
Fonctions de contraste avec référence

5.1 Présentation

Ce chapitre se place dans le cadre des approches séquentielles qui ont déjà été abordées dans les chapitres 3 et 4. Nous nous concentrerons plus particulièrement sur l'extraction d'une source par un filtre MISO. L'extension à des méthodes MIMO est ensuite possible par utilisation d'un procédé de déflation.

L'intérêt majeur de ce chapitre est de proposer des fonctions de contraste dont la dépendance est *quadratique* en les paramètres. La maximisation du critère est alors fortement simplifiée : il n'est plus nécessaire de recourir à une optimisation itérative et il en résulte une charge de calcul considérablement allégée en comparaison de celle requise par la méthode du chapitre 3. Cet allègement se traduit par une accélération remarquable du processus de séparation.

L'idée centrale qui a permis l'élaboration des contrastes est celle de signal de *référence*. Ainsi, notre approche pourrait, dans un premier temps, être qualifiée de semi-aveugle même si son utilité s'est par la suite avérée dans le cas aveugle. La notion de signal de référence, ainsi qu'un rapide tour d'horizon des travaux en lien avec les nôtres, sont présentés dans un premier temps. Nous démontrons ensuite la validité des critères de contrastes dans le cas i.i.d. et non i.i.d. Enfin, nous détaillons les possibles difficultés de mise en œuvre avant de donner quelques résultats de simulation.

5.1.1 Signal de référence

Notations Dans ce chapitre, nous considérons l'extraction MISO d'une source à partir des observations. Nous reprenons par conséquent des notations identiques à celles du chapitre 3 et nous nous plaçons dans le cas sans bruit : $(y(n))_{n \in \mathbb{Z}}$ est ainsi la sortie du filtre séparateur $\mathbf{w}[z]$ et, en introduisant le filtre ligne global $\mathbf{g}[z] := \mathbf{w}[z]\mathbf{M}[z]$, on a dans le cas non bruité $y(n) = \mathbf{w}[z]\mathbf{x}(n) = \mathbf{g}[z]\mathbf{s}(n)$. Les fonctions de transfert $\mathbf{w}[z]$ et $\mathbf{g}[z]$ correspondent à une même ligne des filtres respectifs $\mathbf{W}[z]$ (filtre séparateur) et $\mathbf{G}[z]$ (filtre global).

Référence L'approche ici développée s'apparente aux approches *semi-aveugles* dans la mesure où nous supposons que nous disposons d'une première approximation notée $\mathbf{w}_r[z]$

5.1.3 Retour sur la définition d'un contraste

5.1.3-a) Dissymétrie possible

Dans le chapitre 3, la définition 5 d'une fonction de contraste supposait une invariance par translation temporelle, c'est-à-dire que la valeur du contraste était la même pour la sortie $(y(n))_{n \in \mathbb{Z}}$ et ses translatées $(y(n-l))_{n \in \mathbb{Z}}$ où l est fixé quelconque dans \mathbb{Z} . Cette invariance, valable de fait pour tous les contrastes considérés jusqu'à présent ne nous est toutefois d'aucune utilité pour la séparation de sources : un critère de séparation qui privilégierait l'obtention d'une version retardée particulière de l'une des sources pourrait en effet tout à fait nous convenir. C'est, en l'occurrence, ce qui advient ici lorsque le contraste considéré dépend non plus des seules statistiques de la sortie, mais des statistiques conjointes de la sortie et du signal de référence. Pour cette raison, nous utiliserons dans ce chapitre la définition suivante :

Définition 7 (contraste MISO, cas i.i.d. sans invariance temporelle) Une fonction de contraste MISO (ou contraste MISO) est une fonction réelle de la sortie $(y(n))_{n \in \mathbb{Z}}$ du filtre ligne global $\mathbf{g}[z]$ qui vérifie :

(i) Il existe $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour toute sortie globale $(y(n))_{n \in \mathbb{Z}}$ on ait :

$$\mathcal{C}((y(n))_{n \in \mathbb{Z}}) \leq \mathcal{C}((s_{i_0}(n-l_0))_{n \in \mathbb{Z}}). \quad (5.3)$$

(ii) L'égalité dans l'équation (5.3) ci-dessus n'est possible que si le filtre MISO global est du type $\mathbf{g}[z] = (0, \dots, 0, \alpha z^{-d}, 0, \dots, 0)$, où α est un scalaire non nul, $d \in \mathbb{Z}$ et la position de l'élément non nul αz^{-d} correspond à un indice i_0 qui assure la condition (i).

La définition ci-dessus prend en compte les variations éventuelles de la valeur d'un contraste pour des retards différents de la source reconstruite. En revanche, dans le cas de sources non linéaires (et non i.i.d.) aucune adaptation de la définition n'est nécessaire. En effet, la définition 6 prend en compte le fait qu'une source non i.i.d. ne peut être reconstruite qu'à un filtrage scalaire près, ce qui inclut l'ensemble des décalages temporels possibles.

5.1.3-b) Normalisation en puissance

Nous rappelons le cadre de base mis en place au chapitre 3 pour l'étude des approches séquentielles, et que l'on retrouve ici. Nous conservons l'hypothèse H.10 qui s'écrit pour mémoire :

H.10 Pour tout $j \in \{1, \dots, N\}$, la fonction d'autocorrélation de la $j^{\text{ème}}$ source, notée $(\gamma_j(k))_{k \in \mathbb{Z}} := (\mathbb{E}\{s_j(n)s_j^*(n-k)\})_{k \in \mathbb{Z}}$, est définie positive. De plus, chaque source est de puissance unité ($\gamma_j(0) := \mathbb{E}\{|s_j(n)|^2\} = 1$).

Dès lors, pour tout indice $j \in \{1, \dots, N\}$ et tout filtre SISO $h[z]$, nous définissons la norme :

$$\|h\|_j := \left(\sum_{(k,l) \in \mathbb{Z}^2} h(k)h^*(l)\gamma_j(l-k) \right)^{\frac{1}{2}}. \quad (5.4)$$

et, pour tout filtre global $\mathbf{g}[z] := (g_1[z], \dots, g_N[z])$, nous définissons sa norme ℓ^2 pondérée :

$$\|\mathbf{g}\| := \left(\sum_{j=1}^N \|g_j\|_j^2 \right)^{\frac{1}{2}}. \quad (5.5)$$

Enfin, rappelons que nous notons :

$$\forall i \in \{1, \dots, N\} \quad \mathcal{G}_i := \{ \mathbf{g}[z] = (g_1[z], \dots, g_N[z]) \mid \|g_j\|_j = \delta_{i-j}, \forall j \} \quad (5.6)$$

l'ensemble des filtres globaux de norme unité dont toutes les composantes sont nulles, sauf la $i^{\text{ème}}$.

De même qu'au chapitre 3, nous normaliserons la sortie globale afin qu'elle soit de puissance unité : $E\{|y(n)|^2\} = 1$. Compte tenu de la puissance unité des sources, ceci est équivalent à travailler sur l'ensemble des filtres globaux de norme unité qui a été noté :

$$\mathcal{F}_1 := \{ \mathbf{g}[z] \mid \|\mathbf{g}\| = 1 \}. \quad (5.7)$$

Tous les critères seront considérés sous contrainte que le filtre global soit de norme un ; par conséquent, le scalaire α introduit dans les définitions 6 et 7 vérifie $|\alpha| = 1$. Enfin, nous supposons :

H.15 Le filtre global de référence $\mathbf{t}[z]$ est stable (*i.e.* sa réponse impulsionnelle est absolument sommable) et de norme unité : $\|\mathbf{t}\| = 1$.

Notons que dans le cas de sources i.i.d. de variance 1, la contrainte de norme 1 s'écrit pour le filtre global et le filtre de référence :

$$\|\mathbf{g}\|^2 = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} |g_i(k)|^2 = 1 \quad \text{et} \quad \|\mathbf{t}\|^2 = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} |t_i(k)|^2 = 1. \quad (5.8)$$

5.2 Contrastes MISO dans le cas convolutif

5.2.1 Cas de sources i.i.d.

5.2.1-a) Contraste quadratique à base de cumulants croisés d'ordre 4

Dans ce paragraphe, nous supposons disposer d'un signal de référence $(z(n))_{n \in \mathbb{Z}}$ et nous définissons le cumulante croisé suivant :

$$\kappa_4\{y(n), z(n)\} := \text{Cum}\{y(n), y^*(n), z(n), z^*(n)\} \quad (5.9)$$

Nous considérons alors le critère suivant :

$$\mathcal{C}_z\{y(n)\} := |\kappa_4\{y(n), z(n)\}| \quad (5.10)$$

Nous allons étudier sous quelles conditions sur le signal ou filtre de référence, le critère (5.10) constitue un contraste. Le résultat sera donné par la proposition 21 avant que les conditions soient discutées dans le paragraphe 5.2.1-b). Nous pouvons dans un premier temps affirmer :

Lemme 4 *On suppose vérifiée la contrainte de puissance $E\{|y(n)|^2\} = \|\mathbf{g}\|^2 = 1$. Il existe $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour toute sortie globale $(y(n))_{n \in \mathbb{Z}}$ on ait : $\mathcal{C}_z\{y(n)\} \leq \mathcal{C}_z\{s_{i_0}(n - l_0)\}$ (équation (5.3)).*

Preuve: Par multilinéarité des cumulants, on peut écrire :

$$\kappa_4\{y(n), z(n)\} = \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 |t_j(k)|^2 \kappa_{s_j}(0, 0, 0) \quad (5.11)$$

Définissons

$$\forall (j, k) \in \{1, \dots, N\} \times \mathbb{Z}, \quad \mathcal{M}_j(k) := |t_j(k)|^2 \kappa_{s_j}(0, 0, 0) \quad (5.12)$$

et :

$$\mathcal{M}_{\max} := \max_{j=1}^N \left(\sup_{k \in \mathbb{Z}} |\mathcal{M}_j(k)| \right) \quad (5.13)$$

En raison de la stabilité de $\mathbf{t}[z]$, la famille $|\mathcal{M}_j(k)|, k \in \mathbb{Z}$ est sommable. Ainsi, la borne supérieure qui a été définie est atteinte pour un couple $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$. Il vient alors :

$$\mathcal{C}_z\{y(n)\} = |\kappa_4\{y(n), z(n)\}| = \left| \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 \mathcal{M}_j(k) \right| \quad (5.14)$$

$$\leq \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 |\mathcal{M}_j(k)| \quad (5.15)$$

$$\leq \mathcal{M}_{\max} \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2. \quad (5.16)$$

En utilisant successivement l'hypothèse de norme unité et le fait que la borne supérieure est atteinte, on en déduit que :

$$\mathcal{C}_z\{y(n)\} \leq \mathcal{M}_{\max} = \mathcal{C}_z\{s_{i_0}(n - l_0)\}, \quad (5.17)$$

d'où le résultat annoncé. ■

Le lemme suivant donne une condition nécessaire et suffisante pour que le majorant ci-dessus soit atteint :

Lemme 5 *La fonction $\mathcal{C}_z\{\cdot\}$ atteint sa borne supérieure (i.e. $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$) si et seulement si :*

$$\sum_{(j,k) \in \mathfrak{I}} |g_j(k)|^2 = 1 \quad (5.18)$$

où \mathfrak{I} est l'un des ensembles :

$$\mathfrak{I}_+ := \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid \mathcal{M}_j(k) = \mathcal{M}_{\max}\} \quad (5.19)$$

$$\mathfrak{I}_- := \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid \mathcal{M}_j(k) = -\mathcal{M}_{\max}\} \quad (5.20)$$

Cette condition revient de façon équivalente à dire que les coefficients $g_j(k)$ sont identiquement nuls en dehors de l'ensemble \mathfrak{I}_+ (ou \mathfrak{I}_-).

Preuve: Supposons $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$. D'après l'égalité (5.14), il existe $\epsilon \in \{-1, +1\}$ tel que :

$$\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 \mathcal{M}_j(k) = \epsilon \mathcal{M}_{\max} \quad (5.21)$$

Puisque le filtre $\mathbf{g}[z]$ est de norme unité, l'égalité ci-dessus peut s'écrire :

$$\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 (\mathcal{M}_j(k) - \epsilon \mathcal{M}_{\max}) = 0 \quad (5.22)$$

Comme $(\mathcal{M}_j(k) - \epsilon \mathcal{M}_{\max})$ est de signe constant, nous en déduisons que pour tout $(j, k) \in \{1, \dots, N\}$ nous avons soit $|g_j(k)|^2 = 0$, soit $\mathcal{M}_j(k) = \epsilon \mathcal{M}_{\max}$. L'égalité (5.18) découle alors immédiatement si l'on se souvient que $\mathbf{g}[z]$ est de norme unité.

Réciproquement, si (5.18) est vérifiée, il découle de l'égalité $\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 = 1$ que $\forall (j, k) \notin \mathcal{I}, g_j(k) = 0$ et l'égalité $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$ est évidente. ■

Nous sommes à ce stade en mesure de démontrer la propriété suivante qui caractérise les conditions sous lesquelles la fonction $\mathcal{C}_z\{.\}$ est un contraste :

Proposition 21 *Sous contrainte de norme 1, la fonction $\mathcal{C}_z\{.\}$ est un contraste si et seulement si les ensembles \mathcal{I}_+ et \mathcal{I}_- contiennent chacun au plus un élément.*

Preuve: Supposons d'abord que \mathcal{I}_+ et \mathcal{I}_- contiennent au plus un élément. Le lemme 4 assure la majoration nécessaire pour obtenir la propriété de contraste.

Remarquons que $\mathcal{I}_+ \cup \mathcal{I}_-$ contient au moins un élément qui est le couple (i_0, l_0) du lemme 4. Si $\mathcal{C}_z\{.\}$ atteint sa borne supérieure, d'après le lemme 5, $|g_i(k)| = 0$ pour tout $(i, k) \notin \mathcal{I}$, où \mathcal{I} représente l'un des ensembles \mathcal{I}_+ ou \mathcal{I}_- . Comme ces derniers ensembles ont au plus un élément, et que $\mathcal{I}_+ \cup \mathcal{I}_-$ est non vide, c'est donc que \mathcal{I} (c'est-à-dire soit \mathcal{I}_+ soit \mathcal{I}_- selon le cas) a un et un seul élément (j_0, k_0) . Dès lors, nous avons $|g_{j_0}(k_0)| = 1$ et $g_i(k) = 0$ pour tout $(i, k) \neq (j_0, k_0)$, ce qui prouve que le filtre est séparant. La propriété de contraste en découle donc.

Réciproquement, supposons par exemple que \mathcal{I}_+ contienne plus d'un élément. Tout filtre global satisfaisant

$$\begin{cases} \forall (j, k) \notin \mathcal{I}_+ & g_j(k) = 0, \\ \forall (j, k) \in \mathcal{I}_+ & |g_j(k)|^2 = \frac{1}{\text{card}(\mathcal{I}_+)} \end{cases} \quad (5.23)$$

où card signifie cardinal de l'ensemble, rend le critère maximal sans pour autant être séparant au sens de la définition 7. La condition sur les ensembles \mathcal{I}_+ et \mathcal{I}_- est donc non seulement suffisante, mais aussi nécessaire. ■

5.2.1-b) Discussion des conditions de validité du contraste

Nous discutons ici la condition de validité du critère de contraste avec référence donnée par la proposition 21.

Condition sur le système de référence seul En premier lieu, nous pouvons noter que la condition sur les ensembles \mathcal{I}_+ et \mathcal{I}_- est une condition conjointe sur le système de référence et les caractéristiques des sources. Les valeurs des coefficients $\mathcal{M}_j(k)$ définis à l'équation (5.12) dépendent en effet simultanément des coefficients $t_j(k)$ du système de référence et des autocumulants $\kappa_{s_j}(0, 0, 0)$ des sources. Il peut paraître maladroit d'exprimer la condition de validité du contraste comme dépendant de caractéristiques inconnues des sources ; ce point de vue permet cependant d'exprimer des conditions générales.

Le choix du système de référence est libre et il peut être plus naturel d'exprimer une condition sur lui seul. Nous supposons donc les caractéristiques des sources fixées quelconques, dans la mesure toutefois où l'une des sources au moins a un cumulant d'ordre quatre non nul :

$$\exists j \in \{1, \dots, N\}, \quad \kappa_{s_j}(0, 0, 0) \neq 0. \quad (5.24)$$

En toute généralité et par convention, nous supposons que la source numéro 1 a un cumulant plus grand en valeur absolue que tous les autres cumulants des sources. Ceci revient à écrire :

$$\forall j \in \{1, \dots, N\} \quad |\kappa_{s_j}(0, 0, 0)| \leq |\kappa_{s_1}(0, 0, 0)| \quad (5.25)$$

Il est alors immédiat de vérifier que si le système de référence satisfait :

$$\exists l \in \mathbb{Z} \mid \forall (j, k) \in \{1, \dots, N\} \times \mathbb{Z} \quad (j \neq 1 \text{ ou } k \neq l) \Rightarrow |t_j(k)| < |t_1(l)| \quad (5.26)$$

nous avons pour tout $(j, k) \neq (1, l)$

$$|\mathcal{M}_j(k)| = |t_j(k)| |\kappa_{s_j}(0, 0, 0)| < |t_1(l)| |\kappa_{s_1}(0, 0, 0)| = |\mathcal{M}_1(l)| \quad (5.27)$$

et $\mathfrak{J}_+ \cup \mathfrak{J}_-$ se réduit exactement au singleton $\{(1, l)\}$. \mathfrak{J}_+ et \mathfrak{J}_- ont donc chacun au plus un élément et \mathcal{C}_z est un contraste.

Une interprétation possible Une lecture possible de la condition donnée par l'équation (5.26) est la suivante : parmi les coefficients du système de référence global, l'un d'entre eux, $t_1(l)$ est en même temps le plus grand en valeur absolue et correspond à la première source, c'est-à-dire celle dont le cumuland est le plus grand. D'après la preuve de la proposition 21, le critère de contraste avec référence qui vérifie cette condition mène alors à la séparation de cette source avec un retard temporel l donné par le coefficient le plus grand.

De façon plus générale, nous pouvons interpréter la condition de la proposition 21 comme la nécessité de présence majoritaire de l'une des versions retardée des sources dans le signal de référence. Cette présence majoritaire peut être liée soit à une valeur plus importante du cumuland de la source en question, soit à un système de référence privilégiant la source en question. Ainsi, cette approche donne bien une condition moins restrictive pour la validité des contrastes que si cette dernière est exprimée sur le seul filtre de référence, comme nous venons de le faire.

Choix du système de référence Nous avons jusqu'à présent supposé le système de référence donné. Connaissant maintenant les conditions que ce dernier doit respecter, nous pouvons examiner la façon de le choisir.

Supposons que les sources aient toutes les mêmes caractéristiques et qu'il existe un couple d'indices $(i_0, k_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour tout $(i, k) \neq (i_0, k_0)$ on ait $|t_i(k)| < |t_{i_0}(k_0)|$. Par le même raisonnement qu'au paragraphe ci-dessus, il est immédiat de remarquer alors que $\mathfrak{J}_+ \cup \mathfrak{J}_-$ est un singleton et que donc $\mathcal{C}_z\{\cdot\}$ est un contraste. Il suffit donc que l'un des coefficients du filtre de mélange majore strictement les autres pour que le contraste soit valable. Plus généralement, nous pouvons énoncer la proposition suivante, dont la preuve est évidente :

Proposition 22 *Une condition suffisante pour que la fonction $\mathcal{C}_z\{\cdot\}$ soit un contraste (sous contrainte de norme unité) est :*

$$\exists (j_0, k_0) \in \{1, \dots, N\} \times \mathbb{Z} \mid \forall (j, k) \neq (j_0, k_0) \quad |\mathcal{M}_j(k)| < |\mathcal{M}_{j_0}(k_0)| \quad (5.28)$$

La condition de validité d'un contraste avec référence apparaît donc comme très peu contraignante. Plus généralement, nous pouvons affirmer :

Proposition 23 *Supposons qu'il existe une source de cumuland d'ordre quatre non nul (équation (5.24)). Si le système de référence est de réponse impulsionnelle finie et ses coefficients sont tirés aléatoirement selon une loi conjointe continue, alors la condition de la proposition 21 sur les ensembles \mathfrak{J}_+ et \mathfrak{J}_- est presque sûrement vérifiée.*

Preuve: En effet, l'hypothèse sur les sources assure que \mathcal{M}_{\max} est non nul. De plus, comme indiqué dans la preuve de la proposition 21, $\mathcal{J}_+ \cup \mathcal{J}_-$ est non vide et ainsi il existe un couple (i_0, l_0) (le couple du lemme 4) tel que $\mathcal{M}_{\max} = \epsilon \mathcal{M}_{i_0}(l_0)$ où $\epsilon \in \{-1, +1\}$.

Un couple $(j, k) \neq (i_0, l_0)$ appartient donc à \mathcal{J}_+ si et seulement si $\mathcal{M}_j(k) = \epsilon \mathcal{M}_{i_0}(l_0)$, ce qui s'écrit aussi $|t_j(k)|^2 \kappa_{s_j}(0, 0, 0) = \epsilon |t_{i_0}(l_0)|^2 \kappa_{i_0}(0, 0, 0)$. Selon les valeurs de ϵ et $\kappa_{s_j}(0, 0, 0)$, ceci est ou bien impossible, ou alors arrive avec une probabilité nulle lorsque les coefficients $(t_j(k))$, $(j, k) \in \{1, \dots, N\} \times \mathbb{Z}$ sont tirés selon une loi continue. Le même raisonnement peut être mené sur \mathcal{J}_- , ce qui prouve que presque sûrement, les ensembles \mathcal{J}_+ et \mathcal{J}_- ont au plus un élément. ■

5.2.1-c) Généralisation pour des cumulants à tous ordres

Nous expliquons dans ce paragraphe comment les résultats du paragraphe précédents peuvent se généraliser à d'autres cumulants d'ordre supérieur ou égal à 3. Par ailleurs, il est possible de choisir librement de conjuguer ou non les signaux dans les cumulants. Par commodité, utilisons le symbole $\mathbf{y}(n)$ pour désigner au choix le signal $y(n)$ ou son conjugué $y^*(n)$. De même, $\mathbf{z}(n)$ désigne au choix $z(n)$ ou $z^*(n)$. Nous pouvons alors définir pour tout $R \geq 3$ la forme générale du cumulant pouvant servir à l'élaboration d'un contraste :

$$\kappa_R\{\mathbf{y}(n), \mathbf{z}(n)\} := \text{Cum}\{\mathbf{y}(n), \mathbf{y}(n), \underbrace{\mathbf{z}(n), \dots, \mathbf{z}(n)}_{R-2 \text{ fois}}\} \quad (5.29)$$

Les propriétés du paragraphe précédent se généralisent alors au critère :

$$\mathbf{C}_z\{\mathbf{y}(n)\} := |\kappa_R\{\mathbf{y}(n), \mathbf{z}(n)\}| \quad (5.30)$$

Comme le cumulant ci-dessus n'est plus nécessairement réel, il convient au lieu des ensembles \mathcal{J}_+ et \mathcal{J}_- de considérer l'ensemble :

$$\mathcal{J} = \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid |\mathcal{M}_j(k)| = \mathcal{M}_{\max}\} \quad (5.31)$$

Les propositions du paragraphe précédent se généralisent alors aisément et permettent d'affirmer :

Proposition 24 *Sous contrainte de norme 1, la fonction $\mathbf{C}_z\{\cdot\}$ est un contraste si et seulement si l'ensemble \mathcal{J} contient un élément et un seul.*

La démonstration de cette proposition est quasiment identique à celle déjà faite. Comme le seul cas pratique considéré en simulation est celui des cumulants d'ordre quatre, nous ne démontrons pas ce résultat. Remarquons aussi que pour des signaux sources complexes circulaires, il convient pour obtenir un contraste que le nombre de conjugués et de non conjugués s'équilibre dans (5.29). A défaut, nous aurions en effet $\mathcal{M}_{\max} = 0$ et $\mathcal{J} = \{1, \dots, N\} \times \mathbb{Z}$.

5.2.2 Cas de sources non i.i.d.

Nous étudions à présent le cas de sources non linéaires (et non i.i.d.). Nous allons d'abord montrer que, sous certaines hypothèses techniques, les critères qui viennent d'être introduits constituent des contrastes pour des sources non i.i.d. au sens de la définition 6. L'étude des hypothèses techniques et la recherche de conditions suffisantes pour qu'elles soient satisfaites nous préoccupera dans un second temps.

5.2.2-a) Validité des contrastes avec référence dans le cas non i.i.d.

La démonstration de la validité des critères de contraste avec référence présente de fortes similitudes avec celle de la proposition 9 du chapitre 3. Pour tout i , nous introduisons ainsi une borne supérieure du critère lorsque seule la $i^{\text{ème}}$ source est présente :

$$\forall i \in \{1, \dots, N\} \quad \mathcal{M}_i := \sup_{\mathbf{g} \in \mathcal{G}_i} \mathcal{C}_z\{\mathbf{g}[z]\mathbf{s}(n)\} \quad (5.32)$$

Contrairement au cas i.i.d. il n'est pas assuré que cette borne supérieure soit atteinte. De même qu'au chapitre 3, nous poserons donc cette hypothèse technique :

H.16 Pour tout $i \in \{1, \dots, N\}$, la borne supérieure définie à l'équation (5.33) est atteinte pour un filtre noté $\mathbf{g}_i^\# \in \mathcal{G}_i$.

Enfin, par analogie avec le cas i.i.d. du paragraphe 5.2.1, nous définissons :

$$\mathcal{M}_{\max} := \max_{j=1}^N \left(\mathcal{M}_j \right) = \max_{j=1}^N \left(\sup_{\mathbf{g} \in \mathcal{G}_j} \mathcal{C}_z\{\mathbf{g}[z]\mathbf{s}(n)\} \right) \quad (5.33)$$

La proposition suivante peut alors être énoncée :

Proposition 25 *Sous contrainte de norme 1 et sous l'hypothèse H.16, la fonction $\mathcal{C}_z\{\cdot\}$ est un contraste si et seulement si l'ensemble*

$$\mathfrak{J}' := \{j \in \{1, \dots, N\} \mid \mathcal{M}_j = \mathcal{M}_{\max}\} \quad (5.34)$$

contient un unique élément.

Preuve: De même qu'à l'équation (3.11), la sortie globale $y(n)$ peut s'écrire :

$$y(n) = \sum_{j=1}^N g_j[z]s_j(n) = \sum_{j=1}^N \|g_j\|_j \tilde{g}_j[z]s_j(n) = \sum_{j=1}^N \|g_j\|_j \tilde{y}_j(n) \quad (5.35)$$

où l'on a posé :

$$\tilde{g}_j[z] := \begin{cases} \frac{g_j[z]}{\|g_j\|_j} & \text{si } \|g_j\|_j \neq 0 \\ 0 & \text{sinon} \end{cases} \quad (5.36)$$

et :

$$\tilde{y}_j(n) := \tilde{g}_j[z]s_j(n) \quad (5.37)$$

Par multilinéarité des cumulants et indépendance des $(\tilde{y}_j(n))_{n \in \mathbb{Z}}$, $j \in \{1, \dots, N\}$, il vient :

$$\kappa_4\{y(n), z(n)\} = \sum_{j=1}^N \|g_j\|_j^2 \kappa_4\{\tilde{y}_j(n), z(n)\} \quad (5.38)$$

Nous pouvons alors effectuer les majorations suivantes :

$$\mathcal{C}_z\{y(n)\} = |\kappa_4\{y(n), z(n)\}| \leq \sum_{j=1}^N \|g_j\|_j^2 |\kappa_4\{\tilde{y}_j(n), z(n)\}| \quad (5.39)$$

$$\leq \sum_{j=1}^N \|g_j\|_j^2 \mathcal{C}_z\{\tilde{y}_j(n)\} \quad (5.40)$$

et donc, se rappelant (5.32), (5.37) et que chacun des filtres scalaires $\tilde{g}_j[z]$ est de j -norme unité :

$$\mathcal{C}_z\{y(n)\} \leq \sum_{j=1}^N \|g_j\|_j^2 \mathcal{M}_j \quad (5.41)$$

$$\leq \mathcal{M}_{\max} \sum_{j=1}^N \|g_j\|_j^2 \leq \mathcal{M}_{\max} \quad (5.42)$$

La propriété de majoration nécessaire à la validité du contraste est donc prouvée.

Par ailleurs, en cas d'égalité ci-dessus (ce qui se produit pour au moins un filtre global compte tenu de l'hypothèse H.16), en exploitant l'égalité entre (5.41) et (5.42), il vient :

$$\sum_{j=1}^N \|g_j\|_j^2 (\mathcal{M}_{\max} - \mathcal{M}_j) = 0 \quad (5.43)$$

et ainsi :

$$\forall j \in \{1, \dots, N\} \quad \begin{cases} \|g_j\|_j = 0 \\ \mathcal{M}_j = \mathcal{M}_{\max}, \end{cases} \quad \text{ou bien :} \quad (5.44)$$

ce qui donne encore $\forall j \notin \mathcal{J}', \|g_j\|_j = 0$. Si \mathcal{J}' contient un unique élément j_0 , nous avons alors $\|g_j\|_j = 0$ si $j \neq j_0$ et $\|g_{j_0}\|_{j_0} = 1$ compte tenu de la norme unité du filtre global. Ce dernier est donc séparant, ce qui prouve la propriété de contraste.

Dans le cas où \mathcal{J}' contient en revanche plusieurs éléments dont entre autres j_1 et j_2 , le filtre global défini par $\frac{1}{\sqrt{2}}(\mathbf{g}_{j_1}^\# [z] + \mathbf{g}_{j_2}^\# [z])$ atteint le maximum du critère sans être pour autant séparant. La condition sur \mathcal{J}' est donc nécessaire. ■

Remarque 13: L'hypothèse H.16 peut en réalité être affaiblie en :

H.17 Pour tout $i \in \mathcal{J}'$, la borne supérieure définie à l'équation (5.33) est atteinte pour un filtre noté $\mathbf{g}_i^\# \in \mathcal{G}_i$.

Au cours de la démonstration, nous utilisons en effet le fait que la borne supérieure est atteinte uniquement pour le filtre séparant correspondant à un indice dans l'ensemble \mathcal{J}' .

5.2.2-b) Condition suffisante de validité du contraste

Nous cherchons dans ce paragraphe une condition suffisante pour pouvoir appliquer la proposition 25. Comme nous allons le constater, la généralisation de la proposition 22 au cas non i.i.d. est plus délicate et l'interprétation de son équivalent est moins immédiate. Aussi, le lecteur pourra-t-il sauter ce paragraphe au cours d'une première lecture.

Examinons d'abord le cas particulier où la référence consiste en la filtrée scalaire de l'une des sources, par exemple la $j_0^{\text{ème}}$. Soit $j \in \{1, \dots, N\}, j \neq j_0$, et soit un filtre global $\mathbf{g} \in \mathcal{G}_j$ dont seule la $j^{\text{ème}}$ composante est non nulle. La sortie globale $y(n) = \mathbf{g}[z]\mathbf{s}(n)$, qui ne dépend que de $(s_j(n)), n \in \mathbb{Z}$, est alors indépendante de la référence, qui ne dépend que de $(s_{j_0}(n))_{n \in \mathbb{Z}}$. Le contraste (5.10) est donc nul et on en déduit par conséquent que pour tout $j \neq j_0$, $\mathcal{M}_j = 0$. Ainsi, l'ensemble \mathcal{J}' de la proposition 25 ne contient qu'un élément, si bien sûr les cumulants d'ordre quatre des sources sont non nuls. La condition de la proposition 25 est donc vérifiée lorsque la référence est une filtrée scalaire de l'une des sources et ce résultat est conforme à nos attentes.

Nous donnons maintenant une autre condition suffisante, généralisant la proposition 22 :

Proposition 26 Une condition suffisante pour que la condition de la proposition 25 soit vérifiée est qu'il existe $j_0 \in \{1, \dots, N\}$ tel que :

$$\forall j \neq j_0 \quad \|t_{j_0}\|_\infty^2 |\kappa_{s_{j_0}}(0, 0, 0)| > (\|t_{j_0}\|_1^2 - \|t_{j_0}\|_\infty^2) \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \\ + (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1, \quad (5.45)$$

où pour toute famille de nombres $(a(k))_{k \in \mathcal{K}}$ on a défini les normes :

$$\|a\|_1 := \sum_{k \in \mathcal{K}} |a(k)| \\ \|a\|_\infty := \sup_{k \in \mathcal{K}} |a(k)|$$

et où pour chacune des sources $(s_j(n))_{n \in \mathbb{Z}}$, $j \in \{1, \dots, N\}$ on a supposé la densité spectrale minorée par $\Gamma_j^{\min} > 0$.

Preuve: La démonstration est quelque peu fastidieuse et se déroule en trois temps : minoration des \mathcal{M}_j , majoration des \mathcal{M}_j et enfin conclusion sur la proposition.

Minoration des \mathcal{M}_j Par multilinéarité des cumulants et indépendance des sources, on a

$$\kappa_4\{\tilde{y}_j(n), z(n)\} = \\ \sum_{k_1, k_2, l_1, l_2} \tilde{g}_j(k_1) \tilde{g}_j^*(k_2) t_j(l_1) t_j^*(l_2) \text{Cum}\{s_j(n - k_1), s_j^*(n - k_2), s_j(n - l_1), s_j^*(n - l_2)\} \\ = \sum_{k_1, k_2, l_1, l_2} \tilde{g}_j(k_1) \tilde{g}_j^*(k_2) t_j(l_1) t_j^*(l_2) \kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2). \quad (5.46)$$

Puisque (5.32) s'écrit aussi $\mathcal{M}_j := \sup_{\|\tilde{g}_j\|=1} \mathcal{C}_z\{\tilde{g}_j[z]s_j(n)\}$, en prenant le filtre particulier de réponse impulsionnelle $(\delta_{n-p})_{n \in \mathbb{Z}}$, on en déduit que pour tout $p \in \mathbb{Z}$:

$$\mathcal{M}_j \geq \left| \sum_{k_1, k_2, l_1, l_2} \delta(k_1 - p) \delta(k_2 - p) t_j(l_1) t_j^*(l_2) \kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2) \right| \\ = \left| \sum_{l_1, l_2} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right| \\ \geq |t_j(p)|^2 |\kappa_{s_j}(0, 0, 0)| - \left| \sum_{(l_1, l_2) \neq (p, p)} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right|.$$

De plus, nous pouvons écrire :

$$\left| \sum_{(l_1, l_2) \neq (p, p)} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right| \\ \leq \sup_{(l_1, l_2) \neq (p, p)} |\kappa_{s_j}(0, p - l_1, p - l_2)| \sum_{(l_1, l_2) \neq (p, p)} |t_j(l_1)| |t_j(l_2)| \\ = \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| (\|t_j\|_1^2 - |t_j(p)|^2).$$

Les deux inégalités précédentes mènent, pour tout $p \in \mathbb{Z}$, à :

$$\mathcal{M}_j \geq |t_j(p)|^2 \left(|\kappa_{s_j}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \right) - \|t_j\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \quad (5.47)$$

et en prenant la borne supérieure sur l'ensemble des $|t_j(p)|$, $p \in \mathbb{Z}$:

$$\mathcal{M}_j \geq \|t_j\|_\infty^2 \left(|\kappa_{s_j}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \right) - \|t_j\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \quad (5.48)$$

Majoration des \mathcal{M}_j D'après (5.46), nous pouvons écrire :

$$\begin{aligned} |\kappa_4\{\tilde{y}_j(n), z(n)\}| &\leq \sum_{k_1, k_2, l_1, l_2} |\tilde{g}_j(k_1)| |\tilde{g}_j^*(k_2)| |t_j(l_1)| |t_j^*(l_2)| |\kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2)| \\ &\leq \|t_j\|_\infty^2 \sum_{k_1, k_2} |\tilde{g}_j(k_1)| |\tilde{g}_j(k_2)| \sum_{l_1, l_2} |\kappa_{s_j}(k_1 - k_2, l_1, l_2)| \end{aligned} \quad (5.49)$$

En changeant de variable $k = k_1, p = k_1 - k_2$ et en réarrangeant, il vient :

$$|\kappa_4\{\tilde{y}_j(n), z(n)\}| \leq \|t_j\|_\infty^2 \sum_{l_1, l_2} \sum_p |\kappa_{s_j}(p, l_1, l_2)| \sum_k |\tilde{g}_j(k)| |\tilde{g}_j(k - p)|$$

et comme $\sum_k |\tilde{g}_j(k)| |\tilde{g}_j(k - p)| \leq \sum_k |\tilde{g}_j(k)|^2 \leq (\Gamma_j^{\min})^{-1} \|\tilde{g}_j\|_j^2 = (\Gamma_j^{\min})^{-1}$:

$$|\kappa_4\{\tilde{y}_j(n), z(n)\}| \leq \|t_j\|_\infty^2 \sum_{l_1, l_2} \sum_p |\kappa_{s_j}(p, l_1, l_2)| = (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1 \quad (5.50)$$

La borne supérieure du terme à gauche ci-dessus vérifie donc :

$$\mathcal{M}_j \leq (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1. \quad (5.51)$$

Preuve proposition 26 Une condition suffisante pour que \mathfrak{J}' de l'équation (5.34) se réduise au singleton $\{j_0\}$ est :

$$\begin{aligned} (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1 &< \|t_{j_0}\|_\infty^2 \left(|\kappa_{s_{j_0}}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \right) \\ &\quad - \|t_{j_0}\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \end{aligned} \quad (5.52)$$

Le membre de gauche est en effet celui de la majoration de \mathcal{M}_j dans (5.51) et celui de droite est celui de la minoration de \mathcal{M}_{j_0} dans (5.48). Il s'ensuit donc $\mathcal{M}_j < \mathcal{M}_{j_0}$ pour tout $j \neq j_0$, ce qui assure que l'ensemble \mathfrak{J}' de l'équation est un singleton. On vérifie alors immédiatement l'équivalence entre (5.45) et (5.52). ■

Comme le système de référence a été supposé stable (hypothèse H.15), les différentes normes $\|\cdot\|_1$, $\|\cdot\|$ et $\|\cdot\|_\infty$ sur le système de référence existent et de plus, il existe $p^\sharp \in \mathbb{Z}$ tel que $|t_{j_0}(p^\sharp)| = \|t_{j_0}\|_\infty$. La condition (5.45) de la proposition 26 peut s'interpréter alors comme le fait que $t_{j_0}(p^\sharp)$ est le terme «dominant» dans $(t_j(k))_{1 \leq j \leq N, k \in \mathbb{Z}}$. Cette même condition (5.45) requiert également que $|\kappa_{s_{j_0}}(0, 0, 0)|$ ne soit pas trop petit par rapport à $\sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)|$, ce qui n'est pas nécessairement vérifiée pour tout signal non i.i.d. comme l'a prouvé le contre-exemple du paragraphe 2.3.2-b). Nous constatons ainsi que la condition obtenue de validité du contraste avec référence peut s'interpréter de façon analogue au cas i.i.d. : il s'agit d'une condition conjointe sur les sources et le système de référence, qui assure la présence majoritaire de l'une des sources dans le signal de référence.

Remarque 14: Dans le cas du paragraphe précédent où les $(s_j(n))_{n \in \mathbb{Z}}$ forment des suites i.i.d., on a $\Gamma_j^{\min} = 1$ et $\kappa_{s_j}(k_1, k_2, k_3) = 0$ pour tout $(k_1, k_2, k_3) \neq (0, 0, 0)$. Par suite, $\|\kappa_{s_j}\|_1 = |\kappa_{s_j}(0, 0, 0)|$ et (5.45) se simplifie en :

$$\forall j \neq j_0 \quad \|t_{j_0}\|_\infty^2 |\kappa_{s_{j_0}}(0, 0, 0)| > \|t_j\|_\infty^2 |\kappa_{s_j}(0, 0, 0)|. \quad (5.53)$$

Nous retrouvons donc bien la condition (5.28) obtenue dans le cas i.i.d.

De plus, puisque (5.45) implique que (5.53) est vérifiée, nous constatons que la condition de validité obtenue dans le cas non i.i.d. est plus restrictive que celle du cas i.i.d. (avec équivalence dans le cas de sources i.i.d.).

5.3 Éléments pour la mise en œuvre

5.3.1 Expression des formes quadratiques : critère et contrainte

De même que dans le paragraphe 3.4.2, nous supposons maintenant que le filtre séparant est RIF. Sa réponse impulsionnelle est donnée par $\mathbf{w}(0), \dots, \mathbf{w}(D-1)$ et nous définissons le vecteur suivant de taille $1 \times QD$:

$$\underline{\mathbf{w}} := (\mathbf{w}(0) \quad \mathbf{w}(1) \quad \dots \quad \mathbf{w}(D-1)) \quad (5.54)$$

ainsi que pour tout $n \in \mathbb{Z}$ le vecteur colonne $\underline{\mathbf{x}}(n) := (\underline{x}_1(n), \dots, \underline{x}_{QD}(n))^T$ de taille $QD \times 1$:

$$\underline{\mathbf{x}}(n) := (\mathbf{x}(n)^T \quad \mathbf{x}(n-1)^T \quad \dots \quad \mathbf{x}(n-D+1)^T)^T \quad (5.55)$$

La sortie du filtre global s'exprime alors :

$$y(n) = \underline{\mathbf{w}} \underline{\mathbf{x}}(n) \quad (5.56)$$

Les observations étant données, $\underline{\mathbf{x}}(n)$ est donné et $y(n)$ s'exprime donc en fonction de $\underline{\mathbf{w}}$ seulement. Dès lors, la méthode proposée pour l'extraction d'une source se ramène, d'après la proposition 21, au problème d'optimisation :

$$\text{maximiser } \mathcal{C}_z\{y(n)\} \quad \text{sous la contrainte : } \mathbb{E}\{|y(n)|^2\} = 1, \quad (5.57)$$

où la variable à optimiser est le vecteur $\underline{\mathbf{w}}$. En effet, à condition de définir les notations qui suivent, les expressions apparaissant ci-dessus peuvent être données explicitement en fonction de $\underline{\mathbf{w}}$. Définissons

$$\mathbf{R} := \mathbb{E}\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\} \quad (5.58)$$

et $\mathbf{C} = (C_{ij})_{(i,j) \in \{1, \dots, QD\}^2}$ la matrice dont les éléments sont donnés par :

$$\forall (i, j) \in \{1, \dots, QD\}^2 \quad C_{ij} := \text{Cum}\{\underline{x}_i(n), \underline{x}_j^*(n), z(n), z^*(n)\}. \quad (5.59)$$

Les définitions (5.58) et (5.59) permettent alors d'écrire :

$$\mathbb{E}\{|y(n)|^2\} = \underline{\mathbf{w}} \mathbf{R} \underline{\mathbf{w}}^H \quad \kappa_4\{y(n), z(n)\} = \underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H \quad (5.60)$$

et donc aussi :

$$\mathcal{C}_z\{y(n)\} := |\kappa_4\{y(n), z(n)\}| = |\underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H|. \quad (5.61)$$

Remarquons au passage que les matrices \mathbf{R} ainsi que \mathbf{C} sont hermitiennes, ce qui est d'ailleurs en conformité avec le fait que les quantités ci-dessus sont réelles. A ce stade, le problème (5.57) se réduit au problème d'optimisation suivant :

$$\max |\underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H| \quad \text{sous la contrainte : } \underline{\mathbf{w}} \mathbf{R} \underline{\mathbf{w}}^H = 1. \quad (5.62)$$

Il s'agit d'un problème dont le critère est une valeur absolue d'une fonction quadratique et dont la contrainte est quadratique. Nous allons montrer ci-dessous qu'il est aisé d'en obtenir une solution.

5.3.2 Restriction du problème

5.3.2-a) Définie positivité de \mathbf{R} ou positivité seulement ?

Dans le cas où \mathbf{R} est définie positive, l'ensemble des vecteurs $\underline{\mathbf{w}}$ tels que $\underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = 1$ est un compact et, grâce à la continuité, nous sommes assurés de l'existence d'une borne supérieure finie au critère que nous souhaitons maximiser. Cette borne est par ailleurs atteinte pour au moins un vecteur, qui constitue alors une solution au problème donné à l'équation (5.62). L'existence d'une solution au problème est donc acquise lorsque \mathbf{R} est définie positive.

Cependant, la matrice \mathbf{R} n'est en général que positive non définie. C'est en particulier le cas lorsque :

- le mélange est instantané et il y a strictement plus de capteurs que de sources. Dans ce cas en effet, $\underline{\mathbf{x}}(n) = \mathbf{x}(n) = \mathbf{M}(0)\mathbf{s}(n)$ et ainsi $\mathbf{R} = \mathbf{E}\{\mathbf{x}(n)\mathbf{x}(n)^H\} = \mathbf{M}(0)\mathbf{M}(0)^H$ n'est pas de rang plein.
- une des sources est identiquement nulle. C'est, dans l'idéal, ce qui arrive après une étape de déflation. Ce cas sera étudié au paragraphe 5.3.4.
- de façon générale \mathbf{R} n'est pas de rang plein dès que les nombres de capteurs et de sources vérifient $Q > N + 1$ et que la longueur du filtre séparant est choisie comme la longueur minimale assurant l'existence d'un inverse ($D = N(L - 1)$).

Compte tenu de ce fait, rien ne nous assure *a priori* que le critère à maximiser est borné ni que le problème posé admet une solution.

5.3.2-b) Non unicité du filtre séparant

Nous nous plaçons maintenant dans le cas où \mathbf{R} est positive non définie. Compte tenu de sa définition, ceci survient lorsqu'il existe une relation de dépendance linéaire entre les signaux $\underline{x}_i(n), i \in \{1, \dots, QD\}$ contenus dans $\underline{\mathbf{x}}(n)$. Il s'ensuit alors que le filtre séparant n'est pas unique.

Supposons en effet $\underline{\mathbf{w}}_0^H \in \text{Ker } \mathbf{R}$. Alors, $\underline{\mathbf{w}}_0\mathbf{R}\underline{\mathbf{w}}_0^H = \mathbf{E}\{\underline{\mathbf{w}}_0\mathbf{x}(n)\mathbf{x}(n)^H\underline{\mathbf{w}}_0^H\} = 0$ et le signal $\underline{\mathbf{w}}_0\underline{\mathbf{x}}(n)$ est alors identiquement nul. Ceci montre que deux filtres $\underline{\mathbf{w}}$ et $\underline{\mathbf{w}} + \underline{\mathbf{w}}_0$ avec $\underline{\mathbf{w}}_0^H \in \text{Ker } \mathbf{R}$ donnent le même signal de sortie au niveau du séparateur.

5.3.2-c) Contrainte supplémentaire

Compte tenu des remarques précédentes, nous imposons par conséquent dans le problème d'optimisation la contrainte supplémentaire suivante :

$$\underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp \quad (5.63)$$

et le problème d'optimisation à résoudre s'écrit finalement :

$$\max |\underline{\mathbf{w}}\mathbf{C}\underline{\mathbf{w}}^H| \quad \text{sous les contraintes : } \begin{cases} \underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = 1 \text{ et :} \\ \underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp. \end{cases} \quad (5.64)$$

L'ensemble des vecteurs $\underline{\mathbf{w}}$ satisfaisant les deux contraintes ci-dessus est un compact et l'existence d'une solution est ainsi garantie. La discussion du paragraphe précédent assure par ailleurs qu'il est suffisant de considérer cet ensemble pour décrire l'ensemble des sorties possibles au niveau du séparateur.

5.3.3 Maximisation exacte par décomposition en valeurs singulières

Bien que les résultats soient connus, nous expliquons maintenant pour des raisons d'exhaustivité et de clarté, la façon simple dont le problème d'optimisation (5.64) peut se résoudre. La méthode est basée sur le lemme suivant :

Lemme 6 *Soit $p \in \mathbb{N}^*$ et $\tilde{\mathbf{C}} \in \mathbb{C}^{p \times p}$ une matrice symétrique hermitienne ($\tilde{\mathbf{C}}^H = \tilde{\mathbf{C}}$). Les vecteurs de $\mathbb{C}^{p \times 1}$ solutions du problème d'optimisation*

$$\max |\tilde{\mathbf{x}}^H \tilde{\mathbf{C}} \tilde{\mathbf{x}}| \quad \text{sous la contrainte : } \tilde{\mathbf{x}}^H \tilde{\mathbf{x}} = 1 \quad (5.65)$$

sont les vecteurs appartenant à l'intersection de la sphère unité ($\tilde{\mathbf{x}}^H \tilde{\mathbf{x}} = 1$) et de la réunion des sous-espaces propres de $\tilde{\mathbf{C}}$ relatifs aux valeurs propres maximales en valeur absolue.

Voyons la façon dont nous nous ramenons au lemme précédent. \mathbf{R} étant hermitienne, elle est diagonalisable dans une base orthogonale. \mathbf{R} étant positive, mais non nécessairement définie, nous supprimons dans cette diagonalisation les valeurs propres nulles et les colonnes correspondantes et l'écrivons sous la forme $\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{U}^H$. Dans cette écriture, \mathbf{S} est carrée, diagonale. De plus, sa taille est celle de \mathbf{R} uniquement dans le cas où cette dernière est définie positive. De même, \mathbf{U} est semi-unitaire (carrée unitaire uniquement si \mathbf{R} est définie positive). Nous définissons alors :

$$\mathbf{P} := \mathbf{S}^{\frac{1}{2}}\mathbf{U}^H \quad \text{et : } \mathbf{Q} := \mathbf{U}\mathbf{S}^{-\frac{1}{2}} \quad (5.66)$$

\mathbf{P} et \mathbf{Q} sont pseudo-inverses l'une de l'autre au sens de Moore-Penrose. Par ailleurs, l'espace engendré par \mathbf{Q} est $\text{Im } \mathbf{Q} = \text{Im } \mathbf{R} = (\text{Ker } \mathbf{R})^\perp$. La contrainte $\underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp$ du problème (5.64) est donc prise en compte si l'on écrit :

$$\underline{\mathbf{w}}^H = \mathbf{Q}\tilde{\mathbf{x}} \quad (\text{ce qui revient à poser : } \tilde{\mathbf{x}} = \mathbf{P}\underline{\mathbf{w}}^H). \quad (5.67)$$

On a alors :

$$\underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = \underline{\mathbf{w}}\mathbf{U}\mathbf{S}\mathbf{U}^H\underline{\mathbf{w}}^H = \underline{\mathbf{w}}\mathbf{P}^H\mathbf{P}\underline{\mathbf{w}}^H = \tilde{\mathbf{x}}^H\tilde{\mathbf{x}}, \quad (5.68)$$

et en posant $\tilde{\mathbf{C}} := \mathbf{Q}^H\mathbf{C}\mathbf{Q}$, nous sommes bien ramenés au problème du lemme 6.

Par conséquent, une solution exacte au problème peut être obtenue par une diagonalisation (ou décomposition en valeurs singulières) de la matrice $\tilde{\mathbf{C}}$. L'avantage que l'on en tire est double : d'une part, la maximisation du contraste ne requiert qu'un nombre fini d'opérations. La rapidité de la maximisation est ainsi fortement accrue et il n'est pas nécessaire de recourir à un algorithme itératif d'optimisation qui se révèle bien plus lourd. D'autre part, le maximum *global* est obtenu et la méthode est donc insensible à l'existence de maxima locaux potentiellement non séparants.

Remarque 15: Le problème d'optimisation défini par (5.62) peut aussi être abordé au travers de l'étude du Lagrangien. On tombe alors sur le problème de recherche de valeurs propres généralisées λ et de vecteurs propres généralisés $\underline{\mathbf{w}}^H$, qui s'écrit :

$$\mathbf{C}\underline{\mathbf{w}}^H = \lambda\mathbf{R}\underline{\mathbf{w}}^H \quad (5.69)$$

La résolution de ce problème est équivalente à l'approche qui a été présentée.

5.3.4 Procédé séquentiel de déflation

Disposant maintenant de tous les éléments pour extraire une source, nous considérons le problème d'une séparation MIMO de l'ensemble des sources. La solution la plus simple consiste à utiliser une procédure de déflation telle qu'elle a été présentée dans le paragraphe 3.2.1.

Nous avons expliqué qu'il fallait dans la méthode d'optimisation tenir compte du rang de la matrice \mathbf{R} définie à l'équation (5.58). Examinons l'effet de la suppression d'une source sur le rang de cette matrice. Nous définissons pour cela $\underline{\mathbf{s}}(n)$ le vecteur colonne de dimension $N(L + D - 1)$:

$$\underline{\mathbf{s}}(n) := (\mathbf{s}(n)^T \quad \mathbf{s}(n-1)^T \quad \dots \quad \mathbf{s}(n-L-D+2)^T)^T \quad (5.70)$$

et \mathcal{T}_D la matrice de Sylvester suivante de taille $QD \times N(L + D - 1)$:

$$\mathcal{T}_D(\mathbf{M}) := \begin{pmatrix} \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) & 0 & \dots & 0 \\ 0 & \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) \end{pmatrix}. \quad (5.71)$$

Le vecteur $\underline{\mathbf{x}}(n)$ défini à l'équation (5.55) s'exprime alors par :

$$\underline{\mathbf{x}}(n) = \mathcal{T}_D(\mathbf{M})\underline{\mathbf{s}}(n) \quad (5.72)$$

$$(5.73)$$

ce qui mène à l'expression suivante de la matrice \mathbf{R} :

$$\mathbf{R} := E\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\} = \mathcal{T}_D(\mathbf{M})E\{\underline{\mathbf{s}}(n)\underline{\mathbf{s}}(n)^H\}\mathcal{T}_D(\mathbf{M})^H \quad (5.74)$$

La suppression de la contribution d'une source au cours d'une étape de déflation est équivalente, dans le cas idéal, à considérer que simultanément sont mis à zéros l'une des composantes du vecteur $\mathbf{s}(n)$ et la colonne correspondante du filtre $\mathbf{M}[z]$. Ceci entraîne donc la mise à zéro simultanément de $L + D - 1$ composantes de $\underline{\mathbf{s}}(n)$ ou de $L + D - 1$ colonnes de $\mathcal{T}_D(\mathbf{M})$, ce qui entraîne une chute équivalente du rang de la matrice \mathbf{R} .

Nous en concluons qu'il convient de prendre en compte dans les procédures précédentes le fait que le rang de \mathbf{R} a chuté de $(D+L-1)$ fois le nombre de contributions déjà soustraites des observations. Les valeurs singulières de \mathbf{R} les plus petites sont donc considérées en nombre correspondant comme nulles dans la description de la procédure de maximisation du paragraphe 5.3.3.

5.4 Simulations

Nous présentons les simulations en considérant d'abord l'extraction d'une seule source. La possibilité d'itérer par processus de déflation sera ensuite vérifiée.

5.4.1 Extraction d'une source à l'aide d'une référence

5.4.1-a) Utilisation de la source comme référence

Un premier test a été mené dans le cas idéal où la source réelle sert de référence. Les résultats sont donnés sur la figure 5.2. Les erreurs quadratiques moyennes sur la source

reconstituée ont été tracées en fonction du nombre d'échantillons disponibles. Ces résultats proviennent d'un moyennage sur 1000 réalisations où les filtres de mélanges, de longueur 3, ont des coefficients tirés aléatoirement selon une loi gaussienne centrée et normalisée. Les simulations ont été réalisées pour des systèmes de mélange à 2 entrées/3 sorties et 3 entrées/4 sorties, ainsi que pour deux types de sources distinctes : i.i.d. uniformes et i.i.d. PAM4. L'observation des EQM moyennes de la figure 5.2 est encourageante et nous invite à engager une étude plus approfondie de la méthode et à considérer le problème de l'obtention d'une référence dans le cadre aveugle.

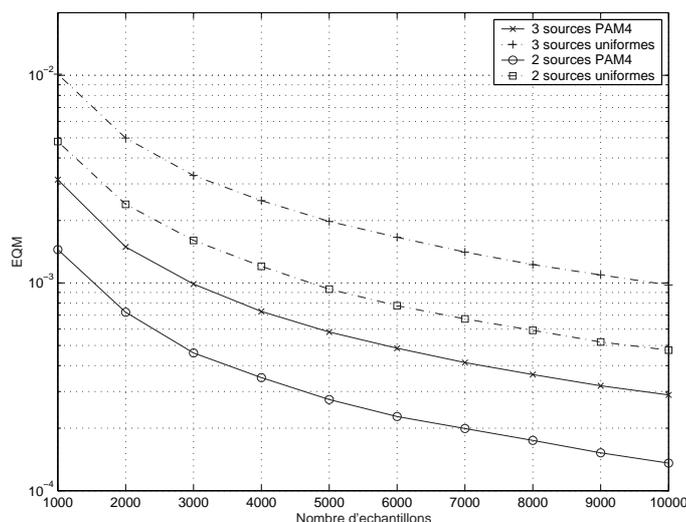


FIG. 5.2 – EQM moyenne sur les sources reconstituées en fonction du nombre d'échantillons. Les valeurs ont été obtenues avec des filtres de mélange de longueur 3, et à l'aide d'un moyennage sur 1000 réalisations.

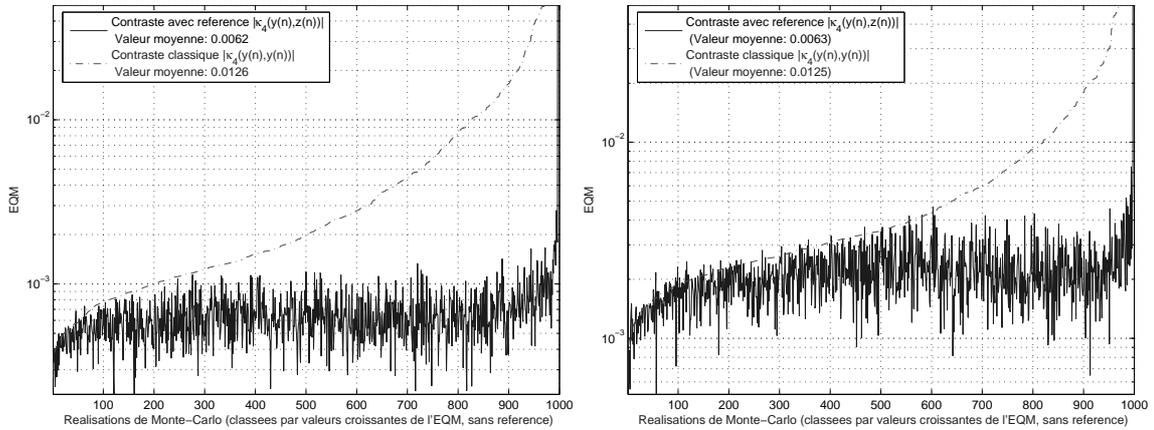
5.4.1-b) Utilisation initiale d'un contraste sans référence

Convergence de l'optimisation complète du contraste sans référence Pour obtenir les résultats précédents (figure 5.2), nous avons artificiellement supposé connue l'une des sources que nous avons utilisée comme signal de référence. Cette méthode n'est naturellement pas applicable en pratique et nous proposons de considérer le cas où la référence utilisée est une source extraite à l'aide d'un contraste MISO tel que ceux présentés au chapitre 3.

Les résultats sont donnés pour les deux types de sources précédemment considérées (i.i.d. uniformes ou PAM4) sur les graphiques de la figure 5.3. Ils représentent, pour chacune des 1000 réalisations de Monte-Carlo, l'EQM sur la source reconstituée pour la méthode sans référence et pour la méthode avec. Pour des raisons de lisibilité, les réalisations ont été classées par ordre croissant de l'EQM obtenue par la méthode sans référence.

Nous constatons que l'utilisation d'un contraste avec référence améliore sensiblement la qualité de la séparation. Ceci illustre un premier intérêt des développements de ce chapitre.

L'apport apparaît toutefois plus nettement si l'on compare les charges de calculs induites par chacune des méthodes : à titre indicatif, pour un mélange avec un filtre 3 sources/4 capteurs, de longueur 3 et 5000 échantillons, la même séparation dure de l'ordre



(a) Sources i.i.d. PAM4.

(b) Sources i.i.d. uniformes.

FIG. 5.3 – EQM sur les sources reconstituées. Comparaison sur 1000 réalisations de Monte-Carlo (3 sources, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons) du résultat avec le contraste $\mathcal{C}_z\{\cdot\}$ et le contraste classique sans référence $|\kappa_4\{y(n), y(n)\}|$. Les réalisations ont été classées par ordre croissant de l'EQM pour le contraste sans référence.

de 20 secondes pour les 1000 itérations nécessaires à une bonne convergence d'un contraste classique. En comparaison, l'étape d'optimisation du contraste avec référence prend, dans les mêmes conditions, au plus 20 *centièmes* de secondes. Le gain en temps est donc considérable et constitue l'atout majeur de la méthode. De plus, la méthode ne se trouve que faiblement handicapée par un nombre accru d'échantillons : l'estimation des matrices notées \mathbf{C} et \mathbf{R} dans le paragraphe 5.3 ne se réalise en effet qu'une seule fois et l'optimisation s'effectue alors en ne faisant appel qu'à ces seules matrices dont la taille est réduite. Ainsi, pour 10000 et 50000 échantillons, les temps respectifs d'optimisation passent à 44 et 221 secondes pour un contraste classique contre 23 et 93 centièmes de secondes pour le contraste avec référence. Ce point explique par ailleurs le nombre très élevé d'échantillons qui a pu être utilisé dans la suite de nos simulations.

Non convergence de l'optimisation du contraste sans référence Compte tenu des résultats précédents, il est apparu intéressant de regarder l'effet de l'utilisation comme référence du résultat fourni par un contraste classique mais dont l'optimisation itérative a été interrompue avant convergence. Les résultats sont donnés à la figure 5.4. Il apparaît une grande majorité des cas où le contraste par référence offre une performance très honorable malgré une référence manifestement éloignée de l'une des sources. Cependant, pour environ 5% des cas, le résultat du contraste classique n'est pas convenable (en raison d'une optimisation inachevée) et la méthode par référence ne parvient pas non plus à une solution acceptable. L'étude du paragraphe 5.4.2 fournira des éléments d'explication. Le gain en temps de calcul demeure toutefois élevé dans la mesure où l'on s'épargne la nécessité d'attendre la convergence de l'algorithme d'optimisation itératif.

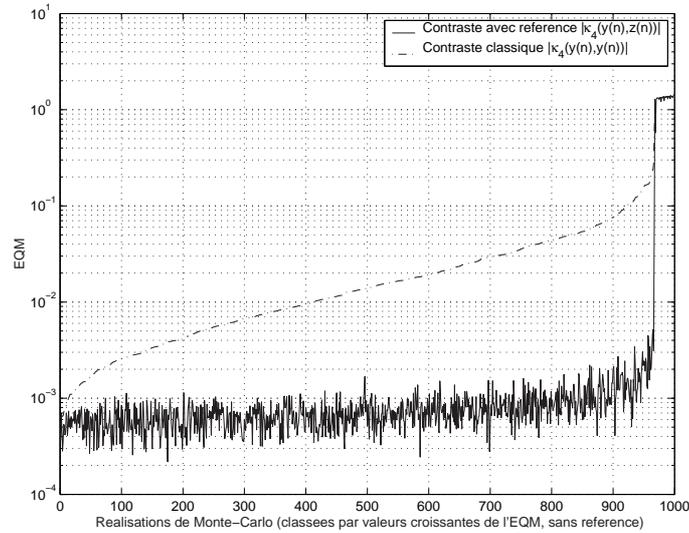


FIG. 5.4 – EQM sur les sources reconstituées (3 sources de type PAM4, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons). L’optimisation du contraste sans référence n’a pas été menée jusqu’à la convergence et le résultat obtenu a servi de référence. Les réalisations ont été classées par ordre croissant de l’EQM pour le contraste sans référence.

5.4.1-c) Cas de sources non i.i.d.

Les simulations précédentes ont été brièvement reprises pour des sources non i.i.d. afin d’effectuer une vérification de principe de la validité de la méthode dans ce cas. Les résultats sont donnés dans le tableau 5.1 pour des sources CPM d’indices de modulation 0.25, 0.4 et 0.6.

	Contraste sans réf, 10 itérations	Avec réf. colonne préc.	Contraste sans réf, 100 itérations	Avec réf. colonne préc.
Sources CPM, 5000 échantillons	0.1993	0.0052	0.0068	0.0010
Sources CPM, 2000 échantillons	0.1917	0.0123	0.0075	0.0022

TAB. 5.1 – Valeur moyenne du critère τ_1 sur 1000 réalisations. Le mélange était tiré aléatoirement, à coefficients complexes avec 3 sources et 4 capteurs.

5.4.2 Tirage aléatoire du système de référence

Au cours des simulations précédentes, l’obtention d’un signal de référence a constitué la difficulté majeure dans l’utilisation des techniques de ce chapitre. Il est cependant apparu dans les propositions 22 et 23 que les conditions sur le système de référence étaient relativement faibles. En particulier, n’importe quel contraste avec référence obtenu à l’aide d’un système de référence tiré aléatoirement devrait d’après la proposition 23 mener à la séparation. Cette affirmation doit cependant être modulée par le fait que nous ne disposons que de statistiques estimées. Nous examinons maintenant ce point au travers de simulations.

Parallèlement, il nous est venu à l'idée d'itérer le processus de séparation avec référence et d'obtenir ainsi un algorithme de type «point fixe». Plus précisément, la méthode mise en œuvre et testée a été la suivante :

- choisir par tirage aléatoire un vecteur $\underline{\mathbf{w}}^0$ correspondant à un système séparant de référence.
- pour n variant de 1 à n_{\max} :
 - fixer la référence à $z(n) = \underline{\mathbf{w}}^{n-1} \underline{\mathbf{x}}(n)$.
 - optimiser $\mathcal{C}_z\{\cdot\}$, ce qui donne le filtre séparateur $\underline{\mathbf{w}}^n$.
- le filtre séparateur retenu est celui correspondant à $\underline{\mathbf{w}}^{n_{\max}}$ et la source reconstituée s'écrit $\underline{\mathbf{w}}^{n_{\max}} \underline{\mathbf{x}}(n)$.

Les résultats fournis par cette méthode sur 100 réalisations sont livrés dans les tableaux 5.2, 5.3 et 5.4 pour un nombre d'itération maximal $n_{\max} = 5$. Le nombre de sources a été fixé à 3, le nombre de capteurs à 4. Les filtres de mélange étaient de longueur 3 et ont été tirés aléatoirement (loi gaussienne centrée normalisée sur les coefficients).

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.2687	18	0	0
2	0.0435	88	73	16
3	0.0087	96	96	81
4	0.0009	100	99	91
5	0.0006	100	100	96

TAB. 5.2 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 5000 échantillons.

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.1662	80	5	0
2	0.0069	99	90	69
3	0.0003	100	100	99
4	0.0003	100	100	100
5	0.0003	100	100	100

TAB. 5.3 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 10000 échantillons.

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.0191	89	60	4
2	3.5620e-5	100	100	100
3	2.8186e-5	100	100	100
4	2.8186e-5	100	100	100
5	2.8186e-5	100	100	100

TAB. 5.4 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 100000 échantillons.

Nous constatons dans le tableau 5.4 que, pour un nombre d'échantillons grand, le

résultat de séparation est en général bon dès la première maximisation du contraste avec référence. Un examen plus attentif indique toutefois que les cas de séparation insatisfaisante à la première étape sont liés à l'existence de deux valeurs propres relativement proches dans la matrice notée $\tilde{\mathbf{C}}$ au paragraphe 5.3, donnant ainsi une possibilité de se prémunir contre de tels cas. Par ailleurs, l'augmentation du nombre d'échantillons permet de fortement limiter le nombre de cas non séparant dès la première étape, comme l'indique l'évolution des valeurs dans les tableaux 5.2, 5.3 et 5.4.

En conclusion, en présence d'un nombre limité d'échantillons, les erreurs d'estimation des matrices \mathbf{C} et \mathbf{R} peuvent conduire à un écart de performances par rapport au comportement théorique prédit par la proposition 23. Néanmoins, il est tout à fait positif de constater que, dans les cas où la séparation n'est pas effective dès la première maximisation du contraste avec référence, le fait de recommencer cette maximisation permet de converger vers l'une des sources. Ainsi, la méthode proposée semble pouvoir s'appliquer dans un cadre aveugle.

5.4.3 Séparation MIMO par déflation

La séparation MIMO a été testée avec 3 sources, 4 capteurs, des filtres de mélange tirés aléatoirement de longueur 3 et des sources i.i.d. PAM4. Les sources ont été extraites en utilisant la méthode de maximisations successives d'un contraste avec référence décrite au paragraphe précédent. Le passage à l'extraction MIMO a été mené par des étapes de déflation successives décrites au paragraphe 3.2.1 et en prenant en compte la remarque du paragraphe 5.3.4.

Les résultats provenant de 100 réalisations de Monte-Carlo sont regroupés dans les tableaux 5.5, 5.6 et 5.7. Nous pouvons constater que la méthode offre des résultats satisfaisants.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	0.0006	100	100	100
2	0.0569	96	95	87
3	0.0251	96	91	30

TAB. 5.5 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 5000 échantillons et, pour chaque extraction, 15 itérations de maximisation d'un contraste avec référence.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	0.0003	100	100	100
2	0.0145	98	98	97
3	0.0251	98	98	93

TAB. 5.6 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 10000 échantillons et, pour chaque extraction, 10 itérations de maximisation d'un contraste avec référence.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	3.0512e-5	100	100	100
2	0.0095	99	99	99
3	0.0053	99	99	99

TAB. 5.7 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 100000 échantillons et, pour chaque extraction, 5 itérations de maximisation d'un contraste avec référence.

5.5 Conclusion

Partant de la notion de signal de référence, nous avons introduit une classe de contrastes MISO utilisables *a priori* dans un cas semi-aveugle. La validité de ces critères de contraste a été établie aussi bien dans le cas de sources i.i.d. que dans le cas non i.i.d. Compte tenu des conditions de validité de ces contrastes, les connaissances nécessaires à l'obtention d'un signal de référence se sont révélées cependant relativement faibles. Les simulations ont ainsi mis en évidence la possibilité d'utiliser ces contrastes dans un cadre aveugle. Enfin, l'avantage majeur des critères introduits réside dans leur dépendance quadratique vis à vis des paramètres. Il en résulte la possibilité de les optimiser par une méthode rapide et efficace. Ainsi, ces contrastes MISO offrent une méthode performante d'extraction d'une source. Combinés avec des étapes de déflations, nos simulations laissent augurer d'excellentes perspectives pour la séparation de mélanges MIMO.