

SAUD, un algorithme d'ICA par déflation semi-algébrique

Laurent ALBERA^{1,2}, Pierre COMON³, Hong XU³

¹INSERM, U 642, Rennes, F-35000 France

²Université de Rennes 1, LTSI, F-35000 France

³I3S, UNSA, CNRS UMR6070, Algorithmes-Euclide-B, Sophia-Antipolis, F-06903 France

laurent.albera@univ-rennes1.fr, pcomon@i3s.unice.fr, hongxu@i3s.unice.fr

Résumé – Nous proposons dans cet article une nouvelle technique d'ICA baptisée SAUD (Semi-Algebraic Unitary Deflation) et basée sur une procédure de déflation unitaire semi-algébrique. SAUD permet ainsi d'identifier une à une les composantes indépendantes sans souffrir des inconvénients des approches adaptatives (lenteur de convergence, etc.). Des simulations numériques en contexte biomédical (EEG) illustrent le bon comportement de SAUD face à COM2, DEFA et FastICA.

Abstract – We propose in this paper a new ICA method, namely SAUD, based on a Semi-Algebraic Unitary Deflation procedure. SAUD is able to identify one by one the independent components avoiding the drawbacks of an adaptive solution (slow convergence, etc.). Computer results, representative of the biomedical context (EEG), show the good behavior of SAUD compared with other algorithms such as COM2, DEFA and FastICA.

1 Introduction

L'intérêt pour l'Analyse en Composantes Indépendantes (ICA) n'a cessé de grandir et ce depuis les travaux précurseurs de C. Jutten et al. [10]. Nombre d'approches ont vu le jour, chacune visant à traduire la notion d'indépendance statistique sous une forme algorithmiquement exploitable [1, 3, 5, 8, 11]. Parmi elles, certaines ont établi un lien entre l'ICA et la décomposition canonique de tableaux cumulants d'ordre supérieur, proposant alors de nouveaux outils de décomposition tensorielle [1, 3, 5].

Si la séparation de sources rencontre un grand succès, c'est en partie dû au fait qu'elle trouve sa place dans grand nombre d'applications telles que les télécommunications, le tatouage, le traitement de la parole ou bien encore le domaine biomédical. Ainsi, des travaux visant l'élaboration d'un système ambulatoire multi-varié permettant d'explorer, sur une longue durée et à domicile, les fonctions neurologiques pour le diagnostic de différents troubles du sommeil, ont récemment été menés [12]. Ces derniers ont conduit à l'élaboration d'un nouveau système d'acquisition composé de seulement quatre électrodes et d'une référence en lieu et place des systèmes standards d'enregistrement composés au minimum d'une dizaine de capteurs. Cette réduction du nombre d'électrodes a cependant nécessité l'utilisation de l'ICA afin d'identifier les composantes d'intérêt pour le clinicien.

Néanmoins, et tel est le cas pour l'application précédente, il n'est pas toujours nécessaire d'extraire l'ensemble des composantes, d'où l'émergence des techniques d'ICA dites de déflation visant à identifier les composantes l'une après l'autre. La première méthode à voir le jour, que nous nommerons DEFA (DEFlation Approach for ICA),

fut proposée à la fin du siècle dernier [7]. D'autres suivront telles que la version non symétrique de FastICA [8] ou bien, plus récemment, l'approche RobustICA [14]. Si DEFA [7] est remarquable d'un point de vue théorique, elle souffre en pratique des inconvénients liés à l'utilisation de l'algorithme du gradient.

Nous proposons dans cet article une nouvelle technique d'ICA baptisée SAUD (Semi-Algebraic Unitary Deflation) et basée sur une procédure de déflation unitaire semi-algébrique. SAUD permet ainsi d'identifier une à une les composantes indépendantes sans souffrir des inconvénients des approches adaptatives (lenteur de convergence, etc.). Des simulations numériques en contexte biomédical illustrent le bon comportement de SAUD face à COM2, DEFA et FastICA.

2 Préliminaires et problématique

L'ICA peut être interprétée comme une extension de l'Analyse en Composantes Principales (PCA) qui pour sa part n'impose l'indépendance statistique des composantes qu'à l'ordre deux. Avant de définir plus en détails l'ICA, rappelons la définition d'un contraste [5]. Celle-ci nécessite d'introduire quelques notations. En particulier, notons \mathcal{F} l'espace vectoriel des vecteurs aléatoires à valeur dans \mathbb{R}^P , \mathcal{S} le sous-espace vectoriel de \mathcal{F} dont les éléments ont tous des composantes mutuellement indépendantes, \mathcal{H} l'espace des matrices réelles de taille $(P \times P)$, $\mathcal{H} \cdot \mathcal{S}$ le sous-espace de \mathcal{F} des vecteurs aléatoires obtenus par une transformation affine issue de \mathcal{H} d'un vecteur de \mathcal{S} , \mathbf{I}_P la matrice identité de \mathcal{H} , et \mathcal{T} le sous-espace de \mathcal{H} des matrices *triviales* (i.e. de la forme $\mathbf{D}\mathbf{\Pi}$ où \mathbf{D} est une matrice diagonale de \mathcal{H} et $\mathbf{\Pi}$ une matrice de permutation de \mathcal{H}).

Définition 1 Un contraste ψ est une application de $\mathcal{H} \times \mathcal{H} \cdot \mathcal{S}$ dans \mathbb{R} satisfaisant les deux propriétés suivantes :

- i) la valeur du contraste est maximale pour tout vecteur de \mathcal{S} , i.e. $\forall \mathbf{s} \in \mathcal{S}, \forall \mathbf{A} \in \mathcal{H}, \psi(\mathbf{A}, \mathbf{A}\mathbf{s}) \leq \psi(\mathbf{I}_P, \mathbf{s})$;
- ii) deux valeurs de contraste sont égales si et seulement si les deux matrices de \mathcal{H} concernées sont égales à une matrice triviale près, i.e. $\forall \mathbf{s} \in \mathcal{S}, \psi(\mathbf{A}, \mathbf{A}\mathbf{s}) = \psi(\mathbf{I}_P, \mathbf{s}) \Leftrightarrow \mathbf{A} \in \mathcal{T}$.

Nous pouvons dès lors définir l'ICA d'un vecteur aléatoire à valeur dans \mathbb{R}^P .

Définition 2 L'ICA d'un vecteur aléatoire \mathbf{x} de longueur N et de matrice de covariance $\mathbf{R}_{\mathbf{x}}$ est donnée par le couple de matrices (\mathbf{A}, Δ) vérifiant les propriétés suivantes :

- i) la matrice de covariance $\mathbf{R}_{\mathbf{x}}$ se décompose sous la forme $\mathbf{R}_{\mathbf{x}} = \mathbf{A}\Delta\mathbf{A}^T$ où Δ est une matrice semi-définie positive et où \mathbf{A} est une matrice de rang colonne plein ;
- ii) le vecteur \mathbf{x} se décompose sous la forme $\mathbf{x} = \mathbf{A}\mathbf{s}$ où \mathbf{s} est un vecteur aléatoire de longueur P , de matrice de covariance Δ et dont les composantes sont le plus indépendantes possibles au sens de la maximisation du contraste choisi.

Des deux définitions précédentes découle l'indétermination suivante, qui, rappelons-le, n'est en général pas contraignante d'un point de vue applicatif.

Propriété 1 Soit (\mathbf{A}, Δ) le couple de matrices caractérisant l'ICA d'un vecteur aléatoire \mathbf{x} , il en est de même pour le couple (\mathbf{A}', Δ') vérifiant $\mathbf{A}' = \mathbf{A}\mathbf{T}$, où \mathbf{T} est une matrice de \mathcal{T} .

Le choix du contraste est donc crucial dans la réalisation d'une ICA. L'information mutuelle, qui n'est autre que la distance de Kullback entre la densité de probabilité conjointe des composantes estimées et le produit des densités marginales des mêmes composantes, est l'un des contrastes les plus naturels [5]. Cependant sa mise en oeuvre n'en reste pas moins difficile de par l'estimation des densités qu'il requiert. Un contraste plus facilement exploitable peut être obtenu au moyen des cumulants d'ordre 4 du vecteur aléatoire \mathbf{x} dont on souhaite réaliser l'ICA (voir [2, section III] pour un rappel sur les cumulants d'ordre 4). En pratique, ce contraste est plutôt construit en fonction des cumulants d'ordre 4 du vecteur blanchi de \mathbf{x} .

Définition 3 On appelle blanchiment spatial (ou standardisation) d'un vecteur aléatoire \mathbf{x} l'opération qui consiste à identifier la matrice de passage Θ telle que le vecteur $\tilde{\mathbf{x}} = \Theta\mathbf{x}$ ait une matrice de covariance égale à l'identité. On dit que $\tilde{\mathbf{x}}$ est le vecteur blanchi de \mathbf{x} .

Le blanchiment du vecteur \mathbf{x} considéré conduit tout simplement à réaliser une ACP de \mathbf{x} . En outre, le blanchiment a pour utilité, non seulement de réduire la dimension du vecteur \mathbf{x} si celle-ci était supérieure au nombre de composantes principales de \mathbf{x} , mais également d'orthogonaliser la matrice \mathbf{A} de la définition (2), facilitant sa recherche.

Propriété 2 Soit $\tilde{\mathbf{x}} = \Theta\mathbf{x}$ le vecteur aléatoire \mathbf{x} blanchi et $(\tilde{\mathbf{A}}, \Delta)$ le couple de matrices caractérisant l'ICA de $\tilde{\mathbf{x}}$. La matrice $\tilde{\mathbf{A}}$ est alors orthogonale et le couple $(\Theta^{-1}\tilde{\mathbf{A}}, \Delta)$ caractérise l'ICA de \mathbf{x} .

Le contraste employé dans COM2 [5], basé sur les cumulants d'ordre 4 du vecteur blanchi de \mathbf{x} est décrit ci-dessous.

Théorème 1 L'application ψ définie par :

$$\psi(\mathbf{U}, \mathbf{s}) = \sum_{p=1}^P (\mathcal{C}_{p,p,p,p,\mathbf{s}})^2 \quad (1)$$

où $\mathcal{C}_{n_1, n_2, n_3, n_4, \mathbf{s}}$ désigne la (n_1, n_2, n_3, n_4) -ième composante du tableau cumulant d'ordre 4, $\mathcal{C}_{4, \mathbf{s}}$, du vecteur $\mathbf{s} = \mathbf{U}^T \tilde{\mathbf{x}}$ où $\tilde{\mathbf{x}}$ est un vecteur blanchi de \mathcal{F} et \mathbf{U} une matrice orthogonale, est un contraste [5].

Nous allons dans la section suivante présenter la méthode SAUD qui consiste à réaliser par déflation l'ICA du vecteur blanchi $\tilde{\mathbf{x}}$ en exploitant le fait que la matrice de facteurs $\tilde{\mathbf{A}}$ recherchée est orthogonale.

3 La méthode SAUD

Soit $\tilde{\mathbf{A}}(:, \sigma(1))$ l'un des vecteurs colonnes de $\tilde{\mathbf{A}}$ recherché lors de la première étape de SAUD, et soit $s_{\sigma(1)}$ la composante indépendante associée définie par la relation $s_{\sigma(1)} = \tilde{\mathbf{A}}(:, \sigma(1))^T \tilde{\mathbf{x}}$, où σ est une permutation de $\{1, 2, \dots, P\}$. Considérons à présent le lemme suivant [13] :

Lemme 1 Tout vecteur ligne normalisé \mathbf{q} de \mathbb{R}^P dont la dernière composante est strictement positive, peut être représenté comme la dernière ligne d'une matrice orthogonale \mathbf{Q}_{β} donnée par $\mathbf{Q}_{\beta} = \mathbf{Q}_{\beta_1} \cdots \mathbf{Q}_{\beta_{P-1}}$ où les $P-1$ composantes du vecteur $\beta = [\beta_1, \dots, \beta_{P-1}]^T$ sont des angles appartenant à l'intervalle $]-\pi/2, \pi/2[$ et où les $P-1$ matrices \mathbf{Q}_{β_p} sont des matrices de rotation de Givens données par :

$$\mathbf{Q}_{\beta_p} = \begin{bmatrix} \mathbf{I}_{p-1} & 0 & \cdots & 0 \\ 0 & \cos \beta_p & 0 & -\sin \beta_p \\ \vdots & 0 & \mathbf{I}_{P-1-p} & 0 \\ 0 & \sin \beta_p & 0 & \cos \beta_p \end{bmatrix} \quad (2)$$

D'après le lemme 1, sous l'hypothèse que la dernière composante de $\tilde{\mathbf{A}}(:, \sigma(1))$ est non nulle, il existe une matrice $\mathbf{Q}_{\beta^{(1)}}$ orthogonale dont la dernière ligne correspond au vecteur $\text{Sgn}\{\tilde{\mathbf{A}}(P, \sigma(1))\} \tilde{\mathbf{A}}(:, \sigma(1))^T / \|\tilde{\mathbf{A}}(:, \sigma(1))\|$ où $\text{Sgn}\{\cdot\}$ désigne la fonction signe. La méthode DEFA consiste à déterminer le vecteur de paramètres $\beta^{(1)}$, et donc le vecteur $\tilde{\mathbf{A}}(:, \sigma(1))$, à un scalaire près, à l'aide du théorème suivant :

Théorème 2 La maximisation du critère suivant :

$$\psi(\beta, \mathbf{s}) = (\mathcal{C}_{1,1,1,1,\mathbf{s}})^2 \quad (3)$$

où $\mathcal{C}_{1,1,1,1,\mathbf{s}}$ est le cumulants marginaux d'ordre 4 de la variable s définie par $s = \mathbf{Q}_{\beta}(P, \cdot) \tilde{\mathbf{x}}$, permet l'identification du vecteur de paramètres $\beta^{(1)}$.

En d'autres termes, (3) est un contraste au sens défini dans [6, ch.3, p.78]. La méthode DEFA réalise cette optimisation à l'aide de l'algorithme du gradient. Notons que la maximisation de (3) assure l'indépendance mutuelle de $s_{\sigma(1)}$ et des $P-1$ autres composantes extraites lors des étapes suivantes de DEFA.

3.1 Identification du premier facteur

Une alternative à l'optimisation de (3) consiste à exploiter le théorème suivant [5, p.294] :

Théorème 3 *Soit \mathbf{s} un vecteur aléatoire de composantes mutuellement indépendantes, parmi lesquelles au plus une est gaussienne, et dont les densités de probabilité ne sont pas réduites à une masse ponctuelle. Soit \mathbf{T} une matrice orthogonale de taille $(P \times P)$ et \mathbf{z} le vecteur défini par $\mathbf{z} = \mathbf{T}\mathbf{s}$. Alors les trois propriétés suivantes sont équivalentes : i) les composantes de \mathbf{z} sont indépendantes deux à deux, ii) les composantes de \mathbf{z} sont mutuellement indépendantes, et iii) \mathbf{T} est une matrice triviale.*

Au vu de l'indétermination de l'ICA (voir propriété 1), ce théorème nous permet de relâcher la contrainte d'indépendance mutuelle entre la composante $s_{\sigma(1)}$ et les $P-1$ autres composantes au profit d'une indépendance deux à deux. Par conséquent, une identification séquentielle des $P-1$ paramètres $\beta_p^{(1)}$ du vecteur $\beta^{(1)}$, contrairement à l'identification globale faite par DEFA, est alors possible. En effet, chaque paramètre $\beta_p^{(1)}$ peut être obtenu en maximisant l'indépendance entre $s_{\sigma(1)}$ et l'une des $P-1$ autres composantes. Ainsi, le premier paramètre $\beta_{P-1}^{(1)}$ peut être identifié en maximisant la fonction de coût suivante :

$$\psi(\beta, \mathbf{s}^{(1,1)}) = (\mathcal{C}_{P,P,P,P,\mathbf{s}})^2 + (\mathcal{C}_{P-1,P-1,P-1,P-1,\mathbf{s}^{(1,1)}})^2 \quad (4)$$

où le vecteur $\mathbf{s}^{(1,1)}$ est défini par $\mathbf{s}^{(1,1)} = \mathbf{Q}_{\beta_{P-1}} \tilde{\mathbf{x}}$. Une fois $\beta_{P-1}^{(1)}$ identifié, on affecte le vecteur $\mathbf{Q}_{\beta_{P-1}} \tilde{\mathbf{x}}$ à $\mathbf{s}^{(1,p)}$. Le p -ième ($1 < p \leq P-1$) paramètre $\beta_{P-p}^{(1)}$ est quant-à-lui identifié en maximisant la fonction de coût suivante :

$$\psi(\beta, \mathbf{s}^{(1,p)}) = (\mathcal{C}_{P,P,P,P,\mathbf{s}})^2 + (\mathcal{C}_{P-p,P-p,P-p,P-p,\mathbf{s}^{(1,p)}})^2 \quad (5)$$

où le vecteur $\mathbf{s}^{(1,p)}$ est défini par $\mathbf{s}^{(1,p)} = \mathbf{Q}_{\beta_{P-p}} \mathbf{s}^{(1,p-1)}$. Une fois $\beta_{P-p}^{(1)}$ identifié, on affecte le vecteur $\mathbf{Q}_{\beta_{P-p}} \mathbf{s}^{(1,p-1)}$ à $\mathbf{s}^{(1,p)}$. Cette récurrence permet ainsi d'identifier les $P-1$ paramètres $\beta_p^{(1)}$ du vecteur $\beta^{(1)}$. Cependant, la p -ème étape de récurrence a pour inconvénient de modifier les résultats obtenus lors des $p-1$ précédentes étapes. Par conséquent, nous proposons, à l'image de la méthode de Jacobi [5], d'effectuer plusieurs balayages de la récurrence décrite précédemment. Plus précisément, une fois les $P-1$ paramètres $\beta_p^{(1)}$ estimés, il est conseillé d'affiner l'estimation de la matrice $\mathbf{Q}_{\beta^{(1)}}$ en effectuant la récurrence précédente une nouvelle fois. Le critère (4) est alors à nouveau maximisé en remplaçant $\tilde{\mathbf{x}}$ par le vecteur $\mathbf{Q}_{\beta^{(1)}} \tilde{\mathbf{x}}$. En pratique, si \mathbf{x} est de longueur N , environ \sqrt{N} balayages sont

suffisants pour converger vers la matrice $\mathbf{Q}_{\beta^{(1)}}$ optimale. Le vecteur $\tilde{\mathbf{A}}(:, \sigma(1))$ est alors obtenu à un scalaire près comme la transposée du dernier vecteur de ligne de $\mathbf{Q}_{\beta^{(1)}}$. Quant à la composante $s_{\sigma(1)}$, elle est donnée, à une variable déterministe près, par la dernière composante du vecteur $\mathbf{s}^{(1,p)}$ issu du dernier balayage.

3.2 Identification des autres facteurs

La seconde étape de l'algorithme SAUD consiste à identifier un deuxième vecteur colonne de $\tilde{\mathbf{A}}$, noté $\tilde{\mathbf{A}}(:, \sigma(2))$, ainsi que la composante $s_{\sigma(2)}$ associée. Pour cela, considérons la matrice $\mathbf{Q}'_{\beta^{(1)}}$ obtenue comme le bloc supérieur de taille $(P-1 \times P)$ de la matrice $\mathbf{Q}_{\beta^{(1)}}$. Cette matrice $\mathbf{Q}'_{\beta^{(1)}}$ peut être vue comme le projecteur sur l'espace vectoriel orthogonal à l'espace engendré par le vecteur $\tilde{\mathbf{A}}(:, \sigma(1))$. Ainsi multiplier le vecteur $\tilde{\mathbf{x}}$ à gauche par $\mathbf{Q}'_{\beta^{(1)}}$ équivaut à éliminer la contribution de la composante $s_{\sigma(1)}$ du vecteur $\tilde{\mathbf{x}}$. Nous noterons $\tilde{\mathbf{x}}^{(2)} = \mathbf{Q}'_{\beta^{(1)}} \tilde{\mathbf{x}}$ ce nouveau vecteur aléatoire ainsi obtenu. En utilisant le lemme 1, il est possible de modéliser le vecteur $\mathbf{q} = \tilde{\mathbf{A}}(:, \sigma(2))^\top \mathbf{Q}'_{\beta^{(1)}}^\top$ de \mathbb{R}^{P-1} , après normalisation par la quantité $\text{Sgn}\{q(P-1)\} \|\mathbf{q}\|$, comme la dernière ligne d'une matrice $\mathbf{Q}_{\beta^{(2)}}$ de taille $(P-1 \times P-1)$ où $\beta^{(2)}$ est un vecteur de paramètres de longueur $P-2$. La récurrence décrite dans la section 3.1 et appliquée initialement au vecteur $\tilde{\mathbf{x}}^{(2)}$ permet d'identifier la matrice $\mathbf{Q}_{\beta^{(2)}}$ et donc le vecteur $\mathbf{q} = \tilde{\mathbf{A}}(:, \sigma(2))^\top \mathbf{Q}'_{\beta^{(1)}}^\top$ à un scalaire près. Le second facteur $\tilde{\mathbf{A}}(:, \sigma(2))$ est alors donné par l'expression $\tilde{\mathbf{A}}(:, \sigma(2)) = \mathbf{Q}'_{\beta^{(1)}}^\top \mathbf{Q}_{\beta^{(2)}} (P-1, :)^T$. La composante $s_{\sigma(2)}$ est donnée, à une variable déterministe près, par $s_{\sigma(2)} = \mathbf{q} \tilde{\mathbf{x}}^{(2)}$.

La p -ième ($p < P$) étape de SAUD se déduit assez simplement de la deuxième étape décrite dans le paragraphe précédent. Considérons ainsi la matrice $\mathbf{Q}'_{\beta^{(p-1)}}$ obtenue comme le bloc supérieur de taille $(P-p \times P-p+1)$ de la matrice $\mathbf{Q}_{\beta^{(p-1)}}$. Nous cherchons à présent à identifier une matrice $\mathbf{Q}_{\beta^{(p)}}$ de taille $(P-p \times P-p)$ dont la dernière ligne est colinéaire au vecteur $\tilde{\mathbf{A}}(:, \sigma(p))^\top \mathbf{Q}'_{\beta^{(1)}}^\top \mathbf{Q}'_{\beta^{(2)}}^\top \dots \mathbf{Q}'_{\beta^{(p-1)}}^\top$. Pour cela il suffit d'effectuer la récurrence décrite dans la section 3.1 en l'appliquant initialement au vecteur $\tilde{\mathbf{x}}^{(p)} = \mathbf{Q}'_{\beta^{(p-1)}} \tilde{\mathbf{x}}^{(p-1)}$.

4 Simulations numériques

Nous avons choisi de tester et de comparer les méthodes COM2 [5], FastICA [8], DEFA [7], et SAUD au travers d'une simulation numérique illustrant une application de la séparation de sources au domaine biomédical. On distingue deux implantations de DEFA : l'une à l'aide du gradient stochastique et notée SG-DEFA, et l'autre notée SG-DEFA, dans laquelle l'algorithme du gradient stochastique est remplacé par une approche déterministe où l'espérance mathématique est estimée sur un bloc d'échantillons et où le pas fixe est remplacé par un pas adaptatif. L'application biomédicale ciblée concerne l'extraction d'activités électriques intracérébrales physiologiquement

indépendantes à partir d'observations issue d'enregistrements ElectroEncéphaloGraphiques (EEG) de scalp. Pour cela, nous avons positionné deux sources de courant dans le cerveau aux positions respectives $\rho_1 = [3, 3, 3]^T$ et $\rho_2 = [1, -3, -6]^T$ ainsi que cinq électrodes sur le scalp aux positions respectives $r_1 = [4.6, 4.6, 6.5]^T$, $r_2 = [4.6, -4.6, 6.5]^T$, $r_3 = [-6.2, 6.2, -2.8]^T$, $r_4 = [-6.2, -6.2, -2.8]^T$ et $r_5 = [-5.4, 0, 7.4]^T$. Nous avons choisi de manière aléatoire l'orientation des deux sources. Sous l'hypothèse d'un milieu de diffusion constitué de trois sphères concentriques de conductivités homogènes, nous avons alors été en mesure d'écrire analytiquement l'expression du mélange A en fonction des positions des sources, des positions des électrodes, des conductivités et des rayons des trois sphères [2, section II]. Ainsi ce mélange représente physiologiquement la diffusion des sources de courant intracérébrales vers la surface. En ce qui concerne la génération des activités électriques de profondeur, autrement dit les réalisations du vecteur $s = [s_1, s_2]^T$, nous avons utilisé le modèle de Jansen décrit dans [9]. Nous avons ainsi généré deux sources de courant présentant un certain nombre de "pointes" durant l'intervalle de temps observé. Le critère employé pour apprécier au mieux les résultats d'extraction de la source p ($1 \leq p \leq 2$) pour une méthode donnée est le rapport signal sur bruit plus interférence maximal associé à la source p , plus connu sous le dénominateur de SINRMp [1, section V]. Les résultats de simulation ont été moyennés sur 200 réalisations. La figure 1 représente le critère SINRM1 en sor-

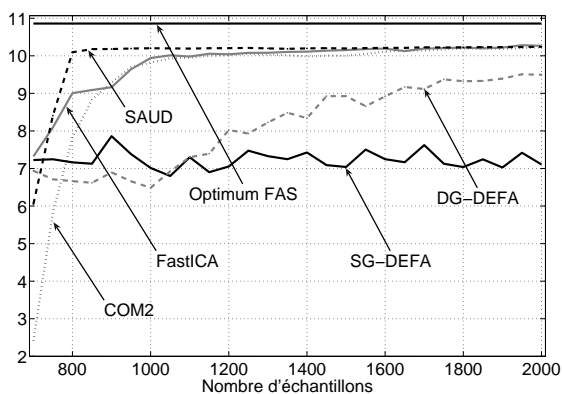


FIG. 1 – SINRM1 pour un SNR de 15 dB

tie des méthodes COM2, FastICA, SG-DEFA, DG-DEFA et SAUD en fonction du nombre d'échantillons pour un SNR de 15 dB. La méthode SAUD présente de bons résultats d'extraction de la première source (des résultats identiques sont obtenus pour la seconde source) et semble être la plus rapide en termes de vitesse de convergence. L'algorithme DG-DEFA semble nécessiter un plus grand nombre d'échantillons afin d'atteindre la qualité d'extraction de COM2, FastICA et SAUD. Quant à la méthode SG-DEFA, elle souffre à n'en pas douter des limitations de l'algorithme du gradient à pas fixe. En effet, si la valeur retenue pour le pas ($\mu = 10^{-2}$) convient pour un faible nombre d'échantillons cela ne semble pas être le cas lorsque le nombre d'échantillons est élevé.

5 Conclusion

On revisite l'algorithme DEFA [7] d'ICA par déflation, ce qui permet de remplacer l'algorithme du gradient à pas fixe par une solution semi-algébrique à chaque étape. La convergence de cette nouvelle méthode, baptisée SAUD, est de ce fait beaucoup plus rapide. En outre, SAUD permet d'établir un lien entre DEFA et la méthode COM2 [4]. En effet, en échangeant la boucle sur les balayages et celle sur les lignes dans la technique du balayage de Jacobi [4], on aboutit à l'algorithme de déflation SAUD.

Références

- [1] L. ALBERA, A. FERREOL, P. CHEVALIER, and P. COMON, "ICAR, a tool for blind source separation using fourth order statistics only," *IEEE Trans. On Sig. Proc.*, vol. 53, no. 10, pp. 3633–3643, Oct. 2005.
- [2] L. ALBERA, A. FERREOL, D. COSANDIER-RIMELE, I. MERLET, and F. WENDLING, "Brain source localization using a fourth order deflation scheme," *IEEE Trans. On Biomedical Engineering*, to appear.
- [3] J.-F. CARDOSO and A. SOULOUMIAC, "Blind beamforming for non-gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, December 1993.
- [4] P. COMON, "Analyse en composantes indépendantes et identification aveugle," *Trait. du Sig.*, vol. 7, June, 1990.
- [5] —, "Independent Component Analysis, a new concept?" *Sig. Proc., Elsevier*, vol. 36, no. 3, pp. 287–314, April 1994.
- [6] P. COMON and C. JUTTEN, *Séparation de Sources : Concepts de base et Analyse en Composantes Indépendantes*, ser. Collection IC2. Hermès, 2007, vol. 1.
- [7] N. DELFOSSE and P. LOUBATON, "Adaptive blind separation of independent sources : a deflation approach," *Sig. Proc., Elsevier*, vol. 45, pp. 59–83, 1995.
- [8] A. HYVARINEN, J. KARHUNEN, and P. OJA, *Independent component analysis*, ser. Wiley interscience. Simon Haykin, 2001.
- [9] B. JANSEN and V. RIT, "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns?" *Biological Cybernetics*, vol. 73, no. 4, pp. 357–366, September 1995.
- [10] C. JUTTEN and J. HERAULT, "Une solution neuromimétique au problème de séparation de sources," *Traitement du Signal*, vol. 5, no. 6, pp. 389–403, 1988.
- [11] T. W. LEE, M. GIROLAMI, and T. J. SEJNOWSKI, "Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources," *Neural Computation*, vol. 11, no. 2, pp. 417–441, 1999.
- [12] F. POREE, A. KACHENOURA, H. GAUVRIT, C. MORVAN, G. CARRAULT, and L. SENHADJI, "Blind source separation for ambulatory sleep recording," *IEEE Trans. on Information Technology in Biomedicine*, vol. 10, no. 2, pp. 293–301, 2006.
- [13] P. A. REGALIA, "An adaptive unit norm filter with applications to signal analysis and karhunen-loeve transformations," *IEEE Trans. on Circuits and Systems*, vol. 37, no. 5, pp. 646–649, May 1990.
- [14] V. ZARZOZO, P. COMON, and M. KALLEL, "How fast is FastICA," in *EUSIPCO 06*, Florence, Italy, September 4-8 2006.