

Stabilité de l'algorithme de séparation de sources de Jutten et Héroult

Stability of the source separation algorithm of Jutten and Héroult



Jean-Claude FORT

Laboratoire de Statistiques Médicale,
URA CNRS 1323, Université Paris V,
45 rue des Saints-Pères, 75006 Paris

Maître de Conférences à l'Université Paris V. Après une formation en probabilités et statistiques à l'Université d'Orsay, il débute ses recherches en 1977 par l'utilisation des Grandes Déviations dans l'étude et la simulation des algorithmes stochastiques. A travers l'algorithme de Kohonen il découvre en 1983 les « réseaux de neurones ». Depuis, c'est dans ce domaine qu'il travaille.

RÉSUMÉ

Nous étudions l'algorithme de séparation de sources proposé par Jutten et Héroult, dans le cas de deux sources. Nous nous intéressons plus particulièrement à la stabilité des points de convergence lorsqu'ils sont solutions du problème et nous mettons en évidence une large classe de sources pour lesquelles l'algorithme ne conduit pas à la séparation des sources : il converge vers des points parasites. Ces résultats sont confirmés par des simulations.

MOTS CLÉS

Séparation de sources, algorithme stochastique, équation différentielle moyenne.

ABSTRACT

We study the algorithm proposed by Jutten and Héroult in order to separate linear mixing of two independent sources. In particular we obtain results about the stability of the limit points when they are solutions of this problem. Furthermore we point out that for a large class of sources the algorithm converges toward spurious points which are not solutions of the problem. These results are illustrated by computer simulations.

KEY WORDS

Source separation, stochastic algorithm, ordinary differential equation.

Introduction

En 1985, Ch. Jutten et J. Héroult [1] ont présenté un algorithme de séparation de sources fondé a priori sur une

Note : Au moment de la soumission de cet article, nous avons eu connaissance du travail de E. Sorouchyari « Separation of sources, stability analysis ». Cet article développe de façon élégante et plus avancée le problème de la stabilité de l'« algorithme moyen » ; notre étude montre la possibilité de divergence vers l'infini et explore plus précisément l'algorithme stochastique lui-même par simulation.

approche « neuromimétique » du problème. Les contraintes fortes engendrées par cette approche (voir § 1) les ont conduits à un algorithme de mise en œuvre très simple et dont les résultats présentés montrent l'efficacité. De surcroît, son caractère neuromimétique autorise la construction de machines très rapides dédiées à cette tâche.

Ceci explique l'intérêt que nous avons porté à cet algorithme et notre volonté d'éclaircir quelques aspects importants, peu abordés jusqu'à présent : l'existence de solutions parasites, les conditions de stabilité des solutions, parasites

ou non, la possibilité de fuite à l'infini, même dans les « bons » cas. Quelques points supplémentaires font également l'objet de remarques.

Bien que limité au cas de deux sources indépendantes, ce travail donne lieu aux conclusions essentielles suivantes, qui sans nul doute subsistent en dimension supérieure :

- la nature des résultats obtenus par l'algorithme dépend du type des sources utilisées : gaussiennes, sous-gaussiennes, sur-gaussiennes. Malheureusement, on ne peut identifier ces types que si ils sont communs à toutes les sources ;
- il se peut que l'algorithme ne converge pas. Même s'il converge, on est assuré de la convergence vers une bonne solution seulement dans le cas sous-gaussien. Nous proposons un palliatif dans le cas sur-gaussien, le cas où les sources ne sont pas de même type restant en suspens ;
- par ailleurs, l'initialisation de l'algorithme ne peut être quelconque, puisqu'il peut exister des points de fuite qui le conduisent à diverger. Fort heureusement, il apparaît que la condition initiale nulle est une bonne condition au moins en moyenne comme nous l'observons expérimentalement.

Le reste de cet article s'organise en trois sections : tout d'abord la position du problème, sa réduction et le rappel de la solution de Héroult et Jutten. Nous étudions dans la deuxième section l'équation différentielle moyenne classiquement associée à l'algorithme. Enfin, dans la section 3, nous faisons quelques remarques d'ordre pratique sur la conduite de l'algorithme.

1. Position du problème

On considère p sources (processus) indépendantes $X_1(t), \dots, X_p(t)$, inconnues, pour $t \in [0, T]$. De ces sources, on observe p combinaisons linéaires $E_1(t), \dots, E_p(t)$, pour $t \in [0, T]$. Notant A la matrice (indépendante du temps) dite « de mélange », on a

$$(1.1) \quad E(t) = A \cdot X(t).$$

On suppose ici que la matrice A est inversible. A est inconnue et l'on cherche à retrouver les sources initiales, c'est-à-dire à calculer A^{-1} , ou plus précisément, une matrice B telle que

$$BAX(t) = \Delta \Sigma X(t)$$

où Σ est une matrice de permutation et Δ une matrice diagonale.

La contrainte que se sont imposée a priori Héroult et Jutten est le caractère « neuromimétique », c'est-à-dire local de l'algorithme de recherche. Il doit pouvoir être réalisé par un réseau de neurones formels, chaque neurone utilisant l'information provenant des autres à travers une intensité « synaptique » ou force de liaison (c_{ij}) ($i \neq j$).

Notant $S(t)$ la sortie du réseau et $C(t)$ la matrice des liaisons c_{ij} à l'instant t , on a alors :

$$AX(t) - C(t)S(t) = S(t)$$

Si la matrice $I + C(t)$ est inversible, l'équation précédente s'écrit

$$S(t) = [I + C(t)]^{-1} AX(t).$$

(L'observation du comportement de cet algorithme nous montre que cette condition d'inversibilité n'est pas très astreignante : en pratique les trajectoires de l'algorithme « traversent » la variété d'équation $\det(I + C) = 0$.)

Cela revient donc à chercher une matrice B de la forme $(I + C)^{-1}$.

Pour réaliser ce programme, des considérations de minimisation d'« énergie » ou de « bruit » conduisent Jutten et Héroult à proposer des algorithmes de la forme suivante : on se donne une matrice initiale $C(0)$, et au pas t entier, on calcule $C(t + 1)$ par

$$(1.2) \quad C_{ij}(t + 1) = C_{ij}(t) + \varepsilon \tilde{h}[S_i(t)] \tilde{k}[S_j(t)]$$

où h et k sont des fonctions impaires et régulières, $\tilde{h}(S)$ et $\tilde{k}(S)$ sont les versions centrées de $h(S)$ et $k(S)$, ε est un « petit » paramètre. Leur étude porte essentiellement sur le cas où $h(x) = x^3$ et $k(x) = \text{Arctg } x$.

Le cas où $h(x)$ et $k(x)$ valent x est écarté : il ne conduit qu'à des matrices C symétriques ce qui, en général, ne peut pas séparer les sources. Par conséquent, le cas le plus simple est celui où $h(x) = x^3$ et $k(x) = x$. C'est celui que nous envisageons ici ; l'étude de l'algorithme est alors celle de l'équation :

$$(1.3) \quad C_{ij}(t + 1) = C_{ij}(t) + \varepsilon S_i(t)^3 S_j(t)$$

si nous supposons X centré, ce que nous ferons dans toute la suite.

Cet algorithme est bien « local », puisque la modification de la liaison « synaptique » C_{ij} ne dépend que de la sortie des « neurones » i et j .

L'algorithme (1.2) est bien évidemment une version stochastique et discrète de l'équation, dite équation différentielle moyenne

$$(1.4) \quad \frac{d}{dt} C_{ij} = E \{ \tilde{h}(S_i) \tilde{k}(S_j) \}$$

(1.3) étant la version de

$$(1.5) \quad \frac{d}{dt} C_{ij} = E \{ S_i^3 S_j \}.$$

Si il existe C telle que les différentes composantes S_j soient indépendantes, alors $E(S_i^3 \cdot S_j) = 0$ et cette matrice C est solution de (1.5).

Dans l'équation (1.3) où ε est petit, on estime $E(S_i^3 \cdot S_j)$ par une somme pondérée de $S_i(t)^3 S_j(t)$, les fluctuations individuelles se compensant du fait de l'évolution « lente » de C et de la stationnarité des sources.

Par ailleurs, il est clair que la convergence d'un tel algorithme est due uniquement à la propriété d'indépendance des sources ; il en résulte deux conséquences importantes numériquement :

• l'algorithme ne peut en aucun cas retrouver l'ordre des différentes sources. Si Σ est une matrice de permutation ($p \times p$), on peut aussi écrire (1.1) sous la forme :

$$(1.6) \quad E(t) = (A\Sigma^{-1}) \cdot Y(t), \quad \text{avec} \quad Y(t) = \Sigma X(t)$$

• l'algorithme ne peut pas non plus retrouver l'échelle de chaque $X_i(t)$:

Si Δ est une matrice diagonale, la même observation peut provenir du système (1.1) ou du système

$$(1.7) \quad E(t) = (A\Delta^{-1}) \cdot Y(t), \quad \text{avec} \quad Y(t) = \Delta \cdot X(t).$$

Cette seconde remarque justifie le choix fait d'imposer à tous les coefficients diagonaux de l'estimation de A d'être égaux à 1. Cela permet de supposer (tout au moins pour l'étude théorique) que, si l'on note a_{ij} les coefficients de la matrice A , on a $a_{ii} = 1$.

2. Équation différentielle moyenne

Une première approche de l'évolution de l'algorithme (1.3) consiste en l'étude de l'équation différentielle moyenne (1.5). On sait en effet que, sous réserve que le processus $X(t)$ soit stationnaire et mélangeant, les lignes de champs de cette équation décrivent correctement le comportement moyen de (1.3), ainsi que son comportement limite quand ε tend vers zéro. Les comportements « aberrants » pouvant résulter des fluctuations autour de ces lignes de champ pourraient être décrits de manière plus fine au moyen de théorèmes d'invariance et de grandes déviations (cf. [3, 4]).

Nous menons ici l'étude de l'équation différentielle moyenne dans le cas de la dimension 2 ($p = 2$). Compte tenu des remarques faites ci-dessus, notons alors les matrices A et C de la façon suivante :

$$A = \begin{pmatrix} 1 & b \\ c & 1 \end{pmatrix} \quad \text{et} \quad I + C(t) = \begin{pmatrix} 1 & u(t) \\ v(t) & 1 \end{pmatrix}.$$

Nous supposons bien sur l'existence de moments de X jusqu'à l'ordre 4, et l'équation (1.5) s'écrit alors

$$\begin{aligned} du/dt = & (1 - uv)^{-4} [(1 - cu)^3 (c - v) EX_1^4 + \\ & + (b - u)^3 (1 - bv) EX_2^4 \\ & + 3(1 - cu)(b - u) \{(1 - cu)(1 - bv) \\ & + (b - u)(c - v)\} EX_1^2 EX_2^2] \end{aligned}$$

$$\begin{aligned} dv/dt = & (1 - uv)^{-4} [(1 - cu)(c - v)^3 EX_1^4 + \\ & + (b - u)(1 - bv)^3 EX_2^4 \\ & + 3(1 - bv)(c - v) \{(1 - cu)(1 - bv) \\ & + (b - u)(c - v)\} EX_1^2 EX_2^2]. \end{aligned}$$

Comme ce sont des équations de degré 4, il peut y avoir jusqu'à 16 points stationnaires, i.e. annulant simultanément du/dt et dv/dt . Cependant, les seuls points pour lesquels il y a bien séparation sont

$$P_1 = (b \quad c), \quad \text{qui correspond à } S_1 = X_1 \text{ et } S_2 = X_2$$

et

$$P_2 = (1/c \quad 1/b), \quad \text{qui correspond à } S_1 = -bX_2 \text{ et } S_2 = -cX_1$$

STABILITÉ LOCALE AUX POINTS P_1 ET P_2

La première question intéressante est la nature de ces points P_1 et P_2 . Nous étudions donc la stabilité de (1.5) au voisinage de ces points par linéarisation.

En ce qui concerne P_1 , posons

$$\begin{aligned} u &= b + \beta \\ v &= c + \gamma \end{aligned}$$

On a alors

$$\begin{aligned} E(S_1^3 \cdot S_2) = & (1 - bc)^{-1} \times \\ & \times (3 - \beta EX_1^2 EX_2^2 - \gamma EX_1^4) + O(\beta^2 + \gamma^2) \end{aligned}$$

et

$$\begin{aligned} E(S_1 \cdot S_2^3) = & (1 - bc)^{-1} \times \\ & \times (-3 \gamma EX_1^2 EX_2^2 - \beta EX_2^4) + O(\beta^2 + \gamma^2). \end{aligned}$$

On est amené à considérer les valeurs propres de la matrice

$$-\frac{1}{1 - bc} \begin{pmatrix} 3 EX_1^2 EX_2^2 & EX_1^4 \\ EX_2^4 & 3 EX_1^2 EX_2^2 \end{pmatrix}.$$

Posons

$$G = EX_1^4 \cdot EX_2^4 - 9[EX_1^2 \cdot EX_2^2]^2.$$

Si G est positif, les deux valeurs propres sont de signes contraires et le point P_1 est instable.

Lorsque G est négatif, les deux valeurs propres sont de même signe :

si $1 - bc$ est positif, les deux valeurs propres sont négatives et P_1 est stable ;

si $1 - bc$ est négatif, les deux valeurs propres sont positives et P_1 est répulsif suivant les directions principales.

En ce qui concerne P_2 , posant $v = 1/b + \beta$ et $u = 1/c + \gamma$, on est conduit à considérer les valeurs propres de la matrice

$$-\frac{1}{(1 - bc) b^3 c^3} \begin{pmatrix} 3 b^2 c^2 EX_1^2 EX_2^2 & b^4 EX_1^4 \\ c^4 EX_2^4 & 3 b^2 c^2 EX_1^2 EX_2^2 \end{pmatrix}.$$

Si G est positif, les deux valeurs propres sont de signes contraires et le point P_2 est instable.

Lorsque G est négatif, les deux valeurs propres sont de même signe :

si $bc(1 - bc)$ est négatif, P_2 est stable

si $bc(1 - bc)$ est positif, P_2 est antistable.

La discussion selon le signe de G correspond en un certain sens à la comparaison des lois de X_1 et X_2 à la loi gaussienne. Introduisons quelques définitions nécessaires à la classification des cas que l'on peut rencontrer :

Définition 2.1 : Soit V une variable aléatoire réelle, notons $C_4(V)$ le cumuland d'ordre 4 de V :

$$C_4(V) = E \{V^4\} - 3[E \{V^2\}]^2.$$

Il a les propriétés suivantes :

- 1) si V suit une loi gaussienne, $C_4(V) = 0$,
- 2) si $(V_i)_{i=1, \dots, k}$ sont indépendantes et $C_4(V_i) > 0$ (resp. $C_4(V_i) < 0$, $C_4(V_i) = 0$) alors toute combinaison linéaire des V_i , $W = \sum \alpha_i V_i$ vérifie également $C_4(W) > 0$ (resp. $C_4(W) < 0$, $C_4(W) = 0$).

Définition 2.2 : Nous dirons qu'une variable V est de type « gaussien » si $C_4(V) = 0$, « surgaussien » si $C_4(V) > 0$, « sous-gaussien » si $C_4(V) < 0$.

Par conséquent quel que soit le nombre de sources (X_i) si elles sont toutes gaussiennes, sous-gaussiennes ou surgaussiennes, cela reste vrai pour les variables observées (E_i) et les sorties (S_i) . Ainsi le signe de G est indépendant du mélange et peut être connu après estimation des moments des E_i ou bien des S_i .

Par contre, lorsque les sources (X_i) ne sont pas toutes de même type, le signe de G dépend des rapports

$$a_i = \frac{EX_i^4}{3(EX_i^2)^2}$$

et ne peut être connu par la seule observation de E ou de S .

Définition 2.3 : Pour un couple de variables nous appelons

- PG (« pseudo-gaussien ») le cas $G = 0$, i.e. $a_1 a_2 = 1$
- PSG (« pseudo-surgaussien ») le cas $G > 0$, i.e. $a_1 a_2 > 1$
- PsG (« pseudo-sousgaussien ») le cas $G < 0$, i.e. $a_1 a_2 < 1$.

Cas pseudo-sousgaussien :

Dans ce cas (et a fortiori dans le cas sousgaussien), il existe toujours une solution stable au problème posé. Plus

précisément, il existe toujours une unique solution stable dans le « diamant » $|uv| < 1$.

Si de plus $bc < 0$, les deux solutions sont stables ; il se peut donc qu'une solution stable n'appartienne pas au diamant (cf. fig. 2.2).

Il nous faut cependant noter que dans le cas où il n'y a qu'une solution stable (i.e. $bc > 0$), il existe en outre deux solutions instables, points de fuite, comme le montre la figure 2.1 représentant un cas typique.

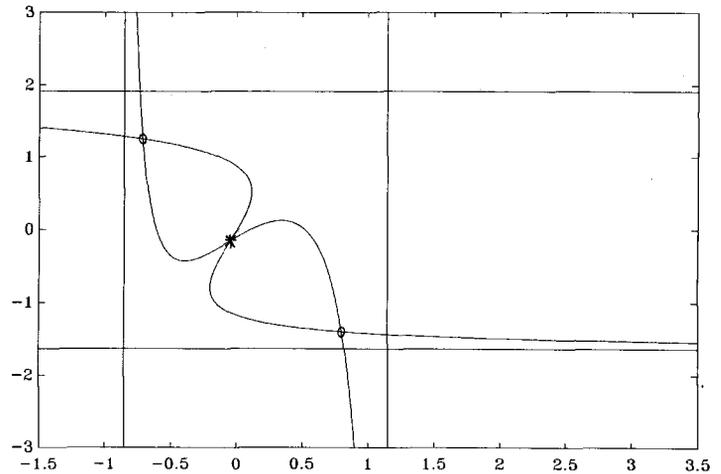


Fig. 2.2. — $bc < 0$. Deux solutions stables \circ séparées par une solution instable $*$.

Cas pseudo-gaussien :

L'équation (1.5) se réduit alors à

$$\begin{aligned} du/dt &= 3(1-uv)^{-4} [(1-cu)^2 EX_1^2 + (b-u)^2 a_2 EX_2^2] H(u, v) \\ dv/dt &= 3(1-uv)^{-4} [(c-v)^2 EX_1^2 + (1-bv)^2 a_2 EX_2^2] H(u, v) \end{aligned}$$

où

$$H(u, v) = a_1(1-cu)(c-v) EX_1^2 + (b-u)(1-bv) EX_2^2$$

En conséquence l'ensemble des points stationnaires est une hyperbole. Il faut remarquer que cette hyperbole ne coïncide avec l'hyperbole de décorrélation de S_1 et S_2

$$H(u, v) = (1-cu)(c-v) EX_1^2 + (b-u)(1-bv) EX_2^2$$

que si l'on se trouve dans le cas gaussien.

Cas pseudo-surgaussien :

Les deux points d'intérêt sont instables et par conséquent inaccessibles même en inversant le sens des trajectoires. En revanche il existe deux points parasites stables, comme le montrent les tracés de quelques trajectoires de (1.4) dans deux cas représentatifs (cf. fig. 2.3 et 2.4).

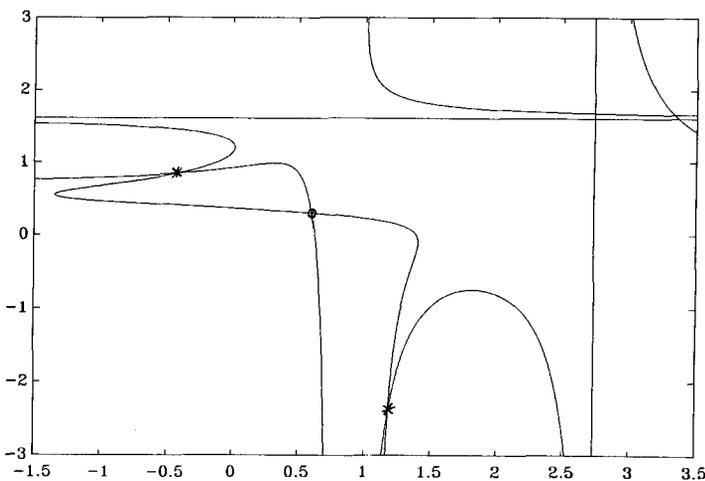


Fig. 2.1. — $bc > 0$. Outre la solution stable $(0,6, 0,3)$ (notée \circ), il y a deux solutions parasites (notées $*$), qui sont de surcroît des points de fuite vers l'infini.

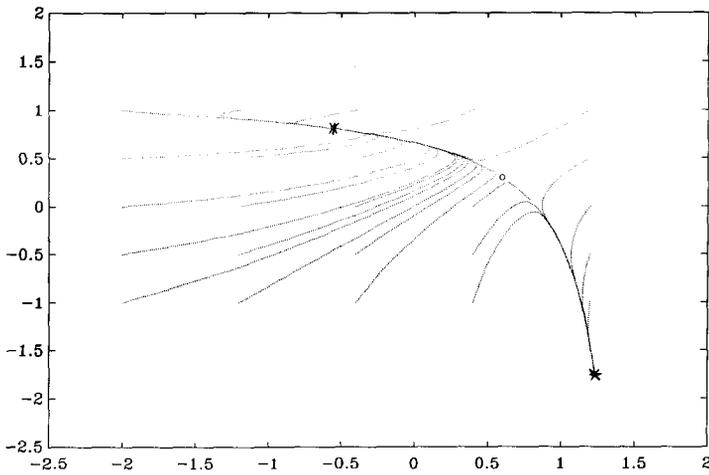


Fig. 2.3. — Cas PSG, $b = 0,6$, $c = 0,3$. Les trajectoires se dirigent vers les points parasites stables * en évitant le point d'intérêt \circ de coordonnées (0,6, 0,3).

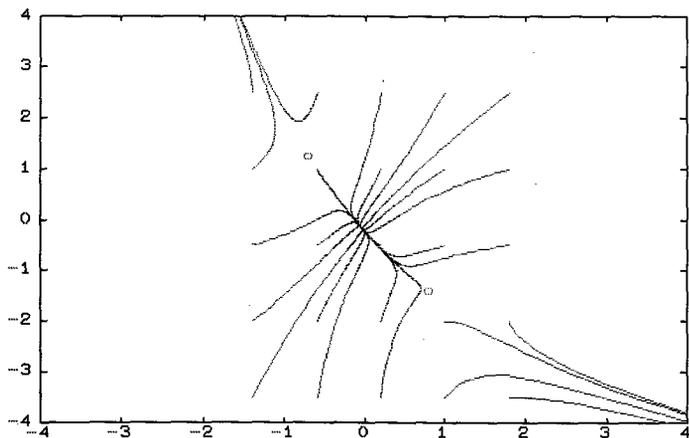


Fig. 2.4. — PSG, $b = 0,8$, $c = -1,4$, $bc < 0$. Les trajectoires évitent les points d'intérêt \circ et s'enfuient ou se dirigent vers le point parasite stable.

En résumé, il ressort donc que dans tous les cas l'algorithme converge pour certaines conditions initiales. En pratique, il s'agit de savoir si il converge vers les solutions recherchées ou vers des solutions parasites. Pour cela, on peut estimer les moments centrés d'ordres 2 et 4 des mélanges observés : si ces deux mélanges sont sousgaussiens, on est assuré que les points limites observés correspondent à de bonnes solutions ; s'ils sont tous les deux surgaussiens, ce sont des solutions parasites. Malheureusement, lorsque les deux mélanges ne sont pas de même nature, il n'est pas possible de prévoir le type des points limites. Une étude des moments croisés d'ordre supérieurs peut néanmoins permettre de savoir si les sources sont réellement indépendantes. Cependant la conclusion reste délicate puisque les sources présentent une apparente indépendance jusqu'à l'ordre 4.

Remarques :

1) Il est à noter que les points « parasites », c'est-à-dire instables ou points de fuite dans le cas PsG, et les points stables parasites dans le cas PSG sont alignés avec 0. En effet, on remarque que la transformation sur les paramètres

$$(u, v) \rightarrow (1/v, 1/u)$$

échange S_1 et S_2 en vS_2 et uS_1 , ce qui, bien sûr, conduit aux mêmes points stables dans (1.4).

2) Bien que nous n'en fassions pas la démonstration théorique, il nous semble bon de souligner que, dans le cas PsG, le point 0, qui correspond à l'initialisation par la matrice identique, se trouve pour l'équation (1.5) toujours dans le bassin d'attraction du point stable situé dans le « diamant » $|uv| < 1$. La figure 2.5 le montre clairement.

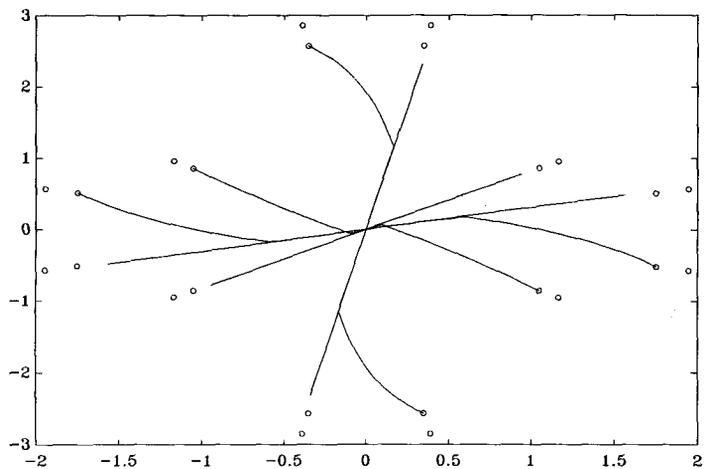


Fig. 2.5. — Pour différentes valeurs de b et c telles que $|bc|$ proche de 1, les trajectoires issues de 0 se dirigent vers (b, c) .

3. L'algorithme stochastique

Il s'agit de l'algorithme défini par l'équation (1.2), c'est-à-dire un algorithme à pas constant. Il ne peut prétendre à converger au sens propre du terme vers une solution du problème, mais à osciller faiblement autour de cette solution. Bien entendu l'existence de points de fuite et la théorie des grandes déviations [4] montrent que cela ne peut durer indéfiniment, cependant nous parlerons de convergence.

1) Dans un premier temps, nous avons confronté le comportement de l'algorithme à celui prévu par l'équation (1.4). Les sources sont des suites de variables aléatoires i.i.d. Nous avons choisi des combinaisons linéaires de variables uniformes sur $[0,1]$ pour représenter le cas sous-gaussien et des combinaisons linéaires de variables exponentielles pour le cas surgaussien. Bien évidemment le comportement de (1.3) ($h = x^3$, $k = x$) est conforme à la

description donnée au § 2. Quant à l'algorithme utilisé par Jutten et Héroult ($h = x^3$, $k = \text{Arctg } x$) l'allure de ses trajectoires est qualitativement la même (cf. fig. 3.1) : dans le cas PsG, les points stables sont les mêmes et l'existence de points de fuite demeure. Dans le cas PSG, les points stables parasites sont encore présents, bien que différents.

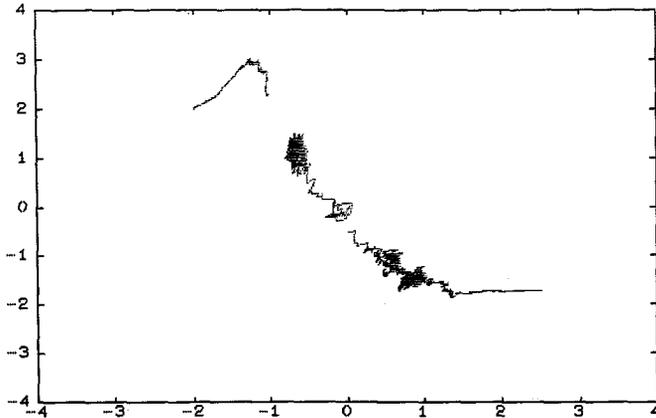


Fig. 3.1 a. — Cas sousgaussien $b = 0,8$, $c = -1,4$, équation (1.3), 3 500 pas, $\epsilon = 4.10^{-3}$. Les trajectoires aléatoires se dirigent vers les solutions stables.

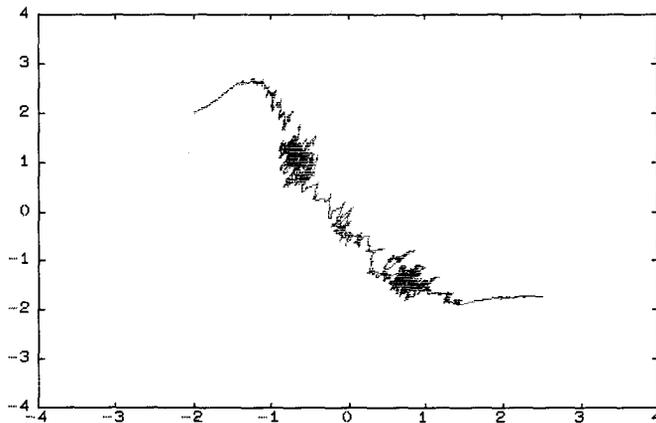


Fig. 3.1 b. — Cas sousgaussien $b = 0,8$, $c = -1,4$, équation (1.2), 3 500 pas, $\epsilon = 4.10^{-3}$. Les trajectoires aléatoires se dirigent vers les solutions stables.

Dans le cas surgaussien, les trajectoires ne se dirigent pas vers la bonne solution, mais vers un point parasite (cf. fig. 3.2).

2) La conduite de l'algorithme est justifiable de la remarque suivante : si au lieu d'utiliser brutalement (1.2) ou (1.3), on calcule une estimation autorégressive des moments croisés d'ordre 4 de S_1 et S_2 dans (1.3),

$$M_{ij}(t+1) = (1 - \alpha) M_{ij}(t) + \alpha S_1^i(t+1) S_2^j(t+1)$$

ou bien des moments croisés de $h(S_1)$ et $k(S_2)$ dans (1.2), on obtient des trajectoires un peu moins chaotiques sans détruire le comportement moyen donné par (1.4) et (1.5).

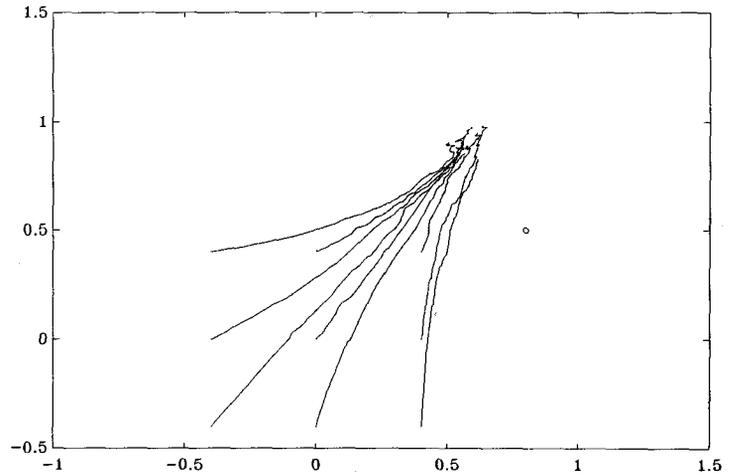


Fig. 3.2 a. — Cas surgaussien, $b = 0,8$, $c = 0,5$, équation (1.2), $\epsilon = 4.10^{-3}$.

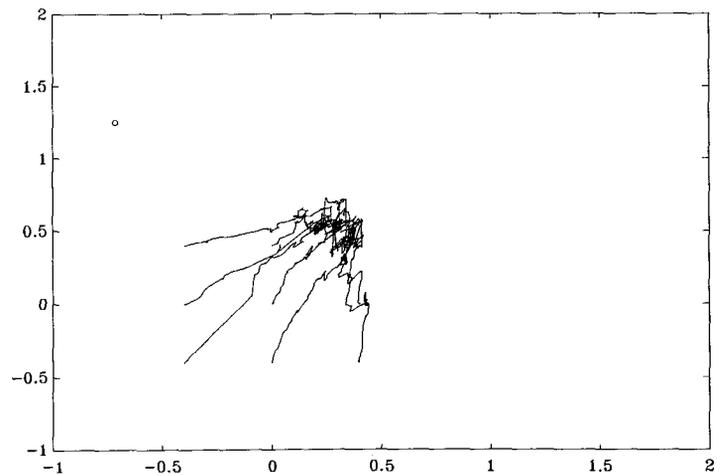


Fig. 3.2 b. — Cas surgaussien, $b = 0,8$, $c = -1,4$, équation (1.2), $\epsilon = 4.10^{-3}$.

Les quelques exemples suivants (fig. 3.3) l'illustrent bien. Le fait de conserver « en mémoire » l'estimation antérieure n'est pas fondamentalement contraire à l'esprit « réseau

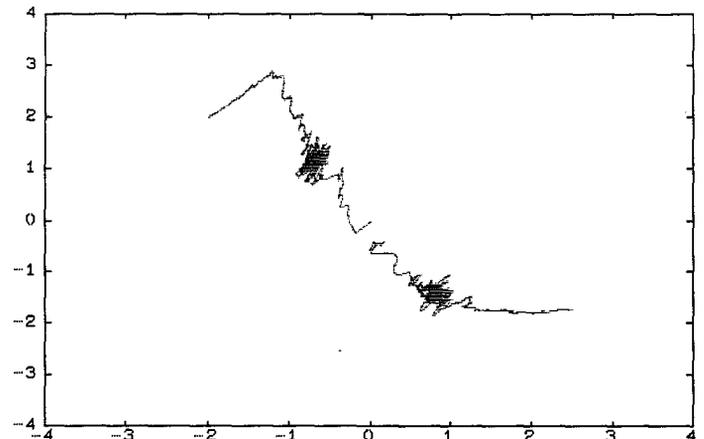


Fig. 3.3. — Équation (1.2), 3 500 pas, $\epsilon = 4.10^{-3}$, $\alpha = 0,1$ lorsque la mémoire est grande (α petit), les trajectoires sont plus lisses (comparer aussi avec la figure 3.1).

de neurones». Cela suppose simplement que l'état à l'instant t est combinaison linéaire de l'état à l'instant $t - 1$ et de la nouvelle information parvenue à l'instant t .

3) Dans le premier article de Jutten-Hérault [1] il semble que la nature des sources soit de peu d'importance quant au résultat obtenu, tout au moins tant qu'elles sont indépendantes. Ce n'est en fait pas le cas. Nous avons fait quelques simulations où les sources ne sont plus des suites de variables i.i.d. et l'on constate que les résultats peuvent être perturbés.

1^{er} cas : $X_i(t)$ est un processus autorégressif stationnaire

$$X_i(t + 1) = \alpha_i X_i(t) + \eta_i(t + 1)$$

où les η_i forment une suite i.i.d. de variables sous-gaussiennes. Lorsque α_i croît la qualité de l'estimation du mélange s'altère. Pour des α_i de l'ordre de 0,8, l'algorithme ne peut pratiquement pas converger vers les solutions espérées, comme le montre la figure 3.4.

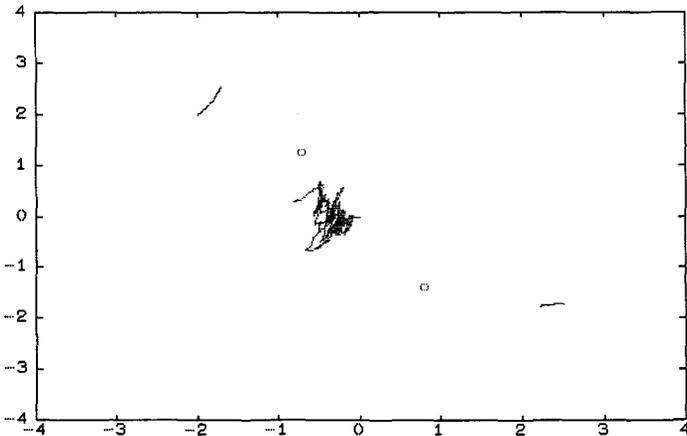


Fig. 3.4. — Équation (1.2), $h(x) = x^3$, $k(x) = \text{Arctg } x$, 5 000 pas, $\epsilon = 0,02$, $\alpha_1 = \alpha_2 = 0,8$.

2^e cas : les deux sources sont déterministes (a) ou bien légèrement bruitées (b).

(a) Si les deux sources sont déterministes mais ressemblent suffisamment à des processus indépendants, les moments croisés empiriques tendent vers zéro quand t tend vers l'infini, alors il y a encore convergence. C'est le cas pour l'exemple suivant (fig. 3.5).

(b) Si les deux sources sont déterministes dont la corrélation empirique est non nulle, auxquelles on a ajouté un bruit blanc sous gaussien, si le rapport signal/bruit est trop élevé, l'information fournie par le bruit ne suffit pas pour assurer la convergence. En voici un exemple (fig. 3.6).

4) Dans le cas pseudo surgaussien, on peut mettre au point un algorithme amenant le point courant à se déplacer sur l'hyperbole de décorrélation en remontant le champ précédent. Néanmoins, ce champ n'étant pas exactement tangent à cette hyperbole, cet algorithme se révèle inefficace comme le montre la figure 3.7 où le point cherché n'est pas stable.

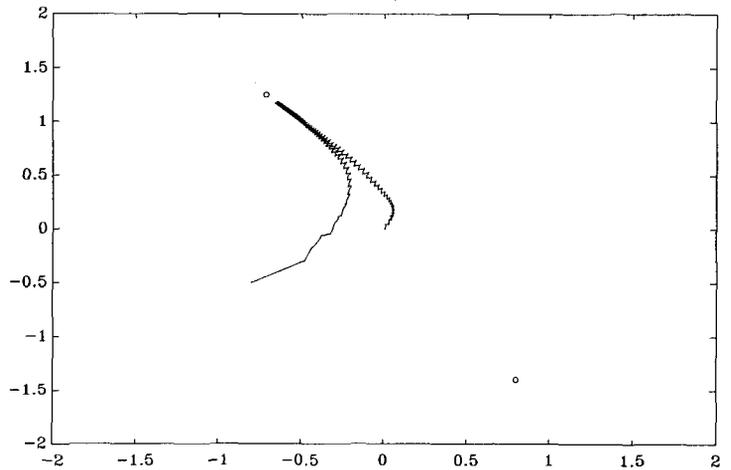


Fig. 3.5. — Équation (1.2), $h(x) = x^3$, $k(x) = \text{Arctg } x$. Sources déterministes : $X_1 = \sin 2 \pi t/50$, $X_2 = \sin 2 \pi (t - 17)/31$.

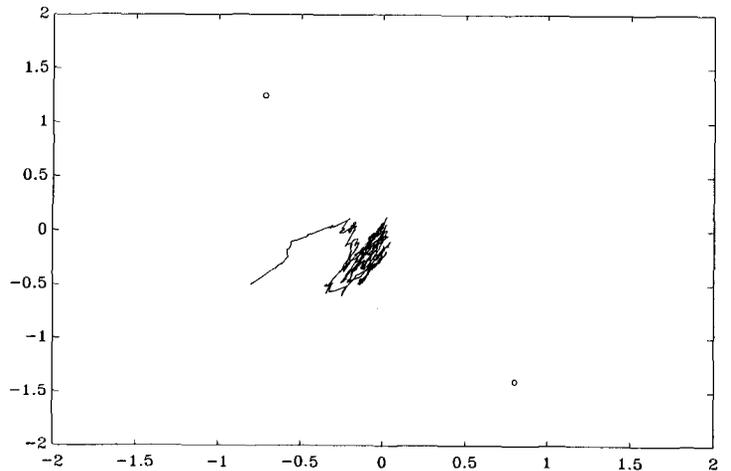


Fig. 3.6. — Équation (1.2), $h(x) = x^3$, $k(x) = \text{Arctg } x$. $X_1 = \sin 2 \pi t/50 + \eta_1(t)$, $X_2 = \sin 2 \pi (t - 17)/25 + \eta_2(t)$ où $\eta_1(t)$ et $\eta_2(t)$ sont deux bruits blancs sous-gaussiens de variance 0,1.

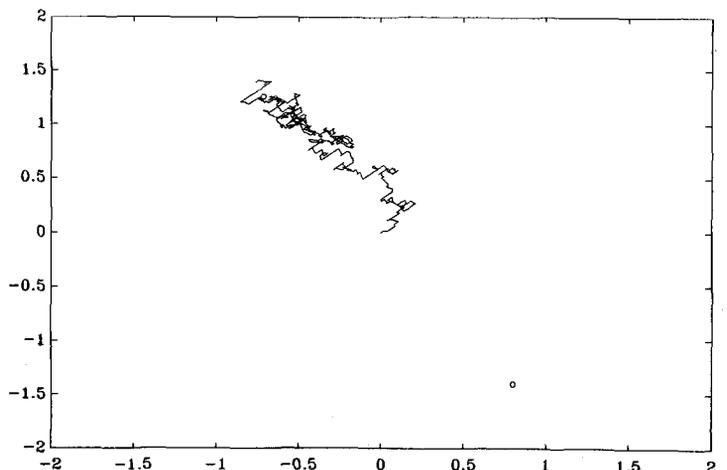


Fig. 3.7. — Les sources sont des suites de v.a. exponentielles recentrées.

4. Conclusion

Cet article présente un début d'étude de l'algorithme proposé par Hérault et Jutten. Les résultats confirment leurs travaux, avec cependant une restriction importante dont il faudra tenir compte dans la mise en œuvre effective de l'algorithme : de la nature des sources dépend la convergence vers une solution convenable.

Nos résultats sont encore limités puisqu'ils ne concernent que le cas de deux sources. Néanmoins, ils sont symptomatiques du cas général et nous présenterons bientôt une étude pour un nombre quelconque de sources, où nous débouchons sur une conclusion analogue.

En tout état de cause, il serait souhaitable de proposer un algorithme de même type susceptible de traiter simultanément le cas « pseudo-sousgaussien » et le cas « pseudo-sur-gaussien », dont nous avons montré qu'il échappe à

l'algorithme de Hérault et Jutten. Le cas « gaussien », lui, ne relève pas de cet algorithme.

Premier manuscrit reçu le 9 mai 1990.

Mes remerciements vont à B. Prum pour son aide et ses encouragements.

BIBLIOGRAPHIE

- [1] C. JUTTEN, J. HÉRAULT, Une solution neuromimétique au problème de séparation de sources. *Traitement du signal*, vol. 5, N° 6 (1988) p. 389-403.
- [2] C. JUTTEN, J. HÉRAULT, P. COMON, E. SOROUCHYARI, Separation of sources. Part I, II, III à paraître dans *Signal Processing* (1990).
- [3] A. BENVENISTE, M. METIVIER, P. PRIOURET, *Algorithmes adaptatifs et approximations stochastiques*. Masson (1987).
- [4] M. COTTRELL, J. C. FORT, G. MALGOUYRES, Large deviations and rare events in the study of stochastic algorithms. *IEEE-AC* 28 (1983).