



HAL
open science

Conditions d'acceptabilité dans les problèmes de séparation de sources

Vincent Vigneron, Sébastien Lagrange, Christian Jutten

► **To cite this version:**

Vincent Vigneron, Sébastien Lagrange, Christian Jutten. Conditions d'acceptabilité dans les problèmes de séparation de sources. XXXVèmes journées de Statistiques, Jun 2003, Paris, France. pp.883-886. hal-00233001

HAL Id: hal-00233001

<https://hal.science/hal-00233001>

Submitted on 3 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conditions d'*acceptabilité* pour les problèmes de séparation de sources

V. Vigneron^{1,2} S. Lagrange², C. Jutten²

¹MATISSE-SAMOS UMR 8595
90, rue de Tolbiac
75634 Paris cedex 13

²LIS
avenue Félix Viallet
38031 Grenoble cedex

Résumé

L'algorithme proposée par Herault-Jutten (HJ) utilise un réseau de neurones linéaires totalement interconnecté. Il est capable de mettre en œuvre ce qui est baptisé la séparation de sources aveugle (BSS) sans connaître les caractéristiques des entrées ni celles des canaux de transmissions. L'algorithme d'apprentissage s'appuie sur une loi de Hebb généralisée telle que chaque sortie soit proportionnelle à seulement une source en annulant l'influence des autres sources. Dans cet article, nous montrons comment les conditions théoriques de stabilité peuvent permettre l'estimer les paramètres du mélange et restaurer les sources inconnues au moyen du calcul ensembliste.

1 Introduction

Considérons le réseau de neurones de HJ présenté Figure 1 [2]. Les signaux observés $x_i(t)$ et $x_2(t)$ sont constitués à partir des combinaisons linéaires de 2 processus aléatoires d'entrées (sources primaires) $s_1(t)$ et $s_2(t)$, *i.e.* :

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t), \quad x_2(t) = a_{21}s_1(t) + a_{22}s_2(t), \quad (1)$$

qui peut être écrite sous une forme plus compacte $X(t) = AS(t)$, où $X(t) = (x_1(t), x_2(t))$, $S = (s_1(t), s_2(t))$ et $A \in \mathbb{R}^{2 \times 2}$ est une matrice constante régulière. Herault et Jutten [2] imposent que les sources soient centrées, *i.e.* $\mathbb{E}[s_1] = \mathbb{E}[s_2] = 0$, stationnaires et indépendants. Les sorties du système sont données (en omettant l'index t) par :

$$y_1 = x_1 - w_{12}y_2, \quad y_2 = x_2 - w_{21}y_1. \quad (2)$$

Les poids du réseau w_{12} et w_{21} sont adaptés de façon à ce que : (i) chaque signal de sortie soit proportionnel à seulement une source primaire en annulant l'influence des autres sources, (ii) les signaux de sortie $y_1(t)$ et $y_2(t)$ sont statistiquement indépendants (au terme du processus d'adaptation). En éliminant x_1 et x_2 des équation (1), on trouve à partir de l'Eq.(2) :

$$y_1 = \frac{1}{1 - w_{12}w_{21}}((a_{11} - w_{12}a_{21})s_1 + (a_{12} - w_{12}a_{22})s_2), \quad (3)$$

$$y_2 = \frac{1}{1 - w_{12}w_{21}}((a_{21} - w_{21}a_{11})s_1 + (a_{22} - w_{21}a_{12})s_2). \quad (4)$$

L'indépendance des signaux signifie que ces signaux doivent être statistiquement décorrelés *i.e.* $\mathbb{E}[s_1s_2] = 0$ et par suite $\mathbb{E}[y_1y_2] = 0$.

Cette dernière condition n'est pas suffisante pour définir la règle d'apprentissage. L'indépendance statistique des signaux de sortie y_1 et y_2 implique aussi $\mathbb{E}[f(y_1)g(y_2)] = 0$, où f et g sont des fonctions non-linéaires, typiquement $f(x) = x^3$, $g(x) = x$. Herault et Jutten ont proposé la règle d'adaptation suivante pour les coefficients w_{ij} [2]

$$\frac{dw_{12}}{dt} = \mu y_1^3 y_2, \quad \frac{dw_{21}}{dt} = \mu y_2^3 y_1, \quad (5)$$

où $\mu > 0$ est une constante. On peut montrer que $\mathbb{E}[\frac{dw_{ij}}{dt}] \rightarrow 0 \Rightarrow \mathbb{E}[f(y_i)g(y_j)] \rightarrow 0$ ($i \neq j$). L'algorithme HJ fonctionne car il fait appel à des fonctions non-linéaires $f(\cdot)$ et $g(\cdot)$ introduisant ainsi la contribution de moments d'ordre supérieur dans la fonctionnelle à annuler.

Nous donnons dans la suite les condition de stabilité qui restreignent la mise en œuvre de l'algorithme de séparation. Un algorithme de séparation basé sur le calcul par intervalle est proposé. et les zones d'équilibre sont examinées.

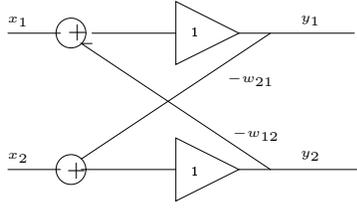


FIG. 1 – Réseau de Herault-Jutten à 2 canaux.

2 Conditions de stabilité

Sorouchyari a proposé dans [6] une analyse précise de la stabilité du réseau au regard de la théorie de la stabilité de Lyapunov. Les points d'équilibre du réseau HJ sont solutions du système :

$$\begin{pmatrix} 1 & w_{12} \\ w_{21} & 1 \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (6)$$

La première solution s'écrit $w_{12} = \frac{a_{12}}{a_{22}}$ et $w_{21} = \frac{a_{21}}{a_{11}}$, la seconde $w_{12} = \frac{a_{11}}{a_{21}}$ et $w_{21} = \frac{a_{22}}{a_{12}}$ [6]. Notons que le circuit Fig. 1 se comporte comme un réseau récursif pour lequel la *stabilité* est garantie seulement si $w_{12}w_{21} < 1$. Comon *et al.* [1] et Sorouchyari [6] ont étudié les propriétés de convergence de l'algorithme dans le cas d'un réseau 2 entrées/2 sorties. Ils ont démontré qu'il existe quatre paires de points d'équilibre dans le sens où si le point (a, b) est un point d'équilibre, alors le point $(\frac{1}{a}, \frac{1}{b})$ est une autre solution (voir [6]). L'étude de stabilité réalisée par Sorouchyari consiste à ajouter une petite perturbation au point stationnaire et à analyser le comportement du système autour de ce point [4]. Les conditions de stabilité proposées sont alors :

$$w_{12}w_{21} < 1 \quad \text{and} \quad \mathbb{E}[y_1^4]\mathbb{E}[y_2^4] > 9(\mathbb{E}[y_1^2]y_2^2)^2 \quad (7)$$

Sorouchyari [6] montre qu'un seul de ces points stationnaires est une solution séparatrice stable.

Le critère de stabilité de l'Eq. (7), basé sur la structure jacobienne de la matrice de mélange et qui décide de l'*admissibilité* des paramètres w_{12} et w_{21} , a plusieurs inconvénients : (i) le choix des valeurs initiales est largement aléatoire, (ii) il n'y a aucune garantie de convergence vers l'optimum global, (iii) plus que la valeur optimale des paramètres, on aimerait caractériser l'ensemble des valeurs *acceptables* de ces paramètres, (iv) l'incertitude sur les estimateurs n'est évaluée que sur la base d'hypothèses asymptotiques, donc aucune évaluation valable n'est fournie sur la valeur des estimées.

3 Maximum de *consistence*

C'est pourquoi on s'intéresse à l'ensemble des modèles acceptables. La première étape consiste à lister toutes les propriétés que le modèle doit satisfaire pour être acceptable. L'*Acceptabilité* d'un modèle est définie par un ensemble de contraintes que doivent satisfaire les paramètres. Lorsque ces conditions d'acceptabilité sont satisfaites, on souhaite caractériser – approximativement mais de manière garantie – l'ensemble de *toutes* les valeurs \mathbb{W} (parfois appelées *ensemble de vraisemblance*) de w_{12} et w_{21} qui sont *consistantes*/admissibles avec les données, *i.e.* l'ensemble de toutes les valeurs qui sont consistantes avec l'ensemble des valeurs possibles et qui satisfait *toutes* les conditions d'acceptabilité. Ceci est réalisé par l'algorithme fourni à la section 4.

Arithmétique des intervalles Un intervalle $[x]$ est un sous-ensemble fermé de \mathbb{R} . L'ensemble de tous les intervalles \mathbb{R} est noté \mathbb{IR} . Soit $[x] = \{x \in \mathbb{R} \mid \bar{x} \leq x \leq \underline{x}, \bar{x} \in \mathbb{R}, \underline{x} \in \mathbb{R}\}$, où \bar{x} et \underline{x} sont les bornes supérieure et inférieure de $[x]$. Un nombre réel est un intervalle tel que $\bar{x} = \underline{x} = x$. Les opérations de base sur les nombre réels et les vecteurs de réels telles que $+$, $-$, $*$, $/$, \sin , \exp , \dots s'étendent aux intervalles de façon naturelle. Par exemple

$$[x] \times [y] = \{x \times y \mid x \in [x], y \in [y]\} = [\min(\underline{x} \cdot \underline{y}, \underline{x} \cdot \bar{y}, \bar{x} \cdot \underline{y}, \bar{x} \cdot \bar{y}), \max(\underline{x} \cdot \underline{y}, \underline{x} \cdot \bar{y}, \bar{x} \cdot \underline{y}, \bar{x} \cdot \bar{y})].$$

Une *boite* $[\mathbf{x}]$ de \mathbb{R}^p est le produit cartésien de p intervalles, *i.e.* $[\mathbf{x}] = [x_1] \times \dots \times [x_p]$.

$[f]$ est une *fonction d'inclusion* de la fonction vectorielle f si, pour chaque boite $[\mathbf{x}]$, $[f]([\mathbf{x}])$ est aussi une boite telle que

$$\forall [\mathbf{x}], f([\mathbf{x}]) \triangleq \{f(\mathbf{x}) \mid \mathbf{x} \in [\mathbf{x}]\} \subset [f]([\mathbf{x}]). \quad (8)$$

$f(\mathbf{x})$, i.e. la boîte englobante de $f([\mathbf{x}])$, voir Fig. 2.a. Le calcul par interval permet d'utiliser les fonctions d'inclusion pour une large classe de fonctions non-linéaires.

L'union intervalle $[\mathbf{x}] \sqcup [\mathbf{y}]$ est la plus petite boîte qui contient l'union de deux boîtes $[\mathbf{x}] \cup [\mathbf{y}]$.

Dans le modèle neuronal HJ, nous souhaitons caractériser l'ensemble des paramètres $\mathbb{W} \subset \mathbb{I}\mathbb{R}^2$ tel que $(\mathbb{W}, \mathbb{X}) = f^{-1}(\mathbb{Y})$ à partir de la connaissance de l'ensemble des paires $\mathbb{Y} = \{(y_{1_i}, y_{2_i})\}_{i=1}^T \in \mathbb{R}^2$ et de la fonction (supposée inversible) f . Soit \mathbb{X} l'ensemble des processus recherchés appelés sources. Alors, le problème de séparation de source peut se formuler comme un problème d'inversion ensembliste, à résoudre de manière globale. L'ensemble des paramètres (w_{12}, w_{21}) à trouver sera calculé en construisant des sous-ensembles de boîtes sans chevauchement de largeur non-nulle. L'exploration est limitée à une boîte de départ, notée $[\mathbf{w}]^{(0)}$, qui est partagée par l'algorithme en boîtes plus petites autant de fois que nécessaires jusqu'à ce que, soit une conclusion soit atteinte, soit la largeur de la boîte soit plus petite qu'un paramètre de tolérance ϵ . L'analyse par intervalle fournit ainsi deux tests de base pour décider si une boîte considérée $[f]([\mathbf{x}], [\mathbf{w}])$ appartient à \mathbb{Y} :

$$[f]([\mathbf{x}], [\mathbf{w}]) \subset [\mathbf{y}] \Rightarrow [\mathbf{w}] \in \mathbb{W} \quad \text{i.e. } [\mathbf{w}] \text{ est faisable,} \quad (9)$$

$$[f]([\mathbf{x}], [\mathbf{w}]) \cap [\mathbf{y}] = \emptyset \Rightarrow [\mathbf{w}] \cap \mathbb{W} = \emptyset \quad \text{i.e. } [\mathbf{w}] \text{ est infaisable.} \quad (10)$$

Dans tous les autres cas, la boîte $[\mathbf{w}]$ est *indéterminée* et sa largeur est plus *grande* qu'un paramètre de précision ϵ . Alors elle doit être bissectée en deux sous-pavés notés W^- qui contient toutes les boîtes démontrées *faisables*, et W^+ composé de toutes les boîtes *indéterminées*. Le test sera appliqué *recursivement* à ces pavés nouvellement générés (voir Figure 2.b). Il est facile à partir de ces sous-pavés d'encadrer l'ensemble d'intérêt $W \subset \mathbb{W}$ contenu dans $[\mathbf{w}]^{(0)}$ en remarquant que :

$$W^- \subset [\mathbf{w}]^{(0)} \cap W \subset W_{\text{out}} := W^- \cup W^+. \quad (11)$$

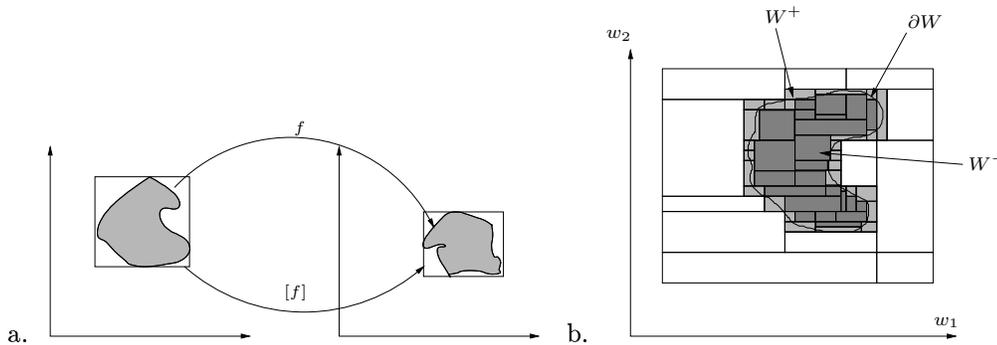


FIG. 2 – (a) Fonction d'inclusion minimale $[f]$ d'une fonction f . (b) Pavage faisable.

W_{out} est l'union finie des boîtes garantie contenir la portion \mathbb{W} d'intérêt. Une pile de type FIFO est utilisée pour conserver en mémoire l'ensemble des pavés examinés. L'initialisation est réalisée en forçant la pile $\text{stack} = \emptyset$, et $W^+ = \emptyset$. L'algorithme requiert que la recherche débute dans une boîte "suffisamment" large $[\mathbf{w}]^{(0)} =]-\infty; \infty[$ dans laquelle W est garantie appartenir.

Au terme du processus récursif, la consistance de W étant donné les données est maximisée et aucun pavé indéterminé n'aura de largeur supérieure à ϵ [3]. On peut admettre que W^+ et W_{out} convergeront vers W (respectivement par l'intérieur et par l'extérieur) quand $\epsilon \rightarrow 0$ [5]. A la fin, on peut afficher dans l'espace des paramètres l'ensemble W^- de tous les pavés retenus qui ont été prouvés faisables (voir Figure 2.b).

Conditions d'acceptabilité des solutions d'un problème de séparation de source Un premier exemple de conditions d'acceptabilité est donné par les résidus entre les données $(y_1, y_2)^T$ et les sorties correspondantes du modèle : s'ils sont bornés par des valeurs connues, ils expriment l'intervalle de confiance attaché aux mesures. Un tel critère est satisfait indirectement en minimisant les corrélations nonlinéaires (voir section 2) : $f([y_{1_i}])g([y_{2_i}]) \in]0; \infty[, \forall i \in \{1, \dots, T\}$. On peut également considérer d'autres conditions d'acceptabilité, par

exemple celles données par les contraintes de stabilité optimales. Une condition nécessaire et suffisante pour que le modèle HJ soit asymptotiquement stable est que les équations (7) soient satisfaites, *i.e.* ssi

$$\left(9\overline{[y_1]^2[y_2]^2} - \overline{[y_1]^4[y_2]^4}\right) \in]-\infty; 0[\text{ et } (1 - [w_{12}][w_{21}]) \in]0; \infty[, \quad (12)$$

où

$$\overline{[y_1]^2[y_2]^2} = \frac{1}{T} \sum_i [y_{1i}]^2 [y_{2i}]^2, \quad \overline{[y_2]^4} = \frac{1}{T} \sum_i [y_{1i}]^4 \text{ et } \overline{[y_2]^4} = \frac{1}{T} \sum_i [y_{2i}]^4. \quad (13)$$

4 Simulations

En guise d'illustration, considérons le modèle à temps discret dans lequel les données ont été générées par simulation du modèle suivant $k = 1, \dots, 500$:

$$\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} = \begin{pmatrix} 1 & 0,6 \\ 0,3 & 1 \end{pmatrix} \begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix}, \quad (14)$$

où $s_1(k) = \sin(7,3kT_e)$, $s_2(k) = \sin(4kT_e)$, $T_e = 0,2$ et en ajoutant à $(x_1(k), x_2(k))$ un bruit blanc aléatoire distribué uniformément dans l'intervalle $[-10^{-2}; 10^{-2}]$.

ALGORITHME DE CONSISTANCE	
entrées :	$([x_1(k)], [x_2(k)]), i = 1, \dots, T;$ (supposé centré)
init :	$[y_{1i}] := [y_{2i}] :=]-\infty; \infty[, i = 1, \dots, T;$ $[w_{12}] := [w_{21}] :=]-5, 5[;$ $[\delta^+] := [0; \infty]; [\delta_-] :=]-\infty; 0]; [\kappa] := [0; \infty];$
répéter	<div style="padding-left: 20px;"> Pour $i := 1$ à T <div style="padding-left: 20px;"> <ol style="list-style-type: none"> 1 $[w_{12}] := [w_{12}] \cap \left(\frac{1}{y_{2i}}([x_{1i}] - [y_{1i}])\right);$ 2 $[w_{21}] := [w_{21}] \cap \left(\frac{1}{y_{1i}}([x_{2i}] - [y_{2i}])\right);$ 3 $[y_{1i}] := [y_{1i}] \cap ([x_{1i}] - [w_{12}] \times [y_{2i}]);$ 4 $[y_{2i}] := [y_{2i}] \cap ([x_{2i}] - [w_{21}] \times [y_{1i}]);$ 5 $[\kappa] := [\kappa] \cap ([y_{1i}] \times [y_{1i}] \times [y_{1i}] \times [y_{2i}]);$ 6 $[\delta^+] := [\delta^+] \cap (1 - [w_{12}] \times [w_{21}]);$ FinPour </div> </div>
tant que la contraction est significative	<div style="padding-left: 20px;"> <ol style="list-style-type: none"> 7 $[\delta_-] := [\delta_-] \cap (\overline{[y_1]^4[y_2]^4} - 9\overline{[y_1]^2[y_2]^2})$ </div>
sortie :	$[w_{12}], [w_{21}];$

Ainsi, l'intervalle $([x_1(k)], [x_2(k)])$ est assuré contenir les vraies données inconnues. Les domaines de départ sont pris égaux à $[w_{12}](0) := [w_{21}](0) :=]-5, 5[$ et pour les $[y_i](k)$ égaux à $[-10, 10]^2$. Le problème à résoudre est le suivant : étant donné un ensemble de contraintes d'acceptabilité sur le réseau HJ, il s'agit de calculer le domaine d'appartenance des estimées des sources y_{1i}, y_{2i} et w_{12}, w_{21} .

Results Après calcul, les intervalles contractés des Figures 3.b and c incluent les vraies valeurs des paramètres. Un grand nombre de bisections a été réalisé.

Conclusion For the first time, this paper studies the application of interval analysis to parameter estimation in BSS problems. An algorithm to enclose efficiently all consistent values for the unknown parameter vector inside a box has been presented in a error bounded context. An illustrative example has shown the efficiency of the approach.

Références

- [1] P. Comon, C. Jutten and J. Herault, "blind separation of sources, Part 2 : Problems statement", *Signal Processing*, 24, pp. 11-20, 1991.

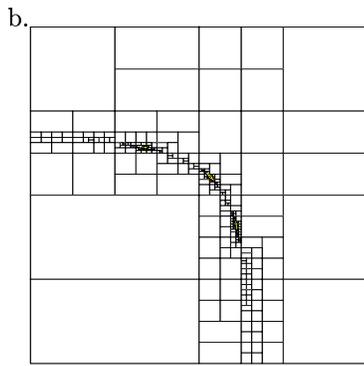


FIG. 3 – (a) True values of the parameters w_{12}, w_{21} . (b-c) Contracted domains for $([w_{12}], [w_{21}])$ without/including stability conditions.

- [2] J. Herault and C. Jutten, “Space ot time adaptative sinal processing by neural network models”, Neural networks for computing, Snowbird, 1986.
- [3] L. Jaulin and E. Walter, “Guaranteed nonlinear parameter estimation from bound-error data via interval analysis”, Mathematics and Computers in Simuation, **5**, pp. 123-137,1993.
- [4] N. Minorsky, “Nonlinear oscillations”, Van Nostrand, Reinhold, New York, 1962.
- [5] J.P. Norton, “Special issue on bounded-error estimation : Issue 1”. In *International Symposium Control and Signal Processing*, **8(1)**, pp. 1-118, 1994.
- [6] E. Sorouchyari, “Blind source separation, Part 3 : Stabiity analysis”, *Signal Processing*, **24**, pp. 21-29, 1991.