

Borne Bayésienne pour les systèmes large-échelle : cas d'un bruit à queue lourde

Virginie OLLIER^{1,2}, Rémy BOYER², Mohammed Nabil EL KORSO³, Pascal LARZABAL¹

¹SATIE/ENS Cachan/Université Paris-Saclay, 61 av. du Président Wilson, 94235 Cachan cedex, France

²L2S/Université Paris-Sud, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France

³LEME/Université Paris-Ouest, 50 rue de Sèvres, 92410 Ville d'Avray, France

virginie.ollier@satie.ens-cachan.fr

Résumé – Nous nous intéressons dans ce travail à l'étude de bornes de Cramér-Rao Bayésiennes dans le cadre de l'estimation robuste de paramètres pour un bruit non-Gaussien. En effet, les mesures observées peuvent être perturbées par des réalisations de bruit aberrantes dont une modélisation *ad-hoc* est une loi de Student. A l'aide de la théorie des grandes matrices aléatoires, nous dérivons de manière analytique dans ce travail la borne Bayésienne de Cramér-Rao. Enfin, nous étudions l'effet du degré de liberté de la loi de Student sur leur comportement.

Abstract – We are discussing in this work Bayesian lower bounds for multi-parameter robust estimation in non-Gaussian noise environment. Indeed, the data model can be affected by the presence outliers, leading to sparse noise following a Student's t distribution for instance. Using some results from the large random matrix theory, we derive closed-form expressions of the Bayesian Cramér-Rao bound asymptotically and study the influence of the degree of freedom.

1 Introduction

En estimation robuste de paramètres [1], les mesures acquises sont supposées être potentiellement perturbées par la présence de réalisations de bruit dites aberrantes, contrairement au cas classique où la perturbation est supposée issue d'une distribution gaussienne. Considérer une loi de Student à faible degré de liberté prend en compte efficacement la présence d'aberrations [2], tandis que nous pouvons nous ramener au cas Gaussien pour un degré de liberté tendant vers l'infini. Nous introduisons ici un modèle Bayésien hiérarchique du bruit, qui présente l'avantage de prendre en compte un grand nombre de distributions et qui a été considéré dans un grand nombre d'applications telles que la détection d'anomalies pour les images hyperspectrales [3] ou la segmentation de séries temporelles en astronomie [4].

Etudier la BCRB (borne de Cramér-Rao Bayésienne) fournit une borne inférieure sur l'erreur quadratique moyenne d'un estimateur [5]. Dans ce travail, nous analysons la BCRB de l'amplitude et de l'hyperparamètre du bruit à l'aide de résultats de la théorie des matrices aléatoires [6, 7, 8], pour un régime dit asymptotique. Plus précisément, le nombre de mesures et le nombre de paramètres d'intérêt tendent tous deux vers l'infini, mais avec un rapport fini.

Dans ce papier, nous utilisons le symbole $\text{Tr}[\cdot]$ pour la trace, \mathbf{I}_K et $\mathbf{0}_{K \times 1}$ pour, respectivement, la matrice identité de taille

$K \times K$ et le vecteur nul de taille $K \times 1$. La densité de probabilité (ddp) d'une variable aléatoire u est notée $p(u)$. Plus précisément, la loi gaussienne s'écrit $\mathcal{N}(\cdot, \cdot)$, paramétrée par une moyenne et une matrice de covariance, $\mathcal{G}(\cdot, \cdot)$ est la loi Gamma décrite par un paramètre de forme et l'inverse du paramètre d'échelle, tandis que $\mathcal{IG}(\cdot, \cdot)$ est la loi inverse-Gamma. Si $u \sim \mathcal{G}(a, b)$, alors $p(u|a, b) = \frac{b^a u^{a-1} e^{-bu}}{\Gamma(a)}$, où $\Gamma(\cdot)$ est la fonction Gamma. Si $u \sim \mathcal{IG}(a, b)$, alors $p(u|a, b) = \frac{b^a u^{-a-1} e^{-\frac{b}{u}}}{\Gamma(a)}$. En outre, la loi de Student peut être généralisée à l'aide de trois paramètres telle que si $u \sim \mathcal{S}(\mu, \sigma^2, \nu)$, alors $p(u|\mu, \sigma^2, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\sigma^2}} (1 + \frac{1}{\nu} \frac{(u-\mu)^2}{\sigma^2})^{-\frac{\nu+1}{2}}$. La loi Gamma-normale est une distribution bivariable telle que si $(u, w) \sim \text{GammaNormal}(\mu, \lambda, a, b)$, alors $p(u, w|\mu, \lambda, a, b) = \frac{b^a \sqrt{\lambda}}{\Gamma(a)\sqrt{2\pi}} w^{a-\frac{1}{2}} e^{-bw} e^{-\frac{\lambda w(u-\mu)^2}{2}}$. Pour finir, le symbole $\xrightarrow{a.s.}$ signifie convergence presque sûre, la notation grand O est $O(\cdot)$, $\lambda_i(\cdot)$ est la i -ème valeur propre d'une matrice et l'espérance par rapport à $p(\mathbf{u}|\mathbf{w})$ se note $\mathbb{E}_{\mathbf{u}|\mathbf{w}}$.

2 Modèle d'observation en présence de bruit non-Gaussien

2.1 Définition du modèle

Notons \mathbf{y} le vecteur de mesures de taille $N \times 1$. Le modèle Bayésien linéaire hiérarchique s'écrit comme suit

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \quad (1)$$

Ces travaux sont financés par le projet MAGELLAN (ANR-14-CE23-0004-01).

où chaque élément $[\mathbf{A}]_{i,j}$ de la matrice \mathbf{A} de taille $N \times K$, avec $K < N$, est une variable indépendante et identiquement distribuée selon une loi gaussienne de moyenne nulle et de variance $1/N$ [8, 9].

L'amplitude inconnue \mathbf{x} est obtenue comme suit

$$\mathbf{x} = [x_1, \dots, x_K]^T \sim \mathcal{N}(\mathbf{0}_{K \times 1}, \sigma_x^2 \mathbf{I}_K), \quad (2)$$

avec σ_x^2 la variance supposée connue. Notons que les mesures sont perturbées par un vecteur bruit \mathbf{e} supposé indépendant de \mathbf{x} .

2.2 Modèle hiérarchique du bruit

Chaque élément du bruit e_i est supposé être une variable conditionnellement indépendante sachant γ et identiquement distribuée selon une loi gaussienne circulaire centrée, soit

$$e_i | \gamma \sim \mathcal{N}\left(0, \frac{\sigma^2}{\gamma}\right), \quad (3)$$

où γ est un hyperparamètre inconnu et σ^2 un paramètre d'échelle fixé. Lorsque l'hyperparamètre suit une loi Gamma, i.e.,

$$\gamma \sim \mathcal{G}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad (4)$$

où ν est le degré de liberté, la loi conjointe de (e_i, γ) est une loi Gamma-normale [10], telle que

$$(e_i, \gamma) \sim \text{GammaNormal}\left(0, \frac{1}{\sigma^2}, \frac{\nu}{2}, \frac{\nu}{2}\right). \quad (5)$$

La loi marginale de e_i consiste alors à intégrer la loi conjointe selon γ , ce qui s'écrit [11, 12]

$$\mathcal{S}(e_i | 0, \sigma^2, \nu) = \int_0^\infty \mathcal{N}\left(e_i | 0, \frac{\sigma^2}{\gamma}\right) \mathcal{G}\left(\gamma | \frac{\nu}{2}, \frac{\nu}{2}\right) d\gamma, \quad (6)$$

soit $e_i \sim \mathcal{S}(0, \sigma^2, \nu)$, qui est la généralisation de la loi de Student, définie par trois paramètres. Lorsque $\nu \rightarrow \infty$, cette loi tend vers une gaussienne de moyenne nulle et de variance σ^2 . Avec (3) et (4), et sachant que $\frac{1}{\gamma} \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, il est possible d'exprimer la variance notée σ_e^2 de chaque élément de \mathbf{e} , ce qui donne, pour $\nu > 2$,

$$\sigma_e^2 = \mathbb{E}_\gamma \mathbb{E}_{e_i | \gamma} \{e_i^2\} = \sigma^2 \mathbb{E}_\gamma \left\{ \frac{1}{\gamma} \right\} = \sigma^2 \frac{\nu}{\nu - 2}. \quad (7)$$

3 BCRB pour un bruit distribué selon une loi de Student

Le vecteur de paramètres inconnus est le suivant

$$\boldsymbol{\theta} = [\mathbf{x}^T, \gamma]^T. \quad (8)$$

Supposons \mathbf{x} et γ indépendants, de sorte que la ddp jointe $p(\mathbf{y}, \boldsymbol{\theta})$ se décompose comme suit

$$p(\mathbf{y}, \boldsymbol{\theta}) = p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) = p(\mathbf{y} | \boldsymbol{\theta}) p(\mathbf{x}) p(\gamma). \quad (9)$$

Si nous notons $\hat{\boldsymbol{\theta}}$ l'estimation du vecteur $\boldsymbol{\theta}$, l'EQM (erreur quadratique moyenne) vérifie l'inégalité suivante

$$\text{EQM}(\boldsymbol{\theta}) = \text{Tr} \left[\mathbb{E}_{\mathbf{y}, \boldsymbol{\theta}} \left\{ (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \right\} \right] \geq \text{Tr} [\mathbf{C}], \quad (10)$$

où \mathbf{C} est la matrice BCRB de taille $(K + 1) \times (K + 1)$, soit l'inverse de la matrice d'information Bayésienne notée \mathbf{J} , dont la structure est la suivante

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathbf{x}, \mathbf{x}} & \mathbf{0}_{K \times 1} \\ \mathbf{0}_{1 \times K} & \mathbf{J}_{\gamma, \gamma} \end{bmatrix}, \quad \mathbf{C} = \mathbf{J}^{-1} = \begin{bmatrix} \mathbf{C}_{\mathbf{x}, \mathbf{x}} & \mathbf{0}_{K \times 1} \\ \mathbf{0}_{1 \times K} & \mathbf{C}_{\gamma, \gamma} \end{bmatrix}. \quad (11)$$

Nous supposons que des conditions d'identifiabilité et de régularité sont respectées dans le modèle [5] de sorte que la matrice \mathbf{J} puisse se décomposer comme suit

$$\mathbf{J} = \mathbb{E}_{\boldsymbol{\theta}} \left\{ \mathbf{J}_D^{(\boldsymbol{\theta}, \boldsymbol{\theta})} \right\} + \mathbf{J}_P^{(\boldsymbol{\theta}, \boldsymbol{\theta})} + \mathbf{J}_{HP}^{(\boldsymbol{\theta}, \boldsymbol{\theta})}, \quad (12)$$

avec

$$[\mathbf{J}_D^{(\boldsymbol{\theta}, \boldsymbol{\theta})}]_{i,j} = \mathbb{E}_{\mathbf{y} | \boldsymbol{\theta}} \left\{ -\frac{\partial^2 \log p(\mathbf{y} | \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right\}, \quad (13)$$

$$[\mathbf{J}_P^{(\boldsymbol{\theta}, \boldsymbol{\theta})}]_{i,j} = \mathbb{E}_{\mathbf{x}} \left\{ -\frac{\partial^2 \log p(\mathbf{x})}{\partial \theta_i \partial \theta_j} \right\}, \quad (14)$$

$$[\mathbf{J}_{HP}^{(\boldsymbol{\theta}, \boldsymbol{\theta})}]_{i,j} = \mathbb{E}_\gamma \left\{ -\frac{\partial^2 \log p(\gamma)}{\partial \theta_i \partial \theta_j} \right\} \quad (15)$$

où $(i, j) \in \{1, \dots, K + 1\}^2$. Remarquons alors que nous pouvons écrire

$$\mathbf{J}_{\mathbf{x}, \mathbf{x}} = \mathbb{E}_\gamma \left\{ \mathbf{J}_D^{(\mathbf{x}, \mathbf{x})} \right\} + \mathbf{J}_P^{(\mathbf{x}, \mathbf{x})}, \quad (16)$$

$$\mathbf{J}_{\gamma, \gamma} = \mathbb{E}_\gamma \left\{ \mathbf{J}_D^{(\gamma, \gamma)} \right\} + \mathbf{J}_{HP}^{(\gamma, \gamma)}. \quad (17)$$

Nous cherchons à préciser les expressions intervenant dans (16) et (17). Pour cela, remarquons que le vecteur d'observation \mathbf{y} suit une loi gaussienne conditionnellement à $\boldsymbol{\theta}$, i.e.,

$$\mathbf{y} | \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{R}), \quad (18)$$

où $\boldsymbol{\mu} = \mathbf{A}\mathbf{x}$ et $\mathbf{R} = \frac{(\nu-2)\sigma_e^2}{\nu\gamma} \mathbf{I}_N$. En utilisant la formule de Slepian-Bangs [13],

$$[\mathbf{J}_D^{(\boldsymbol{\theta}, \boldsymbol{\theta})}]_{i,j} = \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_i} \right)^T \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_j} + \frac{1}{2} \text{Tr} \left[\frac{\mathbf{R}}{\partial \theta_i} \mathbf{R}^{-1} \frac{\mathbf{R}}{\partial \theta_j} \mathbf{R}^{-1} \right], \quad (19)$$

nous obtenons directement

$$\mathbf{J}_D^{(\mathbf{x}, \mathbf{x})} = \frac{\nu\gamma}{(\nu-2)\sigma_e^2} \mathbf{A}^T \mathbf{A}. \quad (20)$$

En outre, d'après (2) où les amplitudes x_i sont supposées indépendantes, nous pouvons écrire

$$-\log p(\mathbf{x}) = \sum_{i=1}^K \left(\frac{1}{2} \log(2\pi\sigma_x^2) + \frac{x_i^2}{2\sigma_x^2} \right), \quad (21)$$

ce qui donne

$$\mathbf{J}_P^{(\mathbf{x}, \mathbf{x})} = \frac{1}{\sigma_x^2} \mathbf{I}_K. \quad (22)$$

Sachant que $\mathbf{R}^{-1} = \frac{\nu}{\sigma_e^2} \mathbf{I}_N$, nous avons

$$\mathbf{J}_D^{(\gamma, \gamma)} = \frac{\sigma^4}{2\gamma^4} \text{Tr} [\mathbf{R}^{-2}] = \frac{N}{2\gamma^2}. \quad (23)$$

Pour finir, le dernier terme de $J_{\gamma,\gamma}$ s'écrit

$$J_{HP}^{(\gamma,\gamma)} = \mathbb{E}_\gamma \left\{ -\frac{\partial^2 \log p(\gamma)}{\partial \gamma^2} \right\} = \frac{\nu - 2}{2} \mathbb{E}_\gamma \left\{ \frac{1}{\gamma^2} \right\}. \quad (24)$$

Le moment d'ordre deux d'une variable selon une loi inverse Gamma est connu, i.e.,

$$\mathbb{E}_\gamma \left\{ \frac{1}{\gamma^2} \right\} = \frac{\nu^2}{(\nu - 2)(\nu - 4)}, \quad (25)$$

avec $\nu > 4$, ce qui donne

$$J_{\gamma,\gamma} = \frac{N\nu^2}{2(\nu - 2)(\nu - 4)} + \frac{\nu^2}{2(\nu - 4)}. \quad (26)$$

Pour tenir compte des performances globales, nous nous intéressons plus particulièrement à la trace normalisée de la matrice BCRB, notée simplement \mathcal{B} . Dans le cas de l'amplitude \mathbf{x} , nous obtenons

$$\mathcal{B}(\mathbf{x}) = \frac{\text{Tr}[\mathbf{C}_{\mathbf{x},\mathbf{x}}]}{K} \text{ avec } \mathbf{C}_{\mathbf{x},\mathbf{x}} = \sigma_x^2 (r\mathbf{A}^T\mathbf{A} + \mathbf{I}_K)^{-1}, \quad (27)$$

où $r = \text{RSB} \frac{\nu}{\nu-2}$ et $\text{RSB} = \frac{\sigma_x^2}{\sigma_\varepsilon^2}$ (rapport signal sur bruit).

4 Régime asymptotique

D'après (26), la BCRB de l'hyperparamètre en régime asymptotique s'écrit

$$C_{\gamma,\gamma} = \frac{1}{J_{\gamma,\gamma}} \xrightarrow{N \rightarrow \infty} C_{\gamma,\gamma}^\infty, \quad (28)$$

où

$$C_{\gamma,\gamma}^\infty = \frac{2(\nu - 4)}{\nu^2}. \quad (29)$$

4.1 Contexte des grandes matrices aléatoires

Dans le contexte des grandes matrices aléatoires, nous considérons le cas où $K, N \rightarrow \infty$ avec $\frac{K}{N} \rightarrow \beta \in]0, 1[$. Ainsi, nous obtenons l'expression de la BCRB en régime asymptotique, soit

$$\mathcal{B}(\mathbf{x}) \xrightarrow{a.s.} \mathcal{B}^\infty(\mathbf{x}). \quad (30)$$

Avec (27) et [7, p. 11], cela donne

$$\mathcal{B}^\infty(\mathbf{x}) = \sigma_x^2 \left(1 - \frac{f(r, \beta)}{4r\beta} \right) \quad (31)$$

où $f(r, \beta) = \left(\sqrt{r(1 + \sqrt{\beta})^2 + 1} - \sqrt{r(1 - \sqrt{\beta})^2 + 1} \right)^2$.

4.2 Analyses et comparaisons

4.2.1 Régimes limites

• Lorsque $\beta \ll 1$, i.e., $K \ll N$, en négligeant les termes d'ordre supérieur ou égal à $O(\beta^2)$, nous obtenons

$$f(r, \beta) \approx \frac{4\beta r^2}{r + 1}. \quad (32)$$

Une expression asymptotique de la BCRB pour l'amplitude est donc

$$\mathcal{B}^\infty(\mathbf{x}) \approx \frac{\sigma_x^2}{r + 1} = \frac{(\nu - 2)\sigma_x^2}{\nu(1 + \text{RSB}) - 2}. \quad (33)$$

• Lorsque r est petit, ce qui signifie à faible RSB, d'après la décomposition en série de Neumann [14], nous avons $(r\mathbf{A}^T\mathbf{A} + \mathbf{I}_K)^{-1} \approx \mathbf{I}_K - r\mathbf{A}^T\mathbf{A}$ lorsque la valeur propre maximale vérifie $\lambda_{\max}(r\mathbf{A}^T\mathbf{A}) < 1$. Remarquons que $r\lambda_{\max}(\mathbf{A}^T\mathbf{A}) \xrightarrow{a.s.} r(1 + \sqrt{\beta})^2$ [6, 7, 8]. En outre, pour $\text{RSB} \ll (\nu - 2)/(4\nu)$, nous obtenons

$$\begin{aligned} \mathcal{B}(\mathbf{x}) &\approx \frac{\sigma_x^2}{K} (\text{Tr}[\mathbf{I}_K] - r\text{Tr}[\mathbf{A}^T\mathbf{A}]) \\ &\xrightarrow{a.s.} \sigma_x^2(1 - r) = \frac{\sigma_x^2}{\nu - 2}(\nu - 2 - \nu\text{RSB}). \end{aligned} \quad (34)$$

• Pour un grand r , soit à fort RSB, on obtient [6, 7, 8]

$$\begin{aligned} \mathcal{B}(\mathbf{x}) &\approx \frac{\sigma_x^2}{rK} \left(\text{Tr}[(\mathbf{A}^T\mathbf{A})^{-1}] - \frac{1}{r}\text{Tr}[(\mathbf{A}^T\mathbf{A})^{-2}] \right) \\ &\xrightarrow{a.s.} \frac{\sigma_x^2}{r} \left(\frac{1}{1 - \beta} - \frac{1}{r} \frac{1}{(1 - \beta)^3} \right) \\ &= \frac{(\nu - 2)\sigma_x^2}{\nu\text{RSB}(1 - \beta)} \left(1 - \frac{\nu - 2}{\nu\text{RSB}(1 - \beta)^2} \right). \end{aligned} \quad (35)$$

Nous cherchons désormais à étudier l'influence de ν sur le comportement de $\mathcal{B}^\infty(\mathbf{x})$.

4.2.2 Influence du degré de liberté à RSB fixé

Dans ce but, nous introduisons deux modèles différents

$$(M_0) : \mathbf{y}_0 = \mathbf{A}\mathbf{x} + \mathbf{e}_0 \text{ avec } e_{i_0} \sim \mathcal{S}(0, \sigma_0^2, \nu_0), \quad (36)$$

$$(M_1) : \mathbf{y}_1 = \mathbf{A}\mathbf{x} + \mathbf{e}_1 \text{ avec } e_{i_1} \sim \mathcal{S}(0, \sigma_1^2, \nu_1). \quad (37)$$

Nous notons (M_0) le modèle de référence et (M_1) le modèle alternatif. D'après (31), l'expression asymptotique de la BCRB pour le k -ème modèle où $k \in \{0, 1\}$ est la suivante

$$\mathcal{B}_k^\infty(\mathbf{x}) = \sigma_x^2 \left(1 - \frac{f(r_k, \beta)}{4r_k\beta} \right) \quad (38)$$

avec $r_k = \text{SNR}_k \frac{\nu_k}{\nu_k - 2}$ et $\text{SNR}_k = \frac{\sigma_x^2}{\sigma_k^2}$. Pour pouvoir comparer les bornes, nous imposons $\text{SNR}_0 = \text{SNR}_1$, ce qui signifie que $r_1 = \frac{\nu_1(\nu_0 - 2)}{\nu_0(\nu_1 - 2)}r_0$. Un cas particulier à étudier est celui où ν_0 est faible, ce qui correspond à la loi de Student, et $\nu_1 \rightarrow \infty$ ce qui donne une gaussienne [2] et fournit l'expression suivante

$$\mathcal{B}_1^\infty(\mathbf{x}) \stackrel{\nu_1 \rightarrow \infty}{=} \sigma_x^2 \left(1 - \frac{\nu_0 f\left(\frac{\nu_0 - 2}{\nu_0}r_0, \beta\right)}{4(\nu_0 - 2)r_0\beta} \right). \quad (39)$$

5 Simulations numériques

Nous considérons le cas où $N = 100$ et $K = 10$ de sorte que $\beta \ll 1$. La variance de l'amplitude σ_x^2 est fixée à 1. Sur la figure 1, nous traçons la BCRB du vecteur amplitude \mathbf{x} , donnée par (27) et (31) (régime asymptotique), (33) (faible β), (34)

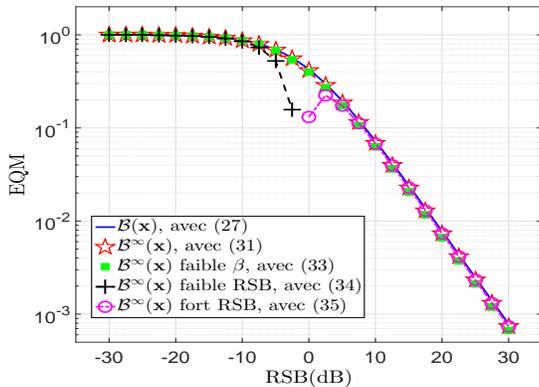


FIGURE 1 – $\mathcal{B}(\mathbf{x})$ en fonction du RSB en dB avec les approximations correspondant aux régimes limites de la section 4.2.1.

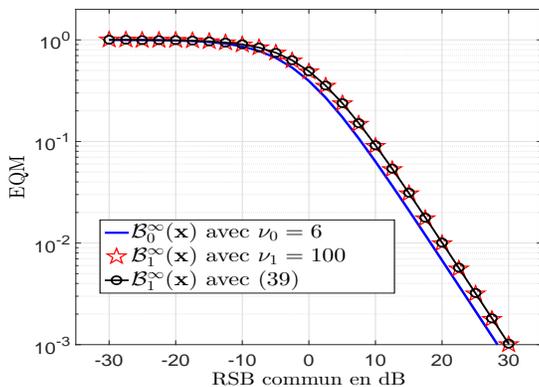


FIGURE 2 – Expressions des bornes asymptotiques pour deux modèles (M_0) et (M_1) à RSB fixé.

(faible RSB) et (35) (fort RSB), en fonction du RSB en dB pour $\nu = 6$.

Nous remarquons que $\mathcal{B}(\mathbf{x})$ coïncide exactement avec son expression asymptotique en (31). La théorie des matrices aléatoires permet donc de décrire avec précision le comportement de la BCRB lorsque $K, N \rightarrow \infty$ avec $\frac{K}{N} \rightarrow \beta$ et permet d'obtenir une expression analytique. Dans notre cas, $\beta = 0.1 \ll 1$ donc l'expression en (33) est une bonne approximation de la borne. Enfin, remarquons qu'à faible et fort RSB, les courbes collent à l'expression de la BCRB.

Sur la figure 2, comme expliqué en section 4.2.2, nous étudions deux modèles pour différentes valeurs de ν . Nous remarquons que la borne est plus faible pour $\nu_0 = 6$, plus particulièrement à faible bruit, que pour $\nu_1 = 100$. En outre, l'approximation donnée en (39) est valable puisque ν_1 est relativement élevé. Rappelons qu'un faible degré de liberté ν est bien adapté pour modéliser la présence de réalisations aberrantes, notamment à l'aide de distributions telles que la loi de Student, particulièrement robuste [15]. En revanche, un large degré de liberté tend vers le cas Gaussien. D'après la figure 2, on peut espérer atteindre de meilleures performances si nous favorisons des

TABLE 1 – $C_{\gamma,\gamma}^\infty$ en fonction de ν .

ν	10	30	50	70	100
$C_{\gamma,\gamma}^\infty$ avec (29)	0.1	0.06	0.04	0.03	0.02

distributions dites à queue lourde telles que la loi de Student.

Pour finir, le tableau 1 représente l'évolution de $C_{\gamma,\gamma}^\infty$ en fonction de ν . Contrairement à l'amplitude, nous remarquons qu'une BCRB plus faible est obtenue pour l'hyperparamètre lorsque ν augmente, i.e., lorsque la distribution du bruit tend vers une gaussienne.

6 Conclusion

Dans cette étude, nous caractérisons de manière analytique les performances d'estimation minimales dans le contexte Bayésien pour un système linéaire dit large-échelle et un bruit à queue lourde. Le modèle de bruit est de type hiérarchique selon une loi Gamma-normale favorisant l'apparition de réalisations dites aberrantes. En se basant sur la théorie des grandes matrices aléatoires, nous fournissons ici des expressions analytiques simples de la BCRB pour l'estimation d'un vecteur aléatoire et de l'hyperparamètre du bruit, dans plusieurs régimes limites typiques.

Références

- [1] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing : A tutorial-style treatment of fundamental concepts," *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 61–80, 2012.
- [2] S. Kotz and S. Nadarajah, *Multivariate t-distributions and their applications*. Cambridge University Press, 2004.
- [3] G. E. Newstadt, A. O. Hero, and J. Simmons, "Robust spectral unmixing for anomaly detection," in *IEEE Workshop on Statistical Signal Processing (SSP)*, Gold Coast, VIC, 2014, pp. 109–112.
- [4] N. Dobleon, J.-Y. Tournet, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series : Bayesian sampling with a hierarchical model," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 414–423, 2007.
- [5] H. L. Van Trees and K. L. Bell, *Bayesian bounds for parameter estimation and nonlinear filtering/tracking*. New York : Wiley-IEEE Press, 2007.
- [6] J. W. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate analysis*, vol. 54, no. 2, pp. 175–192, 1995.
- [7] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Foundations and Trends in Communications and Information Theory. Now Publishers Inc., 2004, vol. 1, no. 1.
- [8] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.
- [9] V. V. Buldygin and Y. V. Kozachenko, *Metric characterization of random variables and random processes*. American Mathematical Soc., 2000, vol. 188.
- [10] J. M. Bernardo and A. F. M. Smith, *Bayesian theory*. New York : J. Wiley, 1994.
- [11] M. Svensén and C. M. Bishop, "Robust Bayesian mixture modelling," *Neurocomputing*, vol. 64, pp. 235–252, 2005.
- [12] J. Christmas, "Bayesian spectral analysis with Student-t noise," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2871–2878, 2014.
- [13] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson Prentice Hall, Upper Saddle River, NJ, 2005.
- [14] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists, sixth edition*. Academic press, 2005.
- [15] K. L. Lange, R. J. A. Little, and J. M. G. Taylor, "Robust statistical modeling using the t distribution," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 881–896, 1989.