ON THE MAP ESTIMATION IN THE CONTEXT OF ELLIPTICAL DISTRIBUTIONS

S. Zozor [1] *and C. Vignat* [2]

[1] GIPSA-Lab, Departement Image et Signal, 961 Rue de la Houille Blanche, 38402 St Martin d'Hères, France
email: steeve.zozor@gipsa-lab.inpg.fr

[2] Institut Gaspard Monge, Université de Marne-la-Vallée, 5 Boulevard Descartes, 77454 Marne-la-Vallée, France
email: vignat@univ-mlv.fr

ABSTRACT

The purpose of this paper is to study the estimation problem of a multivariate elliptically symmetric random variable corrupted by a multivariate elliptically symmetric noise. In this study, the maximum a posteriori (MAP) approach is presented, extending recent works by Alecu *et al.* [1] and Selesnick [2, 3]: (i) the estimation is performed in a multivariate context, (ii) the corrupting noise is not limited to be Gaussian. This paper also extends our previous work that dealt with the minimum mean square error (MMSE) approach [4]. The MMSE is briefly recalled and the MAP is derived. Then the practical use of the MAP in a general setting is discussed and compared to that of the MMSE and of the Wiener estimator. Several examples illustrate the behaviors of these estimators and exhibit their performances.

1. INTRODUCTION

Investigations on elliptically symmetric distributions appear recurrently in the signal processing literature; very recent contributions such as [1, 2, 3] are examples. The interest of elliptically symmetric random vectors lies in the fact that they can be viewed as natural extensions of multivariate Gaussian random vectors. As a consequence, various processings used in the Gaussian case can be naturally extended to the class of elliptically symmetric distributions.

Among the first works on elliptically symmetric distributions in signal processing, one must cite the two almost simultaneous papers by Chu [5] and Yao [6]; the first one deals with the subclass of Gaussian scale mixture (GSM) distributions in the frameworks of optimal estimation, filtering and stochastic control. The second paper provides several statistical results for estimation and detection in the more general case of elliptically symmetric distributions. A more recent contribution can be found for example in [7, Ch. 2].

The study of such random vectors gave rise to various applications such as radar processing [8], image processing [9, 2], blind source separation [10] or parameter estimation [11] to cite recent contributions.

This paper aims at extending recent results on estimation in the elliptical context. The first study is that of Alecu *et al.* which dealt with what they call the Gaussian transform of a symmetrically distributed scalar random variable [1]. The Gaussian transform of a GSM is nothing but the probability density function (pdf) of the square of its mixing variable. In their paper, Alecu *et al.* revisited the denoising problem in the context of a one-dimensional GSM corrupted by a Gaussian noise. The second study is that of Selesnick, in which the denoising problem is presented in the context of a spherically invariant radial exponential or Laplacian random vector corrupted by a Gaussian white noise [2, 3]. Laplacian or radial exponentially distributed random vector also enter into the class of elliptically (more specifically spherically) distributed random vector and are GSM. As in [1, 2, 3], this work can find applications in image denoising, or even in radar data processing since radar clutter is often modeled by GSM processes [8].

In a previous work, we have revisited the denoising problem in the general case of a (wide sense) GSM, corrupted by a (wide sense) GSM noise [4]. The proposed approach was based on the minimal mean square error estimator (MMSE). The purpose of this proposal is to extend as far as possible the denoising problem of a multivariate elliptically symmetric random variable corrupted by a multivariate elliptically symmetric noise even when one or both vectors cannot be expressed as a (wide sense) GSM. As in [1, 3], we concentrate here on the maximum a posteriori estimator (MAP).

In section 2 we briefly recall the basics about elliptically invariant random vectors we need for our approach. Then, section 3 is devoted to the denoising problem where both the vector to estimate and the corrupting noise are elliptically distributed. In this section we will revisit the MAP estimator in this elliptically distributed framework. We will show that under an additional assumption on the covariance matrices of X and Z, the MAP estimator reduces to a one dimensional problem. We also give some conditions which ensure the existence of a MAP estimator, and a possible recursive approach to determine the MAP. Finally, the shape of the MAP estimator and its performances are exhibited, and compared to the MMSE and to the Wiener estimator ones.

2. RECALLS ON ELLIPTICAL DISTRIBUTIONS

A d-dimensional random vector X is elliptically distributed if its probability density function[1] p_X is a function of the quadratic form $(x-\mu)^t R^{-1}(x-\mu)$, that is

$$p_X(x) = |R_X|^{-1/2} d_X\left((x-\mu_X)^t R_X^{-1}(x-\mu_X)\right)$$

where d_X is a function $\mathbb{R}_+ \to \mathbb{R}_+$, called density generator, where μ_X is a location parameter, and where R_X is a symmetric definite positive matrix called characteristic matrix [5, 6, 7]. The density p_X is also said elliptically contoured, in the sense that the constant probability hypersurfaces are ellipsoidal. When defined, μ_X is the mean of X. Matrix R_X is defined up to a scaling factor, however, the indeterminacy can be removed by imposing a constraint on R_X. As an example, when X has a covariance, imposing $\int_{\mathbb{R}_+} r^{d+1} d_X\left(r^2\right) dr = \Gamma(d/2+1)/\pi^{\frac{d}{2}}$, where Γ is the Euler

[1] More rigorously such vector are defined via the characteristic function [7]. The elliptical property is preserved by Fourier transform [12, 6, 4], and in this paper we restrict to vectors that admit a pdf

Gamma function, removes this indeterminacy, implying that R_X is the covariance matrix of X. Finally, note that for any invertible matrix M, vector MX remains elliptically distributed. In particular, the elliptical vector $R_X^{-1/2}X$ has a characteristic matrix equal to the $(d \times d)$ identity, and is said spherically or orthogonally invariant (or distributed).

In the sequel, without loss of generality, we will focus on the situation where $\mu_X = 0$.

In the case where function $s \mapsto d_X(s)$ admits an inverse Laplace transform (see [5]), the pdf p_X can be expressed as

$$p_X(x) = \int_0^{+\infty} f_a(a) \mathscr{G}(x|_{a^2 R_X})\, da \tag{1}$$

where $\mathscr{G}(x|_R) = (2\pi)^{-\frac{d}{2}} |R|^{-\frac{1}{2}} \exp\left(-\frac{x^t R^{-1} x}{2}\right)$ denotes the d-dimensional zero-mean Gaussian pdf with covariance matrix R. Integrating p_X over x shows that, provided that integration signs can be exchanged, function f_a sums to 1. Moreover, under the condition that d_X is a completely monotone function, function f_a is nonnegative [6, 13, 14] and hence is a pdf. In terms of random vector, X has then the stochastic representation of a Gaussian scale mixture (GSM)[2] $X \stackrel{d}{=} AN$ where $A \sim f_a$ and N is a Gaussian vector independent of A, of covariance matrix R_X, $N \sim \mathscr{G}(\cdot|_{R_X})$. Equality is in the sense of distributions. When the decomposition (1) holds but f_a is not a pdf, we will call X wide sense GSM.

Many properties of elliptically distributed random vector can be found in [5, 6, 7].

3. RANDOM VECTOR ESTIMATION IN THE ELLIPTICALLY FRAMEWORK

We consider the generalization of the problem of [1, 3] of estimating a d-dimensional random vector X elliptically distributed and with covariance matrix R_X, from a noisy observation

$$Y = X + Z \tag{2}$$

where the noise Z is assumed independent of X and elliptically distributed with covariance matrix R_Z. As in [1, 3], we assume that both covariance matrices R_X and R_Z and pdfs p_X and p_Z are known. This problem generalizes [1] in the sense that dimension d is arbitrary and that Z can be non-Gaussian. It also generalizes [2, 3] which also restricted Z to be Gaussian and X to be radially exponential or Laplacian respectively.

Multiplying the observation by an adequate matrix, as done in [4], the problem can be restricted to the case where $R_Z = I$ (pre-whitening of Z) and $R_X = \Delta$ diagonal (adequate rotation after the pre-whitening of Z).

3.1 Recall on the Minimum Mean Square Estimation

The well-known Minimum Mean Square Error (MMSE) estimator of X based on the observation Y, *i.e.* the vector $\widehat{X}$ that minimizes the quadratic error $E[\|\widehat{X} - X\|^2]$, is given by the conditional mean [15], $\widehat{X}_{\text{mmse}} = E[X|Y]$. We have previously shown that, when both X and Z are wide sense GSM, with mixing functions f_a and f_b respectively, the MMSE estimator writes

$$\widehat{X}_{\text{mmse}} = \frac{\displaystyle\int_{\mathbb{R}_+^2} \mathscr{G}(y|_{a^2\Delta + b^2 I})\, a^2 f_a(a) f_b(b) (a^2\Delta + b^2 I)^{-1}\, da\, db}{\displaystyle\int_{\mathbb{R}_+^2} \mathscr{G}(y|_{a^2\delta + b^2 I}) f_a(a) f_b(b)\, da\, db} \Delta y \tag{3}$$

which requires integrations over $\mathbb{R}_+^2$ instead of the integrations over $\mathbb{R}^d$ required in general.

When both X and Z are Gaussian, one recovers the Wiener estimator

$$\widehat{X}_{\text{w}} = (\Delta + I)^{-1} \Delta y = \left(I + \Delta^{-1}\right)^{-1} y \tag{4}$$

which is also the best linear estimator in the mean square error (MSE) sense, for any uncorrelated vectors X and Z [15].

3.2 Maximum a posteriori estimation

The MMSE estimator in its form (3) can be used only when X and Z are wide sense GSM. When at least one of these vectors does not satisfy this assumption, the previous derivations do not hold anymore and the MMSE estimation can become much more complicated to derive. Such a situation occurs *e.g.* in the case of pdfs with bounded support. Indeed, for such pdfs, the density generator has obviously bounded support, and hence cannot admit an analytic continuation – otherwise it would be identically zero [16, identity theorem p. 65]. However, in order to admit an inverse Laplace transform, a function must be analytic [14, II.5 th. 5a]. The MMSE can also be difficult the implement if a mixing function exists, but has no analytical expression as in the generalized Gaussian case [4].

In such situations, the maximum a posteriori (MAP) approach is a good alternative, as shown in [1, 2, 3]. The MAP consists in seeking the vector $\widehat{X}$ that maximizes the posterior density $p_{X|Y}(x,y)$, what yields to

$$\begin{aligned} \widehat{X}_{\text{map}}(y) &= \arg\max_X p_X(x) p_Z(y - x) \\ &= \arg\max_X d_X(\|x\|^2) d_Z(\|\Delta^{-\frac{1}{2}}(y - x)\|^2) \end{aligned} \tag{5}$$

by the Bayes rule. Vanishing the gradient $\nabla_x d_X(\|x\|^2) d_Z(\|\Delta^{-\frac{1}{2}}(y-x)\|^2)$, one finally finds that the MAP is solution in x of the nonlinear equation

$$x = \left(I + \frac{d'_X\left(\|\Delta^{-\frac{1}{2}}x\|^2\right)}{d_X\left(\|\Delta^{-\frac{1}{2}}x\|^2\right)} \frac{d_Z\left(\|y - x\|^2\right)}{d'_Z\left(\|y - x\|^2\right)} \Delta^{-1} \right)^{-1} y \tag{6}$$

This expression looks very similar to the Wiener filter, except that the matrix factor now depends on the solution x itself.

In the particular case where $\Delta = \sigma^2 I$, the rotated vector $C_\theta^t \widehat{X}_{\text{map}}(C_\theta y)$ is solution of (6) for any rotation matrix C_θ. Thus, $\widehat{X}_{\text{map}}(y)$ can be sought under the form

$$\widehat{X}_{\text{map}}(y) = \overline{X}_{\text{map}}\left(\|y\|^2\right) y \tag{7}$$

where $\overline{X}_{\text{map}}$ is solution of

$$\overline{x} = \frac{\sigma^2}{\sigma^2 + \frac{d'_X\left(\frac{\overline{x}^2\|y\|^2}{\sigma^2}\right)}{d_X\left(\frac{\overline{x}^2\|y\|^2}{\sigma^2}\right)} \frac{d_Z\left((1-\overline{x})^2\|y\|^2\right)}{d'_Z\left((1-\overline{x})^2\|y\|^2\right)}} \tag{8}$$

[2] $\sim$ means "distributed according to"

Note that in this context, the MMSE and Wiener estimators (3)-(4) can also be written under the form $\widehat{X}_{\text{mmse,w}} = \overline{X}_{\text{mmse,w}}.y$ [4]: in this situation, we will call $\overline{X}$ "estimator magnitude". One can easily see that $\widehat{X}_{\text{w}} \in [0\,;\,1]$, and that in the strict sense GSM, the MMSE estimator magnitude is in $[0\,;\,1]$ as well.

In the general case, solving (8) is not a trivial task. Nevertheless, under additional assumptions, equation (8) can be numerically solved.

1. If both d_X and d_Z are continuously differentiable and monotonic (necessarily decreasing), then equation (8) has at least one solution, which, from (8), is necessarily in $[0\,;\,1]$. Indeed, writing $\overline{x} = G(\overline{x})$ where $G(\overline{x})$ is the right hand-side of (8), it is immediate that $G(0) \geq 0$ and $G(1) \leq 1$, hence, from the continuity assumption, there exists at least one $\overline{x}$ so that $\overline{x} = G(\overline{x})$.
2. If, additionally, both d_X and d_Z admit a second derivative and are log-convex (implying that $d_{X,Z}d''_{X,Z} - d'^2_{X,Z} \geq 0$), the fixed-point method $\overline{x}_{k+1} = G(\overline{x}_k)$ with $\overline{x}_0 \in [0\,;\,1]$ always converges to a solution of (8), which is a maximum of the posterior distribution (except if it is initialized at a minimum). Deriving the expression of G versus $\overline{x}$ and using both the fact that $\overline{x} \in [0\,;\,1]$, the negativity of $d'_{X,Z}$, and the log-convexity assumption shows that G is an increasing function. Moreover, using additional algebra, one can show that at the maxima of the posterior density versus $\overline{x}$, $G'(\overline{x}) \leq 1$: the maxima are the stable fixed points of G and the series $\overline{x}_{k+1} = G(\overline{x}_k)$ is necessarily monotonic and bounded by the stable maxima.

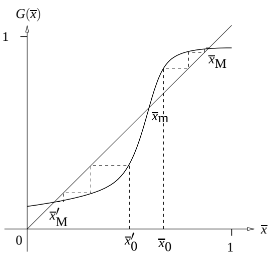

Figure 1: Illustration of eq. (8), when both d_X and d_Z are decreasing and log-convex. Points $\overline{x}'_\text{M}$ and $\overline{x}_\text{M}$ correspond to two local maxima of the MAP equation, while the second point, $\overline{x}_\text{m}$ corresponds to a minimum. The dotted line illustrates the behavior of the possible sequence $\overline{x}_{k+1} = G(\overline{x}_k)$.

In the case where both X and Z are GSM in the strict sense, it is clear from (1) that $d'_{X,Z} \leq 0$. Moreover, by applying the Cauchy-Schwartz inequality to $d'_{X,Z}(r)$ written under the form (1) yields to the log-convexity of $d_{X,Z}$: in the GSM context, both previous conditions are thus satisfied.

Finally, for $\Delta \not\propto I$, one can still write $\widehat{X}_{\text{map}} = \overline{X}_{\text{map}} \odot y$, where $\odot$ is the Hadamard (or entrywise) product. Under the monotony assumption, one has $\overline{X}_{\text{map}} \in [0\,;\,1]^d$ and under the log-convexity assumption, the fixed point method converges to a local maximum. The proof is similar, using the partial derivative in the components of $\overline{x}$.

Note that since there may exist several maxima, this fixed-point method does not guarantee that the sequence converges to a global maximum. Hence, it can be preferable to consider more efficient optimization methods (*e.g.* such as the simulated annealing method).

The estimation problem addressed in [3], where a radially exponential or a Laplace random vector X corrupted by a Gaussian noise Z is to be estimated, is a special case of the previous study. Hence, the search for the MAP proposed in [2, 3] is a particular case of the previous result.

The next subsection aims at illustrating both the behavior of the MAP estimator and its performances versus the signal to noise ratio. The advantage and some drawbacks of the MAP approach will be discussed, and the MAP will be compared with the MMSE and Wiener estimators.

3.3 Some illustrative examples

In the following, we give three illustrations where $\Delta = \sigma^2 I$ and a fourth one where $\Delta \not\propto I$. The performances in terms MSE of both estimators $\widehat{X}_{\text{map}}$, $\widehat{X}_{\text{mmse}}$ and $\widehat{X}_w$ will be shown versus the signal-to-noise ratio $SNR = \frac{E[\|X\|^2]}{E[\|Z\|^2]} = \frac{\text{Trace}(\Delta)}{d} = \sigma^2$.

The first illustration deals with a vector X following an exponential power distribution [17, 18] $X \sim p_X(x) \propto e^{-(\gamma x^t x)^{\frac{p}{2}}}$ where $\gamma = \frac{\Gamma\left(\frac{d+2}{p}\right)}{d\,\Gamma\left(\frac{d}{p}\right)}$, corrupted by a Gaussian random vector. A d-variate exponential power random vector is a GSM, however the mixing function is expressed via the pdf $\mathcal{G}_\alpha$ of an α-stable variable of stability index α and totally skewed to the right (with skew parameter $\beta = 1$) [4], $f_a(a) \propto a^{d-3}\,\mathcal{G}_{\frac{p}{2}}\left(a^{-2}(2\gamma)^{-1}\cos^{-\frac{2}{p}}(p\pi/4)\right)$. Density $\mathcal{G}_\alpha$ has no analytical expression [4] except in the particular case of the Lévy distribution given by $\alpha = \frac{1}{2}$. This case corresponds to $p = 1$, and the distribution is known as the radially exponential distribution [2].

Figure 2(left) describes the behavior of $\overline{X}_{\text{map}}$ versus $\|y\|$, for $SNR = -5$ dB, in dimension $d = 5$, and for $p = .7$ (top) and $p = 1$ (bottom). The solid line corresponds to the MAP, while the dashed line represents the MMSE and the dotted line depicts the Wiener estimator. The MMSE is numerically computed via (3), and the code available at the following address [19] is used to numerically compute f_a when $p \neq 1$ [4]. Furthermore, in the case $p = 1$, one can easily show that the MAP has an explicit form [1, 2, 3]: $\overline{X}_{\text{map}} = \left(1 - \frac{\sqrt{d+1}}{\sigma\|y\|}\right)\mathbb{1}_{[0\,;\,1]}\left(1 - \frac{\sqrt{d+1}}{\sigma\|y\|}\right)$ where $\mathbb{1}_A$ denotes the indicator function of set A. For $p = .7$, the MAP was determined by a exhaustive search of the maximum of $d_X(\overline{x}^2\|y\|^2/\sigma^2)\,d_Z((1-\overline{x})^2\|y\|^2)$ for $\overline{x} \in [0\,;\,1]$.

In this figures, one can observe that the Wiener estimator differs from the MAP and from the MMSE, while these two last estimators behave similarly, and tend to be identical when $\|y\|$ is large. One can also observe that for small $\|y\|$ the MAP's magnitude is null, while it tends to unity for large $\|y\|$. This can be understood considering the tails of the distributions of X and Z. Indeed, the tails of p_X are heavier than those of p_Z: large (resp. small) values of the observation Y are more probably due to X (resp. Z). Finally, one can observe a discontinuity for the case $p = .7$. Whatever p, for small $\|y\|$, the posterior density, function of $\overline{x}$, is decreasing and the MAP is null. However, one can observe that 0 is a fixed point of G, although the derivative of the posterior density is not zero. Indeed, d'_X/d_X is infinite for $\overline{x} = 0$ and the division by this quantity to obtain G allows to recover this solution here. As $\|y\|$ increases, one observes the "emergence" of a local maximum. This local maximum becomes global for large $\|y\|$. For small p the location of the local maximum is "far" from zero, explaining the observed discontinuity; this is not the case for "large" p. This is illustrated in figure 2(middle).

The right figures 2 depict the behavior of the MSE normalized by the variance of vector X, $\frac{MSE}{d\sigma^2}$, of the MAP estimator, as a function of the SNR, compared to that of the Wiener estimator and to the minimal MSE. One can observe that although the Wiener estimator highly differs from the MMSE, surprisingly their performances are similar. Conversely, the MAP's performances differ from that of the MMSE. This is particularly true for small SNR, and for "small" p. An explanation comes from the fact that for large SNR, X "dominates" the observation: thus both the MMSE and the MAP tend to be unity even for small values of $\|y\|$.

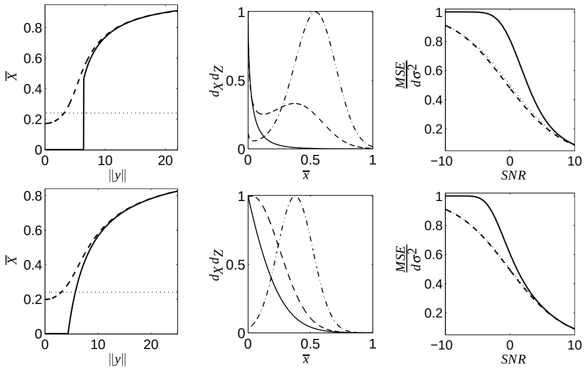

Figure 2: Left: Estimators' magnitudes $\overline{X}_{\text{map}}$ (solid line), $\overline{X}_{\text{mmse}}$ (dashed line) and $\overline{X}_{\text{w}}$ dotted line) as a function of $\|y\|$; Middle: Shape of the posterior density (arbitrarily rescaled) versus $\bar{x}$ for some values of $\|y\|$. Right: Mean Square Error normalized by the variance of X, $\frac{MSE}{d\sigma^2}$ versus the SNR σ^2; X is exponentially power distributed corrupted by Gaussian noise, the SNR is -5dB, the dimension is $d = 5$ and $p = .7$ (upper curves) or $p = 1$ (lower curves). $\|y\| = 3$ (solid line) 6 (dashed line) and 7 (dashed-dotted line) for $p = .7$ and $\|y\| = 3, 6, 7$ for $p = 1$.

The second case concerns a vector X with $d-$variate Student-t distribution embedded in Student-t noise Z. In this situation, the density generators write $d_{X,Z}(r) \propto \left(1 + \frac{r}{k_{x,z}}\right)^{-\frac{d+\nu_{x,z}}{2}}$ where $\nu_{x,z}$ are the degrees of freedom and $k_{x,z} = \nu_{x,z} - 2$ if $\nu_{x,z} > 2$, and $k_{x,z} = 1$ otherwise[3]. X and Z are GSM and their mixing variable are square root of Gamma distributed [4].

Although this case can be treated in the general setting previously presented, the MAP simplifies again. Indeed, one can show that the estimator magnitude is solution of the cubic equation $(\nu_x + \nu_z + 2d)\,\|y\|^2 \bar{x}^3 - (2\nu_x + \nu_z + 3d)\,\|y\|^2 \bar{x}^2 + (\sigma^2 k_x(\nu_z + d) + (k_z + \|y\|^2)(\nu_x + d))\bar{x} - \sigma^2 k_x(\nu_z + d) = 0$. In this particular situation the MAP can hence be sought using Cardano's approach that finds the roots of a third order polynomial. Moreover, the case where X (resp. Z) is Gaussian can be recovered by dividing the cubic polynomial by ν_x (resp. ν_z) and letting ν_x and $k_x = \nu_x - 2$ (resp. ν_z and k_z) tend to infinity.

Figure 3 describes the behavior of the MAP magnitude versus $\|y\|$ (left) and its performances compared to the MMSE (right). The MAP is sought using the above-mentioned Cardano's approach, while the MMSE magnitude is implemented from the numerical integration of (3) and using polar coordinates that simplify the integrations over $\mathbb{R}_+^2$ to integrations over $[0\,;\pi/2]$. The same analysis as in the first example holds concerning the estimators' magnitudes. But in this example, the Wiener estimator gives performances similar to the MAP's ones. The losses of performance of these two estimators compared to the MMSE are large, especially for small SNR. The only advantage of the MAP here is its simplicity compared to the MMSE (in spite of the possible existence of local maxima).

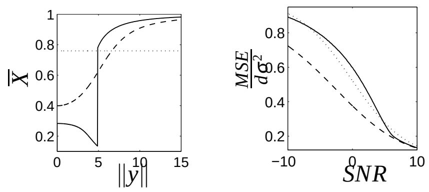

Figure 3: Same as figures 2, but for X and Z Student-t distributed with $\nu_X = 2.5$, $\nu_Z = 10$, $d = 5$ and for $SNR = +5$dB.

The third example aims to show that the previous results still hold even if X, although elliptically distributed, is not a (wide-sense) GSM. To this end, we consider the example of a Student-r random vector X, corrupted by a Gaussian noise. The density generator of X writes $d_X(r) \propto \left(1 - \frac{r}{\nu+2}\right)^{\frac{\nu-d}{2}} \mathbb{1}_{[0\,;\,\nu+2]}(r)$. In this example, d_X does not admit a Gaussian scale mixture form as mentioned in the introduction of the MAP estimation. Thus, since developing the MMSE estimator is difficult, even computationally, in this example, only the MAP and Wiener estimation will be illustrated. Moreover, in this situation again the MAP can be sought easily since it satisfies the cubic equation $\|y\|^2 \bar{x}^3 - \|y\|^2 \bar{x}^2 - \big((\nu+2)s^2 + \nu - d\big)\bar{x} + (\nu+2)s^2 = 0$, which can be solved via Cardano's approach.

The results are exhibited in figure 4. The same interpretation of the estimator's magnitude behavior holds (but here, the noise pdf has heavier tails than that of the pdf of the vector to estimate). In this example, the performances of the MAP are highly better than that of the Wiener estimator. Moreover an important advantage of the MAP in this example is that the MMSE cannot be implemented in the general setting of [4].

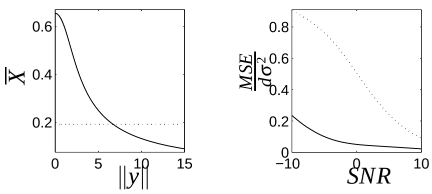

Figure 4: Same as figure 2, but for the non wide-sense GSM above-presented. The parameters are $\nu = 6$, $d = 5$ and $SNR = -5$dB.

The last example concerns the estimation of a generalized Laplacian vector of density generator $d_X(r) \propto r^{\frac{\nu-d}{4}} K_{\frac{\nu-d}{2}}(\sqrt{\nu r})$ [12], embedded in Gaussian noise, but in the context where $\Delta = \frac{\sigma^2}{3}\left(\begin{smallmatrix}1 & 0\\ 0 & 5\end{smallmatrix}\right)$. Figures 5 depict the behavior of each components of $\overline{X}_{\text{map}}$ versus y. These figures clearly show that the estimators do not remain functions of a quadratic form since the constant $\overline{X}$ levels seem to be ellipsoidal, but with eccentricities that depend on the level.

[3] in this case X has infinite covariance; for $\nu = 1$ one finds a Cauchy random vectors

Mainly, the same behavior with the same interpretation as for the previous examples hold.

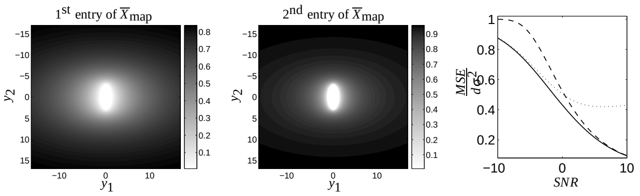

Figure 5: MAP estimator magnitudes (first and second component) as a function of y, in the context of a 2-variate random Laplacian vector ($\nu = 3$), corrupted by Gaussian noise. $\sigma = 1$, so that the SNR is 0dB.

4. DISCUSSION

In this paper we have revisited the estimation of an elliptically distributed random vector corrupted by an elliptically distributed random noise. This study extends: (i) the approach of [1], that deals only with scalar GSM (and with a Gaussian corrupting noise); (ii) the study of [2, 3] where the dimension is not restricted to 1, but where the noise is still considered Gaussian and the vectors to estimate restricted to Laplacian or radially exponential distributed. Here, the vector to be estimated and the noise are both assumed elliptically distributed, without restricting their statistics, and the MAP approach is developed. As for the MMSE, previously studied in [4], the MAP simplifies when both the vector to estimate and the noise are spherically invariant (or with proportional covariance matrices). The MAP is solution of a nonlinear equation, that has not always an analytical solution. But under the assumption that the density generators are decreasing and log-convex, a fixed-point approach allows an easy numerical determination. A drawback of this approach is that it can converge to a local maximum for the posterior density instead of the global one. Thus, more refined approaches can be envisaged, but such studies are left as perspectives.

Obviously, the MMSE estimator is the approach to consider in order to ensure minimum mean square error. Moreover, the integral form of this estimator can still exist even if the MSE does not exist (*e.g.* vector with infinite variance). The interest of the MMSE is that it can be systematically implemented when the mixing densities are known or can be numerically evaluated. However, the MMSE suffers from limitations. In situations where no mixing functions exists, such as in the Student-r case (bounded support pdf), the MMSE cannot be implemented in the general setting of [4] and its evaluation can become much more complicated. Instead, the use of the MAP allows to solve the estimation problem in a systematic way. Moreover, implementing the MAP does not need numerical integrations. Thus, the computational cost can be reduced (provided a fast algorithm is used to solve the nonlinear equation that gives the MAP).

In terms of performances, the various examples show that, apart the MMSE, there is no universal "best" estimator between the MAP and the Wiener estimator.

REFERENCES

[1] T. I. Alecu, S. Voloshynovskiy, and T. Pun, "The Gaussian transform of distributions: Definition, computation and application," *IEEE Trans. on Sig. Proc.*, vol. 54, pp. 2976–2985, Aug. 2006.

[2] F. Shi and I. W. Selesnick, "An elliptically contoured exponential mixture model for wavelet based image denoising," *Appl. Comput. Harmo. Anal.*, vol. 23, pp. 131–151, July 2007.

[3] I. W. Selesnick, "The estimation of Laplace random vectors in additive white Gaussian noise," *IEEE Trans. on Sig. Proc.*, vol. 56, pp. 3482–3496, Aug. 2008.

[4] S. Zozor and C. Vignat, "Revisiting the denoising problem in the context of elliptical distributions," in *EUSIPCO*, (Lausanne, Switzerland), 25-29 Aug. 2008.

[5] K.-C. Chu, "Estimation and decision for linear systems with elliptical random processes," *IEEE Trans. on Auto. Control*, vol. 18, pp. 499–505, Oct. 1973.

[6] K. Yao, "A representation theorem and its applications to spherically-invariant random processes," *IEEE Trans. on Info. Theory*, vol. 19, pp. 600–608, Sept. 1973.

[7] K. T. Fang, S. Kotz, and K. W. Ng, *Symmetric multivariate and related distributions*. Monographs on stat. and proba. 36, London: Chapman & Hall, 1990.

[8] M. Rangaswamy, D. Weiner, and A. Öztürk, "Non-Gaussian random vector identification using spherically invariant random processes," *IEEE Trans. on Aero. and Elect. Syst.*, vol. 26, pp. 111–124, Jan. 1993.

[9] J. Portilla, V. Strela, J. Wainwright, and R. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. on Image Proc.*, vol. 12, pp. 1338–1351, Nov. 2003.

[10] K. Todros and J. Tabrikian, "Blind separation of independent sources using Gaussian mixture model," *IEEE Trans. on Sig. Proc.*, vol. 55, pp. 3645–3658, July 2007.

[11] Y. Chitour and F. Pascal, "Exact maximum likelihood estimates for SIRV covariance matrix: Existence and algorithm analysis," *IEEE Trans. on Sig. Proc.*, vol. 56, pp. 4563–4573, Oct. 2008.

[12] R. Lord, "The use of the Hankel transform in statistics I. General theory and examples," *Biometrika*, vol. 41, pp. 44–55, June 1954.

[13] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2. New-York: Wiley & Sons, 1971.

[14] D. V. Widder, *The Laplace Transform*. Princeton Mathematical Series, Princeton University Press, 1946.

[15] S. M. Kay, *Fundamentals for Statistical Signal Processing: Estimation Theory*. vol. 1, Upper Saddle River, NJ: Prentice Hall, 1993.

[16] G. F. Carrier, M. Krook, and C. E. Pearson, *Function of a Complex Variable: Theory and Technique*. Philadelphia: SIAM, 2005.

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

[18] T. J. Kozubowski, "On the vertical density of the multivariate exponential power distribution," *Statistics*, vol. 36, no. 3, pp. 219–221, 2002.

[19] J. P. Nolan, "Numerical calculation of stable densities and distribution functions," *Stochastic Models*, vol. 13, no. 4, pp. 759–774, 1997. http://academic2.american.edu/∼jpnolan.