

Estimation MAP pour des variables elliptiquement distribuées

Steeve ZOZOR¹, Christophe VIGNAT²

¹GIPSA-Lab, Département Image est Signal
961 rue de la Houille Blanche, B.P. 46, 38402 Saint Martin d'Hères Cedex, France

²Institut Gaspard Monge, Equipe SYSCOM
5 Boulevard Descartes, 77454 Marne-la-Vallée, France
steeve.zozor@gipsa-lab.inpg.fr, vignat@univ-mlv.fr

Résumé – Dans cet article nous revisitons le problème de l'estimation d'un vecteur aléatoire d -dimensionnel noyé dans du bruit, dans le contexte où les vecteurs sont elliptiquement distribués. L'approche proposée est basée sur l'estimation par maximum a posteriori (MAP). Ce travail étend des études récentes sur ce problème d'estimation, études restreintes au cas d'une variable de type mélange d'échelle Gaussien bruitée par une Gaussienne [1–3] : nous nous plaçons dans le cadre où le vecteur à estimer et le bruit sont elliptiquement distribués quelconques. Nous montrons que le MAP est solution d'une équation non linéaire et déterminons des conditions sous lesquelles une solution peut être calculée par un algorithme du point fixe.

Abstract – In this paper we revisit the estimation problem of a d -variate random vector corrupted by a noise in the elliptically distributed context. The proposed approach lies on the maximum a posteriori (MAP) estimation procedure. This work extends the recent works [1–3], which are restricted to Gaussian scale mixtures corrupted by a Gaussian noise. In our study, we assume that both the vector to estimate and the noise are elliptically distributed, possibly non-Gaussian. We then show that the MAP is solution of a nonlinear equation and we determine conditions to ensure that the solutions can be searched via a fixed point algorithm.

1 Introduction

Des études sur les vecteurs elliptiquement distribués paraissent régulièrement en traitement du signal, en témoignent les articles récents [1–5]. Un des intérêts de la classe des vecteurs elliptiquement distribués, ou plus précisément de la sous-classe des mélanges d'échelle Gaussien, est qu'elle étend naturellement la classe des vecteurs Gaussiens ; ainsi, de nombreux traitements effectués dans le cadre Gaussien peuvent être étendus plus ou moins facilement à cette classe de vecteurs.

L'étude de vecteurs elliptiquement distribués en traitement du signal est assez ancienne [6, 7], mais donne toujours lieu à de nombreuses applications [2, 4, 5, 8]. . . . Dans une précédente étude [9], nous avons revisité le problème d'estimation d'un mélange d'échelle Gaussien (au sens large), bruité par un vecteur non Gaussien, lui aussi mélange d'échelle Gaussien (au sens large). L'approche choisie était celle du critère de l'erreur quadratique moyenne minimale (EQMM), ce travail étendant au cas multivarié et au bruit non Gaussien les travaux d'Alecu *et al.* [1]. Toutefois, tout vecteur elliptiquement distribué n'est pas forcément mélange d'échelle Gaussien, et en dehors de ce cadre, l'approche proposée ne peut plus être mise en œuvre.

Dans cet article, nous proposons d'étudier ce problème d'estimation dans le cadre elliptique général, en adoptant une approche maximum a posteriori (MAP). Cette étude généralise à la fois [1] et les plus récentes contributions [2, 3].

2 Les distributions elliptiques en bref

Un vecteur aléatoire X de dimension d est dit elliptiquement distribué si sa densité de probabilité (ddp) p_X est fonction de

la forme quadratique positive $(x - \mu_X)^t R_X^{-1} (x - \mu_X)$:

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

où la fonction $d_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est appelée générateur de la densité de probabilité, le vecteur μ_X est un paramètre de position (moyenne si elle existe) et la matrice symétrique définie positive R_X est appelée matrice caractéristique [6, 7, 10]. Cette matrice est définie à un facteur scalaire près, mais en normalisant d_X de telle sorte que $\int_{\mathbb{R}_+} r^{d+1} d_X(r^2) dr = \Gamma(d/2 + 1)/\pi^{d/2}$ (à condition que l'intégrale converge), R_X coïncide avec la matrice de covariance de X . Si R_X est proportionnelle à la matrice identité I , X est dite sphériquement distribuée, ou sphériquement symétrique. Dans la suite, nous supposons sans perte de généralité que $\mu_X = 0$. Par transformation de Fourier de la ddp p_X , on montre également que la fonction caractéristique de X est fonction de la forme quadratique positive $x^t R_X x$ uniquement [7, 11, 12]. Un vecteur elliptiquement distribué peut être défini par sa fonction caractéristique, sans toutefois admettre une densité de probabilité [10, 13]. Dans cet article, nous nous restreindrons au cas de vecteurs admettant une ddp (et donc un générateur de densité).

Lorsque l'application $s \mapsto d_X(s)$ admet une transformée de Laplace inverse, la ddp p_X peut être exprimée comme une convolution multiplicative (ou de type Mellin) entre une densité de probabilité Gaussienne et une fonction f_a dite fonction de mélange [6]. Cette fonction de mélange f_a est d'intégrale unité. Si de plus d_X est complètement monotone ($(-1)^n d_X^{(n)} \geq 0$), f_a est non négative et peut donc être vue comme une ddp [7, 14, 15] ; on peut alors écrire $X \stackrel{d}{=} aN$ avec a variable aléatoire scalaire de ddp f_a , indépendante de N , vecteur Gaussien centré, de covariance R_X , l'égalité étant en distribution.

Le vecteur X est alors appelé mélange d'échelle Gaussien. Si de plus f_a ne dépend pas de la dimension, le mélange est dit consistant [11]. Si f_a existe mais n'est pas une ddp, on appellera X mélange d'échelle Gaussien au sens large. Cette dernière classe (incluant les mélanges Gaussiens au sens strict) est elle-même strictement incluse dans la classe des vecteurs elliptiquement distribués. En effet, un vecteur de loi Student- t , pour lequel $d_X(r) \propto (1-r/(\nu+2))^{(\nu-d)/2} \mathbb{1}_{[0;1]}(r/(\nu+2))$, avec $\nu \geq d-2$, n'est pas mélange d'échelle Gaussien, même au sens large, puisque la fonction d_X n'admet pas de prolongement analytique dans le plan complexe (étant nulle sur un domaine, elle devrait être nulle partout), et donc ne possède pas de transformée de Laplace inverse.

3 Estimation MAP

Dans cet article, nous proposons de revisiter le problème de l'estimation d'un vecteur d -dimensionnel X , elliptiquement distribué, noyé dans du bruit Z , également elliptiquement distribué et indépendant de X [1–3] à partir de l'observation

$$Y = X + Z$$

On supposera connues les ddp p_X de X et p_Z de Z , et donc les générateurs de densité respectifs d_X et d_Z ainsi que les matrices de covariance R_X et R_Z . Sans perte de généralité, on se restreint au cas où Z est de covariance identité I (par pré-blanchiment de Z). On supposera également sans perte de généralité que la matrice de covariance de X est diagonale Δ (par une rotation adéquate qui ne modifie pas la blancheur de Z ; plus de détails peuvent être trouvés dans [9]).

Nous avons précédemment montré que dans le cas où X et Z sont tous les deux mélanges d'échelle Gaussien au sens large, avec f_a et f_b pour fonctions de mélange respectives, l'estimateur $\hat{X}_{\text{eqm}}(Y)$ minimisant l'erreur quadratique moyenne (EQM) $E[\|\hat{X} - X\|^2]$ (i.e. l'espérance conditionnelle $\hat{X}_{\text{eqm}} = E[X|Y]$ [16]), peut être exprimé sous une forme simple faisant appel à $d+1$ intégrations sur \mathbb{R}_+^2 (au lieu d'intégrations sur \mathbb{R}^d). Cet estimateur peut être vu comme une double convolution multiplicative entre un estimateur de Wiener (pour lequel X et Z seraient Gaussiens de covariances $a^2\Delta$ et b^2I), les fonctions de mélange $f_a(a)$ et $f_b(b)$ et une ddp Gaussienne de covariance $a^2\Delta + b^2I$. Cette expression peut encore être simplifiée dans le cas où $\Delta \propto I$ (i.e. si X et Z peuvent être conjointement blanchis); bien entendu, dans le cas où X et Z sont Gaussiens, on retrouve l'estimateur de Wiener $\hat{X}_w = (\Delta + I)^{-1}\Delta y$. On trouvera plus de détails dans [9].

Toutefois, l'étude de [9] n'est plus valide lorsque X et/ou Z n'est pas mélange d'échelle Gaussien au sens large. D'autre part, les intégrations numériques nécessaires peuvent s'avérer coûteuses. Dans ces cas, une approche de type maximum a posteriori (MAP) est une bonne alternative. Rappelons que le MAP consiste à sélectionner la valeur X qui maximise la ddp a posteriori $p_{X|Y}(x, y)$, la loi de Bayes conduisant à $\hat{X}_{\text{map}}(y) = \arg \max_x p_X(x)p_Z(y-x)$. C'est précisément l'approche adoptée dans [1–3].

En exprimant les ddp de X et Z à l'aide de leurs générateurs respectifs d_X et d_Z , puis en annulant le gradient en x de la densité a posteriori, on aboutit facilement à l'estimateur MAP

comme solution de l'équation non linéaire

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

Remarquons que cette expression ressemble à l'estimateur de Wiener, à la différence que la matrice de régression dépend de l'observation y et de la solution x elle-même.

De même que pour l'estimateur EQMM, le MAP peut être simplifié dans le cas où $\Delta = \sigma^2 I$ (σ^2 est alors le rapport signal-sur-bruit (RSB)) : on voit facilement que pour toute rotation C_θ , $C_\theta^t \hat{X}_{\text{map}}(C_\theta y)$ est solution de (1) et donc que $\hat{X}_{\text{map}}(y)$ s'écrit

$$\hat{X}_{\text{map}}(y) = \bar{X}_{\text{map}}(\|y\|^2) y \quad (2)$$

avec \bar{X}_{map} solution \bar{x} de

$$\bar{x} = \frac{\sigma^2}{\sigma^2 + \frac{d'_X(\frac{\bar{x}^2\|y\|^2}{\sigma^2})}{d_X(\frac{\bar{x}^2\|y\|^2}{\sigma^2})} \frac{d_Z((1-\bar{x})^2\|y\|^2)}{d'_Z((1-\bar{x})^2\|y\|^2)}} \quad (3)$$

On appellera \bar{X} amplitude de l'estimateur. Cette amplitude contient à elle seule la non linéarité de l'estimateur.

La détermination du MAP (3) n'est pas aisée, mais on peut toutefois en proposer une approche récursive basée sur les principes suivants :

1. Si d_X et d_Z sont continûment dérivables et monotones (nécessairement décroissantes), l'équation (3) possède au moins une solution qui appartient forcément à l'intervalle $[0; 1]$. Ce domaine apparaît clairement au vu de (3). De plus, en réécrivant (3) sous la forme $\bar{x} = G(\bar{x})$, on a $G(0) \geq 0$ et $G(1) \leq 1$: il existe donc au moins un \bar{x} tel que $\bar{x} = G(\bar{x})$.
2. Si de plus d_X et d_Z sont deux fois dérivables et log-convexes ($d_{X,Z} d''_{X,Z} - d'^2_{X,Z} \geq 0$), la méthode du point fixe $\bar{x}_{k+1} = G(\bar{x}_k)$ initialisée en $\bar{x}_0 \in [0; 1]$ converge vers un maximum de la densité a posteriori. Ceci est démontré en dérivant deux fois la densité a posteriori; puis, utilisant le fait que la solution appartient à l'intervalle $[0; 1]$ et que $d'_{X,Z} \leq 0$ ainsi que la log-convexité. On montre que G est croissante, et que les maxima de la densité a posteriori sont les points fixes stables de G . Si a contrario d_X et d_Z sont log-concaves, G est décroissante et donc admet un point fixe unique. Toutefois, un algorithme de type point fixe ne convergera pas forcément vers le point fixe (on peut obtenir une suite oscillante).

Ceci est illustré par la figure 1.

Comme le générateur de densité d'un mélange d'échelle Gaussien est décroissant et log-convexe, la convergence de la méthode du point fixe est assurée dans le contexte de mélanges d'échelle Gaussiens. Enfin, l'approche du point fixe peut être naturellement étendue au cas $\Delta \not\propto I$; il suffit d'écrire l'estimateur sous la forme $\hat{X}_{\text{map}} = \bar{X}_{\text{map}} \odot y$, où \odot est le produit de Hadamard (ou terme à terme), et de raisonner composante à composante.

La résolution par la méthode du point fixe est celle employée dans [3], dans le cas où $\Delta \propto I$ et où Z est Gaussien. On vient de voir qu'elle s'étend à un cadre beaucoup plus large. Toutefois, elle souffre du fait que le maximum trouvé peut être local.

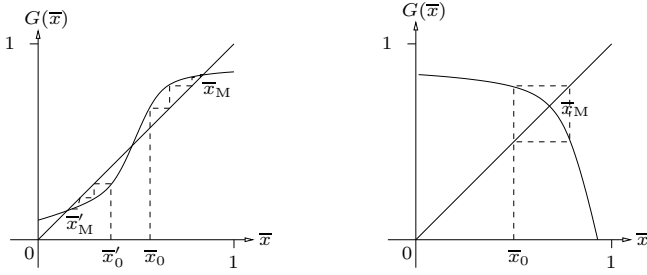


FIG. 1 – Illustration de la fonction G de (3), quand d_X et d_Z sont décroissantes et log-convexes (gauche) ou log-concaves (droite). Dans le cas log-convexe, une suite $\bar{x}_{k+1} = G(\bar{x}_k)$ sera monotone, comme illustré en pointillés (G est trait plein), et converge vers un point fixe stable. Dans le cas log-concave, le point fixe est unique, mais l’algorithme précédent ne converge pas forcément vers le point fixe.

Ceci apparaît de façon flagrante dans [3] : en effet la densité Laplacienne de X diverge à l’origine (singularité intégrable). Il s’en suit que le MAP devrait être nul, quelle que soit l’observation Y . Le point fixe obtenu dans [3], solution de (3), est précisément un maximum local (cependant, on pourrait se poser la question du sens même d’un estimateur identiquement nul). Une méthode d’optimisation plus fine, du type recuit simulé, peut alors être préférable pour déterminer le MAP. Ce type d’approche est laissée comme perspective.

Afin d’illustrer l’estimation MAP dans le contexte elliptiquement distribué, considérons comme premier exemple le cas d’un vecteur X Student- t , bruité par un vecteur Student- t

$d_{X,Z}(r) \propto \left(1 + \frac{r}{k_{x,z}}\right)^{-\frac{\nu_{x,z}+d}{2}}$, de degrés de liberté $\nu_{x,z} > 0$ et avec $k_{x,z} = \nu_{x,z} - 2$ si $\nu_{x,z} > 2$ et $k_{x,z} = 1$ sinon. Ce type de situation permet l’utilisation de l’approche EQMM sous sa forme générique de [9] ; toutefois, l’approche MAP permet de s’affranchir des intégrations numériques qui peuvent être coûteuses. De plus, en se plaçant dans le cadre $\Delta = \sigma^2 I$, la résolution de (3) se ramène à la recherche de la solution de l’équation polynomiale de degré 3 $(\nu_x + \nu_z + 2d)\|y\|^2 \bar{x}^3 - (2\nu_x + \nu_z + 3d)\|y\|^2 \bar{x}^2 + (\sigma^2 k_x(\nu_z + d) + (k_z + \|y\|^2)(\nu_x + d))\bar{x} - \sigma^2 k_x(\nu_z + d) = 0$ pour laquelle on peut appliquer la méthode de Cardan ; il suffira de retenir les solutions réelles, de les injecter dans la densité a posteriori et de retenir celle qui conduit au maximum. On peut donc ici s’affranchir de l’approche point-fixe. Notons que la densité a posteriori peut être décroissante, et donc la dérivée ne jamais s’annuler : il faut donc ajouter la solution $\bar{x} = 0$ aux racines trouvées avant injection dans la densité a posteriori. Les figures 2(a) et 2(b) représentent l’amplitude du MAP (3) et celle des estimateurs d’EQMM et de Wiener, en fonction de $\|y\|$, pour deux RSB différents. La figure 2(c) décrit l’EQM de ces estimateurs (EQM normalisée par la variance de X , ici $d\sigma^2$) en fonction du rapport signal-sur-bruit $RSB = \sigma^2$ en dB. La densité a posteriori, fonction de $\bar{x} = x/\|y\|$, est également tracée pour plusieurs valeurs de $\|y\|$ en figures 2(d) et 2(e), afin d’illustrer le problème de l’existence d’éventuels maxima locaux. On peut constater sur les amplitudes que le MAP et l’estimateur EQMM ont le même type de comportement : faibles pour $\|y\|$ faible et tendant vers 1 pour $\|y\|$ grand. Dans le cas de grandes valeurs de $\|y\|$, ceci s’explique par le fait que X est “dominant”

dans l’observation Y car les queues de la distribution p_X sont plus lourdes que celle de p_Z . D’autre part, la discontinuité du MAP qu’on aperçoit figure 2(a) est une conséquence de l’existence de maxima locaux. En effet, pour $\|y\|$ faible, la densité a posteriori ne fait apparaître qu’un maximum qui est global. A mesure que $\|y\|$ augmente, un maximum local émerge, “loin” du maximum global, et finit par devenir le maximum global. Ceci est illustré figure 2(d) pour un RSB de 5 dB. Pour un RSB plus fort, cet effet disparaît car le second maximum local qui “tend à apparaître” est proche du premier, et que ce soit avant ou lors de la “transition” il n’émerge pas : il y a toujours un seul maximum. C’est seulement à $\|y\|$ plus fort qu’on retrouve la situation à deux maxima, mais bien après “transition”. Ceci est illustré figure 2(e) pour un RSB de 6.2 dB. A noter enfin que l’estimateur de Wiener est très différent des deux autres ; toutefois, en termes d’EQM 2(c), sur cet exemple précis, ce dernier et le MAP ont les mêmes performances. Mettre en œuvre une approche MAP n’est donc pas toujours la meilleur alternative à l’estimateur EQMM.

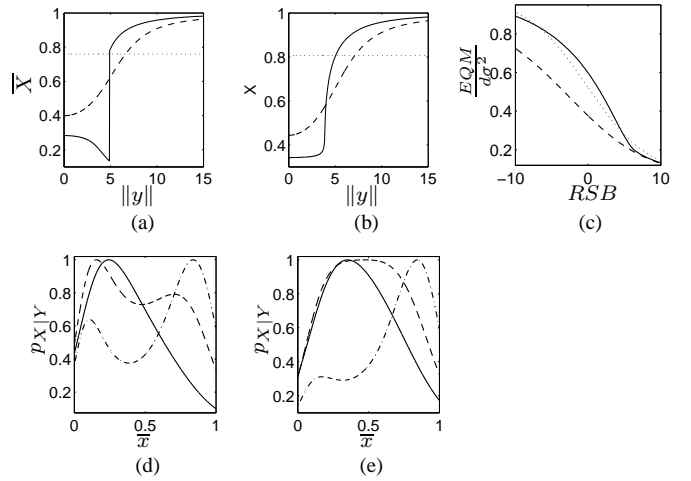


FIG. 2 – Cas où X et Z sont Student- t ($\nu_x = 2.5$ et $\nu_z = 10$), en dimension $d = 5$. (a) et (b) : amplitude des estimateurs MAP (trait plein), d’EQMM (tirets) et de Wiener (pointillés) en fonction de $\|y\|$ et pour un RSB de 5dB et 6.2dB respectivement. (b) : EQM normalisée des estimateurs en fonction du RSB. (c) et (d) : densité a posteriori (amplitude arbitraire) pour un RSB de 5dB et 6.2dB respectivement, et dans les cas $\|y\| = 3$ (trait plein), $\|y\| = 4.5$ ou 3.9 (tirets (c) et (d) respectivement) et $\|y\| = 5.5$ (pointillés).

Le second exemple, figure 3, concerne le cas où X est Student- r ($\nu_x = 6$) noyé dans du bruit Gaussien. Cet exemple est typique du cas où l’approche EQMM de [9] ne peut plus être mise en œuvre. Ici encore le MAP conduit à résoudre une équation polynomiale de degré 3, $\|y\|^2 \bar{x}^3 - \|y\|^2 \bar{x}^2 - (\sigma^2(\nu_x + 2) + \nu_x - d)\bar{x} + \sigma^2(\nu_x + 2) = 0$ (toute configuration où X et Z sont tels que $d'_{X,Z}/d_{X,Z}$ sont polynomiaux – Student, Gaussien... – conduira à chercher les racines d’un polynôme). L’interprétation de l’amplitude du MAP (figure 3(a)) est similaire à la précédente, avec cette fois Z ayant des queues de distribution plus lourdes que X . En terme d’EQM (figure 3(b)), contrairement au cas précédent, le MAP est bien meilleur que l’estimateur de Wiener et donc montre son intérêt. La figure

3(c) représente $G(\bar{x}) = \frac{\sigma^2}{\sigma^2 + \frac{d'_X \left(\frac{\bar{x}^2 \|y\|^2}{\sigma^2} \right)}{d_X \left(\frac{\bar{x}^2 \|y\|^2}{\sigma^2} \right)}} \frac{d'_Z \left((1-\bar{x})^2 \|y\|^2 \right)}{d_Z \left((1-\bar{x})^2 \|y\|^2 \right)}$ dont le

MAP est point fixe (3). Dans le cas Student-r considéré, d_X est monotone mais log-concave tandis que d_Z est aussi monotone mais log-convexe. Cette figure illustre que dans ce cadre, G n'est plus forcément croissante et qu'une approche de type point fixe ne convergera pas forcément vers le point fixe.

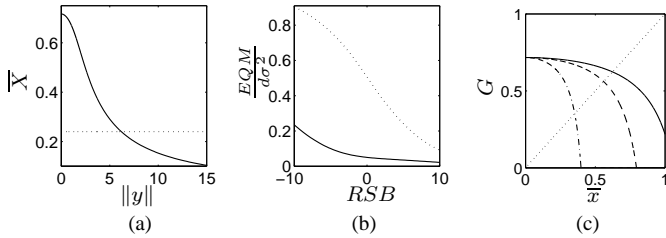


FIG. 3 – Cas où X est Student-r ($\nu = 6$) noyé dans du bruit Z Gaussien, en dimension $d = 5$ et pour un RSB de -5 dB. (a) : amplitude des estimateurs MAP (trait plein), EQMM (tirets) et de Wiener (pointillés) en fonction de $\|y\|$. (b) : EQM normalisée des estimateurs en fonction du RSB. (c) : Fonction $G(\bar{x})$ dont le MAP est point fixe (3), pour $\|y\| = 1.5$ (trait plein), $\|y\| = 2$ (tirets) et $\|y\| = 4$ (trait mixte). Les pointillés représentent la bissectrice.

4 Conclusion

Dans cet article, nous avons revisité l'estimation d'un vecteur elliptiquement distribué par un bruit également elliptiquement distribué, mais pas forcément Gaussien. Ce travail étend les récentes contributions du cadre scalaire de [1], et du cadre d'un signal Laplacien d -varié de [3], le bruit étant Gaussien dans ces deux précédentes études.

Nous avons précédemment montré que, dans le cadre de mélanges Gaussiens au sens large, une approche EQMM pouvait être mise en œuvre numériquement à moindre coût. Dans le cas contraire (Student-r par exemple), ou pour s'affranchir d'intégrations numériques, l'approche MAP est une bonne alternative. Elle ne fait appel qu'à des fonctions scalaires, mais le MAP est solution d'une équation non linéaire par forcément soluble analytiquement. Sous certaines conditions (monotonie et log-convexité), il est possible de rechercher la solution à moindre coût par un algorithme de type point-fixe. Toutefois, une telle approche ne résout pas le problème d'existence possible de maxima locaux. Ainsi, des méthodes plus élaborées peuvent être envisagées (recuit simulé, essais particuliers, ...). De telles approches sont laissées en perspectives.

On notera enfin que le MAP ressemble à l'estimateur de Wiener, la matrice de régression dépendant en particulier du MAP lui-même. Des approximations du MAP peuvent alors être envisagées. En fonction de la configuration du problème, on peut approcher cette matrice par un développement de Taylor à l'ordre 0 en $\bar{x} = 0$ (faible RSB, queues de distribution du bruit plus lourdes que celles du signal...) ou autour de $\bar{x} = 1$ (cas opposé). Cette approche permettrait de s'affranchir de la mise en œuvre d'un algorithme d'optimisation, potentiellement coûteux en temps de calcul. De telles approximations et l'analyse des performances qui en découlent sont également laissées

en perspective.

Références

- [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] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] S. Zozor and C. Vignat, "Revisiting the denoising problem in the context of elliptical distributions," in *EU-SIPCO*, (Lausanne, Switzerland), 25-29 Aug. 2008.
- [10] K. T. Fang, S. Kotz, and K. W. Ng, *Symmetric multivariate and related distributions*. Monographs on stat. and proba. 36, London : Chapman & Hall, 1990.
- [11] Y. Kano, "Consistency property of elliptic probability density functions," *Journal of Multivariate Analysis*, vol. 51, pp. 139–147, October 1994.
- [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] S. Cambanis, S. Huang, and G. Simons, "On the theory of elliptically contoured distributions," *Journal of Multivariate Analysis*, vol. 11, pp. 368–385, September 1981.
- [14] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2. New-York : John Wiley & Sons, Inc, 1971.
- [15] D. V. Widder, *The Laplace Transform*. Princeton Mathematical Series, Princeton University Press, 1946.
- [16] S. M. Kay, *Fundamentals for Statistical Signal Processing : Estimation Theory*. vol. 1, Upper Saddle River, NJ : Prentice Hall, 1993.