# DÉCONVOLUTION DE PROCESSUS IMPULSIONNELS AVEC CALCUL EXACT D'UN CRITÈRE MAP

**Yves Goussard et Guy Demoment**

Laboratoire des Signaux et Systèmes (CNRS / ESE / UPS), École Supérieure d'Électricité
Plateau de Moulon, 91192 GIF-SUR-YVETTE Cédex, France

et Gréco Traitement du Signal et Image

## RÉSUME

**Cet article présente une méthode itérative de déconvolution de processus Bernoulli-gaussiens par maximum *a posteriori*. Ce problème de détection-estimation est reformulé comme celui d'un changement de conditions initiales dans un problème de moindres carrés. Ceci conduit à un algorithme de structure très simple et qui permet le calcul exact du critère MAP. La méthode résultante est de mise en œuvre aisée, numériquement peu coûteuse tout en restant très proche de l'optimalité.**

## SUMMARY

**This article presents an iterative maximum *a posteriori* deconvolution method for Bernoulli-Gaussian processes. This detection-estimation problem is formulated as that of a change of initial conditions in a least-squares problem. This yields an algorithm with a very simple structure which allows the evaluation of the MAP criterion without any approximation. The resulting method is easy to implement, computationally inexpensive and remains nearly optimal.**

## 1. Introduction.

Le sujet de cette communication est la restauration de trains d'impulsions peu denses distordus par un système linéaire et affectés d'un bruit d'observation. Ce problème se rencontre dans de nombreux domaines des sciences pour l'ingénieur, et par exemple en exploration géophysique, en échographie médicale ultra-sonore, en contrôle non-destructif ou plus généralement lorsque l'on cherche à caractériser un milieu de propagation non-homogène à partir de mesures effectuées à sa surface. Dans ce cas, les observations peuvent être considérées en première approximation comme le produit de convolution de l'ondelette (forme de l'onde incidente) et de la réflectivité (dérivée de l'impédance acoustique) qui caractérise le milieu. Sous l'hypothèse d'un milieu composé de strates homogènes, la réflectivité est nulle partout sauf aux frontières entre strates, et le problème est donc de restaurer cette réflectivité à partir des observations et des informations disponibles sur l'ondelette et les caractéristiques du bruit. Il s'agit en général d'un problème de *déconvolution myope*, car ni l'ondelette, ni les caractéristiques statistiques du bruit ne sont connues précisément. Nous nous limiterons ici à la *déconvolution simple* dans laquelle la seule inconnue est la réflectivité à restaurer.

La déconvolution simple de processus impulsionnels présente les difficultés habituelles de résolution des problèmes mal-posés, auxquelles viennent s'ajouter d'autres obstacles liés à la nature de l'entrée et aux contraintes de mise en œuvre particulières aux domaines d'application visés. Quelques méthodes basées sur l'utilisation de normes $L_p$ [1], de critères d'entropie minimale [2] ou de techniques de codage multi-impulsionnel [3] ont été proposées. Celles-ci ne prennent pas en compte explicitement la nature particulière de l'entrée et le caractère mal posé du problème. Mendel *et al.* [4-6], suivi par d'autres auteurs [7-9], a proposé de décrire la réflectivité comme un processus Bernoulli-gaussien (B-G) et d'utiliser un critère du type vraisemblance *a posteriori* pour en effectuer la restauration. Toutes les méthodes de ce type présentent un certain degré de sous-optimalité. En effet, la restauration d'un processus B-G est un problème de détection-estimation qu'il est impossible de résoudre exactement en raison de la taille usuelle des signaux (plusieurs centaines à plusieurs milliers d'échantillons). Certaines de ces méthodes ont une structure récurrente afin de pouvoir traiter les données en ligne. Elles souffrent généralement d'un manque de robustesse, sensible surtout lorsque le contenu spectral de l'ondelette est pauvre. D'autres sont itératives, mais la lourdeur des calculs en rend la mise en œuvre et l'emploi difficiles.

La méthode proposée ci-après a pour objet de corriger en partie ce défaut. Il s'agit d'une procédure itérative de déconvolution de signaux B-G, basée sur une représentation par moyenne ajustée (MA) de l'ondelette. Les développements qui suivent sont très proches de ceux de Mendel *et al.*, ce qui n'est pas étonnant dans la mesure où tant les modèles des signaux que le critère à optimiser sont très proches. Par contre, on met ici en évidence une structure algorithmique très simple, du type *moindres carrés récurrents*, qui n'a pas été exploitée par les précédents auteurs. On est ainsi en mesure d'effectuer simultanément les opérations de détection et d'estimation, et de réduire le volume des calculs en même temps que la complexité de mise en œuvre.

## 2. Formulation du problème.

Sous les hypothèses de linéarité des phénomènes et d'additivité du bruit d'observation, l'équation entrée-sortie du système s'écrit

$$z(k) = \sum_{i=0}^{n} h(i)\, x(k-i) + n(k) \qquad 1 \le k \le P \tag{1}$$

où $z$, $x$ et $n$ représentent respectivement les observations, la reflectivité inconnue et le bruit d'observation. De manière à simplifier les calculs, on suppose que l'ondelette $h$ ne varie pas au cours du temps et que son support est de $n$+1 échantillons. Les résultats s'étendent facilement au cas d'une ondelette non-stationnaire. Par concaténation des échantillons de $z$, $n$ et $x$ dans des vecteurs $\boldsymbol{z}$, $\boldsymbol{n}$ et $\boldsymbol{x}$ de dimensions respectives $P$, $P$ et $N$, (1) peut être réécrite sous forme matricielle

$$\boldsymbol{z} = \boldsymbol{H}\, \boldsymbol{x} + \boldsymbol{n} \tag{2}$$

où $\boldsymbol{H}$ contient les échantillons décalés de l'ondelette $h$. Le bruit $n$ est supposé indépendant de $x$, blanc, gaussien centré et de variance $r^n$. Le modèle Bernoulli-gaussien de l'entrée $x$ s'exprime comme suit

$$\begin{aligned} &x(k)\text{: VA gaussienne centrée de variance } r^x\, t(k) \\ &t(k)\text{: VA de Bernoulli} \begin{cases} p\{t(k) = 1\} = \lambda \\ p\{t(k) = 0\} = 1-\lambda \end{cases} \end{aligned} \tag{3}$$

$x$ et $t$ étant des processus blancs. Lorsque $\lambda$ est petit devant un, le modèle (3) définit bien une suite d'impulsions peu dense. On remarque aussi que conditionnellement à la connaissance de la séquence de Bernoulli $t$, $x$ est un processus gaussien non-stationnaire. Contrairement à Mendel *et al.* [4, 5], on ne fait pas intervenir dans le modèle (3) de processus aléatoire gaussien auxiliaire $w$ tel que $x = t\, w$. On simplifie ainsi les données du problème d'estimation et la robustesse de la méthode se trouve améliorée [10].

La restauration de $\boldsymbol{x}$ se fait par maximisation du critère de vraisemblance *a posteriori*

$$J(\boldsymbol{t}, \boldsymbol{x}) \triangleq p\{\boldsymbol{t}, \boldsymbol{x} \mid \boldsymbol{z}\} \propto p\{\boldsymbol{z} \mid \boldsymbol{t}, \boldsymbol{x}\}\, p\{\boldsymbol{x} \mid \boldsymbol{t}\}\, p\{\boldsymbol{t}\} \tag{4}$$

qui, compte tenu de (3), prend la forme logarithmique suivante

$$\begin{aligned} J(\boldsymbol{t}, \boldsymbol{x}) \propto & -\frac{(\boldsymbol{z} - \boldsymbol{H x})'\,(\boldsymbol{z} - \boldsymbol{H x})}{r^n} - \frac{\boldsymbol{x}'\, \boldsymbol{T}\, \boldsymbol{x}}{r^x} \\ & - N_e\left(Ln\left(2\pi r^x\right) + 2Ln\left(\frac{1-\lambda}{\lambda}\right)\right) \end{aligned} \tag{5}$$

où $\boldsymbol{T}$ est définie par $\boldsymbol{T} \triangleq \mathrm{Diag}\{t(k)\}_{1 \le k \le N}$ et où $N_e$ est le nombre de uns de la séquence de Bernoulli. Un caractère ' indique une transposition. A séquence de Bernoulli fixée, les quantités $\boldsymbol{T}$ et $N_e$ sont connues et la maximisation de $J$ conduit à estimer $\boldsymbol{x}$ par

$$\hat{\boldsymbol{x}} = \boldsymbol{P}\, \boldsymbol{H}'\, (r^n)^{-1}\, \boldsymbol{z} \tag{6a}$$

$$\boldsymbol{P} = \boldsymbol{\Pi} - \boldsymbol{\Pi}\, \boldsymbol{A}\, \boldsymbol{\Pi} \tag{6b}$$

$$\boldsymbol{A} = \boldsymbol{H}' \left(\boldsymbol{H}\boldsymbol{\Pi}\boldsymbol{H}' + r^n \boldsymbol{I_d}\right)^{-1} \boldsymbol{H} \tag{6c}$$

$$\boldsymbol{\Pi} = r^x\, \boldsymbol{T} \tag{6d}$$

On reconnait dans (6) les formules classiques d'estimation d'une grandeur gaussienne par maximum *a posteriori*, et dans $\boldsymbol{\Pi}$ et $\boldsymbol{P}$ respectivement les matrices de covariance d'estimation *a priori* et *a posteriori*. Les équations (6) traduisent simplement le fait qu'à séquence de Bernoulli donnée, $\boldsymbol{x}$ est gaussien. La structure linéaire de l'estimateur et le fait que $\boldsymbol{t}$ intervienne dans (6) comme une condition initiale sont les points-clé de l'établissement de l'algorithme itératif.

## 3. Algorithme itératif.

La maximisation exacte de $J$ défini par (5) nécessiterait de calculer $\hat{x}$ et la valeur correspondante du critère pour les $2^N$ réalisations possibles de la séquence de Bernoulli. Comme mentionné plus haut, un tel calcul est inenvisageable, et même en n'essayant qu'une partie des réalisations possibles de $t$, on souhaite ne pas avoir à faire un calcul complet des $\hat{x}$ et $J$ correspondants. Pour cela, on définit la notion de *séquences voisines* de la manière suivante: deux séquences de Bernoulli sont voisines si elle diffèrent par une et une seule de leurs composantes. Puis, pour explorer l'ensemble des réalisations possibles de $t$, on procède de manière *séquentielle*, en passant d'une séquence de Bernoulli à une séquence qui lui est voisine. Une telle manière de faire se justifie seulement si on dispose d'équations simples liant les valeurs de $\hat{x}$ et $J$ correspondant à deux séquences voisines. De telles équations sont établies ci-après. Il s'agit là d'une étape indispensable pour l'application de cette méthode à des données de dimensions réalistes.

Se pose enfin la question de la stratégie d'exploration des différentes réalisations possibles de $\boldsymbol{t}$. La plus simple consiste, en partant d'une valeur initiale $\boldsymbol{t}_0$, à essayer toutes les séquences $\boldsymbol{t}$ qui lui sont voisines, à retenir celle qui rend $J$ maximum, et à recommencer jusqu'à atteindre un maximum relatif. C'est très exactement la technique proposée par Kormylo et Mendel avec le détecteur SMLR [4], et c'est celle qui est adoptée ici. Son principal inconvénient est que la convergence vers le maximum global de $J$ n'est pas garantie. Pour y remédier, on peut penser maximiser le critère par des techniques stochastiques, mais ce point ne sera pas abordé ici. On établit ci-dessous les équations liant les valeurs des estimées et des critères de deux séquences voisines.

### *3.1. Relation entre estimées et critères de deux séquences voisines.*

Dans la suite, on indexe respectivement par 0 et $k$ les grandeurs relatives à la séquence de Bernoulli initiale et à celle qui en diffère par sa composante d'indice $k$, et on appelle $\boldsymbol{v}_k$ le vecteur de dimension $N$ dont toutes les composantes sont nulles sauf celle d'indice $k$ qui est égale à 1. Les équations (6) montrent que pour trouver la relation entre $\hat{\boldsymbol{x}}_k$ et $\hat{\boldsymbol{x}}_0$, il suffit de trouver celle qui lie $\boldsymbol{P}_k$ à $\boldsymbol{P}_0$. On pourrait penser que, par application répétée du lemme d'inversion de matrice et en faisant intervenir les inverses des matrices $\boldsymbol{P}_k$ et $\boldsymbol{\Pi}_k$, le problème trouve une solution immédiate. Malheureusement, $\boldsymbol{P}_k$ et $\boldsymbol{\Pi}_k$ ne sont généralement pas inversibles en raison des zéros de la séquence $\boldsymbol{t}_k$, ce qui interdit une telle démarche. On est donc contraint de travailler sur les quantités elles-mêmes et non sur leurs inverses, ce qui constitue la principale difficulté pour l'établissement des équations. On ne donne dans ce qui suit que les grandes lignes de la démonstration. Pour les détails, se reporter à [11].

Partant de

$$\Pi_k = \Pi_0 \pm v_k r^x v_k' \tag{7}$$

où le signe ± correspond à l'addition ou la suppression d'un un à la séquence $t_0$, par application du lemme d'inversion de matrice à (6c), il est aisé d'établir que

$$A_k = A_0 - A_0 v_k \rho^{-1} v_k' A_0 \tag{8a}$$

$$\rho_k = \pm (r^x)^{-1} + v_k' A_0 v_k \tag{8b}$$

L'expression de $P_k$ est ensuite obtenue en utilisant (6b). Les calculs sont fastidieux et ne sont pas présentés ici. En notant

$$a_k \triangleq \Pi_0 A_0 v_k \tag{9}$$

on parvient à l'expression suivante

$$P_k = P_0 + (a_k - v_k)\, \rho_k^{-1}\, (a_k - v_k)' \tag{10}$$

En reportant cette dernière équation dans (6a) et en utilisant l'identité

$$\Pi_0 H' \left( H \Pi_0 H' + r^n I_d \right)^{-1} = P_0 H' (r^n)^{-1} \tag{11}$$

on obtient finalement

$$\hat{x}_k = (a_k - v_k)\, \rho_k^{-1}\, v_k' H' (r^n)^{-1} \left( z - H \hat{x}_0 \right) \tag{12}$$

La relation entre les estimées correspondant à deux séquences voisines peut donc être exprimée à l'aide des équations

$$\hat{x}_k = (a_k - v_k)\, \rho_k^{-1}\, v_k' H' (r^n)^{-1} \left( z - H \hat{x}_0 \right) \tag{13a}$$

$$\rho_k = \pm (r^x)^{-1} + v_k' A_0 v_k \tag{13b}$$

$$a_k = \Pi_0 A_0 v_k \tag{13c}$$

$$A_k = A_0 - A_0 v_k \rho^{-1} v_k' A_0 \tag{13d}$$

L'algorithme (13) a une structure simple qui rappelle fortement celle des algorithmes du type moindres carrés récurrents. Cependant, celui-ci ne fait pas intervenir explicitement les matrices de covariance d'erreur d'estimation $P_0$ et $P_k$. En effet, si $a_k$ peut être exprimé en fonction de $P_0$ , il n'en est pas de même pour $\rho_k$.. Cela traduit le caractère en général non inversible de $P_0$. On est donc contraint de remplacer une récurrence sur $P$ par une récurrence sur $A$, ce qui a des effets négatifs sur le volume des calculs: $P$ a une dimension effective ($N_e$ x $N_e$) alors que $A$ est de dimension ($N$ x $N$). L'algorithme (13) est néanmoins intéressant dans le contexte de la déconvolution itérative, car l'équation de Riccati (13d) qui sert à la mise à jour de $A$ n'est utilisée que lorsque l'on retient la séquence $t_k$ de préférence à $t_0$, ce qui se produit rarement, surtout si l'on démarre la procédure à proximité de la solution.

En ce qui concerne le critère, la technique la plus simple et la moins pénalisante d'un point de vue volume des calculs consiste à l'évaluer pour chaque valeur de $\hat{x}_k$ en utilisant (5).

*3.2. Procédure itérative.*

Comme indiqué plus haut, la stratégie adoptée pour l'exploration des séquences possibles consiste, en partant d'une séquence initiale, à essayer toutes les séquences voisines, à retenir celle qui maximise le critère correspondant et à recommencer jusqu'à ce que la convergence soit atteinte. La croissance du critère à chaque itération est ainsi garantie et comme le nombre de réalisations possibles de la séquence de Bernoulli est fini, la convergence en un nombre fini (peut-être très grand) d'itérations est elle aussi assurée. La structure de la procédure ainsi qu'une évaluation du nombre de multiplications réelles sont précisées ci-dessous. Les chiffres relatifs à la phase d'initialisation correspondent à une séquence $t_0$ nulle.

(1) INITIALISATION.

| | |
|---|---|
| Donnée de $t_0, z, H, r^x, r^n, \lambda$. | 0 mult. |
| Calcul de $\hat{x}_0$ | 0 mult. |
| Calcul de $A_0$ | $n^3$+1 mult. |
| Calcul de $z - H\hat{x}_0$ | 0 mult. |
| Calcul de $J_0$ | $N$+1 mult. |

(2) ITÉRATION — Pour $k$ variant de 1 à $N$:

| | |
|---|---|
| Modification de la valeur de $t(k)$ | 0 mult. |
| Calcul de $\hat{x}_k$ par les trois premières équations de (13) | |
|     Calcul de $a_k$ | $N_e$ mult. |
|     Calcul de $\rho_k$ | 1 mult. |
|     Calcul de $\hat{x}_k$ | $N+N_e+2$ mult. |
| Calcul de $z - H\hat{x}_k$ | $N\,N_e$ mult. |
| Calcul de $J_k$ | $N+N_e+5$ mult. |

(3) TEST DE CONVERGENCE.

| | |
|---|---|
| Calcul de $J_1 = \mathrm{Max}\{J_k\}$ | 0 mult. |
| Si $J_1 \leq J_0$ | |
|     CONVERGENCE | |
| Sinon | |
|     Mise à jour de $t_0$ | 0 mult. |
|     Mise à jour de $\hat{x}_0$ | $N+2N_e+5$ mult. |
|     Mise à jour de $A_0$ | $2N^2$ mult. |
|     Mise à jour de $z - H\hat{x}_0$ | $N\,N_e$ mult. |
|     Mise à jour de $J_0$ | 0 mult. |
|     Retour en (2) | 0 mult. |

On constate donc que le volume de calcul est tout à fait raisonnable et qu'il est compatible avec l'utilisation de moyens de calcul légers du type station de travail. Le facteur le plus pénalisant est la nécessité de stocker la matrice $A$. Cependant, les caractéristiques des stations de travail actuellement disponibles permettent le stockage en mémoire vive de toutes les grandeurs nécessaires au fonctionnement de la procédure.

## 4. Simulations.

De manière à évaluer l'efficacité de la procédure, des tests ont été effectués sur données synthétiques. Deux résultats de simulation sont présentés ici. Le premier correspond à un exemple relativement facile déjà traité par de nombreux auteurs [4, 8, 9], et permet donc une comparaison aisée avec d'autres méthodes existantes. L'ondelette, de contenu spectral relativement riche, est présentée figure 1. Le signal observé est présenté figure 2. Le rapport signal-à-bruit sur les observations, défini comme le rapport entre la puissance moyenne des observations et la variance du bruit, est égal à 10 dB. Sur la figure 3, les cercles représentent la réflectivité à restaurer et le résultat fourni par la méthode apparait en traits pleins. Les groupes d'impulsions très proches sont correctement restitués, ce qui indique un excellent fonctionnement de la méthode. En partant d'une séquence de Bernoulli identiquement nulle, moins de 20 itérations sont nécessaires pour obtenir le résultat.

Le deuxième exemple est représentatif des difficultés rencontrées en exploration géophysique. L'ondelette, qui correspond à une source sismique réelle, est représentée figure 4. Son contenu spectral est pauvre, ce qui apparait clairement sur le signal observé présenté figure 5. Ici aussi, le rapport signal-à-bruit est égal à 10 dB. Le résultat de la déconvolution est donné figure 6. Dans ces conditions plus difficiles, l'algorithme continue à se comporter de manière tout à fait satisfaisante tant du point de vue de la vitesse de convergence que de la qualité des résultats.

## 5. Conclusion.

Nous avons présenté une procédure itérative de déconvolution de processus Bernoulli-gaussiens qui apporte les améliorations suivantes aux méthodes existantes: *i*) L'algorithme de détection-estimation a une structure simple du type moindres carrés récurrents, ce qui permet un calcul exact du critère MAP et une mise en œuvre aisée. *ii*) Les étapes de détection et d'estimation sont effectuées simultanément. *iii*) Le volume des calculs est modéré. Ces caractéristiques sont dues essentiellement au modèle MA utilisé pour représenter l'ondelette. Comme l'indiquent les résultats de simulation, les performances de la méthode sont très satisfaisantes. L'utilisation de techniques stochastiques pour maximiser le critère permettrait certainement de les améliorer encore.

## 6. Références.


[1] R. Yarlagadda, J.B. Bednar and T.L. Watt, "Fast Algorithms for Lp Deconvolution," *IEEE Trans. Acoustics, Speech & Signal Processing*, ASSP-33, 174-182, 1985.

[2] R.A. Wiggins, "Minimum Entropy Deconvolution," *Proc. IEEE Intern. Symp. Computer-Aided Seismic Analysis & Discrimination*, Falmouth, MA, 7-14, 1977.

[3] Y. Goussard, G. Demoment and Y. Grenier, "Déconvolution de processus impulsionnels par algorithmes rapides: Approches AR et MA," *Proc. 11° colloque GRETSI*, Nice, France, 1987.

[4] J. Kormylo and J.M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Information Theory*, **IT-28**, 482-488, 1982.

[5] J.M. Mendel, "Optimal seismic deconvolution," Academic Press, New York, 1983.

[6] J. Goutsias and J.M. Mendel, "Maximum-likelihood deconvolution: An optimization theory perspective," *Geophysics*, **51**, 1206-1220, 1986.

[7] A.K. Mahalanabis, S. Prasad and K.P. Mohandas, "Recursive decision-directed estimation of reflection coefficients for seismic data deconvolution," *Automatica*, **18**, 721-726, 1982.

[8] S.D. Kollias and C.C. Halkias, "An instrumental variable approach to minimum-variance seismic deconvolution," *IEEE Trans. Geoscience & Remote Sensing*, **GE-23**, 778-788, 1985.

[9] Y. Goussard and G. Demoment, "Recursive deconvolution of Bernoulli-Gaussian processes using a MA representation," *To appear in IEEE Trans. Geophysics & Remote Sensing*, **GE-27**, 1989.

[10] F. Monfront, "Déconvolution de signaux Bernoulli-gaussiens par maximum de vraisemblance," Rapport de stage de DEA, Université de Paris-Sud, Orsay, 1988.

[11] Y. Goussard, "Déconvolution de processus Bernoulli-gaussiens. Un algorithme itératif.," *Rapport interne LSS # GPI - 89/02*, 1989.


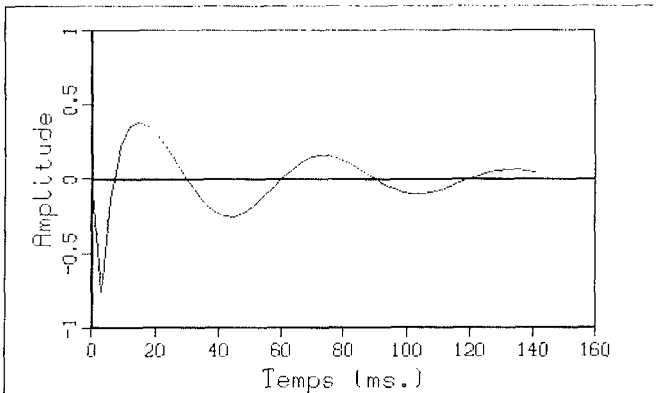

FIGURE 1: ONDELETTE

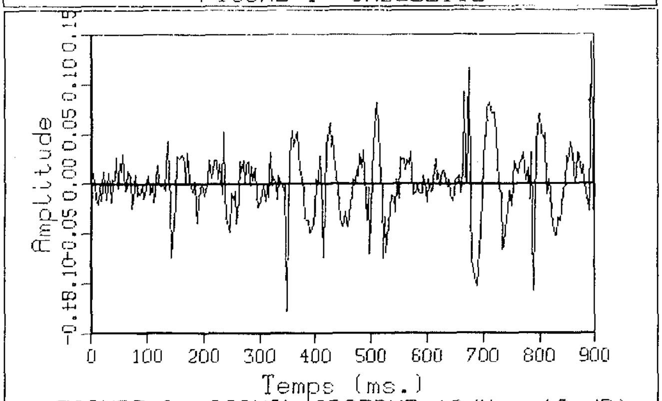

FIGURE 2: SIGNAL OBSERVE (S/N = 10 dB)

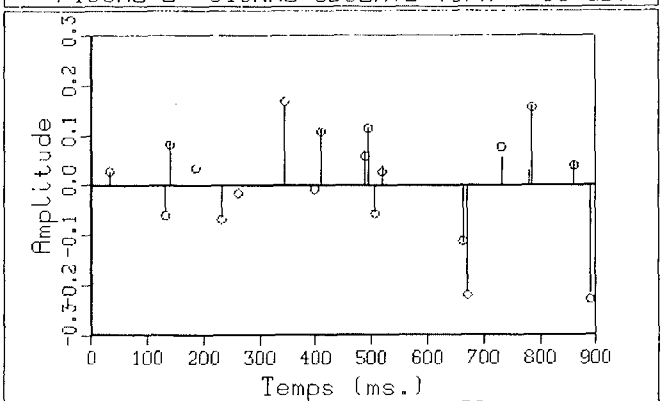

FIGURE 3: RESULTATS

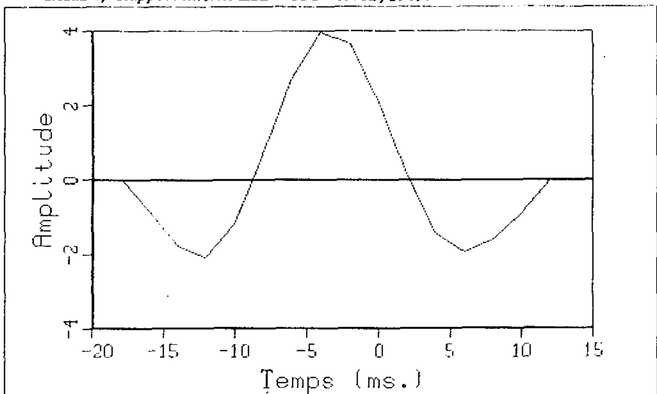

FIGURE 4: ONDELETTE

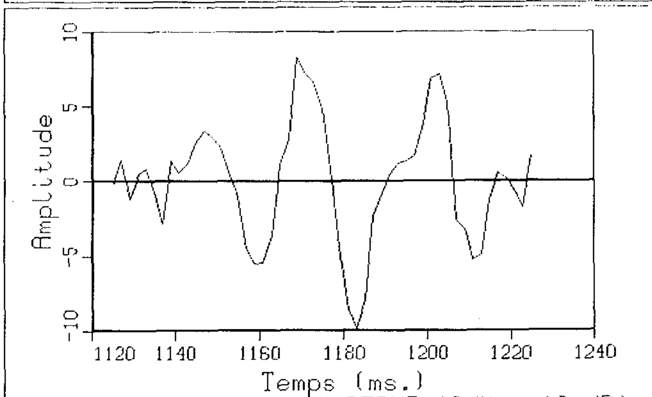

FIGURE 5: SIGNAL OBSERVE (S/N = 10 dB)

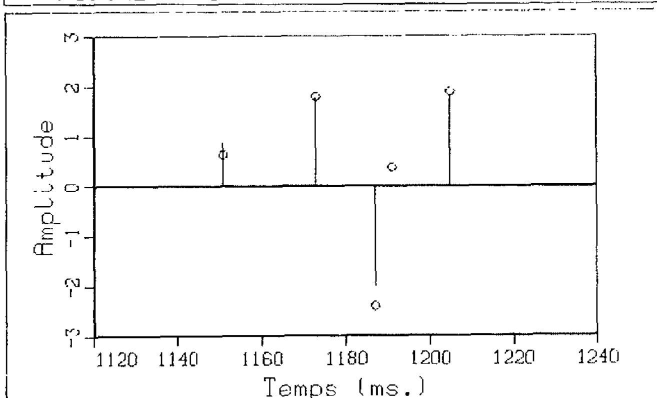

FIGURE 6: RESULTATS