HAL
open science

Borne de Cramér-Rao intrinsèque pour la matrice de covariance des distributions elliptiques complexes

Arnaud Breloy, Alexandre Renaux, Guillaume Ginolhac, Florent Bouchard

To cite this version:

Arnaud Breloy, Alexandre Renaux, Guillaume Ginolhac, Florent Bouchard. Borne de Cramér-Rao intrinsèque pour la matrice de covariance des distributions elliptiques complexes. GRETSI 2017 - XXVIème Colloque francophone de traitement du signal et des images, Sep 2017, Juan-Les-Pins, France. hal-01586939

HAL Id: hal-01586939

https://centralesupelec.hal.science/hal-01586939v1

Submitted on 13 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Borne de Cramér-Rao intrinsèque pour la matrice de covariance des distributions elliptiques complexes

Arnaud BRELOY[1], Alexandre RENAUX[2], Guillaume GINOLHAC[3], Florent BOUCHARD[4]

[1]LEME, Université Paris Nanterre, France

[2]L2S, Université Paris-Sud, France

[3]LISTIC - Université Savoie Mont-Blanc, France

[4]GIPSA-lab - Grenoble-INP, France

arnaud.breloy@u-paris10.fr, alexandre.renaux@u-psud.fr
guillaume.ginolhac@univ-smb.fr, florent.bouchard@gipsa-lab.grenoble-inp.fr

Résumé – Dans cette communication, nous proposons une expression de la Borne de Cramér-Rao (BCR) intrinsèque concernant la matrice de covariance pour des données issues d'une distribution elliptique complexe générale. Par rapport aux travaux récents sur ce thème, nous nous appuyons sur la propriété géométrique des matrices de covariance d'appartenir à une variété Riemannienne. La dérivation s'appuie sur le calcul d'une nouvelle métrique de Fisher associée aux distributions elliptiques sur cette variété. Ensuite, nous pouvons en déduire une nouvelle BCR ainsi qu'une nouvelle distance entre deux matrices permettant de définir une Erreur Quadratique Moyenne (EQM) liée à une distance naturelle sur la variété Riemannienne. En simulation, nous étudions les performances de différents estimateurs sur des données issues de la t-distribution.

Abstract – In this paper, we propose a new Cramér-Rao Bound of the covariance matrix for data drawn from a complex elliptical distribution. Compared to other references of these last years, we use the geometrical property that the covariance matrices lie in a Riemann manifold. The derivation is based on the computation of a new metric associated to elliptical distributions on this manifold. Next, a new CRB as well as a new distance between two matrices allowing to define a Mean Square Error (MSE) linked to the natural distance on the Riemann manifold are deduced. In a simulation section, different estimators of covariance matrix are studied for data drawn from a t-distribution.

1 Introduction

La matrice de covariance est un élément important dans la plupart des algorithmes de traitement de signal comme la détection, l'estimation par maximum de vraisemblance, la classification, ... Dans la plupart des cas, elle est inconnue et doit donc être estimée à partir d'un ensemble de données. Dans le cas Gaussien, l'estimateur du maximum de vraisemblance est bien connu et s'appelle la *Sample Covariance Matrix* (SCM). Afin de traiter le cas non Gaussien, il est maintenant usuel d'utiliser les distributions elliptiques complexes pour modéliser ce type de données [1] [1]. Dans ce cadre, il est possible de construire des estimateurs du maximum de vraisemblance (MLE) adaptés aux paramètres de la loi. Néanmoins, il existe d'autres d'estimateurs qui ne sont pas issues de la théorie du maximum de vraisemblance et qui sont connus pour avoir de bonnes performances en particulier en terme robustesse. Les plus connus sont l'estimateur de Huber [2] et celui de Tyler [3, 4] (même si cet estimateur est aussi l'estimateur du maximum de vraisemblance d'une sous famille des distributions elliptiques, les Spherically Invariant Random Vectors). Ces estimateurs sont aussi très intéressants d'un point de vue pratique car ils n'utilisent pas certains paramètres de la distribution elliptique qui sont souvent inconnus en pratique.

Dans tout problème d'estimation, il est intéressant de comparer les performances des estimateurs à une borne de performance ultime comme par exemple la borne de Cramér-Rao (BCR). Pour les matrices de covariances, on peut trouver un résultat pour la BCR en Gaussien [5]. Cette formule est souvent appelée formule de Slepian-Bang. Pour des données non Gaussiennes, les travaux sont assez récents et ont considéré tout d'abord des sous familles des distributions elliptiques comme la K-distribution [6, 7] ou alors le cas de la matrice de covariance seul et à moyenne nulle pour les CES [8]. La référence [9] a permis de généraliser ces résultats au cas des distributions elliptiques à moyenne et matrice de covariance paramétrées.

Néanmoins, on peut constater que tous les travaux susmentionnés ne considèrent pas que les matrices de covariance sont des matrices symétriques hermitiennes et donc appartiennent à une variété Riemannienne. Or, le papier [10] (on peut aussi s'intéresser à une extension de ce travail dans [11]) a montré dans le cas de données complexes Gaussiennes qu'en utilisant

1. Par abus de langage, nous utiliserons le terme de matrice de covariance alors qu'en fait les distributions elliptiques dépendent d'une matrice de dispersion

la distance naturelle sur la variété Riemannienne, on peut trouver une nouvelle BCR qui permet de mieux mesurer le comportement des estimateurs. En particulier, elle permet de montrer que la SCM souffre de mauvaises performances avec peu de données ce qui n'est pas mis en avant lorsqu'on s'intéresse à la BCR classique. Dans ce papier, nous proposons donc d'étendre les travaux de [10] au cas des distributions elliptiques complexes. En simulation, nous montrerons sur une t-distribution les performances du MLE et de l'estimateur de Tyler.

Notations : Une lettre en italique est une quantité scalaire, les caractères gras en minuscule (majuscule) sont des vecteurs (matrices) et $(.)^H$ correspond à l'opérateur transposé conjugué. $\mathbf{I}_N$ est la matrice identité de taille $N \times N$. $Tr(\mathbf{R})$ représente l'opérateur de trace de la matrice $\mathbf{R}$ et $\mathbb{E}$ est l'opérateur d'espérance.

2 Distribution Elliptique Complexe

Dans le cas à moyenne nulle, une donnée $\mathbf{z} \in \mathbb{C}^M$ à distribution elliptique complexe s'écrit de la manière suivante [1] :

$$\mathbf{z} = \sqrt{\mathcal{Q}}\mathbf{R}^{1/2}\mathbf{u} \tag{1}$$

avec

$$\mathbf{u} \sim \mathcal{U}\left(\mathbb{C}S^M\right) \tag{2}$$

et la factorisation

$$\mathbf{R}^{1/2}\mathbf{R}^{H/2} = \mathbf{R} \tag{3}$$

Dans ce cadre, la densité de probabilité est de la forme

$$f\left(\mathbf{z}|\mathbf{R}, g\right) \propto |\mathbf{R}|^{-1} g\left(Tr\left\{\mathbf{R}^{-1}\mathbf{Z}\right\}\right) \tag{4}$$

où g est appelé la fonction génératrice (l'exponentielle, par exemple, pour le cas Gaussien) et $\mathbf{Z} = \mathbf{z}\mathbf{z}^H$.

En ignorant les termes constants, la log-vraisemblance de l'ensemble des données $\{\mathbf{z}_k\}_{k=1,\ldots,K}$ (contracté à $\{\mathbf{z}_k\}$) est

$$l\left(\{\mathbf{z}_k\}|\mathbf{R}, g\right) = -K\log|\mathbf{R}| + \sum_{k=1}^{K}\log\left(g\left(Tr\left\{\mathbf{R}^{-1}\mathbf{Z}_k\right\}\right)\right) \tag{5}$$

où $\mathbf{Z}_k = \mathbf{z}_k\mathbf{z}_k^H$. On utilisera par la suite la fonction suivante $\phi(t)$:

$$\phi(t) = \frac{g'(t)}{g(t)} \tag{6}$$

3 Borne de Cramer Rao Intrinsèque

Le calcul de la BCR intrinsèque se déroule en plusieurs étapes [10] :

1. Calcul de la métrique d'information de Fisher
2. Calcul de la Matrice d'Information de Fisher
3. Calcul de la distance associé à la nouvelle métrique d'information de Fisher et inégalité entre BCR intrinsèque et erreur quadratique moyenne dans la distance naturelle

3.1 Calcul de la Métrique d'Information de Fisher

Soit $\partial\mathbf{R}$ la différentielle de la matrice de covariance de $\mathbf{R}$. La métrique d'information de Fisher pour les distributions elliptiques complexes est donnée dans la proposition suivante.

Proposition 3.1. Métrique d'Information de Fisher
Pour un ensemble de données $\{\mathbf{z}_k\}$ distribué selon une loi elliptique complexe de matrice de covariance $\mathbf{R}$ et de générateur g, la métrique d'information de Fisher sur la variété Riemannienne est donnée par :

$$\begin{aligned} g_{fim}\left(\partial\mathbf{R}, \partial\mathbf{R}\right) &= K\left(\alpha Tr\left\{\left(\partial\mathbf{R}\mathbf{R}^{-1}\right)^2\right\}\right. \\ &\quad \left.+\beta Tr^2\left\{\partial\mathbf{R}\mathbf{R}^{-1}\right\}\right) \end{aligned} \tag{7}$$

avec

$$\begin{cases} \alpha &= -\left(1 + \dfrac{2}{M}\mathbb{E}\left[\mathcal{Q}_k\phi\left(\mathcal{Q}_k\right)\right]\right) \\ \beta &= -\dfrac{2}{M(M+1)}\mathbb{E}\left[\mathcal{Q}_k^2\phi'\left(\mathcal{Q}_k\right)\right] \end{cases} \tag{8}$$

Démonstration. Par manque de place, nous ne présentons que les étapes principales de la démonstration. Le détail des calculs seront disponibles dans un prochain papier de revue.

Pour commencer il est nécessaire d'effectuer un développement de Taylor à l'ordre 2 de l'Eq. (5). Ensuite, en s'inspirant de la démarche effectuée dans le cas Gaussien dans [10], on utilise la définition suivante :

$$g_{fim}\left(\partial\mathbf{R}, \partial\mathbf{R}\right) = -\mathbb{E}\left[\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}l\left(\{\mathbf{z}_k\}|\mathbf{R} + t\partial\mathbf{R}\right)\right|_{t=0}\right] \tag{9}$$

Notons que pour certaines étapes du calcul et en particulier des termes α et β, nous avons utilisé une démarche similaire à celle présentée dans [9]. □

Dans le cas Gaussien, la métrique équivalente à celle de l'Eq. (7) se réduit au premier terme $KTr\left\{\left(\partial\mathbf{R}\mathbf{R}^{-1}\right)^2\right\}$.

3.2 Calcul de la Matrice d'Information de Fisher

Proposition 3.2. Matrice d'Information de Fisher
Pour un ensemble de données $\{\mathbf{z}_k\}$ distribué selon une loi elliptique complexe de matrice de covariance $\mathbf{R}$ et de générateur g, la matrice d'information de Fisher $\mathbf{G}_{fim}$ est la suivante :

$$\mathbf{G}_{fim} = K\left(\alpha\mathbf{I}_{M^2} + \beta\begin{bmatrix}\mathbf{1}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{M^2-M}\end{bmatrix}\right) \tag{10}$$

où α et β sont définis dans l'Eq (8).

Démonstration. Nous ne présentons encore une fois que les étapes principales de la démonstration. Le détail des calculs seront disponibles dans un prochain papier de revue. Pour cette preuve, nous avons encore utilisé la même démarche que celle présentées dans [10] pour le cas Gaussien.

Il faut tout d'abord polarisé la métrique (7) :

$$g_{fim}(\mathbf{R}_1, \mathbf{R}_2) = \frac{1}{4}\left(g_{fim}(\mathbf{R}_1 + \mathbf{R}_2, \mathbf{R}_1 + \mathbf{R}_2) - g_{fim}(\mathbf{R}_1 - \mathbf{R}_2, \mathbf{R}_1 - \mathbf{R}_2)\right) \quad (11)$$

où $\mathbf{R}_1$ et $\mathbf{R}_2$ sont des matrices de covariances. On choisit ensuite une base orthonormale pour l'espace tangent à la matrice $\mathbf{R}$:

$$\mathbf{D}_{ij}^h = \mathbf{R}^{1/2}\mathbf{D}_{ij}^{h0}\mathbf{R}^{1/2} \quad (12)$$

où h correspond au cas réel r ou imaginaire im (avec le cas particulier des éléments diagonaux $\mathbf{D}_{ii}$ uniquement en réels). Les matrices $\mathbf{D}_{ij}^{h0}$ composent donc la base orthonormale pour l'espace tangent à la matrice $\mathbf{I}$:

$$\begin{array}{lcl} \mathbf{D}_{ii}^{0} & = & 1 \text{ pour élément (i,i) et } 0 \text{ partout ailleurs} \\ \mathbf{D}_{ii}^{r0} & = & 2^{-\frac{1}{2}} \text{ pour élément (i,j) et (j,i) et } 0 \text{ partout ailleurs} \\ \mathbf{D}_{ii}^{im0} & = & 2^{-\frac{1}{2}}\sqrt{-1} \text{ pour élément (i,j) et} \\ & & -2^{-\frac{1}{2}}\sqrt{-1} \text{ pour élément (j,i) et } 0 \text{ partout ailleurs} \end{array} \quad (13)$$

Finalement, on peut calculer la matrice d'information de Fisher à partir de l'Eq. (11) et de la base $\{\mathbf{D}_{ij}^h\}$ définie ci-dessus. □

Remarque 3.1. Cas particulier Gaussien
Dans le cas Gaussien nous avons le résultat suivant :

$$\mathbf{G}_{fim} = K\mathbf{I}_{M^2} \quad (14)$$

En comparant ce résultat avec celui obtenu dans l'Eq. (10), on note que ce dernier conduit à un résultat plus structuré ainsi que l'intégration des paramètres α et β (qui sont fonctions des paramètres de la loi comme la fonction génératrice $g(t)$).

3.3 Inégalité entre Borne de Cramer Rao Intrinsèque et Erreur Quadratique Moyenne dans la distance naturelle

Tout d'abord pour définir la relation reliant la BCR intrinsèque à l'Erreur Quadratique Moyenne (EQM) sur la variété Riemannienne, il est nécessaire de définir la distance entre 2 matrices quand les données sont distribuées selon une loi elliptique complexe. A partir de la métrique calculée dans la proposition 3.1, nous calculons la distance associée dans la proposition suivante.

Proposition 3.3. Distance sur la variété Riemannienne pour des données elliptiques complexes
Soient $\mathbf{R}_1$ et $\mathbf{R}_2$ deux matrices de covariance issues d'une distribution elliptique complexe de même générateur g, la distance naturelle entre ces deux matrices est donnée par :

$$d_{natural}(\mathbf{R}_1, \mathbf{R}_2) = \alpha \sum_{i=1}^{M} \log^2 \lambda_i + \beta \left(\sum_{i=1}^{M} \log \lambda_i\right)^2 \quad (15)$$

où α et β sont définis dans l'Eq (8) et $\{\lambda_i\}$ sont les valeurs propres de la matrice $\mathbf{R}_1^{-1}\mathbf{R}_2$

Démonstration. La preuve reprend les étapes du cas Gaussien pour passer de la métrique à la distance. Tout d'abord, on vérifie que la géodésique utilisée dans le cas Gaussien est aussi valable dans le cas elliptique et ensuite le démonstration est directe [12, 13, 14]. □

Dans le cas Gaussien, nous retrouvons la distance bien connue comprenant seulement le premier terme avec $\alpha = 1$.

Finalement nous obtenons la proposition finale ci-dessous.

Proposition 3.4. Lien entre BCR intrinsèque et erreur quadratique moyenne sur la distance naturelle
La BCR intrinsèque et l'erreur quadratique moyenne sur la distance naturelle entre $\mathbf{R}$ et un estimateur non biaisé $\hat{\mathbf{R}}$ sont reliées par l'inégalité suivante :

$$\mathbb{E}[d_{natural}^2(\mathbf{R}, \hat{\mathbf{R}})] \geq Tr(\mathbf{G}_{fim}^{-1}) \quad (16)$$

où $\mathbf{G}_{fim}$ est défini dans l'Eq (10).

Démonstration. La preuve est quasi identique à celle du théorème 4 de [10]. □

4 Simulations

Pour valider nos résultats et comparer les performances d'estimateurs, nous utilisons la t-distribution définie par son degré de liberté d qui fait partie de la famille des distributions elliptiques complexes. Dans ce cas, des formules analytiques sont disponibles pour les paramètres α et β de l'Eq. (8) :

$$\begin{cases} \alpha & = 1 \\ \beta & = -\dfrac{2}{M+d+1} \end{cases} \quad (17)$$

Pour cette distribution, nous avons aussi une forme simple pour l'estimateur du maximum de vraisemblance de la matrice :

$$\hat{\mathbf{R}}_{MLE} = \frac{1}{K}\sum_{k=1}^{K} \frac{2M+d}{d + 2\mathbf{z}^H \hat{\mathbf{R}}_{MLE}^{-1}\mathbf{z}_k}\mathbf{z}_k\mathbf{z}_k^H \quad (18)$$

Nous comparerons ses performances à celles de la SCM, $\hat{\mathbf{R}}_{SCM} = (1/K)\sum_{k=1}^{K}\mathbf{z}_k\mathbf{z}_k^H$ et de l'estimateur de Tyler :

$$\hat{\mathbf{R}}_{Tyl} = \frac{M}{K}\sum_{k=1}^{K}\frac{\mathbf{z}_k\mathbf{z}_k^H}{\mathbf{z}^H\hat{\mathbf{R}}_{Tyl}^{-1}\mathbf{z}_k} \quad (19)$$

La vraie matrice de covariance a pour élément $\mathbf{R}(i,j) = \rho^{|i-j|}$. Nous calculons ainsi la CRB en utilisant la formule $Tr(\mathbf{G}_{fim}^{-1})$ et l'Eq (10). Nous estimons ensuite les EQMs naturelles $\mathbb{E}[d_{natural}^2(\mathbf{R}, \hat{\mathbf{R}})]$ pour les 3 estimateurs en effectuant 10000 tirages de Monte-Carlo.

Sur la figure 1, le degré de liberté est choisi très fort et donc nous sommes proches d'une distribution Gaussienne. Nous montrons la BCR naturelle ainsi que les EQMs naturelles pour les 3 estimateurs, SCM, MLE et Tyler. On peut noter que la BCR est bien une borne inférieure et que la SCM et le MLE ont des

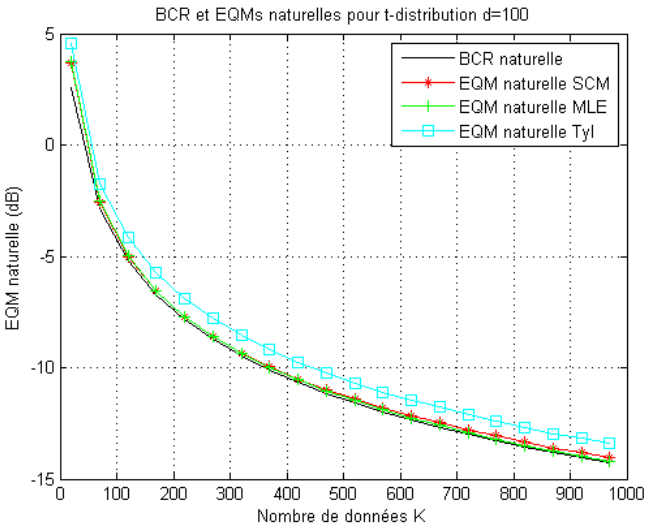

FIGURE 1 – CRB naturelle, EQMs naturelles pour la SCM, le MLE et Tyler. $m = 6$, $\rho = 0.9$ et $d = 100$.

performances très proches ce qui est tout à fait attendu. En revanche, l'estimateur de Tyler a de moins bonnes performances ce qui rejoint les analyses faites par exemple dans [15].

Sur la figue 2, nous montrons les mêmes courbes pour un cas plus hétérogène (non Gaussien) avec $d = 10$. Le MLE a dans ce cas toujours de bonnes performances tandis que celles de la SCM sont fortement dégradées. Au contraire, les performances de l'estimateur de Tyler sont un peu meilleures que dans le cas quasi-Gaussien.

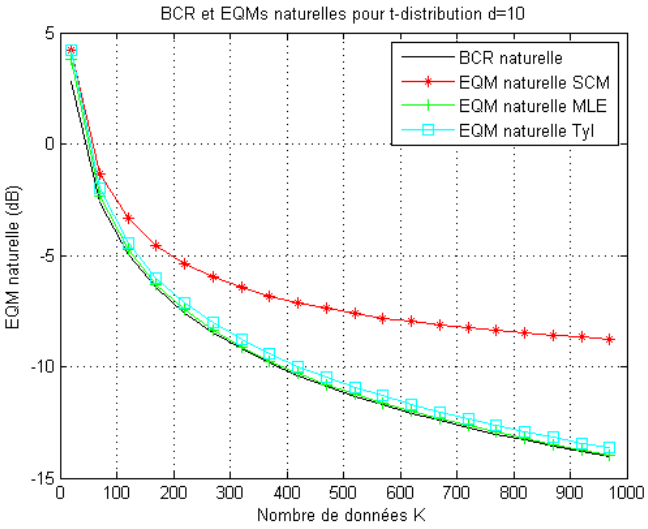

FIGURE 2 – CRB naturelle, EQMs naturelles pour la SCM, le MLE et Tyler. $m = 6$, $\rho = 0.9$ et $d = 10$.

5 Conclusion

Dans ce papier, nous avons proposé une expression de la borne de Cramér-Rao intrinsèque pour la matrice de covariance issue de données distribuées selon une loi elliptique complexe. Cette dérivation s'appuie sur la propriété géométrique que les matrices de covariance appartiennent à une variété Riemannienne. Ce calcul nous a permis aussi de définir une nouvelle distance naturelle sur cette variété associée aux distributions elliptiques complexes. Les simulations sur des données t-distribuées ont permis de valider notre résultat de BCR et de retrouver certaines conclusions de la littérature sur le comportement de la SCM et de l'estimateur de Tyler.

En perspectives, nous analyserons plus finement nos résultats afin d'exhiber les différences avec la BCR développée dans [9] ce qui devrait permettre de mesurer l'apport de la géométrie Riemannienne dans l'étude des bornes inférieures de l'EQM. Finalement, nous essaieront d'étendre ce calcul aux BCR pour des projecteurs et/ou de smatrices structurées.

Références

[1] E. Ollila, D. Tyler, V. Koivunen, and H. Poor, "Complex elliptically symmetric distributions : Survey, new results and applications," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5597–5625, 2012.

[2] P.J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 73 – 101, 1964.

[3] D. Tyler, "Robustness and efficiency properties of scatter matrices," *Biometrika*, vol. 70, no. 2, pp. 411, 1983.

[4] F. Pascal, Y. Chitour, J.P. Ovarlez, P. Forster, and P. Larzabal, "Existence and characterization of the covariance matrix maximum likelihood estimate in spherically invariant random processes," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 34 – 48, 2008.

[5] S.M. Kay, *Fundamentals of statistical signal processing - Detection theory - Vol.2*, Prentice Hall PTR., 1998.

[6] F. Pascal and A. Renaux, "Statistical analysis of the covariance matrix MLE in K-distributed clutter," *Elsevier Signal Processing*, vol. 90, pp. 1165 – 1175, 2010.

[7] M.N. El Korso, A. Renaux, and P. Forster, "CRLB under K-distributed observation with parameterized mean," in *IEEE Sensor Array Multichannel Workshop SAM-2014*, June 2014, A Coruña, Spain.

[8] M.V. Greco and F. Gini, "Cramér-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 6401 ?6409, December 2013.

[9] O. Besson and Y. Abramovich, "On the fisher information matrix for multivariate elliptically contoured distributions," *Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, Nov 2013.

[10] S. T. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1610–1630, May 2005.

[11] N. Boumal, "On intrinsic Cramér-Rao bounds for Riemannian submanifolds and quotient manifolds," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1809–1821, April 2013.

[12] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, USA, 2008.

[13] G. Meyer, *Geometric optimization algorithms for linear regression on fixed-rank matrices*, Ph.D. thesis, University of Liège, 2011.

[14] S. Lang, *Fundamentals of differential geometry*, Springer, 1999.

[15] F. Pascal, P. Forster, J.P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2206 – 2217, June 2008.