A semi-parametric approach to risk management

N.H. Bingham†and Rüdiger Kiesel‡¶ and Rafael Schmidt§

† Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH. E-mail: nick.bingham@brunel.ac.uk

‡ Department of Financial Mathematics, University of Ulm, Helmholtzstr 18, 89069 Ulm and Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE. E-mail: kiesel@mathematics.uni-ulm.de

§ Dept. of Number Theory and Probability Theory, University of Ulm, 89069 Ulm, Germany. E-mail: rafael.schmidt@mathematik.uni-ulm.de

Abstract. The benchmark theory of mathematical finance is the Black-Scholes-Merton (BSM) theory, based on Brownian motion as the driving noise process for stock prices. Here the distributions of financial returns of the stocks in a portfolio are multivariate normal. Risk management based on BSM underestimates tails. Hence estimation of tail behaviour is often based on extreme value theory (EVT). Here we discuss a semi-parametric replacement for the multivariate normal involving normal variance-mean mixtures. This allows a more accurate modelling of tails, together with various degrees of tail dependence, while (unlike EVT) the whole return distribution can be modelled. We use a parametric component, incorporating the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and a non-parametric component, which we can think of as a density on $[0, \infty)$, modelling the shape (in particular the tail-decay) of the distribution. We work mainly within the family of elliptically-contoured distributions, focussing particularly on normal variance mixtures with self-decomposable mixing distributions. We discuss efficient methods to estimate the parametric and non-parametric components of our model and provide an algorithm for simulating from such a model. We fit our model to several financial data series. Finally, we calculate Value at Risk (VaR) quantities for several portfolios and compare these VaRs to those obtained from simple multivariate normal and parametric mixture models.

Submitted to: *Quantitative Finance*

¶ To whom correspondence should be addressed

1. Introduction

Standard portfolio selection (based on Markowitz mean-variance theory of risk and return, from 1952 on, and the Sharpe-Lintner-Mossin capital asset pricing model (CAPM) of 1964-66) relates to the concepts of risk and return - that is, of the covariance matrix $\mathbf{\Sigma}$ of our vector of asset prices as well as its mean vector $\boldsymbol{\mu}$. The CAPM allows us to work with the covariance of each asset with some weighted average, or 'index', serving as a proxy for the market as a whole. However, the role of the (multivariate) normal, or Gaussian, model for distributions of financial returns (or log-prices) has been questioned on two principal grounds: the empirical evidence shows that most financial data show both pronounced asymmetry and much heavier tail behaviour than is consistent with normality. Given the importance of appropriate tail modelling in the context of modern risk management based on risk measures such as *Value at Risk (VaR)* this drawback has become even more relevant in recent years. Furthermore the role of the covariance matrix as the only measure of dependence has been questioned.

Because of this, there has been much interest in developing financial models which preserve at least some of the advantages of the classical normal case - familiarity, tractability, interpretability of mean vectors $\boldsymbol{\mu}$ and covariance matrices $\mathbf{\Sigma}$ - without the drawbacks of being unable to handle asymmetry or heavy tails - *skewness* and *kurtosis*. The normal-based theory remains, of course, extremely important, as a benchmark theory against which any theory must first be judged.

Other parametric models have been proposed - principally the stable and hyperbolic models (see §2). Reality is always more complicated than can be captured by just a few parameters, and so a non-parametric approach may also be adopted. The thesis of this paper is that the right way to seek to combine the advantages of the parametric and non-parametric approaches is to use a *semi-parametric* model. We use a *parametric* component, incorporating the mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$, and a *non-parametric* component, modelling the *shape* of the distribution - specifically, questions of tail-decay (kurtosis). Here, shape - which we can think of as a density on $[0, \infty)$ - incorporates what remains when we work up to location and scatter - that is, modulo affine transformations - while $(\boldsymbol{\mu}, \mathbf{\Sigma})$ represents the affine part. Especially in higher dimensions we observe that the shapes of most well-known parametric distributions are too restrictive for an appropriate fit to real financial data.

In §2 below we summarize the theoretical background, largely following Bingham and Kiesel (2002), to which this paper forms a sequel. We turn to dependence questions in §3, estimation in §4 and simulation in §5. We apply our results to a number of real financial data sets in §6 and close in §7 with an outline of further research directions.

2. The semi-parametric model

2.1. Normal variance-mean mixtures

One framework in which one can escape the limitations of the Black-Scholes-Merton framework while still retaining some of the convenience of the Gaussian case is that of *normal variance-mean mixtures* (NVMM)), for a good review of which see Barndorff-Nielsen, Kent, and Sørensen (1982). If F is a law on $[0, \infty)$, we sample U from F, and

$$\boldsymbol{X}|(U = u) \sim N_r(\boldsymbol{\mu} + u\boldsymbol{\beta}, u\Delta) \qquad (NVMM)$$

- where $\boldsymbol{\mu}, \boldsymbol{\beta}$ are r-vectors and Δ is a symmetric positive-definite $r \times r$ matrix with determinant one - then $\boldsymbol{X}$ is called a *normal variance-mean mixture* with *mixing law* F. If F has Laplace-Stieltjes transform Φ -

$$\Phi(s) := \int_0^\infty e^{-su} dF(u) \quad (s > 0)$$

- then $\boldsymbol{X}$ has characteristic function

$$\psi_{\boldsymbol{X}}(\boldsymbol{t}) = \exp\{i\boldsymbol{t}'\boldsymbol{\mu}\}\, \Phi\left(\frac{1}{2}\boldsymbol{t}'\Delta\boldsymbol{t} - i\boldsymbol{t}'\boldsymbol{\beta}\right). \qquad (NVMM')$$

2.2. Normal variance mixtures

If we specialize to $\boldsymbol{\beta} = 0$ in the above, we obtain the class of *normal variance mixtures*:

$$\boldsymbol{X}|(U = u) \sim N_r(\boldsymbol{\mu}, u\boldsymbol{\Sigma}); \quad \psi_{\boldsymbol{X}}(\boldsymbol{t}) = \exp\{i\boldsymbol{t}'\boldsymbol{\mu}\}\, \Phi\left(\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right). \qquad (NVM)$$

In this case the law of $\boldsymbol{X}$ is isotropic, and we may - and do - drop the restriction to unit determinant, imposed above for reasons of identifiability.

The class NVM of normal variance mixtures has important advantages over NVMM. One that concerns us is that structural information on the mixing law F transfers to the mixture law. For example, if F is *infinitely divisible* ($F \in ID$), so too is the law of $\boldsymbol{X}$, as follows immediately from (NVM).

Subclasses of ID are also important here. Call the law of $\boldsymbol{X}$ *self-decomposable* if for each $\rho \in (0, 1)$ there exists a characteristic function ψ_ρ for which

$$\psi_{\boldsymbol{X}}(\boldsymbol{t}) = \psi_{\boldsymbol{X}}(\rho\boldsymbol{t}) \cdot \psi_\rho(\boldsymbol{t}). \qquad (SD)$$

Then self-decomposable laws are infinitely divisible:

$$SD \subset ID,$$

and self-decomposability of F again transfers to self-decomposability of (the law of) $\boldsymbol{X}$. We write

$$SDNVM$$

for the resulting class of self-decomposable normal variance mixtures.
Examples.
1. Multivariate normal. The law $N_r(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is in SDNVM, with

$$\Phi(u) = \exp(-u/2).$$

2. Simple mixtures. As a simple first approach on can use a weighted average of normal distribution. Here the mixing law is discrete as e.g. in Hull and White (1998), who used a Bernoulli distribution.
3. Generalised hyperbolic laws. A particular choice of mixing law in the above - the *generalised inverse Gaussian* laws, GIG (a three-parameter family) - gives rise to the *generalised hyperbolic* laws, GH. Here Φ involves the quotient of Bessel functions K_λ of the third kind. For details and background and references, see Bingham and Kiesel (2001a). The class GH, and more generally NVM, has been used in many dimensions in the context of portfolio theory by Bingham and Kiesel (2001b) and Bauer (2000), or in particular Eberlein (2001), Eberlein and Prause (2002) and Eberlein and Özkan (2003) and references therein.

The common feature of the above examples is the use of a mixing distribution from a parametric family of distributions. We propose to estimate the mixing distribution non-parametrically, which allows us to use the full flexibility of mixture models (see §2.5 below for a detailed discussion).

We will be using the theory of elliptically-contoured distributions (below). It is more tractable but less general than NVMM. By contrast, Korsholm (2000) used NVMM for financial modelling in one dimension, where he was able to handle asymmetry.

2.3. Elliptically-contoured distributions

An r-dimensional distribution is *spherically symmetric* if it is invariant under the action of the orthogonal group $O(r)$. It is *elliptically symmetric*, or *elliptically contoured*, if it is the image of a spherically symmetric law under an affine transformation. We confine attention for convenience to the absolutely continuous, full-rank, L_2 case. Then the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ exist; $\boldsymbol{\Sigma}$ is invertible; the density $f(\boldsymbol{x})$ exists, and is a function of the quadratic form

$$Q := (\boldsymbol{x} - \boldsymbol{\mu}')\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) : \quad f(\boldsymbol{x}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}}\, g(Q) \qquad (EC)$$

for some scalar function g, called the *density generator* of $\boldsymbol{X}$. We write $f \sim EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. The main building block of this paper refers to the following *semi-parametric* elliptically-contoured model: The parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represent the **parametric part** of our model (mean and covariance, as in the Markovitz theory), while the density generator g is modelled as the **non-parametric part**. In particular, we are not restricted to the limited flexibility of any parametric density generator (as in the multivariate normal or generalized hyperbolic case). The analysis of real financial data in §6 reveals that this limitation might be crucial, especially in higher dimensions.

Note that the density generator contains information on the *shape* of the distribution - tail decay, etc., while escaping the curse of dimensionality. The definition of elliptically-contoured distributions does not require a distribution to belong to the L_2 or L_1 space. In that case $\boldsymbol{\Sigma}$ denotes some scaling matrix and $\boldsymbol{\mu}$ is a location vector.

The characteristic function has the form

$$\psi_{\boldsymbol{X}}(\boldsymbol{t}) = \exp\{i\boldsymbol{t}'\boldsymbol{X}\} \cdot \phi\left(\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right) \qquad (EC')$$

for some function ϕ called the *characteristic generator*. For background, see Fang, Kotz, and Ng (1990), Ch. 2, Bingham and Kiesel (2002). Observe that $NVM \subset EC$, with $\phi(u) = \Phi(u)$ in the notation above.

The class EC of elliptically-contoured laws shares many of the structural and closure properties of the multivariate normal. For our portfolio analysis we need the fact that linear combinations of elliptically-contoured random vectors are elliptically contoured. The mean vectors and covariance matrices behave as always under linear transformations $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$:

$$\boldsymbol{\mu} \mapsto \boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \quad \boldsymbol{\Sigma} \mapsto \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}',$$

and the density generator and characteristic generator are unchanged.

2.4. Self-decomposable elliptically-contoured laws.

There are a number of theoretical advantages in combining the properties of self-decomposability with being elliptically contoured: call the resulting class

$$SDEC.$$

(i) Self-decomposable distributions are absolutely continuous and unimodal (Sato (1999), Th. 28.4 and Th. 53.1). So for SDEC, the density generator *exists and is decreasing*. This provides a far-reaching and easily visualised generalisation of the Gaussian case, where

$$g(u) = \exp(-u/2)/(2\pi)^{\frac{1}{2}r}.$$

It also allows us to incorporate the structural information in our estimation procedure (see §4).

(ii) Self-decomposability is well adapted to modelling the *dynamic* (time-series) aspects of portfolios, as well as the *static* (distributional) aspects above. For, the defining property (SD) corresponds well to autoregressive time-series models

$$\boldsymbol{X}_t = \rho\boldsymbol{X}_{t-1} + \epsilon_t$$

with ϵ_t the innovation or error term. This link is developed in, e.g., Barndorff-Nielsen, Jensen, and Sørensen (1998).

(iii) In addition, for the normal variance-mixture case, self-decomposability transfers from the mixing law to the mixture law, as noted above:

$$NVM \subset EC, \quad SDNVM \subset SDEC.$$

Examples. 1. The generalised hyperbolic laws GH are self-decomposable (Halgreen (1979)), and - in higher dimensions - elliptically contoured:

$$GH \in SDEC.$$

See e.g. Bingham and Kiesel (2001b) for background and references.
2. The *variance gamma* laws are also in SDEC. For background, see Madan and Seneta (1990), Carr, Chang, and Madan (1998).
3. The multivariate t-distributions are in SDEC; see Fang, Kotz, and Ng (1990), §3.3.6.

2.5. Role of the density generator g

The density generator g relates to the shape of the data-generating distribution. In particular, the tail-decay of g reflects the tail-decay of the data itself, one of the prime features under investigation here. Thus, the tail behaviour of g relates most strongly to risk-management questions involving value at risk (VaR), for background on which we refer to Dowd (1998) and Jorion (2000).

The motivating data sets for us are time series of financial asset returns. Here the distributional aspects, and tail behaviour in particular, depend on the frequency with which returns are calculated (see e.g. Pagan (1996)). Since returns over long periods are sums of returns over shorter periods, the central limit theorem will give a tendency towards normality – 'aggregational Gaussianity'. This is indeed observed, for returns over longer periods than about 16 trading days – monthly returns, say – and accords with benchmark Black-Scholes-Merton theory. At the other extreme, very high-frequency data – returns over the order of minutes, say – display power-law decay, for reasons involving self-similarity and scaling arguments as in physics; see e.g. Dacorogna, Müller, and Pictet (1998) and Schmitt, Schertzer, and Lovejoy (1999) for background. Intermediate between these are, say, daily returns, for which the log-linear tails of the hyperbolic/NIG model are often used, etc.

For power-law decay, one may appeal also to the methods of extreme-value theory (EVT: for background, see e.g. Embrechts, Klüppelberg, and Mikosch (1997)), and in particular the Hill-type estimators (for a recent application to finance see Kiesel, Perraudin, and Taylor (2003)). However, EVT is only suited to model the tails of the distribution, while the semi-parametric approach allows to model the whole return distribution.

The tail behaviour of g is best regarded in context. What really matters is the behaviour of the whole curve g itself, regarded as a member of the continuum of curves obtained as the return interval varies.

Several important questions, however, relate not to g but to the parametric part $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ of our model. For example a decision on whether or not to include an extra asset in our portfolio is best answered by looking at $\boldsymbol{\mu}, \boldsymbol{\Sigma}$: a good candidate asset should show strong growth, and strong negative correlation with existing assets. Preferences between individual assets can also be settled by looking at $\boldsymbol{\mu}, \boldsymbol{\Sigma}$: see Embrechts, McNeil, and Straumann (2001), §3.4.

To summarize: g focuses on properties of the portfolio as a whole, while $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ focuses on properties of the individual assets within the portfolio. Thus, if we wish to rank our preferences between the individual assets in a fixed basket of assets, our preferences are the same in any fixed elliptically-contoured world as in the normal/Gaussian world (since $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ characterize any coherent risk measures in such settings, see Embrechts, McNeil, and Straumann (2001), §3.4.). By contrast, if we compare the g of a typical 'old economy' (of Germany's DAX, the USA's DOW JONES, etc.) with that of a typical 'new economy' portfolio (of Germany's NEMAX (Neuer Markt), the USA's NASDAQ etc.), the extra riskiness of the second is reflected in the fatter tail of g (as well as in an increased volatility).

3. Dependence

The present section discusses the overall dependence structure of the semi-parametric model. In particular, we outline a new result about dependencies of extreme events for elliptically-contoured distributions, which will be utilized in our empirical analysis (see §6).

3.1. Copulae

It is usually more convenient to separate out $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ from the dependence structure, and this is most conveniently done by using the theory of *copulae*, for textbook accounts of which see e.g. Joe (1997), Nelsen (1999). One works to within type, and (by use of the probability integral transformation) transforms all marginal distributions to uniform $(0, 1)$. The joint distribution on $[0, 1]^r$ now encodes all information on dependence structure unencumbered by information on marginals (in particular, by $\boldsymbol{\mu}, \boldsymbol{\Sigma}$), and is called the *copula* (it *couples* the marginal laws to give the joint law).

In particular if $\boldsymbol{X} = (X_1, \ldots, X_r)'$ has joint distribution F with continuous marginals $F_1, \ldots, F_r$, then the distribution function of the transformed vector

$$(F_1(X_1), \ldots, F_r(X_r))$$

is a copula C, and

$$F(x_1, \ldots, x_r) = C(F_1(x_1), \ldots, F_r(x_r)).$$

Hence the t-copula is expressed by:

$$C^t_{\nu,R}(\boldsymbol{u}) = t^r_{\nu,R}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_r)),$$

where $t^r_{\nu,R}$ denotes the distribution function of an r-variate t-distributed random vector with parameter $\nu > 2$ and linear correlation matrix R. Furthermore, t_ν^{-1} is the inverse of the univariate t-distribution function with parameter ν.

For general EC laws, it is immediate that the mean vector μ is no copula parameter. Further, the covariance matrix Σ enters the dependence structure only via the linear correlation matrix R. Although the principal criticisms of elliptically-contoured distributions concerns their *too symmetric* distributional structure, the copula or dependence structure of EC laws turns out to be quite flexible in comparison to many other copulae, like the well-known Archimedean copulae. There are at least $r(r-1)/2$ parameters belonging to the correlation matrix R (even in our semi-parametric model) that govern the overall (linear) dependence structure of an r-dimensional EC law.

Turning to dependencies of extreme events, an important extremal dependence measure is given by tail dependence. This measure arises in the context of VaR quantification for asset portfolios. A bivariate random vector (X_1, X_2) is called *tail dependent* with *tail-dependence coefficient* λ_U if

$$\lambda_U = \lim_{\alpha \to 0} I\!P(X_2 \leq VaR_\alpha(X_2) | X_1 \leq VaR_\alpha(X_1)) > 0, \tag{1}$$

i.e. the limit exists and is greater than zero. Consequently, (X_1, X_2) is called *tail independent* if $\lambda_U = 0$. Similarly we define the *lower-tail dependence coefficient* λ_L. Tail dependence of bivariate random vectors describes the amount of asymptotic dependence in the upper-right quadrant or lower-left quadrant of a bivariate distribution. Thus extreme events (for example large portfolio losses or gains) are no longer independent if the corresponding bivariate marginal distributions exhibits tail-dependence. The degree of tail dependence is completely determined by the bivariate copula, i.e.:

$$\lambda_U = \lim_{u \to 1} (1 - 2u + C(u,u))/(1-u) \quad \text{and} \quad \lambda_L := \lim_{u \to 0} C(u,u)/u. \tag{2}$$

The bivariate Gaussian copulae with correlation $\rho \in (-1, 1)$ have no lower and upper-tail dependence, while t_ν copulae have. In fact, in case of the t_ν copulae, we have increasing tail dependence with decreasing parameter ν.

3.2. Tail dependence for elliptically-contoured distributions

It is natural to ask whether there exists a general characterization of tail dependence for elliptically-contoured distributions. Schmidt (2002) shows that tail dependence within the latter class of distributions is *almost equivalent* to density generators possessing a power law decay. The characterization requires the following well-known definitions: i) A measurable function $f : I\!R_+ \to I\!R_+$ is called *regularly varying* (at ∞) with index $\alpha \in I\!R$ if for any $t > 0 : \lim_{x \to \infty} f(tx)/f(x) = t^\alpha$. ii) A measurable function $f : I\!R_+ \to I\!R_+$ is called *O-regularly varying* (at ∞) if for any $t \geq 1 : 0 < \liminf_{x \to \infty} f(tx)/f(x) \leq \limsup_{x \to \infty} f(tx)/f(x) < \infty$. Thus, regularly varying functions behave asymptotically like power functions and O-regularly varying functions have growth (or decay) bounded between powers. For more details regarding regular variation and O-regular variation we refer the reader to Bingham (1987), pp. 16, and pp. 61, and Resnick (1987), pp. 12.

An elliptically-contoured random vector $\boldsymbol{X} \in EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, $r \geq 2$, with positive-definite covariance matrix $\boldsymbol{\Sigma}$ has tail-dependent bivariate margins if its density generator g is regularly varying with index $-\alpha/2 - 1$, $\alpha > 0$. On the other hand, all bivariate margins of an elliptically-contoured random vector with positive-definite covariance matrix $\boldsymbol{\Sigma}$ are tail-independent (under a minor additional condition) if its density generator is eventually decreasing and not O-regularly varying. As all well-known density generators are either regularly varying or not even O-regularly varying, the above statements are almost necessary and sufficient.

Suppose $(X_i, X_j)'$ is the bivariate margin for an elliptically-contoured random vector $\boldsymbol{X} = (X_1, \dots, X_r)' \in EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ having a regularly varying density generator with index $-\alpha/2 - 1,\ \alpha > 0$. Then the tail-dependence coefficient $\lambda_{ij} := \lambda_U^{ij} = \lambda_L^{ij}$ is given by the following formula:

$$\lambda_{ij} = \frac{\displaystyle\int_0^{h(\rho_{ij})} \frac{u^\alpha}{\sqrt{1-u^2}} du}{\displaystyle\int_0^1 \frac{u^\alpha}{\sqrt{1-u^2}} du}, \tag{3}$$

where $h(\rho_{ij}) = \sqrt{1 + \frac{(1-\rho_{ij})^2}{1-\rho_{ij}^2}}$ and $\rho_{ij} := \boldsymbol{\Sigma}_{ij} / \sqrt{\boldsymbol{\Sigma}_{ii}\boldsymbol{\Sigma}_{jj}}$ denotes the correlation coefficient, if it exists.

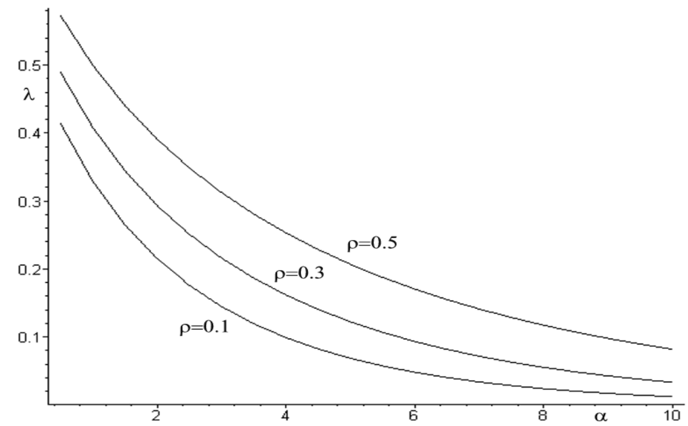

Figure 1. Tail dependence coefficient λ_U versus regular variation index α for $\rho = 0.5,\ 0.3,\ 0.1$ for elliptically-contoured distributions

Figure 1 shows for the bivariate case the interplay of the tail dependence coefficient λ and the regular variation index α for various correlation coefficients ρ. An immediate consequence for the GH law is tail-independence because of the exponential decay of the corresponding density generator. For a multivariate extension of the tail-dependence concept see Schmidt and Stadtmüller (2002).

3.3. Partial symmetry and copulae

The present section is devoted to several extensions of elliptically-contoured distributions and the corresponding dependence structures, respectively. Note that if we work within *type* - reduce to the standardised case $\boldsymbol{\mu} = 0, \boldsymbol{\Sigma} = \boldsymbol{I}$ - we have no reason to distinguish one component from another, and obtain *exchangeability* (for background on exchangeability, see e.g. Aldous (1985), and in the context of multivariate dependence, Joe (1997)). In the context of portfolio theory in finance, this is too restrictive: what is needed instead is *partial exchangeability*. Thus, our vector $\boldsymbol{X} = (X_1, \ldots, X_r)$ of r stock prices (or their returns) will represent the evolution of r individual assets, but these will typically be chosen so as to balance our holdings between the various economic sectors in which we choose to invest (petrochemicals, pharmaceuticals, motor manufacturing, etc.). We need a model which gives exchangeability *within* sectors (to a first approximation - as the relevant companies operate within a similar economic environment), but not *between* sectors (since we wish to balance our holdings between sectors by Markowitzian diversification).

Models for partial exchangeability have been studied in considerable generality. Structure theorems have been obtained (Aldous (1985), §14), which may be loosely summarised as giving non-linear analogues of the familiar Linear Model of statistics. The simplest non-trivial example (row-column exchangeability for infinite arrays) has a representation of the form

$$X_{ij} = f(\alpha, \xi_i, \eta_j, \lambda_{ij})$$

where f is a deterministic (non-linear!) function, α represents an overall effect, ξ_i, η_j row and column effects and λ_{ij} an individual random effect. Here the $\xi_i, \eta_j, \lambda_{ij}$ are independent, and can be taken as uniform on $(0,1)$. For finite arrays, as here, this set-up and its relatives provide ways (sufficient but not necessary) to model the types of partial exchangeability relevant to portfolio theory. Here of course the function f contains all the structural information - $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, as well as the dependence structure.

Families of copulae displaying the partial exchangeability structure one requires may be readily constructed. Perhaps the simplest such construction generalises the idea of an *Archimedean copula* (Joe (1997), §4.2). To keep the notation simple, consider a trivariate case (to model a portfolio of three assets). A trivariate copula of the form

$$C(\mathbf{u}) := \psi(\psi^{-1} \circ \phi(\phi^{-1}(u_1) + \phi^{-1}(u_2)) + \psi^{-1}(u_3))$$

(Joe (1997), §4.7 has, for suitably restricted ϕ and ψ, $(1,2)$ bivariate marginals of the form

$$C(u_1, u_2) = \phi(\phi^{-1}(u_1) + \phi^{-1}(u_2))$$

and $(1,3)$ and $(2,3)$ bivariate marginals of the same form but with ψ for ϕ. A quadrivariate version, with

$$C(\mathbf{u}) := \psi(\psi^{-1} \circ \phi(\phi^{-1}(u_1) + \phi^{-1}(u_2)) + \psi^{-1} \circ \zeta(\zeta^{-1}(u_3) + \zeta^{-1}(u_4)))$$

(with ϕ, ψ, ζ suitably restricted) has $1,2$ exchangeable and $3,4$ exchangeable, but $1,2$ not exchangeable with $3,4$. Higher-dimensional analogues may be constructed similarly.

The gains and losses here are obvious. The gain is that one can readily model whatever partial summetry structure reflects the construction of one's portfolio. The loss is that the complexity of the model ramifies rapidly with dimension - as inevitably it must. In particular, one is immediately exposed to the *curse of dimensionality*, to quote the late Richard Bellman's famous and telling phrase. The class $\mathbb{C}_r$ of all copulae on the unit r-cube is very rich for r at all large - too rich for convenience. The obvious way to escape the curse of dimensionality is to seek parametric families of copulae, adapted to modelling partial symmetry (Joe (1997), §5.3) and negative dependence (Joe (1997), §5.4). A considerable body of theory exists here, and there is much scope for empirical work. However, even in the fully symmetric case of elliptically-contoured distributions, we know that parametric treatments are insufficiently flexible to accommodate the full variety encountered in practice, and this is all the more true in the more complicated setting of partial symmetry. What is worse, the computational complications arising from the curse of dimensionality are unavoidable even in principle. There is clearly much scope for, and need of, both non-parametric modelling of copulae in this setting of financial portfolios, and empirical work in fitting models to data.

4. Estimation

Fitting data to the semi-parametric model requires estimation of the parametric location and scatter parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and the non-parametric shape component g. We present robust and appropriate estimators fulfilling the prerequisite of decreasing density generators g.

4.1. Estimating the mean and the covariance

Natural estimators for the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ of a multivariate elliptically-contoured distribution are the sample mean and the Pearson product-moment covariance. Although the last estimator yields almost surely a positive-definite covariance matrix, which avoids unpleasant matrix transformations, both estimators are vulnerable to heavy-tailed and contaminated data: a single grossly aberrant reading can destroy the accuracy of both. Several robust multivariate estimation techniques for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ have been proposed in the literature. Among them we mention:

(i) Multivariate trimming (Hahn, Mason, and Weiner 1991),

(ii) Minimum-volume ellipsoid estimator for $\boldsymbol{\Sigma}$ (Rousseeuw and van Zomeren 1990),

(iii) Median estimator for $\boldsymbol{\mu}$ and LS-estimator for $\boldsymbol{\Sigma}$ (described below).

The third set of estimators has much to recommend it. First, they provide robust estimators utilizing exactly the elliptical structure of EC laws. Second, they are applicable even if the covariance matrix or the mean vector do not exist. In the following we will give a short description of the LS-estimator, which applies a well-known stochastic representation of elliptical distributions (cf. Fang, Kotz, and Ng (1990) Theorem 2.14). Every r-dimensional elliptically distributed random vector $\boldsymbol{X} \in EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (positive-definite matrix) can be stochastically represented by

$$\boldsymbol{X} \overset{d}{=} \boldsymbol{\mu} + R_r \boldsymbol{A}' \boldsymbol{U}^{(r)} \tag{4}$$

with Cholesky decomposition $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{\Sigma}$. The random variable $R_r \geq 0$ is independent of the r-dimensional random vector $\boldsymbol{U}^{(r)}$ which is uniformly distributed on the unit sphere in $\mathbb{R}^r$. We refer to R_r as the *random variate* of an elliptical distribution. According to equation (4) we obtain

$$R_r \boldsymbol{U}^{(r)} \stackrel{d}{=} (\boldsymbol{A}')^{-1}(\boldsymbol{X}-\boldsymbol{\mu}) \quad \text{and} \quad \boldsymbol{U}^{(r)} \stackrel{d}{=} \frac{(\boldsymbol{A}')^{-1}(\boldsymbol{X}-\boldsymbol{\mu})}{||(\boldsymbol{A}')^{-1}(\boldsymbol{X}-\boldsymbol{\mu})||_2}.$$

For a bivariate EC law the covariance or scaling matrix $\boldsymbol{\Sigma}$ is now estimated in a two-step procedure. Let $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n$ be a random sample, distributed according to an $EC_2(\boldsymbol{\mu},\boldsymbol{\Sigma};g)$ law. First the location or mean vector is estimated by the median estimator $\hat{\boldsymbol{\mu}} := \hat{\boldsymbol{\mu}}_n = \bar{x}_{0.5}$. Second an LS-estimator is defined by

$$\hat{\boldsymbol{A}}_*' := \operatorname{argmin}_{\boldsymbol{C}'\in\Gamma} \sum_{i=1}^{n} \Big(D_i(\boldsymbol{C}') - \frac{2\pi}{n}\Big)^2 \text{ with } D_i(\boldsymbol{C}') = \min_{j:j\neq i} (\arccos < \boldsymbol{U}_i^{(2)}, \boldsymbol{U}_j^{(2)} >),$$

where $\boldsymbol{U}_i^{(2)} \stackrel{d}{=} (\boldsymbol{C}')^{-1}(\boldsymbol{X}_i-\hat{\boldsymbol{\mu}})/||(\boldsymbol{C}')^{-1}(\boldsymbol{X}_i-\hat{\boldsymbol{\mu}})||_2$, Γ denotes the set of all lower triangle matrices with determinate one and $< \cdot,\cdot >$ is the scalar-product. $2\pi/n$ matches the minimal angle between n perfectly uniform distributed vectors on the 2-dimensional unit-sphere. Note that the corresponding estimate $\hat{\boldsymbol{\Sigma}}_* = \hat{\boldsymbol{A}}_*'\hat{\boldsymbol{A}}_*$ is not necessarily a covariance matrix but some scaling matrix. In order to obtain an estimate $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{A}}'\hat{\boldsymbol{A}}$ of the covariance matrix $\boldsymbol{\Sigma}$, if it exists, we must rescale $\hat{\boldsymbol{\Sigma}}_*$ by $\hat{\mu}_{R^2}/2$ with $\hat{\mu}_{R^2}$ being some estimator of $E(R_*^2) := E(||(\hat{\boldsymbol{A}}_*')^{-1}(\boldsymbol{X}-\hat{\boldsymbol{\mu}})||_2)$. The LS estimator has the advantage of being robust, having a low MSE (see (Frahm, Junker, and Schmidt 2002) for statistical investigations), and providing an almost-surely positive-definite estimate of the covariance-matrix. The estimation process in higher dimensions is described in Frahm and Junker (2003).

4.2. Estimating the density generator g

For the non-parametric estimation of the density generator we utilize the transformed data

$$\boldsymbol{p}_i = (\boldsymbol{x}_i-\hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{x}_i-\hat{\boldsymbol{\mu}}), \;\; i=1,\ldots,n,$$

where $\boldsymbol{x}_i$ is a realisation of $\boldsymbol{X}_i$, $i=1,\ldots,n$. The data points $\boldsymbol{p}_i$ are approximate realisations of the distribution induced by the squared random variate R_r^2. In particular the latter data come from a density $f_{R_r^2}$ which relates to the unknown density generator g as follows (cf. Fang, Kotz, and Ng (1990), Remark to Theorem 2.9)

$$f_{R_r^2}(u) = \frac{\pi^{r/2}}{\Gamma(r/2)} u^{r/2-1} g(u). \tag{DQ}$$

For the non-parametric estimation of the density $f_{R_r^2}$ given in (DQ), we use standard results on optimal bandwidth selection, see e.g. Härdle (1990b), Chapter 4, or Wand and Jones (1995), Chapter 3. In particular, we use the variants of direct plug-in rules (Wand and Jones (1995), Section 3.6.).

Let h be some arbitrary density function which is unknown. According to the above references, the asymptotic integrated mean-squared error (A-MISE) optimal bandwidth for kernel density estimation is given by

$$b_0 = \left(\frac{\|K\|_2^2}{\|h''\|_2^2 \, (\mu_2(K))^2 n} \right)^{\frac{1}{5}}$$

(K is a kernel, $\|.\|_2$ is the standard L_2-norm, $\mu_2(k) = \int x^2 K(x)dx$). The problematic feature here is that the second derivative of the unknown density h is involved. A first idea is to replace h'' by the corresponding value of a reference distribution. In our applications we use the Gaussian kernel and as reference distribution the normal distribution. This leads to (see Härdle (1990b) (4.1.1)) $\hat{b}_0 = (4\hat{\sigma}^5/(3n))^{1/5}$ with $\hat{\sigma}$ an estimator for the standard deviation of the unknown distribution. Since in the context of financial data we are dealing with heavy-tailed distributions, we want to make the choice of bandwidth insensitive to outliers. Thus instead of $\hat{\sigma}$ we use the interquartile range $\hat{R} = X_{[0.75n]} - X_{[0.25n]}$ as a more robust estimate for the scale parameter. Thus we choose the bandwidth according to

$$\hat{b}_0 = 1.06 \min\{\hat{\sigma}, \hat{R}/1.34\} n^{-\frac{1}{5}}$$

(note that for the Gaussian data $\hat{R} = \hat{\sigma} \cdot 1.34$).

A sophistication of the above idea is to use multi-stage plug-in estimators: Instead of using a reference distribution immediately, one uses a kernel estimate of the density functional $\|h''\|_2^2$, which will again be sensitive to a bandwidth choice. This bandwidth then can be chosen via a reference distribution, or another estimate of a relevant density functional can be employed. Motivated by our own simulation study and detailed comparison studies of data-driven bandwidth selection principles, e.g. Park and Marron (1990) and Wand and Jones (1995), §3.8. for data from heavy-tailed distributions, we employed a two-stage plug-in estimator suggested by Sheather and Jones (1991) and outlined in Wand and Jones (1995), §3.6 for our estimation problem.

A similar problem arises in estimating confidence intervals. Asymptotic confidence intervals are given by

$$\left[\hat{h}_b(x) - n^{-\frac{2}{5}} \left(\frac{c^2}{2} h''(x)\mu_2(K) + d_\alpha \right), \hat{h}_b(x) - n^{-\frac{2}{5}} \left(\frac{c^2}{2} h''(x)\mu_2(K) - d_\alpha \right) \right],$$

where $d_\alpha = u_{1-\alpha/2}\sqrt{c^{-1}h(x)\,\|K\|_2^2}$, $u_{1\alpha/2}$ is the $1-\alpha/2$-quantile of the standard Normal distribution, and c is obtained from the relation $b_0 = cn^{-\frac{1}{5}}$.

Again this formula involves the unknown density and its second derivative. We use estimates $\hat{h}_{b_1}$ and $\hat{h}''_{b_1}$ with a slightly larger bandwidth b_1.

Since our data come from a distribution with non-negative support, we face the problem of density estimation near boundary points. Several possible modifications to improve performance near boundaries are discussed in Härdle (1990a), §4.4, Wand and Jones (1995),§2.11 and in the articles Müller (1991) and Jones (1993).

In the context of elliptically-contoured distribution the problem is discussed in Stute and Werner (1991), who study the estimation of the d-dimensional density f, and Hodgson, Linton, and Vorkink (2002). Like these two studies our simulation results confirm that a modification suggested by Schuster (1985) works well. Applying Schuster's modification to our setting we incorporate the additional information of positive support of the underlying density by adding a mirror image term to the standard kernel estimator. That is, we use as an estimator

$$\hat{h}_n(u) = \frac{1}{n\hat{b}_0} \sum_{i=1}^{n} \left[K\left(\frac{\boldsymbol{x} - \boldsymbol{x}_i}{\hat{b}_0}\right) + K\left(\frac{\boldsymbol{x} + \boldsymbol{x}_i}{\hat{b}_0}\right)\right].$$

According to Section 2.4, one of the main structural features we require, is the monotonicity of the density generator g. For that we transform the estimated density generator via a monotone regression method suggested in Härdle (1990a), Section 8.1. However, this regression method assumes an unimodal density function which has to be verified (compare also Anderson, Fang, and Hsu (1986) or Fang and Zhang (1990), Section 4.1). The reason for unimodality is because we mirror the observations on a possible mode point and then use the monotone regression algorithm of the partially mirrored data.

Remark. In §6 we compare the non-parametric estimates of the density generator g (in our semi-parametric model) with various parametric estimates. In the latter context, we estimate the parameters of the corresponding density $f_{R_r^2}$ (see equation (DQ)) via maximum-likelihood techniques, utilizing the approximate realizations $\boldsymbol{q}_i = ||\hat{\boldsymbol{\Sigma}}^{-1/2}(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}})||_2,\ i = 1, \ldots, n$.

5. Simulation

Here we address the generation of multivariate random vectors from our semi-parametric model. Efficient and fast multivariate random-number generators are indispensable for modern portfolio investigations. Monte-Carlo simulations for Value at Risk (VaR) and pricing calculations require many random vectors, which have to be sampled in a reasonable time-frame. For elliptical distributions the multidimensional simulation problem boils down to a one-dimensional simulation task.

The simulation from the semi-parametric law $EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \hat{g})$, with estimated density generator $\hat{g}$ via the non-parametric methods introduced in §4, is performed by utilizing formula (DQ). Observe that the corresponding estimate of the density function $\hat{f}_{R_r^2}$ is discretely evaluated at points $u_1, \ldots, u_m$ and the entire function estimate $\hat{f}_{R_r^2}$ is obtained via linear interpolation. This construction of the density function $f_{R_r^2}$ leads to a straightforward random-number generator for the random variate R_r based on the well-known Inversion Transformation Method (see for example Niederreiter (1993)).

Pseudo algorithm for generating multivariate random numbers from the semi-parametric $EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ law:

(i) Set $\boldsymbol{\Sigma} = \boldsymbol{A}'\boldsymbol{A}$, via Cholesky decomposition.

(ii) Sample a random number from R_r (as described above).

(iii) Sample r independent random numbers $Z_1, \ldots, Z_r$ from a univariate standard-normal distribution $N_1(0, 1)$.

(iv) Set $\boldsymbol{Z} = (Z_1, \ldots, Z_r)$.

(v) Set $\boldsymbol{U}^{(r)} = ||\boldsymbol{Z}||^{-1} \cdot \boldsymbol{Z}$.

(vi) Return $\boldsymbol{X} = \boldsymbol{\mu} + R_r \cdot \boldsymbol{A}'\boldsymbol{U}^{(r)}$.

6. Application to financial data

Our financial time series consists of 6 series of major stock indices covering the (sub−) periods between January 1987 to December 2002 (data obtained from Bloomberg Financial Services). We investigated the density generators of various portfolios of indices. In particular we use

- Mature markets (from 1987 to 2002): DOW JONES, DAX, CAC40, NIKKEI;
- New economy markets (from 1998 to 2002): NASDAQ, NEMAX.

In addition, we used 6 major US stocks, namely Ford, Boeing, General Motors, Dell Computers, Cisco Systems, Microsoft during the period 1990 to 2002 to investigate portfolios of stocks.

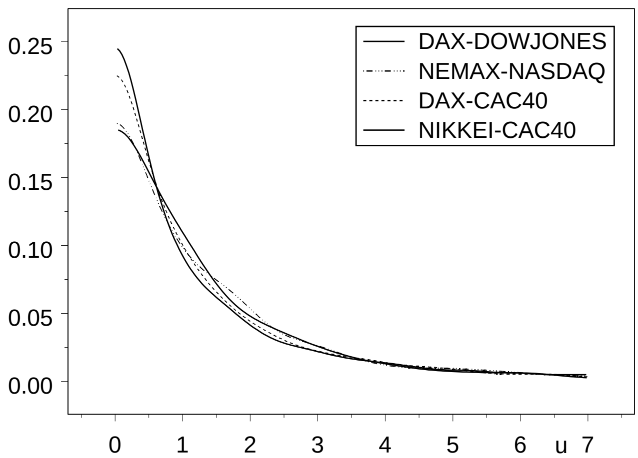

Figure 2. Density generators of bivariate log-returns for four asset series

6.1. Fitting the data

We start with several visual tests for elliptical symmetry of the underlying distribution. Plots of the four time-series (Figure 3), contour plots of the estimate of the density (Figure 4) (obtained by using the function kde2d from Venables and Ripley (1999) in S-Plus), and the estimate of the density generator g (Figure 2, 6, 7 and 8) are provided.

For our two-dimensional examples we used the DAX-DOW JONES, DAX-CAC40, NIKKEI-CAC40, and NEMAX-NASDAQ series; however, the results are representative for all combinations we tried.

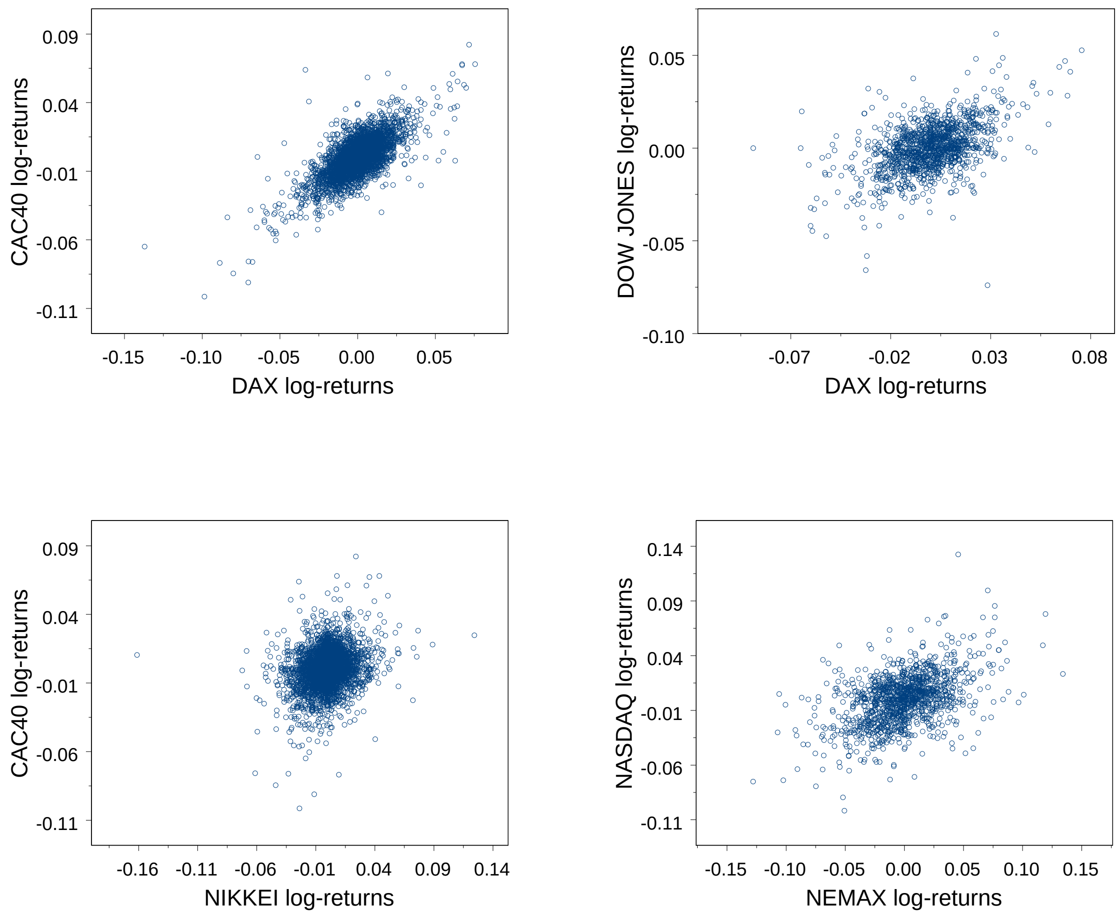

Figure 3. Scatter plots of bivariate log-returns of four asset series.

A more formal investigation is the use of QQ-plots as discussed in the next section. Figure 5 displays such plots for the above index combinations. All plots indicate a satisfactory fit of our underlying model.

Finally we proceeded to estimate density generators using our favorite two-step selection principle. Plots for two- and high-dimensional portfolios are displayed in Figures 6, 7 and 8

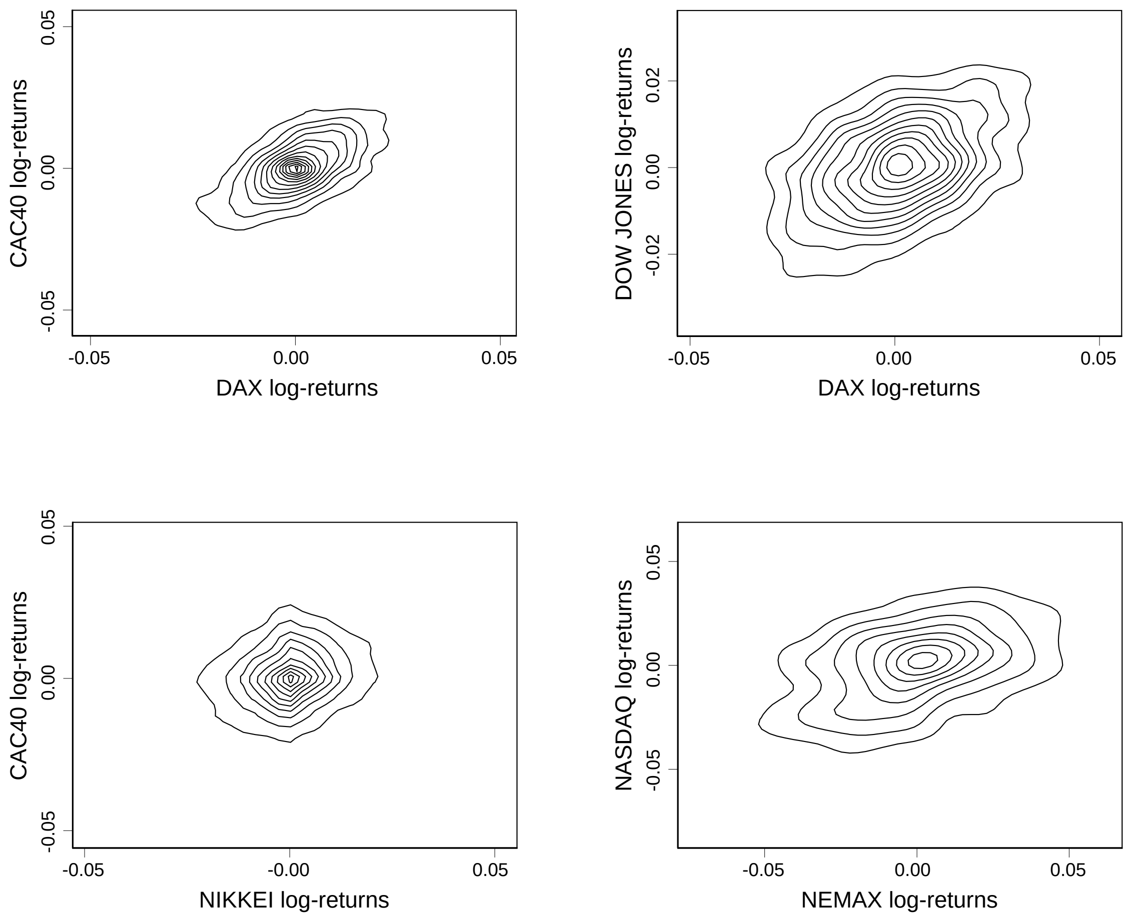

Figure 4. Contour-plots of bivariate log-returns of four asset series.

6.2. Tests for elliptical symmetry

To test for spherically symmetric and elliptically symmetric distributions we use a simple graphical method suggested by Li, Fang, and Zhu (1997). Alternatively, one could also use one of the tests suggested in e.g. Baringhaus (1991), Beran (1979) or Manzotti, Pérez, and Quiroz (2002).

This method is based on use of a statistic $t(\boldsymbol{x})$ which is invariant under orthogonal transformations. From Fang, Kotz, and Ng (1990), Theorem 2.22, we know that the distribution of such a statistic remains invariant for $\boldsymbol{X} \sim S_r^+(\phi)$, i.e. r-dimensional spherical distributions with no probability mass at the origin.

For our purpose we use the standard t-statistic: For a one-dimensional random sample $(X_1, X_2, \ldots, X_r)$ define $\bar{X} := \frac{1}{r}\sum_{i=1}^{r} X_i$ and $s^2 := \frac{1}{r-1}\sum_{i=1}^{r}(X_i - \bar{X})^2$; then

$$t(X_1, X_2, ..X_r) = \frac{\sqrt{r}\bar{X}}{s} \sim t_{r-1}$$

if $(X_1, X_2, ..X_r) \sim N(\mathbf{0}, \boldsymbol{I}_r)$. By the above invariance principle one can replace the multivariate normal by any appropriate spherical distribution, and indeed, after

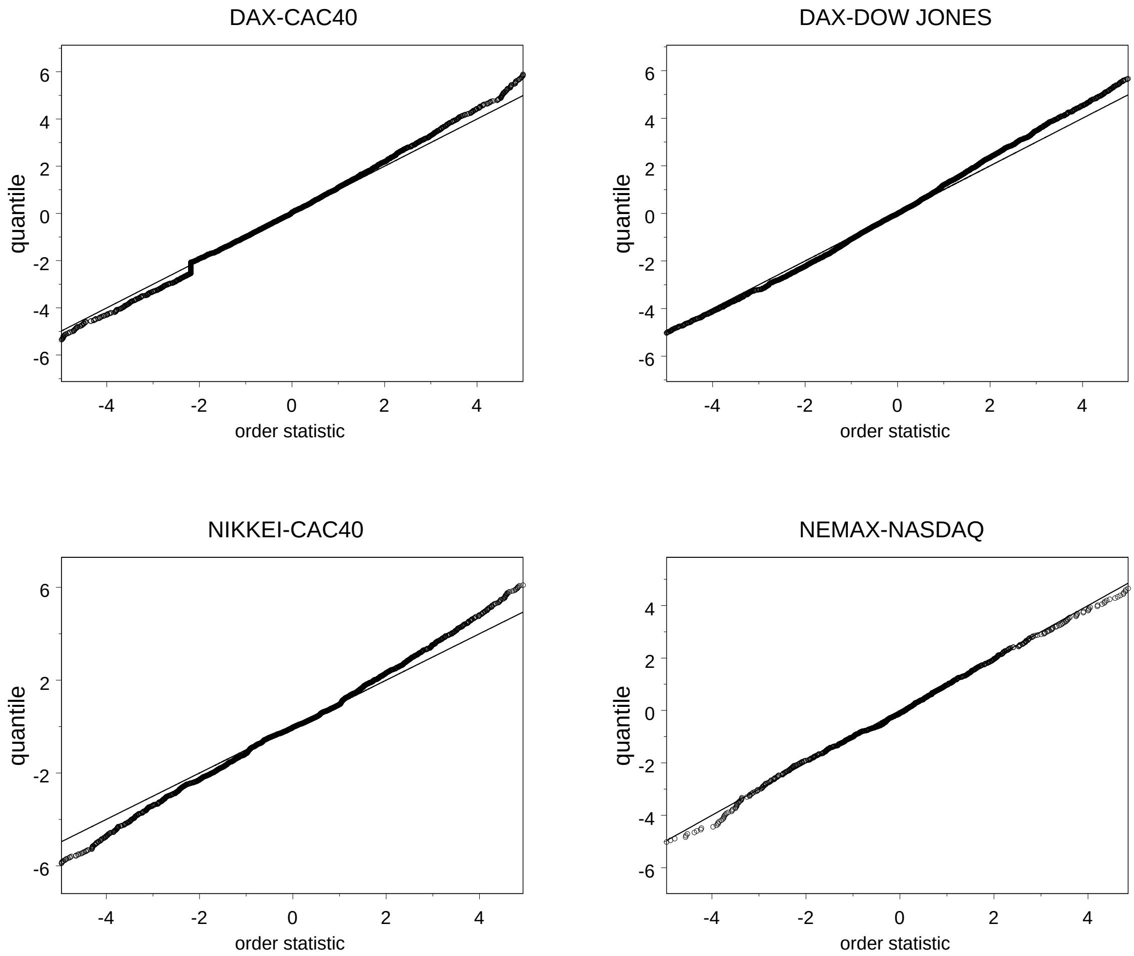

Figure 5. QQ-plots of bivariate log-returns of four asset series

normalisation, elliptical distribution. Therefore a QQ-plot (Figure 5) of a sample against the appropriate t-distribution should indicate possible deviations from the class of spherical (ellipical) distributions.

Further possible graphical estimators are outlined in Li, Fang, and Zhu (1997), with underlying theory given in Fang, Kotz, and Ng (1990), §2.7.

6.3. Multidimensional portfolios

In Figure 8 several multidimensional portfolios are considered in order to stress the applicability of the semi-parametric model for higher dimensional asset-return modelling. First we construct two four-dimensional asset-return vectors consisting of the indices DAX - DOWJONES - NIKKEI - CAC40 and DAX - NEMAX - DOWJONES - NASDAQ, the first being a pure 'old' economy portfolio and the second a mixture of 'new' and 'old' economy indices. We provide the QQ-plots to justify applying the semi-parametric model and the corresponding density generator. Again we observe that the density-generator of the t-distribution fits the non-parametric density generator much

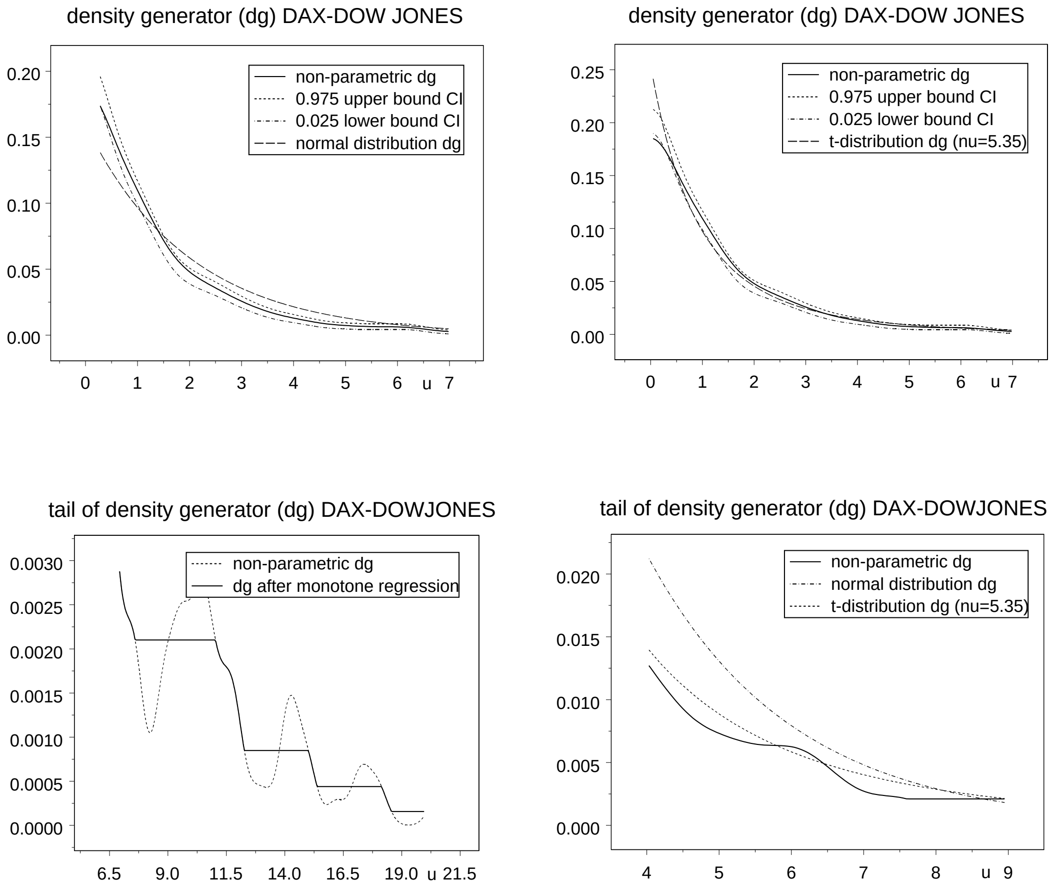

Figure 6. Density generators of bivariate log-returns for DAX-DOWJONES from 1998 until 2002

better than the normal distribution. However, the density generator of the t-distribution decreases as a power-law in the entire range of the data, which is actually not observed in the data. Finally we construct a six-dimensional portfolio consisting of the stocks: FORD - BOEING - GM - DELL - CISCO - MICROSOFT. Here we emphasize that the semi-parametric model we propose has much to recommend it for high-dimensional portfolio modelling. For the data we investigated, the curse of dimensionality is avoided in a satisfactory way through the non-parametric density generator.

6.4. Portfolio analysis

The Value at Risk (VaR) of an asset portfolio with r assets and portfolio-value functions $p_t : I\!R^r \to I\!R,\ t \in \mathbb{N}$ at some discrete time point t is defined by

$$VaR_t^\alpha = \mathrm{argsup}_{y \in I\!R} I\!P(\Delta p_t \le y) \le \alpha, \tag{5}$$

for confidence level $\alpha > 0$. Here the random vectors $\boldsymbol{X}_t,\ t \in \mathbb{N}$ contain the nominal asset values at time t and the absolute portfolio return is defined by $\Delta p_t := p_t - p_{t-1}$. We assume that $\Delta p_t(\boldsymbol{X}_t, \boldsymbol{X}_{t-1})$ depends only on the componentwise relative asset-returns

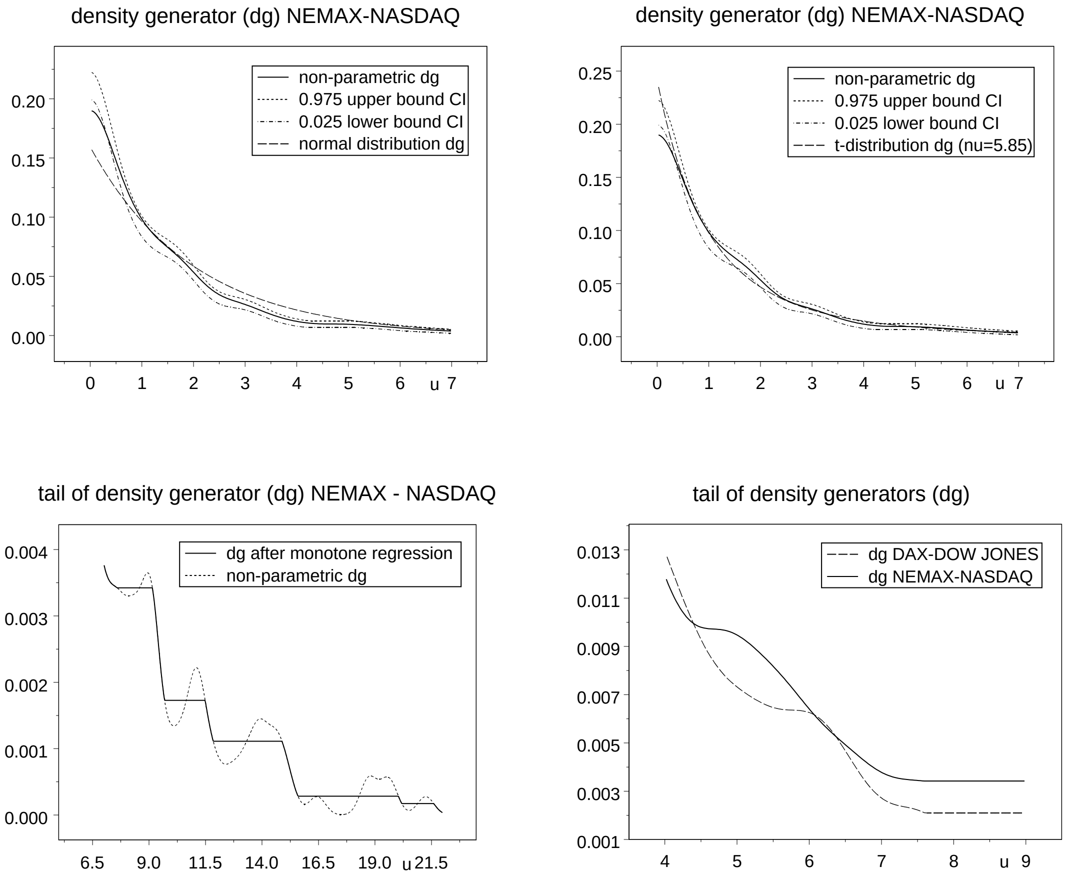

Figure 7. Density generators of bivariate log-returns for NEMAX-NASDAQ from 1998 until 2002

at time t and the nominal asset values $\boldsymbol{X}_{t-1}$. In particular we define the random variable V_t^p describing the portfolio's absolute return at time t by

$$V_t^p := \Delta p_t(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}) = f_t(\Delta_{rel}\boldsymbol{X}_t, \boldsymbol{X}_{t-1})$$

for some functions $f_t : I\!R^r \times I\!R^r \to I\!R$ and $\Delta_{rel}\boldsymbol{X}_t := (\log(\boldsymbol{X}_t^i) - \log(\boldsymbol{X}_{t-1}^i))_{i=1,\ldots,r}$. For modelling reasons we suppose the log-returns $\Delta_{rel}\boldsymbol{X}_t,\ \ t = 1,\ldots,n$, to be iid observations. The main intention of this section is now formulated by the assumption that $\Delta_{rel}\boldsymbol{X}_t$ follows an EC law:

$$\Delta_{rel}\boldsymbol{X}_t \sim EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g). \tag{6}$$

For reasons of simplicity we do not utilize the innovations of an autoregressive-type process or an ARCH-type process, as in our context the results turned out to be quite similar.

In general, Value at Risk determination within modern portfolio theory utilizes one of four approaches: i) explicit VaR calculation, ii) structured Monte-Carlo simulation, iii) variance-covariance approximation, or iv) historical simulation. One of the main

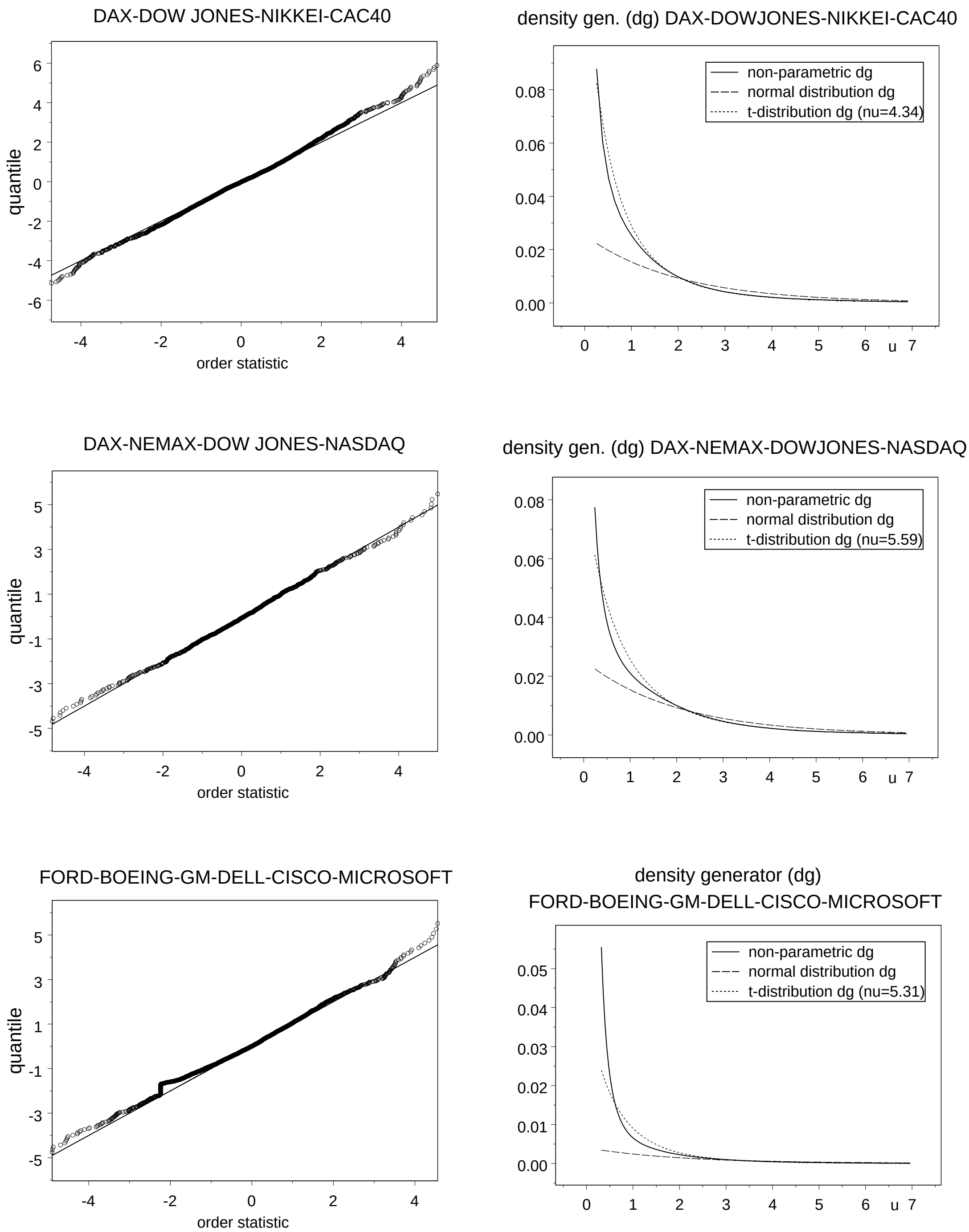

Figure 8. Density generators for multidimensional log-returns of DAX-DOWJONES-NIKKEI-CAC40 from 1987 until 2002 (first level), of DAX-NEMAX-DOWJONES-NASDAQ from 1998 until 2002 (second level) and of FORD-BOEING-GM-DELL-CISCO-MICROSOFT from 1990 until 2002 (third level).

advantages of assumption (6) is that the portfolio's VaR can be explicitly calculated for every linear asset portfolio. Precisely, for a linear asset portfolio we obtain

$$V_t^p = \boldsymbol{X}_{t-1}' \Delta_{rel} \boldsymbol{X}_t.$$

The portfolio's VaR for confidence level α, under a given realisation $\boldsymbol{x}_{t-1}$ of the asset-value vector and elliptically-contoured asset-returns $\Delta_{rel}\boldsymbol{X}_t \in EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ with positive-definite $\boldsymbol{\Sigma}$, is given by

$$VaR_t^\alpha = \boldsymbol{x}_{t-1}'\boldsymbol{\mu} - h(\alpha)\sqrt{\boldsymbol{x}_{t-1}'\boldsymbol{\Sigma}\boldsymbol{x}_{t-1}}, \tag{7}$$

with $h(\alpha)$ being defined below. Formula (7) can be readily shown via the stochastic representation (4) of elliptically-contoured asset-return vectors, i.e.,

$$\Delta_{rel}\boldsymbol{X}_t \stackrel{d}{=} \boldsymbol{\mu} + R_r \boldsymbol{A}'\boldsymbol{U}^{(r)}.$$

Then

$$V_t^p = \boldsymbol{x}_{t-1}'\Delta_{rel}\boldsymbol{X}_t \stackrel{d}{=} \boldsymbol{x}_{t-1}'\boldsymbol{\mu} + R_r(\boldsymbol{A}\boldsymbol{x}_{t-1})'\boldsymbol{U}^{(r)}$$

is distributed according to an $EC_1(\boldsymbol{x}_{t-1}'\boldsymbol{\mu}, \boldsymbol{x}_{t-1}'\boldsymbol{\Sigma}\boldsymbol{x}_{t-1}; g)$ law. Define now $VaR_t^\alpha := \boldsymbol{x}_{t-1}'\boldsymbol{\mu} - h(\alpha)\sqrt{\boldsymbol{x}_{t-1}'\boldsymbol{\Sigma}\boldsymbol{x}_{t-1}}$ with $h(\alpha)$ the $1-2\alpha$ quantile of the positive random variable R_rB, where B^2 is $Beta(1/2, (r-1)/2)$ distributed and independent of R_r. Observe that

$$\begin{aligned} &I\!P(\boldsymbol{x}_{t-1}'\Delta_{rel}\boldsymbol{X}_t \leq VaR_t^\alpha) = I\!P(\boldsymbol{x}_{t-1}'\boldsymbol{\mu} + R_r(\boldsymbol{A}\boldsymbol{x}_{t-1})'\boldsymbol{U}^{(r)} \leq VaR_t^\alpha) \\ &= I\!P(R_rBU^{(1)} \leq -h(\alpha)) = \frac{1}{2}I\!P(R_rB \geq h(\alpha)) = \alpha, \end{aligned}$$

where the second-to-last equation follows by Theorem 2.15 in Fang, Kotz, and Ng (1990). Now formula (7) follows.

Tables 1 and 2 show various VaR figures for some asset portfolio we considered previously. First we present the empirical VaRs for linear portfolios with equally weighted assets for confidence levels $\alpha = 0.01,\ 0.025,\ 0.05$. Further we calculate the VaRs analytically via different density generators, namely via the non-parametric density generator, and the density generators of the fitted t-distribution, the normal distribution and generalized hyperbolic distribution (GH). Finally a simulation study shows the finite-sample properties of the corresponding VaR estimations. We conclude that even in the case of VaR considerations, where we are particularly interested in the tail behavior of the density generator, the parametric approach does not outperform the non-parametric estimation. The maximum-likelihood estimates utilized for the parametric density generators yield an overall fit including all available data and usually underestimates the tail, whereas VaR primarily examines the tail behavior and tail dependence of a multivariate portfolio distribution.

Data analysis. The analytical VaR figures in Table 1 reveal that the normal distribution underestimates the 0.01 and 0.025 VaR's considerably whereas the 0.05 VaR is estimated reasonably. By contrast, the semi-parametric EC-law and the t-distribution perform better for the higher 0.01 and 0.025 quantiles than for the 0.05 VaR. The GH-law underestimates the VaR for nearly every asset combination. For the simulated VaR figures in Table 2 we obtain a similar picture. As a result we suggest the semi-parametric model as a suitable substitute for the above parametric EC-laws in the context of portfolio VaR calculations.

Comparison of 'old' and 'new' economy. We would expect (as we know from the collapse of the dotcom bubble) that 'high-tech', 'new' economy stocks (indices NEMAX-NASDAQ) are riskier that traditional 'old' economy stocks (indices DAX-DOW JONES). We would thus expect fatter tails - slower decay of the density generator g - in the 'new' case then in the 'old'. Indeed, in Figure 7 we clearly see this, when the tails of the density generator are graphed: 'new' lies above 'old' in the tails. Note however that the comparison is reversed in the bulk of data, where in particular the degrees of freedom (df) of the approximating t-distributions are determined. Thus 'new' has df 5.85 while 'old' has 5.35. Within the t-family, higher df means thinner tails; the crossing of the graphs of the density generator explains this reversal.

Note also the much bigger 1% VaR, 61.4 to 36.8, of 'new' against 'old'. Referring to (7), note that 'new' and 'old' have similar $h(\alpha)$; the difference is accounted for by the different covariance matrices $\mathbf{\Sigma}$.

Table 1. Empirical and analytical VaR's to the confidence levels $\alpha = 0.01,\ 0.025,\ 0.05$ for various linear portfolios with equally weighted assets.

Assets	empirical VAR			semi-parametric VAR			normal VAR		
	$\alpha = 0.01$	0.025	0.05	$\alpha = 0.01$	0.025	0.05	$\alpha = 0.01$	0.025	0.05
DAX-DOWJONES	-36.76	-28.37	-22.19	-35.63	-26.75	-20.82	-31.41	-26.48	-22.24
NEMAX-NASDAQ	-61.41	-46.81	-38.28	-61.05	-47.27	-36.71	-53.85	-45.43	-38.20
DAX-CAC40	-38.47	-27.61	-19.37	-32.91	-24.96	-19.57	-29.66	-24.96	-20.92
NIKKEI-CAC40	-29.95	-22.86	-18.50	-29.14	-22.30	-17.23	-25.68	-21.64	-18.17
DAX-DOWJONES-NIKKEI-CAC40	-28.13	-19.83	-15.13	-23.53	-17.19	-12.19	-22.25	-18.73	-15.70
DAX-NEMAX-DOWJONES-NASDAQ	-44.05	-35.78	-28.31	-44.25	-32.32	-23.45	-39.93	-33.68	-28.31
FORD-BOEING-GM-DELL-CISCO-MICROSOFT	-42.65	-34.15	-27.69	-44.97	-32.34	-20.39	-38.93	-32.68	-27.30

Assets	empirical VAR			t VAR			GH VAR		
	$\alpha = 0.01$	0.025	0.05	$\alpha = 0.01$	0.025	0.05	$\alpha = 0.01$	0.025	0.05
DAX-DOWJONES	-36.76	-28.37	-22.19	-34.95	-26.93	-21.24	-32.87	-25.74	-20.14
NEMAX-NASDAQ	-61.41	-46.81	-38.28	-59.45	-46.28	-36.66	-58.33	-45.52	-35.79
DAX-CAC40	-38.47	-27.61	-19.37	-33.65	-25.18	-19.47	-31.47	-24.30	-18.79
NIKKEI-CAC40	-29.95	-22.86	-18.50	-29.19	-21.71	-16.68	-28.51	-21.92	-16.91
DAX-DOWJONES-NIKKEI-CAC40	-28.13	-19.83	-15.13	-25.23	-18.89	-14.62	-22.58	-17.52	-13.64
DAX-NEMAX-DOWJONES-NASDAQ	-44.05	-35.78	-28.31	-44.25	-34.31	-27.17	-42.10	-32.80	-25.37
FORD-BOEING-GM-DELL-CISCO-MICROSOFT	-42.65	-34.15	-27.69	-43.43	-33.26	-26.04	-47.64	-36.50	-28.23

Table 2. Mean (standard deviation) of 100 simulated VaR-figures based on 2000 equally weighted asset-return data.

Assets	semi-parametric VAR $\alpha = 0.01$	0.025	0.05	normal VAR $\alpha = 0.01$	0.025	0.05
DAX-DOWJONES	-35.93 (2.21)	-26.76 (1.17)	-21.05 (0.73)	-31.49 (1.16)	-26.44 (0.84)	-22.21 (0.63)
NEMAX-NASDAQ	-60.64 (3.41)	-47.32 (2.46)	-36.90 (1.47)	-54.09 (1.93)	-45.66 (1.40)	-38.27 (1.21)
DAX-CAC40	-33.85 (1.86)	-26.03 (1.33)	-20.28 (0.71)	-29.76 (1.14)	-24.96 (0.71)	-21.00 (0.57)
NIKKEI-CAC40	-29.24 (1.90)	-22.26 (1.08)	-17.18 (0.77)	-25.83 (0.90)	-21.64 (0.58)	-18.18 (0.55)
DAX-DOWJONES-NIKKEI-CAC40	-21.10 (0.97)	-16.95 (0.61)	-13.48 (0.46)	-22.19 (0.71)	-18.78 (0.50)	-15.74 (0.41)
DAX-NEMAX-DOWJONES-NASDAQ	-41.34 (1.74)	-33.25 (1.34)	-26.65 (1.02)	-40.04 (1.27)	-33.80 (0.92)	-28.41 (0.73)
FORD-BOEING-GM-DELL-CISCO-MICROSOFT	-40.28 (1.20)	-32.27 (1.22)	-25.89 (0.89)	-38.93 (1.43)	-32.57 (1.10)	-27.21 (0.92)
Assets	t VAR $\alpha = 0.01$	0.025	0.05	GH VAR $\alpha = 0.01$	0.025	0.05
DAX-DOWJONES	-35.31 (2.14)	-27.12 (1.33)	-21.37 (0.84)	-34.56 (1.64)	-27.12 (1.14)	-21.30 (0.73)
NEMAX-NASDAQ	-59.83 (2.90)	-46.39 (2.02)	-36.91 (1.37)	-57.27 (2.86)	-44.91 (1.99)	-35.56 (1.34)
DAX-CAC40	-33.74 (2.42)	-25.13 (1.26)	-19.43 (0.73)	-31.97 (1.56)	-24.76 (0.96)	-19.24 (0.67)
NIKKEI-CAC40	-29.73 (2.30)	-21.81 (1.13)	-16.72 (0.69)	-29.56 (1.73)	-22.71 (1.03)	-17.65 (0.72)
DAX-DOWJONES-NIKKEI-CAC40	-25.41 (1.84)	-18.89 (1.02)	-14.56 (0.62)	-22.88 (1.15)	-17.55 (0.69)	-13.74 (0.51)
DAX-NEMAX-DOWJONES-NASDAQ	-44.65 (2.44)	-34.35 (1.62)	-27.29 (1.06)	-43.30 (2.22)	-33.83 (1.42)	-26.49 (1.08)
FORD-BOEING-GM-DELL-CISCO-MICROSOFT	-43.97 (2.45)	-33.51 (1.54)	-26.22 (1.03)	-46.69 (2.77)	-35.73 (1.69)	-27.36 (1.25)

7. Conclusion

The problem addressed here is the modelling of stock-price and asset-return distributions in higher dimensions, motivated by questions of portfolio selection and risk management in finance. The model proposed here is semi-parametric, and uses elliptically-contoured distributions, specifically normal variance mixtures with self-decomposable mixing distributions, with particular emphasis on the density generator g. In our view, this approach has much to recommend it: it has very nice theoretical properties, is easy and convenient to implement in simulation studies, and provides a good fit to a range of real financial data sets.

Many interesting theoretical and practical questions remain. We single out two: extension from normal variance mixtures to normal mean-variance mixtures (which leads beyond the elliptically-contoured framework), and integration of our approach with the most closely related recent and contemporary work, that of Barndorff-Nielsen, Jensen, and Sørensen (1998), Korsholm (2000), Hodgson, Linton, and Vorkink (2002), Barndorff-Nielsen and Shephard (2001), Barndorff-Nielsen and Shephard (2002).

We close by reminding ourselves, and the reader, that the real world of finance poses many questions even more basic than those of the selection and management of portfolios of stocks that motivate us here. The crux is questions of *asset allocation*, between stocks (of whatever kind) and bonds (of whatever kind). Here fund managers daily make highly committal choices when investing the contributions from our pension funds and yours, for instance. Dwarfing the 'tactical' choices involved such routine operations is the underlying 'strategic' question of the irrational exuberance of such markets, to quote the immortal words of the President of the Federal Reserve Bank, Mr. Alan Greenspan. The real questions here involve market psychology, which are inherently harder, because less amenable to quantitative treatment, than the questions treated here (for an entertaining account see Shiller (2000)).

Aldous, J., 1985, *Exchangeability and related topics* vol. 1117 of *École d' éte de probabilités de Saint Flour XIII, Lecture Notes in Math.* (Springer, Berlin).

Anderson, T.W., K.-T. Fang, and H. Hsu, 1986, Maximum-likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions, *Can. J. Statist.* 14, 55–59.

Baringhaus, L., 1991, Testing for spherical symmetry of a multivariate distribution, *Ann. Stat.* 19, 899–917.

Barndorff-Nielsen, O.E., J.L Jensen, and M. Sørensen, 1998, Some stationary processes in discrete and continuous time, *Adv. Appl. Probab.* 30, 989 – 1007.

Barndorff-Nielsen, O.E., J.T. Kent, and M. Sørensen, 1982, Normal variance-mean mixtures and z-distributions, *Internat. Stat. Review* 50, 145–159.

Barndorff-Nielsen, O.E., and N. Shephard, 2001, Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, *J. R. Statist. Soc. B* 63, 167–241.

Barndorff-Nielsen, O.E., and N. Shephard, 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *J. R. Statist. Soc. B* 64, 253–280.

Bauer, C., 2000, Value at risk using hyperbolic distributions, *J. Economics and Business* 52, 455–467.

Beran, R., 1979, Testing for ellipsoidal symmetry of a multivariate density, *Ann. Statist.* 7, 150–162.

Bingham, N.H., Goldie C.M. Teugels J.L., 1987, *Regular Variation.* (Cambridge: Cambridge University Press).

Bingham, N.H., and R. Kiesel, 2001a, Hyperbolic and semiparametric models in finance, in P. Sollich, A.C.C. Coolen, L.P. Hughston, and R.F. Streater, eds.: *Disordered and complex systems* (Amer. Inst. of Physics,).

Bingham, N.H., and R. Kiesel, 2001b, Modelling asset return with hyperbolic distributions, in J. Knight, and S. Satchell, eds.: *Asset return distributions* (Butterworth-Heinemann,).

Bingham, N.H., and R. Kiesel, 2002, Semi-parametric modelling in finance: theoretical foundation, *Quantitative Finance 2* pp. 241–250.

Carr, P., E. Chang, and D.B. Madan, 1998, The Variance-Gamma process and option pricing, *European Finance Review* 2, 79–105.

Dacorogna, M.M., U.A. Müller, and O.V. Pictet, 1998, *Heavy tails in high-frequency financial data*eds.: Adler, R.J. and Feldmann, R.E. and Taqqu, M.S., in A practical guide to heavy tails. Statistical techniques and applications, Springer pp. 55–77.

Dowd, K., 1998, *Beyond Value at Risk: The New Science of Risk Management.* (John Wiley & Sons, Chichester).

Eberlein, E., 2001, Applications of generalized hyperbolic Lévy motions to finance, in O.E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds.: *Lévy processes: Theory and Applications* (Birkhäuser, Boston,).

Eberlein, E., and F. Özkan, 2003, Time consistency of Lévy models, *Quantitative Finance* 3, 40–50.

Eberlein, E., and K. Prause, 2002, The generalized hyperbolic model: Financial derivatives and risk measures, in H. Geman, D. Madan, S.R. Pliska, and T. Vorst, eds.: *Mathematical Finance: Bachelier Congress 2000* (Springer,).

Embrechts, P., C. Klüppelberg, and T. Mikosch, 1997, *Modelling extremal events.* (Springer, New York Berlin Heidelberg).

Embrechts, P., A. McNeil, and D. Straumann, 2001, Correlation and dependency in risk management: Properties and pitfalls, in M. Dempster, and H.K. Moffat, eds.: *Risk Management: Value at Risk Beyond* (Cambridge University Press,).

Fang, K.-T., S. Kotz, and K.-W. Ng, 1990, *Symmetric multivariate and related distributions.* (Chapman & Hall, London).

Fang, K.-T., and Y.-T. Zhang, 1990, *Generalized multivariate analysis.* (Science Press, Beijing; Springer-Verlag, Heidelberg).

Frahm, G., and M. Junker, 2003, Generalized elliptical distributions – models and estimation, *Research Center Caesar, Bonn Germany, preprint.*

Frahm, G., M. Junker, and R. Schmidt, 2002, Estimating the tail-dependence coefficient, *www.mathematik.uni-ulm.de/finmath/*.

Hahn, M.G., D.M. Mason, and D.C. Weiner (eds.), 1991, *Sums, trimmed sums and extremes.* (Birkhäuser, Boston).

Halgreen, C., 1979, Self-decomposability of the generalized inverse Gaussian and hyperbolic distribution functions, *Z. Wahrschein.* 47, 13–18.

Härdle, W., 1990a, *Applied nonparametric regression* vol. 19 of *Econometric Society Monographs.* (Cambridge University Press).

Härdle, W., 1990b, *Smoothing techniques: With implementation in S.* (Springer-Verlag).

Hodgson, D.J., O. Linton, and K. Vorkink, 2002, Testing the capital asset pricing model efficiently under elliptical symmetry: a semi-parametric approach, .

Hull, J., and A. White, 1998, Value at risk when daily changes in market variables are not normally distributed, *J.Derivatives* 5, 9–19.

Joe, H., 1997, *Multivariate models and dependence concepts* vol. 73 of *Monographs on Statistics and Applied Probability.* (Chapman & Hall, London).

Jones, M.C., 1993, Simple boundary correction for kernel density estimation, *Statist. Comput.* 3, 135–146.

Jorion, P., 2000, *Value at Risk.* (McGraw-Hill, New York) 2nd. edn.

Kiesel, R., W. Perraudin, and A.P. Taylor, 2003, An extremes analysis of VaRs for emerging market benchmark bonds, in *Proceedings of the Econometrics Workshop in Karlsruhe* (Springer Verlag,).

Korsholm, L., 2000, The semi-parametric normal mean-variance mixture model, *Scand. J. Statist.* 27, 227–261.

Li, R., K.-T. Fang, and L. Zhu, 1997, Some Q-Q probability plots to test spherical and elliptical symmetry, *J. Comp. Graph. Statist.* 6, 435–450.

Madan, D.B., and E. Seneta, 1990, The Variance-Gamma (VG) model for share market returns, *Journal of Business* pp. 511–524.

Manzotti, A., F.J. Pérez, and A.J. Quiroz, 2002, A statistic for testing the null hypothesis of elliptical symmetry, *Journal of Multivariate Analysis* 81, 274–285.

Müller, H.G., 1991, Smooth optimum kernel estimators near endpoints, *Biometrika* 78, 521–530.

Nelsen, R.B., 1999, *An introduction to copulas* vol. 139 of *Springer Lecture Notes in Statistics.* (Springer).

Pagan, A.R., 1996, The econometrics of financial markets, *J. Empirical Finance* 3, 15–102.

Park, B.U., and J.S. Marron, 1990, Comparison of data-driven bandwidth selectors, *J. Am. Statist. Ass.* 85, 66–73.

Resnick, S., 1987, *Extreme Values, Regular Variation, and Point Processes.* (Springer New York).

Rousseeuw, P.J., and B.C. van Zomeren, 1990, Unmasking multivariate outliers and leverage points, *Journal of American statistics association* 85, 633–639.

Sato, K.-I., 1999, *Lévy processes and infinite divisibility* vol. 68 of *Cambridge studies in advanced mathematics.* (Cambridge Univ. Press).

Schmidt, R., 2002, Tail dependence for elliptically contoured distributions, *Math. Methods of Operations Research* 55, 301–327.

Schmidt, R., and U. Stadtmüller, 2002, Nonparametric estimation of tail dependence, *www.mathematik.uni-ulm.de/finmath/*.

Schmitt, F., D. Schertzer, and S. Lovejoy, 1999, Multifractal analysis of foreign exchange data, *Appl. Stoch. Models & Data Analysis* 15, 29–53.

Schuster, E.F., 1985, Incorporating support constraints into nonparametric estimators of densities, *Comm. Statist. Theor. Math.* 14, 1123–1136.

Sheather, S.J., and M.C. Jones, 1991, A reliable data-based bandwidth selection method for kernel density estimation, *J.R. Statist. Soc. B* 53, 683–690.

Shiller, R.J., 2000, *Irrational Exuberance.* (Princeton University Press).

Stute, W., and U. Werner, 1991, Nonparametric estimation of elliptically contoured densities, in G. Roussas, eds.: *Nonparametric functional estimation and related topics* (Kluwer Academic

Publishers,).

Venables, W.N., and B.D. Ripley, 1999, *Modern Applied Statistics with S-PLUS.* (Springer-Verlag New York Berlin Heidelberg) 3rd edn.

Wand, M.P., and M.C. Jones, 1995, *Kernel Smoothing.* (Chapman & Hall, London).

Niederreiter, H., 1993, *Random number generation and quasi-Monte Carlo methods.* (SIAM Philadelphia).