



HAL
open science

Acceptability conditions for blind source separation problems

Vincent Vigneron, Sébastien Lagrange, Christian Jutten

► **To cite this version:**

Vincent Vigneron, Sébastien Lagrange, Christian Jutten. Acceptability conditions for blind source separation problems. 11th European Symposium on Artificial Neural Networks (ESANN 2003), Apr 2003, Bruges, Belgium. pp.301-306. hal-00233005

HAL Id: hal-00233005

<https://hal.science/hal-00233005>

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.

Acceptability conditions for BSS problems

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

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

²LIS
avenue Félix Viallet
38040 Grenoble cedex

Abstract. The Herault-Jutten (HJ) algorithm is a neuromimetic structure capable to perform *blind source separation* (BSS) of a linear mixture from an array of sensors without knowing the transmission characteristics of the channels, nor the inputs. The learning algorithm developed by Herault and Jutten is based on the generalized Hebb's rule in such a way that each output signal will be proportional to only one source by cancelling the influence of the other source. In this article, we show how theoretic stability conditions can be used for parameter estimation to restore the primary sources via interval computations.

1 Introduction

Consider the HJ network presented in Figure 1 [2]. Suppose the observation signals $x_1(t)$ and $x_2(t)$ are linear combinations of the input signals (primary sources) $s_1(t)$ and $s_2(t)$, *i.e.*:

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

which can be written in a more compact matrix form $X(t) = AS(t)$, where $X(t) = (x_1(t), x_2(t))$, $S = (s_1(t), s_2(t))$ and $A \in \mathbb{R}^{2 \times 2}$ is a regular constant matrix. The sources are assumed to be zero-mean, *i.e.* $\mathbb{E}[s_1] = \mathbb{E}[s_2] = 0$, stationary and independent. The outputs of the system are given (omitting time index t) by:

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

The network contains adaptative weights w_{12} and w_{21} which must be adjusted in such a way that: (*i*) each output signal will be proportional to only one primary sources by canceling the influence of the other sources (*ii*) the output signals $y_1(t)$ and $y_2(t)$ are statistically independent (after the adaptation process). Eliminating x_1 and x_2 from equations (1), we obtain from Eq.(2):

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

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

The independence of the signals mathematically means that these signals must be at least decorrelated, *i.e.* $\mathbb{E}[s_1 s_2] = 0$ and therefore $\mathbb{E}[y_1 y_2] = 0$.

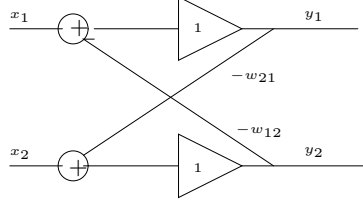


Figure 1: The two-channels Herault-Jutten network.

The latter condition is not sufficient to find the learning rule. Statistical independence of the output signals y_1 and y_2 implies $\mathbb{E}[f(y_1)g(y_2)] = 0$, where f and g are nonlinear functions, typical examples are $f(x) = x^3, g(x) = x$. On the basis of the above requirements, Herault and Jutten have proposed to consider the following adaptation rule for the coefficients w_{ij} [2]

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

where $\mu > 0$ is a constant. It is expected that $\mathbb{E}[\frac{dw_{ij}}{dt}] \rightarrow 0 \Rightarrow \mathbb{E}[f(y_i)g(y_j)] \rightarrow 0$ ($i \neq j$). The functions $f(\cdot)$ and $g(\cdot)$ introduce higher-order moments.

The study is carried out first by examining stability conditions. An interval algorithm for BSS is then derived and equilibrium point zones are examined.

2 Stability conditions

A rigorous stability analysis of the network was given by Sorouchyari in terms of Lyapunov stability theory [6]. Equilibrium points of the H-J network are solutions of the system:

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

The first solution reads $w_{12} = \frac{a_{12}}{a_{22}}$ and $w_{22} = \frac{a_{21}}{a_{11}}$, the second reads $w_{12} = \frac{a_{11}}{a_{21}}$ and $w_{22} = \frac{a_{22}}{a_{12}}$ [6]. It should be noted that the circuit Fig. 1 behaves as a feedback network which is *stable* under the condition $w_{12} w_{21} < 1$. Comon *et al.* [1] and Sorouchyari [6] investigate the convergence properties of the algorithm and perform a stability analysis for a 2 inputs/2 outputs network. They demonstrate that there are 4 paired equilibrium points in the sense that if the point (a, b) is an equilibrium point, then the point $(\frac{1}{a}, \frac{1}{b})$ is also a solution (see [6]).

Stability analysis performed by Sorouchyari consists to introduce a small perturbation on the stationary points and to consider the further behaviour of the system around these points [4]. The stability conditions proposed are therefore:

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

Sorouchyari [6] shows that only one of the stationary points will be a stable separating solution.

The stability criteria in Eq. (7) which are based on the jacobian structure of the mixing matrix and which decide if the parameters w_{12} and w_{21} are *admissible*, have several drawbacks: (i) the choice of the initial value relies on guesswork, (ii) no guarantee of convergence to the global optimum can be provided, (iii) we are not interested in *the* optimal value, we rather like to characterize the set of *all* the acceptable values, (iv) uncertainty on the estimate is evaluated on the base of asymptotic assumptions, so no reliable evaluation is provided of the precision with which the estimated value is obtained.

3 Maximum consistency

This is why we shall look for the set of all models that are acceptable. The first step is then to list all the properties that the model should have to be acceptable. *Acceptability* will be defined here by a set of inequalities to be satisfied by the parameters. Once these conditions of acceptability have been defined, we wish to characterize – approximately but in a guaranteed way – the set of *all* values \mathbb{W} (sometimes called *likelihood set*) of w_{12} and w_{21} that are *consistent*/admissible with the data, *i.e.* the set of all values that are consistent with the prior feasible set and that satisfy all conditions of acceptability. This will be performed by the algorithm given in section 4.

Interval arithmetical An interval $[x]$ is a closed and connected subset of \mathbb{R} . The set of intervals of \mathbb{R} will be denoted \mathbb{IR} . Let define $[x] = \{x \in \mathbb{R} \mid \bar{x} \leq x \leq \underline{x}, \bar{x} \in \mathbb{R}, \underline{x} \in \mathbb{R}\}$, where \bar{x} and \underline{x} are the upper and under bounds of $[x]$. A real number is an interval such that $\bar{x} = \underline{x} = x$. Basic operations on real numbers and vectors such as $+$, $-$, $*$, $/$, \sin , \exp , \dots extend to intervals in a natural way. For instance $[x] \times [y] = \{x \times y \mid x \in [x], y \in [y]\} = [\min(\underline{x} \cdot \underline{y}, \underline{x} \cdot \bar{y}, \bar{x} \cdot \underline{y}, \bar{x} \cdot \bar{y}), \max(\underline{x} \cdot \bar{y}, \underline{x} \cdot \bar{y}, \bar{x} \cdot \underline{y}, \bar{x} \cdot \bar{y})]$. A *box* $[\mathbf{x}]$ of \mathbb{R}^p is a cartesian product of p intervals, *i.e.* $[\mathbf{x}] = [x_1] \times \dots \times [x_p]$.

$[f]$ is an *inclusion function* of the vector function f if, for any box $[\mathbf{x}]$, $[f]([\mathbf{x}])$ is also a box such that

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

$[f]([\mathbf{x}])$ is thus the smallest box that contains $f(\mathbf{x})$, *i.e.* the envelopping box of $f([\mathbf{x}])$, see Fig. 2.a. Interval computation makes it possible to obtain inclusion functions of a large class of nonlinear functions.

The interval union $[\mathbf{x}] \sqcup [\mathbf{y}]$ is the smallest box which contains the union of two

boxes $[\mathbf{x}] \cup [\mathbf{y}]$.

In the HJ neural-like system, we aim at characterizing the parameter set $\mathbb{W} \subset \mathbb{R}^2$ such that $(\mathbb{W}, \mathbb{X}) = f^{-1}(\mathbb{Y})$ from the knowledge of the set of pairs datum $\mathbb{Y} = \{(y_{1i}, y_{2i})\}_{i=1}^T \in \mathbb{R}^2$ and the vector function (supposedly invertible) f . Let \mathbb{X} be the set of unknown primary sources. Then the BSS problem is formulated here as a problem of set inversion, which must be solved globally. The analysis of the parameter space of interest (w_{12}, w_{21}) will be performed by building sets of non-overlapping boxes with nonzero width. Exploration is limited to an initial box of interest, say $[\mathbf{w}]^{(0)}$, which is split by the algorithm into smaller boxes whenever needed until either a conclusion can be reached or the width of the box considered becomes smaller than some tolerance parameter ϵ . Interval analysis provide us two basic tests for deciding whether the given box $[f]([\mathbf{x}], [\mathbf{w}])$ is included in \mathbb{Y} :

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

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

In all other cases, $[\mathbf{w}]$ is *indeterminate* and has a width *greater* than the precision parameter ϵ . Then it should be bisected into two subpavings namely W^- containing all boxes that were proved *feasible*, and W^+ consisting of all *indeterminate* boxes and the test should be *recursively* applied to these newly generated boxes (see Figure 2.b). From these subpavings, it is easy to bracket the portion of $W \subset \mathbb{W}$ contained in $[\mathbf{w}]^{(0)}$ as: $W^- \subset [\mathbf{w}]^{(0)} \cap W \subset W_{\text{out}} := W^- \cup W^+$.

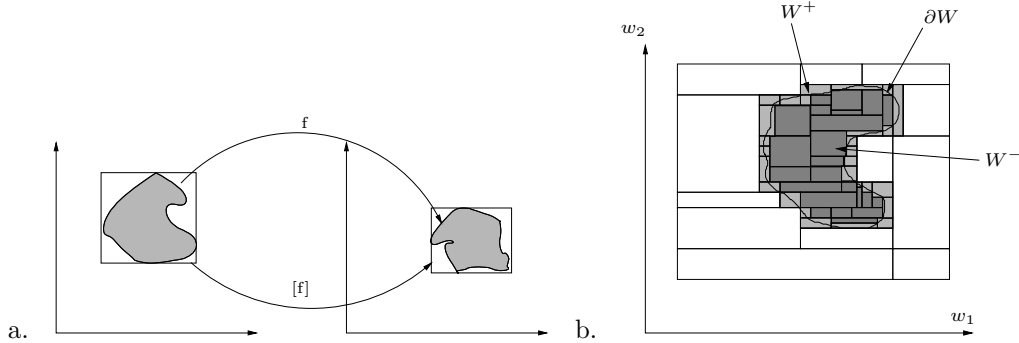


Figure 2: (a) Minimal inclusion function $[f]$ of a function f . (b) Feasibility of boxes.

W_{out} is a finite union of boxes guaranteed to contain the portion of \mathbb{W} of interest. A stack will be used to store the boxes still under considerations. Initialisation is performed by setting $\text{stack} = \emptyset, W^+ = \emptyset$. The algorithm requires a very large search box $[\mathbf{w}]^{(0)} =]-\infty; \infty[$ to which W is guaranteed to belong. Upon completion of this algorithm, the consistency of W given the data is maximized and no indeterminate box will have a width larger than ϵ [3]. Under a

few realistic technical conditions W^+ and W_{out} will tend to W (respectively from within and from without) when $\epsilon \rightarrow 0$ [5]. At the end, the set W^- of all boxes that have been proved to be feasible can be plotted in the parameter space (see Figure 2.b).

Conditions of acceptability of BSS solutions A first example of acceptability conditions is that the residuals between the data $(y_1, y_2)^T$ and corresponding model output lie between known bounds that express the confidence interval attached to individual measurements. Such criterion is performed indirectly in minimizing nonlinear correlations (see section 2): $f([y_{1_i}])g([y_{2_i}]) \in]0; \infty[, \forall i \in \{1, \dots, T\}$. Other acceptability conditions not directly related to the errors could be considered as well, for instance the coputation of guaranteed stability domains. A necessary and sufficient condition for the HJ model to be asymptotically stable is that equations (7) are satisfied, *i.e.* iff $(9\overline{[y_1]^2[y_2]^2} - \overline{[y_1]^4[y_2]^4}) \in]-\infty; 0[$ and $(1 - [w_{12}][w_{21}]) \in]0; \infty[$, where $\overline{[y_1]^2[y_2]^2} = \frac{1}{T} \sum_i [y_{1_i}]^2 [y_{2_i}]^2$, $\overline{[y_1]^4} = \frac{1}{T} \sum_i [y_{1_i}]^4$ and $\overline{[y_2]^4} = \frac{1}{T} \sum_i [y_{2_i}]^4$.

4 Test simulation

As an illustration, consider the discrete time model in which the data have been generated by simulating for $k = 1, \dots, 500$: $\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} = \begin{pmatrix} 1 & 0,6 \\ 0,3 & 1 \end{pmatrix} \begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix}$, where $s_1(k) = \sin(7, 3kT_e)$, $s_2(k) = \sin(4kT_e)$, $T_e = 0, 2$ and by adding to $(x_1(k), x_2(k))$ a random white noise with a uniform distribution in the interval $[-10^{-2}; 10^{-2}]$.

CONSISTENCY ALGORITHM	
inputs:	$([x_1(k)], [x_2(k)]), i = 1, \dots, T$; (assumed to be zero-mean)
init:	$[y_{1_i}] := [y_{2_i}] :=]-\infty; \infty[, i = 1, \dots, T$; $[w_{12}] := [w_{21}] :=]-5, 5[$; $[\delta^+] := [0; \infty]$; $[\delta_-] :=]-\infty; 0]$; $[\kappa] := [0; \infty]$;
repeat	
	for $i := 1$ to T
1	$[w_{12}] := [w_{12}] \cap (\frac{1}{y_{2_i}}([x_{1_i}] - [y_{1_i}]))$;
2	$[w_{21}] := [w_{21}] \cap (\frac{1}{y_{1_i}}([x_{2_i}] - [y_{2_i}]))$;
3	$[y_{1_i}] := [y_{1_i}] \cap ([x_{1_i}] - [w_{12}] \times [y_{2_i}])$;
4	$[y_{2_i}] := [y_{2_i}] \cap ([x_{2_i}] - [w_{21}] \times [y_{1_i}])$;
5	$[\kappa] := [\kappa] \cap ([y_{1_i}] \times [y_{1_i}] \times [y_{1_i}] \times [y_{2_i}])$;
6	$[\delta^+] := [\delta^+] \cap (1 - [w_{12}] \times [w_{21}])$;
	endfor
7	$[\delta_-] := [\delta_-] \cap (\overline{[y_1]^4[y_2]^4} - 9\overline{[y_1]^2[y_2]^2})$
	while the contraction is significant
output:	$[w_{12}], [w_{21}]$;

We are thus certain that the interval data $([x_1(k)], [x_2(k)])$ contain the un-

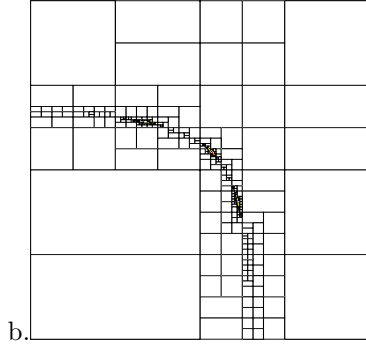


Figure 3: (a) True values of the parameters w_{12}, w_{21} . (b-c) Contracted domains for $([w_{12}], [w_{21}])$ without/including stability conditions.

known true data. The prior domains (frame boxes) are $[w_{12}](0) := [w_{21}](0) :=] - 5, 5[$ and for the $[y_i](k)$'s are all taken equal to $[-10, 10]^2$. The problem to solve is: given conditions of acceptability of the HJ system, compute accurate interval enclosure for the unknown true values for the y_{1_i}, y_{2_i} 's and w_{12}, w_{21} 's.

Results After completion, the contracted intervals in the Figures 3.b and c include the true values of the parameters. A large number of bisections have to be performed. The computing time is about 0,5 seconds for both cases.

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.

References

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