Numerical Analysis and Stochastic Modeling in Mathematical Finance

A dissertation presented for partial fulfillment of
the requirements for the degree Doktor Ingeniør

Lars Oswald Dahl
Norwegian University of Science and Technology,
Department of Mathematical Sciences
and
Storebrand Investments

August 2002

Acknowledgements

I am very grateful to **Fred Espen Benth** for the enormous quality of his supervision. He has remarkable abilities and profound sincerity both towards mathematics and supervision. **Hege**, thank you for pointing out the important aspects. You have illuminated my path. I am grateful for the support I have received from Storebrand Investments and especially **Jørgen Larssen** and **Age Bakker** who have encouraged me to keep up the pace. I also want to express my gratitude to **Helge Holden** who has facilitated the finish as well as the startup. I want to thank my cooperates at Storebrand Investments; **Lars Mouland**, **Kjell Sverre Hatlen** and **Snorre Storseth**, who have motivated me an contributed to a pleasant working environment. The group around and arrangements under direction of **Bernt Øksendal** have been an important part of my apprenticeship and laid a foundation for my acquirement of knowledge in the field. **My parents** indirect influence through their enthusiasm and excitement for their own research during my childhood and adolescence have nursed my interests in this similar direction. Last but not least I want to thank **Bjørn Magnussen** and **Bjørn Erling Vembe** for their mentorship and edifying talks during my first fumbling year as a dr. student. You have all contributed to the completion of the project and a joyful period.

The first year of the study was done with a grant from the Norwegian Research Council, while the major part has been carried out with financial support from Storebrand Investments as part of my employment there. Supervisors have been **Helge Holden**, Department of Mathematical Sciences – Norwegian University of Science and Technology and **Fred Espen Benth**, Department of Mathematics – University of Oslo.

Oslo, August 2002
Lars O. Dahl

Contents

Chapter 1

Introduction

The main goal of this thesis has been to study and develop faster and more accurate methods for pricing and hedging exotic options. This has involved work on the models describing prices and hedges as well as the stochastics driving them. We have also put effort into algorithmic interpretation and implementation of the models to enable efficiency measurement with regards to computing time. In some of the articles we have aspired to find criteria to decide whether the pricing methods we have developed can be expected to perform well, enabling practicians to find a good numerical method for their given pricing/hedging problem easier. However, the most optimistic reader must be warned: We have not found one single method that works best for all types of option pricing problems, and we do not think that such a method exists. Pricing and hedging of exotic options involve thorough knowledge of the problem at hand, and the mastering of a tool box of numerical methods from which a suitable one can be picket. We believe, however, that the thesis contributes some to the enlargement of the tool box.

The fundamentals for developing derivatives and their pricing methods is the theory of mathematical finance. This theory gives a consistent model for the market, with well defined assumptions and broadly studied shortcomings. The shortcomings of the current theory and the methods based on it, are the background for striving for new and more realistic treatment of the instruments and events in the market. This is done both by searching for extensions of the theory itself to improve its ability to embrace the abundance of the market and its random behavior, and by developing improved models accepting the assumptions of the current theory. The research presented in this thesis is mainly in the last category.

In very short terms I will sum up the research area of the thesis and give the motivation for the chosen themes and the main results. In this introduction chapter I will also try to highlight the red thread of the work. The foundation on which all our work rests is as follows: The theoretical market is denoted $\mathcal{M}$. The consistency of the market is achieved through formulation of a set of rather abstract properties and limitations, which I will not go through in detail. It can be found in, e.g., [KS98]. It is, however, important to notice that the model of the price processes of the basic assets in the market is very confined by $\mathcal{M}$. That is; in order to fulfill basic intuitive

principles in the market, such as no arbitrage, we have limited freedom in choosing the models and the stochastics of the prices of the assets in the market. Our research uses the standard market as a basis, and the formulation of the derivative contracts we work with is consistent with this. A derivative instrument is a construction on $\mathcal{M}$ to achieve a special exposure towards the basic assets contained in $\mathcal{M}$. In the standard market all derivatives are actually redundant, in that their exposure can be constructed by trading in the basic assets (which includes the risk free asset). It can, however, be cumbersome to achieve this in practice, and therefore a marketplace for such redundant exposures have emerged. In addition, this market is continuously developing towards more specialized needs, challenging both practicians and researchers.

The modeling of the price processes of the basic assets in $\mathcal{M}$ is central for the development of effective methods for pricing the derivative instruments. In the first five articles[1] of the thesis the price processes of the $N+1$ basic assets are modeled by the geometric Brownian motion leading to the expressions

$$\begin{aligned} S_0(t) &= \mathrm{e}^{rt} \\ S_n(t) &= S_n(0)\exp\big((r-\frac{1}{2}\sum_{d=1}^{N}\sigma_{nd}^2)t+\sum_{d=1}^{N}\sigma_{nd}W_0^{(d)}(t)\big), \quad n=1,\ldots,N\ , \end{aligned}$$

where r is the risk free rate, and the volatility matrix $\sigma \in \mathbb{R}^{N\times N}$ is such that $\sigma\sigma^T \equiv \sigma^2$ is the covariance matrix of the returns of the assets. The stochastics of the prices are modeled by the Brownian motion $W_0(t)$. This process is an essential part of $\mathcal{M}$. Let $\theta \in \mathbb{R}^N$ be given by the relation $b - r\underline{1} = \sigma\theta$, where $b \in \mathbb{R}^N$ is the vector of drift coefficients for the assets of the market. θ is called the market price of risk. By using the Girsanov theorem (see [Øks98]) it can be shown that

$$W_0(t) = W(t) + \int_0^t \theta ds, \quad \forall t \in [0,T]$$

is an N-dimensional Brownian motion under the risk free measure P_0 relative to the filtration $\{\mathcal{F}(t)\}$ of $W(t)$. The process $W(t)$ is the Brownian motion observed for the assets in the market under the market induced probability measure P. For a more comprehensive survey of these aspects, see [KS98, Ch. 1]. Actually there are several properties of the Brownian motion $W_0(t)$ that can be utilized to construct fast pricing algorithms for the derivatives. We particularly take advantage of this in the articles where we price so called Asian options.

Next, to recapitulate for the reader the notion of a derivative instrument, let us look at the pricing of a European call option contract on one asset in the market. The holder of a call option contract is entitled to receive the underlying asset by paying the

[1] In the last article of the thesis, we model the prices of the underlying processes by a Schwartz mean reverting model. This does, however, not conflict with the construction of a well behaving market $\mathcal{M}$. Actually, analogous modeling aspects need to be performed in this setting to achieve speed up, as in the GBM setting. Please refer to the article in chap. 7 for details.

strike price[2] q at the maturity time T. He has earned money if the asset price $S(T)$ at T is above the strike price. If, however, the price is below the strike, a rational actor will do nothing. The question we face at time $t = 0$, is how much this option contract should cost. In $\mathcal{M}$, a unique solution to this problem exists. It is given by

$$V(0) = \mathrm{e}^{-r(T)}\mathrm{E}_0[f(S(T))] \,, \tag{1.1}$$

where f in this simple case is given by $f(X) = (X - q)^+$. Most methods looked at throughout the thesis will be valid for more general f, and in most of the articles we consider a more general argument X to f than in (1.1). Actually, this simple expression can be used and modified to elaborate on the finance semantics of the field; We have call options and put options. If we use $f(X) = (q-X)^+$ we get a put option. That is the right to sell the underlying instead of buying it. If $X = \sum_{n=1}^{N} S_n(T)$, we have a basket option. If $X = \sum_{k=1}^{K}(S(t_k))$ we get an arithmetic average Asian option contract, and the obvious combination gives an Asian basket option. Throughout the thesis we will mainly look at these types of contracts in the numerical examples, even if the methods we develop can be applied in more general settings. We do of course give a more rigorous definition of the contracts as we use them in the articles. The first article of the thesis elaborates on the question: How wrong is it to collapse the basket and treat it as one asset, with a corresponding deduced volatility. This would lead to a simpler pricing procedure[3] compared to the apparatus needed to price the full basket option. The findings are interesting both from a practical and theoretical point of view.

An interesting and sophisticated symbiosis exists between the option pricing problems and the complex field of random number generation. In order to be able to price many of the option contracts we have to use so called Monte Carlo (MC) and quasi-Monte Carlo (QMC) methods. In the thesis we look exclusively on these types of methods, and especially try to identify and optimize the speedup that can be achieved by employing the latter. The use and success of these methods are based on basic properties of probability theory. The conventional mean-value estimator of the expected value can, by the strong law of large numbers, be shown to converge to the true expected value with probability one as the number of samples (n) goes to infinity. This implies that the error of the Monte Carlo estimator, which is the process defined by the difference between the expected value and the mean value, is normal distributed with mean 0 and variance σ_f^2/n, where σ_f^2 is the second moment of f. The convergence speed of the MC method in its native form is therefore $\sigma_f/\sqrt{n}$. In order to put these methods to play in an optimal way in the option pricing problem, we have to look into the price process as modeled in (1.1). By definition, the expected value can be written as an integral. This is convenient, since we then can think of the problem as an integration problem. The integration error stated above is large

[2] The strike price is labeled K in the article "On Derivatives of Claims in Commodity and Energy Markets Using a Malliavin Approach" in chap. 7

[3] Although the deduction of this one asset option pricing formula is by far not trivial, and in fact highly celebrated. It leads to the famous Black and Scholes option pricing formula, for which its originators received the Nobel prize in economics in 1994

compared to deterministic methods if the goal is to evaluate integrals in one or two dimensions. However, the convergence rate of the MC method remains the same as the dimension of the problem is increased, while the convergence rate of any deterministic integration method will degenerate with dimension, leaving the MC approach superior as the dimension passes four or five. Since the challenge in option pricing often occurs as a consequence of increased problem dimension, the MC and QMC methods have gained broad popularity. Furthermore, the methods are easy to apply, and benefit from the increasing power of computers. In the cases above where X is defined as a sum, each element adds noise. But X_k is seldom independent of $X_j \, \forall \, k, j$. Here the symbiosis between option pricing and random number generation come to play, because in order to optimize the use of the quasi random number generators of the QMC methods, we need to identify the orthogonal vector space of noise in the problem. Actually the definition of $\mathcal{M}$, and the properties of the Brownian motion guarantee that such a space exists with full rank. In the articles we go through the necessary modeling aspects concerning the prices, and give some background on the QMC methods and the low-discrepancy sequences they are based upon. Especially the second article; "Valuation of Asian Basket Options with QMC Techniques and Singular Value Decomposition", has a section devoted to this. The third article of the thesis looks into questions concerning qualities of the orthogonal vector space of noise in the Asian basket option problem. We wanted to find and compare the lengths of the vectors, and investigate the bias introduced to the price by eliminating the set of shortest vectors. For many types of problems it turned out that the complexity could be reduced substantially without altering the accuracy of the prices. The increased awareness of these mechanisms leads to the search for methods to take advantage of it in other ways. In the fourth and fifth article we study and develop adaptive integration methods to encompass these characteristics of the multidimensional option pricing problems. The work of these articles depend on the formulation of the problem as an integral, and the methods developed can be used with success also for integration problems with no relation to finance. A considerable part of these articles deal with aspects concerning implementation of the methods. In order to achieve reliable results and to pinpoint pro and cons of the different techniques, we had to formulate comparable efficiency measures for the different methods. We performed simulations summing up to about 6000 CPU-hours on a cluster of personal computers with average CPU-speed of 1 GHz.

The last article of the thesis looks into the problem of hedging various option contracts in the commodity and energy markets. Hedging of option contracts amounts to neutralizing the risk inherent in the contract by trading in the underlying instruments or other basic instruments which have suitable characteristics. In the standard market $\mathcal{M}$ this is possible to achieve. It can, however, involve continuous trading, which is not possible in practice. The risk of holding a derivative can be expressed by the sensitivities of the price of the derivative contract with respect to the input parameters (initial price of the underlying assets, volatility of underlying, etc.). By comparing these sensitivities to the corresponding sensitivities of other alternative investments (including the underlying instruments of the derivative contract) one can construct a trading strategy that removes or reduces the risks of the derivative contract. This is

important for an issuer of such contracts since he often do not want to be exposed to the risks involved in the contract.

It is harder to formulate hedging problems as expected values and integrals than it is for pricing problems, and a substantial part of the last article is devoted to this. We have put particular focus on using the Malliavin approach developed in [FLL$^+$99] and [FLL$^+$01] to find tangible expressions for the sensitivities of various contracts in the commodity and energy market. Once such formulations are reached, the methods developed in the previous articles can be used to produce numerical solutions. In the last part of the article we compare different numerical methods used on different problem formulations. We show that the adaptive method developed and presented in the articles of chap. 5 and 6 have very good performance for these problems.

Chapter 2

Valuation of European Call Options on Multiple Underlying Assets by Using a Quasi-Monte Carlo Method. A Case with Baskets from Oslo Stock Exchange[1]

Author(s): Lars O. Dahl

Keywords: Basket options, Quasi-Monte Carlo methods, low-discrepancy sequences, multi dimensional integration

Abstract: The Black & Scholes formula will give a wrong price if used directly on basket options, and most often the Black & Scholes price underestimates the real price. Depending on the parameters of the underlying assets the error could be substantial. This paper briefly outlines the theoretical background for the above statement, and discusses the use of simulation methods to solve the problem. The focus will be on a quasi-Monte Carlo method which employ a low discrepancy Sobol sequence of numbers in $[0,1]^s$ - s being the number of assets in the basket. Some baskets consisting of the largest stocks on Oslo Stock Exchange are constructed, and the prices of these baskets calculated by the quasi-Monte Carlo method are compared to the corresponding Black & Scholes prices.

2.1 Introduction

When deducing the Black & Scholes formula one uses that the price process of the underlying asset follows a Geometric Brownian motion (GBM). If one assumes that each asset in a basket has a GBM price process, the basket itself can not have a GBM price process because the sum of GBM is not GBM. This is equivalent to the fact that the sum of log-normal distributed random variables is not a log-normal random

[1]Published in the AFIR 2000 proceedings. Full reference: [Dah00]

variable. This means that we can not, in theory, use the Black & Scholes formula to calculate the price of derivatives with several underlying assets.

It is known [KS98, Ch. 2.2] that the price of a European contingent claim can be expressed as an expected value of an expression in which the price processes of the assets are present. This expected value can be simulated by the use of Monte Carlo methods. The conventional Monte Carlo method, however, is rather slow, and we therefore apply quasi Monte Carlo methods (QMC) to the basket option pricing problem. This leads to the need for a rewriting of the original integral, which is over the domain $\mathbb{R}^N$, to an integral over the domain $[0,1]^N$.

Section 2.2 describes the market briefly and section 2.3 introduces the expressions for the value of the claim. In section 2.4 we give a short motivation for the use of QMC, and performs the rewriting of the integral which describes the value of the claim. In section 2.5 we present some simulation results on the convergence speed of the QMC compared to the standard Monte Carlo method. Section 2.6 gives the results on the pricing of some baskets of assets from Oslo Stock Exchange.

2.2 The Market

We operate in the context of a complete, standard financial market $\mathcal{M}$, with constant risk-free rate r and volatility matrix σ. The price processes of the assets in this market are governed by a set of stochastic differential equations (SDEs). There are $N+1$ assets in the market, one risk free asset and N risky assets. The model for the risky assets is the so called GBM. For a comprehensive survey of the assumptions and properties of the market see [KS98]. The solution to the SDEs is achieved by the development of a risk free measure P_0 and straightforward use of the Itô Formula. This leads to the following expressions for the price processes:

$$S_0(t) = e^{rt}, \text{ (because } S_0(0) = 1) \tag{2.1}$$

$$S_n(t) = S_n(0) \exp\left[(r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t + \sum_{d=1}^{N} \sigma_{nd} W_0^{(d)}(t)\right], \qquad n = 1, \ldots, N\ . \tag{2.2}$$

The volatility matrix σ is an $N \times N$ matrix such that $\sigma\sigma^T \equiv \sigma^2$ is the covariance matrix of the returns of the assets. The stochastic process $W_0(t)$ is an N dimensional Brownian motion under the risk free measure P_0. In the following analysis it is convenient to write the price process for the stocks like

$$S_n(u) = h_n(u - t, S(t), \sigma(W_0(u) - W_0(t))), \qquad 0 \le t \le u \le T\ , \tag{2.3}$$

where $h : [0,\infty) \times \mathbb{R}_+^N \times \mathbb{R}^N \to \mathbb{R}_+^N$ is the function defined by

$$h_n(t,p,y) \stackrel{\triangle}{=} p_n \exp\left[(r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t + y_n\right], \qquad n = 1, \ldots, N\ . \tag{2.4}$$

The process $W_0(t)$ is an essential part of the market $\mathcal{M}$. When we use constant coefficients as in our case, there is a constant $\theta \in \mathbb{R}^N$ called the market price of risk. θ is given by the relation $b - r\underline{1} = \sigma\theta$, where $b \in \mathbb{R}^N$ is the vector of drift coefficients for the assets of $\mathcal{M}$. By using the Girsanov theorem it can be shown that

$$W_0(t) = W(t) + \int_0^t \theta \, ds, \qquad \forall t \in [0, T] \tag{2.5}$$

is an N-dimensional Brownian motion under the risk free measure P_0 relative to the filtration $\{\mathcal{F}(t)\}$ of $W(t)$. The process $W(t)$ is the Brownian motion observed for the assets in the market under the measure P induced by the market. For a more general survey see [KS98, Ch. 1]. The Brownian motion W_0 has the property that $\Delta W_0(t) \sim \mathcal{N}_N(0, \sqrt{\Delta t})$, and this is exploited when we deduce explicit formulas and algorithms suited for the simulations in the next sections.

2.3 The Value of the Claim

A European contingent claim (ECC) is defined as a cumulative income process. Without going into the details and proofs which can be found in [KS98, Ch. 2.4], we state that the value at time t of European call options (ECO), which is the class of claims we are looking at, can be expressed as

$$V^{ECO}(t) = e^{-r(T-t)} E_0[\varphi(S(T))|\mathcal{F}(t)] \, . \tag{2.6}$$

The function $\varphi(S(T))$ is depending on the construction of the individual ECO. We are developing a pricing formula for an ECO on a portfolio or basket of assets where the different assets can be correlated. This contract is to be understood as the right (not obligation) to get the basket consisting of an amount $H_n(0)$ of each of the underlying assets at the expiration time T, by paying the strike price q_P. Thus the function $\varphi(\cdot)$ in this particular case is given by

$$\varphi(S(T)) = (\sum_{n=1}^{N} H_n(0) S_n(T) - q_P)^+, \qquad n = 1, \ldots, N \, , \tag{2.7}$$

and this leads to the price process of the particular ECO we are considering

$$V^{ECO}(t) = e^{-r(T-t)} E_0[(\sum_{n=1}^{N} H_n(0) S_n(T) - q_P)^+ | \mathcal{F}(t)] \tag{2.8}$$

$$= e^{-r(T-t)} E_0[(\sum_{n=1}^{N} H_n(0) h_n(T-t, S(t), \tag{2.9}$$

$$\sigma(W_0(T) - W_0(t))) - q_P)^+ | \mathcal{F}(t)] \, . \tag{2.10}$$

For simplicity we prefer to consider the value at $t = 0$, which is the value at the time you are buying the ECO. The expression then becomes

$$V^{ECO}(0) = e^{-rT} E_0[(\sum_{n=1}^{N} H_n(0) h_n(T, S(0), \sigma(W_0(T))) - q_P)^+] \, . \tag{2.11}$$

From elementary probability theory we have that this expected value can be expressed as an integral over $\mathbb{R}^N$, see for instance [Øks98, Ch. 2.1].

$$V^{ECO}(0) = e^{-rT} \int_{\mathbb{R}^N} (\sum_{n=1}^{N} H_n(0) h_n(T, S(0), \sigma z) - q_P)^+ \frac{\exp\left(-\|z\|^2/(2T)\right)}{(2\pi T)^{N/2}} \, dz \, . \tag{2.12}$$

2.4 Quasi-Monte Carlo Methods

The motivation for the development of quasi-Monte Carlo methods is that the conventional Monte Carlo method only converges at an order of $\mathcal{O}(1/\sqrt{K})$, where K is the number of simulations. The idea with quasi-Monte Carlo methods is to increase this rate. The literature on quasi-Monte Carlo methods yields an integration error that is proportional to the discrepancy of the used point sequence. The discrepancy is a measure of nonuniformity of this point sequence. An upper bound for the discrepancy is $\mathcal{O}((\log K)^s/K)$, where s is the dimension of the problem. See e.g. [MC98], [MC94] or [Nie92] on this subject. The error bound for the quasi-Monte Carlo methods is thus better than that of the conventional Monte Carlo method as $K \to \infty$, but it is evident that K needs to be extremely large for reasonable sized s in order for the benefit to appear. In practice however one experiences better performance than the given theoretical bounds when employing quasi-Monte Carlo methods, see e.g. [KW97].

The uniform distributed sequences are the basis of the quasi-Monte Carlo methods, and are objects of extensive research [Nie92], [Owe99], [Owe98]. A low discrepancy sequence which is rather simple to implement yet with good performance is the Sobol sequence. It was first presented in [Sob67], and an improved implementation of it was presented in [AS79]. Computational tests were carried out in [KW97] and indicates that the Sobol method works well, also for higher dimensional problems. In [Pas97] it is reported that they had success with the Sobol method also for very high dimensions, but other literature like [CMO97] report that the so called effective dimension d_s for the problem is actually lower than the real dimension s.

One problem with the Sobol sequence approach is that it is difficult to find the sequence for dimensions $s > 45$, but for the purpose of this paper in which $s < 13$ the Sobol method is well suited. For an extension of the dimension to include a larger portfolio, other techniques may be employed. See [Owe98] on this subject.

If we were to employ a standard Monte Carlo method for the evaluation of the expressions (2.11) or (2.12) we could simply use the fact that $W_0(T) \sim \mathcal{N}_N(0, \sqrt{T})$ in (2.11) and use some known algorithm like the Polar Marsaglia or the Marsaglia Bray

to simulate values from the standard normal distribution, see [Rip87, Ch. 3] for an overview of such methods. These methods are known as rejection methods, which means that some combinations of the uniform distributed variables used in the algorithms are rejected. When using quasi-Monte Carlo methods however, we need to ensure that we do not reject any of the uniform distributed numbers. This is because we have to maintain the low discrepancy characteristics of the uniform distributed sequence. To cope with this we have to rewrite the integral in (2.12) to be over the unit cube $[0,1]^N$ instead of $\mathbb{R}^N$. The integral is rewritten

$$V^{ECC}(0) = e^{-rT} \int_{\mathbb{R}^N} \varphi(h(T, S(0), \sigma z\sqrt{T})) \frac{\exp\left(-\|z\|^2/2\right)}{(2\pi)^{N/2}}\, dz \tag{2.13}$$

$$= e^{-rT} \int_{[0,1]^N} \varphi(h(T, S(0), \sigma\Psi^{-1}(y)\sqrt{T}))\, dy\ , \tag{2.14}$$

where $\Psi^{-1} : [0,1]^N \to \mathbb{R}^N$ is a vector of inverse cumulative normal distribution functions with mean 0 and variance 1: $\Psi^{-1}(y) = (\Psi_1^{-1}(y_1), \ldots, \Psi_N^{-1}(y_N))$. In fact we use that $\Psi_n^{-1}(y_n) = \mathrm{erf}^{-1}(2y_n - 1)$, where

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt\ , \tag{2.15}$$

in order to calculate the elements of $\Psi^{-1}(\cdot)$. We do the evaluation of the expression in (2.15) by a suitable rational approximation (the same as the one used in MATLAB from The MathWorks Inc.). Given this conversion we approximate the integral by using samples from the Sobol sequence as input to $\Psi^{-1}(\cdot)$, and calculate the mean value of the integrand in the domain of integration which is $[0,1]^N$. The value of the ECO-price in (2.12) is thus approximated by

$$V^{ECO}(0) \approx \frac{e^{-rT}}{K} \sum_{j=1}^{K} (\sum_{n=1}^{N} H_n(0) h_n(T, S(0), \sigma\Psi^{-1}(\theta_j)\sqrt{T}) - q_P)^+\ , \tag{2.16}$$

where $\theta_j \in [0,1]^N$ is the jth vector in the low discrepancy sequence used. As K is increased in (2.16) we get better approximations. In the next section we do some simulations to investigate how K should be chosen in order to get trustworthy approximations.

2.5 Convergence of the Simulations

In order to test the convergence speed of the quasi-Monte Carlo method compared to the standard Monte Carlo method we simulate the price of an ECO for different number of simulations for both the methods and compare the prices. The plot of the results in figure 2.1 indicates that the convergence rate is much faster for the quasi-Monte Carlo method than for the conventional method. We see that the faster convergence is obtained both for baskets with two and twelve assets. Whether this

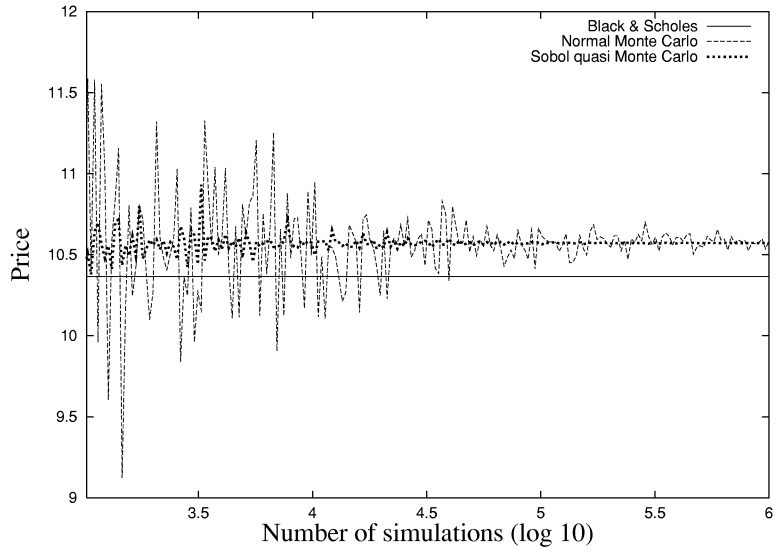
large difference in convergence rate is maintained as N is increased is not investigated, but other literature report that the quasi-Monte Carlo methods looses some of the advance as N grows larger. Note that we have used logarithmic scaling on the x-axis of the plots to specify number of simulations. The solid horizontal lines are the corresponding Black & Scholes prices for each of the baskets, and is included to illustrate the size of the simulation error compared to the difference between the simulated prices and the Black & Scholes price. We emphasize that the Monte Carlo and quasi-Monte Carlo methods converge to the correct price of the option given the assumptions on the market $\mathcal{M}$.

2.6 Results

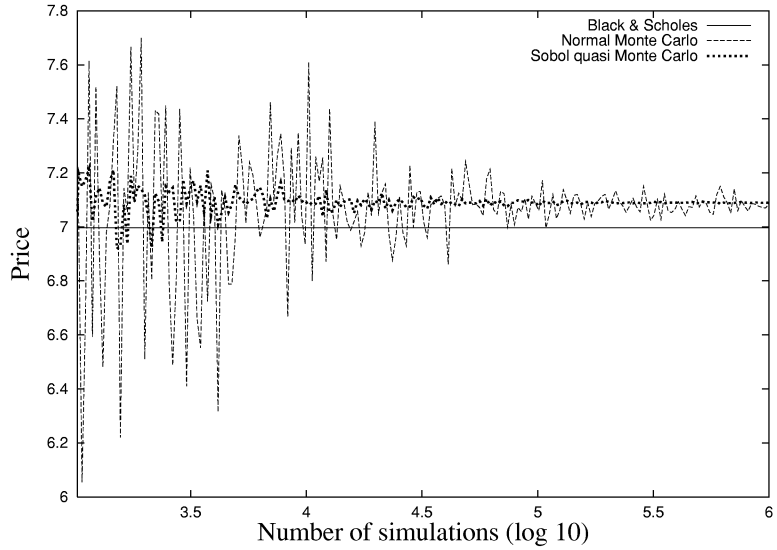
The price of the ECO depends on the risk free rate r, the volatility matrix σ and the strike price q_P of the basket. For special structures of σ the price calculated with the conventional Black & Scholes and the QMC method will differ significantly. See [Dah99]. We wanted to investigate the corresponding price difference for the correlation structure between the largest companies on OSE. We construct a portfolio consisting of assets from two to twelve of the largest companies on OSE. The sum of the market cap of the twelve companies constitute about half of the total market cap on OSE. We have constructed each of the baskets in such a way that their market value at $t = 0$ is 100. We set time to maturity to one year ($T = 1$) and vary the strike price from 50 to 150. The figure 2.2 gives the graphs of the findings by plotting the difference in percent between the Black & Scholes price and the QMC price as a function of the strike - stock price ratio for different baskets. We see that the differences grows as the call option goes from in the money to out of the money, but that the difference is not directly affected by the number of companies in the basket. The difference is more affected by the correlation structure between the assets. This can be read from the second plot in figure 2.2 in which the basket constructed with four securities give an error of up to 11 percent.

2.7 Conclusion

The use of Quasi Monte Carlo methods gives faster convergence than the use of conventional Monte Carlo methods for the problem we are studying. Furthermore the Black & Scholes formula will give a too low price for the ECO, and more sophisticated methods must be used. We have suggested a method involving the use of low discrepancy sequences of Sobol type resulting in a quasi-Monte Carlo numerical integration. For the actual baskets we are considering, the difference in price calculated with the Black & Scholes method and the quasi-Monte Carlo method is up to 11 percent when the European call option is out of the money.

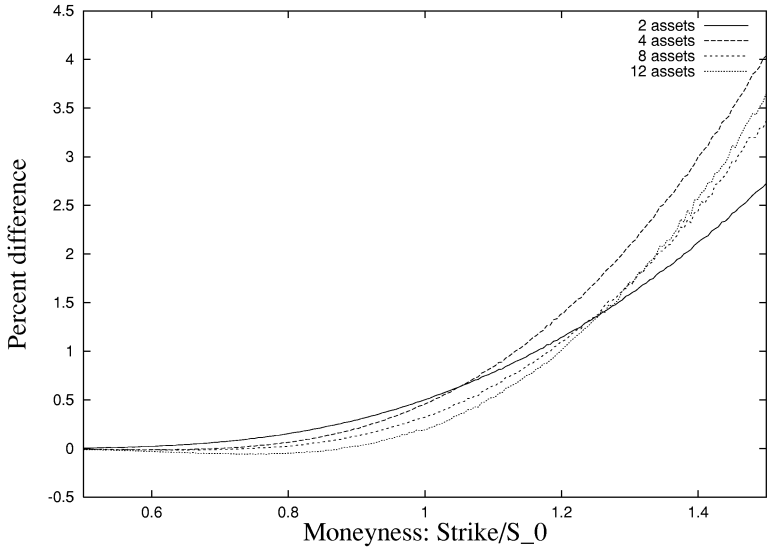


Basket consisting of two assets.

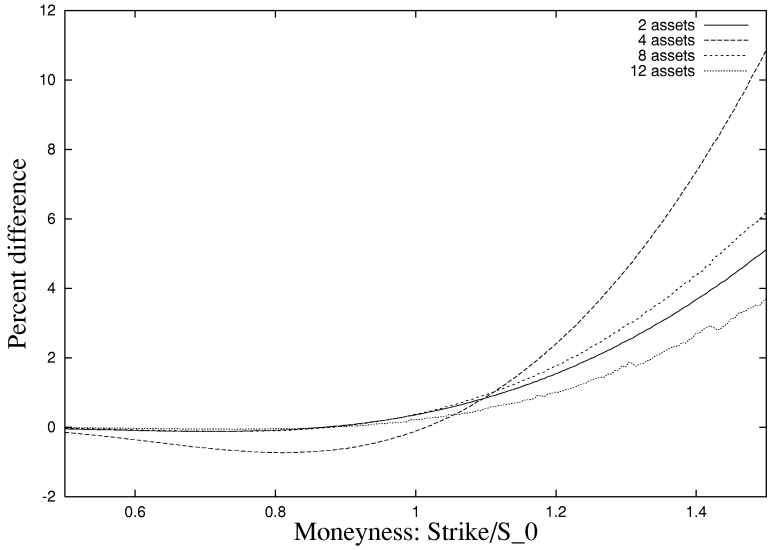


Basket consisting of twelve assets.

Figure 2.1: Convergence rate for conventional Monte Carlo and quasi-Monte Carlo simulation methods for an ECO on a basket of assets. We have set the initial price $S(0)$ of the basket to 100, and the strike price q_P to 125.



Assets are ordered descending by market cap; NHY, RCL, ORK, PGS, STB, etc.



Assets are ordered in a way that gives bigger difference between the Black & Scholes price and the QMC price; MED, NYA, PGS, NSG, TOM, etc.

Figure 2.2: Difference in percent between prices calculated with the Black & Scholes method and the quasi-Monte Carlo Method. We have used $5E5$ simulations for each point in the graph. The market value of the baskets at $t = 0$ is 100. We set time to maturity to one year ($T = 1$) and vary the strike price q_P from 50 to 150.

Acknowledgements

I want to thank my supervisor Fred Espen Benth, Department of Mathematics - University of Oslo, for very motivating and amending talks. Furthermore I want to thank my employer Storebrand Investments. The work could not have been done without their support.

Chapter 3

Valuation of Asian Basket Options with Quasi-Monte Carlo Techniques and Singular Value Decomposition[1]

Author(s): Lars O. Dahl[2], Fred E. Benth[3]

Keywords: Asian options, Asian basket options, quasi-Monte Carlo methods, Brownian Bridge, Singular Value Decomposition

Abstract: We propose pricing methods for European-style Asian arithmetic average basket options in a Black-Scholes framework based on a QMC method. The nature of QMC methods enables us to enhance the accuracy by decomposing the correlation structure of the noise in the problem using singular value decomposition. This leads to optimal utilization of the low discrepancy sequence, and gives several orders of magnitude enhanced performance over conventional QMC and standard MC methods.

3.1 Introduction

There are no closed form pricing formulas for the European-style Asian arithmetic average options (hereafter Asian options), neither the single asset option nor the basket option. Both problems must be solved by numerical solution methods, and are computer intensive tasks. The option price is given by an expected value, and the pricing is therefore an integration problem. In this paper we formulate the pricing problems explicitly as multi-dimensional integrals, which enables us to use quasi-Monte Carlo (QMC) methods to approximate their values. The main goal is the

[1]Submitted for publication February 2001 to Journal of Computational Finance. Full reference: [DB01]

[2]The first author is grateful for financial support from Storebrand Investments

[3]The second author is grateful to MaPhySto and the Norwegian Computing Centre for financial support. MaPhySto - Centre for Mathematical Physics and Stochastics, is funded by a grant from the Danish National Research Foundation.

pricing of the Asian basket option, but the single asset option is also discussed as an introduction to the basket case and for comparative studies between different path discretization schemes.

It is well known that Asian options and other path dependent options hold certain properties that can be exploited to increase the convergence rate when calculating their values with QMC methods. This is done by combining the QMC method with variance reduction techniques. Singular value decomposition (SVD) of the noise term in the problem is suggested. We propose to use a combination of SV-decomposition of the covariance matrix of the Brownian paths, *and* a representation of the volatility matrix for the assets in the basket by using SV-decomposition of the covariance matrix of asset returns. We demonstrate that this approach leads to considerably better convergence properties than in the case where conventional discretization is used on the Brownian paths, and Cholesky-decomposition is used to create the volatility matrix from the covariance matrix of the returns. We have also included the Brownian Bridge method for the discretization of the Brownian paths in order to compare with the SVD approach for the single asset case, and we show that the proposed SVD method work better than the Brownian Bridge method as well.

For Asian basket options the number of dimensions in the problem may grow significantly as it is the product of the number of assets in the basket and the number of time discretization points. We show how to avoid performing an SV-decomposition of the full problem by using the direct matrix product to combine the decompositions of each of the covariance structures (the path and the basket) into a matrix that describes the full system. This reengineering of the problem enables us to exploit the QMC method better. This is because low discrepancy sequences, which is the basic part of the QMC approach, often have the property that some elements of the sample vector have better discrepancy characteristics than other, and by reengineering the problem we adapt it to this property.

Section 3.2 describes the financial market in which we will do our analysis, and section 3.3 gives some background information on multi-dimensional Brownian motion which is useful in our context. In section 3.4 we give a short motivation for the use of QMC methods and give arguments for why the methods we use in our approach do work. Section 3.5 introduces the general expression for the value of the claim. In section 3.6 we point out the properties of the asset price process we have to use in order to formulate the Asian basket option problem as an integral. The section focuses on the *conventional* QMC approach to outline the general concept and to show how we construct the algorithms we compare the SVD methods with. We describe the Brownian Bridge technique in section 3.7, and the SVD in section 3.8. In section 3.9 we present numerical results comparing the different methods, and finally we conclude in section 3.10.

3.2 The Market

We operate in the context of a complete, standard financial market $\mathcal{M}$, with constant risk-free rate r and volatility matrix σ. The price processes of the assets in this market are governed by a set of stochastic differential equations (SDEs). There are $N+1$ assets in the market, one risk free asset and N risky assets. The model for the risky assets is the so called geometric Brownian motion. For a comprehensive survey of the assumptions and properties of the market see [KS98]. The solution to the SDEs is achieved by the development of a risk free measure P_0 and straightforward use of the Itô Formula. This leads to the following expressions for the price processes:

$$S_0(t) = e^{rt} \tag{3.1}$$

$$S_n(t) = S_n(0) \exp\big((r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t + \sum_{d=1}^{N} \sigma_{nd} W_0^{(d)}(t)\big), \quad n = 1, \ldots, N\,. \tag{3.2}$$

The volatility matrix $\sigma \in \mathbb{R}^{N\times N}$ is such that $\sigma\sigma^T \equiv \sigma^2$ is the covariance matrix of the returns of the assets. Note that the relation defining σ is not unique, and this is a key feature for the enhancement of the numerical methods. The stochastic process $W_0(t)$ is an N dimensional Brownian motion under the risk free measure P_0. In the following analysis it is convenient to write the price process for the risky assets like

$$S_n(u) = h_n(u-t, S(t), \sigma(W_0(u) - W_0(t))), \quad 0 \le t \le u \le T\,, \tag{3.3}$$

where $h : [0,\infty) \times \mathbb{R}_+^N \times \mathbb{R}^N \to \mathbb{R}_+^N$ is the function defined by

$$h_n(t,p,w) \equiv p_n \exp\big((r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t + w_n\big), \quad n = 1, \ldots, N\,. \tag{3.4}$$

The process $W_0(t)$ is an essential part of the market $\mathcal{M}$. $\theta \in \mathbb{R}^N$ is given by the relation $b - r\underline{1} = \sigma\theta$, where $b \in \mathbb{R}^N$ is the vector of drift coefficients for the assets of $\mathcal{M}$. θ is called the market price of risk. By using the Girsanov theorem it can be shown that

$$W_0(t) = W(t) + \int_0^t \theta\, ds, \quad \forall t \in [0,T] \tag{3.5}$$

is an N-dimensional Brownian motion under the risk free measure P_0 relative to the filtration $\{\mathcal{F}(t)\}$ of $W(t)$. The process $W(t)$ is the Brownian motion observed for the assets in the market under the market induced probability measure P. For a more comprehensive survey of these aspects, see [KS98, Ch. 1].

3.3 Useful Properties of the Brownian Motion

In this section we present some well-known properties of Gaussian processes which are useful for our approach. The Brownian motion $W_0(t) \in \mathbb{R}^N$ is a Gaussian process,

which means that the random variable $Z = (W_0(t_0), \ldots, W_0(t_K) \in \mathbb{R}^{N(K+1)}$ has a normal distribution. The covariance matrix of Z is given by

$$C_Z^2 = \begin{bmatrix} t_0 I & t_0 I & \ldots & t_0 I \\ t_0 I & t_1 I & \ldots & t_1 I \\ \vdots & \vdots & \ddots & \vdots \\ t_0 I & t_1 I & \ldots & t_K I \end{bmatrix}, \tag{3.6}$$

where I is the $N \times N$ identity matrix, see e.g. [Øks98]. Let C^2 be the covariance matrix of Z for the case where $N = 1$. If we construct the process $X = (\sigma W_0(t_0), \ldots, \sigma W_0(t_K)) \in \mathbb{R}^{N(K+1)}$, where $W_0(t_k) \in \mathbb{R}^N$ and $\sigma^2 \in \mathbb{R}^{N \times N}$ is positive semidefinite, the covariance matrix $C_X^2 \in \mathbb{R}^{N(K+1) \times N(K+1)}$ of the process X is given by

$$C_X^2 = \begin{bmatrix} t_0\sigma^2 & t_0\sigma^2 & \ldots & t_0\sigma^2 \\ t_0\sigma^2 & t_1\sigma^2 & \ldots & t_1\sigma^2 \\ \vdots & \vdots & \ddots & \\ t_0\sigma^2 & t_1\sigma^2 & \ldots & t_K\sigma^2 \end{bmatrix}. \tag{3.7}$$

In the notation of the direct product of matrices we can write $C_X^2 = C^2 \otimes \sigma^2$. To see that C_X^2 is given by (3.7), consider the process $\hat{Z}_0 = \sigma W_0(t_0) \in \mathbb{R}^N$: We know that $\hat{Z}_0 \sim \mathcal{N}_N(0, t_0 \sigma\sigma^T)$. By using this for each of the N processes $\hat{Z}_K$ contained in $Z \in R^{N(K+1)}$ we get to the expression (3.7). It can be shown that the eigenvalues of C_X^2 are found directly from the eigenvalues of the matrix σ^2 and the eigenvalues of C^2 by the relation $\Lambda_{C_X^2} = \Lambda_{C^2} \otimes \Lambda_{\sigma^2} \in \mathbb{R}_+^{N(K+1) \times N(K+1)}$. See [Lam91] and [Lan69] for a full treatment of the direct matrix product. The eigenvalue property enables us to find an ordering of the total set of eigenvalues. In section 3.4 we will look into the QMC method, and reveal the advantage of knowing the eigenvalues and their ordering in finance problems.

3.4 Quasi-Monte Carlo Methods

In this section we give a brief survey of the QMC-technique. The goal is to evaluate an integral of the form

$$\int_{[0,1]^s} f(y)\, dy \approx \frac{1}{L+1} \sum_{l=0}^{L} f(y^l)\,. \tag{3.8}$$

The sequence $\{y^l\}$ of vectors $y^l = (y_1^l, \ldots, y_s^l) \in [0,1]^s,\ l = 0, \ldots, L$ used for the approximation can be generated by a systematic combinatorial approach, giving a conventional grid-based numerical integration algorithm. The problem with this approach is that the complexity grows exponentially with the dimension s, leading to practically useless algorithms for $s > 5$. If, on the other hand, the sequence is created

by letting each vector y^l be independent uniform random variables in $[0,1]^s$, we get the conventional MC approach. This enables us to increase the number of evaluation points in a smooth manner, filling the domain of integration gradually. QMC methods keep this nice feature of the MC approach, but uses number sequences which are not random. These number sequences are constructed with the intention of filling the domain of integration as evenly as possible, resulting in methods where the approximation of the integral can be obtained with even fewer integrand evaluations than in the conventional MC Methods. Conventional MC method only converges at an order of $\mathcal{O}(1/\sqrt{L})$, but QMC methods are able to increase this rate. The employment of QMC methods are closely linked to the formulation of the problem as a multi dimensional integral, and strict control of the use of the number sequence $\{y^l\}$ in the construction of the distributions used. QMC methods are based on the approach of removing randomness from the generation of sampling sequences. The idea is to look for fixed sequences that perform better than random sequences in a well defined sense. The measurement of this behavior is not trivial in general, and these uniform distributed sequences are objects of extensive research, see e.g. [NX98], [Owe99], [Owe98]. The discrepancy of the sequence is used to measure how well distributed the samples are, see e.g. [JBT96], [PT97], [Pas97] for more details. Discrepancy is defined as follows

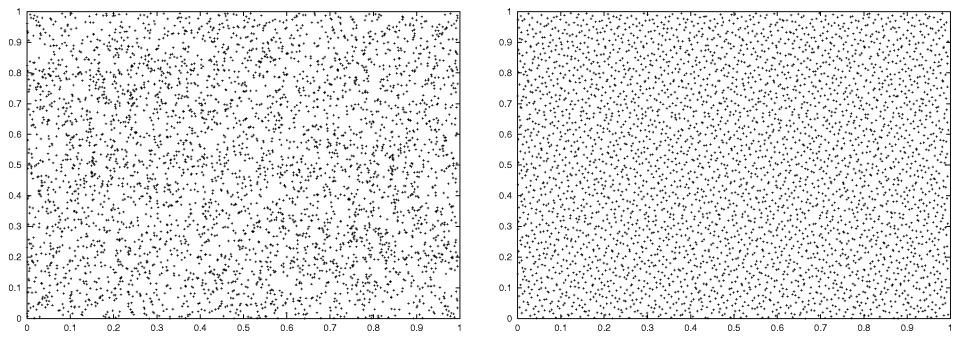


Figure 3.1: An illustration of the ability to fill the domain $[0,1]^2$ uniformly by the use of the conventional pseudo-random numbers (left), and the Halton leaped low discrepancy sequence (right).

Definition 3.4.1. *Let $\mathcal{B}$ be a family of shapes which are subsets of $[0,1]^s$. Given a sequence $\{y^l\}$ of sample points. The discrepancy of $\{y^l\}$ with respect to $\mathcal{B}$ is*

$$D_L(B,\{y^l\}) = \sup_{B\in\mathcal{B}} \left| \frac{\#\{y^l \in B\}}{L} - \lambda(B) \right| , \tag{3.9}$$

where $\lambda(B)$ is the volume of B and y^l, $l = 1,\ldots,L$ are elements of the sequence $\{y^l\}$.

The definition says that we are finding the maximum difference between the fraction of points inside one of the shapes and the volume of the shape. When the set of shapes

$\mathcal{B}$ is the set of boxes with a corner at the origin, this is called the star discrepancy $D_L^*(\{y^l\})$.

The Koksma-Hlawka theorem gives an upper bound for the error in QMC methods, see e.g. [KN74], [Nie78], [Nie87]. It is given as the product of the variance of the function that is integrated and the discrepancy.

Theorem 3.4.1 (Koksma-Hlawka theorem). *Let $\hat{I}$ be the estimator of the integral I over the domain Ω. Then an upper bound for the error is*

$$\left| I - \hat{I} \right| \leq V(f) D_L^*(\{y^l\}) \,, \tag{3.10}$$

where $V(f)$ is the total variation of the function f over Ω in the sense of Hardy and Krause.

See [MC94], [MC95] or [Nie92] for extensive surveys on this subject, and [Woz91] and [MC94] for the alternative approach involving the so called Woźniakowski's identity. The theorem says that QMC methods yields an integration error that is proportional to the discrepancy of the point sequence used. In s dimensions, it is possible to find sequences $\{y^l\}$ such that

$$D_L^*(\{y^l\}) = \mathcal{O}(\frac{(\log L)^{s-1}}{L}) \,. \tag{3.11}$$

The given error bound is thus better than that of the conventional MC method as the number of simulations L grows to infinity, but it is evident that L needs to be very large for reasonable sized s (the dimension of the problem) in order for the benefit to appear. In practice, however, the theoretical bounds for QMC methods are conservative, (see [KW97] for a general survey or [Dah00] for an example involving European basket options). The measure of discrepancy is mainly used as a criteria for constructing good low discrepancy sequences rather than finding error bounds of integration rules.

In conventional MC methods it is common to use some known algorithm like the Polar-Marsaglia or the Marsaglia-Bray to simulate values from the standard normal distribution, (see [Rip87, Ch. 3] for an overview of such methods). These methods are known as rejection methods, which means that some combinations of the uniform distributed variables used in the algorithms are rejected. When using QMC methods however, we need to ensure that we do not reject any of the uniform distributed numbers. This is because we have to maintain the low discrepancy characteristics of the uniform distributed sequence. Standard rejection methods can therefore not be used directly in QMC methods. Smoothed rejection methods developed in [CM95] and [Caf98] can however be used with QMC methods, but are harder to implement.

A low discrepancy sequence which is rather simple to implement is the Halton sequence. It was first presented in [Hal60]. In this paper we are going to use an extension of the Halton sequence denoted the Halton leaped sequence. It was presented in [KW97], together with good leap values. We have used the leap value 31

for the numerical experiments in this paper. We are not going to dwell on the choice of sequence here, but mention that other types exist; We have the Sobol sequence first introduced in [Sob67], with an improved implementation presented in [AS79]. We also have the Faure sequence [Fau82] and the Van der Corput sequence [Pag92]. These named sequences have been shown to belong to a generalized family of (t, s)-sequences for which [Nie92] is a comprehensive reference. An other family of sequences are produced by so called lattice methods for which [SJ94] is the definitive reference. The research in this field is rich and extensive, both in comparative studies between sequences (and their extensions) for a variety of dimension and integrands (e.g. in [KW97], [MC94] and [LL00b]), in connection with variance reduction techniques (e.g. [Owe98] and [LL00a]), and in solution of concrete problems in finance (e.g. [Pas97], [PT95], [MC95], [LL98], [BG96] and [PAG98]).

In many finance problems the so called effective dimension d_s for the problem is actually lower than the real dimension s. (See e.g [CMO97] and [CM96] for finance problems, and [SW98] for a general discussion). This property is present both for path dependent option problems and multi-asset options. The problem of pricing Asian basket options has a mix of both, and some of the challenge is to pinpoint the effective dimensions of the problem. The concept of effective dimension is closely linked to the so called ANOVA decomposition. (See e.g. [Hoe48], [ES81] or [Owe98, Owe99]). It is used to find a representation of the integrand as a sum $\mathcal{F}_{\mathcal{A}}$ of orthogonal functions. If each of these orthogonal functions depends only on a distinct subset of the coordinates, the integrand can be written as a sum of integrals of functions of lower dimension, and the complexity of the problem has been reduced with regards to the integral dimension. Even if we are not able to reduce the dimension of the original integrand by this approach, we can find that some of the orthogonal functions in $\mathcal{F}_{\mathcal{A}}$, say $\mathcal{F}_{\mathcal{A}^C}$, have little effect on the value of the integral. Then if $\mathcal{F}_{\mathcal{A}} - \mathcal{F}_{\mathcal{A}^C}$ have dimension d_s, and d_s is lower than the dimension of original integral, but estimates the true value within acceptable limits (ε), we say that the original problem has effective dimension d_s. In finance problems we can often achieve a representation involving matrices describing the connection between the different variables linearly as arguments to the exponential function, i.e $f(x) = \exp\left(\sum_i c_i x_i\right),\ c_i \leq c_{i+1},\ \forall\, i < s$. This is the case in the problem we are studying, and we achieve this by using the SV-decomposition. If we truncate the sum $\sum_i c_i x_i$ at some point d, where $c_d < \hat{\varepsilon} \ll c_0$ we will get a good approximation of the original problem by evaluating the integral over this lower dimensional integrand. The effective dimension found by the SV-decomposition approach and the effective dimension from the ANOVA approach are compatible, since we can write the exponential function as a sum of polynomials through a series expansion. This means that QMC-methods are well suited for integrals of functions with low effective dimension. Especially if we can find the dimensions having effect, and are able to employ a low discrepancy sequence $\{y^l\}$ for which we know the elements y_i^l having the lowest discrepancy. A numerical test to find the effective dimension of the single asset Asian option problem is performed in subsection 3.9.1.

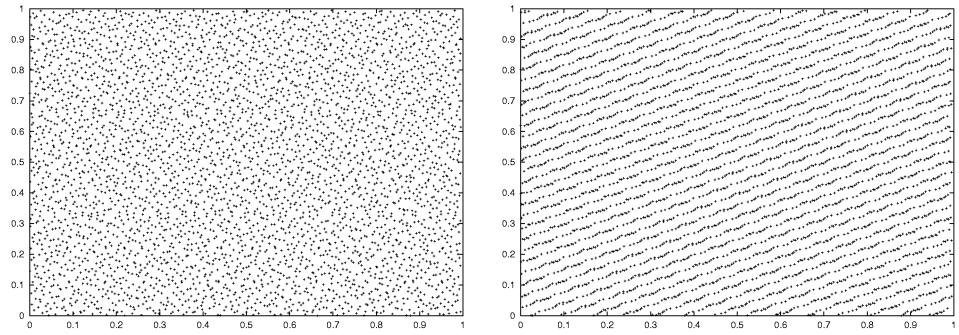


Figure 3.2: The initial elements have better abilities to fill the domain than the one further back in the low discrepancy vector. Projections onto $[0,1]^2$ of Halton leaped. $(0, 1$ and $42, 43)$.

3.5 The Value of the Claim

An Asian option is actually a special type of European contingent claim, which is defined as a cumulative income process. Without going into details (which can be found in [KS98, Ch. 2.4]), we state that the value at time t of an Asian option is

$$V(t) = e^{-r(T-t)}\mathrm{E}_0[\varphi(\Upsilon(\mathcal{T}))|\mathcal{F}(t)] \,, \tag{3.12}$$

where $\varphi(\cdot)$ is a Borel measurable function. This function can for example be given by

$$\varphi(\Upsilon(\mathcal{T})) = (\Upsilon(\mathcal{T}) - q)^+ \,, \tag{3.13}$$

resulting in the European-style Asian option. A variety of different option contracts fits into this framework by choosing different functions $\Upsilon(\mathcal{T}) : \mathbb{R}^{K+1} \to \mathbb{R}$, $\varphi(\cdot) \in \mathbb{R}$ and $\mathcal{T} \in \mathbb{R}^{K+1}$. All of them, however, are European contingent claims.

The theoretical definition of the Asian option is

$$\text{(For single asset option)} \quad \Upsilon_1(t_0, T) = \int_{t_0}^{T} S(u)\, \mu(du) \tag{3.14}$$

$$\text{(For basket option)} \quad \Upsilon_N(t_0, T) = \int_{t_0}^{T} \sum_{n=1}^{N} S_n(u)\, \mu(du) \,, \tag{3.15}$$

for some Borel measure μ on $[t_0, T]$. Our formulation is rather general, but the measure μ is usually given by $\mu(du) = (T - t_0)^{-1}du$. Other candidates can, however, easily be handled by our setup. If we for example choose $\mu(du) = \delta_T(du)$, where δ_T is the Dirac point mass at T, we get a European call option. Other examples are given in, e.g. [RS95]. Note also that contracts often are specified with $t_0 = 0$, but in our discussion we only need $t_0 < T$.

In real applications the integrals in the formulations of $\Upsilon_1(t_0, T)$ and $\Upsilon_N(t_0, T)$ must be approximated, and often these approximations are specified in the contracts by specifying the number of sampling points along the path. For this purpose let $\mu(du) = (T - t_0)^{-1}du$ and $\mathcal{T} = (t_0, t_1, \ldots, t_K)$, $t_K = T$, and specify the number $K + 1$ of sampling points. The length of the intervals $t_k - t_{k-1}$ need not be equal, but we shall assume this here for simplicity. Approximations of (3.14) and (3.15) can then be carried out by using the expressions

$$\text{(For single asset option)} \quad \hat{\Upsilon}_1(\mathcal{T}) = \frac{1}{K+1}\sum_{k=0}^{K} S(t_k) \tag{3.16}$$

$$\text{(For basket option)} \quad \hat{\Upsilon}_N(\mathcal{T}) = \frac{1}{K+1}\sum_{k=0}^{K}\sum_{n=1}^{N} S_n(t_k)\,. \tag{3.17}$$

Note that by choosing $N = 1$ in (3.17), the basket option is a single asset option. These types of approximations are also necessary in order to apply the MC and QMC framework. We will briefly discuss the convergence of (3.14) to (3.16) and (3.15) to (3.17) in section 3.9. For simplicity we prefer to consider $V(t)$ at $t = 0$, which is the value at the time the option is bought. Note that this does *not* imply $t_0 = 0$. The expression we are going to use throughout the rest of the paper for the value of the Asian basket option thus becomes

$$V_N(0) = e^{-rT}\mathrm{E}_0[\varphi(\hat{\Upsilon}_N(\mathcal{T}))]\,. \tag{3.18}$$

In section 3.6 we will show the *conventional* way of how the price processes for the risky assets can be expressed as models where the the noise is implemented as independent stochastic processes. This independence formulation enables us to express the expected value in expression (3.18) as an integral over $\mathbb{R}^{N(K+1)}$, see for instance [Øks98, Ch. 2.1]. Furthermore we will show that in the particular case of integrals involving distribution functions, we can convert the integral from $\mathbb{R}^{N(K+1)}$ to $[0,1]^{N(K+1)}$.

3.6 The Conventional Way of Formulating E_0 as an Integral

In the standard formulation of the asset price processes given by (3.3) and (3.4), the Brownian motion used as the driving noise has a built-in correlation structure. In order to formulate the Asian basket option pricing problem as an integral we have to model the price processes in terms of independent stochastic variables. This can be done in several ways, resulting in algorithms with different properties when used together with QMC methods. In this section we outline the *conventional* way of doing this. This approach exploits that the increments $\Delta W_0(t_k)$ are independent,

and results in the following expressions:

$$\sum_{k=0}^{K} S_n(t_k) = \sum_{k=0}^{K} h_n(t_k, S(0), \sigma W_0(t_k)) \tag{3.19}$$

$$= S_n(0) h_n(t_0, 1, \bar{\epsilon}_0)\big(1 + h_n(\Delta t_1, 1, \bar{\epsilon}_1)\big(1 + h_n(\Delta t_2, 1, \bar{\epsilon}_2)\big(1 + \cdots\big)\big)\big) \tag{3.20}$$

$$= S_n(t_0)\delta^n_{K,1}(\bar{\epsilon}_1, \ldots, \bar{\epsilon}_K)\ , \tag{3.21}$$

where $\Delta t_k = t_k - t_{k-1}$, $\bar{\epsilon}_k \sim \mathcal{N}_N(0, \Delta t_k \sigma^2)$, $0 < k \leq K$ are independent of $\mathcal{F}_{t_{k-1}}$, while $\bar{\epsilon}_0 \sim \mathcal{N}_N(0, t_0\sigma^2)$. The notation including the δ function is achieved by letting

$$\delta^n_{K,K}(\bar{\epsilon}_K) = 1 + h_n(\Delta t_K, 1, \bar{\epsilon}_K) \tag{3.22}$$

$$\delta^n_{K,k}(\bar{\epsilon}_k, \ldots, \bar{\epsilon}_K) = 1 + h_n(\Delta t_k, 1, \bar{\epsilon}_k)\delta^n_{K,k+1}(\bar{\epsilon}_{k+1}, \ldots, \bar{\epsilon}_K), \tag{3.23}$$

$$k = K-1, \ldots, 1\ . \tag{3.24}$$

Note furthermore that the most common method of finding σ is by use of the Cholesky decomposition of σ^2, even if other types like the SV-decomposition exists, and are more suited in finance problems. In section 3.9, where we compare different approaches, we have used the Cholesky decomposition when calculating values of the Asian basket option by the *conventional* approach, while we have used the SV-decomposition in the SVD approach described in section 3.8.

With the expression (3.21) we can write the value of the Asian basket option as

$$V_N(0) = e^{-rT}\mathrm{E}_0[\varphi\big(\hat{\Upsilon}_N(\mathcal{T})\big)]$$

$$= e^{-rT}\int_{\mathbb{R}^{N(K+1)}} \varphi\Big(\frac{1}{K+1}\sum_{n=1}^{N} S_n(t_0)\delta^n_{K,1}(x)\Big)\psi(x)\,dx \tag{3.25}$$

$$= e^{-rT}\int_{[0,1]^{N(K+1)}} \varphi\Big(\frac{1}{K+1}\sum_{n=1}^{N} S_n(t_0)\delta^n_{K,1}(\Psi^{-1}(y))\Big)\,dy\ , \tag{3.26}$$

where $\psi : \mathbb{R}^{N(K+1)} \to \mathbb{R}^{N(K+1)}$ is the density of an $N(K+1)$ - dimensional centered Gaussian random variable with covariance matrix equal to the identity, and $\Psi^{-1} : [0,1]^{N(K+1)} \to \mathbb{R}^{N(K+1)}$ is a vector of inverse cumulative normal distribution functions with mean 0 and variance 1: $\Psi^{-1}(y) = (\Psi_1^{-1}(y_1), \ldots, \Psi^{-1}_{N(K+1)}(y_{N(K+1)}))$. In the following it is convenient to introduce the notation $\hat{\varphi}(Y) \equiv \varphi(\frac{Y}{K+1})$ in order to simplify the expressions. The transformation of the integral over $\mathbb{R}^{N(K+1)}$ in (3.25) to $[0,1]^{N(K+1)}$ in (3.26) is due to the mapping performed by the function $\Psi^{-1}(\cdot)$, and is valid for any inverse cumulative distribution function. In order to find the inverse of $\Psi(\cdot)$ we use that $\Psi_n^{-1}(y_n) = \mathrm{erf}^{-1}(2y_n - 1)$. We do the evaluation by a rational approximation suggested in [Mor95]. Other types of methods for calculating $\Psi^{-1}(\cdot)$ could be employed, but caution must be taken when used together with QMC. (See comments in section 3.4 on this.) It is evident at this stage that we can approximate the option price by making use of a convenient set of points $\{y\} \subseteq [0,1]^{N(K+1)}$. In the *conventional* QMC approach this is done by using samples y^l from the Halton

leaped sequence as input to $\Psi^{-1}(\cdot)$, and calculating the mean value of the integrand in the domain of integration. The value for an Asian basket option for this approach is thus approximated by

$$V_N(0) \approx \frac{e^{-rT}}{L} \sum_{l=1}^{L} \hat{\varphi}\Big(\sum_{n=1}^{N} S_n(t_0)\delta_{K,1}^n(\Psi^{-1}(y^l))\Big) , \tag{3.27}$$

where $y^l \in [0,1]^{N(K+1)}$ is the l'th vector in the low discrepancy sequence. As L is increased in (3.27) we get better approximations. The approach in this section, however, does not take into account any special structures of the integrand. This must be done in order to exploit that some of the elements of the low discrepancy vector in the QMC method are more evenly distributed than others. In section 3.4 we have discussed methods of reengineering the integrand to reduce the complexity and to improve convergence rates. The next sections are devoted to the specification of two types of decompositions, the Brownian Bridge approach and the SVD approach.

3.7 The Brownian Bridge Approach

The Brownian Bridge approach is presented for the pricing of Asian options on a single underlying asset only. This is because the coupling between the asset dependency and the time dependency in the basket case is hard to define for the Brownian Bridge approach. We will show in the next section that this coupling can be handled rather easily in the SVD approach. The inclusion of the Brownian Bridge approach is solely for reason of comparison. In section 3.9 we present results showing that the SVD approach turns out to be better among the two.

When evaluating path dependent options we have to simulate the path – one way or the other. In the formulation leading to the conventional approach, involving the $\delta(\cdot)$ function, this is done by using the independent increment property of Brownian motion. We can, however, achieve a representation of the path by using another approach – the so called Brownian Bridge approach:

1. Before entering the simulation loop: Choose T, and the number of equal time steps $K = 2^p$. Set $\Delta t = \frac{T}{K}$ and $t_k = k\Delta t,\ k = 1, \ldots, K$.

2. Inside the simulation loop: Generate Gaussian independent variables ϵ_j^l, for each of the L turns in the simulation loop, distributed according to

$$\epsilon_j^l = \sqrt{\hat{t}_j}\Psi^{-1}(y_j^l) \sim \mathcal{N}(0, \hat{t}_j),\ j = 0, \ldots, K-1 \tag{3.28}$$

$$\text{where } \hat{t}_0 = T \text{ and } \hat{t}_j = \frac{T}{2^{2+\lfloor \log_2 j \rfloor}},\ j = 1, \ldots, K-1 \tag{3.29}$$

3. The Wiener path $w_l(t)$ is sampled at each t_k as

$$w_l(t_0) = 0 \tag{3.30}$$
$$w_l(t_K) = \sigma\epsilon_0^l \qquad \left(\sim \mathcal{N}(0, \sigma^2 T)\right) \tag{3.31}$$
$$w_l(t_{K/2}) = \frac{1}{2}(w_l(t_0) + w_l(t_K)) + \sigma\epsilon_1^l \qquad \left(\sim \mathcal{N}(0, \sigma^2 \frac{T}{2})\right) \tag{3.32}$$
$$w_l(t_{K/4}) = \frac{1}{2}(w_l(t_0) + w_l(t_{K/2})) + \sigma\epsilon_2^l \qquad \left(\sim \mathcal{N}(0, \sigma^2 \frac{T}{4})\right) \tag{3.33}$$
$$w_l(t_{3K/4}) = \frac{1}{2}(w_l(t_{K/2}) + w_l(t_K)) + \sigma\epsilon_3^l \qquad \left(\sim \mathcal{N}(0, \sigma^2 \frac{T}{4})\right) \tag{3.34}$$
$$\vdots$$

4. The price path is then calculated by using

$$S^l(t_k) = S(t_0) \exp\left((r - \frac{1}{2}\sigma^2)t_k + w_l(t_k)\right) \tag{3.35}$$

5. The price of the option is found by averaging:

$$V_1(0) \approx \frac{\mathrm{e}^{-rT}}{L} \sum_{l=1}^{L} \hat{\varphi}\Big(\sum_{k=0}^{K} S^l(t_k)\Big) \tag{3.36}$$

Although the total variance in this representation is the same as in the standard discretization, much more of the variance is contained in the first few steps of the Brownian Bridge formula. This reduces the effective dimension of the simulation and increases the effect of the low discrepancy sequence used. It turns out that this decomposition is not optimal, and the optimal decomposition is given by the SVD method presented next.

3.8 The SVD Approach

A random variable $Y \sim \mathcal{N}_N(0, \Sigma\Sigma^T)$ can be written $Y = \Sigma X$ where $X \sim \mathcal{N}_N(0, I)$, and I is the $N \times N$ identity matrix. In section 3.3 we discussed properties of multidimensional Brownian motion, and concluded in expression (3.7) with the covariance matrix of the process. Given a covariance matrix $\Sigma^2 = \Sigma\Sigma^T$ there are several alternatives of finding the matrix Σ. The Cholesky decomposition produces a Σ matrix which is triangular, while the Σ matrix from the SV-decomposition can be written as $E\sqrt{\Lambda}$, where E contains the eigenvectors of Σ^2 and Λ is a diagonal matrix with the corresponding eigenvalues in decreasing order on the diagonal. We will use the SV-decomposition *both* for the Wiener path along the time dimension *and* to find a volatility matrix σ used in the modeling of the price process of the underlying assets. The properties of Brownian motion enables us to perform two separate SV-decompositions instead of one large: One for the covariance matrix

$C^2 \in \mathbb{R}^{(K+1)\times(K+1)}$ given in section 3.3, describing the path-dependencies, and one for the covariance matrix $\sigma^2 \in \mathbb{R}^{N\times N}$ for the underlying assets. The eigenvalues of C^2 and σ^2 can then be combined by the direct matrix product into an ordering $O : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ of the total set of eigenvalues for the full problem to give us $\lambda_{O(\cdot,\cdot)}$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N(K+1)}$. This method enables us to allocate specific elements y_i^l from the low discrepancy vector y^l to the different orthogonal noise generators (represented by the eigenvalues λ) of the full problem. If y^l is a vector from the Halton leaped sequence, the noise term with the biggest eigenvalue is mapped to y_1^l the next biggest to y_2^l and so forth. In order for this approach to be effective, the discrepancy of y_1^l should be lower than the discrepancy of y_2^l etc. This is a property of many low discrepancy sequences, and the Halton leaped sequence seems to have this characteristic. (See e.g. [KW97]).

The principles of the SVD method for the basket option problem given as a list of tasks are as follows:

1. Before entering the simulation loop: Find $\sigma \in \mathbb{R}^{N\times N}$ by performing an SV-decomposition of the covariance matrix σ^2, and $C \in \mathbb{R}^{(K+1)\times(K+1)}$ by an SV-decomposition of the covariance matrix C^2. Find a relation $O(\cdot,\cdot)$ between the time discretization point k, the asset n and the ordering of the eigenvalues λ by sorting the output from the direct matrix product in reverse order:

$$\lambda = \lambda_{\sigma^2} \otimes \lambda_{C^2} \tag{3.37}$$

2. Inside the simulation loop: Create a low discrepancy vector $y^l \in [0,1]^{N(K+1)}$ for each of the L turns in the simulation loop.

3. Find the corresponding inverse cumulative normal values

$$\epsilon^l_{O(n,k)} = \Psi^{-1}(y^l_{n,k})\ n = 1,\ldots,N, \quad k = 0,\ldots,K \tag{3.38}$$

4. Find the asset price for each of the N assets in each of the $K+1$ points along the time line. This is done by

$$S_n^l(t_k) = S_n(t_0)\exp\big((r - \frac{1}{2}\sum_{d=1}^{N}\sigma_{nd}^2)t_k + \sum_{d=1}^{N}\sigma_{nd}\sum_{j=0}^{K}C_{kj}\epsilon^l_{O(d,j)}\big) \tag{3.39}$$

5. Find the average of all the asset prices computed in (3.39) and evaluate $\hat{\varphi}(\hat{\Upsilon}_N)$.

6. The option price is approximated by performing the described loop L times, averaging the L results of $\hat{\varphi}(\cdot)$, and discounting by e^{-rT}.

The full expression for the approximate value of the Asian basket option by the SVD

approach is therefore given by

$$V_N(0) \approx \frac{\mathrm{e}^{-rT}}{L} \sum_{l=1}^{L} \hat{\varphi}\Big(\sum_{n=1}^{N} \sum_{k=0}^{K} S_n(t_0) \exp\big((r - \frac{1}{2} \sum_{d=1}^{N} \sigma_{nd}^2) t_k + \sum_{d=1}^{N} \sigma_{nd} \sum_{j=0}^{K} C_{kj} \epsilon^l_{O(d,j)} \big) \Big) \tag{3.40}$$

where σ_{nd} and C_{kj} are elements of the matrices resulting from the SV-decompositions. The matrices σ and C together with the function $O(\cdot,\cdot)$ are the essential parts of this approach.

3.9 Numerical Results

We present the numerical results from simulations of prices of both single asset Asian call options and Asian basket call options. The simulations for single asset options is performed to show the difference between the convergence of the conventional recursive approach, the Brownian Bridge and the SVD method, while the basket option simulations only compare the conventional recursive approach with the SVD approach. We will also look briefly into the convergence of the sum in (3.16) to the value of the integral (3.14) as the number of evaluation points K along the path increases. For the numerical calculations we specify $\varphi(\cdot)$ to be the payoff function of a call option. This specification also effects $\hat{\varphi}(\cdot)$, giving:

$$\varphi(Y) \equiv \big(Y - q\big)^+ \tag{3.41}$$

$$\hat{\varphi}(Y) \equiv \big(\frac{Y}{K+1} - q\big)^+ = \frac{1}{K+1}\big(Y - \hat{q}\big)^+ , \tag{3.42}$$

where $\hat{q} \equiv q(K+1)$. In addition we let $t_0 = 0$ in the numerical examples.

3.9.1 Convergence of $\hat{\Upsilon}$

The integrand in the expression for $\hat{\Upsilon}_1(T)$ must be approximated when calculating the value of the theoretical expression for the Asian option, and therefore it is interesting to investigate how fast the convergence of the sum in (3.16) to the value of the integral (3.14) is achieved. We do this numerically by looking at the expressions:

$$\mathrm{e}^{-rT}\mathrm{E}_0[\varphi\big(\int_0^T S(u)T^{-1}\,du\big)] = \lim_{K\to\infty} \mathrm{e}^{-rT}\mathrm{E}_0[\hat{\varphi}\big(\sum_{k=0}^{K} S(t_k)\big)] \tag{3.43}$$

$$= \lim_{\substack{K\to\infty \\ L\to\infty}} \frac{\mathrm{e}^{-rT}}{L} \sum_{l=1}^{L} \hat{\varphi}\big(\sum_{k=0}^{K} S^l(t_k)\big) \ . \tag{3.44}$$

When performing the calculation we use the SVD method since this have the lowest variance, i.e we can keep L smaller than for the other methods. Figure 3.3 gives the level of accuracy for a given K compared to the true value of the integral. We see that the convergence is rather fast, and that for $K = 10$, we are well within 0.1% of the value of the integral. This indicates that an effective dimension of $d_s = 10$ for this problem is a conservative estimate. Notice that we do not know the real value of the integral, and therefore these convergence results are purely indicative. Note that $\sigma = 0.3$ for the underlying asset in this case, while the other parameters are equivalent to those given in subsection 3.9.2.

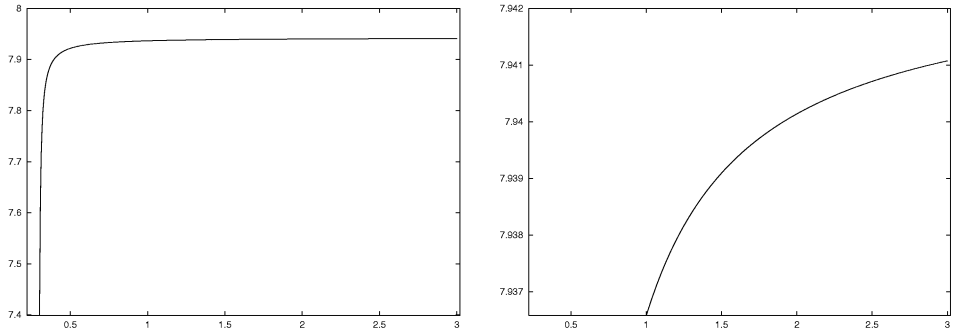


Figure 3.3: The convergence of the option price as $K \to \infty$

3.9.2 Single Asset Asian Options

We will simulate prices for options where time to maturity is one year $T = 1.0$, initial price $S(0) = 100$ and strike $q = 100$. Furthermore the risk free rate in the market is $r = 0.05$, the volatility is constant $\sigma = 0.4$, and the assets pay no dividends. With this setup we calculate the price for $K = 2^p$, $p \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, by using L number of simulations in the range $L \in [10^3, 10^5]$. The standard MC method (STA), the conventional QMC method by the use of the recursive $\delta(\cdot)$ function (REC), the Brownian Bridge method (BB) and the SVD method (SVD) are compared by use of a set of graphs. The variance of the resulting series are also given as a measure of convergence speed. Notice that each new point in each of the graphs are calculated by using non-overlapping sequences of low discrepancy vectors. The results is shown as graphs where the price of the option is on the Y-axis and the X-axis show the number of simulations on $\log_{10}$ scale. See Figure 3.4 and Figure 3.5.

It is important to quantify the performance of the different approaches, and we have done this by simply calculating the variances of the graphs. Although the prices have been calculated by a deterministic approach, and we therefore can not truly trust statistical measures on the behavior, we believe that the used measure will give some insights. The result of this measurement for an Asian option on a single underlying asset is given for $K = 2^p$, $p = 1, \ldots, 9$ in Table 3.1. We have done two sets of

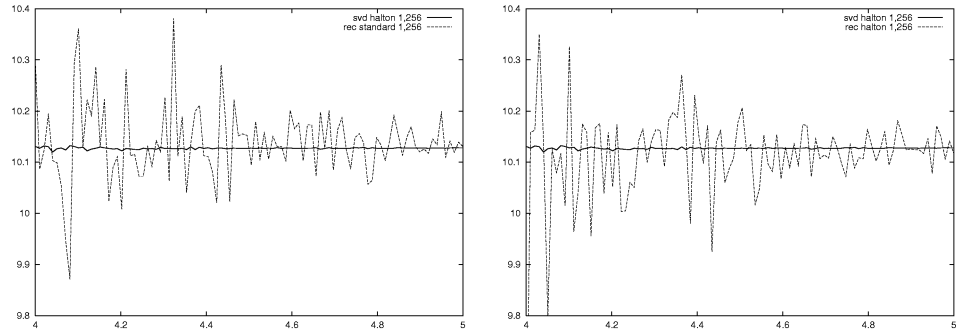


Figure 3.4: Comparing *SVD* method with *STA* and *REC*. For these simulations, $N = 1$ and $K = 256$, and *REC* does not perform better than *STA* for dimensions this high.

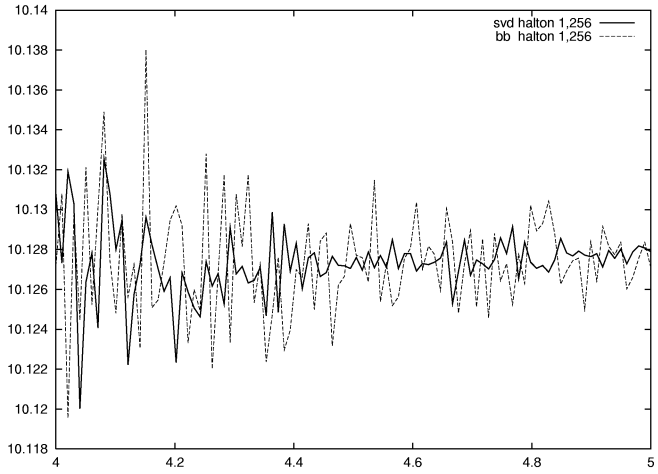


Figure 3.5: Comparing *SVD* with *BB*. We see that *SVD* is slightly better than *BB*. $N = 1$ and $K = 256$.

simulations, one for $L \in [10^3, 10^4]$ and one for $L \in [10^4, 10^5]$.

When we plot these results (for $p \leq 9$) we get the rather illustrative picture in Figure 3.6, showing that the SVD method has close to constant variance as the number of dimensions 2^p are increased, while the variance of the conventional method increases linearly (note that the scale on both axis are logarithmic). The standard (non QMC-approach) has constant high variance, and the Brownian Bridge approach increases a bit in the start but stabilizes on a lower level than the SVD method.

Table 3.1: The variances are calculated for each series of 100 prices calculated by simulating for L in the range $[10^3, 10^4]$ and $[10^4, 10^5]$. The mapping onto these ranges are logarithmic, i.e there are fewer samples from the end of the interval than the beginning. $N = 1$ and $K = 2^p$.

L	p	STANDARD	REC	BB	SVD
$10^3 \to 10^4$	1	0.0750028	0.000164428	0.000234812	0.000349644
$10^3 \to 10^4$	2	0.0880283	0.000484648	0.000313312	0.000217205
$10^3 \to 10^4$	3	0.0943216	0.00216916	0.00030832	0.000221485
$10^3 \to 10^4$	4	0.074528	0.00749465	0.000311825	0.000220451
$10^3 \to 10^4$	5	0.0789701	0.021403	0.000303524	0.000221139
$10^3 \to 10^4$	6	0.0750914	0.01854	0.000307054	0.000221403
$10^3 \to 10^4$	7	0.0805887	0.0351491	0.000305817	0.000221582
$10^3 \to 10^4$	8	0.0490915	0.258421	0.000309338	0.00022173
$10^3 \to 10^4$	9	0.053857	0.407136	0.00031079	0.000221813

L	p	STANDARD	REC	BB	SVD
$10^4 \to 10^5$	1	0.00611862	2.32071e-006	2.16537e-006	3.09931e-006
$10^4 \to 10^5$	2	0.00542252	6.30376e-006	3.75607e-006	2.78211e-006
$10^4 \to 10^5$	3	0.00637498	3.73341e-005	5.16935e-006	2.79055e-006
$10^4 \to 10^5$	4	0.00827774	0.000298189	6.74235e-006	2.77861e-006
$10^4 \to 10^5$	5	0.00709202	0.000459984	7.33869e-006	2.78717e-006
$10^4 \to 10^5$	6	0.00766036	0.000888891	7.40184e-006	2.80059e-006
$10^4 \to 10^5$	7	0.00605674	0.00134237	7.23032e-006	2.8036e-006
$10^4 \to 10^5$	8	0.00528101	0.00532122	7.38271e-006	2.80411e-006
$10^4 \to 10^5$	9	0.0059849	0.0312871	7.45808e-006	2.80269e-006

3.9.3 Basket Asian Options

When we calculate the value of the basket option, an additional element concerning the N assets in the basket comes into consideration. In section 3.3 we described how to find the eigenvalues of the full system, and in section 3.8 we showed how to utilize this to optimize the use of the low discrepancy sequence. In this section we will use a setup of the simulation similar to the one used for the single asset option, but in addition we will let the number of assets vary: $N \in \{2, 4, 8, 16, 32, 64\}$.

The results are given in Table 3.2 and illustrated in Figure 3.8. In the illustrations we have kept the number of sampling points K constant and increased the number of assets N in the basket. The different methods are labeled *REC* for the conventional QMC method, *SVD1* for the full SVD method including an ordering of the total noise in the problem by the use of the $O(\cdot, \cdot)$ function, and *SVD2* for an SVD method where we have decomposed both time and asset dimensions, but not combined them into

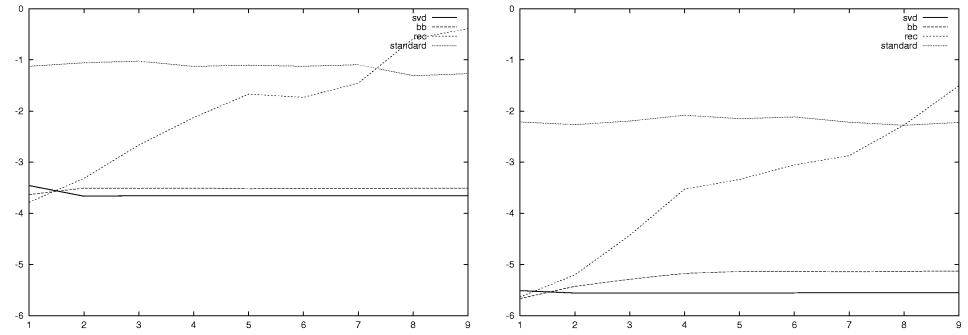


Figure 3.6: The plots have p on the X-axis, and $\log_{10}$ of the variance on the Y-axis. There are 9 estimated variance values for each approach, one for each p. Each series of prices contains 100 values, and are created by simulating prices for L in the range $[10^3, 10^4]$ (left) and $[10^4, 10^5]$ (right). The mapping onto these ranges are logarithmic, i.e there are fewer samples from the end of the interval than the beginning. $N = 1$ and $K = 2^p$.

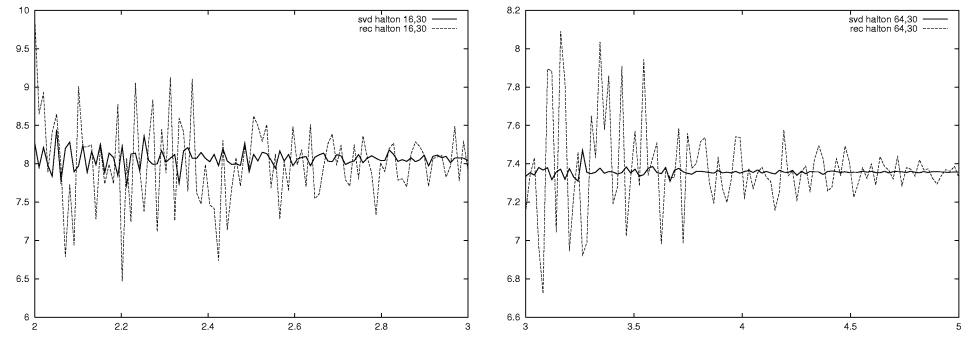


Figure 3.7: Comparison of the methods for different baskets. N and K given in the labels of the plots.

an overall ordering. In *SVD2* the N first elements of the low discrepancy vector are used for the noise in the problem stemming from the time discretization point giving the biggest contribution, and these N elements are used in an ordering according to the contributions from the different assets. This should theoretically give the *SVD1*-method best performance, but for the example we have tested, this conclusion can not be drawn. In the numerical studies of the Asian basket option, we have estimated the matrix σ^2 by using asset return time-series from the Oslo Stock Exchange in Norway to get a realistic case.

Table 3.2: The variance of the option price as the number of assets $N = 2^i$ in the basket is increased. Here $K = 2^5 = 32$.

L	i	REC	SVD2	SVD1
$10^3 \to 10^4$	1	0.00727686	0.000156036	0.000141992
$10^3 \to 10^4$	2	0.0108797	0.000231383	0.000234575
$10^3 \to 10^4$	3	0.00988478	0.000276212	0.00019724
$10^3 \to 10^4$	4	0.0240513	0.000312554	0.000224188
$10^3 \to 10^4$	5	0.0374931	0.000709355	0.00073982
$10^3 \to 10^4$	6	0.0486269	0.00028011	0.000391925

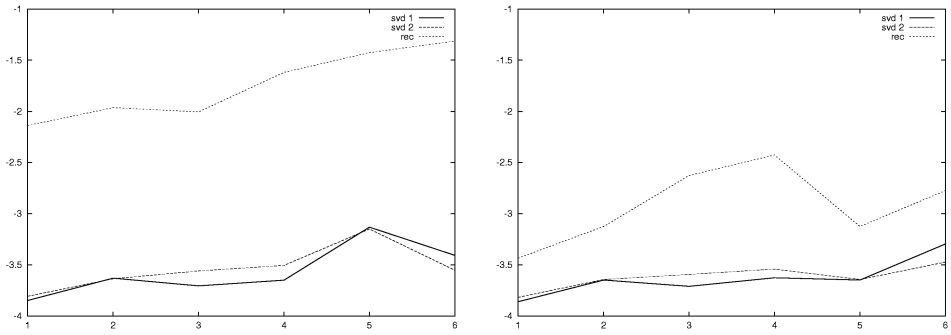


Figure 3.8: The variance of the option price as the number of assets in the basket is increased. $N = 2^i$, i on the X-axis. Time discretization is $K = 2^5 = 32$ (left) and $K = 2^2 = 4$ (right).

3.10 Conclusions

The use of QMC methods gives faster convergence than conventional MC Methods for both single asset and basket Asian options. By using the low discrepancy sequence more effectively, the examples we have simulated show a large performance gain compared to the conventional QMC method. Furthermore we get better results when using the SVD approach than the Brownian Bridge approach for single asset Asian options. The benefit of the SVD approach increases as the number of sampling points in the time interval increase due to the fact that the conventional QMC method becomes less effective, while the SVD method maintains its efficiency. The conventional QMC method is actually outperformed even by the conventional MC method for very high dimensions ($K \geq 256$), while the QMC method based on SVD remains very good also for high dimensions. The problem clearly has low effective dimension, but while we can estimate this to about 10 for the single asset Asian option, the effective dimension of the basket Asian option will depend on the covariance structure of the assets in the basket, and can therefore vary among different baskets.

The numerical tests show that as we increase the number of assets in the basket, the difference between the conventional QMC method and the SVD method is seemingly constant, or slightly increasing as N becomes large.

Chapter 4

Fast Evaluation of the Asian Basket Option by Singular Value Decomposition[1]

Author(s): Lars O. Dahl, Fred E. Benth

Keywords: Effective dimension, quasi-Monte Carlo methods, low-discrepancy sequences, multi dimensional integration

Abstract: We investigate the use of singular value decomposition of the noise term in the Asian basket option problem. By performing this decomposition the problem can be formulated as an integral. We find a criterium for deciding the effective dimension of the integrand in the framework of the singular value decomposition. The resulting integration problem is calculated by a suited quasi Monte Carlo method. The simulation results show that the proposed criterium works well, and that the computing time can be reduced significantly compared to the full problem.

4.1 Introduction

It is well known that many finance problems hold certain properties that can be exploited to increase the convergence rate when calculating their values with quasi Monte Carlo (QMC) methods. This is done by combining the QMC method with variance reduction techniques. Singular value decomposition (SVD) of the noise term in the problem is suggested. In the pricing of the European-style Asian arithmetic average basket option (hereafter Asian basket option), the correlation structure of the noise is a mix of the correlation structure of the Brownian paths and the correlation structure of the assets in the basket. We have in [DB01] developed a method to decompose the full problem into orthogonal factors. For the single asset Asian option the correlation structure is static, and the effective dimension only depends on the chosen time discretization. This problem is often referred to in the QMC literature,

[1] Published in the MCQMC2000 proceedings. Full reference: [DB02]

see e.g. [PAG98]. The Asian basket option, however, has a more complex, non-static, correlation structure – depending also on the basket. The effective dimension will therefore vary among different baskets. The goal of this paper is to show how the decomposition of the full Asian basket option problem is performed, and to quantify the effective dimension for the single asset Asian option and for different Asian basket options. We discuss the link of our approach to the notion of the ANOVA decomposition discussed in [Owe98] and [Owe99].

The outline of the article is as follows: We give some required properties of the Brownian motion in sec. 4.2. In sec. 4.3 we formulate the Asian basket option pricing problem as a multi-dimensional integration problem. In sec. 4.4 we implement the SVD solution into this framework. We develop a criterium for finding the effective dimension of the integrand, and formulate the pricing problem in accordance with the reduced dimension as an approximate solution in sec. 4.5. Numerical results illustrating convergence and accuracy are presented in sec. 4.6, and we conclude in sec. 4.7.

4.2 Useful Properties of the Brownian Motion

In this section we present some well-known properties of Gaussian processes required to formulate the Asian basket option problem as an integral in the SVD framework. The Brownian motion $W_0(t) \in \mathbb{R}^N$ is a Gaussian process, which means that the random variable $Z = (W_0(t_0), \ldots, W_0(t_K)) \in \mathbb{R}^{N(K+1)}$ has a normal distribution. The covariance matrix of Z is given by

$$C_Z^2 = \begin{bmatrix} t_0 I & t_0 I & \ldots & t_0 I \\ t_0 I & t_1 I & \ldots & t_1 I \\ \vdots & \vdots & \ddots & \vdots \\ t_0 I & t_1 I & \ldots & t_K I \end{bmatrix}, \tag{4.1}$$

where I is the $N \times N$ identity matrix, see e.g. [Øks98]. Let C^2 be the covariance matrix of Z for the case where $N = 1$. If we construct the process $X = (\sigma W_0(t_0), \ldots, \sigma W_0(t_K)) \in \mathbb{R}^{N(K+1)}$, where $W_0(t_k) \in \mathbb{R}^N$ and $\sigma^2 \in \mathbb{R}^{N \times N}$ is positive semi-definite, the covariance matrix $C_X^2 \in \mathbb{R}^{N(K+1) \times N(K+1)}$ of the process X is given by

$$C_X^2 = \begin{bmatrix} t_0\sigma^2 & t_0\sigma^2 & \ldots & t_0\sigma^2 \\ t_0\sigma^2 & t_1\sigma^2 & \ldots & t_1\sigma^2 \\ \vdots & \vdots & \ddots & \\ t_0\sigma^2 & t_1\sigma^2 & \ldots & t_K\sigma^2 \end{bmatrix}. \tag{4.2}$$

In the notation of the direct product of matrices we can write $C_X^2 = C^2 \otimes \sigma^2$. The eigenvalues of C_X^2 can be found directly from the eigenvalues of the matrix σ^2 and the eigenvalues of C^2 by the relation $\Lambda_{C_X^2} = \Lambda_{C^2} \otimes \Lambda_{\sigma^2} \in \mathbb{R}_+^{N(K+1) \times N(K+1)}$. See [Lam91]

and [Lan69] for a full treatment of the direct matrix product. The eigenvalue property enables us to find an ordering of the total set of eigenvalues.

4.3 The Integral Formulation of the Asian Basket Option

We operate in the context of a complete, standard financial market with constant risk-free rate r and volatility matrix σ. The price processes in the market are modeled by:

$$S_0(t) = e^{rt} \tag{4.3}$$

$$S_n(t) = S_n(0) \exp\left((r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t + \sum_{d=1}^{N} \sigma_{nd} W_0^{(d)}(t)\right), \quad n = 1, \ldots, N\ . \tag{4.4}$$

The volatility matrix $\sigma \in \mathbb{R}^{N\times N}$ is such that $\sigma\sigma^T \equiv \sigma^2$ is the covariance matrix of the returns of the assets. Note that the relation defining σ is not unique, and this is a key feature for the enhancement of the numerical methods. The stochastic process $W_0(t)$ is an N dimensional Brownian motion under the risk free measure P_0.

An Asian option is actually a special type of European contingent claim, which is defined as a cumulative income process. Without going into details (which can be found in [KS98]), we state that the value at time t of an Asian option is

$$V(t) = e^{-r(T-t)} \mathrm{E}_0[\varphi(\Upsilon(\mathcal{T}))|\mathcal{F}(t)]\ , \tag{4.5}$$

where $\varphi(\cdot)$ is a Borel measurable function. In our study we will use the function $\varphi(Y) = (Y - q)^+$, resulting in the European-style Asian option. Furthermore we are interested in $V(t)$ at $t = 0$, i.e. at the time the option is bought. In the theoretical definition of the Asian option the function $\Upsilon(\cdot)$ is given by

$$\Upsilon_N(t_0, T) = \int_{t_0}^{T} \sum_{n=1}^{N} S_n(u)\, \mu(du)\ , \tag{4.6}$$

for some Borel measure μ on $[t_0, T]$. We will use the measure $\mu(du) = (T - t_0)^{-1}du$, and specify $t_0 = 0$. For a discussion of a more general setup, see e.g. [RS95]. The integral must be discretized in order to do calculations with QMC. Note that this discretization often is specified in the contract by specifying the number of sampling points along the path. For this purpose let $\mathcal{T} = (t_0, t_1, \ldots, t_K)$, $t_K = T$, and specify the number $K + 1$ of sampling points. The length of the intervals $t_k - t_{k-1}$ need not be equal, but we shall assume this here for simplicity. The approximations of (4.6) can then be carried out by using the expression

$$\hat{\Upsilon}_N(\mathcal{T}) = \frac{1}{K+1} \sum_{k=0}^{K} \sum_{n=1}^{N} S_n(t_k)\ . \tag{4.7}$$

Note that by choosing $N = 1$ in (4.7), the basket option is a single asset option. The convergence of of (4.6) to (4.7) is discussed in [DB01].

The expression for the Asian basket option problem thus becomes:

$$V_N(0) = e^{-rT}\mathrm{E}_0[\varphi(\hat{\Upsilon}_N(\mathcal{T}))] \,. \tag{4.8}$$

In order to replace the expected value E_0 in (4.8) with an integral over $[0,1]^{N(K+1)}$, we need to formulate $\varphi(\hat{\Upsilon}_N(\mathcal{T}))$ in terms of independent stochastic processes. For path dependent options, this can be done in many ways, resulting in different methods (see [DB01]). Here we outline the basic principle of how to obtain an integral representation. We introduce the notation $S_n(t_k, \varepsilon_{nk})$ to be able to see the *independent* stochastic processes explicitly. The value of the claim can then be written

$$V_N(0) = e^{-rT} \int_{\mathbb{R}^{N(K+1)}} \varphi\Big(\frac{1}{K+1}\sum_{k=0}^{K}\sum_{n=1}^{N} S_n(t_k, x)\Big)\psi(x)\,dx \tag{4.9}$$

$$= e^{-rT} \int_{[0,1]^{N(K+1)}} \varphi\Big(\frac{1}{K+1}\sum_{k=0}^{K}\sum_{n=1}^{N} S_n(t_k, \Psi^{-1}(y))\Big)\,dy \tag{4.10}$$

$$\approx \frac{e^{-rT}}{L}\sum_{l=1}^{L} \varphi\Big(\frac{1}{K+1}\sum_{k=0}^{K}\sum_{n=1}^{N} S_n^l(t_k, \Psi^{-1}(y^l))\Big) \,, \tag{4.11}$$

where $\psi : \mathbb{R}^{N(K+1)} \to \mathbb{R}^{N(K+1)}$ is the density of an $N(K+1)$ - dimensional centered Gaussian random variable with covariance matrix equal to the identity, and $\Psi^{-1} : [0,1]^{N(K+1)} \to \mathbb{R}^{N(K+1)}$ is a vector of inverse cumulative normal distribution functions with mean 0 and variance 1: $\Psi^{-1}(y) = (\Psi_1^{-1}(y_1), \ldots, \Psi_{N(K+1)}^{-1}(y_{N(K+1)}))$. The vector $y^l \in [0,1]^{N(K+1)}$ is the l'th sample from the low discrepancy sequence $\{y^l\}$ of the QMC method. We do the evaluation of $\Psi^{-1}(\cdot)$ by a rational approximation suggested in [Mor95]. The specification of (4.11) will be performed in sec. 4.4.

4.4 The Singular Value Decomposition

A random variable $Y \sim \mathcal{N}_N(0, \Sigma\Sigma^T)$ can be written $Y = \Sigma X$ where $X \sim \mathcal{N}_N(0, I)$, and I is the $N \times N$ identity matrix. In our problem, and other finance problems, the covariance matrix is often given or observable. With a known covariance matrix $\Sigma^2 = \Sigma\Sigma^T$, there are several alternative ways of finding the matrix Σ. Since Σ^2 is a positive semi-definite matrix, the SVD produces Σ such that $\Sigma = E\sqrt{\Lambda}$, where E contains the eigenvectors of Σ^2 and Λ is a diagonal matrix with the corresponding eigenvalues in decreasing order on the diagonal. We will use the SVD *both* for the Wiener path along the time dimension *and* to find a volatility matrix σ used in the modeling of the price process of the underlying assets. The total collection of noise in the Asian basket option problem can by the use of the properties of Brownian motion be produced by two separate SVDs instead of one large: One for the covariance matrix $C^2 \in \mathbb{R}^{(K+1)\times(K+1)}$ given in sec. 4.2, describing the path-dependencies, and one for

the covariance matrix $\sigma^2 \in \mathbb{R}^{N\times N}$ for the returns of the underlying assets. The eigenvalues of C^2 and σ^2 can then be combined by the direct matrix product into an ordering $O : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ of the total set of eigenvalues for the full problem to give us $\lambda_{O(\cdot,\cdot)}$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N(K+1)}$. This method enables us to allocate specific elements y_i^l from the low discrepancy vector y^l to the different orthogonal noise generators (represented by the eigenvalues λ) of the full problem. The noise term with the biggest eigenvalue is mapped to y_1^l the next biggest to y_2^l and so forth. In order for this approach to be effective, the equidistribution properties of y_1^l should be better than that of y_2^l etc. This is a property of many low discrepancy sequences, see e.g. [KW97].

The principles of the SVD method for the basket option problem given as a list of tasks are as follows:

1. Before entering the simulation loop: Find $\sigma \in \mathbb{R}^{N\times N}$ by performing an SVD of the covariance matrix σ^2, and $C \in \mathbb{R}^{(K+1)\times(K+1)}$ by an SVD of the covariance matrix C^2. Find a relation $O(\cdot,\cdot)$ between the time discretization point k, the asset n and the ordering of the eigenvalues λ by sorting the output from the direct matrix product in reverse order:
$$\lambda = \lambda_{\sigma^2} \otimes \lambda_{C^2} \ . \tag{4.12}$$
2. Inside the simulation loop: Create a low discrepancy vector $y^l \in [0,1]^{N(K+1)}$ for each of the L turns in the simulation loop.
3. Find the corresponding inverse cumulative normal values
$$\varepsilon^l_{O(n,k)} = \Psi^{-1}(y^l_{nk}) \ n = 1,\ldots,N, \quad k = 0,\ldots,K \ . \tag{4.13}$$
4. Find the asset price for each of the N assets in each of the $K+1$ points along the time line. This is done by
$$S^l_n(t_k) = S_n(t_0) \exp\big((r - \frac{1}{2}\sum_{d=1}^{N}\sigma^2_{nd})t_k + \sum_{d=1}^{N}\sigma_{nd}\sum_{j=0}^{K} C_{kj}\varepsilon^l_{O(d,j)}\big) \ . \tag{4.14}$$
5. Find the average of all the asset prices computed in (4.14) and evaluate $\hat{\varphi}(\hat{\Upsilon}_N)$.
6. The option price is approximated by performing the described loop L times, averaging the L results of $\hat{\varphi}(\cdot)$, and discounting by e^{-rT}.

The full expression for the approximate value of the Asian basket option by the SVD approach is therefore given by

$$\begin{aligned} V_N(0) \approx & \frac{e^{-rT}}{L}\sum_{l=1}^{L}\hat{\varphi}\Big(\sum_{n=1}^{N}\sum_{k=0}^{K} S_n(t_0) \\ & \exp\big((r - \frac{1}{2}\sum_{d=1}^{N}\sigma^2_{nd})t_k + \sum_{d=1}^{N}\sigma_{nd}\sum_{j=0}^{K} C_{kj}\varepsilon^l_{O(d,j)}\big)\Big) \ , \end{aligned} \tag{4.15}$$

where σ_{nd} and C_{kj} are elements of the matrices resulting from the SVDs. The matrices σ and C together with the function $O(\cdot,\cdot)$ are the essential parts of this approach.

To show the effectivity of the SVD method compared to a standard QMC approach in which Cholesky factorization is used and no ordering of noise terms is performed, we have priced a basket option with the two methods and plotted the result in Fig. 4.1. A way of quantifying the difference is to measure the variance of the graphs as the number of simulations is increased. According to this measure the SVD method has a variance of about 1 percent of the standard method for the given configuration. I.e. the standard method need 10^2 times more simulations to achieve the same accuracy. See e.g. [DB01] for other configurations and more details. The low discrepancy

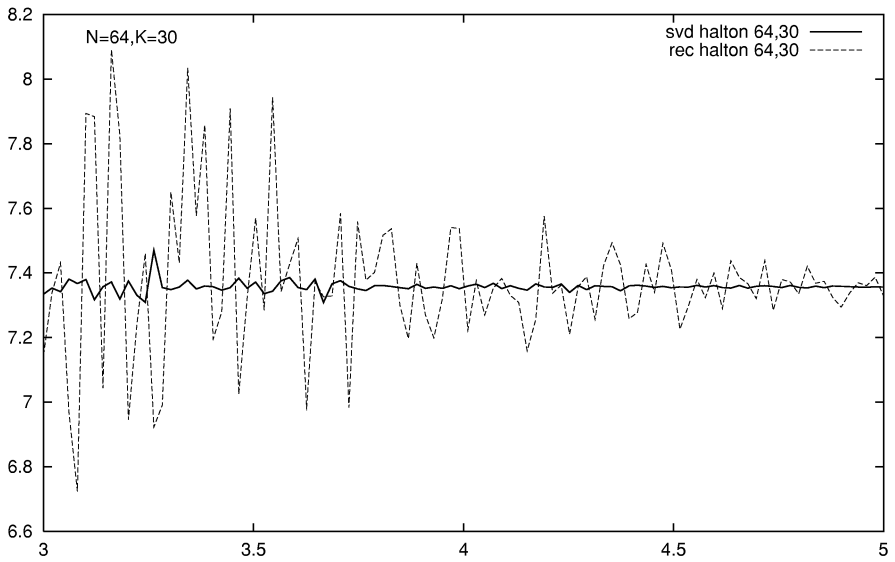


Figure 4.1: Comparison of the SVD method and the standard method (labeled REC) for a basket option where $N = 64$ and $K = 30$. The price of the option is given on the Y-axis, while the number of drawings in $\log_{10}$ scale is given on the X-axis for each simulated price. The option is at-the-money with strike $q = 100$, $T = 1$ (one year to maturity), and is priced in a market with risk free rate $r = 0.05$

sequence used in this simulation, and in the calculations performed in sec. 4.6, is the Halton leaped sequence with leap number 31. This sequence was presented in [KW97], together with other good leap values. In the next section we show how we can exploit the SVD method to further reduce the problem complexity and required computing time for a given accuracy.

4.5 Reducing the Problem Complexity

In sec. 4.4 we presented a method to find an ordering of the noise components of the problem, and this ability together with the mentioned property of many low discrepancy sequences that the equidistribution of some components are better than other, can be exploited to achieve faster convergence. The questions we want to elaborate on in this section are how we can measure this increased convergence, how we can reduce the problem complexity by using the SVD technique, and which criteria we can use to perform this reduction. By using a general concept, labeled the ANOVA decomposition, it is possible to find a representation of the integrand as a sum $\mathcal{F}$ of orthogonal functions. If each of these orthogonal functions depends only on a distinct subset of the coordinates, the integrand can be written as a sum of integrals of functions of lower dimension, and the complexity of the integral problem has been reduced with regards to the dimension. Even if we are not able to reduce the dimension of the original integrand by this approach, it may be that some of the orthogonal functions in $\mathcal{F}$, say $\mathcal{G} \subset \mathcal{F}$, have little effect on the value of the integral. Then if $\mathcal{F} - \mathcal{G}$ have dimension M, and M is lower than the dimension NK of the original integrand, but estimates the true value of the integral within acceptable limits α when integrated, we say that the original integrand has effective dimension M.

It is hard to use this concept directly on the Asian basket option problem, because we have a complex function where both $\hat{\phi}(\cdot)$ and $\exp(\cdot)$ are present. However, it turns out that we can find an approximation that works very well: Consider the noise term in (4.15);

$$\zeta_{nk}^{l} = \sum_{d=1}^{N} \sigma_{nd} \sum_{j=0}^{K} C_{kj} \varepsilon_{O(d,j)}^{l} \ . \tag{4.16}$$

The matrices C and σ are results of SVDs, and can be written $E_{C^2}\sqrt{\Lambda_{C^2}}$ and $E_{\sigma^2}\sqrt{\Lambda_{\sigma^2}}$ respectively, where $|E_{\sigma^2}| = 1$ and $|E_{C^2}| = 1$. The variance of the noise term can therefore be written

$$\mathrm{Var}[\zeta_{nk}] = \sum_{d=1}^{N} \sum_{j=0}^{K} \big((E_{\sigma^2}\sqrt{\Lambda_{\sigma^2}})_{nd}(E_{C^2}\sqrt{\Lambda_{C^2}})_{kj}\big)^2 \tag{4.17}$$

$$= \sum_{d=1}^{N} \sum_{j=0}^{K} (E_{\sigma^2})_{nd}^2 (E_{C^2})_{kj}^2 (\lambda_{\sigma^2})_d (\lambda_{C^2})_j \ . \tag{4.18}$$

If the product of the eigenvalues in (4.18) is small for a pair of indexes (d, j), the contribution to $\mathrm{Var}[\zeta_{nk}]$ from this element of the sum is small, and can be neglected in an approximate solution. Note that $\lambda = \lambda_{\sigma^2} \otimes \lambda_{C^2}$ are entries of the matrix $\Lambda_{C_{\hat{X}}^2}$ defined in sec. 4.2, and the ordering of them is given by $O(\cdot, \cdot)$. Our hypothesis is that by choosing an error tolerance α for the integration, we can find an approximate

effective dimension M of the problem by the relation

$$\inf_{M}\Big[\sum_{m=1}^{M}\lambda_m\Big] \geq (1-\alpha)\sum_{m=1}^{NK}\lambda_m \ , \tag{4.19}$$

and by the mapping $O(\cdot,\cdot)$ we find which pairs (d_m, j_m) of the noise term we should use in the approximate integration to minimize the work-load. The resulting approximation of the noise term is therefore given by

$$\zeta_{nk} \approx \sum_{m=1}^{M}\sigma_{nd_m}C_{kj_m}\varepsilon_m \ . \tag{4.20}$$

Actually the proposed approach in (4.19) will eliminate the smallest noise terms in a problem defined by a linear function: Let $f(\zeta) = \sum_{n=1}^{N}\sum_{k=0}^{K}\zeta_{nk}$, then f has variance

$$\mathrm{Var}[f(\zeta)] = \sum_{d=1}^{N}\sum_{j=0}^{K}\Big(\sum_{n=1}^{N}\sum_{k=0}^{K}\sigma_{nd}C_{kj}\Big)^2 \tag{4.21}$$

$$= \sum_{d=1}^{N}\sum_{j=0}^{K}(\lambda_{\sigma^2})_d(\lambda_{C^2})_j \ , \tag{4.22}$$

and the smallest entries can easily be found by considering the eigenvalues alone.

In sec. 4.6 we do calculations on the integrand of the Asian basket option problem to show that (4.19) represents a good method to reduce the problem complexity within known accuracy constraints, and that the method produces a good estimate for the effective dimension. Furthermore we fix $\alpha = 0.01$ and find the effective dimension $M(N,K)$ for different sized baskets by using (4.19). The results are given in tabular and plotted form.

4.6 Numerical Results

We have chosen to test the method in (4.19) for three different configurations. All of them, however, are at-the-money options with strike $q = 100$, $T = 1$ (one year left to maturity), and priced in a market with risk free rate $r = 0.05$. All simulation results are obtained by using $L = 10^4$ number of drawings. The first result is given in Fig. 4.2 for a basket option consisting of $N = 32$ assets using $K = 32$ time discretization points. The next two is given in Figs. 4.3 and 4.4 for basket options with $N = 2, K = 256$ and $N = 64, K = 2$, respectively. The number of assets and time discretization points are specified in the upper left corner of each figure. We need some notation to describe the plots: Let $v_M(\alpha)$ be the value of the option by using the M biggest noise terms, M is given by specifying α in (4.19), and let $\hat{\varphi}_M(\alpha)$ be the integrand by the same logic. The plots include five graphs each: The *Value fraction* which is the ratio $v_M(\alpha)/v_{NK}(1)$, the *Var fraction* which is the ratio

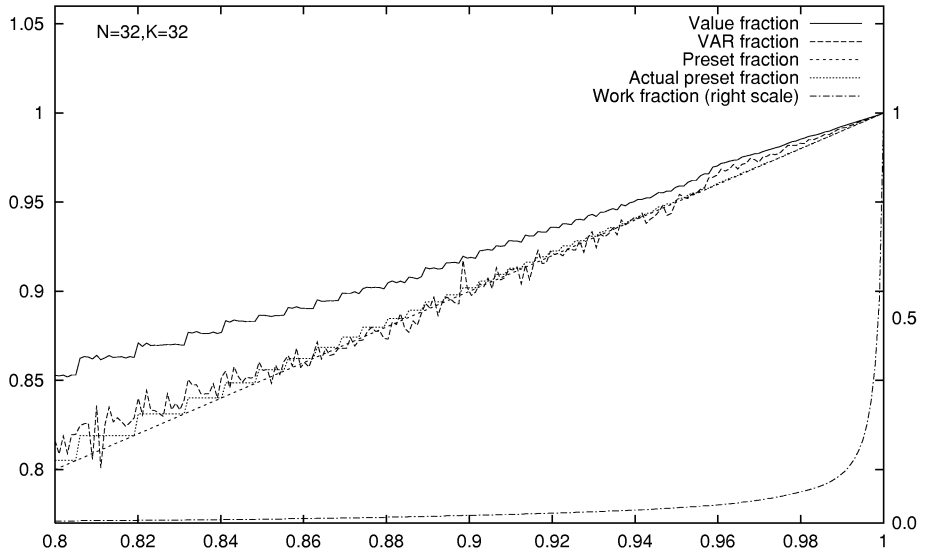


Figure 4.2: Different ratios (explained in sec. 4.6) between calculation with full dimension and calculation for dimension M as a function of the preset fraction $(1-\alpha)$. The basket contains $N = 32$ assets, and the time discretization is given by $K = 32$

$\mathrm{Var}[\hat{\varphi}_M(\alpha)]/\mathrm{Var}[\hat{\varphi}_{NK}(1)]$, the *Preset fraction* which is simply $(1-\alpha)$, the *Actual preset fraction* which is $\sum_{m=1}^{M}\lambda_m/\sum_{m=1}^{NK}\lambda_m$ and the *Work fraction* which is M/NK. Note that the work fraction uses the right Y-axis as scale. The work fraction is also given in Table 4.1 and in Fig. 4.5 for a broader range of configurations. The different baskets contains assets from Oslo stock exchange, and therefore represents a realistic correlation structure for this part of the noise in the problem.

An interesting special case of the above calculations is the single asset Asian option. This is because the results are static, as opposed to basket options where the results can be different for different equal-sized baskets. In this special case we have calculated the effective dimension for a broader range of K by the use of (4.19). The result is given in Fig. 4.6. Here we have included the analytic function $r(x) = \mathrm{ceil}(13.5\,\mathrm{atan}(0.07x))$ that mimics the behavior of the effective dimension curve, and can serve as a rule of thumb. The equivalent results for the single asset Asian option as the ones presented in Figs. 4.2, 4.3 and 4.4 looks very much like those in Fig. 4.3, that is the *Value fraction* and the *Var fraction* is close to, but above, the *Actual preset fraction*. This means that the effective dimension plotted in Fig. 4.6 can be used as a close but conservative estimate. Furthermore the *Work fraction* is below 20 percent for $\alpha = 0.005$, and about 8 percent for $\alpha = 0.01$.

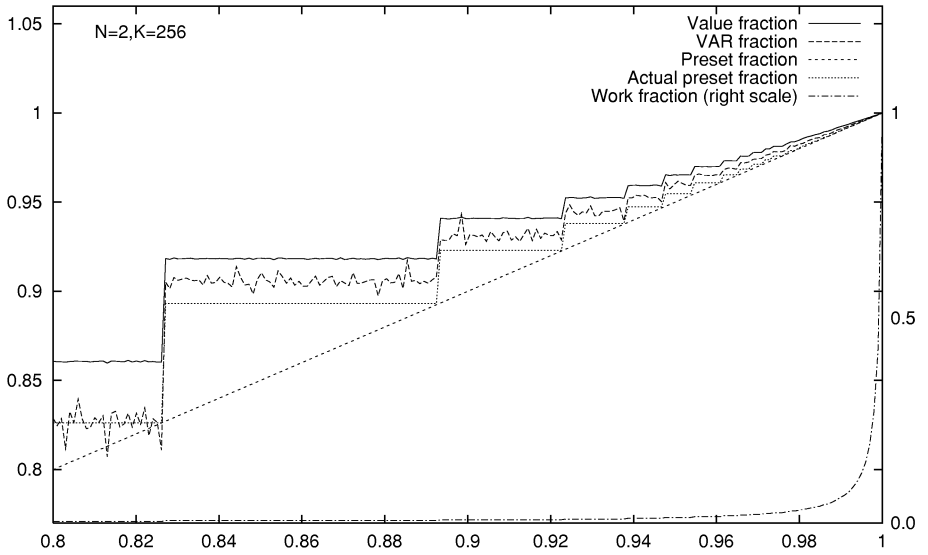


Figure 4.3: Different ratios (explained in sec. 4.6) between calculation with full dimension and calculation for dimension M as a function of the preset fraction $(1-\alpha)$. The basket contains $N = 2$ assets, and the time discretization is given by $K = 256$

Table 4.1: Effective dimension for different basket sizes and discretization schemes. The entries in the table are $\frac{M}{NK}$ where M is found for $\alpha = 0.01$ and NK is the dimension of the full problem. The ratio is the work-load needed for the approximate simulation in each case

N \ K	8	16	24	32	40	48	56	64
8	$\frac{25}{64}$	$\frac{35}{128}$	$\frac{43}{192}$	$\frac{51}{256}$	$\frac{56}{320}$	$\frac{60}{384}$	$\frac{63}{448}$	$\frac{65}{512}$
16	$\frac{41}{128}$	$\frac{52}{256}$	$\frac{61}{384}$	$\frac{69}{512}$	$\frac{77}{640}$	$\frac{84}{768}$	$\frac{89}{896}$	$\frac{92}{1024}$
24	$\frac{65}{192}$	$\frac{81}{384}$	$\frac{90}{576}$	$\frac{99}{768}$	$\frac{107}{960}$	$\frac{116}{1152}$	$\frac{123}{1344}$	$\frac{129}{1536}$
32	$\frac{86}{256}$	$\frac{104}{512}$	$\frac{114}{768}$	$\frac{123}{1024}$	$\frac{132}{1280}$	$\frac{140}{1536}$	$\frac{148}{1792}$	$\frac{156}{2048}$
40	$\frac{105}{320}$	$\frac{127}{640}$	$\frac{138}{960}$	$\frac{147}{1280}$	$\frac{156}{1600}$	$\frac{164}{1920}$	$\frac{173}{2240}$	$\frac{181}{2560}$
48	$\frac{120}{384}$	$\frac{145}{768}$	$\frac{157}{1152}$	$\frac{167}{1536}$	$\frac{176}{1920}$	$\frac{184}{2304}$	$\frac{193}{2688}$	$\frac{201}{3072}$
56	$\frac{142}{448}$	$\frac{173}{896}$	$\frac{187}{1344}$	$\frac{197}{1792}$	$\frac{206}{2240}$	$\frac{215}{2688}$	$\frac{223}{3136}$	$\frac{231}{3584}$
64	$\frac{169}{512}$	$\frac{209}{1024}$	$\frac{225}{1536}$	$\frac{237}{2048}$	$\frac{246}{2560}$	$\frac{255}{3072}$	$\frac{264}{3584}$	$\frac{272}{4096}$

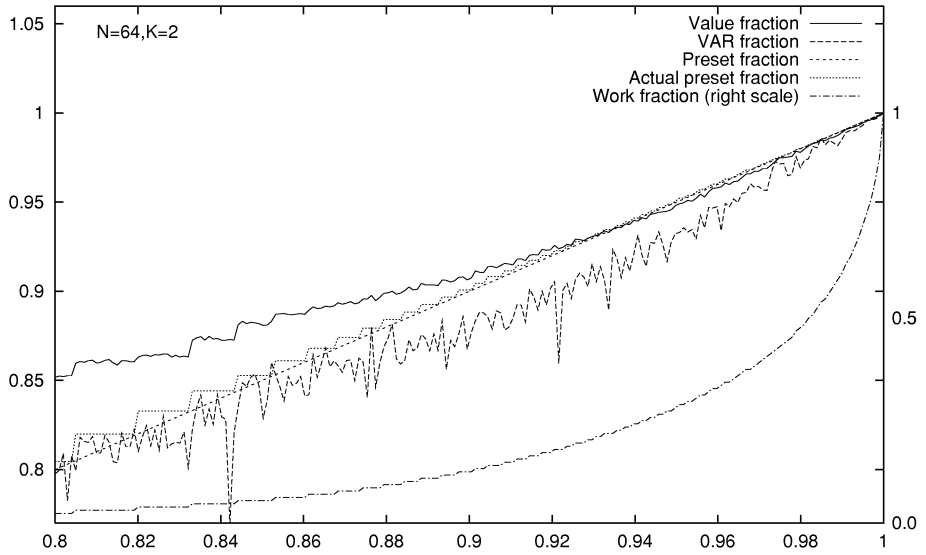


Figure 4.4: Different ratios (explained in sec. 4.6) between calculation with full dimension and calculation for dimension M as a function of the preset fraction $(1-\alpha)$. The basket contains $N = 64$ assets, and the time discretization is given by $K = 2$

4.7 Conclusion

The proposed method gives a rather easy-to-use tool to decide an approximate effective dimension of Asian basket options. From the calculations it is evident that as the number of assets N in the basket is increased, the work-load of the approximate method increases more than by a similar increase in time discretization points K.

Acknowledgements

The first author is grateful for financial support from Storebrand Investments. The second author is grateful to MaPhySto and the Norwegian Computing Centre for financial support. MaPhySto - Centre for Mathematical Physics and Stochastics, is funded by a grant from the Danish National Research Foundation.

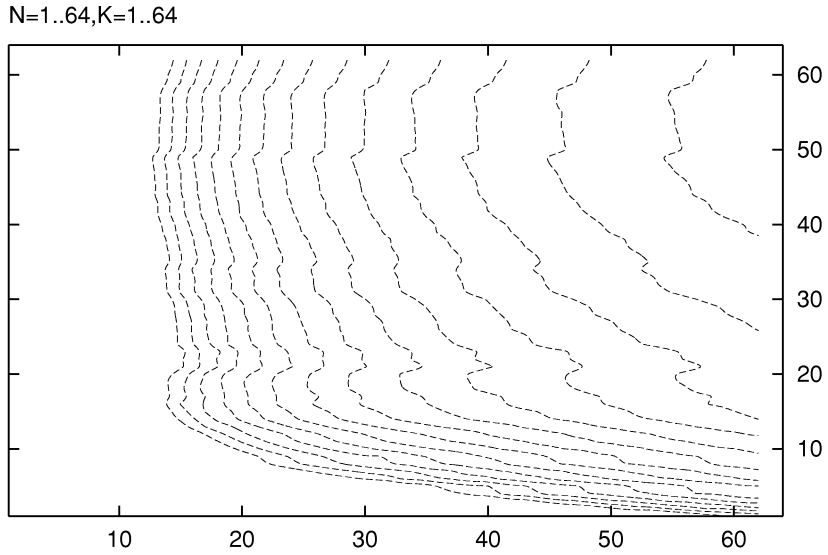


Figure 4.5: A contour map of the ratio between M and NK for different K and N, given $\alpha = 0.01$. The contour lines encapsulate configurations in which the work-load is less than 7–20 percent – starting at 7 percent in the upper right corner. The X-axis represents K and the Y-axis N

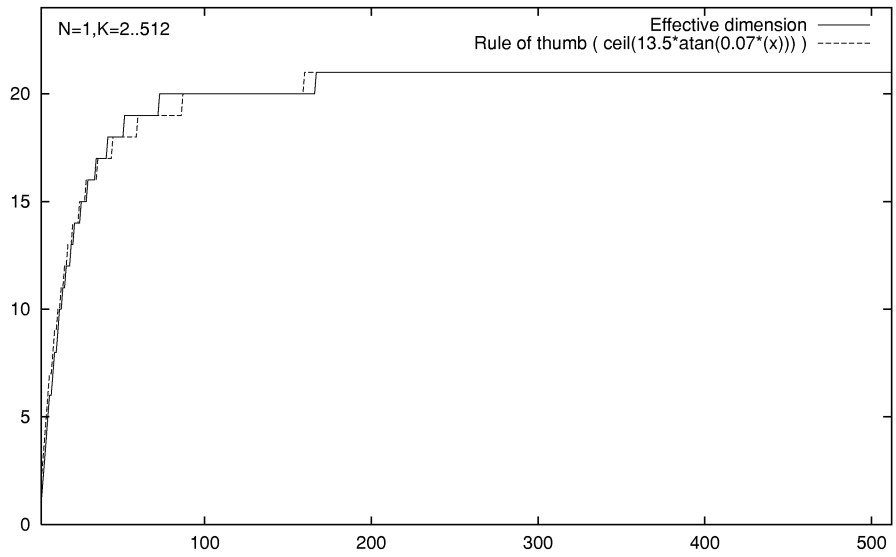


Figure 4.6: The effective dimension M as a function of the dimension K of the single asset Asian option problem. We have included the graph of the function $r(x) = \text{ceil}(13.5\,\text{atan}(0.07x))$, showing that we can use this as a rule of thumb to set the effective dimension for the problem

Chapter 5

An Adaptive Method for Pricing Contingent Claims. Part I[1]

Author(s): Lars O. Dahl

Keywords: Adaptive integration, multi dimensional integration, pricing of contingent claims, quasi-Monte Carlo methods

Abstract: The paper presents an adaptive method for the evaluation of multidimensional integrals over the unit cube. The measure used to partition the domain is suited for integrands which are monotonic in each dimension individually, and is therefore suitable for problems stemming from finance where this is often the case. We use a QMC method for each sub-problem resulting from the partitioning of the domain. The article is part one of a work on this topic, and presents the method together with various local variance reduction techniques. The material is presented with an alignment to option pricing problems. In the companion paper we present an option pricing problem and simulation results on different setups of this. We compare the convergence properties of the adaptive method with the convergence properties of the QMC method used directly on the problem. We find that the adaptive method in many configurations outperform the conventional QMC method, and we develop criteria on the problem for when the adaptive method can be expected to outperform the conventional.

5.1 Introduction

The fair value of contingent claims can be expressed as an expected value, which in turn can be written as an integral. In many situations these integrals are multidimensional. The value of the integrand can vary significantly in the domain of integration, and in some cases there can be only small parts of the domain in which the integrand

[1]Submitted for publication August 2002 to Journal of Applied and Theoretical Finance. Full reference: [Dah01]

is non-zero. Since Monte Carlo methods and QMC methods distribute the evaluation points as evenly as possible, they will waste calculations on regions which are not important. An adaptive method tries to allocate the resources to the important parts of the domain.

In option pricing problems where the option is written on several underlying assets and/or the option price is path dependent in time, the price is found by calculating a multidimensional integral. QMC methods are often deployed for these problems, and additional methods for reducing the variance of the estimator for the price of the option can be found. The adaptive technique incorporates some common variance reduction principles implicitly, but the modifications of the integrand are not exactly the same. In particular the adaptive method is closely related to importance sampling and stratified sampling. References on these methods are [GHS99a], [GHS98], [GHS99b], [GHS00], [OZ00], [PF90], [VAD98]. The purpose of these methods, as the adaptive, is to use the samples from the QMC generator in a more effective way than distributing them evenly in the domain. Some knowledge of the integrand must be present, or collected, for the success of these methods. Importance sampling can be dangerous in that one is not guaranteed lower variance. The variance can actually blow up for special cases. The advantage of the adaptive method is that information is collected as the calculation is done, and function evaluations of the integrand is used to guide the distribution of the points in the domain.

The challenge in getting the adaptive method to perform well is to find good measures of the behavior of the integrand in local domains, and to find a data structure that minimizes overhead from the process of analyzing the domain. The basic algorithmic approach in the article is partly based on the work presented in [BEG91], but the approach is modified in some essential aspects concerning the use of function evaluations in local domains. In addition to presenting the principles of the adaptive method, we develop and formulate variance reduction techniques applied to the local domains produced by the sub-division process of the adaptive method. This, together with the use of MC and QMC for the evaluation of the sub-problems resulting from the adaptive process is to our knowledge not investigated before.

In the companion paper [Dah02a] we perform numerical tests and find criteria for deciding when the adaptive method can be expected to perform well.

5.2 QMC Integration

The goal is to evaluate multidimensional integrals over the region $\Omega = [0,1]^D$ accurately, effectively and robust. When $D > 4$ the MC and QMC methods are competitive with any advanced numerical integration method for general integrands. For large D they are the only realistic alternatives. The estimator most commonly used for the integral is given by

$$F = \int_\Omega f(x)\,dx \approx \frac{1}{L}\sum_{l=1}^{L} f(y_l)\ , \tag{5.1}$$

where $\{y_l\}$ is a sequence of vectors from a pseudo random number generator or a low discrepancy sequence. The MC method is based on drawing random or pseudo random numbers as arguments to f, while the QMC methods use sequences of arguments to f designed to be as evenly distributed as possible in Ω. In this article we use the Mersenne Twister as pseudo random number generator and an extension of the Halton[2] sequence as basis for the QMC method. References on low discrepancy sequences are e.g. [NX98], [Owe99], [Owe98].

5.3 The Adaptive Method

It is easy to construct a QMC estimator for the integral over a part of Ω, and we therefore can construct a method to evaluate the integral over all of Ω as a sum of such estimated values. Let $\Omega = \cup_i \omega_i$, $\cap_i \omega_i = \emptyset$, $i = 1, \ldots, P$. Then

$$F = \sum_{i=1}^{P} \int_{\omega_i} f(x)\, dx \tag{5.2}$$

$$\approx \sum_{i=1}^{P} \frac{|\omega_i|}{\Delta L_i} \sum_{l=L_{i-1}}^{L_i - 1} f(y_l)\ , \tag{5.3}$$

where $\Delta L_i = L_i - L_{i-1}$ and y_l is scaled such that $y_l \in \omega_i$ when $l \in [L_{i-1}, L_i]$. $|\omega_i|$ is to be understood as the volume of ω_i. The adaptive algorithm should decide on the number of sub-domains and their sizes. That is P and $|\omega_i|\,, \forall i$. Furthermore, the algorithm has to pick the best set of sub-domains, and how many simulation points ΔL_i to use in each of them. Alternative approaches use information from the integrand to develop approximations of the integral in sub-domains with a deterministic approach rather than with simulation. This is done in e.g. [BEG91], [GC97], [Coo97], [PF90].

We have chosen to use a binary tree to represent the domain decomposition. Each node in the tree corresponds to a distinct part of the domain, and when we expand the tree we divide the domain represented by a node in two parts (not necessarily of equal size). The criteria we use to decide on division are twofold: First we find an estimate of how much the domain contributes to the overall variance of the integral, and if this is more than a preset amount we split the domain. In order to avoid an explosive increase in partitions, we only create two new subdomains from each domain that meets the variance criterion. They are produced by dividing the domain across the axis corresponding to the dimension along which the integrand has largest variability according to a well behaving measure. The divide and conquer algorithm is terminated when the estimated variability in all sub-domains are less than a preset limit. This approach could ideally produce an answer with guaranteed accuracy, but the measure we use for the variance contribution is only indicative. The variance

[2] We have in previous work ([DB01] and [DB02]) found that the Halton leaped sequence with leap number 31, presented in [KW97], performs well compared to other easily implemented sequences

contribution from a node is calculated by the following approach: Consider a sub-domain ω_i. Place one point p_0^i in the middle of ω_i and two points p_{j1}^i, p_{j2}^i on each line l_j going through the middle of ω_i, parallel with the axis A_j. The points $p_{j\cdot}^i$ are located on each of the borders of ω_i. For each point, find the value $f(p_{j\cdot}^i)$ of the integrand. This setup is illustrated in fig. 5.3 for a two dimensional example. We estimate the contribution to the variance from each sub-domain by the expression

$$\delta_i = \frac{|\omega_i|}{2D}\Big(\sum_{j=1}^{D} \frac{f(p_{j1}^i) + f(p_{j2}^i)}{2} - f(p_0^i)\Big) \ , \tag{5.4}$$

where the volume of the sub-domain is calculated by $|\omega_i| = \prod_{j=1}^{D}(p_{j2}^i - p_{j1}^i)$. The axis we divide in order to create a finer partitioning is found by one of the expressions

$$A_i = \sup_j\{j; \big|f(p_{j2}^i) - f(p_0^i)\big| + \big|f(p_{j1}^i) - f(p_0^i)\big|\} \tag{5.5}$$

$$\text{or } A_i = \sup_j\{j; \big|f(p_{j2}^i) + f(p_{j1}^i) - 2f(p_0^i)\big| \ , \tag{5.6}$$

depending on the problem. These calculations are performed recursively as given in alg. 1. In [BEG91] they use a fourth difference operator, but we find that the rule (5.5), which is a second difference operator, suits our setup better and gives better overall performance for the types of problems we have tested.

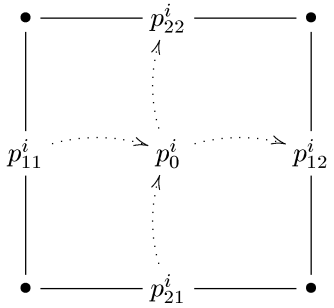


Figure 5.1: Illustration of the placement of points in ω_i in order to calculate the axis of division and the estimated variance of the sub-domain. Note that $\Omega = [0,1]^2$ in this illustration

If the adaptive algorithm performs perfectly in accordance with the assumptions, the contribution from each sub-domain to the overall variance should be equal. Therefore $\sigma_i\,|\omega_i| = c$, $\forall i$ ideally. But even if the adaptive process aspire to use the simulation points as effectively as possible, we get some sub-domains in which the measured variability is close to the preset limit, and some where the variability is considerably lower than the limit. To circumvent this behavior we use less simulation points in the sub-domains where the variability is low. Theoretically, the fraction for the optimal

allocation of points in each sub-domain can be shown to be

$$q_i^* = \frac{r_i \sigma_i}{\sum_{l=1}^{P} r_l \sigma_l},\ i = 1, \ldots, P\ , \tag{5.7}$$

where r_i is the probability for a point to be contained in each bin represented by ω_i. Therefore $r_i = |\omega_i|$ in our setting. This leads to the allocation of simulation points by the relation

$$\Delta L_{i+1} = L \frac{\sigma_i \, |\omega_i|}{\sum_{l=1}^{P} \sigma_l \, |\omega_l|}\ , \tag{5.8}$$

where L is the total budget of simulation points. This approach, however, assume knowledge of all σ_i, and the adaptive approach does not provide this knowledge at the stage in the process where the contributions to the value of the total integral are calculated. Instead, we have chosen to implement a simpler approach to finding the number of simulation points in each bin. We use the relation

$$\Delta L_{i+1} = L \frac{\sigma_i}{\delta}\ , \tag{5.9}$$

where δ is the stopping criterion for the adaptive process on the variance estimates. This approach avoid the overhead by traversing the tree to collect the σ_l values, and in our tests the approach works well compared to using the same number of simulations in each bin.

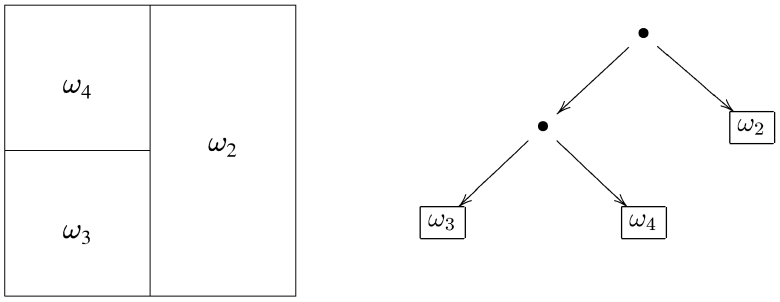


Figure 5.2: Illustration of method for partitioning the domain $\Omega = \cup_i \omega_i$. The framed boxes in the tree cover the domain without overlapping

The adaptive approach is of course infested with overhead compared to plain simulation. For the adaptive method to be competitive with regards to computing time, one must focus on building good data structures and use effective algorithms for the administration of the simulation. In addition to the overhead connected to the creation and traversing of the binary tree holding the information about the segmentation of the domain, extra integrand evaluations must be performed in order to guide the partitioning of the domain. See fig. 5.2 for an illustration. We are unfortunately not able

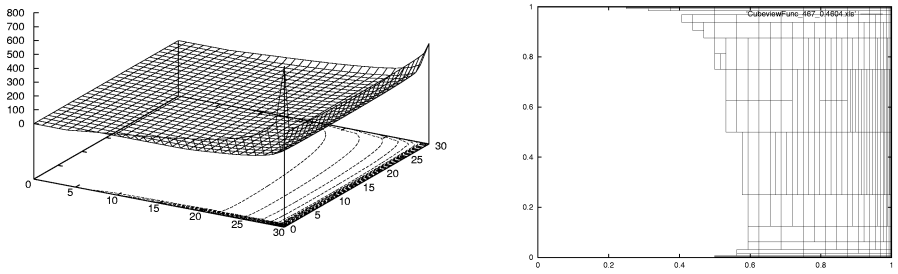


Figure 5.3: An example (in the right part of the figure) of how the adaptive method would partition the domain for the payoff function of a two asset basket option graphed in the left figure

to use the calculated points in the adaptive process in the estimation of the integral, because a bias will be introduced. In many classes of problems, however, the domain can be reused when finding other parameters of the problem. An example of this can be found in finance where one often need to calculate the hedges (derivatives with respect to certain parameters) of an option contract as well as the price, see [BG96] on simulation of hedging parameters, and [FLL$^+$99] and [FLL$^+$01] for a formulation of the hedges by Malliavin calculus. In the adaptive setting the hedges can be calculated effectively without having to recreate the partitioning of the domain, see e.g. [DBK02] or [Dah02b] for formulation and calculation of hedging parameters using this framework. The extra computer work involved in the adaptive method must therefore be accounted for in the comparison with the conventional method. In many classes of problems the adaptive method has a far better performance, even when all aspects of increased time consumption are taken into account. For a full discussion of the criteria on the problem for this to be the case, look at the companion paper [Dah02a].

In addition to simply sub-dividing the problem, there are several techniques that can be applied to the integrand in each sub-domain to improve the performance further. We notice, however, that the performance is not always increased. An investigation and discussion of these aspects are also carried out in the companion paper [Dah02a]. In the next section we present the techniques for variance reduction on the estimator of the integral, and focus on methods which can be applied to each sub-problem from the domain-decomposition in the adaptive process.

5.4 Local Variance Reduction

The estimator F for the unknown integral is sensitive to the choice of sampling points. This is the background for the possibility to find estimators and techniques delivering

a low variance estimate. Variance reduction techniques may involve using calculation methods, such as QMC, designed to remove uncertainty. A supplementary approach often used together with the QMC principle, is to find alternative estimators for F than the mean value. The only requirement on the new estimator is that it has no, or neglectable, bias. Control variates (CV), weighted uniform sampling (WUS), stratified sampling (SS) and importance sampling (IS) are categories of techniques to provide such estimators. In our adaptive setting, we aim at employing such techniques also locally in each sub-domain. If we denote by G_ω the exact integral of a chosen function $g_\omega(x)$ in a sub-domain ω, the estimator F_ω of the unknown integral in a local domain can be found by the modified estimators;

$$\text{CV:} \quad F_\omega = \sum_{l=L_i}^{L_{i+1}} (f_\omega(y_l) - g_\omega(y_l)) + G_\omega \tag{5.10}$$

$$\text{WUS:} \quad F_\omega = \frac{\sum_{l=L_i}^{L_{i+1}} f_\omega(y_l)}{\sum g_\omega(y_l)} G_\omega \tag{5.11}$$

$$\text{IS:} \quad F_\omega = \sum_{l=L_i}^{L_{i+1}} \frac{f_\omega(z_l)}{g_\omega(z_l)} G_\omega \ , \tag{5.12}$$

where $y \sim \omega$ uniformly, while $z \sim \frac{g(y)}{G_\omega}$ (which is also in ω, but not uniformly). A description of how to generate z for a general g is given in e.g. [Ö99]. Stratification is somewhat different than the techniques formulated in (5.10)-(5.12), and aims at distributing the arguments to f evenly, so as to guarantee that the average is actually representable for f. We explain the connection between SS and the adaptive method in the next section. In the following sections we handle CV and IS. WUS is very similar to CV and the same considerations apply. We therefore refrain from covering WUS in addition to CV.

5.4.1 Stratification

Stratification is a powerful technique and is related with the principles of the QMC method. The goal is to create a distribution of points in the domain for which a guaranteed fraction lies in specific bins. This is actually what is going on in special types of low discrepancy sequence generators such as the Latin hyper-cube sampling method. The adaptive method is an advanced sort of stratification in that we adapt the need for stratification as the integrand is changing. To gain further benefit from the stratification, it is for some problems possible to modify the integrand so that one can stratify the dimensions of the problem contributing the most to the variance of the integral estimator. In the option pricing problem one can re-engineer the problem so that the dimensions in the resulting integration problem correspond to independent stochastic processes in the original problem of finding an estimated expected value. By doing this, we can find the dimensions of the problem having the biggest contribution to the uncertainty of the final answer. By using more sampling points

along the most important dimensions and by guaranteeing that these sampling points are evenly distributed, we can reduce the uncertainty, while keeping the number of sampling points low. Stratification is in the option pricing problem taken care of by the use of an SV-decomposition of the covariance matrix observed from the market or derived from the problem, and a rearranging of the independent noise components according to the size of the corresponding eigenvalues. When we then use a QMC method as the engine for simulating the values, stratification is, depending of the low discrepancy sequence used, more or less taken care of implicitly. In addition to change the integration problem to be able to identify the biggest noise components, we have locally experimented with rearranging the noise components according to the axis along which the integrand changes most rapid, but this does not seem to give consistent gain in overall variance levels of the final estimator.

5.4.2 Control Variates

One method of achieving further variance reduction is to use the so called Control variate technique. We will employ this technique locally in each sub-domain. It must be adjusted to the limitations in number of integrand evaluations we need to impose on our selves to maintain speed, and we therefore use the integrand values in the points p_{j1}, p_{j2}, p_0, which we have already calculated, to find a function $g_i(x)$ replicating the integrand f as closely as possible in each domain ω_i. This approach enables us to use the sampling points in the QMC approach to evaluate a modified integral with lower value, and thus smaller variance. The unbiased estimator of the integral in the domain ω_i is obtained by adding the known deterministic value G_i of the integral of the function g_i to the integral of the difference between f and g_i. Lets introduce some notation to state this exact: Pricing an option corresponds to evaluating the integral $\int_{[0,1]^D} f(x)\,dx$. In each sub-domain we evaluate $\int_{\omega_i} f(x)\,dx$. The idea is to find an approximate function $g_i(x)$ for $f(x)$ in each sub-domain ω_i, having the property that the integral G_i of g_i in this domain can be found deterministic. That is,

$$F_i = \int_{\omega_i} f(x)\,dx \tag{5.13}$$

$$= \int_{\omega_i} f(x) - g_i(x)\,dx + G_i\ . \tag{5.14}$$

If the integral of the difference between the functions f and g_i is smaller than the integral of f in ω_i, this approach will give an estimate of F_i with lower variance than evaluating the integral of $f(x)$ in ω_i directly.

Finding g_i and G_i have to be relatively cheap for this approach to be competitive. We have developed three different function classes that meet these demands. In the following we drop the subscripts i and ω_i on the functions for cleaner notation. Note, however, that all parameters, functions and integrals in the remaining part of the

section are found for a general local domain. The three function classes are given by

$$g(x) = \sum_{j=1}^{D} a_j (x_j - h_j)^{e_1} + \sum_{j=1}^{D} b_j (x_j - h_j)^{e_2} + c \tag{5.15}$$

$$g(x) = \exp(\sum_{j=1}^{D} a_j (x_j - h_j)) - \exp(\sum_{j=1}^{D} b_j (x_j - h_j)) + c \tag{5.16}$$

$$g(x) = \sum_{j=1}^{D} a_j (\exp\left(b_j (x_j - h_j)\right) - 1) + c\,, \tag{5.17}$$

where $h_j = (p_{j1} + p_{j2})/2$. These functions are accommodating with regard to the placement of the points in the sub-domains of the adaptive algorithm, since we as part of the adaptive process already have evaluated f in these points. To find the parameters a_j, b_j and c we only need to solve D equations in two variables and one equation for c. First we consider the polynomial function in (5.15). The solution of the equations for $a_j, b_j\ \forall j$ for this function has the structure:

$$a_j = \frac{\bar{h}_j^{e_2}(f(d_j) - c) - (-\bar{h}_j)^{e_2}(f(u_j) - c)}{\bar{h}_j^{e_1}(-\bar{h}_j)^{e_2} - \bar{h}_j^{e_2}(-\bar{h}_j)^{e_1}} \tag{5.18}$$

$$b_j = \frac{(-\bar{h}_j)^{e_1}(f(d_j) - c) + \bar{h}_j^{e_1}(f(u_j) - c)}{\bar{h}_j^{e_1}(-\bar{h}_j)^{e_2} - \bar{h}_j^{e_2}(-\bar{h}_j)^{e_1}}\,, \tag{5.19}$$

where $\bar{h}_j = (p_{j2} - p_{j1})/2$ and $c = f(h_1, h_2, \ldots, h_D)$. The value of the integral of g over a sub-domain is

$$G = \prod_{k=1}^{D} 2(\bar{h}_k)\Big(\frac{1}{e_1 + 1}\sum_{j=1}^{D} \frac{a_j\big((\bar{h}_j)^{e_1+1} - (-\bar{h}_j)^{e_1+1}\big)}{2\bar{h}_j} + \frac{1}{e_2 + 1}\sum_{j=1}^{D} \frac{b_j\big((\bar{h}_j)^{e_2+1} - (-\bar{h}_j)^{e_2+1}\big)}{2\bar{h}_j} + c\Big)\,. \tag{5.20}$$

It is not difficult to see that if $(e_1 + 1)$ is divisible by two, the first part does not contribute to the integral, and similarly for the second part if $(e_2 + 1)$ is divisible by two. The relation between e_1 and e_2 must be controlled in order for the equations to have solution: Without loosing generality assume $a_1 \geq a_2$. We must have $a_1 = a_2 + (2n + 1)$, where $n \in \{0, 1, \ldots\}$.

The class of functions where we have a sum of exponential functions, given in (5.16), is also rather easy to handle: Let $r_{j1} = f(p_{j1}) - c$, $r_{j2} = f(p_{j2}) - c$ and $\bar{h}_j = (p_{j2} - p_{j1})/2$. Then the parameters of this class are given by

$$a_j = -\frac{r_{j1} r_{j2}}{r_{j2} + r_{j1}} \qquad\qquad b_j = \frac{\ln(-r_{j1}/r_{j2})}{\bar{h}_j}\,, \tag{5.21}$$

and the value of the integral is

$$G = \left(\prod_{j=1}^{D} 2\bar{h}_j \right) \left[\sum_{j=1}^{D} a_j \left(\frac{\exp(b_j \bar{h}_j) - \exp(-b_j \bar{h}_j)}{2 b_j \bar{h}_j} - 1 \right) + c \right] . \tag{5.22}$$

It is, however, evident that this function class is not very adapted to the option pricing problem in general, because the sum of variables in the argument to the exponential function is not taken into account. If we look at the last function class, given in (5.17), this is taken better care of. We can find the parameters of this class by solving a set of equations as before. Using the same notation as above, the parameters are

$$a_j = \frac{\ln(\frac{1}{2r_{j1}}(r_{j1}r_{j2} \pm \sqrt{r_{j1}^2 r_{j2}^2 - 4r_{j1}r_{j2}}))}{\bar{h}_j} \tag{5.23}$$

$$b_j = \frac{\ln(-r_{j2} + \frac{1}{2r_{j1}}(r_{j1}r_{j2} \pm \sqrt{r_{j1}^2 r_{j2}^2 - 4r_{j1}r_{j2}}))}{\bar{h}_j} , \tag{5.24}$$

where we pick the root which ensures that we get a positive argument to the $\ln(\cdot)$ function. Note however, that we are not able to find a solution to all a_j, b_j for general f, and the function class therefore is not sufficiently robust to be used alone. To be employed in applications, it must be combined with one of the other classes for the dimensions resulting in a negative root as argument to the $\ln(\cdot)$ function. The value of the integral of the function in (5.17) is given by

$$G = \prod_{j=1}^{D} \frac{\exp(a_j \bar{h}_j) - \exp(-a_j \bar{h}_j)}{a_j} - \prod_{j=1}^{D} \frac{\exp(b_j \bar{h}_j) - \exp(-b_j \bar{h}_j)}{b_j} + c \prod_{j=1}^{D} 2\bar{h}_j . \tag{5.25}$$

The function class capturing the mix of the two exponential function classes, (5.16) and (5.17), is given by

$$g(x) = \exp\left(\sum_{j \in \hat{D}} a_j (x_j - h_j)\right) - \exp\left(\sum_{j \in \hat{D}} b_j (x_j - h_j)\right) + \sum_{j \notin \hat{D}} a_j (\exp\left(b_j (x_j - h_j)\right) - 1) + c , \tag{5.26}$$

where $\hat{D} \subset \{1, \ldots, D\}$ is the set of dimension where the root is positive. The parameters are given by (5.23) and (5.24) for respectively a_j and b_j for $j \in \hat{D}$, and by (5.21) for $j \in \{1, \ldots, D\} \setminus \hat{D}$. c is equal to the value of f in the mid-point of the domain. The integral is the sum of the expressions given in (5.22) and (5.25) for the appropriate sets of indices.

As a closure of this section I will briefly mention aspects concerning a common control variate technique in Asian option pricing; namely to use the geometric average Asian option as a control variate (for which there exists an easier calculation routine) when pricing the arithmetic average Asian option. In the adapted routine this would not serve as a local variance reduction technique, because it would not take advantage of the gathered information about the local behavior. Instead, the parameters used throughout would be static, amounting to alter the problem only in a global fashion.

5.4.3 Importance Sampling

We have been looking at the possibility to perform importance sampling locally for each sub-domain by finding general distribution functions suited for this approach. As we shall see however, the attempts turned out to produce functions which were to hard to calculate as part of the adaptive process. Therefore we have only implemented a version in which the importance sampling is applied to the globally defined function before the adaptive procedure is started. The main result of these calculations with regards to the adaptive method, is that the adaptive method performs equally well together with importance sampling compared to the non-adaptive method. It is, however, not always improvements tied to the use of either importance sampling nor the adaptive method. For details on these results, consult the companion paper [Dah02a]. References on importance sampling for finance problems are [GHS99a], [GHS98], [KP92]. The first part of this section investigates local IS, and show that it is difficult to apply.

In order to formulate the inverse cumulative distribution function appearing in the general IS framework, we use the approach described in [Ö99] on the basis of the function classes (5.16) and (5.17). Let y be the low discrepancy sequence in $[0,1]^D$, and let $g(x_1,\ldots,x_D)$ be a probability density function and $G(x_1,\ldots,x_D)$ its distribution function. Denote by $g_{1,\ldots,k}(x_1,\ldots,x_k)$ the marginal density function for $x_1,\ldots,x_k$, in general given by

$$g_{1,\ldots,k}(x_1,\ldots,x_k) = \int_{[0,1]^{D-k}} g(x_1,\ldots,x_D)\, dx_{k+1}\ldots d_D \ . \tag{5.27}$$

The distribution function $G_k(x_k)$ associated with the conditional density for x_k given $x_1,\ldots,x_{k-1}$, is then given by

$$G_k(x_k) = \int_0^{x_k} \frac{g_{1,\ldots,k}(x_1,\ldots,x_k)}{g_{1,\ldots,k-1}(x_1,\ldots,x_{k-1})}\, dx \ . \tag{5.28}$$

To obtain a g-distributed sequence z in $[0,1]^D$, we need to find the function $G_k^{-1}(\cdot)$. Then $z_n = (z_{n1},\ldots,z_{nD})$ where $z_{ni} = G_i^{-1}(y_{ni})$. This principle can theoretically be applied to each sub-domain by changing the function $g(x_1,\ldots,x_D)$ to mimic the function f in each domain. For the function class (5.17), the integration in (5.27),

without the integration limits, yields

$$g_{1,\ldots,k}(x_1,\ldots,x_k) = \sum_{i=1}^{k} a_i \exp(b_i(x_i - h_i)) \prod_{j=k+1}^{D} x_j + \sum_{i=k+1}^{D} \left[\frac{a_i}{b_i} \exp(b_i(x_i - h_i)) \prod_{\substack{j=k+1 \\ j\neq i}}^{D} x_j\right] + \left(c - \sum_{i=1}^{D} a_i\right) \prod_{j=k+1}^{D} x_j \ , \quad (5.29)$$

where D is the total dimension of the integrand. When the integration limits are inserted for the local domain, and the expression for $G_k(x_k)$ is calculated, we get a complicated expression in terms of the constants. This expression can, however, be reduced to a rather tangible function in one variable when the constants are collected, giving

$$G_k(x_k) = c_{k1} \exp(c_{k2}x_k - c_{k3}) + c_{k4}x_k \ , \quad (5.30)$$

where c_{ki} are constants. We therefore have to solve an equation of the form (5.30) to find the g-distributed sequence $\{z_l\}$ for this class of functions. The solution is

$$x_k = a_{k1} W(a_{k2} \exp(a_{k3}y + a_{k4})) + a_{k5} \ , \quad (5.31)$$

where W is the so called Lambert W function and a_{ki} are constants. References on the Lambert W function are e.g. [CGH+96] and [CJK97].

In a similar manner we can develop the distribution function for the class given in eq. (5.16);

$$g_{1,\ldots,k}(x_1,\ldots,x_k) = \frac{\exp(\sum_{i=1}^{s} a_i(x_i - h_i))}{\prod_{i=k+1}^{s} a_i} - \frac{\exp(\sum_{i=1}^{s} b_i(x_i - h_i))}{\prod_{i=k+1}^{s} b_i} + c \prod_{i=k+1}^{s} x_i \ , \quad (5.32)$$

giving

$$G_k(x_k) = c_{k1} \exp(c_{k2}x_k - c_{k3}) - c_{k4} \exp(c_{k5}x_k - c_{k6}) + c_{k7} \ , \quad (5.33)$$

where c_{ki} are general constants. The inversion of this function is given by

$$x_k = a_{k1} R(a_{k2} \exp(Z) + a_{k3} \exp(a_{k4}Z + a_{k5}) + a_{k6} + a_{k7}y) + a_{k8} \ , \quad (5.34)$$

where the function R means that we have to find the roots with regards to Z of the argument to R.

Both the Lambert W function and the root in (5.34) can probably be implemented, but would be rather slow to use in the adaptive setting as they require numerical estimation or series expansion approximation. I have also developed the $G^{-1}(\cdot)$ function

for the class of polynomial functions described in eq. (5.15). In order to apply the IS technique with this function class as a base, we need to find the roots of a polynomial of degree $(e1 + 1)$, and in order for the method to be effective we need $e1 > 1$. The consequence is that we have to find the roots of at least a cubic polynomial.

The previous discussion shows that it is hard to find a general function class suited for the IS technique together with the adaptive method. An alternative is to try to change the measure under the constraint that distribution is still normal, but with other parameters. This amounts to choosing the function g to be a version of the normal distribution function. Following the deduction in e.g. [GHS99a] we look at functions $g(z) > 0 \Rightarrow h(z) > 0, \forall z \in \Omega$. Then

$$\mathrm{E}_g[G(Z)] = \int_\Omega G(z)g(z)\,dz \tag{5.35}$$

$$= \int_\Omega G(z)\frac{g(z)}{h(z)}h(z)\,dz \equiv \mathrm{E}_h[G(Z)\frac{g(Z)}{h(Z)}]\,, \tag{5.36}$$

where the subscript of the expectation indicates which measure it is taken under. The factor $g(Z)/h(Z)$ is the Radon-Nikodym derivative. The function h now gives a degree of freedom with respect to minimizing the variance.

$$\int_\Omega \left(G(z)\frac{g(z)}{h(z)} - \mathrm{E}_g[G(Z)]\right)^2 h(z)\,dz\,. \tag{5.37}$$

In our setting g is a normal distribution, and we want h to be a normal distribution as well, with the same variance structure as g. The importance sampling estimator then becomes

$$\mathrm{E}_g[G(Z)] = \mathrm{E}_h[G(Z)\exp(-\mu^T Z + \frac{1}{2}\mu^T\mu)] \tag{5.38}$$

$$= \mathrm{E}_g[G(Z+\mu)\exp(-\mu^T Z - \frac{1}{2}\mu^T\mu)]\,. \tag{5.39}$$

In [GHS99a] a detailed discussion of constructing the optimization problem for finding the best μ for h is presented. In [GHS99b] a simpler interpretation is given. It states that if we let $G(z) \equiv \exp(F(z))$, and F is approximately linear near μ, then $F(Z+\mu)$ in (5.39) is approximately $F(\mu) + \nabla F(\mu)Z$. The substitution yields

$$\exp\left(F(Z+\mu) - \mu^T Z - \frac{1}{2}\mu^T\mu\right) \approx$$

$$\exp\left(F(\mu) + \nabla F(\mu)Z - \mu^T Z - \frac{1}{2}\mu^T\mu\right)\,. \tag{5.40}$$

In order to make this expected value a non-stochastic variable we need to ensure that

$$\nabla F(\mu)Z - \mu^T Z \equiv 0 \tag{5.41}$$

$$\Rightarrow \nabla F(\mu) = \mu\,. \tag{5.42}$$

The solution to (5.42) can under conditions stated in [GHS99a] be found from a fix-point equation. We limit ourselves to stating that the method is working in the setting of our test problems. Numerical results are presented in the accompanying article [Dah02a].

Acknowledgements

The author is grateful for financial support from Storebrand Investments. I also want to thank F. E. Benth for valuable talks and for applying amendments to the manuscript.

Appendix A

In this appendix we give the pseudo-code for the adaptive subdivision algorithm. Note that the value of the integral can be calculated as the tree is created, and no additional traversion is necessary.

Algorithm 1: Adaptive subdivision

Description: The basic element of the algorithm is a node Nω representing a sub-domain. Nω has one left and one right pointer which are able to point to other nodes. Each node know its volume, can hold an estimate of the contributed variance, and knows the coordinates of its corners.
Input: A pointer ptrNω pointing to a node Nω
Output: A binary tree where the leaf nodes cover the domain exactly. The sum of the values in the leaf nodes is the value of the integral.

```
Dc(ptrNω)
if FindAxisAndVariance(ptrNω) > MaxVar
    Create(ptrNω.Left)
    Dc(ptrNω.Left)
    Create(ptrNω.Right)
    Dc(ptrNω.Right)
```

Chapter 6

An Adaptive Method for Evaluating Multidimensional Contingent Claims. Part II[1]

Author(s): Lars O. Dahl

Keywords: Adaptive integration, multi dimensional integration, Asian options, quasi-Monte Carlo methods

Abstract: This is part two of a work on adaptive integration methods aimed at multidimensional option pricing problems in finance. It presents simulation results of an adaptive method developed in the companion article [Dah01] for the evaluation of multidimensional integrals over the unit cube. The article focuses on a rather general test problem constructed to give insights in the success of the adaptive method for option pricing problems. We establish a connection between the decline rate of the ordered eigenvalues of the pricing problem and the efficiency of the adaptive method relative to the non-adaptive. This gives criteria for when the adaptive method can be expected to outperform the non-adaptive for other pricing problems. In addition to evaluating the method for different problem parameters, we present simulation results after adding various techniques to enhance the adaptive method itself. This includes using variance reduction techniques for each sub-problem resulting from the partitioning of the integration domain. All simulations are done with both pseudo-random numbers and quasi-random numbers (low discrepancy sequences), resulting in Monte Carlo (MC) and quasi-Monte Carlo (QMC) estimators and the ability to compare them in the given setting.

The results show that the adaptive method can give performance gains in the order of magnitudes for many configurations, but it should not be used incautious, since this ability depends heavily on the problem at hand.

[1]Submitted for publication August 2002 to Journal of Applied and Theoretical Finance. Full reference: [Dah02a]

6.1 Introduction

The fair value of contingent claims can be expressed as an expected value, which in turn can be written as an integral. In many situations these integrals are multidimensional. The value of the integrand can vary significantly in the domain of integration, and in some cases there can be only small parts of the domain in which the integrand is non-zero. Since MC and QMC methods distribute the evaluation points as evenly as possible, they will waste calculations on regions which are not important. An adaptive method tries to allocate the resources to the important parts of the domain.

In option pricing problems where the option is written on several underlying assets and/or the option price is path-dependent in time, the price is found by calculating a multidimensional integral. QMC methods are often deployed for these problems, and additional methods for reducing the variance of the estimator for the price of the option can be found. The adaptive technique incorporates some common variance reduction principles implicitly, but the modifications of the integrand are not exactly the same. In particular, the adaptive method is closely related to importance sampling and stratified sampling. References on these methods are [GHS99a], [GHS98], [GHS99b], [GHS00], [OZ00], [PF90], [VAD98]. The purpose of these methods, as the adaptive, is to use the samples from the QMC generator in a more effective way than distributing them evenly in the domain. Some knowledge of the integrand must be present, or collected, for the success of these methods. Importance sampling can be dangerous in that one is not guaranteed lower variance. The variance can actually blow up for special cases. The advantage of the adaptive method is that information is collected as the calculation is done, and function evaluations of the integrand is used to guide the distribution of the points in the domain.

In this article we design a parameterized test problem which has similarities with option pricing problems. We do a number of numerical simulations on this problem for various parameters of the problem, and various adjustments of the adaptive method. The purpose is to find good criteria for when the adaptive method can be expected to outperform a conventional method for different problems. In particular we experiment with local control variates in each sub-domain from the adaption process. Furthermore, we investigate the effect of using the Halton Leaped low discrepancy sequence compared to the Mersenne Twister pseudo random sequence as an engine for the resulting QMC respectively MC methods, and look at the impact these methods have on the adaptive method compared to the non-adaptive method. We test the methods on the problem after applying importance sampling to investigate if the performance of the methods are affected by this. The use of variance reduction techniques in local domains and the use of MC and QMC for the evaluation of the sub-problems resulting from the adaption process is to our knowledge not investigated before. The numerical tests show that the adaptive method has good performance compared to non-adaptive methods for many problems. In particular we get good results for out-of-the-money options and general functions having different behavior in different parts of the domain. Furthermore we are able to identify easy implementable criteria for when the adaptive method is preferable over the non-adaptive method.

In sec. 6.2 we describe a finance problem suited as an example to test the algorithms. In sec. 6.3 we generalize this problem by parameterizing it, and describe how to find the introduced parameter for general option pricing problems with the same structure. We show that the introduced parameter can be used to characterize real world problems in sec. 6.4, and we introduce two efficiency measures for the methods in sec. 6.5. In sec. 6.6 and sec. 6.7 we give numerical results, and in sec. 6.8 we elaborate further on the findings by looking at two functions with less resemblance to finance.

6.2 Pricing of Exotic Options

The integrand $f(\cdot)$ in our problem is a function in $C^0([0,1]^s)$ and is, in its initial form, monotone in each variable. The different variance reduction techniques described below can alter the monotonic behavior, but the continuity is maintained. Motivated from applications of finance, we may assume in this paper that the function has only one hyper-plane in which it is not differentiable, and that it in the rest of the domain is in $C^\infty([0,1]^s)$. Note that the method is applicable for more general integrands, but this would involve a more complicated and time consuming routine to investigate the integration domain to insure a proper subdivision. The method used in this paper is tailored for finance problems, and enables us to get proper subdivisions by using relatively few resources.

We have chosen to investigate the pricing problem of an Asian option on multiple underlying assets as an example (called an Asian basket option). The adaptive method is also suited for other option contracts which can be formulated as an integration problem, and is especially successful for problems where it is possible to identify a subset of dimensions/variables for which the integrand is most dependent, and problems where the integrand is zero or constant in parts of the domain. We give a brief description of the Asian basket option pricing problem below.

We operate in the context of a complete, standard financial market with constant risk-free rate r and constant volatility, see e.g. [KS98] for a definition of such a market. A stock in this market is modeled by

$$S_n(t) = S_n(0) \exp\left(P_n + \sum_{d=1}^{N} \sigma_{nd} W_0^d(t)\right),\ n = 1, \ldots, N\ , \tag{6.1}$$

where σ is the volatility matrix of the N assets in the market and the stochastic process $W_0(t)$ is an N dimensional Brownian motion under the risk free measure. The term P_n is introduced for easier notation and is given by $(r - \frac{1}{2}\sum_{d=1}^{N} \sigma_{nd}^2)t$. Without going into details (which can be found in [KS98]), we state that the value at time t of an Asian option is

$$V(t) = e^{-r(T-t)} \mathrm{E}_0[\varphi(\Upsilon(\mathcal{T}))|\mathcal{F}(t)]\ , \tag{6.2}$$

where $\varphi(\cdot)$ is a Borel measurable function given by $\varphi(Y) = (Y - q)^+$, resulting in the European-style Asian option with strike q. Furthermore we are interested in $V(t)$ at

$t = 0$, i.e. at the time the option is bought. The argument $\mathcal{T}$ is a vector of discrete points in time $\mathcal{T} = (t_0, t_1, \ldots, t_K)$, $t_K = T$, where the number of sampling points $K + 1$ is specified. The length of the intervals $t_k - t_{k-1}$ need not be equal. For a discussion of an even more general setup, see e.g. [RS95]. The Υ function for the Asian basket option is given by

$$\Upsilon(\mathcal{T}) = \frac{1}{N(K+1)} \sum_{k=0}^{K} \sum_{n=1}^{N} S_n(t_k) \; . \tag{6.3}$$

The price of the Asian basket option thus becomes:

$$V(0) = e^{-rT} \mathrm{E}_0[\varphi\big(\Upsilon(\mathcal{T})\big)] \; . \tag{6.4}$$

In order to replace the expected value in (6.4) with an integral over $[0,1]^D$, where $D = N(K+1)$, and to be able to solve the integral effectively with QMC methods and apply the adaptive approach, we must formulate $\varphi\big(\Upsilon(\mathcal{T})\big)$ in terms of independent stochastic processes. This can be done by constructing the σ matrix from an SV-decomposition of the covariance matrix of returns for the assets over some period, and by performing an SV-decomposition of the covariance structure of the discretized Brownian motion to simulate the $W(t_k)$ process. See e.g. [Øks98] on the Brownian motion or [DB01] for the full treatment with regards to the Asian basket option. We then obtain a formulation where we have identified the eigenvalues of σ^2, and are able to give priority to the parts of the problem corresponding to the biggest eigenvalues. The unveiling and ordering of the eigenvalues leads to a formulation where the integrand is most volatile along the first dimensions, with decreasing volatility for consecutive dimensions. Note that there are other approaches to achieve a formulation in terms of independent stochastic processes for path dependent options (see [DB01]), but the method described here is the only one for which we are able to identify the most important sources of noise for the full problem, and therefore the best suited model for the adaptive approach. To expose the independent stochastic processes we introduce the notation $S(t_k, \varepsilon_k)$. The price of the option at $t = 0$ can then be written

$$V(0) = e^{-rT} \int_{\mathbb{R}^D} \varphi\Big(\frac{1}{D} \sum_{k=0}^{K} \sum_{n=1}^{N} S_n(t_k, x)\Big) \psi(x)\, dx \tag{6.5}$$

$$= e^{-rT} \int_{[0,1]^D} \varphi\Big(\frac{1}{D} \sum_{k=0}^{K} \sum_{n=1}^{N} S_n(t_k, \varepsilon(y))\Big)\, dy \; , \tag{6.6}$$

where $\psi : \mathbb{R}^D \to \mathbb{R}^D$ is the density of a D - dimensional centered Gaussian random variable with covariance matrix equal to the identity, and $\varepsilon : [0,1]^D \to \mathbb{R}^D$ is a vector of inverse cumulative normal distribution functions with mean 0 and variance 1: $\varepsilon(y) = \langle \varepsilon_1(y_1), \ldots, \varepsilon_D(y_D) \rangle$. We do the evaluation of $\varepsilon(\cdot)$ by a rational approximation suggested in [Mor95]. The final problem formulation in (6.6) is suitable for the adaptive algorithm. Note that the independent stochastic variables in the problem represents the coordinate axis of the integration problem, and the identification and

ordering of the orthogonal noise components leads to a problem formulation where the integrand is "most active" in an area of the domain as small as possible. Furthermore the activity is localized to as few quadrants as possible, which also are neighbors. These characteristics are all favorable to the adaptive method. In the next section we will look at how the eigenvalues of the single asset Asian problem is altered by changing the time discretization, and use this to formulate a more general test problem suited for testing the performance of the adaptive method for general option pricing problems.

6.3 Parameterizing the Noise Structure of the Option Problem

We know from previous work in e.g. [BEG91], and from tests of our own algorithm, that the adaptive approach is not always competitive with conventional MC and QMC methods. The integrand need to have special characteristics as the dimension becomes high for this to be the case. The adaptive method should therefore not be used incautious. We want to formulate criteria for deciding when the adaptive method is competitive with conventional QMC for general option pricing and hedging problems. These problems are often on a similar form as the problem described in sec. 6.2, and the model of the underlying assets for a range of problems can generally be written as

$$S_n = S_n \exp\left(P_n + C_n\varepsilon\right) ; 1 \leq n \leq N \ , \tag{6.7}$$

where C_n is the n'th row of a positive semidefinite matrix C and ε is vector of independent normal distributed variables with length N. The matrix C can be constructed by setting $C = \Sigma\Lambda^{1/2}$, where Λ is the eigenvalue matrix of C^2, and Σ is the matrix of eigenvectors. The matrix C^2 is a parameter depending on the market and the type of option to be priced. In the single asset Asian option problem C^2 is only depending on the form of the time discretization, and we can therefore use this problem to alter the contents of C, and to find a criterion based on the eigenvalues of C^2 to measure the relative efficiency of the adaptive approach.

We have formulated a functional behavior of the time discretization as

$$f(t,a) = t^a \ , t \in [0,1] \text{ and } a \in [0.001, 1] \ . \tag{6.8}$$

When the time discretization given by this function is applied to the pricing of the asian option, each value of a give a different set of eigenvalues and eigenvectors for the matrix C^2. In fig. 6.1 we have plotted the square root of the normed eigenvalues in descending order for a typical configuration together with the function

$$l(i) = i^b, \ i = 1, \ldots, D, \ b \leq 0 \ , \tag{6.9}$$

where the parameter b is found by a least squares fit of the function to the square root of the eigenvalues in descending order. As can be seen from the graphs, the fitted

function is rather similar to the graphed square root of the eigenvalues. To extend the discussion to the range of $b > -1.37$ in (6.9) not captured by the adjustment of a in (6.8) we have constructed an artificial parameterized problem class where we start out with the Asian option problem with D equidistributed time discretization points, and the parameter a equal to 1. The resulting C matrix of this problem is changed by exchanging the Λ matrix with a new $\hat{\Lambda}$ matrix where the diagonal elements are $\lambda_1 l(i)$, λ_1 being the largest eigenvalue of the original C^2 matrix. By simulating on this artificial problem for $b \in [-1.4, 0]$ we can extend the range in which we can evaluate the relative performance of the adaptive method. The reason for switching approach is that we are not able to produce eigenvalue dependence in the range $b < 1.37$ for any a in (6.8). As a is increased above one, b will increase rather than decrease further. In fig. 6.1 we have illustrated the dependence between a and b.

Since option pricing problems often share the structure described above, and the behavior of the square root of the eigenvalues can be matched satisfactorily with the function $l(i)$, the conclusion of our numerical tests can be used to decide when to employ the adaptive method for new problems. The graphed results to illustrate the relative performance of the adaptive method will be done as a function of b rather than a, because the parameter b is the one that can be extracted from general option pricing problems.

6.4 Comparison with Actual Pricing Problems

The parameterization in sec. 6.3 turned out to give a very close relationship between the rate of decline of the square root of the ordered eigenvalues and the function given in $l(i) = i^b$ for a fitted b. To argue that the survey we have performed can be put to use, we give the eigenvalues and the appurtenant b for some real world problems.

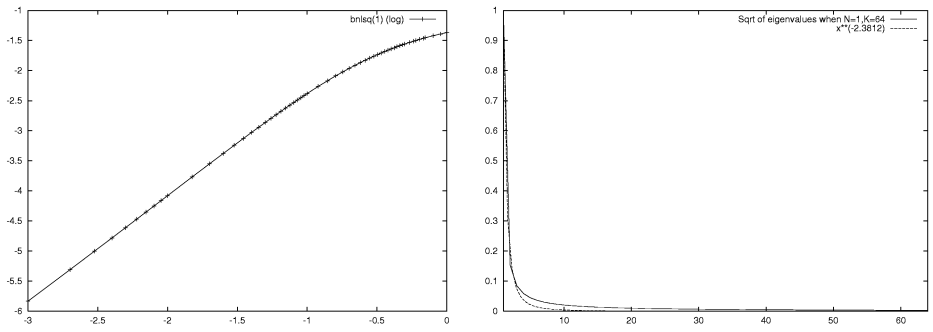


Figure 6.1: In the left plot the parameter b is given as a function of $\log_{10}(a)$. The dependence is close to linear. In the right plot a the square root of the eigenvalues (for $a = 0.1$) and func. 6.9 (for the fitted b) is given for $D = K = 64$ eigenvalues

These are graphed in fig. 6.2. These plots show that for noise structures emerging

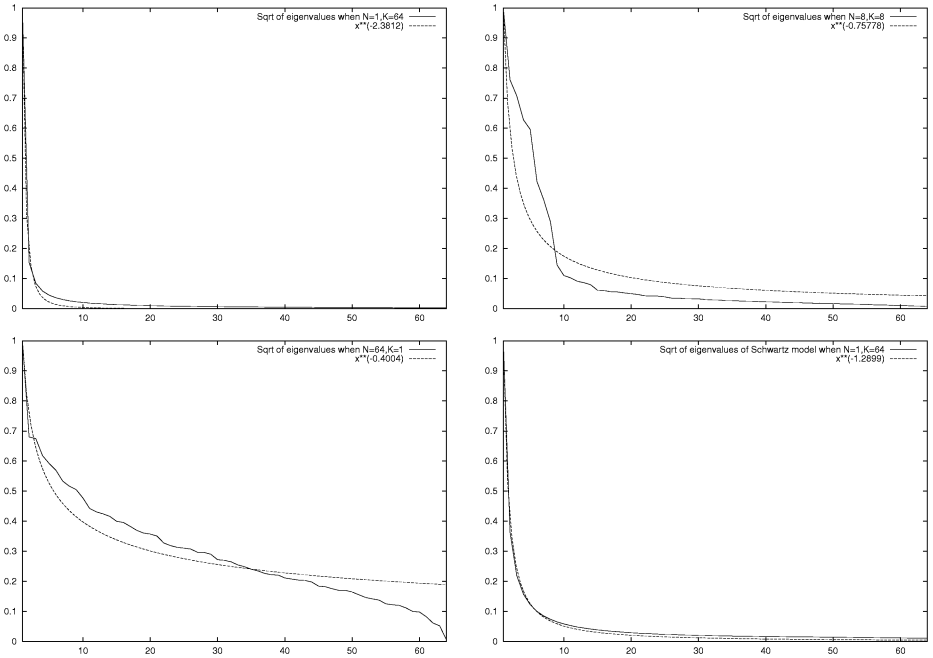


Figure 6.2: The eigenvalues for the noise structure for a set of examples. The first three cases are for the Asian basket option case with $(N = 1, K = 64)$, $(N = 8, K = 8)$ and $(N = 64, K = 1)$. The last example is from an asian option on a forward contract in the energy market, and $(N = 1, K = 64)$. We have used $a = 0.1$ in eq. 6.8 for the propagation of time. As before N is number of assets, while K is number of time steps. We have included the appurtenant least-squares fit of func. 6.9. On the x-axis is the dimension corresponding to the eigenvalues. On the y-axis is the square root of the normed eigenvalues

from a discretized Brownian motion, like in the Asian option cases, the fit is very close. For noise structures extracted from market data, the fit is not that good. It is, however, the decline rate of the largest eigenvalues that influences the convergence rate most, and this shape is captured well by the functional relationship including b. Note furthermore that the eigenvectors should not have a large impact on the convergence rate as they have unit length.

6.5 Efficiency Measures

In order to compare the performance of the adaptive method with the conventional we need suitable measures. We focus on two such measures; one where the extra time consumption in the adaptive approach is accounted for, and one based purely on the difference in convergence speed in terms of variance. In detail I have measured the variance of 100 simulations of the integral by the adaptive and the non-adaptive for each combination of the parameters. The series of estimators for the adaptive method depend both on the distribution of points in each sub-domain from the adaption process and on the ability of the method to sub-divide the domain well. In order for the variance v_a of the 100 samples from the adaptive method and the variance v_n of the conventional method to be comparable, the point estimates of the price are calculated by using different values δ for the subdivision accuracy in the adaptive method. This is to avoid using the exact same subdivision of the domain over and over, which would give the adaptive method an advantage when comparing with the non-adaptive. A specification of $\delta = 10^{-3}$ as input to the simulation, therefore results in the adaptive method using $\delta \in [10^{-3}, 2{\cdot}10^{-3}]$ descending linearly as the simulations are performed to force different subdivisions for each estimated value. Furthermore it must be mentioned that the estimators are based on drawings from consecutive parts of the pseudo- and quasi-random number sequences for respectively the MC and QMC methods.

The first measure of efficiency, labeled DE (for definite efficiency) is constructed by the relation

$$\mathrm{DE} = \log_{10}(v_a T_a) - \log_{10}(v_n T_n) \ , \tag{6.10}$$

where T_a and T_n are the computing times for resp. the adaptive and the conventional method. The DE-measure gives us a scale for the relative efficiency of the adaptive method to the conventional, where a negative number indicates preference for the adaptive method, while the same positive number gives an equivalent preference for the conventional. The scale is logarithmic, so a number of -1 would indicate that the adaptive was ten times better than the conventional, and so on. To see that the product of computing time and variance gives a good measure of efficiency for a method, consider the estimator $\hat{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ with X_i i.i.d., $\mathrm{E}[X_i] = X$ and $\mathrm{Var}[X_i] = \sigma_X^2 < \infty$. The central limit theorem asserts that as the number of replications N increases we have

$$\sqrt{N}(\hat{X}_N - X) \to \mathcal{N}(0, \sigma_X^2) \ . \tag{6.11}$$

Suppose that we use a fixed amount of computing time τ for each replication, then the number of replications we can calculate given a computational budget $\mathcal{T}$ is $\lfloor \mathcal{T}/\tau \rfloor$. This gives

$$\sqrt{\mathcal{T}}(\hat{X}_N - X) \to \mathcal{N}(0, \sigma_X^2 \tau) \ , \tag{6.12}$$

which suggests that, asymptotically, we should prefer an estimator having the lowest value of the product $\sigma_X^2 \tau$. It should be mentioned that the above arguments are

valid for truly stochastic variables only, and applicable for pseudo random variables. The same arguments may therefore not be used incautious to crate a measure for QMC methods (which are deterministic). QMC methods and their low discrepancy sequences are constructed with the purpose of improving the convergence beyond the $\sqrt{N}$ rate, An extensive literature on these subjects exist, see e.g. [NX98], [Owe99], [Owe98]. It can be shown that when a low discrepancy sequence is used, the convergence of an integral to its true value is proportional to the star discrepancy of the sequence multiplied by the variation of the integrand in the sense of Hardy and Krause. This is formalized through the Koksma-Hlawka theorem (see e.g. [KN74], [Nie78], [Nie87]). If we assume that the Koksma-Hlawka theorem can be interpreted similarly for the adaptive and the non-adaptive method, we can use the arguments above to justify the use of the DE-measure of efficiency also in the QMC setting. We can not, however, use the DE measure to compare calculation methods with differing underlying sequences. Therefore we will always construct the DE measure with common sequences for the underlying estimators.

The second measure of efficiency we have chosen to present, does not include the computing time. The background for presenting this measure, is that, although we have coded the algorithms as effectively as we could, there could still be ways to make them faster. The second measure therefore serves as comparison between the variances of a set of 100 replications of the adaptive method compared to the conventional. This measure is in the graphs denoted by F (for fractional efficiency). It is calculated as

$$F = \log_{10}(v_a) - \log_{10}(v_n) \ . \tag{6.13}$$

A further aspect worth noting is that the efficiency-measures we have chosen are indifferent of the level of the integral itself: If we in order to normalize the integral, multiply by a factor of d, the variance will change by a factor of d^2. The estimators, however, are not affected since $\log_{10}(d^2 v_a) - \log_{10}(d^2 v_n) \ = \log_{10}(v_a) - \log_{10}(v_n)$. Therefore the changes in the measures are due to the change of parameters, not other trivial effects. This is important, because the value of the integral will change as we change parameters of the option contract, and changes of parameters have to be done to reveal the characteristics of the integration methods.

6.6 Numerical Results of the Parameterized Problem

In this section we will draw attention, and try to explain the main findings of the numerical simulations. We have simulated the price of the options by using different configurations in order to compare the performance of the adaptive algorithm with the conventional[2] non-adaptive. For the market, we have used a risk free rate of

[2]Note that the label conventional is here used on a highly sophisticated method using principal component analysis (SVD) to find the most important dimensions of the noise sources combined, and using a low discrepancy sequence as uniform number generator

0.05 and a volatility of 0.5331 for the asset in the single asset Asian option (the volatility of Kværner[3] in 2000). The varied parameters of the model are the number of time discretization steps (K): $\{2, 64\}$, moneyness (m): $\{-25, 0, 25, 50\}$, sequence generator: (method) {H,M} (Halton Leaped, Mersenne Twister), Control variate technique: (CV) {Y,N} (yes, no), importance sampling: (IS) {Y,N} (yes,no), the parameter a in eq. 6.8: (a): $[0.001, 1]$ – leading to a decline rate for the square root of the eigenvalues in the range $[-5.88, -1.37]$ (see sec. 6.3 on the relation between a and b and how we find an extended range for b). In the figures these different parameters are varied and plotted against each other. We have also varied the variance estimator limit δ, which controls the subdivision coarseness, (δ): $\{10^{-2}, 10^{-3}, 10^{-4}\}$. Note that when the DE - efficiency measure is negative, the adaptive method has the best performance. In the relevant plots, we have included the horizontal line at 0 to ease the reading of the graphs.

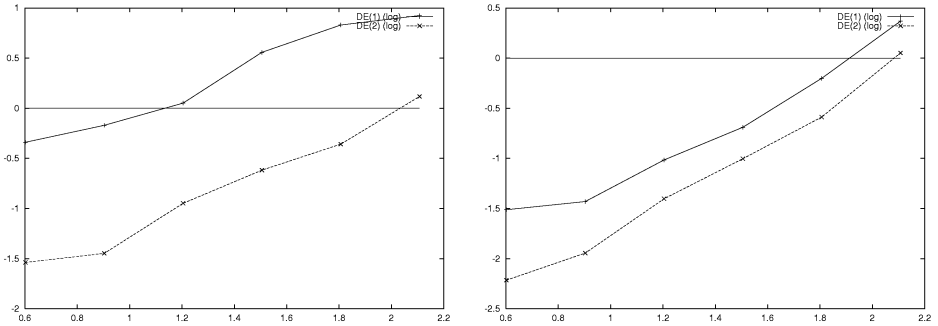


Figure 6.3: Plot of DE-efficiency vs. $\log_{10}(D)$ (dimension) for $b = -1.55$ (1) and $b = -2.38$ (2) (equivalent to $a = 0.5$ and $a = 0.1$). In the left figure is the Halton Leaped sequence used, while the Mersenne Twister is used in the right

First of all, the dimension of the problem is important for the relative performance, and the ability of the adaptive method to outperform is better for lower dimensions (fig. 6.3). Furthermore the adaptive method is relatively better for problems where the eigenvalues decline rapid, and for out-of-the-money options (figs. 6.4, 6.5). These findings are connected, and the explanation for them are twofold: first it is evident that the adaptive method is able to exploit that the integrand has value zero in large areas of the domain. This explains why out-of-the-money option are easier for the adaptive method than in-the-money options (fig. 6.8). A rapid declining rate of the eigenvalues tells us that the contribution to the variance of the integral is limited to a small set of dimensions of the integrand, and the adaptive method is able to exploit this. Further arguments for the connection between the effective dimension of the problem and the behavior of the eigenvalues can be found in [DB02] where

[3]Kværner is one of the largest companies in the main index of Oslo stock exchange, and had a rather high volatility in 2000

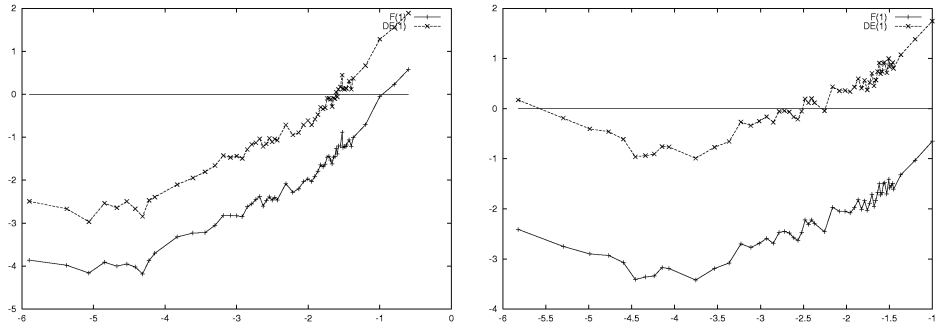


Figure 6.4: Plot of both efficiency measures (F and DE) vs. $\log_{10}(b)$ (eigenvalue decline rate) for a typical case. In the left figure $D = 16$ and in the right $D = 128$

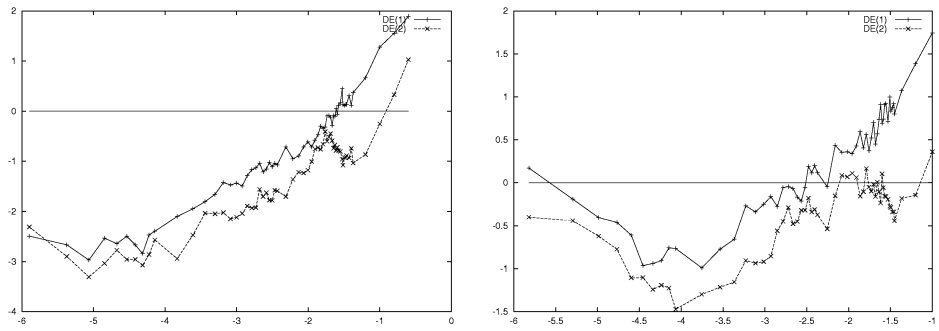


Figure 6.5: Plot of DE-efficiency vs. b for different moneyness – (1) corresponds to moneyness = 0, while (2) corresponds to moneyness = 50. In the left figure $D = 16$ and in the right $D = 128$

we have discussed and measured the relation between the eigenvalues of the noise of the problem and the number of contributing dimensions, and concluded with a close relationship.

Next we turn to the question; why does the ability to outperform decrease with increasing dimension (fig. 6.3)? The answer to this lies in the curse of dimensionality: The hyper-surface of the cube increases as the dimension increases while the volume is constant equal to 1. This results in the fraction of volume where the integrand has value zero decreases as the dimension of the problem increases if the fraction of important dimensions does not decrease. This is because the fraction of the volume lying close to the borders of the domain is increasing as the number of dimensions increases. To concretize; the volume fraction of the $[0.01, 0.99]^2$ cube is about 0.96 relative to the $[0, 1]^2$ cube, while the volume fraction of the $[0.01, 0.99]^{64}$ cube relative

to the $[0,1]^{64}$ cube is about 0.27. The same argument in dimension 1024, which is a realistic case with 32 assets and 32 time discretization points, gives a fraction of about 10^{-9}. The nested exponential functions in the integrand in option pricing problems results in dominance for a variable as it goes to the border, so by the argument above we have a situation where the adaptive method looses an important advantage as the number of contributing dimensions increase.

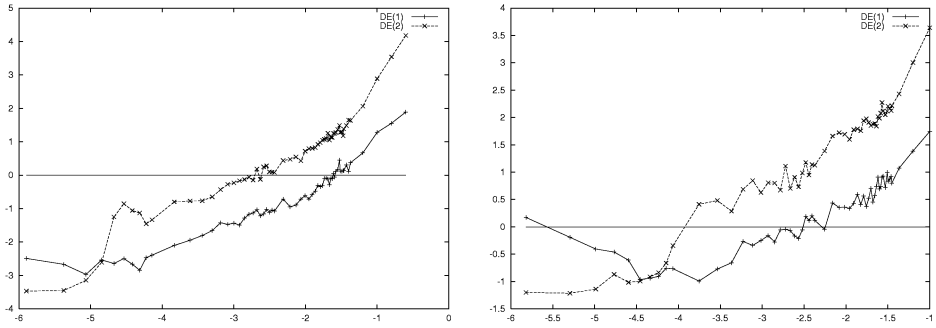


Figure 6.6: Plot of DE-efficiency vs. b **with** the control variate technique (label (2)) and **without** (label (1)). In the left figure $D = 16$ and in the right $D = 128$

The local control variates technique enhance the performance only for problems having the most rapid decline of the eigenvalues, (fig. 6.6). This is probably due to the failure of generating a polynomial function behaving like the original function for the option pricing integrand, even for very local domains, as the dimension is increased. Note, however, that for less complicated integration problems we have experienced that the control variate technique gives a huge performance gain. In sec. 6.8 we give some numerical results showing this.

There can be various reasons for choosing a plain MC method when pricing options, and the adaptive method can be used in this setting as well. The findings of our numerical tests are that the relative performance of the adaptive method is better in the MC setting than in the QMC setting (fig.6.7). Note, however, that the QMC estimators gives the most accurate results for both the adaptive and the conventional method, but because the non-adaptive MC method has considerably lower performance than the corresponding QMC method the relative performance is better in the MC setting. See e.g. [Dah00] or [DB01] for illustrations on the performance differences between MC and QMC.

It is interesting to note that as the variance estimator limit C, which controls the subdivision coarseness, is changed, the measures of relative efficiency is changed in a different way when the Halton sequence is used and when the Mersenne Twister sequence is used. In the QMC setting the relative efficiency only changes for the set of problems with a very high decline rate ($b < -4.5$), while it gets better for all problems in the MC setting. The second observation reveals that the adaptive method gains

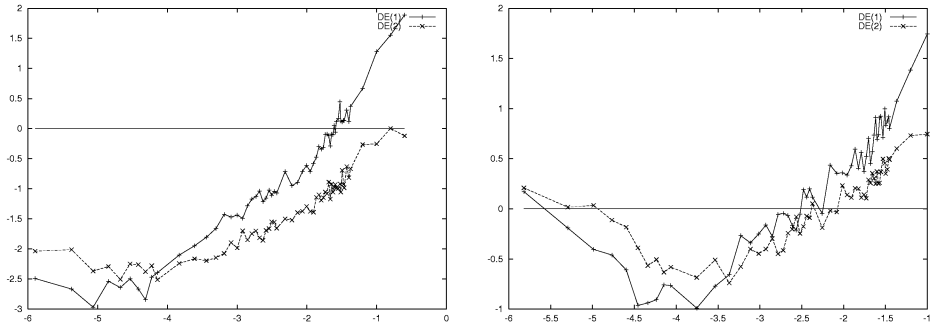


Figure 6.7: Plot of DE-efficiency vs. b by using different uniform number generators – (1) corresponds to the Halton Leaped sequence, while (2) corresponds to the Mersenne Twister. In the left figure $D = 16$ and in the right $D = 128$

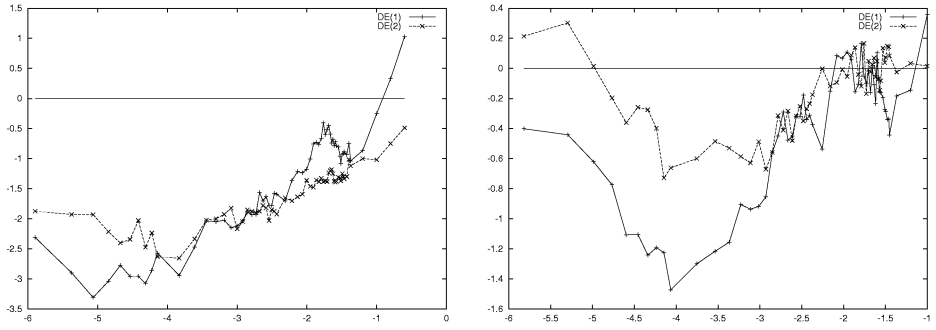


Figure 6.8: The same comparison as in fig. 6.7, but now we do the plots for moneyness $= 50$ rather than moneyness $= 0$. In the left figure $D = 16$ and in the right $D = 128$

more from adding computational work than the non-adaptive when a pseudo random sequence is used, while the first observation means that this is only the case for the QMC setting for $b < -4.5$. For $b > -4.5$, the relative efficiency is the same. An illustration of these results is given in fig. 6.9.

To investigate the effect of applying importance sampling, we have to modify the efficiency measures. The background for this is that by applying importance sampling we modify the integrand in such a way that the adaptive method will act differently on the problem. This results in a different subdivision and a different number of simulation points than for the simulation on the non-altered problem. The DE-measure is a compound measure and can still be used in the importance sampling (IS) setting to find the effect of applying the adaptive method to an IS modified problem, but we must use less compound measures if we want an answer to whether we have bene-

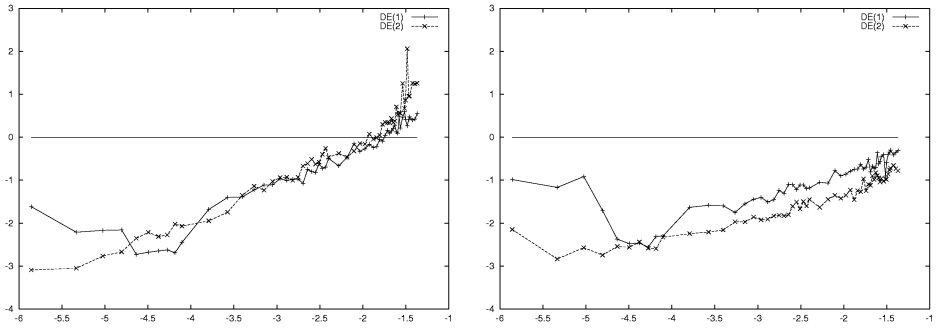


Figure 6.9: Plot of DE-efficiency vs. b for $\delta = 10^{-3}$ in label (1) and $\delta = 10^{-4}$ in label (2). The dimension of the problem is $D = 32$. In the left figure the Halton Leaped sequence is used, while in the right the Mersenne Twister

fitted by applying importance sampling in the first place. This situation is different from what was the case when applying the control variate technique, which only affected the treatment of the integrand in the adaptive setting. We therefore introduce the non-relative measures $E_a = \log_{10}(Lv_a)$ and $E_n = \log_{10}(Lv_n)$ for respectively the adaptive and non-adaptive method. They can be used to compare the effects of changing parameters and variance reduction techniques, but not to compare the effect of changing method. The main finding is that the performance of the adaptive method is virtually unchanged by applying importance sampling. The case for the non-adaptive method, however, is different. Here the insertion of the IS technique has a positive effect for $D \leq 32$. Furthermore, the variance of the estimator did not increase by applying importance sampling for any of our simulations. The time consumption, however, increased as a consequence of the fixed-point iteration. It is interesting to note that the adaptive method had the ability to neutralize the effect of importance sampling. Note that the importance sampling technique, which is applied to the problem globally, is only practical for integration problems of probability density functions and can not be applied to the functions in sec. 6.8.

6.7 Numerical Results of Actual Pricing Problems

In sec. 6.4 we found that many option pricing problems have an eigenvalue structure that can be imitated by the simple function $l(i) = i^b$, for some b. In this section we are going to simulate the value of a European basket option with $N = 8$ underlying assets to check if the resulting values for (b, DE) are in accordance with the ones we have predicted by simulating on the corresponding parameterized Asian option problem with $K = 8$ time discretization points. The results are given in fig. 6.11, and shows a good match.

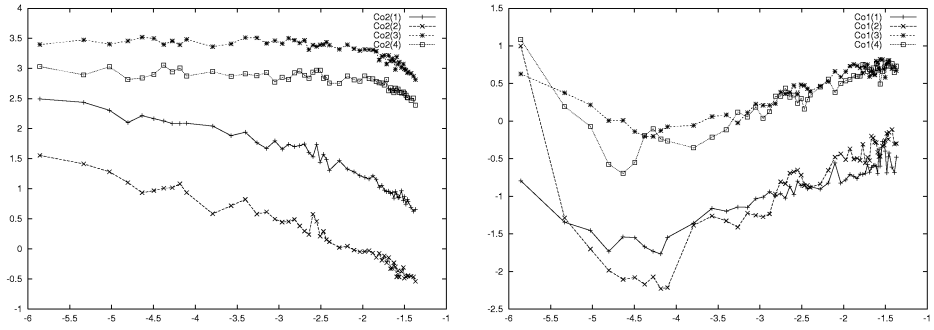


Figure 6.10: Plot of E_n-efficiency vs. b for respectively no importance sampling and QMC (1), importance sampling and QMC (2), no importance sampling and MC (3) and importance sampling and MC(4). In the right plot is the corresponding graphs for the measure E_a (for the adaptive method). The dimension is $D = 32$

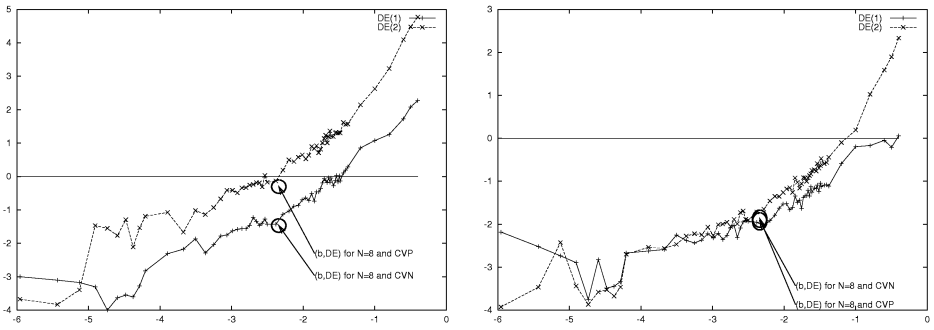


Figure 6.11: Plot of DE-efficiency vs. b **with** the control variate technique in label (2) and **without** in (1). Included in the figures are the coordinates of a basket option problem simulated with the same methods. The dimensions of the problems are $D = 8$. In the left figure the Halton Leaped sequence is used, while in the right the Mersenne Twister

6.8 Additional Test Functions

We have included two additional test functions, with less resemblance with the option pricing problem. This is to show that the control variate technique can be a valuable method for various integration problems, even if it did not perform very well in general for the option pricing problems. Also for these functions a parameter b is added in order to control fluctuations of the integrand for increasing dimensions. Note, however, that there are no close relationship between the interpretation of b for these functions and for the problem in sec. 6.3, except that it is used to parameterize

the behavior of the integrand in different dimensions. The functions are given by

$$f_1(x) = \sum_{i=1}^{D} i^b (1 + \mathrm{e}^{(\pi x_i)}),\ x \in [0,1] \tag{6.14}$$

$$f_2(x) = D \exp\left(\frac{1}{D} \sum_{i=1}^{D} i^b 6 x_i\right),\ x \in [0,1]\,, \tag{6.15}$$

where $b \in [-4,-1]$. Simulations on these functions shows that the control variate technique enhance the adaptive method considerably (figs. 6.12 and 6.13), and that the gain from applying CV is increasing as b is decreasing.

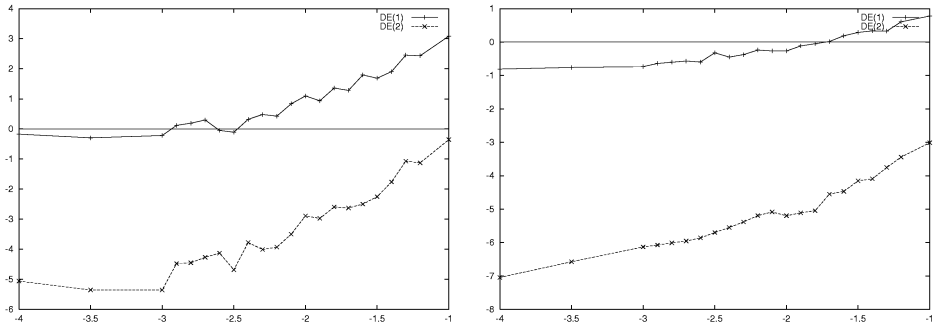


Figure 6.12: Plot of DE-efficiency vs. b **with** the control variate technique in label (2) and **without** in (1) for the function f_2. The dimensions of the problem is $D = 32$. In the left figure the Halton Leaped sequence is used, while in the right the Mersenne Twister

6.9 Conclusion and Discussion

It is rather obvious that an adaptive QMC integration method for some problems will outperform a non-adaptive. And the integrands for which this will be the case have to be sufficiently nice in parts of the integration domain to allow the more sophisticated method to come out ahead. We wanted to find the set of option pricing problems where this was the case, and criteria for pinpointing these problems in advance. The criteria were formulated through a parameterized function fitted to the sorted eigenvalues in descending order. Furthermore we established an indirect method for altering the parameter. This gave us an apparatus to investigate the performance of the adaptive method compared to the non-adaptive for different parameters of the problem, and for various sophistication levels of the methods. The major part of this article focuses on presenting and elaborating on these results. In the companion article [Dah01] the method itself and the sophistication levels are developed. We do not include any

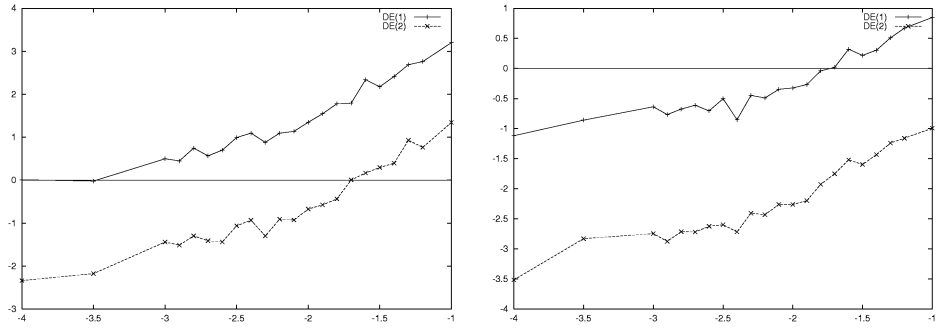


Figure 6.13: Plot of DE-efficiency vs. b **with** the control variate technique in label (2) and **without** in (1) function for f_1. The dimensions of the problem is $D = 32$. In the left figure the Halton Leaped sequence is used, while in the right the Mersenne Twister

of the detailed findings in this section, as most of the article is dedicated to this, but limit ourself to stating that we have found that the adaptive method can give considerable performance gains for many problems. In general it will be the method of choice for low dimensional problems, and for higher dimensional problems it can give considerable performance gains for problems with certain characteristics. It should, however, not be used incautious for high dimensional problems, since the performance also can deteriorate for problems not having these characteristics.

It should be mentioned that a vast set of simulations has been performed, summing up to about 6000 hours of computing time on a cluster of PC's with average CPU speed of about 1Ghz. Additional and deepening results can therefore be obtained by contacting the author.

Acknowledgements

The author is grateful for financial support from Storebrand Investments. I also want to thank F. E. Benth for valuable talks and for applying amendments to the manuscript.

Chapter 7

On Derivatives of Claims in Commodity and Energy Markets Using a Malliavin Approach[1]

Author(s): Lars O. Dahl, Fred E. Benth, Kenneth H. Karlsen

Keywords: Options in commodity and energy markets, sensitivity measures, hedging, Malliavin derivative, quasi-Monte Carlo simulation

Abstract: In this paper we investigate the recently introduced Malliavin approach compared to more classical approaches to find sensitivities of options in commodity and energy markets. The Malliavin approach has been developed in the paper [FLL$^+$99] and [FLL$^+$01]. In commodity and energy markets, some special dynamics for the underlying security and some new products different from Black & Scholes markets are encountered. In addition to investigating the numerical values of the expressions by conventional Monte Carlo (MC) and quasi-Monte Carlo (QMC) methods, we apply an adaptive approach developed in the papers [Dah01] and [Dah02a]. This adaptive method is also applied to the so called Localized Malliavin approach developed in the paper [FLL$^+$01]. The numerical results show that we can get substantial variance reduction in estimators by choosing sophisticated methods for the simulations, and that the Malliavin approach is a very powerful tool for formulating the sensitivity estimators.

7.1 Introduction

In commodity and energy markets, the underlying product of a derivative contract may be either the spot or the forward/futures contract on the spot. A much used model for spot prices in commodity and energy markets is Schwartz' mean-reverting

[1] In preprint as working paper May 2002. Full reference: [DBK02]

model (see [Sch97] and [KH00]). Formulated in a risk-neutral world it has the dynamics,

$$dS(t) = \alpha(\mu - \lambda - \ln S(t))S(t)\,dt + \sigma S(t)\,dW(t)\ . \tag{7.1}$$

Here, α is the mean-reversion rate, σ the volatility, e^{μ} the long-term level for the spot price and λ the market price of risk. $W(t)$ is a standard Brownian motion defined on a complete probability space $\big(\Omega, \mathcal{F}, P\big)$, where $\mathcal{F}_t$ is the augmentation with respect to P of the filtration generated by W, $0 \le t \le T < \infty$. If we introduce $\gamma = \alpha(\mu - \lambda) - \sigma^2/2$, we may write $S(t) = \exp(X(t))$ for the Ornstein-Uhlenbeck process

$$dX(t) = (\gamma - \alpha X(t))\,dt + \sigma dW(t)\ , \tag{7.2}$$

with $X(0) = \ln x, S(0) = x$. The process $X(t)$ has an analytical expression

$$X(t) = \mathrm{e}^{-\alpha t}\ln x + \gamma(1 - \mathrm{e}^{-\alpha t}) + \sigma\int_0^t \mathrm{e}^{-\alpha(t-s)}\,dW(s)\ .$$

This analytical expression is useful when calculating derivatives of claims.

Prices of forward instruments can be derived in an arbitrage-free way from the spot price (see e.g. [CS00, Ben02, Pil98]). However, motivated from the Heath-Jarrow-Morton approach in interest rate theory, one may instead write down the risk-neutral dynamics of the forward price directly. We assume the dynamics of the forward contract on the spot is given in the risk-neutral world as

$$dF(t,T) = \sigma(t,T)F(t,T)\,dW(t), \quad F(0,T) = x(T)\ , \tag{7.3}$$

where $x(T)$ is todays forward curve. We assume σ is an integrable function such that $\int_0^T \sigma^2(t,T)\,dt < \infty$, which means that $t \to F(t,T)$ is a martingale. An explicit representation of $F(t,T)$ is

$$F(t,T) = x(T)\exp\big(-\frac{1}{2}\int_0^t \sigma^2(s,T)\,ds + \int_0^t \sigma(s,T)\,dW(s)\big)\ . \tag{7.4}$$

A frequently used volatility structure in commodity and energy markets is given by

$$\sigma(t,T) = \sigma\mathrm{e}^{-\alpha(T-t)}\ . \tag{7.5}$$

This specification is motivated from the mean-reverting model of Schwartz for spot prices, which implies this volatility structure for the forward price (see e.g. [CS00, Ben02]). When considering claims on spot prices, the dynamics in (7.1) will be assumed. When we on the other hand analyze claims on the forward, we use the forward dynamics (7.3) with volatility structure as in (7.5).

We focus in this paper on derivatives of the price of different claims with respect to different parameters in the underlying. In sec. 7.2 we consider derivatives of European options on spot prices of the commodity and energy market. In particular we find expressions for delta, gamma and vega. In sec. 7.3 we find the derivatives of European options on forward prices. The same parameters as for the spot are calculated. We

advance in sec. 7.4 by looking at path dependent options, in particular the European-style arithmetic average Asian option, which produces a multidimensional problem. The sensitivities we find by differentiating the prices of such claims are extensively used in the process of hedging contracts of this type. Practicians need to have a well developed intuition of the dependence of their position on the movements and events in the market, and a range of literature give interpretations of the parameters. See e.g. [Hul97], [Ave97], [Tal97].

The derivatives can be expressed in various ways, depending on how they are deduced and the assumptions made in the deduction. We will mainly focus on the Malliavin approach recently introduced in the papers [FLL$^+$99] and [FLL$^+$01]. The approach uses the Malliavin derivative together with properties of the Skorohod integral to produce formulas for the derivatives of options. These formulas are expressed in terms of the expectation of the option's payoff multiplied with some random variable which is (usually) a function of the underlying. A neat feature of the Malliavin approach is that this random variable is not dependent on the actual option (that is, f), but on the underlying product. This means that Monte Carlo based algorithms for numerically evaluating derivatives can be made for general options, and not specifically for each option. This is in contrast to the direct method (also called infinitesimal perturbation analysis (IPA)), where one must deduce individual expressions for each payoff function and underlying contract since the derivative is expressed in terms of the differential of the payoff function. See e.g. [Gla91] for an overview of the direct method or [BG96] for deduction of sensitivity expressions in the geometric Brownian motion (GBM) setting. Another conventional method frequently used is the so-called density approach, which relies on the existence of a density of the underlying product. This method expresses the derivative in terms of the option's payoff multiplied with a random variable, very much similar to the Malliavin approach. However, the density method deduces *one* such random variable, while the Malliavin method provides a flexible class of variables. Also, when for instance the option is of Asian-type, there exists no density, and the density approach fails. The Malliavin approach handles this type of products, demonstrating the flexibility of the method.

To compare the Malliavin approach and the alternative methods, we deduce sensitivities by the direct method and the density approach, whenever these methods can be used. In this way we are able to illustrate both the flexibility gained with the Malliavin approach in that it can be applied with success where other methods fail, and to investigate how different numerical methods apply to the different derivative approaches. A further improvement of the Malliavin approach is the localized version introduced in [FLL$^+$99, FLL$^+$01]. The Localized Malliavin approach uses the Malliavin methods around point where the payoff function is not smooth, and direct method outside.

The contribution of the paper is twofold. First, we present formulas for derivatives of options in commodity and energy markets based on the Malliavin approach and compare these with the corresponding expression found by conventional methods. A large portion of the present paper consists of such formulas. Secondly, we investigate effective numerical methods for estimating the sensitivities based on the different formulas

derived in the first part. We have implemented a quasi-Monte Carlo method based on the Halton[2] low discrepancy sequence as a basis for this exploration. Furthermore we have adjusted an adaptive method developed by the authors in [Dah01] and [Dah02a] to the current problem, resulting in an adaptive QMC method. The numerical tests are performed both with and without the adaptive method in order to investigate the effect of applying this. Furthermore we investigate the difference in numerical stability and convergence speed for the estimators deduced by the three approaches; Malliavin, Localized Malliavin and Forward Difference. We know from previous work that the adaptive method is able to perform very well for low dimensional integrals, and the results of this paper show that also in the current setting it gives enormous speedup for many of the problems. The numerical results furthermore verify that the Malliavin approach is the best alternative for finding sensitivities when the payoff function of the option is discontinuous. Our numerical results, somewhat surprisingly, also show that the Local Malliavin approach does not give an estimator with lower variance than the Malliavin approach, but almost identical. However, we have not numerically tested the Localized Malliavin approach to calculate the gamma, and it is likely that the Localized Malliavin approach is able to perform better in this setting.

We emphasize that the adaptive method in this context is not "competing" with the Malliavin approach, but a supplement used to refine the use of Monte Carlo sampling points also in the Malliavin and Local Malliavin context. The results of the simulations are collected in sec. 7.7.

7.2 Derivatives of European Options on Commodity and Energy Spots

Consider a European option with maturity T and payoff $f(S(T))$, where $f \in L^2(\mathbb{R})$ and $\mathbb{E}\big[f(S(T))^2\big] < \infty$. The price of the option is

$$u = \mathrm{e}^{-rT}\mathbb{E}\big[f(S(T))\big] \; . \tag{7.6}$$

Recall that the spot is formulated directly in the risk-neutral setting. We shall use the notation $u(x)$ when we consider the price as a function of the strike spot price x, and $u(\sigma)$ when we consider the price as a function of the volatility σ. For simplicity, we will assume throughout the rest of the paper that the risk-free interest rate is zero, i.e. $r = 0$.

We are interested in calculating the delta, $u'(x)$, of the option price. Before we go to the Malliavin deduction, let us present the direct approach, also called infinitesimal perturbation analysis (IPA). Provided we can move differentiation into the expectation (sufficient conditions for this are given in e.g. [L'E91], [Gla91], [BG96]), we can

[2]The Halton sequence was first presented in [Hal60]. In this paper we are using an extension of the Halton sequence denoted the Halton leaped sequence. It was presented in [KW97], together with good leap values. We have used the leap value 31 in the numerical experiments.

simply write the delta as

$$\begin{aligned} u'(x) &= \mathbb{E}\big[f'(S(T))\frac{dS(T)}{dx}\big] \\ &= \mathbb{E}\big[f'(S(T))S(T)\frac{\mathrm{e}^{-\alpha T}}{x}\big] \ . \end{aligned}$$

When $f(x) = (x-K)^+$, we have $f'(x) = \mathbf{1}_{x>K}$. This method however gets into trouble if f is discontinuous, for example $f(x) = \mathbf{1}_{x>K}$. By the same argument, the direct method is not suited if we want to find the second derivative (e.g. gamma) of this f. Furthermore, the algorithmic treatment of this approach is depending on the specific payoff function f, resulting in individual implementations for each payoff function and each instrument. The method is therefore not very flexible in this context. As we shall see, The Malliavin approach circumvent all these limitations. An other approach that has the ability to circumvent this, but gives a bias, is the finite difference (FD) method. The derivative is then simply found by the estimator

$$\frac{u(x+h)-u(x)}{h} \ ,$$

where we use the same Brownian trajectories for both function evaluations to reduce the variance of the estimator. The parameter h should be small to reduce the bias of the estimator, but a smaller h results in an estimator with larger variance. We use h in the range $[0.1, 0.001]$ percent of x. See [BG96] for a discussion on finding the optimal h for the FD estimator.

Next we turn to the Malliavin approach. However, before we can state the propositions on the sensitivities , we need to introduce the Malliavin derivative and state some useful properties of the Skorohod integral.

7.2.1 Some Results from the Malliavin Calculus

Let $\mathcal{C}$ be the set of cylinder functions on the probability space, e.g. the set of random variables of the form

$$G = g\big(\int_0^\infty h_{1(t)}\,dW(t), \ldots, \int_0^\infty h_{n(t)}\,dW(t)\big) \ ,$$

where $g \in \mathcal{S}(\mathbb{R}^n)$, the Schwartz space of rapidly decreasing and infinitely differentiable functions on $\mathbb{R}^n$, and $h_i \in L^2(\Omega \times \mathbb{R})$. The Malliavin derivative of $G \in \mathcal{C}$ is the process $D_t X$ defined as

$$D_t G = \sum_{i=1}^n \frac{\partial g}{\partial x_i} h_i(t) \ .$$

Introducing the Banach space $\mathbf{D}^{1,2}$ as the completion of $\mathcal{C}$ with respect to the norm

$$\|G\|_{1,2}^2 = \mathbb{E}\big[G^2\big] + \mathbb{E}\big[\int_0^T (D_t G)^2\,dt\big] \ ,$$

we can extend D to be a closed linear operator defined in $\mathbf{D}^{1,2}$. If Y is an Itô-integrable process, then the Malliavin derivative of $\int_0^T Y(s)\,dW(s)$ is

$$D_t \int_0^T Y(s)\,dW(s) = Y(t)\mathbf{1}_{t<T} \ .$$

Furthermore, if $Y \in \mathbf{D}^{1,2}$ and g is a continuously differentiable function with bounded derivative, then $g(Y) \in \mathbf{D}^{1,2}$, and the chain rule holds for the Malliavin derivative:

$$D_t g(Y) = g'(Y) D_t Y \ .$$

We proceed with some results on the Skorohod integral, a stochastic integral for a class of anticipating stochastic processes $Y(t)$ which we denote $\int Y(s)\,\delta W(s)$. It is defined as the adjoint operator of D in the following manner: Let Y be a stochastic process. Then Y is said to be Skorohod integrable if for any $G \in \mathbf{D}^{1,2}$ we have

$$\mathbb{E}\big[\int_0^T Y(t) D_t G\,dt\big] \leq C\|G\|_{1,2} \ ,$$

where C is a constant depending on Y. The Skorohod integral of Y, $\int Y(s)\,\delta W(s)$, is defined by the following duality relation: For any $G \in \mathbf{D}^{1,2}$

$$\mathbb{E}\big[G \int_0^T Y(t)\,\delta W(t)\big] = \mathbb{E}\big[\int_0^T Y(t) D_t G\,dt\big] \ .$$

We state two basic properties of the Skorohod integral, which will be used frequently in what follows. The first proposition tells us that Skorohod integration is a true generalization of the Itô integral:

Proposition 7.2.1. *Let Y be an Itô-integrable stochastic process. Then, Y is integrable in the sense of Skorohod and*

$$\int_0^T Y(t)\,\delta W(t) = \int_0^T Y(t)\,dW(t) \ .$$

The Skorohod integral possesses an integration-by-parts property:

Proposition 7.2.2 (Integration-by-parts). *Let $G \in \mathbf{D}^{1,2}$ be an $\mathcal{F}_T$-adapted random variable. Then, for any Skorohod integrable stochastic process Y*

$$\int_0^T GY(t)\,\delta W(t) = G\int_0^T Y(t)\,\delta W(t) - \int_0^T Y(t) D_t G\,dt \ .$$

The proofs of the above propositions can be found in e.g. [Nua95], where a complete account of the Malliavin Calculus can be found.

7.2.2 Derivatives of Options on Spot

We now turn our attention to the computation of expressions of option derivatives using the Malliavin approach. We remark that many of the results below follow from the general results in [FLL$^+$99, FLL$^+$01]. For the sake of clarity we have chosen to derive the expressions for the specific models we have in mind. Introduce the set of functions

$$\Gamma_T = \{a \in L^2([0,T]) | \int_0^T a(t)\,dt = 1\} .$$

Then,

Proposition 7.2.3 (Delta by the Malliavin approach). *The delta of $u(x)$ can be represented as*

$$u'(x) = \frac{1}{\sigma x}\mathbb{E}\big[f(S(T))\int_0^T a(t)e^{-\alpha t}\,dW(t)\big] ,$$

where $a(t) \in \Gamma_T$.

Proof. Assume first that f is continuously differentiable with bounded derivative. It can then be shown that differentiation and expectation commutes, and thus

$$\begin{aligned} u'(x) &= \mathbb{E}\big[f'(S(T))\frac{\partial}{\partial x}S(T)\big] \\ &= x^{-1}\mathrm{e}^{-\alpha T}\mathbb{E}\big[f'(S(T))S(T)\big] . \end{aligned}$$

We have used that $\partial S(T)/\partial x = x^{-1}\mathrm{e}^{-\alpha T}S(T)$. The Malliavin derivative of the spot price is

$$D_t S(T) = \mathrm{e}^{X(T)}D_t X(T) = S(T)\sigma\mathrm{e}^{-\alpha(T-t)}\mathbf{1}_{\{t<T\}} .$$

Choose a function $a(t) \in \Gamma_T$. Integrating both sides above give

$$S(T) = \sigma^{-1}\mathrm{e}^{\alpha T}\int_0^T a(t)\mathrm{e}^{-\alpha t}D_t S(T)\,dt .$$

Using the properties of the Skorohod integral and the Malliavin derivative, this yields,

$$\begin{aligned} u'(x) &= \frac{1}{x\sigma}\mathbb{E}\big[\int_0^T f'(S(T))D_t S(T)a(t)\mathrm{e}^{-\alpha t}\,dt\big] \\ &= \frac{1}{x\sigma}\mathbb{E}\big[\int_0^T D_t f(S(T))a(t)\mathrm{e}^{-\alpha t}\,dt\big] \\ &= \frac{1}{x\sigma}\mathbb{E}\big[f(S(T))\int_0^T a(t)\mathrm{e}^{-\alpha t}\,dW(t)\big] . \end{aligned}$$

By a density argument the formula can be extended to $f \in L^2$. See [FLL$^+$99] for details on this. □

In all the following deductions of option derivatives using the Malliavin approach we shall use the method above with first assuming smooth payoff functions, and then passing to the limit by a density argument. This will from now on be done without being explicitly stated.

If we choose $a(t) = \mathrm{e}^{2\alpha t}/\int_0^T \mathrm{e}^{2\alpha t}\,dt = 2\alpha \mathrm{e}^{2\alpha t}/(\mathrm{e}^{2\alpha T}-1)$, we get

$$u'(x) = \frac{2\alpha}{\sigma x(\mathrm{e}^{2\alpha T}-1)}\mathbb{E}\big[f(S(T))\int_0^T \mathrm{e}^{\alpha t}\,dW(t)\big]\ .$$

But

$$\begin{aligned}\int_0^T \mathrm{e}^{\alpha t}\,dW(t) &= \sigma^{-1}\mathrm{e}^{\alpha T}\cdot\sigma\int_0^T \mathrm{e}^{-\alpha(T-t)}\,dW(t)\\ &= \sigma^{-1}\mathrm{e}^{\alpha T}\big(X(T)-\mathrm{e}^{-\alpha T}\ln x-\gamma(1-\mathrm{e}^{-\alpha T})\big)\\ &= \sigma^{-1}\big(\mathrm{e}^{\alpha T}X(T)-\ln x-\gamma(\mathrm{e}^{\alpha T}-1)\big)\ ,\end{aligned}$$

where $\gamma = \alpha(\mu-\lambda)-\sigma^2/2$. Hence,

$$u'(x) = \mathbb{E}\big[f(S(T))\frac{2\alpha}{x\sigma^2(\mathrm{e}^{2\alpha T}-1)}\big(\mathrm{e}^{\alpha T}\ln S(T)-\ln x-\gamma(\mathrm{e}^{\alpha T}-1)\big)\big]\ . \tag{7.7}$$

If we differentiate the delta of $u(x)$, we find the gamma:

Proposition 7.2.4 (The Malliavin approach). *The gamma of $u(x)$ can be represented as*

$$u''(x) = \frac{1}{\sigma^2x^2}\mathbb{E}\big[f(S(T))\big\{Z(T)^2-\sigma Z(T)-\int_0^T a^2(t)e^{-2\alpha t}\,dt\big\}\big]\ ,$$

where $Z(T) = \int_0^T a(t)e^{-\alpha t}\,dW(t)$ and $a(t)\in\Gamma_T$.

Proof. From Prop. 7.2.3 we have

$$u'(x) = \frac{1}{\sigma x}\mathbb{E}\big[f(S(T))Z(T)\big]$$

for

$$Z(T) = \int_0^T a(t)\mathrm{e}^{-\alpha t}\,dW(t)\ .$$

Hence,

$$\begin{aligned}u''(x) &= \frac{d}{dx}\frac{1}{\sigma x}\mathbb{E}\big[f(S(T))Z(T)\big]\\ &= \frac{-1}{\sigma x^2}\mathbb{E}\big[f(S(T))Z(T)\big]+\frac{1}{\sigma x}\mathbb{E}\big[f'(S(T))\frac{\partial}{\partial x}S(T)\cdot Z(T)\big]\ .\end{aligned}$$

We investigate the second expectation: Using that $\frac{\partial}{\partial x}S(T) = \mathrm{e}^{-\alpha T}S(T)/x$, and

$$S(T) = \sigma^{-1}\mathrm{e}^{\alpha T}\int_0^T a(t)\mathrm{e}^{-\alpha t}D_t S(T)\,dt\ ,$$

we obtain

$$\begin{aligned}\mathbb{E}\big[f'(S(T))\frac{\partial}{\partial x}S(T)\cdot Z(T)\big] &= \frac{1}{\sigma x}\mathbb{E}\big[\int_0^T f'(S(T))D_t S(T)a(t)\mathrm{e}^{-\alpha t}Z(T)\,dt\big]\\ &= \frac{1}{\sigma x}\mathbb{E}\big[\int_0^T D_t f(S(T))a(t)\mathrm{e}^{-\alpha t}Z(T)\,dt\big]\\ &= \frac{1}{\sigma x}\mathbb{E}\big[f(S(T))\int_0^T a(t)\mathrm{e}^{-\alpha t}Z(T)\delta W(t)\big]\ ,\end{aligned}$$

where δW mean the Skorohod integral, which is present since $Z(T)$ is anticipating. By the integration-by-parts formula for Skorohod integrals,

$$\begin{aligned}\int_0^T a(t)\mathrm{e}^{-\alpha t}Z(T)\delta W(t) &= Z(T)\int_0^T a(t)\mathrm{e}^{-\alpha t}\,dW(t) - \int_0^T a(t)\mathrm{e}^{-\alpha t}D_t Z(T)\,dt\\ &= Z(T)^2 - \int_0^T a^2(t)\mathrm{e}^{-2\alpha t}\,dt\ .\end{aligned}$$

This proves the result.

□

Note that

$$\mathbb{E}\big[Z(T)^2\big] = \int_0^T a^2(t)\mathrm{e}^{-2\alpha t}\,dt\ ,$$

by the Itô isometry. Consider the specific choice $a(t) = 2\alpha\mathrm{e}^{2\alpha t}/(\mathrm{e}^{2\alpha T}-1)$: Then

$$\begin{aligned}Z(T) &= \frac{2\alpha}{\mathrm{e}^{2\alpha T}-1}\int_0^T \mathrm{e}^{\alpha t}\,dW(t)\\ &= \frac{2\alpha}{\sigma(\mathrm{e}^{\alpha T}-1)}\big(\mathrm{e}^{2\alpha T}\ln S(T) - \ln x - \gamma(\mathrm{e}^{\alpha T}-1)\big)\ ,\end{aligned}$$

where $\gamma = \alpha(\mu-\lambda) - \sigma^2/2$. Furthermore,

$$\int_0^T a^2(t)\mathrm{e}^{-2\alpha t}\,dt = \frac{2\alpha}{(\mathrm{e}^{2\alpha T}-1)}\ .$$

The FD estimator for the gamma is given by

$$\frac{u(x+h)-2u(x)+u(x-h)}{h^2}\ ,$$

where the same considerations to h as for the delta apply. In sec. 7.7 numerical tests are presented for the Malliavin approach and the FD approach to compare convergence speeds.

It is possible to derive the delta and the gamma by the density approach since the probability density of $X(t)$ is known. We state the result for the delta only:

Proposition 7.2.5 (Delta by the density approach). *The delta of $u(x)$ can be represented as*

$$u'(x) \quad = \quad \frac{1}{x} E\big[f(S_T)\frac{2\alpha}{\sigma^2(e^{2\alpha T}-1)}\big(e^{\alpha T}\ln S(T) \;-\; \ln x \;-\; \gamma(e^{\alpha T} \;-\; 1)\big)\big] \quad . \tag{7.8}$$

Proof. Since

$$X(T) = \mathrm{e}^{-\alpha T}\ln x - \gamma(1-\mathrm{e}^{-\alpha T}) + \sigma\int_0^T \mathrm{e}^{-\alpha(T-s)}\,dW(s)\ ,$$

we have that $X(T)$ is normally distributed with expectation $\mathrm{e}^{-\alpha T}\ln x - \gamma(1-\mathrm{e}^{-\alpha T})$ and variance $\sigma^2(1-\mathrm{e}^{-2\alpha T})/2\alpha$. Denoting the density by $\phi(z;x)$ (as a function of z), we find by straightforward differentiation with respect to x

$$\frac{\partial\phi}{\partial x}(z;x) = \phi(z;x)\frac{1}{x}\mathrm{e}^{-\alpha T}\frac{z-\gamma(1-\mathrm{e}^{-\alpha T})-\mathrm{e}^{-\alpha T}\ln x}{\frac{\sigma^2}{2\alpha}(1-\mathrm{e}^{-2\alpha T})}\ .$$

Since differentiation and expectation commute in this case, we find

$$\begin{aligned} u'(x) &= \frac{d}{dx}\int f(\mathrm{e}^z)\phi(z;x)\,dz \\ &= \int f(\mathrm{e}^z)\frac{\partial\phi}{\partial x(z;x)}\,dz \\ &= \int f(\mathrm{e}^z)\frac{1}{x}\mathrm{e}^{-\alpha T}\frac{z-\gamma(1-\mathrm{e}^{-\alpha T})-\mathrm{e}^{-\alpha T}\ln x}{\frac{\sigma^2}{2\alpha}(1-\mathrm{e}^{-2\alpha T})}\phi(z;x)\,dz\ , \end{aligned}$$

which yields the desired result. □

Note that the density approach leads to the same formula as in (7.7), which was derived using the Malliavin approach with a specific choice of the weight function $a(t)$.

We consider the vega for the European option on spot using the Malliavin approach:

Proposition 7.2.6 (Vega by the Malliavin approach). *The vega of $u(\sigma)$ can be represented as*

$$u'(\sigma) = \sigma^{-1}E\big[f(S_T)\Big\{Z(T)\int_0^T e^{-\alpha(T-t)}\,dW_t - Z(T)\sigma(1-e^{-\alpha T}) - 1\Big\}\big]\ , \tag{7.9}$$

where $Z(T) = \int_0^T a(t)e^{\alpha(T-t)}\,dW_t$ and $a(t) \in \Gamma_T$.

Proof. The Malliavin derivative of S_T is given by

$$D_t S_T = S_T \sigma \mathrm{e}^{-\alpha(T-t)} \mathbf{1}_{t<T} \, .$$

By multiplying with a weight function $a(t) \in \Gamma_T$, and integrating each side, we find (after rearranging)

$$S_T = \sigma^{-1} \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} D_t S_T \, dt \, .$$

The expression for the vega is now found by:

$$\begin{aligned}
u'(\sigma) &= \mathrm{E}\big[f'(S_T)\frac{dS_T}{d\sigma}\big] \\
&= \mathrm{E}\big[f'(S_T) S_T \big\{ \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \sigma(1-\mathrm{e}^{-\alpha T}) \big\}\big] \\
&= \mathrm{E}\big[\int_0^T f'(S_T) D_t S_T \sigma^{-1} \mathrm{e}^{\alpha(T-t)} a(t) \big\{ \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \sigma(1-\mathrm{e}^{-\alpha T}) \big\} \, dt\big] \\
&= \sigma^{-1} \mathrm{E}\big[\int_0^T D_t f(S_T) \mathrm{e}^{\alpha(T-t)} a(t) \big\{ \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \sigma(1-\mathrm{e}^{-\alpha T}) \big\} \, dt\big] \\
&= \sigma^{-1} \mathrm{E}\big[f(S_T) \int_0^T \mathrm{e}^{\alpha(T-t)} a(t) \big\{ \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \sigma(1-\mathrm{e}^{-\alpha T}) \big\} \, \delta W_t\big] \, .
\end{aligned}$$

The last stochastic integral δW_t is the Skorohod integral. Using the integration-by-parts formula for Skorohod integration, we get

$$\begin{aligned}
&\int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s \, \delta W_t \\
&= \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \, dW_t \cdot \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} D_t \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s \, dt \\
&= \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \, dW_t \cdot \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \mathrm{e}^{-\alpha(T-t)} \mathbf{1}_{t<T} \, dt \\
&= \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \, dW_t \cdot \int_0^T \mathrm{e}^{-\alpha(T-s)} \, dW_s - 1 \, .
\end{aligned}$$

Hence,

$$\begin{aligned}
u'(\sigma) = \sigma^{-1} \mathrm{E}\big[f(S_T) \Big\{ &\int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \, dW_t \cdot \int_0^T \mathrm{e}^{-\alpha(T-t)} \, dW_t \\
&- \sigma(1-\mathrm{e}^{-\alpha T}) \int_0^T a(t) \mathrm{e}^{\alpha(T-t)} \, dW_t - 1 \Big\}\big] \, .
\end{aligned}$$

□

Choosing the weight function

$$a(t) = 2\alpha \mathrm{e}^{-2\alpha(T-t)}/(1-\mathrm{e}^{-2\alpha T}) \ ,$$

yields, after some calculations,

$$\int_0^T a(t)\mathrm{e}^{\alpha(T-t)}\,dW_t \cdot \int_0^T \mathrm{e}^{-\alpha(T-t)}\,dW_t = \frac{2\alpha}{\sigma^2(1-\mathrm{e}^{-2\alpha T})}\big(\ln S_T - \mathrm{e}^{-\alpha T}\ln S_0 - \gamma(1-\mathrm{e}^{-\alpha T})\big)^2$$

and

$$\sigma(1-\mathrm{e}^{-\alpha T})\int_0^T a(t)\mathrm{e}^{\alpha(T-t)}\,dW_t = \frac{2\alpha(1-\mathrm{e}^{-\alpha T})}{(1-\mathrm{e}^{-2\alpha T})}\big(\ln S_T - \mathrm{e}^{-\alpha T}\ln S_0 - \gamma(1-\mathrm{e}^{-\alpha T})\big) \ .$$

Note that the chosen $a(t)$ gives

$$\begin{aligned} Z(T) &= \frac{2\alpha}{(1-\mathrm{e}^{-2\alpha T})}\int_0^T \mathrm{e}^{-\alpha(T-t)}\,dW_t \\ &= \frac{2\alpha}{(1-\mathrm{e}^{-2\alpha T})}\sigma^{-1}\big(\ln(S_T) - \mathrm{e}^{-\alpha T}\ln(S_0) - \gamma(1-\mathrm{e}^{-\alpha T})\big) \ , \end{aligned}$$

where we have used that $\int_0^T \mathrm{e}^{-\alpha(T-t)}\,dW_t = \sigma^{-1}(\ln(S_T) - \mathrm{e}^{-\alpha T}\ln(S_0) - \gamma(1-\mathrm{e}^{-\alpha T}))$. Repeated use of this in (7.9), and insertion of $Z(T)$ gives a computable expression for the vega.

The FD estimator for the vega is given analogous to the delta as

$$\frac{u(\sigma+h)-u(\sigma)}{h} \ .$$

7.3 Derivatives of European Options on Commodity and Energy Forwards

Consider a European option with maturity $\tau < T$ and payoff $f(F(\tau,T))$. The price of this option is

$$u = \mathbb{E}\big[f(F(\tau,T))\big] \ . \tag{7.10}$$

Like for derivatives of spot options, we shall use the notation $u(x(T))$ and $u(\sigma)$ to emphasize the parameters of interest. First, we are interested in calculating the delta of u, that is, the derivative with respect of $x(T)$. Strictly speaking, $x(T)$ is a function of T, the maturity of the forward contract, and the derivative should be interpreted as a functional derivative. However, we keep T fixed here, and therefore we may

treat $du(X(T))/dx(T)$ as a standard derivative with respect to the variable $x(T)$. We denote this derivative $u'(x(T))$, which measures the sensitivity of u with respect to the initial forward price $x(T)$.

By the direct approach we find the expression (under the assumption that f is sufficiently regular so that differentiation can be moved inside the expectation, see e.g. [L'E91], [Gla91], [BG96] for conditions):

Proposition 7.3.1 (The direct approach). *The delta of $u(x(T))$ can be represented as*

$$u'(x(T)) = \frac{1}{x(T)}\mathbb{E}\Big[f'(F(\tau,T))F(\tau,T)\Big] .$$

Proof. Direct differentiation gives (assuming sufficient regularity such that differentiation and expectation interchange),

$$u'(x(T)) = \mathbb{E}\Big[f'(F(\tau,T))\frac{d}{dx(T)}F(\tau,T)\Big] .$$

It is easily seen that $\frac{d}{dx(T)}F(\tau,T) = x^{-1}(T)F(\tau,T)$. □

The density of $F(\tau,T)$ is known, which means that we can differentiate with respect to the density function instead: By the density approach we find:

Proposition 7.3.2 (The density approach). *The delta of $u(x(T))$ can be represented as*

$$u'(x(T)) = \frac{1}{x(T)}\mathbb{E}\Big[f(F(\tau,T))\Big(\frac{\ln(F(\tau,T)/x)}{\int_0^\tau \sigma^2(t,T)\,dt} + \frac{1}{2}\Big)\Big] .$$

Proof. We write $F(\tau,T)$ as

$$F(\tau,T) = \exp\big(\ln x(T) - \frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt + \int_0^\tau \sigma(t,T)\,dW(t)\big) .$$

Since $\int_0^\tau \sigma(t,T)\,dW(t) \sim \mathcal{N}\big(0, \int_0^\tau \sigma^2(t,T)\,dt\big)$, we have

$$F(\tau,T) = \exp\big(\ln x(T) - \frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt + \epsilon\cdot\sqrt{\int_0^\tau \sigma^2(t,T)\,dt}\big) ,$$

where $\epsilon \sim \mathcal{N}(0,1)$ and the equality is in distribution. Hence,

$$u(x(T)) = \int_{\mathbb{R}} f(\mathrm{e}^z)g(z;x(T))\,dz$$

for the density function

$$g(z;x(T)) = \frac{1}{\sqrt{2\pi\int_0^\tau \sigma^2(t,T)\,dt}}\exp\Big(-\frac{(z-\ln x(T)+\frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt)^2}{2\int_0^\tau \sigma^2(t,T)\,dt}\Big) .$$

Differentiation of $g(z;x(T))$ with respect to $x(T)$ yields,

$$\frac{dg(z;x(T))}{dx(T)} = g(z;x(T))\frac{z-\ln x(T)+\frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt}{x(T)\int_0^\tau \sigma^2(t,T)\,dt}\ .$$

Hence,

$$\begin{aligned} u'(x(T)) &= \int_{\mathbb{R}} f(\mathrm{e}^z)g(z;x(T))\frac{(z-\ln x(T)+\frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt}{x\int_0^\tau \sigma^2(t,T)\,dt}\,dz \\ &= \mathbb{E}\Big[f(F(\tau,T))\frac{\ln(F(\tau,T)/x(T))+\frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt}{x(T)\int_0^\tau \sigma^2(t,T)\,dt}\Big]\ . \end{aligned}$$

Thus, the proposition is proved. □

Finally, using the Malliavin approach we find

Proposition 7.3.3 (The Malliavin approach). *The delta of $u(x(T))$ can be represented as*

$$u'(x(T)) = \frac{1}{x(T)}\mathbb{E}\Big[f(F(\tau,T))\int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)\Big]\ ,$$

where $a \in \Gamma_\tau$.

Proof. We follow the argumentation in Fournié et al [FLL+99, Section 3.2]: Introduce the process $Y(t,T)$ by

$$Y(t,T) = \exp\big(-\frac{1}{2}\int_0^\tau \sigma^2(t,T)\,dt + \int_0^\tau \sigma(t,T)\,dW(t)\big)\ ,$$

which yields the representation $F(t,T) = x(T)Y(t,T)$. Let us do some calculations with the Malliavin derivative of $F(t,T)$: Straightforward application of the Malliavin derivative yields

$$D_tF(\tau,T) = x(T)Y(\tau,T)\sigma(t,T)\mathbf{1}_{t<\tau}\ .$$

Rearranging,

$$Y(\tau,T)\mathbf{1}_{t<\tau} = x^{-1}(T)\sigma^{-1}(t,T)D_tF(\tau,T)\ .$$

Multiplying both sides with a function $a(t) \in \Gamma_\tau$, and then integrating from 0 to τ, gives,

$$Y(\tau,T) = x^{-1}(T)\int_0^\tau D_tF(\tau,T)a(t)\sigma^{-1}(t,T)\,dt\ .$$

Direct differentiation gives

$$\begin{aligned} u'(x(T)) &= \mathbb{E}\big[f'(F(\tau,T))Y(\tau,T)\big] \\ &= x^{-1}(T)\mathbb{E}\big[\int_0^\tau f'(F(\tau,T))D_tF(\tau,T)a(t)\sigma^{-1}(t,T)\,dt\big] \\ &= x^{-1}(T)\mathbb{E}\big[\int_0^\tau D_tf(F(\tau,T))a(t)\sigma^{-1}(t,T)\,dt\big] \\ &= x^{-1}(T)\mathbb{E}\big[f(F(\tau,T))\int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)\big]\,, \end{aligned}$$

where we used the chain rule for Malliavin derivative in the second last equality and the duality between Malliavin differentiation and Skorohod integration in the last. □

We consider different choices of the function $a(t)$: Choose $a(t) = K\sigma(t,T)$, where $K = 1/\int_0^\tau \sigma(t,T)\,dt$. Then

$$u'(x(T)) = x^{-1}(T)\mathbb{E}\Big[f(F(\tau,T))\frac{W(\tau)}{\int_0^\tau \sigma(t,T)\,dt}\Big]\,. \tag{7.11}$$

A different choice could be $a(t) = K\sigma^2(t,T)$, where $K = 1/\int_0^\tau \sigma^2(t,T)\,dt$. Then

$$u'(x(T)) = \mathbb{E}\Big[f(F(\tau,T))\frac{\int_0^\tau \sigma(t,T)\,dW(t)}{\int_0^\tau \sigma^2(t,T)\,dt}\Big]\,, \tag{7.12}$$

which, after a slight rewriting, coincides with the delta obtained using the density method.

We are also interested in calculating the gamma of u, that is, the double derivative with respect of $x(T)$. Similar considerations as for the delta applies.

Proposition 7.3.4 (The Malliavin approach). *The gamma of $u(x(T))$ can be represented as*

$$u''(x(T)) = x(T)^{-2}\mathbb{E}[f(F(\tau,T))\Big\{Z^2(\tau,T) - Z(\tau,T) - \int_0^\tau a^2(t)\sigma^{-2}(t,T)\,dt\Big\}]\,,$$

where $Z(\tau,T) = \int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)$ and $a(t) \in \Gamma_\tau$.

Proof. Write the the delta as

$$u'(x(T)) = x(T)^{-1}\mathbb{E}[f(F(\tau,T))Z(\tau,T)]\,,$$

where $Z(\tau,T) = \int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)$. Then the gamma is given by differentiation

of delta with respect to the initial condition $x(T)$:

$$\begin{aligned} u''(x(T)) &= -x(T)^{-2}\mathbb{E}[f(F(\tau,T))Z(\tau,T)] \\ &\quad + x(T)^{-1}\mathbb{E}[f'(F(\tau,T))\frac{F(\tau,T)}{x(T)}Z(\tau,T)] \\ &= -x(T)^{-2}\mathbb{E}[f(F(\tau,T))Z(\tau,T)] \\ &\quad + x(T)^{-2}\mathbb{E}[\int_0^\tau f'(F(\tau,T))D_tF(\tau,T)a(t)\sigma^{-1}(t,T)Z(\tau,T)\,dt] \\ &= -x(T)^{-2}\mathbb{E}[f(F(\tau,T))Z(\tau,T)] \\ &\quad + x(T)^{-2}\mathbb{E}[f(F(\tau,T))\int_0^\tau a(t)\sigma^{-1}(t,T)Z(\tau,T)\,\delta W(t)]\,, \end{aligned}$$

where we used the chain rule for Malliavin derivative in the second last equality and the duality between Malliavin differentiation and Skorohod integration in the last. Using that

$$Z(\tau,T) = \int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t), \qquad D_tZ(\tau,T) = a(t)\sigma^{-1}(t,T)\mathbf{1}_{t<\tau}\,,$$

the integration-by-parts formula of the Skorohod integral gives

$$\begin{aligned} \int_0^\tau a(t)\sigma^{-1}(t,T)Z(\tau,T)\,\delta W(t) &= Z(\tau,T)\int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t) \\ &\quad - \int_0^\tau a(t)\sigma^{-1}(t,T)D_tZ(\tau,T)\,dt \\ &= Z(\tau,T)^2 - \int_0^\tau a^2(t)\sigma^{-2}(t,T)\,dt\,. \end{aligned}$$

The final formula for the gamma therefore reads

$$\begin{aligned} u''(x(T)) = x(T)^{-2}\mathbb{E}[f(F(\tau,T))\Big(&\big(\int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)\big)^2 \\ &- \int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t) - \int_0^\tau a^2(t)\sigma^{-2}(t,T)\,dt\Big)]\,. \end{aligned}$$

□

In order to get an implementable expression for gamma, choose the weight function $a(t) = \sigma(t,T)/\int_0^T \sigma(t,T)\,dt$. Then

$$u''(x) = x(t)^{-2}\mathbb{E}[f(F(\tau,T))\{W_\tau^2C^2 - W_\tau C - \tau C^2\}]\,,$$

where $C = \alpha e^{-\alpha T}(e^{\alpha\tau}-1)/\sigma$.

7.4 Derivatives of Asian Options on Commodity and Energy Forwards

Define an Asian claim with maturity $\tau < T$,

$$u = \mathbb{E}\big[f\big(\int_0^\tau F(t,T)\,dt\big)\big]\ . \tag{7.13}$$

Assume $f \in L^2(\mathbb{R})$ and $\mathbb{E}\big[f\big(\int_0^\tau F(t,T)\,dt\big)^2\big] < \infty$. Here one of the drawbacks of the density approach becomes evident. The density of $\int_0^T F(t,T)\,dt$ is *not* explicitly known to us, so that the density approach is not applicable. Consider the Malliavin approach.

Proposition 7.4.1 (The Malliavin approach). *The delta of $u(x(T))$ can be represented as*

$$u'(x(T)) = \mathbb{E}\Big[f\big(\int_0^\tau F(t,T)\,dt\big)X(\tau,T)\Big]\ ,$$

where

$$\begin{aligned} X(\tau,T) = \frac{2}{x(T)\int_0^\tau F(t,T)\,dt}\Big\{&\sigma^{-2}(\tau,T)F(\tau,T) - \sigma^{-2}(0,T)x(T) \\ &+ 2\int_0^\tau \sigma_t(t,T)\sigma^{-3}(t,T)F(t,T)\,dt \\ &+ \frac{\int_0^\tau \sigma^{-1}(t,T)F(t,T)\int_t^\tau \sigma(u,T)F(u,T)\,du\,dt}{\int_0^\tau F(t,T)\,dt}\Big\}\ . \end{aligned}$$

Proof. Direct differentiation, and integration-by-parts yield

$$\begin{aligned} u'(x(T)) &= \mathbb{E}\big[f'\big(\int_0^\tau F(t,T)\,dt\big)\int_0^\tau Y(t,T)\,dt\big] \\ &= \mathbb{E}\big[f'\big(\int_0^\tau F(t,T)\,dt\big)2\int_0^\tau Y(t,T)\int_t^\tau Y(s,T)\,ds\,dt\int_0^\tau Y(t,T)\,dt)^{-1}\big] \\ &= \mathbb{E}\big[\int_0^\tau f'\big(\int_0^\tau F(t,T)\,dt\big)2Y(t,T)\int_t^\tau Y(s,T)\,ds\,dt\int_0^\tau Y(t,T)\,dt)^{-1}\,dt\big]\ , \end{aligned}$$

where $F(t,T) = x(T)Y(t,T)$. A straightforward calculation reveals

$$D_tF(s,T) = F(s,T)\sigma(t,T)\mathbf{1}_{\{t<s\}}\ .$$

Thus

$$\int_0^\tau D_tF(s,T)\,ds = x(T)\sigma(t,T)\int_t^\tau Y(s,T)\,ds\ .$$

We therefore have (using the properties of Malliavin derivative)

$$\begin{aligned}
u'(x(T)) &= \mathbb{E}\Big[\int_0^\tau f'\Big(\int_0^\tau F(s,T)\,ds\Big)\int_0^\tau D_t F(s,T)\,ds x^{-1}(T)\sigma^{-1}(t,T)2Y(t,T)\\
&\quad\times\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\,dt\Big]\\
&= \mathbb{E}\Big[\int_0^\tau f'\Big(\int_0^\tau F(s,T)\,ds\Big)D_t\int_0^\tau F(s,T)\,ds x^{-1}(T)\sigma^{-1}(t,T)2Y(t,T)\\
&\quad\times\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\,dt\Big]\\
&= \mathbb{E}\Big[\int_0^\tau D_t f\Big(\int_0^\tau F(s,T)\,ds\Big)2Y(t,T)x^{-1}(T)\sigma^{-1}(t,T)\\
&\quad\times\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\,dt\Big]\\
&= \mathbb{E}\Big[f\Big(\int_0^\tau F(s,T)\,ds\Big)X(\tau,T)\Big]\,,
\end{aligned}$$

where

$$X(\tau,T) = \frac{2}{x(T)}\int_0^\tau Y(t,T)\sigma^{-1}(t,T)\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\,\delta W(t)\,.$$

Let us calculate $X(\tau,T)$: Integration-by-parts for Skorohod integrals:

$$\begin{aligned}
&\int_0^\tau \sigma^{-1}(t,T)Y(t,T)\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\delta W(t)\\
&\qquad= \int_0^\tau \sigma^{-1}(t,T)Y(t,T)\,dW(t)\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\\
&\qquad\qquad- \int_0^\tau \sigma^{-1}(t,T)Y(t,T)D_t\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\,dt\\
&\quad= \int_0^\tau \sigma^{-1}(t,T)Y(t,T)\,dW(t)\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-1}\\
&\qquad\quad+ \int_0^\tau \sigma^{-1}(t,T)Y(t,T)\Big(\int_0^\tau Y(s,T)\,ds\Big)^{-2}\int_t^\tau Y(u,T)\sigma(u,T)\,du\,dt\,.
\end{aligned}$$

Consider $\sigma^{-2}(t,T)F(t,T)$ and assume $\partial\sigma(t,T)/\partial t := \sigma_t(t,T)$ exists. The Itô Formula yields,

$$\begin{aligned}
d\big(\sigma^{-2}(t,T)F(t,T)\big) &= -2\sigma^{-3}(t,T)\sigma_t(t,T)F(t,T)\,dt + \sigma^{-2}(t,T)dF(t,T)\\
&= -2\sigma^{-3}(t,T)\sigma_t(t,T)F(t,T)\,dt + \sigma^{-1}(t,T)F(t,T)\,dW(t)\,.
\end{aligned}$$

Integrating both sides from 0 to τ, and inserting into the expression for $X(\tau,T)$, gives the desired result. □

Let us consider the concrete choice of $\sigma(t,T)$ given in (7.5). In this case it is straightforward to see that

$$X(\tau,T) = \frac{2}{x(T)}\Big\{ \mathrm{e}^{2\alpha T}\frac{\mathrm{e}^{-2\alpha\tau}F(\tau,T) - x(T)}{\sigma^2\int_0^\tau F(t,T)\,dt} + 2\alpha\mathrm{e}^{2\alpha T}\frac{\int_0^\tau \mathrm{e}^{-2\alpha t}F(t,T)\,dt}{\sigma^2\int_0^\tau F(t,T)\,dt} + \frac{\int_0^\tau \mathrm{e}^{\alpha t}F(t,T)\int_t^\tau \mathrm{e}^{-\alpha u}F(u,T)\,du\,dt}{(\int_0^\tau F(t,T)\,dt)^2}\Big\}\,.$$

In practice one is interested in Asian options where the averaging is taken over discrete dates, i.e. the arithmetic average Asian option. The payoff function will be

$$u = \mathbb{E}\big[f\big(\sum_{k=0}^{n} F(t_k,T)\big)\big]\,, \tag{7.14}$$

where $0 = t_0 < t_1 < \ldots < t_n \le T$. The delta and gamma are given as follows:

Proposition 7.4.2 (The Malliavin approach). *The delta and gamma of $u(x(T))$ can be represented as*

$$u'(x(T)) = \frac{1}{x(T)}\mathbb{E}\Big[f\big(\sum_{k=0}^{n} F(t_k,T)\big)Z(t_n,T)\Big]$$

$$u''(x(T)) = \frac{1}{x^2(T)}\mathbb{E}\Big[f\big(\sum_{k=0}^{n} F(t_k,T)\big)\Big\{Z^2(t_n,T) - Z(t_n,T) - \int_0^{t_n} a^2(t)\sigma^{-2}(t,T)\,dt\Big\}\Big]\,,$$

where $Z(t_n,T) = \int_0^{t_n} a(t)\sigma^{-1}(t,T)\,dW(t)$ and $a(t)$ is such that $\int_0^{t_1} a(t)\,dt = 1$, and $\int_{t_k}^{t_{k+1}} a(t)\,dt = 0$ for $k = 1,\ldots,n-1$.

Proof. Direct differentiation leads to

$$u'(x(T)) = \mathbb{E}\big[\sum_k f'\big(\sum_k F(t_k,T)\big)Y(t_k,T)\big]\,,$$

where $Y(t_k,T) = x^{-1}(T)F(t_k,T)$. Since

$$D_t F(t_k,T) = F(t_k,T)\sigma(t,T)\mathbf{1}_{\{t<t_k\}}\,,$$

we find

$$Y(t_k,T)\mathbf{1}_{\{t<t_k\}} = x^{-1}(T)\sigma^{-1}(t,T)D_t F(t_k,T)\,.$$

Introducing a function $a(t) \in \Gamma_{t_k}$ for all $k = 1,\ldots,n$, and integrating both sides after multiplication with this function gives

$$Y(t_k,T) = x^{-1}(T)\int_0^{t_n} a(t)\sigma^{-1}(t,T)D_t F(t_k,T)\,dt\,.$$

Hence,

$$\begin{aligned} u'(x(T)) &= x^{-1}(T)\mathbb{E}\big[\int_0^{t_n} \sum_k f'(\sum_k F(t_k,T)) D_t F(t_k,T) a(t)\sigma^{-1}(t,T)\,dt\big] \\ &= x^{-1}(T)\mathbb{E}\big[\int_0^{t_n} D_t f(\sum_k F(t_k,T)) a(t)\sigma^{-1}(t,T)\,dt\big] \\ &= x^{-1}(T)\mathbb{E}\big[f(\sum_k F(t_k,T)) \int_0^{t_n} a(t)\sigma^{-1}(t,T)\,dW(t)\big]\ . \end{aligned}$$

□

Here is an example of a function $a(t)$ satisfying the property in Prop. 7.4.2:

$$a(t) = \begin{cases} t_1^{-1}, & t \in [0,t_1), \\ t - \frac{1}{2}(t_{k+1}+t_k), & t \in [t_k,t_{k+1}), k = 1,\dots,n-1 \end{cases}\ .$$

For this a, the Itô integral inside the expression for $u'(x(T))$ becomes:

$$\begin{aligned} \int_0^{t_n} a(t)\sigma^{-1}(t,T)\,dW(t) = &\sum_{k=0}^{n-1} c_k \int_{t_k}^{t_{k+1}} \sigma^{-1}(t,T)\,dW(t) \\ &+ \sum_{k=1}^{n-1} \int_{t_k}^{t_{k+1}} t\sigma^{-1}(t,T)\,dW(t)\ , \end{aligned}$$

where

$$c_k = \begin{cases} t_1^{-1}, & k = 0 \\ -\frac{1}{2}(t_{k+1}+t_k), & k > 0 \end{cases}\ .$$

Define $X_k := c_k \int_{t_k}^{t_{k+1}} \sigma^{-1}(t,T)\,dW(t)$ and $Y_k = \int_{t_k}^{t_{k+1}} t\sigma^{-1}(t,T)\,dW(t)$. It is easily seen that $\{X_k\}_k$ are independent random variables. Likewise for $\{Y_k\}_k$. In addition, X_i and Y_j are independent for all $i \neq j$. Both are distributed as

$$X_k \sim \mathcal{N}\big(0, c_k^2 \int_{t_k}^{t_{k+1}} \sigma^{-2}(t,T)\,dt\big),$$

and

$$Y_k \sim \mathcal{N}\big(0, \int_{t_k}^{t_{k+1}} t^2\sigma^{-2}(t,T)\,dt\big).$$

The covariance between X_k and Y_k is

$$\mathrm{Cov}\big(X_k, Y_k\big) = c_k \int_{t_k}^{t_{k+1}} t\sigma^{-2}(t,T)\,dt\ .$$

A natural example to consider is $\sigma(t,T)$ as given in (7.5).

In [FLL$^+$01] they show that the representation of the derivative is of minimal variance if the weight can be written as a functional of the underlying price process. We demonstrate how this is here; Use Itôs Formula on $a(t)\sigma^{-2}(t,T)\ln F(t,T)$ to obtain

$$\begin{aligned}
&d\big(a(t)\sigma^{-2}(t,T)\ln F(t,T)\big) = \\
&\quad \big\{a'(t)\sigma^{-2}(t,T) - 2a(t)\sigma^{-3}(t,T)\sigma_t(t,T)\big\}\ln F(t,T)\,dt \\
&\quad + a(t)\sigma^{-2}(t,T)\frac{1}{F(t,T)}\,dF(t,T) + \frac{1}{2}a(t)\sigma^{-2}(t,T)\frac{-1}{F(t,T)^2}\big(dF(t,T)\big)^2 \\
&= \big\{a'(t)\sigma^{-2}(t,T) - 2a(t)\sigma^{-3}(t,T)\sigma_t(t,T) - \frac{1}{2}a(t)\big\}\ln F(t,T)\,dt \\
&\quad + a(t)\sigma^{-1}(t,T)\,dW(t)\ .
\end{aligned}$$

Hence, integrating both sides and rearranging, we get,

$$\begin{aligned}
\int_0^{t_n} a(t)\sigma^{-1}(t,T)\,dW(t) &= a(t_n)\sigma^{-2}(t_n,T)\ln F(t_n,T) - a(0)\sigma^{-2}(0,T)x(T) \\
&\quad - \int_0^{t_n}\big\{a'(t)\sigma^{-2}(t,T) - 2a(t)\sigma^{-3}(t,T)\sigma_t(t,T) - \frac{1}{2}a(t)\big\}\ln F(t,T)\,dt\ .
\end{aligned}$$

This shows that the weight is a functional of the underlying process $F(t,T)$.

We now look at the vega of an Asian option. Choose the volatility to be $\sigma(t,T) = \sigma\alpha(t,T)$, and consider (7.13) as a function of σ, that is, $u(\sigma)$. We are now interested in calculating the *vega* of $u(\sigma)$, $u'(\sigma)$, using the Malliavin approach. We concentrate the calculation to the discrete Asian case:

$$u(\sigma) = \mathbb{E}\big[f(\sum_{k=0}^{n} F(t_k,T))\big]\ ,$$

where $0 = t_0 < t_1 < \ldots < t_n \le T$.

Proposition 7.4.3 (The Malliavin approach). *The vega of u can be represented as*

$$u'(\sigma) = \mathbb{E}\Big[f(\sum_{k=0}^{n} F(t_k,T))X(\{t_k\},T)\Big]$$

where

$$\begin{aligned}
X(\{t_k\},T) = -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{k=1}^{n}\Big\{\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)\,dW(t) \\
\times\Big(\ln F(t_k,T) - \ln F(t_{k-1},T) - \frac{1}{2}\sigma^2\int_{t_{k-1}}^{t_k}\alpha^2(t,T)\,dt\Big)\Big\}\ ,
\end{aligned}$$

and $a(t)$ is a function such that $\int_{t_{k-1}}^{t_k} a(t)\,dt = 1$ for $k = 1,\ldots,n$.

Proof. Direct differentiation gives

$$u'(\sigma) = \mathbb{E}\big[f'(\sum_{k=0}^{n} F(t_k,T)) \sum_{k=1}^{n} F(t_k,T) Z(t_k,T)\big] \ ,$$

where

$$Z(t_k,T) = \int_0^{t_k} \alpha(t,T)\, dW(t) - \sigma \int_0^{t_k} \alpha^2(t,T)\, dt \ .$$

The Malliavin derivative of $F(t_k,T)$ is given by

$$D_t F(t_k,T) = \sigma\alpha(t,T) F(t_k,T) \mathbf{1}_{\{t<t_k\}} \ .$$

Define as in Fournié et al [FLL$^+$99, Section 3.3]:

$$\beta_a(t) = \sum_{k=1}^{n} a(t)\big(Z(t_k,T) - Z(t_{k-1},T)\big)\mathbf{1}_{\{t_{k-1}<t<t_k\}} \ ,$$

where $\int_{t_{k-1}}^{t_k} a(t)\, dt = 1, k = 1,2,\ldots,n$. This yields (note that $Z(0,T) = 0$),

$$\begin{aligned}
\int_0^T \sigma^{-1}\alpha^{-1}(t,T) D_t F(t_k,T) \beta_a(t)\, dt &= F(t_k,T) \int_0^{t_k} \beta_a(t)\, dt \\
&= F(t_k,T) \sum_{i=1}^{k} \int_{t_{i-1}}^{t_i} a(t)\, dt \big(Z(t_i,T) - Z(t_{i-1},T)\big) \\
&= F(t_k,T) Z(t_k,T) \ .
\end{aligned}$$

Hence,

$$\begin{aligned}
u'(\sigma) &= \sigma^{-1}\mathbb{E}\big[\int_0^T \sum_{k=1}^{n} f'(\sum_{k=0}^{n} F(t_k,T)) D_t F(t_k,T) \alpha^{-1}(t,T)\beta_a(t)\, dt\big] \\
&= \sigma^{-1}\mathbb{E}\big[\int_0^T D_t f(\sum_{k=0}^{n} F(t_k,T)) \alpha^{-1}(t,T)\beta_a(t)\, dt\big] \\
&= \mathbb{E}\big[f(\sum_{k=0}^{n} F(t_k,T)) \sigma^{-1} \int_0^T \alpha^{-1}(t,T)\beta_a(t)\, \delta W(t)\big] \ .
\end{aligned}$$

Note that $\beta_a(t)$ is $\mathcal{F}_{t_k}$-measurable for $t \le t_k$, and thus anticipating. We calculate the

Skorohod integral using integration-by-parts formula:

$$\begin{aligned}
X(\{t_k\},T) &= \sigma^{-1}\int_0^T \alpha^{-1}(t,T)\beta_a(t)\,\delta W(t)\\
&= \sigma^{-1}\sum_{k=1}^{n}\int_0^T \mathbf{1}_{\{t_{k-1}<t<t_k\}}\alpha^{-1}(t,T)a(t)\big(Z(t_k,T)-Z(t_{k-1},T)\big)\,\delta W(t)\\
&= \sigma^{-1}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)\big(Z(t_k,T)-Z(t_{k-1},T)\big)\,\delta W(t)\\
&= \sigma^{-1}\sum_{k=1}^{n}\big(Z(t_k,T)-Z(t_{k-1},T)\big)\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)dW(t)\\
&\quad - \sigma^{-1}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)D_t\big(Z(t_k,T)-Z(t_{k-1},T)\big)\,dt\,.
\end{aligned}$$

But, since

$$Z(t_k,T)-Z(t_{k-1},T)=\int_{t_{k-1}}^{t_k}\alpha(t,T)\,dW(t)-\sigma\int_{t_{k-1}}^{t_k}\alpha^2(t,T)\,dt\,,$$

we find (for $t\in[t_{k-1},t_k]$),

$$D_t\big(Z(t_k,T)-Z(t_{k-1},T)\big)=\alpha(t,T)\,.$$

Hence, the last sum of integrals above becomes

$$\sigma^{-1}\sum_{k=1}^{n}a(t)\,dt=\sigma^{-1}\sum_{k=1}^{n}1=\frac{n}{\sigma}\,.$$

Recalling that

$$\ln F(t_k,T)-\ln F(t_{k-1},T)=\frac{1}{2}\sigma^2\int_{t_{k-1}}^{t_k}\alpha^2(t,T)\,dt+\sigma\int_{t_{k-1}}^{t_k}\alpha(t,T)\,dW(t)\,,$$

we find

$$Z(t_k,T)-Z(t_{k-1},T)=\frac{1}{\sigma}\Big(\ln F(t_k,T)-\ln F(t_{k-1},T)-\frac{1}{2}\sigma^2\int_{t_{k-1}}^{t_k}\alpha^2(t,T)\,dt\Big)\,.$$

Hence,

$$\begin{aligned}
X(\{t_k\},T) &= \frac{1}{\sigma^2}\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)\,dW(t)\\
&\qquad \Big(\ln F(t_k,T)-\ln F(t_{k-1},T)-\frac{1}{2}\sigma^2\int_{t_{k-1}}^{t_k}\alpha^2(t,T)\,dt\Big)-\frac{n}{\sigma}\,.
\end{aligned}$$

□

Now consider a specific choice for $a(t)$; let for $t \in [t_{k-1}, t_k)$

$$a(t) = \alpha^2(t,T) / \int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\ .$$

Then,

$$\begin{aligned}\int_{t_{k-1}}^{t_k} a(t)\alpha^{-1}(t,T)\,dt &= \frac{\int_{t_{k-1}}^{t_k} \alpha(t,T)\,dW(t)}{\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt} \\ &= \frac{1}{\sigma}\big(\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\big)^{-1} \\ &\qquad \times \Big(\ln F(t_k,T) - \ln F(t_{k-1},T) + \frac{1}{2}\sigma^2 \int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\Big)\end{aligned}$$

by using

$$\ln F(t_k,T) - \ln F(t_{k-1},T) = -\frac{1}{2}\sigma^2 \int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt + \sigma \int_{t_{k-1}}^{t_k} \alpha(t,T)\,dW(t)\ .$$

Hence,

$$\begin{aligned}X(\{t_k\},T) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{k=1}^{n}\big(\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\big)^{-1} \\ \times \Big(\ln F(t_k,T) - \ln F(t_{k-1},T) - \frac{1}{2}\sigma^2 \int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\Big) \\ \times \Big(\ln F(t_k,T) - \ln F(t_{k-1},T) + \frac{1}{2}\sigma^2 \int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\Big)\ .\end{aligned}$$

Therefore,

$$X(\{t_k\},T) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{k=1}^{n}\Big(\frac{\big(\ln F(t_k,T) - \ln F(t_{k-1},T)\big)^2}{\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt} - \frac{\sigma^4}{4}\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt\Big)\ .$$

If we furthermore let $\alpha(t,T) = \mathrm{e}^{-\alpha(T-t)}$ for a constant α, we get

$$\int_{t_{k-1}}^{t_k} \alpha^2(t,T)\,dt = \frac{1}{2\alpha}\big(\mathrm{e}^{-2\alpha(T-t_k)} - \mathrm{e}^{-2\alpha(T-t_{k-1})}\big)\ ,$$

which inserted into the expression for $X(\{t_k\},T)$ gives

$$\begin{aligned}X(\{t_k\},T) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{k=1}^{n}\Big(\frac{2\alpha\big(\ln F(t_k,T) - \ln F(t_{k-1},T)\big)^2}{\mathrm{e}^{-2\alpha(T-t_k)} - \mathrm{e}^{-2\alpha(T-t_{k-1})}} \\ - \frac{\sigma^4}{8\alpha}\big(\mathrm{e}^{-2\alpha(T-t_k)} - \mathrm{e}^{-2\alpha(T-t_{k-1})}\big)\Big)\ .\end{aligned}$$

Finally, note that by choosing $n = 1$ we get the vega for a European forward option: Let $t_1 = \tau$;

$$u'(\sigma) = \frac{d}{d\sigma}\mathbb{E}\big[f(F(\tau,T))\big] = \mathbb{E}\big[f(F(\tau,T))X(\tau,T)\big] \ ,$$

where

$$X(\tau,T) = -\frac{1}{\sigma} + \frac{1}{\sigma^2}\Big\{\int_0^\tau a(t)\alpha^{-1}(t,T)\,dW(t) \\ \times \Big(\ln F(\tau,T) - \ln x(T) - \frac{1}{2}\sigma^2\int_0^\tau \alpha^2(t,T)\,dt\Big)\Big\}$$

and $\int_0^\tau a(t) = 1$.

7.5 Localized Malliavin Approach for Call Options

If a localized Malliavin technique is to be applied, we have to specify the payoff function f prior to the deduction. We calculate here expressions using the payoff of a call option, i.e. we choose

$$f(x) = \max\big(x-K,0\big) = \big(x-K\big)^+ \ ,$$

where K is the contracted strike price. As argued in Fournié et al, variance reduction is achieved by using a *localized* Malliavin technique. Instead of using the Malliavin approach to calculate expressions for derivatives globally, they suggest to use the approach only around the singularity of the payoff function. That is, localize the payoff function $f(x)$ around $x = K$, and use the Malliavin approach on this piece of f.

More specifically, introduce a "smoothened" Heaviside function

$$H_a(x) = \begin{cases} 0, & x < K-a \\ \frac{x-(K-a)}{2a}, & K-a \le x \le K+a \\ 1, & x > K+a \end{cases} \ , \tag{7.15}$$

for a constant a (which is not necessarily small!). Introduce

$$g_a(x) = \int_{-\infty}^x H_a(y)\,dy = \begin{cases} 0, & x < K-a \\ \frac{1}{4a}\big(x-(K-a)\big)^2, & x-K \le x \le K+a \\ 1, & x > K+a \end{cases} \ .$$

Note that $g_a'(x) = H_a(x)$. Finally, let

$$f_a(x) = f(x) - g_a(x) = \big(x-K\big)^+ - g_a(x) \ ,$$

and notice that $f_a(x) = 0$ whenever $x < K-a$ or $x > K+a$. We will understand $f_a(x)$ as the *localized* version of $f(x)$.

Proposition 7.5.1. *The delta and gamma of a call on the spot is given by resp.*

$$\begin{aligned}
u'(x) &= \frac{1}{\sigma x}\mathbb{E}\big[f_a(S(T))Z(T)\big] + \frac{1}{x e^{\alpha T}}\mathbb{E}\big[H_a(S(T))S(T)\big] \\
u''(x) &= \frac{1}{\sigma^2 x^2(T)}\mathbb{E}\big[f_a(S(T))\big\{Z^2(T) - \sigma Z(T) - \int_0^T a^2(t)e^{-2\alpha t}\,dt\big\}\big] \\
&\quad + \frac{1}{x^2(T)e^{\alpha T}}\mathbb{E}\big[H_a'(S(T))\tau^2(T)e^{-\alpha T} - H_a(S(T))S(T)(1-e^{-\alpha T})\big]\,,
\end{aligned}$$

where $Z(T) = \int_0^T a(t)e^{-\alpha t}\,dW(t)$ *and* $a(t) \in \Gamma_T$.

Proof. Represent $(x-K)^+ = f_a(x) + g_a(x)$. Use Prop. 7.2.3 on the first term and direct differentiation on the second to obtain

$$\begin{aligned}
u'(x) &= \frac{d}{dx}\mathbb{E}\big[f_a(S(T))\big] + \frac{d}{dx}\mathbb{E}\big[g_a(S(T))\big] \\
&= \frac{1}{\sigma x}\mathbb{E}\big[f_a(S(T))Z(T)\big] \\
&\quad + \mathbb{E}\big[g_a'(S(T))x^{-1}(T)\mathrm{e}^{-\alpha T}S(T)\big]\,.
\end{aligned}$$

In the last equality we used that $\frac{d}{dx}S(T) = x^{-1}(T)\mathrm{e}^{-\alpha T}S(T)$, which proves the formula for the delta.

Using Prop. 7.2.4 on the first term and direct differentiation on the second yields,

$$\begin{aligned}
u''(x) &= \frac{d^2}{d^2x}\mathbb{E}\big[f_a(S(T))\big] + \frac{d^2}{d^2x}\mathbb{E}\big[g_a(S(T))\big] \\
&= \frac{1}{\sigma^2 x^2(T)}\mathbb{E}\big[f_a(S(T))\big\{Z^2(T) - \sigma Z(T) - \int_0^T a^2(t)\mathrm{e}^{-2\alpha t}\,dt\big\}\big] \\
&\quad + \frac{d}{dx}\mathbb{E}\big[H_a(S(T))x^{-1}(T)\mathrm{e}^{-\alpha T}S(T)\big]\,.
\end{aligned}$$

Differentiation of the last expectation yields the formula for the gamma. □

Proposition 7.5.2. *The delta and gamma of a call on the forward is given by resp.*

$$\begin{aligned}
u'(x(T)) &= \frac{1}{x(T)}\mathbb{E}\big[f_a(F(\tau,T))Z(\tau,T)\big] + \frac{1}{x(T)}\mathbb{E}\big[H_a(F(\tau,T))F(\tau,T)\big] \\
u''(x(T)) &= \frac{1}{x^2(T)}\mathbb{E}\big[f_a(F(\tau,T))\big\{Z^2(\tau,T) - Z(\tau,T) \\
&\qquad - \int_0^\tau a^2(t)\sigma^{-2}(t,T)\,dt\big\}\big] + \frac{1}{x^2(T)}\mathbb{E}\big[H_a'(F(\tau,T))F^2(\tau,T)\big]\,,
\end{aligned}$$

where $Z(\tau,T) = \int_0^\tau a(t)\sigma^{-1}(t,T)\,dW(t)$ *and* $a(t) \in \Gamma_\tau$.

Proof. Using the Malliavin expression for delta and direct differentiation we get

$$\begin{aligned} u'(x(T)) &= \frac{d}{dx(T)}\mathbb{E}\big[f_a(F(\tau,T))\big] + \frac{d}{dx(T)}\mathbb{E}\big[g_a(F(\tau,T))\big] \\ &= \frac{1}{x(T)}\mathbb{E}\big[f_a(F(\tau,T))Z(\tau,T)\big] + \mathbb{E}\big[g_a'(F(\tau,T))\frac{d}{dx(T)}F(\tau,T)\big] \ . \end{aligned}$$

But $\frac{d}{dx(T)}F(\tau,T) = x^{-1}(T)F(\tau,T)$, which yields the delta formula.
Using the expression for the gamma we have

$$\begin{aligned} u''(x(T)) &= \frac{d^2}{d^2x(T)}\mathbb{E}\big[f_a(F(\tau,T))\big] + \frac{d^2}{d^2x(T)}\mathbb{E}\big[g_a(F(\tau,T))\big] \\ &= x^{-2}(T)\mathbb{E}\big[f_a(F(\tau,T))\big\{Z^2(\tau,T) - Z(\tau,T) - \int_0^\tau a^2(t)\sigma^{-2}(t,T)\,dt\big\}\big] \\ &\quad + \frac{d}{dx(T)}\Big(\frac{1}{x(T)}\mathbb{E}\big[H_a(F(\tau,T))F(\tau,T)\big]\Big) \ . \end{aligned}$$

Differentiating in the last expression yields the desired result for the gamma of a call on the spot. □

Proposition 7.5.3. *The delta and gamma of a call on the Asian forward is given by resp.*

$$\begin{aligned} u'(x(T)) &= \frac{1}{x(T)}\mathbb{E}\big[f_a\big(\sum_{k=0}^n F(t_k,T)\big)Z(t_n,T)\big] \\ &\quad + \frac{1}{x(T)}\mathbb{E}\big[H_a\big(\sum_{k=0}^n F(t_k,T)\big)\big(\sum_{k=0}^n F(t_k,T)\big)\big] \\ u''(x(T)) &= \frac{1}{x^2(T)}\mathbb{E}\Big[f_a\big(\sum_{k=0}^n F(t_k,T)\big)\Big\{Z^2(t_n,T) - Z(t_n,T) \\ &\quad - \int_0^{t_n} a^2(t)\sigma^{-2}(t,T)\,dt\Big\}\Big] \\ &\quad + \frac{1}{x^2(T)}\mathbb{E}\Big[H_a'\big(\sum_{k=0}^n F(t_k,T)\big)\big(\sum_{k=0}^n F(t_k,T)\big)^2\Big] \ , \end{aligned}$$

where $Z(t_n,T) = \int_0^{t_n} a(t)\sigma^{-1}(t,T)\,dW(t)$ *and* $a(t)$ *is an integrable function on* $[0,t_n]$ *satisfying* $\int_0^{t_1} a(t)\,dt = 1$ *and* $\int_{t_k}^{t_{k+1}} a(t)\,dt = 0$, $k = 1,\dots,n-1$.

Proof. Using Prop. 7.4.2 and the fact that

$$\frac{d}{dx(T)}F(t_k,T) = \frac{1}{x(T)}F(t_k,T) \ ,$$

give the result. (see proofs for standard calls on forwards above). □

7.6 The Adaptive Method

The adaptive method is introduced to enable better utilization of the sampling points from the QMC method. The principle is very simple: Use more simulation points in the parts of the domain where the integrand fluctuates than in parts of the domain where it is zero or flat (linear). In the papers [Dah01] and [Dah02a] the current adaptive method was presented for multidimensional integrals, but in this paper we shall only need it for one-dimensional integration.

We give a brief overview of the adaptive method limited to one-dimensional problems: It is easy to construct a QMC estimator for the integral g over a part of the integration domain $\mathcal{D}$, and we therefore can construct a method to evaluate the integral over all of $\mathcal{D}$ as a sum of such estimated values. Let $\mathcal{D} = \cup_i D_i$, $\cap_i D_i = \emptyset$, $i = 1, \ldots, P$. Then

$$\begin{aligned} A &= \sum_{i=1}^{P} \int_{D_i} g(x)\, dx \\ &\approx \sum_{i=1}^{P} \frac{|D_i|}{\Delta L_i} \sum_{l=L_{i-1}}^{L_i - 1} g(x_l)\ , \end{aligned}$$

where $\Delta L_i = L_i - L_{i-1}$ and x_l is scaled such that $x_l \in D_i$ when $l \in [L_{i-1}, L_i]$. $|D_i|$ is to be understood as the length of D_i for integration of one-dimensional integrands, and the volume for multi-dimensional integrands. The adaptive algorithm should decide on the number of sub-domains and their sizes, that is P and $|D_i|\,, \forall i$. Furthermore, the algorithm has to pick the best set of sub-domains, and how many simulation points ΔL_i to use in each of them. Alternative approaches use information from the integrand to develop approximations of the integral in sub-domains with a deterministic approach rather than with simulation. This is done in e.g. [BEG91], [GC97], [Coo97], [PF90].

We have chosen to use a binary tree to represent the decomposition of the domain. Each node in the tree corresponds to a distinct part of the domain, and when we expand the tree we divide the domain represented by a node in two parts of equal size. For one dimensional integrands, the criterion we use to decide on division is simply to find the parts of the domain that contributes more than a preset amount to the overall variance of the estimator. The divide and conquer algorithm is terminated when the estimated variability in all sub-domains are less than a preset limit. We estimate the contribution to the variance from each sub-domain by the expression

$$C_i = \frac{|D_i|}{2}\Big(\frac{g(p_1^i) + g(p_2^i)}{2} - g(p_0^i)\Big)\ ,$$

where $|D_i|$ is the length of the sub-domain.

If the adaptive algorithm performs perfectly in accordance with the assumptions, the contribution from each sub-domain to the overall variance should be equal. Therefore $\sigma_i\,|D_i| = c$, $\forall i$ ideally. But even if the adaption process aspire to use the simulation points as effectively as possible, we get some sub-domains in which the measured

variability is close to the preset limit, and some where the variability is considerably lower than the limit. To circumvent this behavior we use less simulation points in the sub-domains where the variability is low. Theoretically, the fraction for the optimal allocation of points in each sub-domain can be shown to be

$$q_i^* = \frac{r_i \sigma_i}{\sum_{l=1}^{P} r_l \sigma_l}, \; i = 1, \ldots, P \; , \tag{7.16}$$

where r_i is the probability for a point to be contained in each bin represented by D_i. Therefore $r_i = |D_i|$ in our setting. This leads to the allocation of simulation points by the relation

$$\Delta L_{i+1} = L \frac{\sigma_i \, |D_i|}{\sum_{l=1}^{P} \sigma_l \, |D_l|} \; , \tag{7.17}$$

where L is the total budget of simulation points. This approach, however, assume knowledge of all σ_i, and the adaptive approach does not provide this knowledge at the stage in the process where the contributions to the value of the total integral are calculated. Instead, we have chosen to implement a simpler approach to finding the number of simulation points in each bin. We use the relation

$$\Delta L_{i+1} = L \frac{\sigma_i}{C} \; , \tag{7.18}$$

where C is the stopping criterion for the adaption process on the variance estimates. This approach avoid the overhead by traversing the tree to collect the σ_l values, and in our tests the approach works well compared to using the same number of simulations in each bin.

For more details on the adaptive method, further variance reduction techniques and numerical simulations on multidimensional problems, see [Dah01] and [Dah02a].

7.7 Numerical Examples and Comparison

In the examples presented below, we set the risk-free rate to zero ($r = 0$), use $\sigma = 0.3$ for the constant σ in (7.5), and calculate the different measures for at-the-money call options with three months left to maturity ($T = 0.25$). The options on the forward contracts are calculated on forwards with six months left to expiration. For the parameters of the Schwartz mean-reverting model we use $\alpha = 0.5, \mu = 5, \lambda = 1.6$. We need to find the strike prices giving us at-the-money options: For contracts on the forward this strike is simply given by $K = x(T)$ since $t \to F(t,T)$ is a martingale in the risk-free setting (recall that $x(T)$ is todays forward curve). For options on spot the at-the-money strike is given by

$$\begin{aligned} K = \mathbb{E}\big[S(T)\big] &= \mathbb{E}\big[\exp(X(T))\big] = \exp\left(\mathbb{E}\big[X(T)\big] + \frac{1}{2}\operatorname{Var}[X(T)]\right) \\ &= \exp\left(\mathrm{e}^{-\alpha T} \ln S(0) + \gamma(1 - \mathrm{e}^{-\alpha T}) + \frac{\sigma^2(1 - \mathrm{e}^{-2\alpha T})}{4\alpha}\right) . \end{aligned}$$

In the simulations we use $x(T) = 100$ for options on forwards, and $S(0) = 100$ for options on spot prices.

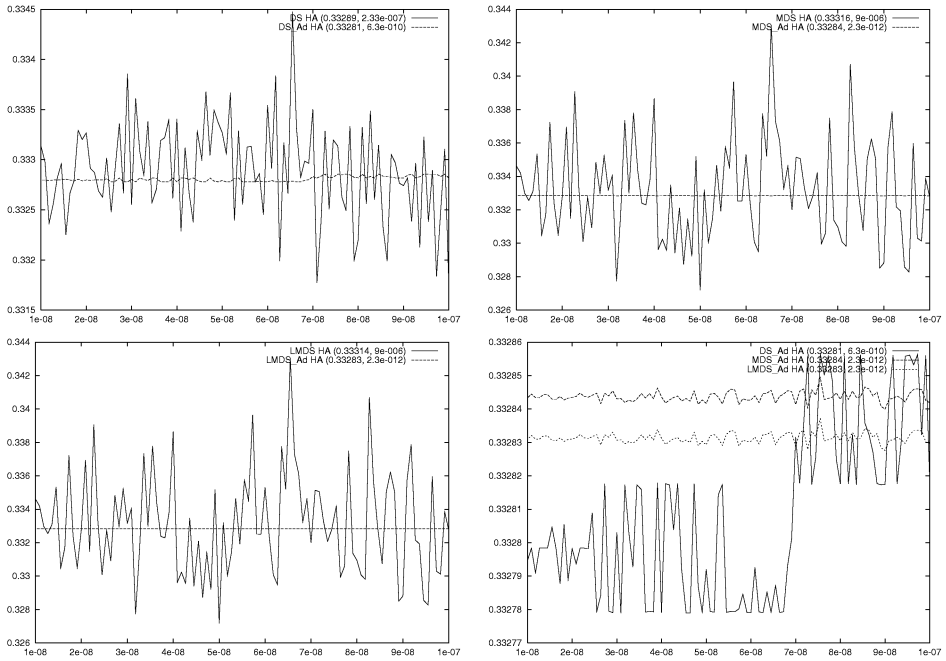


Figure 7.1: Delta for European option on commodity and energy spot

In the label of the plots the two numbers in the parenthesis are mean value and variance of the series of 100 estimators making up the plots. To label the plots we have used the abbreviations; **DS**: Delta Spot, **GS**: Gamma Spot, **VS**: Vega Spot, **DF**: Delta Forward, **GF**: Gamma Forward, **VF**: Vega Forward, **Ad**: Adaptive method is used, **M**: Malliavin approach is used, **LM**: Localized Malliavin approach is used, **HA**: The Halton Leaped method is used as low discrepancy sequence generator (used in all simulations in this article).

Note that when neither the Malliavin or the localized Malliavin method are used, the FD estimator is used. The given abbreviations are combined to indicate the numerical experiment currently investigated. An example of an abbreviation is "LMDF_Ad HA", indicating that the simulation is performed by the localized Malliavin approach for the delta of a forward contract using the adaptive method and the Halton leaped sequence.

The number of samples for each of the 100 estimated values are in the range 800 to 1600, given as a result of the accuracy demanded of the adaptive method (this accuracy is given on the y-axis of the plots, and is in the range [1e-7,1e-8]. For an MC-estimator this is a very low sample size, but for the adaptive method it is

Table 7.1: Sensitivity parameters for European option on commodity and energy spot

Parameter	Method	Value (mean)	Variance
Delta	FD	0.33289	2.33e-007
Delta	FD Adaptive	0.33291	6.30e-010
Delta	Malliavin	0.33316	9.00e-006
Delta	Malliavin Adaptive	0.33284	2.30e-012
Delta	Local Malliavin	0.33314	9.00e-006
Delta	Local Malliavin Adaptive	0.33283	2.30e-012
Gamma	FD	0.01636	6.44e-003
Gamma	FD Adaptive	0.01205	1.29e-004
Gamma	Malliavin	0.015331	5.88e-008
Gamma	Malliavin Adaptive	0.015293	5.47e-014
Vega	FD	14.548	2.31e-003
Vega	FD Adaptive	14.539	5.67e-009
Vega	Malliavin	14.574	4.44e-001
Vega	Malliavin Adaptive	14.54	4.19e-008

Table 7.2: Sensitivity parameters for European option on commodity and energy forward

Parameter	Method	Value (mean)	Variance
Delta	FD	0.52483	7.16e-007
Delta	FD Adaptive	0.52479	1.19e-009
Delta	Malliavin	0.52491	1.86e-005
Delta	Malliavin Adaptive	0.52485	4.17e-012
Delta	Local Malliavin	0.52489	1.86e-005
Delta	Local Malliavin Adaptive	0.52482	1.30e-010
Gamma	FD	0.032778	2.07e-002
Gamma	FD Adaptive	0.027123	3.60e-004
Gamma	Malliavin	0.032031	1.96e-006
Gamma	Malliavin Adaptive	0.032022	1.87e-013
Vega	FD	16.527	2.87e-003
Vega	FD Adaptive	16.526	1.70e-008
Vega	Malliavin	16.531	5.23e-001
Vega	Malliavin Adaptive	16.527	4.98e-008

enough to reach the prescribed accuracy (see [Dah01] for further details on the adaptive method). In [Dah02a] we find an estimator for comparing the efficiency of the adaptive approach with the conventional, where also the extended computing time of

the adaptive method is taken into account. In one-dimensional problems, however, the extended computing time is small. Thus, the method with the lowest variance is preferred in the following cases.

In the last plot of the figs. 7.1 and 7.4, where the accurate estimators resulting from the adaptive method are compared, we see that some of the methods are biased. Based on the fact that the Malliavin approach gives an unbiased estimator, it is evident that the FD and the local Malliavin estimators are biased. This occurs also for the FD estimators of the calculations of vega in figs. 7.2 and 7.6. If we look at figs. 7.3 and 7.5 we see that the FD estimators are not able to capture the discontinuity in the first derivative and give an estimator with large variance.

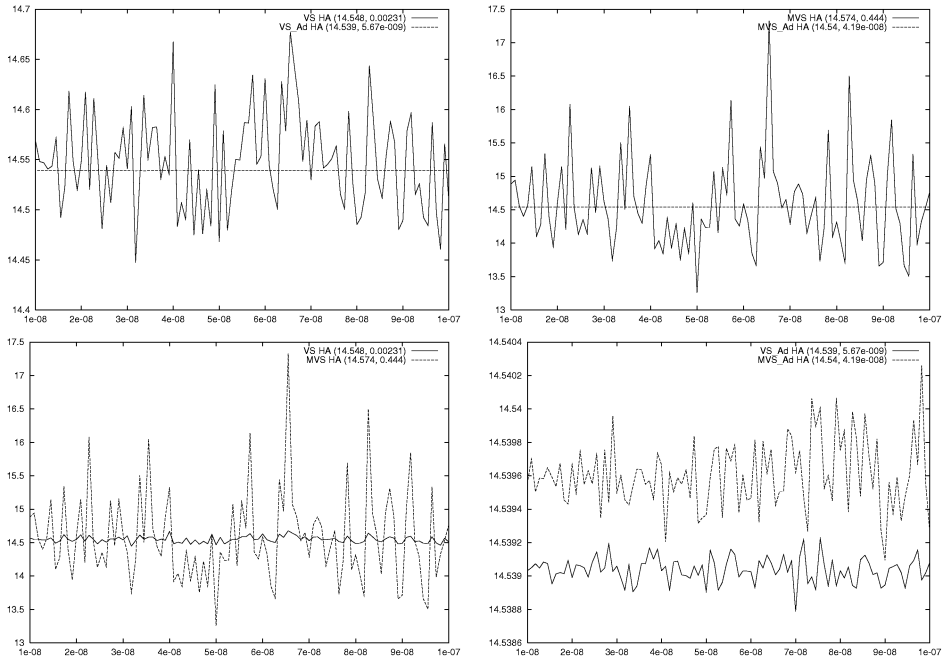


Figure 7.2: Vega for European option on commodity and energy spot

7.8 Conclusion

The article deduces expressions for sensitivities of various derivative instruments in the commodity and energy market. The main focus is on the Malliavin approach, since we by this method can produce unbiased estimators under milder conditions on the payoff functions f than with conventional methods. The numerical results show that there are no apparent sacrifices connected to the Malliavin approach. The

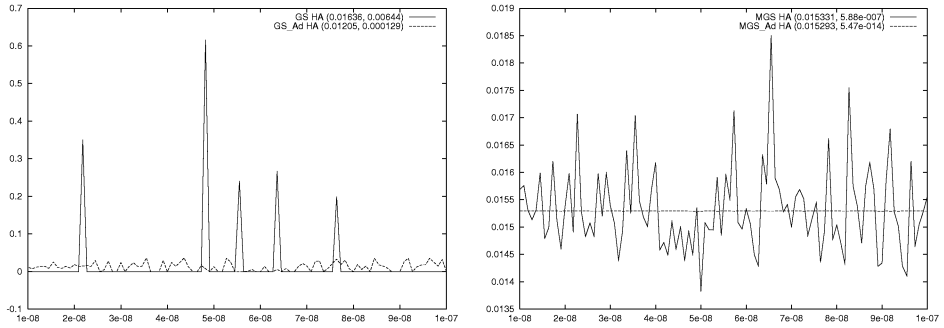


Figure 7.3: Gamma for European option on commodity and energy spot

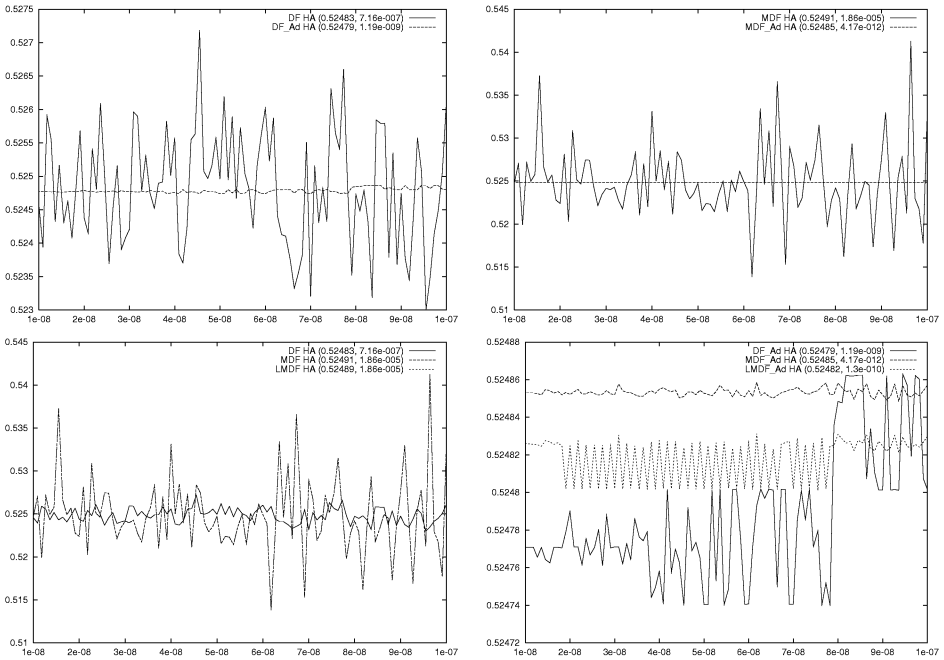


Figure 7.4: Delta for European option on commodity and energy forward

estimators of the Malliavin approach and the conventional FD approach have very similar convergence properties in the cases where they both exist, at least when the adaptive approach is employed to the problems.

The authors are currently working with numerical algorithms to use QMC methods on the estimators for the Asian contracts in the commodity and energy market presented

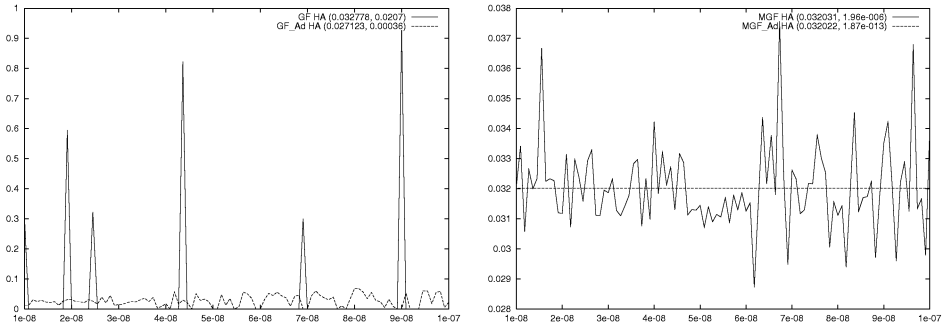


Figure 7.5: Gamma for European option on commodity and energy forward

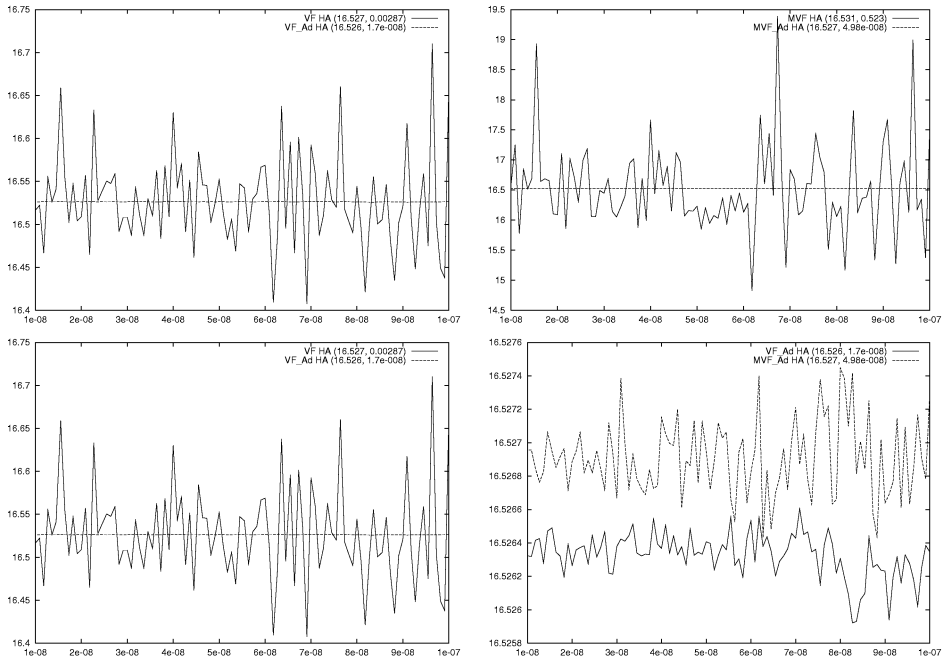


Figure 7.6: Vega for European option on commodity and energy forward

in the paper. Some of the challenges lie in using the low discrepancy sequence in an optimal way. The results of these investigations will be reported elsewhere.

Bibliography

[AS79] I. A. Antonov and V. M. Saleev. An economic method of computing LP_τ-sequences. *USSR Computational Mathematics and Mathematical Physics*, 19:252–256, 1979.

[Ave97] M. Avellaneda. An introduction to option pricing and the mathematical theory of risk. In *Proceedings of the American Mathematical Society Winter Meeting*, 1997.

[BEG91] J. Berntsen, T. O. Espelid, and A. Genz. An adaptive algorithm for the approximate calculation of multiple integrals. *ACM Transactions on Mathematical Software*, 17(4):437–451, 1991.

[Ben02] F. E. Benth. *Matematisk Finans - en Innføring i Opsjonsteori med Stokastisk Analyse*. Universitetsforlaget, 2002.

[BG96] M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42(2):269–285, 1996.

[Caf98] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7:1–49, 1998.

[CGH+96] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5:329–359, 1996.

[CJK97] R. M. Corless, D. J. Jeffrey, and D. E. Knuth. A sequence of series for the Lambert W function. In W. Küchlin, editor, *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, pages 197–204, 1997.

[CM95] R. E. Caflisch and B. Moskowitz. Modified Monte Carlo methods using quasi-random sequences. In H. Niederreiter and P. J. Shieu, editors, *Monte Carlo & Quasi-Monte Carlo Methods in Scientific Computing*, pages 1–16. Springer-Verlag, New York, 1995.

[CM96] R. E. Caflisch and W. Morokoff. Valuation of mortgage backed securities using the quasi-Monte Carlo method. In D. Duffie and D. Tavella, editors, *Proceedings of the First Annual Computational Finance Conference, Stanford*, 1996.

[CMO97] R. E. Caflisch, W. Morokoff, and A. Owen. Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension. *Journal of Computational Finance*, 1:27–46, 1997.

[Coo97] R. Cools. The approximation of low-dimensional integrals: available tools and trends. Technical report, Department of Computer Science, K. U. Leuven, 1997.

[CS00] L. Clewlow and C. Strickland. *Energy Derivatives. Pricing and Risk Management.* Lacima Publications, London, 2000.

[Dah99] L. O. Dahl. Valuation of derivatives consisting of multiple underlying assets by the use of quasi-Monte Carlo methods. Manuscript, 1999.

[Dah00] L. O. Dahl. Valuation of European call options on multiple underlying assets by using a quasi-Monte Carlo method. a case with baskets from Oslo stock exchange. In *Proceedings AFIR 2000*, volume 10, pages 239–248, 2000. (Chap. 2 in this thesis).

[Dah01] L. O. Dahl. An adaptive method for evaluating multidimensional contingent claims. Part I. Submitted to International Journal of Applied and Theoretical Finance. In preprint at http://www.math.uio.no/eprint. (Chap. 5 in this thesis), 2001.

[Dah02a] L. O. Dahl. An adaptive method for evaluating multidimensional contingent claims. Part II. Submitted to International Journal of Applied and Theoretical Finance. In preprint at http://www.math.uio.no/eprint. (Chap. 6 in this thesis), 2002.

[Dah02b] L. O. Dahl. Fast evaluation of derivatives of the Asian basket option by SVD and adaptive integration. Manuscript, 2002.

[DB01] L. O. Dahl and F. E. Benth. Valuation of Asian Basket Options with Quasi-Monte Carlo Techniques and Singular Value Decomposition. Submitted to Journal of Computational Finance. In preprint at http://www.math.uio.no/eprint (Chap. 3 in this thesis), 2001.

[DB02] L. O. Dahl and F. E. Benth. Fast evaluation of the Asian basket option by singular value decomposition. In K. T. Fang, F. J. Hickernell, and H. Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pages 201–214. Springer-Verlag, Berlin, 2002. (Chap. 4 in this thesis).

[DBK02] L. O. Dahl, F. E. Benth, and K. H. Karlsen. On derivatives of claims in commodity and energy markets using a Malliavin approach. In preprint at http://www.math.uio.no/eprint. (Chap. 7 in this thesis), 2002.

[ES81] B. Efron and C. Stein. The jackknife estimator of variance. *Annals of Statistics*, 9:586–596, 1981.

[Fau82] H. Faure. Discrépance de suites associées à un système de numération (en dimension s). *Acta Arithmetica*, 41:337–351, 1982.

[FLL+99] E. Fournié, J. M. Lasry, J. Lebuchoux, P. L. Lions, and N. Touzi. Application of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3:391–412, 1999.

[FLL+01] E. Fournié, J. M. Lasry, J. Lebuchoux, P. L. Lions, and N. Touzi. Application of Malliavin calculus to Monte Carlo methods in finance II. *Finance and Stochastics*, 5:201–236, 2001.

[GC97] A. Genz and R. Cools. An adaptive numerical cubature algorithm for simplices. Technical report, Department of Computer Science, K. U. Leuven, 1997.

[GHS98] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Gaussian importance sampling and stratification: Computational issues. In D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, editors, *Proceedings of the 1998 Winter Simulation Conference*, volume 1, pages 685–693. IEEE Computer Society Press, 1998.

[GHS99a] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance*, 9(2):117–152, 1999.

[GHS99b] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Importance sampling in the Heath-Jarrow-Morton framework. *Journal of Derivatives*, 7(1):32–50, 1999.

[GHS00] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Importance sampling and stratification for value-at-risk. In Y. S. Abu-Mostafa, B. LeBaron, A. W. Lo, and A. S. Weigend, editors, *Computational Finance 1999, Proceedings of the Sixth International Conference on Computational Finance*, pages 7–24. MIT Press, 2000.

[Gla91] P. Glasserman. Gradient estimation via perturbation analysis. Kluwer Academic Press, Norwell, MA, 1991.

[Hal60] J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2:84–90, 1960.

[Hoe48] W. Hoeffding. A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics*, 19:293–325, 1948.

[Hul97] J. C. Hull. *Options Futures and Other Derivatives.* Prentice Hall, third edition, 1997.

[JBT96] C. Joy, P. P. Boyle, and K. S. Tan. Quasi-Monte Carlo methods in numerical finance. *Management Science*, 42(6), 1996.

[KH00] K. Kåresen and E. Husby. A joint state-space model for electricity spot and futures prices. Report 965, Norwegian Computing Center, 2000.

[KN74] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences.* John Wiley & Sons, 1974.

[KP92] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, Berlin, 1992.

[KS98] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance.* Springer-Verlag, Berlin, 1998.

[KW97] L. Kocis and W. J. Whiten. Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software*, 23(2):266–294, 1997.

[Lam91] J. D. Lambert. *Numerical Methods for Ordinary Differential Equations: The Initial Value Problem.* John Wiley & Sons, 1991.

[Lan69] P. Lancaster. *Theory of Matrices.* Academic Press, New York, 1969.

[L'E91] P. L'Ecuyer. An overview of derivative estimation. In *Proceedings of the 1991 Winter Simulation Conference*, pages 207–217. IEEE Press, 1991.

[LL98] C. Lemieux and P. L'Ecuyer. Efficiency improvement by lattice rules for pricing Asian options. In J.S. Carson D.J. Medeiros, E.F. Watson and M.S. Manivannan, editors, *Proceedings of the 1998 Winter Simulation Conference*, 1998.

[LL00a] P. L'Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(9):1214–1235, 2000.

[LL00b] C. Lemieux and P. L'Ecuyer. A comparison of Monte Carlo, lattice rules and other low-discrepancy point sets. In H. Niederreiter and J. Spanier, editors, *Monte Carlo and Quasi-Monte Carlo methods 1998*, pages 326–340. Springer-Verlag, Berlin, 2000.

[MC94] W. J. Morokoff and R. E. Caflisch. Quasi-random sequences and their discrepancies. *SIAM Journal of Scientific Statistical Computing*, 15:1251–1279, 1994.

[MC95] W. J. Morokoff and R. E. Caflisch. Quasi-Monte Carlo integration. *Journal of Computational Physics*, 122:218–230, 1995.

[MC98] W. J. Morokoff and R. E. Caflisch. Quasi-Monte Carlo simulation of random walks in finance. In *Monte Carlo and Quasi-Monte Carlo Methods 1996*, pages 340–352. Springer-Verlag, Berlin, 1998.

[Mor95] B. Moro. The full monte. *Risk*, 8(2):57–58, 1995.

[Nie78] H. Niederreiter. Quasi-Monte Carlo methods and pseudo-random numbers. *Bulletin of American Mathematical Society*, 84:957–1041, 1978.

[Nie87] H. Niederreiter. Point sets and sequences with small discrepancy. *Monatshefte für Mathematik*, 104:273–337, 1987.

[Nie92] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods.* SIAM, 1992.

[Nua95] D. Nualart. *The MAlliavin Calculus and Related Topics.* Springer-Verlag, New York, 1995.

[NX98] H. Niederreiter and C. Xing. The algebraic-geometry approach to low-discrepancy sequences. In *Monte Carlo and Quasi-Monte Carlo Methods 1996*, pages 139–160. Springer-Verlag, Berlin, 1998.

[Ö99] G. Ökten. Error reduction techniques in quasi-Monte Carlo integration. *Mathematical and computer modelling*, 30(7-8):61–69, 1999.

[Øks98] B. Øksendal. *Stochastic Differential Equations, An Introduction with Applications.* Springer-Verlag, Berlin, 1998.

[Owe98] A. B. Owen. Latin supercube sampling for very high-dimensional simulations. *ACM Transactions on Modeling and Computer Simulation*, 8(1):71–102, 1998.

[Owe99] A. B. Owen. Monte Carlo extension of quasi-Monte Carlo. Technical report, Department of Statistics, Stanford University, 1999.

[OZ00] A. Owen and Y. Zhou. Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143, 2000.

[Pag92] G. Pages. Van der Corput sequences, Kakutani transformation and one-dimensional numerical integration. *Journal of Computational and Applied Mathematics*, 44:21–39, 1992.

[PAG98] M. Broadie P. Acworth and P. Glasserman. A comparison of some Monte Carlo and quasi Monte Carlo techniques for option pricing. In *Monte Carlo and Quasi-Monte Carlo Methods 1996.* Springer-Verlag, Berlin, 1998.

[Pas97] S. H. Paskov. New methodologies for valuing derivatives. In S. Pliska and M. Dempster, editors, *Mathematics of Derivative Securities*, pages 545–582. Cambridge University Press, U.K., 1997.

[PF90] W. H. Press and G. R. Farrar. Recursive stratified sampling for multidimensional Monte Carlo integration. *Computers in Physics*, 4(2):190–195, 1990.

[Pil98] D. Pilipović. *Energy Risk. Valuing and Managing Energy Derivatives.* McGraw-Hill, New York, 1998.

[PT95] S. H. Paskov and J. Traub. Faster evaluation of financial derivatives. *Journal of Portfolio Management*, 22(1):113–130, 1995.

[PT97] A. Papageorgiou and J. F. Traub. Faster evaluation of multidimensional integrals. *Computers in Physics*, 11(6):574–578, 1997.

[Rip87] B. D. Ripley. *Stochastic Simulation.* John Wiley & Sons, 1987.

[RS95] L. C. G Rogers and Z. Shi. The value of an Asian option. *Journal of Applied Probability*, 32:1077–1088, 1995.

[Sch97] E. S. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, LII(3):923–973, 1997.

[SJ94] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration.* Oxford Clarendon Press, 1994.

[Sob67] I. M. Sobol. The distribution of points in a cube and approximate evaluation of integrals. *U.S.S.R Computational Mathematics and Mathematical Physics*, 7:86–112, 1967.

[SW98] I. H. Sloan and H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *Journal of Complexity*, 14(1):1–33, 1998.

[Tal97] N. Taleb. *Dynamic Hedging.* John Wiley & Sons, 1997.

[VAD98] F. J. Vázquez-Abad and D. Dufresne. Accelerated simulation for pricing Asian options. In D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, editors, *Proceedings of the 1998 Winter Simulation Conference*, pages 1493–1500, 1998.

[Woz91] H. Wozniakowski. Average case complexity of multivariate integration. *Bulletin of American Mathematical Society*, 24:185–194, 1991.