A new Fourier transform algorithm for value-at-risk

Claudio Albanese*
Ken Jackson†
Petter Wiberg‡

Abstract. In this paper, we introduce a new Fourier method for computing value-at-risk for a portfolios 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 support our approach. We discuss the details of the algorithm. The article 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.

Key words. Value-at-risk, mathematical finance, Fourier analysis, FFT, numerical methods.

AMS subject classifications. 91B28 (65T50,62P20,91B30)

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 unfavorable 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, 3, 4, 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 $\Delta\Pi$:

$$\mathrm{P}\left[\Delta\Pi \le -\mathrm{VaR}\,\right] = 1-\alpha. \tag{1.1}$$

In the special case that $\Delta\Pi$ is linear and the risk factor returns are multivariate normal there is closed form expression for the value-at-risk [23, 19]. 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 non-linear

* Department of Mathematics, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3, and Math Point Ltd., Toronto (www.math-point.com). E-mail: albanese@math.toronto.edu, phone: 416-978-5237.

† Department of Computer Science, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada M5S 3G4. E-mail: krj@cs.toronto.edu, phone: 416-978-7075.

‡ Communicating author. Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mail: wiberg@maths.warwick.ac.uk, phone: +44(0)-7652-2662.

function. Often the local error in linear approximations are too large to give accurate value-at-risk estimates, an effect that is exasperated by dynamic hedging strategies.

Because of the importance of value-at-risk simulations, several methods to compute value-at-risk for portfolios with derivatives have been proposed, see [23, 19, 10] 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 (1.1) can be approximated locally by a second order Taylor approximation,

$$\Delta\widetilde{\Pi} = \Xi + \Delta^T r + \frac{1}{2} r^T \Gamma r, \tag{1.2}$$

where the vector r contains the returns on each of the risk factors. One strategy, studied by several authors, is to use a return model where the characteristic function for $\Delta\widetilde{\Pi}$ can be derived explicitly. The distribution for $\Delta\widetilde{\Pi}$ can then be computed by inverting the Fourier transform numerically. Mina and Ulmer [22] 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's 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 [22, 11, 17]) 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 flavors. Britten-Jones and Schaefer [8] use an approximation to the distribution of a sum of non-central χ^2 random variables. Mina and Ulmer [22] 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 [15] explore a saddle point approximation method. They argue that, similar to 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 ideas with three example distributions. Section 3 presents the fast convolution method. We show that the method is efficient, although this 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. Modeling 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 modeled 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 its 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

$$p(r) = p_1(x_1) \cdots p_n(x_n), \tag{2.1}$$

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 Toronto Stock Exchange (TSX).

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

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

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, 23]. 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 (1.2) and the covariance matrix can be diagonalized simultaneously by solving a generalized eigenvalue problem:

$$\begin{cases} \Lambda = A^T \Gamma A, \\ C = AA^T. \end{cases} \tag{2.2}$$

The matrix A is nonsingular and Λ is a diagonal matrix with diagonal elements λ_i for $i = 1, \ldots, n$. This is a standard problem appearing in many different applications. We refer to [18] for discussion about how to solve eigenvalue problems numerically and to [1] for a discussion about how to take advantage of the special structure presented in the problem above. For the transformed variables $x = A^{-1}(r - \mu)$, the approximation (1.2) is

$$\Delta\widetilde{\Pi} = \widetilde{\Xi} + \sum_{i=1}^{n} \left(\widetilde{\Delta}_i x_i + \frac{\lambda_i}{2} x_i^2\right) \tag{2.3}$$

where

$$\widetilde{\Xi} = \Xi + \mu^T \Delta + \frac{1}{2}\mu^T \Gamma \mu, \tag{2.4}$$

$$\widetilde{\Delta} = A^T(\Delta + \Gamma\mu). \tag{2.5}$$

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 (2.1) and the remaining parameters for each p_i are estimated independently. With this estimation procedure, we can insure that the marginal distributions p_i are accurate in the canonical directions from (2.2).

In the remainder of this paper, we use *admissible model* to refer to a generic return model (2.1) 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 modeled by a product pdf (2.1) is a corner stone of the computational algorithm developed in Section 3. The family of admissible models has several important properties.

1. Historical time series show that returns typically are scattered around a central point in space which suggest that returns can be modeled 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 (2.1).
2. 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 (2.1). In particular, if the covariance matrix of X is the identity then the covariance matrix for the returns R and the sample covariance are the same.
3. 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 with fat tails.

In the value-at-risk literature, many other types of return models 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 (2.1) can be used to get an accurate model for returns for the *particular* portfolio at hand, see [28] 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. [14]). In the admissible models, the independence assumption implicit in (2.1) 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 Toronto35 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 (1.2) 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

	Rejections of hypothesis per confidence level		
	at 1%	at 5%	at 10%
24/05/2001-25/05/1999	18/595 (3.0%)	66/595 (11%)	108/595 (18%)
24/05/2001-25/05/2000	13/595 (2.2%)	39/595 (6.5%)	78/595 (13%)
24/05/2000-25/05/1999	14/595 (2.4%)	59/595 (9.9%)	102/595 (17%)

TABLE 2.1

Binomial test of independence for portfolios with two call options. The options are at-the-money and have 3 months to maturity. The frequencies in the table are the number of rejections of the null hypothesis, that the variables pass the independence test, out of all possible pairs of stocks in the Toronto 35 index.

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

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

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

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

This means that the null hypothesis can be refuted if the difference between q and its sample estimate $\widehat{q}$ is large. The estimates for the probabilities $\widehat{p}_1$, $\widehat{p}_2$ and $\widehat{q}$ are the ratio 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

$$(\widehat{q} - q)\sqrt{\frac{d}{q(1-q)}} \sim N(0, 1) \quad \text{as} \quad t \to \infty. \tag{2.6}$$

For large d, we can use this approximation and refute the null hypothesis, and consequently also independence, if (2.6) is outside a confidence interval.

The result from this statistical test is displayed in Table 2.1. We make two observations. First, the table shows that for our 595 test portfolios, a majority passes the independence test. It is expected that the rate of failure is 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 test 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 (2.1) is reasonable for risk factor returns.

2.2. Three examples of admissible models. There are many possible alternatives that could be used in (2.1). We have chosen to consider three instances.

First, the standard normal model. In this model,[1] the random variable for the transformed returns X is normal $N(\mu, \sigma^2)$ with the probability density function (pdf)

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

[1] 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.

Second, 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, 26]). The pdf is

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

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

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

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

$$\widehat{p}_d(x) = \frac{1}{dh}\sum_{i=1}^{d} K\left(\frac{x-x_i}{h}\right) \tag{2.7}$$

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

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

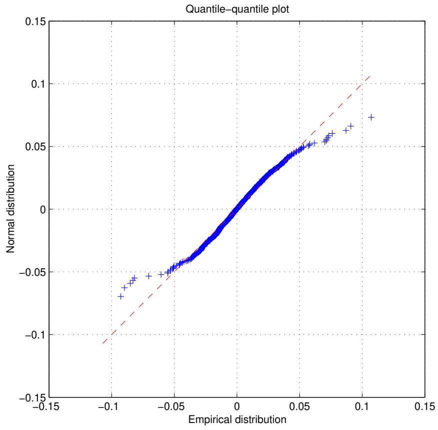
Density estimates of this form were first suggested in [25] and analyzed by Parzen in [24]. The weighting function $K(x)$ above is from [27].

Figure 2.1 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 model 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 [28] for details about this aspect of our implementation.

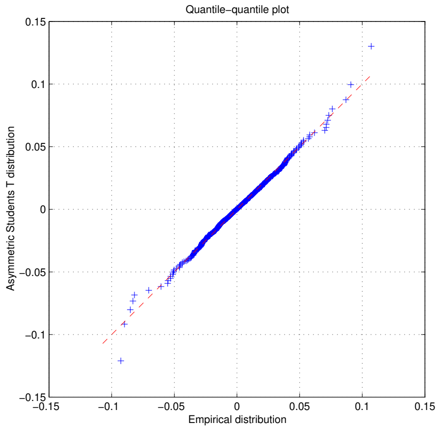
3. Computation of value-at-risk. The value-at-risk is the solution to (1.1) where $\Delta\Pi$, the change in portfolio value, is a non-linear function. To solve (1.1) directly, Monte Carlo methods are the only viable, but typically time consuming, strategy. However, if $\Delta\Pi$ can be approximated locally by a quadratic form (1.2) then we can substitute $\Delta\widetilde{\Pi}$ for $\Delta\Pi$ and approximate the value-at-risk VaR by the solution $\widetilde{\text{VaR}}$ to

$$\text{P}\left[\Delta\widetilde{\Pi} \le -\widetilde{\text{VaR}}\right] = 1-\alpha. \tag{3.1}$$

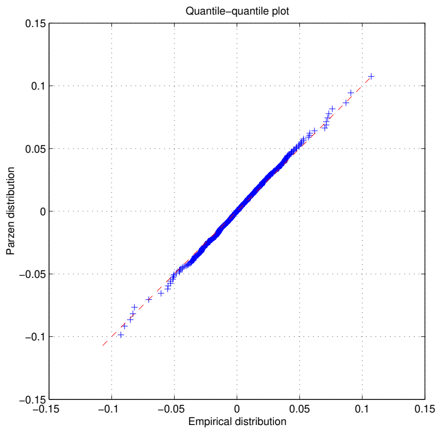
In this section, we introduce a method to solve (3.1). The method is based on the convolution theorem and can be applied to admissible return models. The advantage of our approach is that the method can cope with a wider class of return models,



(i) Normal model



(ii) Asymmetric Student T model



(iii) Parzen estimate

FIG. 2.1. *Quantile-quantile plots for Bell Canada Enterprises. Each plot shows the quantiles of the empirical distribution, the horizontal axis, for 4 years of returns compared to the fitted model, the vertical axis. If the two distributions have the same quantiles all the markers would line up on the dashed reference line.*

compared to the Fourier transform methods discussed in [22, 11, 17]. 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 step of the algorithm.

3.1. Step 2: discretization and convolution. For an admissible return model, (2.3) shows that the local approximation to the dynamics of portfolio value $\Delta\widetilde{\Pi}$ is a sum of quadratic functions,

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

To simplify our notation, we write VaR for $\widetilde{\text{VaR}}$ in (3.1) and similarly for (2.3). A consequence of the construction of the admissible return models is that the random variables π_i are independent. Therefore the pdf for $\Delta\widetilde{\Pi}$ is a convolution product:

$$p_{\Delta\widetilde{\Pi}}(x) = (p_{\pi_1} * \cdots * p_{\pi_n})(x + \Xi). \tag{3.2}$$

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

$$p_{\pi_1} * \cdots * p_{\pi_n} = \left(\prod_{i=1}^{n} \widehat{p_{\pi_i}} \right)^{\vee}.$$

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

The probability density for π_i has a singularity at the critical point of the function. This means that unless the densities are discretized with a scheme that conserves probability the accuracy may be very poor. This effect is clearly illustrate by the example in Figure 3.1. 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 neighborhood of each grid point

$$p_{\pi_i} \approx p_i^D = h \sum_{j=1}^{N-1} p_i^j \delta_j \quad \text{where} \quad p_i^j = \frac{1}{h} \int_{\xi_j - h/2}^{\xi_j + h/2} p_{\pi_i}(u) du. \tag{3.3}$$

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 endpoint 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 transforms 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 $\Delta\widetilde{\Pi}$ can be approximated by

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

where

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

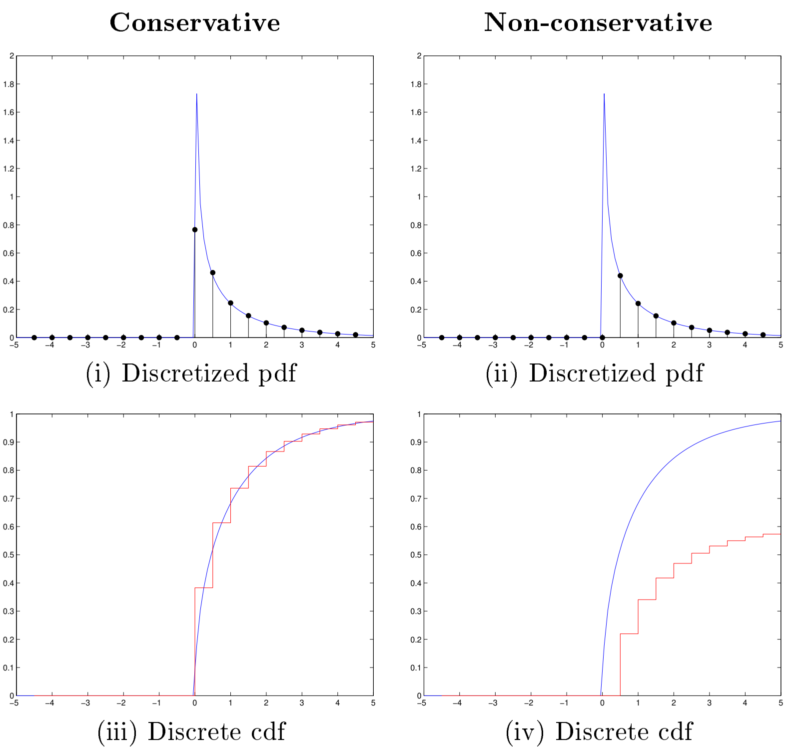


FIG. 3.1. *An example of the importance of a discretization method that conserves probability density. Consider $\pi = X^2$ where X is normal—a χ^2 random variable with one degree of freedom. For the conservative method, the discrete pdf and cdf are close to the pdf and cdf of the χ^2 random variable. 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 the original cdf.*

and $\{P^j\}_{j=0}^{N-1}$ is the discrete Fourier transforms of the coefficients $\{p_j\}_{j=0}^{N-1}$. In his thesis, Wiberg [28] shows that $p^D_{\Delta\widetilde{\Pi}}$ converges weakly to $p_{\Delta\widetilde{\Pi}}$. 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

$$p^D_{\Delta\widetilde{\Pi}} = h \sum_{j=0}^{N-1} p_j \delta_j(\bullet + \Xi),$$

to the pdf of $\Delta\widetilde{\Pi}$. 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:

$$P_{\Delta\widetilde{\Pi}}(\xi_k - \Xi) = h\frac{p_0}{2} + h\sum_{j\le k} p_j \quad \text{and} \quad P_{\Delta\widetilde{\Pi}}(\xi_N - \Xi) = P_{\Delta\widetilde{\Pi}}(a) = hp_0 + h\sum_{j=1}^{N-1} p_j.$$

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

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

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

$$L(p) = \xi_k + \left(\frac{p - P_{\Delta\widetilde{\Pi}}(\xi_k)}{P_{\Delta\widetilde{\Pi}}(\xi_{k+1}) - P_{\Delta\widetilde{\Pi}}(\xi_k)} \right) h, \qquad p \in [P_{\Delta\widetilde{\Pi}}(\xi_k), P_{\Delta\widetilde{\Pi}}(\xi_{k+1})].$$

and the desired approximation to value-at-risk is

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

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 $\mathrm{VaR} = 2\mathrm{VaR}_f - \mathrm{VaR}_c$. 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 (2.2) which takes $O(n^2 \min(n,d))$ floating point operations. For the second step, the discrete Fourier transforms 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 the discrete approximation $p^D_{\Delta\widetilde{\Pi}}$ therefore requires a total of $O(nN \log N)$ floating point operations. The computational work in the final step in negligible compared to 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 3.2 paints a different picture. The plot shows how the computation time increseas 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.

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.

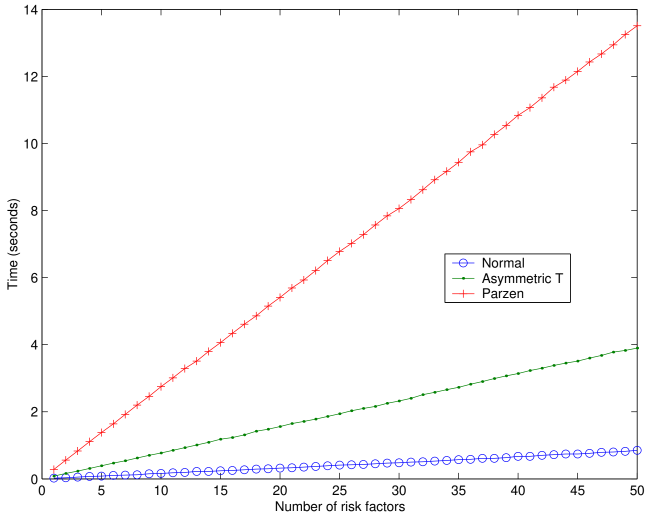


FIG. 3.2. *Computation time (in seconds) versus the number risk factors n for the three return models from Section 2. The number of dates in the time series is $d = 1000$ and the grid has $N = 4096$ points. As the figure shows, the difference between the models is large implying that the parameter estimation and discretization dominate the actual computation time.*

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 3 months to maturity. The Black-Scholes formula gives a price of \$3.23 and \$1.39 for each of the options. We consider the situation of an investor who has sold the portfolio, meaning that Γ in (1.2) is negative,

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

In addition to the options, we assume that the portfolio contains positions in the underlying stocks, i.e, Δ in (1.2) is a parameter where Δ_1 and Δ_2 are proportional to the amount of BCE and CTRa in the portfolio.

To understand how the value-at-risk changes with the quantities of the underlying stocks, we compute the 95% and 99% value-at-risk on a grid with $-10 \leq \Delta_i \leq 10$ where $i = 1, 2$. Figure 4.1 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 $\alpha = 0.99$, 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% value-at-risk, respectively. The differences between the asymmetric T and Parzen model are about 19% and 7% for the 95% and 99% value-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 emphasize the importance of the choice return model for the simulation result. In particular, we note that Figure 2.1 suggest that both the asymmetric T and the Parzen models are reasonable models that capture the tail behavior 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 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 Toronto35 is an index consisting of stocks for 35 major Canadian companies 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.

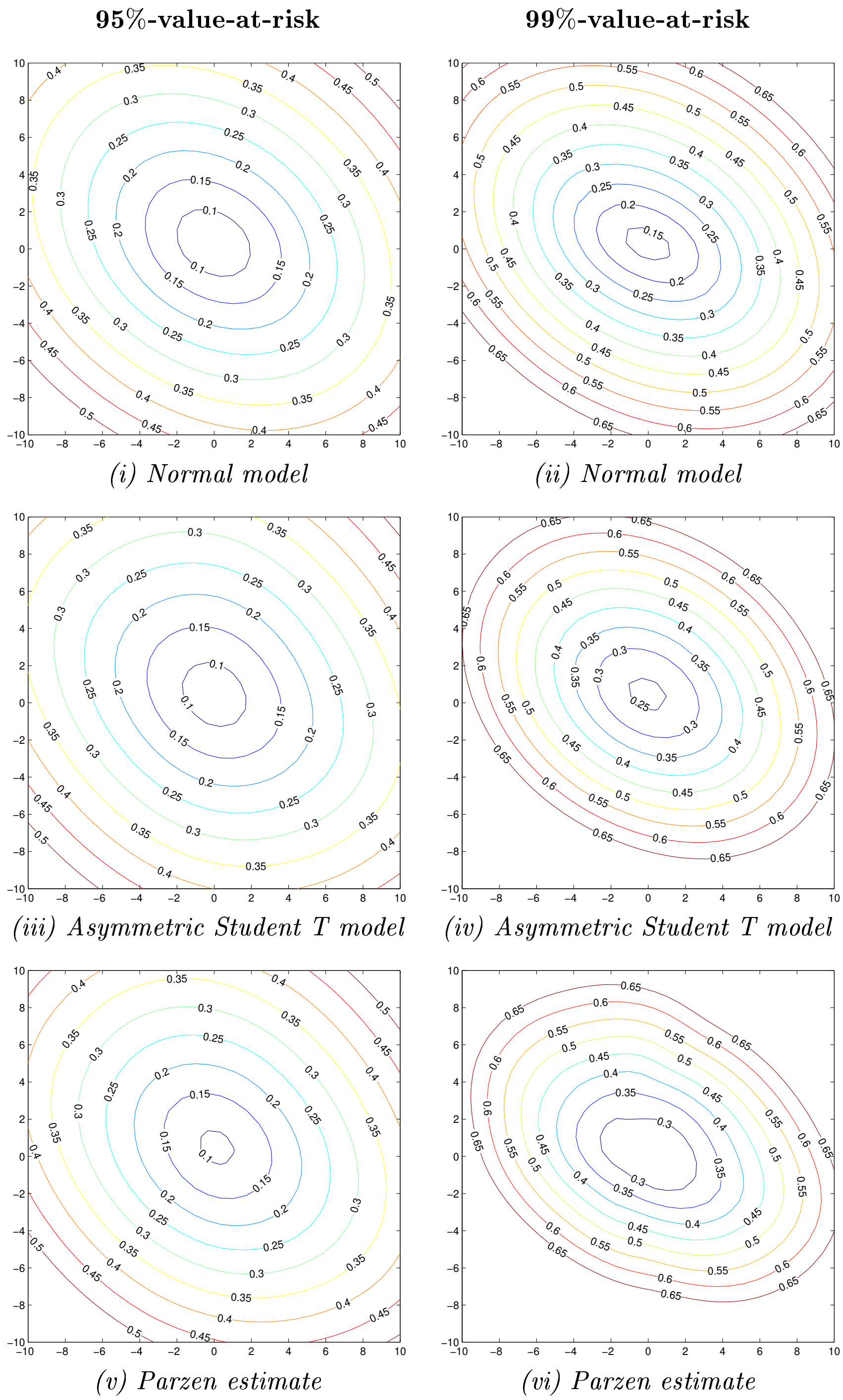
We now consider two alternative models for the local dynamics of the value of the portfolio. In the first model, the risk factors are the two stock prices and the index. The approximation (1.2) is

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

The parameter Θ is the derivative with respect to time which gives $\Xi = \Theta/250$ in (1.2), 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 this has on the value-at-risk surface for the portfolio. We consider a simple model where the relative change in volatility for all three risk factors is 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:

$$\Theta = 0.0396 - 0.2354 \cdot \theta_2$$
$$\Delta = \begin{bmatrix} -24.3797 \\ -10.4725 \\ 564.75 \cdot \theta_1 + 321.71 \cdot \theta_2 \\ -4.1963 + 22.0409 \cdot \theta_2 \end{bmatrix}$$
$$\Gamma = \begin{bmatrix} -99.0967 & 0 & 0 & 0 \\ 0 & -42.6138 & 0 & 0 \\ 0 & 0 & 2233.2 \cdot \theta_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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 4.2. The plots show

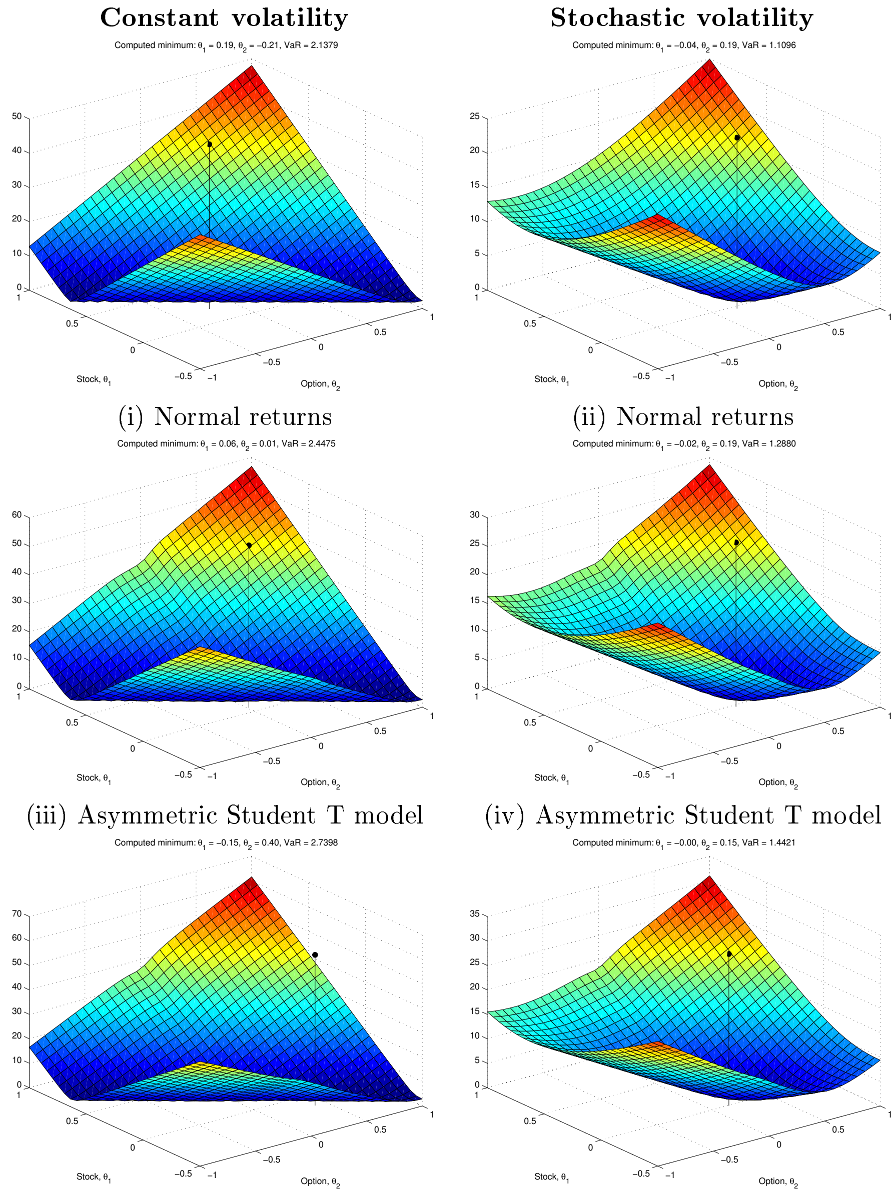


(i) Normal model *(ii) Normal model*

(iii) Asymmetric Student T model *(iv) Asymmetric Student T model*

(v) Parzen estimate *(vi) Parzen estimate*

FIG. 4.1. *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.*



(i) Normal returns (ii) Normal returns

(iii) Asymmetric Student T model (iv) Asymmetric Student T model

(v) Parzen estimate (vi) Parzen estimate

FIG. 4.2. *The 99% value-at-risk surface as a function of θ_1, the position in Toronto35, and θ_2, the position in call options on Toronto35. 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.*

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 to reduce 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 portfolio value can be approximated by a quadratic function, and for return models with fat-tails. The fat-tails are introduced by 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] C. Albanese, K. Jackson, and P. Wiberg. Dimension reduction in the computation of value-at-risk. *Journal of Risk Finance*, 3(4):41–53, Summer 2002.

[2] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards. Technical report, Bank for International Settlements, 1988.

[3] Basel Committee on Banking Supervision. Amendment to the Capital Accord to incorporate market risks. Technical report, Bank for International Settlements, 1996.

[4] Basel Committee on Banking Supervision. Overview of the amendment to the Capital Accord to incorporate market risks. Technical report, Bank for International Settlements, 1996.

[5] Basel Committee on Banking Supervision. The new Basel Capital Accord. Technical report, Bank for International Settlements, 2001.

[6] T.S. Beder. VAR: Seductive but dangerous. *Financial Analysts Journal*, 51(5):12–24, 1995.

[7] W.L. Briggs and V.E. Henson. *The DFT: An Owner's Manual for the Discrete Fourier Transform.* SIAM, 1995.

[8] M. Britten-Jones and S.M. Schaefer. Non-linear value-at-risk. *European Finance Review*, 2(2):161–187, 1998.

[9] J.Y. Campbell, A.W. Lo, and A.C. MacKinlay. *The Econometrics of Financial Markets.* Princeton University Press, 1997.

[10] D. Duffie and J. Pan. An overview of value-at-risk. *Journal of Derivatives*, 4(3):7–49, 1997.

[11] D. Duffie and J. Pan. Analytic value-at-risk with jumps and credit risk. *Finance and Stochastics*, 5(2):155–180, 2001.

[12] L. El-Jahel, W. Perraudin, and P. Sellin. Value at risk for derivatives. *Journal of Derivatives*, 6(3):7–26, 1999.

[13] P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependency in risk management: properties and pitfalls. In M.A.H. Dempster, editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, 1999.

[14] A. Feuerverger. A consistent test for bivariate dependence. *International Statistical Review*, 61(3):419–433, 1993.

[15] A. Feuerverger and A.C.M. Wong. Computation of value-at-risk for non-linear portfolios. *Journal of Risk*, 3(3):37–55, Fall 2000.

[16] C. Gasquet and P. Witomski. *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets.* Springer-Verlag, 1998.

[17] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Portfolio value-at-risk with heavy-tailed risk factors. *Mathematical Finance*, 12(3):239–269, July 2002.

[18] G.H. Golub and C.F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, USA, second edition, 1989.

[19] J.C. Hull. *Options, Futures and Other Derivatives.* Prentice Hall, 4th edition edition, 2000.

[20] P. Jorion. Risk^2: Measuring the risk in value at risk. *Financial Analysts Journal*, 52(6):47–56, 1996.

[21] R.J. Larsen and M.L. Marx. *An Introduction to Mathematical Statistics and Its Applications.* Prentice-Hall, second edition edition, 1986.

[22] J. Mina and A. Ulmer. Delta-Gamma four ways. Technical report, RiskMetrics Group, 1999.

[23] J.P. Morgan/Reuters. *RiskMetrics—Technical Document.* J.P. Morgan, 4th edition edition, 1996.

[24] E. Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33(3):1065–1076, September 1961.

[25] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27(3):832–837, September 1956.

[26] S.M. Ross. *Introduction to Probability Models.* Harcourt Academic Press, 7th edition, 2000.

[27] J.R. Thompson and R.A. Tapia. *Nonparametric Function Estimation, Modeling, and Simulation.* SIAM, 1990.

[28] P. Wiberg. *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, 2002.