THE MULTIVARIATE NORMAL INVERSE GAUSSIAN DISTRIBUTION: EM-ESTIMATION AND ANALYSIS OF SYNTHETIC APERTURE SONAR DATA

Tor Arne Øigård [a], Alfred Hanssen [b], and Roy Edgar Hansen [c]

[a] Department of Mathematics and Statistics, University of Tromsø, NO-9037 Tromsø, Norway
torarne@stat.uit.no
[b] Department of Physics, University of Tromsø, NO-9037 Tromsø, Norway
alfred@phys.uit.no
[c] Norwegian Defence Research Establishment, NO-2027 Kjeller, Norway
Roy-Edgar.Hansen@ffi.no

ABSTRACT

The heavy-tailed Multivariate Normal Inverse Gaussian (MNIG) distribution is a recent variance-mean mixture of a multivariate Gaussian with a univariate inverse Gaussian distribution. Due to the complexity of the likelihood function, parameter estimation by direct maximization is exceedingly difficult. To overcome this problem, we propose a fast and accurate multivariate Expectation-Maximization (EM) algorithm for maximum likelihood estimation of the scalar, vector, and matrix parameters of the MNIG distribution. Important fundamental and attractive properties of the MNIG as a modeling tool for multivariate heavy-tailed processes are discussed. The modeling strength of the MNIG, and the feasibility of the proposed EM parameter estimation algorithm, are demonstrated by fitting the MNIG to real world wideband synthetic aperture sonar data.

1. INTRODUCTION

In most applications of statistical signal processing, Gaussian random processes are assumed since other distributional assumptions often lead to untractable mathematical difficulties. However, in many practical instances the measured probability density function of the random process exhibits much heavier tails than the Gaussian distribution. A number of models have been proposed for such heavy tailed random processes. In the last two decades data with heavy tails have been collected in several fields.

The *multivariate normal inverse Gaussian* (MNIG) is a recent variance-mean mixture of a multivariate Gaussian distribution with an inverse Gaussian mixing distribution. Recently, there has been an increasing interest in such models for financial and signal processing applications, mainly because the resulting distributions are highly suitable to model a large class of non-Gaussian semi heavy-tailed processes which also allows for skewness [4, 9]. In particular, we propose that the MNIG-model may be very useful for modeling multivariate impulsive noise.

Estimation of MNIG parameters is not treated to any great extent in the existing literature. The conventional way of estimating the MNIG parameters has been the maximum likelihood estimation method, but unfortunately, this method is complicated and computationally intensive since slowly converging numerical optimizations routines have to be applied [7]. Also, due to the complexity of the likelihood, direct maximization is difficult. In this paper we will present a multivariate EM-algorithm for estimation of the scalar, vector, and matrix parameters of the MNIG distribution. This EM-algorithm overcomes the numerical complexities, and it is readily implemented.

The paper is organized as follows. In section 2 we discuss the MNIG in detail, and the attractiveness of the MNIG-distribution as a practical tool for multivariate noise modeling is demonstrated. In section 3 we present a multivariate Expectation-Maximization (EM) algorithm for the estimation of the scalar, vector, and matrix parameters of the MNIG. In section 4 we analyze a real world wideband synthetic aperture sonar data set, and we fit the data to the MNIG model by applying our parameter estimation algorithm.

2. MULTIVARIATE NORMAL INVERSE GAUSSIAN DISTRIBUTION

Mixtures of normal distributions play an increasingly important role in the theory and practice of contemporary statistical modeling. Scale mixtures of the normal distribution, which assume that the variance is not fixed for all the members of the population, have been widely used to model heteroscedacity (non constant variance). Extending such models, normal variance-mean mixtures assume that the variance is not fixed but it is also related to the mean [3]. A rich family of distributions with useful properties can arise using this scheme.

An MNIG distributed random variable is a variance-mean mixture of a d-dimensional Gaussian random variable $\mathbf{Y}$ with a univariate inverse Gaussian distributed mixing variable Z. Hence, a MNIG distributed random variable with parameters $\alpha > 0$, $\beta \in \mathbb{R}^d$, $\delta > 0$, $\mu \in \mathbb{R}^d$, and $\mathbf{\Gamma} \in \mathbb{R}^{d \times d}$ can be constructed from

$$\mathbf{X} = \mu + Z\mathbf{\Gamma}\beta + \sqrt{Z}\mathbf{\Gamma}^{1/2}\mathbf{Y}, \tag{1}$$

where $Z \sim \mathrm{IG}\left[\delta^2, \alpha^2 - \beta^T\mathbf{\Gamma}\beta\right]$, $\mathrm{IG}[\chi,\psi]$, $\chi,\psi > 0$, denotes the inverse Gaussian distribution with probability density function [8]

$$f_Z(z) = \left(\frac{\chi}{2\pi z^3}\right)^{1/2} e^{\sqrt{\chi\psi}} \exp\left[-\frac{1}{2}\left(\chi z^{-1} + \psi z\right)\right], \quad z > 0, \tag{2}$$

and $\mathbf{Y} \sim \mathcal{N}_d[\mathbf{0},\mathbf{I}]$. Observe that the random variable $\mathbf{X}|Z$ is Gaussian with mean $\mu + Z\mathbf{\Gamma}\beta$ and variance $Z\mathbf{\Gamma}$, i.e. $\mathbf{X}|Z \sim \mathcal{N}_d\left[\mu + Z\mathbf{\Gamma}\beta, Z\mathbf{\Gamma}\right]$. Thus, we have a stochastic mean and variance when Z is stochastic, and hence the term variance-mean mixture.

From Eq. (1) we can calculate the probability density function of $\mathbf{X}$, and we say that a d-dimensional stochastic column vector $\mathbf{X}$ is MNIG distributed if the probability density function can be written as [5, 9, 16]

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \int f_{\mathbf{X}|Z}(\mathbf{x}|z) f_Z(z)\,dz & (3)\\ &= \frac{\delta}{2^{\frac{d-1}{2}}}\left[\frac{\alpha}{\pi q(\mathbf{x})}\right]^{\frac{d+1}{2}} \exp\left[p(\mathbf{x})\right] K_{\frac{d+1}{2}}\left[\alpha q(\mathbf{x})\right], & (4)\end{aligned}$$

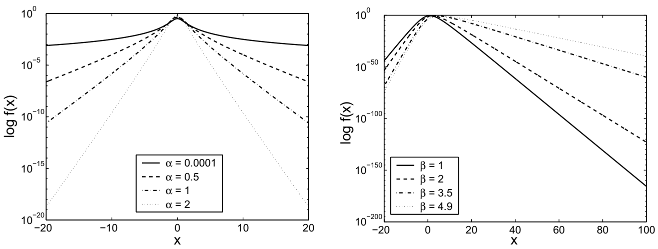
where

$$p(\mathbf{x}) = \delta\sqrt{\alpha^2 - \beta^T\mathbf{\Gamma}\beta} + \beta^T(\mathbf{x} - \mu),$$

and

$$q(\mathbf{x}) = \sqrt{\delta^2 + \left[(\mathbf{x} - \mu)^T\mathbf{\Gamma}^{-1}(\mathbf{x} - \mu)\right]}.$$

We thank Sean Chapman and his team at QinetiQ, UK, for providing the high resolution synthetic aperture sonar data.



(a) (b)

Figure 1: (a) Univariate NIG-density for different values of α. Here, $\beta = \mu = 0$, and $\delta = 1$, (b) NIG-density for different values of β. Here, $\alpha = \delta = 5$, and $\mu = 0$.

Here, $K_d(x)$ is the modified Bessel function of the second kind with index d. As seen from definition in Eq. (4) the shape of the MNIG density is specified by two scalar parameters $\alpha, \delta > 0$, two vector parameters β, and μ and one matrix parameter $\mathbf{\Gamma}$. This parameterization is very flexible indeed, making it possible to model a large variety of unimodal shapes and with various decay rates of the tails.

The parameters of the MNIG-distribution have natural interpretations relating to the overall shape of the density as follows. The parameter α controls the "steepness" of the density, in the sense that the steepness or pointiness of the density increases monotonically with increasing α. This has implications also for the tail behavior by the fact that large values of α implies light tails, while smaller values of α implies heavier tails. The parameter β is a vector skewness parameter, the parameter δ is a scale parameter, and the parameter μ is a vector translation parameter. The structure matrix $\mathbf{\Gamma}$ is assumed to be a positive semidefinite symmetric matrix with unity determinant, $\det \mathbf{\Gamma} = 1$. This matrix is decisive in controlling the degree of correlations between the components of $\mathbf{X}$. Note that the inequality $\alpha^2 > \beta^T \mathbf{\Gamma} \beta$ has to be satisfied for the MNIG to exist [5].

By choosing $\mathbf{\Gamma}$ appropriately, and by choosing the parameters α, β, μ, and δ, one may use the multivariate NIG to model statistically dependent (semi) heavy-tailed data.

In Fig. 1 we show the univariate NIG ($d = 1$) as a special case. The left panel shows the dependency on α for fixed $\beta = \mu = 0$, and $\delta = 1$. Note that the tails become heavier as the value of the α parameter decreases. The right panel shows the dependency on the β parameter for fixed $\alpha = \delta = 1$, and $\mu = 0$. Note that the skewness increases as β increases. Note that the vertical scale is logarithmic to emphasize the tails.

MNIG variables obey several desirable properties that make them suitable for practical multivariate modeling. We will in the following sections demonstrate the attractiveness of the MNIG distribution in terms of some of its properties.

2.1 Moments of the MNIG

The cumulant generating function of the MNIG is given by [9, 16]

$$\Psi_X(\omega) = \delta \left[\sqrt{\alpha^2 - \beta^T \mathbf{\Gamma} \beta} - \sqrt{\alpha^2 - (\beta + j\omega)^T \mathbf{\Gamma} (\beta + j\omega)} \right] + j\omega^T \mu. \quad (5)$$

Note the simple form of the cumulant generating function despite the relatively complex form of the probability density function. Also, from Eq. (5) we note that the MNIG distribution is infinitely divisible.

The mean vector variable can now easily be evaluated to give [9]

$$E\{\mathbf{X}\} = \mu + \frac{\delta \mathbf{\Gamma} \beta}{\sqrt{\alpha^2 - \beta^T \mathbf{\Gamma} \beta}}, \quad (6)$$

and the covariance matrix of $\mathbf{X}$ is given by [9]

$$\mathbf{\Sigma} = \delta \left(\alpha^2 - \beta^T \mathbf{\Gamma} \beta \right)^{-1/2} \left[\mathbf{\Gamma} + \left(\alpha^2 - \beta^T \mathbf{\Gamma} \beta \right)^{-1} \mathbf{\Gamma} \beta \beta^T \mathbf{\Gamma} \right]. \quad (7)$$

Notice that choosing $\mathbf{\Gamma} = \mathbf{I}$ is not sufficient to produce a diagonal covariance matrix. This is because the vector parameter β enters the first and second order statistics in a non-trivial manner. The MNIG is symmetric if and only if $\mathbf{\Gamma} = \mathbf{I}$ and $\beta = \mathbf{0}$. In this case we denote the MNIG distribution a *symmetric multivariate normal inverse Gaussian* (SMNIG) distribution. When $\beta = \mathbf{0}$ and $\mathbf{\Gamma} \neq \mathbf{I}$ the MNIG is *semi-symmetric*.

2.2 Summation property and transformation property

A very attractive and useful property of the MNIG that cannot be overrated, is that it exhibits a certain closedness under summation (but not for weighted summation). This has far reaching consequences when considering sums of MNIG variables.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_M$ be M independent MNIG-variables with common shape parameters α, β, and $\mathbf{\Gamma}$, but having individual location parameters $\mu_1, \ldots, \mu_M$ and individual scale parameters $\delta_1, \ldots, \delta_M$. Then, the sum variable $\mathbf{Y} = \mathbf{X}_1 + \cdots + \mathbf{X}_M$ is also MNIG distributed with parameters $(\alpha, \beta, \mu_{tot}, \delta_{tot})$, where $\mu_{tot} = \sum_{i=1}^M \mu_i$ and $\delta_{tot} = \sum_{i=1}^M \delta_i$.

Now, let $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ be an arbitrary linear transformation of $\mathbf{X}$, where $\mathbf{X} \sim \text{MNIG}[\alpha, \beta, \delta, \mu, \mathbf{\Gamma}]$, $\mathbf{A}$ is a real valued $d \times d$ coefficient matrix, and $\mathbf{b}$ is a $d \times 1$ dimensional real vector. Then, it can be shown that the transformed variable $\mathbf{Y}$ is also MNIG distributed with parameters $(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}, \tilde{\mathbf{\Gamma}})$ [4], where

$$\tilde{\alpha} = \alpha \left| \det \mathbf{A} \right|^{-1/d} \quad (8)$$

$$\tilde{\beta} = \mathbf{A}^{-T} \beta \quad (9)$$

$$\tilde{\delta} = \delta \left| \det \mathbf{A} \right|^{1/d} \quad (10)$$

$$\tilde{\mu} = \mathbf{A}\mu + \mathbf{b} \quad (11)$$

$$\tilde{\mathbf{\Gamma}} = \mathbf{A} \mathbf{\Gamma} \mathbf{A}^T \left| \det \mathbf{A} \right|^{-2/d}. \quad (12)$$

Here $|\det \mathbf{A}|$ denote the magnitude of the determinant of $\mathbf{A}$. Also, for the univariate NIG distribution we have that the parameters $\bar{\alpha} = \delta\alpha$ and $\bar{\beta} = \delta\beta$ are invariant under location-scale changes of the NIG distributed random variable X [5].

2.3 Limiting distributions

The multivariate Gaussian distribution is in fact a limiting distribution for the MNIG, $\mathbf{X} \sim \mathcal{N}_d \left[\mu + \sigma^2 \mathbf{\Gamma} \beta, \sigma^2 \mathbf{\Gamma} \right]$ in the limit $\delta \to \infty$ and $\alpha \to \infty$ but such that $\delta / \alpha = \sigma^2$.

Another important special case for the MNIG is the *multivariate Cauchy* distribution (i.e., the multivariate t-distribution with one degree of freedom [6]). This occurs when $\mathbf{\Gamma} = \mathbf{I}$ and $\alpha \to 0$.

When $\beta = \mathbf{0}$ the MNIG belongs to the class of elliptical distributions [6]. If in addition $\mathbf{\Gamma} = \mathbf{I}$ it belongs to the class of spherical distributions [6].

2.4 Tail behavior

Asymptotically, the Bessel function behaves as [1]

$$K_d(x) \sim \sqrt{\frac{\pi}{2x}} \exp(-x), \quad |x| \to \infty, \quad \forall d. \quad (13)$$

Hence, the tail of the MNIG decays as

$$f_{\mathbf{X}}(\mathbf{x}) \sim \|\mathbf{x}\|_{\mathbf{\Gamma}}^{-\frac{d+2}{2}} \exp\left(\beta^T \mathbf{x} - \alpha \|\mathbf{x}\|_{\mathbf{\Gamma}} \right), \quad (14)$$

for $\alpha \nrightarrow 0$, where $\|\mathbf{x}\|_{\mathbf{\Gamma}} = \left(\mathbf{x}^T \mathbf{\Gamma}^{-1} \mathbf{x}\right)^{1/2}$ is the weighted $\mathbf{\Gamma}$-norm of $\mathbf{x}$.

The multivariate Cauchy limit is obtained for $\alpha \to 0$. In this case the Bessel function asymptotically behaves as [1]

$$K_d(x) \sim |x|^{-d}, \quad |x| \to 0, \tag{15}$$

for which we find that the tail of the MNIG behaves as

$$f_{\mathbf{X}}(\mathbf{x}) \sim \|\mathbf{x}\|_{\mathbf{\Gamma}}^{-(d+1)}, \quad \|\mathbf{x}\| \to \infty. \tag{16}$$

It is interesting to note that in the non-Cauchy limit, the tails exhibit a combined algebraic and exponential decay, and that the exponential component is controlled by the values of α and β.

3. THE MULTIVARIATE EXPECTATION-MAXIMIZATION ALGORITHM

In principle, one could find the maximum-likelihood estimates of the MNIG parameters. In practice, however, the direct maximization of the likelihood function proves to be difficult. The EM algorithm is a powerful algorithm for ML estimation for data containing "missing values" [12, 13]. This formulation is particularly suitable for distributions arising as mixtures since the mixing operation can be considered responsible for producing missing data [11]. The EM algorithm consists of two major steps: an expectation step, followed by a maximization step. The expectation is with respect to the unknown underlying variables, using the current estimate of the parameters and conditioned upon the observations. During the maximization step one maximizes the complete data likelihood using the expectations of the previous step. In [11] an EM-algorithm for estimation of the parameters of the univariate NIG distribution was developed. We will in this section provide an EM-type algorithm for the maximum likelihood estimation of the MNIG parameters. The algorithm is very easy to implement, it is numerically stable, and the convergence rate is reasonably fast. However, if the likelihood contains several modes it may be trapped in local extrema. This can be solved by choosing proper initial values, or by repeating the procedure several times with different initial values.

We assume that the true data $\mathbf{Y}_i = (\mathbf{X}_i, Z_i)$ consist of an observable part $\mathbf{X}_i$ and an unobservable part Z_i. In our case, the unobserved quantities Z_i are simply the realizations of the unobserved IG distributed mixing parameter for each data point [11].

As a helpful tool for the derivation of the EM-algorithm we exploit the fact that the conditional distribution $Z|\mathbf{X}$, where Z is GIG$[\lambda, \delta, \gamma]$ distributed, and $\mathbf{X}$ is the multivariate generalized hyperbolic distributed with parameters $(\lambda, \alpha, \beta, \delta, \mu, \mathbf{\Gamma})$, is also GIG distributed,

$$Z|\mathbf{X} \sim \mathrm{GIG}[\lambda - d/2, q(\mathbf{x}), \alpha]. \tag{17}$$

Note that the conditional distribution is independent of β. In our case it is sufficient to focus on the IG distribution, which is a special case of the GIG distribution for $\lambda = -1/2$ [4, 5].

For densities belonging to the exponential family, the E-step calculates the conditional expectations of the sufficient statistics of the unobserved data. Hence, at the E-step one needs to calculate the conditional expectation of the sufficient statistics for the inverse Gaussian distribution. These are $\sum_i Z_i$ and $\sum_i Z_i^{-1}$ [11]. Thus, in the E-step we need the first order moment and the inverse first moment of $Z|\mathbf{X}$. It can be shown that these are given by

$$\varsigma_i = \mathrm{E}\left\{Z_i | \mathbf{X}_i = \mathbf{x}_i\right\} = \frac{q(\mathbf{x}_i)}{\alpha} \frac{K_{(d-1)/2}[\alpha q(\mathbf{x}_i)]}{K_{(d+1)/2}[\alpha q(\mathbf{x}_i)]} \tag{18}$$

$$\varphi_i = \mathrm{E}\left\{Z_i^{-1} | \mathbf{X}_i = \mathbf{x}_i\right\} = \frac{\alpha}{q(\mathbf{x}_i)} \frac{K_{(d+3)/2}[\alpha q(\mathbf{x}_i)]}{K_{(d+1)/2}[\alpha q(\mathbf{x}_i)]}, \tag{19}$$

for $i = 1, \ldots, N$. Next, compute $\xi = \sum_{i=1}^N \varsigma_i / N$ and $\vartheta = N\left(\sum_{i=1}^N (\varphi_i - \xi^{-1})\right)^{-1}$.

The M-step maximizes the complete likelihood by updating the parameters using the expectations of the sufficient statistics found at the E-step. As mentioned earlier, we have that $\mathbf{X}_i|Z_i$ is Gaussian distributed. Hence, the log-likelihood function is given by

$$L(\mu, \beta, \mathbf{\Gamma} | \mathbf{x}, Z) \propto -\frac{N}{2} \log\det\mathbf{\Gamma} - \frac{1}{2}\sum_{i=1}^{N} \frac{1}{Z_i} \left(\mathbf{x}_i - \mu - Z_i \mathbf{\Gamma}\beta\right)^T \mathbf{\Gamma}^{-1} \left(\mathbf{x}_i - \mu - Z_i \mathbf{\Gamma}\beta\right).$$

From this we get the maximum likelihood equations

$$\frac{\partial L}{\partial \mu} = -\mathbf{\Gamma}^{-1} \sum_i \frac{1}{Z_i} \mathbf{x}_i + N\beta + \mathbf{\Gamma}^{-1} \mu \sum_i \frac{1}{Z_i} = \mathbf{0}, \tag{20}$$

$$\frac{\partial L}{\partial \beta} = -\sum_i \mathbf{x}_i + N\mu + \mathbf{\Gamma}\beta \sum_i Z_i = \mathbf{0}, \tag{21}$$

$$\frac{\partial L}{\partial \mathbf{\Gamma}} = \frac{1}{2}\left[\mathbf{\Gamma}^{-1}(\mathbf{R} - \mathbf{\Gamma})\mathbf{\Gamma}^{-1} - \mathbf{A}\right] = \mathbf{0}, \tag{22}$$

where $\mathbf{R} = N^{-1} \sum_i (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T / Z_i$, and $\mathbf{A} = \beta\beta^T N^{-1} \sum_i 1/Z_i$. Then, by solving for the parameters using Eqs. (20) – (22), and replacing the quantities Z_i and Z_i^{-1} with their respective estimators, we get the following set of estimators

$$\widehat{\delta} = \sqrt{\vartheta} \tag{23}$$

$$\widehat{\beta} = \widehat{\mathbf{\Gamma}}^{-1} \left(\frac{\sum_i \mathbf{x}_i \varphi_i - \bar{\mathbf{x}} \sum_i \varphi_i}{n - \xi \sum_i \varphi_i}\right) \tag{24}$$

$$\widehat{\mu} = \bar{\mathbf{x}} - \xi \widehat{\mathbf{\Gamma}} \widehat{\beta} \tag{25}$$

$$\widehat{\alpha} = \sqrt{\left(\widehat{\delta}/\xi\right)^2 + \widehat{\beta}^T \widehat{\mathbf{\Gamma}} \widehat{\beta}}, \tag{26}$$

where $\bar{\mathbf{x}} = \sum_i \mathbf{x}_i / n$. The $\mathbf{\Gamma}$-matrix can then be estimated as follows

$$\widehat{\mathbf{\Gamma}} = \frac{\mathbf{R}(\mathbf{B} + \mathbf{I})^{-1}}{\left\{\det\left[\mathbf{R}(\mathbf{B} + \mathbf{I})^{-1}\right]\right\}^{1/d}}, \tag{27}$$

where

$$\mathbf{B} = \frac{1}{2}\left[-\mathbf{I} + \mathbf{W}\left(\mathbf{I} + 4\mathbf{D}_{AR}\right)^{1/2} \mathbf{W}^{-1}\right], \tag{28}$$

and $\mathbf{D}_{AR} = \mathbf{W}^{-1} \mathbf{ARW}$, where $\mathbf{D}_{AR}$ is a diagonal matrix of eigenvalues, and $\mathbf{W}$ is a matrix whose columns are the corresponding eigenvectors so that $\mathbf{ARW} = \mathbf{WD}_{AR}$. The algorithm is iterated between these two steps until a reasonable convergence criterion has been reached. Observe that $\det\widehat{\mathbf{\Gamma}} = 1$, as required. Also, we see that in the symmetric case ($\beta = \mathbf{0}$), we have that the matrix $\mathbf{B} = \mathbf{0}$. Thus, in this particular case the $\mathbf{\Gamma}$ estimator reduces to a *weighed empirical correlation matrix estimator* where the weights are the φ_i's. The algorithm yields very accurate estimates, even for small data sets, and it converges fast to a stable solution. The convergence of the univariate version of the EM algorithm is studied thoroughly in [11].

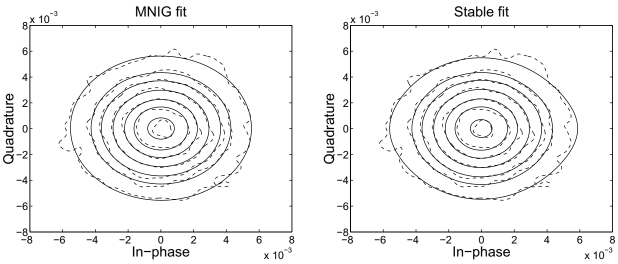
4. WIDEBAND SYNTHETIC APERTURE SONAR DATA

In this section we study the statistics of randomly selected pixels in a wideband synthetic aperture sonar image of the seafloor [10]. The sonar operates with 64 kHz bandwidth around 150 kHz center frequency, and the length of the synthetic aperture is 15 m. This gives a theoretical resolution of 1.2 cm x 1.5 cm in the processed image. The data for the statistical analysis are taken from a selected area of the processed image. The area contains no visible features on the seafloor, i.e., only background acoustical noise and reverberation is assumed to be present.

We fitted a bivariate NIG to the complex in-phase/quadrature pixel pairs constituting the complex image. The MNIG parameters were found to be

$$\widehat{\alpha} = 1885 \qquad \widehat{\delta} = 0.0066$$
$$\widehat{\beta} = \begin{bmatrix} -3.1982 \\ 23.747 \end{bmatrix} \qquad \widehat{\mu} = \begin{bmatrix} 1.871 \cdot 10^{-3} \\ -6.081 \cdot 10^{-3} \end{bmatrix}$$
$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} 0.9888 & -0.0025 \\ -0.0025 & 1.0113 \end{bmatrix}.$$

The left and right panel of Fig. 2 shows level plots of the MNIG ($d = 2$) model fit and the bivariate Stable model fit, respectively, for the wideband synthetic aperture sonar data. The model fits were compared to a non-parametric bivariate kernel density estimate [12] where the Gaussian kernel was used, and the smoothing parameters were $h_I = 0.029$ and $h_Q = 0.030$. When fitting the complex valued synthetic aperture sonar data to the bivariate Stable model, we assumed independence between the in-phase and the quadrature components. The program STABLE described in [14] was used to estimate the parameters of the time series and to calculate the Stable densities $f_I(x)$ and $f_Q(y)$ for the in-phase and quadrature components, respectively. The bivariate Stable density was calculated by means of $f(x,y) = f_I(x) f_Q(y)$. We see that both the MNIG fit and the Stable fit appears to model the radar data with a high degree of accuracy, both around the mode and in the tails.



(a) (b)

Figure 2: Level plots of the (a) bivariate NIG model fit (solid) and (b) bivariate Stable model fit (solid), and the bivariate kernel density estimate (dashed) from the wideband synthetic aperture sonar data. Smoothing parameters were $h_I = 0.029$, and $h_Q = 0.030$, and the smoothing kernel was Gaussian.

5. CONCLUSION

In this paper we reviewed the recent multivariate normal inverse Gaussian (MNIG) distribution. We demonstrated that the parametrization of the MNIG-distribution allows for a very flexible formulation of multivariate leptokurtic data. By adjusting the parameters of the MNIG distribution, we argued that one may capture the behavior of a large number of non-Gaussian multivariate distributions with tails ranging from the multivariate Gaussian to the multivariate Cauchy distribution.

Among the many advantages of the MNIG-distribution, we recall that (i) the density is given in closed form, (ii) MNIG variables are closed under summation, and (iii) the parametrization of the model allows for a large range of probability density shapes with an arbitrary degree of multivariate skewness.

We introduced a fast and efficient EM-algorithm for maximum likelihood estimation of the MNIG parameters. The proposed EM-algorithm has several advantages over direct maximization of the likelihood function. E.g., the EM-algorithm is readily implemented, it is numerically stable, and it is very accurate.

We fitted synthetic aperture sonar data to the MNIG distribution, and we found that the MNIG captured the inherent impulsiveness of these data with great accuracy. We believe that the MNIG distribution will prove to be useful for the modeling of impulsive signals, noise, and interference in multivariate sonar, radar, and communication systems.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1974.

[2] O. E. Barndorff-Nielsen, "Exponentially decreasing distributions for the logarithm of particle size," *Proc. Roy. Soc. Lond.*, Vol. A353, pp. 401–419, 1977.

[3] O. E. Barndorff-Nielsen, J. Kent and M. Sørensen, "Normal variance-mean mixtures and z-distributions," *International Statistical Review*, Vol. 50, pp. 145-159, 1982.

[4] O. E. Barndorff-Nielsen, "Normal inverse Gaussian processes and the modeling of stock returns," Research Report No. 300, Dept. of Theoretical Statistics, Inst. of Mathematics, Univ. of Aarhus, Denmark, 1995.

[5] O. E. Barndorff-Nielsen, "Normal inverse Gaussian distributions and stochastic volatility modeling," *Scand. J. Statist.*, Vol. 24, pp. 1–13, 1997.

[6] M. Bilodeau, and D. Brenner, *Theory of Multivariate Statistics*, New York, NY: Springer, 1999.

[7] E. Bølviken and F. E. Benth, "Quantification of risk in Norwegian stocks via the normal Gaussian distribution," Research Report Sand/04/00, Norwegian Computing Center, Oslo 2000.

[8] J. L. Folks and R. S. Chhikara, "The inverse Gaussian distribution and its statistical application — A review," *J. R. Statist. Soc. B*, Vol. 40, pp. 263-289, 1978.

[9] A. Hanssen and T. A. Øigård, "The normal inverse Gaussian distribution as a flexible model for heavy-tailed processes", in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Proccessing*, Baltimore, MD, June 2001, 5 pp.

[10] A. Hanssen, J. Kongsli, R.E. Hansen, and S. Chapman, "Statistics of synthetic aperture sonar images," in *Proc. MTS/IEEE Oceans 2003 Marine Technology and Ocean Science Conf.*, San Diego, CA, Sept. 22 - 26, 2003, pp. 2635 - 2640.

[11] D. Karlis, "An EM type algorithm for maximum likelihood estimation for the normal inverse Gaussian distribution," *Statistics and Probability Letters*, Vol. 57, pp. 43-52, 2002.

[12] W. L. Martinez and A. R. Martinez, *Computational Statistics Handbook with MATLAB*, Boca Raton, FL: Chapman & Hall/CRC, 2002.

[13] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, Vol. 13, pp. 47-60,Nov 1996.

[14] J. P. Nolan, "Fitting data and assessing goodness of fit with stable distributions," in *Proceedings of the Conference of Applications of Heavy tailed Distributions in Economics, Engineering and Statistics*, American University, Washington, DC, June 1999.

[15] T. H. Rydberg, "The normal inverse Gaussian Lévy process: Simulation and approximation," *Commun. Statist.–Stochastic Models*, Vol. 34, pp. 887 – 910, 1997.

[16] T. A. Øigård and A. Hanssen, "The multivariate normal inverse Gaussian heavy-tailed distribution: Simulation and estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, USA, May 2002, Vol. 2, pp. 1489-1492.