

Precise option pricing by the COS method—How to choose the truncation range

Gero Junike*, Konstantin Pankrashkin

Carl von Ossietzky Universität, Institut für Mathematik, Oldenburg 26111, Germany

ARTICLE INFO

Article history:

Received 24 August 2021

Revised 20 November 2021

Accepted 6 January 2022

Available online 23 January 2022

2010 MSC:

65T40

42A10

60E10

91G20

Keywords:

COS method

Cosine expansion

Option pricing

Truncation range

Markov's inequality

ABSTRACT

The Fourier cosine expansion (COS) method is used for pricing European options numerically very fast. To apply the COS method, a truncation range for the density of the log-returns need to be provided. Using Markov's inequality, we derive a new formula to obtain the truncation range and prove that the range is large enough to ensure convergence of the COS method within a predefined error tolerance. We also show by several examples that the classical approach to determine the truncation range by cumulants may lead to serious mispricing. Usually, the computational time of the COS method is of similar magnitude in both cases.

© 2022 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY license
(<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

In mathematical finance the logarithmic price of a stock is usually modeled by a random variable X . In many financial models the probability density function f of X exists but its precise structure is unknown. On the other hand the characteristic function φ of X (that is, the Fourier transform of f) is often given explicitly, e.g. see models discussed in Fang and Oosterlee [8]. In this setting, it is necessary to compute integrals of the form

$$\int_{\mathbb{R}} v(x) f(x) dx \quad (1)$$

numerically as fast as possible, where v is a function describing an insurance contract on a stock, like a *call* or *put option*, which for example protects the holder against a fall in the price of the stock. The integral is interpreted as the price of such an insurance contract. How can we compute the integral without knowing f ? A straightforward, efficient and robust method to retrieve the density of a random variable from its characteristic function and to compute prices of call or put options is the *COS method* proposed by Fang and Oosterlee in their seminal work [8].

It is of utmost importance to price call and put options very fast because stock price models are typically calibrated to given prices of liquid call and put options by minimizing the mean-square-error between model prices and given market

* Corresponding author.

E-mail addresses: gerojunike@uol.de (G. Junike), konstantin.pankrashkin@uol.de (K. Pankrashkin).

prices. During the optimization routine, model prices of call and put options need to be evaluated very often for different model parameters.

Under suitable assumptions, the COS method exhibits exponential convergence and compares favorably to other Fourier-based pricing techniques, see Fang and Oosterlee [8]. The COS method is widely applied in mathematical finance, see for instance [2,9,10,12,13,26,28], see Leitao et al. [16], Liu et al. [17], [18] for an application of the COS method in a data-driven approach.

The main idea of the COS method is to approximate the density f with infinite support on a finite range $[a, b]$. The truncated density is then approximated by a (finite) cosine expansion. The integral (1) can then be approximated highly efficiently if v describes the payoff of a put or call option and the characteristic function of f is given in closed-form.

However, it is an open question how to choose the range $[a, b]$ in practice. In this article, we aim to give an answer. More precisely, given some tolerance $\varepsilon > 0$, we derive the minimal length of the range $[a, b]$ such that the absolute difference of the approximation by the COS method and the integral in (1) is less than the tolerance.

Fang and Oosterlee [8] in Eq. (49), see also Eq. (6.44) in Oosterlee and Grzelak [22], proposed some rule of thumb based on cumulants to get an idea how to choose $[a, b]$. In particular, they suggested

$$[a, b] = \begin{cases} [c_1 \pm 12\sqrt{c_2}] & , n_c = 2 \\ [c_1 \pm 10\sqrt{c_2 + \sqrt{c_4}}] & , n_c = 4 \\ [c_1 \pm 10\sqrt{c_2 + \sqrt{c_4 + \sqrt{c_6}}}] & , n_c = 6, \end{cases} \quad (2)$$

where c_1, c_2, c_4, c_6 are the first, second, fourth and sixth cumulants of X . The parameter n_c may be chosen by the user. If not stated otherwise, we use $n_c = 4$. In Section 4.3, we provide several examples where the COS method leads to serious mispricing if the truncation range is based on Eq. (2). Note that there are also Fourier pricing techniques based on wavelets, see Ortiz-Gracia and Oosterlee [23], [24], which do not rely on an a-priori truncation range.

This article is structured as follows: after reviewing the COS method in Section 2, we provide a new proof of convergence of the COS method using elementary tools of Fourier analysis in Section 3. Using Markov's inequality, the proof allows us to derive a minimal length for the range $[a, b]$ given a pre-defined error tolerance. Numerical experiments and applications to model calibration can be found in Section 4. Section 5 concludes.

2. Review of the COS method

Let f be a probability density. The characteristic function φ of f is defined by

$$\varphi(u) = \int_{\mathbb{R}} f(x) e^{iux} dx. \quad (3)$$

Throughout the article, we assume f to be centered around zero, that is $\int_{\mathbb{R}} xf(x)dx = 0$, and we make no additional hypothesis on the support of f . This assumption is mainly made to keep the notation simple. We summarize the approximation by the COS method proposed by Fang and Oosterlee [8], see also [22] for detailed explanations. For $L > 0$ let $f_L = f1_{[-L,L]}$. We define the following basis functions

$$e_k^L(x) := 1_{[-L,L]}(x) \cos\left(k\pi \frac{x+L}{2L}\right), \quad k = 0, 1, 2, \dots$$

and the classical cosine-coefficients of f_L

$$a_k^L := \frac{1}{L} \int_{-L}^L f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx, \quad k = 0, 1, 2, \dots$$

Let $N \in \mathbb{N}$. The density f is approximated in three steps:

$$f(x) \approx f_L(x) \approx \sum_{k=0}^N 'a_k^L e_k^L(x) \approx \sum_{k=0}^N 'c_k^L e_k^L(x), \quad (4)$$

where $\sum'$ indicates that the first summand (with $k = 0$) is weighted by one-half, and we approximate the classical cosine-coefficients a_k^L by an integral over the whole real line

$$\begin{aligned} a_k^L &\approx \frac{1}{L} \int_{\mathbb{R}} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \\ &= \frac{1}{L} \operatorname{Re} \left\{ \varphi\left(\frac{k\pi}{2L}\right) e^{i\frac{k\pi L}{2}} \right\} \\ &=: c_k^L, \quad k = 0, 1, 2, \dots \end{aligned}$$

The coefficients c_k^L can be computed directly if the characteristic function of f is given in closed-form.

Let $v: \mathbb{R} \rightarrow \mathbb{R}$ be a (at least locally integrable) function. For $0 < M \leq L$ denote

$$v_k^M := \int_{-M}^M v(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx, \quad k = 0, 1, 2, \dots \quad (5)$$

Then [8] observed that for M large enough and replacing f by its approximation (4) it holds

$$\int_{\mathbb{R}} f(x) v(x) dx \approx \int_{-M}^M \sum_{k=0}^N c_k^L e_k^L(x) v(x) dx = \sum_{k=0}^N c_k^L v_k^M$$

and called the approximation of the integral *COS method*.

Thus the density f is approximated by a sum of cosine functions making use of the characteristic function to evaluate the cosine coefficients analytically. Working with logarithmic prices, call and put options can be described by truncated exponential functions and the cosine coefficients v_k^M of call and put options can be obtained in explicit form as well. Therefore, option prices can be computed numerically highly efficiently.

3. A new framework for the COS method

In this section, we revisit the convergence of the COS method. The proof allows us to derive a minimal length for the finite range $[-L, L]$, given some error tolerance between the integral and its approximation by the COS method.

Let $\mathcal{L}^1$ and $\mathcal{L}^2$ denote the sets of integrable and square integrable real-valued functions on $\mathbb{R}$, and by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|_2$ we denote the scalar product and the norm on $\mathcal{L}^2$. We denote by $x \vee y$ the maximum value of two real numbers x, y .

Definition 1. A function $f \in \mathcal{L}^1$ is called *COS-admissible*, if

$$B(L) := \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{\mathbb{R} \setminus [-L, L]} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 \rightarrow 0 \text{ as } L \rightarrow \infty.$$

Proposition 2. Assume that $f \in \mathcal{L}^1 \cap \mathcal{L}^2$ with

$$\int_{\mathbb{R}} |xf(x)|^2 dx < \infty,$$

then

$$B(L) \leq \frac{2}{3} \frac{\pi^2}{L^2} \int_{\mathbb{R} \setminus [L, L]} |xf(x)|^2 dx$$

and f is COS-admissible.

Proof. We have $B(L) \leq 4(S_L + \tilde{S}_L)$ with

$$\begin{aligned} S_L &:= \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_L^{\infty} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2, \\ \tilde{S}_L &:= \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{-\infty}^{-L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2, \end{aligned}$$

and it is sufficient to show $\lim_{L \rightarrow \infty} \tilde{S}_L = \lim_{L \rightarrow \infty} S_L = 0$. We will prove it for S_L only, the proof for $\tilde{S}_L$ being almost identical.

Let $j \in \mathbb{N}$. Using the classical cosine expansion of f on the interval $[2jL - L, 2jL + L]$ and Parseval's identity we obtain

$$\begin{aligned} \int_{2jL-L}^{2jL+L} |f(x)|^2 dx &= \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x-(2jL-L)}{2L}\right) dx \right|^2 \\ &\quad \equiv (-1)^{jk} \cos\left(k\pi \frac{x+L}{2L}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2. \end{aligned} \quad (6)$$

Further, using the Cauchy-Schwarz inequality, we estimate

$$\begin{aligned} \left| \int_L^{\infty} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 &= \left| \sum_{j=1}^{\infty} \frac{1}{j} \cdot j \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 \\ &\leq \underbrace{\left(\sum_{j=1}^{\infty} \frac{1}{j^2} \right)}_{=\pi^2/6} \sum_{j=1}^{\infty} j^2 \left| \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2, \end{aligned}$$

then

$$\begin{aligned} S_L &\leq \frac{\pi^2}{6} \sum_{k=0}^{\infty} \frac{1}{L} \sum_{j=1}^{\infty} j^2 \left| \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 \\ &= \frac{\pi^2}{6} \sum_{j=1}^{\infty} j^2 \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{2jL-L}^{2jL+L} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 \\ &\stackrel{(6)}{=} \frac{\pi^2}{6} \sum_{j=1}^{\infty} j^2 \int_{2jL-L}^{2jL+L} |f(x)|^2 dx. \end{aligned}$$

For $x \in [2jL - L, 2jL + L]$ one has $j \leq \frac{x}{L}$, hence,

$$j^2 \int_{2jL-L}^{2jL+L} |f(x)|^2 dx \leq \frac{1}{L^2} \int_{2jL-L}^{2jL+L} |xf(x)|^2 dx,$$

$$S_L \leq \frac{\pi^2}{6} \sum_{j=1}^{\infty} \frac{1}{L^2} \int_{2jL-L}^{2jL+L} |xf(x)|^2 dx = \frac{\pi^2}{6L^2} \int_L^{\infty} |xf(x)|^2 dx.$$

Hence, the assumption (2) implies $\lim_{L \rightarrow \infty} S_L = 0$. $\square$

Corollary 3. Let the density f be bounded, with finite first and second moments, then f is COS-admissible.

Corollary 3 already shows that the class of COS-admissible densities is very large. Next, we provide further sufficient conditions for COS-admissibility. In particular, we show that the densities of the stable distributions for stability parameter $\alpha \in (\frac{1}{2}, 2]$, including the Normal and the Cauchy distributions, and the density of the Pareto distribution are COS-admissible.

Corollary 4. Let $f \in \mathcal{L}^1$ such that its characteristic function φ defined in (3) has a weak derivative φ' satisfying $|\varphi|^2 + |\varphi'|^2 \in \mathcal{L}^1$. Then f is COS-admissible.

Proof. As known from the Fourier analysis of tempered distributions, e.g. Hörmander [14, Chap. VII], the function φ' is the Fourier transform of $x \mapsto ix f(x)$, and due to Plancherel theorem we have

$$\int_{\mathbb{R}} (|f(x)|^2 + |xf(x)|^2) dx = \frac{1}{2\pi} \int_{\mathbb{R}} (|\varphi(u)|^2 + |\varphi'(u)|^2) du < \infty.$$

Hence, the assumptions of Proposition 2 are satisfied. $\square$

Example 5. The densities of the stable distributions whose characteristic functions are of the form

$$\varphi(u) = \exp[i\mu u - |cu|^\alpha (1 - i\beta\Phi_\alpha(u) \operatorname{sgn} u)],$$

$$\Phi_\alpha(u) = \begin{cases} \tan \frac{\pi\alpha}{2}, & \alpha \neq 1, \\ -\frac{2}{\pi} \log |cu|, & \alpha = 1, \end{cases}$$

for parameters $\alpha \in (\frac{1}{2}, 2]$, $\beta \in [-1, 1]$, $c > 0$ and $\mu \in \mathbb{R}$ are COS-admissible by Corollary 4. Recall that this class includes the Normal and the Cauchy distributions.

Example 6. The density of the Pareto distribution with scale $\beta > 0$ and shape $\alpha > 0$ can be described by $f(x) = \alpha\beta^\alpha x^{-(\alpha+1)}$, for $x \geq \beta$, and is COS-admissible. To see this, let $B(L)$ as in Definition 1. It holds by integration by parts and using $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$ and $\sin(k\pi) = 0$, $k = 1, 2, \dots$, that

$$B(L) = \frac{1}{L} \left| \int_L^{\infty} f(x) dx \right|^2$$

$$+ \sum_{k=1}^{\infty} \frac{1}{L} \left| -\frac{2L}{k\pi} \int_L^{\infty} f'(x) \sin\left(k\pi \frac{x+L}{2L}\right) dx \right|^2$$

$$\leq \frac{\beta^{2\alpha}}{L^{2\alpha+1}} + \frac{4\alpha^2 \beta^{2\alpha}}{\pi^2 L^{2\alpha+1}} \sum_{k=1}^{\infty} \frac{1}{k^2} \rightarrow 0, \quad L \rightarrow \infty.$$

Now we discuss the use of COS-admissible functions for the approximation of some integrals arising in mathematical finance. The next theorem shows that in particular a density with infinite support can be approximated by a cosine expansion.

Theorem 7. Assume $f \in \mathcal{L}^1 \cap \mathcal{L}^2$ to be COS-admissible, then

$$\lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \left\| f - \sum_{k=0}^N c_k^L e_k^L \right\|_2 = 0.$$

Proof. Let f_L , a_k^L , e_k^L , c_k^L , v_k^M as in Section 2. Let $L > 0$. Recall that $\langle e_0^L, e_0^L \rangle = 2L$ and $\langle e_k^L, e_l^L \rangle = L\delta_{k,l}$ for $(k, l) \neq (0, 0)$. Consider the cosine coefficients of the tails of f , defined by

$$\tilde{c}_k^L = \frac{1}{L} \int_{\mathbb{R} \setminus [-L, L]} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx.$$

Then it holds $c_k^L = a_k^L + \tilde{c}_k^L$, and it follows

$$\left\| f - \sum_{k=0}^N c_k^L e_k^L \right\|_2 \leq \underbrace{\left\| f - f_L \right\|_2}_{=: A_1(L)} + \underbrace{\left\| f_L - \sum_{k=0}^N a_k^L e_k^L \right\|_2}_{=: A_2(L, N)} + \underbrace{\left\| \sum_{k=0}^N \tilde{c}_k^L e_k^L \right\|_2}_{=: A_3(L, N)}.$$

Due to $f \in \mathcal{L}^2$ it holds $\lim_{L \rightarrow \infty} A_1(L) = 0$. For each fixed L one has $\lim_{N \rightarrow \infty} A_2(L, N) = 0$ as f is square integrable on each $[-L, L]$. Further,

$$\begin{aligned} A_3(L, N)^2 &= \left\langle \sum_{k=0}^N \tilde{c}_k^L e_k^L, \sum_{k=0}^N \tilde{c}_k^L e_k^L \right\rangle = L \sum_{k=0}^N |\tilde{c}_k^L|^2 \leq L \sum_{k=0}^{\infty} |\tilde{c}_k^L|^2 \\ &= \sum_{k=0}^{\infty} \frac{1}{L} \left| \int_{\mathbb{R} \setminus [-L, L]} f(x) \cos\left(k\pi \frac{x+L}{2L}\right) dx \right|^2 = B(L), \end{aligned}$$

and $\lim_{L \rightarrow \infty} B(L) = 0$ as f is COS-admissible.

Take any $\varepsilon > 0$ and choose L_0 such that $A_1(L) < \frac{\varepsilon}{3}$ and $B(L) < \left(\frac{\varepsilon}{3}\right)^2$ for all $L > L_0$, then $A_3(L, N) < \frac{\varepsilon}{3}$ for all $L > L_0$ and all N . For any $L > L_0$, choose N_L sufficiently large to have $A_2(L, N) < \frac{\varepsilon}{3}$ for all $N > N_L$, then

$$\left\| f - \sum_{k=0}^N c_k^L e_k^L \right\|_2 < \varepsilon \text{ for all } L > L_0 \text{ and all } N > N_L,$$

which finishes the proof. $\square$

The next corollary shows that the integral (1) can be computed very efficiently if the characteristic function of f and the classical cosine coefficients of v are available in analytical form: this justifies the use of the COS method under some mild technical assumptions on the density f and the function v . The most important example in the applications for v is the call option $x \mapsto \max(e^x - K, 0)$ and the put option $x \mapsto \max(K - e^x, 0)$ for some $K \geq 0$. The coefficients (5) can be obtained analytically for call and put options.

Corollary 8. Let $f \in \mathcal{L}^1 \cap \mathcal{L}^2$ be COS-admissible and $v : \mathbb{R} \rightarrow \mathbb{R}$ be locally in $\mathcal{L}^2$, that is,

$$v_M := v \mathbf{1}_{[-M, M]} \in \mathcal{L}^2 \text{ for any } M > 0.$$

Assume that $vf \in \mathcal{L}^1$, then the integral of the product of f and v can be approximated by a finite sum as follows.

Let $\varepsilon > 0$ and let $M > 0$ and $\xi > 0$ be such that

$$\int_{\mathbb{R} \setminus [-M, M]} |v(x)f(x)| dx \leq \frac{\varepsilon}{2}, \quad \|v_M\|_2 \leq \xi. \quad (7)$$

Let $L \geq M$ such that

$$\|f - f_L\|_2 \leq \frac{\varepsilon}{6\xi} \quad (8)$$

and

$$B(L) \leq \left(\frac{\varepsilon}{6\xi}\right)^2. \quad (9)$$

Choose N_L large enough so that

$$\left\| f_L - \sum_{k=0}^N a_k^L e_k^L \right\|_2 \leq \frac{\varepsilon}{6\xi}, \quad N \geq N_L.$$

Then it holds for all $N \geq N_L$

$$\left| \int_{\mathbb{R}} v(x)f(x) dx - \sum_{k=0}^N c_k^L v_k^M \right| \leq \varepsilon. \quad (10)$$

Proof. Let $A_1(L)$ and $A_2(L, N)$ as in the proof of Theorem 7. Due to $v_k^M = \langle v_M, e_k^L \rangle$, for all $N \geq N_L$ and by Theorem 7 one obtains

$$\begin{aligned} & \left| \int_{\mathbb{R}} v(x)f(x) dx - \sum_{k=0}^N c_k^L v_k^M \right| \\ &= \left| \int_{\mathbb{R} \setminus [-M, M]} v(x)f(x) dx + \langle v_M, f \rangle - \sum_{k=0}^N c_k^L \langle v_M, e_k^L \rangle \right| \\ &\leq \left| \int_{\mathbb{R} \setminus [-M, M]} v(x)f(x) dx \right| + \left| \langle v_M, f - \sum_{k=0}^N c_k^L e_k^L \rangle \right| \\ &\leq \int_{\mathbb{R} \setminus [-M, M]} |v(x)f(x)| dx + \|v_M\|_2 \left\| f - \sum_{k=0}^N c_k^L e_k^L \right\|_2 \\ &< \frac{\varepsilon}{2} + \xi \left(A_1(L) + A_2(L, N) + \sqrt{B(L)} \right) \\ &\leq \frac{\varepsilon}{2} + \xi \left(\frac{\varepsilon}{6\xi} + \frac{\varepsilon}{6\xi} + \frac{\varepsilon}{6\xi} \right) = \varepsilon. \end{aligned}$$

$\square$

In the next corollary, we apply [Corollary 8](#) to bounded functions v and, given some $\varepsilon > 0$, obtain explicit formulae for M and L such that the Inequalities (7), (8) and (9) hold. For example, put options are bounded. To find a lower bound for M , we need to estimate the tail sum of f . We apply Markov's inequality to estimate the tail sum using the n th-moment of f .

If f has semi-heavy tails, i.e. the tails of f decay exponentially, all moments of f exist and can be obtained by differentiating the characteristic function n times. In a financial context, log-returns are often modeled by semi-heavy tails or lighter distributions: for instance the distribution of log-returns in the Heston model is between the exponential and the Gaussian distribution, see Drăgulescu and Yakovenko [7]. See Schoutens [27] for an overview of Lévy models with semi-heavy tails. In particular, the Lévy process CGMY or the generalized hyperbolic processes, see [1], have semi-heavy tails.

We can find a suitable value of L with the help of the bound for $B(L)$ in [Proposition 2](#). For some Lévy models, the density f is given in closed-form, see Schoutens [27], and $B(L)$ can be estimated directly. In the next corollary, we propose to approximate the tails of f by a density λ , which is given in closed-form. We suggest to use the Laplace density, which decays exponentially just like a density with semi-heavy tails. The (central) Laplace density has only one free parameter describing the variance. One can use a moment-matching method to calibrate the Laplace density λ , setting the variances corresponding to f and λ equal.

The Laplace density with mean zero and variance σ^2 ($\sigma > 0$) is given by

$$\lambda_\sigma(x) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}\frac{|x|}{\sigma}}, \quad x \in \mathbb{R}. \quad (11)$$

Corollary 9 (COS method, Markov range). *Let both the density f and the function v be bounded, with $|v(x)| \leq K$ for all $x \in \mathbb{R}$ and some $K > 0$. For some even natural number $n \geq 2$ assume the n th-moment of f exists and denote it by μ_n , i.e.*

$$\mu_n = \int_{\mathbb{R}} x^n f(x) dx < \infty. \quad (12)$$

Let $\varepsilon > 0$ and

$$M = \sqrt[n]{\frac{2K\mu_n}{\varepsilon}}. \quad (13)$$

Assume there is some $\sigma > 0$ such that

$$f(x) \leq \lambda_\sigma(x), \quad |x| \geq M. \quad (14)$$

Set

$$L = M \vee \left(-\frac{\sigma}{2\sqrt{2}} \log \left(\frac{\sqrt{2}\sigma\varepsilon^2}{72MK^2} \frac{12}{\pi^2} \left(\frac{\sigma^2}{M^2} + \frac{2\sqrt{2}\sigma}{M} + 4 \right)^{-1} \right) \right) \\ \vee -\frac{\sigma}{2\sqrt{2}} \log \left(\frac{2\sqrt{2}\sigma\varepsilon^2}{72MK^2} \right). \quad (15)$$

Choose N_L large enough. Then it holds for all $N \geq N_L$, that

$$\left| \int_{\mathbb{R}} v(x)f(x)dx - \sum_{k=0}^N c_k^L v_k^M \right| \leq \varepsilon. \quad (16)$$

Proof. It holds by Markov's inequality and the definition of M

$$\int_{\mathbb{R} \setminus [-M, M]} |v(x)f(x)| dx \leq K \int_{\mathbb{R} \setminus [-M, M]} f(x) dx \leq K \frac{\mu_n}{M^n} \leq \frac{\varepsilon}{2}.$$

Hence [Eq. \(7\)](#) is satisfied. Next we show Inequality (8) holds. Let $\xi = \|v_M\|_2$ as in [Corollary 8](#). As v is bounded by K it holds $\xi^2 \leq 2MK^2$. Using the upper bound (14) and the expressions (11) it follows

$$\|f - f_L\|_2^2 = \int_{\mathbb{R} \setminus [-L, L]} f^2(x) dx \\ \leq \frac{1}{2\sigma^2} \int_{\mathbb{R} \setminus [-L, L]} e^{-2\sqrt{2}\frac{|x|}{\sigma}} dx \\ = \frac{1}{2\sqrt{2}\sigma} e^{-2\sqrt{2}\frac{L}{\sigma}} \leq \frac{\varepsilon^2}{72MK^2} \leq \left(\frac{\varepsilon}{6\xi} \right)^2.$$

The second last inequality follows by the definition of L and the last inequality holds as $\xi^2 \leq 2MK^2$. Hence Inequality (8) is satisfied. To show that also Inequality (9) is respected, we use [Proposition 2](#). It holds

$$B(L) \leq \frac{2}{3} \frac{\pi^2}{L^2} \int_{\mathbb{R} \setminus [-L, L]} |xf(x)|^2 dx$$

Table 1

Advanced stock price models. Parameters for the first fourteen models are as in Fang and Oosterlee [8]. The parameters of the models M1, M2, M3 and M4 are specified in Section 4.3. ε describes the error tolerance and the columns n_c and n_M , describe the number of cumulants, respectively the number of moments, used to determine the truncation range. The columns L_c , L_M , N_c , N_M , t_c , t_M describe the truncation range, the minimal value of steps N to ensure convergence of the COS method, and the computational time in microseconds, respectively for the cumulants range and the Markov range.

Model	Para-meters	ε	n_c	n_M	L_c	L_M	N_c	N_M	t_c	t_M
Heston	$T = 1$	10^{-7}	4	8	3.4	9.4	220	580	216	387
Heston	$T = 1$	10^{-4}	4	4	3.4	12	120	420	171	318
Heston	$T = 10$	10^{-7}	4	8	11.1	28	130	280	187	319
Heston	$T = 10$	10^{-4}	4	4	11.1	39.6	80	260	154	269
VG	$T = 0.1$	10^{-7}	4	8	0.8	2.3	630	950	242	326
VG	$T = 0.1$	10^{-4}	4	4	0.8	2.9	80	360	89	167
VG	$T = 1$	10^{-7}	4	8	1.9	4.3	100	190	80	103
VG	$T = 1$	10^{-4}	4	4	1.9	6.9	40	150	68	88
CMGY	$Y = 0.5$	10^{-7}	4	8	5.6	12.5	110	230	113	153
CMGY	$Y = 0.5$	10^{-4}	4	4	5.6	21.1	60	220	87	136
CGMY	$Y = 1.5$	10^{-7}	4	8	13.4	32.9	40	100	89	100
CGMY	$Y = 1.5$	10^{-4}	4	4	13.4	62.4	30	130	75	108
CGMY	$Y = 1.98$	10^{-7}	4	8	98	254.6	40	100	81	102
CGMY	$Y = 1.98$	10^{-4}	4	4	98	484.3	30	140	71	108
MJD	M1	10^{-7}	4	8	0.9	4	–	390	–	164
MJD	M1	10^{-7}	6	8	2.8	4	270	390	128	164
MJD	M2	10^{-8}	6	8	5.8	18.2	–	5750	–	1444
CGMY	M3	10^{-7}	4	8	1.5	9	–	1990	–	658
Heston	M4	10^{-2}	2	8	1.3	3.7	–	190	–	211

$$\begin{aligned}
&\leq \frac{4}{3} \frac{\pi^2}{L^2} \int_L^\infty x^2 \lambda_\sigma^2(x) dx \\
&= \frac{\pi^2}{12\sqrt{2}\sigma} \left(\frac{\sigma^2}{L^2} + \frac{2\sqrt{2}\sigma}{L} + 4 \right) e^{-2\sqrt{2}\frac{L}{\sigma}} \\
&\leq \frac{\pi^2}{12\sqrt{2}\sigma} \left(\frac{\sigma^2}{M^2} + \frac{2\sqrt{2}\sigma}{M} + 4 \right) e^{-2\sqrt{2}\frac{L}{\sigma}} \\
&\leq \frac{\varepsilon^2}{72MK^2} \leq \left(\frac{\varepsilon}{6\xi} \right)^2,
\end{aligned}$$

by the definition of L . By Corollary 8, Inequality (16) is obtained. $\square$

Remark 10. The original proof of the convergence of the COS method given in Fang and Oosterlee [8] is somewhat more restrictive compared to the results stated in Corollary 8 because it is assumed that the sum over the cosine coefficients of the payoff function is finite, i.e.,

$$\sum_{k=N}^{\infty} |\nu_k^M| < \infty, \quad (17)$$

see Lemma 4.1 in Fang and Oosterlee [8]. In particular, the cosine coefficients of a put option decay as $\frac{1}{k}$, see Appendix A, and do not satisfy Eq. (17). On the other hand, according to Corollary 9, put options can be priced very well by the COS method.

Remark 11. For small ε , the parameter M is of order $\varepsilon^{-1/n}$, while the other terms in (15) are of order $\log \varepsilon$. Hence, $L = M$ for ε small enough, i.e. the formula for L simplifies considerably. Indeed, for the numerical experiments reported in Table 1 we observed $L = M$.

Remark 12. The smoother f , the smaller N_L may be chosen to obtain a good approximation, see Fang and Oosterlee [8] and references therein.

Remark 13. The COS method has also been applied to price exotic options like Bermudan, American or discretely monitored barrier and Asian options, see Fang and Oosterlee [9, 10], Zhang and Oosterlee [28]. One of the central idea when pricing these exotic options is the approximation of the density of the log returns by a cosine expansion as in Theorem 7. But some care is necessary when applying Corollary 8 to obtain a truncation range, because the truncation range also depends on the payoff function itself.

The COS method has been extended to the multidimensional case, see Ruijter and Oosterlee [26], using a heuristic truncation range similar to Eq. (2). In a future research, we would like to generalize our results, in particular generalize Definition 1, Theorem 7 and Corollary 9, and derive a truncation range in a multidimensional setting.

4. Applications

In this section, we apply Corollary 9 to some stock price models. We use the following setting: let $(\Omega, \mathcal{F}, Q)$ be a probability space. Q is a risk-neutral measure. Let S_T be a positive random variable describing the price of the stock at time $T > 0$. The price of the stock today is denoted by S_0 . We assume there is a bank account paying continuous compounded interest $r \in \mathbb{R}$. Let

$$X := \log(S_T) - E[\log(S_T)] \quad (18)$$

be the centralized log-returns. $E[\log(S_T)]$ is the expectation of the log-returns under the risk-neutral measure and can be obtained from the characteristic function of $\log(S_T)$. Because the density of X is centered around zero, it is justified to approximate the density of X by a symmetric interval $[-L, L]$. In [8] a slightly different centralization of the log-returns has been considered. The characteristic function of X is denoted by φ_X and the density of X is denoted by f_X . In this setting, the price of a put option is given by

$$e^{-rT} E[(K - S_T)^+] = e^{-rT} \int_{\mathbb{R}} v(x) f_X(x) dx, \quad (19)$$

where

$$v(x) = (K - e^{x+E[\log(S_T)]})^+, \quad x \in \mathbb{R}. \quad (20)$$

We approximate the integral at the right-hand-side of Eq. (19) by the COS method. The cosine coefficients v_k^M of v can be obtained in explicit form, see Appendix A or [8]. The price of a call option is given by the put-call parity.

There are two formulae to choose the truncation range $[-L, L]$ for the COS method: the *cumulants range*, based on Eq. (2) and the *Markov range*, based on Corollary 9. To obtain the former, one needs to compute the first, second and forth cumulants. Regarding the second, one need to compute the n th-moment. We will see that $n = 4$, $n = 6$ or $n = 8$ represent a reasonable choice. The n th-moment and the n th-cumulant are similar concepts and can be obtained from the n th-derivative of the characteristic function. In Sections 4.1–4.4 we compare both formulae under the following aspects:

- How does the choice of the truncation range affect the number of terms N ? Certainly the larger the range, the larger we have to set N to obtain a certain precision.
- Are there (important) examples where the COS method fails using the Markov or the cumulants truncation range?
- How can we avoid to evaluate the n th-derivative of the characteristic function to compute the n th-moment or cumulant of the log-returns, each time we apply the COS method, for performance optimization?

In addition, we provide insights which moment to use for the Markov range and we will apply the Markov range to all models discussed in Fang and Oosterlee [8]. All numerical experiments are carried out on a modern laptop (Intel i7-10750H) using the software R and vectorized code without parallelization.

4.1. Black–Scholes and Laplace model

In the Black–Scholes model, see Black and Scholes [3], X is normally distributed. In the Laplace model, see Madan [20], X is Laplace distributed, see Eq. (11) for the density of the Laplace distribution. Both distributions are simple enough such that the quantile functions are given explicitly. There are also closed-form solutions for the prices of call and put options under both models.

Fig. 1 compares M , defined by the n th-moment of X , see Eq. (13), to the quantiles of X . The higher n , the sharper Markov's inequality. At least for the Normal and Laplace distributions, good values for M are obtained for $n \geq 6$. Using higher than the 8th-moment only marginally improves M . If not stated otherwise, we will therefore use the 8th-moment to obtain the Markov truncation range for the COS method.

Consider an at-the-money call option on a stock with price $S_0 = 100$ today and with one year left to maturity. Assume the interest rates are zero. The left panel of Fig. 2 displays L over the error tolerance ε for the Markov and the cumulants range. The figure also shows the minimal value for N such that the absolute difference of the true price of the option and the approximation of the price using the COS method is below the error tolerance ε . The computational time to compute the option price by the COS method using N steps is also indicated.

For a reasonable error tolerance, e.g. $\varepsilon \in (10^{-3}, 10^{-8})$, the minimal value of N to ensure the COS method is close enough to its reference price is at most twice as large using the Markov range instead of the cumulants range.

So using the Markov range based on the 8th-moment instead of the well-established cumulants range, at most doubles the computational time of the COS method for the same level of accuracy. We see a similar pattern for advanced stock price models, see Table 1.

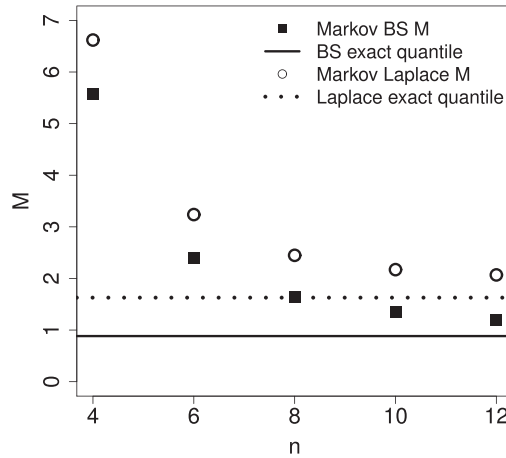


Fig. 1. Moments for Markov range. The variable M to build the truncation range, see Eq. (13), is shown for different moments and $K = 1$ and $\varepsilon = 10^{-5}$ for a Normal (BS) and Laplace distribution with mean zero and standard deviation 0.2, respectively. The solid and dotted lines correspond to the quantiles, i.e. the minimal value for M such that Eq. (7) is satisfied (setting $\nu = 1$ in Eq. (7)).

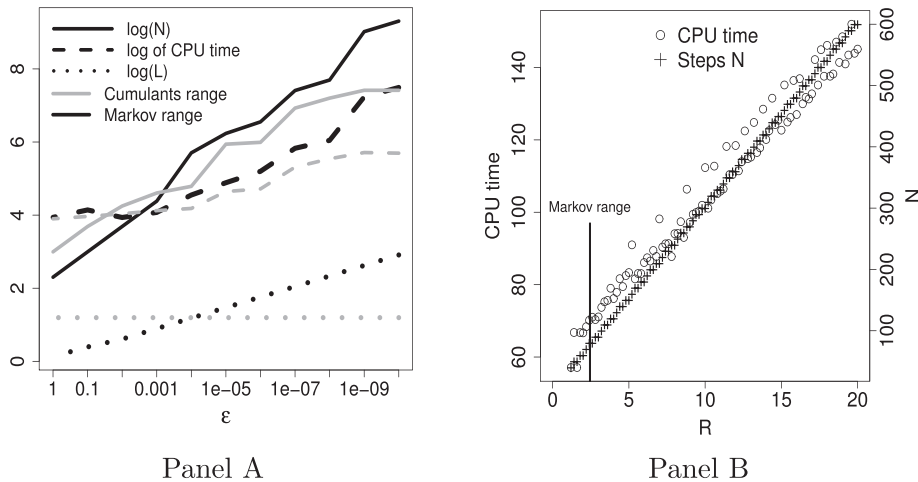


Fig. 2. Convergence of the COS method for a call option in the Laplace model with volatility 0.2. Panel A: The variable L to build the truncation range is shown for the Markov range and the cumulants range over different ε . The minimal number of steps N to obtain convergence of the COS method and the computational time (in microseconds) is illustrated as well. For the Markov range, the 8th-moment is used to obtain L , see Eq. (15). For the cumulants range four cumulants are used, see Eq. (2). Panel B: The minimal number of steps N to obtain convergence of the COS method and the computational time (in microseconds) is illustrated for different truncation ranges $[-R, R]$ for $\varepsilon = 10^{-3}$.

The right panel of Fig. 2 shows the effect of size of the truncation range, which is between one and a hundred times the volatility $\sigma = 0.2$, on the computational time of the COS method. Setting $\varepsilon = 10^{-3}$, we see a linear relationship between the size of the truncation range and the minimal value for N to ensure convergence of the COS method for the Laplace model. The computational time is directly related to N . Setting the truncation range too large, increases the computational time unnecessarily. Thus Corollary 9 may help to save computational time as well. For example, using the Markov range instead of $[-100\sigma, 100\sigma]$, where σ is the volatility of the log-returns, reduces the computational time by more than a factor two.

4.2. Advanced models

Fang and Oosterlee, see Fang and Oosterlee [8], applied the COS method with the cumulants range to three advanced stock price models with different parameters, namely the Heston model, see Heston [11], the Variance Gamma model (VG), see Madan et al. [19], and the CGMY model, see Carr et al. [5]. We repeat the empirical study using the Markov range instead.

Table 1 shows the minimal value of N to ensure the approximation of the price of an option by the COS method is close enough to its reference price, which is taken from Fang and Oosterlee [8], for both truncation ranges. For those models, we conclude that using the Markov range based on the 8th moment and an error tolerance of $\varepsilon = 10^{-7}$ instead of the cumulants range, increases L by about the factor 2.5 and N by about the factor 2.2. (Using the 4th moment for the Markov range and an

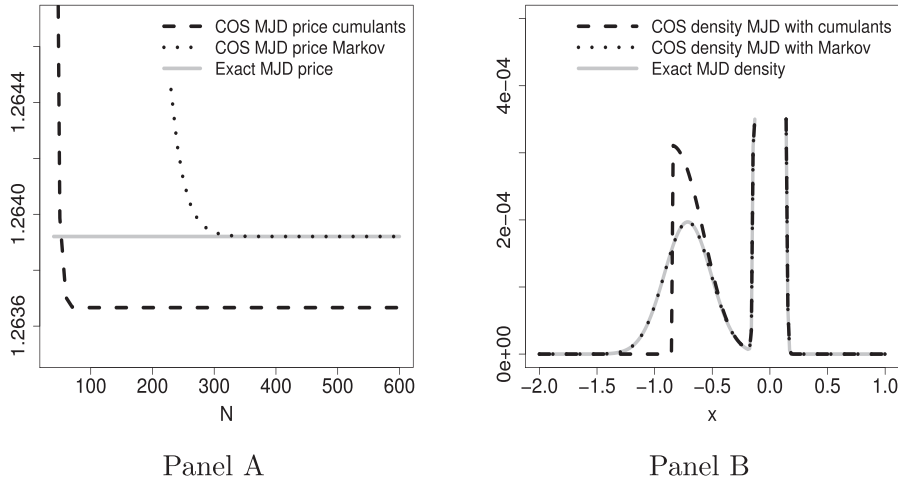


Fig. 3. COS method for MJD model. Panel A: convergence of COS call prices for model M1. The approximation of the call option price by the COS method, where the truncation range is based on four cumulants, is 1.263666 but the reference price, i.e. the exact price, is 1.263921. Panel B: MJD density and COS approximations. The density approximation by the COS method using four cumulants for the truncation range does not change no matter how large we choose N . We set $\varepsilon = 10^{-7}$.

error tolerance of $\varepsilon = 10^{-4}$ increases L and N by about the factor four). The terms N directly determine the computational time of the COS method.

4.3. Examples where COS method diverges

In the Merton jump diffusion model (MJD), see Merton [21], which is a generalization of the Black-Scholes model, the stock price is modeled by a jump-diffusion process: the number of jumps are modeled by a Poisson process with intensity $\eta > 0$, i.e. the expected number of jumps in the time interval $[0, T]$ is ηT . The instantaneous variance of the returns, conditional on no arrivals of jumps, is given by $\sigma^2 > 0$. The jumps are log-normal distributed. The expected percentage jump-size is described by $\kappa \in (-1, \infty)$. The variance of the logarithm of the jumps is described by $\delta^2 > 0$.

For model M1 we choose

$$T = 0.1, \quad \sigma = 0.1, \quad \eta = 0.001, \quad \kappa = -0.5, \quad \delta = 0.2$$

and for model M2, we set

$$T = 0.01, \quad \sigma = 0.1, \quad \eta = 0.00001, \quad \kappa = e^{-6.98} - 1 \approx -0.999, \quad \delta = 0.2.$$

For both models we set $S_0 = 100$ and $r = 0$ and analyze a call option with strike $K = 100$.

Under the Merton jump diffusion model the characteristic function and the density of the log-returns and pricing formulae for put and call option are given in closed-form in terms of an infinite series. We use the first one hundred terms of that series to obtain reference prices for model M1 and M2 and we also apply the Carr-Madan formula, see Carr and Madan [6], to confirm the reference prices.

The left panel of Fig. 3 shows the reference price of a call option under model M1 and the prices using the COS method with the truncation range $[-L, L]$ based on cumulants ($L = 0.85$ using four cumulants) and Markov's inequality ($L = 3.99$ using the 8th-moment). An application of Corollary 9 provides a satisfactory result. However, we clearly see the approximation of the price by the COS method does not converge properly using the cumulants range. The relative error is about two basis points (BPS), a significant difference, independent how large we choose N . The cumulants range is too short and does not fully capture the second mode (the jump) of the MJD density, as shown by the right panel of Fig. 3.

Next we test the cumulants range with six cumulants for the MJD model. The truncation range using six cumulants for model M1 is large enough to ensure the convergence of the COS model, see Table 1. But the cumulants range with six cumulants applied to model M2 is $[-5.8, 5.8]$ and is not large enough to ensure convergence within the required precision ($\varepsilon = 10^{-8}$). Why? If a jump occurs, the expected jump size of the log-returns is equal to

$$\log(\kappa + 1) - \frac{1}{2}\delta^2 = -7,$$

which is not inside the interval $[-5.8, 5.8]$. Hence, again, the second mode of the density, i.e. the jump, is not fully captured by the truncation interval based on six cumulants. We report the prices for model M2 using different numeric approximations

$$\pi_{\text{MJD}} = 0.3989455935507185$$

Table 2

CGMY model. Parameters $C = 0.005$ and $G = M = Y = 1.5$. Choose $\varepsilon = 10^{-7}$. The reference price is 1.02168477497..., which we obtained using an approximating range ten times larger than the Markov range, i.e. $[-89, 89]$, and $N = 10^7$. Increasing the truncation range or N does not change the first 10 decimal digits of the reference price anymore. We set the difference between the reference price and the COS approximation to zero if the first 10 decimal digits coincide.

Terms N	Markov abs. error	Cumul. abs. error	Markov rel. error in BPS	Cumul. rel. error in BPS
1000	4.8×10^{-4}	1.07×10^{-4}	4.74	1.04
2000	8.4×10^{-8}	1.07×10^{-4}	8.2×10^{-4}	1.04
4000	0	1.07×10^{-4}	0	1.04
8000	0	1.07×10^{-4}	0	1.04

$$\pi_{\text{Carr-Madan}} = 0.3989455935506932$$

$$\pi_{\text{Markov}} = 0.3989455935506925$$

$$\pi_{\text{Cumulants}} = 0.3989454898987361$$

The price π_{MJD} is obtained by the closed-form solution of the call price in terms of an infinite series using the first one hundred terms. $\pi_{\text{Carr-Madan}}$ is obtained by the Carr-Madan formula where the damping-factor is set to $\alpha = 0.1$, we use $N = 2^{17}$ points and the truncated Fourier domain is set to $[0, 1200]$. π_{Markov} is obtained by the COS method with $N = 10^6$ terms and using the 8th-moment and $\varepsilon = 10^{-8}$ to obtain the truncation range. $\pi_{\text{Cumulants}}$ is obtained by the COS method also with $N = 10^6$ terms and using six cumulants to obtain the truncation range.

Model M3 is a CGMY model with parameters $C = 0.005$ and $G = M = Y = 1.5$. Consider an at-the-money call option on a stock with price $S_0 = 100$ today and with 0.1 years left to maturity. Assume the interest rates are zero. Using the cumulative range with four cumulants, the relative error of the approximation by the COS method and the reference price is about one basis point, a significant difference, see Table 2, and does not improve when increasing N .

Model M4 is the Heston model with the following parameters: speed of mean-reversion $\kappa = 1$, level of mean-reversion $\eta = 0.05$, vol of vol $\theta = 2$, initial vol $v_0 = 0.01$ and correlation $\rho = -0.75$. Consider an at-the-money call option on a stock with price $S_0 = 100$ today and with 0.5 years left to maturity. Assume the interest rates are zero. Set $\varepsilon = 10^{-2}$.

Using the cumulative range based only on the second cumulant, the truncation range is $[-1.33, 1.33]$ and the price of the option by the COS method is 1.709.

On the other hand, 1.738 is the price of the option based on the Markov range, which is $[-3.71, 3.71]$. We used $N = 1000$. Using a larger N does not change the first three digits of the prices anymore. We also applied the Carr-Madan formula to confirm the price 1.738.

4.4. Application to model calibration

Usually, one proceeds as follows to calibrate a stock price model, like the Heston model, to real market data: given a set of market prices of put and call options, minimize the mean square error between market prices of the options and the corresponding prices predicted by the model. During the optimization phase, model prices need to be computed very often, e.g. by the COS method.

We assume the model is described by m parameters. Let $0 < T_0 \leq T_1$ be the smallest and largest maturity of the put and call options, respectively let $\Theta \subset \mathbb{R}^m$ be the space of feasible parameters of the model.

Let X_T^θ be the centralized log returns for the parameter $\theta \in \Theta$ and maturity $T \in [T_0, T_1]$. To compute the price of a put or call option with strike K by the COS method via Corollary 9, we need to estimate the n th-moment

$$\mu_n^{\theta, T} := E \left[\left(X_T^\theta \right)^n \right],$$

to obtain an estimate for the truncation range $[-L, L]$ of the density of X_T^θ . The n th-moment could directly be determined by differentiating the characteristic function X_T^θ exactly n -times using a computer-algebra-system.

In general, evaluating the n th-derivative of a characteristic function each time the COS method is called, might slow down the total calibration time significantly because the n th-derivative of the characteristic function of some models can be very involved. For fixed $n \in \mathbb{N}$, we therefore let

$$\begin{aligned} h : \Theta \times [T_0, T_1] &\rightarrow [0, \infty) \\ (\theta, T) &\mapsto \mu_n^{\theta, T}. \end{aligned}$$

In the case of the Heston model, the function h is continuous, see Lemma 1 in Ruckdeschel et al. [25]. We propose to identify a function $\hat{h}$ as an approximation of h upfront before the calibration procedure. The evaluation of $\hat{h}$ is expected to be fast. One might for example obtain $\mu_n^{\theta, T}$ for a (large) sample in $\Theta \times [T_0, T_1]$ and defined $\hat{h}$ by a non-linear regression.

Training $\hat{h}$ to the sample takes some time but need to be done only once. The idea is to use $\hat{h}(\theta, T)$ as an approximation of $\mu_n^{\theta, T}$ to obtain the truncation range via Corollary 9. Even if $\hat{h}$ is only a rough estimate of the n th-moment, we expect this

approach to work well because Markov's inequality usually overestimates the tail sum, which provides us with a certain "safety margin".

We illustrate this idea for the Heston model. First, we define Θ for the Heston model, which has five parameters: the speed of mean reversion κ , the mean level of variance η , the volatility of volatility θ , the initial volatility ν_0 and the correlation ρ . We assume

$$\Theta \times [T_0, T_1] = (10^{-3}, 10) \times (10^{-3}, 2)^3 \times (-1, 1) \times \left(\frac{1}{12}, 2\right).$$

We randomly choose 5×10^5 values in $\Theta \times [T_0, T_1]$ and compute $\mu_8^{\theta, T}$ for each of those values by a Monte Carlo simulation. (The 8th-derivative of the characteristic function of the Heston model is extremely involved).

Then we train a small random forest, see Breiman [4], consisting of 50 decisions trees. We choose a random forest for interpolation because the calibration is straightforward. We are confident that other interpolation methods, e.g. based on neural networks, produce similar results.

Next, we define a test set and choose 1000 values in $\Theta \times [T_0, T_1]$ randomly. Junike et al. [15] calibrated the Heston model to a time series of 100 time points in Summer 2017 of real market data of put and call options, including calm and more volatile trading days. We add those 100 parameters to our test set with a maturity of half a year.

For each parameter set, we compute a reference price¹ of three call options with strikes $K \in \{75, 100, 125\}$. We choose $S_0 = 100$, $r = 0$ and $\varepsilon = 10^{-4}$.

We obtain very satisfactory results approximating prices of call options by the COS method if the truncation range is obtained via Corollary 9, where the 8th-moment is estimated by the random forest $\hat{h}$. The maximal absolute error for the market data test set over all options is less than ε for $N \geq 500$. The maximal absolute error for the random test set is less than ε for $N \geq 2000$.

Last we comment on the CPU time: computing one option price by the COS method for $N = 1000$ takes about 700 microseconds using the software R and vectorized code without parallelization. To evaluate $\hat{h}$ on 10^4 parameter sets using R's package *randomForest*, also without parallelization, takes on average 60 microseconds per set.

5. Conclusions

The COS method is used to compute certain integrals appearing in mathematical finance by efficiently retrieving the probability density function of a random variable describing some log-returns from its characteristic function. The main idea is to approximate a density with infinite support on a finite range by a cosine series.

We provided a new framework to prove the convergence of the COS method, which enables us to obtain an explicit formula for the minimal length of the truncation range, given some maximal error tolerance between the integral and its approximation by the COS method.

The formula for the truncation range is based on Markov's inequality and it is assumed that the density of the log-returns has semi-heavy tails. To obtain the truncation range, we need the n th-moment of the (centralized) log-returns. The larger n , the sharper Markov's inequality. From numerical experiments, we concluded that $n = 8$, $n = 6$ or even $n = 4$ is a reasonable choice.

The n th-moment could directly be determined from the n th-derivative of the characteristic function. In the case of the Heston model, the characteristic function is too involved and instead we successfully employed a machine learning approach to estimate the 8th-moment.

Acknowledgments

We thank two anonymous referees for many valuable comments and suggestions enabling us to improve the quality of the paper.

Appendix A. Cosine-coefficients for a put option

Define X as in Eq. (18) to compute the price of a put option with strike K by the COS method using Markov's inequality for the truncation range. The characteristic function of X is given by

$$\varphi_X(u) = \varphi_{\log(S_T)}(u) \exp(-iuE[\log(S_T)]),$$

where the expectation $E[\log(S_T)]$ can be computed from the characteristic function of S_T by

$$E[\log(S_T)] = -i\varphi'_{\log(S_T)}(0).$$

¹ We computed the reference price by the COS method using an approximating range two times larger than the Markov range and $N = 10^6$ terms. We verified the reference price by the Carr–Madan formula, where the damping-factor is set to $\alpha = 0.1$, we use $N = 2^{17}$ points and the truncated Fourier domain is set to $[0, 1200]$.

Compute n th-moment of X with tricks from Section 4.4 or by

$$\mu_n = \frac{1}{i^n} \left. \frac{\partial^n}{\partial u^n} \varphi_X(u) \right|_{u=0}, \quad n \in \{2, 8\}.$$

Choose $\sigma \geq \sqrt{\mu_2}$ and set M and L as in Eqs. (13) and (15). Let

$$d := \min(\log(K) - E[\log(S_T)], M).$$

Compute the cosine-coefficients v_k^M of v , defined in Eq. (20), by

$$\begin{aligned} v_k^M &= \int_{-M}^M (K - e^{x+E[\log(S_T)]})^+ \cos\left(k\pi \frac{x+L}{2L}\right) dx \\ &= K \underbrace{\int_{-M}^d \cos\left(k\pi \frac{x+L}{2L}\right) dx}_{=:\Psi_0} \\ &\quad - e^{E[\log(S_T)]} \underbrace{\int_{-M}^d e^x \cos\left(k\pi \frac{x+L}{2L}\right) dx}_{=:\Psi_1}. \end{aligned}$$

Ψ_0 and Ψ_1 can be computed easily:

$$\Psi_0 = \begin{cases} \frac{2L}{k\pi} \left(\sin\left(k\pi \frac{d+L}{2L}\right) - \sin\left(k\pi \frac{-M+L}{2L}\right) \right), & k > 0 \\ d + M, & k = 0 \end{cases}$$

and

$$\begin{aligned} \Psi_1 &= \left[e^d \left(\frac{k\pi}{2L} \sin\left(k\pi \frac{d+L}{2L}\right) + \cos\left(k\pi \frac{d+L}{2L}\right) \right) \right. \\ &\quad \left. - e^{-M} \left(\frac{k\pi}{2L} \sin\left(k\pi \frac{-M+L}{2L}\right) + \cos\left(k\pi \frac{-M+L}{2L}\right) \right) \right] \frac{1}{1 + \left(\frac{k\pi}{2L}\right)^2}. \end{aligned}$$

Choose N large enough and set

$$c_k^L := \frac{1}{L} \operatorname{Re} \left\{ \varphi_X\left(\frac{k\pi}{2L}\right) e^{i\frac{k\pi}{2}} \right\}, \quad k = 0, 1, \dots, N.$$

Then it holds for the price of a European put option

$$e^{-rT} E[(K - S_T)^+] \approx e^{-rT} \sum_{k=0}^N {}' c_k^L v_k^M,$$

where $\sum'$ indicates that the first term in the summation is weighted by one-half. The price of a call option can be computed using the put-call parity.

References

- [1] J. Albin, M. Sundén, On the asymptotic behaviour of Lévy processes, part I: subexponential and exponential processes, *Stoch. Process. Appl.* 119 (1) (2009) 281–304.
- [2] C. Bardgett, E. Gourier, M. Leippold, Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets, *J. Finance Econ.* 131 (2019) 593–618.
- [3] F. Black, M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.* 81 (3) (1973) 637–654.
- [4] L. Breiman, Random forests, *Mach. Learn.* 45 (1) (2001) 5–32.
- [5] P.P. Carr, H. Geman, D.B. Madan, M. Yor, The fine structure of asset returns: an empirical investigation, *J. Bus.* 75 (2) (2002) 305–332.
- [6] P.P. Carr, D.P. Madan, Option valuation using the fast Fourier transform, *J. Comput. Finance* 2 (4) (1999) 61–73.
- [7] A.A. Drăgulescu, V.M. Yakovenko, Probability distribution of returns in the Heston model with stochastic volatility, *Quant. Finance* 2 (6) (2002) 443–453.
- [8] F. Fang, C.W. Oosterlee, A novel pricing method for European options based on Fourier-cosine series expansions, *SIAM J. Sci. Comput.* 31 (2009) 826–848.
- [9] F. Fang, C.W. Oosterlee, Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions, *Numer. Math.* 114 (2009) 27–62.
- [10] F. Fang, C.W. Oosterlee, A Fourier-based valuation method for bermudan and barrier options under Heston's model, *SIAM J. Financ. Math.* 2 (2011) 439–463.
- [11] S. L. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Rev. Financ. Stud.* 6 (2) (1993) 327–343.
- [12] L.A. Grzelak, C.W. Oosterlee, On the Heston model with stochastic interest rates, *SIAM J. Financ. Math.* 2 (2011) 255–286.
- [13] A. Hirs, Computational Methods in Finance, CRC Press, 2012.
- [14] L. Hörmander, The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis, second ed., Springer-Verlag, 1990.

- [15] G. Junike, W. Schoutens, H. Stier, Performance of advanced stock price models when it becomes exotic: an empirical study, *Ann. Financ.* (2021), doi:10.1007/s10436-021-00396-2.
- [16] A. Leita, C.W. Oosterlee, L. Ortiz-Gracia, S.M. Bohte, On the data-driven COS method, *Appl. Math. Comput.* 317 (2018) 68–84.
- [17] S. Liu, A. Borovykh, L.A. Grzelak, C.W. Oosterlee, A neural network-based framework for financial model calibration, *J. Math. Ind.* 9 (1) (2019) 1–28.
- [18] S. Liu, C.W. Oosterlee, S.M. Bohte, Pricing options and computing implied volatilities using neural networks, *Risks* 7 (1) (2019) 1–22.
- [19] D.B. Madan, P.P. Carr, E.C. Chang, The variance gamma process and option pricing, *Rev. Finance* 2 (1) (1998) 79–105.
- [20] D.B. Madan, Adapted hedging, *Ann. Finance* 12 (3) (2016) 305–334.
- [21] R.C. Merton, Option pricing when underlying stock returns are discontinuous, *J. Financ. Econ.* 3 (1–2) (1976) 125–144.
- [22] C.W. Oosterlee, L.A. Grzelak, *Mathematical Modeling and Computation in Finance: With Exercises and Python and Matlab Computer Codes*, World Scientific, 2020.
- [23] L. Ortiz-Gracia, C.W. Oosterlee, Robust pricing of European options with wavelets and the characteristic function, *SIAM J. Sci. Comput.* 35 (5) (2013) B1055–B1084.
- [24] L. Ortiz-Gracia, C.W. Oosterlee, A highly efficient Shannon wavelet inverse Fourier technique for pricing European options, *SIAM J. Sci. Comput.* 38 (1) (2016) B118–B143.
- [25] P. Ruckdeschel, T. Sayer, A. Szimayer, Pricing American options in the Heston model: a close look at incorporating correlation, *J. Deriv.* 20 (3) (2013) 9–29.
- [26] M.J. Ruijter, C.W. Oosterlee, Two-dimensional Fourier cosine series expansion method for pricing financial options, *SIAM J. Sci. Comput.* 34 (2012) B642–B671.
- [27] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley, 2003.
- [28] B. Zhang, C.W. Oosterlee, Efficient pricing of European-style Asian options under exponential Lévy processes based on Fourier cosine expansions, *SIAM J. Financ. Math.* 4 (2013) 399–426.