

Performances des filtres de Wiener spatio-temporels entraînés: le cas des grandes dimensions.

G.T PHAM, P LOUBATON

Université Paris-Est Marne-la-Vallée
LIGM, UMR CNRS 8049
5 Boulevard Descartes
77454 Marne la Vallée Cedex 2, France

gia-thuy.pham@univ-mlv.fr, loubaton@univ-mlv.fr

Résumé – Dans cet article, nous évaluons le rapport signal à interférence plus bruit (SINR) produit par un filtre de Wiener spatio-temporel régularisé estimé par le biais d’une séquence d’apprentissage. Dans le régime asymptotique des grands systèmes défini par le fait que le nombre de capteurs et la taille de la séquence d’apprentissage tendent vers l’infini au même rythme, nous montrons que le SINR converge vers une quantité déterministe. Son analyse permet de montrer tout l’intérêt de la régularisation et d’un bon choix de la taille de fenêtre temporelle utilisée pour estimer le signal utile.

Abstract – In this paper, we evaluate the signal to interference plus noise ratio (SINR) produced by a trained regularized Wiener spatio-temporal filter. In the asymptotic regime of large systems where the number of receivers and the number of snapshots converge to infinity at the same rate, we establish that the SINR converges towards a deterministic quantity. The study of the limit SINR demonstrates the potential of a relevant regularization and of a good choice of the smoothing factor.

1 Introduction

Dans cet article, nous nous intéressons à des situations dans lesquelles l’observation $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ est une série temporelle de dimension M définie comme une version bruitée de la sortie d’un filtre 1 entrée / M –sorties inconnu, à réponse impulsionnelle finie, attaqué par une séquence scalaire $(s_n)_{n \in \mathbb{Z}}$ d’intérêt, supposée indépendante identiquement distribuée. Dans divers contextes applicatifs, il est pertinent de chercher à estimer chaque s_n par $(\mathbf{g}^{(L)})^* \mathbf{y}_n^{(L)}$ où $\mathbf{y}_n^{(L)}$ est le vecteur défini par $\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n}, \dots, \mathbf{y}_{1,n+L-1}, \dots, \mathbf{y}_{M,n}, \dots, \mathbf{y}_{M,n+L-1})^T$ et où le vecteur $\mathbf{g}^{(L)}$ minimise l’erreur quadratique moyenne entre s_n et $(\mathbf{g}^{(L)})^* \mathbf{y}_n^{(L)}$. Il convient alors d’estimer le vecteur $\mathbf{g}^{(L)}$. Pour cela, il est bien entendu possible d’utiliser des approches de déconvolution aveugle, mais il est fréquent que l’on dispose des valeurs prises par N échantillons consécutifs de la suite $(s_n)_{n \in \mathbb{Z}}$ grâce à une phase d’apprentissage préalable; nous supposons sans restriction qu’il s’agit des échantillons s_n pour $n = 1, \dots, N$. Dans ces conditions, le vecteur $\mathbf{g}^{(L)}$ est classiquement estimé par l’estimateur des moindres carrés $\hat{\mathbf{g}}^{(L)}$ défini comme l’action de l’inverse de la matrice de covariance spatio-temporelle empirique (i.e. la matrice $\frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*}$) sur le vecteur d’intercorrélation empirique entre $(s_n)_{n=1, \dots, N}$ et les vecteurs $(\mathbf{y}_n^{(L)})_{n=1, \dots, N}$. Le but de cet article est d’analyser le rapport signal à interférence plus bruit produit par l’estimateur $(\hat{\mathbf{g}}^{(L)})^* \mathbf{y}_n^{(L)}$ de s_n dans le cas où la dimension M de l’observation, qui représente souvent un nombre de capteurs, est "grande". Cette étude est classiquement menée dans le ré-

gime asymptotique dans lequel $N \rightarrow +\infty$ et M fixe, mais cela n’est pas très intéressant en pratique car on s’arrange en général pour que la taille de la séquence d’apprentissage N soit à peine plus grande que ML . Nous allons donc nous placer dans le régime asymptotique dit "des grandes dimensions" dans lequel M et N tendent vers l’infini au même rythme, et établir que le rapport signal à interférence plus bruit associé à l’estimateur $(\hat{\mathbf{g}}^{(L)})^* \mathbf{y}_n^{(L)}$ se comporte comme une quantité déterministe (i.e. ne dépendant que des statistiques du bruit additif perturbant l’observation dans la phase d’apprentissage) que l’on peut analyser pour obtenir des informations très utiles sur les valeurs de L qu’il convient de choisir. Nous nous intéressons également au cas où la matrice de covariance spatio-temporelle est régularisée, et en déduisons qu’il y a énormément à gagner à choisir de bonnes valeurs de L et du paramètre de régularisation. Faute de place, nous n’aborderons pas la question d’estimer ces paramètres, et nous nous contentons d’affirmer que cela est possible. Les résultats correspondants seront présentés dans une publication ultérieure.

A notre connaissance, le problème qui vient d’être évoqué n’a pas encore été considéré dans la littérature dans le contexte des grandes dimensions. Nous pouvons toutefois noter que [4], qui a inspiré ce travail, s’est intéressé à la même question dans le régime des grandes dimensions, mais lorsque le filtre RIF produisant l’observation est réduit à un vecteur constant **connu** (pas de séquence d’apprentissage), et que le paramètre L est égal à 1. [4] utilise des résultats portant sur le comportement

asymptotique de certaines fonctionnelles de la matrice $\mathbf{Y}\mathbf{Y}^*/N$ où \mathbf{Y} est la matrice $M \times N$ définie par $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$. \mathbf{Y} est la somme d'un terme déterministe dû à la séquence d'apprentissage et d'une matrice aléatoire dont les éléments sont indépendants identiquement distribués si on suppose que le bruit additif est blanc temporellement et spatialement. Dans notre contexte, \mathbf{Y} est remplacée par la matrice $ML \times N$ $\mathbf{Y}^{(L)} = (\mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(L)})$ qui a une structure Hankel par bloc dès que $L > 1$, et dont les entrées présentent donc des dépendances dont il s'agit d'analyser l'influence. La conclusion essentielle du présent travail est que tout se passe comme si la matrice $\mathbf{Y}^{(L)}$ ne présentait aucune dépendance.

2 Présentation du problème

D'une façon plus formelle, nous supposons que les vecteurs observés $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ sont donnés par

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n, n = 1, \dots, N \quad (1)$$

où $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ représente la fonction de transfert de taille $M \times 1$ du filtre 1 entrée / M sorties inconnu, et où $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ constitue une suite de i.i.d. de vecteurs aléatoires gaussiens complexes de matrice de covariance $\sigma^2 \mathbf{I}$. Bien que $\mathbf{h}(z)$ soit inconnu, nous supposons que P est connu, ce qui, en pratique, signifie que l'on connaît une borne supérieure du support de la réponse impulsionnelle associée à $\mathbf{h}(z)$.

Soit L un entier, et $\mathbf{y}_n^{(L)}$ le vecteur de taille ML défini par

$$\mathbf{y}_n^{(L)} = (\mathbf{y}_{1,n}, \dots, \mathbf{y}_{1,n+L-1}, \dots, \mathbf{y}_{M,n}, \dots, \mathbf{y}_{M,n+L-1})^T$$

Nous rappelons que $(s_n)_{n \in \mathbb{Z}}$ est connu de $n = 1$ jusqu'à $n = N$, et nous allons à présent exprimer sous forme matricielle les relations existant entre $(\mathbf{y}_n^{(L)})_{n=1, \dots, N}$, $(\mathbf{v}_n^{(L)})_{n=1, \dots, N}$ et $(s_n)_{n=1, \dots, N}$. Pour cela, nous rappelons que $\mathbf{Y}^{(L)}$ est la matrice $ML \times N$ définie par $\mathbf{Y}^{(L)} = (\mathbf{y}_1^{(L)}, \dots, \mathbf{y}_N^{(L)})$. Il est facile de vérifier que

$$\mathbf{Y}^{(L)} = \mathbf{H}^{(L)} \mathbf{S}^{(L)} + \mathbf{V}^{(L)} \quad (2)$$

où la matrice $\mathbf{S}^{(L)}$ est la matrice de Hankel de taille $(P+L-1) \times N$ définie par $(\mathbf{S}^{(L)})_{i,n} = s_{n+i-P}$, et la matrice $\mathbf{H}^{(L)}$ est définie par $\mathbf{H}^{(L)} = (\mathbf{H}_1^{(L)T}, \dots, \mathbf{H}_M^{(L)T})^T$ avec $\mathbf{H}_m^{(L)}$ représentant la matrice de Toeplitz de taille $L \times (P+L-1)$ correspondant à la convolution du signal $(s_n)_{n \in \mathbb{Z}}$ avec la séquence $(\mathbf{h}_{m,p})_{p=0, \dots, P-1}$. Dans la suite, nous notons \mathbf{Y} , \mathbf{H} , \mathbf{S} , et \mathbf{V} les matrices $\mathbf{Y}^{(1)}$, $\mathbf{H}^{(1)}$, $\mathbf{S}^{(1)}$ et $\mathbf{V}^{(1)}$. Il convient de noter que la matrice $\mathbf{S}^{(L)}$ dépend en principe de valeurs de s_n pour $-(P-1) \leq n \leq 0$ et $N < n \leq N+L-1$ du fait d'effets de bord induits par le filtre $\mathbf{h}(z)$. Dans la suite, nous convenons de mettre à 0 les valeurs de s_n correspondant à des indices qui ne sont pas compris entre 1 et N .

Le vecteur $\hat{\mathbf{g}}^{(L)}$ est évidemment défini comme le minimiseur du critère des moindres carrés $\sum_{n=1}^N |\hat{\mathbf{g}}^{(L)*} \mathbf{y}_n^{(L)} - s_n|^2$, et est donné par

$$\hat{\mathbf{g}}^{(L)} = \left(\frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^{(L)} s_n^* \right) \quad (3)$$

et sa version régularisée $\hat{\mathbf{g}}_\lambda^{(L)}$ est définie par

$$\hat{\mathbf{g}}_\lambda^{(L)} = \left(\frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^{(L)} \mathbf{y}_n^{(L)*} + \lambda \mathbf{I}_{ML} \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^{(L)} s_n^* \right) \quad (4)$$

Les estimateurs $\hat{\mathbf{g}}^{(L)}$ et $\hat{\mathbf{g}}_\lambda^{(L)}$ sont destinés à être utilisés pour estimer des valeurs de s_n pour $n > N$, et l'un des façons d'évaluer les performances qu'ils produisent est de calculer le rapport signal à interférence plus bruit qu'ils produisent. Plus précisément, si $n > N$,

$$\hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{y}_n^{(L)} = \hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{h}_P^{(L)} s_n + \hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{H}_1^{(L)*} \mathbf{s}_{(-n)}^{(L)} + \hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{v}_n^{(L)}$$

où $\mathbf{h}_P^{(L)}$ est la colonne P de $\mathbf{H}^{(L)}$, $\mathbf{H}_1^{(L)}$ est la matrice $\mathbf{H}^{(L)}$ privée de sa P -ième colonne et

$$\mathbf{s}_{(-n)}^{(L)} = (s_{n-P+1}, \dots, s_{n-1}, s_{n+1}, \dots, s_{n+L-1})^T$$

En utilisant le fait que la suite (s_n) est i.i.d., on obtient immédiatement que

$$\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) = \frac{|\hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{h}_P^{(L)}|^2}{\hat{\mathbf{g}}_\lambda^{(L)*} \mathbf{H}_1^{(L)} \mathbf{H}_1^{(L)*} \hat{\mathbf{g}}_\lambda^{(L)} + \sigma^2 \|\hat{\mathbf{g}}_\lambda^{(L)}\|^2} \quad (5)$$

Remarque : On note que la régularisation du type Ledoit-Wolf [5] nous mène également à la même expression du SINR via une factorisation simple.

Cette quantité est aléatoire dans la mesure où elle dépend des éléments de la matrice $\mathbf{Y}^{(L)}$, qui sont eux-mêmes aléatoires du fait de la présence de la matrice $\mathbf{V}^{(L)}$. Lorsque N tend vers l'infini et que M et L restent fixes, il est facile de vérifier que si $\lambda = 0$, le filtre $\hat{\mathbf{g}}^{(L)}$ tend vers le filtre de Wiener $(\mathbf{H}^{(L)} \mathbf{H}^{(L)*} + \sigma^2 \mathbf{I})^{-1} \mathbf{h}_P^{(L)}$, et que $\text{SINR}(\hat{\mathbf{g}}^{(L)})$ se comporte comme le SINR correspondant qui est égal à

$$\gamma = \frac{\mathbf{h}_P^{(L)*} (\mathbf{H}^{(L)} \mathbf{H}^{(L)*} + \sigma^2 \mathbf{I})^{-1} \mathbf{h}_P^{(L)}}{1 - \mathbf{h}_P^{(L)*} (\mathbf{H}^{(L)} \mathbf{H}^{(L)*} + \sigma^2 \mathbf{I})^{-1} \mathbf{h}_P^{(L)}} \quad (6)$$

Dans le cas où $\lambda > 0$, l'analyse est également élémentaire. Lorsque M et N sont du même ordre de grandeur, le comportement de $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ est évidemment bien différent, et nécessite un travail beaucoup plus conséquent. A partir de maintenant, nous supposons que nous nous plaçons dans le régime des grandes dimensions suivant :

$$M, N \rightarrow +\infty, \text{ le rapport } c_N = \frac{M}{N} \rightarrow c > 0, \text{ et } P \text{ et } L \text{ n'augmentent pas avec } N.$$

Afin de simplifier les notations, $N \rightarrow +\infty$ devra être compris comme le régime asymptotique qui vient d'être défini. Dans la suite, nous allons énoncer (sans preuve) des résultats permettant de montrer que $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ converge presque sûrement vers une quantité déterministe dont nous donnons l'expression.

3 Etude spectrale de la matrice spatio-temporelle empirique

Afin de simplifier les notations, on note $\mathbf{W}^{(L)}$, $\mathbf{U}^{(L)}$, $\Sigma^{(L)}$ les matrices définies par $\mathbf{W}^{(L)} = \mathbf{V}^{(L)} / \sqrt{N}$, $\mathbf{U}^{(L)} = \mathbf{S}^{(L)} / \sqrt{N}$

et $\Sigma^{(L)} = \mathbf{H}^{(L)}\mathbf{U}^{(L)} + \mathbf{W}^{(L)}$. Ce papier est basé sur des résultats techniques (voir par exemple [3]) montrant que tout se passe comme si les éléments de la matrice $\mathbf{W}^{(L)}$ étaient i.i.d. Avant de présenter ces résultats techniques, nous rappelons que la loi Marchenko-Pastur (voir par exemple [1]) de paramètres (σ^2, d) est la loi de probabilité μ_d (la dépendance en σ^2 n'est pas mentionnée afin de simplifier les notations) définie par :

$$d\mu_d(x) = \delta_0[1 - d^{-1}]_+ + \frac{\sqrt{(x - x^-)(x^+ - x)}}{2\sigma^2 d \pi x} \mathbb{1}_{[x_d^-, x_d^+]}(x) dx$$

avec $x_d^- = \sigma^2(1 - \sqrt{d})^2$ et $x_d^+ = \sigma^2(1 + \sqrt{d})^2$. On note $m_d(z)$ la transformée de Stieltjes de μ_d définie par $m_d(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z}$ et $\tilde{m}_d(z)$ la fonction $\tilde{m}_d(z) = dm_d(z) - (1-d)/z$. On note aussi $\mathbf{Q}_W(z)$ et $\tilde{\mathbf{Q}}_W(z)$ les résolvantes des matrices $\mathbf{W}^{(L)}\mathbf{W}^{(L)*}$ et $\mathbf{W}^{(L)*}\mathbf{W}^{(L)}$ définies par $\mathbf{Q}_W(z) = (\mathbf{W}^{(L)}\mathbf{W}^{(L)*} - z\mathbf{I}_{ML})^{-1}$ et $\tilde{\mathbf{Q}}_W(z) = (\mathbf{W}^{(L)*}\mathbf{W}^{(L)} - z\mathbf{I}_N)^{-1}$. Alors, quand $N \rightarrow +\infty$, les résultats suivants sont vérifiés :

Proposition 1. *La distribution des valeurs propres de la matrice $\mathbf{W}^{(L)}\mathbf{W}^{(L)*}$ converge presque sûrement vers la loi de Marcenko-Pastur μ_{cL} . Si $\mathbf{a}_N, \mathbf{b}_N$ sont des vecteurs unitaires déterministes de dimension ML , alors pour tout $z \in \mathbb{C}^+$, on a la propriété suivante :*

$$\mathbf{a}_N^* (\mathbf{Q}_W(z) - m_{cL}(z)\mathbf{I}) \mathbf{b}_N \rightarrow 0 \text{ p.s.} \quad (7)$$

De même, si $\tilde{\mathbf{a}}_N, \tilde{\mathbf{b}}_N$ sont des vecteurs unitaires déterministes de dimension N , alors pour tout $z \in \mathbb{C}^+$, on a :

$$\tilde{\mathbf{a}}_N^* (\tilde{\mathbf{Q}}_W(z) - \tilde{m}_{cL}(z)\mathbf{I}) \tilde{\mathbf{b}}_N \rightarrow 0 \text{ p.s.} \quad (8)$$

De plus, pour tout $z \in \mathbb{C}^+$, on a

$$\mathbf{a}_N^* (\mathbf{Q}_W(z)\mathbf{W}_N^{(L)}) \tilde{\mathbf{b}}_N \rightarrow 0 \text{ p.s.} \quad (9)$$

Nous rappelons de façon informelle que la convergence de la distribution des valeurs propres de $\mathbf{W}^{(L)}\mathbf{W}^{(L)*}$ vers la distribution de Marcenko-Pastur μ_{cL} signifie que les histogrammes des valeurs propres de cette matrice tendent à se rapprocher de la densité de probabilité de la distribution de μ_{cL} . La Proposition 1 est bien connue quand $L = 1$. A l'exception de (8) et (9), la Proposition 1 est une conséquence des résultats de [3].

4 Comportement asymptotique du SINR

Avant de calculer le SINR, on note $\mathbf{Q}(z) = (\Sigma^{(L)}\Sigma^{(L)*} - z\mathbf{I}_{ML})^{-1}$, et on remarque que $\hat{\mathbf{g}}_\lambda^{(L)} = \mathbf{Q}(-\lambda)\Sigma^{(L)}\mathbf{u}^*$, avec $\mathbf{u} = \frac{1}{\sqrt{N}}(s_1, \dots, s_N)$ qui est la P-ième ligne de $\mathbf{U}^{(L)}$. Pour évaluer le comportement du SINR, il faut étudier $|\mathbf{h}_P^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}|^2$, $\|\mathbf{H}_1^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}\|^2$, et $\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2$. Pour ceci, il est essentiel d'exprimer \mathbf{Q} en fonction de \mathbf{Q}_W car le comportement des formes bilinéaires de \mathbf{Q}_W est donné par la Proposition 1. Afin de simplifier les notations, nous allons omettre l'exposant (L) dans les différentes matrices et ne pas mentionner la dépendance par rapport à la variable complexe z . Nous posons également $d = cL$. \mathbf{Q} peut s'exprimer sous la forme

$$\mathbf{Q} = [\mathbf{W}\mathbf{W}^* + \mathbf{H}\mathbf{U}\mathbf{W}^* + \mathbf{W}\mathbf{U}^*\mathbf{H}^* + \mathbf{H}\mathbf{U}\mathbf{U}^*\mathbf{H}^* - z\mathbf{I}_{ML}]^{-1}$$

Des calculs standards (c.f [2], chapitre 7) permettent d'établir que :

$$\mathbf{Q} = \mathbf{Q}_W - \mathbf{Q}_W(\mathbf{H}, \mathbf{W}\mathbf{U}^*) \begin{pmatrix} \mathbf{U}\mathbf{U}^* & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \times$$

$$\left[\mathbf{I} + \begin{pmatrix} \mathbf{H}^* \\ \mathbf{U}\mathbf{W}^* \end{pmatrix} \mathbf{Q}_W(\mathbf{H}, \mathbf{W}\mathbf{U}^*) \begin{pmatrix} \mathbf{U}\mathbf{U}^* & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \right]^{-1} \begin{pmatrix} \mathbf{H}^* \\ \mathbf{U}\mathbf{W}^* \end{pmatrix} \mathbf{Q}_W$$

En remarquant que $\mathbf{W}^*\mathbf{Q}_W\mathbf{W} = \tilde{\mathbf{Q}}_W\mathbf{W}^*\mathbf{W} = \mathbf{I} + z\tilde{\mathbf{Q}}_W$, et que $\mathbf{U}\mathbf{U}^* \rightarrow \mathbf{I}$, p.s, la proposition 1 permet d'aboutir à

$$\begin{pmatrix} \mathbf{H}^* \\ \mathbf{U}\mathbf{W}^* \end{pmatrix} \mathbf{Q}_W(\mathbf{H}, \mathbf{W}\mathbf{U}^*) \xrightarrow{p.s} \begin{pmatrix} m_d(z)\mathbf{H}^*\mathbf{H} & \mathbf{0} \\ \mathbf{0} & (1 + z\tilde{m}_d(z))\mathbf{I} \end{pmatrix}$$

En posant $\omega(z) = \frac{1}{zm_d(z)\tilde{m}_d(z)}$, on aboutit au résultat suivant :

Proposition 2. *Quand $N \rightarrow +\infty$, les trois termes $\|\mathbf{h}_P^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}\|^2$, $\|\mathbf{H}_1^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}\|^2$ et $\|\hat{\mathbf{g}}_\lambda^{(L)}\|^2$ peuvent être approchés (i.e. ont le même comportement presque sur) par les quantités déterministes suivantes :*

$$\|\mathbf{h}_P^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}\|^2 \simeq (\mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)})^2 \quad (10)$$

$$\begin{aligned} \|\mathbf{H}_1^{(L)*}\hat{\mathbf{g}}_\lambda^{(L)}\|^2 &\simeq \omega(-\lambda)\mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P^{(L)} + \\ &\mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)} \times \\ &\left(1 - \mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)}\right) \quad (11) \end{aligned}$$

$$\begin{aligned} \|\hat{\mathbf{g}}_\lambda^{(L)}\|^2 &\simeq \frac{\sigma^2 d}{\omega(-\lambda) - \sigma^4 d} \left(1 - \mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)}\right) \\ &+ \frac{\omega(-\lambda)(\sigma^2 d + \omega(-\lambda))}{\omega(-\lambda) - \sigma^4 d} \mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-2}\mathbf{h}_P^{(L)} \quad (12) \end{aligned}$$

En posant $\alpha(\lambda) = \mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)}$, $\beta(\lambda) = \mathbf{h}_P^{(L)*}(\mathbf{H}^{(L)}\mathbf{H}^{(L)*} - \omega(-\lambda)\mathbf{I})^{-1}\mathbf{h}_P^{(L)}$, on peut alors établir que

$$\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)}) - \phi(\lambda) \rightarrow 0 \quad (13)$$

où $\phi(\lambda)$ est défini par

$$\phi(\lambda) = \frac{\alpha(\lambda)^2}{[1 - \alpha(\lambda)][(\alpha(\lambda) + \frac{\sigma^4 d}{\omega(-\lambda)^2 - \sigma^4 d}] + \frac{\omega(-\lambda)(\omega(-\lambda) + \sigma^2)}{\omega(-\lambda) - \sigma^4 d} \beta(\lambda)} \quad (14)$$

Dans le cas où $\lambda = 0$, la formule donnant $\phi(0)$ est particulièrement intéressante puisque, en utilisant l'identité $\omega(0) = -\sigma^2$, on obtient immédiatement que

$$\phi(0) = \gamma \frac{(1 - cL)\gamma}{\gamma + cL} \quad (15)$$

où γ est le SINR correspondant au vrai filtre de Wiener (voir formule (6)). Par conséquent, la pénalité en terme de SINR due à phase d'estimation du filtre de Wiener est égale à $(1 - cL)\frac{\gamma}{\gamma + cL}$, et apparaît donc très importante quand L augmente. En absence de régularisation, il est donc en général très pénalisant de choisir des valeurs de L aux alentours de la valeur

maximale égale à la partie entière de c^{-1} . Ainsi que les évaluations numériques qui suivent vont le montrer, la régularisation permet d'améliorer considérablement les choses. Il convient de noter que si $\lambda > 0$, il n'y a aucune raison de supposer que $ML/N < 1$ et que $cL < 1$.

Il est clair que la fonction $\lambda \rightarrow \phi(\lambda)$ dépend de la matrice \mathbf{H} qui est inconnue. Par conséquent, la Proposition 2 ne permet pas d'estimer la valeur optimale du terme de régularisation λ . Cependant, il est tout à fait possible d'estimer $\phi(\lambda)$ de façon consistante pour tout λ à partir des observations. Faute de place, nous ne présentons pas ici les résultats correspondants, et nous contentons d'illustrer dans la section suivante le comportement de la fonction $\lambda \rightarrow \phi(\lambda)$ et tout l'intérêt qu'il y a à choisir les bonnes valeurs de L et de λ .

5 Résultats numériques

Dans cette section, nous allons tout d'abord évaluer dans quelle mesure l'approximation de $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ par $\phi(\lambda)$ est raisonnable. Afin de générer $\mathbf{H} = (\mathbf{h}_0, \dots, \mathbf{h}_{P-1})$, nous considérons la matrice $(\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_P))$, avec le vecteur $\mathbf{a}(\theta) = \frac{1}{M}(1, \dots, e^{i(M-1)\theta})^T$ où les angles $(\theta_p)_{p=0, \dots, P-1}$ sont tirés aléatoirement sur $[0, 2\pi]$. \mathbf{H} est obtenue en normalisant cette matrice par la racine carrée de sa norme de Frobenius. La séquence $(s_n)_{n=1, \dots, N}$ est une réalisation d'une séquence i.i.d prenant des valeurs dans ± 1 avec la probabilité $1/2$. Dans ces conditions, le rapport signal sur bruit SNR est égal à $\frac{1}{\sigma^2}$. Enfin, nous prenons $N = 200$, $M = 40$ et $P = 5$. Dans la figure 1, SNR est égal à 8 dB. Nous évaluons $\text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$ par simulations de Monte-Carlo, et représentons le graphe de la fonction $\lambda \rightarrow \phi(\lambda)$ ainsi que 2 courbes représentant les bornes inférieure et supérieure d'un intervalle de confiance à 95 pour cent de $\lambda \rightarrow \text{SINR}(\hat{\mathbf{g}}_\lambda^{(L)})$. Ainsi que nous pouvons le constater, les 3 courbes sont relativement proches l'une de l'autre.

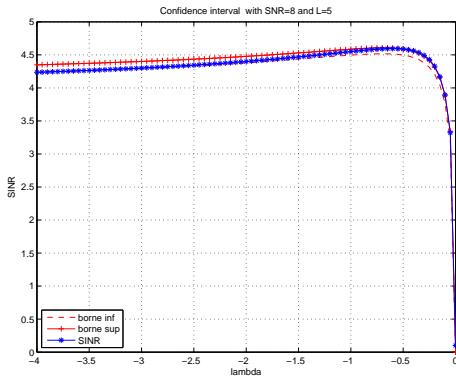


FIGURE 1 – Région de confiance du SINR et la courbe de SINR asymptotique en fonction de λ

Dans les figures 2, nous gardons les mêmes paramètres, mais faisons varier L de 1 à 8 pour évaluer la meilleure valeur de L . A l'exception du point $\lambda = 0$ qui est particulier, il apparaît clairement que la valeur optimale de L est $L = P = 5$. Par ailleurs, le fait de choisir λ convenablement permet d'améliorer les performances de façon conséquente.

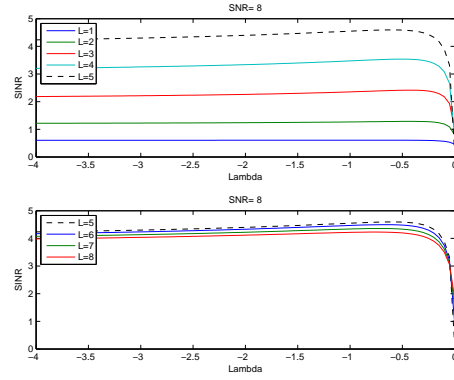


FIGURE 2 – SINR asymptotique en fonction de L et λ

Dans le contexte des figures 3, le SNR vaut 14dB, et la valeur optimale de L est égale à 6. Lorsque le SNR augmente,

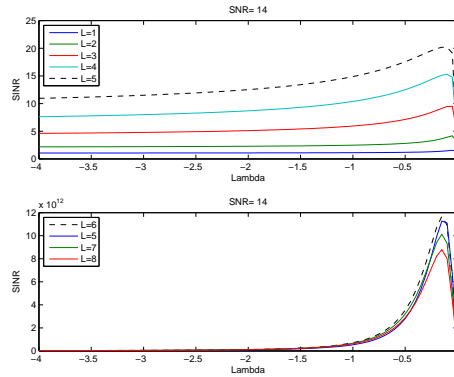


FIGURE 3 – SINR asymptotique en fonction de L et λ

il semble que la valeur optimale de L soit toujours égale à 6, donc ne tend pas à augmenter.

Références

- [1] L.A Pastur, M.Shcherbina, "Eigenvalue Distribution of Large Random Matrices", American Mathematical Society, 2011
- [2] R. Horn, C.R. Johnson, "Matrix Analysis", 2nd. Ed., Cambridge Univ. Press, 2013.
- [3] P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices", to appear on J. of Th. Prob., ArXiv :1405.2006.
- [4] X. Mestre, M. A. Lagunas. "Finite sample size effect on Minimum Variance beamformers : optimum diagonal loading factor for large arrays". IEEE Transactions on Signal Processing, vol. 54, no. 1, pp. 69-82, January 2006.
- [5] O.Ledoit, M.Wolf "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices", Journal of Multivariate Analysis, Volume 88, Issue 2, February 2004, pages 365-411.