On the data-driven COS method

Álvaro Leitao[a,b,*], Cornelis W. Oosterlee[b,a], Luis Ortiz-Gracia[c], Sander M. Bohte[b]

[a] *TU Delft, Delft Institute of Applied Mathematics, Delft, The Netherlands*
[b] *CWI-Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*
[c] *University of Barcelona, School of Economics, Barcelona, Spain*

Abstract

In this paper, we present the data-driven COS method, ddCOS. It is a Fourier-based financial option valuation method which assumes the availability of asset data samples: a characteristic function of the underlying asset probability density function is not required. As such, the presented technique represents a generalization of the well-known COS method [1]. The convergence of the proposed method is $\mathcal{O}(1/\sqrt{n})$, in line with Monte Carlo methods for pricing financial derivatives. The ddCOS method is then particularly interesting for density recovery and also for the efficient computation of the option's sensitivities Delta and Gamma. These are often used in risk management, and can be obtained at a higher accuracy with ddCOS than with plain Monte Carlo methods.

Keywords: The COS method, Density estimation, data-driven approach, Greeks, Delta-Gamma approach, the SABR model

1. Introduction

In quantitative finance, statistical distributions are commonly used for the valuation of financial derivatives and within risk management. The underlying assets are often modeled by means of stochastic differential equations (SDEs). Except for the classical and most simple asset models, the corresponding *probability density function* (PDF) and *cumulative distribution function* (CDF) are typically not known and need to be approximated.

In order to compute option prices, and to approximate statistical distributions, Fourier-based methods are commonly used numerical techniques. They are based on the connection between the PDF and the *characteristic function* (ChF), which is the Fourier transform of the probability density. The ChF is

*Corresponding author.
URL: A.LeitaoRodriguez@tudelft.nl (Álvaro Leitao), c.w.oosterlee@cwi.nl (Cornelis W. Oosterlee), luis.ortiz-gracia@ub.edu (Luis Ortiz-Gracia), S.M.Bohte@cwi.nl (Sander M. Bohte)

often available, and sometimes even in closed form, for the broad class of regular diffusions and also for Lévy processes. Some representative efficient Fourier pricing methods include those by Carr and Madan [2], Boyarchenko and Levendorskii [3], Lewis [4] and Fang and Oosterlee [1]. Here, we focus on the COS method from [1], which is based on an approximation of the PDF by means of a cosine series expansion.

Still, however, the asset dynamics for which the ChF are known is not exhaustive, and for many relevant asset price processes we do not have such information to recover the density. In recent years several successful attempts have been made to employ Fourier pricing methods without the explicit knowledge of the ChF. In Grzelak and Oosterlee [5], for example, a hybrid model with stochastic volatility and stochastic interest rate was linearized by means of expectation operators to cast the approximate system of SDEs in the framework of affine diffusions. Ruijter and Oosterlee [6] discretized the governing asset SDEs first and then worked with the ChF of the discrete asset process, within the framework of the COS method. Borovykh et al. [7] used the Taylor expansion to derive a ChF for which they could even price Bermudan options highly efficiently. In this work, we extend the applicability of the COS method to the situation where only data (samples from an unknown distribution) are available.

The density estimation problem, using a *data-driven* PDF, has been intensively studied in the last decades, particularly since it is a component in the *machine learning* framework [8]. Basically, density estimators can be classified into parametric and non-parametric estimators. The first type relies on the fact that prior knowledge is available (like moments) to determine the relevant parameters, while for non-parametric estimators the parameters need to be determined solely from the samples themselves. Within this second type of estimators we can find histograms, kernel density and orthogonal series estimators. A thorough description of these estimators is provided in [9]. More recently, some applications in finance have also appeared, see [10, 11, 12], for example.

For the valuation of financial derivatives, we will combine density estimators with Fourier-based methods, so orthogonal series form a natural basis. We will focus on the framework of *statistical learning*, see [13]. In statistical learning, a regularization is employed to derive an expression for the data-driven empirical PDF. By representing the unknown PDF as a cosine series expansion, a closed-form solution of the regularization problem is known [13], which forms the basis of the *data-driven COS method* (ddCOS). However, in order to employ the COS method machinery, underlying risk-neutral asset samples are required, i.e. they need to be generated according to some underlying model. This fact implies that the technique presented here results in a hybrid Monte Carlo-Fourier method.

The use of the COS method gives us expressions for option prices and, in particular, for the *option sensitivities or Greeks*. These option Greeks are the derivatives of option price with respect to a variable or parameter. The efficient computation of the Greeks is a challenging problem when only asset samples are available. Existing approaches are based on Monte Carlo-based techniques, like on finite-differences (bump and revalue), pathwise or likelihood ratio techniques, for which details can be found in [14], chapter 7. Several extensions

and improvements of these approaches have appeared, for example, based on adjoint formulations [15], the ChF [16, 17], Malliavin calculus [18, 19], algorithmic differentiation [20, 21] or combinations of these [22, 23, 24]. Intuitively, the ddCOS method follows a similar approach as likelihood ratio method, i.e. it relies on the differentiation of the (recovered) density function. On the other hand, our method can be also related with the improved methodologies employing the so-called *Malliavin derivative*, since it introduces a sample-based *weighted coefficients* that multiply the payoff coefficients. For both techniques, the differentiation of the payoff function (or payoff coefficients) is avoided.

All in all, the computation of the Greeks can be quite involved. The ddCOS method is not directly superior to Monte Carlo methods for option valuation, but it is competitive for the computation of the corresponding sensitivities. We derive simple expressions for the Greeks Delta and Gamma. The importance of Delta and Gamma in dynamic hedging and risk management is well-known. A useful application is found in the *Delta-Gamma approach* [25] to quantify market risk. The approximation of risk measures like *Value-at-Risk* (VaR) and *Expected Shortfall* (ES) under the Delta-Gamma approach is still nontrivial. Next to Monte Carlo methods, Fourier techniques have been employed in this context, when the ChF of the change in the value of the option portfolio is known (see [26, 27]). For example, the COS method has been applied in [28] to efficiently compute the VaR and ES under the Delta-Gamma approach. The ddCOS method may generalize the applicability to the case where only data is available.

This paper is organized as follows. The ddCOS method, and the origins in statistical learning and Fourier-based option pricing, are presented in Section 2. Variance reduction techniques can also be used within the ddCOS method, providing an additional convergence improvement. We provide insight and determine values for the method's open parameters in Section 3. Numerical experiments, with a focus on the option Greeks, are presented in Section 4. We conclude in Section 5.

2. The data-driven COS method

In this section we will discuss the ddCOS method, in which aspects of the Monte Carlo method, density estimators and the COS method are combined to approximate, in particular, the option Greeks Delta and Gamma. We will focus on European options here.

The COS method in [1] is a Fourier-based method by which option prices and sensitivities can be computed for various options under different models. The method relies heavily on the availability of the ChF, i.e., the Fourier transform of the PDF. In the present work, we assume that only asset samples are available, not the ChF, resulting in the data-driven COS method. It is based on regularization in the context of the statistical learning theory, presented briefly in Section 2.2. The connection with the COS method is found in the fact that the data-driven PDF appears as a cosine series expansion.

2.1. The COS method

The starting point for the well-known COS method is the risk-neutral option valuation formula, where the value of a European option at time t, $v(x,t)$, is an expectation under the risk neutral pricing measure, i.e.,

$$v(x,t) = \mathrm{e}^{-r(T-t)}\mathbb{E}\left[v(y,T)|x\right] = \mathrm{e}^{-r(T-t)}\int_{\mathbb{R}} v(y,T)f(y|x)\mathrm{d}y, \tag{1}$$

with r the risk-free rate, T the maturity time, and $f(y|x)$ the PDF of the underlying process, and $v(y,T)$ represents the option value at maturity time, being the payoff function. Typically, x and y are chosen to be scaled variables,

$$x := \log\left(\frac{S(0)}{K}\right) \quad \text{and} \quad y := \log\left(\frac{S(T)}{K}\right),$$

where $S(t)$ is the underlying asset process at time t, and K is the strike price.

Density $f(y|x)$ is unknown in most cases and in the COS method it is approximated, on a finite interval $[a,b]$, by a cosine series expansions, i.e.,

$$f(y|x) = \frac{1}{b-a}\left(A_0 + 2\sum_{k=1}^{\infty} A_k(x)\cdot\cos\left(k\pi\frac{y-a}{b-a}\right)\right),$$
$$A_0 = 1, \;\; A_k(x) = \int_a^b f(y|x)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y, \quad k = 1,2,\ldots.$$

By substituting this expression in Equation (1), interchanging the summation and integration operators using Fubini's Theorem, and introducing the following definition,

$$V_k := \frac{2}{b-a}\int_a^b v(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y,$$

we find that the option value is given by

$$v(x,t) \approx \mathrm{e}^{-r(T-t)}\sum_{k=0}^{\infty}{}' A_k(x)V_k, \tag{2}$$

where $'$ indicates that the first term is divided by two. So, the product of two real-valued functions in Equation (1) is transformed into the product of their cosine expansion coefficients, A_k and V_k. Density coefficients A_k can be computed by the ChF and V_k is known analytically (for many types of options).

Closed-form expressions for the option Greeks can also be derived. From the

COS option value formula, Δ and Γ are obtained by

$$\begin{aligned}\Delta &= \frac{\partial v(x,t)}{\partial S} = \frac{1}{S(0)}\frac{\partial v(x,t)}{\partial x} \approx \exp(-r(T-t))\sum_{k=0}^{\infty}{}'\frac{\partial A_k(x)}{\partial x}\frac{V_k}{S(0)},\\ \Gamma &= \frac{\partial^2 v(x,t)}{\partial S^2} = \frac{1}{S^2(0)}\left(-\frac{\partial v(x,t)}{\partial x}+\frac{\partial^2 v(x,t)}{\partial x^2}\right)\\ &\approx \exp(-r(T-t))\sum_{k=0}^{\infty}{}'\left(-\frac{\partial A_k(x)}{\partial x}+\frac{\partial^2 A_k(x)}{\partial x^2}\right)\frac{V_k}{S^2(0)}.\end{aligned} \tag{3}$$

Due to the rapid decay of the coefficients, $v(x,t)$, Δ and Γ can be approximated with high accuracy by truncating the infinite summation in Equations (2) and (3) to N terms. Under suitable assumptions, exponential convergence is proved and numerically observed.

2.2. Statistical learning theory for density estimation

In the setting of this paper, we assume a vector of n independent and identically distributed (i.i.d.) samples, $X_1, X_2, \ldots, X_n$. Based on these samples, we wish to find an accurate approximation of the PDF estimator, $f_n(x)$, which should approximate density $f(x)$.

By definition, the PDF is related to its CDF $F(x)$,

$$\int_{-\infty}^{x} f(y)\mathrm{d}y = F(x). \tag{4}$$

Function $F(x)$ is approximated by the empirical approximation,

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n}\eta(x-X_j), \tag{5}$$

where $\eta(\cdot)$ is a step function. This approximation converges to the "true CDF" with rate $\mathcal{O}(1/\sqrt{n})$ since, according to the central limit theorem, the estimation error defined as $\sqrt{n}\,(F_n(x)-F(x))$ follows the asymptotically normal distribution with zero mean (further details in [29]).

Rewriting Equation (4) as a linear operator equation, gives us,

$$Cf = F \approx F_n,$$

where the operator $Ch := \int_{-\infty}^{x} h(z)\mathrm{d}z$.

As explained in [13], this operator equation represents an ill-posed problem, and therefore a *risk functional* should be constructed, with a regularization term, as follows

$$R_{\gamma_n}(f, F_n) = L^2_{\mathcal{H}}(Cf, F_n) + \gamma_n W(f), \tag{6}$$

where $L_{\mathcal{H}}$ is a metric of the space $\mathcal{H}$ and $\gamma_n > 0$ is a parameter which gives a weight to the regularization term $W(f)$. The solution of $Cf = F_n$ belongs to

$\mathcal{D}$, the domain of definition of $W(f)$. Functional $W(f)$ takes real non-negative values in $\mathcal{D}$. Furthermore, $\mathcal{M}_c = \{f : W(f) \leq c\}$ is a compact set in the space where the solution exists and is unique.

The solution f_n, minimizing the functional in Equation (6), converges almost surely to the desired density. For the ill-posed density estimation problem, other conditions imposed for consistency (see details in [13], chapter 7), include

$$\begin{aligned} &\gamma_n \to \infty \quad \text{as} \quad n \to \infty, \text{ and} \\ &\frac{n}{\log n}\gamma_n \to \infty \quad \text{as} \quad n \to \infty. \end{aligned} \tag{7}$$

2.3. Regularization and Fourier-based density estimators

A relation exists between the regularization approach in Equation (6) and Fourier-based density approximation, more specifically, cosine series expansion estimators. By specific choices for the metric and the regularization term in Equation (6), i.e., $L_{\mathcal{H}} = L_2$, and

$$W(f) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathcal{K}(z-x) f(x) \mathrm{d}x \right)^2 \mathrm{d}z,$$

with kernel $\mathcal{K}(z-x)$, the functional reads

$$R_{\gamma_n}(f, F_n) = \int_{\mathbb{R}} \left(\int_0^x f(y)\mathrm{d}y - F_n(x) \right)^2 \mathrm{d}x + \gamma_n \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathcal{K}(z-x) f(x) \mathrm{d}x \right)^2 \mathrm{d}z. \tag{8}$$

Denoting by $\hat{f}(u)$, $\hat{F}_n(u)$ and $\hat{\mathcal{K}}(u)$ the Fourier transforms of $f(x)$, $F_n(x)$ and $\mathcal{K}(x)$, respectively, an expression for $\hat{F}_n(u)$ can be derived by applying Fourier transformation to Equation (5),

$$\begin{aligned} \hat{F}_n(u) &= \frac{1}{2\pi} \int_{\mathbb{R}} F_n(x) \mathrm{e}^{-iux} \mathrm{d}x \\ &= \frac{1}{2n\pi} \int_{\mathbb{R}} \sum_{j=1}^{n} \eta(x - X_j) \mathrm{e}^{-iux} \mathrm{d}x = \frac{1}{n} \sum_{j=1}^{n} \frac{\exp(-iuX_j)}{iu}, \end{aligned}$$

where $i = \sqrt{-1}$ is the imaginary unit.

By employing the *convolution theorem* and *Parseval's identity*, Equation (8) can be rewritten as

$$R_{\gamma_n}(f, F_n) = \left\| \frac{\hat{f}(u) - \frac{1}{n}\sum_{j=1}^{n} \exp(-iuX_j)}{iu} \right\|_{L_2}^2 + \gamma_n \left\| \hat{\mathcal{K}}(u)\hat{f}(u) \right\|_{L_2}^2.$$

As the functional $R_{\gamma_n}(f, F_n)$ is quadratic with respect to $\hat{f}$, the condition for its minimum (see [13], for details) is given by,

$$\frac{\hat{f}(u)}{u^2} - \frac{1}{nu^2} \sum_{j=1}^{n} \exp(-iuX_j) + \gamma_n \hat{\mathcal{K}}(u)\hat{\mathcal{K}}(-u)\hat{f}(u) = 0,$$

which entails that the Fourier transform of the solution (in terms of n) has the expression

$$\hat{f}_n(u) = \left(\frac{1}{1 + \gamma_n u^2 \hat{\mathcal{K}}(u)\hat{\mathcal{K}}(-u)} \right) \frac{1}{n} \sum_{j=1}^{n} \exp(-iuX_j). \tag{9}$$

Once the Fourier transform of the solution for the general regularization problem has been derived, we then find the connection with the series estimators, particularly in the form of cosine series expansion. For this, we further assume that the kernel $\mathcal{K}$ is the p-th derivative of the Dirac delta function, i.e., $\mathcal{K}(x) = \delta^{(p)}(x)$, and the desired PDF, $f(x)$, belongs to the class of functions whose p-th derivative ($p \geq 0$) belongs to $L_2(0, \pi)$, the risk functional becomes

$$R_{\gamma_n}(f, F_n) = \int_0^\pi \left(\int_0^x f(y)\mathrm{d}y - F_n(x) \right)^2 \mathrm{d}x + \gamma_n \int_0^\pi \left(f^{(p)}(x) \right)^2 \mathrm{d}x. \tag{10}$$

Given a series expansion in *orthonormal functions*, $\psi_1(\theta), \ldots, \psi_k(\theta), \ldots, \theta \in (0, \pi)$, the approximation to the unknown PDF will be of the form

$$f_n(\theta) = \frac{1}{\pi} + \frac{2}{\pi} \sum_{k=1}^{\infty} \tilde{A}_k \psi_k(\theta), \tag{11}$$

with $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_k, \ldots$ expansion coefficients, defined as $\tilde{A}_k =< f_n, \psi_k >$.

We need to compute the expansion coefficients so that the functional in Equation (10) is minimized. The coefficients $\tilde{A}_k$ cannot be directly computed from the definition since the unknown PDF, f_n, is implicitly involved in the expression, i.e.,

$$\begin{aligned} \tilde{A}_k &=< f_n, \psi_k >=< \hat{f}_n, \hat{\psi}_k > \\ &= \int_0^\pi \left(\left(\frac{1}{1 + \gamma_n u^2 \hat{\mathcal{K}}(u)\hat{\mathcal{K}}(-u)} \right) \frac{1}{n} \sum_{j=1}^{n} \exp(-iu\theta_j) \right) \cdot \hat{\psi}_k(u)\mathrm{d}u. \end{aligned}$$

Using cosine series expansions, i.e., $\psi_k(\theta) = \cos(k\theta)$, it is well-known that $\hat{\psi}_k(u) = \frac{1}{2}(\delta(u - k) + \delta(u + k))$. This facilitates the computation of the series coefficients, $\tilde{A}_k$, avoiding the calculation of the integral. Thus, the minimum of

the functional using cosine series expansions is obtained when

$$\begin{aligned}\tilde{A}_k &= \frac{1}{2n}\left(\left(\frac{1}{1+\gamma_n(-k)^2\hat{\mathcal{K}}(-k)\hat{\mathcal{K}}(k)}\right)\sum_{j=1}^{n}\exp(ik\theta_j)\right.\\ &\quad\left.+\left(\frac{1}{1+\gamma_n k^2\hat{\mathcal{K}}(k)\hat{\mathcal{K}}(-k)}\right)\sum_{j=1}^{n}\exp(-ik\theta_j)\right)\\ &= \frac{1}{1+\gamma_n k^2\hat{\mathcal{K}}(k)\hat{\mathcal{K}}(-k)}\frac{1}{n}\sum_{j=1}^{n}\cos(k\theta_j)\\ &= \frac{1}{1+\gamma_n k^{2(p+1)}}\frac{1}{n}\sum_{j=1}^{n}\cos(k\theta_j),\end{aligned} \tag{12}$$

where $\theta_j \in (0,\pi)$ are given samples of the unknown distribution. In the last step, $\hat{\mathcal{K}}(u) = (iu)^p$ is used.

Assuming that the samples are given, the solution contains two free parameters: *regularization parameter* γ_n, and *smoothing parameter* p.

In Section 3, we will discuss the impact of the regularization parameter on the convergence to the density in terms of the number of samples. We will use $p = 0$ here.

Smoothing parameter example

In order to give an insight in the influence of parameter p on the approximation, in Figure 1 standard normal densities obtained for several values of p are shown. With p increasing, the densities get increasingly smooth. The choice $p = 0$ (regularizing the density itself and not imposing regularization upon its derivatives) appears most appropriate in our context of smooth densities.

[Figure 1 about here.]

2.4. The ddCOS method

We are now ready to present the ddCOS method, where we employ the series expansion coefficients from the regularization approach. We replace the A_k-coefficients from Equation (2) by those coefficients based on data, $\tilde{A}_k$ in Equation (12).

So, suppose we have risk neutral samples (or values) from an underlying asset at a future time t, i.e., $S_1(t), S_2(t), \ldots, S_n(t)$. We compute the value of a European option with maturity time T and strike price K, and require therefore the samples $S_j(T)$. With a logarithmic transformation, we have

$$Y_j := \log\left(\frac{S_j(T)}{K}\right).$$

Before employing these samples in the regularization approach and because the solution is defined in $(0, \pi)$, we need to transform the samples by the following change of variables,

$$\theta_j = \pi \frac{Y_j - a}{b - a},$$

where the boundaries a and b are defined as

$$a := \min_{1 \le j \le n} (Y_j), \quad b := \max_{1 \le j \le n} (Y_j).$$

The A_k coefficients in Equation (2) are replaced by the data-driven $\tilde{A}_k$ in Equation (12),

$$A_k \approx \tilde{A}_k = \frac{\frac{1}{n} \sum_{j=1}^{n} \cos\left(k\pi \frac{Y_j - a}{b-a}\right)}{1 + \gamma_n k^{2(p+1)}}.$$

The ddCOS pricing formula for European options based on risk neutral data is now obtained as

$$\begin{aligned} \tilde{v}(x,t) &= \mathrm{e}^{-r(T-t)} \sum_{k=0}^{\infty}{}' \frac{\frac{1}{n} \sum_{j=1}^{n} \cos\left(k\pi \frac{Y_j - a}{b-a}\right)}{1 + \gamma_n k^{2(p+1)}} \cdot V_k \\ &= \mathrm{e}^{-r(T-t)} \sum_{k=0}^{\infty}{}' \tilde{A}_k V_k. \end{aligned} \tag{13}$$

As in the original COS method, we must truncate the infinite sum in Equation (13) to a finite number of terms N, i.e.,

$$\tilde{v}(x,t) = \mathrm{e}^{-r(T-t)} \sum_{k=0}^{N}{}' \tilde{A}_k V_k, \tag{14}$$

which completes the ddCOS pricing formula.

The samples Y_j should originate from one initial state, i.e. the dependency on the state x is implicitly assumed. In the case of European options this is typically fulfilled. In the Monte Carlo method, for example, all simulated asset paths depart from the same point $S(0)$, so that $x := \log\left(\frac{S(0)}{K}\right)$.

Regarding the Greeks, we can also derive data-driven expressions for the Δ and Γ sensitivities. We first define the corresponding sine coefficients as

$$\tilde{B}_k := \frac{\frac{1}{n} \sum_{j=1}^{n} \sin\left(k\pi \frac{Y_j - a}{b-a}\right)}{1 + \gamma_n k^{2(p+1)}}.$$

Taking derivatives in Equation (14) w.r.t the samples, Y_j, and following the COS expression for the sensitivities in Equation (3), the data-driven Greeks, $\tilde{\Delta}$

and $\tilde{\Gamma}$, can be obtained by

$$\tilde{\Delta} = \mathrm{e}^{-r(T-t)} \sum_{k=0}^{N}{}' \tilde{B}_k \cdot \left(-\frac{k\pi}{b-a}\right) \cdot \frac{V_k}{S(0)},$$

$$\tilde{\Gamma} = \mathrm{e}^{-r(T-t)} \sum_{k=0}^{N}{}' \left(\tilde{B}_k \cdot \frac{k\pi}{b-a} - \tilde{A}_k \cdot \left(\frac{k\pi}{b-a}\right)^2\right) \cdot \frac{V_k}{S^2(0)}.$$

The obtained sample-based expressions for the Greeks keep the payoff coefficients invariant, while the density coefficients are differentiated. This fact again suggest a connection with the methods relying on the Malliavin derivative, where the payoff function is smartly weighted in order to compute the sensitivities.

2.4.1. Application of variance reduction

Because of the focus on asset path data, the ddCOS method is related to the Monte Carlo method. Variance reduction in Monte Carlo methods is typically achieved by the use of variance reduction techniques. The ddCOS method also admits an additional variance reduction, in this case, for the computation of the expansion coefficients, $\tilde{A}_k$. We show how to introduce *antithetic variates (AV)* to our method. Since one of the assumptions for the regularization approach is that the samples are i.i.d., an immediate application of AV is not possible. Therefore, if we assume that antithetic samples, Y_i', to the original samples Y_i, can be computed without any serious computational effort, a new estimator for the coefficients can be defined as

$$\bar{A}_k := \frac{1}{2}\left(\tilde{A}_k + \tilde{A}_k'\right),$$

where we denote by $\tilde{A}_k'$ the corresponding "antithetic coefficients", obtained by Y_i'. By a similar derivation as for the standard AV technique, it can be proved that the use of coefficients $\bar{A}_k$ will give us a variance reduction compared to using the $\tilde{A}_k$ coefficients. Other variance reduction techniques may also be considered for the ddCOS method under the assumption of i.i.d. samples.

In order to reduce the variance of any estimator, additional information may be introduced. A well-known property to fulfill is the martingale property. To preserve this property, a simple transformation of the samples can be made by

$$\begin{aligned} S(T) &= S(T) - \frac{1}{n}\sum_{j=1}^{n} S_j(T) + \mathbb{E}[S(T)], \\ &= S(T) - \frac{1}{n}\sum_{j=1}^{n} S_j(T) + S(0)\exp(rT). \end{aligned}$$

As this modification is performed over the samples, it can also be used in the context of the ddCOS method.

3. Choice of Parameters in ddCOS Method

In this section, the selection and the influence of the regularization parameter γ_n in the ddCOS method is studied.

3.1. Regularization parameter γ_n

The regularization parameter γ_n plays an important role in the empirical PDF f_n. Without the inclusion of the regularization term, the density approximation provided by Equation (10), would give us a standard orthogonal series estimator. The choice of parameter γ_n impacts the efficiency of the data-driven COS method, since it is related to the required number of data samples, and by reducing the number of samples, the overall computational cost can be reduced.

The first option for choosing the regularization parameter, γ_n, which was proposed in [13], is given by the following rule,

$$\gamma_n = \frac{\log\log n}{n}. \tag{15}$$

As proved in [13], this rule provides a robust asymptotic rate of convergence under the assumption of a compactly supported density. It implies, with probability one, uniformly converging approximations f_n to the unknown density. Note, however, that the regularization parameter does not satisfy the second condition in Equation (7).

Although Equation (15) ensures an optimal asymptotic convergence in terms of n, it may not be the optimal γ_n-value for density estimation with a given fixed amount of samples. For that purpose we can exploit the relation between the empirical CDF, $F_n(x)$, and the unknown CDF, $F(x)$. This relation can be modeled by means of different *statistical laws*. Some examples include the Kolmogorov-Smirnov, Anderson-Darling, Kuiper and the Smirnov-Cramér-von Mises laws, by which a measure of the distance, or, *goodness-of-fit*, between $F_n(x)$ and $F(x)$ can be defined.

We are interested in a statistic which has a distribution, independent of the actual CDF and the number of samples n, and consider the Smirnov-Cramér-von Mises (SCvM) statistic [30, 31], defined as

$$\omega^2 := n \int_{\mathbb{R}} \left(F(x) - F_n(x)\right)^2 \mathrm{d}F(x).$$

Based on an approximation of the desired PDF, f_{γ_n} (depending on γ_n) and thus the CDF, F_{γ_n}, we choose the regularization parameter such that F_{γ_n} satisfies the SCvM statistic optimally, by solving (in terms of γ_n) the following equation:

$$\int_{\mathbb{R}} \left(F_{\gamma_n}(x) - F_n(x)\right)^2 \mathrm{d}F(x) = \frac{m_{\omega^2}}{n},$$

where m_{ω^2} is the mean of the SCvM statistic, ω^2. In the one-dimensional case, a simplified expression can be derived [32, 31], i.e.,

$$\sum_{j=1}^{n} \left(F_{\gamma_n}(\bar{X}_j) - \frac{i-0.5}{n}\right)^2 = m_{\omega^2} - \frac{1}{12n}, \tag{16}$$

with $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_n$, the ordered array of samples $X_1, X_2, \ldots, X_n$. It can be proved (details in [13]) that, by solving Equation (16) under the assumption that the solution is in the form of a cosine series expansion, a regularization parameter can be determined, which provides an almost optimal rate of convergence in n towards the desired density function.

Next to the improvement of the method's convergence, the quality of the density approximation, in terms of the considered expansion coefficients, is also influenced by the regularization parameter.

In order to assess the impact of γ_n on the quality of approximation, we employ the well-known *Mean Integrated Squared Error* (MISE), which is commonly used in density estimation (also known as the *risk function*). The formal definition reads

$$\mathbb{E}\left[\|f_n - f\|_2^2\right] = \mathbb{E}\left[\int_{\mathbb{R}} \left(f_n(x) - f(x)\right)^2 \mathrm{d}x\right].$$

In our case, the MISE can be decomposed into two terms (see [33]), as follows,

$$\mathbb{E}\left[\int_0^\pi \left(f_n(x) - f(x)\right)^2 \mathrm{d}x\right] = \sum_{k=1}^{N} \mathbb{V}\mathrm{ar}\left[\tilde{A}_k\right] + \sum_{k=N+1}^{\infty} A_k^2$$

where the A_k are the "true" coefficients from Equation (2), and the $\tilde{A}_k$ are from Equation (12). The MISE, as it is defined, is the summation of the bias and the variance of the estimator. In the Fourier cosine expansions context, an increasing N implies smaller bias (but bigger variance). The opposite also holds i.e. small N produces more bias and smaller variance. We need to compute the variance of the data-driven coefficients, $\tilde{A}_k$. By Equation (12), we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\tilde{A}_k\right] &= \mathbb{V}\mathrm{ar}\left[\frac{1}{1+\gamma_n k^{2(p+1)}} \frac{1}{n} \sum_{j=1}^{n} \cos(k\theta_j)\right] \\ &= \frac{1}{\left(1+\gamma_n k^{2(p+1)}\right)^2} \frac{1}{n^2} \sum_{j=1}^{n} \mathbb{V}\mathrm{ar}\left[\cos(k\theta_j)\right].\end{aligned}$$

By basic trigonometric properties, this variance can be computed as

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\cos(kx)\right] &= \mathbb{E}\left[\cos^2(kx)\right] - \left(\mathbb{E}\left[\cos(kx)\right]\right)^2 \\ &= \int_0^\pi \cos^2(kx) f(x) \mathrm{d}x - \left(\int_0^\pi \cos(kx) f(x) \mathrm{d}x\right)^2 \\ &= \int_0^\pi \left(\frac{1+\cos(2kx)}{2}\right) f(x) \mathrm{d}x - A_k^2 = \frac{1}{2} + \frac{A_{2k}}{2} - A_k^2,\end{aligned}$$

where the definition of the expansion coefficients is used in steps 2 and 3.

The expression for the variance of $\tilde{A}_k$ then reads

$$\mathbb{V}\mathrm{ar}\left[\tilde{A}_k\right] = \frac{1}{\left(1+\gamma_n k^{2(p+1)}\right)^2} \frac{1}{n} \left(\frac{1}{2} + \frac{1}{2} A_{2k} - A_k^2\right),$$

and the MISE is thus given by

$$\mathrm{MISE} = \frac{1}{n}\sum_{k=1}^{N}\frac{1}{\left(1+\gamma_n k^{2(p+1)}\right)^2}\left(\frac{1}{2}+\frac{1}{2}A_{2k}-A_k^2\right)+\sum_{k=N+1}^{\infty}A_k^2. \qquad (17)$$

Example

The error measure defined in Equation (17) is employed to analyze the influence of the regularization. We use the standard normal distribution as a reference test case. The coefficients that are based on the available analytic solution are replaced by the corresponding data-driven coefficients that depend on γ_n.

In Figure 2a, we present the convergence results for different regularization parameters. Next to the rules suggested by Equations (15) and (16), we also include the case $\gamma_n = 0$ to highlight the benefits of employing the regularization approach. The obtained results confirm the improvements provided by both γ_n-rules, with the almost optimal γ_n given by the SCvM statistic.

[Figure 2 about here.]

A second aspect which is influenced by γ_n is the accuracy with respect to the number of expansion terms N in Equation (14). For this, in Figure 2b we present the MISE for the standard normal distribution when different coefficients A_k are employed: A_k by the γ_n-rule (15) (red lines), A_k by the SCvM (16) (blue lines) and, as a reference, the A_k-coefficients obtained by the ChF (black dashed line). We notice that, when $\gamma_n = 0$ is used in the MISE formula (all dashed lines), for increasing value of N, the approximation deteriorates, resulting in increasing approximation errors. In contrast, when the corresponding γ_n is used (regular lines), the error stabilizes. Since the number of expansion coefficients is typically chosen high, this property of the regularization approach is useful.

3.1.1. Optimal N-values

As mentioned, with an increasing number of series expansion coefficients N, the approximation based on the regularization approach does not improve any further. This fact indicates that we need to determine an *optimal* value of N, i.e. the smallest value of N for which the MISE stabilizes, see Figure 2b. Using small N-values is important within the data-driven methodology, since parameter N considerably affects the performance of the method.

We propose an empirical procedure to compute the optimal value of N. The MISE in Equation (17) depends on the number of samples n, the number of coefficients N, and the coefficients themselves A_k. Since we wish to compute no more coefficients than necessary, we focus on the parameters n and N. Regarding the influence of the number of samples, see also the curves in Figure 3a, higher n-values also require higher N-values to ensure stabilizing errors in the MISE curve. To determine a relation between n and N, we need to simplify the MISE formula as we desire a closed-form expression. We discard the second part in Equation (17), as it goes to zero when N increases. Within the first part,

we approximate the quantity $\left(\frac{1}{2} + \frac{1}{2}A_{2k} - A_k^2\right) \approx \frac{1}{2}$. Then, the approximate formula for the MISE is found to be

$$\text{MISE} \approx \frac{1}{n}\sum_{k=1}^{N} \frac{\frac{1}{2}}{\left(1+\gamma_n k^{2(p+1)}\right)^2} =: \text{MISE}_N,$$

where n and N are directly connected, but also γ_n appears. It is possible to prove that the above MISE proxy, MISE_N, is an upper bound for the first part in Equation (17).

From Figure 3b, we observe two important facts: the MISE proxy provides a highly satisfactory approximation for the first addend of the MISE, which converges to the MISE when N increases. By combining these two observations, we will employ the MISE proxy to determine the optimal number of terms N. Since the computation of γ_n by Equation (16) involves N, we use the case where γ_n is determined by Equation (15) (which only depends on n). Figure 3b shows that the MISE_N (to a different level of accuracy) is very similar in both cases, where the γ_n rule appears conservative, i.e. bigger N-values are required to reach the non-decreasing error region. The proposed procedure iteratively determines whether or not we reached error stabilization by checking the differences in MISE_N between two consecutive N-values. When this difference is less than a predefined tolerance, ϵ, we have approximated the optimal N-value. Since N should grow with n, we propose to use $\epsilon := \frac{1}{\sqrt{n}}$, i.e. the expected order of accuracy for the density approximation should be given in terms of the number of samples.

By collecting all described components, the approximately optimal N-value becomes a function only of n. The iterative methodology is described in Algorithm 1. In Figure 4, we observe that the resulting optimal N function is an increasing staircase function (with a predefined floor of $N = 5$), see also [33].

[Figure 3 about here.]

Algorithm 1: Optimal N-values.

Data: n, γ_n
$N_{min} = 5$
$N_{max} = \infty$
$\epsilon = \frac{1}{\sqrt{n}}$
$\text{MISE}_* = \infty$
for $N = N_{min} : N_{max}$ **do**
  $\text{MISE}_N = \frac{1}{n}\sum_{k=1}^{N} \frac{\frac{1}{2}}{\left(1+\gamma_n k^{2(p+1)}\right)^2}$
  $\epsilon_N = \frac{|\text{MISE}_N - \text{MISE}_*|}{|\text{MISE}_N|}$
  if $\epsilon_N > \epsilon$ **then**
    $N_{op} = N$
  else
    Break
  $\text{MISE}_* = \text{MISE}_N$

[Figure 4 about here.]

Now, we have described the techniques to determine values for the regularization parameter γ_n, and for the number of coefficients, N. By these, the ddCOS method is defined with only Monte Carlo samples as the input.

4. Applications of the ddCOS method

In this section, we present some applications of the ddCOS method. The first application is an option pricing experiment, where we show the method's convergence. Subsequently, we present the performance regarding the computation of the Greeks, where ddCOS exhibits a stable convergence and can be employed with involved models, as we only need asset samples. We also compute the Greeks under the SABR model. Once the Greeks have been efficiently approximated, they can be used for the computation of the VaR and ES risk measures within the Delta-Gamma approach. All steps in this methodology can be performed by the ddCOS method.

The experiments have been carried out on a computer system with the following characteristics: CPU Intel Core i7-4720HQ 2.6GHz and RAM memory of 16GB RAM. The employed software package is Matlab R2016b.

4.1. Option valuation and Greeks

First of all, we numerically test the convergence of the ddCOS method in an option valuation experiment. The *Geometric Brownian Motion* (GBM) asset dynamics are employed, since a reference value for the option value is available by the Black-Scholes formula. The regularization parameter γ_n is set as in Equation (15), as for option valuation experiments the difference between this rule and γ_n based on the SCvM statistic in Equation (16) is not significant. Moreover, the use of the γ_n rule provides faster ddCOS estimators.

As is common in Monte Carlo experiments, the *Mean Squared Error* (MSE) is considered as the error measure. In the convergence tests, the reported values are computed as the average of 50 experiments.

The expected order of convergence for the option values is $\mathcal{O}(1/\sqrt{n})$, according to the convergence of the empirical CDF towards the true CDF in Equation (5). In Section 2.4.1, the application of antithetic variates in the ddCOS framework has been presented. In Figure 5, we confirm that this variance reduction technique provides a similar improvement in terms of precision as when it is applied to the plain Monte Carlo method. Another observation is that, under this particular setting, the estimators (both ddCOS and Monte Carlo) of the put option value result in smaller variances than the call option estimators. In terms of accuracy, it is thus worth computing the put value and then use the *call-put parity* formula for call options. In addition, the use of the put together with the call-put parity is recommended since call payoff functions grow exponentially and may give rise to cancellation errors.

[Figure 5 about here.]

We have empirically shown in Figure 5 that the ddCOS method converges to the true price with the expected convergence rate $\mathcal{O}(1/\sqrt{n})$, which resembles the plain Monte Carlo convergence. However, by the ddCOS method, not only the option value but also the sensitivities can readily be obtained. This is an advantage w.r.t Monte Carlo-based methods for estimating sensitivities, where often, additional simulations, intermediate time-steps or prior knowledge are required.

Thus, a similar convergence test is performed for the Δ and Γ sensitivities, see Figure 6. As Monte Carlo-based method for the Greeks calculation we consider the *Finite Difference* method (bump and revalue, denoted as MCFD). We have chosen MCFD for the comparison because it is flexible and it does not require prior knowledge. MCFD may require one or two extra simulations, and the choice of optimal *shift* parameter may not be trivial. The reference Delta and Gamma are given by the Black-Scholes formula. In Figure 6 we observe the expected convergence and the reduction in the variance due to the use of AV. In both experiments, while the Δ is very well approximated by the ddCOS and MCFD methods, the second derivative, Γ, appears more complicated for the MCFD method. This fact was already pointed out by Glasserman in [14]. The ddCOS estimator, however, is accurate and stable as it is based on the data-driven PDF and the COS machinery. That implies that our method does not suffer from the instabilities and potential important errors (specially in the second derivative) generated by the finite difference approximation. Therefore, the ddCOS provides a fast convergence with reduced variance, easily further improved by applying variance reduction techniques.

[Figure 6 about here.]

Using $n = 10^5$, in Table 1 we now compare the Δ and Γ estimations obtained under the GBM dynamics for several strikes. The performance of the ddCOS method is very satisfactory as it is accurate, with small *Relative Error* (RE, averaged over K) and reproduces the reference values very well. The difficulties of the MCFD estimating Γ are more clearly visible.

[Table 1 about here.]

We wish to test the ddCOS method in a more complex situation, by adding jumps in the form of a Merton jump-diffusion asset price process. To accurately compute the option sensitivities in this case gives rise to difficulties for Monte Carlo-based methods. We perform a similar experiment as before, where now the underlying asset follows the Merton jump-diffusion model, and the obtained Δ and Γ are presented in Table 2. In this case, the reference value is provided by the COS method at a high accuracy.

[Table 2 about here.]

In terms of computational cost, the ddCOS method is a competitive alternative, as additional simulations are not needed. Notice that in these latter experiments AV techniques are not employed. Under the Merton dynamics, the

ddCOS method takes 0.1813 seconds and MCFD 0.3149 seconds. In this case, the use of the ddCOS method reduces the computational costs, as the cost of an individual simulation by the Merton model is significantly higher than for GBM asset dynamics.

4.2. The SABR model

The SABR model [34] is interesting within the ddCOS framework since the ChF is not known and, furthermore, the asset path Monte Carlo simulation is not trivial. The model is a stochastic-local volatility model which is widely used in FX modeling, and is given by

$$\begin{aligned} \mathrm{d}S(t) &= \sigma(t)S^{\beta}(t)\mathrm{d}W_S(t), \quad S(0) = S_0 \exp\left(rT\right), \\ \mathrm{d}\sigma(t) &= \alpha\sigma(t)\mathrm{d}W_\sigma(t), \qquad\quad \sigma(0) = \sigma_0, \end{aligned} \tag{18}$$

where $S(t) = \bar{S}(t)\exp\left(r(T-t)\right)$ is the forward value of the underlying $\bar{S}(t)$, with r the interest rate, S_0 the spot price and T maturity time. The stochastic volatility process is denoted by $\sigma(t)$, with $\sigma(0) = \sigma_0$, $W_S(t)$ and $W_\sigma(t)$ are two correlated Brownian motions with correlation ρ (i.e. $W_S W_\sigma = \rho t$). The parameters of the SABR model are $\alpha > 0$ (the volatility of the volatility), $0 \leq \beta \leq 1$ (the elasticity) and ρ (the correlation coefficient).

In [34], the authors provided a closed-form approximation formula for the *implied volatility* under the SABR dynamics, which is often used within the calibration. However, the closed-form expression is derived by perturbation theory, and therefore the formula is not accurate for small strike values, for long time to maturity options or for high volatilities (see, for example, [35, 36]).

The calculation of the Greeks under the SABR model becomes challenging but can be addressed by the ddCOS method. To employ the method, we need samples of the underlying asset at time T. Here, we make use of the one time-step SABR Monte Carlo simulation introduced by Leitao et al. in [35]. This one time-step SABR simulation is based on the expression for the CDF of the conditional SABR process [37]. For $S(0) > 0$, the conditional CDF of $S(t)$ with an absorbing boundary at $S(t) = 0$, given the volatility $\sigma(t)$, and the conditional *time-integrated variance* $\int_0^t \sigma^2(s)\mathrm{d}s|\sigma(t)$, reads

$$Pr\left(S(t) \leq K | S(0) > 0, \sigma(t), \int_0^t \sigma^2(s)\mathrm{d}s\right) = 1 - \chi^2(a; b, c), \tag{19}$$

where

$$\begin{aligned} a &= \frac{1}{\nu(t)}\left(\frac{S(0)^{1-\beta}}{(1-\beta)} + \frac{\rho}{\alpha}\left(\sigma(t) - \sigma(0)\right)\right)^2, & c &= \frac{K^{2(1-\beta)}}{(1-\beta)^2\nu(t)}, \\ b &= 2 - \frac{1 - 2\beta - \rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}, & \nu(t) &= (1-\rho^2)\int_0^t \sigma^2(s)\mathrm{d}s, \end{aligned}$$

and $\chi^2(x; \vartheta, \xi)$ is the non-central chi-square CDF.

This formula is *exact* for the case $\rho = 0$ and results in an *approximation* otherwise. So, to apply the one time-step Monte Carlo simulation for the SABR dynamics, we perform the following steps (with the terminal time T):

- *Simulation of SABR's volatility.* From Equation (18), the volatility in the SABR model is governed by the well-known log-normal distribution.

- *Simulation of SABR's time-integrated variance, conditional on the terminal value of the volatility, i.e.,* $\int_0^T \sigma^2(s)\mathrm{d}s|\sigma(T)$. In [35], the authors proposed a combination of Fourier- and copulas-based techniques, resulting in a fast and accurate sampling procedure.

- *Simulation of SABR's forward asset process.* The forward dynamics are obtained by inverting the CDF in Equation (19). This inverse SABR distribution has to be calculated by means of some numerical approximation. The efficient inversion in [38] is the choice here.

Thus, the ddCOS method will be combined with the one time-step SABR simulation to efficiently compute Δ and Γ under the SABR dynamics.

For the numerical experiments, we consider two parameter settings. First of all, a basic parameter set is taken, where the SABR formula is valid and can be used as a reference. The results are presented in Table 3. For the second test we use a more difficult set of parameters (i.e., Set III in [35]), where the SABR formula does not provide accurate results anymore. In Table 4, we observe that the ddCOS provides accurate Δ-values in this case, without any problems. The reference value has been computed by the MCFD in combination with the SABR Monte Carlo simulation in [36], with a large number of Monte Carlo paths ($n = 10,000,000$) and time steps ($4T$). The convergence in n of the ddCOS Δ estimator under the SABR dynamics is shown in Figure 7a. The calculation of Γ when the underlying is governed by the SABR model is again involved and the MCFD estimation is not reliable. In Figure 7b, the convergence of the ddCOS Γ estimator is presented, where we observe convergence, with impressive variance reduction.

[Table 3 about here.]

[Table 4 about here.]

[Figure 7 about here.]

4.3. VaR, ES and the Delta-Gamma approach

In the evaluation of market risk, the computation of risk measures is important, and even mandatory for regulatory purposes to estimate the risk of large losses. With the risk factors denoted by S and a time horizon Δt, we define the change in S at time Δt by ΔS. The variation in S directly affects the value of a portfolio $V(S,t)$, containing derivatives of S. We denote the changes in the value of the portfolio by ΔV, so that the definition of the loss in interval $[t, \Delta t]$ is given by

$$L := -\Delta V = V(S,t) - V(S + \Delta S, t + \Delta t).$$

In order to manage possible large losses, we are interested in the distribution of L, specifically in the CDF, $F_L(x) = \mathbb{P}(L < x)$, which can be employed to compute the risk measures VaR or ES. The formal definition of the VaR reads

$$\mathbb{P}(\Delta V < \text{VaR}(q)) = 1 - F_L(\text{VaR}(q)) = q,$$

with q a predefined confidence level, whereas, given the VaR, the ES measure is computed as

$$\text{ES} := \mathbb{E}[\Delta V | \Delta V > \text{VaR}(q)].$$

So, VaR is given as a quantile of the loss distribution, while ES is the average of the largest possible losses.

Although simple in definition, the practical computation of these risk measures is a challenging and computationally expensive problem, especially when the changes in V cannot be assumed linear in S. Then, VaR and ES estimation is often performed by means of an Monte Carlo method. In order to find a balance between accuracy and tractability, one of the employed methodologies is the Delta-Gamma approximation which combines Monte Carlo path generation, a second-order Taylor expansion and the sensitivities to reduce the computational cost and capture the non-linearity in portfolio changes. The delta-gamma approximation of ΔV (in the case of only one risk factor) is given by

$$\Delta V \approx \sum_{i=1}^{M} w_i \frac{\partial v_i}{\partial S} \Delta S + \frac{1}{2} \sum_{i=1}^{M} w_i \frac{\partial^2 v_i}{\partial S^2} (\Delta S)^2,$$

with M the number of assets depending on risk factor S, w_i and v_i the amount and the value of asset i, respectively. The partial derivatives are evaluated at initial time t. In the case of options contracts, these partial derivatives correspond to the Δ and Γ sensitivities. It is usually assumed that the distribution of ΔS is known (normal, Student's $\mathbf{t}$, etc). Then, by applying the Delta-Gamma approach, the distribution of the losses, F_L, and therefore the VaR and the ES are easily calculated.

The use of the ddCOS method in the context of the Delta-Gamma approach generalizes its applicability. Since the use of the ChF is not longer required, we can assume non-trivial dynamics for ΔS, where the use of Fourier inversion methods (as in [28]) would be a limitation (it may be impossible to obtain a ChF). As we have seen, by employing the ddCOS method, Δ and Γ can be computed at once and, therefore, be directly employed within the Delta-Gamma approximation. The ddCOS method can thus be used to recover the distribution of ΔV, whenever samples are available. This can be useful when historical data is employed, and no particular distribution is assumed.

In order to show the performance of the ddCOS method within the Delta-Gamma approach, we first repeat the experiments from [28]. Two portfolios are considered, both with the same composition (one European call and half a European put under the same underlying asset, maturity 60 days and strike $K = 101$) but different time horizons, i.e. 1 day and 10 days. We denote them by Portfolio 1 and Portfolio 2, respectively. The underlying asset follows a GBM

with $S(0) = 100$, $r = 0.1$ and $\sigma = 0.3$. Change ΔS is assumed to be normally distributed.

In Figure 8 the recovered densities by the COS and ddCOS methods are depicted. An almost perfect fit is observed, with the expected small-sized oscillations in the data-driven approach. Since the computational domain is also driven by data, the ddCOS recovered density remains within the defined domain, avoiding incorrect estimations outside the domain (see the COS curve in Figure 8b).

[Figure 8 about here.]

We also employ the ddCOS method to compute the risk measures VaR and ES. The convergence of VaR and ES in terms of n is presented in Figure 9, with reference values provided by [28]. As expected, the convergence rate for both estimators is $\mathcal{O}(1/\sqrt{n})$.

[Figure 9 about here.]

4.3.1. Smoothing the density of L

As seen in Figure 8, the densities estimated by the ddCOS method exhibit some artificial oscillations due to the lack of data in particular regions and the so-called *Gibbs phenomenon*. Two possibilities to avoid the appearance of these oscillations are increasing smoothing parameter p, and the application of so-called spectral filters within the ddCOS formula (13). By parameter p we can include derivatives of the PDF into the regularization (see Equation (10)). We analyze the use of $p = 1$. Filtering was already successfully applied in the context of the Delta-Gamma approximation in [28], based on the work by Ruijter et al. [39], and we refer to the references for the filter details. Adding hte filter is almost trivial as it merely implies a multiplication with a specific filter term. Based to the references, we here choose the so-called 6-*th order exponential filter* within the ddCOS formula.

In Figure 10, the resulting densities from the application of both alternatives are presented. Whereas both smoothing techniques give highly satisfactory results for Portfolio 1, the spectral filters are superior in the case of Portfolio 2. Based on these tests, we suggest the use of a spectral filter to obtain smooth densities. Note, however, that the application of these smoothing procedures does not give us an improvement in the convergence, which is still dominated by the order of convergence in Equation (5).

[Figure 10 about here.]

4.3.2. Delta-Gamma approach under the SABR model

In order to further test the ddCOS method in the context of the Delta-Gamma approach, we now assume the dynamics of the underlying asset and ΔS to be governed by the SABR dynamics. In Figure 11a, the obtained VaR and ES when varying n are presented. No reference is available here, since the MCFD is unstable for the Γ computation under the SABR model. We observe

that, already for $n = 10^3$, a stable Γ-value is found and, even more important, the variance is negligible. By the ddCOS method, a closed-form expression for the loss distribution is also obtained. The recovered F_L and the corresponding PDF f_L are depicted in Figure 11b (for $n = 10^5$), where we also include the resulting densities when employing the exponential spectral filter.

[Figure 11 about here.]

In Table 5, the VaR and ES under SABR are presented for several choices of q, ranging from 10% to 90%. Again, the results seem to be coherent.

[Table 5 about here.]

5. Conclusions

In this work, the ddCOS method has been introduced. The method extends the COS method applicability to cases when only data samples of the underlying asset are available. The method exploits a closed-form solution, in terms of Fourier cosine expansions, of a density. The use of the COS machinery in combination with density estimation allowed us to develop a data-driven method which can be employed for option pricing and risk management. The ddCOS method particularly results in an efficient method for the Δ and Γ sensitivities computation, based solely on the samples. Therefore, it can be employed within the Delta-Gamma approximation for calculating risk measures. Through several numerical examples, we have empirically shown the convergence of our method. In some cases, in order to get monotonic densities, it may be beneficial to add a filter term to the ddCOS method.

A possible future extension may be the use of other basis functions. Haar wavelets are for example interesting since they provide positive densities and allow an efficient treatment of dynamic data.

Acknowledgment

We would like to thank prof. Yuying Li (U. Waterloo, Canada) for the fruitful discussions about the Statistical Learning Theory and its applications.

References

[1] F. Fang, C. W. Oosterlee, A novel pricing method for European options based on Fourier-cosine series expansions, SIAM Journal on Scientific Computing 31 (2008) 826–848.
URL `http://oai.cwi.nl/oai/asset/13283/13283A.pdf`

[2] P. Carr, D. B. Madan, Option valuation using the Fast Fourier transform, Journal of Computational Finance 2 (1999) 61–73.

[3] S. Boyarchenko, S. Levendorskii, Efficient variations of the Fourier transform in applications to option pricing, Journal of Computational Finance 18 (2) (2014) 57–90.

[4] A. L. Lewis, A Simple Option Formula for General Jump-Diffusion and Other Exponential Levy Processes, available at SSRN (2001).
URL https://ssrn.com/abstract=282110

[5] L. A. Grzelak, C. W. Oosterlee, An equity-interest rate hybrid model with stochastic volatility and the interest rate smile, Journal of Computational Finance 15 (4) (2012) 45–77.

[6] M. J. Ruijter, C. W. Oosterlee, Numerical Fourier Method and Second-order Taylor Scheme for Backward SDEs in Finance, Applied Numerical Mathematics 103 (C) (2016) 1–26. doi:10.1016/j.apnum.2015.12.003.
URL http://dx.doi.org/10.1016/j.apnum.2015.12.003

[7] A. Borovykh, A. Pascucci, C. W. Oosterlee, Pricing Bermudan Options Under Local Lvy Models with Default, To appear in Journal of Mathematical Analysis and Applications.
URL https://ssrn.com/abstract=2771632

[8] C. M. Bishop, Pattern Recognition and Machine Learning, Information Science and Statistics, Springer, 2006.
URL https://books.google.nl/books?id=kTNoQgAACAAJ

[9] B. W. Silverman, Density estimation for statistics and data analysis, Chapman & Hall, London, 1986.

[10] B. V. Es, P. Spreij, H. V. Zanten, Nonparametric volatility density estimation, Bernoulli 9 (3) (2003) 451–465.

[11] P. Malliavin, M. E. Mancino, A Fourier transform method for nonparametric estimation of multivariate volatility, The Annals of Statistics 37 (4) (2009) 1983–2010.

[12] R. Sen, C. Ma, Forecasting density function: Application in finance, Mathematical Finance 5 (2015) 433–447.

[13] V. N. Vapnik, Statistical Learning Theory, Wiley-Interscience, 1998.

[14] P. Glasserman, Monte Carlo Methods in Financial Engineering, Applications of Mathematics : Stochastic Modelling and Applied Probability, Springer, 2004.
URL https://books.google.es/books?id=e9GWUsQkPNMC

[15] M. B. Giles, Smoking adjoints: fast Monte Carlo Greeks, Risk Magazine 19 (1) (2006) 88–92.

[16] P. Glasserman, Z. Liu, Sensitivity estimates from characteristic functions, Operations Research 58 (6) (2010) 1611–1623. doi:10.1287/opre.1100.0837.
URL http://dx.doi.org/10.1287/opre.1100.0837

[17] P. Glasserman, Z. Liu, Estimating Greeks in simulating Lvy-driven models, Journal of Computational Finance 14 (2) (2011) 3–56.

[18] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, N. Touzi, Applications of Malliavin calculus to Monte Carlo methods in finance, Finance and Stochastics 3 (4) (1999) 391–412. doi:10.1007/s007800050068.
URL http://dx.doi.org/10.1007/s007800050068

[19] M. H. Davis, M. P. Johansson, Malliavin Monte Carlo Greeks for jump diffusions, Stochastic Processes and their Applications 116 (1) (2006) 101–129. doi:http://dx.doi.org/10.1016/j.spa.2005.08.002.
URL //www.sciencedirect.com/science/article/pii/S0304414905001146

[20] L. Capriotti, Fast Greeks by algorithmic differentiation, Journal of Computational Finance 14 (3) (2011) 3–35.

[21] J. du Toit, U. Naumann, Adjoint Algorithmic Differentiation Tool Support for Typical Numerical Patterns in Computational Finance, To appear in Journal of Computational Finance.
URL https://www.nag.co.uk/doc/techrep/pdf/tr3_14.pdf

[22] N. Chen, P. Glasserman, Malliavin Greeks without Malliavin calculus, Stochastic Processes and their Applications 117 (11) (2007) 1689–1723, recent Developments in Mathematical Finance: Special issue based on the {CCCP} Meeting, April 2006, New York, NY. doi:http://dx.doi.org/10.1016/j.spa.2007.03.012.
URL //www.sciencedirect.com/science/article/pii/S0304414907000877

[23] M. B. Giles, Vibrato Monte Carlo Sensitivities, in: P. L. Ecuyer, A. B. Owen (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2008, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 369–382. doi:10.1007/978-3-642-04107-5_23.
URL http://dx.doi.org/10.1007/978-3-642-04107-5_23

[24] L. Capriotti, M. B. Giles, Algorithmic differentiation: adjoint Greeks made easy, Risk Magazine 25 (10).

[25] M. Britten-Jones, S. M. Schaefer, Non-Linear Value-at-Risk, Review of Finance 2 (2) (1999) 161–187.
URL http://EconPapers.repec.org/RePEc:oup:revfin:v:2:y:1999:i:2:p:161-187.

[26] J. V. Siven, J. T. Lins, A. Szymkowiak-Have, Value-at-Risk computation by Fourier inversion with explicit error bounds, Finance Research Letters 6 (2) (2009) 95–105. doi:http://dx.doi.org/10.1016/j.frl.2008.12.002.
URL //www.sciencedirect.com/science/article/pii/S1544612308000652

[27] R. Chen, L. Yu, A novel nonlinear Value-at-Risk method for modeling risk of option portfolio with multivariate mixture of normal distributions, Economic Modelling 35 (2013) 796–804. doi:http://dx.doi.org/10.1016/j.econmod.2013.09.003.
URL //www.sciencedirect.com/science/article/pii/S0264999313003623

[28] L. Ortiz-Gracia, C. W. Oosterlee, Efficient VaR and Expected Shortfall computations for nonlinear portfolios within the delta-gamma approach, Applied Mathematics and Computation 244 (2014) 16–31. doi:http://dx.doi.org/10.1016/j.amc.2014.06.110.
URL //www.sciencedirect.com/science/article/pii/S0096300314009540

[29] A. W. van der Vaart, Asymptotic statistics, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1998.

[30] H. Cramér, On the composition of elementary errors, Scandinavian Actuarial Journal 1928 (1) (1928) 13–74. arXiv:http://dx.doi.org/10.1080/03461238.1928.10416862, doi:10.1080/03461238.1928.10416862.
URL http://dx.doi.org/10.1080/03461238.1928.10416862

[31] N. V. Smirnov, Theory of Probability and Mathematical Statistics (selected works), Nauka, Moscow, 1970.

[32] T. W. Anderson, On the Distribution of the Two-Sample Cramér-von Mises Criterion, The Annals of Mathematical Statistics 33 (3) (1962) 1148–1159.

[33] R. A. Kronmal, M. E. Tarter, The estimation of probability densities and cumulatives by Fourier series sethods, Journal of the American Statistical Association 63 (323) (1968) 925–952. arXiv:http://www.tandfonline.com/doi/pdf/10.1080/01621459.1968.11009321, doi:10.1080/01621459.1968.11009321.
URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1968.11009321

[34] P. S. Hagan, D. Kumar, A. S. Lesniewski, D. E. Woodward, Managing smile risk, Wilmott Magazine (2002) 84–108.

[35] A. Leitao, L. A. Grzelak, C. W. Oosterlee, On a one time-step Monte Carlo simulation approach of the SABR model: Application to European options, Applied Mathematics and Computation 293 (2017) 461 – 479. doi:http://dx.doi.org/10.1016/j.amc.2016.08.030.

URL http://www.sciencedirect.com/science/article/pii/S0096300316305252

[36] A. Leitao, L. A. Grzelak, C. W. Oosterlee, On an efficient multiple time-step Monte Carlo simulation of the SABR model, To appear in Quantitative Finance.
URL AvailableatSSRN:http://ssrn.com/abstract=2764908

[37] O. Islah, Solving SABR in exact form and unifying it with LIBOR market model, available at SSRN (2009).
URL http://ssrn.com/abstract=1489428

[38] B. Chen, C. W. Oosterlee, H. van der Weide, A low-bias simulation scheme for the SABR stochastic volatility model, International Journal of Theoretical and Applied Finance 15 (2) (2012) 1250016–1 – 1250016–37.

[39] M. J. Ruijter, M. Versteegh, C. W. Oosterlee, On the application of spectral filters in a Fourier option pricing technique, Journal of Computational Finance 19 (1) (2014) 75–106.

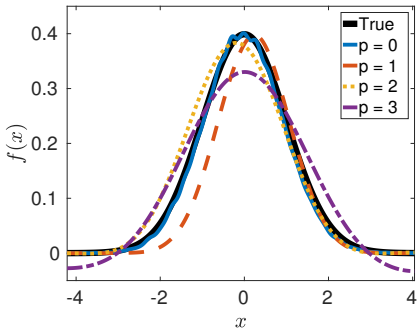
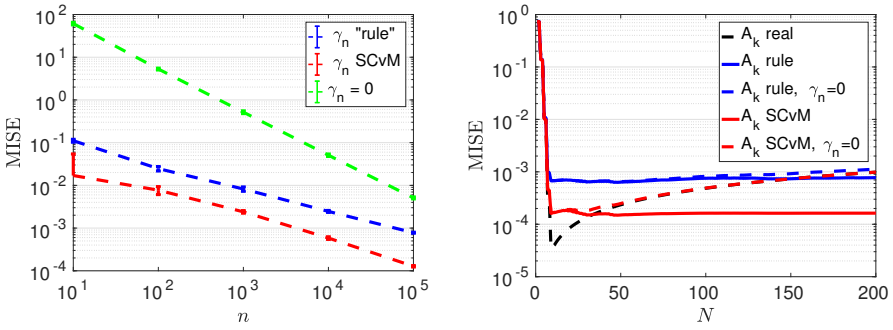


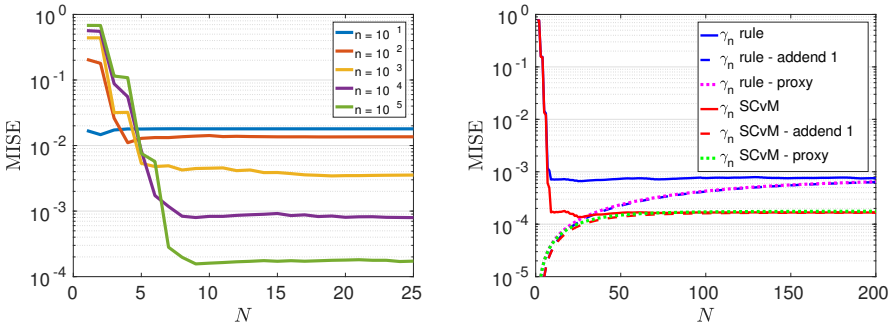
Figure 1: Smoothing of the density approximation in relation to parameter p.



(a) Convergence in terms of n.

(b) Accuracy in terms of N.

Figure 2: Influence of γ_n on the convergence w.r.t. the number of samples n and the number of terms N.



(a) MISE for several values of n.

(b) Comparison between MISE and MISE_N.

Figure 3: Influence of the number of samples n, and the number of coefficients N in the MISE and in the MISE proxy, MISE_N.

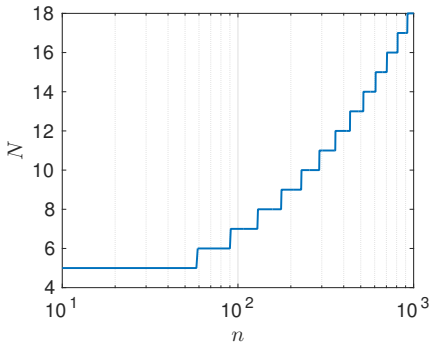
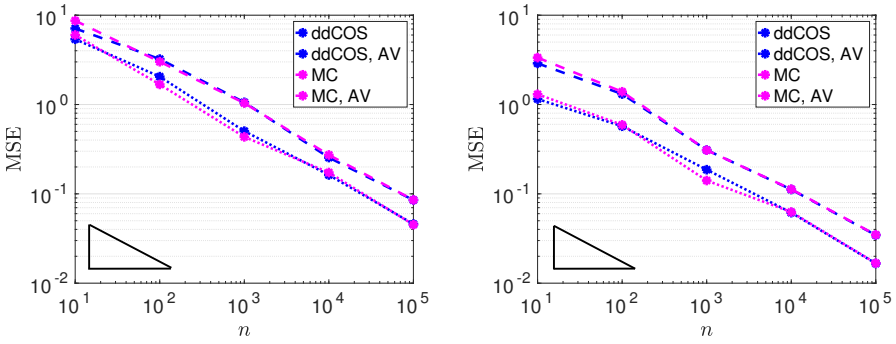


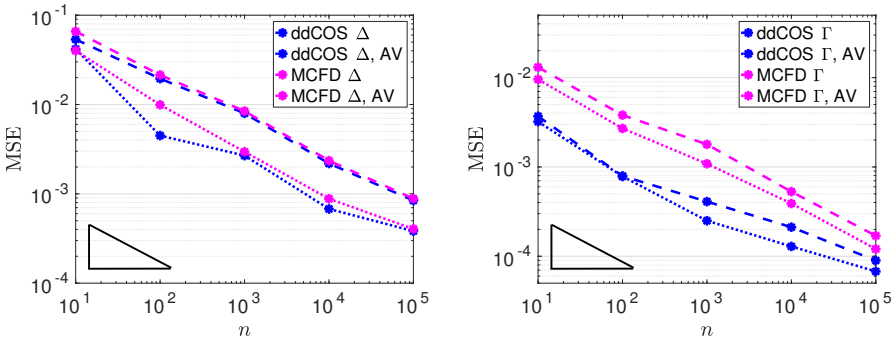
Figure 4: Optimal N-values in terms of n, computed by Algorithm 1.



(a) Call: Strike $K = 100$.

(b) Put: Strike $K = 100$.

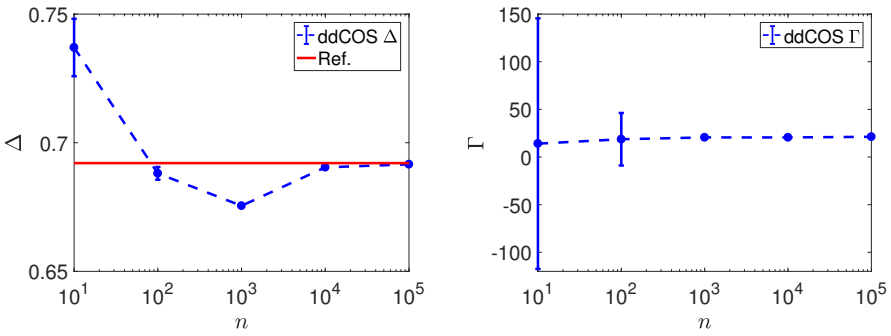
Figure 5: Convergence in prices of the ddCOS method: Antithetic Variates (AV); GBM, $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$.



(a) Δ (Call): Strike $K = 100$.

(b) Γ: Strike $K = 100$.

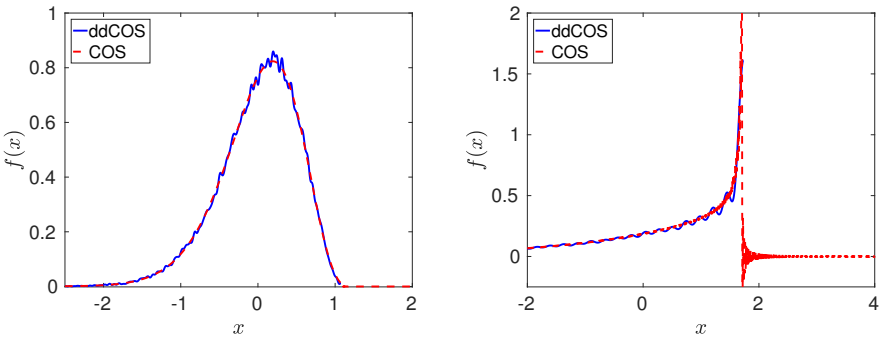
Figure 6: Convergence in Greeks of the ddCOS method: Antithetic Variates (AV); GBM, $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$.



(a) Δ: Strike $K = 0.04$.

(b) Γ: Strike $K = 0.04$.

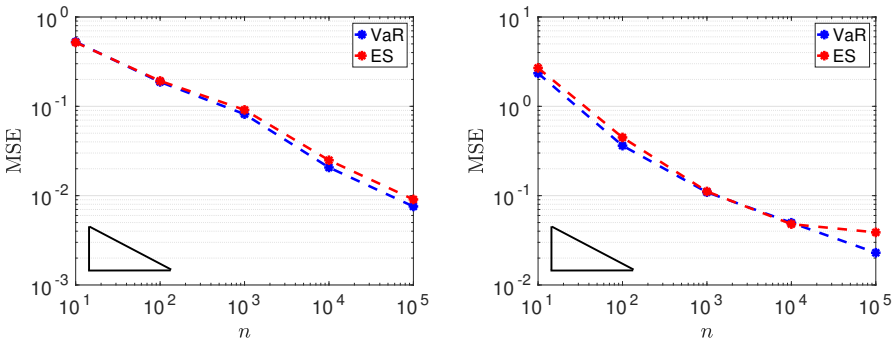
Figure 7: The ddCOS method: Greeks convergence test.



(a) Density Portfolio 1.

(b) Density Portfolio 2.

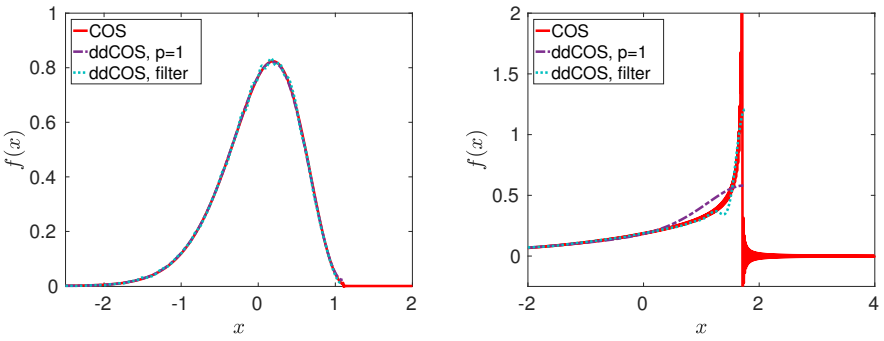
Figure 8: Recovered densities of L: ddCOS vs. COS.



(a) Portfolio 1: $q = 99\%$.

(b) Portfolio 2: $q = 90\%$.

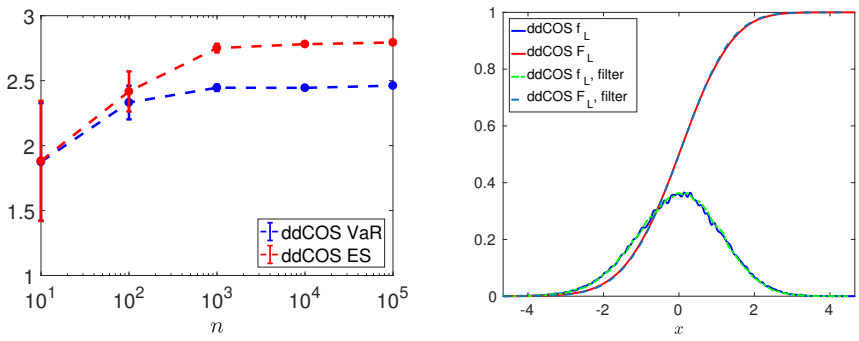
Figure 9: VaR and ES convergence in n.



(a) Density Portfolio 1.

(b) Density Portfolio 2.

Figure 10: Smoothed densities of L.



(a) VaR and ES: $q = 99\%$.

(b) F_L and f_L.

Figure 11: Delta-Gamma approach under the SABR model. Setting: $S(0) = 100$, $K = 100$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$, $T = 2$, $q = 99\%$ and $\Delta t = 1/365$.

K (% of $S(0)$)	80%	90%	100%	110%	120%
0.1			Δ		
Ref.	0.8868	0.8243	0.7529	0.6768	0.6002
ddCOS	0.8867	0.8240	0.7528	0.6769	0.6002
RE	1.1012×10^{-4}				
MCFD	0.8876	0.8247	0.7534	0.6773	0.6006
RE	7.5168×10^{-4}				
			Γ		
Ref.	0.0045	0.0061	0.0074	0.0085	0.0091
ddCOS	0.0045	0.0062	0.0075	0.0084	0.0090
RE	8.5423×10^{-3}				
MCFD	0.0045	0.0059	0.0071	0.0079	0.0083
RE	4.9554×10^{-2}				

Table 1: GBM call option Greeks: $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$.

K (% of $S(0)$)	80%	90%	100%	110%	120%
			Δ		
Ref.	0.8385	0.8114	0.7847	0.7584	0.7328
ddCOS	0.8383	0.8113	0.7846	0.7585	0.7333
RE	2.7155×10^{-4}				
MCFD	0.8387	0.8118	0.7850	0.7586	0.7330
RE	3.1265×10^{-4}				
			Γ		
Ref.	0.0022	0.0024	0.0027	0.0029	0.0030
ddCOS	0.0022	0.0024	0.0027	0.0029	0.0030
RE	8.2711×10^{-3}				
MCFD	0.0023	0.0026	0.0028	0.0031	0.0033
RE	6.118×10^{-2}				

Table 2: Merton jump-diffusion call option Greeks: $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$, $\mu_j = -0.2$, $\sigma_j = 0.2$ and $\lambda = 8$ and $T = 2$.

K (% of $S(0)$)	80%	90%	100%	110%	120%
	Δ				
Ref.	0.9914	0.9284	0.5371	0.0720	0.0058
ddCOS	0.9916	0.9282	0.5363	0.0732	0.0058
RE	5.2775×10^{-3}				
MCFD	0.9911	0.9279	0.5368	0.0737	0.0058
RE	5.5039×10^{-3}				

Table 3: Call option Greek Δ under the SABR model: $S(0) = 100$, $r = 0$, $\sigma_0 = 0.3$, $\alpha = 0.4$, $\beta = 0.6$, $\rho = -0.25$ and $T = 2$.

K (% of $S(0)$)	80%	90%	100%	110%	120%
	Δ				
Ref.	0.8384	0.7728	0.6931	0.6027	0.5086
ddCOS	0.8364	0.7703	0.6902	0.6006	0.5084
RE	2.7855×10^{-3}				
Hagan	0.8577	0.7955	0.7170	0.6249	0.5265
RE	3.1751×10^{-2}				

Table 4: Δ under SABR model. Setting: Call, $S(0) = 0.04$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$ and $T = 2$.

q	10%	30%	50%	70%	90%
VaR	-1.4742	-0.5917	-0.0022	0.5789	1.3862
ES	0.1972	0.5345	0.8644	1.2517	1.8744

Table 5: VaR and ES under SABR model. Setting: $S(0) = 100$, $K = 100$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$, $T = 2$, and $\Delta t = 1/365$.