

Récepteur aveugle basé sur la décomposition PARAFAC pour des canaux de communication non-linéaires

A.Y. KIBANGOU¹, G. FAVIER¹, M.M. HASSANI²

¹Laboratoire I3S/CNRS/UNSA
2000 Route des lucioles, B.P. 121, 06903 Sophia Antipolis Cedex, France

²Laboratoire d'Electronique et Instrumentation
B.P. 2390, 40000 Marrakech, Maroc
kibangou@i3s.unice.fr, favier@i3s.unice.fr
hassani@ucam.ac.ma

Résumé – Dans cet article, nous proposons une nouvelle approche d'égalisation aveugle basée sur une modélisation tensorielle d'un système de communication non-linéaire SIMO du type Wiener. Le tenseur cubique formé par les signaux reçus s'exprime comme une décomposition du type PARAFAC avec comme dimensions le temps, l'espace et la redondance introduite par un précodage. Les conditions d'unicité de cette décomposition sont établies et une solution d'égalisation aveugle sans ambiguïté est proposée.

Abstract – In this paper, we present a new approach for blind equalization of Wiener SIMO nonlinear channels. The received signals form a cubic tensor the dimensions of which are time, space and redundancy induced by a linear precoder. The equalization process is based on the PARAFAC decomposition of this three-way array. The uniqueness conditions are provided, then a solution for blind equalization without ambiguity is derived.

1 Introduction

Depuis la fin des années 90, quelques solutions ont été apportées au problème de l'égalisation aveugle de canaux de communication non-linéaires. Outre les méthodes basées sur l'utilisation de réseaux neuronaux, trois principales approches ont été considérées. Elles étendent aux systèmes non-linéaires des résultats établis dans le cas linéaire. La première est déterministe [1]. Les auteurs y montrent que, sous certaines conditions, des égaliseurs linéaires de type FIR peuvent réaliser une égalisation du type forçage à zéro de canaux SIMO non-linéaires. La méthode proposée dans ce travail pionnier présente cependant un certain nombre de limitations liées au rang et à la structure de la matrice associée au canal, d'une part, et à la mémoire du noyau de Volterra associé au canal, d'autre part. Certaines de ces restrictions ont été récemment relaxées dans la seconde approche basée sur l'exploitation des statistiques du second ordre [2]. Toutefois, comme pour la plupart des algorithmes d'égalisation basés sur l'exploitation des statistiques du second ordre, cette approche requiert une complexité de calcul importante. Enfin, la troisième approche [3] est aussi basée sur les statistiques du second ordre, mais requiert que chaque sous-canal soit linéarisé par un filtre de Volterra d'ordre et de mémoire connus. Cette condition est très restrictive. Dans les trois principales approches sus-citées, les systèmes de communication considérés sont de type SIMO, cette structure étant obtenue soit en utilisant un réseau d'antennes soit en sur-échantillonnant la sortie du système. Il est bien connu que dans ce cas, afin de pouvoir déterminer un égaliseur, les différents sous-canaux ne doivent avoir aucun facteur commun.

Dans cet article, nous proposons une nouvelle approche d'éga-

lisation aveugle basée sur une modélisation tensorielle d'un système de communication SIMO du type Wiener, dont les différents sous-canaux ont un facteur commun.

Depuis le début des années 2000, une attention particulière est accordée à l'utilisation des techniques de décomposition tensorielle dans les télécommunications. Ainsi, différents algorithmes d'estimation aveugle, de canal et/ou du signal, basés sur une décomposition tensorielle du type PARAFAC sont proposés dans la littérature [4, 5]. Ceux-ci se restreignent au cas de systèmes linéaires. A notre connaissance, l'approche que nous proposons est une première application de la décomposition PARAFAC pour des fins d'égalisation de canaux non-linéaires.

Dans la section 2, nous traitons de la modélisation tensorielle du système de communication considéré, des conditions d'unicité de la décomposition PARAFAC ainsi que de l'estimation de ses différents facteurs. La méthode est ensuite illustrée par un exemple dans la section 3, avant de conclure.

2 Modélisation et estimation

Nous considérons un système de communication non-linéaire, à temps discret, d'entrée x et de sorties y_k , $k = 0, 1, \dots, K_a - 1$, où K_a désigne le nombre de capteurs utilisés au niveau du récepteur. Le signal émis par une station mobile, une fois reçu au niveau d'un des capteurs, est tout d'abord converti en fréquences optiques, au niveau d'un point d'accès radio (RAP, *Radio Access Point*), puis transmis via une fibre optique à une station de base où s'effectue le traitement des données en bande de base (Figure 1).

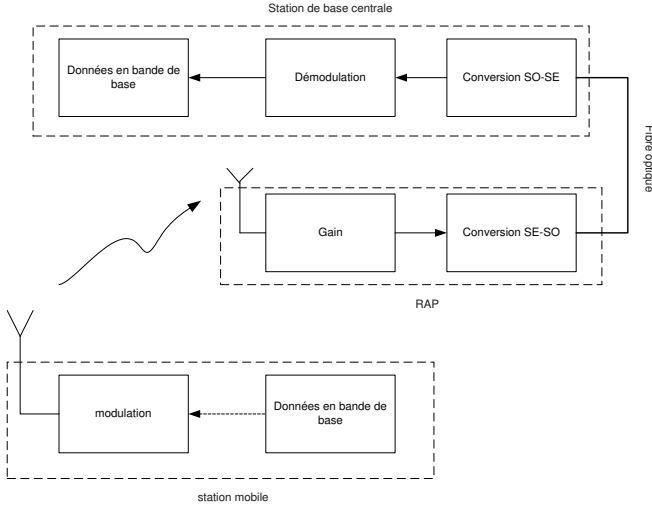


FIG. 1 – Schéma fonctionnel du système.

Lorsque la longueur de la fibre est relativement courte (quelques kilomètres) et la fréquence radio de quelques GHz, les effets de dispersion de la fibre et le bruit inhérent au laser sont négligeables [6]. Dans ce cas, la distorsion non-linéaire introduite par la conversion du signal électrique (SE) en un signal optique (SO) devient prépondérante [7]. Sachant que plusieurs capteurs sont utilisés, le système peut alors être vu comme un système SIMO du type Wiener, chaque sous canal k étant constitué par la mise en cascade d'un canal linéaire h_k et d'une non-linéarité statique représentée par un polynôme P_k de degré P et de coefficients $c_{p,k}$, $p = 1, \dots, P$. Sous l'hypothèse que la source est à bande étroite par rapport aux antennes de réception, la réponse impulsionnelle du canal de transmission h_k est donnée par [5] : $h_k(i) = \alpha_k h(i)$, où α_k est le gain complexe de l'antenne considérée. Le modèle de Volterra décrivant ce système s'écrit alors :

$$y_k(nM - i) = \sum_{p=1}^P \sum_{i_1, \dots, i_p=0}^{M-1} h_{p,k}(i_1, \dots, i_p) \times \prod_{j=1}^p x(nM - i - i_j) + w_k(nM - i), \quad (1)$$

$$i = 0, \dots, M - 1,$$

où $n = 0, 1, \dots, N - 1$ désigne le numéro du bloc de symboles transmis, chaque bloc comportant M symboles, et :

$$h_{p,k}(i_1, \dots, i_p) = \beta_{k,p} \prod_{j=1}^p h(i_j), \quad \beta_{k,p} = c_{p,k} \alpha_k^p. \quad (2)$$

En supposant que la séquence transmise $x(\cdot)$ soit obtenue à partir de la séquence informative via un précodage simple tel que pour $i = 0, \dots, M - 1$, le symbole $x(nM - i)$ ne soit non nul et égal à $u(n)$ que lorsque $i = M - 1$ [8], la relation (1) devient :

$$y_k(nM - i) = \sum_{p=1}^P h_{p,k}(M - 1 - i, \dots, M - 1 - i) u^p(n) + w_k(nM - i), \quad i = 0, \dots, M - 1. \quad (3)$$

Ce précodage correspond à une insertion de $(M - 1)$ symboles nuls entre chaque symbole informatif. Les paramètres du mo-

dèle de Volterra (1) se réduisent donc aux seuls coefficients diagonaux. Notons que l'information contenue dans les signaux reçus $y_k(nM - i)$, $i = 0, \dots, M - 1$, est redondante car ces signaux dépendent de la même entrée $u(n)$.

En notant :

$$h_{p,k,i} = h_{p,k}(M - 1 - i, \dots, M - 1 - i),$$

$$y_{k,n,i} = y_k(nM - i),$$

$$w_{k,n,i} = w_k(nM - i),$$

pour $k = 0, 1, \dots, K_a - 1$, $n = 0, 1, \dots, N - 1$ et $i = 0, 1, \dots, M - 1$, où N représente le nombre de blocs transmis, nous avons :

$$y_{k,n,i} = \sum_{p=1}^P \beta_{k,p} h^p(M - 1 - i) u^p(n) + w_{k,n,i}. \quad (4)$$

Nous pouvons reconnaître en (4) l'écriture scalaire de la décomposition PARAFAC du tenseur cubique \mathbb{Y} , d'éléments $y_{k,n,i}$, le nombre de facteurs de cette décomposition étant égal au degré P du polynôme de la non-linéarité statique. Dans le cas sans bruit, en fixant respectivement une des dimensions du tenseur, l'écriture matricielle correspondante est :

$$\mathbf{Y}_{k..} = \mathbf{H} D_k(\mathbf{B}) \mathbf{U}, \quad k = 0, \dots, K_a - 1, \quad (5)$$

$$\mathbf{Y}_{..n} = \mathbf{B} D_n(\mathbf{U}^T) \mathbf{H}^T, \quad n = 0, \dots, N - 1, \quad (6)$$

$$\mathbf{Y}_{..i} = \mathbf{U}^T D_i(\mathbf{H}) \mathbf{B}^T, \quad i = 0, \dots, M - 1, \quad (7)$$

où $D_k(\cdot)$ est l'opérateur qui forme une matrice diagonale avec la $(k + 1)$ -ème ligne de la matrice en argument, $\mathbf{Y}_{k..}$, $\mathbf{Y}_{..n}$ et $\mathbf{Y}_{..i}$ sont de dimensions respectives $(M \times N)$, $(K_a \times M)$ et $(N \times K_a)$, \mathbf{H} et \mathbf{U} sont des matrices de Vandermonde de dimensions respectives $(M \times P)$ et $(P \times N)$, et \mathbf{B} la matrice de dimensions $(K_a \times P)$, telles que :

$$\mathbf{H} = \begin{pmatrix} h(M-1) & h^2(M-1) & \dots & h^P(M-1) \\ h(M-2) & h^2(M-2) & \dots & h^P(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ h(0) & h^2(0) & \dots & h^P(0) \end{pmatrix}$$

$$\mathbf{U} = \begin{pmatrix} u(0) & u(1) & \dots & u(N-1) \\ u^2(0) & u^2(1) & \dots & u^2(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ u^P(0) & u^P(1) & \dots & u^P(N-1) \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} \beta_{0,1} & \beta_{0,2} & \dots & \beta_{0,P} \\ \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,P} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{K_a-1,1} & \beta_{K_a-1,2} & \dots & \beta_{K_a-1,P} \end{pmatrix}$$

En tenant compte des conditions d'unicité de cette décomposition [9] et de la structure de Vandermonde des matrices \mathbf{H} et \mathbf{U} , nous déduisons le théorème suivant :

Théorème : Soit $k_{\mathbf{B}}$ le rang de Kruskal de \mathbf{B} , i.e. le nombre maximal de colonnes de \mathbf{B} tel que tout ensemble de $k_{\mathbf{B}}$ colonnes de \mathbf{B} soient linéairement indépendantes. Si le signal u prend au moins $P \leq N$ valeurs distinctes non nulles et si le canal est tel qu'au moins $P \leq M$ de ses coefficients $h(i)$ soient distincts, alors la décomposition PARAFAC est unique, à une matrice de permutation et une matrice diagonale près, si $k_{\mathbf{B}} \geq 2$.

Preuve : Il est établi que l'unicité de la décomposition PARAFAC est garantie lorsque $k_{\mathbf{B}} + k_{\mathbf{H}} + k_{\mathbf{U}^T} \geq 2P + 2$ [9, 10]. En tenant compte de la structure Vandermonde, le fait qu'au moins P coefficients $h(i)$ de la matrice \mathbf{H} soient distincts, $P \leq M$, garantit qu'elle soit de rang colonne complet. Par conséquent, on a : $k_{\mathbf{H}} = P$. De même, si le signal u prend au moins $P \leq N$ valeurs distinctes non nulles, cela implique que $k_{\mathbf{U}^T} = P$. Par suite, la condition d'unicité se ramène à : $k_{\mathbf{B}} \geq 2$.

On peut noter que suivant la construction de la matrice \mathbf{B} , il est nécessaire d'utiliser au moins deux capteurs différents parmi les K_a capteurs pour réaliser la condition $k_{\mathbf{B}} \geq 2$.

L'unicité démontrée dans le théorème est à une matrice de permutation et un facteur d'échelle près. En d'autres termes, tout autre ensemble de matrices $\tilde{\mathbf{H}}$, $\tilde{\mathbf{B}}$ et $\tilde{\mathbf{U}}^T$ pouvant reconstituer les matrices $\mathbf{Y}_{k..}$, $k = 0, \dots, K_a - 1$ est relié à l'ensemble de matrices \mathbf{H} , \mathbf{B} et \mathbf{U}^T par :

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{\Pi}\mathbf{\Delta}_1, \quad \tilde{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Delta}_2, \quad \tilde{\mathbf{U}}^T = \mathbf{U}^T\mathbf{\Pi}\mathbf{\Delta}_3,$$

où $\mathbf{\Pi}$ est une matrice de permutation et les $\mathbf{\Delta}_i$, $i = 1, 2, 3$ sont des matrices diagonales telles que

$$\mathbf{\Delta}_1\mathbf{\Delta}_2\mathbf{\Delta}_3 = \mathbf{I}.$$

Il est à noter que P étant fini, l'ensemble des matrices de permutation est aussi fini. Cependant, toute permutation sur les lignes de \mathbf{U} rend ambiguë l'estimation de la séquence informative émise ; d'où la nécessité de trouver un moyen pour lever cette ambiguïté. En tenant compte de la structure de la matrice \mathbf{U} , nous définissons le critère suivant :

$$\rho_p = \frac{\min |\mathbf{U}_{p.}|}{\max |\mathbf{U}_{p.}|}, \quad (8)$$

où $\max |\mathbf{U}_{p.}|$ et $\min |\mathbf{U}_{p.}|$ désignent respectivement le plus grand et le plus petit des modules des éléments de la p -ème ligne de \mathbf{U} . Ce critère décroît strictement avec l'ordre de la non-linéarité et est invariant par rapport à tout facteur d'échelle affectant une ligne donnée. En conséquence, une fois la matrice \mathbf{U} estimée, le calcul de ρ_p correspondant à chacune des lignes permet de définir la matrice de permutation à appliquer à \mathbf{U} de sorte que ses lignes soient dans un ordre qui assure la décroissance de ρ_p . Par suite, la décomposition PARAFAC est unique à une matrice diagonale près. En d'autres termes, tout autre ensemble de matrices $\tilde{\mathbf{H}}$, $\tilde{\mathbf{B}}$ et $\tilde{\mathbf{U}}^T$ pouvant reconstituer les matrices $\mathbf{Y}_{k..}$, $k = 0, \dots, K_a - 1$ est relié à l'ensemble de matrices \mathbf{H} , \mathbf{B} et \mathbf{U}^T par :

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{\Delta}_1, \quad \tilde{\mathbf{B}} = \mathbf{B}\mathbf{\Delta}_2, \quad \tilde{\mathbf{U}}^T = \mathbf{U}^T\mathbf{\Delta}_3,$$

avec $\mathbf{\Delta}_1\mathbf{\Delta}_2\mathbf{\Delta}_3 = \mathbf{I}$. En utilisant un correcteur de gain, toute ambiguïté sur l'estimation de \mathbf{U} peut être complètement levée [4], permettant ainsi l'extraction de la séquence informative formée par la première ligne de \mathbf{U} .

L'estimation des matrices \mathbf{H} , \mathbf{U} et \mathbf{B} peut être effectuée en utilisant l'algorithme des moindres carrés alternés (*Alternating Least Squares (ALS)*) dont le principe est d'estimer, au sens des moindres carrés, un sous ensemble de paramètres en utilisant les estimés antérieurs des autres sous ensembles de paramètres, puis de réitérer le processus pour un autre sous ensemble de paramètres jusqu'à la convergence de l'algorithme. Les matrices $\mathbf{Y}_{k..}$ étant données, leur concaténation permet d'écrire :

$$\begin{pmatrix} \mathbf{Y}_{0..} \\ \vdots \\ \mathbf{Y}_{K_a-1..} \end{pmatrix} = \begin{pmatrix} \mathbf{H}D_0(\mathbf{B}) \\ \vdots \\ \mathbf{H}D_{K_a-1}(\mathbf{B}) \end{pmatrix} \mathbf{U}. \quad (9)$$

En notant $\tilde{\mathbf{Y}}_{k..}$ la version bruitée de $\mathbf{Y}_{k..}$, et en considérant les estimés précédents de \mathbf{H} et de \mathbf{B} , notés $\hat{\mathbf{H}}$ et $\hat{\mathbf{B}}$, le problème des moindres carrés est alors formulé comme suit :

$$\min_{\mathbf{U}} \left\| \begin{pmatrix} \tilde{\mathbf{Y}}_{0..} \\ \vdots \\ \tilde{\mathbf{Y}}_{K_a-1..} \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{H}}D_0(\hat{\mathbf{B}}) \\ \vdots \\ \hat{\mathbf{H}}D_{K_a-1}(\hat{\mathbf{B}}) \end{pmatrix} \mathbf{U} \right\|_F^2, \quad (10)$$

où $\|\cdot\|_F$ est la norme de Frobenius. La solution à ce problème est donnée par :

$$\hat{\mathbf{U}} = \begin{pmatrix} \hat{\mathbf{H}}D_0(\hat{\mathbf{B}}) \\ \vdots \\ \hat{\mathbf{H}}D_{K_a-1}(\hat{\mathbf{B}}) \end{pmatrix}^\dagger \begin{pmatrix} \tilde{\mathbf{Y}}_{0..} \\ \vdots \\ \tilde{\mathbf{Y}}_{K_a-1..} \end{pmatrix}, \quad (11)$$

où \dagger indique la pseudo-inverse.

De même, les estimés au sens des moindres carrés des matrices \mathbf{B} et \mathbf{H} sont donnés par :

$$\hat{\mathbf{H}}^T = \begin{pmatrix} \hat{\mathbf{B}}D_0(\hat{\mathbf{U}}^T) \\ \vdots \\ \hat{\mathbf{B}}D_{N-1}(\hat{\mathbf{U}}^T) \end{pmatrix}^\dagger \begin{pmatrix} \tilde{\mathbf{Y}}_{.0} \\ \vdots \\ \tilde{\mathbf{Y}}_{.N-1} \end{pmatrix}, \quad (12)$$

$$\hat{\mathbf{B}}^T = \begin{pmatrix} \hat{\mathbf{U}}^T D_0(\hat{\mathbf{H}}) \\ \vdots \\ \hat{\mathbf{U}}^T D_{M-1}(\hat{\mathbf{H}}) \end{pmatrix}^\dagger \begin{pmatrix} \tilde{\mathbf{Y}}_{.0} \\ \vdots \\ \tilde{\mathbf{Y}}_{..M-1} \end{pmatrix}. \quad (13)$$

3 Résultats de simulation

Considérons le système caractérisé par une non-linéarité représentée par le polynôme P de degré trois tel que

$$P(u) = u + 0.3u^2 + 0.1u^3$$

et une partie linéaire définie par $h_k(i) = \alpha_k h(i)$ avec

$$h = (1 \quad 0.5 \quad -0.2)^T.$$

Pour différentes valeurs de α_k , i.e. différents capteurs, et $N = 500$ blocs de symboles du type PAM-4, la figure 2 montre le taux d'erreur par symbole Pe en fonction du rapport signal à bruit (RSB), pour trois valeurs différentes de K_a . L'évaluation des performances a été réalisée à l'aide d'une simulation du type Monte-Carlo avec 100 expériences indépendantes. Nous pouvons noter que le taux d'erreur décroît lorsque le nombre de capteurs croît. Dans un environnement fortement bruité, la diversité introduite par les capteurs n'améliore que faiblement le taux d'erreur. Cette amélioration devient significative lorsque le niveau de bruit diminue.

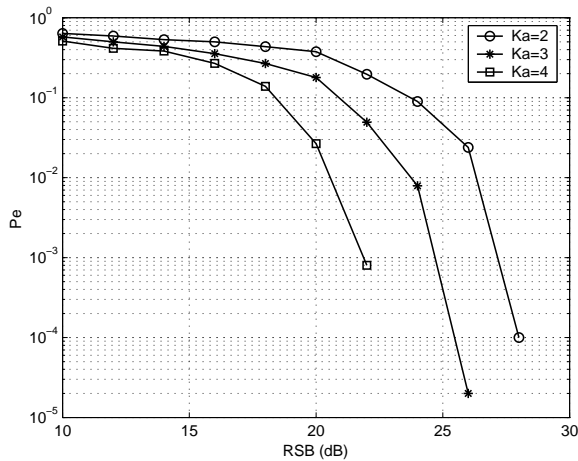


FIG. 2 – Taux d’erreur d’estimation des symboles en fonction du RSB ($N=500$).

Nous comparons ensuite la méthode d’égalisation proposée avec un égaliseur non aveugle, i.e. en supposant le canal connu, du type forçage à zéro de l’erreur d’égalisation. En utilisant trois antennes de réception, nous pouvons noter que les performances sont relativement proches (Figure 3).

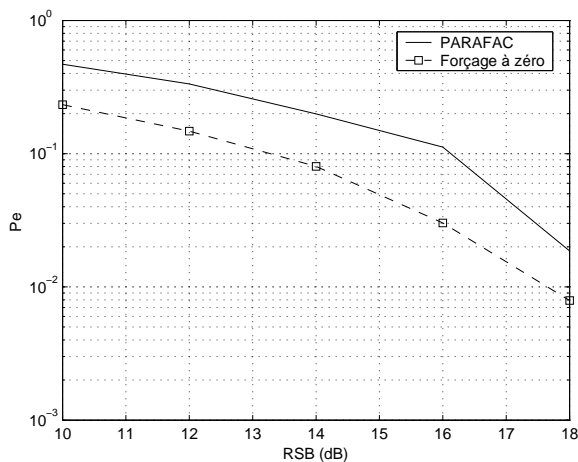


FIG. 3 – Taux d’erreur d’estimation des symboles en fonction du RSB ($N=500$), $K_a = 3$.

4 Conclusion

Dans cet article, nous avons montré qu’en utilisant un pré-codage simple, un système non-linéaire SIMO du type Wiener peut être modélisé comme un tenseur cubique dont les dimensions sont le temps, l’espace et la diversité liée à la redondance introduite par le pré-codage. Le tenseur formé par les signaux reçus s’exprime naturellement comme une décomposition du type PARAFAC. Les conditions garantissant l’unicité de cette décomposition, à des facteurs d’échelle près, ont été établies. La séquence d’information transmise peut ainsi être estimée sans ambiguïté. L’écueil majeur de l’approche proposée est la réduction du taux d’information transmise. Une perspective de ce travail vise à rechercher un pré-codage plus efficace permet-

tant un taux de transmission plus élevé tout en assurant une décomposition PARAFAC du tenseur des signaux reçus.

Références

- [1] G. B. Giannakis and E. Serpedin. Linear multichannel blind equalizers of nonlinear FIR Volterra channels. *IEEE Trans. Signal Processing*, 45(1) :67–81, January 1997.
- [2] R. Lopez-Valcarce and S. Dasgupta. Blind equalization of nonlinear channels from second-order statistics. *IEEE Trans. Signal Processing*, 49(12) :3084–3097, 2001.
- [3] G. V. Raz and B. Van Veen. Blind equalization and identification of nonlinear and IIR systems- A least squares approach. *IEEE trans. Signal Processing*, 48(1) :192–200, January 2000.
- [4] N.D. Sidiropoulos, G.B. Giannakis, and R. Bro. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Trans. Signal Processing*, 48(3) :810–823, March 2000.
- [5] A. De Baynast and L. De Lathauwer. Détection autodidacte pour des systèmes à accès multiples basée sur l’analyse PARAFAC. In *Proc. of GRETSI’03*, Paris, September 2003.
- [6] W.I. Way. Optical fiber based microcellular systems. An overview. *IEICE Trans. Commun.*, E76-B(9) :1091–1102, September 1993.
- [7] X.N. Fernando and A. B. Sesay. Higher order adaptive filter based predistortion for nonlinear distortion compensation of radio over fiber links. In *Proc. of the Intern. Conf. on Communications (ICC’ 2000)*, vol. 1/3, pages 367–371, New-Orleans, LA, USA, June 2000.
- [8] R. Lopez-Valcarce, X. Song, and S. Dasgupta. Blind equalization of nonlinear channels from second-order statistics using precoding and channel diversity. In *Proc. ICASSP*, vol. 4, pages 2073–2076, Salt-Lake city, UT, USA, 2001.
- [9] J.B. Kruskal. Three-way arrays : rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Applicat.*, 18 :95–138, 1977.
- [10] T. Jiang and N. Sidiropoulos. Kruskal permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints. *IEEE Trans. on Signal Processing*, 52(9) :2625–2636, September 2004.