

A new Fourier transform algorithm for value-at-risk

Claudio Albanese¹, Ken Jackson² and Petter Wiberg^{3,4}

¹ Department of Mathematics, University of Toronto, 100 St George Street, Toronto, ON, M5S 3G3, Canada

² Department of Computer Science, University of Toronto, 10 King's College Road, Toronto, ON, M5S 3G4, Canada

³ Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

E-mail: albanese@math.toronto.edu, krj@cs.toronto.edu and wiberg@maths.warwick.ac.uk

Received 6 June 2003, in final form 2 February 2004

Published 4 March 2004

Online at stacks.iop.org/Quant/4/328 (DOI: 10.1088/1469-7688/4/3/008)

Abstract

In this paper, we introduce a new Fourier method for computing value-at-risk for a portfolio with derivatives and for return models with fat tails. The new method does not assume that the characteristic function for the return model is known explicitly. We define a class of admissible models for returns and present statistical evidence that supports our approach. We discuss the details of the algorithm. The paper concludes with two applications of value-at-risk. Both examples illustrate the effect that changes in the models for portfolio value and for risk factor returns have on the value-at-risk surface.

1. Introduction

In this paper, we present a new method for computing value-at-risk. The method is intended for portfolios that include options and other derivatives and for returns with fat tails. Value-at-risk (VaR) is a measure of the market risk, the chance of a loss in a portfolio caused by unfavourable changes in prices and rates. Minimum risk management standards for banks are set and enforced by national regulators. As part of these standards, banks are encouraged to use mathematical models for value-at-risk to estimate their exposure to market risk. Minimum capital requirements are determined from such simulations and the quality of the models used (see [2–5]).

For a given portfolio, value-at-risk is defined as the maximum loss over a specified holding period and for a given confidence level α . In other words, if $\Delta\Pi$ is the change in portfolio value during the holding period, value-at-risk is the $(1 - \alpha)$ -percentile of the distribution of $\Delta\Pi$:

$$P[\Delta\Pi \leq -\text{VaR}] = 1 - \alpha. \quad (1)$$

In the special case where $\Delta\Pi$ is linear and the risk factor returns are multivariate normal there is a closed form expression for the value-at-risk [19, 24]. However, this simple model suffers from two serious limitations. First, real world returns have fatter tails than the normal distribution, which may lead to too conservative estimates for the risk. Second, for portfolios with derivatives, the change in value is a nonlinear function. Often the local errors in linear approximations are too large to give accurate value-at-risk estimates, an effect that is exacerbated by dynamic hedging strategies.

Because of the importance of value-at-risk simulations, several methods for computing value-at-risk for portfolios with derivatives have been proposed; see [10, 19, 24] for an overview. The method presented in this paper is based on Fourier transforms. We assume that the change in portfolio value $\Delta\Pi$ in (1) can be approximated locally by a second-order Taylor approximation,

$$\Delta\tilde{\Pi} = \Xi + \Delta^T r + \frac{1}{2} r^T \Gamma r, \quad (2)$$

where the vector r contains the returns on each of the risk factors. The vector Δ and the matrix Γ contain the first and second derivatives of $\Delta\Pi$ with respect to the risk factor returns

⁴ Author to whom any correspondence should be addressed.

r , and Ξ is a constant. One strategy, studied by several authors, is to use a return model where the characteristic function for $\Delta\tilde{\Pi}$ can be derived explicitly. The distribution for $\Delta\tilde{\Pi}$ can then be computed by inverting the Fourier transform numerically. Mina and Ulmer [23] consider the case where returns are normal. Duffie and Pan [11] explore this idea further, and they extend the model to a fat-tailed mixture model and show how to include credit risk. Glasserman *et al* [17] generalize the method to returns with a multivariate Student- t distribution. To distinguish our method from other Fourier methods, we call it the *fast convolution method*. The fast convolution method is different from existing methods (see [11, 17, 23]) in that it does not require explicit knowledge about the characteristic function. It therefore enjoys greater flexibility in the type of return models that can be handled.

Two popular, alternative approaches to value-at-risk algorithms are Monte Carlo methods and analytic approximation methods. Monte Carlo methods are very general; they can be used with a variety of return models and general functions for the portfolio value. However, Monte Carlo methods are computationally intensive and the rate of convergence is often slow, although this can be improved with sophisticated variance reduction techniques. Also Monte Carlo methods may give different results for each simulation, which makes using these methods in value-at-risk optimization problems difficult. Analytic methods come in many different flavours. Britten-Jones and Schaefer [8] use an approximation to the distribution of a sum of non-central χ^2 random variables. Mina and Ulmer [23] discuss two parametric methods: the Johnson family of distributions and the Cornish–Fisher expansion. El-Jahel *et al* [12] propose a method of moments algorithm that can handle fat-tailed returns. Feuerverger and Wong [14] explore a saddle point approximation method. They argue that, similarly to the case for other methods for normal returns, fat tails can be incorporated with a mixture model.

In section 2, we introduce the class of admissible return models in the fast convolution method and illustrate this idea with three example distributions. Section 3 presents the fast convolution method. We show that the method is efficient, although it can be further improved with dimension reduction for large portfolios [1]. In section 4, we present two numerical examples. In the first example we study the impact of the distribution assumption on an example portfolio. In the second example, we study the difference between a value-at-risk model with constant and with stochastic volatility and show that the two value-at-risk surfaces have different characteristics. In section 5 we conclude with a summary of the results and a discussion of future work.

2. Modelling risk factor returns

In this section, we discuss the class of return models admissible in the fast convolution method. Suppose that the evolution of the risk factors is modelled by a standard random walk. The values of the risk factors over time are described by a stochastic process $\{S_i\}_{i=1}^{\infty}$ where the j th component of the vector S_i is the value of the j th risk factor at time i . In

the simplest form, the relative returns $(S_i - S_{i-1})/S_{i-1}$ are independent and identically distributed, and we let R denote a generic random variable of this type. More general models have been studied extensively in econometrics [9], but this simple model is sufficiently general for our purposes.

Although the relative returns in the model are independent over time, the dependence structure for the components of the vector R is very important for the accuracy of value-at-risk models. In the fast convolution method, the dependence structure in the risk factor returns is approximated using a product pdf

$$p(r) = p_1(x_1) \cdots p_n(x_n), \quad (3)$$

where x is an affine transformation of the original returns r . The affine transformation is determined as a part of the parameter estimation procedure that fits the return model to a time series $\{r_i\}_{i=1}^d$ of past returns. For simplicity, we assume that the time steps for the returns agree with the time steps of the random walk model. In our examples, the time series consists of historical returns computed from daily closing prices for stocks on the Toronto Stock Exchange (TSX).

The mean and covariance matrix for the returns can be estimated with standard statistics

$$\mu = \frac{1}{d} \sum_{i=1}^d r_i \quad \text{and} \quad C = \frac{1}{d-1} \sum_{i=1}^d (r_i - \mu)(r_i - \mu)^T;$$

see for example [21]. Because the properties of financial time series tend to change over time, it might be preferable to use an estimator that gives more weight to recent observations [19, 24]. The use of such estimators has little effect on the computational procedure presented, but it is known that it may lead to significant changes in the value-at-risk estimates [6, 20].

The approximation (2) and the covariance matrix can be diagonalized simultaneously by solving a generalized eigenvalue problem:

$$\begin{aligned} \Lambda &= A^T \Gamma A, \\ C &= A A^T. \end{aligned} \quad (4)$$

The matrix A is non-singular and Λ is a diagonal matrix with diagonal elements λ_i for $i = 1, \dots, n$. This is a standard problem appearing in many different applications. We refer the reader to [18] for a discussion of how to solve eigenvalue problems numerically and to [1] for a discussion of how to take advantage of the special structure presented in the problem above. For the transformed variables $x = A^{-1}(r - \mu)$, the approximation (2) is

$$\Delta\tilde{\Pi} = \tilde{\Xi} + \sum_{i=1}^n \left(\tilde{\Delta}_i x_i + \frac{\lambda_i}{2} x_i^2 \right) \quad (5)$$

where

$$\tilde{\Xi} = \Xi + \mu^T \Delta + \frac{1}{2} \mu^T \Gamma \mu, \quad (6)$$

$$\tilde{\Delta} = A^T (\Delta + \Gamma \mu). \quad (7)$$

If a multivariate normal distribution is used to model the returns, μ and C determine the model uniquely. Otherwise, the factorization of the portfolio gives the coordinate system

for (3) and the remaining parameters for each p_i are estimated independently. With this estimation procedure, we can ensure that the marginal distributions p_i are accurate in the canonical directions from (4).

In the remainder of this paper, we use *admissible model* to refer to a generic return model (3) fitting into the framework for parameter estimation introduced in this section. In section 2.1, we examine the arguments for using admissible models in value-at-risk simulations. In section 2.2, we give three admissible models used in the computational examples in section 4.

2.1. The case for admissible models

The assumption that the risk factor returns can be modelled by a product pdf (3) is a cornerstone of the computational algorithm developed in section 3. The family of admissible models has several important properties.

- (i) Historical time series show that returns typically are scattered around a central point in space which suggests that returns can be modelled with a unimodal distribution. Therefore, multivariate normal distributions are a commonly used model in finance. The normal model is a member of the family of admissible models, but more general models can be formulated by using different distributions p_i in (3).
- (ii) The dependence structure for risk factors is important for the accuracy of value-at-risk models. In the family of admissible models, the dependence structure is approximated by choosing an appropriate coordinate system for x in (3). In particular, under the transformation $X = A^{-1}(R - \mu)$ from (4), the covariance matrix for X is the identity while the covariance matrix of R is equal to the sample covariance C .
- (iii) Real world time series show that large returns, positive and negative, are more common than in a normal distribution, i.e. the distribution has fat tails. In the framework of admissible models, fat tails can be captured with one-dimensional distributions p_i .

In the value-at-risk literature, return models of many other types have been proposed. Two important classes of such models are copula models [13] and elliptical distributions. In general, the admissible models are different from both copula and elliptical distributions. The admissible models are constructed taking the portfolio structure into account. By exploiting the portfolio composition, the simple structure (3) can be used to get an accurate model for returns for the *particular* portfolio at hand; see [29] for more details and examples.

The current lack of consensus about what is the ‘correct’ model for returns is partly due to the difficulty of performing statistical tests in high dimensions (see e.g. [15]). In the admissible models, the independence assumption implicit in (3) can be tested. We conclude this section with a small statistical test that gives confidence that the class of models proposed is a reasonable approximation. Although the question of independence is somewhat misguided (e.g. it does

not tell us about the accuracy for value-at-risk) it is unlikely that it is accurate if the independence is not valid.

To test the independence assumption, we use a simple binomial test, as discussed for example in [21]. We consider returns for pairs of stocks in the Toronto 35 index and portfolios containing one call option for each stock. This gives us a total of 595 two-dimensional portfolios. Using the standard Black–Scholes model, we compute an approximation of the form (2) and factor it as described above. This gives two one-dimensional quadratic functions $\pi_1(x_1)$ and $\pi_2(x_2)$ where x_1 and x_2 are the returns in the coordinate system of the factorization. In the binomial test, we consider the event

$$B = (A_1 \cap A_2) \cup (A_1^c \cap A_2^c)$$

where

$$A_i = \{\pi_i(x_i) : \pi_i(x_i) > \mu_{\pi_i}\} \quad \text{for } i = 1, 2,$$

and μ_{π_i} is the sample mean of π_i . The null hypothesis is that the probability of B , denoted by q , is

$$q = \hat{p}_1 \hat{p}_2 + (1 - \hat{p}_1)(1 - \hat{p}_2).$$

This means that the null hypothesis can be refuted if the difference between q and its sample estimate \hat{q} is large. The estimates for the probabilities \hat{p}_1 , \hat{p}_2 and \hat{q} are the ratios of the occurrences of A_1 , A_2 and B to the total number of samples d . Treating each sample as an independent Bernoulli trial, the random variable

$$(\hat{q} - q) \sqrt{\frac{d}{q(1-q)}} \sim N(0, 1) \quad \text{as } d \rightarrow \infty. \quad (8)$$

For large d , we can use this approximation and refute the null hypothesis, and consequently also independence, if (8) is outside a confidence interval.

The result from this statistical test is displayed in table 1. We make two observations. First, the table shows that for our 595 test portfolios, a majority passes the independence test. The rate of failure is expected to be equal to the confidence level. The table shows that the actual rate is higher, but in the overwhelming majority of cases the test passes. Second, the rate of failed tests is higher when two years of data are used than for the shorter time series of one year. In conclusion, the statistical test supports the view that an admissible model (3) is reasonable for risk factor returns.

2.2. Three examples of admissible models

There are many possible alternatives that could be used in (3). We have chosen to consider three instances.

First, we consider the standard normal model. In this model⁷, the random variable for the transformed returns X is normal $N(\mu, \sigma^2)$ with the probability density function (pdf)

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

⁷ In the case of log returns, this model is often called the log-normal model; stock prices are log-normally distributed. We caution the reader that some authors use the terminology ‘normal model’ for a model where stock price changes are normal.

Table 1. A binomial test of independence applied to portfolios with two call options. The test is performed for all 595 pairs of stocks in the Toronto 35 index. The options are at-the-money and have three months to maturity. The frequencies in the table are the number of rejections of the null hypothesis for three different confidence levels. The corresponding percentage of rejected portfolios is reported in parentheses.

Time period	Rejection of the null hypothesis		
	at 1%	at 5%	at 10%
24.05.01–25.05.99	18/595 (3.0%)	66/595 (11%)	108/595 (18%)
24.05.01–25.05.00	13/595 (2.2%)	39/595 (6.5%)	78/595 (13%)
24.05.00–25.05.99	14/595 (2.4%)	59/595 (9.9%)	102/595 (17%)

Second, we consider the asymmetric Student- t model. Compared to the normal distribution, the Student- t distribution has an additional parameter ν that controls the fatness of the tails (see e.g. [21, 27]). The pdf is

$$p_T(x; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

In the asymmetric Student- t model, we generalize the Student- t distribution slightly. The density is divided into two parts relative to the median m , and each half has a parameter, ν_+ and ν_- , that controls the fatness of the tail. Furthermore, a parameter ρ balances the contribution to the overall variance σ^2 from the two halves. The random variable X has the pdf

$$p(x) = \begin{cases} \frac{p_T\left(\sqrt{\frac{\nu_-}{\nu_- - 2}} \frac{(x-m)}{\sigma\sqrt{\rho}}; \nu_-\right) \sqrt{\frac{\nu_-}{\nu_- - 2}}}{\sigma\sqrt{\rho}} & \text{if } x \leq m, \\ \frac{p_T\left(\sqrt{\frac{\nu_+}{\nu_+ - 2}} \frac{(x-m)}{\sigma\sqrt{1-\rho}}; \nu_+\right) \sqrt{\frac{\nu_+}{\nu_+ - 2}}}{\sigma\sqrt{1-\rho}} & \text{if } x > m. \end{cases}$$

Third, we consider a non-parametric model where X has the pdf given by a density estimate. In this model, the pdf is estimated by

$$\hat{p}_d(x) = \frac{1}{dh} \sum_{i=1}^d K\left(\frac{x - x_i}{h}\right) \quad (9)$$

where $\{x_i\}_{i=1}^d$ is the time series of returns and

$$K(x) = \frac{15}{16}(1 - x^2)^2 \quad \text{for } |x| \leq 1.$$

Density estimates of this form were first suggested in [26] and analysed by Parzen in [25]. The weighting function $K(x)$ above is from [28].

Figure 1 shows quantile–quantile plots for returns on Bell Canada Enterprises (BCE). The three plots show that all three instances capture the overall shape of the empirical distribution. However, the tails of the normal model do not agree with the empirical distribution and underestimate the probability of large changes in the stock price. The asymmetric Student- t and the non-parametric Parzen type models perform better. In all our examples, the parameters in the parametric models are computed with standard statistical estimators; see [21] for a general discussion of estimation and [29] for details about this aspect of our implementation.

3. Computation of value-at-risk

The value-at-risk is the solution to (1) where $\Delta\Pi$, the change in portfolio value, is a nonlinear function. To solve (1) directly, Monte Carlo methods are the only viable strategy, but are typically time consuming. However, if $\Delta\Pi$ can be approximated locally by a quadratic form (2) then we can substitute $\tilde{\Delta\Pi}$ for $\Delta\Pi$ and approximate the value-at-risk VaR by the solution $\tilde{\text{VaR}}$ to

$$P[\tilde{\Delta\Pi} \leq -\tilde{\text{VaR}}] = 1 - \alpha. \quad (10)$$

In this section, we introduce a method for solving (10). The method is based on the convolution theorem and can be applied to admissible return models. The advantage of our approach is that the method can cope with a wider class of return models, compared to the Fourier transform methods discussed in [11, 17, 23]. The basic method has three main steps: (1) estimate the parameters of the model for risk factor returns; (2) discretize the probability densities and use the fast Fourier transform to compute a discrete approximation to the distribution of $\Delta\Pi$; and (3) use inverse interpolation to approximate the value-at-risk. The first step is discussed in section 2. In the next two subsections, we discuss the details of the second and third steps of the algorithm.

3.1. Step 2: discretization and convolution

For an admissible return model, (5) shows that the local approximation to the dynamics of portfolio value $\tilde{\Delta\Pi}$ is a sum of quadratic functions,

$$\tilde{\Delta\Pi} = \Xi + \sum_{i=1}^n \pi_i(x_i) \quad \text{where } \pi_i = \Delta_i x_i + \frac{\lambda_i}{2} x_i^2.$$

To simplify our notation, we write VaR for $\tilde{\text{VaR}}$ in (10) and similarly for (5). A consequence of the construction of the admissible return models is that the random variables π_i are independent. Therefore the pdf for $\tilde{\Delta\Pi}$ is the convolution product

$$p_{\tilde{\Delta\Pi}}(x) = (p_{\pi_1} * \dots * p_{\pi_n})(x + \Xi) \quad (11)$$

where p_{π_i} is the pdf of π_i . To approximate the pdf for $\tilde{\Delta\Pi}$, we can discretize the densities p_{π_k} and use the discrete Fourier transform, since their continuous counterparts satisfy [16]

$$p_{\pi_1} * \dots * p_{\pi_n} = \left(\prod_{i=1}^n \hat{p}_{\pi_i}\right)^\vee.$$

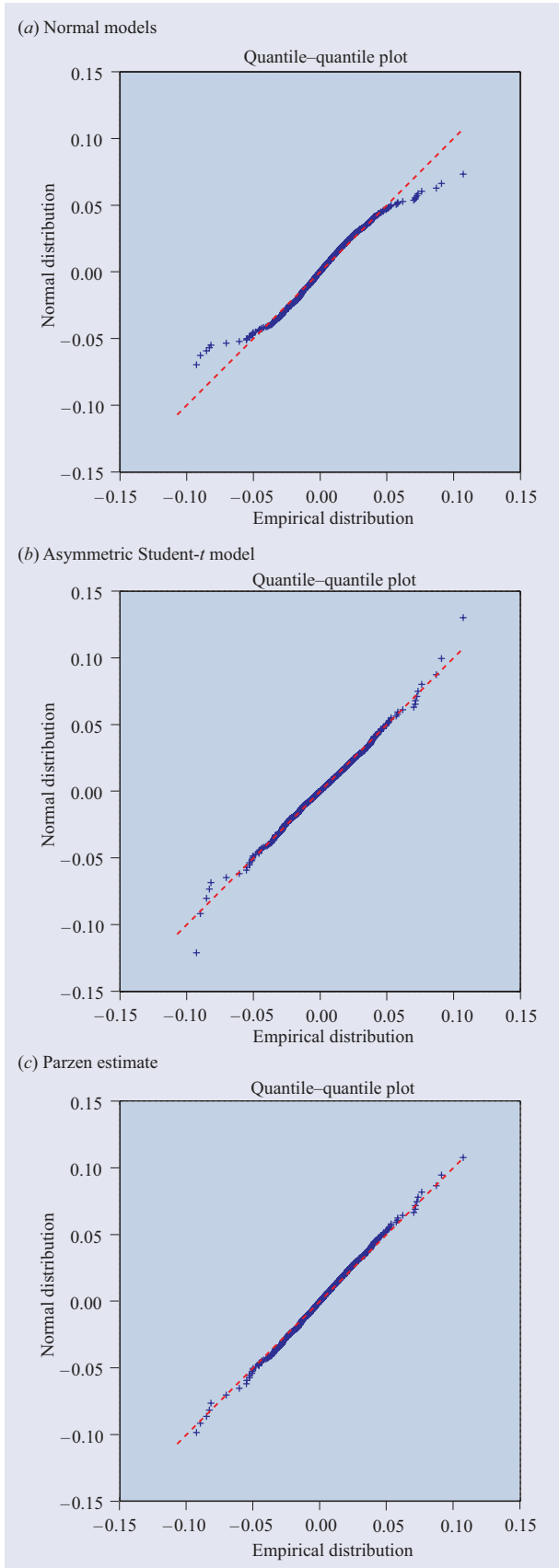


Figure 1. Quantile–quantile plots for Bell Canada Enterprises. Each plot shows the quantiles of the empirical distribution, the horizontal axis, for four years of returns compared to the fitted model, the vertical axis. If the two distributions had the same quantiles all the markers would line up on the dashed reference line.

The functions \widehat{f} and f^\vee are the Fourier transform and the inverse Fourier transform of f .

The probability density p_{π_i} has a singularity at the critical point of the quadratic function π_i . This means that unless the densities are discretized with a scheme that conserves probability the accuracy may be very poor. This effect is clearly illustrated by the example in figure 2. Let $[-a, a]$ be an interval and let $\{\xi_j = -a + jh\}_{j=0}^{N-1}$ where $h = 2a/N$. We discretize p_{π_i} by integrating over the neighbourhood of each grid point

$$p_{\pi_i} \approx p_i^D = h \sum_{j=1}^{N-1} p_i^j \delta_j \quad \text{where } p_i^j = \frac{1}{h} \int_{\xi_j-h/2}^{\xi_j+h/2} p_{\pi_i}(u) du. \tag{12}$$

Here δ is the delta function and $\delta_j = \delta(\xi_j - x)$. In our implementation, we chose to associate the grid point x_0 with both the left and the right end-points of the interval and to integrate over the intervals $[-a, -a + h/2]$ and $[a - h/2, a]$.

Let $\{P_i^j\}_{j=0}^{N-1}$ be the discrete Fourier transform of $\{p_i^j\}_{j=0}^{N-1}$ that defines $p_i^D(x)$. Provided that the interval $[-a, a]$ is large enough, the distribution of $\Delta\tilde{\Pi}$ can be approximated by

$$(p_{\pi_1} * \dots * p_{\pi_n})(x + \Xi) \approx p_{\Delta\tilde{\Pi}}^D(x) = h \sum_{j=0}^{N-1} p_j \delta_j(x + \Xi)$$

where

$$P^j = (-1)^{j(n-1)} h^{n-1} \prod_{k=1}^n P_k^j$$

and $\{P_j\}_{j=0}^{N-1}$ is the discrete Fourier transform of the coefficients $\{p_j\}_{j=0}^{N-1}$. In his thesis, Wiberg [29] shows that $p_{\Delta\tilde{\Pi}}^D$ converges weakly to $p_{\Delta\tilde{\Pi}}$. He has also proved that the error from approximating the continuous convolution with the circulant convolution is proportional to the probability outside $[-a, a]$ and the tail decay of the functions.

3.2. Step 3: inverse interpolation

In the previous subsection, we showed how to compute a discrete approximation

$$p_{\Delta\tilde{\Pi}}^D(x) = h \sum_{j=0}^{N-1} p_j \delta_j(x + \Xi),$$

to the pdf of $\Delta\tilde{\Pi}$. Given this approximation, the value-at-risk can be computed with inverse interpolation. First, add the coefficients to compute the cumulative distribution function at the grid points:

$$P_{\Delta\tilde{\Pi}}(\xi_k - \Xi) = h \frac{p_0}{2} + h \sum_{j \leq k} p_j$$

and

$$P_{\Delta\tilde{\Pi}}(\xi_N - \Xi) = P_{\Delta\tilde{\Pi}}(a) = hp_0 + h \sum_{j=1}^{N-1} p_j.$$

Since the first point of the grid corresponds to the interval $[-a, -a + h/2] \cup [a - h/2, a]$, we choose to assign half of the

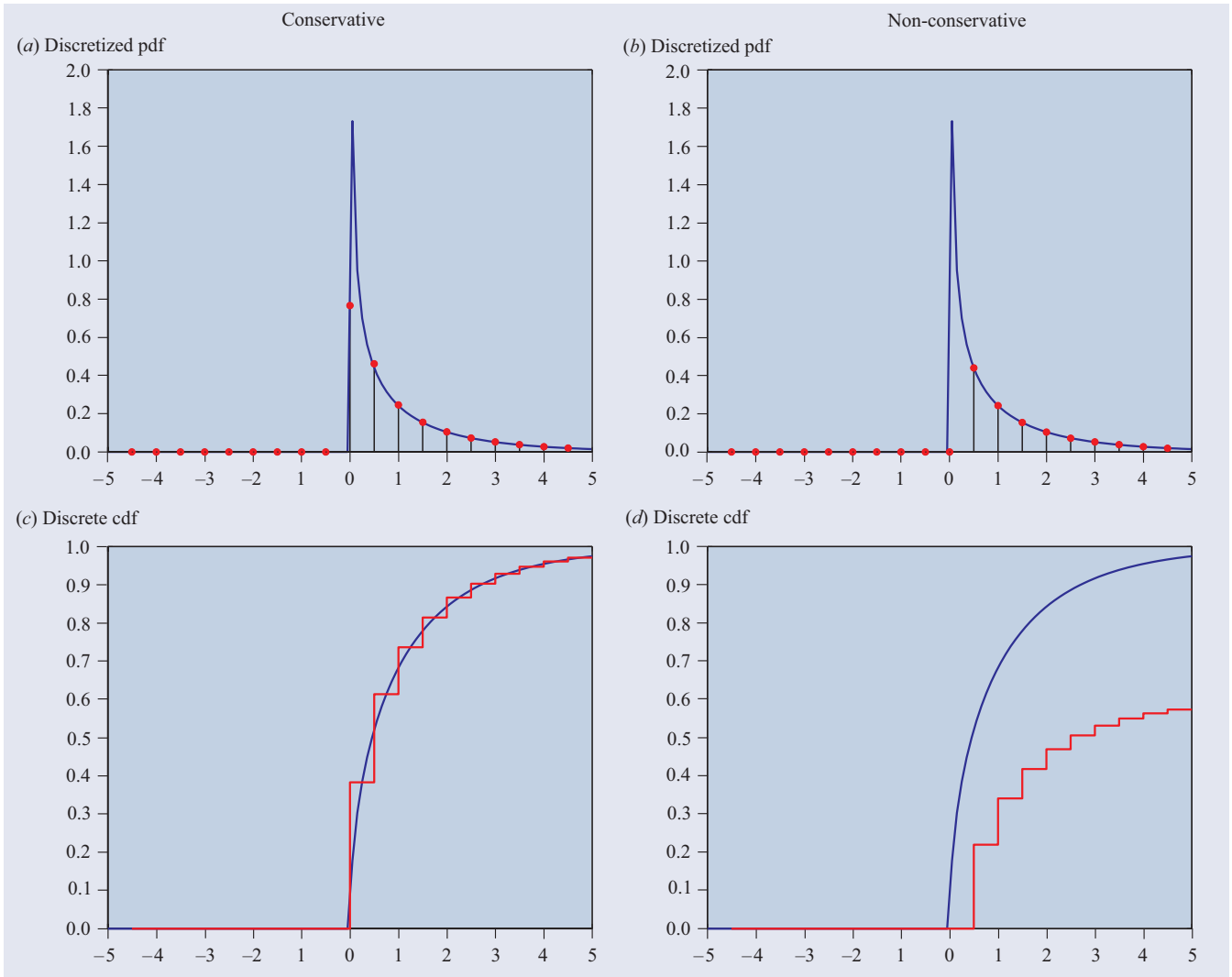


Figure 2. An example of the importance of a discretization method that conserves probability density. Consider $\pi = X^2$ where X is normal—a χ^2 random variable with one degree of freedom. For the conservative method, the discrete pdf and cdf are close to the pdf and cdf of the χ^2 random variable. In the non-conservative method, the pdf is sampled at the grid points to get a discrete approximation. As the graph for the cdf shows, the distribution function for the discrete approximation is not close to the original cdf.

density to the right grid end-point a and half to the left grid end-point $-a$. Second, find the index k such that

$$P_{\Delta\tilde{\pi}}(\xi_k) \leq 1 - \alpha \leq P_{\Delta\tilde{\pi}}(\xi_{k+1})$$

meaning that $-\text{VaR} - \Xi$ is in the interval $[\xi_k, \xi_{k+1}]$. Third, the linear interpolant to the inverse cdf is

$$L(p) = \xi_k + \left(\frac{p - P_{\Delta\tilde{\pi}}(\xi_k)}{P_{\Delta\tilde{\pi}}(\xi_{k+1}) - P_{\Delta\tilde{\pi}}(\xi_k)} \right) h,$$

$$p \in [P_{\Delta\tilde{\pi}}(\xi_k), P_{\Delta\tilde{\pi}}(\xi_{k+1})]$$

and the desired approximation to value-at-risk is

$$\text{VaR} = -L(1 - \alpha) - \Xi.$$

As an extension to the basic algorithm, the accuracy can be improved with Richardson extrapolation. The idea is to compute two solutions VaR_f and VaR_c on a fine grid with step size h and a coarse grid with step size $2h$, respectively. Then value-at-risk is taken to be $\text{VaR} = 2\text{VaR}_f - \text{VaR}_c$.

This procedure almost doubles the computational effort, but the error is reduced to close to $O(h^2)$ instead of $O(h)$, which means that a coarser grid size can be used to achieve the desired accuracy.

3.3. Computational complexity

The computational effort for all three steps depends on three parameters: (1) the number of dates d in the time series used for parameter estimation; (2) the dimension n of the model for changes in the risk factors; and (3) the number of grid points N . For the parameter estimation step, the computational work is dominated by the solution of the generalized eigenvalue problem (4) which takes $O(n^2 \min(n, d))$ floating point operations. For the second step, the discrete Fourier transform is the operation of highest complexity. For a sequence with N points, the discrete Fourier transform and its inverse can be computed with the fast Fourier transform (FFT) in $O(N \log N)$ floating point operations [7, 16]. To compute

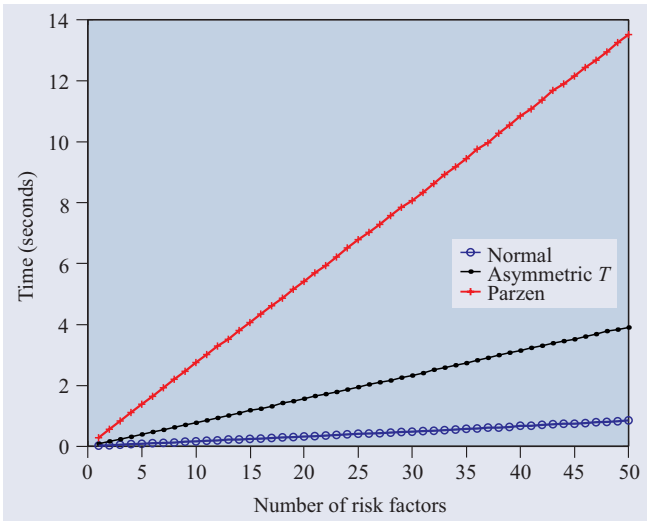


Figure 3. Computation time (in seconds) versus the number of risk factors n for the three return models from section 2. The number of dates in the time series is $d = 1000$ and the grid has $N = 4096$ points. As the figure shows, the difference between the models is large implying that the parameter estimation and discretization dominate the actual computation time.

the discrete approximation $p_{\Delta\pi}^D$ therefore requires a total of $O(nN \log N)$ floating point operations. The computational work in the final step is negligible compared to that in the first two steps. In the analysis, we have not accounted for the amount of computational work for estimating parameters and for discretizing different return models.

In theory, the complexity is dominated by the linear algebra in the first two steps, but the experiments presented in figure 3 paint a different picture. The plot shows how the computation time increases with the number of risk factors. The difference in computation time for the three return models lets us conclude that the majority of the computational work is in estimating the parameters and in discretizing the one-dimensional probability densities. So, in practice the computational work increases linearly with the number of risk factors and the rate of growth depends on the type of return model used.

Much of the interest in algorithms for value-at-risk is motivated by the high computational cost of Monte Carlo methods. It is well known that for Monte Carlo methods the error decreases as $O(1/\sqrt{m})$ where m is the number of samples (see, for example, Liu [22] for details). Figure 4 illustrates the different rates of convergence. The plot shows how the relative error decreases with increasing computation time, for a test portfolio with ten dimensions and assuming normally distributed returns. We are interested in the accuracy to which the methods resolve the approximate model (10), and therefore all methods use (2) to approximate the change in portfolio value. The figure shows that the relative error, as a function of CPU time, for the fast convolution method is smaller, and decreases at a faster rate, than the error for the Monte Carlo method. The figure also shows the additional accuracy gained by using Richardson extrapolation.

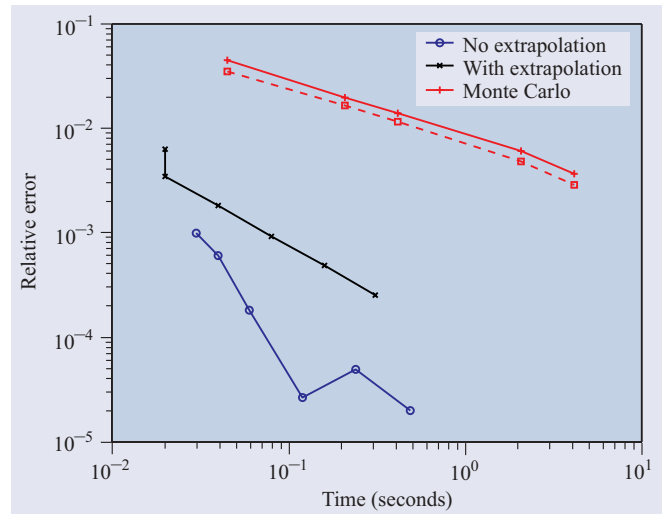


Figure 4. Computation time (in seconds) versus the relative error in the 99% value-at-risk for a sample portfolio with ten normally distributed risk factors. For the fast convolution method, the size of the grid ranges from 256 to 16 384 points and the lines show the relative error with and without extrapolation. For the Monte Carlo method, the number of samples ranges from 10^3 to 10^5 . The solid line shows sample estimates for the standard deviation; and the dashed line shows the sample mean for the relative error. The sample estimates are computed from 100 runs of the Monte Carlo method. For all three methods, the relative errors are approximated using a reference solution computed with the fast convolution method with extrapolation on a very fine grid.

4. Two numerical examples

We conclude with a discussion of two applications that illustrate the performance of the algorithm proposed in this article. In the first example, we consider a simple portfolio with derivatives and study the difference caused by using the three different return models from section 2.2. In the second example, we show that the value-at-risk surface for a model with volatility risk is qualitatively different from a model with constant volatility.

4.1. Comparison of the three models

Consider a portfolio containing one European call option on each of Bell Canada Enterprises (BCE) and Canadian Tire (CTRa), two large corporations with stocks traded on the Toronto Stock Exchange (TSX). Let the options be at-the-money and have three months to maturity. The Black–Scholes formulae give prices of \$3.23 and \$1.39 for the two options. We consider the situation of an investor who has sold the portfolio, meaning that Γ , the second derivative in (2) computed from the Black–Scholes formulae, is negative:

$$\Gamma = - \begin{pmatrix} 99.0967 & 0 \\ 0 & 42.6138 \end{pmatrix}.$$

In addition to the options, we assume that the portfolio contains positions in the underlying stocks; i.e., Δ in (2) is a parameter where Δ_1 and Δ_2 are proportional to the amount of BCE and CTRa in the portfolio.

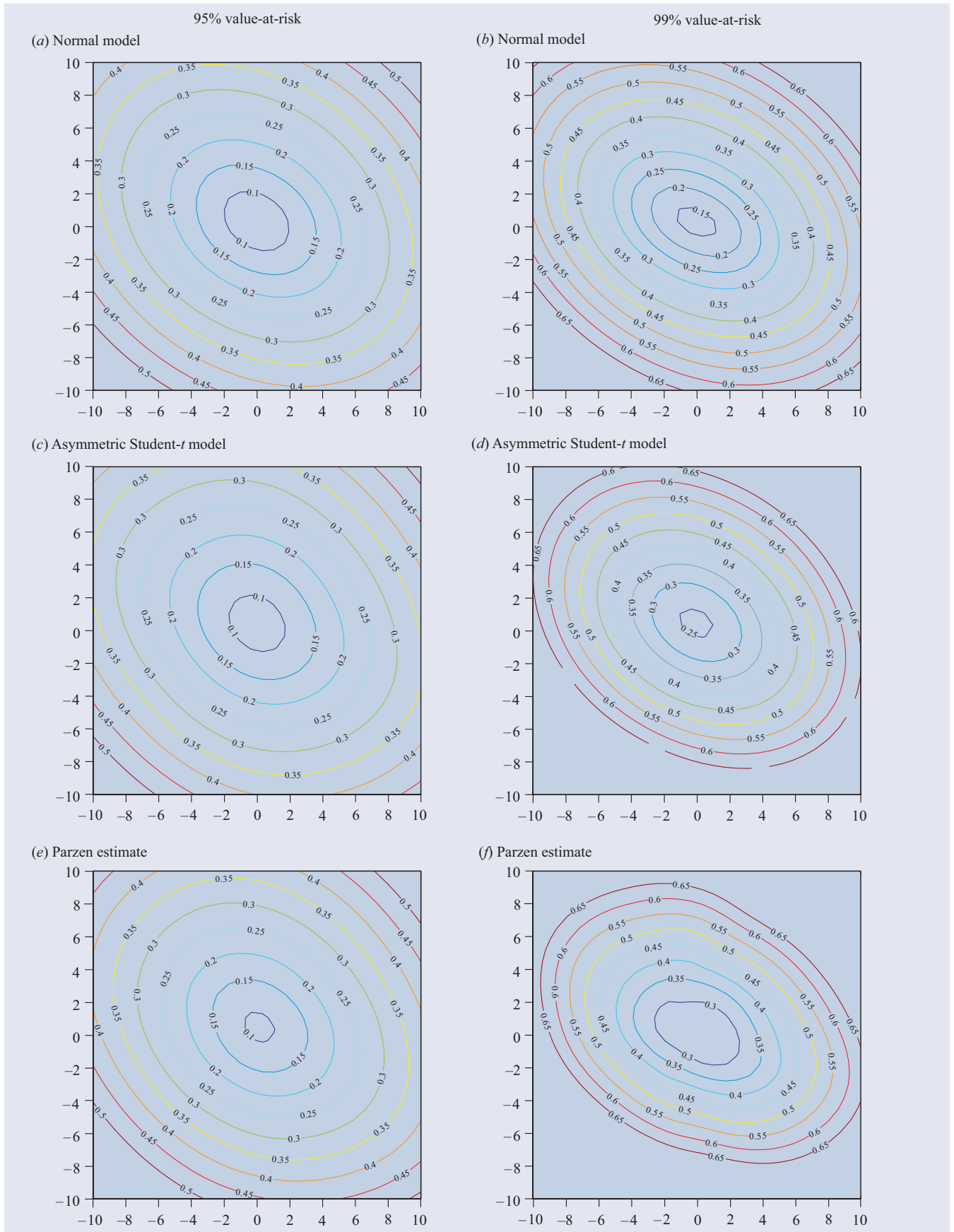


Figure 5. Value-at-risk for a portfolio, with two European call options and positions in the two underlying stocks, as a function of Δ_1 and Δ_2 . The horizontal axis is Δ_1 , the linear position in BCE; and the vertical axis is Δ_2 , the linear position in CTRa.

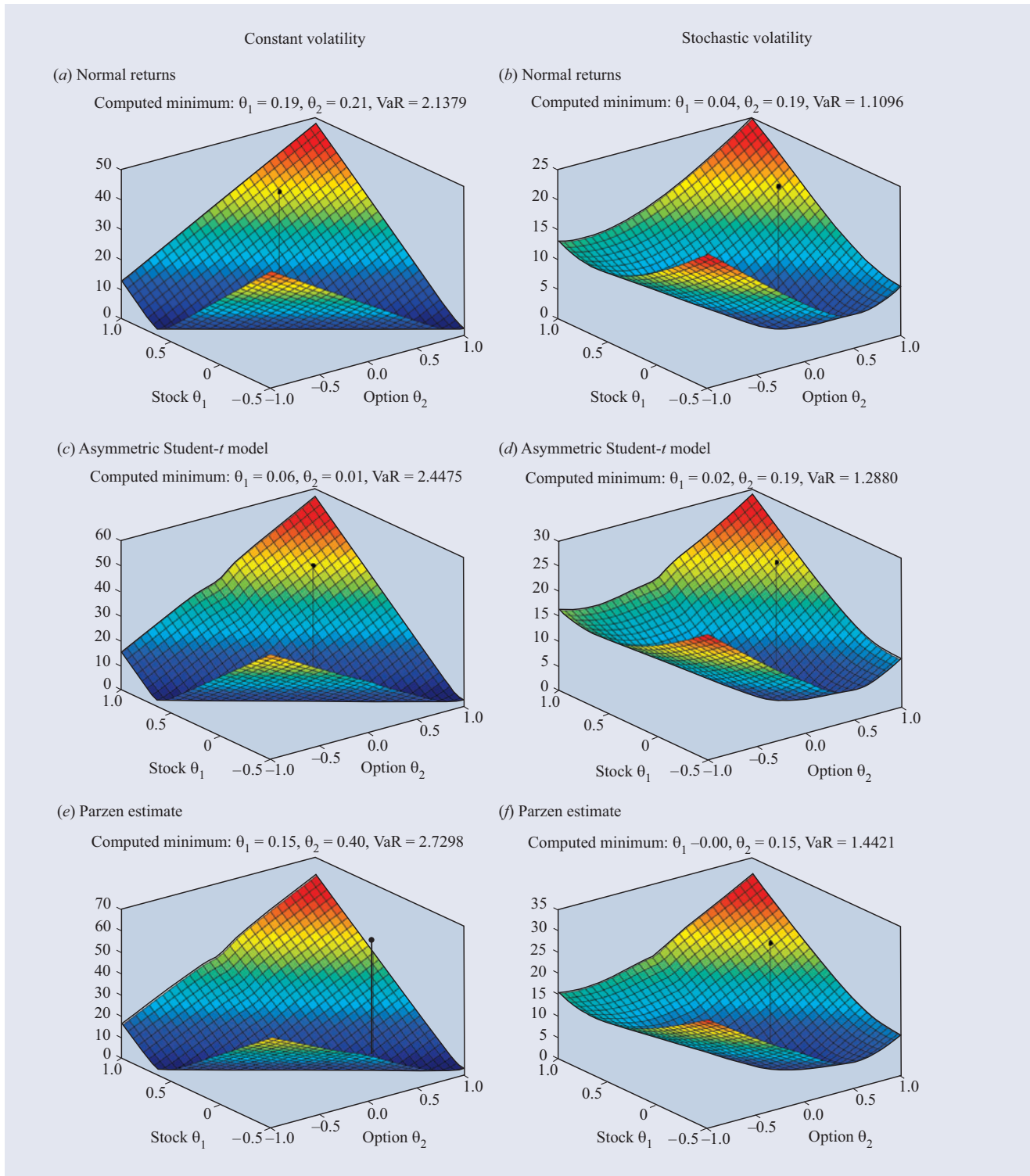


Figure 6. The 99% value-at-risk surface as a function of θ_1 , the position in Toronto 35, and θ_2 , the position in call options on Toronto 35. The point marked is the computed solution to the optimization problem. Volatility as a risk factor changes the shape of the value-at-risk surface and makes the minimum well defined.

To understand how the value-at-risk changes with the quantities of the underlying stocks, we compute the 95% and 99% values-at-risk on a grid with $-10 \leq \Delta_i \leq 10$ where $i = 1, 2$. Figure 5 shows the level sets for value-at-risk as a function of Δ . The figure shows value-at-risk for two different confidence levels, $\alpha = 0.95$ and 0.99 , and for each of the three

return models from section 2.2. Each graph is computed with a 30×30 grid for Δ and with $N = 4096$ grid points in the value-at-risk algorithm.

In the figure, we can see some interesting differences. The value-at-risk is smaller for the normal model compared to the

asymmetric T and Parzen models. For $\Delta = 0$, the relative differences are approximately 16% (38%) and 80% (92%) for the asymmetric T model (Parzen model) and for the 95% and 99% values-at-risk, respectively. The differences between the asymmetric T and Parzen model are about 19% and 7% for the 95% and 99% values-at-risk, respectively. The Parzen model gives a larger value-at-risk compared to both the other models. The level sets of the normal and asymmetric T models are elliptical whereas the level sets for the Parzen model display less symmetry. This example emphasizes the importance of the choice of return model for the simulation result. In particular, we note that figure 1 suggests that both the asymmetric T and the Parzen models are reasonable models that capture the tail behaviour of the data. At the same time, they give rather different estimates for the value-at-risk in this example.

4.2. Influence of volatility risk

In the second example, we want to compare the difference between a model with and without a risk factor for changes in the volatility. Similar to the case for previous examples we consider the portfolio of an investor who has written two European call options on BCE and Canadian Tire. The details are the same as for the first example. The Toronto 35 is an index consisting of stocks for 35 major Canadian companies that trade on the TSX. The index includes both BCE and CTRa and the returns have positive correlation. Suppose that to hedge the portfolio the investor can use both the index itself and an at-the-money call option on the index to hedge the portfolio. Let θ_1 and θ_2 be the number of index and call options on the index in the portfolio.

We now consider two alternative models for the local dynamics of the value of the portfolio. In the first model, the risk factors are the two stock prices and the index. Using the Black–Scholes model, the approximation (2) is

$$\begin{aligned} \Theta &= 0.0396 - 0.2354\theta_2, \\ \Delta &= \begin{pmatrix} -24.3797 \\ -10.4725 \\ 564.75\theta_1 + 321.71\theta_2 \end{pmatrix}, \\ \Gamma &= \begin{pmatrix} -99.0967 & 0 & 0 \\ 0 & -42.6138 & 0 \\ 0 & 0 & 2233.2\theta_2 \end{pmatrix}. \end{aligned}$$

The parameter Θ is the derivative with respect to time which gives $\Xi = \Theta/250$ in (2), assuming 250 trading days per year. In the second model, we add a fourth risk factor by making volatility random. Our goal is to study the qualitative effect that this has on the value-at-risk surface for the portfolio. We consider a simple model where the relative changes in volatility for all three risk factors are the same and we use the Black–Scholes equation to extract $\Delta\tilde{\Pi}$. Although this model is too simple to be of use in practice, it captures the main property of a stochastic volatility model. In particular, it introduces a risk factor that can only be hedged using the option. If we include first-order volatility risk this gives a new $\Delta\Gamma$ approximation for the portfolio:

$$\Theta = 0.0396 - 0.2354\theta_2,$$

$$\begin{aligned} \Delta &= \begin{pmatrix} -24.3797 \\ -10.4725 \\ 564.75\theta_1 + 321.71\theta_2 \\ -4.1963 + 22.0409\theta_2 \end{pmatrix}, \\ \Gamma &= \begin{pmatrix} -99.0967 & 0 & 0 & 0 \\ 0 & -42.6138 & 0 & 0 \\ 0 & 0 & 2233.2\theta_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

We computed the 99% value-at-risk for a range of values for the parameters θ_1 and θ_2 . The result for the experiment is displayed in figure 6. The plots show the value-at-risk surfaces for each of the three return models from section 2.2. The marker indicates the portfolio (θ_1, θ_2) that minimizes value-at-risk. We can draw several interesting conclusions. The addition of stochastic volatility changes the form of the value-at-risk surface. The constant volatility surface has a flat fold and the minimum value-at-risk portfolio is degenerate. Hence, hedging the portfolio using both the index and an index option does not lead to any significant reduction in the value-at-risk, compared to using the index alone. In the stochastic volatility model, the degenerate minimum is replaced by a well defined minimum. Furthermore, the portfolio with optimal value-at-risk combines a position in the index and the index option for reducing the risk, and using only the index gives a less efficient hedge.

5. Concluding remarks

In this paper, we have introduced a new Fourier algorithm for value-at-risk. The method can be used for portfolios with derivatives, provided that the local dynamics of the portfolio value can be approximated by a quadratic function, and for return models with fat tails. The fat tails are introduced via a product pdf. We discussed the details of the algorithm, emphasizing the importance of a discretization scheme that conserves probability. The performance of the method was demonstrated with the help of two applications. In the first example, we study the difference that the choice of return model has on the estimate for value-at-risk. In the second example, we compare a constant volatility model to a stochastic volatility model and show that the resulting value-at-risk surfaces have important qualitative differences.

References

- [1] Albanese C, Jackson K and Wiberg P 2002 Dimension reduction in the computation of value-at-risk *J. Risk Finance* **3** 41–53
- [2] Basel Committee on Banking Supervision 1988 International convergence of capital measurement and capital standards *Technical Report* Bank for International Settlements
- [3] Basel Committee on Banking Supervision 1996 Amendment to the Capital Accord to incorporate market risks *Technical Report* Bank for International Settlements
- [4] Basel Committee on Banking Supervision 1996 Overview of the amendment to the Capital Accord to incorporate market risks *Technical Report* Bank for International Settlements
- [5] Basel Committee on Banking Supervision 2001 The new Basel Capital Accord *Technical Report* Bank for International Settlements
- [6] Beder T S 1995 VAR: seductive but dangerous *Financial Analysts J.* **51** 12–24

- [7] Briggs W L and Henson V E 1995 *The DFT: an Owner's Manual for the Discrete Fourier Transform* (Philadelphia, PA: SIAM)
- [8] Britten-Jones M and Schaefer S M 1998 Non-linear value-at-risk *Eur. Finance Rev.* **2** 161–87
- [9] Campbell J Y, Lo A W and MacKinlay A C 1997 *The Econometrics of Financial Markets* (Princeton, NJ: Princeton University Press)
- [10] Duffie D and Pan J 1997 An overview of value-at-risk *J. Derivatives* **4** 7–49
- [11] Duffie D and Pan J 2001 Analytic value-at-risk with jumps and credit risk *Finance Stochastics* **5** 155–80
- [12] El-Jahel L, Perraudin W and Sellin P 1999 Value at risk for derivatives *J. Derivatives* **6** 7–26
- [13] Embrechts P, McNeil A and Straumann D 1999 Correlation and dependency in risk management: properties and pitfalls *Risk Management: Value at Risk and Beyond* ed M A H Dempster (Cambridge: Cambridge University Press) pp 176–223
- [14] Feuerverger A and Wong A C M 2000 Computation of value-at-risk for non-linear portfolios *J. Risk* **3** 37–55
- [15] Feuerverger A 1993 A consistent test for bivariate dependence *Int. Stat. Rev.* **61** 419–33
- [16] Gasquet C and Witomski P 1998 *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets* (Berlin: Springer)
- [17] Glasserman P, Heidelberger P and Shahabuddin P 2002 Portfolio value-at-risk with heavy-tailed risk factors *Math. Finance* **12** 239–69
- [18] Golub G H and Van Loan C F 1989 *Matrix Computations* 2nd edn (Baltimore, MD: John Hopkins University Press)
- [19] Hull J C 2000 *Options, Futures and Other Derivatives* 4th edn (Englewood Cliffs, NJ: Prentice-Hall)
- [20] Jorion P 1996 Risk²: measuring the risk in value at risk *Financial Analysts J.* **52** 47–56
- [21] Larsen R J and Marx M L 1986 *An Introduction to Mathematical Statistics and Its Applications* 2nd edn (Englewood Cliffs, NJ: Prentice-Hall)
- [22] Liu J S 2001 *Monte Carlo Strategies in Scientific Computing* (Berlin: Springer)
- [23] Mina J and Ulmer A 1999 Delta–gamma four ways *Technical Report RiskMetrics Group*
- [24] J P Morgan/Reuters 1996 *RiskMetrics—Technical Document* 4th edn <http://www.riskmetrics.com/rmconv.html>
- [25] Parzen E 1961 On estimation of a probability density function and mode *Ann. Math. Statistics* **33** 1065–76
- [26] Rosenblatt M 1956 Remarks on some nonparametric estimates of a density function *Ann. Math. Statistics* **27** 832–7
- [27] Ross S M 2000 *Introduction to Probability Models* 7th edn (New York: Academic)
- [28] Thompson J R and Tapia R A 1990 *Nonparametric Function Estimation, Modeling, and Simulation* (Philadelphia, PA: SIAM)
- [29] Wiberg P 2002 Computation of value-at-risk: the fast convolution method, dimension reduction and perturbation theory *PhD Thesis* Department of Computer Science, University of Toronto, Toronto, Canada