

Maximum de vraisemblance paramétrée (PML) : application à la localisation spatio-temporelle radar

Parameterized maximum likelihood method (PML) : application to space-time radar localization

par Nicolas BERTAUX⁽¹⁾, Philippe LARZABAL⁽¹⁾, Claude ADNET⁽²⁾ et Eric CHAUMETTE⁽²⁾

⁽¹⁾LESIR-ENS Cachan. 61 avenue du président Wilson. 94235 Cachan Cedex.

Tel : 01.47.40.27.06, Fax : 01.47.40.27.08

E-mail : nicolas.beraux@lesir.ens-cachan.fr

⁽²⁾THOMSON-CSF AIRSYS. 7-9 avenue des Mathurins. 92220 Bagneux.

résumé et mots clés

Nous présentons une technique du maximum de vraisemblance qui localise des sources dont les formes d'ondes sont identiques et connues. Il s'agit d'une localisation radar conjointe direction-retard qui est une extension aux cas de signaux cohérents des méthodes actuellement utilisées et exploitant la connaissance des signaux émis. Les résultats obtenus s'appliquent de la même manière aux cas des communications mobiles pour lesquelles on veut identifier le canal de propagation à l'aide d'un signal connu. Un estimateur exact du maximum de vraisemblance est présenté. Les variances d'estimation ainsi que les bornes de Cramer-Rao sont établies. Des résultats de simulations viennent illustrer le comportement des algorithmes pour lesquels les performances sont comparées à la borne de Cramer-Rao.

Traitement d'antenne actif, haute résolution, performances.

abstract and key words

We present a maximum likelihood method for the localization of sources with known waveforms. It's a joint space time radar localization which is a generalisation of recent methods to coherent signal. The obtained results are usefull in wireless communications for the identification of propagation channel with a pilot signal. An exact maximum likelihood method is presented. Variances of estimation and related Cramer Rao Bound are established. Simulations results illustrate the behaviour of the algorithm.

Actif array signal Processing, high resolution, performances.

1. introduction

Les techniques de traitement d'antenne ont connu un formidable essor ces dernières années. Beaucoup de travaux ont été effectués pour améliorer leur performance. Dans ces développements, le

signal reçu a souvent été considéré inconnu. Dans ce cadre là, des algorithmes du type sous espace MUSIC [Schmidt81], [Bi-venvenu83], ESPRIT [Roy89], et bien d'autres [Kumaresan83], [Reddi79], [Viberg91], ont été développés sur la seule base de considérations statistiques sur le signal reçu. De la même manière, l'estimation des temps de retard a été traitée dans [Figueiredo83], [Lichtenstein68], [Manabe92], [Manickam94]. Il s'agit là d'es-

timations disjointes. Plus performantes, des méthodes du maximum de vraisemblance déterministes [Stoica90], [Bohme86], [Ottersten93] et stochastiques [Bohme84], [Stoica89], [Stoica90] ont vu le jour. Elles modélisent le signal inconnu soit comme des paramètres déterministes inconnus, soit comme un processus stochastique. Il apparaît ainsi que la méthode du maximum de vraisemblance déterministe MVD (situation passive) est sous optimale en présence de cibles seules, car nous disposons d'une information *a priori* très forte non utilisée : le signal émis est connu. La localisation de cibles radar, et non de brouilleurs bien entendu, doit donc inclure cette information dans le processus de localisation.

Ainsi un certain intérêt a été porté aux algorithmes de localisation qui font des *a priori* sur les signaux afin d'améliorer les performances en localisation. Cet intérêt est motivé par des applications pour lesquelles il est raisonnable de supposer que le signal est en partie connu. Une des premières applications concerne les communications mobiles où le message reçu contient un signal de référence connu qui peut être exploité afin d'améliorer les performances des algorithmes de localisation. Le problème du radar actif est très similaire dans la mesure où le signal réfléchi par les cibles peut être modélisé à l'aide de l'impulsion émise. La motivation principale du présent travail concerne une localisation spatio-temporelle radar à haute résolution.

Malgré son importance peu de travaux ont été effectués par la communauté scientifique à ce sujet. Récemment [Li93], [Li95], [Viberg97] ont proposé des algorithmes de localisation angulaire tenant compte des formes d'onde. Malgré des améliorations proposées dans [Gounon94], [Grouffaud97], [Cedervall97] ces algorithmes sont mis en échec dans le cas de sources cohérentes. Une amélioration fondée sur un lissage spatio-fréquentiel a cependant été introduite dans [Goncalves97]. A notre connaissance seules les références [Wax97] et [Graffoulière97] traitent le problème qui nous intéresse. Le premier ne propose pas un estimateur exact du maximum de vraisemblance et le second n'analyse que le cas monocible (ou cibles très éloignées). L'objectif du présent article est de proposer un estimateur exact du maximum de vraisemblance, une mise en œuvre raisonnable ainsi qu'une étude de performances.

Nous proposons pour cela de paramétrer les signaux reçus à partir de répliques retardées de l'impulsion émise. Il vient ainsi, contrairement au MVD passif, que le nombre de paramètres recherché est maintenant $k \cdot M$, (M étant le nombre de cibles et k le nombre de paramètres décrivant le modèle). Le modèle introduit dépendra du modèle de rétrodiffusion des cibles. Le nombre de paramètres à estimer est alors constant et indépendant du nombre de snapshots. Ce type d'estimateur sera nommé PML (Parameterized Maximum Likelihood method).

Présentation du document

Dans la section 2 nous introduisons le modèle paramétré utilisé pour le signal rétrodiffusé par les cibles. La log-vraisemblance, ainsi que l'estimateur PML associés sont introduits et analysés

dans la section 3. La mise en œuvre de la méthode PML est présentée et commentée dans la section 4. Les performances théoriques, biais, variance, ainsi que les bornes de Cramer-Rao associées sont calculées dans le paragraphe 5. Des simulations, caractéristique d'un contexte radar, viennent confirmer et illustrer les résultats escomptés. Les améliorations introduites par rapport au MVD et au classique filtre adapté sont ainsi mises en évidence. Une conclusion et des perspectives de recherche sont esquissées dans le paragraphe 7.

2. modélisation paramétrique du signal reçu

Nous supposons dans cette étude que les signaux reçus s'expriment comme une combinaison linéaire d'un certain jeu de fonctions de base connues. Ces fonctions de base sont des répliques retardées d'un même signal. Dans le cas idéal de cibles ponctuelles, nous pouvons supposer que la réflexion provenant de chaque cible est une version atténuée et retardée du signal transmis. Nous négligerons l'effet Doppler. Le travail que nous avons mené dans cette étude concerne l'estimation des angles d'arrivées, notés θ , et des temps de retard, noté τ , des signaux réfléchis par les cibles. Cela, à l'aide d'une antenne de géométrie quelconque composée de N capteurs identiques et omnidirectionnels. Les signaux seront prélevés régulièrement sur T échantillons.

On notera α_m le scalaire complexe modélisant la SER (Surface équivalent radar) de la cible m , $\mathbf{a}(\theta_m)$ le vecteur ($N \times 1$) directionnel associé à la cible indiquée m , $s(t)$ l'impulsion émise et $\mathbf{n}(t)$ le vecteur ($N \times 1$) de bruit additionnel supposé blanc, gaussien, circulaire, à moyenne nulle et de puissance σ^2 .

Soit \mathbf{X} la matrice ($N \times T$) des observations définie par : $\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_T)]$, $\bar{\mathbf{X}}$ le vecteur ($N \cdot T \times 1$) issu de la concaténation des colonnes de \mathbf{X} : $\bar{\mathbf{X}} = \text{vec}(\mathbf{X})$, soit \mathbf{N} la matrice ($N \times T$) des bruits définie par : $\mathbf{N} = [\mathbf{n}(t_1), \mathbf{n}(t_2), \dots, \mathbf{n}(t_T)]$, $\bar{\mathbf{N}}$ le vecteur ($N \cdot T \times 1$) issu de la concaténation des colonnes de \mathbf{N} : $\bar{\mathbf{N}} = \text{vec}(\mathbf{N})$, et soit $\mathbf{s}(\tau_m)$ le vecteur de dimension ($T \times 1$) correspondant à l'impulsion retardée de τ_m : $\mathbf{s}(\tau_m) = [s(t_1 - \tau_m), s(t_2 - \tau_m), \dots, s(t_T - \tau_m)]^T$. Alors les signaux arrivant sur le réseau s'écrivent :

$$\mathbf{x}(t) = \sum_{m=1}^M \alpha_m \cdot \mathbf{a}(\theta_m) \cdot s(t - \tau_m) + \mathbf{n}(t). \quad (1)$$

On obtient l'écriture matricielle concaténée suivante :

$$\bar{\mathbf{X}} = \sum_{m=1}^M \alpha_m \cdot (\mathbf{s}(\tau_m) \otimes \mathbf{a}(\theta_m)) + \bar{\mathbf{N}}, \quad (2)$$

où \otimes dénote le produit de Kronecker. Notons $\Psi(\theta, \tau)$ la matrice complexe $(T \cdot N \times M)$ des vecteurs « retard-direction » définie de la manière suivante :

$$\Psi(\theta, \tau) = [[\mathbf{s}(\tau_1) \otimes \mathbf{a}(\theta_1)], [\mathbf{s}(\tau_2) \otimes \mathbf{a}(\theta_2)], \dots, [\mathbf{s}(\tau_M) \otimes \mathbf{a}(\theta_M)]]$$

où θ est le vecteur réel $(M \times 1)$ des directions d'arrivées et τ le vecteur réel $(M \times 1)$ des retards.

Finalement les signaux arrivant sur l'antenne s'écrivent :

$$\bar{\mathbf{X}} = \Psi(\theta, \tau) \cdot \alpha + \bar{\mathbf{N}}. \quad (3)$$

où α est le vecteur complexe $(N \times 1)$ des coefficients de réflexion des cibles.

3. Log-vraisemblance et estimateur PML

3.1. Log-vraisemblance des observations

La vraisemblance des observations s'obtient de manière très similaire à celle du MVD [Stoica90]. Celle-ci s'écrit :

$$\begin{aligned} \mathcal{L}(\bar{\mathbf{X}}|\theta, \tau, \alpha, \sigma^2) \\ = -T \cdot N \cdot \log(\pi \cdot \sigma^2) - \frac{1}{\sigma^2} |\bar{\mathbf{X}} - \Psi(\theta, \tau) \cdot \alpha|^2. \end{aligned} \quad (4)$$

La suite de cette étude sera consacrée à l'estimation de $\hat{\theta}$, $\hat{\tau}$, $\hat{\alpha}$ et $\hat{\sigma}^2$, maximisant la log-vraisemblance (4).

3.2. estimateur PML

La forme de la log-vraisemblance permet une détermination analytique, pour tout θ et τ , des estimateurs de la puissance de bruit σ^2 ainsi que du vecteur des coefficients de réflexion α . Ces estimateurs s'obtiennent en annulant les dérivées respectivement par rapport à σ^2 et à α . Ils ont pour expression :

$$\hat{\sigma}^2 = \frac{1}{N \cdot T} |\bar{\mathbf{X}} - \Psi(\theta, \tau) \cdot \alpha|^2 \quad (5)$$

et

$$\hat{\alpha} = (\Psi^H(\theta, \tau) \cdot \Psi(\theta, \tau))^{-1} \cdot \Psi^H(\theta, \tau) \cdot \bar{\mathbf{X}}. \quad (6)$$

En injectant les valeurs de $\hat{\sigma}^2$ et $\hat{\alpha}$ données par (5) et (6) dans la log-vraisemblance (4), on obtient les estimations $\hat{\theta}$ et $\hat{\tau}$. Il vient :

$$\begin{aligned} \hat{\theta}, \hat{\tau} = \operatorname{argmin}_{\theta, \tau} \\ \left| \bar{\mathbf{X}} - \Psi(\theta, \tau) \cdot (\Psi^H(\theta, \tau) \cdot \Psi(\theta, \tau))^{-1} \cdot \Psi^H(\theta, \tau) \cdot \bar{\mathbf{X}} \right|^2. \end{aligned}$$

Notant $\Pi_{\Psi}^{\perp}(\theta, \tau)$ le projecteur orthogonal sur l'image de $\Psi(\theta, \tau)$:

$$\Pi_{\Psi}^{\perp}(\theta, \tau) = \mathbf{I} - \Psi(\theta, \tau) \cdot (\Psi^H(\theta, \tau) \cdot \Psi(\theta, \tau))^{-1} \cdot \Psi^H(\theta, \tau) \quad (7)$$

les estimations $\hat{\theta}$ et $\hat{\tau}$ sont données par :

$$\begin{aligned} \hat{\theta}, \hat{\tau} &= \operatorname{argmin}_{\theta, \tau} \left| \Pi_{\Psi}^{\perp}(\theta, \tau) \cdot \bar{\mathbf{X}} \right|^2 \\ &= \operatorname{argmin}_{\theta, \tau} \left\{ \bar{\mathbf{X}}^H \cdot \Pi_{\Psi}^{\perp}(\theta, \tau) \cdot \bar{\mathbf{X}} \right\}. \end{aligned} \quad (8)$$

Nous noterons par la suite $L(\theta, \tau)$ le critère introduit dans l'équation (8) :

$$L(\theta, \tau) = \bar{\mathbf{X}}^H \cdot \Pi_{\Psi}^{\perp}(\theta, \tau) \cdot \bar{\mathbf{X}} \quad (9)$$

4. mise en œuvre du PML

4.1. méthode du second ordre

L'étude réalisée sur le MVD [Bertaux98b] a permis de conclure qu'une méthode itérative de descente du second ordre était performante pour minimiser le critère du maximum de vraisemblance dans une application radar. En effet, ce type d'algorithme propose un temps de convergence réduit et il est facile d'assurer la convergence vers un minimum ($H > 0$). De plus, notre travail étant axé sur l'étude des performances statistiques, la vitesse de convergence d'un algorithme du second ordre s'avère appréciable pour réaliser des simulations de Monte-Carlo. L'expression du critère (9) autorise de plus un calcul aisé du gradient et du hessien, ainsi l'utilisation d'un algorithme de Newton semble être une solution légitime. Notons \mathbf{p} le vecteur $(2 \cdot M \times 1)$ des paramètres de localisation spatio-temporelle : $\mathbf{p} = [\theta^T, \tau^T]^T$. L'algorithme itératif de Newton est alors le suivant :

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - [\mathbf{H}(L(\hat{\mathbf{p}}_k))]^{-1} \cdot \mathbf{grad}(L(\hat{\mathbf{p}}_k)) \quad (10)$$

4.2. calcul du gradient

Le gradient est calculé en annexe 8,2, il a pour expression :

$$\begin{aligned} \mathbf{grad}(\mathbf{p}) &= \frac{\partial L(\mathbf{p})}{\partial \mathbf{p}} = -2 \cdot \operatorname{Re} \\ &\left\{ \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes (\bar{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}))^T \right) \odot \left(\mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \bar{\mathbf{X}} \right) \right\} \end{aligned} \quad (11)$$

où $\Psi^{\#}(\mathbf{p})$ est la pseudo-inverse de $\Psi(\mathbf{p})$:

$$\Psi^{\#}(\mathbf{p}) = (\Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}))^{-1} \cdot \Psi^H(\mathbf{p})$$

et \mathbf{D} la matrice ($T \cdot N \times 2 \cdot M$) des dérivées des vecteurs « retard-direction » :

$$\mathbf{D} = \left[\frac{\partial \Psi(p_1)}{\partial p_1}, \dots, \frac{\partial \Psi(p_{2M})}{\partial p_{2M}} \right] \quad (12)$$

4.3. calcul du hessien

Dans l'algorithme du second ordre (10), le hessien de l'espérance du critère sera implanté. Pour cela il sera calculé en supposant être au voisinage de la solution effective $\hat{\mathbf{p}} \simeq \mathbf{p}$. En effet, il y a plusieurs raisons à cela. Le hessien a une expression très lourde et, pouvant être défini négatif, il peut nuire à la convergence de l'algorithme. De plus, à fort RSB, le hessien tend vers son espérance et, au voisinage de la solution, l'algorithme (10) s'écrit au premier ordre en $\frac{1}{RSB}$ selon (13) (annexe 8.3). Le hessien de l'espérance du critère est (annexe 8.4) défini non négatif, assurant ainsi la convergence de l'algorithme vers un minimum.

L'algorithme implanté est finalement de la forme :

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - \zeta \cdot [\mathbf{H}(E\{L(\hat{\mathbf{p}}_k)\})]^{-1} \cdot \mathbf{grad}(L(\hat{\mathbf{p}}_k)) \quad (13)$$

où le coefficient réel ζ permet d'assurer la décroissance du critère (9). En effet $\mathbf{H}(E\{L(\hat{\mathbf{p}}_k)\})$ étant toujours défini non-négatif, il existe alors un coefficient $0 < \zeta \leq 1$ tel que $L(\hat{\mathbf{p}}_{k+1}) \leq L(\hat{\mathbf{p}}_k)$. On fixe ζ à 1 tant que $L(\hat{\mathbf{p}}_{k+1}) \leq L(\hat{\mathbf{p}}_k)$, dans de cas contraire on diminue ζ jusqu'à ce que l'inégalité soit respectée. On obtient :

$$\mathbf{H}(E\{L(\mathbf{p})\}) = 2 \cdot \text{Re} \left\{ \text{diag} \left(\begin{bmatrix} \alpha^* \\ \alpha \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} \right) \right\} \quad (14)$$

L'estimation du hessien est finalement donnée par :

$$\mathbf{H}(E\{\widehat{L}(\mathbf{p})\}) = 2 \cdot \text{Re} \left\{ \text{diag} \left(\begin{bmatrix} \hat{\alpha}^*(\mathbf{p}) \\ \hat{\alpha}(\mathbf{p}) \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \hat{\alpha}(\mathbf{p}) \\ \hat{\alpha}^*(\mathbf{p}) \end{bmatrix} \right) \right\} \quad (15)$$

où $\hat{\alpha}(\mathbf{p}) = \Psi^{\#}(\mathbf{p}) \cdot \bar{\mathbf{X}}$.

4.4. écriture matricielle réductrice de la charge de calcul

L'intérêt principal de l'écriture proposée jusqu'à maintenant réside dans l'exploitation théorique qui peut être menée. On constate en revanche que le formalisme introduit engendre des vecteurs et des matrices de dimensions importantes. Comme cela est démontré dans [Bertaux98a], il est possible de revenir à des écritures matricielles de dimensions raisonnables pour les traitements à mettre en œuvre lors de l'implantation des algorithmes

proposés sur un système opérationnel. Les nouvelles écritures pour le critère, gradient et hessien sont les suivantes [Bertaux98a] :

$$L(\mathbf{p}) = \bar{\mathbf{X}}^H \cdot \bar{\mathbf{X}} - \text{diag}(\mathbf{A}^H(\theta) \cdot \mathbf{X} \cdot \mathbf{S}^*(\tau))^H \cdot \left(\mathbf{S}^H(\tau) \cdot \mathbf{S}(\tau) \odot \mathbf{A}^H(\theta) \cdot \mathbf{A}(\theta) \right)^{-1} \cdot \text{diag}(\mathbf{A}^H(\theta) \cdot \mathbf{X} \cdot \mathbf{S}^*(\tau)) \quad (16)$$

$$\frac{\partial L(\mathbf{p})}{\partial \mathbf{p}} = -2 \cdot \text{Re}$$

$$\left\{ \begin{array}{l} \left[\begin{array}{l} \mathbf{V}^*(\mathbf{p}) \\ \mathbf{V}(\mathbf{p}) \end{array} \right] \odot \\ \left[\begin{array}{l} \text{diag}(\partial \mathbf{A}^H(\theta) \cdot \mathbf{X} \cdot \mathbf{S}^*(\tau)) \\ - (\mathbf{S}^H(\tau) \cdot \mathbf{S}(\tau) \odot \partial \mathbf{A}^H(\theta) \cdot \mathbf{A}(\theta)) \cdot \mathbf{V}(\mathbf{p}) \\ \text{diag}(\mathbf{A}^H(\theta) \cdot \mathbf{X} \cdot \partial \mathbf{S}^*(\tau)) \\ - (\partial \mathbf{S}^H(\tau) \cdot \mathbf{S}(\tau) \odot \mathbf{A}^H(\theta) \cdot \mathbf{A}(\theta)) \cdot \mathbf{V}(\mathbf{p}) \end{array} \right] \end{array} \right\} \quad (17)$$

où l'on a posé :

$$\mathbf{V}(\mathbf{p}) = \Psi^{\#}(\mathbf{p}) \cdot \bar{\mathbf{X}} = (\mathbf{S}^H(\tau) \cdot \mathbf{S}(\tau) \odot \mathbf{A}^H(\theta) \cdot \mathbf{A}(\theta))^{-1} \cdot \text{diag}(\mathbf{A}^H(\theta) \cdot \mathbf{X} \cdot \mathbf{S}^*(\tau))$$

et où $\partial \mathbf{A}(\theta)$ est la matrice des dérivées des vecteurs directionnels et $\partial \mathbf{S}(\tau)$ est la matrice des dérivées des vecteurs signaux (ou impulsions retardées).

$$\mathbf{H}(E\{\widehat{L}(\mathbf{p})\}) = 2 \cdot \text{Re}$$

$$\left\{ \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \otimes (\mathbf{V}^*(\mathbf{p}) \cdot \mathbf{V}^T(\mathbf{p})) \right) \odot \left(\begin{array}{l} \left[\begin{array}{l} \mathbf{S}^H \cdot \mathbf{S} \odot \partial \mathbf{A}^H \cdot \partial \mathbf{A} \quad \mathbf{S}^H \cdot \partial \mathbf{S} \odot \partial \mathbf{A}^H \cdot \mathbf{A} \\ \partial \mathbf{S}^H \cdot \mathbf{S} \odot \mathbf{A}^H \cdot \partial \mathbf{A} \quad \partial \mathbf{S}^H \cdot \partial \mathbf{S} \odot \mathbf{A}^H \cdot \mathbf{A} \end{array} \right] - \\ \left[\begin{array}{l} \mathbf{S}^H \cdot \mathbf{S} \odot \partial \mathbf{A}^H \cdot \mathbf{A} \\ \partial \mathbf{S}^H \cdot \mathbf{S} \odot \mathbf{A}^H \cdot \mathbf{A} \end{array} \right] \cdot (\mathbf{S}^H \cdot \mathbf{S} \odot \mathbf{A}^H \cdot \mathbf{A})^{-1} \cdot \\ \left[\begin{array}{l} \mathbf{S}^H \cdot \mathbf{S} \odot \partial \mathbf{A}^H \cdot \mathbf{A} \\ \partial \mathbf{S}^H \cdot \mathbf{S} \odot \mathbf{A}^H \cdot \mathbf{A} \end{array} \right]^H \end{array} \right) \right\} \quad (14)$$

où l'on a omis les dépendances en τ de \mathbf{S} et $\partial \mathbf{S}$, et en θ de \mathbf{A} et $\partial \mathbf{A}$.

5. performances théoriques de l'estimateur

5.1. calcul de la borne de Cramer-Rao

Les calculs des bornes de Cramer-Rao (BCR) permettent de déterminer les limites inférieures de la variance d'estimation de tout estimateur non biaisé. Ces limites sont : $\text{var}(\hat{\mathbf{q}}) \geq \text{BCR} = \mathbf{J}_{\mathbf{q}}^{-1}$ où $\mathbf{J}_{\mathbf{q}}$ est la matrice d'information

de Fisher : $\mathbf{J}_{\mathbf{q}} = E \left\{ \frac{\partial \mathcal{L}(\mathbf{q})}{\partial \mathbf{q}} \cdot \left(\frac{\partial \mathcal{L}(\mathbf{q})}{\partial \mathbf{q}} \right)^H \right\}$, et où \mathbf{q} est le vecteur complexe des paramètres concaténés à estimer : $\mathbf{q} = [\sigma^2, \alpha^T, \alpha^H, \mathbf{p}^T = [\boldsymbol{\theta}^T, \boldsymbol{\tau}^T]]^T$, où \mathbf{p} est le vecteur des paramètres de localisation. α, α^* et σ^2 jouent ici le rôle de paramètres de nuisance. Le choix de la notation complexe pour le calcul de la matrice de Fisher [Bos94], a été motivé par le fait qu'elle permet d'obtenir aisément une forme réduite du bloc $[\mathbf{J}_{\mathbf{q}}^{-1}]_{\mathbf{p}}$ des paramètres de localisation.

La matrice $\mathbf{J}_{\mathbf{q}}$ a la forme bloc suivante :

$$\mathbf{J}_{\mathbf{q}} = \begin{bmatrix} J_{\sigma^2} & \mathbf{J}_{\sigma^2 \alpha} & \mathbf{J}_{\sigma^2 \alpha^*} & \mathbf{J}_{\sigma^2 \mathbf{p}} \\ [\mathbf{J}_{\sigma^2 \alpha}]^H & \mathbf{J}_{\alpha} & \mathbf{J}_{\alpha \alpha^*} & \mathbf{J}_{\alpha \mathbf{p}} \\ [\mathbf{J}_{\sigma^2 \alpha^*}]^H & [\mathbf{J}_{\alpha \alpha^*}]^H & \mathbf{J}_{\alpha^*} & \mathbf{J}_{\alpha^* \mathbf{p}} \\ [\mathbf{J}_{\sigma^2 \mathbf{p}}]^H & [\mathbf{J}_{\alpha \mathbf{p}}]^H & [\mathbf{J}_{\alpha^* \mathbf{p}}]^H & \mathbf{J}_{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} J_{\sigma^2} & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times 2N} \\ \mathbf{0}_{N \times 1} & \mathbf{J}_{\alpha} & \mathbf{0}_{N \times N} & \mathbf{J}_{\alpha \mathbf{p}} \\ \mathbf{0}_{N \times 1} & \mathbf{0}_{N \times N} & \mathbf{J}_{\alpha^*} & \mathbf{J}_{\alpha^* \mathbf{p}} \\ \mathbf{0}_{2N \times 1} & [\mathbf{J}_{\alpha \mathbf{p}}]^H & [\mathbf{J}_{\alpha^* \mathbf{p}}]^H & \mathbf{J}_{\mathbf{p}} \end{bmatrix},$$

où l'expression de chaque bloc est donnée en annexe 8.5.1.

Une expression réduite de la borne de Cramer-Rao relatif aux paramètres de localisation \mathbf{p} , est obtenue à l'aide une inversion par blocs de la matrice $\mathbf{J}_{\mathbf{q}}$. On abouti au résultats suivant :

$$\mathbf{BCR}(\mathbf{p}) = [\mathbf{J}_{\mathbf{q}}^{-1}]_{\mathbf{p}} = \frac{\sigma^2}{2} \cdot \left[\text{Re} \left\{ \text{diag} \left(\begin{bmatrix} \alpha^* \\ \alpha^* \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \right\} \right]^{-1},$$

où la démonstration est donnée en annexe 8.5.2. .

Cette expression générale de la BCR relative aux paramètres \mathbf{p} permet de retrouver le cas particulier monocible [Graffoulière97]. En effet, dans ce cas on vérifie que $[\mathbf{J}_{\mathbf{q}}^{-1}]_{\mathbf{p}} = \mathbf{J}_{\mathbf{p}}^{-1}$, car la matrice $\mathbf{J}_{\mathbf{q}}$ devient bloc diagonale.

5.2. variance théorique

En posant $\hat{\mathbf{p}} = \mathbf{p} + \tilde{\mathbf{p}}$, l'erreur d'estimation $\tilde{\mathbf{p}}$ est donnée de manière classique par : $\tilde{\mathbf{p}} \simeq -[\mathbf{H}(\mathbf{p})]^{-1} \cdot \mathbf{grad}(\mathbf{p})$. La matrice de covariance des estimations est alors donnée par l'équation suivante :

$$[\mathbf{H}(L(\mathbf{p}))]^{-1} \text{var}(\hat{\mathbf{p}}) \simeq E \left\{ [\mathbf{H}\{E(L(\mathbf{p}))\}]^{-1} \cdot \mathbf{grad}(L(\mathbf{p})) \cdot \mathbf{grad}^T(L(\mathbf{p})) \cdot [\mathbf{H}(L(\mathbf{p}))]^{-1} \right\}, \quad (20)$$

dans l'hypothèse où le biais est faible.

Il est apparemment difficile de simplifier l'équation (20). Le calcul proposé ici détermine la matrice de covariance des estimations au

premier ordre en $\frac{1}{RSB}$. Un calcul plus précis à un ordre supérieur, et valable pour de faible RSB, est ardu, mais envisageable.

Un développement limité de l'équation (20) permet d'obtenir (annexe 8.6) :

$$\text{var}(\hat{\mathbf{p}}) \simeq \sigma^2 \cdot [\mathbf{H}(E\{L(\mathbf{p})\})]^{-1} \quad (21)$$

où $\mathbf{H}(E\{L(\mathbf{p})\})$ est donné par l'équation (14).

On constate alors que la Borne de Cramer-Rao relatif aux paramètres de localisation donné par (19) est identique à l'expression, au premier ordre, de la variance théorique précédemment déterminée. Ainsi, on constate que l'estimateur PML est efficace à fort RSB.

5.3. biais théorique

Le biais de l'estimateur PML est nul à fort RSB mais est non nul au premier ordre en $\frac{1}{RSB}$ (annexe 8.7). Les simulations effectuées au paragraphe 6 montrent que, pour des cas difficiles où les cibles sont distantes de moins de $\frac{1}{4}$ de lobe en gisement et en retard, le biais est faible jusqu'à des RSB de $-10dB$. Il est important de conserver ce résultat à l'esprit car la comparaison de la variance d'un estimateur par rapport à la BCR n'a de sens que si le biais est négligeable.

6. simulations

Les simulations mettent en œuvre un réseau linéaire uniforme de 16 capteurs omnidirectionnels. La largeur du lobe à $-3dB$, notée $\Delta\theta_{3dB}$, donnée lors d'une formation de faisceau est de 4.4° pour un angle de gisement de 0° . L'impulsion émise a une largeur de bande de $2MHz$, la durée de l'impulsion est de $25\mu s$ et la fréquence d'échantillonnage est de $5MHz$, la période d'échantillonnage est ainsi de $0.2\mu s$. La largeur temporelle du lobe à $-3dB$ donnée par le filtre adapté à l'impulsion émise est de $0.42\mu s$. Notons qu'un réseau de ce type fournirait un gain équivalent, en sortie d'une formation de faisceaux (formation de voies) suivi d'un filtre adapté, d'environ $33dB$.

La figure 1a montre la partie réelle de l'impulsion de référence définie sur 128 échantillons. Les figures 1b et 1c sont le résultat d'une simulation présentant deux cibles séparées angulairement de 1° et temporellement de $0.8\mu s$. Le RSB est de $-5dB$ pour chaque signal. La figure 1b représente l'évolution temporelle du module de la sortie d'une formation de faisceaux orienté sur la bissectrice des deux cibles (angle moyen). La figure 1c représente le module de la sortie du filtre adapté. On peut constater que le filtre adapté permet, dans cette configuration simple, de séparer temporellement les cibles (points brillants). On comprend

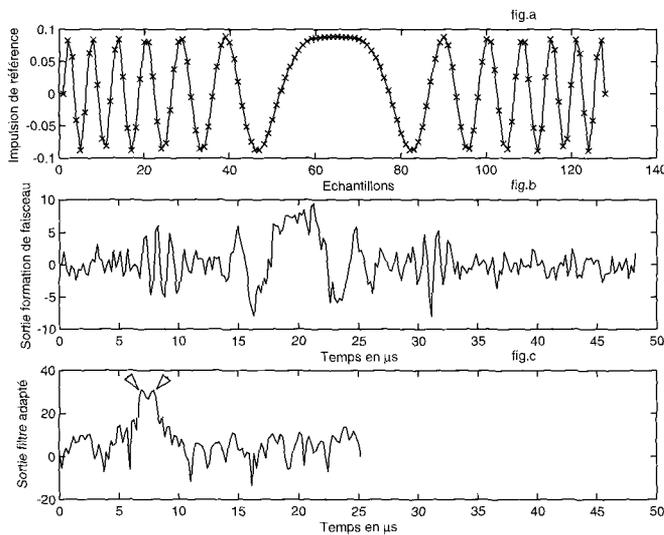


Figure 1. – a. Partie réelle de l’impulsion d’émission référence, b. et c. module du signal en sortie de la formation de faisceau, et du filtre adapté

aisément qu’il s’agit là d’une limite de fonctionnement et que si le scénario devenait plus exigeant la méthode échouerait.

Les simulations proposées permettent d’effectuer la comparaison des performances de l’estimateur PML à celles de l’estimateur (FF+CI) composé d’une Formation de Faisceaux suivie d’une Compression d’Impulsion (filtre adapté). Ce dernier estimateur est identique à l’estimateur du maximum de vraisemblance (avec connaissance de la réplique) dans l’hypothèse **mono-cible**. Cette méthode utilise un maillage, de l’axe direction, sur lequel on forme les faisceaux. Le signal issu de chaque faisceau est ensuite traité par un filtre adapté. L’estimation des directions et des retards consiste à rechercher les maxima d’un spectre spatio-directionnel. Un exemple où les cibles sont très éloignées en angle et en retard est donné par la figure 2.

Nous allons maintenant étudier le comportement statistique de diverses méthodes sur des simulations mettant en œuvre deux cibles. Toutes les simulations proposées mettront en œuvre deux cibles de même coefficient de SER égale à 1 (et donc de même RSB), ceci pour ne pas compliquer les interprétations. Chaque point de calcul est obtenu à partir de 500 tirages indépendants du bruit pour lesquels l’algorithme PML a effectivement localisé deux cibles distinctes. Les tirages ayant donné une solution unique (problème de convergence de l’algorithme de Gauss-Newton) ont été réalisés à nouveau. Les écarts-types des paramètres angles et retards sont normalisés par rapport aux largeurs de lobe respectives.

Par exemple, pour l’angle θ_1 : $STD_{\theta_1} = \frac{\text{écart type sur } \theta_1}{\Delta\theta_{3dB}}$.

Pour information et à titre de comparaison, nous avons reporté la BCR déterministe, correspondant à un modèle où les signaux sont inconnus mais déterministe, sur les planches où sont reportés les écarts types et les BCR de l’estimateur PML. Le MVD n’estimant que les directions d’arrivée, seules les planches représentant les écarts-types sur les angles sont concernées.

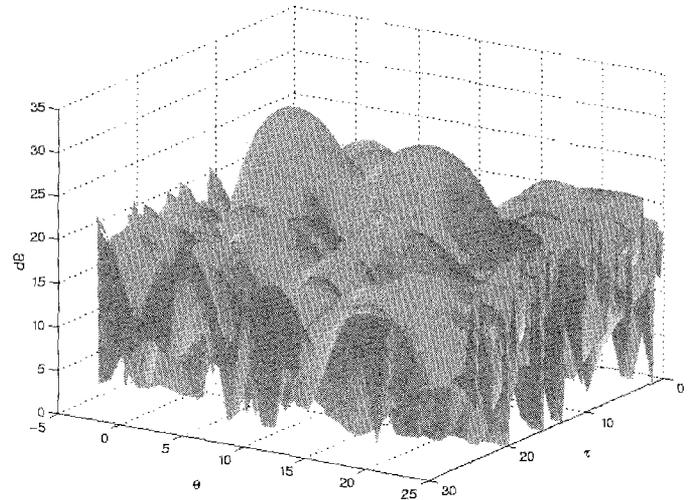


Figure 2. – Exemple de spectre 2D (angle/retard) dans le cas de cibles très éloignées.

6.1. comparaisons avec les méthodes actuelles : scénarii basse résolution

L’objet de cette section est d’illustrer les améliorations apportées par la méthode proposée par rapport aux techniques actuellement utilisées (FF+CI). Dans cette optique, les scénarii retenus doivent être compatibles avec le bon fonctionnement de la méthodes (FF+CI). Par conséquent, ceci explique les domaines d’études retenus. L’étude des performances haute résolution de l’estimateur proposé est donc reportée à la section suivante.

La figure 3 permet l’étude des performances en fonction du retard entre les deux cibles. Dans cette simulation, les cibles sont séparées angulairement de 5° et le RSB est fixé à $-5dB$. La figure 3a affiche les écarts-types mesurés, théoriques, ainsi que les bornes de Cramer-Rao de l’estimateur PML sur un paramètre angulaire. La figure 3c trace la moyenne de chaque position estimée, permettant ainsi d’évaluer le biais des estimateurs. Les figures 3b et 3d sont analogues sur l’axe temporel aux figures a et c. Le retard entre les signaux est normalisé, de manière analogue à STD_{θ_1} , par rapport à la largeur de lobe à $-3dB$ du filtre adapté. On peut constater que l’écart type mesuré correspond parfaitement à l’écart type théorique. Les figures 3b et 3d permettent de mettre en évidence un résultat classique : le filtre adapté ne fonctionne plus pour des cibles proches de moins d’une largeur de lobe.

La figure 4 expose, de manière similaire à 3, les performances obtenues en fonction de la séparation angulaire. Dans ce cas les deux cibles sont séparées temporellement de $0.5\mu s$. Le RSB est fixé à $-5dB$. Comme pour la figure 3, l’estimateur PML ne fait pas apparaître de biais significatif. L’écart type théorique rend parfaitement compte de l’écart type réel.

Une comparaison en fonction du RSB est donnée par la figure 5. Dans ce scénario les cibles sont séparées de 5° en gisement et de $0.5\mu s$ temporellement. Dans cette simulation, seul le RSB des cibles varie. Dans cette configuration l'estimateur PML n'a pas de biais significatif et les écarts-types théoriques et mesurés sont confondus aux BCR. Dans cette configuration particulière, relativement peu contraignante, l'estimateur PML est alors efficace. Comme nous pouvions nous y attendre les performances de la méthode FF+CI sont inférieures à celles du PML. Le biais est important quel que soit le RSB et la solution n'est plus acceptable en dessous de $-15dB$.

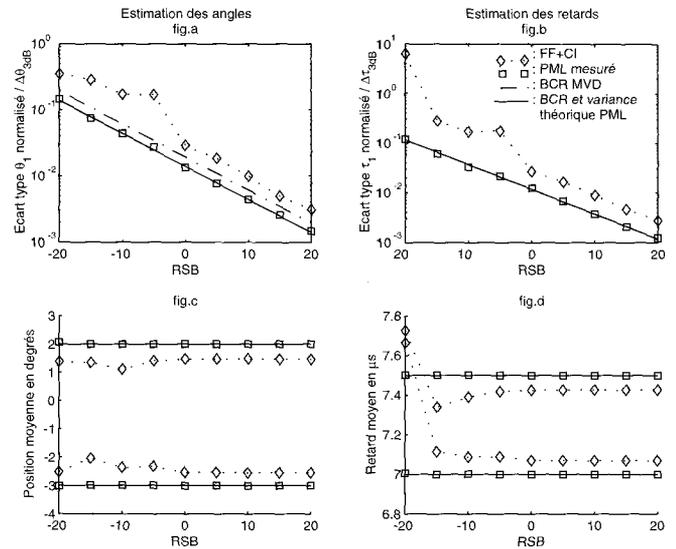


Figure 5. – Comparaisons de performances en fonction du RSB.

6.2. étude des performances de l'estimateur proposé : scénarii haute résolution

Les simulations suivantes, figures 6, 7 et 8, permettent l'étude des performances de l'estimateur PML lors d'un scénario plus contraignant. Dans le cas particulier de ces simulations, les deux cibles seront séparées angulairement ou temporellement de moins d'un quart de largeur de lobe. Il est clair que dans ce type de scénario, la majorité des estimateurs actuels sont inutilisables, en particulier la méthode comparée précédemment ou encore l'estimateur du maximum de vraisemblance déterministe (MVD).

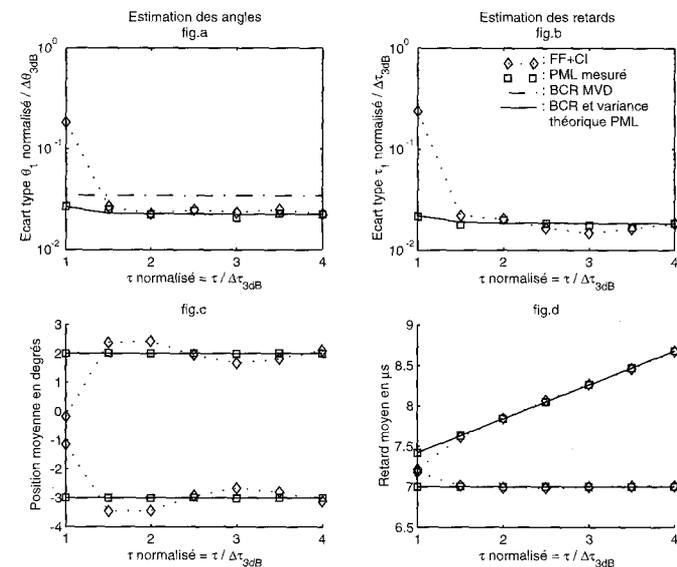


Figure 3. – Comparaison des performances en fonction du retard.

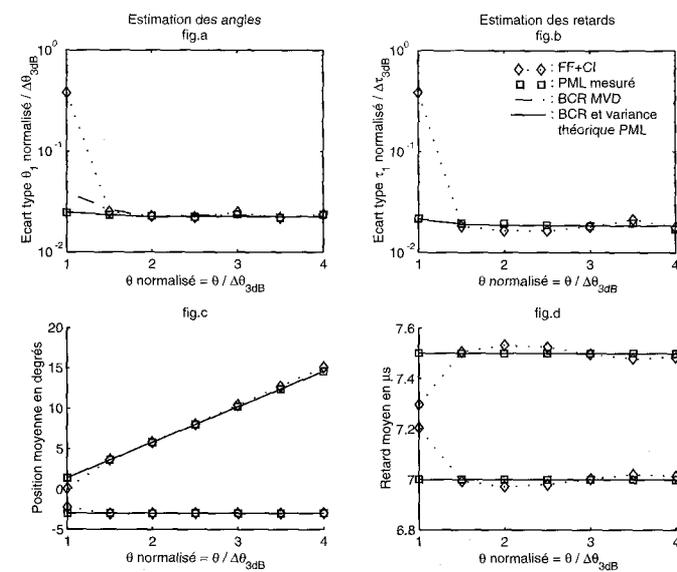


Figure 4. – Comparaison des performances en fonction de la séparation angulaire

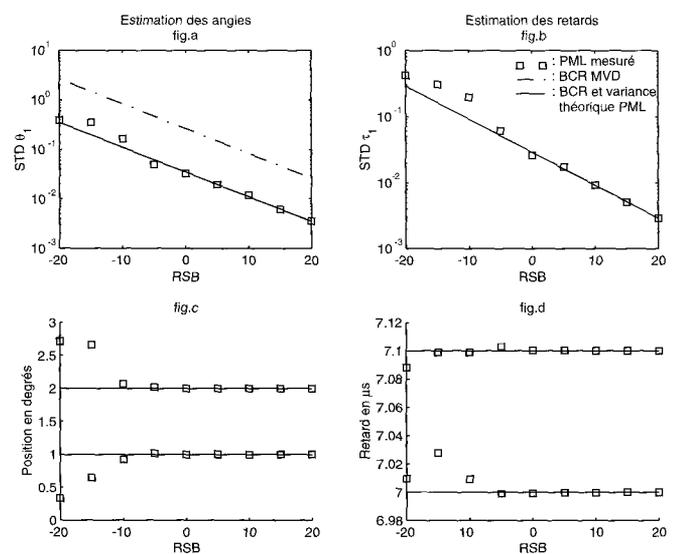


Figure 6. – Performances de l'estimateur PML en fonction du RSB.

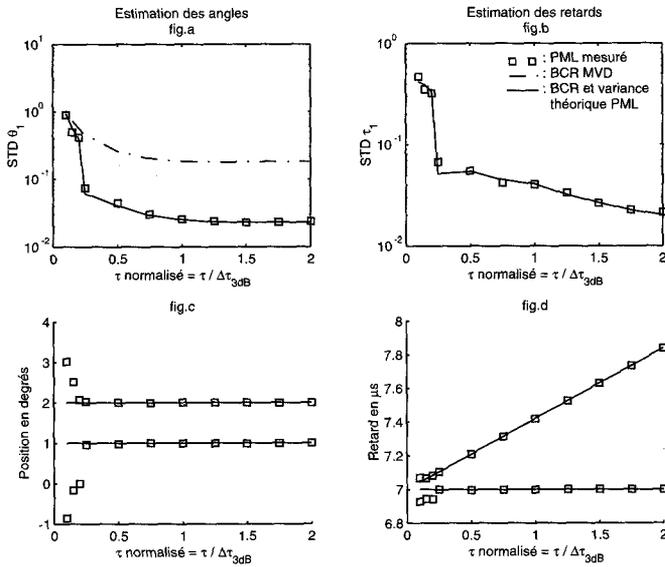


Figure 7. – Performances de l'estimateur PML en fonction du retard entre les signaux.

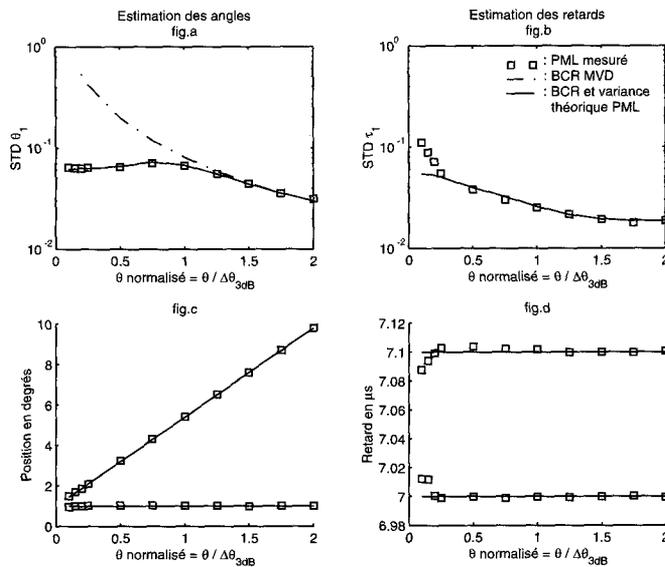


Figure 8. – Performances de l'estimateur PML en fonction de l'écart angulaire entre les cibles.

En effet, dans ce cas, la méthode classique (FF+CI) est alors en limite de résolution, et l'estimateur du MVD est à RSB trop faible. Ainsi, seules les caractéristiques de l'estimateur PML seront reportées sur ces différentes figures.

La figure 6 donne les performances en fonction du RSB. Les cibles sont séparées angulairement de 1° et temporellement de $0.1\mu s$ (notons ici que dans le cas précis de notre application, la période d'échantillonnage est de $0.2\mu s$). On constate qu'à partir de $-10dB$ un biais apparaît sur l'estimation PML et qu'au delà de $-15dB$ ce biais n'est plus négligeable. Les bornes de Cramer-Rao ne sont donc plus significatives en dessous de $-10dB$.

Remarquons aussi que l'estimateur PML, qui était jusqu'à présent consistant, ne l'est plus en dessous de $-10dB$ de RSB. On constate de plus que l'écart type théorique calculé à l'ordre 1 en $\frac{1}{RSB}$, même si il donne encore une estimation acceptable, ne correspond plus à l'écart type réel (mesuré) au delà de $-10dB$. Ceci explique le choix de $-5dB$ de RSB pour toutes les simulations en fonctions de θ et τ . Cela motive aussi un calcul ultérieur de la variance à des ordres plus élevés que dans (21).

La figure 7 met en œuvre deux cibles séparées de 1° angulairement pour un RSB fixé à $-5dB$. Quel que soit le retard entre les signaux, l'écart type théorique coïncide parfaitement avec l'écart type mesuré. L'estimateur PML ne se trouve biaisé que pour des retards inférieurs à $\frac{1}{4}$ de largeur de lobe. L'effet de seuillage au voisinage de $\frac{1}{4}$ de largeur de lobe, déjà observé dans [Pallas86], est à rapprocher, via son autocorrélation, du signal émis.

La figure 8 représente deux cibles séparées de $0.1\mu s$ et le RSB est toujours fixé à $-5dB$. Comme pour la figure précédente, on constate que l'estimateur PML n'est plus efficace lorsque l'écart angulaire ou le retard entre les cibles diminue. L'apparition d'un biais dans l'estimation des paramètres n'est détectable que pour des scénarii où les cibles sont proches de moins $\frac{1}{4}$ de lobe en gisement et en retard.

On constate sur les figures (fig.a) que les performances du PML sont toujours meilleures que la borne de Cramer-Rao déterministe. Cela se comprend aisément dans la mesure où nous avons introduit la connaissance du signal dans la vraisemblance. Le gain de performances est à apparenter au gain de compression d'impulsion, dans le cas monocible, apporté par le filtre adapté.

7. conclusions et perspectives

Conclusions

Le travail présenté dans ce rapport est une contribution au traitement radar multicible à haute résolution. Nous avons supposé dans cette étude qu'il n'y avait pas de brouilleurs et que l'effet du Doppler était négligeable.

Dans un premier temps nous avons replacé nos travaux dans le cadre strict, mais unificateur des méthodes du maximum de vraisemblance. Il en ressort que l'originalité principale de notre approche réside dans le fait que nous proposons d'utiliser la forme du signal émis afin de déterminer un estimateur optimal de la position des cibles à l'aide d'une méthode du maximum de vraisemblance.

Nous avons calculé la vraisemblance correspondante au modèle retenu pour la réflexion et introduit de manière originale un algorithme dénommé PML. L'atout de cet algorithme est qu'il incorpore de manière optimale un modèle des signaux rétrodiffusés par les cibles. Les variances théoriques des estimateurs ont été calculées. Les résultats ainsi obtenus sont parfaitement vérifiés par les variances observées à l'aide des simulations de Monte-Carlo. Les variances théoriques calculées mettent parfaitement en évidence la présence d'un effet de seuillage sur les performances lorsque les écarts temporels diminuent trop. Ainsi, pour deux contributeurs identiques, l'algorithme fournit des estimations jugées satisfaisantes pour des cibles au moins séparées d'un quart du lobe spatial et temporelle. Nous montrons que les performances obtenues sont toujours meilleures que la borne de Cramer-Rao correspondant au modèle déterministe (passif).

Les bornes de Cramer-Rao ont été établies afin de quantifier la qualité de l'estimateur proposé. Nous utilisons la notation complexe pour le calcul des bornes de Cramer-Rao pour montrer que ce dernier peut être simplifié afin d'éviter l'inversion de la matrice de Fisher relative à tous les paramètres. Nous montrons, de plus, que l'estimateur PML est efficace à fort rapport signal à bruit.

Les simulations ont mis en évidence les performances des algorithmes introduits. La retombée pratique immédiate est une nette augmentation des performances en termes de détection, biais, et variance des estimées. La méthode proposée est plus performante que les méthodes actuellement utilisées. La méthode proposée s'adapte à toute sorte d'impulsion mais les performances obtenues dépendent du temps de cohérence (gain du filtre adapté) de l'impulsion.

En contexte opérationnel il y aura simultanément des brouilleurs et des cibles. A partir de la vraisemblance complète, incluant les brouilleurs et les cibles, nous avons développé une méthode du maximum de vraisemblance qui estime la positions des cibles de manière optimale en présence de brouilleurs corrélés spatialement et temporellement. Cette étude en cours d'achèvement sera prochainement soumise à publication.

Perspectives

- Il reste encore beaucoup de travail afin de valider ces algorithmes. Premièrement, il faudra tester la robustesse de la méthode PML vis à vis des erreurs de calibration et de la corrélation spatiale du bruit. Cette méthode sera probablement beaucoup plus robuste que le MVD.
- Cette étude doit être généralisée à des scénarii de cibles de SER différentes, en particulier les cas d'une source d'amplitude inférieure au lobe secondaire d'une autre cible.
- Une perspective fort intéressante consiste à développer un estimateur de type PML pour lequel le signal émis dépendrait de la direction scrutée. Les gains obtenus devraient être considérables. Dans une telle situation, il y aura moins d'interférences entre les différentes directions scrutées.

8. annexes

8.1. notations et abréviations

- N nombre de capteurs
- M nombres de cibles
- T nombre d'échantillons
- θ angle d'arrivée (gisement)
- $\boldsymbol{\theta}$ vecteur des angles d'arrivée
- τ retard d'arrivée
- $\boldsymbol{\tau}$ vecteur des retards d'arrivée
- α coefficient de réflexion
- $\boldsymbol{\alpha}$ vecteur des coefficients de réflexion
- \mathbf{X} matrice ($N \times T$) des observations
- \mathbf{n} vecteur de bruit
- \mathbf{N} matrice ($N \times M$) de bruit
- σ^2 puissance du bruit
- \mathbf{a} vecteur directionnel
- \mathbf{s} vecteur impulsion d'émission
- \mathbf{A} matrice ($N \times M$) des vecteurs directionnels
- $\partial\mathbf{A}$ matrice ($N \times M$) des dérivées des vecteurs directionnels
- \mathbf{S} matrice ($T \times M$) des signaux sources : matrice des impulsions retardées
- $\partial\mathbf{S}$ matrice ($T \times M$) des dérivées des signaux sources
- $\boldsymbol{\Psi}$ matrices ($T \cdot N \times M$) des vecteurs « retard-direction »
- $\tilde{\mathbf{A}}$ vecteur résultat de la concaténation des vecteurs colonnes de la matrice \mathbf{A} , $\mathbf{A} = \text{vec}(\mathbf{A})$
- \mathbf{x}^T transposée de \mathbf{x}
- \mathbf{x}^H transposée conjuguée de \mathbf{x}
- \mathbf{x}^* conjuguée de \mathbf{x}
- $\mathbf{A}^\#$ pseudo-inverse de \mathbf{A}
- $\Pi_{\boldsymbol{\Psi}}$ projecteur sur l'espace engendré par les colonnes de la matrice $\boldsymbol{\Psi}$
- $\Pi_{\boldsymbol{\Psi}}^\perp$ projecteur orthogonal
- $\text{diag}(\mathbf{Q})$ diagonale de la matrice \mathbf{Q} sous forme de vecteurs colonnes
- \otimes produit de Kronecker
- \odot produit terme à terme, produit d'Hadamard
- \mathbf{e}_i est un vecteur colonne qui n'a que son $i^{\text{ème}}$ élément de non nul et égal à 1
- $\hat{\mathbf{p}}$ vecteur estimé
- $\tilde{\mathbf{p}}$ vecteur erreur d'estimation
- $E\{\mathbf{Q}\}$ espérance de \mathbf{Q}
- $\text{var}(\mathbf{p})$ variance de \mathbf{p}
- \mathbf{J} matrice d'information de Fisher
- \mathbf{H} hessien
- grad gradient
- $\underline{\mathbf{H}}$ matrice normalisée par rapport à $\|\boldsymbol{\alpha}\|$: $\underline{\mathbf{H}} = \frac{1}{\|\boldsymbol{\alpha}\|} \cdot \mathbf{H}$
- PML** : Parameterized Maximum Likelihood.
- BCR** : Borne de Cramer-Rao.
- MVD** : Maximum de Vraisemblance Déterministe.

FF+CI : Formation de Faisceaux (formation de voies) suivie d'une Compression d'Impulsion (ou filtre adapté)
RSB : Rapport Signal sur Bruit.

8.2. calcul du gradient

Déterminons le $i^{\text{ème}}$ élément du gradient en dérivant $L(\mathbf{p})$ (donnée par (9)), par rapport à p_i , cela donne :

$$\frac{\partial L(\mathbf{p})}{\partial p_i} = \bar{\mathbf{X}}^H \cdot \frac{\partial \Pi_{\Psi}^{\perp}(\mathbf{p})}{\partial p_i} \cdot \bar{\mathbf{X}}. \quad (22)$$

Pour calculer le terme $\frac{\partial \Pi_{\Psi}^{\perp}(\mathbf{p})}{\partial p_i}$, utilisons l'expression (7) du projecteur orthogonal. On obtient :

$$\begin{aligned} \frac{\partial \Pi_{\Psi}^{\perp}(\mathbf{p})}{\partial p_i} &= \left\{ -\frac{\partial \Psi(\mathbf{p})}{\partial p_i} \cdot \Psi^{\#}(\mathbf{p}) + \Psi^{\#H}(\mathbf{p}) \cdot \Psi^H(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_i} \cdot \Psi^{\#}(\mathbf{p}) \right\} + \{\dots\}^H \\ &= -\left\{ \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_i} \cdot \Psi^{\#}(\mathbf{p}) \right\} - \{\dots\}^H \end{aligned} \quad (23)$$

En remplaçant (23) dans l'équation (22), on arrive au résultat :

$$\frac{\partial L(\mathbf{p})}{\partial p_i} = -2 \cdot \text{Re} \left\{ \bar{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \bar{\mathbf{X}} \right\} \quad (24)$$

Le passage à une écriture matricielle ne pose aucune difficulté majeure. Pour aider à la compréhension lors du passage sous forme matricielle, il est important de constater que la matrice $\frac{\partial \Psi(\mathbf{p})}{\partial p_i}$, de par la forme de $\Psi(\mathbf{p})$, n'a que la $i^{\text{ème}}$ colonne de non nulle, elle peut donc s'écrire :

$$\frac{\partial \Psi(\mathbf{p})}{\partial p_i} = \left[\frac{\partial \Psi(\mathbf{p})}{\partial p_i} \right]_i \cdot \mathbf{e}_i^T = \frac{\partial \Psi(p_i)}{\partial p_i} \cdot \mathbf{e}_i^T$$

où \mathbf{e}_i est un vecteur colonne qui n'a que son $i^{\text{ème}}$ élément non nul et égal à 1 :

$$\mathbf{e}_i = [0 \quad \dots \quad 0 \quad 0 \quad 0 \quad \dots \quad 0]_i^T \quad (25)$$

Ainsi, pour obtenir l'écriture matricielle de (24), il suffit de remarquer que la matrice des vecteurs « retard-direction » dérivés a $2 \cdot M$ colonnes. Il est donc nécessaire de dupliquer le vecteur $\bar{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p})$, on obtient finalement :

$$\frac{\partial L(\mathbf{p})}{\partial \mathbf{p}} = -2 \cdot \text{Re} \left\{ \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes (\bar{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}))^T \right) \odot \left(\mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \bar{\mathbf{X}} \right) \right\} \quad (26)$$

8.3. étude au premier ordre de l'algorithme de Newton

Nous allons ici justifier de l'introduction de l'espérance dans l'algorithme de Newton (13). Afin de simplifier les notations, fixons la puissance du bruit : $\|\mathbf{n}\| = 1$. Ainsi, l'étude en fonction du RSB se ramène à l'étude en fonction du module du coefficient de SER, $\|\alpha\|$. Rappelons l'expression du critère (9) :

$$L(\theta, \tau) = \bar{\mathbf{X}}^H \cdot \Pi_{\Psi}^{\perp} \cdot \bar{\mathbf{X}} \text{ avec } \bar{\mathbf{X}} = \Psi \cdot \alpha + \bar{\mathbf{N}}.$$

Il peut se mettre sous la forme suivante :

$$L = \alpha^H \cdot \Psi^H \cdot \Pi_{\Psi}^{\perp} \cdot \Psi \cdot \alpha + \|\alpha\| \cdot \epsilon(\alpha) \quad (27)$$

où l'on constate aisément que

$$\alpha^H \cdot \Psi^H \cdot \Pi_{\Psi}^{\perp} \cdot \Psi \cdot \alpha = E\{L\}$$

à une constante additive près (31), et

$$\lim_{\|\alpha\| \rightarrow \infty} \epsilon(\alpha) = \text{cte.}$$

Notons $\underline{\alpha}$ le vecteur normalisé tel que $\underline{\alpha} = \frac{\alpha}{\|\alpha\|}$, l'expression (27) devient :

$$L = \|\alpha\|^2 \cdot \left(\underline{\alpha}^H \cdot \Psi^H \cdot \Pi_{\Psi}^{\perp} \cdot \Psi \cdot \underline{\alpha} + \frac{1}{\|\alpha\|} \cdot \epsilon(\alpha) \right)$$

où l'on notera par la suite :

$$E\{L\} = \underline{\alpha}^H \cdot \Psi^H \cdot \Pi_{\Psi}^{\perp} \cdot \Psi \cdot \underline{\alpha}.$$

La linéarité de la dérivation permet d'écrire l'inverse du hessien du critère L sous la forme :

$$\mathbf{H}^{-1}(L) = \frac{1}{\|\alpha\|^2} \cdot \left[\mathbf{H}(E\{L\}) + \frac{1}{\|\alpha\|} \cdot \mathbf{H}(\epsilon(\alpha)) \right]^{-1}.$$

Un développement limité lorsque $\frac{1}{\|\alpha\|} \rightarrow 0$ (c'est-à-dire à fort RSB) du terme en inverse de l'équation précédente permet d'écrire :

$$\begin{aligned} \mathbf{H}^{-1}(L) &\simeq \frac{1}{\|\alpha\|^2} \cdot \left(\mathbf{I} - \frac{1}{\|\alpha\|} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{H}(\epsilon(\alpha)) \right) \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \\ &\simeq \frac{1}{\|\alpha\|^2} \cdot \underbrace{\left[\mathbf{H}(E\{L\}) \right]^{-1}}_{\mathcal{A}} - \\ &\quad \frac{1}{\|\alpha\|^3} \cdot \underbrace{\left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{H}(\epsilon(\alpha)) \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1}}_{\mathcal{B}} \end{aligned} \quad (28)$$

Puisque le terme \mathcal{A} est indépendant du module $\|\alpha\|$ et que le terme \mathcal{B} tend vers une constante indépendante de $\|\alpha\|$ lorsque

$\|\alpha\| \rightarrow \infty$, alors l'inverse du hessien de L , lorsque le RSB croît, tend vers :

$$\mathbf{H}^{-1}(L) \simeq \frac{1}{\|\alpha\|^2} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} = [\mathbf{H}(E\{L\})]^{-1}. \quad (29)$$

Pour déterminer une approximation de l'algorithme de Newton écrivons l'équation (10) en développant le gradient grâce à l'équation (27) :

$$\Delta \hat{\mathbf{p}}_{k+1} = -[\mathbf{H}(L(\hat{\mathbf{p}}_k))]^{-1} \cdot \left\{ \|\alpha\|^2 \cdot \mathbf{grad}(E\{L(\hat{\mathbf{p}}_k)\}) + \|\alpha\| \cdot \mathbf{grad}(\epsilon(\alpha)) \right\}$$

puis remplaçons le hessien par son approximation à fort RSB (28). Une fois développée, nous obtenons :

$$\begin{aligned} \Delta \hat{\mathbf{p}} &\simeq - \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{grad}(E\{L\}) \\ &+ \underbrace{\frac{1}{\|\alpha\|} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{H}(\epsilon(\alpha)) \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{grad}(E\{L\})}_{\mathcal{U}} \\ &- \underbrace{\frac{1}{\|\alpha\|} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{grad}(\epsilon(\alpha))}_{\mathcal{V}} \\ &+ \underbrace{\frac{1}{\|\alpha\|^2} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{H}(\epsilon(\alpha)) \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{grad}(\epsilon(\alpha))}_{\mathcal{W}} \end{aligned}$$

En observant cette dernière équation on remarque que le terme \mathcal{W} est d'ordre supérieur aux termes \mathcal{U} et \mathcal{V} , il devient donc négligeable lorsque le RSB augmente. Le terme \mathcal{U} est lui aussi d'ordre supérieur au terme \mathcal{V} lorsque le RSB augmente et lorsque l'on est proche de la solution. En effet, proche de la solution et à fort RSB alors

$$\left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{grad}(E\{L\}) \approx \Delta \hat{\mathbf{p}},$$

ainsi \mathcal{U} est peu différent de

$$\mathcal{U} \approx \frac{1}{\|\alpha\|} \cdot \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot \mathbf{H}(\epsilon(\alpha)) \cdot \Delta \hat{\mathbf{p}}$$

qui se trouve être du deuxième ordre. \mathcal{U} peut donc être négligé devant \mathcal{V} .

Finalement, l'algorithme de Newton pour des signaux à fort RSB, se simplifie, près de la solution par la relation suivante :

$$\begin{aligned} \Delta \hat{\mathbf{p}} &\simeq - \left[\mathbf{H}(E\{L\}) \right]^{-1} \\ &\quad \cdot \left(\mathbf{grad}(E\{L\}) + \frac{1}{\|\alpha\|} \cdot \mathbf{grad}(\epsilon(\alpha)) \right) \\ &\simeq - \left[\mathbf{H}(E\{L\}) \right]^{-1} \cdot (\mathbf{grad}(L)) \\ \Delta \hat{\mathbf{p}} &\simeq - [\mathbf{H}(E\{L(\hat{\mathbf{p}})\})]^{-1} \cdot \mathbf{grad}(L(\hat{\mathbf{p}})) \end{aligned}$$

Le fait que cette approximation ne soit valable que proche de la solution et à fort RSB n'a pas de conséquences gênantes puisque, comme nous le verrons, le hessien de l'espérance du critère est par construction défini non-négatif. On pourra ainsi assurer la convergence de l'algorithme vers un minimum qui ne sera, bien sûr, le minimum global que si l'initialisation est suffisamment proche.

8.4. calcul du hessien

Dans ce paragraphe, la notation suivante sera utilisée : $\hat{\mathbf{p}}$ dénote la valeur estimée de la vraie valeur \mathbf{p} . Pour calculer l'expression de $\mathbf{H}(E\{L(\hat{\mathbf{p}})\})$, commençons par déterminer l'espérance de $L(\hat{\mathbf{p}})$. Les équations (9) et (3) nous donnent :

$$\begin{aligned} L(\hat{\mathbf{p}}) &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \\ &+ \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \tilde{\mathbf{N}} \\ &+ \tilde{\mathbf{N}}^H \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \\ &+ \tilde{\mathbf{N}}^H \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \tilde{\mathbf{N}} \end{aligned} \quad (30)$$

Le premier terme de (30) est déterministe, les deuxième et troisième ont une espérance nulle, ainsi :

$$\begin{aligned} E\{L(\hat{\mathbf{p}})\} &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \\ &+ E\{\tilde{\mathbf{N}}^H \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \tilde{\mathbf{N}}\} \\ &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \\ &+ E\left\{ \text{tr} \left\{ \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \tilde{\mathbf{N}} \cdot \tilde{\mathbf{N}}^H \right\} \right\} \\ &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \\ &+ \text{tr} \left\{ \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot E\{\tilde{\mathbf{N}} \cdot \tilde{\mathbf{N}}^H\} \right\} \\ &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha + \sigma^2 \cdot \text{tr} \left\{ \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \right\} \\ E\{L(\hat{\mathbf{p}})\} &= \alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha + (N - M) \cdot \sigma^2. \end{aligned} \quad (31)$$

Posons maintenant $\mathbf{r}(\hat{\mathbf{p}}) = \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha$ alors :

$$\begin{aligned} \mathbf{H}(E\{L(\hat{\mathbf{p}})\})_{i,j} &= 2 \cdot \text{Re} \left\{ \frac{\partial^2 \mathbf{r}^H(\hat{\mathbf{p}})}{\partial \hat{p}_i \cdot \partial \hat{p}_j} \cdot \mathbf{r}(\hat{\mathbf{p}}) + \frac{\partial \mathbf{r}^H(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \frac{\partial \mathbf{r}(\hat{\mathbf{p}})}{\partial \hat{p}_j} \right\}. \end{aligned} \quad (32)$$

En faisant tendre $\hat{\mathbf{p}}$ vers \mathbf{p} alors le vecteur $\mathbf{r}(\hat{\mathbf{p}})$ devient le vecteur nul :

$$\lim_{\hat{\mathbf{p}} \rightarrow \mathbf{p}} (\Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha) = \vec{0}$$

ainsi, le premier terme de l'équation (32) disparaît et il ne reste que le terme des dérivées premières. En utilisant l'équation (23), le terme des dérivées premières devient :

$$2 \cdot \operatorname{Re} \left\{ \frac{\partial \mathbf{r}^H(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \frac{\partial \mathbf{r}(\hat{\mathbf{p}})}{\partial \hat{p}_j} \right\} = 2 \cdot \operatorname{Re} \left\{ \begin{array}{l} \underbrace{\left(\alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \frac{\partial \Psi(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \Psi^{\#}(\hat{\mathbf{p}}) \right)}_{\mathcal{A}} \\ + \underbrace{\left(\alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Psi^{\#H}(\hat{\mathbf{p}}) \cdot \frac{\partial \Psi^H(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \right)}_{\mathcal{B}} \\ \cdot \underbrace{\left(\Psi^{\#H}(\hat{\mathbf{p}}) \cdot \frac{\partial \Psi^H(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \right)}_{\mathcal{C}} \\ + \underbrace{\left(\Pi_{\Psi}^{\perp}(\hat{\mathbf{p}}) \cdot \frac{\partial \Psi(\hat{\mathbf{p}})}{\partial \hat{p}_i} \cdot \Psi^{\#}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p}) \cdot \alpha \right)}_{\mathcal{D}} \end{array} \right\}$$

Pour les mêmes raisons que pour le terme en dérivée seconde, les termes \mathcal{A} et \mathcal{C} sont nuls lorsque $\hat{\mathbf{p}}$ tend vers \mathbf{p} . Les termes \mathcal{B} et \mathcal{D} se simplifient lorsque $\hat{\mathbf{p}}$ tends vers \mathbf{p} puisque :

$$\lim_{\hat{\mathbf{p}} \rightarrow \mathbf{p}} (\Psi^{\#}(\hat{\mathbf{p}}) \cdot \Psi(\mathbf{p})) = \mathbf{I}.$$

Finalement, l'expression du hessien de l'espérance du critère est la suivante :

$$\mathbf{H}_{i,j}(E\{L(\mathbf{p})\}) = 2 \cdot \operatorname{Re} \left\{ \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_j} \cdot \alpha \right\} \quad (33)$$

Les mêmes remarques que pour le calcul du gradient nous permettent de passer à une écriture matricielle. On obtient alors :

$$\mathbf{H}(E\{L(\mathbf{p})\}) = 2 \cdot \operatorname{Re} \left\{ \operatorname{diag} \left(\begin{bmatrix} \alpha^* \\ \alpha^* \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \mathbf{D} \cdot \operatorname{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \right\}$$

8.5. calculs des bornes de Cramer-Rao

8.5.1. calcul de la matrice de Fischer

Le calcul des bornes de Cramer-Rao (BCR) nécessite le calcul de la matrice d'information de Fisher. L'expression de l'élément i, j de la matrice de Fisher $\mathbf{J}_{\mathbf{q}}$ a pour forme :

$$\mathbf{J}_{\mathbf{q}} \, i,j = E \left\{ \frac{\partial \mathcal{L}(\mathbf{q})}{\partial q_i} \cdot \left(\frac{\partial \mathcal{L}(\mathbf{q})}{\partial q_j} \right)^* \right\},$$

où l'on rappelle l'expression de la log-vraisemblance \mathcal{L} :

$$\mathcal{L}(\mathbf{q}) = -T \cdot N \cdot \log(\pi \cdot \sigma^2) - \frac{1}{\sigma^2} \cdot |\bar{\mathbf{X}} - \Psi(\mathbf{p}) \cdot \alpha|^2 \quad (34)$$

On obtient l'élément de la matrice de Fisher relatif à la puissance du bruit en dérivant (34) par rapport à σ^2 , on obtient de manière classique :

$$J_{\sigma^2} = \frac{T \cdot N}{\sigma^4}. \quad (35)$$

Les blocs relatifs à la puissance du bruit et aux autres jeux de paramètres se calculent de manière évidente. On obtient les expressions suivantes :

$$\mathbf{J}_{\sigma^2 \beta} = \mathbf{0}_{1 \times N}, \quad \mathbf{J}_{\sigma^2 \alpha} = \mathbf{0}_{1 \times N} \text{ et } \mathbf{J}_{\sigma^2 \alpha^*} = \mathbf{0}_{1 \times 2N}. \quad (36)$$

Le terme général de $[\mathbf{J}_{\alpha}]_{i,j}$ s'obtient à partir de la dérivée de (34) par rapport à α_i (et α_j) :

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = -\frac{1}{\sigma^2} \cdot (\alpha^H \cdot \Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}) \cdot \mathbf{e}_i - \bar{\mathbf{X}} \cdot \Psi(\mathbf{p}) \cdot \mathbf{e}_i) \quad (37)$$

il vient naturellement, en prenant l'espérance :

$$[\mathbf{J}_{\alpha}]_{i,j} = \frac{1}{\sigma^2} \cdot \mathbf{e}_j^T \cdot \Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}) \cdot \mathbf{e}_i$$

Cette dernière expression donne sous forme matricielle :

$$\mathbf{J}_{\alpha} = \frac{1}{\sigma^2} \cdot (\Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}))^T = \frac{1}{\sigma^2} \cdot (\Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}))^* \quad (38)$$

Un calcul similaire donne l'expression de \mathbf{J}_{α^*} :

$$\mathbf{J}_{\alpha^*} = \frac{1}{\sigma^2} \cdot \Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p}) \quad (39)$$

Pour les blocs extra-diagonaux $\mathbf{J}_{\alpha \alpha^*}$ les calculs sont là encore identiques et on obtient le résultat suivant :

$$\mathbf{J}_{\alpha \alpha^*} = \mathbf{0}_{N \times N}. \quad (40)$$

Pour calculer les termes relatifs aux paramètres α et aux paramètres de localisation \mathbf{p} , dérivons l'équation (34) par rapport à p_j :

$$\frac{\partial \mathcal{L}}{\partial p_j} = -\frac{1}{\sigma^2} \cdot \left(\alpha^H \cdot \frac{\partial}{\partial p_j} (\Psi^H(\mathbf{p}) \cdot \Psi(\mathbf{p})) \cdot \alpha - \bar{\mathbf{X}}^H \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_j} \cdot \alpha - \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_j} \cdot \bar{\mathbf{X}} \right).$$

Nous obtenons, grâce à cette dernière expression et à l'expression (37), le terme i, j de $\mathbf{J}_{\alpha \mathbf{p}}$:

$$\begin{aligned} [\mathbf{J}_{\alpha \mathbf{p}}]_{i,j} &= -E \left\{ \frac{\partial \mathcal{L}}{\partial \alpha_i} \cdot \left(\frac{\partial \mathcal{L}}{\partial p_j} \right)^* \right\} \\ &= \frac{1}{\sigma^2} \cdot \left(\alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_j} \cdot \Psi(\mathbf{p}) \cdot \mathbf{e}_i \right), \end{aligned}$$

qui s'écrit sous forme matricielle, simplement en prenant garde aux dimensions, de la manière suivante :

$$\mathbf{J}_{\alpha\mathbf{p}} = \frac{1}{\sigma^2} \cdot \left(\Psi^H \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \right)^* \quad (41)$$

Le calcul est identique pour le terme :

$$\mathbf{J}_{\alpha^*\mathbf{p}} = \frac{1}{\sigma^2} \cdot \Psi^H \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \quad (42)$$

Le bloc de la matrice de Fisher relatif aux paramètres d'intérêts \mathbf{p} s'obtient suivant le même procédé. L'élément i, j du bloc $\mathbf{J}_{\mathbf{p}}$ s'écrit :

$$[\mathbf{J}_{\mathbf{p}}]_{i,j} = \frac{2}{\sigma^2} \cdot \text{Re} \left(\alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_j} \cdot \alpha \right).$$

Pour écrire cette matrice sous forme matricielle, remarquons que, de part la forme de $\Psi(\mathbf{p})$, la matrice dérivée ne possède que la $i^{\text{ème}}$ colonne de non nulle. On a alors l'identité suivante :

$$\frac{\partial \Psi(\mathbf{p})}{\partial p_i} \cdot \alpha = \frac{\partial \Psi(p_i)}{\partial p_i} \cdot \alpha_i.$$

Ainsi, l'élément i, j de la matrice $\mathbf{J}_{\mathbf{p}}$, a pour expression :

$$\mathbf{J}_{\mathbf{p}}_{i,j} = \frac{2}{\sigma^2} \cdot \text{Re} \left(\alpha_i^* \cdot \frac{\partial \Psi^H(p_i)}{\partial p_i} \cdot \frac{\partial \Psi(p_j)}{\partial p_j} \cdot \alpha_j \right)$$

qui donne simplement sous forme matricielle :

$$\mathbf{J}_{\mathbf{p}} = \frac{2}{\sigma^2} \cdot \text{Re} \left(\text{diag} \left(\begin{bmatrix} \alpha^* \\ \alpha^* \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \right) \quad (43)$$

8.5.2. expression compacte de BCR(p)

La matrice de Fisher globale à donc la forme suivante :

$$\mathbf{J}_{\mathbf{q}} = \begin{bmatrix} J_{\sigma^2} & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times 2N} \\ \mathbf{0}_{N \times 1} & \mathbf{J}_{\alpha} & \mathbf{0}_{N \times N} & \mathbf{J}_{\alpha\mathbf{p}} \\ \mathbf{0}_{N \times 1} & \mathbf{0}_{N \times N} & \mathbf{J}_{\alpha^*} & \mathbf{J}_{\alpha^*\mathbf{p}} \\ \mathbf{0}_{2N \times 1} & [\mathbf{J}_{\alpha\mathbf{p}}]^H & [\mathbf{J}_{\alpha^*\mathbf{p}}]^H & \mathbf{J}_{\mathbf{p}} \end{bmatrix}.$$

Puisque : $\mathbf{J}_{\alpha\mathbf{p}} = (\mathbf{J}_{\alpha^*\mathbf{p}})^*$ et $\mathbf{J}_{\alpha} = (\mathbf{J}_{\alpha^*})^*$, alors, à l'aide d'une formule d'inversion par blocs, on obtient :

$$\begin{aligned} [\mathbf{J}_{\mathbf{q}}^{-1}]_{\mathbf{p}} &= [\mathbf{J}_{\mathbf{p}} - [(\mathbf{J}_{\alpha^*\mathbf{p}})^* \quad \mathbf{J}_{\alpha^*\mathbf{p}}^H]] \\ &\quad \cdot \left[\begin{array}{c} (\mathbf{J}_{\alpha^*}^{-1})^* \quad \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} \quad \mathbf{J}_{\alpha^*}^{-1} \end{array} \right] \cdot \left[\begin{array}{c} (\mathbf{J}_{\alpha^*\mathbf{p}})^* \\ \mathbf{J}_{\alpha^*\mathbf{p}} \end{array} \right]^{-1} \\ &= [\mathbf{J}_{\mathbf{p}} - 2 \cdot \text{Re} \{ \mathbf{J}_{\alpha^*\mathbf{p}}^H \cdot \mathbf{J}_{\alpha^*}^{-1} \cdot \mathbf{J}_{\alpha^*\mathbf{p}} \}]^{-1} \end{aligned}$$

La borne de Cramer-Rao relative aux paramètres de localisation \mathbf{p} s'écrit finalement de manière compacte :

$$\begin{aligned} \text{BCR}(\mathbf{p}) &= [\mathbf{J}_{\mathbf{q}}^{-1}]_{\mathbf{p}} \\ &= \frac{\sigma^2}{2} \cdot \left[\text{Re} \left\{ \text{diag} \left(\begin{bmatrix} \alpha^* \\ \alpha^* \end{bmatrix} \right) \cdot \mathbf{D}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \mathbf{D} \cdot \text{diag} \left(\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} \right) \right\} \right]^{-1}. \end{aligned}$$

8.6. calcul de la variance théorique de l'estimateur PML

La matrice de covariance des estimations est donnée par l'équation (20). Le développement limité réalisé au paragraphe §8.3 donne l'expression (28) qui peut se mettre sous la forme :

$$\mathbf{H}^{-1}(L) \simeq \frac{1}{\|\alpha\|^2} \cdot \left[\mathbf{H} \left(\underline{E}\{L\} \right) \right]^{-1} - \frac{1}{\|\alpha\|^3} \cdot \epsilon_{\mathbf{H}}(\alpha) \quad (44)$$

où la matrice $\epsilon_{\mathbf{H}}(\alpha)$ tend vers une matrice constante lorsque $\|\alpha\| \rightarrow \infty$. De manière similaire, étudions le terme $\mathbf{grad}(L(\mathbf{p})) \cdot \mathbf{grad}^T(L(\mathbf{p}))$. L'élément i, j du gradient donné par l'équation (24) permet d'écrire :

$$\begin{aligned} [\mathbf{grad} \cdot \mathbf{grad}^T]_{i,j} &= 2 \cdot \text{Re} \left\{ \tilde{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \tilde{\mathbf{X}} \right\} \\ &\quad + 2 \cdot \text{Re} \left\{ \tilde{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_j} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \tilde{\mathbf{X}} \right\} \end{aligned}$$

où l'on rappelle que $\tilde{\mathbf{X}} = \Psi \cdot \alpha + \tilde{\mathbf{N}}$. De la même manière que pour le hessien, après développement et normalisation par rapport à $\|\alpha\|$, la matrice se met sous la forme :

$$\mathbf{grad} \cdot \mathbf{grad}^H = \|\alpha\|^2 \cdot \underline{\nabla \nabla}_1 + \|\alpha\| \cdot \underline{\nabla \nabla}_2 + \underline{\nabla \nabla}_3 \quad (45)$$

où l'on explicitera que la matrice :

$$\begin{aligned} [\underline{\nabla \nabla}_1]_{i,j} &= 2 \cdot \text{Re} \\ &\quad \left\{ \begin{array}{c} \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \tilde{\mathbf{N}} \cdot \tilde{\mathbf{N}}^H \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_j} \cdot \alpha + \\ \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \tilde{\mathbf{N}} \cdot \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_j} \cdot \Pi_{\Psi}^{\perp}(\mathbf{p}) \cdot \tilde{\mathbf{N}} \end{array} \right\} \quad (46) \end{aligned}$$

Finalement, l'expression (20) de la matrice de covariance s'écrit à l'aide des équation (44) et (45) :

$$\begin{aligned} \text{var}(\hat{\mathbf{p}}) &\simeq E \left\{ \begin{array}{c} \left(\frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} - \frac{1}{\|\alpha\|^3} \cdot \epsilon_{\mathbf{H}}(\alpha) \right) \\ \cdot \left(\|\alpha\|^2 \cdot \underline{\nabla \nabla}_1 + \|\alpha\| \cdot \underline{\nabla \nabla}_2 + \underline{\nabla \nabla}_3 \right) \\ \cdot \left(\frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} - \frac{1}{\|\alpha\|^3} \cdot \epsilon_{\mathbf{H}}(\alpha) \right) \end{array} \right\} \\ &\simeq E \left\{ \frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} \cdot \underline{\nabla \nabla}_1 \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} \right. \\ &\quad \left. + \frac{1}{\|\alpha\|^3} \cdot (\dots) + \frac{1}{\|\alpha\|^4} \cdot (\dots) + \dots \right\}. \end{aligned}$$

Seul le terme d'ordre le plus élevé nous intéresse ici, il reste alors :

$$\begin{aligned} \text{var}(\hat{\mathbf{p}}) &\simeq \frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} \cdot E \{ \underline{\nabla \nabla}_1 \} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} \\ &= \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1} \cdot E \{ \underline{\nabla \nabla}_1 \} \cdot \underline{\mathbf{H}} \underline{\mathbf{E}} L^{-1}. \end{aligned}$$

L'expression (46) nous donne immédiatement :

$$[E\{\nabla\nabla_1\}]_{i,j} = \sigma^2 \cdot 2 \cdot \text{Re} \left\{ \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^\perp(\mathbf{p}) \cdot \frac{\partial \Psi(\mathbf{p})}{\partial p_j} \cdot \alpha \right\}$$

qui avec l'expression (33) permet de trouver :

$$\text{var}(\hat{\mathbf{p}}) \simeq \sigma^2 \cdot [\mathbf{H}(E\{L(\mathbf{p})\})]^{-1}$$

8.7. remarque sur le biais de l'estimateur PML

L'expression du biais est la suivante :

$$E\{\hat{\mathbf{p}}\} = E\left\{[\mathbf{H}(L(\mathbf{p}))]^{-1} \cdot \text{grad}(L(\mathbf{p}))\right\}. \quad (47)$$

L'équation (24) permet d'écrire :

$$\begin{aligned} \frac{\partial L(\mathbf{p})}{\partial p_i} &= -2 \cdot \text{Re} \left\{ \tilde{\mathbf{X}}^H \cdot \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^\perp(\mathbf{p}) \cdot \tilde{\mathbf{X}} \right\} \\ &= -2 \cdot \text{Re} \left\{ \alpha^H \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^\perp(\mathbf{p}) \cdot \tilde{\mathbf{N}} \right\} \\ &\quad -2 \cdot \text{Re} \left\{ \tilde{\mathbf{N}}^H \cdot \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^\perp(\mathbf{p}) \cdot \tilde{\mathbf{N}} \right\} \\ &= [\nabla_1]_i + [\nabla_2]_i, \end{aligned}$$

ainsi en normalisant, on obtient :

$$\text{grad} = \|\alpha\| \cdot \nabla_1 + \nabla_2. \quad (48)$$

Les expressions (47),(44) et (48) permettent alors d'écrire :

$$\begin{aligned} E\{\hat{\mathbf{p}}\} &\simeq E\left\{\left(\frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}}\underline{\mathbf{E}}\underline{\mathbf{L}}^{-1} - \frac{1}{\|\alpha\|^3} \cdot \epsilon_{\mathbf{H}}(\alpha)\right) \cdot (\|\alpha\| \cdot \nabla_1 + \nabla_2)\right\} \\ &\simeq E\left\{\underbrace{\frac{1}{\|\alpha\|} \cdot \underline{\mathbf{H}}\underline{\mathbf{E}}\underline{\mathbf{L}}^{-1} \cdot \nabla_1}_{\mathcal{A}} + \underbrace{\frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}}\underline{\mathbf{E}}\underline{\mathbf{L}}^{-1} \cdot \nabla_2}_{\mathcal{B}} - \underbrace{\frac{1}{\|\alpha\|^2} \cdot \epsilon_{\mathbf{H}}(\alpha) \cdot \nabla_1 + \frac{1}{\|\alpha\|^3} \cdot (\dots)}_{\mathcal{C}}\right\}. \end{aligned}$$

Etudions les termes \mathcal{A} , \mathcal{B} et \mathcal{C} . Le terme \mathcal{A} a une espérance nulle, en effet :

$$E\{\mathcal{A}\} = \frac{1}{\|\alpha\|} \cdot \underline{\mathbf{H}}\underline{\mathbf{E}}\underline{\mathbf{L}}^{-1} \cdot E\{\nabla_1\},$$

et il est aisé de voir que $E\{\nabla_1\} = 0$, alors $E\{\mathcal{A}\} = 0$. De même, le terme \mathcal{B} s'exprime :

$$E\{\mathcal{B}\} = \frac{1}{\|\alpha\|^2} \cdot \underline{\mathbf{H}}\underline{\mathbf{E}}\underline{\mathbf{L}}^{-1} \cdot E\{\nabla_2\}.$$

Puisque

$$E\{[\nabla_2]_i\} = -\sigma^2 \cdot 2 \cdot \text{Re} \left\{ \text{tr} \left\{ \Psi^{\#H}(\mathbf{p}) \cdot \frac{\partial \Psi^H(\mathbf{p})}{\partial p_i} \cdot \Pi_{\Psi}^\perp(\mathbf{p}) \right\} \right\},$$

et sachant que $\Psi^{\#}(\mathbf{p}) \cdot \Pi_{\Psi}^\perp(\mathbf{p}) = \mathbf{0}$, alors $E\{\mathcal{B}\} = 0$. Le terme \mathcal{C} , qui a pour expression :

$$E\{\mathcal{C}\} = \frac{1}{\|\alpha\|^2} \cdot E\{\epsilon_{\mathbf{H}}(\alpha) \cdot \nabla_1\},$$

est considéré *a priori* non nul tant qu'il n'aura pas été étudié plus précisément. Finalement, le biais de l'estimateur PML est donc nul en $\frac{1}{RSB^{1/2}}$, le premier terme non nul étant *a priori* en $\frac{1}{RSB}$.

BIBLIOGRAPHIE

- [Bertaux98a] N. Bertaux et P. Larzabal, «Maximum de Vraisemblance Paramétrée (PML)» Rapport interne LESiR ENS-Cachan, Janvier 1998.
- [Bertaux98b] N. Bertaux, P. Larzabal et P. Forster, «Méthodes du Maximum de Vraisemblance et Subspace Fitting», Partie 1, Chapitre 4, «Méthodes à Haute Résolution pour le Traitement d'Antenne et l'Analyse Spectrale». Ouvrage collectif publié sous l'égide du GDR-PRC-ISIS du CNRS, Edition Hermes Février 1998
- [Bienvenu83] G. Bienvenu et L. Kopp, «Optimality of high resolution array processing using the eigensystem approach», IEEE Trans. ASSP, Vol.31, No.5, pp.1235-1247. October 1983.
- [Bohme84] J.F. Bohme, «Estimation of source parameters by maximum likelihood and non linear regression», ICASSP 84, pp.731-734, 1984.
- [Bohme86] J.F. Bohme, «Estimation of spectral parameters of correlated signals in wavefields.» Signal Processing, Vol. 10, pp.329-337, 1986.
- [Bos94] A. van den Bos, «A Cramer-Rao Lower Bound for Complex Parameters», IEEE Trans. SP, Vol.42, No. 42, pp. 2859, October 1994.
- [Cedervall97] M. Cedervall and R. Moses, «Efficient maximum likelihood DOA estimation for signals with known waveforms in presence of multipath», IEEE Trans.SP, Vol.45, No.3, pp.808-810, March 1997.
- [Figueiredo83] R. Figueiredo and A. Gerber, «Separation of superimposed signals by a cross-correlation method», IEEE Trans.ASSP, Vol.31, pp.1084-1089, October 1983.
- [Goncalves97] D. Concalves et P. Gounon, «Déconvolution de sources large bande et Traitement actif à haute résolution», Colloque GRETSI 1997, pp.71-74, Grenoble, Septembre 1997.
- [Gounon94] P. Gounon, «Analyse spatio-temporelle haute résolution à l'aide d'une antenne active» Traitement du Signal, Vol.11, No.5, pp.351-360, 1994.
- [Graffoulière97] P. Graffoulière, «Méthodes actives spatio-temporelles large bande. Techniques et performances. Applications en sonar», Thèse de doctorat de l'INPG, Mars 1997.
- [Grouffaud97] J. Grouffaud, «Identification spatio-temporelle de canaux de propagation à trajets multiples» Thèse de doctorat de l'Ecole Normale Supérieure de Cachan, Juin 1997
- [Kumaresan83] R. Kumaresan and D. Tufts, «Estimation angle-of-arrival of multiple plane waves», IEEE Trans.AES, Vol.19, pp.134-139, January 1983.
- [Li93] J. Li and R.T. Compton, «Maximum Likelihood Angle Estimation for Signals with Known Waveforms». IEEE trans.SP. Vol.41, No. 9, pp.2850-2862, September 1993.
- [Li95] J. Li, P. Stoica, M. Viberg and T. Kailath, «Computationally efficient angle estimation for signals with known waveforms», IEEE Trans.SP, Vol.43, pp.2154-2163, September 1995.

- [Lichtenstein68] M. Lichtenstein and T. Young, «*The resolution of closely spaced signals*», IEEE Trans.IT, Vol.14, pp.288-293, 1968.
- [Manabe92] T. Manabe and H. Takai, «*Superresolution of multipath delay profile measured by pn correlation method.*», IEEE Trans.SP, Vol.40, No.5, pp.177-183, May 1992.
- [Manickam94] T. Manickam, R. Vaccaro and D. Tufts, «*A least squares algorithm for multipath time delay estimation*», IEEE Trans.SP, Vol.42, No.11, pp.3223-3229, November 1994.
- [Ottersten93] B. Ottersten, M. Viberg, P. Stoica and A. Nehorai, «*Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing*», Chapter 4 in «*Radar Array Processing*» Simon Haykin (ed.), Springer-Verlag.
- [Pallas86] M.A. Pallas and G. Jourdain, «*Joint estimation of close delays and application to underwater acoustics*», Proceeding of Eusipco-86, Vol.2, pp.973-976, September 1986.
- [Reddi79] S. Reddi, «*Multiple source location. A digital approach*», IEEE Trans. AES, pp.95-105, 1979.
- [Roy89] R. Roy and T. Kailath, «*ESPRIT-Estimation of signal parameters via rotational invariance techniques*», IEEE Trans.ASSP, Vol.37, No.7, July 1989.
- [Schmidt81] R.O. Schmidt, «*A signal subspace approach to multiple emitter location and spectral estimation*» PhD thesis, Stanford Univ, Stanford, CA, November 1981
- [Stoica89] P. Stoica and A. Nehorai, «*MUSIC, maximum likelihood and Cramer-rao bound*», IEEE Trans ASSP, Vol.37, No.12, pp.720-741, May 1989.
- [Stoica90] P. Stoica and A. Nehorai, «*MUSIC, Maximum likelihood, and Cramer-Rao bound : Further results and compararisons*», IEEE Trans.ASSP, Vol.38, No.12, pp.2140-2150, December 1990.
- [Viberg91] M. Viberg and Ottersten, «*Sensor array processing based on subspace fitting*», IEEE Trans.ASSP, Vol.39, pp.1110-1121, May 1991.
- [Viberg97] M. Viberg, P. Stoica and B. Ottersten, «*Maximum likelihood array processing in spatially correlated noise fields using parameterized signals*» IEEE Trans.SP, Vol.45, No.10, pp.2477-2484, October 1997., IEEE Trans.SP, Vol.45, No.4, pp.996-1004, April 1997.
- [Wax97] M. Wax and A. Leshem, «*Joint Estimation of Delays and Directions of Arrival of Multiple Reflections of a Known Signal*» IEEE Trans.SP, Vol.45, No.10, pp.2477-2484, October 1997.

Manuscrit reçu le 6 avril 1998.

LES AUTEURS

Nicolas BERTAUX

Nicolas Bertaux est né à Reims en 1970. Ancien élève de l'Ecole Normale Supérieure de Cachan, il fait son monitorat à l'ENS de cachan et poursuit sa troisième année de thèse au Laboratoire d'Electricité, Signaux et Robotique (LESiR) de l'ENS de Cachan, en collaboration avec Thomson-CSF AirSys situé à Bagneux.

Claude ADNET

Claude Adnet a soutenu, en 1990, une thèse en traitement du signal au CEPHAG (INP Grenoble). Depuis 1991, il est ingénieur en recherche et développement à THOMSON CSF/AIRSYS dans le service algorithmes et nouveaux concepts.

Pascal LARZABAL

Pascal Larzabal est né en 1962 à Saint Jean de Luz au pays basque. Ancien élève de l'Ecole Normale Supérieure de Cachan, il est actuellement Maître de Conférences à l'IUT de Cachan où il enseigne l'électronique. Il est responsable de l'Equipe Traitement d'Antenne du Laboratoire d'Electricité, Signaux et Robotique (LESiR) de l'ENS Cachan et son thème de recherche concerne l'estimation en traitement d'antenne pour l'identification d'un front d'onde.

Eric CHAUMETTE

Eric Chaumette est né en 1965 à Chartres. Depuis 1990, il est ingénieur en recherche et développement à THOMSON CSF/AIRSYS dans le service algorithmes et nouveaux concepts. Il est membre du Collège Scientifique de Thomson depuis 1997.