# Approximations of option price elasticities for importance sampling 

Armin Müller ${ }^{11}$

Discussion paper No. 499

November 24, 2016

Diskussionsbeiträge der Fakultät für Wirtschaftswissenschaft der FernUniversität in Hagen Herausgegeben vom Dekan der Fakultät<br>Alle Rechte liegen bei dem Verfasser


#### Abstract

Importance sampling is a powerful instrument to reduce the standard error of Monte-Carlo estimators. Different importance sampling approaches have in common that the optimal importance sampling probability density is unknown. To approximate this unknown density, in this article we will analyze approximations of option price elasticities. The considered importance sampling approach involves adding an additional drift term. For models with stochastic volatility and for pathdependent options, we show that several approaches exist to achieve considerable variance reduction.


[^0]
## 1 Introduction

Importance sampling allows to significantly decrease the standard error of a MonteCarlo estimator. For a given sample size $N$, unbiased estimators can be calculated with a reduced empirical variance compared to the naive estimator. Most approaches to importance sampling have in common that an approximate knowledge of the distribution of the quantity of interest is required in order to efficiently conduct the simulation [1, 2, 3, 4, 5].

An intuitive interpretation of a formula presented by Singer [5, 6] is the following: In a plain Monte-Carlo simulation of a European call, several simulated paths of the underlying $S$ terminate below the strike price $K$. These realizations with a final value of the pay-off function of zero do not contribute to the Monte-Carlo estimator. However, they do increase its variance. On the one hand, the additional drift introduced by importance sampling helps to "push" an increased amount of paths above the strike price $K$. On the other hand, the additional drift term should not be too high, because then the final values of the simulated paths would be too widely dispersed, again yielding an increased empirical variance. The final value of the trajectory is then adjusted for the likelihood of the modified path to occur under the original measure of probability. Thus, one obtains an unbiased Monte-Carlo estimator with reduced variance.

More specifically, to apply Singer's formula [5], knowledge of the elasticity

$$
\begin{equation*}
\epsilon=\frac{S}{C} \frac{\partial C}{\partial S} \tag{1}
\end{equation*}
$$

of the option price $C$ with respect to the underlying $S$ is required. Generally, in option pricing one conducts Monte-Carlo simulations to estimate $C$. Therefore, ex anteriori $\epsilon$ is an unknown quantity in many cases.

To conduct Monte-Carlo simulations improved by importance sampling applying the mentioned formula, approximations of the option price elasticity $\epsilon$ are required. Option price elasticities, i.e. the ratio between the percentage change of the option price and the related percentage change of an independent variable, are not as broadly discussed in literature as the partial derivatives of option prices (the so-called "Greeks" [7, 8]). This paper will systematically analyze several approximation techniques for option price elasticities. The intention is to give a broad overview of different possible approximations that can be employed in different applications.

It has been previously shown that the approach discussed in this paper can be very well combined with other variance reduction techniques generating synergies
[9]. Furthermore, other approaches, which have been previously discussed in literature, e.g., [2, 4], are outperformed for the pricing of European and arithmetic Asian options [10].

The article is organized as follows: In section 2, we introduce the basic concepts of importance sampling and explicitly derive the variance reduction formula applied in this analysis. Then, we present several approaches to approximate option price elasticities in section 3 . Section 4 summarizes the central results of this study and discusses the need for further research.

## 2 Importance sampling

Transformation of the probability measure Hammersley and Morton (1954) [11] first applied importance sampling techniques to efficiently simulate chain reactions in nuclear reactors. They used a transformation of the probability measure to shift the weight of the simulated sample paths to areas of high interest.
Let $X$ be a random variable with probability density $p$. The expectation value of a function $h: \mathbb{R}^{d} \rightarrow \mathbb{R}, X \rightarrow h(X)$ is defined as

$$
\begin{equation*}
\alpha=\mathbb{E}_{p}[h(X)]=\int h(x) p(x) \mathrm{d} x . \tag{2}
\end{equation*}
$$

This quantity can be estimated using the unbiased Monte-Carlo estimator

$$
\begin{equation*}
\hat{\alpha}_{p}=\frac{1}{n} \sum_{i=1}^{N} h\left(X_{i}\right) \tag{3}
\end{equation*}
$$

for i.i.d. realizations $X_{1}, \ldots, X_{N}$ of the random variable $X$.
It can be recast as

$$
\begin{equation*}
\alpha=\int h(x) p(x) \mathrm{d} x=\int h(x) \frac{p(x)}{p^{\prime}(x)} p^{\prime}(x) \mathrm{d} x=\mathbb{E}_{p^{\prime}}\left[h(X) \frac{p(X)}{p^{\prime}(X)}\right] \tag{4}
\end{equation*}
$$

using any other probability density $p^{\prime}$. The ratio $p / p^{\prime}$ is referred to as Radon-Nikodým derivative or likelihood ratio [3, 12].
The empirical variance of the unbiased Monte-Carlo estimator

$$
\begin{equation*}
\hat{\alpha}_{p^{\prime}}=\frac{1}{n} \sum_{i=1}^{N} h\left(X_{i}\right) \frac{p\left(X_{i}\right)}{p^{\prime}\left(X_{i}\right)} \tag{5}
\end{equation*}
$$

is calculated as

$$
\begin{equation*}
\operatorname{Var}_{p^{\prime}}\left[h(X) \frac{p(X)}{p^{\prime}(X)}\right]=\mathbb{E}_{p^{\prime}}\left[\left(h(X) \frac{p(X)}{p^{\prime}(X)}\right)^{2}\right]-\mathbb{E}_{p^{\prime}}\left[\left(h(X) \frac{p(X)}{p^{\prime}(X)}\right)\right]^{2} . \tag{6}
\end{equation*}
$$

The aim of importance sampling is to minimize this variance term. For a nonnegative function $h$, the density

$$
\begin{equation*}
p^{\prime}(x) \propto h(x) p(x) \tag{7}
\end{equation*}
$$

leads to a vanishing variance term in equation (6). By normalizing, the product $h(x) p(x)$ can be transformed into a probability density. To calculate the normalization factor, the integral

$$
\begin{equation*}
\alpha=\int h(x) p(x) \mathrm{d} x \tag{8}
\end{equation*}
$$

must be calculated. However, our initial goal in equation (2) is exactly the calculation of this quantity. This means, that $\alpha$ generally is a unknown quantity. Notwithstanding, significant reduction of the estimator's empirical variance can already be achieved by approximating this normalization factor [3].

Determination of the additional drift term In this paper, we will further analyze an importance sampling approach introduced by Singer in $2014[5]^{2}$. We present this approach in this section. Our aim is to determine a variance reduced estimator of the Feynman-Kac formule $\sqrt{3}^{3}$

$$
\begin{equation*}
C(X(t), t)=\mathbb{E}\left[e^{-\int_{t}^{T} r(X(\tau), \tau) \mathrm{d} \tau} h[X(T)] \mid X(t)=X_{t}\right] . \tag{9}
\end{equation*}
$$

A more common nomenclature would be to write the option price $C$ as a function of the underlying price $S(t)$ and time $t$, i.e. $C=C(S(t), t)$. Here however, we consider the more general case $C=C(X(t), t)$, where $X(t)$ is a multivariate stochastic process.
Apart from containing the underlying price $S(t)$, this process $X(t)$ could also consist of a stochastic volatility element $\sigma(t)$. We analyze this type of bivariate stochastic process in a variance reduced importance sampling Monte-Carlo simulation of European calls in the Scott model with stochastic volatility.

Also, the stochastic process $X(t)$ can contain the integral of the underlying price $S(t)$ with respect to time as an additional component, i.e. the element $Y(t)=$

[^1]$\int_{0}^{t} S(\tau) \mathrm{d} \tau$. The analysis of arithmetic Asian call options in the Black-Scholes model examines this type of bivariate process ${ }^{4}$.
Going forward, we will assume a constant interest rate $r$ :
\[

$$
\begin{align*}
C(X(t), t) & =e^{-r(T-t)} \mathbb{E}\left[h[X(t)] \mid X(t)=X_{t}\right] \\
& =e^{-r(T-t)} \int h[X(T)] p(X(T), T \mid X(t), t) \mathrm{d} X(T) \tag{10}
\end{align*}
$$
\]

In the univariate case with $X(t)=S(t)$ this is the solution of the Black-Scholes differential equation [15, 16]

$$
\begin{equation*}
\frac{\partial C}{\partial t}+r S \frac{\partial C}{\partial S}+\frac{1}{2} \sigma^{2} S^{2} \frac{\partial^{2} C}{\partial S^{2}}-r C=0 \tag{11}
\end{equation*}
$$

We suppose that the random process $X(t)$ follows an Itō differential equation of the following type:

$$
\begin{equation*}
d X(t)=f[X(t)] \mathrm{d} t+g[X(t)] \mathrm{d} W(t) . \tag{12}
\end{equation*}
$$

Here, $f$ is the drift vector, $t$ the time, $g$ the diffusion coefficient and $W(t)$ a multivariate Wiener process.
This stochastic differential equation can be discretized $\sqrt[5]{5}$. We select a grid $\tau_{k}=$ $t+k \Delta \tau$ with $n$ steps, choosing $\tau_{n}=T$ and $X_{n}=X(T)$ which involves $\tau_{0}=t$ and $X_{0}=X(t)$. With $\Delta \tau=\tau_{k+1}-\tau_{k}$ and $\Delta X=X_{k+1}-X_{k}$, the following Euler-Maruyama scheme results for equation $12 \sqrt{6}$.

$$
\begin{equation*}
X_{k+1}=X_{k}+f_{k} \Delta \tau+g_{k} \Delta W_{k} \tag{13}
\end{equation*}
$$

For small $\Delta \tau$, the transition density between $\tau_{k}$ and $\tau_{k+1}$ can be approximated by normally distributed Euler transition kernels. Its expectation value yields $f_{k} \Delta \tau$

[^2]and its variance $\left|\Omega_{k} \Delta \tau\right|$ with $\Omega_{k}=g_{k} g_{k}{ }^{T}$ :
\[

$$
\begin{align*}
& p\left(X_{k+1}, \tau_{k+1} \mid X_{k}, t_{k}\right) \\
& \approx \frac{1}{\sqrt{2 \pi\left|\Omega_{k} \Delta \tau\right|}} \exp \left\{-\frac{1}{2}\left(\Delta X-f_{k} \Delta \tau\right)^{T}\left|\Omega_{k} \Delta \tau\right|^{-1}\left(\Delta X-f_{k} \Delta \tau\right)\right\} \\
& =\frac{1}{\sqrt{2 \pi\left|\Omega_{k} \Delta \tau\right|}}
\end{aligned} \begin{aligned}
& \exp \left\{-\frac{1}{2} \Delta X^{T}\left|\Omega_{k} \Delta \tau\right|^{-1} \Delta X\right\}  \tag{14}\\
& \times \exp \left\{f_{k}^{T}\left|\Omega_{k}\right|^{-1} \Delta X-\frac{1}{2} f_{k}^{T}\left|\Omega_{k}\right|^{-1} f_{k} \Delta \tau\right\} .
\end{align*}
$$
\]

Consequently, the transition density $p(X(T), T \mid X(t), t)$ in equation (10) is a product of Gaussian Euler transition kernels:

$$
\begin{equation*}
p(X(T), T \mid X(t), t) \approx \int \ldots \int \prod_{k=0}^{n-1} p\left(X_{k+1}, \tau_{k+1} \mid X_{k}, \tau_{k}\right) \mathrm{d} X_{n-1} \ldots \mathrm{~d} X_{1} \tag{15}
\end{equation*}
$$

Therefore, equation (10) can be approximated as

$$
\begin{align*}
C(X(t), t)=C\left(X_{0}, \tau_{0}\right) \approx & e^{-r\left(\tau_{n}-\tau_{0}\right)} \int \ldots \int h\left(X_{n}\right) p\left(X_{n}, \tau_{n} \mid X_{n-1}, \tau_{n-1}\right) \\
& \quad \ldots \times p\left(X_{1}, \tau_{1} \mid X_{0}, \tau_{0}\right) \mathrm{d} X_{n} \ldots \mathrm{~d} X_{1}  \tag{16}\\
= & : e^{-r\left(\tau_{n}-\tau_{0}\right)} \tilde{C}\left(X_{0}, \tau_{0}\right) .
\end{align*}
$$

This representation allows for numerical calculations.
The density (15) solves the Kolmogorov backward equation ${ }^{[7]}$ with the differential operator $\mathcal{L}^{8}$ :

$$
\begin{align*}
\frac{\partial p}{\partial \tau_{0}} & =-\left[f_{\alpha} \frac{\partial}{\partial X_{0 \alpha}}+\frac{1}{2} \Omega_{\alpha \beta} \frac{\partial^{2}}{\partial X_{0 \alpha} \partial X_{0 \beta}}\right] p  \tag{17}\\
& =-\mathcal{L} p
\end{align*}
$$

In accordance with equation (7), for the optimal density $p^{\prime}=p^{\text {opt }}$ we define

$$
\begin{align*}
p^{\mathrm{opt}} & =\frac{p h}{\mathbb{E}_{p}[h]}  \tag{18}\\
& =\frac{p h}{\tilde{C}} .
\end{align*}
$$

[^3]Going forward, we assume that $h$ does not explicitly depend on $\tau_{0}$ and $X_{0}$. The central element of the derivation of the additional drift term is the assumption, that $p^{\text {opt }}$ fulfills an alternative Kolmogorov backward equation with the differential operator $\mathcal{L}^{\mathrm{opt}}$, drift vector $f^{\text {opt }}$ and an identical diffusion matrix as $p$, i.e. $g^{\mathrm{opt}} \equiv g$ :

$$
\begin{equation*}
p^{\mathrm{opt}} \tau_{\tau_{0}}=-\mathcal{L}^{\mathrm{opt}} p^{\mathrm{opt}} \tag{19}
\end{equation*}
$$

Here, the abbreviation $\frac{\partial p^{\text {opt }}}{\partial \tau_{0}}:=p^{\text {opt }}{ }_{\tau_{0}}$ has been introduced which will be used going forward. Using $p_{\tau_{0}}=-\mathcal{L} p$ and consequently $\tilde{C}_{\tau_{0}}=-\mathcal{L} \tilde{C}$, we obtain

$$
\begin{equation*}
p^{\mathrm{opt}}{ }_{\tau_{0}}=h\left[-\frac{1}{\tilde{C}} \mathcal{L} p+\frac{1}{\tilde{C}^{2}} p \mathcal{L} \tilde{C}\right]=-h \mathcal{L}^{\text {opt }}\left(\frac{p}{\tilde{C}}\right) \tag{20}
\end{equation*}
$$

and after canceling out $h$ and rearranging terms

$$
\begin{equation*}
\frac{1}{\tilde{C}} \mathcal{L} p-\frac{1}{\tilde{C}^{2}} p \mathcal{L} \tilde{C}=\mathcal{L}^{\mathrm{opt}}\left(\frac{p}{\tilde{C}}\right) \tag{21}
\end{equation*}
$$

Applying the quotient rule for $L^{9}$ yields

$$
\begin{equation*}
\mathcal{L}^{\mathrm{opt}}\left(\frac{p}{\tilde{C}}\right)=\frac{1}{\tilde{C}} \mathcal{L}^{\mathrm{opt}} p-\frac{p}{\tilde{C}^{2}} \mathcal{L}^{\mathrm{opt}} \tilde{C}+\Omega_{\alpha \beta}\left(-\frac{1}{\tilde{C}^{2}} p_{, \alpha} \tilde{C}_{, \beta}+\frac{p}{\tilde{C}^{3}} \tilde{C}_{, \alpha} \tilde{C}_{, \beta}\right) \tag{22}
\end{equation*}
$$

and by defining $\mathcal{L}^{\text {opt }}-\mathcal{L}:=\delta f_{\alpha} \frac{\partial}{\partial X_{0_{\alpha}}}$ we obtain

$$
\begin{equation*}
\delta f_{\alpha}\left(\frac{p_{, \alpha}}{p}-\frac{\tilde{C}_{, \alpha}}{\tilde{C}}\right)+\Omega_{\alpha \beta}\left(-\frac{p_{, \alpha}}{p}+\frac{\tilde{C}_{, \alpha}}{\tilde{C}}\right) \frac{\tilde{C}_{, \beta}}{\tilde{C}}=0 . \tag{23}
\end{equation*}
$$

The additional drift term

$$
\begin{equation*}
\delta f=\Omega \frac{\nabla \tilde{C}}{\tilde{C}}=\Omega \frac{\nabla C}{C}=\Omega \nabla \ln C \tag{24}
\end{equation*}
$$

solves this equation.
A new stochastic differential equation with additional drift term similar to equation (12) results:

$$
\begin{align*}
d X(t) & =f[X(t)]^{\text {opt }} \mathrm{d} t+g[X(t)] \mathrm{d} W(t) \\
& =\left(f[X(t)]+\Omega \frac{\nabla C}{C}\right) \mathrm{d} t+g[X(t)] \mathrm{d} W(t) \tag{25}
\end{align*}
$$

[^4]In the univariate case we obtain

$$
\begin{equation*}
d S(t)=\left(f[S(t)]+\frac{g[S(t)]^{2}}{C} \frac{\partial C}{\partial S}\right) \mathrm{d} t+g[S(t)] \mathrm{d} W(t) \tag{26}
\end{equation*}
$$

In the Black-Scholes model with $f=r S$ and $g=\sigma S$ and using the abbreviation for the option price elasticity from equation (1), this equation simplifies to

$$
\begin{equation*}
d S(t)=\left(r+\epsilon \sigma^{2}\right) S(t) \mathrm{d} t+\sigma S(t) \mathrm{d} W(t) \tag{27}
\end{equation*}
$$

In the previous formulae, $C$ is the Feynman-Kac formula from equation (10). The problem from equation (8) materializes again: knowledge of $C$ is required to describe the optimal stochastic differential equation. Approaches how to cope with this problem are the focus of section 3 .

Calculation of the Radon-Nikodým derivative In order to evaluate importance sampling estimators, the Radon-Nikodým derivative introduced in equation (4) must be calculated. From equation (14) we obtain

$$
\begin{align*}
\frac{p}{p^{\mathrm{opt}}=} & \exp \left\{f_{k}^{T}\left|\Omega_{k}\right|^{-1} \Delta X-\frac{1}{2} f_{k}^{T}\left|\Omega_{k}\right|^{-1} f_{k} \Delta \tau\right\} \\
& \times \exp \left\{-f_{k}^{\text {opt } T}\left|\Omega_{k}\right|^{-1} \Delta X+\frac{1}{2} f_{k}^{\text {opt } T}\left|\Omega_{k}\right|^{-1} f_{k}^{\text {opt }} \Delta \tau\right\} \\
= & \exp \left\{\left(f_{k}-f_{k}^{\text {opt }}\right)^{T}\left|\Omega_{k}\right|^{-1}\left[\Delta X-\frac{1}{2}\left(f_{k}+f_{k}^{\text {opt }}\right) \Delta \tau\right]\right\}  \tag{28}\\
= & \exp \left\{-\delta f_{k}^{T}\left|\Omega_{k}\right|^{-1}\left[f_{k}^{\text {opt }} \Delta \tau+g_{k} \Delta W_{k}-\frac{1}{2}\left(f_{k}+f_{k}^{\text {opt }}\right) \Delta \tau\right]\right\} \\
= & \exp \left\{-\frac{1}{2} \delta f_{k}^{T}\left|\Omega_{k}\right|^{-1} \delta f_{k} \Delta \tau-\delta f_{k}^{T}\left|\Omega_{k}\right|^{-1} g_{k} \Delta W_{k}\right\}
\end{align*}
$$

Importance sampling of European calls in the Scott model In section 3, we will conduct a profound analysis of importance sampling applied to Monte-Carlo simulations of European call options in a model with stochastic volatility.

The call considered has the pay-off function

$$
\begin{equation*}
C_{T}=\left(S_{T}-K\right)^{+} \equiv \max \left(S_{T}-K, 0\right) . \tag{29}
\end{equation*}
$$

The dynamics of stock prices follow a model introduced by Scott in 1987 [18]. Two stochastic differential equations describe not only the price of the underlying
itself but also its volatility:

$$
\begin{align*}
\mathrm{d} S & =r S \mathrm{~d} t+\sigma S \mathrm{~d} W_{1} \\
\mathrm{~d} \sigma & =\lambda(\bar{\sigma}-\sigma) \mathrm{d} t+\eta \mathrm{d} W_{2} \tag{30}
\end{align*}
$$

The first of these two equations corresponds to the stochastic differential equation of the underlying price in the Black-Scholes model. The second equation describes the volatility itself as a stochastic processes. $W_{1}$ and $W_{2}$ are Wiener processes, $\lambda$ and $\eta$ are additional parameters that specify the dynamics of the stochastic volatility process.

This bivariate stochastic process can be written as a vector process:

$$
\underbrace{\left[\begin{array}{c}
\mathrm{d} S  \tag{31}\\
\mathrm{~d} \sigma
\end{array}\right]}_{\mathrm{d} X}=\underbrace{\left[\begin{array}{c}
r S \\
\lambda(\bar{\sigma}-\sigma)
\end{array}\right]}_{f} \mathrm{~d} t+\underbrace{\left[\begin{array}{cc}
\sigma S & 0 \\
0 & \eta
\end{array}\right]}_{g} \underbrace{\left[\begin{array}{c}
\mathrm{d} W_{1} \\
\mathrm{~d} W_{2}
\end{array}\right]}_{\mathrm{d} W}
$$

Adding the additional drift term as given in equation (25) yields

$$
\left[\begin{array}{c}
\mathrm{d} S  \tag{32}\\
\mathrm{~d} \sigma
\end{array}\right]=(\left[\begin{array}{c}
r S \\
\lambda(\bar{\sigma}-\sigma)
\end{array}\right]+\underbrace{\left[\begin{array}{c}
\sigma^{2} S^{2} C_{S} / C \\
\eta^{2} C_{\sigma} / C
\end{array}\right]}_{\delta f}) \mathrm{d} t+\left[\begin{array}{cc}
\sigma S & 0 \\
0 & \eta
\end{array}\right]\left[\begin{array}{c}
\mathrm{d} W_{1} \\
\mathrm{~d} W_{2}
\end{array}\right] .
$$

For the option price elasticity with respect to the underlying, we use the abbreviation

$$
\begin{equation*}
\epsilon:=\frac{S}{C} \frac{\partial C}{\partial S} \tag{1}
\end{equation*}
$$

The abbreviation

$$
\begin{equation*}
\chi:=\frac{1}{C} \frac{\partial C}{\partial \sigma} \tag{33}
\end{equation*}
$$

is not a price elasticity in the strict sense as there appears no $\sigma$ in the numerator of the first fraction. Therefore, going forward we will refer to $\chi$ as a "pseudoelasticity".

Using these two abbreviations, one obtains

$$
\left[\begin{array}{c}
\mathrm{d} S  \tag{34}\\
\mathrm{~d} \sigma
\end{array}\right]=\left[\begin{array}{c}
r S+\sigma^{2} \epsilon S \\
\lambda(\bar{\sigma}-\sigma)+\eta^{2} \chi
\end{array}\right] \mathrm{d} t+\left[\begin{array}{cc}
\sigma S & 0 \\
0 & \eta
\end{array}\right]\left[\begin{array}{l}
\mathrm{d} W_{1} \\
\mathrm{~d} W_{2}
\end{array}\right] .
$$

For the purpose of Monte-Carlo simulations, this vector equation must be dis-
cretized. As before, we choose an Euler-Maruyama approximation and obtain

$$
\left[\begin{array}{c}
S_{k+1}  \tag{35}\\
\sigma_{k+1}
\end{array}\right]=\left[\begin{array}{l}
S_{k} \\
\sigma_{k}
\end{array}\right]+\left[\begin{array}{c}
r S_{k}+\sigma_{k}^{2} \epsilon_{k} S_{k} \\
\lambda\left(\bar{\sigma}-\sigma_{k}\right)+\eta^{2} \chi_{k}
\end{array}\right] \Delta \tau+\left[\begin{array}{cc}
\sigma_{k} S_{k} & 0 \\
0 & \eta
\end{array}\right]\left[\begin{array}{c}
\Delta W_{1 k} \\
\Delta W_{2 k}
\end{array}\right] .
$$

Using equations (5), (15) and (28), for the importance sampling Monte-Carlo simulation with sample size $N$ and $n$ discretization steps, we obtain the variance reduced estimator with bivariate additional drift

$$
\begin{align*}
\hat{C}_{\text {European Call }}^{\text {Scott model }} & \\
=\frac{1}{N} \sum_{i=1}^{N} \exp & \left\{-\sum_{k=0}^{n-1}\left[\sigma \epsilon_{i k} \Delta W_{1 i k}+\frac{1}{2} \sigma^{2} \epsilon_{i k}^{2} \Delta \tau+\eta \chi_{i k} \Delta W_{2 i k}+\frac{1}{2} \eta^{2} \chi_{i k}^{2} \Delta \tau\right]\right\} \\
& \times e^{-r(T-t)}\left(S_{i n}-K\right)^{+} \tag{36}
\end{align*}
$$

The following special cases can be considered:

- By setting $\epsilon_{i k} \equiv 0$ and $\chi_{i k} \equiv 0$ for all $i, k$, one recovers the plain MonteCarlo estimator corresponding to the original stochastic differential equation of $X(t)$ without additional drift term.
- By setting only $\chi_{i k} \equiv 0$ for all $i, k$, one obtains an importance sampling Monte-Carlo estimator with univariate additional drift term. Only the first component of the stochastic differential equation of $X(t)$, i.e. the equation of the underlying $S(t)$, is modified.
- One might also choose $\epsilon_{i k} \equiv 0$ for all $i, k$, but allow $\chi_{i k} \neq 0$. This would imply adding a drift term to the volatility component only and keeping unchanged the price component. However, this approach is not pursued further in this article.


## Importance sampling of arithmetic Asian calls in the Black-Scholes model

Analogously, we derive an importance sampling Monte-Carlo estimator for arithmetic Asian calls in the Black-Scholes model.

The pay-off function ${ }^{10}$ of the considered long call position is

$$
\begin{equation*}
C_{T}=\left(\bar{S}_{T}-K\right)^{+} \text {with } \bar{S}_{T}=\frac{1}{n} \sum_{l=1}^{n} S_{l} \text {. } \tag{37}
\end{equation*}
$$

Using the same abbreviation as for European calls in the Scott model, we obtain the following stochastic differential equation for the underlying price:

$$
\begin{equation*}
\mathrm{d} S=\left(r+\sigma^{2} \epsilon\right) S \mathrm{~d} t+\sigma S \mathrm{~d} W \tag{38}
\end{equation*}
$$

Discretization yields

$$
\begin{equation*}
S_{k+1}=S_{k}+\left(r+\sigma^{2} \epsilon_{k}\right) S_{k} \Delta \tau+\sigma S_{k} \Delta W_{k} . \tag{39}
\end{equation*}
$$

We obtain the estimator

$$
\begin{align*}
& \hat{C}_{\text {Arithmetic Asian Call }}^{\text {Black-Scholes model }}(S(t), t) \\
& =\frac{1}{N} \sum_{i=1}^{N} \exp  \tag{40}\\
& \left.\times-\sum_{k=0}^{n-1}\left[\sigma \epsilon_{i k} \Delta W_{i k}+\frac{1}{2} \sigma^{2} \epsilon_{i k}^{2} \Delta \tau\right]\right\} \\
&
\end{align*}
$$

Again, we can set $\epsilon_{i k} \equiv 0$ for all $i, k$ and recover the plain Monte-Carlo estimator. A distinction between a univariate and bivariate additional drift is not applicable in this case, as a constant volatility is assumed.

In the next section, we examine different approximation approaches for the elasticity $\epsilon$ and the pseudo-elasticity $\chi$.

## 3 Approximations of option price elasticities

### 3.1 General remarks

Ex anteriori, the option price and its elasticities are unknown quantities. The aim of Monte-Carlo simulations is the actual determination of a specific option price. Generally, for the determination of the option price elasticity, knowledge of a relevant part of the option price curve and its first partial derivative is required. However, in order to conduct variance reduced Monte-Carlo simulations using equations (35) or

[^5](39) and to calculate estimators as given in equations (36) or (40), specific values for $\epsilon$ and $\chi$ values are needed.

Therefore, we approximate the option price elasticity $\epsilon$ and the pseudo-elasticity $\chi$. In this chapter, we present a range of such approximations suitable for importance sampling. We present numerical results in terms of variance reduction ratios, i.e. the ratios between empirical variances of plain Monte-Carlo estimators and the empirical variances of variance reduced Monte-Carlo estimators. Also, we analyze the $\epsilon$ and $\chi$ surfaces as a function of the two independent variables $S$ and the time until expiration $T-t$.

### 3.2 Approximation by constant values

The least sophisticated approach is to approximate both $\epsilon$ and $\chi$ by constant values for all $S$ values and all discretization steps [5, 19, 20]. This approach satisfies the first element of the intuitive interpretation discussed in the introduction of this article: By increasing the drift term using a constant elasticity approximation, more sample paths are pushed above the strike price $K$. Thus, an increased amount of sample paths contributes to the option price estimator. One might argue that the higher the additional drift term, the more paths will terminate above the strike price $K$. However, there are two arguments against very high additional drift terms: First, the additional drift increases the exponential growth component of the stochastic process. This increases the dispersion of simulation results counteracting the intended effect of importance sampling. Second, for a given discretization with $n$ steps, an increasing additional drift term also increases the discretization bias inherent to Monte-Carlo simulations of discretized Euler-Maruyama schemes ${ }^{11}$,
A numerical example (see figure 1) reflects this line of thought: As shown in subfigure (a), for a European call simulated in the Scott model, a constant additional term with $\epsilon=1.0$ leads to only modest variance reduction. It is interesting to observe that for high volatility environments the constant approximation delivers higher variance reduction over a broad range of starting values of the underlying $S(0)$ than for low volatility environments.
For a higher additional drift term choosing $\epsilon=10.0$, the observations are different (see subfigure (b)). For low and intermediate volatilities, a strong variance reduction is observed for out-of-the-money options. However, for higher starting values of the underlying the variance is increased. This is plausible, as for small $S(0)$ values the high additional drift term helps to push more trajectories above the strike price $K$. On the other hand, for high $S(0)$ values this effect is not required

[^6]
## Variance reduction: approximation by constant values



Figure 1: (a) European call, Scott model, $\epsilon=1.0$ : Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating a constant option price elasticity $\epsilon=1.0 . r=0.05$, $T=1, K=10, n=100, N=10,000, \lambda=0.25$ and $\eta=0.01$. Red line: $\bar{\sigma}=$ 0.1. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. Solid line: univariate additional drift, i.e. $\chi=0$. Dashed line: bivariate additional drift with $\chi=0.1$. Black line: reference line where the variance reduction equals 1.0. (b) European call, Scott model, $\epsilon=10.0$ : As (a), but $\epsilon=10.0$. (c) Asian call, Black-Scholes model, $\epsilon=1.0$ : Variance reduction factor on a logarithmic scale for an arithmetic Asian call in the Black-Scholes model achieved by importance sampling approximating a constant option price elasticity $\epsilon=1.0 . r=0.05, T=1, K=10, n=100$ and $N=10,000$. Red line: $\sigma=0.1$. Blue line: $\sigma=0.2$. Green line: $\sigma=0.5$. Black line: reference line where the variance reduction equals 1.0. (d) Asian call, Black-Scholes model, $\epsilon=10.0$ : As (c), but $\epsilon=10.0$.

## Elasticities approximated by constant values

(a) $\varepsilon$-surface

(b) $\chi$-surface


Figure 2: (a) $\epsilon$-surface: Surface of a constant option price elasticity $\epsilon$ with respect to the underlying price $S$ as a function of $S$ and the remaining time until expiration $T-t$. (b) $\chi$-surface: Surface of a constant pseudo-elasticity $\chi$ with respect to the volatility $\sigma$ as a function of $S$ and the remaining time until expiration $T-t$.
any more and the additional drift term only leads to more disperse contributions to the Monte-Carlo estimator. In high volatility environments, this approach does not seem to work at all. A very shaky variance reduction curve indicates numerical instabilities, likely resulting from a discretization bias favored by the high increments per discretization step due to the additional drift term.

In both cases (subfigures (a) and (b)), the addition of a bivariate drift term does not appear to add any value. Both for low and for high starting values of the underlying no significant additional variance reduction is observed (dashed line) vis-à-vis the univariate case (solid line). So, the additional computational effort introduced by calculating a bivariate likelihood ratio from equation (28) does not appear to be justified.

For arithmetic Asian options (see subfigures (c) and (d) of figure 1), the same observations basically hold. A low $\epsilon$ value leads to moderate variance reduction across the whole range of $S(0)$ values, while a comparably high $\epsilon$ value leads only to variance reduction for out-of-the-money options. In the latter case, the approach does not work well for high volatility environments. As the simulation was conducted in the Black-Scholes model, no bivariate additional drift is considered.

For subfigures (b) and (d), the $\epsilon$ - and the $\chi$-surface resp. the $\epsilon$-surface are shown

Variance reduction: approximation by step function


Figure 3: (a) European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating $\epsilon$ and $\chi$ by step functions with $\epsilon_{\text {low }}=1.0$ and $\chi_{\text {low }}=0$ for $S(t) \leq K e^{-r(T-t)}$ and $\epsilon_{\text {high }}=10.0$ and $\chi_{\text {high }}=10.0$ otherwise. $r=0.05, T=1$, $K=10, n=100, N=10,000, \lambda=0.25$ and $\eta=0.01$. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. Solid line: univariate additional drift, i.e. $\chi=0$. Dashed line: bivariate additional drift. Black line: reference line where the variance reduction equals 1.0. (b) Asian call, Black-Scholes model: Variance reduction factor on a logarithmic scale for an arithmetic Asian call in the Black-Scholes model achieved by importance sampling approximating $\epsilon$ by a step function with $\epsilon_{\text {low }}=1.0$ for $\bar{S}(t) \leq K e^{-r(T-t)}$ and $\epsilon_{\text {high }}=10.0$ otherwise. $r=0.05, T=1, K=10, n=100$ and $N=10,000$. Red line: $\sigma=0.1$. Blue line: $\sigma=0.2$. Green line: $\sigma=0.5$. Black line: reference line where the variance reduction equals 1.0.
in figure 2

### 3.3 Approximation by step function

A first extension compared to the approximation by a constant value is the application of a step function: For a constant approximation, we observe that a comparably high $\epsilon$ value works well for out-of-the-money options, while for in-the-money options the variance is increased. In this regime, a comparably low $\epsilon$ value turns out to be more appropriate. A natural enhancement is therefore to approximate the options price elasticity by a step function with two plateaus: a higher one for

## Elasticities approximated by step function

(a) $\varepsilon$-surface

(b) $\chi$-surface


Figure 4: (a) $\epsilon$-surface: Surface of the option price elasticity $\epsilon$ with respect to the underlying price $S$ approximated by a step function as a function of $S$ and the remaining time until expiration $T-t$ as defined in equation (41). (b) $\chi$-surface: Surface of the constant pseudo-elasticity $\chi$ with respect to the volatility $\sigma$ approximated by a step function as a function of $S$ and the remaining time until expiration $T-t$ as defined in equation $(41)$.
$S(t) \leq K e^{-r(T-t)}$ and a lower one for $S(t)>K e^{-r(T-t)}$. We introduce a step function approximation using the Heaviside function $\theta(x)^{12}$ as indicated in figure 4.

$$
\begin{align*}
\epsilon & =\epsilon_{\text {low }}+\theta\left(K-S e^{-r(T-t)}\right) \cdot\left(\epsilon_{\text {high }}-\epsilon_{\text {low }}\right) \\
\chi & =\chi_{\text {low }}+\theta\left(K-S e^{-r(T-t)}\right) \cdot\left(\chi_{\text {high }}-\chi_{\text {low }}\right) \tag{41}
\end{align*}
$$

As shown in figure 3, subfigure (a), for European calls in the Scott model we obtain a "best of" of the subfigures (a) and (c) in figure 1. For out-of-the-money options, the higher $\epsilon$ value leads to considerable variance reduction, while for in-the-money-options the lower $\epsilon$ value still achieves a modest variance reduction. Again, for high volatility environments, the results of variance reduced Monte-Carlo simulations become numerically unstable. As for the constant approximation, the choice of a bivariate drift term does not justify the additional computational effort.

The same results hold true for the arithmetic Asian call simulated in the Black-

[^7]Scholes model (see subfigure (b)). Interestingly, for these path-dependent options for which no analytical formula is known, approximating $\epsilon$ by a step function yields the best results in terms of variance reduction. This method outperforms more complex and numerically more demanding approaches discussed in the following subsections 3.4-3.7.

Variance reduction: approximation by lower bound


Figure 5: (a) European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating $\epsilon$ by the option's lower bound. $r=0.05, T=1, K=10$, $n=100, N=10,000, \lambda=0.25$ and $\eta=0.01$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10]$. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. No bivariate simulation was conducted, as the partial derivative of a European call option's lower bound with respect to $\sigma$ is zero and therefore $\chi=0$. Black line: reference line where the variance reduction equals 1.0. (b) Asian call, Black-Scholes model: Variance reduction factor on a logarithmic scale for an arithmetic Asian call in the Black-Scholes model achieved by importance sampling approximating $\epsilon$ by the option's lower bound. $r=0.05, T=1, K=10, n=100$ and $N=10,000$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10]$. Red line: $\sigma=0.1$. Blue line: $\sigma=0.2$. Green line: $\sigma=0.5$. Black line: reference line where the variance reduction equals 1.0.

### 3.4 Approximation by lower bound

A further extension is the approximation of the option price elasticity $\epsilon$ using the option's lower bound [19]

$$
\begin{equation*}
C_{\mathrm{lb}}=\left(S-K e^{-r(T-t)}\right)^{+} \tag{42}
\end{equation*}
$$

The inequality $C>\left(S-K e^{-r(T-t)}\right)^{+}$was introduced by Merton in 1973 [7, 21]. Especially for high volatilities $\sigma$ and long times to maturity $T-t$, for at-
$\varepsilon$-surface approximated by lower bound


Figure 6: Surface of the logarithmized option price elasticity $\epsilon$ with respect to the underlying price $S$ for European calls approximated by the option's lower bound as a function of $S$ and the remaining time until expiration $T-t$.
the-money options the difference between to the true option value and the lower bound $C_{\mathrm{lb}}$ can become considerable. Furthermore, the lower bound and its the partial derivative with respect to the underlying vanish for $S(t) \leq K e^{-r(T-t)}$. In this sense, for out-of-the-money options and for at-the-money options, the approximation of $\epsilon$ employing the lower bound $C_{\mathrm{lb}}$ is a very crude one. However, for deep-in-the-money options the lower bound and its partial derivative with respect to $S$ asymptotically approximate the option price curve.
For the option price elasticity of European calls $\epsilon$ we obtain

$$
\epsilon=\frac{S}{C_{\mathrm{lb}}} \frac{\partial C_{\mathrm{lb}}}{\partial S}= \begin{cases}\left(1-\frac{K e^{-r(T-t)}}{S(t)}\right)^{-1} & \text { for } S(t)>K e^{-r(T-t)}  \tag{43}\\ \frac{0}{0} & \text { for } S(t) \leq K e^{-r(T-t)}\end{cases}
$$

This expression is not defined for $S(t) \leq K e^{-r(T-t)}$. For this regime, we make the ad hoc assumption that the elasticity $\epsilon$ should at least push the trajectories to the level $S(T)=K$ [19]. We obtain

$$
\begin{equation*}
S(t) e^{\left(r+\sigma^{2} \epsilon\right)(T-t)} \stackrel{!}{=} K \Leftrightarrow \epsilon=\frac{1}{\sigma^{2}(T-t)} \ln \left[\frac{K e^{-r(T-t)}}{S(t)}\right] . \tag{44}
\end{equation*}
$$

Combining equations (43) and (44) we obtain

$$
\epsilon(S(t), t)= \begin{cases}\left(1-\frac{K e^{-r(T-t)}}{S(t)}\right)^{-1} & \text { for } S(t)>K e^{-r(T-t)}  \tag{45}\\ \frac{1}{\sigma^{2}(T-t)} \ln \left[\frac{K e^{-r(T-t)}}{S(t)}\right] & \text { for } S(t) \leq K e^{-r(T-t)}\end{cases}
$$

The surface of this function is shown in figure 6. Close to the discontinuity at $S(t)=K e^{-r(T-t)}$, we get comparably low $\epsilon$ values for long times to maturity and comparably high $\epsilon$ values for short times to maturity. In subsection 3.2, numerical results illustrate that $\epsilon=1.0$ achieves variance reduction across the whole range of $S(0)$ values. Other empirical studies come to the conclusion that $\epsilon=1.0$ is a lower bound for option price elasticities for effective importance sampling [19]. Additionally, as discussed in subsection 3.2, too high $\epsilon$ values lead to numerical instabilities and to a strong discretization bias. Therefore, in the numerical example presented in figure 5, we truncate $\epsilon$ to the interval [1.0; 10.0]. For instance, in the same simulation we can alternatively allow $\epsilon$ to take on values in the interval [ $1 ; 10,000]$. However, in this case we observe a strong variation of the variance reduction curves and a significant variance increase for several $S(0)$ values. To avoid these numerical instabilities, a smaller upper bound is required.

The formula of the lower bound $C_{\mathrm{lb}}=S-K e^{-r(T-t)}$ does not depend on the volatility $\sigma$. Put in another way, $\sigma$ only influences the difference between $C$ and $C_{\mathrm{lb}}$. Therefore, we obtain the partial derivative $\left(C_{\mathrm{lb}}\right)_{\sigma}=0$. Thus, a meaningful additional bivariate drift term can't be calculated. As a consequence, in this subsection, we limit our analysis to the case of a univariate additional drift term.

For European calls in the Scott model, we observe that for low and intermediate volatility environments significant variance reduction is achieved. Especially for in-the-money options, compared to the constant approximation or the step-function approximation the performance can be significantly improved. In high volatility environments, the additional drift term again introduces numerical instabilities that over a broad range lead to an increased empirical variance of the Monte-Carlo estimator.

For arithmetic Asian calls, we obtain a slightly different formula for the option price elasticity [20]. Due to the dependency of the pay-off function (37) on discrete values $S_{k}$ of the underlying's trajectory, we can give only a discrete formula for $\epsilon$. However, this does not effect the Monte-Carlo simulation, as it is carried out on a discretized time-grid as well.

We define the lower bound of the arithmetic Asian call as follows:

$$
\begin{equation*}
C_{\mathrm{lb}}\left(S_{k}, \tau_{k}, k\right)=\left(\bar{S}_{k}-K e^{-r\left(\tau_{n}-\tau_{k}\right)}\right)^{+} \text {with } \bar{S}_{k}=\frac{1}{k} \sum_{l=1}^{k} S_{l} \tag{46}
\end{equation*}
$$

For $\bar{S}_{k}>K e^{-r\left(\tau_{n}-\tau_{k}\right)}$ we obtain

$$
\begin{equation*}
\epsilon_{k}=\frac{S_{k}}{C_{\mathrm{lb}}} \frac{\partial C_{\mathrm{lb}}}{\partial S_{k}}=\frac{1}{k} \frac{S_{k} / \bar{S}_{k}}{1-K e^{-r\left(\tau_{n}-\tau_{k}\right)} / S_{k}} . \tag{47}
\end{equation*}
$$

For $\bar{S}_{k}<K e^{-r\left(\tau_{n}-\tau_{k}\right)}$ as for European calls we require ${ }^{13}$

$$
\begin{equation*}
\bar{S}_{k} e^{\left(r+\sigma^{2} \epsilon\right)\left(\tau_{n}-\tau_{k}\right)} \stackrel{!}{=} K \Rightarrow \epsilon=\frac{1}{\sigma^{2}\left(\tau_{n}-\tau_{k}\right)} \ln \left[\frac{K e^{-r\left(\tau_{n}-\tau_{k}\right)}}{\bar{S}_{k}}\right] . \tag{48}
\end{equation*}
$$

In total, for arithmetic Asian options we obtain

$$
\epsilon_{k}=\left\{\begin{array}{ll}
\frac{1}{k} \frac{S_{k} / \bar{S}_{k}}{1-K e^{-r\left(\tau_{n}-\tau_{k}\right) / S_{k}}} & \text { for } \bar{S}_{k}>K e^{-r\left(\tau_{n}-\tau_{k}\right)}  \tag{49}\\
\frac{1}{\sigma^{2}\left(\tau_{n}-\tau_{k}\right)} \ln \left[\frac{K e^{-r\left(\tau_{n}-\tau_{k}\right)}}{S_{k}}\right] & \text { for } \bar{S}_{k} \leq K e^{-r\left(\tau_{n}-\tau_{k}\right)}
\end{array} .\right.
$$

As for European options, we truncate the elasticity values $\epsilon$ to a predefined interval $[1.0 ; 10.0]$.
From figure 5, subfigure (b) we observe similar variance reduction for in-themoney calls as in the previous subsection. For out-of-the-money options, the step function approximation (e.g., $\epsilon=10.0$ ) led to superior results than the lower bound approximation. Therefore, the additional computational effort introduced does not appear to add value ${ }^{14}$

[^8]Variance reduction: approximation by Black-Scholes formula


Figure 7: (a) European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating $\epsilon$ and $\chi$ by the Black-Scholes formula for European call options. $r=0.05, T=1, K=10, n=100, N=10,000, \lambda=0.25$ and $\eta=0.01$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10,000]$ and the absolute value of $\chi$ to the interval $[0 ; 10,000]$. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. Solid line: univariate additional drift, i.e. $\chi=0$. Dashed line: bivariate additional drift. Black line: reference line where the variance reduction equals 1.0. (b) Asian call, Black-Scholes model: Variance reduction factor on a logarithmic scale for an arithmetic Asian call in the Black-Scholes model achieved by importance sampling approximating $\epsilon$ by the Black-Scholes formula for European call options. $r=0.05, T=1, K=10$, $n=100$ and $N=10,000$. We truncate $\epsilon$ to the interval [1;2.5]. Red line: $\sigma=0.1$. Blue line: $\sigma=0.2$. Green line: $\sigma=0.5$. Black line: reference line where the variance reduction equals 1.0.

### 3.5 Approximation by Black-Scholes formula

The previously discussed approximation approaches either neglected the actual shape of the option price curve (which is the case for the constant and the step function approximation) or considered only its asymptotic behavior (which is the case for the lower bound approximation). In this subsection, we employ an explicit option price formula which is related to the European call in the Scott model and the arithmetic Asian call in the Black-Scholes model: We use the Black-Scholes formula for

## Elasticities approximated by Black-Scholes formula

(a) $\varepsilon$-surface

(b) $\chi$-surface


Figure 8: (a) $\epsilon$-surface: Surface of the logarithmized option price elasticity $\epsilon$ with respect to the underlying $S$ calculated by the Black-Scholes formula for European call options as a function of the underlying price $S$ and the remaining time until expiration $T-t$. (b) $\chi$-surface: Surface of the pseudo-elasticity $\chi$ with respect to the volatility $\sigma$ calculated by the Black-Scholes formula for European call options as a function of the underlying price $S$ and the remaining time until expiration $T-t$.

European call options [7, 22]

$$
\begin{align*}
C(S(t), t) & =S(t) \Phi\left(d_{1}\right)-K e^{-r(T-t)} \Phi\left(d_{2}\right) \\
d_{1} & =\frac{\ln \frac{S(t)}{K}+\left(r+\frac{\sigma^{2}}{2}\right)(T-t)}{\sigma \sqrt{T-t}}  \tag{50}\\
d_{2} & =\frac{\ln \frac{S(t)}{K}+\left(r-\frac{\sigma^{2}}{2}\right)(T-t)}{\sigma \sqrt{T-t}}=d_{1}-\sigma \sqrt{T-t}
\end{align*}
$$

to approximate the option price elasticity [5, 19, 20].
To calculate the bivariate additional drift term, knowledge of the partial derivatives of $C$ with respect to the underlying and the volatility, i.e. the "Greeks" $\Delta$ ("Delta") and $\nu$ ("Vega") is required. Derivation yields [8]

$$
\begin{align*}
\Delta & :=\frac{\partial C}{\partial S}=\Phi\left(d_{1}\right)  \tag{51}\\
\nu & :=\frac{\partial C}{\partial \sigma}=S(t) \Phi^{\prime}\left(d_{1}\right) \sqrt{T-t}=S(t) \phi\left(d_{1}\right) \sqrt{T-t}
\end{align*}
$$

Using the definitions (1) and (33), we obtain

$$
\begin{equation*}
\epsilon=\frac{1}{1-\frac{K e^{-r(T-t)}}{S(t)} \frac{\Phi\left(d_{2}\right)}{\Phi\left(d_{1}\right)}} \tag{52}
\end{equation*}
$$

and

$$
\begin{equation*}
\chi=\frac{\phi\left(d_{1}\right) \sqrt{T-t}}{\Phi\left(d_{1}\right)-\frac{K e^{-r(T-t)}}{S(t)} \Phi\left(d_{2}\right)} . \tag{53}
\end{equation*}
$$

Figure 8 shows a plot of these two functions. The asymptotic behavior of $\epsilon$ is of interest:

- $\epsilon$ is obviously strictly positive for call options, as it is the product of the strictly positive elements $S / C$ and $\partial C / \partial S$.
- As $\frac{K e^{-r(T-t)}}{S(t)} \frac{\Phi\left(d_{2}\right)}{\Phi\left(d_{1}\right)}$ consist of only positive elements, we obtain

$$
\begin{equation*}
1-\frac{K e^{-r(T-t)}}{S(t)} \frac{\Phi\left(d_{2}\right)}{\Phi\left(d_{1}\right)}<1 \tag{54}
\end{equation*}
$$

Therefore, we get

$$
\begin{equation*}
\epsilon(S(t), t)>1 \text { for all } S(t), t \tag{55}
\end{equation*}
$$

The previously introduced lower bound for the option price elasticity $(\epsilon \stackrel{!}{>}$ 1.0) receives a theoretical justification.

- In the limit $S(t) \rightarrow \infty$ we obtain $\frac{\Phi\left(d_{2}\right)}{\Phi\left(d_{1}\right)} \rightarrow 1$. Consequently, for high $S(t)$ values the Black-Scholes approximation and the lower bound approximation become asymptotically equivalent.

In figure 7. subfigure (a), we show that for European calls in the Scott model, the Black-Scholes approximation achieves strong variance reduction. In contrast to the previously presented approaches, the employment of a bivariate additional drift appears to add value: At least for out-of-the-money options, and mainly in low volatility environments, a significantly higher variance reduction is achieved than by adding a univariate drift term. For in-the-money options, the achieved variance reductions are similar as the reductions achieved by the lower bound approximation. This is not surprising, as the corresponding $\epsilon$ approximations are asymptotically equivalent for $S \gg K e^{-r(T-t)}$.

For arithmetic Asian options, we use the same approximation formula for $\epsilon$ as for the European call, i.e. we also insert the actual underlying value $S_{k}$ and not $\bar{S}_{k}$ as one might assume. However, as shown in subfigure (b), less variance reduction is achieved than by employing the step function approximation. This is not an intuitive
result, as the shape of the price curve of an arithmetic Asian option comes close to the shape of a European option (for details see subsection 3.6). Therefore, one could assume that the optimal $\epsilon$ might be well approximated by the Black-Scholes formula - or at least better than by a step function.

Attempts to improve the variance reduction by slightly modifying the approximation approach (e.g., the choice of $\bar{S}$ instead of $S$ to calculate $\epsilon$ from the BlackScholes formula, or the truncation to other intervals than [1.0; 2.5]) do not lead to improved results.

### 3.6 Approximation by regression

Approximation by simulation of the entire option price curve Several options have some features in common with European call options in the BlackScholes model. The two options treated in this article, i.e. European calls in the Scott model and arithmetic Asian calls in the Black-Scholes model are examples. It is unlikely that their option price curves have exactly the same shape and position, while it is also unlikely that their curves' shapes look completely different, although the specific position may vary of course.

An alternative approach to approximate the option price elasticity is therefore to use a relatively small sample (with size $N_{1}$ ) to simulate a couple of points of the option price curve to get a rough understanding of its position and shape. Then, the idea is to conduct a non-linear least square regression to fit the curve of a known option price function (e.g., the function of a European call in the Black-Scholes model) to the simulated data. Thus, one obtains a curve from which one can calculate an estimate of the option price elasticity for an importance sampling Monte-Carlo simulation (with sample size $N_{2}{ }^{15}$. In this paper, we use the Black-Scholes formula from equation (50) for non-linear regressions.

We do not present additional plots of the $\epsilon$ - and $\chi$-surface for this approach as they have essentially the same shape as the surfaces presented in figure 8. The specific numerical values may vary of course.
Figure 9 shows a visual representation of the approach: Option price estimators (points) are simulated with a reduced sample size $N_{1}$ and the regression curve (solid line) is estimated according to least square regression. The curve of the Black-Scholes formula (dashed line) is plotted for comparison. It turns out that for European calls in the Scott model, the Black-Scholes formula is a good approximation. Visually, it is not possible to distinguish between the regression line and the Black-Scholes formula's curve. For arithmetic Asian calls the situation is different.

[^9]
## Curve regression



Figure 9: (a) European call, Scott model: Non-linear regression curve of the Black-Scholes formula fitted to 17 estimated points along the option price curve of a European call in the Scott model. For each point, a simulation was conducted with $N_{1}=10 \% \cdot N=1,000, \bar{\sigma}=0.2, r=0.05, T=1, K=10, n=100$, $\lambda=0.25$ and $\eta=0.01$. Solid line: regression curve. Dashed line: curve of Black-Scholes formula using parameters employed in simulation. (b) Asian call, Black-Scholes model: as (a), but for arithmetic Asian call in the Black-Scholes model, $\sigma=0.2$. Here, a significant deviation between the two lines results.

For instance, in the intermediate volatility environment with $\sigma=0.2$, downward adjustments in the strike price, risk-free interest rate and volatility ( $\hat{K}=9.20$, $\hat{r}=-0.0703$ and $\hat{\sigma}=0.130$ instead of $K=10, r=0.05$ and $\sigma=0.2$ ) lead to an regression curve (solid line) below the Black-Scholes formula curve (dashed line).
In the curve regression approach, these estimated parameters are used to calculate $\epsilon$ from equation 52) using $\hat{d}_{1}$ and $\hat{d}_{2}$ as defined in equation 50), but inserting estimated parameters $\hat{K}, \hat{r}$ and $\hat{\sigma}$ instead of the original parameter set. However, for the actual simulation of sample paths and the calculation of the pay-off function, we use the original parameters $K, r$ and $\sigma$.

From figure 10, subfigure (a), we see that for the European call in the Scott model, in the low volatility environment the variance reduction results are approximately the same as for the Black-Scholes formula approximation, using the original set of parameters when calculating $\epsilon$ (compare to figure 7). Especially for the intermediate volatility environment, results appear numerically unstable while for the high volatility environments results are poorer than for the Black-Scholes formula


Figure 10: (a) European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by univariate importance sampling approximating $\epsilon$ by the curve regression approach. $r=0.05$, $T=1, K=10, n=100, N=10,000, N_{1}=1,000, N_{2}=N-N_{1}=9,000$, $\lambda=0.25$ and $\eta=0.01$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10,000]$. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. No bivariate simulation was conducted. Black line: reference line where the variance reduction equals 1.0. (b) Asian call, Black-Scholes model: Variance reduction factor on a logarithmic scale for an arithmetic Asian call in the Black-Scholes model achieved by importance sampling approximating $\epsilon$ by the curve regression approach. $r=0.05, T=1, K=10, n=100, N=10,000$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 2.5]$. Red line: $\sigma=0.1$. Blue line: $\sigma=0.2$. Green line: $\sigma=0.5$. Black line: reference line where the variance reduction equals 1.0.
approach. For European calls in the Scott model, the approach does not appear to justify the additional complexity resulting from the non-linear regression.

Subfigure (b) indicates that for arithmetic Asian options in the Black-Scholes model results are roughly the same for the curve regression approach and for the Black-Scholes formula approach. Despite the significant downward shift of the price curve employed in the calculation of $\epsilon$ (compare figure 9), results have not improved. The approach again performs poorer than the step function approach presented in subsection 3.3

## Variance reduction: approximation by single starting point regression

European call, Scott model


Figure 11: European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating $\epsilon$ by the single starting point regression approach. $r=0.05, T=$ $1, K=10, n=100, N=100,000, N_{1}=10,000, N_{2}=N-N_{1}=90,000$, $\lambda=0.25$ and $\eta=0.01$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10,000]$. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$.

## Approximation by simulation of a single point of the option price curve

A disadvantage of the the presented curve regression method is that to estimate the regression curve several points across the curve have to be estimated. In the previous example, 17 points of the option price curve were estimated - each one with a sample size of $N_{1}=1,000$. Thus, high numerical effort is already required just for the parameter estimation required to calculate the $\epsilon$ values. For instance, the quantity of interest may be an option price for one specific initial value $S(0)$ and not the entire option price curve. Then, it may not appear appropriate to first simulate estimators covering the whole range of the option price curve. Therefore, we present an alternative regression approach which does not require to simulate points across the whole curve. Only trajectories starting with one single $S(0)$ value are required.
In detail, the approach for European calls works as follows:

1. We simulate $N_{1}$ trajectories with $n$ discretization step of the underlying with the same starting value $S(0)$. We obtain a $N_{1} \times(n+1)$ matrix, with $S_{i k}$ being the value for the $k^{\text {th }}$ discretization step of the $i^{\text {th }}$ simulated trajectory

Single starting point regression


Figure 12: Illustrative representation of the starting point regression approach. A multidimensional non-linear least square regression is conducted fitting a BlackScholes formula surface (blue grid) to $N_{1} \cdot(n+1)$ option price estimator data points (black points). In this figure, the sample size is selected artificially low for the purpose of visualization purposes $(N=100)$. With $N=100$ the non-linear regression does not converge. Therefore, no parameter estimation is possible. Consequently, we plot the Black-Scholes option price surface using the original parameters. In this sense, this figure does not show actual data but is rather illustrative.
with $k \in(0, \ldots, n)$ and $i \in(1, \ldots, N)$.
2. We use all the points of the simulated $N_{1} \times(n+1)$ matrix to calculate $N_{1}$. $(n+1)$ option price estimators. E.g., for the $k^{\text {th }}$ discretization step of the $i^{\text {th }}$ simulated trajectory we calculate

$$
\begin{equation*}
\hat{C}_{i k}=e^{-r\left(\tau_{n}-\tau_{k}\right)}\left(S_{i n}-K\right)^{+} . \tag{56}
\end{equation*}
$$

It is important to note that in the bracket term we use $S_{i n}$ and not $S_{i k}$. Thus, we obtain a single option price estimator for each step of the time-grid for each trajectory. The resulting ensemble of $\left(S_{i k}, \tau_{k}, \hat{C}_{i k}\right)$ is plotted in figure 12 (black points).
3. In the next step, we conduct a multidimensional non-linear least square regression using the Black-Scholes formula with $C(S(t), t)$ as dependent variable, $S(t)$ and $t$ as independent variable and $K, r$ and $\sigma$ as parameters. An
option price surface results (blue grid in figure 12).
4. The resulting parameters are then used to calculate $\epsilon$ in an importance sampling Monte-Carlo simulation with sample size $N_{2}$.

The variance reduction results presented in figure 11 are below the results achieved by the approximation employing the Black-Scholes formula with the original set of parameters (compare figure 7). For purposes of convergence of the multidimensional non-linear regression, compared to other simulations we choose a larger sample size $N=N_{1}+N_{2}=100,000$ with $N_{1}=0.1 \cdot N$.

For arithmetic Asian options, this approach does not work due to the path dependency of its pay-off function.

### 3.7 Approximation by integration

The last approximation approach for option price elasticities presented in this article is based on the idea of calculating an estimator for the elasticities by numerical integration.
Integration yields approximations for $C^{\text {int }}, C_{S}^{\text {int }}$ and $C_{\sigma}^{\text {int }}$ at the time $t$ for a given $S(t)$ value. With this information, we can calculate $\epsilon$. To obtain $C^{\text {int }}$, we have to solve the integral

$$
\begin{equation*}
C^{\mathrm{int}}(S(t), t)=e^{-r(T-t)} \int_{-\infty}^{\infty}(S(T)-K)^{+} p[S(T) \mid S(t)] \mathrm{d} S(T) . \tag{57}
\end{equation*}
$$

Using Leibniz' rule for parameter integrals [23], we obtain

$$
\begin{align*}
\left(\frac{\partial C(S(t), t)}{\partial S(t)}\right)^{\mathrm{int}} & =\frac{\partial}{\partial S(t)}\left[e^{-r(T-t)} \int_{-\infty}^{\infty}(S(T)-K)^{+} p[S(T) \mid S(t)] \mathrm{d} S(T)\right] \\
& =e^{-r(T-t)} \int_{-\infty}^{\infty}(S(T)-K)^{+} \frac{\partial}{\partial S(t)} p[S(T) \mid S(t)] \mathrm{d} S(T) \tag{58}
\end{align*}
$$

Generally, the transition density $p[S(T) \mid S(t)]$ is unknown. One approach would be to conduct an approximation using the Markov property and building a product of Euler transition kernels for small $\Delta t$. However, this would leave us with the problem of numerically solving a high-dimensional integral. An alternative is to approximate the unknown density $p[S(T) \mid S(t)]$ by an explicitly known density. For instance, for the Scott model, we can approximate the transition density by the density of the log-normal distribution resulting from a Geometric Brownian Motion, i.e. the Black-Scholes model.

This approximation does not lead to a bias of the option price estimator as we can use any $\epsilon$ to conduct the Monte-Carlo simulations (see section 2). We should just take into account that reasonable $\epsilon$ values should be chosen that lead to an actual variance reduction and not to an increase.
Therefore, we approximate the transition density by

$$
\begin{equation*}
p[S(T) \mid S(t)] \approx \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{S(T)} e^{-\frac{\left(\ln S(T)-\ln S(t)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}} . \tag{59}
\end{equation*}
$$

To numerically calculate the integrals, the integration limits must be transformed to the interval $(-1,1)$ employing an affine transformation [24]:

$$
\begin{equation*}
S(T)=K+\frac{1+y}{1-y} \Rightarrow \int_{K}^{\infty} f[S(T)] \mathrm{d} S(T)=\int_{-1}^{1} \frac{2}{(1-y)^{2}} f\left(K+\frac{1+y}{1-y}\right) \mathrm{d} y \tag{60}
\end{equation*}
$$

This integral can be evaluated by means of Gauss-Legendre quadrature [25]. We approximate

$$
\begin{equation*}
\int_{-1}^{1} g(x) \mathrm{d} x \approx \sum_{i=1}^{n_{\mathrm{GL}}} g\left(x_{i}\right) w_{i} \tag{61}
\end{equation*}
$$

where $x_{i}$ are the roots ("abscissae") of the $n_{\text {GL }}{ }^{\text {th }}$ Legendre polynomial $P_{n_{\mathrm{GL}}}(x)$. The weight $s^{16}$ are defined by

$$
\begin{equation*}
w_{i}=-\frac{2}{\left(1-x_{i}^{2}\right)\left[P_{n_{\mathrm{LL}}}^{\prime}\left(x_{i}\right)\right]^{2}} . \tag{62}
\end{equation*}
$$

For the option price we obtain

$$
\begin{align*}
C^{\text {int }}= & \int_{-\infty}^{\infty}(S(T)-K)^{+} p[S(T) \mid S(t)] \mathrm{d} S(T) \\
= & \int_{K}^{\infty}(S(T)-K) p[S(T) \mid S(t)] \mathrm{d} S(T) \\
= & \int_{K}^{\infty}(S(T)-K) \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{S(T)}  \tag{63}\\
& \times \exp \left\{-\frac{\left(\ln S(T)-\ln S(t)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}\right\} \mathrm{d} S(T) \\
= & \int_{K}^{\infty} f[S(T)] \mathrm{d} S(T) .
\end{align*}
$$

By carrying out the substitution from equation with $S(T)=K+\frac{1+y}{1-y}$ we get

[^10]\[

$$
\begin{align*}
C^{\mathrm{int}}= & \int_{-1}^{1} \frac{2}{(1-y)^{2}} f\left(K+\frac{1+y}{1-y}\right) \mathrm{d} y \\
= & \int_{-1}^{1} \frac{2}{(1-y)^{2}}\left(\frac{1+y}{1-y}\right) \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{K+\frac{1+y}{1-y}}  \tag{64}\\
& \times \exp \left\{-\frac{\left(\ln \left(K+\frac{1+y}{1-y}\right)-\ln S(t)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t)\right)^{2}}{2 \sigma^{2}(T-t)}\right\} \mathrm{d} y
\end{align*}
$$
\]

and by conducting a Gauss-Legendre quadrature as described in equation (61) with the abbreviation

$$
\begin{equation*}
\xi_{i}=\ln \left(K+\frac{1+x_{i}}{1-x_{i}}\right)-\ln S(t)-\left(r-\frac{\sigma^{2}}{2}\right)(T-t) \tag{65}
\end{equation*}
$$

we approximate

$$
\begin{align*}
C^{\mathrm{int}} \approx & \sum_{i=1}^{n_{\mathrm{GL}}} \frac{2}{\left(1-x_{i}\right)^{2}}\left(\frac{1+x_{i}}{1-x_{i}}\right) \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{K+\frac{1+x_{i}}{1-x_{i}}}  \tag{66}\\
& \times \exp \left\{-\frac{\xi_{i}^{2}}{2 \sigma^{2}(T-t)}\right\} \cdot w_{i} .
\end{align*}
$$

Analogously, we obtain

$$
\begin{align*}
\left(\frac{\partial C}{\partial S(t)}\right)^{\mathrm{int}} \approx & \sum_{i=1}^{n_{\mathrm{GL}}} \frac{2}{\left(1-x_{i}\right)^{2}}\left(\frac{1+x_{i}}{1-x_{i}}\right) \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{K+\frac{1+x_{i}}{1-x_{i}}}  \tag{67}\\
& \times \exp \left\{-\frac{\xi_{i}^{2}}{2 \sigma^{2}(T-t)}\right\} \frac{\xi_{i}}{\sigma^{2}(T-t)} \frac{1}{S(t)} \cdot w_{i}
\end{align*}
$$

and

$$
\begin{align*}
\left(\frac{\partial C}{\partial \sigma}\right)^{\text {int }} \approx & \sum_{i=1}^{n_{\mathrm{GL}}} \frac{2}{\left(1-x_{i}\right)^{2}}\left(\frac{1+x_{i}}{1-x_{i}}\right) \frac{1}{\sqrt{2 \pi \sigma^{2}(T-t)}} \frac{1}{K+\frac{1+x_{i}}{1-x_{i}}}  \tag{68}\\
& \times \exp \left\{-\frac{\xi_{i}^{2}}{2 \sigma^{2}(T-t)}\right\} \frac{1}{\sigma}\left[-1-\xi_{i}+\frac{\xi_{i}^{2}}{\sigma^{2}(T-t)}\right] \cdot w_{i} .
\end{align*}
$$

From these approximations, we can calculate $\epsilon$ and $\chi$ for each discretization step of each trajectory using the definitions (1) and (33).

If we employ the presented method to conduct an importance sampling MonteCarlo simulation, for in-the-money options we observe very low variance reductions for medium and high volatility environments or even variance increases for low

Variance reduction: approximation by integration
European call, Scott model


Figure 13: European call, Scott model: Variance reduction factor on a logarithmic scale for a European call in the Scott model achieved by importance sampling approximating $\epsilon$ and $\chi$ by integration. $r=0.05, T=1, K=10, n=10$, $N=10,000, \lambda=0.25$ and $\eta=0.01$. For purposes of numerical stability, we truncate $\epsilon$ to the interval $[1 ; 10,000]$ and the absolute value of $\chi$ to the interval [ $0 ; 10,000]$. Furthermore, for $S>K e^{-r(T-t)}$ during the last 6 discretization steps, we choose the $\epsilon$-value corresponding to the $(n-7)^{\text {th }}$ discretization step. Red line: $\bar{\sigma}=0.1$. Blue line: $\bar{\sigma}=0.2$. Green line: $\bar{\sigma}=0.5$. Solid line: univariate additional drift, i.e. $\chi=0$. Dashed line: bivariate additional drift. Black line: reference line where the variance reduction equals 1.0.
volatility environments. An analysis of the $\epsilon$-surface shown in subfigure (a) of figure 14 yields an explanation for this unexpected result. For high $S$ values and short times to maturity $T-t$, the integration method leads to numerical instabilities. The observed $\epsilon$-peak are not expected and not in accordance with the result calculated directly from the Black-Scholes formula (compare subfigure (a) of 8).

Increasing $n_{\mathrm{GL}}$ could help to mitigate this problem, but this would also further increase the required computational power. A less numerically demanding approach is to introduce an ad hoc modification of $\epsilon$ values as soon as the underlying exceeds the discounted strike price for low times to maturity $T-t$. Therefore, in the numerical example presented in figure 13 we modify the $\epsilon$-surface. For $S>K e^{-r(T-t)}$ during the last 6 discretization steps, we choose the $\epsilon$-value corresponding to the $(n-7)^{\text {th }}$ discretization step. I.e., we "fix" the time during the last few discretization steps for the purpose of calculating $\epsilon$. We thus avoid the described numerical

## Elasticities approximated by integration

(a) $\varepsilon$-surface

(b) $\chi$-surface


Figure 14: (a) $\epsilon$-surface: Surface of the logarithmized option price elasticity $\epsilon$ with respect to the underlying price $S$ calculated by integration as a function of $S$ and the remaining time until expiration $T-t$ for $\bar{\sigma}=0.1$. For in-the-money options, we observe numerically instabilities for low $T-t$ values. (b) $\chi$-surface: Surface of the pseudo-elasticity $\chi$ with respect to the volatility $\sigma$ calculated by integration as a function of $S$ and the remaining time until expiration $T-t$ for $\bar{\sigma}=0.1$.
instabilities which introduce strong variance increases. Of course, this involves a deliberate deviation from the optimal additional drift term. However, this deviation has a far lower impact on the results than the numerical instabilities involved otherwise.

The numerical results in figure 13 show that the significant computational effort required to calculate the elasticities by integration also leads to significant variance reduction, especially for out-of-the-money options.

For out-of-the-money options, the bivariate additional drift yields a slightly higher variance reduction than the univariate case. For European options, especially compared to the Black-Scholes formula approach, this approach appears not to add value, as it is numerically for more demanding.

However, the advantage of the integration approach is, that no knowledge of the options price curve is required to approximate the elasticities. For more complex options, no analytical formula may exist that closely resembles the option price curve. In these cases, the integration approach may be a suitable solution that can yield dramatic variance reduction.

For path-dependent options as an arithmetic Asian call, this approach does not appear advantageous. In our example with $n=100$ discretization steps, two 100dimensional integrals would have to be evaluated for each calculation of $\epsilon$, involving a significant additional computational effort.

## 4 Conclusion

From this analysis, we draw the conclusion that several well-working approaches exist for the approximation of option price elasticities. This a necessary requirement for the presented importance sampling technique. The consideration of several approaches (approximation by constant values, step functions, lower bounds, the Black-Scholes formula, regression, and integration) demonstrates that the approaches vary both in efficacy and in the involved computational effort. For European calls in the Scott model, the achieved variance reduction ranges from less than one order of magnitude for low constant approximations to several orders of magnitude for the Black-Scholes formula approach. Also for Asian calls in the Black-Scholes model, we observe a significant deviation between the efficacy of the different analyzed approaches. We see that the degree of variance reduction depends on several influencing factors: both the volatility and the initial value of the underlying heavily influence the results.
Interestingly, the numerically more demanding approximation procedures are not necessarily the most efficacious ones. For European options we observe, that approximation approaches that closely resemble the results of the Black-Scholes formula for European calls work best. Contrastingly, for arithmetic Asian options, the very simple step function approach which involves only very limited calculations for the approximation of $\epsilon$ delivers the best results. More sophisticated approaches as the regression approach yield a smaller variance reduction. So, there are cases, where additional complexity and numerical effort to approximate elasticities are not justified.

In this article, we introduce several approaches that can be applied even if no or low knowledge of the shape and position of the curve of interest exists ex anteriori, e.g., the constant approach, the step function approach or the integration approach. These approaches may well be used for other applications of Monte-Carlo simulations such as Value at Risk estimations or Portfolio Credit Risk models.
We emphasize that in this article, we consider solely the application of the presented importance sampling approach. The combination with other variance reduction techniques is not considered. The suitability of the presented approach in
conjunction with antithetic variates, control variates, stratified sampling etc. should be further analyzed. For instance, it has been shown that significant variance reduction can be obtained by combining importance sampling with an application of the put-call-parity [9, 10, 27].

One possible criticism regarding some of the presented approximations is the involved considerable numerical effort. However, one can considerably reduce the computing power required for estimating option price elasticities by first calculating the elasticity surface as a function of the underlying $S$ and remaining time $T-t$. When carrying out the variance reduced Monte-Carlo simulation it is then possible to approximate $\epsilon$ (and $\chi$ ) from the previously calculated grid, e.g., by linear intrapolation or by cubic spline intrapolation.

## 5 Acknowledgments

I would like to thank Prof. Dr. Hermann Singer (FernUniversität in Hagen, Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung) for laying the foundation of my research, his interest in my work, and the supervision of my doctoral project. This work was kindly supported by the Konrad Adenauer Foundation, the donor of my graduate scholarship.

## References

[1] Wagner W. Monte Carlo evaluation of functionals of solutions of stochastic differential equations. Variance reduction and numerical examples. Stochastic Analysis and Applications. 1988;6(4):447-468.
[2] Glasserman P, Heidelberger P, Shahabuddin P. Asymptotically Optimal Importance Sampling and Stratification for Pricing Path-Dependent Options. Mathematical Finance. 1999;9(2):117-152.
[3] Glasserman P. Monte Carlo methods in financial engineering. Applications of mathematics ; 53. New York, NY ; Berlin ; Heidelberg [u.a.]: Springer; 2004.
[4] Zhao Q, Liu G, Gu G. Variance Reduction Techniques of Importance Sampling Monte Carlo Methods for Pricing Options. Journal of Mathematical Finance. 2013;3(4):431-436.
[5] Singer H. Importance sampling for Kolmogorov backward equations. AStA Advances in Statistical Analysis. 2014;98(4):345-369.
[6] Singer H. Finanzmarktökonometrie. Wirtschaftswissenschaftliche Beiträge Nr. 171. Physica-Verl.; 1999.
[7] Hull JC. Options, futures, and other derivatives. Pearson Education; 2006.
[8] Yu X, Xie X. On Derivations of Black-Scholes Greek Letters. Research Journal of Finance and Accounting. 2013;4(6):80-85.
[9] Müller A. A joint application of the put-call-parity and importance sampling to variance reduced option pricing. FernUniversität in Hagen - Fakultät für Wirtschaftswissenschaft - Discussion paper No. 496; 2016.
[10] Müller A. Improved Variance Reduced Monte-Carlo Simulation of in-theMoney Options. Journal of Mathematical Finance. 2016 August;6(3):361367.
[11] Hammersley JM, Morton KW. Poor Man's Monte Carlo. Journal of the Royal Statistical Society Series B (Methodological). 1954;16(1):pp. 23-38.
[12] Kloeden PE, Platen E. Numerical solution of stochastic differential equations. Berlin ; New York: Springer-Verlag; 1992.
[13] Melchior M, Öttinger HC. Variance reduced simulations of stochastic differential equations. The Journal of Chemical Physics. 1995;103(21):9506-9509.
[14] Karatzas I, Shreve S. Brownian motion and stochastic calculus. vol. 113. Springer Science \& Business Media; 1988.
[15] Duffie D. Dynamic asset pricing theory. Princeton University Press; 2010.
[16] Zucker M, Singer H. An Alternative Derivation of the Black-Scholes Formula. Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung, FernUniversität in Hagen; 2007.
[17] Risken H. The Fokker-Planck Equation: Methods of Solutions and Applications. 2nd ed. Springer Series in Synergetics. Springer; 1989.
[18] Scott LO. Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. The Journal of Financial and Quantitative Analysis. 1987;22(4):pp. 419-438.
[19] Thimmel B. Varianzreduzierende Verfahren zur Optionspreisberechnung [Diplomarbeit]. FernUniversität in Hagen - Fakultät für Wirtschaftswissenschaft; 2007.
[20] Müller A. Varianzreduzierende Verfahren zur Optionspreisbewertung [Master thesis]. FernUniversität in Hagen - Fakultät für Wirtschaftswissenschaft; 2012.
[21] Merton RC. Theory of Rational Option Pricing. The Bell Journal of Economics and Management Science. 1973;4(1):141-183.
[22] Black F, Scholes M. The Pricing of Options and Corporate Liabilities. The Journal of Political Economy. 1973 May-June;81(3):637-654.
[23] Flanders H. Differentiation under the integral sign. The American Mathematical Monthly. 1973;80(6):615-627.
[24] Antia HM. Numerical Methods for Scientists and Engineers. No. Bd. 1 in Numerical Methods for Scientists and Engineers. Springer Verlag NY; 2002.
[25] Bronstein IN, Semendjajew KA, Musiol G, Mühlig H. Taschenbuch der Mathematik. Verlag Harri Deutsch; 2006.
[26] Kamermans MP. Gaussian Quadrature Weights and Abscissae; 2011. Last accessed: July 27, 2016, 9:21 p.m. http://pomax.github.io/ bezierinfo/legendre-gauss.html.
[27] Reider RL. Two applications of Monte Carlo techniques to finance. University of Pennsylvania; 1994.

Die Diskussionspapiere ab Nr. 183 (1992) bis heute, können Sie im Internet unter http://www.fernuni-hagen.de/wirtschaftswissenschaft/forschung/beitraege.shtml einsehen und zum Teil downloaden.
Ältere Diskussionspapiere selber erhalten Sie nur in den Bibliotheken.

| Nr | Jahr | Titel | Autor/en |
| :---: | :---: | :---: | :---: |
| 420 | 2008 | Stockkeeping and controlling under game theoretic aspects | Fandel, Günter Trockel, Jan |
| 421 | 2008 | On Overdissipation of Rents in Contests with Endogenous Intrinsic Motivation | Schlepütz, Volker |
| 422 | 2008 | Maximum Entropy Inference for Mixed Continuous-Discrete Variables | Singer, Hermann |
| 423 | 2008 | Eine Heuristik für das mehrdimensionale Bin Packing Problem | Mack, Daniel Bortfeldt, Andreas |
| 424 | 2008 | Expected A Posteriori Estimation in Financial Applications | Mazzoni, Thomas |
| 425 | 2008 | A Genetic Algorithm for the Two-Dimensional Knapsack Problem with Rectangular Pieces | Bortfeldt, Andreas Winter, Tobias |
| 426 | 2008 | A Tree Search Algorithm for Solving the Container Loading Problem | Fanslau, Tobias Bortfeldt, Andreas |
| 427 | 2008 | Dynamic Effects of Offshoring | Stijepic, Denis Wagner, Helmut |
| 428 | 2008 | Der Einfluss von Kostenabweichungen auf das NashGleichgewicht in einem nicht-kooperativen Disponenten-Controller-Spiel | Fandel, Günter Trockel, Jan |
| 429 | 2008 | Fast Analytic Option Valuation with GARCH | Mazzoni, Thomas |
| 430 | 2008 | Conditional Gauss-Hermite Filtering with Application to Volatility Estimation | Singer, Hermann |
| 431 | 2008 | Web 2.0 auf dem Prüfstand: Zur Bewertung von InternetUnternehmen | Christian Maaß Gotthard Pietsch |
| 432 | 2008 | Zentralbank-Kommunikation und Finanzstabilität - Eine Bestandsaufnahme | Knütter, Rolf Mohr, Benjamin |
| 433 | 2008 | Globalization and Asset Prices: Which Trade-Offs Do Central Banks Face in Small Open Economies? | Knütter, Rolf <br> Wagner, Helmut |
| 434 | 2008 | International Policy Coordination and Simple Monetary Policy Rules | Berger, Wolfram Wagner, Helmut |
| 435 | 2009 | Matchingprozesse auf beruflichen Teilarbeitsmärkten | Stops, Michael Mazzoni, Thomas |
| 436 | 2009 | Wayfindingprozesse in Parksituationen - eine empirische Analyse | Fließ, Sabine Tetzner, Stefan |
| 437 | 2009 | ENTROPY-DRIVEN PORTFOLIO SELECTION a downside and upside risk framework | Rödder, Wilhelm Gartner, Ivan Ricardo Rudolph, Sandra |
| 438 | 2009 | Consulting Incentives in Contests | Schlepütz, Volker |


| 439 | 2009 | A Genetic Algorithm for a Bi-Objective WinnerDetermination Problem in a Transportation-Procurement Auction" | Buer, Tobias Pankratz, Giselher |
| :---: | :---: | :---: | :---: |
| 440 | 2009 | Parallel greedy algorithms for packing unequal spheres into a cuboidal strip or a cuboid | Kubach, Timo Bortfeldt, Andreas Tilli, Thomas Gehring, Hermann |
| 441 | 2009 | SEM modeling with singular moment matrices Part I: MLEstimation of time series | Singer, Hermann |
| 442 | 2009 | SEM modeling with singular moment matrices Part II: MLEstimation of sampled stochastic differential equations | Singer, Hermann |
| 443 | 2009 | Konsensuale Effizienzbewertung und -verbesserung Untersuchungen mittels der Data Envelopment Analysis (DEA) | Rödder, Wilhelm Reucher, Elmar |
| 444 | 2009 | Legal Uncertainty - Is Hamonization of Law the Right Answer? A Short Overview | Wagner, Helmut |
| 445 | 2009 | Fast Continuous-Discrete DAF-Filters | Mazzoni, Thomas |
| 446 | 2010 | Quantitative Evaluierung von Multi-Level Marketingsystemen | Lorenz, Marina Mazzoni, Thomas |
| 447 | 2010 | Quasi-Continuous Maximum Entropy Distribution Approximation with Kernel Density | Mazzoni, Thomas Reucher, Elmar |
| 448 | 2010 | Solving a Bi-Objective Winner Determination Problem in a Transportation Procurement Auction | Buer, Tobias Pankratz, Giselher |
| 449 | 2010 | Are Short Term Stock Asset Returns Predictable? An Extended Empirical Analysis | Mazzoni, Thomas |
| 450 | 2010 | Europäische Gesundheitssysteme im Vergleich Effizienzmessungen von Akutkrankenhäusern mit DEA - | Reucher, Elmar Sartorius, Frank |
| 451 | 2010 | Patterns in Object-Oriented Analysis | Blaimer, Nicolas Bortfeldt, Andreas Pankratz, Giselher |
| 452 | 2010 | The Kuznets-Kaldor-Puzzle and Neutral Cross-Capital-Intensity Structural Change | Stijepic, Denis Wagner, Helmut |
| 453 | 2010 | Monetary Policy and Boom-Bust Cycles: The Role of Communication | Knütter, Rolf Wagner, Helmut |
| 454 | 2010 | Konsensuale Effizienzbewertung und-verbesserung mittels DEA - Output- vs. Inputorientierung - | Reucher, Elmar Rödder, Wilhelm |
| 455 | 2010 | Consistent Modeling of Risk Averse Behavior with Spectral Risk Measures | Wächter, Hans Peter Mazzoni, Thomas |


| 456 | 2010 | Der virtuelle Peer <br> - Eine Anwendung der DEA zur konsensualen Effizienzbewertung - | Reucher, Elmar |
| :---: | :---: | :---: | :---: |
| 457 | 2010 | A two-stage packing procedure for a Portuguese trading company | Moura, Ana Bortfeldt, Andreas |
| 458 | 2010 | A tree search algorithm for solving the multi-dimensional strip packing problem with guillotine cutting constraint | Bortfeldt, Andreas Jungmann, Sabine |
| 459 | 2010 | Equity and Efficiency in Regional Public Good Supply with Imperfect Labour Mobility - Horizontal versus Vertical Equalization | Arnold, Volker |
| 460 | 2010 | A hybrid algorithm for the capacitated vehicle routing problem with three-dimensional loading constraints | Bortfeldt, Andreas |
| 461 | 2010 | A tree search procedure for the container relocation problem | Forster, Florian Bortfeldt, Andreas |
| 462 | 2011 | Advanced X-Efficiencies for CCR- and BCC-Modell <br> - Towards Peer-based DEA Controlling | Rödder, Wilhelm Reucher, Elmar |
| 463 | 2011 | The Effects of Central Bank Communication on Financial Stability: A Systematization of the Empirical Evidence | Knütter, Rolf <br> Mohr, Benjamin <br> Wagner, Helmut |
| 464 | 2011 | Lösungskonzepte zur Allokation von Kooperationsvorteilen in der kooperativen Transportdisposition | Strangmeier, Reinhard Fiedler, Matthias |
| 465 | 2011 | Grenzen einer Legitimation staatlicher Maßnahmen gegenüber Kreditinstituten zur Verhinderung von Bankenund Wirtschaftskrisen | Merbecks, Ute |
| 466 | 2011 | Controlling im Stadtmarketing - Eine Analyse des Hagener Schaufensterwettbewerbs 2010 | Fließ, Sabine Bauer, Katharina |
| 467 | 2011 | A Structural Approach to Financial Stability: On the Beneficial Role of Regulatory Governance | Mohr, Benjamin Wagner, Helmut |
| 468 | 2011 | Data Envelopment Analysis - Skalenerträge und Kreuzskalenerträge | Wilhelm Rödder Andreas Dellnitz |
| 469 | 2011 | Controlling organisatorischer Entscheidungen: Konzeptionelle Überlegungen | Lindner, Florian Scherm, Ewald |
| 470 | 2011 | Orientierung in Dienstleistungsumgebungen - eine explorative Studie am Beispiel des Flughafen Frankfurt am Main | Fließ, Sabine Colaci, Antje Nesper, Jens |


| 471 | 2011 | Inequality aversion, income skewness and the theory of the welfare state | Weinreich, Daniel |
| :---: | :---: | :---: | :---: |
| 472 | 2011 | A tree search procedure for the container retrieval problem | Forster, Florian Bortfeldt, Andreas |
| 473 | 2011 | A Functional Approach to Pricing Complex Barrier Options | Mazzoni, Thomas |
| 474 | 2011 | Bologna-Prozess und neues Steuerungsmodell - auf Konfrontationskurs mit universitären Identitäten | Jost, Tobias Scherm, Ewald |
| 475 | 2011 | A reduction approach for solving the rectangle packing area minimization problem | Bortfeldt, Andreas |
| 476 | 2011 | Trade and Unemployment with Heterogeneous Firms: How Good Jobs Are Lost | Altenburg, Lutz |
| 477 | 2012 | Structural Change Patterns and Development: China in Comparison | Wagner, Helmut |
| 478 | 2012 | Demografische Risiken - Herausforderungen für das finanzwirtschaftliche Risikomanagement im Rahmen der betrieblichen Altersversorgung | Merbecks, Ute |
| 479 | 2012 | "It's all in the Mix!" - Internalizing Externalities with R\&D Subsidies and Environmental Liability | Endres, Alfred <br> Friehe, Tim <br> Rundshagen, Bianca |
| 480 | 2012 | Ökonomische Interpretationen der Skalenvariablen $u$ in der DEA | Dellnitz, Andreas Kleine, Andreas Rödder, Wilhelm |
| 481 | 2012 | Entropiebasierte Analyse von Interaktionen in Sozialen Netzwerken | Rödder, Wilhelm Brenner, Dominic Kulmann, Friedhelm |
| 482 | 2013 | Central Bank Independence and Financial Stability: A Tale of Perfect Harmony? | Berger, Wolfram Kißmer, Friedrich |
| 483 | 2013 | Energy generation with Directed Technical Change | Kollenbach, Gilbert |
| 484 | 2013 | Monetary Policy and Asset Prices: When Cleaning Up Hits the Zero Lower Bound | Berger, Wolfram Kißmer, Friedrich |
| 485 | 2013 | Superknoten in Sozialen Netzwerken - eine entropieoptimale Analyse | Brenner, Dominic, Rödder, Wilhelm, Kulmann, Friedhelm |
| 486 | 2013 | Stimmigkeit von Situation, Organisation und Person: Gestaltungsüberlegungen auf Basis des Informationsverarbeitungsansatzes | Julmi, Christian Lindner, Florian Scherm, Ewald |
| 487 | 2014 | Incentives for Advanced Abatement Technology Under National and International Permit Trading | Endres, Alfred Rundshagen, Bianca |


| 488 | 2014 | Dynamische Effizienzbewertung öffentlicher <br> Dreispartentheater mit der Data Envelopment Analysis | Kleine, Andreas <br> Hoffmann, Steffen |
| :--- | :--- | :--- | :--- |
| 489 | 2015 | Konsensuale Peer-Wahl in der DEA -- Effizienz vs. <br> Skalenertrag | Dellnitz, Andreas <br> Reucher, Elmar |
| 490 | 2015 | Makroprudenzielle Regulierung - eine kurze Einführung und <br> ein Überblick | Velauthapillai, <br> Jeyakrishna |
| 491 | 2015 | SEM modeling with singular moment matrices <br> Part III: GLS estimation | Singer, Hermann |
| 492 | 2015 | Die steuerliche Berücksichtigung von Aufwendungen für ein <br> Studium - Eine Darstellung unter besonderer <br> Berücksichtigung des Hörerstatus | Meyering, Stephan <br> Portheine, Kea |
| 493 | 2016 | Ungewissheit versus Unsicherheit in Sozialen Netzwerken | Rödder, Wilhelm <br> Dellnitz, Andreas <br> Gartner, Ivan |
| 494 | 2016 | Investments in supplier-specific economies of scope with two <br> different services and different supplier characters: two <br> specialists | Fandel, Günter <br> Trockel, Jan |
| 495 | 2016 | An application of the put-call-parity to variance reduced <br> Monte-Carlo option pricing | Müller, Armin |
| 496 | 2016 | A joint application of the put-call-parity and importance <br> sampling to variance reduced option pricing | Müller, Armin |
| 497 | 2016 | Simulated Maximum Likelihood for Continuous-Discrete <br> State Space Models using Langevin Importance Sampling | Singer, Hermann |
| 498 | 2016 | A Theory of Affective Communication | Julmi, Christian |
| 499 | 2016 | Approximations of option price elasticities for importance <br> sampling | Müller, Armin |


[^0]:    ${ }^{1}$ FernUniversität in Hagen, Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung, D-58084 Hagen, Germany, armin.mueller@fernuni-hagen.de

[^1]:    ${ }^{2}$ A similar approach was presented by Melchior and Öttinger (1995) [13].
    ${ }^{3}$ For details on the Feynman-Kac representation see [14], chapter 5.7 and [15], chapter 5.H and appendix E.

[^2]:    ${ }^{4}$ Note, that the differential equation of the component $Y(t)$ does not include a diffusion term. Therefore, for the bivariate process $X(t)$, a singular diffusion matrix $\Omega$ results. To circumvent this problem in simulations, we treat the stochastic process as a univariate process with just one component $S(t)$. We then calculate the arithmetic average of the underlying values $S(t)$. For more details on arithmetic Asian options, see also [10].
    ${ }^{5}$ Details can be found in Appendix B in [5].
    ${ }^{6}$ For details see [12], chapter 9.1.

[^3]:    ${ }^{7}$ For details see [17], sections 4.2-4.4.
    ${ }^{8}$ Note that in this expression the Einstein summation convention is used, e.g., $f_{\alpha} \frac{\partial}{\partial X_{0, \alpha}}=$ $\sum_{\alpha} f_{\alpha} \frac{\partial}{\partial X_{0, \alpha}}$.

[^4]:    ${ }^{9}$ For details see [5], Appendix C.

[^5]:    ${ }^{10}$ The pay-off function can also be written as a function of the underlying's integral with respect to time, i.e. $Y(t)=\int_{0}^{t} S(\tau) \mathrm{d} \tau$. A bivariate process $X(t)$ with the components $S(t)$ and $Y(t)$ results. For practical reasons, we just consider the univariate process $S(t)$ and calculate its arithmetic average over time. Compare footnote 4 ,

[^6]:    ${ }^{11}$ For details see [3], chapter 6 or [12], chapter 8.

[^7]:    ${ }^{12} \theta(x)=\left\{\begin{array}{l}0 \text { for } x<0 \\ 1 \text { for } x \geq 0\end{array}\right.$

[^8]:    ${ }^{13}$ Note that this expression will not manage to push all trajectories at the level of $K$ as $\epsilon$ only has an effect on contributions to $\bar{S}_{n}$ resulting from the interval $\left[\tau_{k+1}, \tau_{n}\right]$. Contributions from prior instances, i.e. ones that result from the interval $\left[\tau_{0}, \tau_{k}\right]$, are not affected. Therefore, the approximation is biased downwards. Further refinement could lead to additional variance reduction.
    ${ }^{14}$ It should be noted that $\epsilon$ has to be calculated for every discretization step, i.e. in total $N \cdot n$ calculations of $\epsilon$ are required.

[^9]:    ${ }^{15}$ This approach has been suggested in [5, 20] but not yet analyzed.

[^10]:    ${ }^{16}$ Numerical values of weights and abscissae can be found at [26].

