Efficient VaR and Expected Shortfall computations for nonlinear portfolios within the delta-gamma approach

Luis Ortiz-Gracia [a,b,*], Cornelis W. Oosterlee [b,c]

[a] Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, 08193 Bellaterra (Barcelona), Spain
[b] CWI – Centrum Wiskunde & Informatica, NL-1090 GB Amsterdam, The Netherlands
[c] Delft University of Technology, Delft Institute of Applied Mathematics, 2628 CD Delft, The Netherlands

A B S T R A C T

We present four numerical methods to compute the Value-at-Risk and Expected Shortfall risk measure values of portfolios with financial options. The numerical methods are based on either wavelets or Fourier cosine approximations and belong to the class of Fourier inversion methods. We show that the risk measures can be efficiently calculated in terms of accuracy and CPU time. Besides, we provide a theoretical result about the shape of the resulting probability density. This *a priori* knowledge, allows us to enhance the efficiency and effectiveness of the proposed methods. Finally, we assess the accuracy of the approach in the presence of convexity or concavity properties of the financial portfolios.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

A problem of paramount importance in market risk management is the estimation of a profit and loss distribution of a portfolio over a specified time horizon and the associated risk measures. Value-at-Risk (VaR) has become an important measure for estimating and managing portfolio market risk [10]. VaR is defined as a certain quantile of the change in value of a portfolio during a specified holding period. While the basic concept of VaR is simple, many complications can arise in practical use. An important complication is caused by nonlinearity in the portfolio payoff structure. This problem arises for all portfolios that include assets with nonlinear payoffs, such as option positions. For such nonlinear portfolios, VaR can not be computed directly from a risk factor distribution. Instead, the risk factor distribution first needs to be converted into a profit and loss distribution for the portfolio. VaR is then computed from this profit and loss distribution.

The four defining properties of a coherent risk measure are widely treated in [1]. One of these properties is the sub-additivity condition. The VaR measure fails to satisfy this condition, however the measure is widely used in practice. In contrast, the Expected Shortfall (ES) risk measure satisfies the four properties of a coherent risk measure. When distributions are normally, or close to normally, distributed, it can be shown that VaR and ES are quite close and behave similarly. However, as soon as a distribution is characterized by a long tail behavior, the similarity between VaR and ES does not hold any more. In this case, employing the VaR measure may lead to an underestimation of risk.

To evaluate the risk measures, Monte Carlo simulation is often used, i.e. first simulating changes in the risk factors of a portfolio, then the portfolio is re-evaluated and the change in the portfolio's value is estimated. Obtaining accurate VaR estimates can be computationally expensive as there may be a large number of instruments in the portfolio and when the confidence level is high, a large number of simulations may be required to obtain accurate estimates of the tail probability.

Here, we adopt the *delta-gamma* approximation [3,7,13] as an alternative approach to avoid the Monte Carlo method. This approach is based on the assumption that the change in portfolio value is a quadratic function of the changes in the risk factors. Typically, the changes in the risk factors are assumed to be normally distributed, although some authors [8] have considered heavy-tailed risk factors and modeled them by means of the multivariate *t* distribution. This distribution can be used under stress test scenarios, where peak losses can occur more frequently.

Under any of the two assumptions on the changes in the risk factors, the *characteristic function* of the change in portfolio value, i.e. the Fourier transform of the corresponding density, is known in closed form. Fourier techniques to invert this characteristic function in a market risk context have been introduced since the work of [16,5], and were used more recently by [18,2]. In all these works, one has to compute an infinite integral numerically for each point of the probability density function (PDF) (or cumulative distribution function (CDF)).

Our contributions within the delta-gamma approach in the present paper are as follows. On the one hand, we present four numerical methods to efficiently recover the PDF and the CDF from the characteristic function. On the other hand, we derive a useful result about the shape of the density function to be recovered and another result about the maximum profit and loss in our portfolio. We also provide insight in the convexity and concavity features of the portfolio. By this, the density function can be characterized in terms of its unimodal or bimodal behavior. The numerical methods rely on the truncation of the entire real line to perform Fourier inversion. With the *a priori* knowledge of the density shape, this inversion can be done more efficiently. Furthermore, since the delta-gamma approach is a quadratic approximation, convex or concave portfolios are accurately approximated.

Our approach to invert the Fourier transform within the delta-gamma framework is somewhat different than previous approaches in the market risk literature. We consider the COS method [6], based on Fourier cosine expansions, and the Wavelet Approximation (WA) method [11,14,15]. Within the COS approach, we recover the PDF without performing any numerical integration (for the cases that we consider here), so that the speed of the method is impressive. The analytic expression that we obtain allows us to integrate (analytically) the PDF to derive the CDF. The VaR is then computed using a root-finding method and we also provide a formula for the ES. We also consider a second variant of the COS method, called filtered-COS [17], which is sometimes used to remove wiggles associated to Gibbs phenomena that may arise when approximating some functions with discontinuities.

Furthermore, we employ an approach based on wavelets theory (see, for instance, [4]). Regarding the WA, we apply two different approximations, presented in [15], the WA$^{[a,b]}$ and the WA$^{\mathbb{R}}$ methods, where the second one adaptively determines the range of truncation for Fourier inversion. We provide a methodology for computing the VaR and ES from the PDF and also from the CDF in the Wavelet Approximation framework. While in the COS method the coefficients of the density are computed without numerical integration, within the WA approach the density coefficients are recovered by means of finite integrals that are computed numerically. The main advantage of the WA approach, based on compactly supported basis functions, compared to the COS approach, with its global basis functions, is that coefficients can be selectively calculated to compute the VaR value, making the wavelet algorithm very fast. As shown in [11], the VaR can be obtained by computing at most *m* coefficients associated to the CDF, where *m* is the scale of approximation. Summarizing, we present two very accurate and fast approaches based on Fourier inversion that can be combined as well (computing the VaR by the WA method and the ES by the COS method, for example).

The paper is organized as follows. In Section 2 we present the delta-gamma approximation to measure market risk measures for portfolios with financial options. In Section 3 we review the Wavelet Approximation and the COS methods to recover a function from its Fourier transform. In Section 4 we present the methodology to compute the VaR and the ES by means of the inversion methods considered in the earlier section. Furthermore, we study in detail the shape of density functions and some features about convex portfolios. Section 5 is devoted to numerical experiments, and Section 6 concludes.

2. The delta-gamma approximation

Let $F_X(x) := \mathbb{P}(X \leqslant x)$ be the CDF of a random variable X and $f_X(x)$ its PDF.

Suppose the current value of a portfolio is $V(t)$, the holding period is Δt, and the value of the portfolio at time $t + \Delta t$ is $V(t + \Delta t)$. The change in the portfolio value during the holding period is ΔV, where $\Delta V = V(t + \Delta t) - V(t)$. If we hold a short position on the underlying assets, the VaR(α) risk measure, associated with a given probability α, is defined by the relation,

$$\mathbb{P}(\Delta V < \text{VaR}(\alpha)) = \alpha. \tag{1}$$

Typically, in practice, Δt ranges from one day to two weeks and $\alpha \geqslant 0.95$, often $\alpha = 0.99$.

By definition, the Expected Shortfall risk measure at confidence level α is given by,

$$\text{ES}(\alpha) := \mathbb{E}(\Delta V | \Delta V > \text{VaR}(\alpha)),$$

or, alternatively,

$$\text{ES}(\alpha) := \frac{1}{1 - \alpha} \int_{\text{VaR}(\alpha)}^{+\infty} x f_{\Delta V}(x) dx. \tag{2}$$

We assume that there are p risk factors and that $S(t) = (S_1(t), \ldots, S_p(t))^T$ denotes the value of these factors at time t. Define $\Delta S = (S(t + \Delta t) - S(t))$ to be the change in the risk factors during the interval $[t, t + \Delta t]$. Then, the *delta-gamma approximation* is given by,

$$\Delta V \simeq \Delta V_\gamma := \Theta \Delta t + \delta^T \Delta S + \frac{1}{2} \Delta S^T \Gamma \Delta S, \tag{3}$$

where $\Theta = \frac{\partial V}{\partial t}$, $\delta_i = \frac{\partial V}{\partial S_i}$, $\Gamma_{i,j} = \frac{\partial^2 V}{\partial S_i \partial S_j}$ and all partial derivatives are being evaluated at $S(t)$.[1]

If $\mathbb{P}(\Delta V_\gamma - \Theta \Delta t \leqslant x) = \alpha$, then $\mathbb{P}(\Delta V_\gamma \leqslant x + \Theta \Delta t) = \alpha$, where $\alpha \in (0, 1)$. For convenience, we define, $\widetilde{\Delta V}_\gamma := \Delta V_\gamma - \Theta \Delta t$. The following proposition from [12] gives us the characteristic function of $f_{\widetilde{\Delta V}_\gamma}$, under the assumption that $\Delta S \sim \mathcal{N}(0, \Sigma)$.

Proposition 1. *Assume that $\Delta S \sim \mathcal{N}(0, \Sigma)$ for some positive definite matrix Σ. Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of $\Sigma\Gamma$, and let Λ be the diagonal matrix with these eigenvalues on the diagonal. There is a matrix C satisfying $CC^T = \Sigma$ and $C^T \Gamma C = \Lambda$. Let $d = C^T \delta$. Then, the characteristic function corresponding to $f_{\widetilde{\Delta V}_\gamma}$ is given by,*

$$\widehat{f}_{\widetilde{\Delta V}_\gamma}(u) = \mathbb{E}\left(e^{-iu\widetilde{\Delta V}_\gamma}\right) = \exp\left(-\frac{u^2}{2} \sum_{j=1}^{p} \frac{d_j^2}{1 + i\lambda_j u}\right) \prod_{j=1}^{p} (1 + i\lambda_j u)^{-\frac{1}{2}}, \tag{4}$$

where $u \in \mathbb{R}$.

Remark 1. The Black–Scholes model assumes that $\log\left(\frac{S_j(t+\Delta t)}{S_j(t)}\right)$ is normally distributed with mean $\mu_j \Delta t$ and standard deviation $\sigma_j \sqrt{\Delta t}$, for $j = 1, \ldots, p$. Thus, there seems to be an inconsistency between the valuation model and the model used for path simulation. However, for small Δt (as the holding period is),

$$\frac{S_j(t + \Delta t)}{S_j(t)} = 1 + \frac{\Delta S_j}{S_j(t)} \simeq \exp\left(\frac{\Delta S_j}{S_j(t)}\right),$$

which is log-normally distributed if ΔS_j, the jth component of ΔS, is normally distributed. In that case, ΔS_j follows a normal distribution with mean $\mu_j \Delta t$ and standard deviation $S_j(t) \cdot \sigma_j \sqrt{\Delta t}$. We can approximate the mean by zero.

Remark 2. Observe that in the univariate case ($p = 1$), we have,

$$\Delta V_\gamma := \sum_{i=1}^{n} x_i \frac{\partial v_i}{\partial t} \Delta t + \sum_{i=1}^{n} x_i \frac{\partial v_i}{\partial S} \Delta S + \frac{1}{2} \sum_{i=1}^{n} x_i \frac{\partial^2 v_i}{\partial S^2} (\Delta S)^2,$$

where n represents the number of assets in the portfolio, x_i is the amount of asset i and v_i the value of asset i.

In this work we restrict ourselves to the univariate case. The methods presented in this work recover the density function from the Fourier transform and this transform is of dimension one even if we consider the multivariate case, i.e., $p > 1$.

For our numerical techniques insight into the *unimodality* of the distribution to be approximated is very useful.

Definition 1. A distribution with probability density function f is called *unimodal* if a unique $\mathcal{M}$ exists so that f is non-decreasing on $(-\infty, \mathcal{M})$ and non-increasing on $(\mathcal{M}, +\infty)$. The value $\mathcal{M}$ is called the *mode* of the distribution. A distribution is bimodal if it has only two modes.

Note that a distribution with non-increasing (respectively non-decreasing) PDF also falls under unimodal distribution by taking $\mathcal{M}$ to be the left (respectively right) end point of the support of the density function. Under this convention, it is possible for the density function to be infinite, or even undefined at $\mathcal{M}$. Note that the support of the considered distributions need not be finite.

3. Numerical inversion methods

In this section we present the WA$^{[a,b]}$ method [14] and the COS method [6] to recover the density function $f_{\widetilde{\Delta V}_\gamma}$ from its Fourier transform. Recall that based on Proposition 1, the characteristic function for the density $f_{\widetilde{\Delta V}_\gamma}$ is available. For the WA method, we present a second variant called the WA$^\mathbb{R}$ method [15], avoiding the a priori choice of an interval for the approximation. Regarding the COS method we also present another variant called filtered-COS [17]. This alternative method may be helpful when dealing with the so called Gibbs phenomenon.

3.1. The Wavelet Approximation method

3.1.1. The WA$^{[a,b]}$ method

Let us consider a probability density function $f \in L^2(\mathbb{R})$ associated to a certain continuous random variable X, and its Fourier transform, i.e.

---

[1] The delta approximation reads $\Delta V \simeq \Delta V_\delta := \Theta \Delta t + \delta^T \Delta S$.

$$\widehat{f}(w) = \int_{-\infty}^{+\infty} e^{-iwx} f(x) dx. \tag{5}$$

We can expect that the mass in the tails of f decays to zero at infinity, so it can be approximated in a finite interval $[a, b]$, by

$$f^c(x) = \begin{cases} f(x), & \text{if } x \in [a, b], \\ 0, & \text{otherwise}. \end{cases}$$

To determine the interval of integration $[a, b]$, we consider the approximation,

$$[a, b] := \left[\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}\right], \tag{6}$$

where κ_n denotes the nth cumulant[2] of X and $L > 0$, as in [6,15]. Later on, in the numerical examples section, we will give further details about the choice of L and of $[a, b]$.

Let us approximate $f^c(x) \simeq f_m^c(x)$ for all $x \in [a, b]$, where

$$f_m^c(x) = \sum_{k=0}^{2^m - 1} c_{m,k} \phi_{m,k}\left(\frac{x-a}{b-a}\right),$$

with convergence in L^2-norm.

The basis functions used here are $\phi_{m,k} := 2^{m/2} \phi(2^m x - k)$, with ϕ the father Haar wavelet, defined by,

$$\phi(x) = \chi_{[0,1)}(x) = \begin{cases} 1, & \text{if } x \in [0, 1), \\ 0, & \text{otherwise}. \end{cases}$$

The main idea behind the Wavelet Approximation method is to approximate $\widehat{f}$ by $\widehat{f}_m^c$ and then compute the coefficients $c_{m,k}$ by inverting the Fourier transform. Proceeding this way, we have [14],

$$c_{m,0} \simeq \frac{1}{\pi} \int_0^\pi \Re(Q_m(re^{iu})) du, \tag{7}$$

and,

$$c_{m,k} \simeq \frac{2}{\pi r^k} \int_0^\pi \Re(Q_m(re^{iu})) \cos(ku) du, \quad k = 1, \ldots, 2^m - 1, \tag{8}$$

where $r \neq 1$ is a positive real number and,

$$Q_m(z) = \frac{2^{\frac{m}{2}} z^{-\frac{2^m a}{b-a}} \widehat{f}\left(\frac{2^m}{b-a} i \cdot \log(z)\right) \log(z)}{(b-a)(z-1)}.$$

In practice, both integrals in (7) and (8) can be easily computed by means of the Trapezoidal Rule (see [14] for details).

3.1.2. The WA$^{\mathbb{R}}$ method

The main drawback of the WA$^{[a,b]}$ method is that we do not have an estimate of the mass of the density which is lost when truncating the interval. Here, we also discuss an *adaptive method* that allows us to control the mass of the density recovered.

Let f be a probability density function, as in Section 3.1.1. We can approximate $f(x)$ by $f_m(x)$, for all $x \in \mathbb{R}$, where,

$$f_m(x) = \sum_{k \in \mathbb{Z}} c_{m,k} \phi_{m,k}(x), \quad j \geqslant 0,$$

with convergence in the L^2-norm. Note that the coefficients $c_{m,k}$ are different from those in Section 3.1.1. For convenience, we however keep the same notation as in the previous section.

Let us consider the finite sum,

$$f_m^{tr}(z) = \sum_{k=k_1}^{k_2} c_{m,k} \phi_{m,k}(x), \tag{9}$$

where $k_1 < k_2, k_1, k_2 \in \mathbb{Z}$.

Substituting (9) into Fourier transform expression (5), following similar steps as in the previous section and after some algebraic manipulations, we get,

$$c_{m,k_1} \simeq \frac{1}{\pi} \int_0^\pi \Re(Q_m(re^{iu})) du, \tag{10}$$

and,

---

[2] The cumulants are the power series coefficients of the cumulant generating function $\kappa(s) = \log \mathbb{E}(e^{sX})$.

$$c_{m,k_1+k} \simeq \frac{2}{\pi r^k} \int_0^\pi \Re(Q_m(re^{iu})) \cos(ku) du, \quad k = 1, \ldots, k_2 - k_1,$$

where $Q_m(z) = z^{-k_1} \cdot Q_m^*(z)$ with,

$$Q_m^*(z) := \frac{2^{\frac{m}{2}} \widehat{f}(2^m i \cdot \log(z)) \log(z)}{(z-1)}.$$

We need to choose integers k_1 and k_2 to recover the density function without loosing significant accuracy.

To select an appropriate k_1, let us assume first that the density is unimodal (Definition 1). In this case k_1 can be chosen such that, $f_m\left(\frac{k_1}{2^m}\right) < \epsilon_{tol}$, where ϵ_{tol} is a predefined tolerance error so that $f_m(x) < \epsilon_{tol}$ for all $x \leqslant \frac{k_1}{2^m}$.

We can then start the WA$^\mathbb{R}$ algorithm by considering an initial seed, $k_1 = \lfloor 2^m \cdot a \rfloor$, where a is defined in expression (6) and $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. Although we rely on the cumulants to facilitate the work, we could choose a random value as the initial seed. We update value k_1 until the condition $f_m\left(\frac{k_1}{2^m}\right) < \epsilon_{tol}$ is satisfied. The algorithm allows us to immediately calculate the area below the computed density (as a by-product) and we can derive k_2 by computing the coefficients c_{m,k_1+k} until the area is approximately one. In case we can not assume unimodality, we recompute k_1 (if necessary) to go further towards the left side.

3.2. The COS and filtered-COS methods

3.2.1. The COS method

We briefly describe the methodology developed in [6] for solving an inverse Fourier integral.[3] For a function f supported on a finite interval $[a, b] \in \mathbb{R}$, the Fourier-cosine series expansion reads,

$$f(x) = \frac{A_0}{2} + \sum_{k=1}^{+\infty} A_k \cos\left(k\pi \frac{x-a}{b-a}\right),$$

with,

$$A_k = \frac{2}{b-a} \int_a^b f(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx. \tag{11}$$

Since any real function has a cosine expansion when it is finitely supported, the derivation starts with a truncation of the infinite integration range in the inverse Fourier integral expression (5). Due to the conditions for the existence of a Fourier transform, the integrands have to decay to zero at $\pm\infty$ and we can truncate the integration range in a proper way without losing accuracy.

Suppose $[a, b] \in \mathbb{R}$ is chosen such that the truncated integral approximates the infinite counterpart very well, i.e.,

$$\widehat{f}_1^+(w) := \int_a^b e^{iwx} f(x) dx \simeq \int_{\mathbb{R}} e^{iwx} f(x) dx = \widehat{f}^+(w). \tag{12}$$

Here, $\widehat{f}_1^+$ denotes the approximation of the characteristic function on a finite interval.

Comparing equation (12) with the cosine series coefficients of $f(x)$ on $[a, b]$ in (11), we see that,

$$A_k \equiv \frac{2}{b-a} \Re\left(\widehat{f}_1^+\left(\frac{k\pi}{b-a}\right) e^{-i\frac{ka\pi}{b-a}}\right),$$

where $\Re$ denotes the real part of the argument. It then follows from (12) that $A_k \simeq \widetilde{A}_k$ with,

$$\widetilde{A}_k \equiv \frac{2}{b-a} \Re\left(\widehat{f}^+\left(\frac{k\pi}{b-a}\right) e^{-i\frac{ka\pi}{b-a}}\right).$$

In the COS method A_k is replaced by $\widetilde{A}_k$ in the series expansion of $f(x)$ on $[a, b]$, and the series summation is truncated, so that,

$$f_1(x) = \frac{\widetilde{A}_0}{2} + \sum_{k=1}^{N-1} \widetilde{A}_k \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{13}$$

The COS method converges exponentially in N when approximating smooth functions, but many terms are needed when the function or its first derivative presents discontinuities in the domain of approximation.

---

[3] Here,

$$\widehat{f}^+(w) = \int_{\mathbb{R}} e^{iwx} f(x) dx,$$

represents the characteristic function, and hence the Fourier transform of a density function $f(x)$, because the sign of the exponent is different in the definition of the Fourier transform, compared to (5).

3.2.2. The filtered-COS method

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. Although the limit of the partial sums represents the original function exactly, in the finite case there is an overshoot at a jump discontinuity. In the case of a jump discontinuity we may have pointwise convergence almost everywhere, but no uniform convergence. 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.

Filtering may remove the Gibbs phenomenon away from a discontinuity, and the error depends on the distance to the discontinuity. Since the approximation will be smoothened, convergence in the vicinity of a discontinuity will not improve. The technique 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 the coefficients decay faster. Here we provide the definitions of filters of order p_f, [9],

Definition 2 (*Fourier space filter of order p_f*). A real and even $C^\infty([0,1])$ function $\hat{s}(\eta)$ is called a filter of order p_f, if,

1. $\hat{s}(0) = 1$ and $\hat{s}^{(\ell)}(0) = 0$, $1 \leqslant \ell \leqslant p_f - 1$,
2. $\hat{s}(\eta) = 0$ for $|\eta| \geqslant 1$,
3. $\hat{s}(\eta) \in C^{p_f-1}$, $\eta \in (-\infty, \infty)$.

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

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. We will employ a p_fth order spectral filter called the *exponential filter*, which is defined as follows,

$$\hat{s}(\eta) = \exp(-\alpha\eta^{p_f}),$$

where p_f must be even. $\hat{s}(1) = e^{-\alpha}$, so the formal requirements of a p_fth order filter do not hold. However, we use $\alpha = -\log \epsilon_m$, as in [9], where ϵ_m represents the machine epsilon, so that $\hat{s}(1) = \epsilon_m \simeq 0$ within machine precision.

The filtered-COS formula for numerical Fourier inversion [17], equivalent to (13), reads,

$$f_1^{\text{filter}}(x) = \frac{\widetilde{A}_0}{2}\hat{s}(0) + \sum_{k=1}^{N-1} \hat{s}\left(\frac{k}{N-1}\right)\widetilde{A}_k \cos\left(k\pi\frac{x-a}{b-a}\right).$$

For each of the four numerical methods discussed, $WA^{[a,b]}$, $WA^{\mathbb{R}}$, COS and filtered-COS, we have provided criteria to truncate the entire real line to perform the numerical inversion in case that we do not have any knowledge about the shape of the density to be recovered. However, as we will show in the next section by a theoretical study of the density's shape, these algorithms can be enhanced.

4. Value-at-Risk and Expected Shortfall computation

This section is devoted to the computation of the VaR value (1) and the Expected Shortfall (2) by means of the WA and COS methods, within the delta-gamma approximation. For sake of clarity and simplicity, we focus on the $WA^{[a,b]}$ and the COS methods and omit details for $WA^{\mathbb{R}}$ and filtered-COS, since minor modifications in the first two methods lead to these methods.

4.1. Value-at-Risk

We present two variants for computing the VaR value, where one of the variants is based on the probability density function $f_{\widetilde{\Delta V}_\gamma}$ and the other is based on the cumulative distribution function $F_{\widetilde{\Delta V}_\gamma}$. The first variant recovers the density from the characteristic function and then integrates the density for obtaining the VaR. The second directly approximates the CDF. For the COS method, both choices lead to an accurate VaR approximation in similar CPU time, so we present only the approximation based on the PDF. However, we will show in the numerical experiments, that $WA^{[a,b]}$ is more efficient for the second variant.

Let us consider the characteristic function $\widehat{f}_{\widetilde{\Delta V}_\gamma}$ corresponding to the density function $f_{\widetilde{\Delta V}_\gamma}$. Since $\widehat{f}_{\widetilde{\Delta V}_\gamma} \in L^2(\mathbb{R})$ also $f_{\widetilde{\Delta V}_\gamma} \in L^2(\mathbb{R})$, so that $f_{\widetilde{\Delta V}_\gamma}$ can be well approximated in a finite interval $[a,b]$. Taking this into account and integrating by parts, we can consider,

$$\widehat{F}_{\widetilde{\Delta V}_\gamma}(u) := \frac{\widehat{f}_{\widetilde{\Delta V}_\gamma}(u) - e^{-ibu}}{iu},$$

as an approximation to the characteristic function of CDF $F_{\widetilde{\Delta V}_\gamma}$. Once this characteristic function is available, we can recover the CDF directly from it.

4.1.1. VaR computation based on the PDF

We first assume that we have recovered a density function $f_{\widetilde{\Delta V}_\gamma}$ from its characteristic function $\widehat{f}_{\widetilde{\Delta V}_\gamma}$, by means of the $WA_m^{[a,b]}$ method, where m is the scale of approximation. Let f_γ^{wa} be the recovered density (for clarity in the exposition, we drop the subscript $\widetilde{\Delta V}_\gamma$ and use simply γ instead). That is,

$$f_\gamma^{\text{wa}}(x) = \sum_{k=0}^{2^m-1} c_{m,k} \phi_{m,k}\left(\frac{x-a}{b-a}\right).$$

Then, if we denote by F_γ the cumulative distribution function of $\widetilde{\Delta V}_\gamma$, we have,

$$F_\gamma(x) := \int_{-\infty}^{x} f_\gamma(y) dy \simeq F_\gamma^{\text{wa}}(x) := \int_{-\infty}^{x} f_\gamma^{\text{wa}}(y) dy.$$

Expanding the expression above, gives us,

$$F_\gamma^{\text{wa}}(x) = \sum_{k:\, x \notin \text{supp }\phi_{m,k}} c_{m,k} \int_{a_k}^{b_k} \phi_{m,k}\left(\frac{x-a}{b-a}\right) dx + \sum_{k:\, x \in \text{supp }\phi_{m,k}} c_{m,k} \int_{a_k}^{x} \phi_{m,k}\left(\frac{x-a}{b-a}\right) dx, \tag{14}$$

where $\text{supp }\phi_{m,k} = [a_k, b_k)$, with $a_k = a + \frac{b-a}{2^m} \cdot k$ and $b_k = a + \frac{b-a}{2^m} \cdot (k+1)$.

In this case the basis functions do not overlap in the interval of approximation. Then, the second sum in (14) consists only of one term and the expression for the cumulative function can be obtained in a straightforward way,

$$F_\gamma^{\text{wa}}(x) = \frac{b-a}{2^{\frac{m}{2}}} \sum_{k=0}^{\bar{k}-1} c_{m,k} + 2^{\frac{m}{2}} c_{m,\bar{k}}\left(x - a - \frac{\bar{k}(b-a)}{2^m}\right), \tag{15}$$

where $\bar{k}$ is the only value for which $x \in [a_{\bar{k}}, b_{\bar{k}})$.

Let us define $\widetilde{\text{VaR}}_m(\alpha)$ as the VaR value calculated by the $WA_m^{[a,b]}$ method at confidence level α.

Let k_s be the value of k so that $\frac{b-a}{2^{\frac{m}{2}}} \sum_{k=0}^{k_s} c_{m,k}$ is closest to α. Then, we take for the VaR value the midpoint of interval $[a_{k_s}, b_{k_s})$, that is,

$$\widetilde{\text{VaR}}_m(\alpha) = a + \frac{b-a}{2^{m+1}} \cdot (2k_s + 1). \tag{16}$$

Finally, the VaR value associated to the CDF of ΔV_γ reads,

$$\text{VaR}_m(\alpha) := \widetilde{\text{VaR}}_m(\alpha) + \Theta \Delta t. \tag{17}$$

Let now f_γ^{cos} be the recovered density function of $\widetilde{\Delta V}_\gamma$ by using the *COS method* with N terms, that is,

$$f_\gamma^{\text{cos}}(x) = \frac{\widetilde{A}_0}{2} + \sum_{k=1}^{N-1} \widetilde{A}_k \cos\left(k\pi \frac{x-a}{b-a}\right).$$

Following similar steps as before, and after some basic calculus, we find,

$$F_\gamma^{\text{cos}}(x) = \int_a^x f_\gamma^{\text{cos}}(x) dx = \frac{1}{2}\widetilde{A}_0(x-a) + \frac{b-a}{\pi} \sum_{k=1}^{N-1} \frac{\widetilde{A}_k}{k} \sin\left(k\pi \frac{x-a}{b-a}\right). \tag{18}$$

We define $\widetilde{\text{VaR}}_N(\alpha)$ as the VaR value calculated by the COS method with N terms at a confidence level α. We calculate this value by means of a root-finding technique that solves the equation,

$$F_\gamma^{\text{cos}}(\widetilde{\text{VaR}}_N(\alpha)) = \alpha,$$

followed by $\text{VaR}_N(\alpha) := \widetilde{\text{VaR}}_N(\alpha) + \Theta \Delta t$.

4.1.2. VaR computation from the CDF

We here assume that we have recovered the cumulative distribution function $F_{\widetilde{\Delta V}_\gamma}$ from its characteristic function $\Phi_{\widetilde{\Delta V}_\gamma}$, by means of the $WA_m^{[a,b]}$ *method*, with m the scale of approximation. Let F_γ^{wa} be the recovered CDF (for simplicity, we use the same notation as in the previous section), i.e.,

$$F_\gamma^{\text{wa}}(x) = \sum_{k=0}^{2^m-1} c_{m,k} \phi_{m,k}\left(\frac{x-a}{b-a}\right).$$

We can apply a bisection method so that $\widetilde{\text{VaR}}_m(\alpha)$ can be calculated in at most m iterations (see [11] for details), yielding,

$$\widetilde{\text{VaR}}_m(\alpha) = a + \frac{b-a}{2^{m+1}} \cdot (2k_s + 1),$$

for certain $k_s \in \{0, \ldots, 2^m - 1\}$, and the VaR value associated to the CDF of ΔV_γ is found as in (17).

4.2. Expected Shortfall

Also here we distinguish the case where ES is based on the PDF from the case where ES is computed based on the CDF. For the COS method we present only the approach based on the PDF, as in the previous section for VaR computation.

4.2.1. ES computation from the PDF
Let $ES_m(\alpha)$ be the Expected Shortfall computed by means of the $WA_m^{[a,b]}$ method, that is,

$$ES(\alpha) \simeq ES_m(\alpha) := \frac{1}{1-\alpha} \int_{VaR_m(\alpha)}^{+\infty} x f_\gamma^{wa}(x - \Theta\Delta t) dx.$$

$$
\begin{aligned}
ES_m(\alpha) &= \frac{1}{1-\alpha} \left(2^{\frac{m}{2}} c_{m,k_s} \int_{VaR_m(\alpha)}^{b_{k_s}+\Theta\Delta t} x dx + 2^{\frac{m}{2}} \sum_{k=k_s+1}^{2^m-1} c_{m,k} \int_{a_k+\Theta\Delta t}^{b_k+\Theta\Delta t} x dx \right) \\
&= \frac{2^{\frac{m-2}{2}}}{1-\alpha} \left(c_{m,k_s} \left((b_{k_s}+\Theta\Delta t)^2 - (VaR_m(\alpha))^2 \right) + \sum_{k=k_s+1}^{2^m-1} c_{m,k} ((b_k+\Theta\Delta t)^2 - (a_k+\Theta\Delta t)^2) \right).
\end{aligned}
$$

We define $ES_N(\alpha)$ as the Expected Shortfall calculated by the *COS method* with N terms, as

$$
\begin{aligned}
ES(\alpha) \simeq ES_N(\alpha) &:= \frac{1}{1-\alpha} \int_{VaR_N(\alpha)}^{+\infty} x f_\gamma^{cos}(x - \Theta\Delta t) dx = \frac{1}{1-\alpha} \int_{VaR_N(\alpha)}^{b+\Theta\Delta t} x \left(\frac{\widetilde{A}_0}{2} + \sum_{k=1}^{N-1} \widetilde{A}_k \cos\left(k\pi \frac{x - \Theta\Delta t - a}{b-a} \right) \right) dx \\
&= \frac{1}{1-\alpha} \left\{ \frac{\widetilde{A}_0}{4} \left((b+\Theta\Delta t)^2 - (VaR_N(\alpha))^2 \right) + \sum_{k=1}^{N-1} \widetilde{A}_k \left[-\frac{b-a}{k\pi} VaR_N(\alpha) \sin\left(k\pi \frac{VaR_N(\alpha) - \Theta\Delta t - a}{b-a} \right) \right. \right. \\
&\quad \left. \left. + \frac{(b-a)^2}{(k\pi)^2} \left((-1)^k - \cos\left(k\pi \frac{VaR_N(\alpha) - \Theta\Delta t - a}{b-a} \right) \right) \right] \right\}.
\end{aligned}
$$

4.2.2. ES computation from the CDF
If we us consider ES definition (2), and integrate by parts, then,

$$ES(\alpha) = \frac{1}{1-\alpha} \left[x F_{\Delta V}(x)\big|_{VaR(\alpha)}^{+\infty} - \int_{VaR(\alpha)}^{+\infty} F_{\Delta V}(x) dx \right],$$

where $F_{\Delta V}$ is the CDF associated to ΔV.

We define $F_{\Delta V_\gamma}^{wa}(x) := F_\gamma^{wa}(x - \Theta\Delta t)$, where F_γ^{wa} is the CDF calculated in Section 4.1.2 and approximate $F_{\Delta V}$ by F_γ^{wa},

$$ES(\alpha) \simeq ES_m(\alpha) := \frac{1}{1-\alpha} \left[b + \Theta\Delta t - \alpha VaR_m(\alpha) - \frac{b-a}{2^{\frac{m}{2}+1}} c_{m,k_s} - \frac{b-a}{2^{\frac{m}{2}}} \sum_{k=k_s+1}^{2^m-1} c_{m,k} \right].$$

4.3. Suitability of the delta-gamma approximation

We study the suitability of the delta-gamma approach for portfolios of options with several underlying assets. As pointed out by [3], in the one-dimensional case, when the portfolio is convex or concave over the likely range of prices of the underlying then the portfolio value can be accurately approximated by a quadratic function of the price of the underlying and, consequently, the delta-gamma approach performs well. For many option positions, however, the portfolio value may have convex as well as concave regions in the range of likely prices of the underlying. In this case, a quadratic approximation may not provide a satisfactory fit and the delta-gamma approximation may be less reliable.

Let us consider a portfolio $V(S_1(t), \ldots, S_p(t))$, as in Section 2, where the changes in the risk factors are normally distributed. Let D be an open convex set in $\mathbb{R}^p$. Following the theory of convex (concave) functions, if we assume that V is twice continuously differentiable, then V is convex (concave) if and only if $Hess(V)$ is positive (negative) semi-definite in D, with $Hess(V)$ the Hessian associated to V. On the other hand, quadratic form (3) is convex (concave) if and only if Γ is positive (negative) semi-definite. Since $Hess(V) = \Gamma$, we can expect that the delta-gamma approach approximates accurately in those cases in which V is a convex (or concave) portfolio.

The natural questions arising are: what happens when V (and hence the quadratic form) is indefinite? Can we expect accurate VaR and ES values when using the delta-gamma approach in these cases? To answer these questions let us assume that V can be decomposed into a sum of one-dimensional subportfolios,

$$V(S_1(t), \ldots, S_p(t)) = V_1(S_1) + \cdots + V_p(S_p), \tag{19}$$

where each $V_j(S_j)$ is either convex or concave. Observe that if our main portfolio V contains options written on one asset only, decomposition (19) in one-dimensional subportfolios is always possible. The assumption here is regarding the convexity or

concavity. Let us also assume that the changes in the risk factors are *uncorrelated*. We can then decompose the p-dimensional delta-gamma approach (3) as,

$$\Delta V_\gamma = \sum_{j=1}^{p} \Delta V_\gamma^j,$$

where $\Delta V_\gamma^j := \Theta^j \Delta t + \delta^j \Delta S_j + \frac{1}{2}\Gamma^j \Delta S_j^2$, and $\Theta^j = \frac{\partial V_j}{\partial t}, \delta^j = \frac{\partial V_j}{\partial S_j}, \Gamma^j = \frac{\partial^2 V_j}{\partial S_j^2}$. All partial derivatives are evaluated at $S(t)$.

Since $\Delta S = (\Delta S_1, \ldots, \Delta S_p) \sim \mathcal{N}(0, \Sigma)$ and $\Delta S_j, j = 1, \ldots, p$ are uncorrelated, then $\Delta S_j, j = 1, \ldots, p$, are independent. As a consequence, $\Delta V_\gamma^j, j = 1, \ldots, p$, are independent, and

$$f_{\Delta V_\gamma} = f_{\Delta V_\gamma^1} * \ldots * f_{\Delta V_\gamma^p}.$$

We assume that all subportfolios are either convex or concave and then $f_{\Delta V_j}$ can be accurately approximated by $f_{\Delta V_\gamma^j}$, for all $j = 1, \ldots, p$. So,

$$f_{\Delta V_\gamma} = f_{\Delta V_\gamma^1} * \ldots * f_{\Delta V_\gamma^p} \simeq f_{\Delta V_1} * \ldots * f_{\Delta V_p} = f_{\Delta V},$$

where the last equality holds because $S_j, j = 1, \ldots, p$ are independent. We illustrate this fact in the numerical examples section.

Finally, we can conclude that if the one-dimensional subportfolios are convex/concave and the changes in the risk factors are uncorrelated, then the risk measures associated to the main portfolio can be appropriately computed by means of a p-dimensional delta-gamma approach.

4.4. Monotonic features of the delta-gamma density

It is in general a difficult task to check whether a density function is unimodal or not directly from its characteristic function. However, within the delta-gamma approach, the characteristic function $\hat{f}_{\widetilde{\Delta V}_\gamma}$ of density $f_{\widetilde{\Delta V}_\gamma}$ can be seen as the product of the characteristic functions of certain known densities. So, at least in the one-dimensional case, we can extract useful information from characteristic function $\hat{f}_{\widetilde{\Delta V}_\gamma}$, which allows us to know in advance whether density $f_{\widetilde{\Delta V}_\gamma}$ is unimodal or not, thus facilitating the application of the numerical algorithms from Section 3.

Let us start by considering the linear decomposition (see [18] for details),

$$\widetilde{\Delta V}_\gamma = Q_0 + \frac{1}{2}\sum_{j \in J}\lambda_j Q_j,$$

where $J := \{j \in \{1, \ldots, p\}; \lambda_j \neq 0\}, Q_0$ is *normally distributed* with mean $\mu_{Q_0} = -\frac{1}{2}\sum_{j \in J}\frac{d_j^2}{\lambda_j}$ and variance $\sigma_{Q_0} = \sum_{j \notin J}d_j^2$, and Q_j has a *non-central chi-squared distribution* with one degree of freedom and non-centrality parameter $\zeta_j := \frac{d_j^2}{\lambda_j^2}$. Moreover, the variables $Q_0, Q_j, j \in J$ are independent. In this case,

$$f_{\widetilde{\Delta V}_\gamma} = f_{Q_0} * f_{\frac{\lambda_1}{2}Q_1} * \ldots * f_{\frac{\lambda_p}{2}Q_p}, \tag{20}$$

where $f_{Q_0}, f_{\frac{\lambda_j}{2}Q_j}, j \in J$, are the probability density functions of the random variables $Q_0, \frac{\lambda_j}{2}Q_j, j \in J$, respectively.

Following [19], for all $j \in J$, a unique value $\zeta \in (3, +\infty)$ exists, so that f_{Q_j} is decreasing if and only if $\zeta_j \leqslant \zeta$, i.e., f_{Q_j} has a unique mode at zero. Furthermore, f_{Q_j} is bimodal if and only if $\zeta_j > \zeta$, that is, f_{Q_j} has a mode at zero and another mode in $(0, +\infty)$.

Remark 3. The way to compute ζ is explained in [19] in the general case that a random variable is non-central chi-squared distributed with v degrees of freedom and non-centrality parameter ξ. In this case, the expression for its density is given by,

$$f_{v,\xi}(x) = \frac{1}{2}e^{-\frac{x+\xi}{2}}\left(\frac{x}{\xi}\right)^{\frac{v-2}{4}}I_{\frac{v-2}{2}}(\sqrt{\xi x}),$$

where $I_v(x)$ is the modified Bessel function of the first kind given by,

$$I_v(x) = \sum_{k=0}^{\infty}\frac{(x/2)^{2k+v}}{k!\mathcal{G}(v+k+1)}.$$

Here $\mathcal{G}$ denotes the gamma function. Defining $r_v(x) = \frac{I_v(x)}{I_{v-1}(x)}$ and,

$$g_v(\xi) = r_{\frac{v}{2}}\sqrt{\xi(\xi+v-4)} - \frac{\xi-2}{\sqrt{\xi(\xi+v-4)}}, \quad \xi > 4 - v, \ v \in (0, 2),$$

a unique $\xi_v \in (4 - v, +\infty)$ exists, so that $g_v(\xi_v) = 0$. Then, the non-central chi-squared density with $v > 0$ degrees of freedom and non-centrality parameter $\xi \geqslant 0$ decreases if and only if $0 < v \leqslant 2$ and $\xi \leqslant \xi_v$. The density is bimodal if and only if $0 < v < 2$ and $\xi > \xi_v$. The reported values in [19] are $v = 1$ and $\xi_v = 4.217$.

So, we know that $\zeta = 4.217$. From now on, we consider the case $p = 1$. Then, we can state the following theorem and corollary for the one-dimensional case,

Theorem 1. *With the same notation as before, let $f_{\frac{\lambda_1}{2}Q_1}$ be the density function of random variable $\frac{\lambda_1}{2}Q_1$. Then,*

(a) *If $\lambda_1 > 0$ and $\zeta_1 \leqslant \zeta$, then $f_{\frac{\lambda_1}{2}Q_1}$ is decreasing in $(0, +\infty)$, and has only one mode at zero.*
(b) *If $\lambda_1 > 0$ and $\zeta_1 > \zeta$, then $f_{\frac{\lambda_1}{2}Q_1}$ is bimodal and has one mode at zero and another at $(0, +\infty)$.*
(c) *If $\lambda_1 < 0$ and $\zeta_1 \leqslant \zeta$, then $f_{\frac{\lambda_1}{2}Q_1}$ is increasing in $(-\infty, 0)$, and has only one mode at zero.*
(d) *If $\lambda_1 < 0$ and $\zeta_1 > \zeta$, then $f_{\frac{\lambda_1}{2}Q_1}$ is bimodal and has one mode at zero and another at $(-\infty, 0)$.*

Proof 1. Observe that,

$$F_{\frac{\lambda_1}{2}Q_1}(x) = \mathbb{P}\left(\frac{\lambda_1}{2}Q_1 \leqslant x\right) = \mathbb{P}\left(Q_1 \leqslant \frac{2}{\lambda_1}x\right) = F_{Q_1}\left(\frac{2}{\lambda_1}x\right), \quad \text{if} \quad \lambda_1 > 0,$$

and,

$$F_{\frac{\lambda_1}{2}Q_1}(x) = \mathbb{P}\left(\frac{\lambda_1}{2}Q_1 \leqslant x\right) = \mathbb{P}\left(Q_1 \geqslant \frac{2}{\lambda_1}x\right) = 1 - F_{Q_1}\left(\frac{2}{\lambda_1}x\right), \quad \text{if} \quad \lambda_1 < 0.$$

We can derive for the probability density function,

$$f_{\frac{\lambda_1}{2}Q_1}(x) = \frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x\right), \quad \text{if} \quad \lambda_1 > 0,$$

and,

$$f_{\frac{\lambda_1}{2}Q_1}(x) = -\frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x\right), \quad \text{if} \quad \lambda_1 < 0.$$

Since $f_{Q_1} : (0, +\infty) \to \mathbb{R}$, then $f_{\frac{\lambda_1}{2}Q_1} : (0, +\infty) \to \mathbb{R}$, if $\lambda_1 > 0$ and $f_{\frac{\lambda_1}{2}Q_1} : (-\infty, 0) \to \mathbb{R}$, if $\lambda_1 < 0$. If $\zeta_1 \leqslant \zeta$ then f_{Q_1} is decreasing with only one mode at zero. Then, if $0 < x_1 < x_2$ and,

$$f_{\frac{\lambda_1}{2}Q_1}(x_1) = \frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x_1\right) \geqslant \frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x_2\right) = f_{\frac{\lambda_1}{2}Q_1}(x_2), \quad \text{if} \quad \lambda_1 > 0,$$

since f_{Q_1} is decreasing and $\frac{2}{\lambda_1} > 0$. For $\lambda_1 < 0$, if $x_1 < x_2 < 0$, then,

$$f_{\frac{\lambda_1}{2}Q_1}(x_1) = -\frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x_1\right) \leqslant -\frac{2}{\lambda_1}f_{Q_1}\left(\frac{2}{\lambda_1}x_2\right) = f_{\frac{\lambda_1}{2}Q_1}(x_2), \quad \text{if} \quad \lambda_1 > 0,$$

since $\frac{2}{\lambda_1} < 0$. This completes the proof of a) and c).

If $\zeta_1 > \zeta$ then f_{Q_1} is bimodal. We consider the modes of f_{Q_1} located at zero and m_1, where $m_1 > 0$. The density f_{Q_1} is decreasing in the intervals $(0, m_*)$ and $(m_1, +\infty)$ and increasing in (m_*, m_1), where $m_* \in \mathbb{R}, 0 < m_* < m_1$. The proof of (b) and (d) is now straightforward by applying a similar procedure as in (a) and (c) for the intervals defined before. □

Corollary 1. *With the same notation as before, let $f_{\widetilde{\Delta V}_\gamma}$ be the density function of random variable $\widetilde{\Delta V}_\gamma$. Then,*

(a) *If $\lambda_1 > 0$ and $\zeta_1 \leqslant \zeta$, then $f_{\widetilde{\Delta V}_\gamma}$ is decreasing in $(\mu_{Q_0}, +\infty)$, and has only one mode at $\mu_{Q_0} < 0$.*
(b) *If $\lambda_1 > 0$ and $\zeta_1 > \zeta$, then $f_{\widetilde{\Delta V}_\gamma}$ is bimodal and has one mode at $\mu_{Q_0} < 0$ and another at $(\mu_{Q_0}, +\infty)$.*
(c) *If $\lambda_1 < 0$ and $\zeta_1 \leqslant \zeta$, then $f_{\widetilde{\Delta V}_\gamma}$ is increasing in $(-\infty, \mu_{Q_0})$, and has only one mode at $\mu_{Q_0} > 0$.*
(d) *If $\lambda_1 < 0$ and $\zeta_1 > \zeta$, then $f_{\widetilde{\Delta V}_\gamma}$ is bimodal and has one mode at $\mu_{Q_0} > 0$ and another at $(-\infty, \mu_{Q_0})$.*

Proof 2. For $p = 1, f_{\widetilde{\Delta V}_\gamma} = f_{Q_0} * f_{\frac{\lambda_1}{2}Q_1}$, where the density of random variable Q_0 is a Dirac delta function, since $\mu_{Q_0} = -\frac{1}{2}\frac{d_1^2}{\lambda_1}$ and $\sigma_{Q_0} = 0$. Then,

$$f_{\widetilde{\Delta V}_\gamma}(\tau) = (f_{Q_0} * f_{\frac{\lambda_1}{2}Q_1})(\tau) = (f_{\frac{\lambda_1}{2}Q_1} * f_{Q_0})(\tau) = \int_{\mathbb{R}} f_{\frac{\lambda_1}{2}Q_1}(\tau - x)\delta(x - \mu_{Q_0})dx = f_{\frac{\lambda_1}{2}Q_1}(\tau - \mu_{Q_0}).$$

We observe that the convolution product is a translation of the density function $f_{\frac{\lambda_1}{2}Q_1}$, and by Theorem 1 we complete the proof. □

We illustrate the four cases of Corollary 1 in Fig. 1, where $\mu_{Q_0} = -2$ in the cases (a) and (b) and $\mu_{Q_0} = 2$ in (c) and (d).

A consequence of Corollary 1 is that within the delta-gamma framework, the VaR and ES values are bounded below or above depending on the aggregate gamma Greeks, that is, $\Gamma_{1,1} = \frac{\partial^2 V}{\partial S_1^2}$, and S_1 is the only underlying asset in the one-dimensional case. When $\Gamma_{1,1}$ is negative, we know *a priori* that the maximum likely loss in our portfolio is $\mu_{Q_0} + \Theta\Delta t$ (recall that
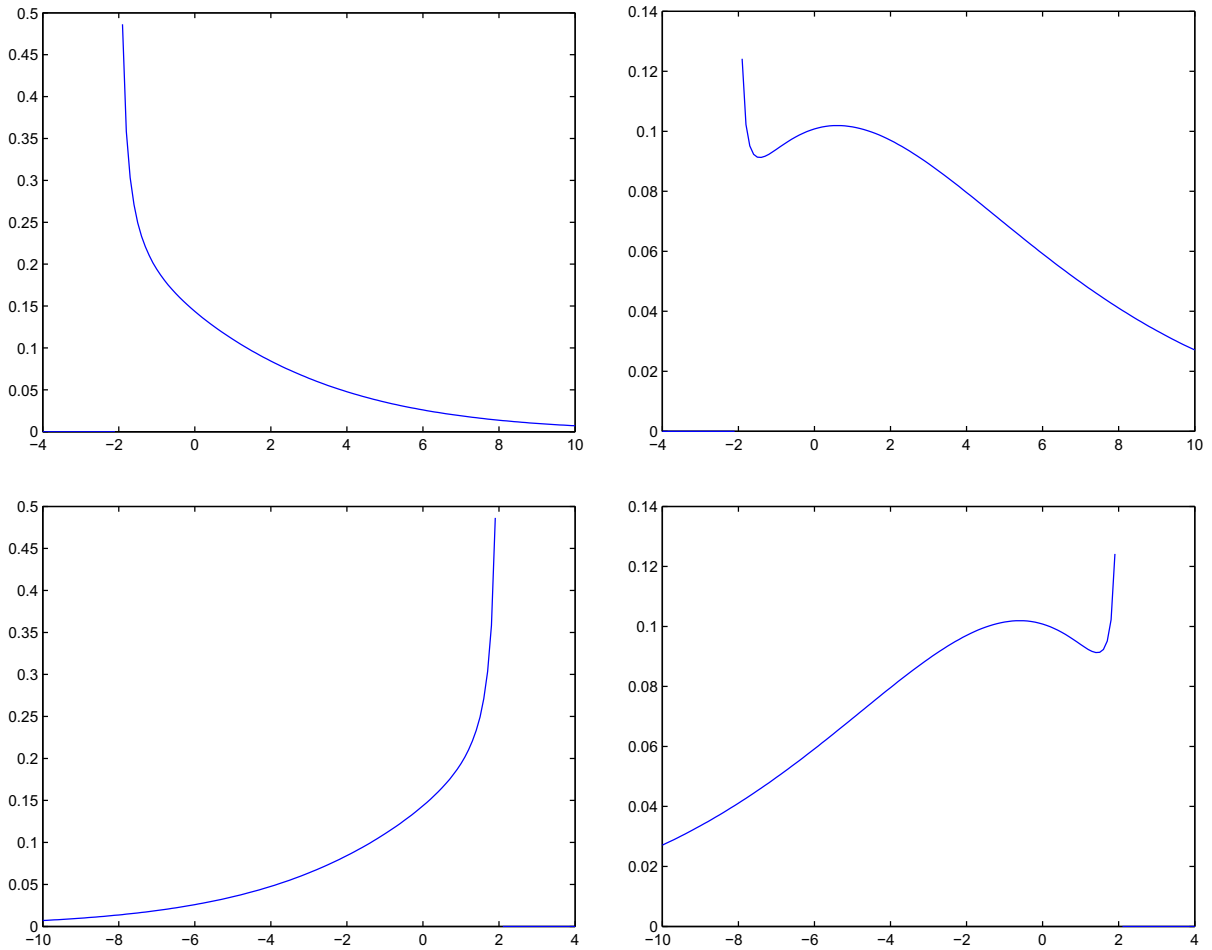
Fig. 1. Illustration of the four cases in Corollary 1. Top left: (a), top right: (b), bottom left: (c) and bottom right: (d).

$\Delta V = V(t + \Delta t) - V(t)$ and then positive values mean losses and negative values mean profits). Furthermore, when $\Gamma_{1,1}$ is positive, then the maximum likely profit is $\mu_{Q_0} + \Theta \Delta t$. We state these facts in the following corollary,

Corollary 2. *Let $\Gamma_{1,1}$ be as in Section 2 and let $VaR(\alpha), ES(\alpha)$ be the risk measures associated to the delta-gamma approach (3), $\alpha \in (0,1)$. Then,*

(a) If $\Gamma_{1,1} > 0$, then $ES(\alpha) > VaR(\alpha) > \mu_{Q_0} + \Theta \Delta t$, for all $\alpha \in (0,1)$;
(b) If $\Gamma_{1,1} < 0$, then $VaR(\alpha) < ES(\alpha) < \mu_{Q_0} + \Theta \Delta t$, for all $\alpha \in (0,1)$.

Proof 3. The proof follows directly from Corollary 1, since λ_1 is the only eigenvalue of $\Sigma\Gamma$ in the delta-gamma approach (3). □

5. Numerical examples

In this section we carry out tests to assess the performance of the methods presented previously. In order to keep a reasonable number of tables and figures, we mainly focus on the COS and WA$^{[a,b]}$ approximations for the numerical experiments. We show the performance of the filtered-COS and WA$^{\mathbb{R}}$ in some difficult cases. Although these last two methods work out well in all the proposed examples, they do not provide additional advantages when dealing with smooth densities.

Here we consider two portfolios of options written on the same underlying asset $S(t)$ as in [3],

Portfolio 1. The portfolio is composed of one short European call and half a short European put with maturity 60 days and time horizon 1 day.

Portfolio 2. The portfolio is composed of one short European call and half a short European put with maturity 60 days and time horizon 10 days.

For both portfolios we assume that the value of the underlying asset at time t is 100 with volatility level $\sigma = 0.3$, interest rate 0.1 and the strike price 101 for each option. We assume a zero dividend yield.

We restrict ourselves to the univariate case, that is, we have only one risk factor ($p = 1$), since the methods presented apply in the same form to the multivariate case (only the characteristic function changes). Hence, we consider here,

$$\Sigma = \left(S(t)\sigma\sqrt{\Delta t}\right)^2, \ \Gamma = \sum_{i=1}^n x_i \frac{\partial^2 v_i}{\partial S^2}, \ C = S(t)\sigma\sqrt{\Delta t}, \ \delta = \sum_{i=1}^n x_i \frac{\partial v_i}{\partial S},$$

where σ is the volatility associated to risk factor $S(t)$.

For the interval of integration $[a, b]$ within the COS and WA$^{[a,b]}$ methods, we use the general rule-of-thumb (6). Following [13] we then have,

$$\kappa_1 = \frac{1}{2}\text{tr}(\Gamma\Sigma), \ \ \kappa_2 = \frac{1}{2}\text{tr}\left((\Gamma\Sigma)^2\right) + \delta^T\Sigma\delta. \tag{21}$$

We could benefit from the information provided by Corollary 1 here, since in the two first portfolios $p = 1$. However, we choose to use the general interval which can be applied also when $p > 1$. We will apply Corollary 1 specifically in combination with the WA$^{\mathbb{R}}$ method.

Remark 4. Regarding the WA$^{\mathbb{R}}$ method, the strategy to recover a density is clearer with the shapes in Fig. 1. In cases (a) and (b) we can start the algorithm by considering $k_1 = \lfloor 2^m \cdot \mu_{Q_0} \rfloor$. For (c) and (d) we can follow the general procedure explained in Section 3.1.2. For these two last cases, we can stop the algorithm either following the general criteria (the area under the function is approximately one) or when $k_2 = \lfloor 2^m \cdot \mu_{Q_0} \rfloor$.

Also for the other numerical methods the choice on integration interval facilitates when the assumptions of the theorem and corollary apply.

We use a bisection method to compute the VaR value by the COS method from the CDF F_γ^{cos} with stopping condition,

$$\left| F_\gamma^{\text{cos}}(\widetilde{\text{VaR}}_N(\alpha)) - \alpha \right| < \epsilon_{cos},$$

where ϵ_{cos} is a pre-defined tolerance error. Regarding the WA$^{[a,b]}$ method, we consider $r = 0.9995$ in (10) and 2^m subintervals when applying the Trapezoidal Rule, where m is the scale of approximation (for a detailed description of the choice of the parameters r and m, we refer the reader to [14]). The number of trading days is fixed to 365. For comparison, we use so-called partial Monte Carlo (PMC) simulations (i.e. MC simulation within the delta-gamma approach) and full Monte Carlo (FMC), to evaluate the whole portfolio, both methods with 10^5 scenarios.[4]

Fig. 2 shows the density for Portfolio 1 recovered by the WA$^{[a,b]}$ method at scale $m = 10$ as well as the COS method with $N = 1024$ terms. Fig. 3 presents the density for Portfolio 2 (left) and a zoom over the vicinity of the non-smooth part (right side). In the right side plot, we have also used a filtered-COS method with exponential filters of orders $p_f = 8, 10$. As we can observe, using the filtered-COS the oscillations are dampened away from the point of irregularity. Moreover, for $p_f \leqslant 6$, the approximation gets too smooth near the conflicting point. We have computed the VaR and the ES values getting more accurate values when using $p_f = 10$. A possible explanation is that we loose a significant part of the density in the peaked zone when using low order filters. Anyway, filtered-COS with $p_f = 10$ has a *similar behavior* as COS without filters when computing the risk measures, concluding that in this case, the oscillations away from the singular point have very little impact on the final results.

Table 1 presents the relative errors when computing the VaR and ES values for Portfolio 1 by means of the COS method with the number of terms (i.e. scale of approximation) ranging from 32 (scale 5) to 256 (scale 8), and using several tolerances for the error when applying the bisection method. We observe significant improvement in the accuracy when decreasing the tolerance error, as expected. We finally choose $\epsilon_{cos} = 10^{-4}$, since a smaller epsilon leads to an increase in CPU time.

Table 2 presents the relative errors when computing the VaR and ES values for Portfolio 1 by means of the WA$^{[a,b]}$ and COS methods with the number of terms ranging from 32 to 256. In the case of the wavelets-based method, we recover the coefficients from the PDF (WA PDF) and directly from the CDF (WA CDF). Regarding the COS method we consider $\epsilon_{cos} = 10^{-4}$. We also provide the CPU time. As we can observe from the table, although both methods are accurate, the WA (CDF) version is extremely fast for computing the VaR value. Despite this, COS is faster in the case that we wish to compute both risk measures. The reason is that the calculation of one coefficient for the WA (CDF) methods is more CPU-time involved than the

---

[4] The programs have been coded in C language and run under Linux OS on a personal computer Intel DG45ID motherboard with Intel Core2Quad Q9550 2.83 GHz processor and 8 GB SDRAM. In the measurement of the CPU time, we also have considered the computation of the Greeks.
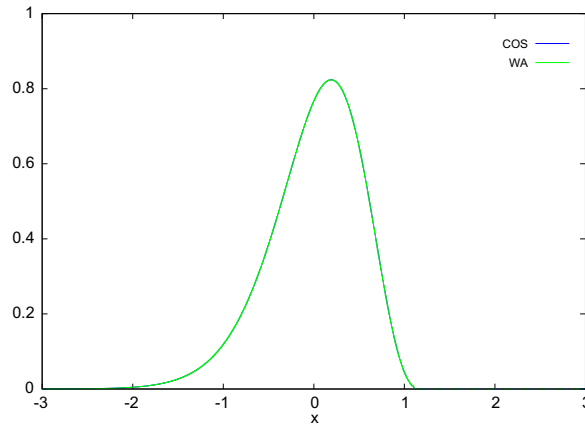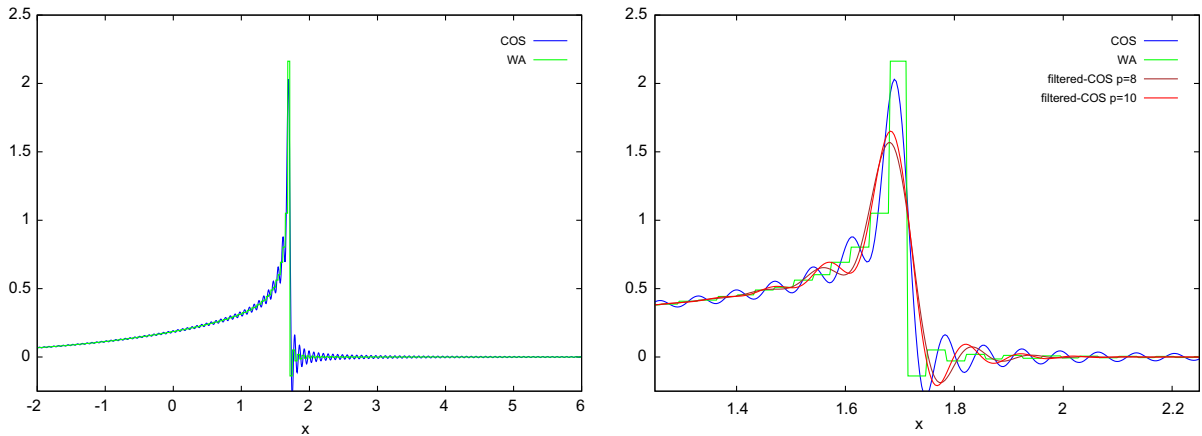
Fig. 2. Density for Portfolio 1.

Fig. 3. Density for Portfolio 2 (left) and zoom (right).

Table 1

Relative errors for Portfolio 1 for VaR and ES values calculated with the COS method at confidence level $\alpha = 0.99$ with $L = 5$. The reference values are: 0.903111819 (VaR), 0.964605146 (ES). These values have been computed as the average between the values obtained with the COS method using $N = 4000, \epsilon_{cos} = 10^{-6}$, and the $WA^{[a,b]}$ method with $m = 12$.

N	$\epsilon_{cos} = 10^{-3}$		$\epsilon_{cos} = 10^{-4}$		$\epsilon_{cos} = 10^{-5}$		$\epsilon_{cos} = 10^{-6}$	
	VaR	ES	VaR	ES	VaR	ES	VaR	ES
32	5.53e−3	5.96e−2	5.80e−4	4.10e−3	2.71e−4	6.92e−4	2.13e−4	5.38e−5
64	5.53e−3	6.12e−2	5.80e−4	5.80e−3	1.16e−4	6.96e−4	5.80e−5	5.94e−5
128	5.53e−3	6.15e−2	5.80e−4	5.94e−3	3.87e−5	8.72e−4	3.87e−5	2.27e−5
256	5.53e−3	6.14e−2	5.80e−4	5.93e−3	3.87e−5	8.67e−4	3.87e−5	1.90e−5

computation of one coefficient for the COS method, and we need many of them when calculating the ES with the WA (CDF) variant.

We show in Table 3 the VaR and ES values for Portfolio 2 evaluated at different percentiles. In this case, we have focused on the WA (CDF) variant and we have simply called it WA. The results show very accurate values at low and high loss levels. We also can observe an overestimation of the risk measured with the VaR and Expected Shortfall within the delta-gamma approach.

The $WA^{\mathbb{R}}$ method can be used in an accurate way to recover the densities associated to Portfolio 1 and Portfolio 2, as shown in Fig. 4. For this purpose we have considered $\epsilon_{tol} = 10^{-4}$. The algorithm will be stopped when the mass of the computed density reaches $1 - \epsilon_{tol}$. Table 4 shows the truncation values k_1 and k_2.

Table 2
Relative errors for Portfolio 1 for VaR and ES values calculated at confidence level $\alpha = 0.99$ with $L = 5$ and $\epsilon_{cos} = 10^{-4}$. The reference values are: 0.903111819 (VaR), 0.964605146 (ES). These values have been computed as the average between the values obtained with the COS method using $N = 4000, \epsilon_{cos} = 10^{-6}$, and the WA$^{[a,b]}$ method with $m = 12$. The CPU times are also provided.

| Number of terms (scale m) | WA (PDF) | | WA (CDF) | | COS | | CPU (milliseconds) | | | | | |
| | | | | | | | WA (PDF) | | WA (CDF) | | COS | |
	VaR	ES	VaR	ES	VaR	ES	VaR	VaR + ES	VaR	VaR + ES	VaR	VaR + ES
32 (5)	2.53e−2	4.00e−1	2.53e−2	1.89e−2	5.80e−4	4.10e−3	0.28	0.35	0.18	0.26	0.30	0.31
64 (6)	6.49e−2	9.06e−1	1.43e−2	3.96e−3	5.80e−4	5.80e−3	0.84	1.13	0.38	0.67	0.61	0.64
128 (7)	5.53e−3	6.75e−2	5.53e−3	1.01e−3	5.80e−4	5.94e−3	2.75	3.83	0.77	1.88	1.23	1.30
256 (8)	1.54e−2	1.80e−1	4.37e−3	3.08e−4	5.80e−4	5.93e−3	9.80	14.06	1.58	5.83	2.49	2.62

Table 3
VaR and ES calculated at several percentiles for Portfolio 2 with $L = 5$. The parameters for COS method are: $N = 256$ and $\epsilon_{cos} = 10^{-4}$. The risk measures with the WA method have been computed directly from the CDF at scale $m = 8$.

| Percentile | WA | | COS | | PMC | | FMC | |
	VaR	ES	VaR	ES	VaR	ES	VaR	ES
10	−2.3523	0.5255	−2.3436	0.5257	−2.3688	0.5239	−2.4785	0.4523
20	−1.0902	0.7975	−1.0891	0.7975	−1.0844	0.8044	−1.1510	0.7288
30	−0.3188	1.0094	−0.3298	1.0093	−0.3333	1.0061	−0.3607	0.9365
40	0.2421	1.1847	0.2219	1.1846	0.2070	1.1798	0.1985	1.1023
50	0.6629	1.3327	0.6535	1.3325	0.6415	1.3305	0.6289	1.2389
60	1.0135	1.4575	1.0020	1.4573	1.0039	1.4565	0.9569	1.3477
70	1.2939	1.5599	1.2863	1.5600	1.2862	1.5592	1.2126	1.4320
80	1.5043	1.6381	1.5081	1.6367	1.5073	1.6371	1.3962	1.4941
90	1.6445	1.6895	1.6508	1.6867	1.6527	1.6874	1.5054	1.5308
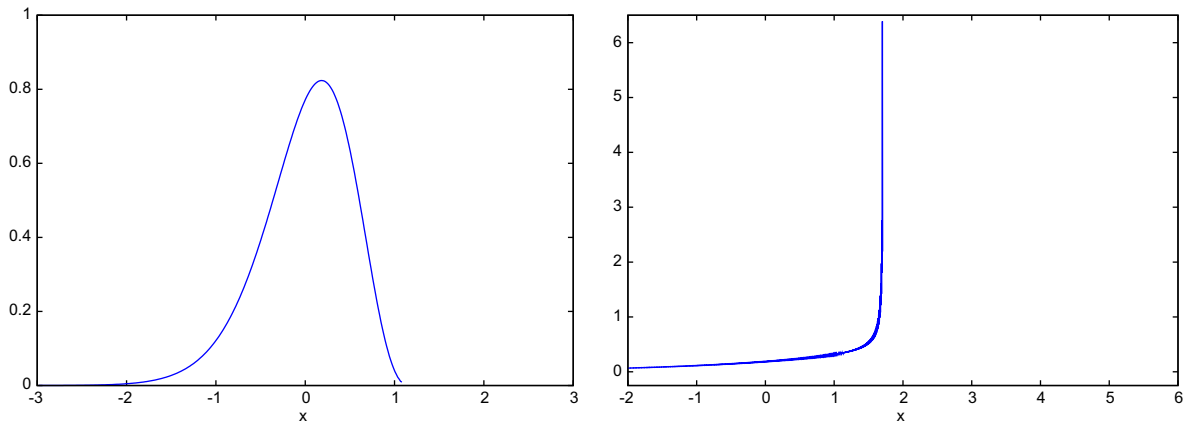
Fig. 4. Density recovered by means of the WA$^{\mathbb{R}}$ for Portfolio 1 (left) and Portfolio 2 (right).

Table 4
Truncation values when using the WA$^{\mathbb{R}}$ method.

Scale	Seed	x	$|f_m(x)|$	k_1	$x = \frac{k_1}{2^m}$	$|f_m(x)|$	k_2	$x = \frac{k_2}{2^m}$	$|f_m(x)|$
Portfolio 1									
6	−167	−2.54	6.05e−4	−196	−3.00	8.87e−5	261	1.08	9.24e−3
Portfolio 2									
8	−2452	−8.91	1.36e−3	−3279	−12.14	9.99e−5	3544	1.70	2.19

Remark 5. Regarding Portfolio 1, we compute $\lambda_1 = -0.1205 < 0, \zeta_1 = 17.1836 > \zeta, \mu_{Q_0} = 1.0355$ as explained in Section 4.4. Although we observe only one peak, following Corollary 1, the density function is bimodal, with one mode in $(-\infty, 1.0355)$ and another at $\mu_{Q_0} = 1.0355$. The second mode is so small that in practice it is not observed and can be neglected in this case. It is worth mentioning that we should add $\mu_{Q_0} + \Theta \Delta t = 1.1025$ to be in the domain of $f_{\Delta V_\gamma}$ (as we did in Fig. 4) instead of $f_{\widetilde{\Delta V_\gamma}}$. If we follow the general criteria (related to the mass of the density) the algorithm stops at $x = 1.08$ and the density evaluated at this point is $9.24 \cdot 10^{-3}$, in perfect agreement with the theory.
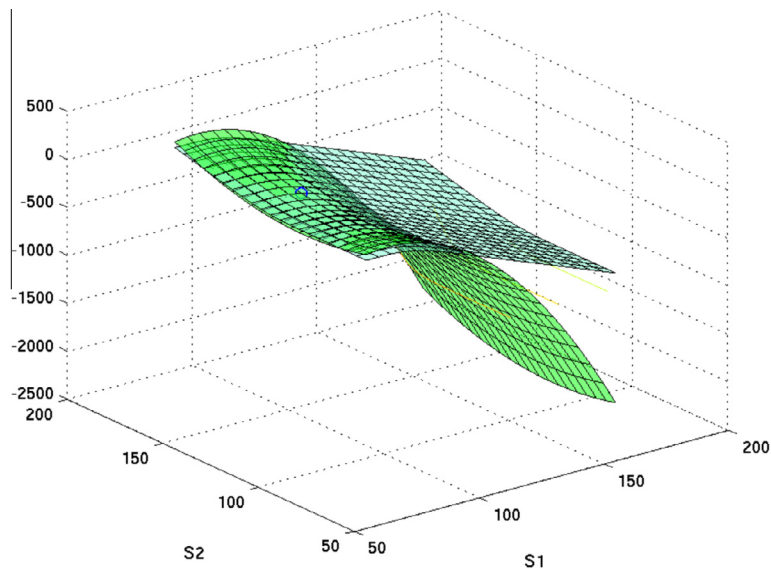
Fig. 5. Portfolio value (green) as a function of the underlying assets and delta-gamma approximation (blue) for Portfolio 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
VaR and ES values for Portfolio 3 at confidence level $\alpha = 0.9$.

Method	VaR	ES
FMC	32.9846	45.3541
PMC	32.4927	44.8006

The second portfolio is a somewhat more interesting case. On the one hand, the peak is sharper than the one observed in Fig. 3. This fact can be interpreted as a more accurate recovery of the density. On the other hand, the algorithm stops at $x = 1.7046$ where the value of the density is 2.19. We have calculated the maximum loss value for the density by means of PMC simulation, obtaining the value 1.7050. Again, all these facts are in accordance with the theory developed in Section 4.4, since $\lambda_1 = -1.2052 < 0, \zeta_1 = 1.7184 < \zeta, \mu_{Q_0} = 1.0355$ and $\mu_{Q_0} + \Theta\Delta t = 1.7050$, and then the density is increasing in $(-\infty, 1.7050)$, once we have added $\Theta\Delta t$.

Let us now consider the following portfolio composed of options written in two different, uncorrelated underlying assets,

Portfolio 3. The portfolio is composed of 10.25 short European calls written on the underlying asset S_1 and 5.5 long European calls written on the underlying asset S_2 with maturity 60 days and time horizon 10 days.

With this example we aim to illustrate the theoretical work done in Section 4.3 about the suitability of the delta-gamma approach. We assume that the value of the underlying assets at time t are $S_1 = 90, S_2 = 130$ both with volatility 0.2, interest rate 0.1 and strike prices $K_1 = 90$ for the options written on S_1 and $K_2 = 125$ for the options written on S_2. We assume a zero dividend yield.

We observe that the portfolio $V(S_1(t), S_2(t)) = V_1(S_1) + V_2(S_2)$ is neither convex nor concave, since the Hessian,

$$\text{Hess}(V) = \begin{pmatrix} -10.25\frac{\partial^2 V}{\partial S_1^2} & 0 \\ 0 & 5.5\frac{\partial^2 V}{\partial S_2^2} \end{pmatrix},$$

is indefinite, as $\frac{\partial^2 V}{\partial S_1^2} > 0$ and $\frac{\partial^2 V}{\partial S_2^2} > 0$ for all (S_1, S_2). However, we observe that V_1 is concave in S_1 and V_2 is convex in S_2 and the delta-gamma approximation is accurate in this case. We plot in Fig. 5 the portfolio value and the delta-gamma approximation and we report in Table 5 the VaR and the ES values at 0.9 confidence level. As we observe, the delta-gamma approximation is very accurate in a neighborhood of the initial value $(S_1, S_2) = (90, 130)$.

6. Conclusions

We have investigated in this work the efficiency of the $WA^{[a,b]}$, COS, $WA^{\mathbb{R}}$ and filtered-COS methods within the delta-gamma framework, to compute the Value-at-Risk and the Expected Shortfall measures in a nonlinear portfolio composed

of options. Regarding the wavelet approaches, we have recovered the risk measures from the probability density function and from the cumulative distribution function. Both procedures give us very accurate results, however, in the second case, the algorithm is faster since fewer coefficients are needed to perform the Fourier inversion. Regarding the COS method, we recover the probability density function and then integrate (analytically) to determine the cumulative distribution function, after which we employ a root-finding method to calculate the risk measures. The filtered-COS method with exponential filter gives us the same accuracy as the COS method, concluding that, in this case, the oscillations away from the singular point have very little impact on the final results. The WA$^{\mathbb{R}}$ method recovers in an adaptive way the density without fixing an interval for the inversion beforehand, although the method is somewhat more CPU time consuming.

We also have stated a useful theorem and two corollaries about the shape of the delta-gamma density and the maximum profit and loss for the one-dimensional case. The numerical experiments carried out with the WA$^{\mathbb{R}}$ method are in accordance with the theory developed in Section 4.4. The problem for the p-dimensional case is more involved and it is subject of future research.

We have seen that if a p-dimensional portfolio is neither convex nor concave, but can be decomposed in several convex (or concave) subportfolios made of options written on one underlying asset, then the delta gamma approach is highly accurate.

Acknowledgments

References

[1] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk, Math. Finance 9 (3) (1999) 203–228.
[2] G. Bormetti, V. Cazzola, D. Delpini, G. Livan, Accounting for risk of non linear portfolios, a novel Fourier approach, Working paper, available at arxiv.org, 2010.
[3] M. Britten-Jones, S.M. Schaefer, Non-linear value-at-risk, Eur. Finance Rev. 2 (1999) 161–187.
[4] I. Daubechies, Ten lectures on wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1992.
[5] D. Duffie, J. Pan, Analytical value-at-risk with jumps and credit risk, Finance Stoch. 5 (2) (2001) 155–180.
[6] F. Fang, C.W. Oosterlee, A novel pricing method for European options based on fourier-cosine series expansions, SIAM J. Sci. Comput. 31 (2008) 826–848.
[7] P. Glasserman, P. Heidelberger, P. Shahabuddin, Importance sampling and stratification for value-at-risk. IBM Research Report, RC 21450, Yorktown Heights, NY, 1999.
[8] P. Glasserman, P. Heidelberger, P. Shahabuddin, Portfolio value-at-risk with heavy-tailed risk factors, Math. Finance 12 (3) (2002) 239–269.
[9] D. Gottlieb, C.W. Shu, On the Gibbs phenomenon and its resolution, SIAM Rev. 39 (4) (1997) 644–668.
[10] P. Jorion, Value at Risk: The New Benchmark for Managing Financial Risk, third ed., McGraw-Hill, 2007.
[11] J.J. Masdemont, L. Ortiz-Gracia, Haar wavelets-based approach for quantifying credit portfolio losses, Quant. Finance (2011), http://dx.doi.org/10.1080/14697688.2011.595731.
[12] A.M. Mathai, S.B. Provost, Quadratic Forms in Random Variables: Theory and Applications, Marcel Dekker Incorporated, 1992.
[13] J. Mina, A. Ulmer, Delta-gamma four ways, RiskMetrics Group LLC, 1999.
[14] L. Ortiz-Gracia, J.J. Masdemont, Peaks and jumps reconstruction with B-splines scaling functions, J. Comput. Appl. Math. 272 (2012) 258–272.
[15] 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.
[16] C. Roubinez (1997). Going Greek with VaR. RISK, February, 57–65.
[17] M.J. Ruijter, M. Versteegh, C.W. Oosterlee, On the application of spectral filters in a Fourier option pricing technique, Working paper, available at www.ssrn.com, 2013.
[18] J.V. Siven, J.T. Lins, A. Szymkowiak-Have, Value-at-risk computation by Fourier inversion with explicit error bounds, Working paper, available at arxiv.org, 2008.
[19] Y. Yu, The shape of the non-central χ^2 density, Working paper, available at arxiv.org, 2011.