Extracting volatility smiles from historical spot data

Master's Thesis

Emil Larsson

Lund, 2017

Financial Economics
Lund University

Abstract

The Black-Scholes model has been the fundamental framework for option pricing since its publication 1973, but it is known to have shortcomings. To correct for this, plenty of research in option pricing theory has been focused on calibrating a stochastic process to match asset behavior in the financial markets better than the geometric Brownian motion that Black-Scholes assume describe asset behaviour justly. A model that has gained popularity in the industry is the SABR volatility model.

In this thesis we develop a numerical option pricing algorithm using the Hedged Monte Carlo method, for which we explore various modifications and additions. Due to its numerical nature, it can be used to price options without assuming a statistical process for the underlying asset. Instead, it estimates option prices based solely on historical data. We evaluate the algorithm with simulated data from the classic Black-Scholes framework and the SABR volatility model to see that the price estimates from our algorithm matches the theoretically correct values. Having validated the algorithm, we apply it on historical FX spot data and obtain empirical volatility smiles that lie close to the smiles observed in the current market.

Keywords: Monte Carlo option pricing, empirical volatility smile
JEL classification: G12

Acknowledgement

First of all, I have to thank quantitative strategists Nicolas Hutchings, Mark Baker and especially Tor Gillberg at Bank of America Merrill Lynch without whom this thesis would never have happened. Thank you for putting of time for me, feeding me with ideas and helping me overcome all the obstacles I have faced.

I am also gratefully indebted to my dear friend and rates salesman Daniel Alavei, for putting me in touch with Bank of America Merrill Lynch and for reviewing my work.

I must not forget to thank Birger Nilsson and Anders Wilhelmsson at Lund University for supervising the thesis and guiding me in the process.

Finally, I must express my very profound gratitude to my family for providing me with unfailing support, not only during my work on this thesis project but throughout my entire education.

Contents

Abbreviations and Glossary **6**

1 Introduction **7**
- 1.1 Background . . . 7
- 1.2 Purpose and contribution . . . 7
- 1.3 Research questions . . . 7
- 1.4 Results . . . 8
- 1.5 Structure of the thesis . . . 8

2 Theory **8**
- 2.1 Black-Scholes model . . . 8
- 2.2 Wiener process . . . 9
- 2.3 Ito's Lemma . . . 9
- 2.4 Geometric Brownian motion . . . 10
- 2.5 SABR volatility model . . . 10
- 2.6 Volatility smiles . . . 10

3 Method **11**
- 3.1 Input data . . . 11
- 3.2 Simulating GBM . . . 12
- 3.3 Simulating SABR . . . 13
- 3.4 A simple Monte Carlo approach . . . 14
- 3.5 Least Squares Monte Carlo (LSM) . . . 15
- 3.6 Hedged Monte Carlo (HMC) . . . 17
- 3.7 Volatiliy smile expansion . . . 17

4 Implementation **19**
- 4.1 LSM implementation . . . 19
- 4.2 HMC implementation . . . 20
  - 4.2.1 Regular HMC . . . 21
  - 4.2.2 Option derivative HMC . . . 22
  - 4.2.3 BS-delta HMC . . . 23
  - 4.2.4 Knots . . . 24
  - 4.2.5 Extrapolated points . . . 24
  - 4.2.6 Final step . . . 25
  - 4.2.7 HMC algorithm . . . 25
  - 4.2.8 Extract volatility . . . 26
- 4.3 Implementation of the smile expansion . . . 26
  - 4.3.1 ”No Move” option pricing . . . 27
  - 4.3.2 Binary option pricing . . . 27
  - 4.3.3 Theoretical BS smile . . . 29

5 Results **29**
5.1 Accuracy of the HMC methods 29
5.1.1 "No Move" option pricing 31
5.1.2 Binary option pricing 32
5.2 Black-Scholes volatility smile 33
5.3 SABR volatility smile 35
5.4 Empirical volatility smile 38
5.5 Discussion 40
5.5.1 HMC versions 40
5.5.2 Empirical smiles 41
5.5.3 Optimizing the performance 41

6 Conclusion **42**
6.1 Suggestions for further research 42

7 Appendix **43**
7.1 Independence of basis functions 43
7.2 Distribution of estimates 45

List of Tables

1 Option and simulated asset properties 30
2 Monte Carlo pricing using BS data. Summary of price statistics. 30
3 Monte Carlo pricing using SABR data. Summary of price statistics. 30
4 "No move" option pricing with the payoff_1 32
5 "No move" option pricing with the Gaussian payoff_2 32
6 Binary option pricing with payoff_1 33
7 Binary option pricing with payoff_2 33
8 Specification for option and BS data 34
9 Specification for option and SABR data 36
10 Calibrated 1W SABR smile for GBPUSD 38

List of Figures

1 Example of time series 12
2 Evaluation scheme 13
3 Call and hedge functions 26
4 "No move" payoff functions 28
5 Binary call payoff functions 28
6 BS HMC smiles 34
7 BS smile expansions 34
8 SABR HMC smiles 36
9 SABR smile expansions 37
10 Convex SABR smiles 37
11 Empirical 1W smiles from HMC 39

12 Empirical 1W smile expansion 39
13 Empirical ON smiles 40
14 Histograms from BS data, 250 time series, 10 hedges 45
15 Histograms from BS data, 1000 time series, 10 hedges 46
16 Histograms from BS data, 250 time series, 100 hedges 47
17 Histograms from SABR data, 250 time series, 10 hedges 48
18 Histograms from SABR data, 1000 time series, 10 hedges 49
19 Histograms from SABR data, 250 time series, 100 hedges 50

Abbreviations and Glossary[1]

At-the-money	Option's strike equals the forward price of the underlying asset
BS	Black-Scholes framework for option pricing
Exotic	A more complex contract than the standard contract
FX	Foreign exchange (currency)
GBM	Geometric Brownian motion
Greek	Partial derivative of the option price, with respect to different factors
HMC	Hedged Monte Carlo
In-the-money	Option with intrinsic value greater than zero
Intrinsic value	Amount an option would pay if it expired today
LSM	Least Squares Monte Carlo
ON	Overnight, tenor equivalent to 1 day
Out-of-the-money	Option with zero intrinsic value
Option delta (Δ)	Price sensitivity with respect to the underlying asset's price
Option vega (ν)	Price sensitivity with respect to the underlying asset's implied volatility
SABR	A stochastic process, short for "Stochastic Alpha, Beta, Rho"
Spot	The current market price of a currency, commodity or security
Straddle	A call option and a put option with the same strike price
Strike	The strike price of an option
Vanilla	A regular contract without additional features (antonym of exotic)
Volatility smile	Implied volatility in BS parametrized by strike
Volatility surface	Implied volatility in BS parametrized by strike and time to maturity

[1] Note that these explanations are valid in the context of this thesis and should not be used as a reference in general.

1 Introduction

1.1 Background

In the foreign exchange (FX) options market, vanilla call and put options are frequently traded instruments. Vanilla contracts are priced in terms of implied Black Scholes volatility. The implied Black Scholes volatility is parametrized by time to maturity and strike to form a volatility surface.

The implied volatility at a specific maturity as a function of strike is referred to as a volatility smile. Market makers use smiles when quoting prices of options. Smiles for contracts with maturity of 1 week and above are quoted frequently, and with small differences between the bid and ask prices. On such tenors, market makers can be sure that they are in line with the market. Shorter maturities are quoted less frequently, and with wider bid-ask spreads, making it difficult to specify the implied volatility accurately for short maturities.

In this project we will develop a numerical method to obtain empirical smiles implied from exchange rate dynamics. With that we hope to accurately estimate volatility smiles for short maturities.

1.2 Purpose and contribution

The primary issue of option pricing is that the true option price is unknown. The Black-Scholes model can used to price options but it rests on various statistical assumptions about the underlying asset that do not necessarily hold. The goal of this thesis is to develop an accurate numerical pricing algorithm that avoids assuming a stochastic process followed by the underlying asset. Instead, we will obtain empirical option prices and empirical volatility smiles from historical asset prices. It can potentially help market makers quote prices of illiquid options or be used in a trading strategy that take advantage of mispricing by the market.

Option pricing and volatility smile modelling are by no means topics that have not been researched before, rather the opposite, and the area has gotten the attention of plenty of smart minds. On the other hand, many papers in the area are written in a quite theoretical manner. We hope to contribute to the field by concretizing and developing the implementation of ideas from other authors. At the same time, we hope to be of practical help to Bank of America Merrill Lynch in their research and trading of FX options.

1.3 Research questions

A major part of the thesis is devoted to the implementation of the Hedged Monte Carlo (HMC) method developed by Bouchaud, Potters and Sestovic in [1], which can be used to estimate option prices from historical spot data. From estimated option prices for a range of strikes, one can find an implied volatility smile. Furthermore, we will explore the usage of the smile expansion formula developed by Bouchaud, Ciliberti, De Leo and Vargas and in [2]. Their formula gives a complete volatility smile from the prices of just a few exotic options, which will explore the feasibility of pricing with the HMC method.

1.4 Results

Our results show the HMC method convergence faster than a simple Monte Carlo method to correct prices in theoretical settings. We are also able to produce realistic empirical smiles with little historical data, a promising sign for our algorithm to be put into practical use. Moreover, the smile expansion formula proves to be a viable addition, it greatly improves the computationally efficiency of obtaining a complete volatility smile with numerical methods, but our results show that it might come at the cost of lost accuracy.

1.5 Structure of the thesis

The remainder of the thesis is organized as follows. Section 2 gives the theoretical background of concepts that are used in the thesis. Section 3 presents methods developed by other authors, which we will use to produce empirical volatility smiles. Section 4 treats the choices we have made in our implementation of those methods. In section 5, we demonstrate and discuss our results. Finally, section 6 wraps up the thesis with a conclusion. In an attempt to improve the reading experience, some material is left for the appendix in section 7 and referenced to throughout the thesis.

2 Theory

This section describe theory that is used in option pricing and we will use to validate our method, including Black-Scholes model, stochastic processes and theoretical volatility smile characteristics.

2.1 Black-Scholes model

Black-Scholes (BS) formula was derived in [3], 1973. It has since then been the most important framework for option pricing. In BS framework the price of a vanilla call option is given by:

$$C(x_0, K, r, T, \sigma) = x_0 N(d_1) - Ke^{-rT} N(d_2) \tag{1}$$

where $N()$ is the Gaussian cumulative distribution function and

$$d_1 = \frac{1}{\sigma\sqrt{T}}[\ln(\frac{x_0}{K}) + (r + \frac{\sigma^2}{2})T] \quad d_2 = d_1 - \sigma\sqrt{T}, \tag{2}$$

x_0 - underlying asset price at $t = 0$
K - strike price
r - continuously compounded risk-free interest rate
T - time to maturity
σ - volatility of the underlying asset

It is common to analyse how options prices change with respect to some of the parameters above by taking partial derivatives of the option price. These measurements are referred to as 'Greeks' as they are denoted in Greek letters. One Greek that will be used in this thesis is Delta (Δ), measuring the price sensitivity with respect to the price of the underlying asset, simply defined as $\Delta = \frac{\partial C}{\partial x}$. In BS framework it is possible to find an explicit expression for Δ by taking the derivative of C and using the chain rule. For a call option this gives (derived in [4]):

$$\Delta = \frac{\partial C}{\partial x_0} = N(d_1) \tag{3}$$

Another option Greek we will use is Vega (ν) which is not really the name of a Greek letter but measures the price sensitivity with respect to the volatility of the underlying asset. For a call option in BS, the Vega come out as (derived in [4]):

$$\nu = \frac{\partial C}{\partial \sigma} = x_0\, N'(d_1)\sqrt{T} \tag{4}$$

2.2 Wiener process

The Wiener process is a cornerstone for many other stochastic processes. A variable W_t follows a Wiener process if the change of the variable, dW_t, during a small period of time dt is (e.g. see [4]):

$$dW_t = \epsilon\sqrt{dt}, \quad \epsilon \sim N(0,1) \tag{5}$$

and dW_t for two different intervals of time are independent.

2.3 Ito's Lemma

Suppose that the asset price x_t follows the Ito process:

$$dx_t = a(x_t,t)dt + b(x_t,t)dW_t \tag{6}$$

where W_t is a Wiener process, $b(x_t,t)$ and $a(x_t,t)$ are functions of x_t and t. Ito's lemma shows that a function G of x_t and t follows the process (e.g. see [4]):

$$dG = (\frac{\partial G}{\partial x_t}a(x_t,t) + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x_t^2}b(x_t,t)^2)dt + \frac{\partial G}{\partial x_t}b(x_t,t)dW_t \tag{7}$$

2.4 Geometric Brownian motion

In BS model it is assumed that the underlying asset x_t follows the Geometric Brownian Motion (GBM), which means it satisfies the stochastic differential equation (e.g. see [4]):

$$dx_t = \mu\, x_t\, dt + \sigma x_t\, dW_t \tag{8}$$

where W follows a Wiener process. As shown in [4], a portfolio can be created containing an option and a fraction of the underlying asset (that follows GBM) such that all risks are eliminated. The value of the option will therefore be independent of investors risk preferences and risk neutral pricing can be used, where $\mu \equiv r$.

2.5 SABR volatility model

An asset is said to follow the Stochastic alpha, beta, rho (SABR) volatility model if the following the stochastic differential equations are satisfied (as presented in [5]):

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta}\, dW_t^1 \\ d\sigma_t &= \alpha\, \sigma_t\, dW_t^2 \\ dW_t^1 dW_t^2 &= \rho dt \end{aligned} \tag{9}$$

where W_t^1 and W_t^2 are (as can be seen) correlated Wiener processes and F_t is the forward price of the asset with settlement at the maturity of the option, observed at time t. The model is calibrated with the characteristic parameters α, β and ρ. These are some interpretations of role that the parameters play in the SABR (more details are given in [5] and [6]):

- $0 \leq \beta \leq 1$: Controls the nature of the asset price distribution. The extreme cases $\beta = 0$ and $\beta = 1$ leads to normally respectively log-normally distributed asset price. The distributions resulting from values in between 0 and 1 do not have any known closed form probability distributions.

- $-1 \leq \rho \leq 1$ controls the correlation between increments in volatility (σ) and forward price (F_t). Typically it takes a negative value when the model is calibrated to match real data, implying that volatility tend to decrease as the forward price increase and vice versa.

- $\alpha > 0$ is the volatility of the volatility.

2.6 Volatility smiles

Volatility smiles for the BS model are flat since, according to the model, volatility should be independent of the strike and all other parameters. However, volatility smiles observed in the financial markets are rarely flat, which highlights a flaw in the BS model. This motivates the interest in studying the behavior of volatility smiles in the field of financial mathematics.

Expressions has been derived for the SABR model to find implied BS volatility, volatility that can be plugged into the BS model to obtain prices that are results of the asset following SABR instead of GBM. Oblój suggests in in [7] that the Taylor expansion of first order for implied BS volatility in the SABR model is (simplified for log-normally distributed asset prices, $\beta = 1$):

$$\begin{aligned}
&\sigma_{BS}(y) = I_0(y)(1 + I_1 T), \\
&I_0|_{\beta=1} = \alpha y \,/\, \ln\Big(\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho}\Big), \\
&I_1|_{\beta=1} = \frac{1}{4}\rho\alpha\sigma_0 + \frac{2 - 3\rho^3}{24}\alpha^2, \text{ where} \\
&y = \ln(x_0/K) \quad \text{and} \quad z = \alpha \ln(x_0/K)/\sigma_0
\end{aligned} \tag{10}$$

By plugging the parameter values into this formula, we get a theoretically correct SABR smile. This smile can be compared to smile estimates from our algorithm based on simulated data from SABR with the same parameters.

3 Method

The Hedged Monte Carlo (HMC) method play an important role in this thesis. Before getting to the HMC method, it is natural to discuss the Least Squares Monte Carlo (LSM) method. These methods have in common that they can be used to price options based solely on historical observations of the underlying asset and they are flexible enough to price some exotic options, such as American options and path-dependent options. Furthermore, we will explore the possibility of combining the HMC method with a smile expansion formula developed by De Leo, Vargas, Ciliberti, and Bouchaud in [2]. Their formula allows one to find the implied BS volatility for any strike, given the prices of a few exotic options. Although at least one these options are rare or even not existing in the markets, the HMC method is flexible enough to price them. Thus, the HMC method together with the smile expansion formula can potentially produce a complete volatility smile by just pricing a few exotic options.

3.1 Input data

We will begin by describing the data (and the notation of it) that is input for the Monte Carlo option pricing methods presented in this thesis. The data consists of a set of time series containing observations of the underlying asset. Every time series should contain equally many observations and each time series should be observed over time span that matches the time to maturity of the option one attempts to price.

Let $x_{l,k}$ denote the price of the underlying asset in the l:th time series at time k. To match the time to maturity of the option, denoted T, each time series $x_{l,0}$,$x_{l,1}$,…,$x_{l,n}$ should be such that the time between the first($x_{l,0}$) and last($x_{l,n}$) observation is T. Additionally, the frequency of the data should be consistent so that the time between two following observations $x_{l,k}$ and $x_{l,k+1}$ is $\Delta t = \frac{T}{n}$.

$$\begin{matrix} x_{0,0} & x_{0,1} & \dots & x_{0,n} \\ x_{1,0} & x_{1,1} & \dots & x_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,0} & x_{m,1} & \dots & x_{m,n} \end{matrix} \tag{11}$$

A data set with $m+1$ data series containing $n+1$ data points each.

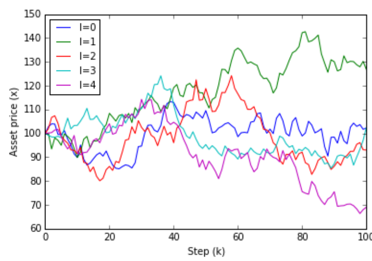

Figure 1: Example of time series. Illustation of 5 time series ($m = 4$) containing 101 data points each ($n = 100$).

A key advantage of the Monte Carlo methods is that the behaviour of the underlying asset is fully described by this data set, meaning one can use historical spot data and avoid any model assumption. However, the 'true' price of a call or put option, that we aspire to obtain, is unknown. On the other hand, if an underlying asset follows the GBM, the exact price is known from BS formula. We will use this fact to evaluate the Monte Carlo methods. By simulating asset data from the GBM and applying the Monte Carlo methods, we can compare the results to the BS price to analyse the accuracy. We will also evaluate the methods with respect to another stochastic volatility model, the SABR volatility model.

We will investigate how the number of time series ($m+1$) and the number of observations per time series ($n+1$) affect the price estimate. In addition to the data set defined above, continuously compounded risk-free interest rate (r), the maturity (T) and the strike (K) of the option are needed to obtain price estimates with the Monte Carlo methods.

3.2 Simulating GBM

To test the methods with BS formula, we need simulated data from the GBM. Take the definition of the GBM for an asset x_t from equation (8). In BS framework, log-normality is assumed for asset prices. To work with log-normal asset prices, set $G(x_t) = \ln(x_t)$ and use Ito's lemma to get:

$$dG = (r - \frac{\sigma^2}{2})dt + \sigma dW_t \tag{12}$$

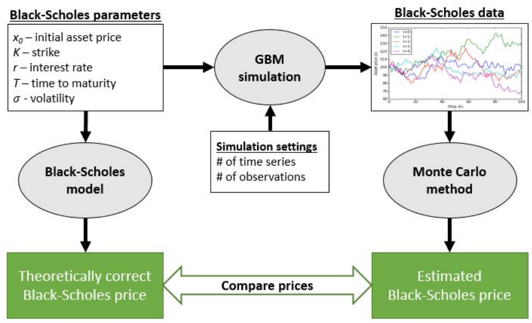

Figure 2: Evaluation scheme. Display of how the methods are evaluated against BS model. The same evaluation scheme is used for SABR, but then with a larger set of parameters.

Solving the equation for x_t with discrete time steps Δt gives:

$$\ln(x_{t+\Delta t}) - \ln(x_t) = (r - \frac{\sigma^2}{2})\Delta t + \sigma\epsilon\Delta t \iff$$
$$\iff x_{t+\Delta t} = x_t \exp\left(\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\epsilon\Delta t\right) \tag{13}$$

where $\epsilon \sim N(0,1)$. It can be seen in equation (13) that $\ln(x_t)$ is normally distributed, so that x_t has a lognormal distribution, and that one can simulate data from the GBM by drawing ϵ from the Gaussian distribution for each time step.

3.3 Simulating SABR

It is complicated to derive an exact simulation scheme for SABR. Chen, Oosterlee & Weide suggest a couple of approximate simulation schemes in [8]. We will use the log-Euler scheme since it preserves positivity of the asset price process:

$$\begin{aligned} F_{t+\Delta t} &= F_t \exp\left(-\frac{\sigma_t^2}{2}\, F_t^{2\beta-2}\Delta t + \sigma_t F_t^{\beta-1}\epsilon_F \Delta t\right) \\ \sigma_{t+\Delta t} &= \sigma_t \exp\left(-\frac{\alpha^2}{2}\,\Delta t + \alpha\epsilon_\sigma\Delta t\right) \end{aligned} \tag{14}$$
$$\text{where } \epsilon_F \sim N(0,1) \text{ and } \epsilon_\sigma \sim N(0,1) \text{ with } \mathrm{Corr}(\epsilon_F, \epsilon_\sigma) = \rho$$

We will be satisfied with a model where $\beta = 1$, with that we are able to use equation (10) to get a theoretical volatility smile to benchmark against in our model analysis and above all, to simplify the simulation of the data. Note that this simplification will just affect the data the we use to evaluate our algorithm, not the algorithm itself. If we also convert the forward price F_t into spot price

x_t ($F_{t+\Delta t}/F_t = x_{t+\Delta t}e^{-r\Delta t}/x_t$) we have:

$$
\begin{aligned}
x_{t+\Delta t} &= x_t \exp\left(-\frac{\sigma_t^2}{2}\ \Delta t + \sigma_t \epsilon_x \Delta t + r\Delta t\right) \\
\sigma_{t+\Delta t} &= \sigma_t \exp\left(-\frac{\alpha^2}{2}\ \Delta t + \alpha \epsilon_\sigma \Delta t\right)
\end{aligned}
\tag{15}
$$

These expressions can be used to simulate an asset that follow the SABR model. The scheme has a known flaw, it can become unstable with x_t diverging to infinity. One can adjust for this by simply re-simulating such data series. To simulate the correlated ϵ_x and ϵ_σ. We draw independently:

$$\epsilon \sim N(0,1) \quad \text{and} \quad \epsilon_x \sim N(0,1) \quad \text{then set} \quad \epsilon_\sigma = \rho \cdot \epsilon_x + \sqrt{1-\rho^2} \cdot \epsilon$$

This gives the desired properties:

$$
\begin{aligned}
&E[\epsilon_\sigma] = E[\rho \cdot \epsilon_x + \sqrt{1-\rho^2} \cdot \epsilon] = \rho E[\epsilon_x] + \sqrt{1-\rho^2} E[\epsilon] = \rho 0 + \sqrt{1-\rho^2} 0 = 0 \\
&V[\epsilon_\sigma] = \rho^2 V[\epsilon_x] + (1-\rho^2) V[\epsilon] = \rho^2 + 1 - \rho^2 = 1 \\
&\text{Corr}[\epsilon_x, \epsilon_\sigma] = \text{Corr}[\epsilon_x, \rho \cdot \epsilon_x + \sqrt{1-\rho^2} \cdot \epsilon] = \rho \cdot \text{Corr}[\epsilon_x, \epsilon_x] = \rho
\end{aligned}
$$

3.4 A simple Monte Carlo approach

Let $C(x_{l,k})$ denote the option price at time k given the underlying asset price $x_{l,k}$. We will at times reduce the notation of the asset price from $x_{l,k}$ to x_k if it is a general expression that holds for any time series.

All Monte Carlo option pricing builds on the fact that given an asset time series, the payoff and thereby the option price is known at maturity. The original and simplest Monte Carlo approach for option pricing (from [9]) estimates the expected payoff and discount it with the risk-free rate (since, as discussed earlier, risk neutrality can be used for option valuation problems).

This is done by by calculating the payoff at maturity for each time series to get $C(x_n)$. For example, for a call option with strike K, $C(x_n) = \max\left[x_n - K\ ,\ 0\right]$. Then discount the average payoff to get the option price, $\hat{C}$:

$$\hat{C} = e^{-rT} \frac{1}{m+1} \sum_{l=0}^{m} C(x_{l,n}) \tag{16}$$

The accuracy of the HMC algorithm will be put in comparison to this estimation. This simple Monte Carlo approach should in theory give correct results but it requires very many time series to produce an accurate estimate, which we will see later. Since historical data is limited and it is desirable to use as recent data as possible, the data demands of this method make it of little practical use, which is why it is desirable to develop less demanding and more accurate methods.

3.5 Least Squares Monte Carlo (LSM)

While one can find a hedging strategy that eliminates risk entirely in the BS framework, for many other stochastic models of asset fluctuations risk in option trading cannot be eliminated and strict arbitrage opportunities do not exist. It is therefore desirable to have a method to compute the option price, hedging strategy and residual risk for any underlying stochastic process. By choosing an optimal trading strategy such that the chosen measure of risk is minimized, an estimate of the option price is obtained using a fair game argument.

One framework using this approach is the "Least Squares Monte Carlo (LSM)" developed by Longstaff and Schwartz in [10], it lays the ground for the HMC method that we will get to later. Let Δt denote the time difference between two steps (observations of the underlying asset price). Define the price of an option with respect to the price of the underlying asset x_k at time $k \cdot \Delta t$, as $C(x_k)$. The LSM algorithm works back in time from maturity, where the option price, $C(x_n)$, is known from the payoff (e.g for a call option $C(x_n) = \max\big[x_n - K\,,\,0\big]$). The option price, $\hat{C} = C(x_0)$, can be estimated by minimizing residual risk at defined as:

$$R = (C(x_n)e^{-rT} - C(x_0))^2 \tag{17}$$

with respect to $C(x_0)$. As touched upon earlier, the option can in theory be replicated in a way to eliminate all risks. As a consequence, risk neutral pricing can be used where (according to [4]) any investment is expected to appreciate with the risk free rate, r:

$$\begin{aligned} E\big[C(x_n)\big] &= C(x_0)e^{rT} \\ \Longleftrightarrow E\big[C(x_n)e^{-rT} - C(x_0)\big] &= 0 \end{aligned} \tag{18}$$

As we can see, the defined risk measure R cannot be negative and it should in theory be zero. It therefore makes sense to find option prices such that R is minimized. Since the asset price can make big jumps over the complete lifespan of the option, the risk can be reduced further by minimization at several points between the start date and maturity. This is done in an iterative manner beginning at maturity, working backwards in time for $k = n-1, \ldots, 0$ and minimizing the local residual risk at each step, defined as:

$$R_k = (C(x_{k+1})e^{-r\Delta t} - C(x_k))^2, \quad \text{where} \quad \Delta t = \frac{T}{n} \tag{19}$$

with respect to $C(x_k)$. By reducing the total risk with many minimization steps, the option price estimate, $C(x_0)$, will (at least in theory) be more accurate.

To solve the minimization problem at any given step, we express $C(x_k)$ as:

$$C(x_k) = \sum_{i=0}^{M} b_i^k f_i(x_k), \tag{20}$$

where $f_i(x_k)$ are basis functions depending on x_k. This turns the minimization problem into a linear one in terms of the coefficients b_i. To minimize R_k, we pick $b_0, b_1, \ldots, b_M$ so that $C(x_k)$ resembles $C(x_{k+1})e^{-r\Delta t}$ as closely as possible. Denote the residuals $C(x_{k+1})e^{-r\Delta t} - C(x_k) = \varepsilon_k$. We will minimize the local residual risk over all time series. As we can see, the problem has been turned into a least squares optimization:

$$\min\Big[R_k\Big] = \min\Big[\sum_{l=1}^{m}\Big(C(x_{l,k+1})e^{-r\Delta t} - C(x_{l,k})\Big)^2\Big] = \min\Big[\sum_{l=1}^{m}(\varepsilon_{l,k})^2\Big] \quad (21)$$

which is solved by the regression:

$$C(x_{k+1})e^{-r\Delta t} = \sum_{i=0}^{M} b_i^k f_i(x_k) + \varepsilon \quad (22)$$

This problem can be expressed in matrix form to solve it like a regular multiple regression:

$$\mathbf{Y_k} = \mathbf{X_k}'\mathbf{B_k} + \varepsilon$$

$$\text{where } \mathbf{Y_k} = \begin{bmatrix} C(x_{0,k+1})e^{-r\Delta t} \\ C(x_{1,k+1})e^{-r\Delta t} \\ \vdots \\ C(x_{m,k+1})e^{-r\Delta t} \end{bmatrix} \quad \mathbf{B_k} = \begin{bmatrix} b_0^k \\ b_1^k \\ \vdots \\ b_M^k \end{bmatrix}$$

$$\mathbf{X_k} = \begin{bmatrix} f_0(x_{0,k}) & f_1(x_{0,k}) & \ldots & f_M(x_{0,k}) \\ f_0(x_{1,k}) & f_1(x_{1,k}) & \ldots & f_M(x_{1,k}) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{m,k}) & f_1(x_{m,k}) & \ldots & f_M(x_{m,k}) \end{bmatrix}$$

for which the least squares estimate given by (as derived in [11]):

$$\mathbf{\hat{B}_k} = (\mathbf{X_k'X_k})^{-1}\mathbf{X_k'Y_k} \quad (23)$$

This illustrates the necessity of many time series $(0, 1, 2 \ldots, m)$, to accurately estimate the regression parameters $\mathbf{B_k}$. In general, the more time series available the better.

With the estimated regression parameters $\mathbf{\hat{B}_k}$, we predict the option prices backwards:

$$\begin{aligned} C(x_{l,k}) &= \sum_{i=0}^{M} \hat{b}_i^k f_i(x_{l,k}), \quad l = 0 \ldots m \iff \\ \iff \mathbf{\hat{C}_k} &= \mathbf{X_k'\hat{B}_k} \end{aligned} \quad (24)$$

We summarize the Least Squares Monte Carlo as follows:

1. Calculate the option payoffs at maturity (time n) to get $\mathbf{\hat{C}_n} = [C(x_{0,n}), C(x_{1,n}), \ldots, C(x_{m,n})]'$

   Iterate the steps 2 to 5 for $k = n-1, n-2, \ldots, 0$:

2. Calculate $\mathbf{Y_k}$ by discounting the option prices from step $k+1$
3. Calculate $\mathbf{X_k}$ using the underlying asset prices from step k
4. Calculate the least squares estimate of $\mathbf{B_k}$:

$$\hat{\mathbf{B}}_\mathbf{k} = (\mathbf{X}'_\mathbf{k}\mathbf{X_k})^{-\mathbf{1}}\mathbf{X}'_\mathbf{k}\mathbf{Y_k} \tag{25}$$

5. Predict the option prices at step k with the regression model:

$$\hat{\mathbf{C}}_\mathbf{k} = \mathbf{X}'_\mathbf{k}\hat{\mathbf{B}}_\mathbf{k} \tag{26}$$

6. Finally we arrive at $\hat{\mathbf{C}}_\mathbf{0}$ from which we estimate the initial option price:

$$\hat{C} = \frac{1}{m+1}\sum_{i=0}^{m} C(x_{i,0}) \tag{27}$$

3.6 Hedged Monte Carlo (HMC)

The HMC method builds on the LSM method. The difference is that a delta hedge is introduced to improve the minimization of local residual risk:

$$R_k = (C(x_{k+1})e^{-r\Delta t} - C(x_k) + \Delta(x_k)[x_k - x_{k+1}e^{-r\Delta t}])^2 \tag{28}$$

This is solved in the same iterative manner as LSM, with a least-squares regression at each step:

$$C(x_{k+1})e^{-r\Delta t} = C(x_k) - \Delta(x_k)[x_k - x_{k+1}e^{-r\Delta t}] + \varepsilon_k \tag{29}$$

While this might not be a perfect hedging strategy, it can vastly reduce the residual risk which hopefully results in more accurate option price estimates. This introduces the problem of how to find $\Delta(x_k)$ $\big($in addition to finding $C(x_k)\big)$. We will discuss and compare different ways to approach this problem in our implementation.

3.7 Volatiliy smile expansion

Next, we will combine the HMC method with the volatility expansion formula derived in [2]. The HMC method can be used to produce volatility smiles alone, but it is quite computationally demanding since it requires pricing of a large number of options with different strikes and converting those prices into implied BS volatility. Additionally, the HMC method estimates prices less accurately for far out-of-the-money and far in-the-money options. These issues could potentially be solved by the smile expansion, which gives the implied BS volatility for any strike from a few exotic at-the-money option prices. The expansion formula for implied BS volatility is derived in [2] and reads as follows:

$$\sigma_{BS}(K) = \sigma(\alpha_T + \beta_T\mathcal{M} + \gamma_T\mathcal{M}^2 + O(\mathcal{M}^3)), \quad \mathcal{M} = \frac{K - x_0}{x_0\sigma\sqrt{T}} \tag{30}$$

where σ is the real volatility of the underlying asset (which can be estimated from the sample) and σ_{BS} is the implied BS volatility, used in BS formula to obtain the market price. It is a "smile expansion" since given the value of all the other variables, one can choose K independently and obtain the implied BS volatility corresponding to each strike. But before one can do this, the unknown coefficients have to be found. They are given by the following expressions:

$$\alpha_T = \sqrt{\frac{\pi}{2}} E[|u_T|] \quad \beta_T = \sqrt{\frac{\pi}{2}}[1 - 2P(u_T > 0)] \quad \gamma_T = \sqrt{\frac{\pi}{2}} p_T(0) - \frac{1}{2\alpha_T} \tag{31}$$

where $u_T = \frac{x_T - x_0}{x_0 \sigma \sqrt{T}}$. An important observation is that the coefficients can be interpreted as the average payoff of some exotic at-the-money options, as explained in [2]:

- $E[|u_T|] = E[|x_T - x_0|] \frac{1}{x_0 \sigma \sqrt{T}}$ can be interpreted as the price of a regular at-the-money straddle divided by $x_0 \sigma \sqrt{T}$, since

$$\begin{aligned} E[|x_T - x_0|] &= E[\max(x_T - x_0, 0) + \max(x_0 - x_T, 0)] = \\ &= E[\max(x_T - x_0, 0)] + E[\max(x_0 - x_T, 0)], \end{aligned} \tag{32}$$

  which is the expected payoffs from a call option and a put option with strike prices $K = x_0$ (i.e. at-the-money).

- $P(u_T > 0)$ can be interpreted as the price of an at-the-money binary call option. To illustrate this, define a Bernoulli random variable, y_T:

$$y_T = \begin{cases} 1 & \text{if } x_T > x_0 \\ 0 & \text{otherwise} \end{cases}$$

  Using this, rewrite $P(u_T > 0)$ in the following way:

$$P(u_T > 0) = P(x_T > x_0) = P(y_T = 1) \cdot 1 + P(y_T = 0) \cdot 0 = E[y_T] \tag{33}$$

  In words, that is the expected payoff from an option that pays off 1 if the underlying asset price exceed its initial price at maturity and 0 otherwise, i.e the expected payoff from an at-the-money binary call option.

- $p_T(0)$ is the value of u_T's density function at 0. It can be interpreted as the price of a "No Move" option, an option that only pays off when the underlying asset ends very close to its initial price. To illustrate this, define a random variable, z_T:

$$z_T = \begin{cases} \frac{1}{2\tau} & \text{if } x_0(1 - \tau\sigma\sqrt{T}) < x_T < x_0(1 + \tau\sigma\sqrt{T}) \\ 0 & \text{otherwise} \end{cases}$$

where τ is a small positive number. Then use the midpoint approximation (e.g. from [12]) and the following rewritings:

$$\begin{aligned}
&p_T(0)\cdot 2\tau \approx P(-\tau < u_T < \tau) \iff \\
\iff\ & p_T(0) \approx \frac{1}{2\tau}P(-\tau < u_T < \tau) = \\
&= \frac{1}{2\tau}P\Big(-\tau < \frac{x_T - x_0}{x_0\sigma\sqrt{T}} < \tau\Big) = \\
&= \frac{1}{2\tau}P\Big(x_0(1-\tau\sigma\sqrt{T}) < x_T < x_0(1+\tau\sigma\sqrt{T})\Big) \\
&= E[z_T]
\end{aligned} \tag{34}$$

which is the expected payoff from a "No Move" option that pays off $\frac{1}{2\tau}$ when the underlying asset ends in the interval $[x_0(1-\tau\sigma\sqrt{T}), x_0(1+\tau\sigma\sqrt{T})]$. Since τ is a small positive number, the term $\tau\sigma\sqrt{T}$ will also be small and positive, and the interval will be narrow (hence the term "No Move"). The approximation will be accurate when $\tau \to 0$. The problem of picking a sufficiently small τ will be discussed in greater detail later on.

Using these interpretations, one can price the suggested options with the HMC method, obtain the coefficients and make use of the smile expansion formula.

4 Implementation

This section describes the implementation and development of the methods presented in the previous section, and the choices we have made in that process.

4.1 LSM implementation

To begin with, one have to pick the basis functions $f_i(x_k)$ used to express:

$$C(x_k) = \sum_{i=0}^{M} b_i^k f_i(x_k), \tag{35}$$

For example, we would obtain a linear model if $M = 1$, $f_0(x_k) = 1$ and $f_1(x_k) = x_k$. However, option prices are solutions to a diffusion equation. As such, they are very smooth functions in x_k. By using the following piecewise quadratic polynomials as basis functions, we ensure a smooth and continuous option price curve $C(x_k)$:

$$\begin{aligned}
f_0(x_k) &= 1 \\
f_1(x_k) &= x_k \\
f_2(x_k) &= x_k^2 \\
f_s(x_k) = (x_k - \xi_{s-3})_+^2 \text{ for } s &= 3, \dots, M
\end{aligned} \tag{36}$$

$$\text{where } (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$

The numbers ξ_s are usually referred to as "knots". It can be shown that the basis functions are independent, we refer to section 7.1 in the Appendix for the proof. It is clear that the first three basis functions make a quadratic polynomial which is smooth. To illustrate that the definition of the basis functions with knots ensures smoothness, let $\xi_0 < \xi_1 < \ldots < \xi_{M-3}$ and take the limits:

$$\begin{aligned} \lim_{x_k \to \xi_i^-} C(x_k) &= b_0^k + b_1^k \xi_i + b_2^k \xi_i^2 + b_3^k (\xi_i - \xi_0)_+^2 + \ldots + b_{i+2}^k (\xi_i - \xi_{i-1})_+^2 \\ \lim_{x_k \to \xi_i^+} C(x_k) &= b_0^k + b_1^k \xi_i + b_2^k \xi_i^2 + b_3^k (\xi_i - \xi_0)_+^2 + \\ &\quad + \ldots + b_{i+2}^k (\xi_i - \xi_{i-1})_+^2 + b_{i+3}^k (\xi_i - \xi_i)_+^2 = \lim_{x_k \to \xi_i^-} C(x_k) \end{aligned} \tag{37}$$

4.2 HMC implementation

The basis functions suggested for LSM in section 4.1 can be used for HMC as well. There are at least three plausible ways of picking the hedge Δ (as suggested in [1, 13]), which we will refer to as different versions of HMC:

1. **Regular HMC:** Express $\Delta(x_k)$ with its own set of basis functions:

$$\Delta(x_k) = \sum_{i=0}^{M} a_i^k g_i(x_k) \tag{38}$$

   This complicates the regressions, compared to LSM, as more parameters have to be estimated.

2. **Option derivative HMC:** Set $\Delta(x_k)$ as the derivative of the option price:

$$a_i = b_i \quad \text{and} \quad g_i(x) = \frac{df_i(x)}{dx} \tag{39}$$

   This will lead to exact results only for Gaussian processes (according to [2]), but by not using an independent set of basis functions it will reduce the computational cost of the "Regular HMC". Without additional parameters the regression matrices will be of the same sizes as in LSM.

3. **BS-delta HMC:** Set $\Delta(x_k)$ to the BS-delta, e.g. in the case of a call option $\Delta(x_k) = N(d_1)$. This delta is just an approximation, as we do not assume that the BS model necessarily is correct. However, this version further simplifies the regressions and ensures a smooth hedge curve, as the BS delta is a smooth function with respect to K. To calculate the BS delta (Δ) one have to guess σ based on the sample.

We will now go into the details of implementing each of these alternatives.

4.2.1 Regular HMC

We want to include the condition defined above in equation (38) in a matrix system for the HMC regression in equation (29). As the left hand side is identical to the one in LSM we will keep:

$$\mathbf{Y_k} = \mathbf{X_k}\mathbf{B_k} + \varepsilon \quad \text{where } \mathbf{Y_k} = \begin{bmatrix} C(x_{0,k+1})e^{-r\Delta t} \\ C(x_{1,k+1})e^{-r\Delta t} \\ \vdots \\ C(x_{m,k+1})e^{-r\Delta t} \end{bmatrix}$$

We find no reason to change the formulation of basis functions to express $C(x_k)$. We have found it reasonable to use 8 knots, but using slightly more or less should not have a major impact on the results. There is no number that can be shown to be optimal in general and an attempt to optimize the number would probably just be over-fitting to the sample. However, we argue that using just 1 or 2 knots would give a jagged curve which would not be ideal and 100 knots would result in very large matrices but not necessarily improve the results significantly.

Since delta is defined as the derivative of the option (with respect to the underlying asset price) and we used a quadratic spline (degree 2) to interpolate the option, we will use a linear spline (degree 1) to interpolate the hedge:

$$\begin{aligned} g_0(x_k) &= 1 \\ g_1(x_k) &= x_k \\ g_s(x_k) &= (x_k - \xi_{s-2})_+ \text{ for } s = 2, \dots, 9 \end{aligned} \tag{40}$$

$$\text{where } (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$

The basis functions we have defined for 8 knots gives us:

$$C(x_k) = \sum_{i=0}^{10} b_i^k f_i(x_k) \quad \text{and} \quad \Delta(x_k) = \sum_{i=0}^{9} a_i^k g_i(x_k)$$

We remind the reader that each hedge adjustment (Δ) is multiplied by $[x_k - x_{k+1}e^{-r\Delta t}]$. Denote $s_k = x_{k+1}e^{-r\Delta t}$. We also remind of the data notation presented in section 3.1; $x_{l,k}$ describes the price of the underlying asset in the l:th time series at time k. The right hand side of the regression model,

$$C(x_{k+1})e^{-r\Delta t} = C(x_k) - \Delta(x_k)[x_k - x_{k+1}e^{-r\Delta t}] + \varepsilon_k, \tag{41}$$

becomes:

$$\begin{aligned} &C(x_k) - \Delta(x_k)[x_k - s_k] + \varepsilon_k = \\ &= \sum_{i=0}^{10} b_i^k f_i(x_k) + \left(\sum_{i=0}^{9} a_i^k g_i(x_k) \right) [s_k - x_k] + \varepsilon_k \end{aligned} \tag{42}$$

To perform the regression, this will be expressed in matrix form with all of the time series:

$$C(x_k) - \Delta(x_k)[x_k - s_k] + \varepsilon = \mathbf{X_{k,1}B_{k,1}} + \mathbf{X_{k,2}B_{k,2}} = \mathbf{X_k B_k} + \varepsilon_\mathbf{k} \tag{43}$$

$$\mathbf{X_k} = \begin{bmatrix}\mathbf{X_{k,1}} & \mathbf{X_{k,2}}\end{bmatrix} \quad \mathbf{B_k} = \begin{bmatrix}\mathbf{B_{k,1}} \\ \mathbf{B_{k,2}}\end{bmatrix}, \text{ where}$$

$$\mathbf{B_{k,1}} = \begin{bmatrix} b_0^k \\ b_1^k \\ \vdots \\ b_{10}^k \end{bmatrix} \; \mathbf{B_{k,2}} = \begin{bmatrix} a_0^k \\ a_1^k \\ \vdots \\ a_9^k \end{bmatrix} \; \mathbf{X_{k,1}} = \begin{bmatrix} 1 & x_{0,k} & x_{0,k}^2 & (x_{0,k}-\xi_0)^2 & \dots & (x_{0,k}-\xi_7)^2 \\ 1 & x_{1,k} & x_{1,k}^2 & (x_{1,k}-\xi_0)^2 & \dots & (x_{1,k}-\xi_7)^2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{m,k} & x_{m,k}^2 & (x_{m,k}-\xi_0)^2 & \dots & (x_{m,k}-\xi_7)^2 \end{bmatrix}$$

$$\mathbf{X_{k,2}} = \begin{bmatrix} (s_{0,k}-x_{0,k}) & x_{0,k}(s_{0,k}-x_{0,k}) & (x_{0,k}-\xi_0)_+(s_{0,k}-x_{0,k}) & \dots & (x_{0,k}-\xi_7)_+(s_{0,k}-x_{0,k}) \\ (s_{1,k}-x_{1,k}) & x_{1,k}(s_{1,k}-x_{1,k}) & (x_{1,k}-\xi_0)_+(s_{1,k}-x_{1,k}) & \dots & (x_{1,k}-\xi_7)_+(s_{1,k}-x_{1,k}) \\ \vdots & \vdots & \vdots & & \vdots \\ (s_{m,k}-x_{m,k}) & x_{m,k}(s_{m,k}-x_{m,k}) & (x_{m,k}-\xi_0)_+(s_{m,k}-x_{m,k}) & \dots & (x_{m,k}-\xi_7)_+(s_{m,k}-x_{m,k}) \end{bmatrix}$$

4.2.2 Option derivative HMC

The simplifications presented in equation (39) turns the basis functions for the hedge into:

$$\begin{aligned} g_0(x_k) &= \frac{d}{dx_k}(1) = 0 \\ g_1(x_k) &= \frac{d}{dx_k}(x_k) = 1 \\ g_2(x_k) &= \frac{d}{dx_k}(x_k^2) = 2x_k \\ g_s(x_k) = \frac{d}{dx}((x_k - \xi_{s-3})^2) &= 2(x_k - \xi_{s-3})_+ \text{ for } s = 3, \dots, 10 \\ \text{and } a_k &= b_k \text{ for all } k = 0, \dots, 10 \end{aligned} \tag{44}$$

With the redefined basis functions g, the right hand side of the regression model becomes

$$\begin{aligned}
&C(x_k) - \Delta(x_k)[x_k - s_k] + \varepsilon_k = \\
&= \sum_{i=0}^{10} b_i^k f_i(x_k) - \Big(\sum_{i=0}^{10} b_i^k g_i(x_k) \Big)[x_k - s_k] + \varepsilon_k = \\
&= \sum_{i=0}^{10} b_i^k \Big(f_i(x_k) - g_i(x_k)[x_k - s_k] \Big) + \varepsilon_k = \\
&= b_0^k \Big(1 - 0 \Big) + b_1^k \Big(x_k - (x_k - s_k) \Big) + b_2^k \Big(x_k^2 - 2x_k(x_k - s_k) \Big) + \\
&\quad + b_3^k \Big((x_k - \xi_0)_+^2 - 2(x_k - \xi_0)_+(x_k - s_k) \Big) + \ldots + \\
&\quad + b_{10}^k \Big((x_k - \xi_7)_+^2 - 2(x_k - \xi_7)_+(x_k - s_k) \Big) + \varepsilon_k = \\
&= b_0^k + b_1^k s_k + b_2^k (2s_k - x_k)x_k + b_3(x_k - \xi_0)_+(2s_k - x_k - \xi_0) + \ldots \\
&\ldots + b_{10}(x_k - \xi_7)_+(2s_k - x_k - \xi_7) + \varepsilon_k
\end{aligned}$$

Just as in section 4.2.1 this is expressed in matrix form for all time series to perform the regression.

$$C(x_k) - \Delta(x_k)[x_k - s_k] + \varepsilon_k = \mathbf{X_k B_k} + \varepsilon_\mathbf{k} \ , \qquad \mathbf{B_k} = \begin{bmatrix} b_0^k \\ b_1^k \\ \vdots \\ b_{10}^k \end{bmatrix}$$

$$\mathbf{X_k} = \begin{bmatrix} 1 & s_{0,k} & (2s_{0,k} - x_{0,k})x_{0,k} & (x_{0,k} - \xi_0)_+(2s_{0,k} - x_{0,k} - \xi_0) & \ldots & (x_{0,k} - \xi_7)_+(2s_{0,k} - x_{0,k} - \xi_7) \\ 1 & s_{1,k} & (2s_{1,k} - x_{1,k})x_{1,k} & (x_{1,k} - \xi_0)_+(2s_{1,k} - x_{1,k} - \xi_0) & \ldots & (x_{1,k} - \xi_7)_+(2s_{1,k} - x_{1,k} - \xi_7) \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & s_{m,k} & (2s_{m,k} - x_{m,k})x_{1,k} & (x_{m,k} - \xi_0)_+(2s_{m,k} - x_{m,k} - \xi_0) & \ldots & (x_{m,k} - \xi_7)_+(2s_{m,k} - x_{m,k} - \xi_7) \end{bmatrix}$$

4.2.3 BS-delta HMC

By using BS-delta the regression becomes even more simple, as the hedge adjustment becomes known. Without any parameters (a_i or b_i) being used to express the hedge, we include the hedge in the known $\mathbf{Y_k}$:

$$\mathbf{Y_k} = \begin{bmatrix} C(x_{0,k+1})e^{-r\Delta t} - \Delta(x_{0,k})(s_{0,k} - x_{0,k}) \\ C(x_{1,k+1})e^{-r\Delta t} - \Delta(x_{1,k})(s_{1,k} - x_{1,k}) \\ \vdots \\ C(x_{m,k+1})e^{-r\Delta t} - \Delta(x_{m,k})(s_{m,k} - x_{m,k}) \end{bmatrix}$$

$\mathbf{B_k}$ and $\mathbf{X_k}$ become identical to the regression matrices for LSM:

$$\mathbf{B_k} = \begin{bmatrix} b_0^k \\ b_1^k \\ \vdots \\ b_{10}^k \end{bmatrix} \quad \text{and} \quad \mathbf{X_k} = \begin{bmatrix} 1 & x_{0,k} & x_{0,k}^2 & (x_{0,k} - \xi_0)^2 & \ldots & (x_{0,k} - \xi_7)^2 \\ 1 & x_{1,k} & x_{1,k}^2 & (x_{1,k} - \xi_0)^2 & \ldots & (x_{1,k} - \xi_7)^2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{m,k} & x_{m,k}^2 & (x_{m,k} - \xi_0)^2 & \ldots & (x_{m,k} - \xi_7)^2 \end{bmatrix}$$

4.2.4 Knots

We have until now only mentioned that 8 knots $[\xi_0, \xi_1, \ldots, \xi_7]$ are used, but not their values. When regressing with a piecewise polynomial these knots determine the points at which the polynomial is allow to change shape. Our polynomials should approximate the option price as a function of the underlying asset and the hedge as a function of underlying asset as good as possible. The knots should therefore ideally be spread out over the data interval with larger density where the curvatures change.

There are two other things to consider; First, we want a method that is flexible and can price a wide range of assets, which means the knots has to be suitable for vastly different data. Second, it is desirable that the knots change with time, as the distribution of time series is likely to change at each time step. Thus, the knots cannot be defined with a fixed value but should rather be defined relative to the asset prices at a given step.

Apart from above mentioned points, we rely on intuition for the placement of nodes. Since the option payoff typically change at the strike price, it is natural to pick one of the nodes equal to the strike, $\xi_0 = K$. Our intuition tells us it should be a good idea to pick the remaining points spread over the interval of asset prices. We find the knots well suited for most regressions if we pick the remaining ξ's in the following in way. Let $\bar{\mathbf{x}}_\mathbf{t}$ denote underlying asset prices of all time series at time t, $\bar{\mathbf{x}}_\mathbf{t} = [x_{0,t}, x_{1,t}, \ldots, x_{m,t}]$

$$\begin{aligned}
\xi_1 &= 0.10 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.10) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_2 &= 0.30 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.30) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_3 &= 0.45 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.45) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_4 &= 0.50 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.50) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_5 &= 0.55 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.55) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_6 &= 0.70 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.70) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}] \\
\xi_7 &= 0.90 \cdot \max[\bar{\mathbf{x}}_\mathbf{t}] + (1 - 0.90) \cdot \min[\bar{\mathbf{x}}_\mathbf{t}]
\end{aligned} \tag{45}$$

We could have distributed the knots more evenly but it will be useful with larger density of knots in the middle of the interval to price at-the-money options.

4.2.5 Extrapolated points

The price and perfect hedge are easy to estimate for an option that is very far into-the-money or out-of-the-money and we will use this fact to further improve the regression. Take a call option as example (the methodology can easily be modified for put or other options); if the underlying asset's price $x > 4K$, the call price is very close to the intrinsic value $C = [x - K]$, since it is very unlikely that the option will become out-of-the-money. As a consequence of this, the hedge becomes $\Delta = \frac{dC}{dx} = \frac{d(x-K)}{dx} = 1$. Similarly for a call option, if the underlying asset is well below the strike price say: $x < 0.25K$, one can be fairly certain the option will expire out-of-the-money and therefore $C = 0$ with $\Delta = 0$.

Adding a few artificial observations of options ($C(x_{l,k})$) far in-the-money and out-of-the-money, with corresponding hedge ($\Delta(x_{l,k})$) and underlying asset price ($x_{l,k}$), will help constrain the regression, forcing the option-asset and hedge-asset curves into this known behaviour far from the strike price.

4.2.6 Final step

A short note on the final step of iterations in the HMC algorithm. Complications arise for regressions at $k = 0$ if the time series are such that the asset price at time 0 is the same for all time series. All x_0 taking the same value makes the matrix $X_0' X_0$ singular and we are thereby unable to find the least-squares solution $\mathbf{\hat{B}_0} = (\mathbf{X_0'X_0})^{\mathbf{-1}}\mathbf{X_0'Y_0}$. This can be solved simply, without a great loss of accuracy, by ending the iteration at $k = 1$ and using the simple MC for the final step $\hat{C} = e^{-r\Delta t}\frac{1}{m+1}\sum_{i=0}^{m}\hat{C}(x_{i,1})$

4.2.7 HMC algorithm

We summarize the steps in the HMC algorithm, which look much like LSM algorithm but with the few adjustments mentioned above:

1. Calculate the option payoffs at maturity to get $\mathbf{\hat{C}_n} = [C(x_{0,n}), C(x_{1,n}), \ldots, C(x_{m,n})]$

   Iterate the steps 2 to 5 for $k = n-1, n-2, \ldots, 1$**:**

2. Calculate $\mathbf{Y_k}$, as defined above for the chosen version.

3. Calculate $\mathbf{X_k}$, as defined above for the chosen version.

4. Calculate the least squares estimate of $\mathbf{B_k}$:

$$\mathbf{\hat{B}_k} = (\mathbf{X_k'X_k})^{\mathbf{-1}}\mathbf{X_k'Y_k} \tag{46}$$

5. Predict the option prices at step k with the regression model:

$$\mathbf{\hat{C}_k} = \mathbf{X_k'\hat{B}_k} \tag{47}$$

6. Finally we arrive at $\mathbf{\hat{C}_1}$ from which we estimate the initial option price:

$$\hat{C} = e^{-r\Delta t}\frac{1}{m+1}\sum_{i=0}^{m} C(x_{i,1}) \tag{48}$$

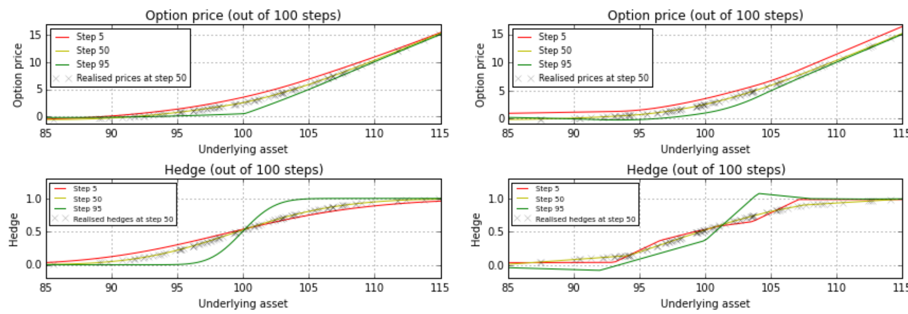

Figure 3: Call and hedge functions. Illustrations of a call price $C(x_k)$ and the hedge $\Delta(x_k)$ at different steps in the algorithm. The call option is at-the-money with x_0=100. The crosses in the upper pictures represent realised call prices ($\hat{\mathbf{C}}_\mathbf{k}$) from equation (47) and the crosses in the lower pictures are realised Δ's from the regression. The plots on the left side come from the "BS-delta HMC" and right side from the "Option derivative HMC".

4.2.8 Extract volatility

The HMC method is used to obtain option prices but the goal of the thesis is to produce volatility smiles, which is the market standard for quoting option prices. It is a one-dimensional optimization problem to go from option price to implied BS volatility. BS formula behave nicely so the Newton method (e.g. see [14]) can be used to estimate the implied BS volatility with fast convergence. Define the function we will try to find a root for as:

$$f(\sigma) = C(x_0, K, r, T, \sigma) - \hat{C}, \tag{49}$$

where $\hat{C}$ is the estimated option price from HMC and $C(x_0, K, r, T, \sigma)$ is the option price given by Black & Scholes formula. We begin by guessing the volatility σ_0. The standard deviation of the sample data would be a good guess, but it is probably more computationally efficient to just pick some other arbitrary value in the interval [0,1]. Starting with at $k = 0$, iterate the Newton method:

$$\sigma_{k+1} = \sigma_k - \frac{f(\sigma_k)}{f'(\sigma_k)} \tag{50}$$

until $f(\sigma_k)$ is sufficiently close to 0 (depending on desired accuracy). We observe that $f'(\sigma_k) = \frac{dC}{d\sigma}$ which corresponds to the Greek Vega. Vega has a closed-form expression for many options in the BS framework. The price and the Vega of a European Call option under the BS framework was presented in section 2.1.

4.3 Implementation of the smile expansion

To obtain the prices of the exotic options needed to use the smile expansion is not trivial. Our attempt to do this as accurately as possible is discussed in the sections to follow.

4.3.1 "No Move" option pricing

An issue arise when pricing the "No Move" option since the probability that the underlying asset end up at the same value as it started is essentially 0 and the density function is unknown. We can, as mentioned earlier, use the approximation:

$$p_T(0) \approx P(-\tau + u_0 < u_T < u_0 + \tau)\frac{1}{2\tau} \tag{51}$$

where τ is a small positive but finite number. With this $p_T(0)$ can be found by pricing an option with the payoff function:

$$\text{"No Move" payoff}_1(u_T) = \begin{cases} \frac{1}{2\tau} & \text{if } -\tau + u_0 < u_T < u_0 + \tau \\ 0 & \text{otherwise} \end{cases} \tag{52}$$

To pick τ one has to carefully extrapolate the results to $\tau = 0$. The disadvantage of this approximation is that in a Monte Carlo simulation the option price will come down to the few number of time series that end in the defined interval. This will likely result in a large variation in price from simulation to simulation. We will evaluate if the variation can be reduced by using another, smoother payoff function:

$$\text{"No Move" payoff}_2(u_T) = exp(-u_T^2/2\delta^2)/\sqrt{2\pi\delta^2}, \tag{53}$$

which is a Gaussian function and δ is a small positive value. One have to find a suitable δ, that does not reduce the accuracy of the price estimate but reduces the variance of our estimates, compared to payoff$_1$. The idea is that payoff$_2$ has tails which makes more time series contribute to the option price with smaller payoffs, see figure 4. The price will therefore depend less on the asset ending up in a tight interval.

4.3.2 Binary option pricing

The binary call option has the payoff function:

$$\text{Binary payoff}_1(u_T) = \begin{cases} 1 & \text{if } u_T > 0 \iff x_T > x_0 \\ 0 & \text{otherwise} \end{cases} \tag{54}$$

Although not as necessary as for the "No Move" option, we might be able to reduce the variance of the binary option estimates by smoothing out its payoff function. An alternative is to use:

$$\text{Binary payoff}_2(u_T) = \frac{1}{1 + e^{-u_T k}} \tag{55}$$

where k is some large number to be calibrated. The idea is the same as for the "No Move" option, the alternative payoff function will result in more time series making smaller contributions.

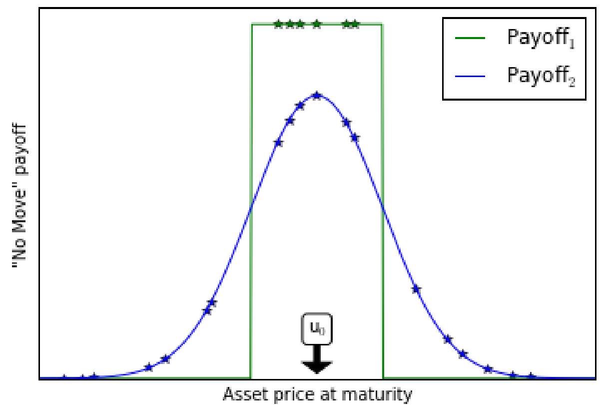

Figure 4: "No move" payoff functions. Illustation of the two payoff functions (52) and (53) for a "No Move" option. The stars are simulated time series that resulted in positive payoffs.

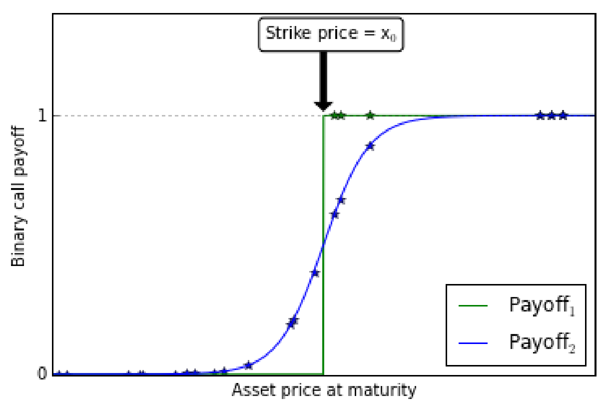

Figure 5: Binary call payoff functions. Illustration of the two payoff functions (54) and (55) for a Binary call option. The stars are simulated time series that resulted in positive payoff.

4.3.3 Theoretical BS smile

According to BS model, the volatility should be independent of all other factors. Thus, if BS assumptions hold, the volatility smile expansion from equation (30) is simplified to:

$$\sigma_{BS}(K) = \sigma \tag{56}$$

This implies that the parameters and their corresponding exotic option prices, if calculated correctly, should take the following values:

$$\begin{aligned}
\alpha_T = 1 &\iff \sqrt{\frac{\pi}{2}}E[|u_T|] = 1 \iff E[|u_T|] = \sqrt{\frac{2}{\pi}} \\
\beta_T = 0 &\iff \sqrt{\frac{\pi}{2}}[1 - 2P(u_T > 0)] = 0 \iff P(u_T > 0) = \frac{1}{2} \\
\gamma_T = 0 &\iff \sqrt{\frac{\pi}{2}}p_T(0) - \frac{1}{2\alpha_T} = 0 \iff p_T(0) = \sqrt{\frac{1}{2\pi}}
\end{aligned} \tag{57}$$

This will come handy when calibrating and testing the smile expansion with data from the GBM.

5 Results

In this section we will test the HMC method's ability to price options and compare the results of the different versions of it. We will also test our implementation of the smile expansion formula and investigate whether it is viable to use in combination with the HMC.

5.1 Accuracy of the HMC methods

The HMC methods are evaluated by pricing assets simulated from the GBM and SABR. The HMC price estimates will be compared to estimates from the simple Monte Carlo approach from section 3.4 and the theoretically correct prices, in line with the evaluation scheme illustrated in figure 2. We will in this section keep the properties seen in table 1 fixed and vary the number of time series and the intervals between the data points in each time series. Shorter intervals between the data points (more steps) in each time series will allow more frequent hedging which could result in higher accuracy. More time series should intuitively give more information which in turn improves the final price estimates.

The following statistics are based on 1000 call option price estimates. The distributions of prices can be seen in the appendix, section 7.2. We are running the methods repeatedly with relatively little data. In the case of our BS data, it appears as if the expected value produced by the HMC methods undershoot the theoretically correct price, although it gets better with more time series. When the asset is simulated from SABR, the variance of the estimates increase

Table 1: Option and simulated asset properties

Property	Notation	Value
Initial asset price	x_0	10
Strike	K	10
Interest rate	r	0.05
Time (year)	T	1
Volatility	σ	0.3
Alpha (SABR)	α	0.5
Rho (SABR)	ρ	0.2
Beta (SABR)	β	1

Table 2: Monte Carlo pricing using BS data. Summary of price statistics.

BS data	250 series, 10 steps		1000 series, 10 steps		250 series, 40 steps	
Model	Mean	SD	Mean	SD	Mean	SD
HMC (Regular)	1.3979	0.0476	1.4164	0.0229	1.3842	0.0775
HMC (Option derivative)	1.4089	0.0455	1.4181	0.0229	1.4099	0.0152
HMC (BS-delta)	1.4227	0.0452	1.4219	0.0226	1.4171	0.0138
Simple MC	1.4234	0.1447	1.4201	0.0751	-	-
Theoretical price	1.4231	-	1.4231	-	1.4231	-

Table 3: Monte Carlo pricing using SABR data. Summary of price statistics.

SABR data	250 series, 10 steps		1000 series, 10 steps		250 series, 100 steps	
Model	Mean	SD	Mean	SD	Mean	SD
HMC (Regular)	1.4207	0.2973	1.4471	0.2477	0.6699	44.2577
HMC (Option derivative)	1.4349	0.2399	1.4435	0.0265	1.3189	0.3926
HMC (BS-delta)	1.4471	0.0542	1.4461	0.0266	1.3479	0.0502
Simple MC	1.4431	0.1758	1.4513	0.0923	-	-
Theoretical price	1.4455	-	1.4455	-	1.4455	-

but the methods still do a good job generally. We will continue by commenting on each of the methods individually.

The "Regular HMC" works decently for our BS data although its mean is furthest from the theoretically correct value out of the different version. It has a tendency to become unstable for our SABR data. The SABR histogram looks decent but outliers (extremely high or negative option prices) distort the mean and the standard deviation. As we have seen earlier, the regression matrices of "Regular HMC" are very large. The inversion of these matrices can produce abnormal results if the sample data is small, the same effect is seen for "Option derivative HMC". The anomaly also seem more pronounced for 100 hedge steps than 10 hedge steps. Many elements in one of the regression matrices of "Regular HMC" contains the term $[x_{k+1} - x_k]$. As the step size decrease, this term becomes smaller. Meanwhile other elements in the matrix, such as x_k, are not affected sizewise. The large differences in size between elements, in combination with many iterations, makes the algorithm prone to errors. This is an argument in favour of using a simplified HMC, in addition to the aspect of computational efficiency.

The instability issue appears for "Option derivative HMC" as well but it less sensitive since it uses with smaller matrices. Generally, "Option derivative HMC" seems to work well.

The "BS-delta HMC" appears to perform best among the HMC methods for BS data, which could be suspected, its hedge is theoretically perfect for such data. It also performs very well on SABR data and there are no apparent issues with instability.

The simple MC's mean values are generally close to the true price but have high variance. This is made especially clear in the histograms in the appendix, section 7.2. In all cases where the HMC methods are stable, they have significantly smaller variance than the simple MC.

A general conclusion we can make is that more time series improves the estimates in all cases, as expected. The consequence of more hedge steps is more unclear, it does not necessarily improve the accuracy of the estimates since it can cause instability. On the other hand, when more steps does not cause instability, as for the "BS-delta HMC", it appears as if standard deviation is reduced but the mean estimate is further from the correct value.

In the following sections we are going to examine the HMC algorithm's ability to produce volatility smiles, which involves pricing options for a range of strikes, unlike the at-the-money tests in this section. Since increasing the number time series seems to be the best way to increase the accuracy, we will test the convergence by running the algorithm with larger amounts of data.

5.1.1 "No Move" option pricing

Before producing volatility smiles with the smile expansion, we will evaluate and calibrate the payoff functions from section 4.3.1 with prices from BS data. For each payoff function and parameter value (δ or τ), we price 100 options and compute the mean and standard deviation of our price estimates. The mean is to be compared with the theoretically correct value presented in section 4.3.3; $p_T(0) = \sqrt{1/2\pi} \approx 0.3989$. The parameters should be extrapolated to zero for the payoff to correspond to that of a "No Move" option. However as the parameters approach 0, the probability that the asset will end in the defined interval approaches 0. As a result, a very narrow payoff function will make the price estimates volatile and prone to error. We want to use an interval for which 1) the expected value is correct and 2) the variance is as low as possible.

Since we do not have formula to calculate BS-delta for a "No Move" option, we have to use either the "Regular HMC" or "Option derivative HMC". Judging from the results in the past section, there is not a huge difference in their performance. We go with the latter for its superior computational efficiency and stability. We begin by testing:

$$\text{"No Move" payoff}_1(u_T) = \begin{cases} \frac{1}{2\tau} & \text{if } -\tau + u_0 < u_T < u_0 + \tau \\ 0 & \text{otherwise} \end{cases} \tag{58}$$

Next, we test:

Table 4: "No move" option pricing with the payoff$_1$

payoff$_1$ - τ	0.01	0.05	0.1	**0.15**	0.2	0.3	0.5	0.7	True price
mean price	0.4296	0.4071	0.3904	0.3986	0.3936	0.3935	0.3851	0.3712	0.3989
SD price	0.1929	0.0921	0.0626	0.0461	0.0376	0.0282	0.0190	0.0098	-

$$\text{"No Move" payoff}_2(u_T) = exp(-u_T^2/2\delta^2)/\sqrt{2\pi\delta^2}, \tag{59}$$

Table 5: "No move" option pricing with the Gaussian payoff$_2$

payoff$_2$ - δ	0.005	0.01	0.05	0.1	**0.15**	0.2	0.3	0.5	True price
mean price	0.3913	0.4018	0.3962	0.3957	0.4017	0.3946	0.3799	0.3595	0.3989
SD price	0.2472	0.1447	0.0663	0.0425	0.0281	0.0294	0.0156	0.0094	-

It can be seen that payoff$_2$ increase the consistency of the estimates, the standard deviations in table 5 are lower than those in table 4 for each $\tau = \delta$. The standard deviations for both versions decrease as parameters increase, which could be expected. 0.15 appears to be a good choice for both parameters, the estimates we get from using 0.15 are close to the theoretical value and the standard deviations are relatively small. For both payoff functions, increasing τ or δ further might reduce the standard deviation but appears to come at the cost of a negative bias in the estimates. Since the Gaussian payoff$_2$ seems to be just as accurate as payoff$_1$, but with lower standard deviation, we decide to use it from here on out.

5.1.2 Binary option pricing

We will now evaluate and calibrate the two different payoff functions used to price the binary option, discussed in section 4.3.2. This will be done in the same manner as for the "No Move" payoffs, by pricing 100 options and comparing the mean and standard deviation. For Binary payoff$_1$ we do not face the issue of picking a parameter value, but it is tested to compare with Binary payoff$_2$. The mean of our estimates is to be compared with the theoretically correct value presented in section 4.3.3; $P(u_T > 0) = 0.5$. Like in the previous section, we will use the "Option derivative HMC". We begin by testing:

$$\text{Binary payoff}_1(u_T) = \begin{cases} 1 & \text{if } u_T > 0 \iff x_T > x_0 \\ 0 & \text{otherwise} \end{cases} \tag{60}$$

Table 6: Binary option pricing with payoff$_1$

True price	0.50000
payoff$_1$ - mean price	0.50054
payoff$_1$ - SD price	0.012477

Next, we test payoff$_2$. As k is increased, this payoff function quickly converge to payoff$_1$ but the smoothing effect will be larger with smaller k values.

$$\text{Binary payoff}_2(u_T) = \frac{1}{1 + e^{-u_T k}} \tag{61}$$

Table 7: Binary option pricing with payoff$_2$

k-value	0.005	0.1	0.5	1	5	10	True price
payoff$_2$ mean price	0.50019	0.50031	0.50146	0.50200	0.50012	0.49883	0.50000
payoff$_2$ SD price	0.00033	0.00073	0.00291	0.00483	0.01132	0.01128	-

By comparing table 6 and 7, we see that the alternative payoff$_2$ seem produce estimates with lower standard deviation. It appears that the expected value is not affected by using the smoothed payoff function, no matter what k-value is used. Although our tests point towards using a very low k-value we will in the coming sections settle for payoff$_2$ and a k-value of 5. By just making a small modification to the geometry of the payoff function, compared to payoff$_1$, we should have a higher chance of succeeding, although it might leave our final result subject to further optimization.

5.2 Black-Scholes volatility smile

In this section we display volatility smiles that are generated from BS data simulated according the specification in table 8. The smiles are generated from call options with a range of strikes, from the strike that gives $\Delta = 0.05$ (far out-of-the-money) to the strike that gives $\Delta = 0.95$ (far in-the-money).

All plots show decent approximations of the BS smile, which is flat $\sigma_{BS}(K) = \sigma = 0.3$. It can be seen that the smiles tend to be less accurate far from at-the-money. For the HMC algorithm alone, this is because the price of a far

Table 8: Specification for option and BS data

Property	Notation	Value
Initial asset price	x_0	10
strike	K	10
Interest rate	r	0
Time	T	1 week (1/52 year)
Volatility	σ	0.3

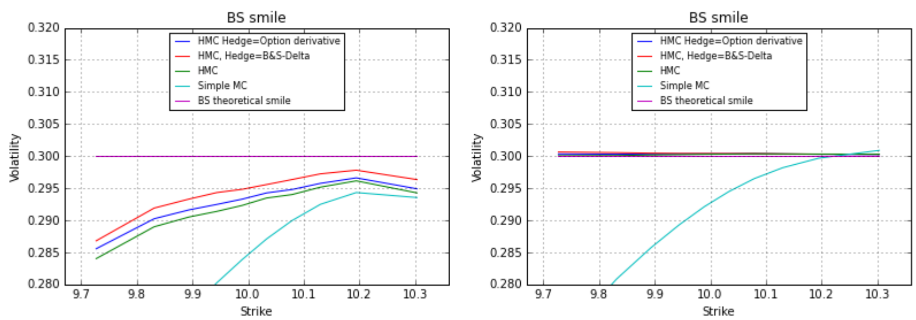

Figure 6: BS HMC smiles. BS volatility smile from the HMC methods (10 hedge steps, 1000 time series (left) and 100 000 time series (right)).

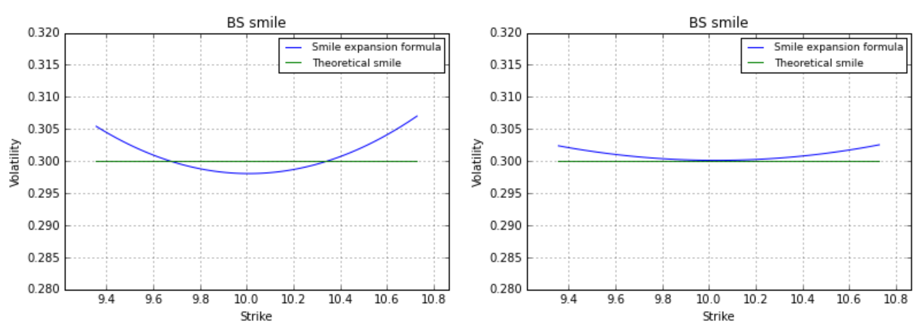

Figure 7: BS smile expansions. BS volatility smile from smile expansion with "Option derivative HMC" with (10 hedge steps, 1000 time series (left) and 100 000 time series (right)).

out-of-the-money option will depend on the few number of time series that produce positive payoffs and similarly, the price of a far in-the-money option will will depend on the few number of time series that does not produce positive payoffs. This dependency makes the estimates volatile. The price of an option that is close to at-the-money does not have the same kind of sensitivity, there is an approximately equal number of time series ending with and without positive payoffs.

The "Option derivative HMC" appears to converge better than the "Regular HMC", judging from the BS-smiles and tests in previous sections, and it was therefore used for the smile expansion. The formula for smile approximation is, as presented in section 3.7:

$$\sigma_{BS}(K) = \sigma(\alpha_T + \beta_T \mathcal{M} + \gamma_T \mathcal{M}^2 + O(\mathcal{M}^3)), \quad \mathcal{M} = \frac{K - x_0}{x_0 \sigma \sqrt{T}} \tag{62}$$

We observe that when the option is at-the-money ($S = K = 10$), $\mathcal{M} = 0$ and therefore $\sigma_{BS}(10) = \sigma\alpha_T = 0.3\alpha_T$ (which should be equal to 0.3). We see in figure 7 (especially in the right graph) that the smile expansion and for $K = 10$ is very close to 0.3, which means that α is estimated accurately.

β_T determines the skew of the smile. The BS smile should be flat and unskewed. We can get an idea of β_T's accuracy by comparing the ends of the volatility smile, which should be level, or by looking at the derivative of the volatility at-the-money, $\sigma'_{BS}(10) = \sigma\beta_T = 0.3\beta_T$ (which should be equal to zero). β_T appears to be accurately estimated, especially with more time series, as seen in figure 7.

γ_T seems to be term that contribute most to the errors. The errors are apparent in both ends of the volatility smile even with 100 000 time series. This is expected for two main reasons; firstly, finding γ_T involves pricing the "No move" option which is a difficult task. Secondly, it is multiplied by a factor 2 term, which makes small deviations from the theoretically correct value (in this case zero) significant for options that are far in-the-money or far out-of-the-money.

The HMC algorithms alone seem to converge better but it is hard to draw any conclusions regarding the relative accuracy of using the smile expansion or not. It is worth noting that the curves in figure 6 and in figure 7 use the same amount of data, but the smile expansion is obtained from pricing just 4 options. The HMC curve, on the other hand, is produced by pricing an option at every point of the volatility curve. Each curve in figure 6 is constructed from 20 option prices, which is 5 times more than the smile expansion curve, and they are still pretty jagged.

5.3 SABR volatility smile

In this section we display volatility smiles that are generated by pricing options with an following asset SABR, simulated with the specification in table 9. The smiles are generated from call options with a range of strikes, from the strike that gives $\Delta = 0.05$ (far out-of-the-money) to the strike that gives $\Delta = 0.95$ (far in-the-money). The produced smiles are compared with the approximate theoretical SABR smile presented in section 2.6.

The convergences are pleasing for SABR data, especially the smile expansion

Table 9: Specification for option and SABR data

Property	Notation	Value
Initial asset price	x_0	10
strike	K	10
Interest rate	r	0
Time	T	1 week (1/52 year)
Volatility at time 0	σ_0	0.1
Alpha (SABR)	α	0.4
Rho (SABR)	ρ	-0.1
Beta (SABR)	β	1

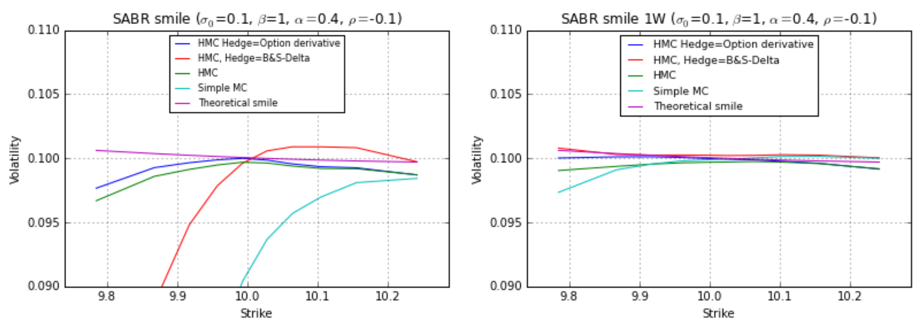

Figure 8: SABR HMC smiles. SABR volatility smile from the HMC methods (10 hedge steps, 10 000 time series (left) respectively 100 000 time series (right)).

seems to produce accurate results. It is worth noting how the "BS-delta HMC" performs well only for options close to being at-the-money in the left picture of figure 8, where 10 000 time series are used. This reveals an important point that we did not see in section 5.1. The outperformance of the HMC methods versus the simple MC is also striking.

We continue by examining the performance for steeper curvature. The following smiles are also for 1W options but with $x_0 = 1$. The SABR parameters are specified each graph. It is clear that the smile expansion does not capture the skewness very well, all of the HMC versions alone do a very job though. Problems arise in the ends of the volatility smile, the same behavior as we saw earlier for smiles from BS data. The simple MC show poor convergence and the "BS-delta HMC" is not as good as the other HMC versions.

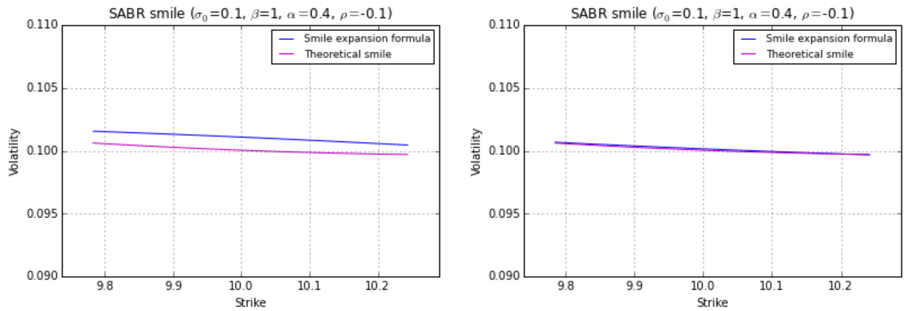

Figure 9: SABR smile expansions. SABR volatility smile from smile expansion with ”Option derivative HMC” (10 hedge steps, 10 000 time series (left) and 100 000 time series (right)).

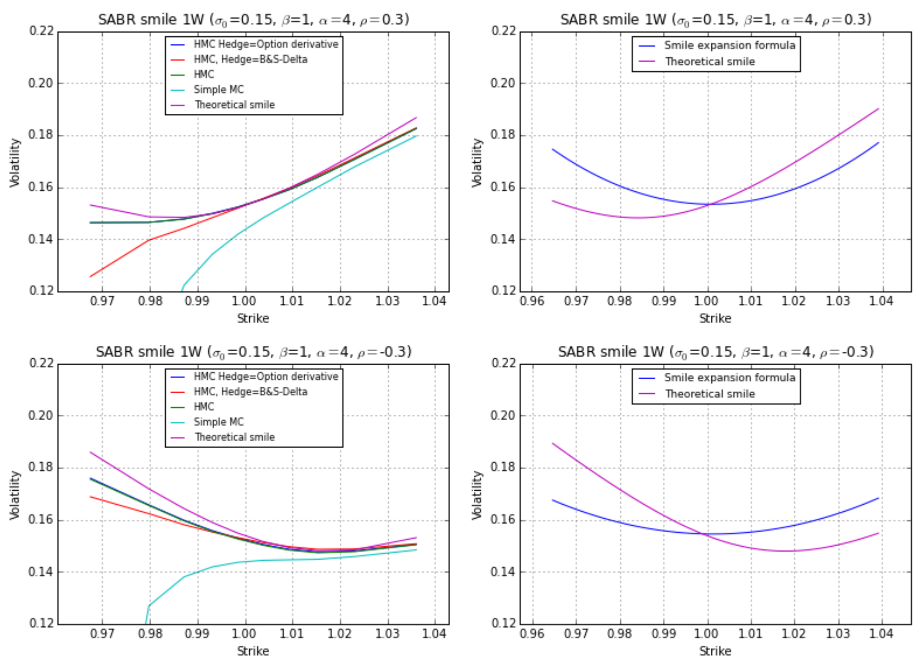

Figure 10: Convex SABR smiles. SABR volatility smiles from the HMC algorithms alone to the left and from smile expansion with ”Option derivative HMC” to the right (10 000 time series and 10 hedge steps).

5.4 Empirical volatility smile

Next, we are going produce empirical volatility smiles for the currency pair GBPUSD (British Pound to US Dollar). We use 10 years of historical spot data for the exchange rate, from November 1, 2006, to November 1, 2016, sampled with bihourly frequency. We will price a 1 week option since the smiles are well defined for that tenor and it is short enough to get a decent amount of time series out of the sample. We clean the data from non-trading days and divide it into time series containing 1 week of trading each (GBPUSD is trading for 122 hours per week). This will give us slighly more than 500 time series (52 weeks per year, times 10 years, minus a few weeks to account for holidays). To simplify things, we re-base the data to x_0=1 for all time series. Like before, volatility is plotted against strike price, from the strike for which $\Delta = 0.05$ to the strike for which $\Delta = 0.95$. Additionally, we plot histograms (corresponding to the y-axis on the right-hand side) displaying the distribution of the data (as x_T/x_0). We compare the results to a SABR smile that is a calibrated (with constraint $\beta = 1$), by Bank of America Merrill Lynch, to approximately match current (as of early December 2016) 1W GBPUSD volatility smiles. The calibrated smile has the following parameters:

Table 10: Calibrated 1W SABR smile for GBPUSD

Parameter	Value
Initial volatility (σ_0)	0.1
Alpha (α)	3.8
Rho (ρ)	-0.09
Beta (β)	1 (constraint)

It should be made clear that the calibrated SABR volatility smile is not a smile we can expect convergence towards, it is itself an approximation and it is reflecting current market conditions. It is neither a good basis for drawing conclusions about the relative performance of the different versions of HMC. Its sole purpose is to give an indication of whether we can produce reasonable empirical smiles with little data.

The "Regular HMC" and the simple MC have some issues pricing the far in-the-money options. Apart from that, all of the smiles are reasonable. It is worth noting how well the convex curvatures are reflected by our algorithm.

The estimated volatility of the 10 year data is slightly higher than σ_0 in table 10, which could help to explain the vertical difference between our empirical smiles and the calibrated SABR.

We conclude the result section by displaying empirical overnight (ON) smiles for the past 10 years. Each smile is produced using 2 years of data, which approximately gives 500 trading days, or equivalently, 500 times series.

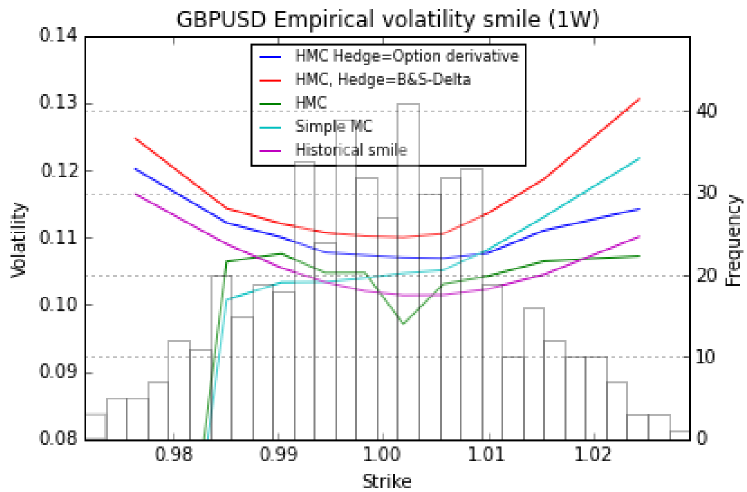

Figure 11: Empirical 1W smiles from HMC. GBPUSD empirical 1 week volatility smile from using only HMC (roughly 500 time series and 60 hedge steps). Histogram shows distribution of underlying asset (x_T/x_0)

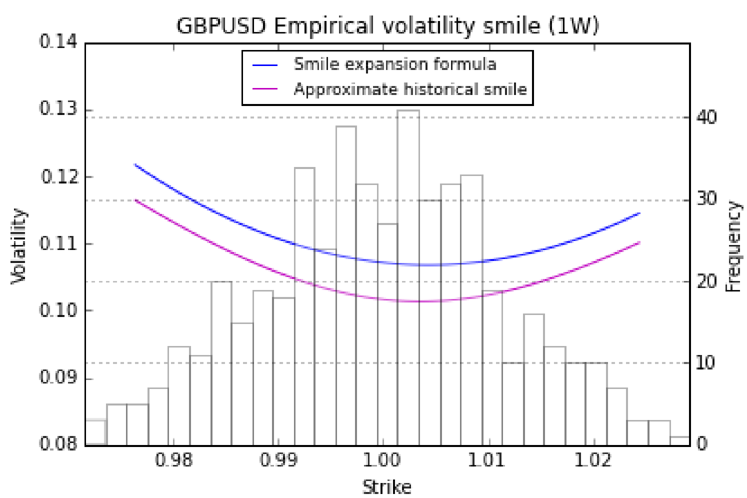

Figure 12: Empirical 1W smile expansion. GBPUSD empirical 1 week volatility smile from HMC with the smile expansion formula (roughly 500 time series and 60 hedge steps). Histogram shows distribution of underlying asset (x_T/x_0)

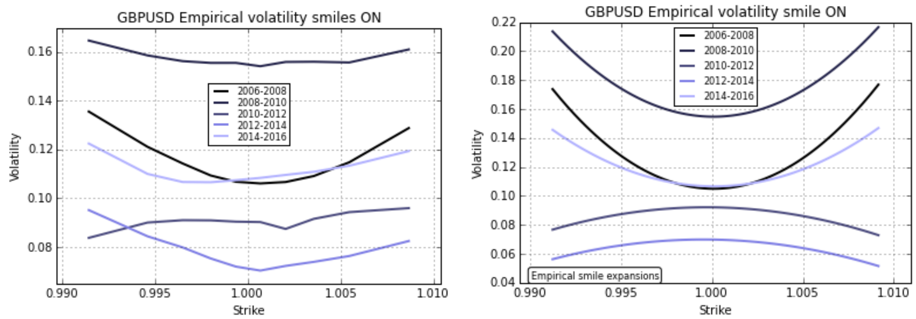

Figure 13: Empirical ON smiles. GBPUSD empirical overnight volatility smiles from the "Option derivative HMC" alone to the left and together with smile expansion to the right (roughly 500 time series and 24 hedge steps).

5.5 Discussion

The HMC algorithm that we develop in this thesis is an empirical model and it is shown to work based on tests. BS model, on the other hand, is a model that hold perfectly well in theory but, as we have touched upon many times before, does not hold perfectly in the real world and neither does any other model for pricing financial assets. With that being said, it can seem as an weakness that all our validations are purely empirical but for a model to be of practical use, empirical evidence can be stronger than theoretical proofs based on questionable assumptions.

5.5.1 HMC versions

The HMC algorithms we have developed and tested show satisfying convergence and clearly outperforms the simple MC. The empirical testing in this thesis have led us to believe that "Option derivative HMC" is the most preferable version of HMC. It does well in all tests, for all types of data. The "BS-delta HMC" is efficient and stable but underperforms relative to the other versions for SABR data, raising doubts about its accuracy for data that is not BS. The "Regular HMC" is not totally reliable, it can become unstable and produce results that differ from the other methods by much, this is seen for both BS and real data. It is also the least computationally efficient version.

The HMC methods show a tendency to undershoot the correct value with few time series (e.g. see figure 16). This can likely be attributed to the residuals being log-normally distributed, like the underlying asset in our BS and SABR simulations. The minimization of least squares residuals relies on normally distributed residuals with a mean value of zero (as stated in [11]) and therefore it only works as an approximation for log-normally distributed residuals. This would explain why the undershooting becomes more pronounced for more hedge steps which results in more least squares minimization. It also becomes clear why the effect is worst for the "Regular HMC", its hedge is entirely based on the log-normally distributed asset prices.

Further, we believe that utilizing the "smile expansion formula" in combination with HMC can make a powerful method, it appears to converge well but with some deviations far from at-the-money. It also greatly improves the computational efficiency of producing a complete volatility smile. The pleasing convergence is a consequence of the HMC method performing well for at-the-money options and that we are able to accurately price "No Move" and binary options. The last part can partially be attributed to our smoothed payoff functions for the exotic options. They proved to successfully reduce the variance of the estimates. However, it seems to struggle with reproducing the skew of asymmetric SABR smiles.

To obtain a smile from the HMC algorithm alone is more reliable and can handle all curvatures. Its major drawback is pricing options that are far from at-the-money, where significantly more data is required in order to obtain the same precision as for at-the-money options.

5.5.2 Empirical smiles

The ultimate goal of the thesis was to develop a method that is able to produce empirical volatility smiles which is accomplished and section 5.4 show results that are promising for the methods to be used in practical settings. It is important to point out that the SABR smile we had as comparison is calibrated with respect to the current market, while the algorithm's empirical smiles are based on the past 10 years. When pricing real options it is important to consider whether the sample data reflect current market conditions. This is not an issue when pricing options from simulated data, as the data is simulated with constant parameters. However in practice, characteristics such as volatility change over time. We might have been able to better match the calibrated SABR smile with more years of data and by only using trading periods that are similar to today's market. An example of a simple adjustment one could make is to divide the sample into high volatility and low volatility parts and use the part that is deemed most suitable given the current market conditions.

Either way, we cannot expect our models to converge exactly to the calibrated SABR, as it is an approximation itself. The important take-aways are that the curvatures are similar and the levels of volatility are roughly the same. The artificial SABR and BS smiles in earlier sections constitutes better model validations.

5.5.3 Optimizing the performance

To use many time series is a crucial factor for the accuracy of the HMC method. Like for any other Monte Carlo method, evaluating more scenarios gives more information from which the right conclusions can be drawn. One could imagine that several hedge steps would improve the accuracy. This was shown to not necessarily be true, but more hedge steps did reduce the variance if instability was avoided. This has led us to suspect that a combination of many time series and many hedge steps could preserve stability and produce the best possible results, but the computational demands have prevented us from thoroughly testing this hypothesis.

We think our results show that the HMC method is a clear step in right direction for using Monte Carlo methods in the field of option pricing, but a few aspects are of concern for it to be put into actual use. The main aspect is probably the amount of data needed for convergence, we saw that it demanded plenty of data for both BS and SABR volatility smiles. Take an example; you want to price a 1-day option. To get a decent approximation of the volatility smile, our tests indicate that you probably want at least 10 000 time series. This means 10 000 trading days which is equivalent to approximately 40 years of data. Firstly, this amount of data does not exist for many assets. Secondly, even if the data exists, it is old and might not fairly reflect today's market conditions.

The HMC is a quite big model with plenty of opportunity to tweak different parameters. To optimize the model with respect to every single one of these is a huge task. As a consequence, we have throughout the thesis taken decisions in what may have seemed like an arbitrary manner to be able to arrive at a final judgement for the HMC method. We therefore wrap this thesis up by pointing out some of these areas in section 6.1, mentioned previously in thesis or not, where more work can be done for someone wanting to improve the implementation of the HMC method.

6 Conclusion

All in all, our implementation and the results produced from it has shown that the HMC method is powerful and that the goal of the thesis, to produce empirical volatility smiles, is achieved. Although the algorithm might require some further tweaking and testing, we think that it is not far from being of practical use in the financial markets. The HMC method is especially suitable for non-standard options with short tenors. Non-standard, because it means that there are not previous quotes to rely on and short tenors, because for such one can divide historical data into a larger amount of time series.

The smile expansion formula is an interesting addition to any numerical option pricing algorithm, predominantly because it greatly improves the computational efficiency for producing a volatility smile. However, from our experience in this thesis, it is not reliable when it comes to producing skewness of the smiles.

6.1 Suggestions for further research

Monte Carlo option pricing is by no means a finished chapter, there are plenty of areas to be further researched. There is obviously the possibility of developing completely new methods, but in terms of the HMC we are able to offer some concrete suggestions on possible things to work on. The points to follow were taken into consideration during our work but left out due to the limited size of the project.

To improve the performance it could be of interest to use other basis functions. We used cubic splines but there are endless of options out there.

Another area of potential improvement are the knots (if splines are used). We thought 8 was a reasonable number, but did not test this decision extensively. With regards to the placement of those knots, we relied on intuition for what should be decent, so there may very well be a better alternative.

An interesting addition would be to extend the least squares minimization with a Lasso regression that would aim to minimize the change in parameter values between each step in the algorithm. The thinking behind this is that the hedge and option prices (as functions of the stock price) should not change drastically between two iterations. We think that this addition could improve the consistency of the estimates and prevent instability.

Moreover, it would be pleasing if the algorithms were validated for more theoretical frameworks, primarily to ensure convergence. It is also relevant to investigate the results of using many time series in combination with many hedge steps. In the real world frequency is not as big of an issue as number of time series, data can be obtained with really high frequency but for many assets existing or relevant historical data is limited. Unfortunately it is very demanding to run HMC with many hedges, doubling the number of hedges doubles the computational requirements which makes the algorithm take twice as long time to run.

In addition to more theoretical optimization of the methods, there is plenty of empirical testing to be done, that part was kept short in this thesis. Because this thesis is written in co-operation with a FX team at Bank of America Merrill Lynch, our empirical testing was also limited to foreign exchange but the method is equally applicable to any other asset. It is left for the future to the test the real world performance more carefully with historical data from a range assets.

7 Appendix

7.1 Independence of basis functions

We will now independence of the basis functions used throughout the thesis, as defined in section 3.5.

$$\begin{aligned} f_0(x) &= 1 \\ f_1(x) &= x \\ f_2(x) &= x^2 \\ f_s(x) = (x-\xi_s)_+^2 \text{ for } s &= 3 \dots M \end{aligned}$$

$$\text{where } (z)_+ = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z > 0 \end{cases}$$

The basis functions are independent if it exists $x_0, x_1, \ldots, x_M$ such that the column vectors

$$\begin{bmatrix} f_0(x_0) \\ f_1(x_0) \\ \vdots \\ f_M(x_0) \end{bmatrix}, \begin{bmatrix} f_0(x_1) \\ f_1(x_1) \\ \vdots \\ f_M(x_1) \end{bmatrix}, \cdots, \begin{bmatrix} f_0(x_M) \\ f_1(x_M) \\ \vdots \\ f_M(x_M) \end{bmatrix}$$

are linearly independent. In our application, all $\xi_k > 0$, although this is not necessary to show independence. Let:
$x_0 = 0$
$0 < x_1 < x_2 < \xi_3$
$\xi_3 < x_3 < \xi_4$
$\xi_4 < x_4 < \xi_5$
$\vdots$
$\xi_M < x_M$
This will make the vectors:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ x_2 \\ x_2^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ * \\ * \\ * \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} 1 \\ * \\ * \\ * \\ * \\ \vdots \\ * \\ * \end{bmatrix}$$

where $*$ is some positive number. It is easy to see that no vector can be expressed as a linear combination of the other vectors and the basis functions must therefore be independent.

7.2 Distribution of estimates

In this section we display the distribution of our tests in section 5.1. The distributions are made up of 1000 price estimates. The different histogram display a varying number of hedge steps and time series (10 or 100 hedges and 250 or 1000 time series).

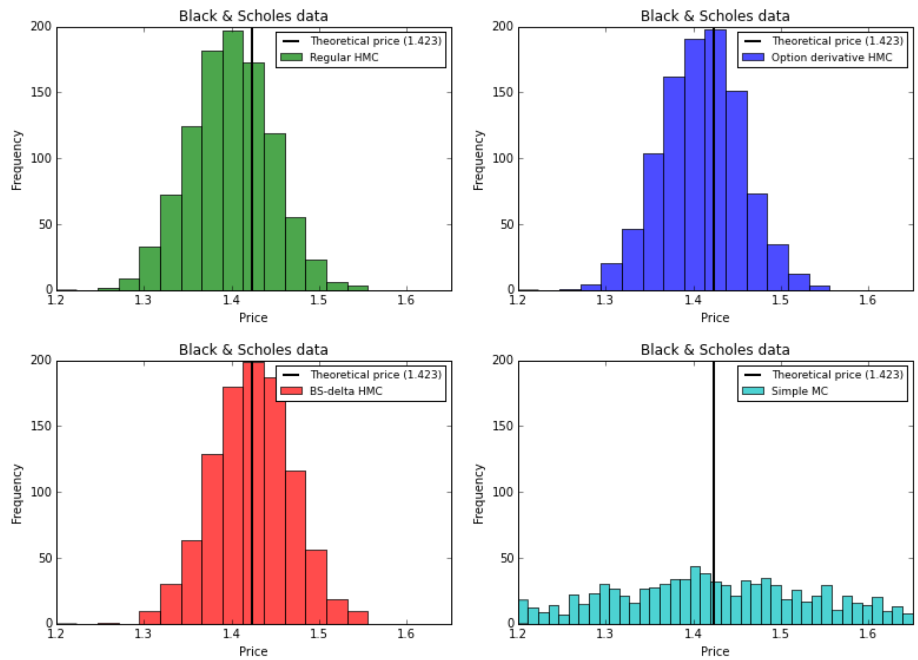

Figure 14: Histograms from BS data, 250 time series, 10 hedges. Distribution of price estimates from Black & Scholes data containing 250 time series, priced with 10 hedge steps.

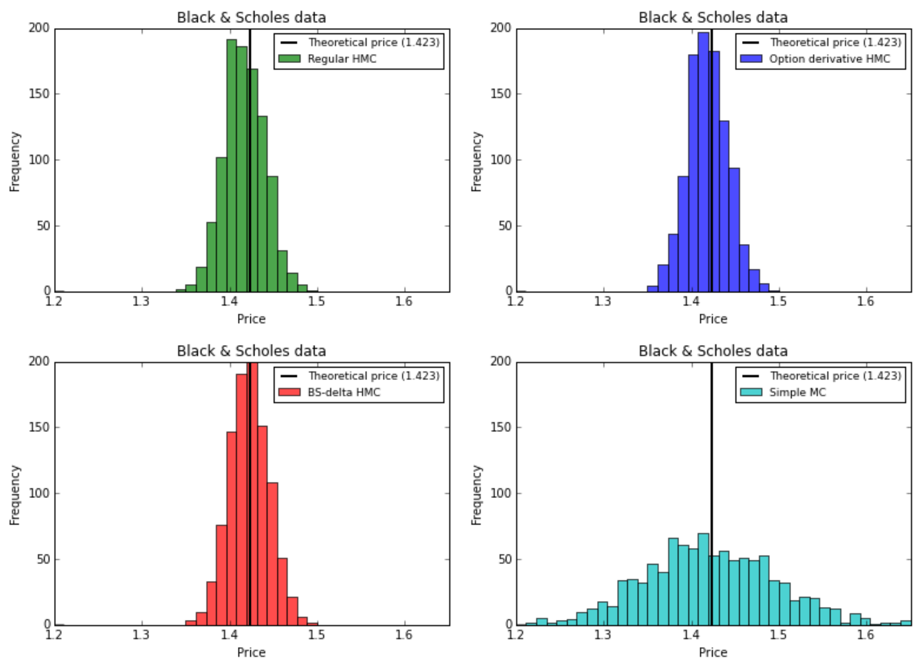

Figure 15: Histograms from BS data, 1000 time series, 10 hedges. Distribution of price estimates from Black & Scholes data containing 1000 time series, priced with 10 hedge steps.

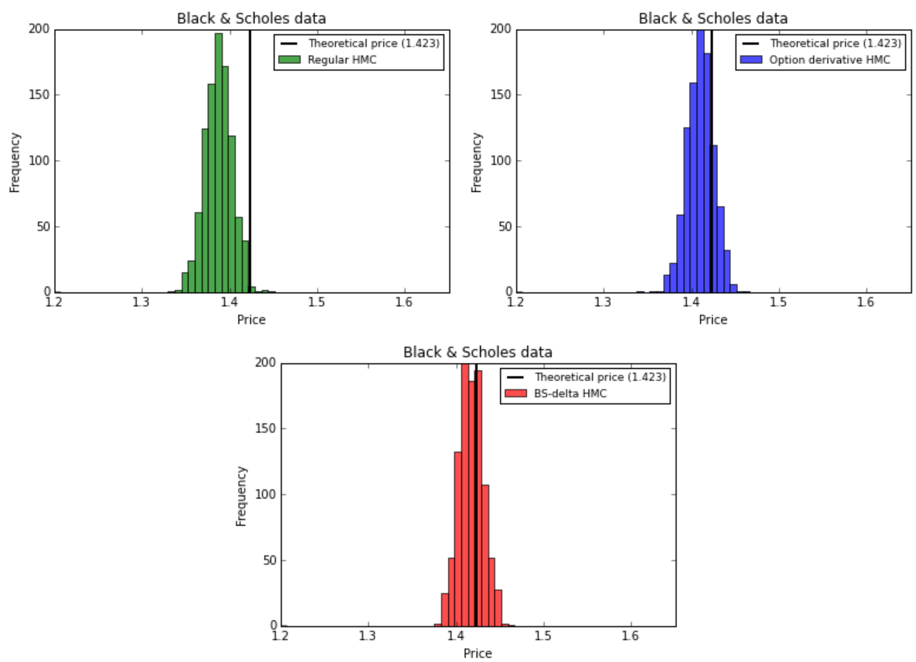

Figure 16: Histograms from BS data, 250 time series, 100 hedges. Distribution of price estimates from Black & Scholes data containing 250 time series, priced with 100 hedge steps.

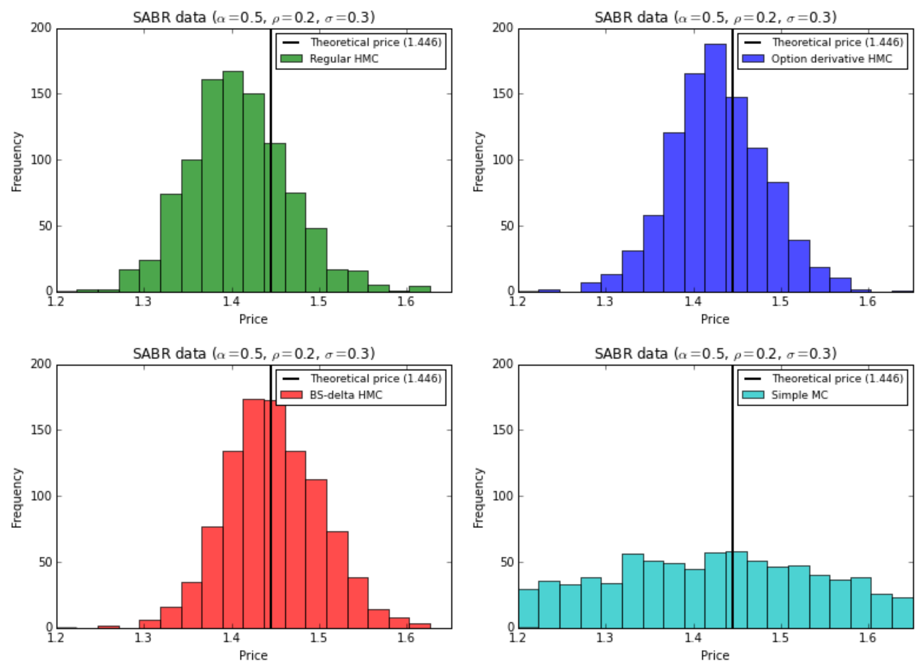

Figure 17: Histograms from SABR data, 250 time series, 10 hedges. Distribution of price estimates from SABR data containing 250 time series, priced with 10 hedge steps.

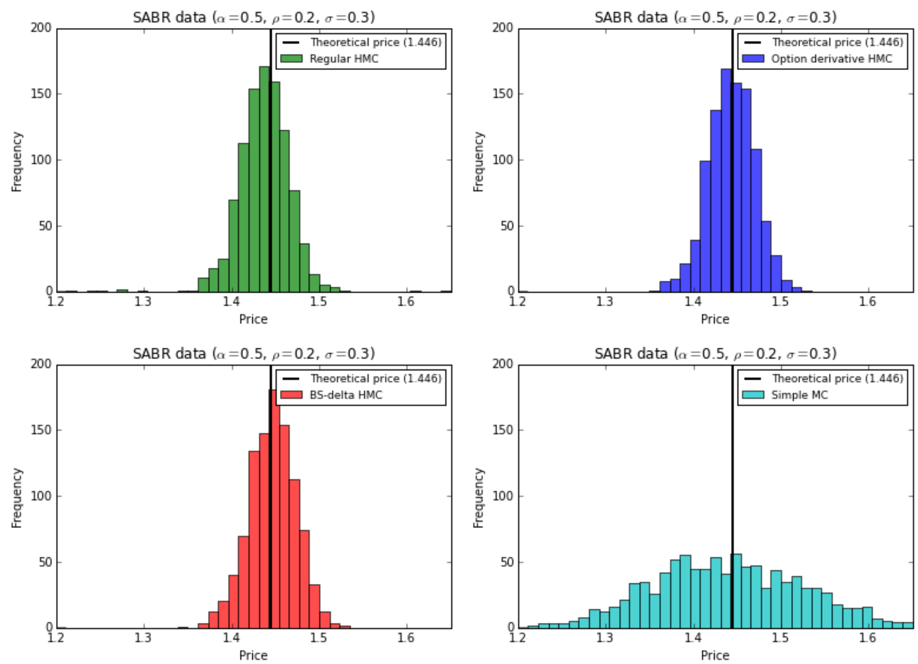

Figure 18: Histograms from SABR data, 1000 time series, 10 hedges. Distribution of price estimates from SABR data containing 1000 time series, priced with 10 hedge steps.

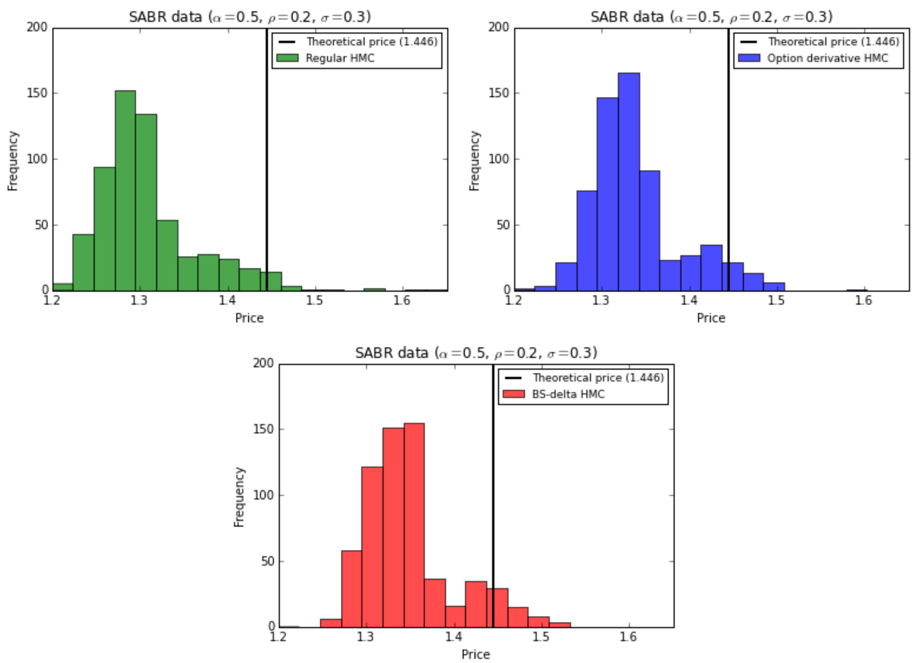

Figure 19: Histograms from SABR data, 250 time series, 100 hedges. Distribution of price estimates from SABR data containing 250 time series, priced with 100 hedge steps.

References

[1] Marc Potters, Jean-Philippe Bouchaud, and Dragan Sestovic. Hedged monte-carlo: low variance derivative pricing with objective probabilities. *Physica A: Statistical Mechanics and its Applications*, 2001.

[2] L. De Leo, V. Vargas, S. Ciliberti, and J.-P. Bouchaud. We've walked a million miles for one of these smiles. *Risk Magazine*, 2012.

[3] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 1973.

[4] J.C. Hull. *Option, Futures, and Other Derivatives*. Prentice Hall, eighth edition, 2012.

[5] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Managing smile risk. *Wilmott Magazine*, 2002.

[6] P. Gauthier and P-Y. H. Rivaille. Fitting the smile. Working Paper, Pricing Partners, 2009.

[7] J. Oblój. Fine-tune your smile: Correction to hagan et al. *Wilmott Magazine*, 2008.

[8] B. Chen, C.W. Oosterlee, and H. van der Weide. Efficient unbiased simulation scheme for the sabr stochastic volatility model. *International Journal of Theoretical and Applied Finance*, page 9, 2011.

[9] P. Boyle. Options: A monte carlo approach. *Journal of Financial Economics*, pages 323–338, 1977.

[10] F. Longstaff and E. Schwartz. Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 2001.

[11] J.O. Rawlings, S.G. Pantula, and D.A Dickey. *Applied Regression Analysis: A Research Tool*. Springer, second edition, 1998.

[12] W. Cheney and D. Kincaid. *Numerical Mathematics and Computing*. Thomson Brooks/Cole, sixth edition, 2008.

[13] L. De Leo, T.-L. Dao, V. Vargas, S. Ciliberti, and J.-P. Bouchaud. Presentation: Smile in the low moments. 2014.

[14] W. Cheney and D. Kincaid. *Numerical Mathematics and Computing*. Thomson Brooks/Cole, sixth edition, 2008.