

Estimation des Directions D'Arrivées incorporant un *a priori*: Algorithmes et Variances Théoriques

Guillaume BOULEUX¹, Rémy BOYER²

¹Laboratoire d'Analyse des Signaux et des Processus Industriels (LASPI)
20 Avenue de Paris, 42300 Roanne, France

²Laboratoire des Signaux et Systèmes (LSS-SUPELEC)
3 rue Joliot-Curie, 91190 Gif-Sur-Yvette, France
guillaume.bouleux@univ-st-etienne.fr, remy.boyer@lss.supelec.fr

Résumé – Dans le contexte de l'estimation des Directions D'Arrivées (DDA), on peut parfois considérer que nous connaissons *a priori* (de manière exacte ou estimée) un ensemble de $M - S$ DDA parmi un total de M . Dans ce contexte, des schémas d'estimation ont été proposés afin de tenir compte de cette connaissance dans le but d'améliorer la localisation des M sources d'intérêt. Ces approches se basent sur la déflation orthogonale du sous-espace signal. Dans [10], nous avons établi et analysé la Borne de Cramer-Rao (BCR) correspondante à ce type de modèle et nous avons montré qu'une connaissance *a priori* d'un ensemble de DDA est bénéfique uniquement pour des sources corrélées et de DDA proches. En particulier, dans le cas de sources non corrélées de DDA proches, les approches basées sur la déflation orthogonale n'améliorent pas l'estimation des DDA d'intérêt. Une solution possible pour résoudre ce problème est d'exploiter une déflation oblique. Selon ce principe, nous proposons deux algorithmes de type MinNorm dont nous caractérisons les performances théoriques.

Abstract – In the context of the Direction-Of-Arrival (DOA) estimation problem, we can sometimes assume that we have *a priori* knowledge (exact or estimated) of a subset of $M - S$ DOA among a total of M . In that situation, some works have proposed to tacking into account of this knowledge to improve the localization of the M unknown sources. The key idea is based on an orthogonal deflation of the signal subspace. In [10], we have derived and analyzed the CRB associated to this model and we have shown that the prior-knowledge of a subset of the DOA is beneficial only for coherent or highly correlated sources associated to closely-spaced DOA. In particular, for uncorrelated sources with closely-spaced DOA, the orthogonal deflation cannot help the estimation of the unknown DOA. Finally, we suggest that a possible solution could be to use oblique projectors. In this paper, we follow this line and we propose two algorithms based on the MinNorm principle and we analyse their theoretical performances.

1 Introduction

L'estimation des DDA de sources bandes étroites est un problème important dans de nombreuses applications et a reçu une attention particulière depuis ces dernières années. En pratique, nous avons parfois la connaissance *a priori* (de manière exacte ou estimée) d'un ensemble de $M - S$ DDA parmi un total de M comme par exemple en traitement RADAR [8], dans le domaine de l'analyse de signaux biomédicaux [4] ou encore en électromagnétique [11]. Plusieurs travaux ont proposé d'intégrer cette connaissance *a priori* dans des algorithmes d'estimation. L'idée clef de ces approches est de projeter les observations bruitées sur le sous-espace engendré par les DDA connues. Ces approches sont donc fortement reliées au problème de déflation du sous-espace signal. Dans [10], les auteurs ont établi et analysé la Borne de Cramér-Rao (BCR) associée à un modèle de signal basé sur le principe de déflation orthogonale. En particulier, les auteurs montrent que la connaissance *a priori* d'un ensemble de DDA est bénéfique uniquement lorsque les sources sont fortement corrélées et de DDA proches. Lorsque les sources sont décorréelées et de DDA proches, la déflation orthogonale n'améliore pas l'estimation des DDA d'intérêt. Afin de résoudre ce problème, nous proposons dans ce travail de récrire le pro-

blème que l'on se pose dans le contexte de projections obliques. Nous proposons donc deux algorithmes de type MinNorm [7] intégrant une connaissance car [9] montre que ce dernier a une meilleur résolution que l'algorithme MUSIC pour un réseau d'antennes linéaires équidistantes et possède un coût de calcul inférieur [3]. Par conséquent, nous proposons deux algorithmes basés sur l'algorithme MinNorm intégrant une connaissance *a priori*, dénommés Prior-MINORM (P-MinNorm) et Weighted P-MinNorm (WP-MinNorm). De plus, nous donnons une nouvelle formulation de l'algorithme Constrained-MinNorm (C-MINORM) de DeGroat *et al.* [8] afin d'établir les variances théoriques de chacun des algorithmes traités dans ce travail. Enfin, nous comparons les différents algorithmes proposés au C-MinNorm de manière théorique et pratique.

2 Représentation Matricielle du problème d'estimation des DDA

2.1 Modèle MIMO (Multi-Input Multi Output)

Supposons qu'il existe M ondes planes à bandes étroites arrivant simultanément sur L capteurs linéairement et uni-

formément répartis. La t -ème observation est donnée par

$$x(t) = Z\alpha(t) + n(t) = [p(\theta_1) \ \dots \ p(\theta_M)] \alpha(t) + n(t) \quad (1)$$

où $x(t) = [x_1(t) \ \dots \ x_L(t)]^T$, $\alpha(t) = [\alpha_1(t) \ \dots \ \alpha_M(t)]^T$ et $x_\ell(t)$ est l'observation sur le ℓ -ème capteur, $\alpha_m(t)$ est la m -ème source, $n(t) = [n_1(t) \ \dots \ n_L(t)]^T$ est un bruit additif gaussien centré, temporellement et spatialement blanc de variance σ^2 . Enfin, $p(\theta)$ est le vecteur de test paramétré par la DDA, θ , et défini selon

$$p(\theta) = \frac{1}{\sqrt{L}} [1 \ e^{-2i\pi(\Delta/c) \sin(\theta)} \ \dots \ e^{-2i\pi(\Delta/c) \sin(\theta)(L-1)}]^T, \quad (2)$$

dans lequel Δ est la distance consécutive entre deux capteurs et c est la longueur d'onde. Le nombre de source M est supposé connu ou préalablement estimé [3]. On supposera que toutes les DDA sont distinctes et que le rang de Z est M (nous supposons ici que $M \leq L$). Finalement, le modèle final pour T observations est

$$X = [x(1) \ \dots \ x(T)] = Z\Lambda + N \quad (3)$$

$$\text{où } \Lambda = [\alpha(1) \ \dots \ \alpha(T)] \text{ et } N = [n(1) \ \dots \ n(T)].$$

2.2 Partitionnement du sous-espace signal

Supposons connaître $M - S$ DDA sur un total de M . Sans perte de généralité, la matrice de Vandermonde Z peut être partitionnée en $Z = [\hat{A} \ B]$, où la matrice \hat{A} de taille $L \times S$ est la matrice composée par les S DDA d'intérêt et B contient les $M - S$ DDA connues *a priori*. Comme $S \leq M$, la matrice A (resp. B) est de rang S (resp. rang $M - S$). Nous nommons $\mathcal{R}(A)$ le sous-espace signal déflaté puisque sa dimension est $M - S$ qui est inférieure à la dimension du sous-espace signal $\mathcal{R}(Z)$. Nous avons alors $\mathcal{R}(A) \subseteq \mathcal{R}(Z) = \mathcal{R}(A) \oplus \mathcal{R}(B)$. De plus, nous supposons par la suite que les sources de DDA connues et d'intérêt sont décorréelées. Par conséquent, la covariance spatiale est bloc-diagonale selon

$$R_X = \mathcal{E}(XX^H) = ZR_\Lambda Z^H = R_A + R_B \quad (4)$$

où $R_A = AR_{\Lambda_A}A^H$, $R_B = BR_{\Lambda_B}B^H$ et $\mathcal{E}(\cdot)$ représente l'espérance mathématique.

3 Algorithmes WP-MinNorm, P-MinNorm et C-MinNorm

Proposition 1 Les critères d'optimisation des algorithmes WP-MinNorm, P-MinNorm et C-MinNorm sont :

$$\arg \min_{\theta} |f_{\text{WP}}(\theta)|^2 \quad \text{où } f_{\text{WP}}(\theta) = 1 - e_1^T E_{(\hat{A} \ B)} p(\theta) \quad (5)$$

$$\arg \min_{\theta} |f_{\text{P}}(\theta)|^2 \quad \text{où } f_{\text{P}}(\theta) = e_1^T P_{\hat{A}}^\perp p(\theta) \quad (6)$$

$$\arg \min_{\theta} |f_{\text{C}}(\theta)|^2 \quad \text{où } f_{\text{C}}(\theta) = e_1^T P_{\hat{Z}}^\perp p(\theta) \quad (7)$$

où $E_{(\hat{A} \ B)}$ est le projecteur oblique [2] sur sous-espace engendré par les DDA d'intérêt, soit $\mathcal{R}(\hat{A})$, le long du sous-espace engendré par les DDA connues, soit $\mathcal{R}(B)$, $P_{\hat{Z}}^\perp =$

$I - ZZ^\dagger$, $P_{\hat{A}}^\perp = I - AA^\dagger$ où \dagger dénote la pseudo inverse de Moore-Penrose et e_1 est la première colonne de la matrice identité.

Preuve : Les trois critères de la proposition 1 découlent de la résolution au sens des moindres carrés pondérés selon le paramètre α . Plus précisément, les critères WP-MinNorm donné en (5) et P-MinNorm donné en (6) sont associés à $\arg \min_{\theta, \alpha} \|p(\theta) - E_{(\hat{A} \ B)} \hat{Z} \alpha\|_W^2$ et le critère C-MinNorm donné en (7) à $\arg \min_{\theta, \alpha} \|p(\theta) - \hat{Z} \alpha\|_W^2$, où $\|\cdot\|_W^2$ est la norme de Frobenius pondérée par $W = e_1 e_1^T$. ■

4 Version polynomiale des algorithmes d'estimation et analyse de performance

La recherche des DDA d'intérêt en parcourant l'ensemble des valeurs possibles du paramètre θ appartenant à une grille discrète est coûteuse. On propose donc une méthode inspirée du root-MinNorm [3] afin de décroître la complexité de la méthode. Dans cette optique, le théorème suivant présente une version polynomiale des algorithmes présentés dans ce papier et met en évidence le lien entre les algorithmes C-MinNorm et WP-MinNorm.

Théorème 1 Soit $z = e^{-2i\pi(\Delta/c) \sin(\theta)}$. Les S racines du polynôme $f_{\text{WP}}(z) = f_{\text{C}}(z) + f_{\text{COR}}(z)$, où $f_{\text{COR}}(z) = e_1^T E_{(B \ \hat{A})} p(z)$ sont les DDA d'intérêt, et $f_{\text{C}}(z)$ est le polynôme associé à l'algorithme C-MinNorm défini par (7).

Preuve : La propriété $I - E_{(\hat{A} \ B)} = P_{\hat{Z}}^\perp + E_{(B \ \hat{A})}$ nous permet d'obtenir l'expression de $f_{\text{COR}}(z)$. Nous introduisons maintenant deux polynômes définis comme suit :

$$f_{\text{C}}(z) = \prod_{m=1}^S (z - z_m) \prod_{m=S+1}^M (z - z_m) \prod_{m=M+1}^L (z - z'_m) \quad (8)$$

où z_m est associée à la m -th DDA et z'_m sont les racines supplémentaires de $f_{\text{C}}(z)$, de plus :

$$f_{\text{COR}}(z) = \prod_{m=1}^S (z - z_m) \prod_{m=S+1}^L (z - z''_m) \quad (9)$$

où z''_m sont les racines supplémentaires de $f_{\text{COR}}(z)$ et donc $f_{\text{WP}}(z)$ se factorise de la manière suivante :

$$\mathcal{C}(z) = Q(z) \prod_{m=1}^S (z - z_m) \quad (10)$$

où $Q(z) = \prod_{m=S+1}^M (z - z_m) \prod_{m=M+1}^L (z - z'_m) + \prod_{m=S+1}^L (z - z''_m)$. Clairement, $Q(z)$ n'a pas de racine triviale, *i.e.*, les DDA connues ou inconnues ne sont pas des racines de $Q(z)$. Par suite et d'après (10), les racines de $f_{\text{WP}}(z)$ sont uniquement les DDA inconnues : $\{\theta_1, \dots, \theta_S\}$. ■

Finalement, la version polynomiale de l'algorithme P-MinNorm est basée sur la recherche des racines du polynôme $f_{\text{P}}(z) = e_1^T P_{\hat{A}}^\perp p(z)$.

4.1 Variances théoriques des algorithmes proposés

Nous étudions le cas où la matrice de covariance n'est pas bien estimée du fait que l'on considère un nombre fini d'observations. Nous définissons alors $\hat{R} = R + \Delta R$ où $R = R_X + \sigma^2 I$ et $\hat{z} = z + \Delta z$. Cette hypothèse permet de déduire les expressions des projecteurs oblique et orthogonaux estimés en fonction de la matrice de covariance estimée. Nous caractérisons ainsi la variance d'estimation théorique des DDA.

L'utilisation du projecteur oblique nécessite d'estimer le sous-espace inconnu $\mathcal{R}(A)$ en fonction de la matrice de covariance R . Une erreur commise sur R induit une erreur sur le sous-espace $\mathcal{R}(A)$. Nous considérons donc le modèle $\hat{A} = A + \Delta A$ comme l'erreur commise sur le sous-espace $\mathcal{R}(A)$.

Par manque de place, nous présentons seulement dans les "grandes lignes" le calcul des variances théoriques. Nous donnons ici l'expression de $E_{(\hat{A} B)}$ en fonction de l'erreur sur $\mathcal{R}(A)$ et de l'erreur sur le projecteur oblique $\delta E_{(\hat{A} B)}$. Cette expression s'écrit :

$$E_{(\hat{A} B)} = E_{(A B)} + (P_Z^\perp + E_{(B A)})\Delta A(A^H P_B^\perp A)^{-1}A^H P_B^\perp + A(A^H P_B^\perp A)^{-1}\Delta A^H P_Z^\perp \quad (11)$$

Après de nombreux traitements algébriques inspirés de [6] nous donnons les expressions des erreurs commises sur des différents projecteurs, soit

$$\delta E_{(\hat{A} B)} = (I - E_{(A B)})\Delta R_A S_A^\dagger E_{(A B)} + A(A^H P_B^\perp A)^{-1}S_A^\dagger \Delta R_A P_Z^\perp \quad (12)$$

où $S_A^\dagger = U_A(\Sigma_A - \sigma^2 I)^{-1}U_A^H$, et

$$\delta E_{(B \hat{A})} = -\Omega S_A^\dagger \Delta R_A P_Z^\perp - E_{(B A)}\Delta R_A S_A^\dagger E_{(A B)} \quad (13)$$

où $\Omega = B(B^H P_A^\perp B)^{-1}B^H$, $S_A^\dagger = U_A(\Sigma_A - \sigma^2 I)U_A^H$. En utilisant la propriété $P_Z^\perp = I - E_{(\hat{A} B)} - E_{(B \hat{A})}$ et les expressions(12) et (13) nous obtenons l'expression de l'erreur commise sur le projecteur orthogonal

$$P_Z^\perp = I - (A(A^H P_B^\perp A)^{-1} - \Omega)S_A^\dagger \Delta R_A P_Z^\perp + P_Z^\perp \Delta R_A S_A^\dagger E_{(A B)}. \quad (14)$$

Lorsque z est une racine, les versions polynomiales des critères (5), (6) et (7) s'annulent. Soit les trois fonctions associées :

$$\begin{aligned} \Pi_{WP}(I - E_{(\hat{A} B)}, \hat{z}) &= e_1^T (I - E_{(\hat{A} B)})p(\hat{z}), \\ \Pi_P(P_A^\perp, \hat{z}) &= e_1^T P_{\hat{A}^\perp} p(\hat{z}), \\ \Pi_C(P_Z^\perp, \hat{z}) &= 1 - e_1^T P_Z^\perp p(\hat{z}). \end{aligned}$$

Nous réalisons maintenant une analyse au premier ordre de ces fonctions, et en utilisant les expressions des erreurs

sur les projecteurs, nous pouvons désormais énoncer le théorème suivant

Théorème 2 *Les variances théoriques des algorithmes Root WP-MinNorm, Root P-MinNorm, and Root C-MinNorm sont données respectivement par :*

$$\begin{aligned} \mathcal{E}(|\theta - \hat{\theta}_{WP}|^2) &= \sigma^2 \frac{(e_1^T (P_Z^\perp + \Omega)e_1) (p(\theta)^H S_A^\dagger R_A S_A^\dagger p(\theta))}{2T |e_1^T E_{(A B)} p'(\theta)|^2}, \\ \mathcal{E}(|\theta - \hat{\theta}_P|^2) &= \sigma^2 \frac{(e_1^T P_A^\perp e_1) (p(\theta)^H S_A^\dagger R_A S_A^\dagger p(\theta))}{2T |e_1^T P_A^\perp p'(\theta)|^2}, \\ \mathcal{E}(|\theta - \hat{\theta}_C|^2) &= \sigma^2 \frac{(e_1^T P_Z^\perp e_1) (p(\theta)^H S_A^\dagger R_A S_A^\dagger p(\theta))}{2T |e_1^T P_Z^\perp p'(\theta)|^2} \end{aligned}$$

où $p'(\theta)$ est la dérivée première de $p(\theta)$.

5 Simulations numériques

Les simulations sont effectuées dans le contexte d'un réseau linéaire de L capteurs uniformément réparties, de demi longueur d'onde et nous considérons $T = 100$ observations. Les performances des algorithmes proposés sont comparées à celle de l'algorithme C-MinNorm [8] au travers du calcul de l'écart-type moyenné sur 500 réalisations du bruit. La Fig. 1-a montre que tous les algorithmes sont équivalents lorsque les DDA sont éloignées. Cependant, Fig. 1-b indique que les algorithmes WP-MinNorm et P-MinNorm ont une variance plus petite que celle de l'algorithme C-MinNorm lorsque les DDA sont proches. A noter également que l'algorithme WP-MinNorm présente la plus petite variance. Enfin, Fig. 1-a et b témoignent de l'adéquation entre les variances théoriques établies par le théorème 2 et les variances pratiques des algorithmes étudiés.

Afin de comparer plus en avant les différents algorithmes, nous formons les deux rapports suivants :

$$\rho_L(\theta) = \frac{\mathcal{E}(|\theta - \hat{\theta}_P|^2)}{\mathcal{E}(|\theta - \hat{\theta}_C|^2)}, \quad \beta_L(\theta) = \frac{\mathcal{E}(|\theta - \hat{\theta}_{WP}|^2)}{\mathcal{E}(|\theta - \hat{\theta}_C|^2)} \quad (15)$$

qui sont respectivement, le rapport entre les variances des algorithmes P-MinNorm et C-MinNorm, et le rapport entre les variances des algorithmes WP-MinNorm et C-MinNorm. Le paramètre θ est la DDA d'intérêt et nous étudions l'évolution des rapports pour plusieurs valeurs de L . Notons que les rapports sont invariants par rapport au niveau de bruit et au paramètre T . Sur les figures Fig. 2-a, b, c et d, nous avons tracé ces rapports pour les cas où les DDA sont éloignées et proches. Les expériences correspondant aux figures Fig. 2-a et b considèrent que θ_1 est la DDA connue et θ_2 est la DDA d'intérêt. Fig. 2-a montre que les variances des algorithmes P-MinNorm et WP-MinNorm sont plus petite que la variance de l'algorithme C-MinNorm pour le cas de DDA proches. En outre, lorsque le nombre de capteurs augmente ou lorsque les DDA sont éloignées la variance de chacun des algorithmes étudiés est équivalente. Le même protocole expérimental est utilisé pour obtenir Fig. 2-b, à l'exception que les DDA

sont très proches. Dans ce cas, il est clair que l'algorithme WP-MinNorm est plus performant que les algorithmes P-MinNorm et C-MinNorm.

Nous considérons ensuite que θ_1 et θ_2 sont les DDA connues et θ_3 est la DDA d'intérêt. Les résultats de simulation dans ce contexte sont représentés sur la Fig. 2-c. On voit alors que les variances des algorithmes WP-MinNorm et P-MinNorm sont plus petites que celle de l'algorithme C-MinNorm mais la différence entre les variances de l'algorithme WP-MinNorm et de P-MinNorm est réduite.

Lorsque les DDA sont plus éloignées, cf. Fig. 2-d où θ_1 , θ_2 , θ_3 sont les DDA connues et θ_4 est la DDA d'intérêt, l'algorithme WP-MinNorm est clairement plus précis que les deux autres algorithmes proposés. Ces simulations nous permettent de conclure que

$$\mathcal{E}(|\theta - \hat{\theta}_{WP}|^2) \leq \mathcal{E}(|\theta - \hat{\theta}_P|^2) \leq \mathcal{E}(|\theta - \hat{\theta}_C|^2). \quad (16)$$

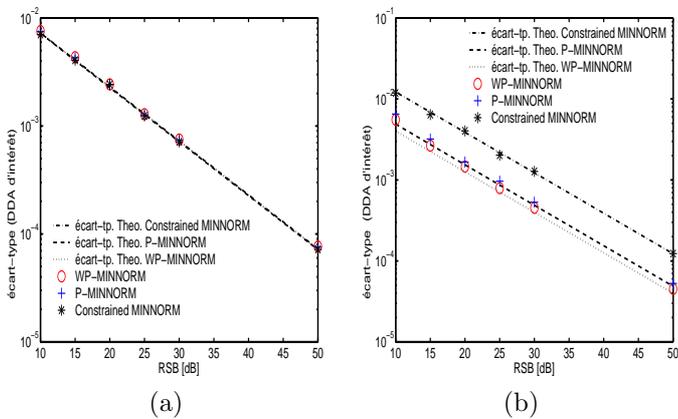


FIG. 1 – écart-type Vs. RSB, (a) $L = 10$ $\theta = [20 \ 2]$, (b) $L = 15$ $\theta = [11 \ 8]$;

6 Conclusion

Dans cet article, nous avons présenté deux algorithmes d'estimation de type MinNorm incorporant un *a priori* permettant ainsi d'estimer S DDA d'intérêt sur un ensemble total de M en utilisant la connaissance de $M - S$ DDA. Contrairement aux méthodes standards basées sur la déflation orthogonale, nous exploitons une déflation oblique du sous-espace signal. Afin d'analyser les algorithmes proposés, nous avons établi leurs variances théoriques. Cette analyse montre que les méthodes basées sur la déflation oblique sont plus performantes que les méthodes basées sur la déflation orthogonale pour des sources décorréliées de DDA proches.

Références

- [1] R.O. Schmidt, *A signal subspace approach to multiple emitter location and spectral estimation*, Ph.D. dissertation, Stanford University, Stanford, CA, 1981.
- [2] R.T. Behrens and L.L. Scharf, "Signal processing applications of oblique projection operators", *IEEE Trans. on Signal Processing*, Vol. 42, Issue 6, June 1994.
- [3] P. Stoica and R.L. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005.

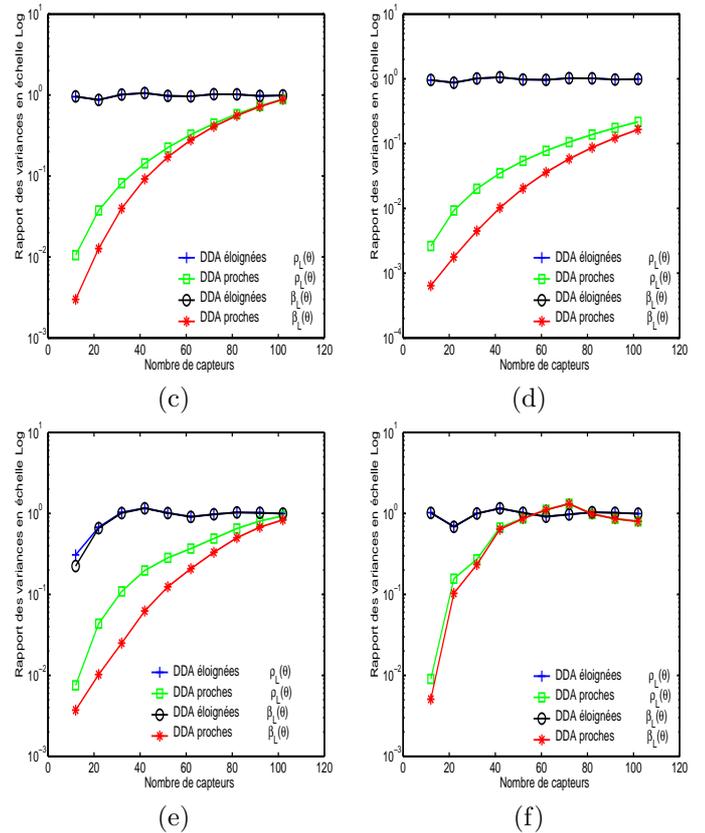


FIG. 2 – Rapport de variance Vs. nombre de capteurs, (c) $\theta = [20 \ 5]$ DDA éloignées et $\theta = [11 \ 10]$ DDA proches, (d) $\theta = [20 \ 5]$ DDA éloignées et $\theta = [11 \ 10.5]$ DDA proches, (e) $\theta = [5 \ 20 \ 30]$ DDA éloignées et $\theta = [18 \ 11 \ 9]$ DDA proches, (f) $\theta = [5 \ 20 \ 30 \ 10]$ DDA éloignées et $\theta = [25 \ 11 \ 9 \ 9.8]$ DDA proches.

- [4] H. Chen, S. Van Huffel, D. Van Ormondt and R. De Beer, "Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain", *Journal of Magnetic Resonance A*, vol. 119, 1996, pp. 225-234.
- [5] P. Stoica and A. Nehorai, "Comparative performance study of element-space and beamspace MUSIC estimators", *Circuits, Systems and Signal Processing*, Vol. 10, no 3, pp285-292, 1991.
- [6] H. Krim, P. Forster and J.G. Proakis, "Operator approach to performance analysis of root-MUSIC and root-min-norm", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 40, Issue 7, July 1992.
- [7] D. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids : Making linear prediction perform like maximum likelihood". *Proc IEEE*, Vol 70, Nf9, Sept 1982.
- [8] R.D. DeGroat, E.M. Dowling and D.A. Linebarger, "The constrained MUSIC problem", *IEEE Trans. on Signal Processing*, Vol 41, Issue 3, March 1993.
- [9] M. Kaveh and A.J. Barabell, "The Statistical Performance of the MUSIC and the Minimum-Nom Algorithms in Resolving Plane Waves in Noise", *IEEE Trans. ASSP*, ASSP-34 :331-341, April 1986.
- [10] R. Boyer and G. Bouleux, "Cramér-Rao Lower Bound for Prior-Subspace Estimation", *Fourth IEEE Workshop on sensor array and multi-channel processing (SAM-2006)*.
- [11] E.M. Dowling, R.D. DeGroat and D.A. Linebarger, "Exponential Parameter Estimation in the Presence of Known Components and Noise", *IEEE Trans. on Antennas and Propagation*, Vol. 42, Nf 5, May 1994.