

# DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

155



NICE du 20 au 24 MAI 1985

MODELISATION AVEC ENTREE MULTI-IMPULSIONNELLE

P. FABRE, C. GUEGUEN

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS  
46, Rue Barrault 75634 PARIS CEDEX 13

## RESUME

La modélisation autorégressive classique a remporté de nombreux succès dans son application à divers types de signaux, mais a également montré certaines limites. Les défauts peuvent être attribués à divers facteurs comme des hypothèses peu réalistes sur l'excitation (bruit blanc) ou sur le modèle (autorégressif seulement), à la suppression de l'information de phase, ou à l'élimination des événements temporels courts. D'où l'idée, née dans le contexte de la parole, d'une modélisation plus fine de l'excitation comme un train d'impulsions convenablement positionnées et dimensionnées. Ces techniques, utilisées initialement en codage de la parole, peuvent résoudre un problème de déconvolution, dont l'application hors du domaine de la parole a été peu tentée.

Cette communication effectue une synthèse des concepts et des méthodes relatives à ce problème. Après une critique des modèles autorégressifs classiques, on expose les principes de l'analyse multi-impulsionnelle classique et on rappelle l'algorithme classique. Divers problèmes d'implantation, ainsi que des variantes intéressantes de la méthode, sont ensuite présentées. En particulier, une technique originale d'estimation récursive avec placement séquentiel des impulsions sur des critères de rupture est ensuite examinée en détail. L'article est illustré de divers simulations sur des signaux synthétiques, permettant d'évaluer méthodes et variantes, en particulier dans un but de déconvolution ; enfin, on propose divers exemples d'application à des signaux réalistes dans le contexte de la parole et des signaux sismiques.

## SUMMARY

Classical linear prediction and autoregressive modelling techniques have widely shown their interest in a large range of applications. Nevertheless, they also exhibit some defects that can be attributed to various reasons : rough hypotheses on the input (white noise) or on the model (autoregressive only), deletion of phase information and short time duration events ... consequently, born in the domain of speech processing, the idea came out of a more accurate modelling of the input as a complex train of pulses of convenient location and size. These techniques can solve an interesting deconvolution problem which occurs in many other domains than speech.

The paper gives a review of the recent methods pertaining to this question. After a critical survey of linear prediction drawbacks, the multipulse modelling approach is summarized and the classical algorithm of ATAL and REMDE is recalled. Various implementation problems and slightly different versions are explored. In particular, a recursive estimation technique locating pulses on jump detection criteria is examined in details. The paper is illustrated by a set of simulations on synthetic signals allowing to rate the outcome of each method ; then, various examples are given on realistic data extracted from speech signals and seismic traces.



1. INTRODUCTION

La prédiction linéaire s'est révélée être une excellente technique de modélisation applicable à différents domaines : analyse et synthèse de la parole, communications numériques, géophysique ... Pourtant, les techniques paramétriques de ce type ne font que modéliser la réponse impulsionnelle du système ou du filtre : canal vocal, canal de transmission, ondelette de source, en supposant que l'entrée est assimilable à un bruit blanc. Cette hypothèse simplificatrice conduit souvent à de bons résultats, mais n'est pas exempte de défauts divers. Ainsi, en analyse de parole, par exemple, on observe que le résiduel (erreur de prédiction) demeure intelligible, ce qui est incompatible avec l'hypothèse de bruit blanc. Après avoir essayé divers raffinements en vue d'améliorer la qualité du modèle (fenêtres très courtes, analyse synchrone, ajout d'une partie MA courte ...), la nécessité d'une entrée élaborée s'est imposée.

En effet, si le modèle AR fournit dans beaucoup d'applications une estimation satisfaisante du spectre de puissance, la phase est par contre ignorée (ou ajustée à sa valeur minimale déduite du modèle). Le résiduel temporel contient par contre une information de phase et conserve les événements courts non prédictibles par le passé. L'estimation spectrale réalisée modélise mal, en l'absence de partie MA, les vallées du spectre. Par ailleurs, l'amortissement estimé est peu réaliste et se révèle en général excessif (surtout en présence de bruit). D'où l'intérêt d'un apport différé d'énergie pour compenser la décroissance temporelle trop rapide de la réponse impulsionnelle. Ceci peut être réalisé par un modèle MA long par rapport à l'ordre AR ; afin de diminuer le nombre de paramètres, ce modèle MA long comportera avantageusement un grand nombre de coefficients nuls. Ceci revient à choisir comme excitation un train (clairsemé) d'impulsions (figure 1).

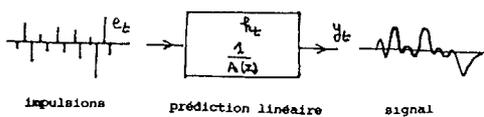


Figure 1

Le diagramme de modélisation typique, associé à une analyse par synthèse correspondra alors à celui de la figure 2, assez similaire au schéma classique. L'entrée est alors une suite d'impulsions dont les positions et amplitudes doivent être judicieusement choisies. Celle-ci remplace donc l'entrée traditionnelle : bruit blanc ou impulsions périodiques. On met ainsi en relief la cohérence de la problématique avec celle de la recherche d'un modèle ARMA, où la partie AR est déterminée de façon classique par filtrage inverse et où la suite d'impulsions est interprétable comme un modèle MA, B(z), clairsemé. L'entrée  $u_t$  est inaccessible en traitement du signal (par opposition de l'automatique). On constate la présence de deux boucles d'ajustement possédant éventuellement leur filtrage de l'erreur propre et leur stratégie de calcul. Ces deux boucles peuvent être éventuellement imbriquées (figure 3).

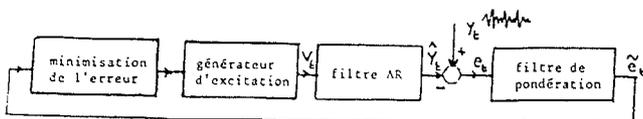


Figure 2

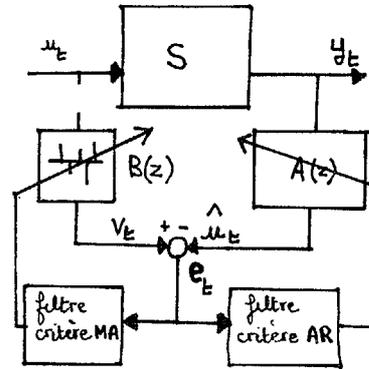


Figure 3

On donnera dans la suite diverses techniques de calcul des paramètres de ce type de modèle, en concentrant particulièrement l'attention sur l'élaboration d'un train d'impulsions. Le paragraphe 2 résume l'approche originale proposée en 1982 par Atal et Remde [1]. Le paragraphe 3 évoque divers problèmes d'implantation de ce type de méthode : choix d'une fonction de positionnement, prise en compte des effets de bords des fenêtres d'analyse, accentuation de l'effet local d'une impulsion, conception d'une salve d'impulsions adjacentes. Diverses solutions originales sont proposées et des approches précédentes se trouvent synthétisées dans une même formulation. Le paragraphe 4 propose une technique récursive qui pose des impulsions au fil du temps sur des critères de rupture et de détection d'évènement, puis en ajuste les amplitudes de manière séquentielle. Les diverses méthodes sont illustrées et comparées sur des signaux synthétiques engendrés par des trains d'impulsions réguliers ou aléatoires. L'article se termine par une évaluation de ces approches sur des signaux réalistes. Si, en effet, ce genre de méthode a trouvé un terrain d'application privilégié dans le codage de la parole avec une amélioration très sensible de la parole synthétique en conservant un débit relativement réduit (voisin de 10 kb/s), ce même type d'approche semble, malgré sa généralité, avoir peu été tenté par ailleurs. On donnera un exemple d'application dans le domaine sismique pour le positionnement des couches géologiques et la détermination de leur coefficient de réflexion (paragraphe 5).

2. DESCRIPTION DE LA METHODE MULTI-IMPULSIONNELLE

2.1 Principe

La méthode multi-impulsionnelle, telle qu'introduite par Atal et Remde en 1982 [1], consiste à déterminer les positions et amplitudes des impulsions qui minimisent un critère lié aux moindres carrés de l'erreur entre le signal original  $y_t$  et le signal synthétique  $\hat{y}_t$  :

$$E = \sum_{t=1}^T (y_t - \hat{y}_t)^2$$

suivant une méthode d'analyse par synthèse illustrée par la figure 2. En effet, si l'on suppose que l'on a placé n impulsions d'amplitudes  $l_1, \dots, l_n$  aux instants  $t_1, \dots, t_n$ , le signal synthétisé s'écrit :

$$\hat{y}_t = h_t * v_t \quad v_t = \sum_i l_i \delta_{t-t_i}$$

où  $h_t$  est la réponse impulsionnelle du modèle AR (avec  $h_t=0$  pour  $t=0, t>N$ ) et  $v_t$  est l'entrée multi-impulsionnelle. L'erreur  $e_t$  correspondant au placement de ces  $n$  impulsions est donc :

$$e_t = y_t - \hat{y}_t = h_t \star (r_t - v_t)$$

où  $r_t$  est le résiduel de prédiction linéaire.

Pour diverses raisons, qui seront explicitées au paragraphe 3, on remplace cette erreur de synthèse par une erreur pondérée, obtenue par filtrage :

$$\tilde{e}_t = W_t \star e_t$$

Dans le domaine de la parole, on choisit pour filtre  $W(z)$ , celui défini par :

$$W(z) = A(z)/\hat{A}(z)$$

Appelé "filtre perceptuel". Il revient à remplacer la réponse impulsionnelle  $h_t$  du modèle AR  $A(z)$ , par celle  $\hat{h}_t$  associée à un modèle  $\hat{A}(z)$

$$\tilde{e}_t = \hat{h}_t \star (r_t - v_t)$$

### 2.2 Détermination des impulsions

Pour déterminer les amplitudes et positions des impulsions  $v_t$  optimales on procède en deux étapes :

Les positions étant supposées choisies, la minimisation du critère des moindres carrés sur  $\tilde{e}_t$  s'effectue en écrivant que la dérivée partielle par rapport aux amplitudes des  $l_i$  est nulle, ce qui conduit par un calcul classique à :

$$\sum_{t=1}^T [\tilde{h}_{t-t_i} \tilde{r}_{t-1} \sum_{j=1}^n \tilde{h}_{t-t_i} \tilde{h}_{t-t_j}] = 0 \text{ pour } i=1, n$$

où  $\tilde{r}_t$  désigne le résiduel  $r_t$  filtré par la réponse impulsionnelle modifiée  $\tilde{h}_t$ . Prenant les notations :

$C : \{c_{ij}\} \quad c_{ij} = \sum_t \tilde{h}_{t-t_i} \tilde{h}_{t-t_j}$  matrice de covariance de  $\tilde{h}_t$  pour les retards  $t_i$  et  $t_j$

$S : \{s_{t_i}\} \quad s_{t_i} = \sum_t \tilde{h}_{t-t_i} \tilde{r}_t$  vecteur d'intercorrélation pour les retards  $t_i$

et  $l$  représentant le vecteur des amplitudes des impulsions aux instants  $t_i$  on a l'équation linéaire :

$$CL = S$$

dont la résolution à des  $t_i$  fixés, fournit les amplitudes cherchées.

La détermination des positions  $t_i$  optimales est, quant à elle, un problème non-linéaire. La solution exacte directe consisterait à calculer l'erreur correspondant à tous les jeux de  $n$  impulsions possibles dans la fenêtre et d'en choisir celui qui rend l'erreur minimale. Cette procédure combinatoire est inextricable. On adopte donc une solution itérative, sous-optimale, consistant à placer successivement les impulsions, grâce à une fonction convenable de positionnement, sans remettre en cause les impulsions précé-

entes. Le signal synthétisé se rapproche du signal original au fur et à mesure que l'on place les impulsions ([1]).

Une solution consiste, après avoir placé  $i$  impulsions, à déterminer la position  $t' = t_{i+1}$  et l'amplitude  $l' = l_{i+1}$ , de sorte que l'énergie de l'erreur pondérée, après placement de cette impulsion, soit minimale. Il est clair que :

$$\tilde{e}_t^{(i+1)} = \tilde{e}_t^{(i)} - l' h_{t-t'}$$

d'où l'énergie de l'erreur :

$$E_{i+1} = E_i - 2l' s_{t'}^{(i+1)} + l'^2 c_{t', t'}$$

et les valeurs optimales sont données par :

$$E_{i+1} = E_i - (s_{t'}^{(i+1)})^2 / c_{t', t'} \text{ pour } l' = (s_{t'}^{(i+1)}) / c_{t', t'}$$

On pourra donc choisir la position optimale  $t' = t_{i+1}$  sur ce critère. Une solution simple approximative consiste à prendre l'emplacement du maximum de  $s_{t'}^{(i)}$ .

### 2.3 Analyse d'un signal synthétique

La méthode classique est appliquée à un signal synthétique constitué par une suite d'impulsions poissonniennes-gaussiennes (figure 4a) filtrée par un modèle AR d'ordre 4 dont le module des pôles est compris entre 0.5 et 0.6 (signal 1). On peut alors constater plusieurs tendances :

- les positions des impulsions synthétiques sont dans l'ensemble correctes.
- les groupes d'impulsions très rapprochées sont remplacés par des impulsions uniques dont la valeur est approximativement égale à la somme des amplitudes des impulsions exactes.
- les impulsions isolées, même de faible amplitude, sont correctement déterminées.
- il apparaît des impulsions parasites (même si le nombre d'impulsions choisi est égal au nombre d'impulsions réel).

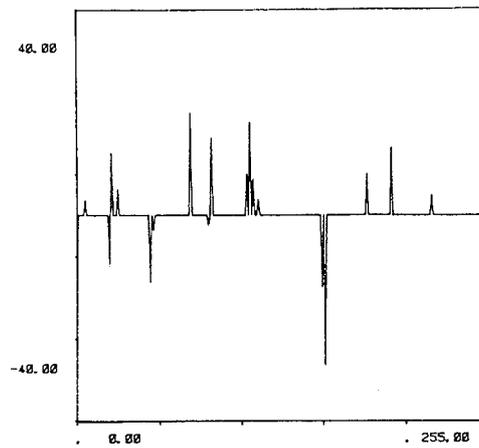


Figure 4a - entrée exacte.

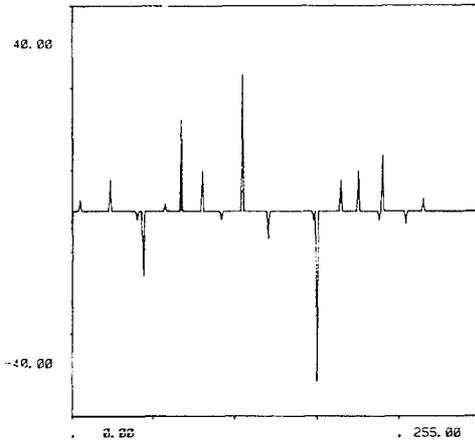


Figure 4b - entrée estimée par la méthode classique.

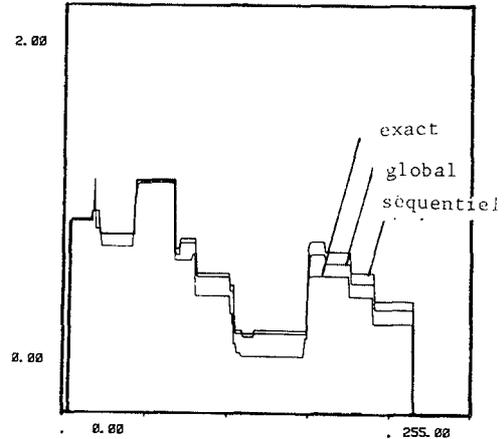


Figure 6

3 Problèmes d'implantation de la méthode multi-impulsionnelle

Divers problèmes se posent lorsqu'on veut utiliser la méthode multi-impulsionnelle : calcul purement séquentiel ou global des impulsions, effets de bord, choix du facteur de pondération [2], [3], [4], [5], [6], [7], [8]. Il faut évaluer la qualité de l'excitation multi-impulsionnelle obtenue : en traitement de la parole, on étudie l'énergie de l'erreur entre le signal synthétisé et le signal original (voir [2], par exemple). Dans le contexte d'autres signaux, et dans un souci de déconvolution, il n'est pas si simple de trouver un critère objectif. Si les impulsions sont placées aux positions exactes, l'erreur quadratique moyenne mesure bien l'écart entre les trains d'impulsions exact et estimé. Ce n'est hélas presque jamais le cas ... ! En s'inspirant des signaux sismiques, où les impulsions représentent des variations d'impédance acoustique entre couches adjacentes, on trouve finalement un critère intéressant : l'erreur quadratique moyenne entre les courbes d'impédance exacte et estimée (en tout état de cause, il s'agit aussi d'un signal synthétisé particulier).

3.1 Calcul global ou séquentiel des impulsions

Le calcul global des impulsions permet d'obtenir les amplitudes optimales au sens des moindres carrés, leurs positions étant fixées, mais nécessite la résolution d'un système linéaire dont la matrice de covariance n'est pas de Toeplitz. La résolution peut se faire par la méthode de Cholesky. La différence entre les amplitudes obtenues est minime (tableau 5). Les courbes d'impédances correspondant aux entrées exacte, estimée séquentiellement, estimée globalement, du signal 1 sont représentées sur la figure 6. On constate que l'estimation globale est toujours meilleure que l'estimation séquentielle ; l'erreur quadratique vaut 2,32 dans le cas de l'estimation séquentielle, et tombe à 1,25 dans le cas de l'estimation globale.

On peut donc éviter le calcul global lorsque sa complexité pose problème - en traitement de la parole, il y a une contrainte de temps réel - mais le calcul global est indispensable dans les autres cas.

	signal 2	signal 2 bruité	signal 3
énergie des impulsions	13.00	13.10	10.80
énergie de l'erreur	$3.54 \cdot 10^{-4}$	$2.0 \cdot 10^{-5}$	$5.8 \cdot 10^{-3}$
rapport	$2.73 \cdot 10^{-5}$	$1.5 \cdot 10^{-6}$	$5.3 \cdot 10^{-4}$

Tableau 5

3.2 Effets de bord des fenêtres d'analyse

Les effets de bord dus au découpage des fenêtres ont un effet notable sur les résultats de la méthode. En effet, on effectue généralement le calcul des coefficients de prédiction linéaire sur des fenêtres de N échantillons (N=128, 256) avec recouvrement de moitié. Si l'on place des impulsions sur toute la fenêtre, on doit alors enlever la contribution de la réponse impulsionnelle des impulsions de la fin de la fenêtre aux échantillons de la fenêtre suivante. C'est la raison pour laquelle on choisit de placer les impulsions exclusivement dans la première demi-fenêtre, ou dans la demi-fenêtre médiane, et de prendre en compte les impulsions ainsi placées dans le calcul de la fenêtre suivante.

La méthode de modélisation utilisée est la méthode d'autocorrélation. En effet, la méthode de covariance est nettement plus complexe et n'améliore pratiquement pas les résultats obtenus ([5]).

3.3 Pondération de l'erreur

Le facteur de pondération  $\gamma$ , présenté au paragraphe 2, se révèle être un paramètre-clé de la méthode. Il a été introduit au départ afin de diminuer l'importance de l'erreur dans les régions formantiques du signal de parole (ATAL [2]). En effet, ce sont les régions du spectre où l'énergie est la plus concentrée, et il y a un effet de masquage. En fait, le facteur de pondération est beaucoup plus qu'un facteur perceptuel, et il se révèle aussi utile hors du domaine de la parole. Le facteur perceptuel peut être considéré comme un facteur de lissage, permettant d'obtenir une erreur pondérée intermédiaire entre le signal analysé et le résiduel, et comme une pondération temporelle, corrigeant le manque d'impulsions en fin de fenêtre.





Le calcul du vecteur des amplitudes,  $L$  se fait en résolvant le système linéaire :

$$RL = S$$

identifique à celui que l'on rencontre dans le cas de la méthode globale, par le lemme d'inversion matricielle.

La méthode décrite précédemment a été mise en oeuvre sur des signaux de parole numérisée. L'algorithme de modélisation est l'algorithme de Ljung-Falconer, à l'ordre 12, avec facteur d'oubli ( $\lambda=0,99$ ), réinitialisé tous les 1000 échantillons environ. Les seuils choisis sont sur l'erreur de prédiction :  $C_o=0,1.\sigma_e^2$  et sur la variation du modèle :  $C_\beta=0,001$ .

La phase du signal est bien reconstituée dans le signal synthétisé, mais il manque malheureusement des hautes fréquences. Les résultats de la méthode récursive sont, en définitive, moins bons que ceux de la méthode globale. Une différence essentielle est que la contribution d'une impulsion placée sur le signal n'est pas réintroduite dans le signal à modéliser : il n'y a pas d'analyse-par-synthèse comme dans la méthode globale.

#### 5. APPLICATION DE LA METHODE MULTI-IMPULSIONNELLE A LA DECONVOLUTION

Il faut maintenant approfondir les caractéristiques de la méthode multi-impulsionnelle en tant que méthode de déconvolution, et comment on l'applique dans un cas réel, celui de signaux sismiques.

##### 5.1 Pouvoir de résolution

La méthode multi-impulsionnelle a tendance à remplacer des groupes d'impulsions très rapprochés, de même signe, par des impulsions uniques (section 2.3). Si l'on considère un doublet (signal 3, paragraphe 3.3), on note que l'écart temporel entre les impulsions est augmenté et que leur amplitude est diminuée, ce qui est gênant. Comment y remédier ? en augmentant le nombre d'impulsions placées. Dans ce cas, des impulsions sont placées pratiquement à chaque échantillon au voisinage du doublet et un calcul global assure l'obtention d'une valeur correcte des amplitudes. Il faut donc, d'une part calculer un nombre assez élevé d'impulsions, et d'autre part faire un calcul global pour bien modéliser la structure fine de l'entrée.

##### 5.2 Comportement en présence de bruit

Il est important d'évaluer le comportement de la méthode multi-impulsionnelle lorsque le signal analysé est entaché d'un bruit dont le niveau est élevé (entre 0 et 20 dB). On peut déjà affirmer que le modèle AR identifié est biaisé (on montre qu'il faut un modèle ARMA pour modéliser correctement un modèle AR bruité), ce qui donne une erreur pondérée initiale élevée. On note expérimentalement que les positions des impulsions sont correctement détectées, mais que les amplitudes estimées sont nettement supérieures en valeur absolue aux amplitudes réelles. De manière générale, on peut se servir de la courbe de l'énergie de l'erreur pondérée en fonction du nombre d'impulsions considérées comme justes par un critère d'arrêt (le reste des impulsions pouvant être considéré comme du bruit).

##### 5.3 Mise en oeuvre sur des signaux sismiques

Le traitement des signaux sismiques a pour but de déterminer la position et la nature des couches du sous-sol afin de savoir s'il y a des minéraux intéressants : pétrole, gaz naturel ... dans le sous-sol sondé. L'onde acoustique produite par une explosion en surface est transmise et réfléchiée par les différentes couches géologiques et enregistrée à diverses profondeurs (sismique verticale). L'onde reçue peut être représentée comme la convolution de l'ondelette de source avec le train d'impulsions constitué par le milieu : la position (temporelle) des impulsions cor-

respond à l'emplacement (spatial) des interfaces et leur amplitude à leur coefficient de réflexion. Il s'agit donc d'un problème de déconvolution en présence de bruit, auquel on a appliqué la méthode multi-impulsionnelle (ordre du modèle  $p=8$ ,  $\gamma=0,7$ ,  $N=256$ ). La figure 11 montre l'analyse d'une partie du sismogramme synthétique défini par MENDEL dans [11]. On observe bien la propagation des ondes descendantes.

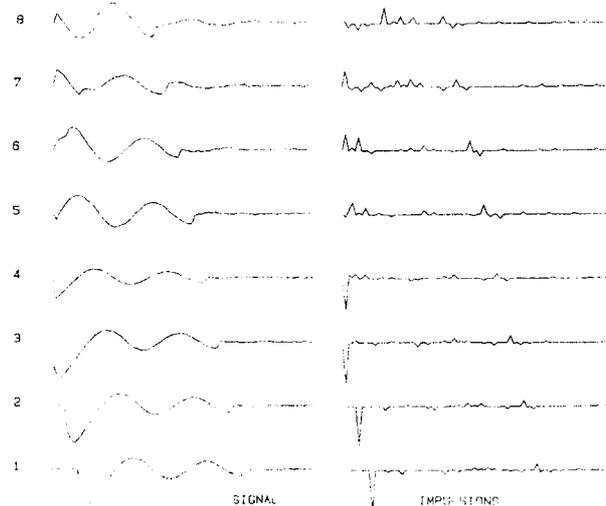


Figure 11

#### Conclusion

A l'origine méthode de codage de la parole, la méthode multi-impulsionnelle a en fait un intérêt plus général et peut s'appliquer à certains problèmes de déconvolution. Son apport réside dans l'introduction de l'analyse par synthèse et d'un filtrage de l'erreur de modélisation. Elle apparaît comme une variante efficace de la modélisation MA abordée sous l'angle des conditions initiales. La méthode doit bien s'appliquer lorsque l'entrée réelle est bien un train clairsemé d'impulsions, ou peut y être ramenée. L'efficacité dans le cas d'une entrée stochastique reste à confirmer ; on peut songer à des variantes intégrant dans l'entrée estimée une partie déterministe et une partie aléatoire, associant l'approche MA traditionnelle et la méthode multi-impulsionnelle.

#### Bibliographie

- [1] B.S. ATAL, J.R. REMDE, "A new model of LPC excitation ...", proc. of ICASSP 82, Paris, pp 614-617.
- [2] S. SINGHAL, B.S. ATAL, "Optimizing LPC filter parameters for multi-pulse excitation", ICASSP 83, 17.2.
- [3] S. SINGHAL, B.S. ATAL, "Improving performance of multi-pulse LPC coders at low bit rates", ICASSP 84, 1.3.1.
- [4] V.K. Jain, "Simplified algorithm for multi-pulse LPC analysis of speech", ICASSP 83, 23.2.1.
- [5] M. BEROUTI et al., "Efficient computation and encoding of the multi-pulse excitation for LPC", ICASSP 84.
- [6] T. ARASEKI et al., "multi-pulse excited speech coder based on maximum crosscorrelation search algorithm", ICASSP 84, 23.3.1.
- [7] G.A. SENENSIEB et al., "A non-iterative algorithm for obtaining multi-pulse excitation for linear-predictive speech coders", ICASSP 84, San Diego, 10.2.1.
- [8] P. KROON, G. DEPREITERE, "Experimental evaluation of different approaches to the multi-pulse coder", proc. of ICASSP 84, San Diego, 10.4.1.
- [9] J. LEROUX, F. GIANNELLA, "Whiteness of the measurement noise as a criterion for ARMA modelization of speech", proc. of ICASSP 82, Paris, pp 1353-1356.
- [10] L. LJUNG, M. MORF, D. FALCONER, "Fast calculation of gain matrices for recursive estimation schemes", Int. J. of Control, vol. 27, n° 1, 1978, 1-19.
- [11] F. HABIBI-ASHRAFI, J.R. MENDEL, "Estimation of parameters in lossless layered media systems", IEEE Trans. on Automatic Control, Vol AC-27, n° 1, Février 1982.