and can reach extremely large values.

1.2 Method

As mentioned above we will first investigate two different simulation techniques. One called the fully implicit Euler scheme. This scheme is generally a simulation method that finds the next value by solving a non-linear equation. In addition to the implicit scheme we will utilize a clever transformation, Section 2.4, to represent the SDE in a simpler way. This transformation turns the original SDE which has multiplicative noise, i.e. the volatility function $v(x)$ in (2.2) depends on the process, in to a new SDE which instead has constant volatility. This will make the simulations more stable. The other technique we examine is called *change-of-time simulation*, Section 2.3.3. Which, in short, changes the time of a Brownian motion so it becomes a solution to the SDE.

In both cases we will then estimate the parameters. With CKLS, we look at the likelihood ratio method, Section 3.2. In this case we have an expression for the parameters in the drift function. While the parameters for the volatility function are estimated using least squares. For the two other SDEs mentioned above we will just use least squares estimation method, Section 3.3.

1.3 Outline

The outline of the thesis is as follows. In Chapter 2, we present useful notation and an overview of the theory we will use. We also give a description of the Euler schemes, *volatility induced stationarity* and introduce the SDEs. In Chapter 3 the methods regarding simulation and statistical estimation of the SDEs are presented. The results are shown with figures in Chapter 4. Chapter 5 is the final chapter containing conclusions.

2
Theory

In this chapter we will introduce the theory that will be used in this thesis. The structure and content presented assumes knowledge of stochastic calculus, hence some theorems will only be presented for completion. While proofs and more details will be referenced. The main book for theory that will we reference for details is [5], but also some specifics can be found in [7, 8].

2.1 Preliminaries

Set $(\Omega, \mathcal{F}, P)$ to be a probability space. Let $\{W_t\}_{t\geq 0}$, $t \in [0, T]$, be a standard Brownian motion and $I = (l, r)$, $-\infty \leq l < r \leq \infty$ an open interval. For measurable functions $b : I \to \mathbb{R}$ and $v : I \to \mathbb{R}_+$, consider the following Itô process

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t v(X_s)\,\mathrm{d}W_s. \tag{2.1}$$

Where $X_0 = \zeta \in I$ is a real random variable independent of the Brownian motion W. The first integral is a Lebesgue integral and the second integral is an Itô integral. For a detailed theory and construction of the Itô integral we reference to [5]. The process X_t (or just X) is said to have a stochastic differential on $t \in [0, T]$,

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + v(X_t)\,\mathrm{d}W_t, \quad X_0 = \zeta, \tag{2.2}$$

called a stochastic differential equation. The SDE (2.2) is only interpreted in the sense of (2.1). We present the famous Itô's lemma. With notation from [5, p. 111].

Lemma 2.1 (Itô's lemma) *Let X_t have a stochastic differential on the form (2.2) for $0 \leq t \leq T$. For $f \in \mathcal{C}^2$ the stochastic differential of the process $f(X_t)$ exists and is given by*

$$\begin{aligned}\mathrm{d}f(X_t) &= f'(X_t)\,\mathrm{d}X_t + \frac{1}{2}f''(X_t)\,\mathrm{d}[X]_t \\ &= \left(f'(X_t)b(X_t) + \frac{1}{2}f''(X_t)v^2(X_t)\right)\,\mathrm{d}t + f'(X_t)v(X_t)\,\mathrm{d}W_t.\end{aligned}$$

Where the quadratic variation of an Itô-process X is defined to be

Definition 2.1 (Quadratic variation, [5]) *The quadratic variation of the process X_t with differential (2.2) for $0 \leq t \leq T$ is given by*

$$[X]_t = \int_0^t \sigma^2(X_s)\,\mathrm{d}s \tag{2.3}$$

For a process to not explode, i.e., reach infinite values in finite time, the following function has to go to infinity,

$$p(x) = \int_{x_0}^{x} S'(y) \left(\int_{x_0}^{y} dm(z) \right) dy$$

This is called the Feller's test for non-explosion, see e.g. [5].

We introduce the following assumptions, which will be assumed throughout the thesis, first some conditions on the drift and volatility functions of (2.2),

Assumption 2.1 *The drift b is continuous and the volatility σ is (strictly) positive and locally Hölder continuous of order $\frac{1}{2}$.*

Secondly, conditions regarding non-explosions,

Assumption 2.2 *We have $p(l^+) = \lim_{x \downarrow l} p(x) = \infty$ and $v(r^-) = \lim_{x \uparrow r} v(x) = \infty$*

By [4] we know that SDEs that fulfill Assumption 2.1 and 2.2 have a unique strong solution.

2.2 Volatility induced stationarity

In this section we present the concept of *volatility induced stationarity*, vis. Which, briefly, states that the stationary behaviour of the process is determined by the volatility of the process. Hence, as the name suggests, it means that the process' stationarity is induced by its volatile behavior.

The traditional interpretation of SDEs is to view the solutions as a noisy ordinary differential equation, ODE. Which means that the SDE is split in two separate parts, one deterministic and one stochastic. And the long term behaviour of the process is determined by the deterministic part, which can be seen as a solution to an ODE. This together with the noisy part of the original equation gives the complete solution process. In cases where the process exhibits vis the deterministic and stochastic parts interact. This means that the traditional interpretation of SDEs on the general form (2.2) where the drift and volatility is separated, and the long-term stationary behaviour is dictated/imposed by the drift alone is not sufficient. It is not sufficient in the sense that the the numerical scheme (2.4) will fail with a non-zero probability, see Proposition 2.1 in Section 2.3.1 with details in [1].

Stochastic calculus has been researched for decades, but the concept of vis has been a recent topic. In 1997 [3] gave the following definition of vis in the context of CKLS, (2.7),

Definition 2.2 ([3]) *The stationary solution to the SDE given by (2.2) is called volatility induced if, for some $x > 0$,*

$$\int_{x}^{\infty} \frac{b(y)}{v(y)} \, \mathrm{d}y > -\infty.$$

This means, for example, that the solution process can reach large values, even in cases when the drift is positive, and still exhibit a mean-reversion effect due to the volatility-induced stationarity.

The (derivative of the) scale measure S and the speed measure m, the latter is known as the non-normalized distribution for the process, is defined to be

$$S'(x) = \exp\left\{-\int_{x_0}^{x} \frac{2\,b(y)\,\mathrm{d}y}{v(y)^2}\right\} \quad \text{and} \quad \mathrm{d}m(x) = \frac{2\,\mathrm{d}x}{v(x)^2 S'(x)} \quad \text{for } x \in I$$

With these measures the following definition, which is a generalized version of Definition 2.2, made by [1] can be stated.

Definition 2.3 ([1]) *The stationary solution to the SDE given by (2.2) has vis at l $[r]$ if $S'(l^+) < \infty$ $[S'(r^-) < \infty]$. In this case we call l $[r]$ a vis-boundary.*

For a process to not explode in finite time, i.e., not reach $\pm\infty$ for $t \in [0, T]$ the function $p(x)$ must go to infinity as $x \to +\infty$,

$$p(x) = \int_{x_0}^{x} S'(y) \left(\int_{x_0}^{y} dm(z)\right) dy.$$

This means that for stationarity to hold the speed measure m must fulfill $m(I) < \infty$ for a general solution process. In the case of $b(x) = 0$ we have that the derivative of the scale measure is constant, $S'(x) = 1$. Here we note that, by inspection of the speed measure, we must have that $v(x) \to \infty$ as $x \to \infty$. This means that as the solution process reaches high values, $v(x)$ will increase and the speed measure will decrease. This makes it more likely to take smaller values as the time spent at large values will decrease. And the process will return back to where it is (locally) unbiased. I.e. the volatility will push the process back to smaller values, [1].

2.3 Numerical simulation schemes

In this section we present why the regular Euler method, which is also the traditional interpretation of SDE, fails with a non-zero probability. The propositions presented here are proved in full detail in [1]. Firstly we set up the notation used in the discretization of the SDE (2.2).

Following the notation in [6], we set $\{t_n\}_{n=0}^{N}$ to be the (equidistant) mesh of the time interval, i.e.,

$$0 = t_0 < \cdots < t_n < t_{n+1} < \cdots < t_N = T.$$

By equidistant mesh we mean that $t_{n+1} - t_n = h$ for $n = 0, \ldots, N-1$. The process X with SDE (2.2) then satisfies

$$X_{t_{n+1}} = X_{t_n} + \int_{t_n}^{t_{n+1}} b(X_s)\mathrm{d}s + \int_{t_n}^{t_{n+1}} v(X_s)\,\mathrm{d}W_s$$

We also set $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ for the time steps $\{t_n\}_{n=0}^N$. We then approximate the integrals above with the following,

$$\int_{t_n}^{t_{n+1}} b\left(X_s\right)\, \mathrm{d}s \approx b\left(X_{t_n}\right) h$$

and

$$\int_{t_n}^{t_{n+1}} v\left(X_s\right)\, \mathrm{d}W_s \approx v\left(X_{t_n}\right) \Delta W_n$$

This gives rise to the Euler-Maruyama scheme, or just The Euler scheme, for $\{Y_n\}_{n=0}^N$ defined by

$$\begin{aligned} Y_0 &= X_0, \\ Y_{n+1} &= Y_n + b(Y_n)h + v(Y_n)\Delta W_n. \end{aligned} \tag{2.4}$$

This is just one of many schemes. In (2.4) only the last value is used to compute the next in the sequence. One can also use the next value as an unknown variable and solve a (non)-linear equation. This is known as the fully implicit Euler scheme, see section Section 2.3.2 below.

2.3.1 Issues with the Euler scheme

The regular Euler scheme is a numerical approximation of the SDE (2.2) where the drift term and the volatility term are treated separately. The iterative algorithm was presented in (2.4). Where each term gives a contribution "independently" of the other. When trying to simulate these SDEs with this method, one notices rather quickly that the scheme breaks down. By breaking down we mean that the algorithm produces infinite values, although the SDE is non-explosive. In this section we present two propositions that tells us that the Euler scheme is not an option for the SDEs we are interested in.

The first proposition we present is for the general case, when the drift is non-zero. This is the case for the SDEs (2.7) and (2.11).

Proposition 2.1 ([1]) *Let $\{Y_n\}_{n\geq 0}$ be the Euler approximation by (2.4) of (2.2) with $\sigma > 0$. If there exists constants $p > 0$ and $\phi > 0$ such that*

$$\lim_{x\to\pm\infty} \frac{|x|^p |b(x)|}{|v(x)|} = 0 \quad \textit{and} \quad \liminf_{x\to\pm\infty} \frac{|b(x)|}{|x|} \geq \phi$$

then

$$\mathbb{P}\left\{ \bigcap_{n\geq 0} \left\{ |Y_{n+1}| \geq (1+\tilde{\phi})\,|Y_n| \right\} \right\} > 0, \quad \textit{for } \tilde{\phi} \in (0, \phi).$$

Proposition 2.1 states that with a positive probability the Euler scheme detailed in (2.4) will fail, i.e., reach $\pm\infty$ in finite time.

The following proposition describes the setting for the local martingale process, that is the solution to the SDE detailed in Section 2.5.2, will fail using the regular Euler

scheme. The reason why we introduce the method of time-changed simulation in Section 2.3.3 is due to the following proposition.

Proposition 2.2 ([1]) *Let $\{Y_n\}_{n\geq 0}$ be the Euler approximation by (2.4) of (2.2) with zero drift, $b = 0$, and $\sigma > 0$. If*

$$\lim_{x\to\pm\infty} \frac{|x|^p}{v(x)} = 0 \quad \text{for some } p > 1,$$

then

$$\mathbb{P}\left\{\bigcap_{n\geq 0}\left\{|Y_{n+1}| \geq (1+\tilde{\phi})\,|Y_n|\right\}\right\} > 0, \quad \text{for } \tilde{\phi} > 0.$$

Mainly, because of these results we will use the fully implicit Euler scheme presented in the next section for the SDEs (2.7) and (2.11). While the method described in Section 2.3.3, *change of time* simulation, will be used to simulate the SDE without drift (2.10).

2.3.2 Implicit Euler scheme

The implicit Euler method is the following method, which gives a simulated path $\{Y_n\}_{n=0}^N$, defined by

$$\begin{aligned} Y_0 &= X_0 \\ Y_{n+1} &= Y_n + \hat{b}(Y_{n+1})h + v(Y_{n+1})\Delta W_n \end{aligned}$$

Where for each time step t_n we have to solve a (non)-linear equation using the next value, Y_{n+1} as the unknown variable. Hence the function takes the form,

$$y = Y_n + (b(y) - v(y)v'(y))\,h + v(y)\Delta W_n$$

and setting the root as the new value Y_{n+1}.

2.3.3 Change-of-time simulation

In this section we give a description of a simulation technique for SDEs where the solution process is a continuous local martingale,

$$\mathrm{d}X_t = v(X_t)\,\mathrm{d}W_t.$$

The simulation method builds on the result from Theorem 2.1. The theorem states, in short, that a continuous (local) martingale is a time-changed Brownian motion. Where time is then scaled by the quadratic variation of the (local) martingale process. Which in practice means that we can get a (weak) solution to a complicated SDE by simulating a Brownian motion instead of other more complex and time consuming methods. To understand this simulation technique we start with defining the following concepts, [5], used in the statement of the theorem.

Definition 2.4 (Stopping time) *A non-negative random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ is called a stopping time with respect to $\mathcal{F}$ if for each t,*

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

Set

$$\tau_t = \inf\{s : [X]_s > t\}$$

To see that $[X]_t$ is stopping times for $\mathcal{F}_{\tau_t}$, note that $\{[X]_s \geq t\} = \{\tau_t \geq s\}$.

Let A_0 bet the set of all càdlàg and non-decreasing functions f defined on $[0, \infty)$ and taking values in $[0, \infty)$ with $f(0) = 0$.

Definition 2.5 (Change of time) *A stochastic process $\{\tau_t\}_{t\in[0,\infty)}$ with trajectories in A_0 is called a change of time (or time change) if the random variable τ_s is a stopping time, for each $s \geq 0$.*

Now the theorem in question is stated in several books but we chose to present it in the same form as [5, p. 204], however they only give an outline of the proof. For a detailed one we reference to [7, p. 181] or [8, p. 64].

Theorem 2.1 (Dambis, Dubins-Schwarz, [5]) *Let X be a continuous martingale, null at zero, such that $[X]_t$ is non-decreasing to ∞, and τ_t defined by $\tau_t = inf\{s : [X]_s > t\}$. Then the process $W_t = X_{\tau_t}$ is a Brownian motion with respect to the filtration $\mathcal{F}_{\tau_t}$. Moreover, $[X]_t$ is a stopping time with respect to this filtration, and the martingale X can be obtained from the Brownian motion W by the change of time $X_t = W_{[X]_t}$. The result also holds when X is a continuous local martingale.*

The algorithm to simulate a SDE with this method is the following, first note that we can approximate the quadratic variation, defined as Definition 2.1, assuming equidistant time step h, with

$$[X]_{t_n} \approx \sum_{i=1}^{n} v^2(X_{i-1})h \tag{2.5}$$

By Theorem 2.1 we have that a (weak) solution to (2.10) is given by

$$X_t = W_{[X]_t}$$

So the iterative method is the following

$$\begin{aligned} X_{t_{n+1}} &= X_{t_{n+1}} - X_{t_n} + X_{t_n} \\ &= W_{[X]_{t_{n+1}}} - W_{[X]_{t_n}} + X_{t_n} \end{aligned}$$

using that $W_t - W_s \sim \mathcal{N}(0, t - s)$ and the approximation (2.5) we arrive at the following algorithm for $n = 1, \ldots, N - 1$

$$\begin{aligned} Y_0 &= X_0 \\ Y_{n+1} &= Z_n + Y_n. \end{aligned}$$

where $Z_n \sim \mathcal{N}\left(0, v^2(Y_n)h\right)$ is independent of earlier iterations. In the case where one would use non-equidistant time-steps, the modification is to replace h with $t_{n+1} - t_n$ above.

2.4 Transformations

We want to find a transformation of the SDE that removes the multiplicative noise of the SDE and represents it in a new form with additive noise instead and where the drift fulfills a one-sided Lipschitz condition.

Definition 2.6 (One-sided Lipschitz) *A function $f : \mathbb{R} \to \mathbb{R}$ is said to fulfill a one-sided Lipschitz condition if there exists a constant C such that*

$$(x - y)(f(x) - f(y)) \leq C|x - y|^2 \quad \text{for } x, y \in I.$$

If Assumption 2.1 holds and the volatility in (2.2) has an absolutely continuous derivative, then the transformation[1]

$$F(x) = \int^x \frac{1}{v(u)} \mathrm{d}u,$$

where the lower point of integration is arbitrary, is strictly increasing and the transformed process $Z_t = F(X_t)$ satisfies the following SDE now with additive noise

$$\mathrm{d}Z_t = \varphi(Z_t)\,\mathrm{d}t + \mathrm{d}W_t, \quad Z_0 = F(\zeta),$$

where

$$\varphi(z) = \frac{b\left(F^{-1}(z)\right)}{v\left(F^{-1}(z)\right)} - \frac{1}{2}v'\left(F^{-1}(z)\right)$$

This expression can be obtained by using Itô's lemma, Lemma 2.1, with $F' = f$

$$f(z) = \frac{1}{v(z)}, \quad f'(z) = -\frac{v'(z)}{v^2(z)}$$

And written out on complete form, the transformed SDE,

$$\mathrm{d}Z_t = \left(\frac{b\left(F^{-1}(Z_t)\right)}{v\left(F^{-1}(Z_t)\right)} - \frac{1}{2}v'\left(F^{-1}(Z_t)\right)\right)\mathrm{d}t + \mathrm{d}W_t, \quad Z_0 = F(\zeta) \tag{2.6}$$

2.5 Stochastic differential equations

Here we give a short introduction to the three different SDEs we will be investigating.

2.5.1 CKLS

This SDE is a generalized version of the CIR model, i.e., we have $\gamma > 1$. The type of CKLS models we want to investigate is where $\alpha > 0$, $\beta \approx \pm\alpha$ and $\gamma > 1$. Specifically we consider parameter values where the volatility elasticity, γ, is large, $\gamma \gg 1$, with example values for $\gamma = 5, 10, 50$. The SDE was originally formulated in the famous

[1] Called the *Lamperti* transform

paper [2]. There are several representations[2] of CKLS, however the one we will use is,

$$\mathrm{d}X_t = (\alpha + \beta X_t)\,\mathrm{d}t + \sigma X_t^{\gamma}\mathrm{d}W_t, \quad X_0 = \zeta > 0. \tag{2.7}$$

The transformation mentioned in Section 2.4 for this model is the following, $f(z) = z^{1-\gamma}$. Hence applying Itô's lemma, Lemma 2.1, the new representation for the SDE (2.7) is,

$$\mathrm{d}Z_t = \varphi(Z_t)\,\mathrm{d}t + (1-\gamma)\sigma\,\mathrm{d}W_t \tag{2.8}$$

where the drift function

$$\varphi(z) = \alpha(1-\gamma)z^{\frac{\gamma}{\gamma-1}} + \beta(1-\gamma)z + \frac{1}{2}\sigma^2\gamma(\gamma-1)z^{-1} \tag{2.9}$$

To solve this SDE with numerical methods, we use the fully implicit Euler scheme, Section 2.3.2, together with the transformation above. Hence the scheme is producing the simulated path $\{Z_n\}_{n=0}^N$, by

$$\begin{aligned} Z_0 &= X_0 \\ Z_{n+1} &= Z_n + \varphi(Z_{n+1})h + (1-\gamma)\sigma\Delta W_n \end{aligned}$$

The non-linear equation to be solved is, $f(Z_{n+1}) = Z_n$, where

$$f(z) = z - \varphi(z)h - (1-\gamma)\sigma\Delta W_n.$$

After we have simulated $\{Z_n\}_{n=0}^N$, we transform back with $f^{-1} = y^{\frac{1}{1-\gamma}}$ to get the path $\{Y_n\}_{n=0}^N$, where $f^{-1}(Z_n) = Y_n$.

2.5.2 Hyperbolic diffusion

This SDE is formulated in [3] as a model for stock prices. Following the notation they use, we start by introducing the function.

$$v^2(x) = \sigma^2 \exp\left[\alpha\sqrt{\delta^2 + (x-\mu)^2} - \beta(x-\mu)\right]$$

where $\alpha > |\beta| \geq 0$, $\delta, \sigma > 0$ and $\mu \in \mathbb{R}$, $v : \mathbb{R} \to (0, \infty)$ is a strictly positive function.

The SDE then takes the form

$$\mathrm{d}X_t = v(X_t)\,\mathrm{d}W_t, \quad X_0 = \zeta \in \mathbb{R} \tag{2.10}$$

The transformation in Section 2.4 is not an option since the transformed SDE will not have a one-sided Lipschitz drift function. The fully implicit Euler scheme is intractable, since we have to solve a very complicated function in each time step. We instead turn to the simulation method detailed in Section 2.3.3 to simulate a solution path to this SDE. We show a sample path using this scheme for $(\alpha, \beta, \delta, \gamma, \sigma) = (5, -4, 1, 7, 0.05)$ of (2.10), in Figure 2.3. Where we, for example, can see how volatile the process is, but it keeps mostly around its stationary level. And the excursions are short and quickly the process comes back up or back down.

[2]For example $dX_t = \theta(\kappa - X_t)dt + \sigma X_t^{\gamma} dW_t$ with $\kappa = -\frac{\alpha}{\beta}$, $\theta = -\beta$.

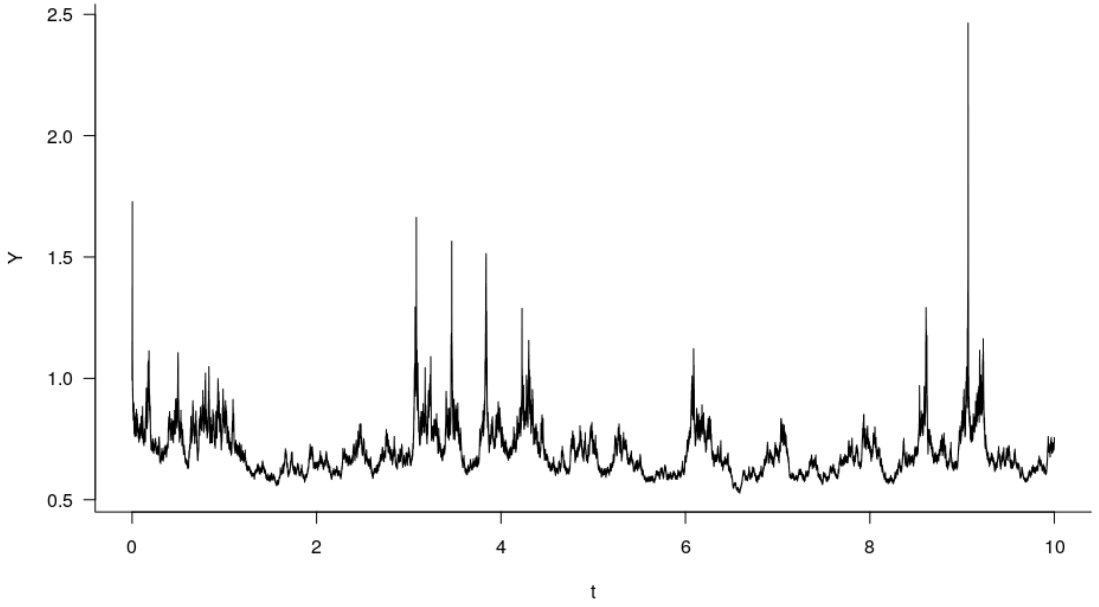

Figure 2.1: *Simulated path of the CKLS, using original representation (2.7)*

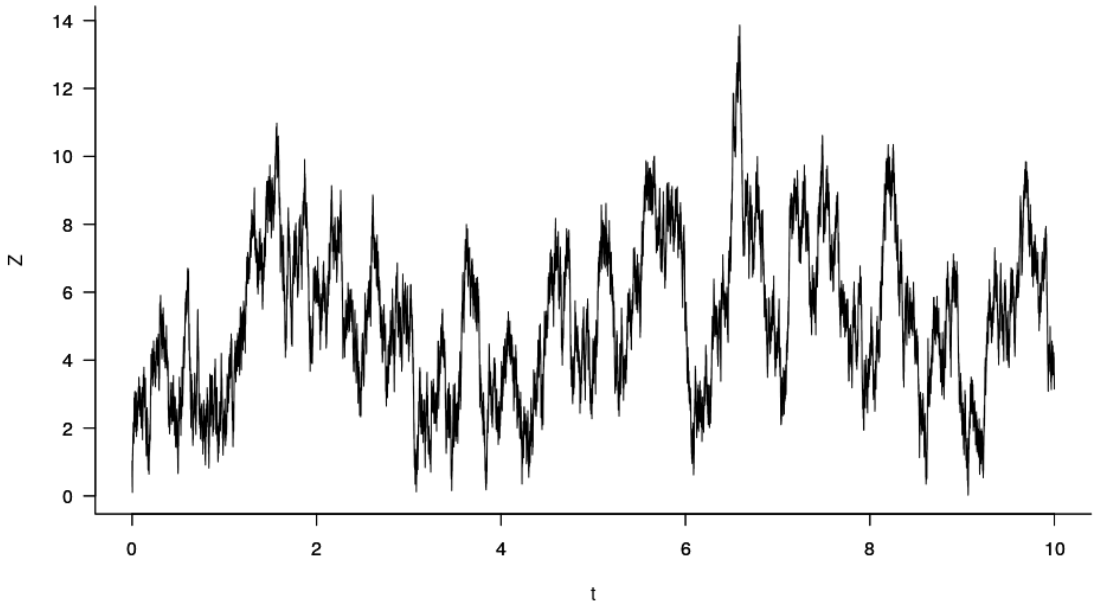

Figure 2.2: *Simulated path of the CKLS, transformed representation (2.8)*

2.5.3 A family of heavy tailed SDE

The last model we will investigate is the following SDE. With only one parameter, $\alpha < \frac{1}{3}$,

$$\mathrm{d}X_t = 3X_t^{\alpha}\mathrm{d}t + 3X_t^{2/3}\mathrm{d}W_t, \quad X_0 = \zeta > 0 \tag{2.11}$$

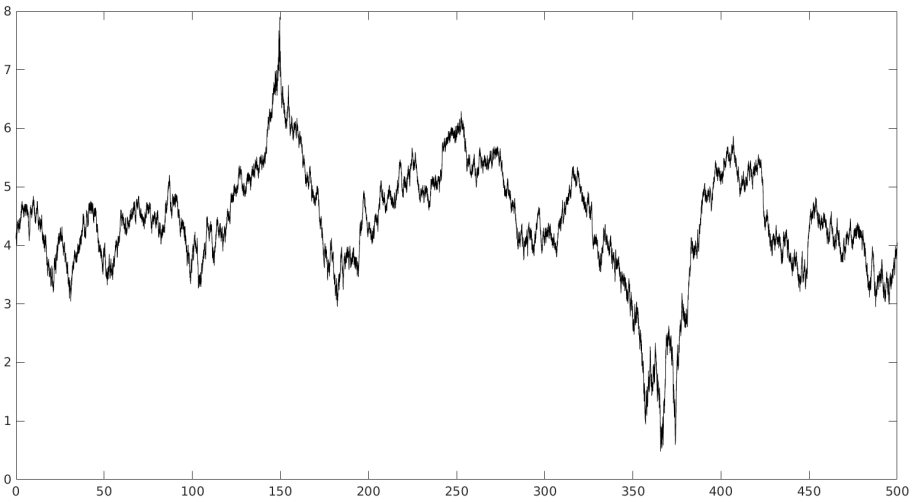

Figure 2.3: *Sample path of (2.10) using the parameter set* $(\alpha, \beta, \delta, \mu, \sigma) = (5, -4, 1, 7, 0.05)$ *the scheme from Section 2.3.3 with* $T = 500$.

We simulate this SDE with the fully implicit Euler scheme, giving the path $\{Y_n\}_{n=0}^N$,

$$\begin{aligned} Y_0 &= X_0 \\ Y_{n+1} &= Y_n + 3Y_{n+1}^{\alpha} h + 3Y_{n+1}^{2/3} \Delta W_n. \end{aligned}$$

Giving the non-linear equation, $f(Y_{n+1}) = Y_n$, where

$$f(y) = y - 3y^{\alpha} h - 3y^{2/3} \Delta W_n$$

A simulated path for this SDE, with $\alpha = -10$ is shown in Figure 2.4. If we let the time interval increase, we see that this SDE can take excursions to very large values, and come back down.

We can use the transform from Section 2.4, $f(z) = z^{\frac{1}{3}}$, with inverse transform, $f^{-1}(y) = y^3$. This gives a representation of the original SDE which has multiplicative noise, in to a SDE with additive noise, i.e.,

$$\mathrm{d}Z_t = \varphi(Z_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$$\mathrm{d}Z_t = \left(Z_t^{3\alpha-2} - Z_t^{-1}\right)\,\mathrm{d}t + \mathrm{d}W_t \tag{2.12}$$

We simulate with the transformed SDE, then go back with the inverse transformation $g(y) = y^3$. In Figure 2.12 we show a path of the transformed SDE.

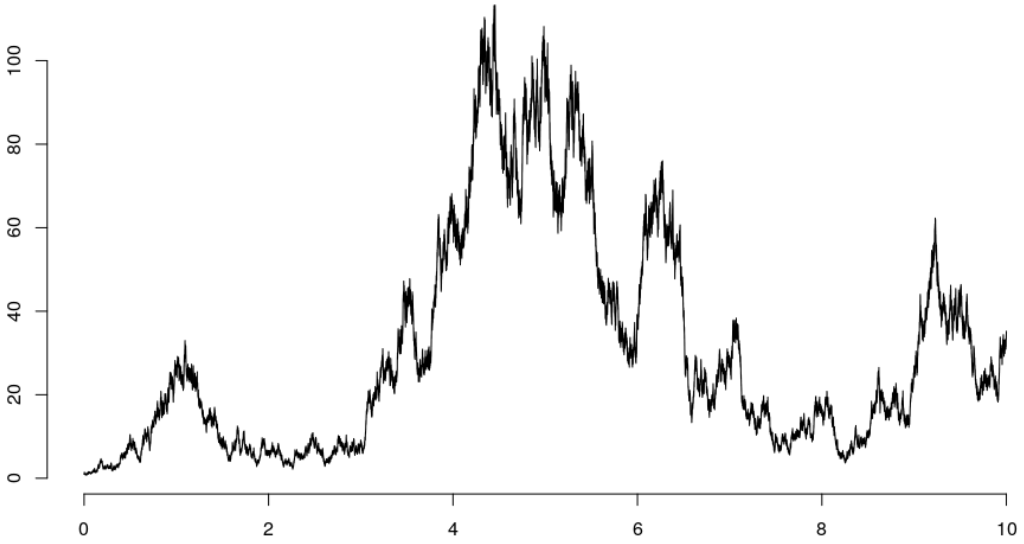

Figure 2.4: *Simulated path of the original representation (2.11)*

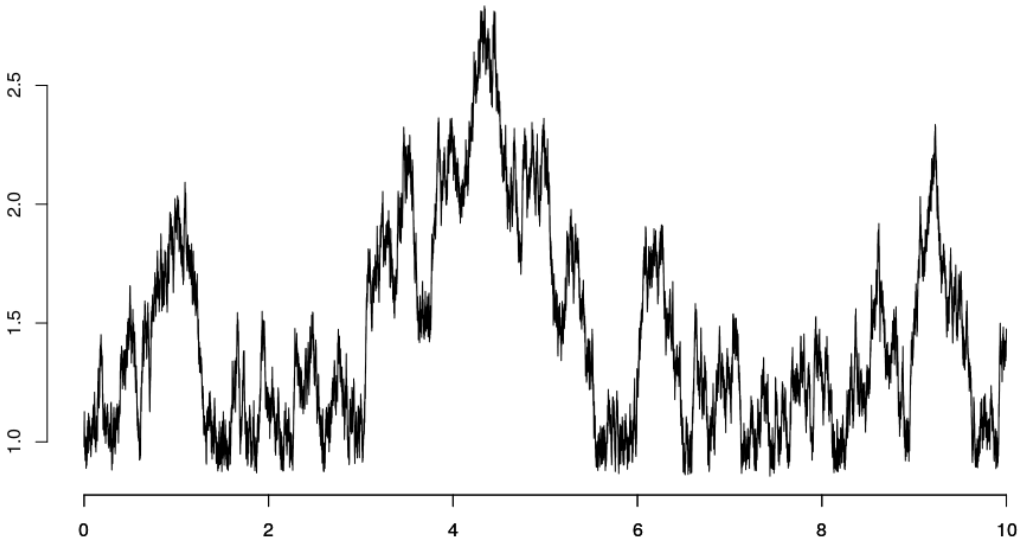

Figure 2.5: *Simulated path of the transformed representation (2.12)*

3
Methods

In this Chapter we will look at some statistical methods for estimation of parameters. Specifically classical maximum likelihood methods. For CKLS, the expressions for the drift parameters the likelihood equations take a complicated form. This method is well detailed in Chapter 10.6 in [5]. While the diffusion parameters in CKLS, Section 2.5.1, together with the parameters for the other two models Section 2.5.2 and 2.5.3, will be estimated using least square methods.

3.1 Simulation

In this section we describe the method used to check convergence of the numerical method. The method works in the following way. First we generate a Brownian motion, $\{W_n\}_{n=0}^N$, with the number of time steps, N,[1]. With size of the time step $h = 1/N$. Then we simulate a path of the SDE using a suitable method, Section 2.3.2 or Section 2.3.3. The next step is to half the number of time steps in the Brownian motion. This is done by adding the adjacent values of the Brownian motion, $\hat{W}_n := W_n + W_{n-1}$. Where $\{\hat{W}_n\}_{n=0}^N$ denotes this new Brownian motion, with $\frac{N}{2}$ number of steps.

After constructing this new Brownian motion, we redo the numerical scheme. This time we simulate a path with $\frac{N}{2}$ time steps. This process is repeated for an appropriate number of iterations. We show an example of the Brownian motions we get from this method, see Figure 3.1. If we see that the numerical simulations are similar, we know that the scheme has converged.

3.2 Likelihood ratio method

Here we describe the likelihood ratio method, LR-method. This method can be used to estimate parameters, specifically we will use the method to estimate the drift parameters for CKLS. Generally speaking it is a ratio between two models, with two different drift functions, b_1 and b_2. Where a decision is made based on if the ratio is larger than one. We will use this method to find an estimate for α and β parameters in the case of CKLS, (2.7). For details of derivation of the likelihood ratio, based on the Radon-Nikodym theorem we reference to [5]. The function in

[1]We used $N \approx 2^{18}$, due to computational limitations. But here we use $N = 2^{12}$ for better visualization.

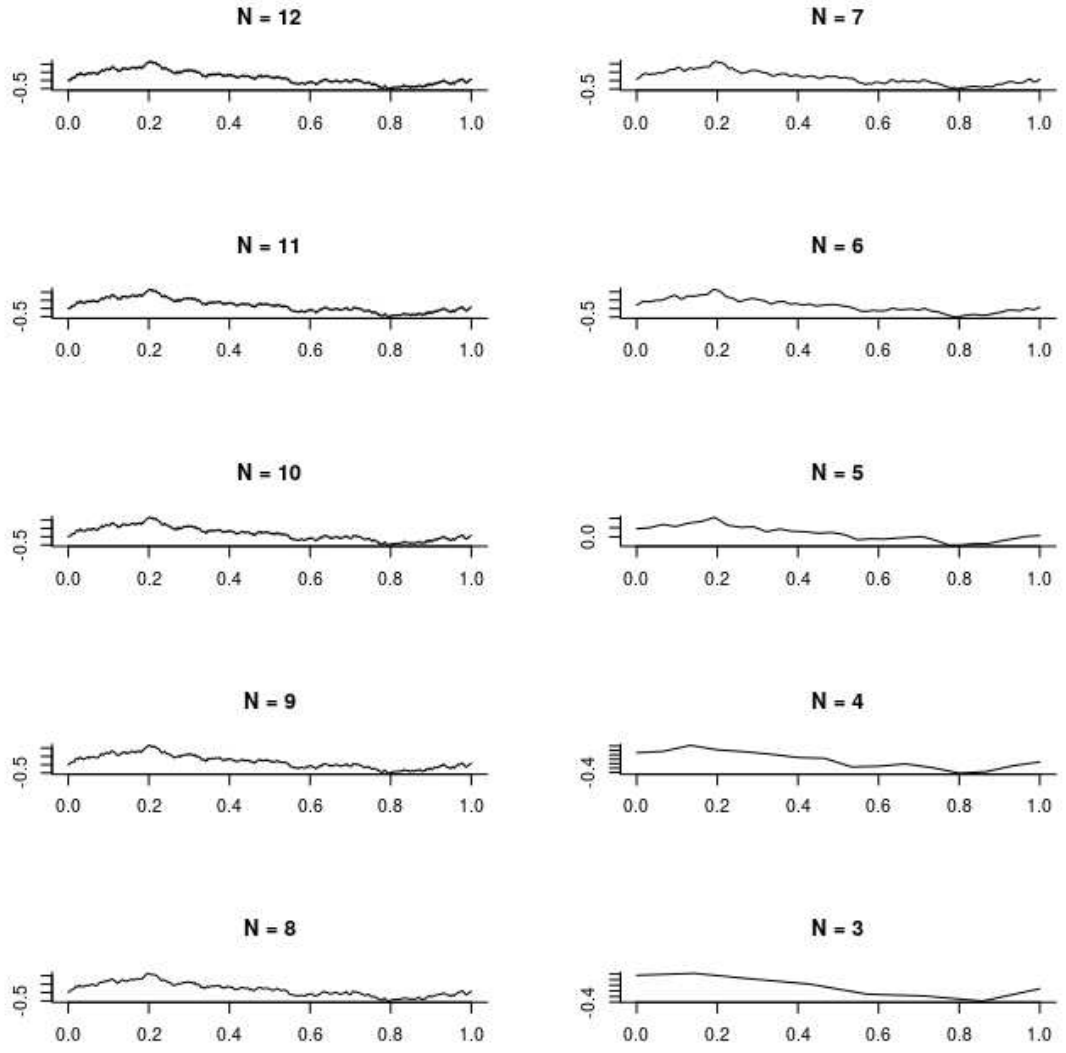

Figure 3.1: *Brownian motion where the steps are decreased from* 2^{12} *down to* 2^3

question is the following, where θ denotes the parameters of interest, which the functions b_1, b_2 and v depends on,

$$\Lambda(\theta, \{X\}_{t=0}^{T}) = \exp\left[\int_0^T \frac{b_2(X_t) - b_1(X_t)}{v^2(X_t)}\,\mathrm{d}X_t - \frac{1}{2}\int_0^T \frac{b_2^2(X_t) - b_1^2(X_t)}{v^2(X_t)}\,\mathrm{d}t\right]$$

We will derive expressions for the drift parameters, first in the original representation, (2.7). Then, as we will simulate the transformed version, we give LR-estimates in for the transformed equation, (2.8). To arrive at the desired equations, we begin to set the b_1 function to zero. We arrive at the following, with $\theta = (\alpha, \beta)$,

$$\Lambda\left(\alpha, \beta, \{X_t\}_{t=0}^{T}\right) = \exp\left[\int_0^T \frac{(\alpha + \beta X_t)}{\sigma^2 X_t^{2\gamma}}\,\mathrm{d}X_t - \frac{1}{2}\int_0^T \frac{(\alpha + \beta X_t)^2}{\sigma^2 X_t^{2\gamma}}\,\mathrm{d}t\right]$$

First we use that the logarithm of a function has the same maximum, and we also expand the parentheses inside the integrals, giving us the following,

$$\lambda(\alpha,\beta,\{X_t\}_{t=0}^T) = \alpha\int_0^T \frac{\mathrm{d}X_t}{X_t^{2\gamma}} + \beta\int_0^T \frac{X_t\,\mathrm{d}X_t}{X_t^{2\gamma}} - \frac{1}{2}\left(\alpha^2\int_0^T \frac{\mathrm{d}t}{X_t^{2\gamma}} + \alpha\beta\int_0^T \frac{X_t\,\mathrm{d}t}{X_t^{2\gamma}} + \beta^2\int_0^T \frac{X_t^2\,\mathrm{d}t}{X_t^{2\gamma}}\right)$$

Now taking derivatives with respect to α and β, respectively, and setting each one to zero we get the following expressions,

$$\hat{\alpha} = \frac{\int_0^T X_t^{-2\hat{\gamma}}\,\mathrm{d}X_t \int_0^T X_t^{2-2\hat{\gamma}}\,\mathrm{d}t - \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}X_t \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t}{\int_0^T X_t^{2-2\hat{\gamma}}\,\mathrm{d}t \int_0^T X_t^{-2\hat{\gamma}}\,\mathrm{d}t - \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t}$$

and

$$\hat{\beta} = \frac{\int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}X_t \int_0^T X_t^{-2\hat{\gamma}}\,\mathrm{d}t - \int_0^T X_t^{-2\hat{\gamma}}\,\mathrm{d}X_t \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t}{\int_0^T X_t^{2-2\hat{\gamma}}\,\mathrm{d}t \int_0^T X_t^{-2\hat{\gamma}}\,\mathrm{d}t - \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t \int_0^T X_t^{1-2\hat{\gamma}}\,\mathrm{d}t}$$

Using the scheme detailed in Section 2.3.2, we simulate a path denoted $\{Y_n\}_{n=0}^N$, which we use to estimate the integrals in (3.2) and (3.2). Specifically, we estimate the integrals by, with appropriate function f,

$$\int_0^T f\left(X_s\right)\,\mathrm{d}s \approx \sum_{n=0}^{N-1} f(Y_n)h,$$

and

$$\int_0^T f\left(X_s\right)\,\mathrm{d}X_t \approx \sum_{n=0}^{N-1} f(Y_n)\left(Y_{n+1}-Y_n\right).$$

The estimate $\hat{\gamma}$ for the γ parameter is found by minimizing the function,

$$(\hat{\sigma},\hat{\gamma}) = \arg\min_{\sigma,\gamma}\left(QV(T) - \sigma^2\int_0^T Y_s^{2\gamma}\,\mathrm{d}s\right)^2$$

where $QV(T)$ is denoting the observed quadratic variation,

$$QV(T) = \sum_{i=1}^{N-1}\left(Y_n - Y_{n-1}\right)^2 \tag{3.1}$$

and the theoretical quadratic variation expressed with the integral in (3.1) is estimated by

$$\sigma^2\int_0^T Y_s^{2\gamma}\,\mathrm{d}s \approx \sigma^2\sum_{n=0}^{N-1} Y_n^{2\gamma}h.$$

The volatility σ parameter is still estimated simultaneously together with the γ parameter, but the estimate $\hat{\sigma}$ is not used here in the estimation for α and β.

Since we instead want to use the representation (2.6), with additive noise, we derive an expression in this setting as well. Now the LR-method takes the following form, continuing to use the function $\varphi(z)$ in (2.9) for more compact notation,

$$\lambda(\alpha, \beta, \{Z_t\}_{t=0}^T) = \exp\left[\int_0^T \frac{\varphi(Z_t)}{\sigma^2(1-\gamma)^2}\,\mathrm{d}Z_t - \frac{1}{2}\int_0^T \frac{\varphi^2(Z_t)}{\sigma^2(1-\gamma)^2}\,\mathrm{d}t\right]$$

Similarly as above, taking the derivatives and setting each expression to zero. Then solving for α and β respectively we get the following equations. Using estimates for γ and σ found from least squares estimate where T is the length for the time interval. Using that $QV(T)$ is the observed quadratic variation, the estimations found by minimizing the following function,

$$(\hat{\sigma}, \hat{\gamma}) = \arg\min_{\sigma,\gamma} \left(QV(T) - \sigma^2(1-\gamma)^2 T\right)^2. \tag{3.2}$$

We denote the following integrals, to make the final expression as compact and readable as possible,

$$A = \int_0^T Z_t^{\frac{\hat{\gamma}}{\hat{\gamma}-1}}\,\mathrm{d}Z_t, \qquad B = \int_0^T Z_t\,\mathrm{d}Z_t, \qquad C = \int_0^T Z_t^{\frac{2\hat{\gamma}}{\hat{\gamma}-1}-1}\,\mathrm{d}t,$$
$$D = \int_0^T Z_t^2\,\mathrm{d}t, \qquad E = \int_0^T Z_t^{\frac{\hat{\gamma}}{\hat{\gamma}-1}+1}\,\mathrm{d}t, \qquad F = \int_0^T Z_t^{\frac{\hat{\gamma}}{\hat{\gamma}-1}-1}\,\mathrm{d}t.$$

The new expression for α and β are the following

$$\hat{\alpha} = \frac{2BE - 2AD + (\hat{\gamma}-1)\hat{\gamma}\hat{\sigma}^2\,(DF - ET)}{2(\hat{\gamma}-1)\,(CD - E^2)}, \tag{3.3}$$

and

$$\hat{\beta} = -\frac{2BC - 2AE + (\hat{\gamma}-1)\hat{\gamma}\hat{\sigma}^2\,(EF - CT)}{2(\hat{\gamma}-1)\,(CD - E^2)}. \tag{3.4}$$

And we estimate the integrals as we mentioned above. The results after estimating the drift parameters using these two expressions are presented in Section 4.2.1.

3.3 Least squares estimation

To estimate the volatility parameters in CKLS we use a least squares estimation. Which we use to minimize the squared difference between the observed quadratic variation, denoted $QV(T)$, and the theoretical quadratic variation, (2.3). This gives the following expression, where we denote the parameter set to be estimated, θ. For the SDE in Section 2.5.2, we use the least squares method to estimate all of the parameters at once.

$$\hat{\theta} = \arg\min_{\theta} \left(QV(T) - [Y]_T(\theta)\right)^2$$

For the family heavy tailed SDE, Section 2.5.3, we also use least squares estimation. However we instead minimize the following squared difference,

$$\hat{\alpha} = \arg\min_{\alpha} \sum_{n=0}^{N-1} \left(Y_{n+1} - Y_n - \left(Y_n^{3\alpha-2} - 1/Y_n\right) h\right)^2$$

4

Results

In this Chapter we present the results. First we show convergence figures for each of the SDEs, which says that our numerical schemes have converged. Secondly, we present the results from the statistical methods from Section 3.2 and 3.3. This is done by the use of box plots.

4.1 Simulation results

Firstly we show the convergence results for CKLS, (2.5.1). We give a figure for the parameter set $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$, but we have tried many different choices and the result is the same.

When we use the transformed representation, (2.9), the result remains. Using the same parameter set and the same seed, the result is shown in Figure 4.2. For the SDE discussed in Section 2.5.3, we see the convergence of the numerical method, Figure 4.3. With the transformed representation, (2.12), the result is the same, as shown in Figure 4.4.

Finally, we show the convergence results for the SDE in Section 2.5.2. In Figure 4.5 we see that the method in Section 2.3.3 results in a convergent numerical scheme.

4.2 Statistical results

Here we present the results from the statistical methods described in Chapter 3.

4.2.1 CKLS

We see in Figure 4.6 that the estimation for the drift parameters is fairly stable around the true value, $(\alpha, \beta = (1, 1)$, with some extreme results. We have simulated 100 different paths using the representation (2.8), by the fully implicit Euler method, Section 2.3.2, with many different combinations of parameters. Here we include four different ones, the parameter sets $(\alpha, \beta, \sigma, \gamma) = (1, \pm 1, 1, 10)$, and $\gamma = 50$, With larger values α or $|\beta|$ the same behaviour remains, hence we show only these results. The estimation for the drift parameters is done by using the likelihood ratio equations (3.3) and (3.4). As we see in Figure 4.6 and Figure 4.7, the estimation is generally stable. With only a few extremes. In Figure 4.7 we instead simulate 100

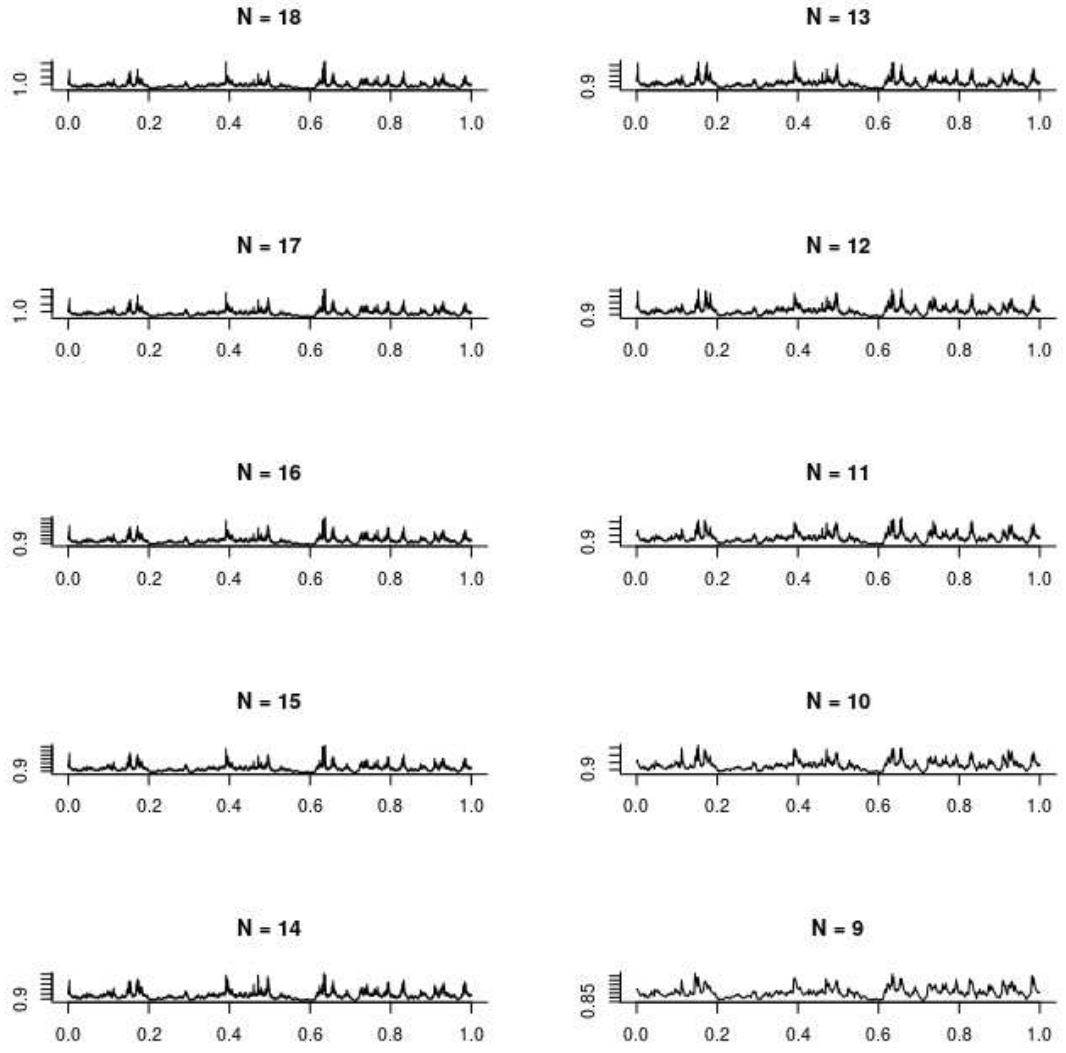

Figure 4.1: *Plot of the convergence results for the CKLS model using the parameter set* $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$.

paths but with negative β. The estimation is still stable. The estimation for the volatility parameters using least squares, (3.2), σ and γ is very precise, with a very small variation around the true value.

While we still have to estimate both the sigma and the gamma parameters in (3.3) and (3.4), the estimation for them are so precise that this does not affect the estimation, compared to using the true values.

We show also results from simulations where we increase the γ parameter to 50, but keep the other parameters the same. Now we see that when both the drift parameters are positive the estimation is very difficult using the likelihood ratio method. As we can see in Figure 4.8, the estimated drift parameters are very spread out. We even get that on average the β parameter is estimated negatively, while the

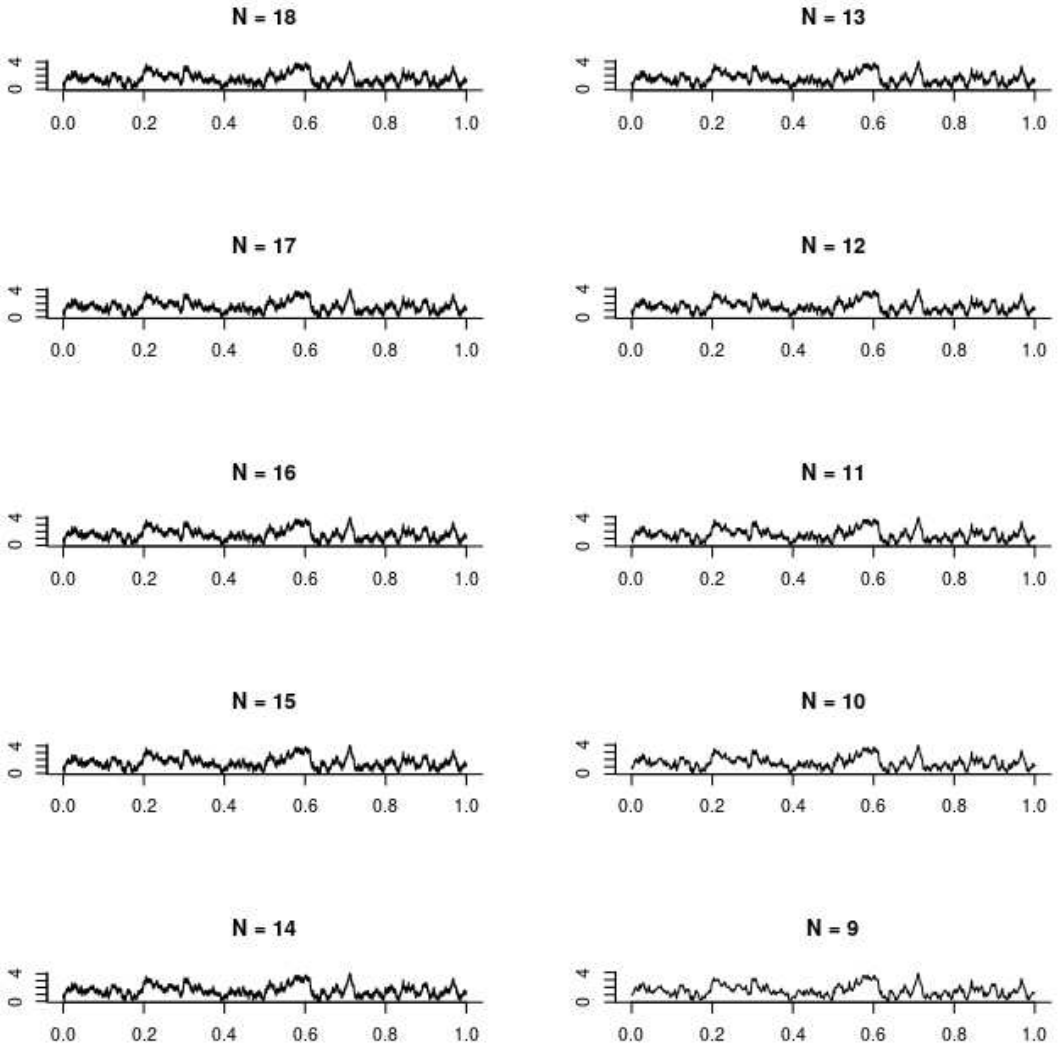

Figure 4.2: *Plot of the convergence results for the transformed CKLS model using the parameter set* $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$.

true is positive. When we instead set the true β parameter to -1, which means we have a mean reversion also in the drift. The estimation is similar to when $\gamma = 10$.

4.2.2 Hyperbolic diffusion

We have tried many different combinations of parameters, however we show only $(\alpha, \beta, \delta, \mu, \sigma) = (7, -6, 1, 7, 0.05)$ as the results is the same. In Figure 4.10 we see that the parameter estimation is fairly stable with the exception of the α parameter, denoted by a in the figure. For every simulation we have run, the α parameter is estimated negatively. We try to investigate where the issues in estimating α is. First we use the real values in place for $(\beta, \delta, \mu, \sigma)$ and only estimate α. When we do this see that estimation is stable, with very small variation between the estimated and the real value. So we try to do this for different combinations where we estimate α

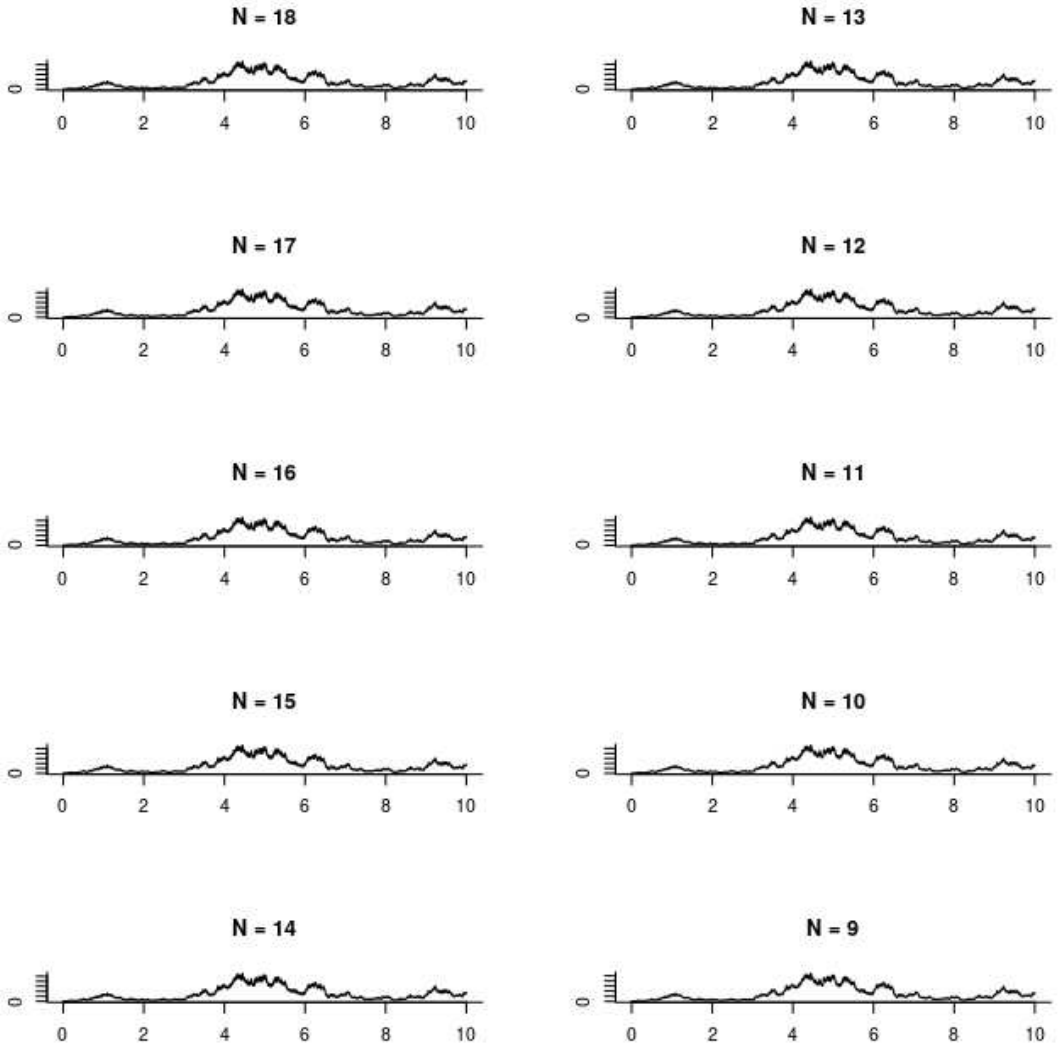

Figure 4.3: *Plot of the convergence results for the family of heavy tailed SDE using* $\alpha = -10$

together with the some of the other parameters, while using the true value for the remaining parameters. What we see is that when we try to combine the estimation with one or more unknown parameter estimations, the estimation for α is very poor and unstable from the true value.

4.2.3 A family of heavy tailed SDE

The least squares estimate for the heavy tailed SDE (2.11) is fairly stable, with only a few paths being estimated to larger negative values, as seen in Figure 4.11. The paths where α is estimated to larger negative values are ones that spend most of the time away for the stationary level. We also note that for positive α, for example with values around $\frac{1}{10}$, the estimation is as stable and gives similar results as for large negative values.

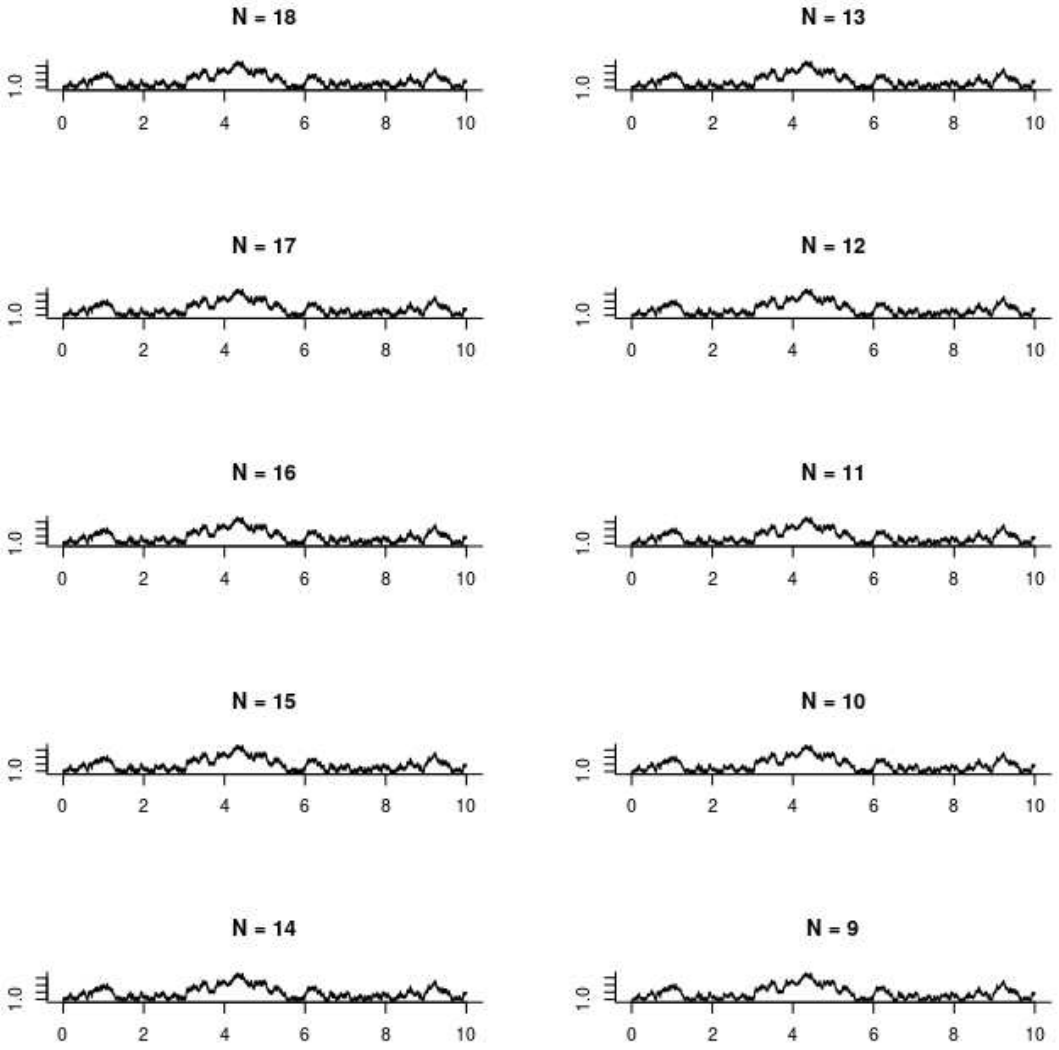

Figure 4.4: *Plot of the convergence results for transformed SDE, (2.12), using* $\alpha = -10$

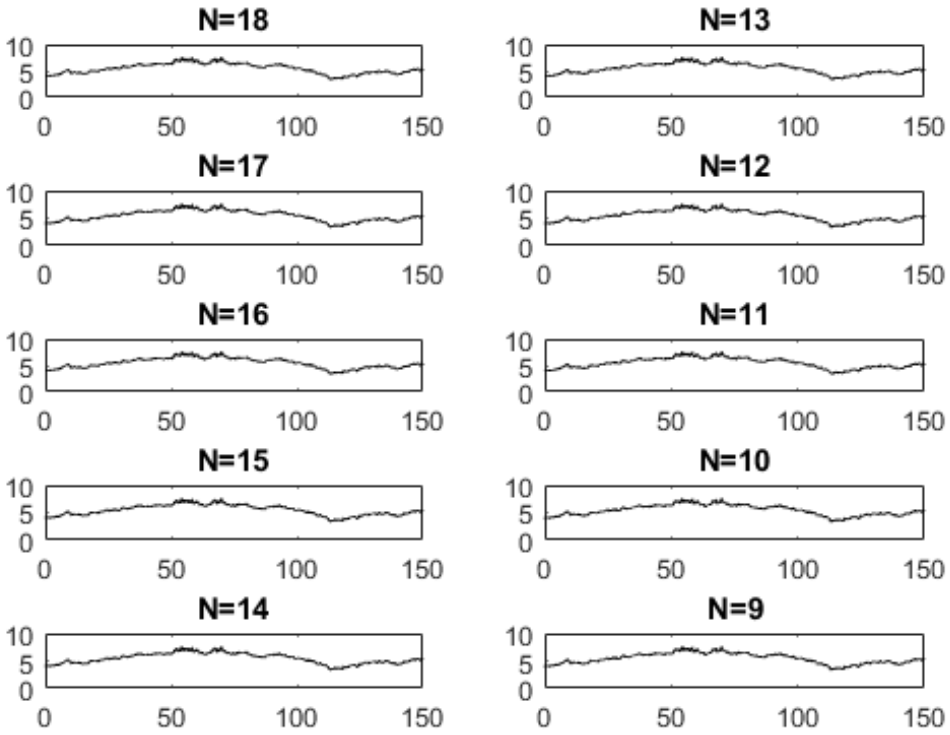

Figure 4.5: *Plot of the convergence results for the hyperbolic diffusion using* $(\alpha, \beta, \delta, \mu, \sigma) = (7, -6, 0.05, 7, 1)$

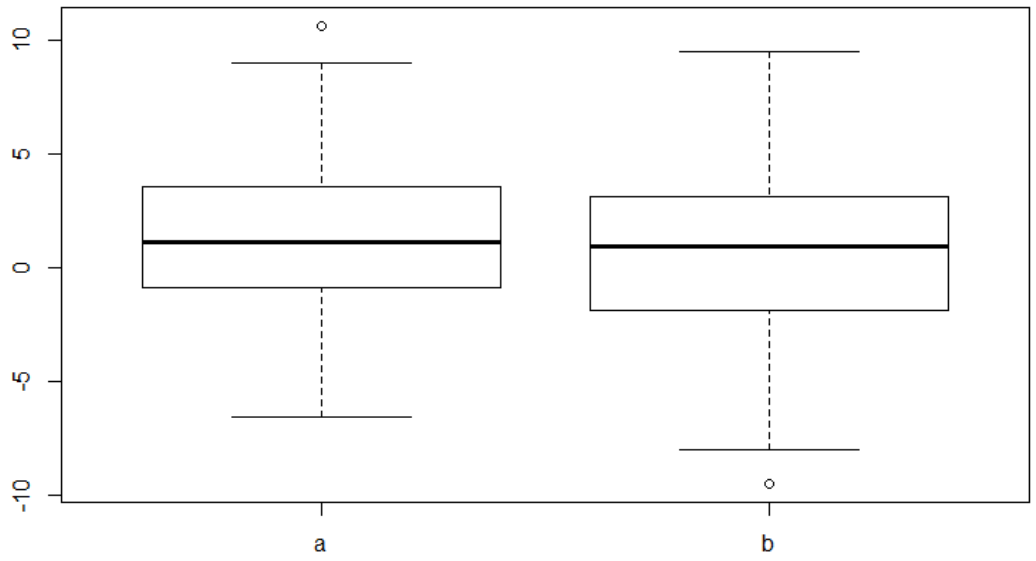

Figure 4.6: *Drift parameters estimated by (3.3) and (3.4) with* $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 10)$.

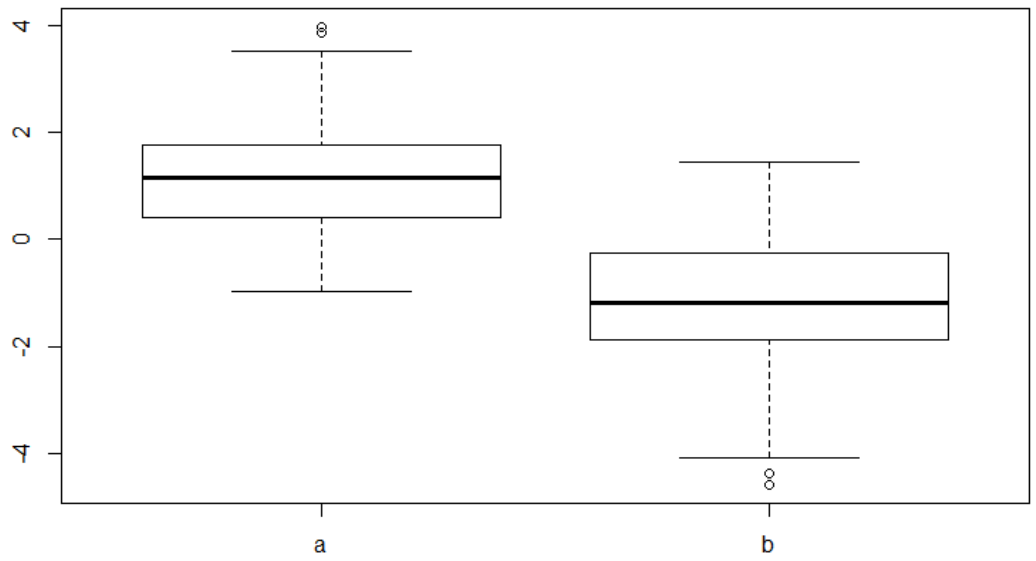

Figure 4.7: *Drift parameters estimated by (3.3) and (3.4) with* $(\alpha, \beta, \sigma, \gamma) = (1, -1, 1, 10)$.

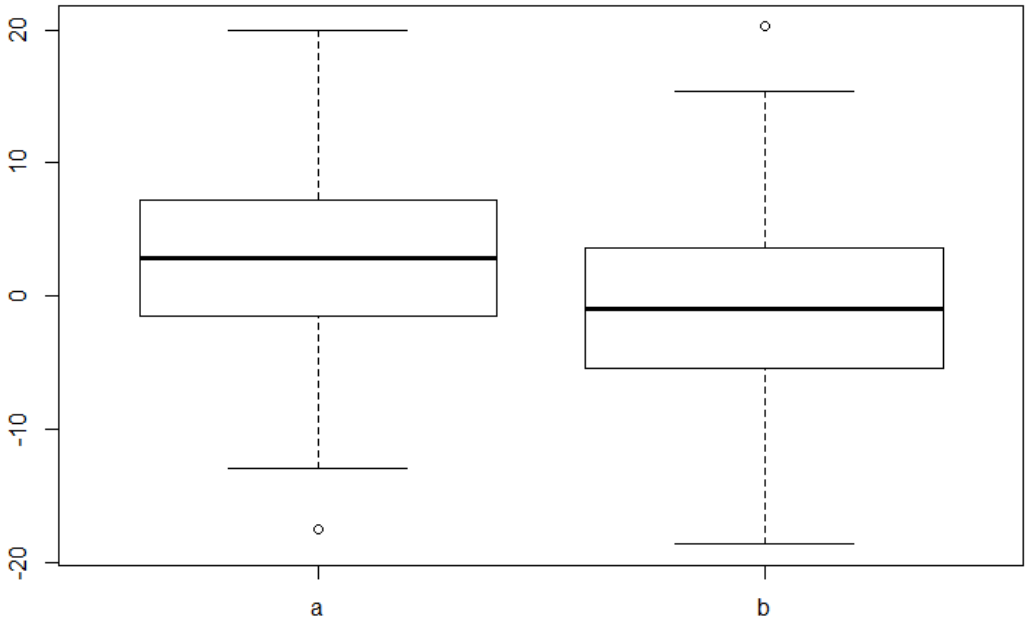

Figure 4.8: *Drift parameters estimated by (3.3) and (3.4) with* $(\alpha, \beta, \sigma, \gamma) = (1, 1, 1, 50)$.

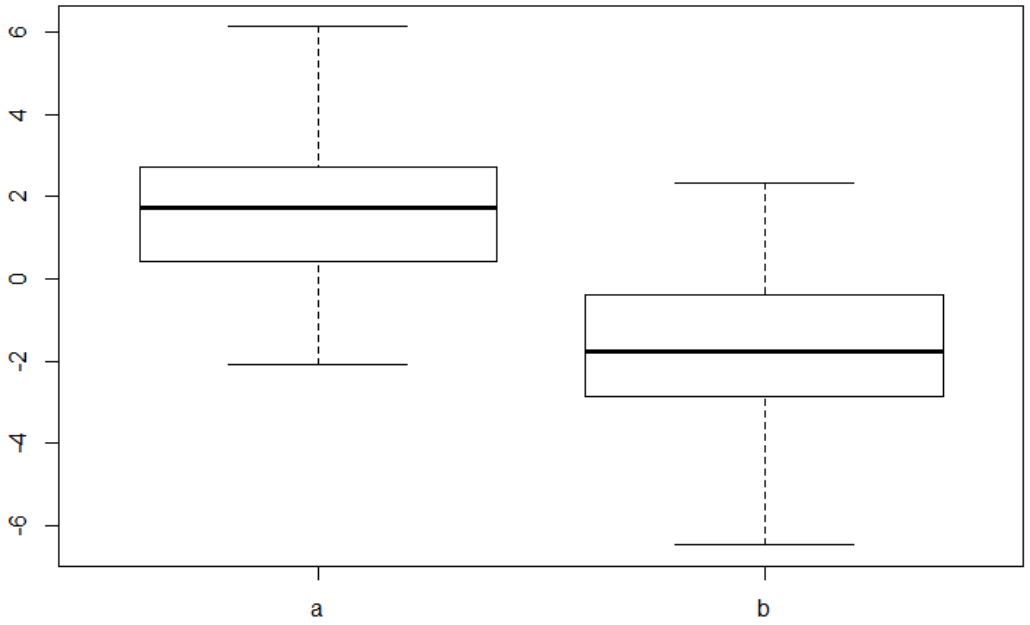

Figure 4.9: *Drift parameters estimated by (3.3) and (3.4) with* $(\alpha, \beta, \sigma, \gamma) = (1, -1, 1, 50)$.

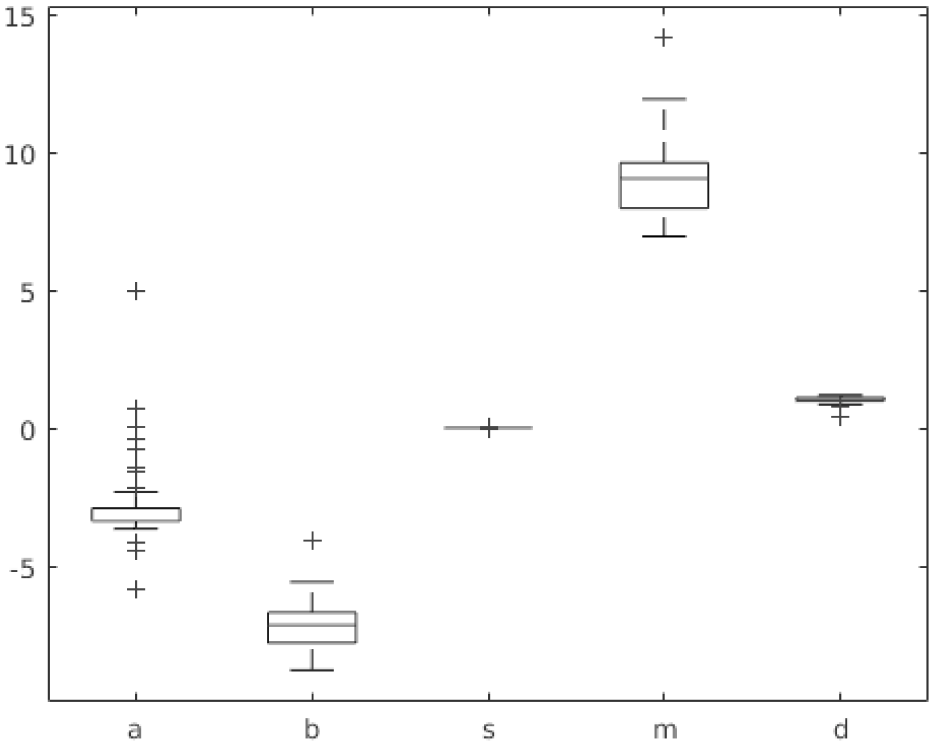

Figure 4.10: *Parameters estimated for the hyperbolic diffusion, Section 2.5.2 using the parameter set* $(\alpha, \beta, \delta, \mu, \sigma) = (7, -6, 1, 7, 0.05)$ *with* $h = 2^{18}$ *and interval length* $T = 50$.

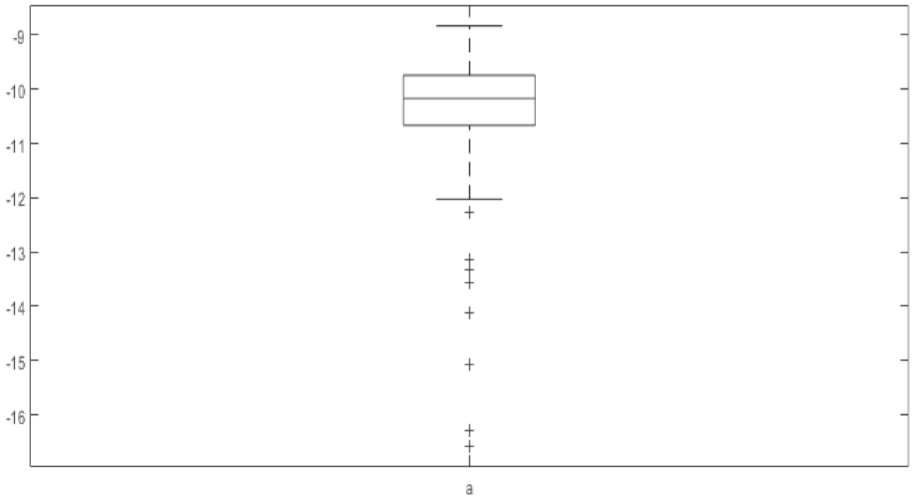

Figure 4.11: *The parameter* α *estimated for the transformed SDE Section 2.11 using step length* $h = 2^{18}$ *and interval length* $T = 10$.

5
Conclusion

5.1 Simulation

The main problem with the simulation was that the methods were time consuming. This forced us to limit the step size and length of the interval. This was the issues for CKLS and the SDE in Section 2.5.3. While for the SDE in Section 2.5.2 we only had to generate random normal numbers, which allowed us to simulate a path over a much longer interval and with shorter time step.

5.2 Statistical methods

For CKLS, Section 2.5.1, the statistical methods in Section 3.2 works well for $\gamma \approx 10$. In this setting the likelihood ratio method is generally stable for the drift parameters, with some exceptions. When γ is increased to around 50, the estimation of the drift parameters is unstable when $\beta > 0$. But similarly stable when $\beta < 0$ as for smaller γ. This make sense, as in this case we have a mean-reversion both in the volatility and the drift, causing the path to behave not as wildly as when $\beta > 0$.

In the case of the SDE from Section 2.5.2, estimation is stable for all parameters except the $\alpha(> 0)$ parameter. This parameter is estimated negatively consistently. We tried to see in what setting the estimation became more stable. But we could only get the estimation to be stable when α was the single unknown parameter to be estimated. As soon as we introduced another unknown, the estimation for α became very difficult.

The estimation using least squares for the SDE in Section 2.5.3 works well, even as the path can take on very large values compared to the stationary level. When we use the original representation the estimation is not as stable as when the transformed SDE is considered.

Bibliography

[1] J.M.P. Albin, B. Astrup Jensen, A. Muszta, and M. Richter. On Volatility induced Stationarity for Stochastic Differential Equations. *D-CAF Working Paper*, 10(2006-10), January 2007.

[2] K. C. Chan, G.A. Karolyi, F. A. Longstaff, and A. B. Sanders. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, 47(3):1209–1227, 1992.

[3] T. Conley, L.P. Hansen, E. G. J. Luttmer, and J. Scheinkman. Short-Term Interest Rates as Subordinated Diffusions. *The Review of Financial Studies*, 10(3):525–577, 04 1997.

[4] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Graduate texts in mathematics: 113. Springer, 1991.

[5] F. Klebaner. *Introduction to stochastic calculus with applications*. Imperial College Press London, 3rd ed. edition, 2012.

[6] S. Larsson. Lecture notes in numerical methods for stochastic odes. Available at: `http://www.math.chalmers.se/Stat/Grundutb/CTH/tms165/1819/NumericalLectures/NumericalLectureNotes.pdf` (visited: 2019-06-15), September 2013.

[7] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion (Grundlehren der mathematischen Wissenschaften)*. Springer-Verlag, 3rd edition, December 1999.

[8] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.