Computation of value-at-risk for nonlinear portfolios

Andrey Feuerverger

Department of Statistics, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada

Augustine C. M. Wong

School of Analytic Studies and Information Technology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada

In this paper, the authors propose saddlepoint approximation methods for fast and accurate computation of *value-at-risk* in large complex portfolios. The method is applicable to portfolios whose value may be estimated by means of a "delta–gamma" approximation based on a large number of underlying risk factors whose random vector of returns has a known multivariate normal distribution for the time period under consideration. This method is not subject to the statistical uncertainty and computational expense of the Monte Carlo method. Some extensions of the method to higher-order portfolio approximations and to nonnormal risk factors are also given.

1. INTRODUCTION

The accurate determination of *value-at-risk* (VaR) is an important problem in modern financial applications. Current practice in this area, together with much relevant background material, is summarized, for example, in the *RiskMetrics Technical Document* (JP Morgan 1996) and by Jorion (1997). For realistic *nonlinear* portfolios, such work is often carried out using Monte Carlo trials. However, such computations can be very time and resource consuming. Furthermore, the accuracy of the method is usually limited to order $1/\sqrt{n}$ in the number n of trials performed.

In this paper, we develop a method for carrying out such computations more accurately and more quickly, without the need to rely upon Monte Carlo trials. Our method is based on analytical formulas derived from the moment generating function which allow us to produce very accurate estimates of VaR. Specifically, the method involves reducing delta–gamma approximations to appropriate quadratic forms to which highly accurate methods of saddlepoint approximation can be applied.

The technical problem is introduced in Section 2 below, while our main analysis is carried out in Section 3. Section 4 details the particulars of the saddlepoint method. The speed and accuracy of the proposed method on high-dimensional problems is demonstrated by numerical examples in Section 5. Finally, various extensions of our methods are indicated in Section 6, particularly extensions for including higher-order terms and for dealing with non-Gaussian risk factors. Some technical derivations are relegated to the Appendix.

Previous work on "analytical" methods for eliminating Monte Carlo trials in VaR work (using delta–gamma portfolio approximations) has been based on Fourier inversion methods.[1] Important contributions in this regard include those of Cardenas *et al*. (1997), Rouvinez (1997), Mina and Ulmer (1999), and Duffie and Pan (1999). In particular, Duffie and Pan extend the Fourier method to include both jumps and credit risk. See also Arvanitis *et al*. (1998), as well as references within the cited papers.[2] Remarks comparing Fourier and saddlepoint approximation methods for computing VaR are given in Section 6.

2. THE TECHNICAL PROBLEM

The technical problem we consider can be described as follows. Let $X = [X_1, \ldots, X_k]^{\mathrm{T}}$ be a random column vector representing the returns, over the single period of time considered, for the k underlying risk factors on the basis of which our portfolio is valued. In cases of interest, k may be quite large (e.g., $k = 500$, or more). It is assumed (initially) that, over the single time period in question, X has a multivariate normal distribution with zero mean vector (since the time interval is typically small) and variance–covariance matrix Σ:

$$X \sim N^k(0, \Sigma). \tag{1}$$

The matrix Σ is constant (in the given time period) and considered to be known. (Estimation of such covariance matrices is described in the *RiskMetrics Technical Document* and may involve GARCH and related methods.)

A large complex portfolio, possibly containing derivative securities, has a random return $g(X)$ (over the same time period) given by some function $g(\cdot)$. The function $g(\cdot)$ is determined by the holdings in the portfolio, and is a function of the returns on the individual assets in the portfolio. The returns on the individual assets are each considered to be known functions of X. Some of these will be simple linear functions, as, for example, when the portfolio has direct holdings in one or more of the k underlying risk factors. Others can be more complex; for example, when derivative securities are held in the portfolio, they may be nonlinear functions of X based on formulas such as the Black–Scholes formula. The function $g(\cdot)$ is, however, considered to be known. As an example, and to help fix ideas, if the portfolio only contains direct holdings in the k "risk-factor assets" (whose vector of returns is given by X) and if a is the column vector giving the dollar amounts invested in these various assets, then

[1] Other approaches that have been tried include Cornish–Fisher expansion and matching moments using the Johnson family of distributions. Of these, the first often lacks accuracy, while the second is not, in general, consistent. See, for example, Pichler and Selitsch (2000), Mina and Ulmer (1999), and references therein.

[2] The website `www.GloriaMundi.org` also refers to much current literature related to VaR.

we will have a linear return $g(X) = a^{\mathsf{T}}X$ for this portfolio. This case may be treated by elementary methods.

The more general problem of interest is this: given the known multivariate normal distribution for the returns X, and the known (but not necessarily linear) return function $g(\cdot)$ for the portfolio, determine the lower αth quantile of the distribution of $g(X)$. This quantile (often with $\alpha = 0.05$ or 0.01) is known as as the *value-at-risk* (VaR) and may be related to regulatory requirements[3] regarding reserve funding requirements. One approach to this problem is to sample X from $N^k(0, \Sigma)$ a large number of times, typically using a Cholesky decomposition of Σ, and to estimate the VaR from the αth quantile of the empirically obtained distribution for $g(X)$. This Monte Carlo approach is theoretically unbiased, but suffers in practice from several drawbacks. For instance, it can be difficult to carry out the Gaussian sampling when k is large, since the matrix Σ needs first to be Cholesky factored (or a square root found by alternate means), and sampling from $N^k(0, \Sigma)$ then requires repeated k-dimensional matrix–vector multiplications. Further, repetitive evaluations of complex g functions are themselves quite time consuming. Another drawback of Monte Carlo methods is that the resulting VaR estimate will itself vary between one "experiment" and the next, i.e., the final answer differs with every set of trials, especially when the number of trials is small. Last, but not least, the number of Monte Carlo trials required for estimating the αth quantile accurately, especially when α is small, can be surprisingly large. To illustrate this last point, Figure 1 shows four histograms which give lower $\alpha = 5\%$ and lower $\alpha = 1\%$ sample percentiles obtained from Monte Carlo samples of sizes $n = 200$ and $n = 1000$, respectively, taken directly from a standard normal distribution. (The histograms shown are each based on 2500 such simulations.) The true values for these percentiles, -1.645 and -2.326, respectively, are shown on the plots by a vertical line. These histograms can be interpreted as the sampling distribution for VaR estimates in the case of a simple portfolio that consists of a single asset having standard Gaussian return. Even in the best of these cases, namely for trials of size $n = 1000$ and the percentile $\alpha = 5\%$, the VaR measures obtained are quite variable, and themselves have a 95% spread range of $(-1.72, -1.57)$; when translated to a typical portfolio, the dollar amount of this Monte Carlo "error" will be quite large.

Evidently, the Monte Carlo approach for determining population percentiles is subject to considerable sampling variability.[4] To assure a VaR figure having a standard error of 0.01 in the present case would require a Monte Carlo sample of approximately 45 000 trials. The problem is even worse if the portfolio is a nonlinear function of the underlying assets, as each iteration involves a complicated function evaluation.

Monte Carlo computations for VaR can be speeded up to some extent, e.g., by using a simplifying approximation to the function g. Common among these

[3] Such as those imposed by the Basle Committee on Banking Supervision.

[4] See Jorion (1996) for a further discussion of this point.

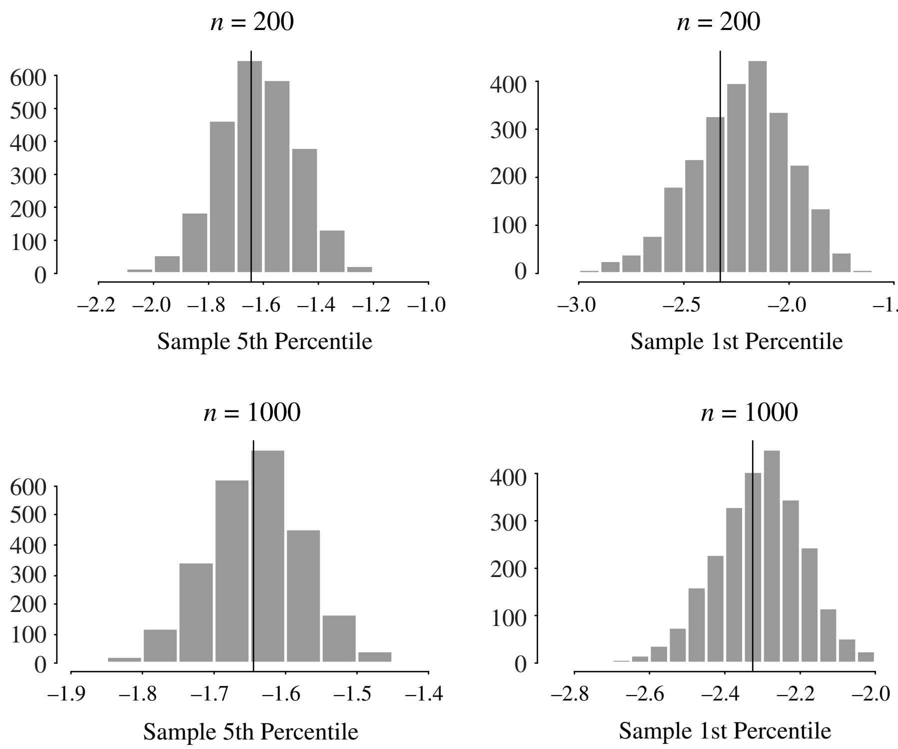

FIGURE 1. Histograms for 2500 lower 5% and lower 1% sample percentiles based on standard normal samples of size $n = 200$ and $n = 1000$.

is the so-called *delta–gamma approximation*. This involves first making Taylor series approximations for the value of each of the assets in the portfolio on which g is based. These component approximations are then summed over all assets in the portfolio to obtain the Taylor approximation for the overall portfolio. Since the components of X (as well as the higher terms of the Taylor approximation) are typically small, keeping only the linear and quadratic terms often yields a sufficiently precise approximation for the overall g function. For historical reasons, the linear terms are called *deltas*, while the quadratic terms are called *gammas*; the second-order Taylor expansion is known as a *delta–gamma approximation*. (When still higher-order terms are used, they too are labeled as *"greeks"*.) Nevertheless, even when using a delta–gamma approximation in place of g, the Monte Carlo approach can still be computationally demanding in portfolios which are large and based on many underlying assets.

3. SOME ANALYSIS FOR DELTA–GAMMA PORTFOLIOS

Let $X = [X_1, \dots, X_k]^{\mathsf{T}}$ be the vector of returns over one time period for our risk factors, and let $g(X)$ be the return for the portfolio of interest over that period. It is assumed that X follows the Gaussian distribution given in (1). A "delta–

gamma" approximation to $g(X)$ may then be written as[5]

$$Y = a_1^{\mathsf{T}} X + X^{\mathsf{T}} B_1 X, \tag{2}$$

where a_1 is a $k \times 1$ column vector, B_1 is a $k \times k$ matrix, and T denotes matrix transposition. (Note that we do not include a factor of $\frac{1}{2}$ with our quadratic term.) Both a_1 and B_1, as well as the covariance Σ in (1), are considered to be constant and known. The matrix B_1 is assumed to be symmetric; otherwise we replace it with $\frac{1}{2}(B_1 + B_1^{\mathsf{T}})$. For Monte Carlo simulation, X may be generated as $X = HZ_{(1)}$ using any H such that

$$\Sigma = HH^{\mathsf{T}}, \tag{3}$$

with $Z_{(1)}$ a $k \times 1$ column vector of *independent* standard normals.[6] For simulation, H is typically chosen to be lower triangular (the Cholesky factorization) to minimize the computations in $X = HZ_{(1)}$, but this is not a requirement below. It follows that (2) can be written as

$$Y = a_1^{\mathsf{T}}(HZ_{(1)}) + (HZ_{(1)})^{\mathsf{T}} B_1 (HZ_{(1)}) = a_2^{\mathsf{T}} Z_{(1)} + Z_{(1)}^{\mathsf{T}} B_2 Z_{(1)}, \tag{4}$$

where

$$a_2 = H^{\mathsf{T}} a_1 \quad \text{and} \quad B_2 = H^{\mathsf{T}} B_1 H. \tag{5}$$

Here, also, B_2 can be assumed to be symmetric. The portfolio is permitted to contain both long and short positions; for this and other reasons, the symmetric matrix B_2 need not be nonnegative definite.[7] It will, however, have real eigenvalues $-\infty < \lambda_1 \leqslant \cdots \leqslant \lambda_k < \infty$, and corresponding real orthonormal right-eigenvectors $P_1, \ldots, P_k$ which may be bound together columnwise to form the orthogonal matrix

$$P = \text{cbind}(P_1, \ldots, P_k).$$

In this notation, the singular-value decomposition for B_2 may be written as

$$B_2 = P \Lambda P^{\mathsf{T}} = \sum_{j=1}^{k} \lambda_j P_j P_j^{\mathsf{T}},$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_k)$ is the diagonal matrix formed from the eigenvalues. We next rewrite (4) as

$$Y = a_2^{\mathsf{T}} P P^{\mathsf{T}} Z_{(1)} + Z_{(1)}^{\mathsf{T}} P \Lambda P^{\mathsf{T}} Z_{(1)} = a^{\mathsf{T}} Z + Z^{\mathsf{T}} \Lambda Z, \tag{6}$$

[5] We are assuming here a zero-mean Gaussian distribution for X. For longer time horizons, it may be appropriate to assume that the distribution of the risk factors is of the form $X + \mu$, with X as in (1) and μ a given vector of means. If $X + \mu$ is used in place of X in (2), the expression is then easily reduced to one having the same form as (2), plus a constant. The remainder of the analysis is then essentially identical.

[6] Note that we will need the matrix H not only to make X independent but also for decomposition of the gamma matrix.

[7] The same assertion also holds true for B_1.

where

$$a = P^{\mathsf{T}} a_2 = P^{\mathsf{T}} H^{\mathsf{T}} a_1 \quad \text{and} \quad Z = P^{\mathsf{T}} Z_{(1)}. \tag{7}$$

Note that $Z = [Z_1, \ldots, Z_k]^{\mathsf{T}}$ also consists of independent standard normal entries. This allows us to write (6) in the equality in distribution form

$$Y \overset{\mathrm{d}}{=} \sum_{j=1}^{k} (a_j Z_j + \lambda_j Z_j^2). \tag{8}$$

Here, a_j are the entries of the vector $P^{\mathsf{T}} H^{\mathsf{T}} a_1$ and λ_j are the eigenvalues of $H^{\mathsf{T}} B_1 H$. The moment generating function of (8) is given by

$$M_Y(t) = \mathsf{E}[\mathrm{e}^{tY}] = \left(\prod_{j=1}^{k}(1 - 2\lambda_j t)\right)^{-1/2} \exp\left(\tfrac{1}{2}\sum_{j=1}^{k}\frac{a_j^2 t^2}{1 - 2\lambda_j t}\right) \tag{9}$$

$$= [\det(I - 2t\Sigma B_1)]^{-1/2} \exp[\tfrac{1}{2} t^2 a_1^{\mathsf{T}} (\Sigma^{-1} - 2tB_1)^{-1} a_1]. \tag{10}$$

See, for example, Mathai and Provost (1992); further details are given in the Appendix. If the maximum eigenvalue $\lambda_k > 0$, we have the constraint $t < (2\lambda_k)^{-1}$; if the minimum eigenvalue $\lambda_1 < 0$, we have the constraint $t > (2\lambda_1)^{-1}$. Altogether, $M_Y(t)$ will always be finite in an interval around the origin; in fact, the region of finiteness will be either a finite or semi-infinite interval, and will include the origin as an interior point.[8] The associated cumulant generating function is then given by

$$K(t) = \log M_Y(t) = -\tfrac{1}{2}\sum_{j=1}^{k}\log(1 - 2\lambda_j t) + \tfrac{1}{2}\sum_{j=1}^{k}\frac{a_j^2 t^2}{1 - 2\lambda_j t} \tag{11}$$

$$= -\tfrac{1}{2}\log\det(I - 2t\Sigma B_1) + \tfrac{1}{2} t^2 a_1^{\mathsf{T}} (\Sigma^{-1} - 2tB_1)^{-1} a_1, \tag{12}$$

while its first two derivatives (which will be required below) are

$$K'(t) = \sum_{j=1}^{k}\frac{\lambda_j}{1 - 2\lambda_j t} + \sum_{j=1}^{k}\frac{a_j^2(t - \lambda_j t^2)}{(1 - 2\lambda_j t)^2} \tag{13}$$

$$= \mathrm{tr}[B_1\Sigma(I - 2tB_1\Sigma)^{-1}] + a_1^{\mathsf{T}}(t\Sigma - t^2\Sigma B_1\Sigma)(I - 2tB_1\Sigma)^{-2} a_1 \tag{14}$$

and

$$K''(t) = \sum_{j=1}^{k}\frac{2\lambda_j^2}{(1 - 2\lambda_j t)^2} + \sum_{j=1}^{k}\frac{a_j^2}{(1 - 2\lambda_j t)^3} \tag{15}$$

$$= 2\,\mathrm{tr}(B_1\Sigma)^2(I - 2t\Sigma B_1)^{-2} + a_1^{\mathsf{T}}\Sigma(I - 2t\Sigma B_1)^{-3} a_1. \tag{16}$$

Further details, together with derivations, are given in the Appendix.

[8] The interval of finiteness can be large or small, but this does not affect our arguments in any way.

4. SADDLEPOINT APPROXIMATIONS FOR DELTA–GAMMA PORTFOLIOS

Consider first the classical problem involving identically and independently distributed random variables $X_1, \ldots, X_n$ drawn from a distribution whose cumulant generating function $\kappa(t)$ is finite on an interval for t that includes 0 in its interior. Then the *saddlepoint approximation* of Lugannani and Rice (1980) for the distribution function of the sample mean $\bar{X} \equiv (1/n) \sum_1^n X_i$ is given by

$$\mathsf{P}[\bar{X} > \bar{x}] = 1 - F_{\bar{X}}(\bar{x}) \sim 1 - \Phi(r) + \varphi(r)\left(\frac{1}{u} - \frac{1}{r}\right), \tag{17}$$

where Φ and φ are, respectively, the cumulative distribution and density functions of a standard normal variable; an alternative approximation, due to Barndorff-Nielsen (1986, 1991), is given by

$$1 - F_{\bar{X}}(\bar{x}) \sim 1 - \Phi\left(r - \frac{1}{r}\log\frac{1}{u}\right). \tag{18}$$

In both cases,

$$r = \pm\sqrt{2n}\,[\hat{\phi}\bar{x} - \kappa(\hat{\phi})]^{1/2} \quad \text{and} \quad u = \hat{\phi}[n\kappa''(\hat{\phi})]^{1/2}, \tag{19}$$

where the saddlepoint $\hat{\phi}$ is defined via the equation

$$\kappa'(\hat{\phi}) = \bar{x}, \tag{20}$$

and the sign of r is chosen to be the same as that of $\hat{\phi}$. Other tail area approximations are given by Daniels (1987). For further background, see also Barndorff-Nielsen and Cox (1979, 1989) and Reid (1996). The saddlepoint approximation[9] to the tail area of $\bar{X}$ is known to be extremely accurate, even for values of n as low as 3, 2, or even 1. Furthermore, it is exact when the underlying distribution is either normal, gamma, or inverse Gaussian. See, for example, Daniels (1980), Hampel (1974), Feuerverger (1989), and Ronchetti and Field (1990). This high degree of accuracy derives from the third-order error structure of the saddlepoint approximation and, specifically, from equalities[10] such as $\mathsf{P}[\bar{X} > \bar{x}] = 1 - \Phi(r) + \varphi(r)(u^{-1} - r^{-1}) + O(n^{-3/2})$.

[9] It is difficult to provide a simple intuitive explanation, based on (17) and (18), for the exceptional effectiveness of saddlepoint approximations. These approximations arise from certain asymptotic mathematical methods. Note, however, that saddlepoint approximations can sometimes be viewed as Edgeworth expansions applied at the mean value of an exponentially tilted density. If $f(x)$ is a density function, an exponentially tilted version is the density $e^{\theta x} f(x) \big/ \int_{-\infty}^{\infty} e^{\theta x} f(x)\,dx$. The parameter θ is chosen so that the mean of the tilted density is at that value x at which the approximation is desired; Edgeworth expansions are most accurate at the mean value. Regularity conditions, under which saddlepoint approximation methods hold, are discussed, for example, by Barndorff-Nielsen and Cox (1989) and Jensen (1995), and are satisfied in the instances we describe. The main requirement is the existence of the cumulant generating function in an interval which includes the origin in its interior.

[10] See, for example, Daniels (1987), Lugannani and Rice (1980), and Barndorff-Nielsen and Cox (1979, 1989).

The quantity (8) in our VaR application does not involve a sample mean or total; nevertheless, it does involve a significant degree of convolution, so that the saddlepoint method is again applicable with a high degree of accuracy. We shall amply demonstrate this point further below. Note, however, that because the convolution (8) does not consist of identically distributed quantities, it is necessary to modify the approximation formulas so that $K(t)$ now plays the role of $n\kappa(t)$. In this new, more relevant, notation, the saddlepoint formulas for the tail areas of (8) continue to be given by (17) and (18), except that (19) is replaced by

$$r = \pm\sqrt{2}\,[\hat{\phi}\bar{x} - K(\hat{\phi})]^{1/2} \quad \text{and} \quad u = \hat{\phi}[K''(\hat{\phi})]^{1/2}, \tag{21}$$

while (20) becomes[11]

$$K'(\hat{\phi}) = \bar{x}. \tag{22}$$

Here, K, K', and K'' are as given in (11), (13), and (15). If it is desired to compute (17) for $\bar{x}$ in the vicinity of the distribution mean (where $\hat{\phi}$ will be close to zero), then r and u will both be close to zero, causing numerical problems when evaluating

$$d = d(u, r) = \frac{1}{u} - \frac{1}{r}.$$

However, following Andrews, Fraser, and Wong (2000), and references therein, near $\hat{\phi} = 0$ we may use the approximation

$$d \sim -\frac{\alpha_3}{6\sqrt{n}} + \frac{\alpha_4 - \alpha_3^2}{24n} r, \tag{23}$$

where α_3 and α_4 are standardized cumulants;[12] alternatively, we may use the linear approximation $d = \hat{a} + \hat{b}r$, with $\hat{a}$ and $\hat{b}$ fitted (near the singularity) by simple linear regression. In the context of our $K(t)$ function, we use $n = 1$ in (23), with α_3 and α_4 as standardized cumulants of $K(t)$. Note that, at the singularity point, (23) gives $d = -\alpha_3/6\sqrt{n}$, leading to the value $\frac{1}{2} - \alpha_3/\sqrt{72\pi n}$ for the right-hand side of (17).

5. NUMERICAL STUDIES

Our computational methods were implemented on an SGI Challenge computer using the S-Plus statistical software (version 5.1) (see, e.g., Becker, Chambers, and Wilks 1988). The form (8) is an arbitrary linear combination of single degree of freedom noncentral chi-squared variates whose coefficients need not

[11] Note that the expressions (21) and (22) involve primarily a change in notation, with $K(t)$ replacing $n\kappa(t)$. Alternately, we may think of these expressions as giving the saddlepoint approximation for the case of a sample of size $n = 1$, but from the convolved distribution defined by $K(t)$.

[12] The jth standardized cumulant α_j is defined by κ_j/σ^j, where κ_j is the jth cumulant, and σ^2 is the second cumulant, i.e., the variance.

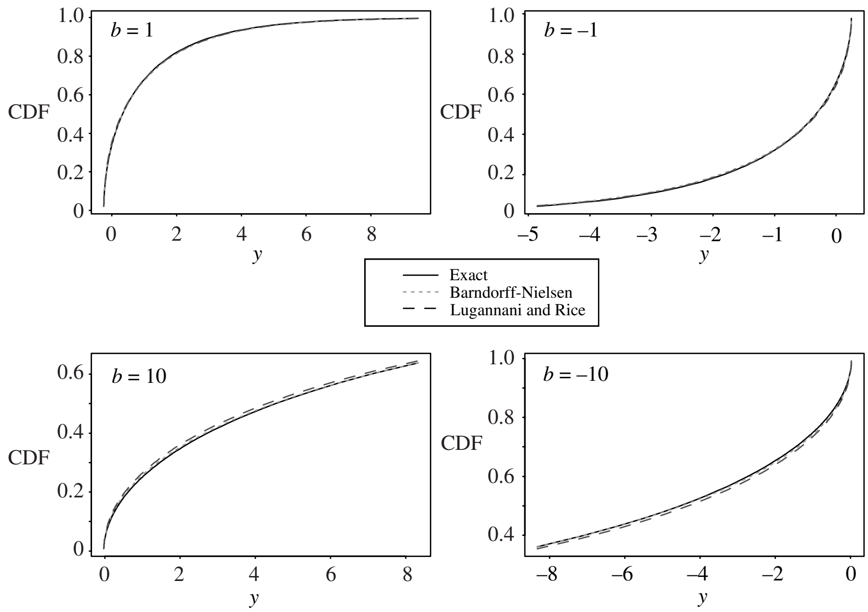

FIGURE 2. Exact cumulative distribution functions and their Barndorff-Nielsen, and Lugannani and Rice saddlepoint approximations for the distribution of $Z + bZ^2$ for $b = \pm 1, \pm 10$, with Z standard normal.

have the same sign. Owing to improvements in accuracy that result from convolution, the worst case scenarios—for quality of the saddlepoint approximations proposed here—will correspond to small values of the number k of terms in (8). In order to study this limitation in accuracy, we examined the worst-case scenario $k = 1$; specifically, we examined the quality of saddlepoint approximations to the distribution of a single term $Z + bZ^2$ for various values of b. Since $Z + bZ^2$ is quadratic, its exact tail probabilities are easily determined. Typical results are given in Figure 2. In this figure, for the indicated values of b, and, in each case, for the segment of the curve where the approximation error is greatest, the exact cumulative distribution function (CDF) is shown as a solid line, and superimposed upon this are the saddlepoint approximated cumulative distribution functions for both the Lugannani and Rice and Barndorff-Nielsen forms of the approximation, these being shown as dashed and dotted lines, respectively. The cumulative distribution function curves in these worst-case scenarios are seen to be extremely close, and are sometimes indistinguishable. Note also that, as b tends either to 0 or to $\pm\infty$, the quantity $Z + bZ^2$ will tend, respectively, towards the normal or chi-squared cases; however, for both of these distributions, the saddlepoint approximation is known to be exact. Furthermore (J. L. Jensen, private communication), since each of the third and higher cumulants of any convolution $U + V$, after standardization, is less than the largest of the corresponding cumulants of U and V, it follows that the normal approximation at the mean-value point for $U + V$ should (in some appropriate sense) be better than the worst of the normal approximations for each of U and V; consequently the saddlepoint approximation for $U + V$ should be no worse than the worst of the saddlepoint approximations for U and V. (Well-behaved distributions, of course, will do much better than this.) Our conclusion, therefore, is that the saddlepoint approximation will have a very high degree of accuracy for any portfolio for which the risk-factor returns are normally distributed with correctly specified covariance and for which a delta–gamma approximation is appropriate.

To demonstrate how our method performs numerically in the context of a large complex portfolio, we randomly generated an arbitrary nonnegative definite covariance matrix Σ of dimension 400×400; we also randomly generated an arbitrary vector a_1 of length 400 and an arbitrary (nonsymmetric) 400×400 random matrix B_1, where a_1 and B_1 correspond to the notation at (2). This numerical complexity corresponds to a large delta–gamma portfolio mapped onto 400 risk factors, all of which contribute nonlinearly to every holding. For this arbitrary "data", computation of the two saddlepoint approximated cumulative distribution functions, shown in Figure 3, took well under one minute of computer time. The two approximations are seen to coincide almost perfectly, and, although the exact distribution cannot be computed in this instance, we know by the foregoing analysis that the true curve should also be in nearly perfect coincidence with its saddlepoint approximants. The required VaR values may be read off (or interpolated) from such saddlepoint-based curves or, alternatively, may be determined very quickly by Newton–Raphson type procedures.

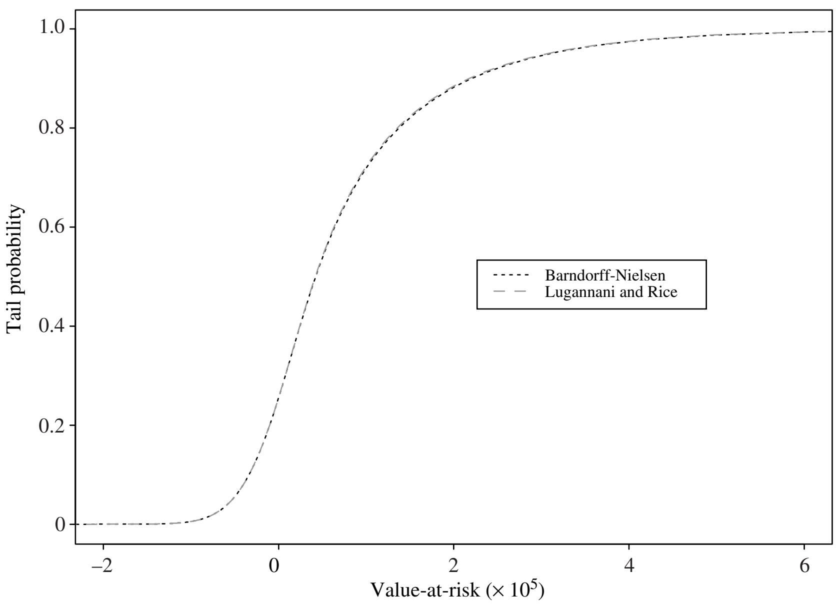

FIGURE 3. Barndorff-Nielsen, and Lugannani and Rice saddlepoint approximations for a (randomly generated) large portfolio.

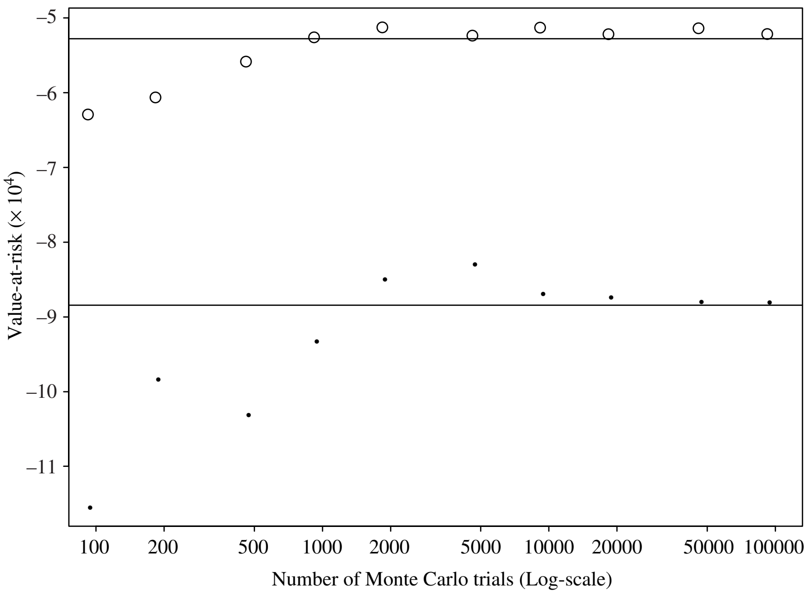

FIGURE 4. Comparison of the empirically determined 1% and 5% VaR values based on Monte Carlo trials (at various sample sizes), with the saddlepoint approximated VaR values (shown as horizontal lines), for our (randomly generated) large portfolio.

Figure 4 is based on the same randomly generated portfolio but involves up to 100 000 Monte Carlo evaluations on this portfolio. As these simulations evolved, we computed the empirical 1% and 5% VaR values for various numbers of trials and plotted them. The horizontal lines (at -5.3×10^4 and -8.8×10^4) give, respectively, the 5% and 1% VaR values as determined by the saddlepoint method, while the dots and circles give, respectively, the corresponding Monte Carlo determined VaR values at the various numbers of trials shown. As can be seen, in the limit of large numbers of Monte Carlo simulations, the empirically determined 5% and 1% VaR values are settling to those that were determined by the saddlepoint method.

6. EXTENSIONS AND REMARKS

In this section, we indicate some extensions of our methods to portfolios with more severe nonlinearities, and to non-Gaussian risk factors. We also comment on connections between the saddlepoint and Fourier methods.

6.1 *Higher-Order Effects*

For a portfolio with nonlinearities that cannot be adequately described by quadratic approximation, one possibility might be to divide it into two subportfolios, the first of which may adequately be regarded as quadratic, and the second of which is dealt with using Monte Carlo or other methods. The two resulting returns distributions (together with some assessment of their dependence) could then be "convolved" using *ad hoc* methods to obtain an approximate VaR estimate for the overall distribution. A more rigorous approach can be based on splitting the portfolio return function into two parts:

$$g(X) = g_1(X) + g_2(X), \tag{24}$$

where $g_1(X)$ is a quadratic approximation to the overall portfolio, while $g_2(X)$—the difference between $g(X)$ and $g_1(X)$, i.e., the error made by the quadratic pricing—is typically only a small part of $g(X)$. The desired cumulative distribution function of $g(X)$ can be written as

$$\mathsf{E}\big[\mathrm{I}\big(g(X) \leqslant c\big)\big] = \mathsf{E}\big[\mathrm{I}\big(g_1(X) \leqslant c\big)\big] + \mathsf{E}\big[\mathrm{I}\big(g(X) \leqslant c\big) - \mathrm{I}\big(g_1(X) \leqslant c\big)\big], \tag{25}$$

where I is the 0–1 indicator function and E is the expectation operator. The first term on the right is determined by the methods we have discussed. The second term on the right involves the expectation of a difference which will usually be 0, and which will only occasionally be $+1$ or -1. Hence, Monte Carlo evaluation of this expectation can be based on a reduced number of trials.[13] This approach estimates the cumulative distribution function of $g(X)$ simultaneously for all c, and smoothing can be applied across values of c to further improve accuracy.

Finally, we remark that higher Taylor series based portfolio approximations such as

$$g(X) = \sum a_i X_i + \sum\sum b_{i,j} X_i X_j + \sum\sum\sum c_{i,j,k} X_i X_j X_k + \sum\sum\sum\sum d_{i,j,k,l} X_i X_j X_k X_l + \cdots \tag{26}$$

can be handled by determining the first few cumulants of such expansions using: (1) linearity in the arguments of multivariate cumulant functions; (2) the Leonov–Shiryaev expansions for multivariate cumulants of products of random variables; and (3) the fact that multivariate cumulants of multivariate normal distributions are zero for cumulants beyond the covariance. See, for example, Brillinger (1975, §2.3) for details of computations of this type. With four (or more) cumulants thus available, we may then substitute the resulting Taylor expansion for the cumulant generating function into the saddlepoint approximation.

The asymptotic accuracy of saddlepoint approximations can be shown to carry over whenever at least four cumulants are used; see, for example, Fraser and Reid (1993). One possibility is to first obtain $K(t)$ using a delta–gamma approximation to the portfolio, and then add to it a polynomial to correct the first four (or more) cumulants. It is also worth remarking that the cumulants of (26) can also be computed for an empirical distribution of the X's (as would be obtained from historical data, for example); furthermore, since nonparametric kernel density estimates are just convolutions of a kernel function[14] with an

[13] For variance reduction techniques, see Fuglsbjerg (2000).

[14] The use of centered Gaussian kernels is obviously preferred here since these possess only a single nonzero cumulant.

empirical distribution, computation of the cumulants of (26) under such densities can be feasible as well.

6.2 Non-Gaussian Risk Factors

We next consider extensions to non-Gaussian risk factors. Since the distribution of portfolio returns under a mixture distribution for the risk factors is, in general, just the corresponding mixture of the portfolio returns under the components of the mixture, and since any distribution can in fact be approximated as a linear combination of Gaussians (a proof of this assertion can be based on Wiener's theorem concerning the closure of translates of functions having nonzero Fourier transform), then substantial generalizations to non-Gaussian risk factors are possible. Indeed, since convolutions are a special case of mixtures, and if the distribution of the risk factors can be modeled as the sum of a multivariate normal and an independent random vector, then it is sometimes possible to apply saddlepoint approximations to the resulting multivariate normal-based components and then to average these appropriately. As a special case of this, note that nonparametric kernel density estimates based on a multivariate normal kernel are just convolutions of the multivariate normal kernel with an empirical multivariate distribution, which in turn is just a mixture of normals with as many mixture components as data points. To illustrate the technique, one common "robustness" distribution involves a mixture of two multivariate normals, the first of which occurs, say 95% of the time, and the second of which occurs the remaining 5% of the time and has covariance matrix, say, 10 times larger than the first. In this case, we can compute the portfolio returns under each of the two normal distributions—using a saddlepoint approximation method in each case—and then "mix" the resulting cumulative distribution function approximations according to the same proportions.[15] Note that it is particularly fast and simple to recompute the saddlepoint approximations when the variance–covariance matrix is changed only by a constant multiple, say from Σ to $s^2\Sigma$, where s is a positive scale quantity. Under this change, H changes to sH, while a_2 and B_2 change to sa_2 and $s^2 B_2$, respectively. The matrix P of column-bound eigenvectors for the new B_2 remains unchanged, but the diagonal matrix Λ of eigenvalues changes to $s^2\Lambda$. Overall, the new representation for (8) involves a_j's that are s times larger, and λ_j's that are s^2 times larger. Consequently these quantities can be obtained essentially without additional computational labor, and so therefore can the associated transform quantities $M_Y(t)$, $K(t)$, and so on. Indeed, the new version of the function $K(t)$ in (11) can be obtained from the old version simply by replacing the argument t by $s^2 t$ and dividing the second term on the right in (11) by s^2. In this way, one can very efficiently obtain saddlepoint approximations for a large number of rescalings of the variance–covariance matrix Σ.

[15] If F_1 and F_2 denote the two resulting distribution function estimates in this case, then the resulting approximation will be $F = 0.95F_1 + 0.05F_2$.

More generally, consider a scale mixture generated by multiplying a multivariate $N^k(0, \Sigma)$ distributed vector X by a common random scaling factor S which has density function $h(s)$ and is independent of X. The corresponding version of (8) becomes

$$Y \stackrel{\mathrm{d}}{=} \sum_{j=1}^{k} \left(a_j S Z_j + \lambda_j S^2 Z_j^2 \right). \tag{27}$$

The moment generating function of this quantity is easily shown to be

$$M_1(t) = \int_0^\infty M_{U,V}(st, s^2 t) h(s) \, \mathrm{d}s, \tag{28}$$

where $M_{U,V}(t, u)$, the bivariate moment generating function of (U, V) with $U = \sum a_j Z_j$ and $V = \sum \lambda_j Z_j^2$, is given by

$$M_{U,V}(t, u) = \mathsf{E}[\mathrm{e}^{tU+uV}] = \left(\prod_{j=1}^{k} (1 - 2\lambda_j u) \right)^{1/2} \exp\left(\tfrac{1}{2} \sum_{j=1}^{k} \frac{a_j^2 t^2}{1 - 2\lambda_j u^2} \right). \tag{29}$$

For certain scale-mixture distributions $h(s)$, moment generating functions of the type (28) can readily be computed either analytically or computationally. If the scale-mixture distribution $h(s)$ is such that $M_1(t)$ is not finite (as would happen, for instance, if we tried to produce a multivariate t-distribution in this way), then the computation (28) can still be carried out provided that characteristic functions are used instead of moment generating functions; the resulting characteristic function can then be inverted by Fourier methods.

6.3 *Comparison with Fourier Methods*

Finally, we point out some comparisons between the saddlepoint approximation methods developed here and Fourier inversion.

1. These methods are very different mathematically. In particular, saddlepoint approximation involves only real-valued functions and elementary operations, while Fourier methods involve the FFT and numerical integration and are therefore somewhat more difficult to implement.
2. Fourier inversion methods can suffer from numerical inaccuracy in the far tails of the distribution; saddlepoint approximation methods do not and can in fact be used to "correct" Fourier inversion results in the tails.
3. Saddlepoint approximation methods require the existence of (but not necessarily full knowledge of) the moment generating functions, while Fourier methods do not.
4. Saddlepoint approximation methods are applicable when only four or more moments are available; for Fourier inversion the full characteristic function is required and cannot be replaced by a Taylor approximation without risk of serious error in the tails.

5. Fourier methods are now well developed, while saddlepoint approximation methods are currently under intensive development.
6. The two methods are best viewed as being complementary.

APPENDIX

In this appendix, we establish some transform characteristics for the distribution corresponding to (8). By direct integration, the joint characteristic function of (Z, Z^2), where Z is a single standard normal, is readily determined to be

$$\mathsf{E}[e^{itZ+iuZ^2}] = \frac{1}{\sqrt{1-2iu}} \exp\left(-\frac{1}{2}\frac{t^2}{1-2iu}\right).$$

It follows that the characteristic function of a single term of the form $aZ + \lambda Z^2$ is given by

$$\mathsf{E}[e^{it(aZ+\lambda Z^2)}] = \frac{1}{\sqrt{1-2i\lambda t}} \exp\left(-\frac{1}{2}\frac{a^2t^2}{1-2i\lambda t}\right),$$

and therefore that the characteristic function of (8) is given by

$$\varphi_Y(t) = \mathsf{E}[e^{itY}] = \left(\prod_{j=1}^{k} \frac{1}{\sqrt{1-2i\lambda_j t}}\right) \exp\left(-\frac{1}{2}\sum_{j=1}^{k}\frac{a_j^2t^2}{1-2i\lambda_j t}\right),$$

or, alternatively, by

$$\varphi_Y(t) = [\det(I - 2itB_2)]^{-1/2} \exp\left(-\frac{1}{2}\sum_{j=1}^{k}\frac{a_j^2t^2}{1-2i\lambda_j t}\right). \tag{A.1}$$

It is worth noting that expressions such as (A.1) can be written in matrix form using the original variables a_1 and B_1 of equation (2). To do this, we first note that

$$\begin{aligned}
\det(I - 2itB_2) &= \det(I - 2itH^{\mathsf{T}}B_1H) \\
&= \det\{H^{\mathsf{T}}[(H^{\mathsf{T}})^{-1}H^{-1} - 2itB_1]H\} \\
&= \det H^{\mathsf{T}} \det[(H^{\mathsf{T}})^{-1}H^{-1} - 2itB_1] \det H \\
&= \det \Sigma \det(\Sigma^{-1} - 2itB_1) \\
&= \det(I - 2it\Sigma B_1) \quad \text{or} \quad \det(I - 2itB_1\Sigma).
\end{aligned}$$

Note, secondly, that

$$\sum_{j=1}^{k} \frac{a_j^2}{1 - 2i\lambda_j t} = a^{\mathsf{T}}(I - 2\mathrm{i}t\Lambda)^{-1}a$$

$$= a^{\mathsf{T}} P^{\mathsf{T}} P(I - 2\mathrm{i}t\Lambda)^{-1} P^{\mathsf{T}} Pa$$

$$= a_2^{\mathsf{T}}[P(I - 2\mathrm{i}t\Lambda)P^{\mathsf{T}}]^{-1} a_2$$

$$= a_2^{\mathsf{T}}(I - 2\mathrm{i}tB_2)^{-1} a_2$$

$$= a_1^{\mathsf{T}} H(I - 2\mathrm{i}tH^{\mathsf{T}} B_1 H)^{-1} H^{\mathsf{T}} a_1$$

$$= a_1^{\mathsf{T}}(\Sigma^{-1} - 2\mathrm{i}tB_1)^{-1} a_1$$

$$= a_1^{\mathsf{T}}(I - 2\mathrm{i}t\Sigma B_1)^{-1}\Sigma a_1 \quad \text{or} \quad a_1^{\mathsf{T}}\Sigma(I - 2\mathrm{i}tB_1\Sigma)^{-1} a_1,$$

where we have used the fact that $\Sigma = HH^{\mathsf{T}}$. Therefore, altogether—in terms of the original variables a_1, B_1, and Σ—we may rewrite (A1) in the matrix form, say,

$$\varphi_Y(t) = [\det(I - 2\mathrm{i}t\Sigma B_1)]^{-1/2} \exp[-\tfrac{1}{2}t^2 a_1^{\mathsf{T}}(\Sigma^{-1} - 2\mathrm{i}tB_1)^{-1} a_1]. \tag{A.2}$$

We remark, in passing, that the distribution of (2) can be determined by numerical Fourier inversion of (A.1), as given, for example, by Feuerverger and McDunnough (1981). See, however, the last paragraph of Section 6. Observe, next, that the moment generating function of (8) is given by (9), or in matrix notation by (10). The associated cumulant generating function is then given by (11), or in matrix notation by (12), while its first two derivatives are as in (13) and (15).

We complete this appendix by showing how equations (13) and (15) can also be written in matrix notation involving only the original quantities a_1, B_1, and Σ. To show this, we consider the five types of terms arising in (13) and (15) and make repeated use of the facts that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, $a = P^{\mathsf{T}} H^{\mathsf{T}} a_1$, $\Sigma = HH^{\mathsf{T}}$, $H^{\mathsf{T}} B_1 H = P\Lambda P^{\mathsf{T}}$, and $P^{\mathsf{T}} P = I$. Within these derivations, we assume H^{T} to be invertible; observe, however, that the form of each final result is such that this invertibility requirement can be eliminated by elementary continuity arguments. Turning now to the five terms arising in (13) and (15), we have first

$$\sum_{j=1}^{k} \frac{\lambda_j}{1 - 2\lambda_j t} = \mathrm{tr}[\Lambda(I - 2t\Lambda)^{-1}] \tag{A.3}$$

$$= \mathrm{tr}[P\Lambda P^{\mathsf{T}} P(I - 2t\Lambda)^{-1} P^{\mathsf{T}}]$$

$$= \mathrm{tr}[H^{\mathsf{T}} B_1 H(I - 2tH^{\mathsf{T}} B_1 H)^{-1}]$$

$$= \mathrm{tr}[B_1 HH^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I - 2tH^{\mathsf{T}} B_1 H)^{-1} H^{\mathsf{T}}]$$

$$= \mathrm{tr}[B_1\Sigma(I - 2tB_1\Sigma)^{-1}]. \tag{A.4}$$

Second,

$$\sum_{j=1}^{k}\frac{a_j^2}{(1-2\lambda_j t)^2} = a^{\mathsf{T}}(I-2t\Lambda)^{-2}a \tag{A.5}$$

$$\begin{aligned}
&= a_1^{\mathsf{T}}HP(I-2t\Lambda)^{-1}P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}H(I-2tP\Lambda P^{\mathsf{T}})^{-1}(I-2tP\Lambda P^{\mathsf{T}})^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}H(I-2tH^{\mathsf{T}}B_1H)^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}HH^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}\Sigma(I-2tB_1\Sigma)^{-2}a_1 \quad \text{or} \quad a_1^{\mathsf{T}}(I-2t\Sigma B_1)^{-2}\Sigma a_1.
\end{aligned} \tag{A.6}$$

Thirdly,

$$\sum_{j=1}^{k}\frac{a_j^2\lambda_j}{(1-2\lambda_j t)^2} = a^{\mathsf{T}}\Lambda(I-2t\Lambda)^{-1}(I-2t\Lambda)^{-1}a \tag{A.7}$$

$$\begin{aligned}
&= a_1^{\mathsf{T}}HP\Lambda P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}HH^{\mathsf{T}}B_1HH^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}\\
&\quad\times(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}\Sigma B_1\Sigma(I-2t\Sigma B_1)^{-2}a_1.
\end{aligned} \tag{A.8}$$

Next,

$$\sum_{j=1}^{k}\frac{\lambda_j^2}{(1-2\lambda_j t)^2} = \operatorname{tr}(\Lambda\Lambda(I-2t\Lambda)^{-1}(I-2t\Lambda)^{-1}) \tag{A.9}$$

$$\begin{aligned}
&= \operatorname{tr}[P\Lambda P^{\mathsf{T}}P\Lambda P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}]\\
&= \operatorname{tr}[(H^{\mathsf{T}}B_1H)(H^{\mathsf{T}}B_1H)(I-2tH^{\mathsf{T}}B_1H)^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}]\\
&= \operatorname{tr}[(B_1H)(H^{\mathsf{T}}B_1H)H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}\\
&\quad\times(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}]\\
&= \operatorname{tr}[(B_1\Sigma)^2(I-2t\Sigma B_1)^{-2}].
\end{aligned} \tag{A.10}$$

And finally,

$$\sum_{j=1}^{k}\frac{a_j^2}{(1-2\lambda_j t)^3} = a^{\mathsf{T}}(I-2t\Lambda)^{-3}a \tag{A.11}$$

$$\begin{aligned}
&= a_1^{\mathsf{T}}HP(I-2t\Lambda)^{-1}P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}P(I-2t\Lambda)^{-1}P^{\mathsf{T}}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}H(I-2tH^{\mathsf{T}}B_1H)^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}HH^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}\\
&\quad\times(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}(H^{\mathsf{T}})^{-1}(I-2tH^{\mathsf{T}}B_1H)^{-1}H^{\mathsf{T}}a_1\\
&= a_1^{\mathsf{T}}\Sigma(I-2t\Sigma B_1)^{-3}a_1.
\end{aligned} \tag{A.12}$$

Substituting (A3)–(A12) in (13) and (15), we thereby obtain the matrix forms given in (14) and (16).

REFERENCES

Andrews, D. F., Fraser, D. A. S., and Wong, A. (2000). Higher order Laplace integration and the hyperaccuracy of recent likelihood methods. Submitted for publication.

Arvanitis, A., Browne, C., Gregory, J., and Martin, R. (1998). A credit risk toolbox. *Risk*, **11**(12), 50–55.

Barndorff-Nielsen, O. E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika*, **73**, 307–322.

Barndorff-Nielsen, O. E. (1991). Modified signed log likelihood ratio. *Biometrika*, **78**, 557–563.

Barndorff-Nielsen, O. E., and Cox, D. R. (1979). Edgeworth and saddlepoint approximations with statistical applications (with discussions). *Journal of the Royal Statistical Society* B, **41**, 279–312.

Barndorff-Nielsen, O. E., and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*. Chapman & Hall, New York.

Becker, R. A., Chambers, J. M., and Wilks, A. R. (1988). *The New S Language*. Wadsworth & Brooks/Cole, Pacific Grove, California.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory*. Holt, Rinehart & Winston, New York.

Cardenas, J., Fruchard, E., Koehler, E., Michel, C., and Thomazeau, I. (1997). VaR: One step beyond. *Risk*, **10**(10), 72–76.

Daniels, H. E. (1980). Exact saddlepoint approximations. *Biometrika*, **67**, 59–63.

Daniels, H. E. (1987). Tail probability approximations. *International Statistical Review*, **55**, 37–48.

Duffie, D., and Pan, J. (1999). Analytical value-at-risk with jumps and credit risk. Working Paper, Graduate School of Business, Stanford University.

Feuerverger, A. (1989). On the empirical saddlepoint approximation. *Biometrika*, **76**, 357–364.

Feuerverger, A., and McDunnough, P. (1981). On efficient inference in symmetric stable laws and processes. In: *Proceedings of the International Symposium on Statistics and Related Topics* (ed. Csorgo, Dawson, Rao, and Saleh), pp. 109–122. North-Holland, Amsterdam.

Fraser, D. A. S., and Reid, N. (1993). Simple asymptotic connections between density and cumulant functions leading to accurate approximations for distribution functions. *Statistica Sinica*, **3**, 67–82.

Fuglsbjerg, B. (2000). Variance reduction techniques for Monte Carlo estimates of value-at-risk. Working Paper, Financial Research Department., SimCorp A/S, Copenhagen, Denmark.

Hampel, F. (1974). Some small sample asymptotics. *Proceedings of the Prague Symposium on Asymptotic Statistics*, Charles University, Prague, 1973, Vol. 2, pp. 109–126.

Jensen, J. L. (1995). *Saddlepoint Approximations*. Oxford University Press.

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

Jorion, P. (1997). *Value at Risk*. McGraw-Hill, New York.

JP Morgan (1996). *RiskMetrics Technical Document*, 4th edn. Morgan Guarantee Trust Company.

Lugannani, R., and Rice, S. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables. *Advances in Applied Probability*, **12**, 475–490.

Mathai, A. M., and Provost, S. B. (1992). *Quadratic Forms in Random Variables, Theory and Applications*. Marcel Dekker, New York.

Mina, J., and Ulmer, A. (1999). Delta-gamma four ways. RiskMetrics Working Paper.

Pichler, S., and Selitsch, K. (2000). A comparison of analytical and VaR methodologies for portfolios that include options. In: *Model Risk, Concepts, Calibration, and Pricing* (ed. R. Gibson), pp. 252–265. Risk Books, London.

Reid, N. (1996). Likelihood and higher order approximations to tail areas: A review and annotated bibliography. *Canadian Journal of Statistics*, **24**, 141–166.

Ronchetti, E., and Field, C. (1990). *Small Sample Asymptotics*, Institute of Mathematical Statistics Lecture Notes, Monograph Series, 13. IMS, Hayward, California.

Rouvinez, C. (1997). Going greek with VaR. *Risk*, **10**(2), 57–66.