On the application of spectral filters in a Fourier option pricing technique

M. J. Ruijter, * M. Versteegh, † C. W. Oosterlee, ‡

May 17, 2013

Abstract

When Fourier techniques are employed to specific option pricing cases from computational finance with non-smooth functions, the so-called Gibbs phenomenon may become apparent. This seriously impacts the efficiency and accuracy of the pricing. For example, the Variance Gamma asset price process gives rise to algebraically decaying Fourier coefficients, resulting in a slowly converging Fourier series. We apply spectral filters to achieve faster convergence. Filtering is carried out in Fourier space; the series coefficients are pre-multiplied by a decreasing filter, which does not add significant computational cost. Tests with different filters show how the algebraic index of convergence is improved.

Keywords: Fourier cosine expansion method, spectral filters, European options, Variance Gamma, portfolio loss distribution, Gibbs phenomenon.

AMS Subject classification: 91G60, 65T40, 94A12, 60E10.

1 Introduction

Fourier techniques have now become well-established in computational finance to efficiently price certain financial instruments, like European options, certain options with early-exercise features and also exotic options, like Asian, multi-asset or barrier options [FO09, RO12, ZO13]. The Fourier techniques belong to the class of numerical integration option pricing methods and they are referred to as "transform methods", because a transformation to the Fourier domain is combined with numerical integration [CM99, FO08, LK07]. Transform methods can readily be used for asset price models for which the characteristic function (ie, the Fourier transform of the probability density function) is available. A specific Fourier pricing technique, which we consider here is the COS method [FO08], is based on Fourier-cosine series expansions. The issues and remedies we address here will however be of relevance for other transform methods as well.

As long as the governing probability density function is sufficiently smooth, an exponential convergence in the number of cosine terms is achieved by the COS method. For certain choices of the asset dynamics, however, the governing probability density function is *not* smooth everywhere. Smoothness issues are encountered, for example, when we model the asset price by the Variance Gamma process [MCC98]. Moreover, in risk management stepwise cumulative distribution functions are typically encountered in portfolio loss modeling with small-sized portfolios, because we

*CWI - Centrum Wiskunde & Informatica, Amsterdam, the Netherlands, email: m.j.ruijter@cwi.nl,

†Delft University of Technology, Delft Institute of Applied Mathematics, Delft, the Netherlands, email: versteegh@gmail.com,

‡CWI - Centrum Wiskunde & Informatica, Amsterdam, the Netherlands, email: c.w.oosterlee@cwi.nl, and Delft University of Technology, Delft Institute of Applied Mathematics, Delft, the Netherlands.

then deal with discrete random variables representing individual obligors that may have a default problem.

When Fourier techniques are employed to specific cases with non-smooth functions, the *Gibbs phenomenon* may become apparent, which seriously impacts the efficiency and accuracy of the valuation. The Gibbs phenomenon reflects the difficulty of approximating a discontinuous function by a finite Fourier series. Although the limit of the partial sums represents the original function exactly, in the finite case there is always an overshoot at a jump discontinuity. The width of this overshoot decreases with the number of Fourier terms, but the height of the maximum does not. In the case of a jump discontinuity we may have pointwise convergence almost everywhere, but no uniform convergence. The Gibbs phenomenon is also related to the principle that the decay of the Fourier coefficients is governed by the smoothness of the function concerned. Functions with a discontinuity in one of the derivatives will have algebraically decaying Fourier coefficients, resulting in a slowly converging Fourier series. The local effect of the Gibbs phenomenon results in oscillations near the jumps, but there also is a global effect: although the error decays away from the jumps, the decay rate is only first order. Thus, the existence of one or more discontinuities drastically reduces the convergence rate over the whole domain, and spectral accuracy is lost [Tad07, GS97].

The research field dealing with Gibbs oscillations is wide and well-established. An excellent overview into the various improvement techniques is given by Tadmor [Tad07]. We will focus on the use of *spectral filters* to deal with the Gibbs phenomenon appearing for non-smooth densities and discrete distribution functions. This is one of the very basic techniques in this field, but we will see that it fits very well to the applications at hand.

Several other techniques have been proposed in the literature to reduce or remove the Gibbs phenomenon. As the optimal filter order for a function with discontinuities is an increasing function of the distance to the nearest discontinuity, the idea of adaptive filtering is to vary the filter order so that it is close to this optimal value. Tadmor [Tad07, Tan06, TT05] describes adaptive filters recovering root-exponential accuracy. This type of filtering is related to superconvergent extraction techniques in finite element methods [vSRV11]. It is however not trivial to efficiently employ an adaptive filter in the context of Fourier option pricing, so we stay with nonadaptive filters here. Mollifiers are a time domain equivalent to filters, in the sense that multiplication by a function in Fourier-space corresponds to a convolution in physical space (which is the basis for mollifiers). Implementation of mollifiers in a Fourier option pricing technique would require reconstruction to the time-domain, which is computationally relatively expensive. The same is true for methods like Gegenbauer polynomial reconstruction, see [Tad07, GS97], Digital Total Variation (DTV) filtering [COS01, Sar06] and for the hybrid methods [Gel00] where a polynomial reconstruction is used only where needed and filtering is used elsewhere.

We start in Section 2 with Fourier series and Fourier-cosine series and explain how we can employ the characteristic function of a random variable to approximate the corresponding density or distribution function. Also the COS method for pricing European options is briefly described. In Section 3 the convergence of the series and improvements by spectral filters are discussed. Extensive numerical experiments are performed in Section 4. Finally, Section 5 concludes.

2 Fourier and Fourier-cosine series expansions

The *Fourier series* of an integrable function $f(y)$ supported on a finite interval $[a, b]$ is defined as ([SS03])

$$f(y) = \sum_{n=-\infty}^{\infty} \hat{f}_n \exp\left(in\pi \frac{2y}{b-a}\right), \tag{2.1}$$

with Fourier coefficients given by

$$\hat{f}_n = \frac{1}{b-a}\int_a^b f(y)\exp\left(-in\pi\frac{2y}{b-a}\right)\,dy. \tag{2.2}$$

The COS method for pricing European options is based on the *Fourier-cosine series* expansion, which is defined by

$$f(y) = \sum_{n=0}^{\infty}{}' \hat{f}_n^c \cos\left(n\pi\frac{y-a}{b-a}\right), \tag{2.3}$$

with Fourier-cosine coefficients given by

$$\hat{f}_n^c = \frac{2}{b-a}\int_a^b f(y)\cos\left(n\pi\frac{y-a}{b-a}\right)\,dy. \tag{2.4}$$

The prime $'$ in the summation indicates division of the first term by two. These cosine series can be seen as a classical Fourier series of a function $f^{ex}(y) = f(|y-a|)$, on an extended interval $[2a-b, b]$, which is mirrored around the midpoint a to make it an even function.

2.1 Recovery density and distribution function

Continuous random variable. In financial option pricing we deal with stochastic asset prices for which the probability density function is usually unknown. The characteristic function is however often known [DPS00] and can be used to approximate the density function, as we explain in this section. Suppose we have a continuous random variable X, with cumulative distribution function $F(.)$ and probability density function $p(.)$. The characteristic function $\varphi(.)$ of X is defined as

$$\varphi(u) := \mathbb{E}\left[e^{iuX}\right]. \tag{2.5}$$

The density and the characteristic function form a Fourier pair, where the characteristic function is the Fourier transform of the density,

$$\varphi(u) = \int_{\mathbb{R}} e^{iuy}\,dF(y) = \int_{\mathbb{R}} e^{iuy}p(y)\,dy, \tag{2.6}$$

and the density is the inverse Fourier transform of the characteristic function:

$$p(y) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iuy}\varphi(u)\,du. \tag{2.7}$$

Its Fourier-cosine series expansion reads

$$p(y) = \sum_{n=0}^{\infty}{}' \hat{p}_n^c \cos\left(n\pi\frac{y-a}{b-a}\right), \tag{2.8}$$

with Fourier-cosine coefficients

$$\hat{p}_n^c = \frac{2}{b-a}\int_a^b p(y)\cos\left(n\pi\frac{y-a}{b-a}\right)\,dy = \frac{2}{b-a}\mathrm{Re}\left\{\int_a^b p(y)\exp\left(in\pi\frac{y-a}{b-a}\right)\,dy\right\}, \tag{2.9}$$

where Re{.} denotes taking the real part. If the density function $p(y)$ decays rapidly to zero for $y \to \pm\infty$, the integration range in Equation (2.6) can be truncated without loss of any significant accuracy:

$$\varphi(u) = \int_{\mathbb{R}} e^{iuy} p(y)\, dy \approx \int_a^b e^{iuy} p(y)\, dy = \varphi_1(u), \tag{2.10}$$

and therefore the characteristic function can be used to efficiently calculate the Fourier-coefficients [FO08]. Combining (2.9) and (2.10) gives us:

$$\hat{p}_n^c = \frac{2}{b-a}\mathrm{Re}\left\{\varphi_1\left(\frac{n\pi}{b-a}\right)e^{in\pi\frac{-a}{b-a}}\right\} \approx \frac{2}{b-a}\mathrm{Re}\left\{\varphi\left(\frac{n\pi}{b-a}\right)e^{in\pi\frac{-a}{b-a}}\right\} := \hat{\varphi}_n^c. \tag{2.11}$$

After truncation of the series summation we end up with the approximation

$$p(y) \approx {\sum_{n=0}^{N}}' \hat{\varphi}_n^c \cos\left(n\pi\frac{y-a}{b-a}\right). \tag{2.12}$$

The distribution function can now be approximated as follows,

$$\begin{aligned} F(y) = \mathbb{P}(X \le y) &\approx \int_a^y p(t)\, dt \approx {\sum_{n=0}^{N}}' \hat{\varphi}_n^c \int_a^y \cos\left(n\pi\frac{t-a}{b-a}\right) dt \\ &= \frac{1}{2}\hat{\varphi}_0^c(y-a) + \sum_{n=1}^{N} \hat{\varphi}_n^c \frac{b-a}{n\pi}\sin\left(n\pi\frac{y-a}{b-a}\right) \\ &= \frac{1}{2}\frac{2}{b-a}(y-a) + \sum_{n=1}^{N}\frac{2}{n\pi}\mathrm{Re}\left\{\varphi\left(\frac{n\pi}{b-a}\right)e^{in\pi\frac{-a}{b-a}}\right\}\sin\left(n\pi\frac{y-a}{b-a}\right). \end{aligned} \tag{2.13}$$

Discrete random variable. If X is a discrete random variable, then a density function does not exist and we use the following Lévy inversion formula connecting the distribution function, $F(.)$, and characteristic function, $\varphi(.)$. For $y-h, y+h \in C(F)$, with $C(F)$ the continuity set of F, there holds ([Gut05])

$$\begin{aligned} \frac{F(y+h)-F(y-h)}{2h} &= \lim_{T\to\infty}\frac{1}{2h}\frac{1}{2\pi}\int_{-T}^{T}\varphi(u)\frac{e^{-iu(y-h)}-e^{-iu(y+h)}}{iu}du \\ &= \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\varphi(u)e^{-iuy}\frac{\sin uh}{uh}du. \end{aligned} \tag{2.14}$$

Suppose F is concentrated on the interval $[a,\infty)$ and $F(a)=0$. Then, we get ($h \ge 0$)

$$F(a+h) = \lim_{T\to\infty}\frac{1}{\pi}\int_{-T}^{T}\varphi(u)e^{-iua}\frac{\sin uh}{u}du = \lim_{T\to\infty}\frac{2}{\pi}\int_0^T \mathrm{Re}\left\{\varphi(u)e^{-iua}\right\}\frac{\sin uh}{u}du. \tag{2.15}$$

Numerical integration with step size Δu results in the approximation ($y \ge a$)

$$\begin{aligned} F(y) &= \lim_{T\to\infty}\frac{2}{\pi}\int_0^T \mathrm{Re}\left\{\varphi(u)e^{-iua}\right\}\frac{\sin u(y-a)}{u}du \\ &\approx \frac{1}{2}\frac{2\Delta u}{\pi}(y-a) + \sum_{n=1}^{N}\frac{2}{n\pi}\mathrm{Re}\left\{\varphi\left(n\Delta u\right)e^{-in\Delta ua}\right\}\sin\left(n\Delta u(y-a)\right). \end{aligned} \tag{2.16}$$

For $\Delta u = \frac{\pi}{b-a}$ this, in fact, corresponds to the formula for the distribution function of a continuous random variable (Equation (2.13)).

2.2 COS method for European options

The COS method, short for Fourier-cosine pricing method, is based on the Fourier-cosine series expansion of a density function to efficiently approximate the expected value of an arbitrary function of random variables. The method has been developed in the first place for pricing financial options, like European [FO08] and Bermudan options [FO09]. The risk-neutral valuation formula for a European option with payoff function $g(.)$ reads

$$v(t_0, x) = e^{-r\Delta t}\mathbb{E}_{\mathbb{Q}}[g(X(T))|X(t_0) = x] = e^{-r\Delta t}\int_{\mathbb{R}} g(y)p(y|x)\,dy. \tag{2.17}$$

Here $X(t)$ is the state process, which can be any monotone function of the underlying asset price $S(t)$. From here it is taken to be the scaled log asset-price, $X(t) := \log(S(t)/K)$, where K is the option's strike price. $\mathbb{E}_{\mathbb{Q}}$ denotes the expectation under risk-neutral measure $\mathbb{Q}$; Δt is the difference between the time of maturity T and the initial time t_0; $p(y|x)$ is the conditional probability density of $X(T)$, given $X(t_0) = x$ and r is the risk-free interest rate.

Explicit expressions for probability density functions $p(y|x)$ encountered in finance are often not known, or involve some mathematical special functions, which make them impractical to calculate. Instead the characteristic function $\varphi(u|x)$ corresponding to $p(y|x)$ is often known [DPS00]. We can replace the density function by its Fourier-cosine series expansion,

$$p(y|x) = \sum_{n=0}^{\infty}{}' \hat{p}_n^c(x)\cos\left(n\pi\frac{y-a}{b-a}\right), \tag{2.18}$$

whose coefficients are approximated by

$$\begin{aligned}\hat{p}_n^c(x) &= \frac{2}{b-a}\int_a^b p(y|x)\cos\left(n\pi\frac{y-a}{b-a}\right)\,dy \\ &\approx \frac{2}{b-a}\mathrm{Re}\left\{\varphi\left(\frac{n\pi}{b-a}\Big|x\right)e^{in\pi\frac{-a}{b-a}}\right\} := \hat{\varphi}_n^c(x).\end{aligned} \tag{2.19}$$

The Fourier-cosine series expansion of the payoff function is given by

$$g(y) = \sum_{n=0}^{\infty}{}' \hat{g}_n^c\cos\left(k\pi\frac{y-a}{b-a}\right), \tag{2.20}$$

with the coefficients,

$$\hat{g}_n^c = \frac{2}{b-a}\int_a^b g(y)\cos\left(n\pi\frac{y-a}{b-a}\right)\,dy. \tag{2.21}$$

For several payoff functions, including plain vanilla puts and calls, the cosine coefficients $\hat{g}_n^c$ are available in closed form [FO08]. Substituting the series in (2.18) into Equation (2.17), replacing the coefficients $\hat{p}_n^c(x)$ by $\hat{\varphi}_n^c(x)$, interchanging summation and integration, and using definition (2.21) we get a formula for $v(t_0, x)$, written as the product of Fourier-cosine coefficients:

$$v(t_0, x) \approx \frac{b-a}{2}e^{-r\Delta t}\sum_{n=0}^{\infty}{}' \hat{\varphi}_n^c(x)\hat{g}_n^c. \tag{2.22}$$

Truncating the series, we obtain the next approximation, the *COS pricing formula*:

$$\hat{v}(t_0, x) := e^{-r\Delta t}\sum_{n=0}^{N}{}' \mathrm{Re}\left\{\varphi\left(\frac{n\pi}{b-a}\Big|x\right)e^{in\pi\frac{-a}{b-a}}\right\}\hat{g}_n^c$$

$$= e^{-r\Delta t} \sum_{n=0}^{N}{}' \operatorname{Re}\left\{ \phi\left(\frac{n\pi}{b-a}\right) e^{in\pi\frac{x-a}{b-a}} \right\} \hat{g}_n^c. \tag{2.23}$$

The last equality holds for processes with independent increments, such as Lévy processes, which include the log-versions of geometric Brownian motion (GBM), Variance Gamma and CGMY models. In that case, the characteristic function can be written as a product of e^{iux} and a part independent of x, that is $\varphi(u|x) = e^{iux}\phi(u)$.

The integration range $[a, b]$ must be chosen carefully to avoid significant errors. An interval which is too small will result in integration range truncation errors and a too wide interval may give rise to cancellation errors. For now we mention the results given by [FO08], in which a rule of thumb for choosing the integration range is given:

$$[a, b] := \left[\kappa_1 - L\sqrt{\kappa_2 + \sqrt{\kappa_4}},\ \kappa_1 + L\sqrt{\kappa_2 + \sqrt{\kappa_4}} \right], \quad L = 10. \tag{2.24}$$

$\kappa_1, ..., \kappa_4$ are the cumulants of the underlying stochastic process $X(t)$, given in [FO08, Table 11].

2.2.1 The option Greeks

The Greeks measure the sensitivities of the option price with respect to a change in its underlying parameters, such as the asset price or the volatility. They are used to hedge the risks in a portfolio. The most well-known Greek parameter is the option Delta, Δ, ie, the first derivative of the option price with respect to underlying asset price $S(t)$. Gamma, Γ, is the second derivative of the option price with respect to the asset price. By the COS pricing formula (2.23) we naturally find the following approximations ($x = \log(S/K)$),

$$\Delta = \frac{\partial v(t_0, x)}{\partial S} \approx \frac{\partial \hat{v}(t_0, x)}{\partial x}\frac{1}{S} = e^{-r\Delta t} \sum_{n=0}^{N}{}' \operatorname{Re}\left\{ \phi\left(\frac{n\pi}{b-a}\right) e^{in\pi\frac{x-a}{b-a}} \frac{in\pi}{b-a} \right\} \hat{g}_n^c \frac{1}{S}, \tag{2.25a}$$

$$\begin{aligned} \Gamma = \frac{\partial^2 v(t_0, x)}{\partial S^2} &\approx \left(\frac{\partial^2 \hat{v}(t_0, x)}{\partial x^2} - \frac{\partial \hat{v}(t_0, x)}{\partial x} \right) \frac{1}{S^2} \\ &= e^{-r\Delta t} \sum_{n=0}^{N}{}' \operatorname{Re}\left\{ \phi\left(\frac{n\pi}{b-a}\right) e^{in\pi\frac{x-a}{b-a}} \left[\left(\frac{in\pi}{b-a}\right)^2 - \frac{in\pi}{b-a} \right] \right\} \hat{g}_n^c \frac{1}{S^2}. \end{aligned} \tag{2.25b}$$

3 Convergence and improvements by spectral filters

In this section, we use, without loss of generality, the interval $[a, b] = [0, 2\pi]$ and we consider the classical Fourier series. As the Fourier-cosine series of a function is equivalent to the symmetrically extended version on an extended domain, the theory here also applies to Fourier-cosine series.

3.1 Convergence of Fourier series

The Fourier series of a function $f(y)$ on $[0, 2\pi]$ is defined as

$$f(y) = \sum_{n=-\infty}^{\infty} \hat{f}_n e^{iny}, \qquad \hat{f}_n = \frac{1}{2\pi}\int_0^{2\pi} f(y)e^{-iny}\, dy. \tag{3.1}$$

The partial sum of the Fourier series is then given by

$$f_N(y) := \sum_{|n|\le N} \hat{f}_n e^{iny}. \tag{3.2}$$

Convergence in L^2-norm is guaranteed when f is square integrable; for $f \in L^2([0, 2\pi])$ we have:

$$\int_0^{2\pi} |f(y) - f_N(y)|^2 \, dy \to 0, \quad \text{as } N \to \infty. \tag{3.3}$$

To make statements about the convergence rate we also need to look at pointwise convergence. If f is a continuous periodic function on $[0, 2\pi]$, so $f(0) = f(2\pi)$, and the Fourier series of f is absolutely convergent, $\sum_{n=-\infty}^{\infty} |\hat{f}_n| < \infty$, then the Fourier series converges uniformly to f. For a non-continuous function there is no uniform convergence and its Fourier series converges to the average of the left- and right-hand limits at a jump discontinuity. In the case that discontinuities of any order are present, the Fourier series expansion only exhibits algebraic convergence. For jump discontinuities, we even encounter 0-th order convergence, which leads to the Gibbs-overshoot.

The algebraic index of convergence of a series a_n is defined as follows: If

$$a_n \sim O(1/n^k), \qquad n >> 1, \tag{3.4}$$

then k is the algebraic index of convergence. This condition means that a constant $c > 0$ and an integer n_0 exist so that

$$|a_n| \leq c/n^k, \qquad \forall n > n_0. \tag{3.5}$$

To measure the algebraic index of convergence of a given series by a numerical experiment, one can look at the log-log plot of a_n. The *upper part* of the slope for $n \to \infty$ equals $-k$. The speed at which the Fourier coefficients decay depends on the smoothness of the function, as is stated in the following theorem:

Theorem 3.1. *(Integration-by-parts coefficient bound)[Boy01]*
If

$$f(0) = f(2\pi),\ f^{(1)}(0) = f^{(1)}(2\pi), \ldots,\ f^{(k-2)}(0) = f^{(k-2)}(2\pi) \tag{3.6}$$

and $f^{(k)}(y)$ is integrable, then the coefficients of the Fourier series have upper bounds $|\hat{f}_n| \leq c/|n|^k$, for some sufficiently large constant c, which is independent of n. So,

$$\hat{f}_n \sim \mathcal{O}\left(|n|^{-k}\right) \quad \text{as} \quad |n| \to \infty. \tag{3.7}$$

An equivalent way of stating the theorem is that, if the two conditions above are satisfied, the algebraic index of convergence is at least as large as k.
Here $f^{(k)}$ denotes the k-th derivative of $f(y)$ and $\mathcal{O}(.)$ is the Landau gauge symbol. The integrability of $f^{(k)}$ requires that $f(y)$, $f^{(1)}(y)$,..., $f^{(k-2)}(y)$ should be continuous.

The absolute error of truncation of the expansion after N terms is denoted by

$$E_{Tr}(N) := |f(y) - f_N(y)| \leq \sum_{n=N+1}^{\infty} |\hat{f}_n + \hat{f}_{-n}| := \sum_{n=N+1}^{\infty} |a_n|.$$

In general the convergence rate of a Fourier series depends on the smoothness of the function on the expansion interval. We refer to [Boy01] for the definitions of algebraically and exponentially converging terms. The following proposition allows us to bound the series truncation error of a geometrically and algebraically converging series.

Proposition 3.1. *(Last coefficient error estimate)[Boy01]*
The truncation error is of the same order of magnitude as the last coefficient retained in the truncation for a series with (at least) geometric convergence. Since the truncation error is a

quantity we can only estimate (in the absence of a known, exact solution), we can loosely speak of the last retained coefficient as being the truncation error, that is:

$$E_{Tr}(N) \sim \mathcal{O}(|a_N|) \sim \mathcal{O}([]\exp(-N\nu)). \tag{3.8}$$

Here, $\nu > 0$ is a constant and [] denotes a factor which varies less than exponentially with N.

If the series has algebraic convergence index $k > 1$, ie, if $a_n \sim \mathcal{O}(1/n^k)$ for large n, then

$$E_{Tr}(N) \sim \mathcal{O}(|Na_N|). \tag{3.9}$$

Note that an algebraically converging series behaves as, ([BO99]),

$$\sum_{n=N+1}^{\infty} \frac{1}{n^k} \sim \mathcal{O}\left(\frac{1}{(k-1)N^{k-1}}\right) = \mathcal{O}(N\frac{1}{N^k}), \quad N >> 1. \tag{3.10}$$

In the numerical experiments in Section 4.1 we will observe that the Fourier coefficients of a $(k-2)$-times continuously differentiable function, $f \in C^{k-2}$, converge with an algebraic index of convergence k, see Theorem 3.1. According to the proposition above we would expect truncation error $E_{Tr}(N)$ to decrease with order $\mathcal{O}(N^{1-k})$. However, we find a faster convergence, of order k. This may be due to the *alternating behavior* of the series. A similar behavior is observed in a well-known series example, presented in Appendix A.

Example for the Gibbs phenomenon: Suppose function f represents a block function, that is $f(y) = \mathbb{I}_{y\in[0.5\pi,1.5\pi]}$, $y \in [0, 2\pi]$. The left-side plot in Figure 3.1 shows function $f(y)$ (green) and its Fourier series approximation with $N = 32$ (red). At the jump discontinuities the limit converges to the average of the values of the function at either side of the jump. The middle plot gives the corresponding error. In the right-side plot we display the convergence of the error at the point $y = \pi$, for increasing values of N (log-log plot). The local effect of the Gibbs phenomenon gives rise to oscillations near the jumps and the partial sum does not converge at the jump, thereby resulting in a lack of uniform convergence. However, there also is a global effect: although the error decays away from the jumps, the algebraic pointwise convergence is only first order, $\mathcal{O}(1/N)$.

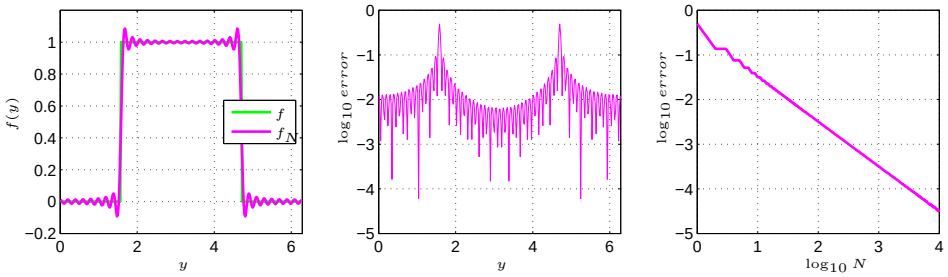


Figure 3.1: Gibbs phenomenon Fourier series ($N = 32$).

3.2 Spectral filter

In this section we introduce the notion of spectral filters, by which we aim to mitigate oscillations related to the Gibbs phenomenon and achieve faster convergence for problems in financial mathematics. Filtering is carried out in Fourier space and the idea is to pre-multiply the expansion coefficients by a decreasing function in such a way that these decay faster. A properly chosen filter will improve the convergence rate away from discontinuities.

Definition 3.1. *[Van91] Any $C^\infty([0,1])$ even function $\hat{s}$, whose support is $[-1,1]$ with $\hat{s}(0)=1$, is called a filter.*

The *filtered* partial sum of a Fourier series is defined by

$$f_N^{\hat{s}}(y) = \sum_{|n|\le N} \hat{s}(n/N)\hat{f}_n e^{iny}. \tag{3.11}$$

We can rewrite this as a convolution in physical space:

$$f_N^{\hat{s}}(y) = \sum_{|n|\le N} \hat{s}(n/N)\hat{f}_n e^{iny} = \frac{1}{2\pi}\int_0^{2\pi} \sum_{|n|\le N} \hat{s}(n/N) e^{iny} e^{-int} f(t)\,dt = \frac{1}{2\pi}\int_0^{2\pi} s(y-t) f(t)\,dt, \tag{3.12}$$

with filter function

$$s(x) = \sum_{|n|\le N} \hat{s}(n/N) e^{inx} = \sum_{|n|\le \infty} \hat{s}(n/N) e^{inx}, \quad x \in [0, 2\pi]. \tag{3.13}$$

Note that $\hat{s}(n/N) = 0$ for $|n| > N$. The filter function is the representation of the spectral filter $\hat{s}(n/N)$ in physical space.

A filter is a continuous function which only modifies high frequency modes, not the low modes. Filtering may remove the Gibbs phenomenon away from a discontinuity, the error depends on the distance to the discontinuity, as we will confirm in Section 3.3. Since the approximation will be smoothened, convergence in the vicinity of a "jump" will not improve. Filtering does not affect the total mass of the resulting approximation (which should be one for a probability density), since the first coefficient is never altered. To be precise,

$$\int_0^{2\pi} f_N^{\hat{s}}(y)\,dy = \sum_{|n|\le N} \hat{s}(n/N)\hat{f}_n \int_0^{2\pi} e^{iny}\,dy = \hat{s}(0/N)\hat{f}_0 2\pi = \int_0^{2\pi} f(y)\,dy. \tag{3.14}$$

The following definition is from [GS97]:

Definition 3.2. *(Fourier space filter of order p) A real and even function $\hat{s}(\eta)$ is called a filter of order p, if*

1. *$\hat{s}(0) = 1$ and $\hat{s}^{(\ell)}(0) = 0, \;\; 1 \le \ell \le p-1$,*
2. *$\hat{s}(\eta) = 0$ for $|\eta| \ge 1$,*
3. *$\hat{s}(\eta) \in C^{p-1}, \quad \eta \in (-\infty, \infty)$.*

Conditions 2 and 3 imply $\hat{s}^{(\ell)}(1) = 0, \;\; 0 \le \ell \le p-1$.

3.2.1 Examples of spectral filters

The following filters are well-known from the literature [Van91, HGG07, GS97]:

- *Fejér filter ([Fej00]):* $\hat{s}(\eta) = 1 - |\eta|$, with order $p = 1$.
- *Lanczos filter ([Lan56]):* $\hat{s}(\eta) = \sin(\pi\eta)/(\pi\eta)$, with order $p = 1$.

- *Raised cosine filter:* $\hat{s}(\eta) = \frac{1}{2}(1 + \cos(\pi\eta))$, with order $p = 2$.

Also general p-th order spectral filters exist:

- *Exponential filter:*

$$\hat{s}(\eta) = \exp(-\alpha\eta^p), \tag{3.15}$$

  where p must be even. $\hat{s}(1) = e^{-\alpha}$, so the formal requirements of a p-th order filter do not hold. However, we will choose $\alpha = -\log\epsilon_m$, where ϵ_m represents the machine epsilon, so that $\hat{s}(1) = \epsilon_m \approx 0$ within machine precision.

- *Vandeven filter ([Van91, Pey02]):*

$$\begin{aligned}\hat{s}(\eta) &= 1 - \frac{(2p-1)!}{((p-1)!)^2}\int_0^{|\eta|} t^{p-1}(1-t)^{p-1}\,dt \\ &= 1 + (-1)^p\frac{(2p-1)!}{((p-1)!)^2}\sum_{j=0}^{p-1}\frac{(-1)^j}{(p-1-j)!j!}\frac{|\eta|^{2p-j-1}}{2p-j-1}.\end{aligned} \tag{3.16}$$

- *Erfc-Log filter ([Boy96]):* Boyd showed that the Vandeven filter can be approximated quite accurately by an analytic function which satisfies all conditions, ie, by the Erfc-Log filter:

$$\hat{s}(\eta) = \frac{1}{2}\mathrm{erfc}\left(2\sqrt{p}(|\eta| - \tfrac{1}{2})\sqrt{\frac{-\log(1-4(|\eta|-1/2)^2)}{4(|\eta|-1/2)^2}}\right), \tag{3.17}$$

  where $\mathrm{erfc}(\cdot)$ is the complimentary Gauss error function.

3.3 Convergence and error analysis

A higher order filter modifies the original function in smooth regions away from a discontinuity and high order accuracy is desirable away from a discontinuity. Low order filtering is however desirable close to a discontinuity, because higher order p values give rise to a highly oscillatory filtered function $f_N^{\hat{s}}$. The following theorem is given in [GS97] and gives a bound on the error. It can be extended in a straightforward way to a function with more points of discontinuity ξ_m.

Theorem 3.2. *Let $f(y)$ be a piecewise $C^p([0,2\pi])$ function with one point of discontinuity ξ. Let $\hat{s}(n/N)$ be a filter of order p. Let now y be a point in $[0,2\pi]$ and denote by $d(y) := \min_{k=-1,0,1}|y - \xi + 2k\pi|$. Then (if $y \neq \xi$),*

$$|f_N^{\hat{s}}(y) - f(y)| = \frac{1}{2\pi}\sum_{\ell=0}^{p-1} s_{l+1}(d(y))\left(f^{(\ell)}(\xi^+) - f^{(\ell)}(\xi^-)\right) + \frac{1}{2\pi}\int_0^{2\pi} s_p(y-t)f^{(p)}(t)\,dt \tag{3.18}$$

$$\leq cN^{1-p}d(y)^{1-p}K(f) + cN^{\frac{1}{2}-p}||f^{(p)}||_{L^2}, \tag{3.19}$$

where

$$K(f) = \sum_{\ell=0}^{p-1} d(y)^\ell\left(f^{(\ell)}(\xi^+) - f^{(\ell)}(\xi^-)\right)\int_{-\infty}^{\infty}|G_\ell^{(p-\ell)}(\eta)|\,d\eta, \tag{3.20a}$$

$$G_\ell(\eta) = \frac{\hat{s}(\eta) - 1}{\eta^\ell}, \tag{3.20b}$$

c is a constant independent of f and N, and

$$s_0(x) = s(x), \qquad s'_\ell = s_{\ell-1}, \quad \ell \geq 1, \qquad \int_0^{2\pi} s_\ell(t)\,dt = 0, \quad \ell \geq 1. \tag{3.21}$$

The error bound decreases with $d(y)$, ie, with the distance to the discontinuity. The filter order determines the rate at which the error remaining after filtering decays. If we have $f \notin C^{p-1}$, ie, if $f(y)$ has a jump discontinuity at one or more points of order smaller than, or equal to, $p-1$, the following estimate holds:

$$|f_N^{\hat{s}}(y) - f(y)| \sim \mathcal{O}(N^{1-p}). \tag{3.22}$$

If $f \in C^{p-1}$, ie, if $f(y)$ is smooth in the sense of possessing at least $p-1$ continuous derivatives, then:

$$|f_N^{\hat{s}}(y) - f(y)| \sim \mathcal{O}\left(N^{\frac{1}{2}-p}\right). \tag{3.23}$$

3.4 Stricter bound error

In the numerical experiments in Section 4.1 we will observe a somewhat *faster convergence* than prescribed by Theorem 3.2. In this section we discuss details of the error analysis.

For the first part in Equation (3.18) the authors in [Van91] prove by induction that

$$s_\ell(x) \sim \mathcal{O}(N^{1-p}), \quad x \in (0, 2\pi),\ 0 \leq \ell \leq p. \tag{3.24}$$

Table 3.1 shows the order of convergence for s_0 and s_1 that we observed by numerical experiments. We tested six different filters: The Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter and the Erfc-Log filter, as described in Section 3.2. Besides we used different filter orders. Here "exp" denotes exponential convergence. The algebraic index of convergence is thus at least filter order p.

Filter	$s_0(x)$	$s_1(x)$
Fejér	1	1
Lanczos	1	2
Raised cosine	2	3
Exponential ($p = 2, 4, 6$)	exp	exp
Vandeven ($p = 1, 3, 5$)	p	p
Vandeven ($p = 2, 4$)	p	$p+1$
Erfc-Log ($p = 1, 2, 3, 4, 5$)	p	p

Table 3.1: Algebraic index of convergence ($x \in (0, 2\pi)$).

For the second part in Equation (3.18) the authors in [Van91, GS97] use the inequalities

$$\left|\frac{1}{2\pi}\int_0^{2\pi} s_\ell(y-t) f^{(\ell)}(t)\,dt\right| \leq \frac{1}{2\pi}\sqrt{\int_0^{2\pi} s_\ell^2(y-t)\,dt}\sqrt{\int_0^{2\pi}(f^{(\ell)}(t))^2\,dt} \tag{3.25}$$

and

$$\frac{1}{2\pi}\int_0^{2\pi} s_\ell^2(t)\,dt = 2\sum_{n=1}^{\infty}\frac{(1-\hat{s}(n/N))^2}{n^{2\ell}} = 2\sum_{n=1}^{N}\frac{(1-\hat{s}(n/N))^2}{n^{2\ell}} + 2\sum_{n=N+1}^{\infty}\frac{1}{n^{2\ell}}. \tag{3.26}$$

They then find

$$\sum_{n=N+1}^{\infty} \frac{1}{n^{2\ell}} \sim \mathcal{O}(N^{1-2\ell}), \tag{3.27a}$$

$$\sum_{n=1}^{N} \frac{(1-\hat{s}(n/N))^2}{n^{2\ell}} \sim \mathcal{O}(N^{1-2\ell}), \quad 1 \leq \ell \leq p. \tag{3.27b}$$

However, for the filters in Table 3.1 we observe by numerical computations that

$$\sum_{n=1}^{N} \frac{(1-\hat{s}(n/N))^2}{n^{2\ell}} \sim \mathcal{O}(N^{\max(-2p,1-2\ell)}), \quad 1 \leq \ell \leq 10. \tag{3.28}$$

The proof for the case $\ell > p$ is open.

3.5 Filtering and the COS method

One of the reasons why the COS method is highly efficient is because pricing formula (2.23) works directly with the coefficients, without a-priori recovery of the functions $p(y|x)$ or $g(y)$. Spectral filters work strictly in the Fourier domain and therefore they can be used directly in the COS pricing formula. Once a suitable filter and order have been chosen one can multiply each of the Fourier-coefficients by a factor $\hat{s}(n/N)$ and work with the COS method as before. This gives us the *filter-COS pricing formula*, as follows

$$v(t_0,x) \approx e^{-r\Delta t} {\sum_{n=0}^{N}}' \hat{s}\left(\frac{n}{N}\right) \mathrm{Re}\left\{\varphi\left(\frac{n\pi}{b-a}\middle|x\right) e^{in\pi\frac{-a}{b-a}}\right\} \hat{g}_n^c, \tag{3.29}$$

where $\hat{s}$ can be any non-adaptive filter. It does not add significant computational costs. Unfortunately, the same does not hold for the adaptive filters [TT05], because if we vary coefficients depending on position we cannot use the orthogonality relation which leads to the COS method anymore.

3.6 Error analysis COS method and filter-COS method

Without filtering: The error of the COS formula without filtering terms is composed of three parts: The integration range truncation error, the series truncation error and the error related to approximating $\hat{p}_n^c$ by $\hat{\varphi}_n^c$. A detailed error analysis was given in [FO08], where it was shown that if $p(y|x) \in \mathbb{C}^\infty$, with a properly chosen truncation range $[a,b]$, then the COS method exhibits geometric convergence.

If the computational domain $[a,b]$ is chosen sufficiently wide, then the so-called *series truncation error* $E_{Tr}^{COS}(N)$, ie,

$$E_{Tr}^{COS}(N) := \frac{b-a}{2} e^{-r\Delta t} \sum_{n=N+1}^{\infty} \hat{\varphi}_n^c(x)\hat{g}_n^c = e^{-r\Delta t} \int_a^b g(y)\left(p(y|x) - p_N(y|x)\right)\, dy, \tag{3.30}$$

dominates the total error. The series truncation error $E_{Tr}^{COS}(N)$ depends on the smoothness of underlying probability density function $p(y|x)$ and payoff function $g(y)$. With Equation (3.30) it follows that if either $\hat{\varphi}_n^c$ or $\hat{g}_n^c$ decreases geometrically, then the error has geometric convergence in N. If however both decay algebraically, then we end up with algebraic convergence. If $\hat{\varphi}_n^c(x) \sim \mathcal{O}(n^{-k_\varphi})$ and $\hat{g}_n^c \sim \mathcal{O}(n^{-k_g})$, Theorem 3.1 gives $E_{Tr}^{COS}(N) \sim \mathcal{O}(N^{-k_f-k_g+1})$. In the numerical

examples we will observe convergence rate $E_{Tr}^{COS}(N) \sim \mathcal{O}(N^{-k_f-k_g})$, which is probably due to the alternating behavior of the series (as in Appendix A).

Payoff functions, like puts and calls, are in general non-smooth, which is the reason for slowly decreasing Fourier coefficients. Asset prices modeled by geometric Brownian motion, jump-diffusion or the Heston model lead to geometric decay of the coefficients $\hat{\varphi}_n^c$, resulting in an exponentially converging COS formula. However, the density functions of the Variance Gamma and CGMY models may be non-smooth. Together with a non-smooth payoff function this will result in rather slow algebraic error convergence in N.

Computing the option Greeks was briefly described in Section 2.2.1. As the coefficients for the Greeks are multiplied by factors $\frac{in\pi}{b-a}$, the algebraic index of convergence is reduced.

With filtering: Applying spectral filter $\hat{s}$ in the COS formula leads to the following relation,

$$E_{Tr}^{filter-COS}(N) = e^{-r\Delta t} \int_a^b g(y) \left(p(y|x) - p_N^{\hat{s}}(y|x)\right) \, dy \tag{3.31}$$

and the convergence depends on the decay of the error $p(y|x) - p_N^{\hat{s}}(y|x)$, which was discussed in Section 3.3 (Theorem 3.2) and Section 3.4.

4 Numerical examples

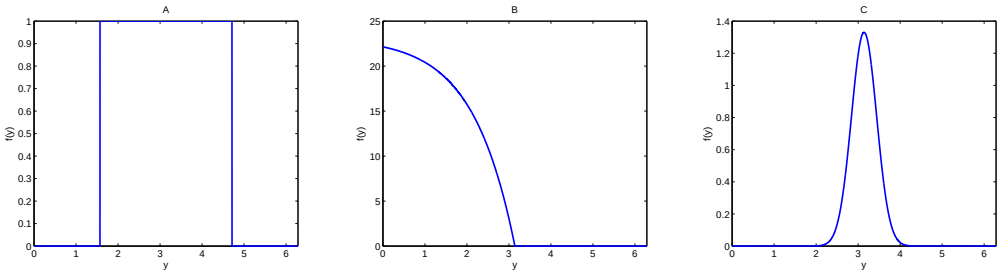
In this section we discuss several numerical experiments. We start in Section 4.1 with three basic test functions $f(y)$. Subsequently, in Section 4.2.1 the density recovery of the Variance Gamma process is studied. Convergence of option prices, computed by the filter-COS method, is discussed in Section 4.2.2 for European-style options and in Section 4.2.3 for Bermudan-style options. An example from portfolio loss modeling, resulting in a staircase distribution function, is presented in Section 4.3.

4.1 Convergence test functions

We perform tests with three different functions $f(y)$ and different filters. The test functions are shown in Figure 4.1, with $[a, b] = [0, 2\pi]$. Function A is governed by a jump discontinuity, function B has a discontinuity in the first derivative and function C is smooth. Figure 4.2 shows the approximation of function A by Fourier series, without (see Figure 3.1) and with exponential filter with order $p = 4$. Function A converges to $f(y) = 1/2$ at the jump discontinuities. The filter improves the error significantly, especially away from the jump discontinuities. The right-side plot shows the convergence of the error at point $y = \pi$, which decreases exponentially in N due to the usage of the filter.

To recover the functions, we test the performance of six filters, ie, the Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter and the Erfc-Log filter (see Section 3.2). We also employ different filter orders. Table 4.1 presents the algebraic index of convergence obtained for the three test functions. Here "exp" denotes exponential convergence. With two numbers, eg, "2, 4", convergence is order $\mathcal{O}(N^{-2})$ on $[0, \pi]$ and $\mathcal{O}(N^{-4})$ on $[\pi, 2\pi]$. A star (*) in the table indicates that the order of convergence was not clearly measurable in our numerical experiments.

Based on the error analysis of Theorem 3.2 and our observations in Section 3.4 we can explain the numbers in Table 4.1 as follows:



(a) $f(y) = \mathbf{1}_{y\in[0.5\pi,1.5\pi]}$. (b) $f(y) = \max(e^{\pi} - e^{y}, 0)$. (c) $f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}\right)$.

Figure 4.1: Three test functions $f(y)$.

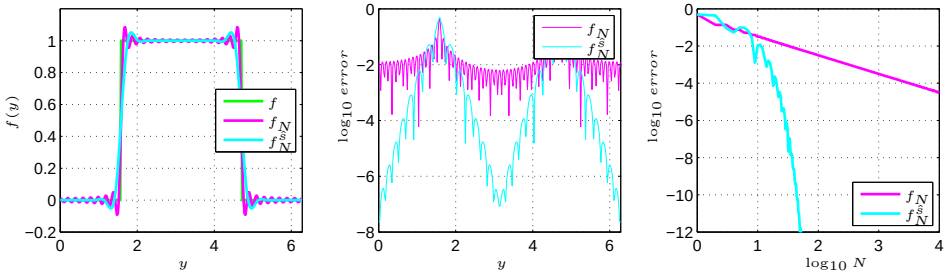


Figure 4.2: Recovery $f(y)$ (A) without and with filter (exponential, order $p = 4$) ($N = 32$).

Filter	Function A	Function B	Function C
No filter	1	1	exp
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	3	2, 3	2
Exponential ($p = 2$)	exp	2, exp	2
Exponential ($p = 4$)	exp	4, exp	4
Exponential ($p = 6$)	exp	*, exp	6
Vandeven ($p = 1, 3, 5$)	p	p	p
Vandeven ($p = 2, 4$)	$p + 1$	p, $p + 1$	p
Erfc-Log ($p = 1, 2, 3, 4, 5$)	p	p	p

Table 4.1: Algebraic index of convergence test functions.

For function A we have

$$|f_N^{\hat{s}}(y) - f(y)| = \sum_m \frac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-)\right) + \frac{1}{2\pi}\int_0^{2\pi} s_1(y-t) f^{(1)}(t)\, dt$$
$$= \sum_m \frac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-)\right). \tag{4.1}$$

The observed decay rates of s_1 are given in Table 3.1 and they correspond to the results in the second column in Table 4.1.

The error for function B reads

$$|f_N^{\hat{s}}(y) - f(y)| = \sum_m \frac{1}{\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-)\right) + \frac{1}{2\pi}\int_0^{2\pi} s_1(y-t) f^{(1)}(t)\, dt$$

$$= \sum_m \frac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-) \right) + \frac{1}{2\pi} \int_0^\pi s_2(y-t) f^{(2)}(t)\, dt$$
$$- \frac{1}{2\pi} s_2(y-\pi) f^{(1)}(\pi) + \frac{1}{2\pi} s_2(y-0) f^{(1)}(0)$$
$$= \ldots = \sum_m \frac{1}{2\pi} s_1(d_m(y)) \left(f(\xi_m^+) - f(\xi_m^-) \right) + \frac{1}{2\pi} \int_0^\pi s_q(y-t) f^{(q)}(t)\, dt$$
$$+ \sum_{\ell=1}^{q-1} \left(-\frac{1}{2\pi} s_{\ell+1}(y-\pi) f^{(\ell)}(\pi) + \frac{1}{2\pi} s_{\ell+1}(y-0) f^{(\ell)}(0) \right). \tag{4.2}$$

So, with the observations in Section 3.4 we obtain an algebraic rate of convergence p. However, for $y \in [\pi, 2\pi]$ we have $\frac{1}{2\pi} \int_0^\pi s_q(y-t) f^{(q)}(t)\, dt \approx 0$, so that we obtain the same index of convergence as s_1.

Function C is approximately smooth, so that

$$|f_N^{\hat{s}}(y) - f(y)| = \frac{1}{2\pi} \int_0^{2\pi} s_q(y-t) f^{(q)}(t)\, dt, \quad \forall q \geq 1. \tag{4.3}$$

Following the arguments in section 3.4 we expect

$$\frac{1}{2\pi} \int_0^{2\pi} s_q(y-t) f^{(q)}(t)\, dt \sim \mathcal{O}\left(N^{\max(-p, \frac{1}{2}-q)} \right). \tag{4.4}$$

This gives us the convergence order $\mathcal{O}(N^{-p})$, as observed.

4.2 The Variance Gamma process

In this section, we discuss various application of the COS method in the context of the Variance Gamma (VG) process [MCC98, Sch03]. We start with the accurate and efficient recovery of the VG density.

4.2.1 Variance Gamma density recovery

In the case of modeling asset prices by a fat-tailed density function, the exponential Variance Gamma jump process can be applied. The Variance Gamma (VG) process is obtained by evaluating a Brownian motion with drift θ and volatility σ at a random time given by a gamma process $\gamma(t)$ with mean rate one and variance rate ν, [MCC98, Sch03]:

$$X(t) = \theta\gamma(t) + \sigma W(\gamma(t)). \tag{4.5}$$

The asset price is then defined as $S(t) = S_0 e^{X(t)}$. The VG process is of bounded variation, has independent increments and is defined by an infinite arrival of jumps. The VG density can be characterized by a *fat tail*: it is suitable to model phenomena where small and relatively large asset values are more probable than would be the case for the lognormal distribution. The characteristic function, $\varphi(u|x) = \mathbb{E}[e^{iuX(T)}|X(t_0) = x] = e^{iux}\phi_{VG}(u)$, is given by [Sch03, MCC98]:

$$\phi_{VG}(u) = \left(1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2 \right)^{-\Delta t/\nu} \sim \mathcal{O}(u^{-2\Delta t/\nu}). \tag{4.6}$$

In [MCC98], the following expression for the VG-density function was derived:

$$p_{VG}(y) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi z}} \exp\left(-\frac{(y-\theta z)^2}{2\sigma^2 z} \right) \frac{z^{\frac{\Delta t}{\nu}-1} \exp(-\frac{z}{\nu})}{\nu^{\frac{\Delta t}{\nu}} \Gamma(\frac{\Delta t}{\nu})}\, dz \quad \text{(with } \Gamma \text{ the Gamma function).} \tag{4.7}$$

It is computationally rather expensive to evaluate (4.7) at each point in the domain of interest. The smoothness of the density function depends on its parameters, to be more precise, with higher values of $\Delta t/\nu$ a larger number of derivatives exist.

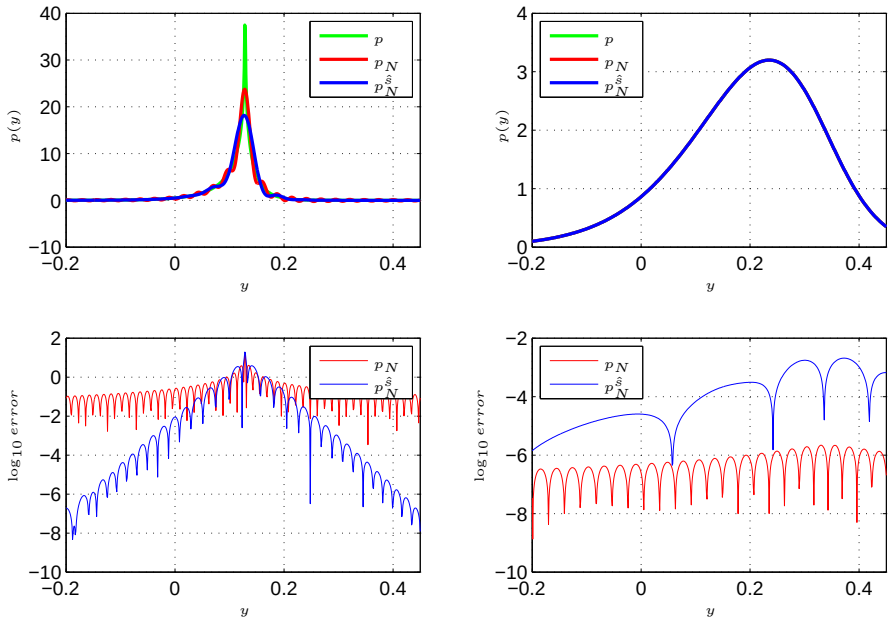


Figure 4.3: VG density and error for different values of T (Left: $T = 0.1$, right: $T = 1$).

The parameters used for our tests here are the same as in [FO08], ie,:

$$K = 90, S(0) = 100, r = 0.1, \sigma = 0.12, \theta = -0.14, \nu = 0.2. \tag{4.8}$$

Figure 4.3 shows the VG-density for two different values of T. For small values of T the peak in the density gets really sharp and is difficult to approximate accurately by a Fourier-cosine series. For increasing ν-values the peak sharpens, while increasing σ-values result in a smaller peak and wider tails, without altering the smoothness properties of the function around the peak significantly.

Figure 4.3 also shows the approximations by Fourier-cosine series, ie, Equation (2.12), with and without exponential filter (order $p = 6$). The characteristic function exhibits an algebraic decay with order $2\Delta t/\nu$, giving rise to an algebraic decay of the Fourier coefficients and therefore slow convergence of the COS method, especially for the density with $T = 0.1$. Filtering works well away from the peak in the VG density, but right at the peak the approximation becomes somewhat worse. This can theoretically be dealt with by *adaptive filtering*, where the filter order depends on the distance to the discontinuity, but also then the approximation at the peak will at best be as accurate as the Fourier approximation.

We evaluate the performance of the filters to recover the VG density. Again six different filters are used and we experiment with different filter orders. We find that lower order filters smooth the sharp peak too much. Table 4.2 shows the algebraic index of convergence for the density recovery. We observe a significant improvement for $T = 0.025$ and $T = 0.1$, whereas for $T = 1$ the "no filter case" is superior.

Filter	$T = 0.025$	$T = 0.1$	$T = 1$
No filter	0.25	1	10
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	2	2	2
Exponential ($p = 2, 4$)	p	p	p
Exponential ($p = 6$)	*	6	6
Vandeven ($p = 1, 2, 3, 4, 5$)	p	p	p
Erfc-Log ($p = 1, 2, 3, 4, 5$)	p	p	p

Table 4.2: Algebraic index of convergence VG density recovery with three values of T.

4.2.2 European options and Greeks under Variance Gamma

We investigate the convergence of the COS and the filter-COS method for pricing options. A put option gives the holder the right, but not the obligation, to sell an underlying asset for a prescribed strike price K at expiration time T. The payoff function thus reads $g(S, T) = \max(K - S(T), 0)$. For a digital (or binary) call option the payoff is either one or zero, with payoff function $g(S, T) = \mathbf{1}_{S(T) \geq K}$.

Reference values for the experiments are obtained by selecting an accurate filter and a large number of terms in the series expansions, see Table 4.3. The convergence of the COS method *without filtering* is of order $2\Delta t/\nu + 2$ for put options and $2\Delta t/\nu + 1$ for digital options. The algebraic index of convergence for the Greeks Delta Δ (Equation (2.25a)), without filtering, is one order lower and for Gamma Γ (Equation (2.25b)) it is two orders lower, see Table 4.4. The absolute value of the series terms gives us $\mathcal{O}\left(1/n^{2\Delta t/\nu+1-2}\right)$ convergence for the Gamma Γ of a digital option. So, the algebraic index of convergence is *non-positive* for expiration times $T = 0.1$ and $T = 0.025$ and the series does not converge ("Div"), because of the Cauchy convergence criterion. The intuition behind this is that for very short expiration times, the Γ converges to the Dirac delta function.

Table 4.5 shows the order of convergence of the European option with different spectral filters. Similar convergence results are obtained for the Greeks Δ and Γ.

We tend to prefer the exponential filter to the Vandeven and Erfc-Log filters. Its implementation is easiest and fastest, although the other filters are not significantly more time-consuming. Furthermore, we observe an exponential convergence for step functions with the exponential filter (see Figure 4.2 and Table 4.1), which is advantageous for recovery of a distribution function for discrete random variables. Therefore, we focus on the exponential filter in the remainder.

For asset price processes with independent increments, like the VG model, we can employ the filter-COS method to compute option values for multiple strike prices simultaneously. For example, for $K \in [80, 120]$ we obtain the same convergence results. However, we observe higher absolute errors in the options for strike prices near the peak value in the VG density function. This can be explained by the smoothing of the peak by a filter.

Figure 4.4 displays the error of the COS formula for the option value and the Greeks for expiration time $T = 0.1$ in a log-log plot. Exponential filters with different orders are used. The use of filters improves the error and convergence order significantly, especially regarding the option Gamma Γ.

	$T = 0.025$		$T = 0.1$		$T = 1$	
	Put	Digital	Put	Digital	Put	Digital
Option value $v(t_0, x)$	0.02435	89.1883	0.09819	86.9759	0.53472	74.7855
Delta Δ	-0.50629	12.8417	-1.82737	40.6550	-5.50365	63.1902
Gamma Γ	11.5565	-313.402	36.5895	-843.107	56.8712	-581.247

Table 4.3: Reference values European options.

	$T = 0.025$		$T = 0.1$		$T = 1$	
	Put	Digital	Put	Digital	Put	Digital
Option value $v(t_0, x)$	2.25	1.25	3	2	12	11
Delta Δ	1.25	0.25	2	1	11	10
Gamma Γ	0.25	Div	1	Div	10	9

Table 4.4: Algebraic index of convergence for European options (no filter).

Filter	Put	Digital
Fejér	1	1
Lanczos	2	2
Raised cosine	2	2
Exponential ($p = 2, 4, 6$)	p	p
Vandeven ($p = 1, 2, 3, 4, 5$)	p	p
Erfc-Log ($p = 1, 2, 3, 4, 5$)	p	p

Table 4.5: Algebraic index of convergence for European options obtained by filter-COS ($T = 0.025$, $T = 0.1$ and $T = 1$).

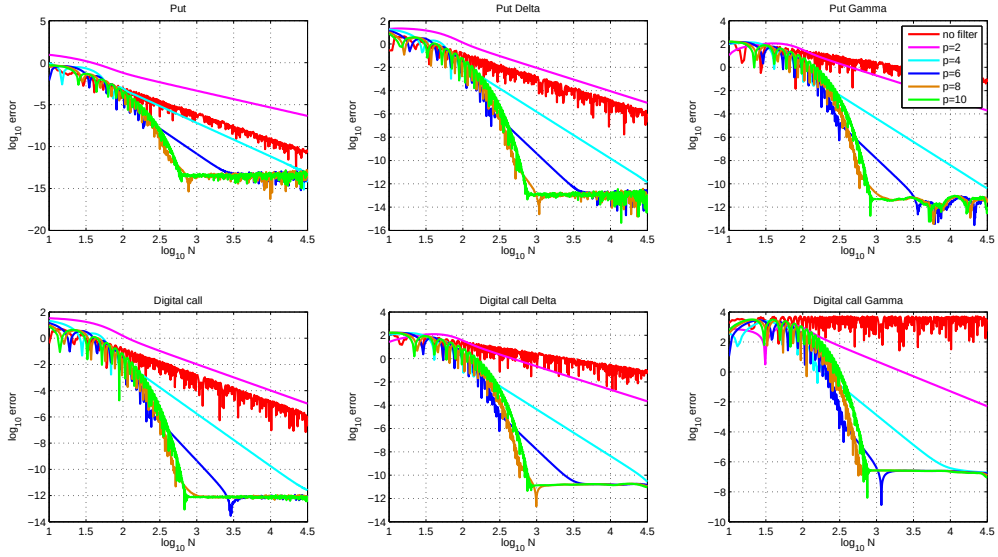


Figure 4.4: Error convergence for European options; put (top) and digital call (bottom) plus the Greeks, with exponential filters ($T = 0.1$).

4.2.3 Bermudan option price under VG

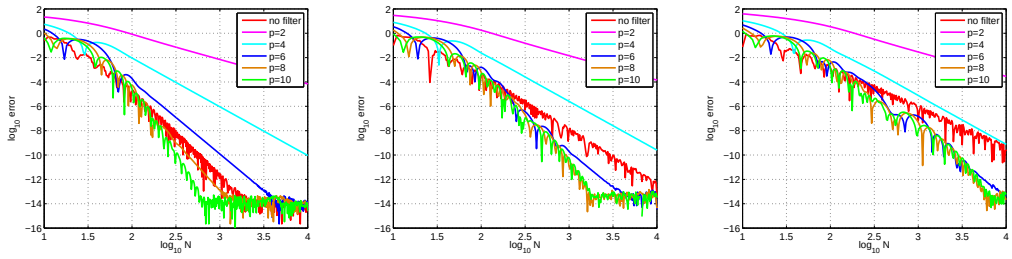
A Bermudan-style option can be exercised at a *set of* $\mathcal{M}$ *early-exercise dates* prior to the expiration time T, $t_0 < t_1 < \ldots t_m < \ldots < t_{\mathcal{M}} = T$, with timestep $\Delta t := t_{m+1} - t_m$. The authors in [FO09]

describe a recursive algorithm, based on the COS method, for pricing Bermudan options backwards in time via Bellman's principle of optimality. We also employ the COS method for Bermudan options here, but replace the coefficients $\hat{\varphi}_n^c$ by the filtered version, ie, $\hat{s}(n/N)\hat{\varphi}_n^c$, similar as in Equation (3.29).

In Table 4.6 and in Figure 4.5 the results for the exponential filter with $T = 1$ and different numbers of early-exercise dates are presented. The COS method with filtering becomes more and more beneficial when a larger number of early-exercise dates is used, meaning a smaller timestep Δt.

	$T = 1$		
Filter	$\mathcal{M} = 2$	$\mathcal{M} = 4$	$\mathcal{M} = 10$
No filter	7	4.5	3
Fejér	1	1	1
Lanczos	2	2	2
Raised cosine	2	2	2
Exponential ($p = 2, 4$)	p	p	p
Exponential ($p = 6$)	6	6	*
Vandeven ($p = 1, 2, 3, 4, 5$)	p	p	p
Erfc-Log ($p = 1, 2, 3, 4$)	p	p	p
Erfc-Log ($p = 5$)	5	5	*

Table 4.6: Algebraic index of convergence for Bermudan put options with filter-COS method ($T = 1$).



(a) $\mathcal{M} = 2$, $v(t_0, x) = 0.64386$. (b) $\mathcal{M} = 4$, $v(t_0, x) = 0.71161$. (c) $\mathcal{M} = 8$, $v(t_0, x) = 0.75220$.

Figure 4.5: Error convergence for Bermudan options with an increasing number of early-exercise dates, filter-COS technique with exponential filters ($T = 1$).

4.3 Portfolio loss distribution

In this section we present a final example, which has a financial background in risk management. The distribution function of interest will be a staircase function, which is difficult to approximate by Fourier series, however the filtering technique will improve the results significantly.

For a bank it is important to manage the risk originating from its business activities. The credit risk underlying a credit portfolio is one of the largest risk portions of a bank. For quantifying losses in credit portfolios one often looks at the Value at Risk (VaR). The VaR of a portfolio at confidence level α is given by the smallest value x, for which the probability that loss L exceeds x is at most $(1 - \alpha)$:

$$\text{VaR}_\alpha = \inf\{x \in \mathbb{R} : P(L > x) \leq (1 - \alpha)\} = \inf\{x \in \mathbb{R} : F_L(x) \geq \alpha\}, \tag{4.9}$$

where F_L is the cumulative loss distribution function.

The Vasicek model [Vas02] is often used to find an approximation to the loss distribution and to compute the VaR. Under this model losses occur when an obligor defaults in a fixed time horizon. Suppose there are $m = 1, \ldots, M$ issuers and X_m represents the individual asset return of issuer m. In accordance with the Vasicek model we then use

$$X_m = \sqrt{\rho_m} Y + \sqrt{1 - \rho_m} Z_m, \quad m = 1, \ldots, M, \tag{4.10}$$

where Y is a common economic factor, Z_m is the idiosyncratic factor for issuer m and ρ_m is the correlation between Y and Z_m. All random variables are assumed to follow a standard normal distribution and Y and Z_m are independent. When the asset return falls below a default threshold c_m, there is a loss ℓ_m. We define the default probability of issuer m by $P_m := \mathbb{P}(X_m < c_m)$. The individual credit loss is defined by $L_m = \ell_m \mathbf{1}_{X_m < c_m}$ and the total portfolio loss reads $L = \sum_{m=1}^{M} L_m$.

The COS method can be used to approximate the cumulative distribution function of the loss by means of its characteristic function, see Equation (2.16), and to calculate the VaR. The probability distribution is based on a discrete set of events, resulting in a *step-wise distribution function*, which causes the COS method to suffer from the Gibbs phenomenon: With a loss distribution which is discontinuous, significant errors appear around the points of discontinuity [Fan12].

Following [MOG11] we take $M = 20$ issuers, with default probability P_m of 1%, asset correlation ρ_m of 50% and exposure $\ell_m = 1$. For the COS method we take $N = 2^{10}$, $[a, b] = [0, \sum_{m=1}^{M} \ell_m]$. The characteristic function of L can be written as

$$\varphi(u) = \mathbb{E}[e^{iuL}] = \mathbb{E}\left[\mathbb{E}\left[e^{iu \sum_{m=1}^{M} \ell_m \mathbf{1}_{X_m < c_m}} | Y\right]\right]. \tag{4.11}$$

An analytic expression is available for the inner conditional expectation. In [Fan12] an integration rule to approximate the outer expectation is employed, whereas in [GRM06] Monte Carlo simulations on Y was used. We test both approaches, with a grid, $y = [-5 : 0.1 : 5]$, for the numerical integration (denoted by COS+NI) and by 5000 simulated values for Y in the Monte Carlo experiment (denoted by COS+MC). With these choices, the computation of the characteristic function, which is the most time-consuming part, is approximately 50 times more expensive for COS+MC.

The loss distribution of this example portfolio is plotted in Figure 4.6. The green line is the result of a full Monte Carlo simulation, with 100.000.000 replications for each X_m, and serves as our reference solution here. The red-dotted and magenta-dotted lines are the COS method approximations with numerical integration and Monte Carlo simulation for Y, respectively. The blue and cyan lines are the filtered-COS results, using an exponential filter with $p = 6$. The COS method without filtering shows a highly oscillatory behavior and does not give accurate results in the tail, where $1 - F_L(L)$ is very small. The COS method with filter *and* with numerical integration, gives however highly accurate results that correspond very well to the full Monte Carlo simulation. The difference gets smaller when the number of Monte Carlo simulations is increased. The results with the filter and Monte Carlo simulations for Y, COS+MC, are sensitive to outliers of draws for Y and these results are not very satisfactory. We would also like to mention that approximations based on Haar wavelets [MOG11] give accurate portfolio loss VaR estimates as well.

Remark 4.1. *The model and computational technique can be extended to higher-dimensional systematic risk factors $Y = [Y_1, \ldots, Y_d]$. In that case, the computation of the outer expectation in Equation* (4.11) *can be performed by adaptive integration, as in [Hua09].*

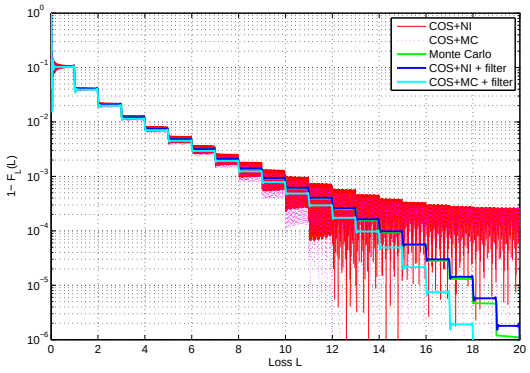


Figure 4.6: Recovery of portfolio loss distribution function using COS method and filtered version (exponential filter, order $p = 6$).

5 Conclusions

The COS method is an option pricing method based on Fourier-cosine expansions which performs very well in general. When the underlying density function is smooth, we can achieve an exponential convergence in the number of cosine coefficients. When the underlying density is not smooth, however, the method may suffer from the Gibbs phenomenon and the convergence is only of algebraic order. A filtering technique to improve the convergence rate for these cases has been discussed in the present paper. In practical cases where the COS method degrades due to discontinuities in functions, the convergence with a filter improves significantly in terms of the number of required Fourier coefficients as well as in CPU time.

Non-adaptive spectral filtering takes place in the Fourier-domain and combines therefore very well with the COS option pricing formula, without adding significant computational costs. The Fourier coefficients are pre-multiplied by a decreasing filter $\hat{s}(n/N)$ so that they decay faster and so that the convergence rate away from a discontinuity is improved. The technique can be used in one-dimensional problems, but also in higher dimensions.

In the numerical examples we tested six different filters, ie, the Fejér filter, the Lanczos filter, the raised cosine filter, the exponential filter, the Vandeven filter and the Erfc-Log filter. Especially the exponential filter gave highly accurate results for step-wise functions.

The plain COS method for option pricing under the Variance Gamma asset price process resulted in algebraic convergence. Our filter-COS method improved the algebraic index of convergence, in particular for short time horizons. Moreover, for the computation of the option Greeks, that suffer from a lower convergence rate without filtering, spectral filters are highly beneficial.

As another example in finance, we discussed portfolio loss modeling. Discrete random variables then give rise to stepwise cumulative distribution functions. We derived a COS formula to recover the distribution, which, of course, also suffered from the Gibbs phenomenon and the resulting oscillations. The approximation was drastically improved by applying spectral filters.

Improved convergence comes without additional computational cost in these applications and the fact that the filtering is easy to implement, even in multiple dimensions, makes the filter-COS method a natural solution for some of the problems described.

Acknowledgment The authors wish to acknowledge F. Fang for sharing knowledge about the

use of the COS method to portfolio loss modeling.

Appendix

A Convergence partial sums Riemann zeta and Dirichlet eta function

The partial sum of terms $1/n^k$ converges to the Riemann zeta function ([BO99])

$$\sum_{n=1}^{N} 1/n^k \to \zeta_R(k), \quad \text{if } N \to \infty \quad (\text{Re}\{k\} > 1). \tag{A.1}$$

Here the series truncation error converges with order $\mathcal{O}(N^{1-k})$. The harmonic series, for $k = 1$, is divergent. The partial sum of alternating terms $(-1)^{n-1}1/n^k$ converges to the Dirichlet eta function ([Son03])

$$\sum_{n=1}^{N}(-1)^{n-1}1/n^k \to \eta_D(k), \quad \text{if } N \to \infty, \tag{A.2}$$

but the series truncation error converges with the higher order $\mathcal{O}(N^{-k})$. To derive this, we rewrite the partial sum as (for N even)

$$\begin{aligned}
\sum_{n=1}^{N}(-1)^{n-1}1/n^k &= \sum_{n=1,n \text{ odd}}^{N}(-1)^{n-1}\frac{1}{n^k} + (-1)^n\frac{1}{(n+1)^k} \\
&= \sum_{n=1,n \text{ odd}}^{N}\frac{1}{n^k} - \frac{1}{(n+1)^k} \\
&= \sum_{m=1}^{N/2}\frac{1}{(2m-1)^k} - \frac{1}{(2m)^k}.
\end{aligned} \tag{A.3}$$

The terms $\frac{1}{(2m-1)^k} - \frac{1}{(2m)^k}$ are positive. We find

$$\begin{aligned}
\sum_{n=1}^{N}(-1)^{n-1}1/n^k &= \sum_{m=1}^{N/2}\frac{(2m)^k - (2m-1)^k}{(2m-1)^k(2m)^k} \\
&= \sum_{m=1}^{N/2}\frac{1}{(2m)^k}\frac{(2m)^k - \left[\sum_{j=0}^{k}\binom{k}{j}(2m)^{k-j}(-1)^j\right]}{\sum_{j=0}^{k}\binom{k}{j}(2m)^{k-j}(-1)^j}.
\end{aligned} \tag{A.4}$$

The truncation error then reads

$$\begin{aligned}
\sum_{n=N+1}^{\infty}(-1)^{n-1}1/n^k &= \sum_{m=N/2+1}^{\infty}\frac{1}{(2m)^k}\frac{k(2m)^{k-1} - \binom{k}{2}(2m)^{k-2} + \ldots}{(2m)^k - k(2m)^{k-1} + \binom{k}{2}(2m)^{k-2} - \ldots} \\
&= \sum_{m=N/2+1}^{\infty}\frac{1}{(2m)^k}\frac{k(2m)^{k-1}}{(2m)^k}, \quad \text{for } m \to \infty \\
&= \sum_{m=N/2+1}^{\infty}\frac{k}{(2m)^{k+1}}.
\end{aligned} \tag{A.5}$$

So, the truncation error converges with order $\mathcal{O}(N^{-k})$ by Theorem 3.1.

References

[BO99] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*. Advanced Mathematical Methods for Scientists and Engineers. McGraw-Hill, 1999.

[Boy96] J. P. Boyd. The Erfc-Log filter and the asymptotics of the Euler and Vandeven sequence accelerations. In A. V. Ilin and L. R. Scott, editors, *Proceedings of the Third International Conference on Spectral and High Order Methods*, pages 267–276. Houston Journal of Mathematics, Houston, 1996.

[Boy01] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Books on Mathematics Series. Dover Publications, New York, 2001.

[CM99] P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2, 1999.

[COS01] T. F. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denoising. *IEEE Transactions on Image Processing*, 10(2):231–241, 2001.

[DPS00] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.

[Fan12] F. Fang. Use COS method to calculate portfolio credit loss. Technical report, Working paper, 2012.

[Fej00] L. Fejér. *Sur les fonctions bornées et intégrables*. 1900.

[FO08] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.

[FO09] F. Fang and C. W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik*, 114(2), 2009.

[Gel00] A. Gelb. A hybrid approach to spectral reconstruction of piecewise smooth functions. *Journal of Scientific Computing*, 15(3):293–322, 2000.

[GRM06] P. Glasserman and J. Ruiz-Mata. Computing the credit loss distribution in the gaussian copula model: a comparison of methods. *Journal of credit risk*, 2(4):33–66, 2006.

[GS97] D. Gottlieb and C.-W. Shu. On the Gibbs phenomenon and its resolution. *SIAM Review*, 39(4):644–668, 1997.

[Gut05] A. Gut. *Probability: A Graduate Course*. Springer Texts in Statistics. Springer Science+Business Media, Incorporated, 2005.

[HGG07] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral Methods for Time-Dependent Problems*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2007.

[Hua09] X. Huang. *Credit Portfolio Losses*. PhD thesis, Delft University of Technology, 2009.

[Lan56] C. Lanczos. *Applied analysis*. Prentice-Hall mathematics series. Prentice-Hall, 1956.

[LK07] R. Lord and C. Kahl. Optimal Fourier inversion in semi-analytical option pricing. *Journal of Computational Finance*, 10(4), 2007.

[MCC98] D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[MOG11] J. J. Masdemont and L. Ortiz-Gracia. Haar wavelet-based approach for quantifying credit portfolio losses. *Quantitative Finance*, iFirst, 2011.

[Pey02] R. Peyret. *Spectral Methods for Incompressible Viscous Flow*. Number v. 148 in Applied Mathematical Sciences. Springer, 2002.

[RO12] M. J. Ruijter and C. W. Oosterlee. Two-dimensional Fourier cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing*, 34(5), 2012.

[Sar06] S. A. Sarra. Digital total variation filtering as postprocessing for Chebyshev pseudospectral methods for conservation laws. *Numerical Algorithms*, 41(1):17–33, 2006.

[Sch03] W. Schoutens. *Lévy processes in finance: pricing financial derivatives*. Wiley Series in Probability and Statistics. Wiley, 2003.

[Son03] J. Sondow. Zeros of the alternating zeta function on the line re(s)= 1. *The American Mathematical Monthly*, 110(5):435–437, 2003.

[SS03] E. M. Stein and R. Shakarchi. *Fourier analysis: an introduction*. Princeton Lectures in Analysis. Princeton University Press, 2003.

[Tad07] E. Tadmor. Filters, mollifiers and the computation of the Gibbs phenomenon. *Acta Numerica*, 16:305–378, 2007.

[Tan06] J. Tanner. Optimal filter and mollifier for piecewise smooth spectral data. *Mathematics of computation*, 75(254):767–790, 2006.

[TT05] E. Tadmor and J. Tanner. Adaptive filters for piecewise smooth spectral data. *IMA journal of numerical analysis*, 25(4):635–647, 2005.

[Van91] H. Vandeven. Family of spectral filters for discontinuous problems. *Journal of Scientific Computing*, 6(2):159–192, 1991.

[Vas02] O. Vasicek. The distribution of loan portfolio value. *Risk*, 15(12):160–162, 2002.

[vSRV11] P. van Slingerland, J.K. Ryan, and C. Vuik. Position-dependent smoothness-increasing accuracy-conserving (SIAC) filtering for improving discontinuous Galerkin solutions. *SIAM Journal on Scientific Computing*, 33(2):802–825, 2011.

[ZO13] B. Zhang and C. W. Oosterlee. Efficient pricing of Asian options under Lévy processes based on Fourier cosine expansions. *SIAM Journal on Financial Mathematics*, to appear in 2013.