Dix-septième colloque GRETSI, Vannes, 13-17 septembre 1999

Sur la résolution de l'identité de Bezout pour l'égalisation autodidacte de systèmes mono-entrée–multi-sorties

Mamadou MBOUP

UFR de Mathématiques et Informatique - CRIP5 – Université René Descartes-Paris V
45, rue des Saints-Pères - 75270 Paris Cedex 06, France
mboup@math-info.univ-paris5.fr

Résumé – Nous présentons une approche polynomiale pour le problème de l'égalisation autodidacte de système multicanal, à réponse impulsionnelle finie. Cette approche utilise une identification entrée-sortie de système, que l'on peut interpréter comme une prédiction linéaire spatio-temporelle puis, une résolution directe de l'identité de Bezout, sous-jacent au problème d'égalisation. En l'abscence de bruit, on montre que le nombre de coefficients non nuls de la réponse combinée canal-égaliseur n'excède pas $N - N' + 1$ où $NetN'$ sont les longueurs du canal et de l'égaliseur. Dans le cas surmodélisé, on montre que cette réponse s'annulle.

Abstract – We propose a polynomial approach to the blind single-input multiple-outputs direct equalization problem. The approach uses an input-output system identification, which may be interpreted as a spatial linear prediction and then, solves directly the Bezout identity which underlies the equalization problem. In the noise-free case, we show that the number of the channel-equalizer combined response nonzero termes does not exceed $N - N' + 1$, where N and N' denote the channel and the equalizer length, respectively. For the overmodeled case, we show that this response vanishes.

1 Introduction

Soit $\{s(n)\}$ une séquence inconnue de symboles émise par une source de communication numérique, à travers un canal de transmission. On note $y_1(n), \ldots, y_L(n)$ les signaux bruités reçus sur un réseau de L capteurs, suivant le modèle

$$\underbrace{\begin{bmatrix} y_1(n) \\ \vdots \\ y_L(n) \end{bmatrix}}_{Y(n)} = \underbrace{\begin{bmatrix} h_1(z) \\ \vdots \\ h_L(z) \end{bmatrix}}_{H(z)} s(n) + \underbrace{\begin{bmatrix} w_1(n) \\ \vdots \\ w_L(n) \end{bmatrix}}_{W(n)}, \qquad (1)$$

où $H(z)$ représente la fonction de transfert du canal, et où $W(n) \in \mathbb{C}^L$ est un bruit additif. Il faut noter que le réseaux de $L > 1$ capteurs n'a pas toujours une réalité physique ; il peut être obtenu de façon virtuelle à partir d'un seul capteur, par un suréchantillonnage de la sortie (voir [1]).

Nous nous intéressons au problème de l'égalisation aveugle d'un tel canal, *i.e.*, la reconstitution du signal émis $s(n)$, à un retard et un facteur d'échelle près, uniquement à partir des observations $Y(n)$. Il s'agit alors de résoudre le

Problème 1. *Etant donnée $\{Y(n)\}$ une séquence d'observations obtenues selon le modèle (1),*
trouver $G(z) = [g_1(z) \ \cdots \ g_L(z)]$ tel que

$$G(z)H(z) = g_1(z)h_1(z) + \cdots + g_L(z)h_L(z) = \alpha z^\nu, \qquad (2)$$

où α est une constante (complexe) et ν, un retard entier.

à toute solution $A(z)$ du problème 1, correspond, en l'abscence de bruit, une sortie

$$\tilde{s}(n) = G(z)Y(n) = G(z)H(z)s(n) = \alpha s(n - \nu)$$

parfaitement égalisée à un retard ν et un facteur d'échelle α près. Notons que ici, on interpréte la variable z (et non z^{-1}) comme l'opérateur retard : $zs(n) = s(n-1)$.

Lorsque le canal a une durée N finie, ce que nous supposons dans la suite, sa fonction de transfert $H(z)$ prend la forme d'un polynome de degré N, à coefficients dans $\mathbb{C}^L$. Pour chaque sous-canal i, $h_i(z)$ est alors un polynome de degré n'excédent pas N. Dans ces conditions, l'équation (2) est une équation de Bezout qui, comme chacun sait, admet une solution si les polynomes $h_i(z)$ $i = 1, L$ sont irréductibles.

Pour résoudre le problème 1 d'égalisation aveugle, une première approche consiste à estimer d'abord le canal $H(z)$ puis à résoudre, généralement de façon implicite, l'identité de Bezout (2). Les méthodes de type sous-espace [1], les méthodes moindre carré [2] et celles utilisant des propriétés de cyclo-stationarité de la sortie privilégient cette approche (voir [3] et les références qu'il contient).

Une deuxième approche, correspondant à la famille des algorithmes de type Module Constante multicanal, tente, par la minimisation d'une fonction coût, de résoudre implicitement l'identité de Bezout, sans passer par une estimation préalable du canal (voir [4]).

De nombreuses autres approches sont présentées dans la littérature (voir [3]).

Dans cet article, nous proposons une approche algébrique, reposant sur une résolution explicite du problème 1.

2 Une approche polynomiale

Cette section décrit, à partir d'une identité de Bezout, les différentes étapes de la construction d'une solution au problème 1.

2.1 Préliminaires

Soient $p_1(z),\ \cdots,\ p_l(z),\ q_1(z),\ \cdots,\ q_m(z)$, $l+m$ polynomes dans l'anneau $\mathbb{C}[z]$. L'ensemble J défini par

$$J = \left\{ \sum_{i=1}^{l} a_i(z)p_i(z) + \sum_{j=1}^{m} b_j(z)q_j(z);\ a_i(z), b_j(z) \in \mathbb{C}[z] \right\}$$

est un idéal principal de $\mathbb{C}[z]$. Dans le cas où les générateurs $p_i(z), i = 1, l$ et $q_j(z), j = 1, m$ sont premiers entre eux, l'idéal J contient le polynome identité et par conséquent, il coïncide avec l'anneau $\mathbb{C}[z]$ tout entier. Ces faits élémentaires sont connus depuis Bezout.

Représentons les générateurs $p_i(z), i = 1, l$ et $q_j(z), j = 1, m$ par les polynomes à valeurs vectorielles, $P(z) \in \mathbb{C}^l[z]$ et $Q(z) \in \mathbb{C}^m[z]$, définis par

$$Q(z) \triangleq \begin{bmatrix} q_1(z) \\ \vdots \\ q_m(z) \end{bmatrix} \text{ et } P(z) \triangleq \begin{bmatrix} p_1(z) \\ \vdots \\ p_l(z) \end{bmatrix} \tag{3}$$

et supposons que $Q(z)$ soit irréductible (les polynomes $q_j(z), j = 1, m$ sont co-premiers). Il existe alors un polynome $Q^\sharp(z) \in \mathbb{C}^{1\times m}[z]$ tel que l'identité de Bezout

$$Q^\sharp(z)Q(z) = 1, \tag{4}$$

soit satisfaite. En vertu de l'irréductibilité de $Q(z)$, le polynome $\begin{bmatrix} P(z) \\ \cdots \\ Q(z) \end{bmatrix} \in \mathbb{C}^{l+m}[z]$ est aussi irréductible et par conséquent, l'équation de Bezout

$$\underbrace{[a_1(z)\ \cdots\ a_l(z)]}_{A(z)} \underbrace{\begin{bmatrix} p_1(z) \\ \vdots \\ p_l(z) \end{bmatrix}}_{P(z)} + \underbrace{[b_1(z)\ \cdots\ b_m(z)]}_{B(z)} \underbrace{\begin{bmatrix} q_1(z) \\ \vdots \\ q_m(z) \end{bmatrix}}_{Q(z)} = \alpha z^\nu, \tag{5}$$

où α et ν sont deux constantes, respectivement complexe et entière, admet au moins une paire $A(z), B(z)$, solution.

Pour la suite, on pose $\deg Q(z) = \kappa = \deg P(z)$ et on considère une solution $Q^\sharp(z)$ de l'équation (4) telle que $\deg Q^\sharp(z) \le \kappa - 1$.

En utilisant l'identité (4), l'équation de Bezout (5) devient

$$\left\{A(z)P(z)Q^\sharp(z) + B(z)\right\} Q(z) = \left\{\alpha z^\nu Q^\sharp(z)\right\} Q(z). \tag{6}$$

On observe maintenant que tout paire $\{A(z), B(z)\}$ satisfaisant à l'égalité des termes entre accolades dans l'éq.(6), *i.e.*,

$$A(z)P(z)Q^\sharp(z) + B(z) = \alpha z^\nu Q^\sharp(z) \tag{7}$$

doit aussi satisfaire l'équation de Bezout (5). Cette remarque est, avec l'équation (7), la clé de la méthode que nous présentons.

Rappelons que $P(z)$ et l'inverse $Q^\sharp(z)$ de $Q(z)$ sont des polynomes à coefficients respectivement dans $\mathbb{C}^l$ et $\mathbb{C}^{1\times m}$, avec $\deg P(z) = \kappa = 1 + \deg Q^\sharp(z)$. La fonction *rationnelle*, $P(z)Q^\sharp(z)$, à coefficients matriciel-$(l \times m)$ est alors un polynome de la forme

$$F(z) \triangleq P(z)Q^\sharp(z) = F_0 + F_1 z + \cdots + F_{2\kappa-1}z^{2\kappa-1}. \tag{8}$$

2.2 Une solution du problème 1.

Nous allons montré comment il est possible de construire une solution de (5) uniquement à partir des coefficients du polynome $F(z)$. Lorsque ces coefficients sont des matrices carrées, $l = m$, on peut démontrer le

Théorème 1. *Soit $P(z)$ et $Q(z)$ deux polynomes, à coefficients dans $\mathbb{C}^m$, premiers entre eux, et tels que $Q(z)$ est irréductible,* $\deg Q(z) = \kappa = \deg P(z)$ *et $P_0 = P(0) \neq \mathbf{0}_m$. Soit $F(z) = P(z)Q^\#(z) = \sum_{i\ge 0} F_i z^i$, avec $F_i \in \mathbb{C}^{m\times m}$ et où le polynome $Q^\#(z)$, de degré $M \le \kappa - 1$ et à coefficients dans $\mathbb{C}^{1\times m}$, vérifie $Q^\#(z)Q(z) = 1$.*

Alors, pour $M = \kappa - 1$ et $\nu = 2\kappa - 1$, les $\mathbb{C}^{1\times m}$-polynomes, $A(z) = \sum_{k=0}^{M} A_k z^k$ et $B(z) = \sum_{k=0}^{M} B_k z^k$, obtenus à partir du système d'équations

$$\underbrace{\begin{bmatrix} F_1^t & F_2^t & \cdots & F_{M+1}^t \\ F_2^t & F_3^t & \cdots & F_{M+2}^t \\ \vdots & \iddots & \iddots & \vdots \\ F_{M+1}^t & F_{M+2}^t & \cdots & F_{2M+1}^t \end{bmatrix}}_{\Gamma_{M+1}} \underbrace{\begin{bmatrix} A_M^t \\ A_{M-1}^t \\ \vdots \\ A_0^t \end{bmatrix}}_{\hat{\mathbf{A}}} = \begin{bmatrix} \mathbf{0}_m \\ \vdots \\ \mathbf{0}_m \\ F_0^t \mathbf{e}_i \end{bmatrix} \tag{9a}$$

$$\begin{bmatrix} A_M & A_{M-1} & \cdots & A_0 \\ A_{M-1} & A_{M-2} & \cdots & \mathbf{0}_{1\times m} \\ \vdots & \iddots & \iddots & \vdots \\ A_0 & \mathbf{0}_{1\times m} & \cdots & \mathbf{0}_{1\times m} \end{bmatrix} \begin{bmatrix} F_0 \\ F_1 \\ \vdots \\ F_M \end{bmatrix} = - \begin{bmatrix} B_M \\ B_{M-1} \\ \vdots \\ B_0 \end{bmatrix} \tag{9b}$$

où $\mathbf{e}_i$ est le ième vecteur de la base canonique de $\mathbb{R}^m$, solvent l'identité de Bezout (5) pour un facteur d'échelle α donné par $\alpha = P_0^t \mathbf{e}_i$.

Remarque 1 : Ce théorème généralise [5, Theorem 8.8.1], donné dans le cas scalaire $m = 1$.

Preuve. Pour un polynome $X(z) = X_0 + X_1 z + \cdots + X_K z^K$, on définit son polynome réciproque par $\hat{X}(z) = X_K + X_{K-1}z + \cdots + X_0 z^K = z^K X(z^{-1})$. Avec cette définition et celle donnée en (8), l'équation (7) peut être réécrite, pour $M = \kappa - 1$ et $\nu = 2\kappa - 1$, sous la forme

$$\hat{A}(z^{-1})F(z) + \hat{B}(z^{-1}) = \alpha z^{M+1} Q^\sharp(z). \tag{10}$$

Cette dernière équation se décompose en

$$[\hat{A}(z^{-1})F(z)]_+ = \alpha z^{M+1} Q^\sharp(z) \tag{11a}$$

$$[\hat{A}(z^{-1})F(z)]_- = -\hat{B}(z^{-1}), \tag{11b}$$

où $[X(z)]_+ = X_1 z + X_2 z^2 + \cdots$ et $[X(z)]_- = \cdots + X_{-1}z^{-1} + X_0$ désignent respectivement la projection strictement causale et la projection anticausale de $X(z)$.

Les équations (9a) et (9b) découlent d'une simple réécriture, sous forme matricielle, respectivement de (11a) et (11b). □

Les équations (9a-9b) montrent que $A(z)$ et $B(z)$ ne dépendent que des coefficients $F_i, i = 0, 2M+1$. Ainsi, la donnée de la fonction (pseudo) rationnelle $F(z)$ permet de résoudre (5). On peut alors reposer le problème 1, pour un retard ν maximal, en terme d'identification de $F(z)$ à partir des observations $Y(n)$. Or, nous établissons le

Lemme 1. *Soit $\{Y(n)\}$ une séquence d'observations obtenue selon le modèle (1) et partitionné, pour $L = 2m$, selon*

$$Y(n) = \begin{bmatrix} Y_1(n) \\ \hdashline Y_2(n) \end{bmatrix} = \begin{bmatrix} h_1(z) \\ \vdots \\ h_m(z) \\ \hdashline h_{m+1}(z) \\ \vdots \\ h_{2m}(z) \end{bmatrix} s(n) + W(n) \tag{12}$$

$$= \begin{bmatrix} P(z) \\ \hdashline Q(z) \end{bmatrix} s(n) + \begin{bmatrix} W_1(n) \\ \hdashline W_2(n) \end{bmatrix}.$$

Si $Q(z)$ est irréductible, alors la fonction de transfert $F(z) = P(z)Q^{\#}(z)$ est donnée par la solution du problème d'identification entrée-sortie, représenté par la minimisation de l'erreur quadratique moyenne

$$\min_{\tilde{F}(z)} E\|Y_1(n) - \tilde{F}(z)Y_2(n)\|_2^2, \tag{13}$$

où $E\{\cdot\}$ désigne l'espérence mathématique.

Preuve. Pour commencer, posons

$$\mathbf{e}(n) \triangleq Y_1(n) - \tilde{F}(z)Y_2(n).$$

En éliminant $s(n)$ des relations

$Y_1(n) = P(z)s(n) + W_1(n)$ et $Y_2(n) = Q(z)s(n) + W_2(n)$,

$Y_1(n)$ apparaît, comme illustré sur le schéma d'identification de la figure 1, comme la sortie bruitée d'un filtre d'entrée bruitée, $Y_2(n)$, et de fonction de transfert $F(z)$.

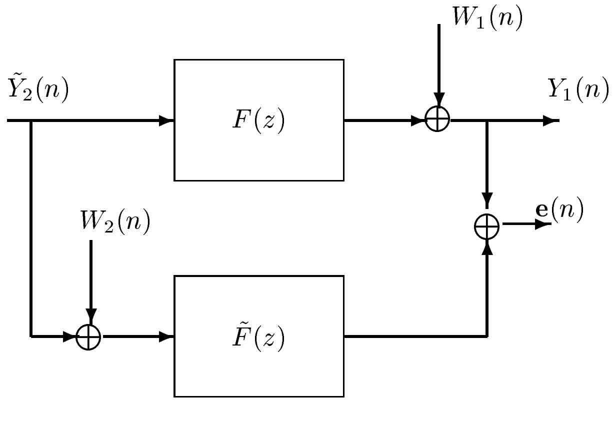

FIG. 1: Schéma d'identification

En effet, puisque $Q(z)$ est irréductible, on peut écrire

$$Q^{\#}(z)[Y_2(n) - W_2(n)] = Q^{\#}(z)Q(z)s(n) = s(n).$$

En reportant cette expression de $s(n)$ dans la première relation, on obtient

$$\begin{aligned} Y_1(n) &= P(z)Q^{\#}(z)(Y_2(n) - W_2(n)) + W_1(n) \\ &= F(z)[Y_2(n) - W_2(n)] + W_1(n), \end{aligned}$$

En notant $\tilde{Y}_2(n) \triangleq Y_2(n) - W_2(n)$, la sortie $Y_2(n)$ en l'absence de bruit, l'erreur d'identification $\mathbf{e}(n)$ prend la forme

$$\mathbf{e}(n) = [F(z) - \tilde{F}(z)]\tilde{Y}_2(n) - \tilde{F}(z)W_2(n) + W_1(n).$$

Ainsi, sous l'hypothèse que le bruit $W(n)$ est *i.i.d*, centré, avec $\mathrm{E}\{W(n)W^{\dagger}(n)\} = \sigma_w^2 \mathbf{I}_{2m}$, la minimisation (13) ci-dessus s'écrit

$$\min_{\tilde{F}(z)} \left\{ \|F(z) - \tilde{F}(z)\|^2_{\mathcal{S}_{\tilde{Y}_2}} + \sigma_w^2(\|\tilde{F}(z)\|_2^2 + m) \right\}, \tag{14}$$

où $\mathcal{S}_{\tilde{Y}_2}(z)$ est la fonction (matricielle) de densité spectrale du processus stationnaire $\tilde{Y}_2$ et $\|\cdot\|_{\mathcal{S}_{\tilde{Y}_2}}$ est la norme $L_2([0,2\pi], \mathcal{S}_{\tilde{Y}_2}(e^{i\theta})d\theta)$. Évidemment, la sortie non bruitée $\tilde{Y}_2(n)$ est supposée être à excitation persistente, pour que $\|\cdot\|_{\mathcal{S}_{\tilde{Y}_2}}$ soit une norme.

Le problème de minimisation (14) montre clairement que la fonction $F(z)$ peut être déterminée de façon simple, par une identification adaptative de système entrée-sortie. Par ailleurs, il convient de noter qu'une contrainte sur $\tilde{F}(z)$, de type norme unitaire permet de corriger le caractère biaisé de cette identification, dû à la présence de bruit. □

Remarque 2 : La minimisation de l'erreur quadratique moyenne (13) s'interprète directement en termes de prédiction linéaire des signaux reçus sur les capteurs 1 à m à partir de ceux des capteurs $m+1$ à $2m$. Il s'agit alors de prédiction spatio-temporelle.

Cette remarque achève la construction d'une solution du problème 1, d'égalisation aveugle.

3 Algorithme et Simulation

Par une approche polynomiale, nous avons proposé dans la section précédente une méthode pour construire une solution au problème d'égalisation autodidacte. Cette section-ci aborde des questions concernant l'implémentation et les propriétés de la méthode.

3.1 Algorithme

Un résumé des différentes étapes développées pour la résolution du problème 1 conduit directement à l'algorithme ci-dessous, que nous proposons pour l'égalisation autodidacte d'un système multicanal comme celui décrit par le modèle de l'équation (1).

1. Estimation de $F(z)$ par une identification adaptative de système à réponse impulsionnelle finie, d'entrée $Y_2(n)$ et de sortie $Y_1(n)$. Il s'agit d'un problème classique qui, aujourd'hui, a atteint un certain degré de maturité.
2. Estimation du degré κ à partir d'une SVD de la matrice bloc Hankel Γ_{M+1}.
3. Obtention de $A(z)$ et $B(z)$ par la résolution de (9a) puis (9b).
4. Égalisation : reconstituer $\tilde{s}_n = A(z)Y_1(n) + B(z)Y_2(n)$.

Nous allons maintenant aborder la question des performances et propriétés de cet algorithme, notamment en ce qui concerne la sous-modélisation et aussi la sur-modélisation. Ceci permettra, du même coup, d'expliciter l'étape 2 de l'algorithme que nous venons de présenter.

On suppose toujours que le canal $H(z)$ a un degré (formel) κ et que on cherche une égalisation avec un retard $\nu = 2\kappa - 1$.

3.1.1 Sous-modélisation

Lorsque le degré, M, choisi pour le filtre d'égalisation $G(z) = [A(z)B(z)]$ est tel que $M < \kappa - 1$, alors la réponse combinée, canal-égaliseur, ciblée–$c(z) = z^{\nu}$–n'est pas atteignable. Dans ce cas, on dit que l'égaliseur $G(z)$ est sous modélisé pour ce retard ν. Dans cette situation, il vient d'après l'équation (9a) que la réponse combinée, qui est un polynome scalaire de degré ν est de la forme

$$c(z) = z^M r(z),$$

où $r(z)$ est par conséquent un polynome de degré $\nu - M$.

3.1.2 Sur-modélisation

La sur modélisation a lieu lorsque $M > \kappa$. Dans ce cas, les seules solutions $A(z)$ de l'équation (11a) (et de façon équivalente (9a)) sont celles pour lesquelles $\hat{\mathbf{A}}$ appartienne au noyau de Γ_{M+1}. Or, ceci implique la nullité du facteur d'échelle α, donc on obtient une réponse combinée nulle, c'est à dire un égaliseur avec une sortie nulle.

3.2 Simulation

Nous présentons maintenant quelques résultats de simulation pour étayer notre propos. On considère une source avec une constellation MAQ-16. Le canal $H(z)$, constitué de $L = 4$ sous-canaux, est représenté par un polynome $\mathbb{R}^4[z]$, de degré $\kappa = 6$, dont les coefficients (obtenus à l'aide de `random`) sont donnés dans la table 1.

TAB. 1: Paramètres du Canal

$\mathbf{h}_0$	$\mathbf{h}_1$	$\mathbf{h}_2$	$\mathbf{h}_3$	$\mathbf{h}_4$	$\mathbf{h}_5$	$\mathbf{h}_6$
-0.06	-1.01	0.61	0.50	1.69	0.59	-0.64
-1.05	1.41	-0.80	0.52	0.22	-0.92	-2.17
-1.33	0.71	1.62	-0.69	0.86	1.25	-1.59
-1.44	0.57	-0.39	0.69	0.81	0.71	1.29

La figure 2 représente les modules des coefficients de la réponse combinée $c(z)$, obtenue par l'algorithme ci-dessus, pour un filtre d'égalisation de degré $M = 5$. La phase d'identification a été réalisée à l'aide de l'algorithme adaptatif des moindres carrées, sans contrainte sur la norme de $\tilde{F}(z)$. La solution est alors biaisée par la présence de bruit. L'exemple de cette figure montre que pour un niveau de bruit relativement faible, on a ici un rapport signal sur bruit de 30 dB, la méthode donne une égalisation satisfaisante du canal. En effet, l'égaliseur ainsi obtenu permet, comme l'illustre la figure 3 représentant la sortie $\tilde{s}(n)$, de reconstituer la constellation des symboles émises.

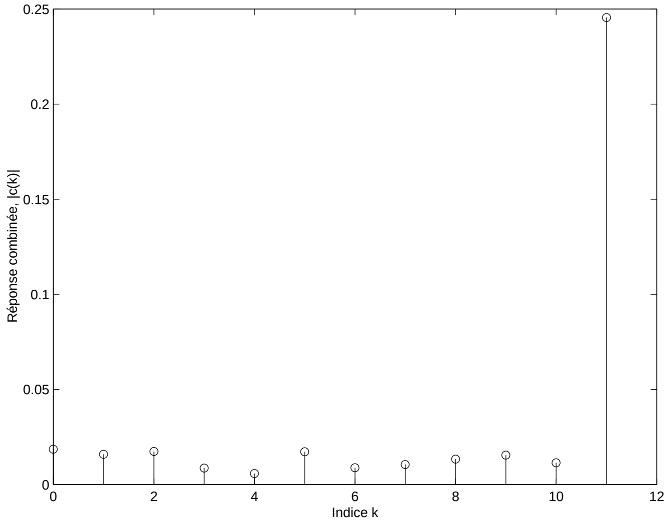

FIG. 2: Égalisation du canal $H(z)$

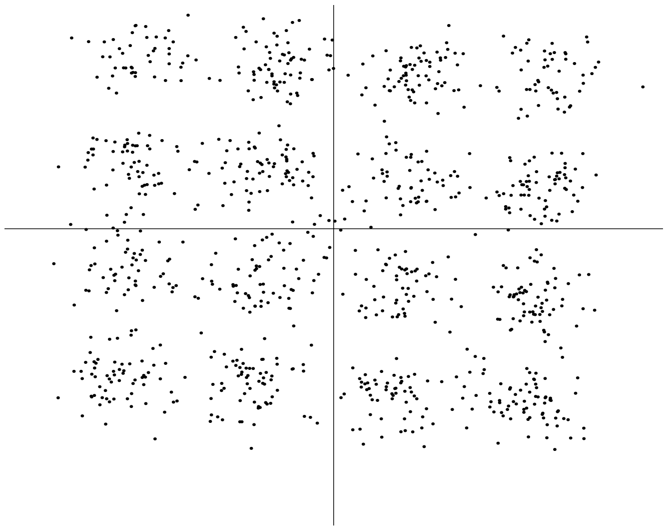

FIG. 3: Reconstruction des symboles

Références

[1] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. on Sig. Proc.*, vol. 43, pp. 515–525, Febrary 1995.

[2] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least squares approach to blind channel identification," *IEEE Trans. on Sig. Proc.*, vol. 43, pp. 340–349, 1995.

[3] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proc. of the IEEE*, vol. 86, pp. 1951–1968, October 1998.

[4] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. of the IEEE*, vol. 86, pp. 1927–1950, 1998.

[5] P. A. Fuhrmann, *A polynomial approach to linear algebra.* Universitext, Springer, 1996.