INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2002-22

Déconvolution Aveugle Bruitée :
Estimation de la Distribution
du Processus Source

E. GAUTHERAT[1]

Juillet 2002

[1] CREST-Laboratoire de Statistique et Université de Reims, email : gauthera@ensae.fr

Déconvolution aveugle bruitée :
estimation de la distribution du processus source

Emmanuelle Gautherat *

July 11, 2002

Résumé :

Cette étude présente des résultats sur l'estimation de la distribution dans le cadre d'un problème de déconvolution aveugle discrète bruitée dont le processus source est de loi discrète finie et inconnue.
Le travail présent établit tout d'abord une version unifiée de plusieurs fonctions de contraste initialement proposées dans le cadre des mélanges de populations ou encore des modèles de convolution.
L'intêret principal de cette étude consiste en la construction d'estimateurs des points de support du processus discret source et des masses les chargeant. On établit la consistance dans le cadre non paramétrique et la vitesse dans le cadre paramétrique de ces estimateurs. Ces résultats sont obtenus sans hypothèse d'indépendance sur le processus.
L'expression de la distribution asymptotique de ces estimateurs met en lumière le rôle du filtre inverse, du nombre de points de support et du choix de la fonction de contraste utilisée.

This paper presents some results on the estimation of the distribution of the signal input in a noisy blind deconvolution problem with a unknown level noise. We focus on the case that the input sequence is drawn from a distribution with finite support (digital deconvolution).
This work unifies different studies using different contrast functions in deconvolution models and mixture population models. But the main interest of this work consists in estimating the finite distribution of the input signal : support and mass. These estimators are strongly consistent in a non-parametric setting. In a parametric context, we obtain the exact asymptotic distribution. These results are obtained also with a dependent structure.
The study of asymptotic distribution gives the precise contribution of the inverse filter, of the cardinality of the support and of the choice of the contrast function considered in the analysis.

*Laboratoire de Statistique du CREST, Université de Reims, e-mail : emmanuelle.gautherat @ensae.fr

1 Introduction

On appelle modèle de déconvolution aveugle, un système où l'on observe le processus $(W_t)_{t=1,\dots,n}$ résultant de la filtration d'un processus entrant $(X_k)_{k\in\mathbb{Z}}$, ou processus source, par un filtre déterministe inconnu $(u_t)_{t\in\mathbb{Z}}$

$$\forall t = 1,\dots,n, \quad W_t = \sum_{k\in\mathbb{Z}} u_k X_{t-k} = (u \star X)_t$$

On s'intéresse ici aux processus sources $(X_k)_{k\in\mathbb{Z}}$ digitaux, c'est-à-dire de distribution discrète. Une série de travaux se focalise sur le cadre plus spécifique des signaux digitaux à support fini. Ces études se scindent en deux groupes. Le premier considère le problème comme un modèle bayésien et s'attache à développer des algorithmes de Monte Carlo afin de calculer les distributions *a posteriori* du signal. On pourra consulter Liu et Chen [LC95] ou encore Chen et Li [CL95]. Le second groupe utilise l'application d'un filtre inverse à la série observée $(W_t)_{t=1,\dots,n}$. Dans cette voie, on peut lire Li [Li95], Gamboa et Gassiat [GG96], Gassiat et Gautherat [GG98] [GG99]. Le travail effectué par Li et Shedden [LS01] prend une place particulière puisqu'il exploite simultanément les deux approches. Le travail présent se situe dans la seconde voie que l'on explicite à présent. Gamboa et Gassiat [GG96] proposent une méthode d'estimation du filtre inverse d'un modèle de déconvolution lorsque le signal source prend ses valeurs dans un alphabet de taille fini, inconnu, dont le cardinal est connu. Cette méthode fut ensuite généralisée par Gassiat et Gautherat [GG98] à l'observation de sorties bruitées additivement par un bruit inobservable $\sigma_0\epsilon$ de niveau σ_0 inconnu. On observe alors

$$\forall t = 1,\dots,n, \quad W_t = (u \star X)_t + \sigma_0 \epsilon_t.$$

Ils proposent un estimateur du filtre et du niveau de bruit. Cette méthode s'applique théoriquement quelque soit le rapport signal sur bruit. Les résultats numériques présentés confirment la théorie. Ces estimateurs semblaient avoir un comportement asymptotiquement meilleur que les autres estimateurs proposés pour ce type de problème dans le sens où ils utilisent *a priori* le caractère discret du signal et de la présence d'un bruit additif. Ces mêmes auteurs présentent dans [GG99] des résultats théoriques de vitesse de convergence du filtre dans le cadre paramétrique (en présence et en absence de bruit). Ces résultats corroborent la conjecture précédement énoncée : on obtient une vitesse plus rapide que les estimateurs n'exploitant pas le caractère discret de la distribution du processus $(X_k)_{k\in\mathbb{Z}}$. La vitesse dans le cadre non bruité peut atteindre une vitesse exponentielle. Dans le cadre non bruité paramétrique, avec un alphabet connu, mais avec un processus entrant éventuellement non stationnaire, Li [Li95] propose également un estimateur du filtre inverse. Tout comme Gassiat et Gautherat dans [GG99], la vitesse de convergence obtenue a une borne supérieure en norme l_1 tronquée du filtre inverse inconnu. Mais lorsque Li teste numériquement la robustesse de cette méthode à la présence de bruit, les résultats se dégradent alors pour un rapport signal sur bruit faible.

Les résultats obtenus jusqu'à présent dans le cadre du modèle de déconvolution bruité par l'application d'un filtre inverse ne concernent que l'identification du filtre inverse inconnu et éventuellement du niveau de bruit. Une méthode de déconvolution complète, c'est-à-dire permettant de restituer le signal entrant $(X_k)_{k\in\mathbb{Z}}$, n'a pas encore été proposée dans la démarche autilisant l'application d'un filtre inverse. Ceci contrairement à l'approche bayésienne qui, elle, travaille sur la restitution effective du signal $(X_k)_{k\in\mathbb{Z}}$ mais bute sur une bonne estimation de la loi du signal $(X_k)_{k\in\mathbb{Z}}$.
L'objet de cet article est de travailler en ce sens. On propose des estimateurs de la distribution du processus entrant ainsi que la loi asymptotique explicite de ces estimateurs. D'une part, dans certaines applications ces résultats suffisent, la déconvolution complète du modèle n'étant pas nécessaire. D'autre part, ces résultats permettront de mettre en oeuvre simplement les méthodes

bayésiennes afin de retrouver le processus $(X_k)_{k\in BBz}$.

L'article est construit de la manière suivante.
Dans le premier paragraphe, on exposera brièvement la méthode de déconvolution aveugle dans le cadre d'un modèle bruité additivement initialement proposée par Gassiat et Gautherat [GG98], dans une version unifiée. On convie le lecteur interessé par les preuves de cette partie à se rapporter à Gautherat [Gau97]. On donne également les résultats de vitesse obtenus dans [GG99] utiles dans la suite de l'article. Les adaptations nécessaires à la présentation unifiée des diférentes fonctions de contraste ne nécessitant pas de preuve substantiellement différente de celles de [GG99], nous ne les présenterons pas.
Dans le second paragraphe, on expose une méthode d'estimation de la loi du processus source $(X_k)_{k\in\mathbb{Z}}$ prenant en compte le caractère dicret et fini de la distribution. On établit la consistance et la distribution asymptotique de ces estimateurs.
Le dernier paragraphe collecte l'ensemble des preuves.

Commençons par préciser le modèle. Soit une suite d'observations réelles $(W_t)_{t\in\mathbb{Z}}$ représentant le signal de sortie d'un système linéaire invariant dans le temps, $\mathcal{U}$, de réponse impulsionnelle $(u_k)_{k\in\mathbb{Z}}$ dans lequel passe le signal inobservable entrant $X=(X_k)_{k\in\mathbb{Z}}$, le tout perturbé par un bruit additif $(\sigma_0\epsilon_k)_{k\in\mathbb{Z}}$ de variance inconnue σ_0^2.

$$\forall t\in\mathbb{Z},\quad W_t=\sum_{j\in\mathbb{Z}} u_j X_{t-j}+\sigma_0\epsilon_t.$$

On applique un système linéaire invariant dans le temps $\mathcal{S}$: $s=(s_k)_{k\in\mathbb{Z}}$ à la sortie $(W_k)_{k\in\mathbb{Z}}$. On travaille à présent sur le processus discret $(W_t(s))_{t\in\mathbb{Z}}$:

$$\forall t\in\mathbb{Z},\quad W_t(s)=\sum_{j\in\mathbb{Z}} s_j W_{t-j}=(s\star u\star X)_t+\sigma_0(s\star\epsilon)_t. \tag{1}$$

Les hypothèses générales sur le modèle sont les suivantes :

- **(M1)** *$(X_k)_{k\in\mathbb{Z}}$ est un processus discret, à valeurs réelles, de loi discrète de support $\{a_1<\ldots<a_p\}$ inconnu, de cardinal $p\geq 2$ connu. Soit $(\pi_i)_{i=1,\ldots,p}$ les masses chargeant $(a_i)_{i=1,\ldots,p}$: $P(X_k=a_i)=\pi_i$ avec $\sum_{i=1}^p \pi_i=1$ et $\forall 1\leq i\leq p,\quad 0<\pi_i<1$.*
- **(M2)** *Soit $U(x):=\sum_{k\in\mathbb{Z}} u_k e^{ikx}$. U est une fonction continue qui ne s'annule pas sur $[0,2\pi[$. $\theta=(\theta_k)_{k\in\mathbb{Z}}$ est le filtre inverse de u. C'est à dire :*

$$(\theta\star u)_k=\sum_{j\in\mathbb{Z}}\theta_j\cdot u_{k-j}=\delta_k,\ k\in\mathbb{Z}$$

  où δ_k désigne le symbole de Kronecker.
- **(M3)** *X est un processus ergodique et stationnaire.*
- **(M4)** *Pour tout entier n et pour tout entier $j_1,\ldots,j_n$ dans $\{1,\ldots,p\}$,*

$$P(X_1=a_{j_1},\ldots,X_n=a_{j_n})\ >\ 0.$$

- **(M5)** *$\epsilon=(\epsilon_k)_{k\in\mathbb{Z}}$ est une suite de variables aléatoires gaussiennes centrées réduites indépendantes entre elles et indépendantes du signal source X; σ_0 est un scalaire réel positif inconnu.*

(M1) permet de considérer toutes les configurations de p points de support. En particulier aucune contrainte, du type treillis par exemple, n'est ici imposée.
(M2) garantit l'existence du filtre inverse θ de u.
(M3) garantit que toute espérance de fonctions intégrables du processus X peut être approximée par sa moyenne empirique. En particulier, on ne demande pas que les (X_k) soient indépendament distribués.
(M4) permet de ne pas surestimer le modèle en considérant que tous les points du support sont chargés par l'ensemble des trajectoires.

On donne ici quelques exemples de processus pour lesquels les hypothèses sont vérifiées :

- Bruit blanc: lorsque les variables X_k, $k \in \mathbb{Z}$, sont indépendament et identiquement distribuées, les hypothèses **(M1), (M3), (M4)** sont vérifiées.
- Processus seuillé : soit $(V_t)_{t\in\mathbb{Z}}$ un processus stationnaire tel que la distribution de toutes ses marginales finies soit continue (par exemple, un processus Gaussien). Soit $m_1 < m_2 \cdots < m_{p-1}$ des nombres réels, avec $m_0 = -\infty$, $m_p = +\infty$. On définit le processus seuillé $(X_k)_{k\in\mathbb{Z}}$ par : $X_k = x_j$ si et seulement si $V_k \in]m_j, m_{j+1}]$. Alors les hypothèses **(M1), (M3), (M4)** sont vérifiées.
- Chaîne de Markov récurrente apériodique : soit $(X_k)_{k\in\mathbb{Z}}$ une chaine de Markov d'espace d'état de cardinal p et une matrice de transition telle que toutes les probabilités de transition soient positives, alors **(M1)** et **(M4)** sont vraies. La chaîne étant apériodique et récurrente, **(M3)** est aussi vérifiée.

La distributivité gaussienne du bruit ϵ a été ainsi choisie pour des raisons de clarté d'écriture. Cependant, tous les résultats de cet article sont également valables lorsque le bruit est de la forme $\sigma_0\eta_k$, où η_k suit une distribution infiniment divisible de classe L de moments connus. On pourra consulter l'ouvrage de Petrov [Pet75] à ce sujet.
On a ainsi pour $s = \theta$,

$$\forall t \in \mathbb{Z}, \quad W_t(\theta) = X_t + \sigma_0(\theta \star \epsilon)_t.$$

Pour $\sigma_0 = 0$, dès que θ est connu, on a résolu le problème. En revanche, en présence d'un bruit additif ($\sigma_0 \neq 0$) les techniques sont totalement différentes selon que le niveau de bruit σ_0 est connu ou non. On se place ici dans le cadre d'un niveau de bruit σ_0 inconnu.

2 Procédure d'estimation.

On donne tout d'abord un premier résultat probabiliste permettant de mesurer la concentration d'une distribution Q sur p points de support inconnus.

Théorème 2.1
Soit ϕ_1 une fonction de $\mathbb{R}$ dans $\mathbb{C}$, soit Q une mesure de probabilité sur $\mathbb{R}$, soit d_Q un vecteur de complexes de taille $(p+1)^2$ tel que $\forall k, j \in \{0, \ldots, p\}$

$$d_Q^{k,j} = \int \phi_1^k(v)\overline{\phi_1^j(v)}dQ(v)$$

Soit $D(d_Q)$ la matrice définie par

$$D(d_Q) = (d_Q^{k,j})_{k,j=0,\ldots,p} = \begin{pmatrix} d_Q^{0,0} & d_Q^{0,1} & \cdots & d_Q^{0,p} \\ d_Q^{1,0} & d_Q^{1,1} & \cdots & d_Q^{1,p} \\ \vdots & \vdots & & \vdots \\ d_Q^{p,0} & d_Q^{p,1} & \cdots & d_Q^{p,p} \end{pmatrix}$$

Enfin, soient $V_0, \ldots, V_p$, $p+1$ variables aléatoires indépendantes de loi Q.
Soit $h(d_Q) = det(D(d_Q))$.
On a alors la représentation suivante:

$$h(d_Q) = \frac{1}{(p+1)!} E\left(\prod_{0 \le j < k \le p} |\phi_1(V_j) - \phi_1(V_k)|^2 \right)$$

On notera indifféremment d_Q et d_V lorsque V est de loi Q.
On tire immédiatement de cette représentation le corollaire suivant.

Corollaire 2.1

$$h(d_Q) \geq 0$$

De plus, si l'on suppose ϕ_1 injective de $\mathbb{R}$ dans $\mathbb{C}$,
$h(d_Q) = 0 \iff Q$ est une mesure de probabilité sur $\mathbb{R}$ discrète avec au plus p points de support.

Exemples :

1. Soit ϕ_1 la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par $\phi_1(z) = z$. On obtient pour $h(d_Q)$ la forme de Hankel.
2. Soit ϕ_1 la fonction de $\mathbb{R}$ dans $\mathbb{C}$ définie par $\phi_1(x) = \exp i2\pi\psi(x)$ où ψ est une fonction de $\mathbb{R}$ dans $]0,1[$ injective. La fonction $h(d_Q)$ est la forme de Toeplitz.
3. On peut considérer par exemple $\phi_1(x) = \exp x$ pour x appartenant à $\mathbb{R}$ ou encore $\phi_1(x) = arctg(x)$.

Remarques :

1. En notant $\phi_k = \phi_1^k$, de telles fonctions ϕ_k forment un système de Tchebychev (T-système). En effet

$$\sum_{k=0}^{p} \phi_k(x) = 0 \iff \sum_{k=0}^{p} \phi_1^k(x) = 0$$

$$\iff \text{on a au plus } p \text{ racines du polynôme en } \phi_1(x) \ : \ \sum_{k=0}^{p} \phi_1^k(x)$$

   De plus ϕ_1 est injective, on a donc au plus p racines x du système.

2. On ne peut considérer ce corollaire comme une conséquence du théorème caractérisant la fonction de contraste exposée dans Gassiat et Gautherat [GG98] utilisant directement les propriétés des T-systèmes. En effet, il n'est pas nécessaire ici de se ramener sur un ensemble inclus dans un compact pour appliquer les propriétés des T-systèmes. Ce point est essentiel, car il permet de travailler avec des signaux dont on ne connait pas *a priori* l'étendue du support.
   De plus l'injectivité de ϕ_1 suffit.
3. Néanmoins, cette représentation ne prend en compte qu'un certain type de fonctions de contraste. Par exemple, elle ne permet pas de traiter le cas de la fonction de contraste du maximum d'entropie sur la moyenne (voir Gamboa et Gassiat [GG96]).

On utilise à présent la divisibilité de la loi gaussienne. Plus particulièrement on utilise la propriété suivante, pour tout $\epsilon \sim \mathcal{N}(0,1)$,

$$\exists \eta, \tilde{\eta} \sim \mathcal{N}(0,1), \ \ \forall 0 \leq \beta, \forall 0 \leq \alpha \leq \beta, \quad \beta\epsilon = \sqrt{\beta^2 - \alpha^2}\, \tilde{\eta} + \alpha\, \eta$$

où $\tilde{\eta}$ et η sont indépendantes entre elles. On obtient ainsi

$$\exists \eta, \tilde{\eta} \sim \mathcal{N}(0,1), \ \ \forall 0 \leq \sigma \leq \sigma_0, \quad W_t(s) = (s \star u \star X)_t + \sqrt{\sigma_0^2 - \sigma^2}(s \star \tilde{\eta})_t + \sigma(s \star \eta)_t$$

On notera pour toute la suite $Z_t(s) = (s \star u \star X)_t + \sqrt{\sigma_0^2 - \sigma^2}(s \star \tilde{\eta})_t$.
On remarque deux points essentiels :

- pour $\sigma = \sigma_0$ et $s = \theta, \quad Z_k(\theta) = X_k$.
- pour tout $s \neq \theta$, pour tout $\sigma \leq \sigma_0$, $Z_t(s)$ a une distribution qui charge plus de p points de support. La preuve de ce résultat peut se consulter dans [GG96].

On construit maintenant $d_Z(s,\sigma)$ sur la loi Q de $Z_0(s)$. D'après les deux remarques précédentes on a pour tout $\sigma \leq \sigma_0$,

$$h(d_Z(s,\sigma)) = 0 \iff \sigma = \sigma_0 \text{ et } s = \theta.$$

On veut construire d_Z à partir de d_W où d_W est construit sur la loi de $W_0(s)$. Pour ce faire, on demande à la fonction ϕ_1, en plus d'être injective, de vérifier la propriété suivante

$$\exists \varphi_1 \exists \varphi_2 \forall x, y \in \mathbb{C}, \quad \phi_1(x+y) = \varphi_1(x) \odot \varphi_1(y)$$

où $\odot$ représente l'addition ou la multiplication et φ_1 et φ_2 sont deux fonctions injectives continuement dérivables telles que $E(\varphi_2(\epsilon)^k \overline{\varphi_2(\epsilon)}^j) < +\infty$ pour tout $k, j \leq p$ avec $\epsilon \sim \mathcal{N}(0, \sigma^2 \|s\|_2^2)$.
Le vecteur d_W s'exprime alors comme suit :

$$\begin{aligned}
\forall s \in l_1(\mathbb{Z}),\ \forall \sigma \leq \sigma_0, \quad (d_W(s))_{k,j} &= E\left(\phi_1(W_0(s))^k \overline{\phi_1(W_0(s))}^j\right) \\
&= E\left(\left(\ \varphi_1(Z_0(s)) \odot \varphi_2(\sigma(s \star \eta)_0)\ \right)^k \overline{\varphi_1(Z_0(s)) \odot \varphi_2(\sigma(s \star \eta)_0)}^j\right).
\end{aligned}$$

De plus, on remarque que les moments conjugués de la fonction φ_2 du bruit filtré par s sont connus car $\sigma(s \star \eta)$ est une variable gaussienne centrée de variance $\sigma^2 \|s\|_2^2$. C'est-à-dire que les coefficients

$$\gamma_{k,j}(s,\sigma) = E\left(\left(\ \varphi_2(\sigma(s \star \epsilon)_1)\ \right)^k \overline{\varphi_2(\sigma(s \star \epsilon)_1)}^j\right)$$

sont connus. Comme φ_2 est continuement dérivable en s et σ, $\gamma_{k,l}(s,\sigma)$ l'est également.

1. **Inversion dans le cas de l'addition.**
   A l'aide de développements du binôme, il est aisé de calculer les moments conjugués d_Z de la variable $\varphi_1(Z_0(s))$ à partir des moments conjugués d_W de $\phi_1(W_0(s))$ et $\gamma_{k,l}(s,\sigma)$. On procède en inversant le système triangulaire suivant :

$$\begin{aligned}
\forall s \in l_1(\mathbb{Z}), \forall \sigma \leq \sigma_0, \quad d_W(s) &= A_{\varphi_2}(\sigma^2 \|s\|_2^2) \left(E(\varphi_1(Z_0(s))^k \overline{\varphi_1(Z_0(s))}^j)\right)_{k,j=0,\ldots,p} \\
&= A_{\varphi_2}(\sigma^2 \|s\|_2^2)\ d_Z(s,\sigma)
\end{aligned}$$

   avec pour tout filtre s et tout niveau de bruit $\sigma \leq \sigma_0$, $A_{\varphi_2}(\sigma^2 \|s\|_2^2)$ une matrice triangulaire inférieure, de taille $(p+1)^2 \times (p+1)^2$, de termes diagonaux égaux à 1, dont les termes non nuls sont des combinaisons linéaires de $\gamma_{k,l}(s,\sigma)$. Ainsi $A_{\varphi_2}(\sigma^2 \|s\|_2^2)$ est différentiable, de différentielle continue par rapport à s et σ.

2. **Inversion dans le cas de la multiplication.**
   On peut encore plus aisément calculer les moments conjugués d_Z de la variable $\varphi_1(Z_0(s))$,

à partir des moments conjugués de $\phi_1(W_0(s))$, d_W, et $\gamma_{k,l}(s,\sigma)$. On procède en inversant le système diagonal suivant :

$$\begin{aligned}\forall s \in l_1(\mathbb{Z}),\ \forall \sigma \leq \sigma_0, \quad d_W(s) &= B_{\varphi_2}(\sigma^2\|s\|_2^2)\left(E\left(\varphi_1(Z_0(s))^k\overline{\varphi_1(Z_0(s))}^j\right)\right)_{k,j=0,\ldots,p}\\ &= B_{\varphi_2}(\sigma^2\|s\|_2^2)\ d_Z(s,\sigma)\end{aligned}$$

où $B_{\varphi_2}(\sigma^2\|s\|_2^2)$ est une matrice carré diagonale non nulle de taille $\left((p+1)^2\right)^2$ dérivable de dérivée continue par rapport à s et σ.

On note $\Delta_{\varphi_2}(\sigma^2\|s\|_2^2)$ la matrice permettant d'inverser le système considéré ($\Delta_{\varphi_2} = A_{\varphi_2}$ dans le cas de l'addition et $\Delta_{\varphi_2} = B_{\varphi_2}$ dans le cas de la multiplication). On remarque que l'on peut toujours définir un vecteur d pour tout $\sigma \in \mathbb{R}+$ par inversion du système

$$\forall s \in l_1(\mathbb{Z}), \forall \sigma \geq 0, \quad d_W(s) = \Delta_{\varphi_2}(\sigma^2\|s\|_2^2)\ d(s,\sigma)$$

où encore $d_W = \Delta_{\varphi_2} d$. Mais pour $\sigma > \sigma_0$, d ne sera pas forcément le moment conjugué d'une variable aléatoire. Aussi appelle t-on d le vecteur des pseudo-moments conjugués et on omet d'indicer par une probabilité ce vecteur.

Jusqu'à présent, on a établi une expression de d en fonction de d_W et Δ_{φ_2}. Il s'agit maintenant de travailler avec les observations $W_1, \ldots, W_n$ d'une part. D'autre part, le filtre u et par là même, le filtre inverse θ de u étant non-paramétrique **(M2)**, il va être nécessaire de tronquer cette suite à l'aide d'un paramètre de troncature $m(n)$, suite croissante d'entiers vérifiant

$$\lim_{n\to+\infty} m(n) = +\infty, \quad \text{et} \quad \lim_{n\to+\infty} \frac{m(n)}{n} = 0 \tag{2}$$

pour estimer le filtre inverse θ.
On définit maintenant tous les objets précédement présentés dans leur version empirique, en prenant soin de n'utiliser que les observations $W_1, \ldots, W_n$ et en gardant l'expression exacte (non-empirique) de Δ_{φ_2} en vertu de la connaissance *a priori* de la nature de la distribution du bruit inobservable ϵ. La filtration de W par s devient donc

$$\forall s \in l_1(\mathbb{Z}),\ \forall t \in \{m(n)+1, \ldots, n-m(n)\} \quad \widehat{W}_t(s) = \sum_{k=-m(n)}^{+m(n)} s_k \cdot W_{t-k}. \tag{3}$$

On définit les estimateurs $\widehat{d}_W$ des moments conjugués de $\phi_1(W_0(s))$ par les moments empiriques des variables tronquées :

$$\left(\widehat{d}_W(s)\right)_{k,j} = \frac{1}{n-2m(n)} \sum_{t=1+m(n)}^{n-m(n)} \phi_1(\widehat{W}_t(s))^k \overline{\phi_1(\widehat{W}_t(s)}^j.$$

On notera $Empi_t$ l'espérance empirique ainsi définie :

$$Empi_t = \frac{1}{n-2m(n)} \sum_{t=1+m(n)}^{n-m(n)}$$

Les estimateurs $\widehat{d}_{k,j}$ des pseudo-moments conjugués $d_{k,j}$ s'expriment comme solutions du système triangulaire :

$$\forall s \in l_1(\mathbb{Z})\ \forall \sigma \geq 0, \quad \widehat{d}_W(s) = \Delta_{\varphi_2}(\sigma^2\|s\|_2^2)\ \widehat{d}(s,\sigma)$$

On définit la fonction de contraste empirique $H_n(s,\sigma)$ exprimant la concentration sur p points de support inconnus des estimateurs des pseudo-moments conjugués $\widehat{d}(s,\sigma)$.

$$\forall s \in l_1(\mathbb{Z}), \forall \sigma > 0, \quad H_n(s,\sigma) = h\left(\widehat{d}(s,\sigma)\right). \tag{4}$$

Ne connaissant ni le comportement, ni le signe de H_n pour $\sigma > \sigma_0$, on considère le carré de H_n d'une part et on introduit un terme de pénalisation $\delta(n)$ du niveau de bruit σ d'autre part. La fonction $J_n(s,\sigma)$ ci-après définie va ainsi prendre ces arguments minimum en les plus petites valeurs de σ.
Soit $\delta(n)$ une suite réelle strictement positive de limite nulle à l'infini.

$$J_n(s,\sigma) = H_n^2(s,\sigma) + \delta(n)^2\,\sigma \tag{5}$$

Dans le but d'établir un résultat de consistance de l'estimateur du filtre inverse θ, on va restreindre l'ensemble dans lequel ce filtre peut évoluer à un compact Θ de $l_1(\mathbb{Z})$. De plus, afin d'assurer l'unicité de la concentration sur p points de support du processus $(\theta \star u \star X)_t$, on fixe dans Θ l'échelle et le décalage dans le temps du filtre inverse θ. C'est-à-dire

$$\forall \theta \in \Theta, \quad (\lambda\theta \star u)_{t-K} = \delta_0 \Longrightarrow K = 0, \ \lambda = 1.$$

Soit maintenant Θ un ensemble compact de $l_1(\mathbb{Z})$ dans lequel l'échelle et le décalage dans le temps sont fixés. On définit Θ_n l'intersection de Θ et des filtres tronqués :

$$\Theta_n = \{(s_k)_{k\in\mathbb{Z}} \in \Theta \ : \ s_k = 0, \ \forall |k| > m(n)\}$$

On définit les estimateurs $(\widehat{\theta},\widehat{\sigma})$ du filtre inverse θ et du niveau de bruit σ_0 par les arguments minimum de J_n sur $\Theta_n \times \mathbb{R}^+$.
On introduit deux hypothèses supplémentaires afin d'obtenir la convergence presque sûre des estimateurs,

- **(M6) :** *Supposons que* $\forall k, j = 1, \ldots, p$,

$$\lim_{n\to\infty} \frac{1}{n\ \delta(n)} \sum_{t=1}^{n} \phi_1(W_t(\theta))^k \overline{\phi_1(W_t(\theta))}^j - E\left(\phi_1(W_0(\theta))^k \overline{\phi_1(W_0(\theta))}^j\right) = 0$$

  presque sûrement.

- **(M7)** *On suppose que*

$$\left(\sum_{|k|>k(n)} |\theta_k|\right)^{\frac{3}{4}} = o(\delta(n))$$

  et que Θ est un sous-ensemble fermé de l'ensemble compact $\{s \in l_1 \ : \ |s_k| \le T\ |k|^{-\alpha}, \ \alpha > 1\}$, où T est un réel strictement positif.

- **(M8)** *Soit ϕ_1 injective de $\mathbb{R}$ dans $\mathbb{C}$ telle que $\phi_1(0) = 0$ dans le cas où $\odot$ représente l'addition et $\phi_1(0) = 1$ dans le cas où $\odot$ représente la multiplication.*

Théorème 2.2
On suppose vérifiées les hypothèses **(M1)** *à* **(M8)**. *Soit $m(n)$ telle que $\lim_{n\to+\infty} \frac{m(n)}{n\delta(n)} = 0$. Alors $\widehat{\theta}$ converge presque sûrement dans l_1 vers θ, et $\widehat{\sigma}$ converge presque sûrement vers σ_0.*

On trouve la preuve de ce résultat dans [GG98] lorsque Φ_1 est l'identité et $\odot$ est l'addition. L'extension au cas $\odot$ multiplicatif n'est pas significativement différente, aussi on ne présentera pas la preuve ici.
Remarque :

- Tant que $\delta(n)^{-1} = o\left(\sqrt{\frac{n}{\log\log n}}\right)$, l'expression $\phi_1(W_t(\theta))^k\overline{\phi_1(W_t(\theta))}^j - E\left(\phi_1(W_0(\theta))^k\overline{\phi_1(W_0(\theta))}^j\right)$ obéit à la loi du log-itéré. Des conditions suffisantes peuvent être trouvées dans [Dou94] ou [HH80].

- On peut alléger l'hypothèse **(M7)** en ne demandant plus que la décroissance de la queue du filtre inverse en $o(\delta(n))$. On obtient alors le théorème (2.2) avec une convergence en probabilité.

3 Restitution de la loi du processus X.

L'objet principal de cet article est la restitution de la distribution du processus émis X. En effet, nous avons jusqu'à présent identifié et estimé le filtre inverse θ de u et le niveau de bruit σ_0 . Or, en présence de bruit ($\sigma_0 \neq 0$), l'identification du filtre inverse θ et du niveau de bruit σ_0 ne permettent pas de restituer le signal $(X_k)_{k\in\mathbb{Z}}$. En effet, les observations filtrées par le filtre inverse, $(W_t(\theta))_{t\in\mathbb{Z}}$, sont des observations bruitées du processus $(X_k)_{k\in\mathbb{Z}}$:

$$\forall t \in \mathbb{Z}, \quad W_t(\theta) = X_t + \sigma_0(\epsilon \star \theta)_t.$$

Afin de tendre vers un résultat de déconvolution totale (restitution des $(X_k)_{k\in\mathbb{Z}}$), on estime la loi du processus $(X_k)_{k\in\mathbb{Z}}$. Il s'agit d'estimer les points de support $a_1, \ldots, a_p$ ainsi que les masses $\pi_1, \ldots, \pi_p$ les chargeant. Il serait théoriquement possible d'utiliser des estimateurs du maximum de vraisemblance des points du support $(a_i)_{1\leq i\leq p}$ et des masses les chargeant $(\pi_i)_{1\leq i\leq p}$. En effet, ayant ici choisi la distribution du bruit ϵ de nature gaussienne, nous avons

$$W_t(\theta) \sim \sum_{j=1}^{p} \pi_j \mathcal{N}(a_j, \sigma_0^2\|\theta\|_2^2)$$

où les $W_t(\theta)$ sont des variables aléatoires dépendantes, de structure de covariance connue (dépendant de θ) lorsque les X_k sont indépendemment distribués. Dans le cas contraire, la nature de la distribution du vecteur $W(\theta) = (W_t(\theta))_{t\in\mathbb{Z}}$ n'est pas accessible, seule ses marges le sont. Cependant, en se cantonnant au cadre plus restrictif où les X_k sont i.i.d., sous réserve que la structure de dépendance des $W_t(\theta)$ ne soit pas trop forte - par exemple, $\theta_k = 0, \quad \forall |k| \geq K$ - on peut utiliser des estimateurs par blocs de taille $2K+1$ du modèle de mélange sachant que grâce à l'hypothèse (**M4**) le modèle n'est ni sous-estimé, ni surestimé.
Mais, outre les restrictions que cette approche impose au modèle en matière de nature du filtre et de structure sur la loi du signal source, la complexité algorithmique générée par le calcul du maximum de vraisemblance se situe très rapidement au delà des limites des capacités calculatoires usuelles. En revanche, on peut alors faire appel à des techniques bayésiennes mais celles-ci vont se heurter au problème de la génération d'une loi π^* suffisament proche de celle de X (voir Li et Shedden [LS01]. On choisit ici d'utiliser directement la nature p-concentrée de la distribution du processus $(X_k)_{k\in\mathbb{Z}}$, sans autre restriction sur le modèle que les hypothèses préalablement posées. On établit des estimations des points de support et des masses. Ces estimateurs sont directement fonction des estimations du filtre inverse θ, du niveau de bruit σ_0 et des pseudo-moments conjugués $d(\theta, \sigma_0)$.
On commence par établir une caractérisation algébrique des points de support $(a_i)_{1\leq i\leq p}$ et des masses $(\pi_i)_{1\leq i\leq p}$. Basée sur cette caractérisation, on définit les estimateurs $(\widehat{a}_i)_{1\leq i\leq p}$ des points

de support et $(\widehat{\pi}_i)_{1\le i\le p}$ des masses. Puis, sans autre hypothèse sur la nature du filtre inverse θ - c'est à dire avec θ éventuellement non-paramétrique - on donne un résultat de consistance de ces estimateurs.
Enfin, lorsque le filtre inverse θ est paramétré, on établit la loi asymptotique de ces estimateurs.

3.1 Caractérisation des points de support et des masses

Il est possible de caractériser les points de support $\{a_i)_{i=1,\dots,p}$ et les masses $(\pi_i)_{i=1,\dots,p}$ les chargeant à partir des seuls moments de $\varphi_1(X_0)$. On utilise pour cela la structure de T-système de l'application φ_1.

On considère la matrice $D(d_X)$ construite sur les $(p+1)^2$ premiers moments conjugués de la variable $\varphi_1(X_0)$. On obtient le résultat algébrique suivant

Lemme 3.1
On suppose **(M1)** *et* **(M8)** *vérifiées. La plus petite valeur propre de $D(d_X)$ est simple et égale à zéro.*

Le théorème suivant exprime les $(a_i)_{1\le i\le p}$ et les $(\pi_i)_{1\le i\le p}$ en fonction seulement des moments de $\varphi_1(X_0)$ de manière algébrique.

Théorème 3.1
On suppose **(M1)** *et* **(M8)** *vérifiées.*

1. *Soit v^* le vecteur propre associé à la plus petite valeur propre de la matrice $D(d_X)$.*
   Soit le polynôme $P_{v^*}(\varphi_1(x)) = \sum_{j=0}^{p} v_j^* \varphi_1(x)^j$.
   Les zéros ordonnés de $f(x) = P_{v^*}(\varphi_1(x))$ sont les points de support $(a_i)_{1\le i\le p}$.
2. *$\{\pi_1,\dots,\pi_p\}$ est l'unique solution du système linéaire en $(x_1,\dots,x_p)$*

$$\forall k=0,\dots,p-1,\quad E(\varphi_1(X_0)^k) = \sum_{i=1}^{p} x_i \varphi_1(a_i)^k$$

Les preuves du lemme 3.1 et du théorème 3.1 se trouvent en annexe.

3.2 Estimation des points de support et des masses

On va établir la version empirique des objets précédemment définis. Pour ce faire, on remarque tout d'abord que

$$d_X = d_Z(\theta,\sigma_0)$$

Soit maintenant $\widehat{v}^*$ le vecteur propre associé à la plus petite valeur propre de la matrice $D(\widehat{d}(\widehat{\theta},\widehat{\sigma}))$ construite sur les pseudo-moments conjugués empiriques.
On définit les estimateurs $(\widehat{a}_1,\dots,\widehat{a}_p)$ comme suit

Définition 3.1
Pour n suffisament grand, $\widehat{a}_1,\dots,\widehat{a}_p$ sont les zéros ordonnés dans $\mathbb{R}$ de $P_{\widehat{v^*}}(\varphi_1(x))$.

On obtient les estimateurs $(\widehat{\pi}_i)_{1\le i\le p}$ des masses chargeant les points du support $(a_i)_{1\le i\le p}$ par inversion du système triangulaire suivant :

Définition 3.2

$$\forall k=0,...,p-1,\quad \widehat{d}(\widehat{\theta},\widehat{\sigma})_{k,0} = \sum_{i=1}^{p} \widehat{\pi}_i\ \varphi_1(\widehat{a}_i)^k$$

Ces définitions utilisent pleinement la structure p-concentrée du signal.
On établit maintenant la consistance de ces estimateurs.

Théorème 3.2
On suppose vérifiées les hypothèses **(M1)** *à* **(M8)** *et* $\lim_{n\to+\infty} \frac{m(n)}{n\delta(n)} = 0$.
Alors, pour tout $i = 1, \dots, p$, $\widehat{a}_i$ *et* $\widehat{\pi}_i$ *sont presque sûrement consistants.*

Avec l'hypothèse (**M7**) allégée, on obtient la consistance en probabilité.
La preuve de ce résultat se trouve en annexe.

3.3 Vitesse de convergence des estimateurs des points de support et des masses.

On cherche maintenant à établir la distribution asymptotique des estimateurs $(\widehat{a}_i)_{1\le i\le p}$ des points de support et des estimateurs $(\widehat{\pi}_i)_{1\le i\le p}$ de leur masse .
On se restreint pour cette étude au cas paramétrique. Plus précisement l'ensemble Θ est représenté par un modèle paramétrique dont le vecteur paramètre ξ évolue dans un compact $\mathcal{K}$ de dimension q, $\xi = (\xi_j)_{j=1\dots q}$:

$$\Theta := \{\theta(\xi),\ \xi \in \mathcal{K}\}$$

la fonction de paramétrisation θ étant connue. Soit ξ_0 la vraie valeur du paramètre.
Nous supposerons que l'hypothèse (**P**) est vérifiée :

- **(P)** L'application $\theta : \xi \to \theta(\xi)$ de $\mathcal{K}$ dans $l_1(\mathbb{Z})$ est deux fois continuement différentiable.
  Pour tout $i = 1, \dots, q$, $(\frac{\partial\theta_k}{\partial\xi_i})_{k\in\mathbb{Z}}$ et $(\frac{\partial^2\theta_k}{\partial\xi_i^2})_{k\in\mathbb{Z}}$ sont dans $l_1(\mathbb{Z})$.
  De plus les vecteurs $(\frac{\partial\theta_k}{\partial\xi_1}(\xi_0))_{k\in\mathbb{Z}}, \dots, (\frac{\partial\theta_k}{\partial\xi_q}(\xi_0))_{k\in\mathbb{Z}}$ et $(\theta(\xi_0)_k)_{k\in\mathbb{Z}}$ sont linéairement indépendants.
  Enfin, le paramétrage vérifie la condition d'identifiabilité suivante :

$$\theta_k(\xi) = \lambda\theta_{k-K}(\xi'),\ \forall k \in \mathbb{Z} \Longleftrightarrow \lambda = 1,\ K = 0 \text{ and } \xi = \xi'.$$

En d'autres termes, **(P)** introduit des conditions sur la fonction de paramétrisation θ permettant de procéder à l'aide de développements à l'ordre 2. La condition d'indépendance linéaire des différentielles partielles premières de $(\theta_k)_{k\in\mathbb{Z}}$ en ξ_0 et de $(\theta(\xi_0)_k)_{k\in\mathbb{Z}}$ est du même type que les conditions de qualification requises pour la recherche d'un optimum sous contraintes (théorème de Kuhn et Tucker). Enfin, comme dans le cadre non paramétrique, il est nécessaire de fixer le décalage K et l'échelle λ afin d'avoir un modèle identifiable par application du filtre inverse $\theta(\xi_0)$.

Pour établir la distribution asymptotique de $(\widehat{a}_i)_{i=1,\dots,p}$ et $(\widehat{\pi}_i)_{i=1,\dots,p}$ on utilise la procédure d'estimation de $\widehat{\xi}$ et $\widehat{\sigma}$ [GG99]. On énonce brièvement les résultats dont nous aurons besoin par la suite. On estime ξ_0 et σ_0, en minimisant $L_n(\xi,\sigma) := J_n(\theta(\xi),\sigma)$ pour tout $\xi \in \mathcal{K}$ et tout $\sigma \in \mathbb{R}^+$. On note $\widehat{\xi}$ et $\widehat{\sigma}$ ces minimum. Par transcription littérale du théorème (2.2) sous les hypothèses **(M1)** à **(M8)** ainsi que l'hypothèse **(P)**, on obtient un résultat de consistance de $\widehat{\xi}$ et $\widehat{\sigma}$. De même les estimateurs $(\widehat{a}_i)_{1\le i\le p}$ et $(\widehat{\pi}_i)_{1\le i\le p}$ sont définis de manière similaire et vérifient la propriété de consistance avec $\widehat{\theta} := \theta(\widehat{\xi})$.

On fixe plus finement les vitesses de convergence relatives du paramètre de troncature $m(n)$, de la pénalisation $\delta(n)$ et du temps n avec l'hypothèse **(M9)** suivante

- **(M9)**

$$\lim_{n\to\infty} \frac{m(n)}{\sqrt{n}} = 0 \quad \lim_{n\to\infty} \sqrt{n} \sum_{|k|>m(n)} |\theta_k| = 0 \quad \lim_{n\to\infty} \sqrt{n}\delta(n) = +\infty$$

$$\lim_{n\to\infty} \sqrt{n}\delta(n)^2 = 0.$$

Enfin, on impose une hypothèse **(M10)** assurant la convergence normale des moments empiriques conjugués standards $\widehat{c}(\theta(\xi^*))$ de $W(\theta(\xi^*))$ et de leur différentielles premières en ξ^*.
On définit les moments empiriques conjugués standarts en ξ par $\widehat{c}(\theta(\xi)) = (\widehat{c}(\theta(\xi))_{k,j})_{(k,j)=\{0,\dots,p\}^2}$ où

$$\widehat{c}(\theta(\xi))_{k,j} = \frac{1}{n}\sum_{t=1}^{n} \phi_1(W_t(\theta(\xi)))^k \overline{\phi_1(W_t(\theta(\xi)))}^j \quad \text{et} \quad c(\theta(\xi))_{k,j} = E\left(\phi_1(W_0(\theta(\xi)))^k \overline{\phi_1(W_0(\theta(\xi)))}^j\right)$$

- **(M10)**

$$\sqrt{n}\begin{pmatrix} \widehat{c}(\theta(\xi^*)) - c(\theta(\xi^*)) \\ D^1_{\xi}\left(\ \widehat{c}\circ\theta(\xi^*) - c\circ\theta(\xi^*)\ \right) \end{pmatrix} \longrightarrow \mathcal{N}(0,\Gamma) \quad \text{en loi}$$

Notons que sous **(M10)**, **(M6)** est vérifiée. Soit Γ_1 la matrice de variance-covariance asymptotique de $\sqrt{n}\,(\widehat{c}(\theta(\xi^*)) - c(\theta(\xi^*)))$. Afin d'alléger l'écriture, nous noterons $\widehat{\theta} = \theta(\widehat{\xi})$ et $\theta = \theta(\xi^*)$ excepté lorsque la dépendance explicite en ξ est requise.
Sous cette hypothèse, l'article [GG99] donne le comportement asymptotique de $\sqrt{n}\,(\widehat{\xi}-\xi^*)$ (preuve du théorème 4.2) qui est tel que :

$$\sqrt{n}\,(\widehat{\xi} - \xi^*) = -\frac{{}^t(D^2_{\xi,\sigma}H_n(\theta(\xi^*),\sigma_0))}{D^1_{\sigma}H_n(\theta(\xi^*),\sigma_0)}\sqrt{n}\ H_n(\theta(\xi^*),\sigma_0)(1+o(1)) \tag{6}$$

Le lemme (3.2) qui suit précise la forme de la différentielle seconde apparaissant dans la loi asymptotique de $\widehat{\xi}$ (6) et une condition suffisante pour que cette différentielle soit non nulle. On introduit quelques notations et une hypothèse utilisée par la suite.
On note $g_{ij} = g(X_0^{(i)}) - g(X_0^{(j)})$ où les $X_0^{(i)}$ sont des copies indépendantes de X_0, et $g_i = g(X_0^{(i)})$.
Soit l'hypothèse **(M11)**

- **(M11)**

$$E\left(|\varphi'_1|^2 \prod_{i<j,\neq(0,1)} |\varphi_{ij}|^2\right) + \frac{1}{2}E\left(\mathcal{R}e\left(\overline{\varphi'}_1\varphi"_1(X_0^{(0)}+X_0^{(1)}) \prod_{i<j,\neq(0,1)} |\varphi_{ij}|^2\right)\right)$$
$$+ \sum_{i<j} E\left(\mathcal{R}e\left(X_0^{(01)}|\varphi'_1|^2\varphi'_1 + X_0^{(0)}|\varphi'_j|^2\varphi'_{01}\right)\overline{\varphi}_{01} \prod_{i'<j',\neq(i,j),\neq(0,1)} |\varphi_{i'j'}|^2\right) \neq 0$$

Lemme 3.2 *Sous les hypothèses* **(M1)** *à* **(M10)**, *on a*

$$D^1_{\sigma}H(\theta(\xi^*),\sigma_0) = \frac{-8\|\theta\|_2^2\sigma_0}{(p-1)!}E\left(|\varphi'_1|^2 \prod_{i<j\neq(0,1)} |\varphi_{ij}|^2\right) \neq 0$$

et avec pour tout l entier relatif,

$$\frac{\partial^2 H(\theta(\xi^*),\sigma_0)}{\partial s_l \partial\sigma} = \frac{-16\sigma_0\theta_l}{(p+1)!}\sum_{i<j} E\left(|\varphi'_j|^2 \prod_{i'<j',\neq(i,j)} |\varphi_{i'j'}|^2\right)$$
$$- \frac{8\sigma_0\|\theta\|_2^2}{(p+1)!}\sum_{i<j} E\left(\mathcal{R}e\left(\overline{\varphi'}_j\varphi"_j(u\star(X^{(i)}+X^{(j)}))_{-l}\right) \prod_{i'<j',\neq(i,j)} |\varphi_{i'j'}|^2\right)$$

$$
\begin{aligned}
- & \frac{4\sigma_0\|\theta\|_2^2}{(p+1)!}\sum_{i<j} E\left(\mathcal{R}e\left((u\star X)_{-l}^{(ij)}|\varphi_j'|^2\sum_{k=i+1,\neq j}^{p}(\overline{\varphi}_{ik}\varphi_i'+|\varphi_{ik}'|^2+\overline{\varphi}_{ik}\varphi_k')\prod_{i'<j'\neq(i,j)\neq(i,k)}|\varphi_{i'j'}|^2\right)\right.\\
- & \mathcal{R}e\left((u\star X)_{-l}^{(ij)}|\varphi_j'|^2\sum_{k=j+1}^{p}(\overline{\varphi}_{jk}\varphi_j'+|\varphi_{jk}'|^2+\overline{\varphi}_{jk}\varphi_k')\prod_{i'<j'\neq(i,j)\neq(j,k)}|\varphi_{i'j'}|^2\right)\\
- & 2\mathcal{R}e\left((u\star X)_{-l}^{(ij)}|\varphi_j'|^2\sum_{k=0}^{i-1}(\overline{\varphi}_{ki}\varphi_i')\prod_{i'<j'\neq(i,j)\neq(k,i)}|\varphi_{i'j'}|^2\right)\\
+ & \left.2\mathcal{R}e\left((u\star X)_{-l}^{(ij)}|\varphi_j'|^2\sum_{k=0,\neq i}^{j-1}(\overline{\varphi}_{kj}\varphi_j')\prod_{i'<j'\neq(i,j)\neq(k,j)}|\varphi_{i'j'}|^2\right)\right)\\
- & \frac{16\sigma_0\|\theta\|_2^2}{(p+1)!}\sum_{i<j}\sum_{i'<j'\neq(i,j)} E\left(\mathcal{R}e\left((u\star X^{(ij)})_{-l}|\varphi_j'|^2\varphi_j'+(u\star X^{(i)})_{-l}|\varphi_{j'}'|^2\varphi_{ij}'\right)\overline{\varphi}_{ij}\prod_{i"<j"}|\varphi_{i"j"}|^2\right)
\end{aligned}
$$

De plus, sous l'hypothèse **(M11)**, *on a* $D^2_{s,\sigma}H(\theta(\xi^*),\sigma_0)\neq 0$

Remarque

- Lorsque l'on considère les fonctions de contraste particulières Hankel ou Toeplitz, on obtient des expression considérablement simplifiées.
- Si l'on utilise la structure de dépendance des $(X_k)_{k\in\mathbb{Z}}$, les termes en $((u\star X)_{-l})_{i,j}$ sont alors explicites. En particulier, dans la cas i.i.d., tous les termes comportant des expressions du type $(u\star X)_{-l})^{(ij)}\prod_{i'<j'}g_{i'j'}$ s'annulent.
- Sous Hankel et Toeplitz, avec un signal i.i.d., **(M11)** est vérifiée quelque soit la nature du filtrage.

Comme nous le verrons avec le théorème (3.3), la distribution asymptotique des estimateurs des points de support et des masses est essentiellement liée à celle des pseudo-moments conjugués $\widehat{d}(\widehat{\theta},\widehat{\sigma})$. C'est pourquoi, dès à présent on explicite totalement cette distribution dans le lemme (3.3) en s'appuyant sur le résultat donné par le lemme (3.2).

Lemme 3.3 *On suppose que les hypothèses* **(M1)** *à* **(M11)** *et* **(P)** *sont vérifiées. Alors,*

$$\lim_{n\to+\infty}\sqrt{n}\left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma})-d(\theta(\xi^*),\sigma_0)\right)=\mathcal{N}(0,\Gamma_d)\qquad \textit{en loi.}$$

où

$$\Gamma_d=\frac{C\Gamma_1\ {}^tC}{D^1_\sigma H(\theta(\xi^*),\sigma_0)^2}$$

avec

$$C=\left[D^1_{\xi,\sigma}d(\theta(\xi^*),\sigma_0)\begin{pmatrix} {}^tD^2_{\xi,\sigma}H(\theta(\xi^*),\sigma_0)\\ 1\end{pmatrix}D^1_d h(d(\theta(\xi^*),\sigma_0))+I_{(p+1)}\right]\neq 0$$

Avec cette distribution asymptotique, on a l'intégralité de la variation stochastique des estimateurs. En effet, comme on peut le voir dans le théorème suivant 3.3, les lois asymptotiques de $\widehat{a}$ et $\widehat{\pi}$ sont ensuite déduites par des manipulations algébriques ne dépendant que de φ_1,p,a,π et en aucun cas de la structure stochastique de $X_k+\sigma_0(\theta\star\epsilon)_k$.

On définit maintenant les objets relatifs aux points de support, à leurs masses et à la fonction de contraste utilisée.

$$A_{\varphi_1} = \begin{pmatrix} 1 & \dots & 1 \\ \varphi_1(a_1) & \dots & \varphi_1(a_p) \\ \vdots & & \vdots \\ \varphi_1(a_1)^{p-1} & \dots & \varphi_1(a_p)^{p-1} \end{pmatrix} \quad A_{\varphi_1'} = \begin{pmatrix} 0 & \dots & 0 \\ & (\pi_j \varphi_1'(a_j) i \varphi_1(a_j)^{i-1}) & \end{pmatrix}_{1 \leq i \leq p-1; 1 \leq j \leq p}$$

$$A = \begin{pmatrix} |\varphi_1'(a_1)|^2 K_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & |\varphi_1'(a_p)|^2 K_p \end{pmatrix}_{1 \leq i \leq p-1; 1 \leq j \leq p} \quad \text{où } K_j = E\left(\prod_{i=1, \neq j}^{p} |\varphi_1(X_0) - \varphi_1(a_i)|^2 \right)$$

$$\begin{array}{rcl} P : \mathbb{C}^{(p+1)^2} & \to & \mathbb{C}^p \\ w & \mapsto & Pw = v = (w_1, w_{(p+2)}, \dots, w_{k(p+1)+1}, \dots, w_{p^2})_{k=0,\dots,(p-1)} \end{array}$$

et $B \in \mathcal{M}_{p,(p+1)^2}$ une matrice définie par colonne par indexation de chacune des $(p+1)^2$ colonnes par le couple $(i,j) \in \{0, \dots, p\}$ et

$$B_{(i,j),.} = {}^t\overline{D_b^1 v_i(\underline{\varphi_1}(a))} v_j^* + {}^t D_b^1 v_j(\underline{\varphi_1}(a)) \overline{v_i^*}$$

On est maintenant en mesure de donner les lois asymptotiques des estimateurs des points de support et des masses les chargeant.

Théorème 3.3 *On suppose* **(M1)** *à* **(M11)** *et* **(P)** *vérifiées,* $\lim_{n \to +\infty} \frac{m(n)}{n\delta(n)} = 0$, *et* $\phi_1(0) = 0$ *dans le cas additif et* $\phi_1(0) = 1$ *dans le cas multiplicatif. On a alors*

$$\sqrt{n}\,(\widehat{a} - a) \longrightarrow \mathcal{N}(0, \Gamma_a) \quad \textit{en loi}$$

et

$$\sqrt{n}(\widehat{\pi} - \pi) \longrightarrow \mathcal{N}(0, \Gamma_\pi) \quad \textit{en loi}$$

où $\Gamma_a = \frac{1}{4|v_p^*|^4} A^{-1} B \frac{C\Gamma_1 \, {}^tC}{D_\sigma^1 H(\theta(\xi^*), \sigma_0)^2} \, {}^t\overline{B} A^{-1}$
et $\Gamma_\pi = A_{\varphi_1}^{-1} \left(A_{\varphi_1'} \frac{A^{-1}}{2|v_p^*|^2} B - P \right) \frac{C\Gamma_1 \, {}^tC}{D_\sigma^1 H(\theta(\xi^*), \sigma_0)^2} \, {}^t(A_{\varphi_1}^{-1} \left(A_{\varphi_1'} \frac{A^{-1}}{2|v_p^*|^2} B - P \right)))$

On donne en quelques mots l'idée de la preuve. Le coeur de celle-ci réside en l'obtention de la vitesse de la plus petite valeur propre $\widehat{\lambda}_0$. La loi asymptotique de $\widehat{\pi}$ se déduit ensuite assez simplement de celle de $\widehat{a}$. Nous avons défini précédemment l'estimateur $\widehat{a}$ des points de support du processus discret X par rapport à la plus petite valeur propre de la matrice $\widehat{D}$. Plus précisemment, les coordonnées de $\widehat{a}$, $(\widehat{a}_1, \dots, \widehat{a}_p)$ sont les p racines du polynôme de degré p en $\varphi_1(x)$: $P_{\widehat{v^*}}(\varphi_1(x)) = \sum_{j=0}^p \widehat{v}_j^* \varphi_1(x)^j$, où $\widehat{v}^*$ est le vecteur propre de la matrice $\widehat{D}$ associé à sa plus petite valeur propre $\widehat{\lambda}_0$. Exprimer la vitesse de convergence de l'estimateur des points de support $\widehat{a}$ revient à relier cette vitesse à celle de la convergence de $\widehat{D}$ vers D. Aussi ne faut-il pas se servir "brutalement" de l'expression du polynôme $P_{\widehat{v^*}}$. Le point clef ici est de saisir le rôle de la plus petite valeur propre $\widehat{\lambda}_0$ de la matrice des estimateurs des pseudo-moments associés $\widehat{D}$. On retrouve ici l'idée développée à plus ou moins grande échelle, d'un indice de concentration de p points de support synthétisé en la plus petite valeur propre d'une matrice de Hankel ou Toeplitz de taille p [LS00], [D.H92], [Mak81]. Comme $\widehat{v}^*$ est le vecteur propre de $\widehat{D}$ associé à la plus petite valeur propre $\widehat{\lambda}_0$, pour tout vecteur v de taille $(p+1)^2$, on a

$$\frac{{}^t\bar{v}\widehat{D}v}{\|v\|_2^2} \geq \frac{{}^t\overline{\widehat{v}^*}\widehat{D}\widehat{v}^*}{\|\widehat{v}^*\|_2^2}$$

Maintenant, on note v l'application bijective associant à tout point $b = (b_1, \ldots, b_p) \in \mathbb{C}^p$ un élément de $\mathbb{C}^{p+1} \cap \{\|\,.\,\|_2 = 1\}$ l'image $v(b)$ définie par :

$$\forall x \in \mathbb{R}, \quad \sum_{j=0}^{p} v_j(b)\varphi_1(x)^j = v_p(b)\prod_{i=1}^{p}(\varphi_1(x) - \varphi_1(b_i))$$

Ainsi, pour tout vecteur $b \in \mathbb{C}^p$, $\exists v \in \mathbb{C}^{(p+1)} \cap \{\|.\|_2 = 1\}$ tel que

$$v = v(b) \quad \text{avec} \quad \frac{{}^t\bar{v}\widehat{D}v}{\|v\|_2^2} = |v_p(b)|^2 \,\mathrm{Empi}_t\left(\prod_{i=1}^{p} |\varphi_1(\widehat{Z}_t(\widehat{\theta};\widehat{\sigma})) - \varphi_1(b_i)|^2\right)$$

On note $\widehat{g}(b)$ cette application. On a pour tout $b \in \mathbb{C}^p$,

$$|v_p(b)|^2 \,\mathrm{Empi}_t\left(\prod_{i=1}^{p} |\varphi_1(\widehat{Z}_t(\widehat{\theta};\widehat{\sigma})) - \varphi_1(b_i)|^2\right) \geq |v_p(b^*)|^2 \,\mathrm{Empi}_t\left(\prod_{i=1}^{p} |\varphi_1(\widehat{Z}_t(\widehat{\theta};\widehat{\sigma})) - \varphi_1(b_i^*)|^2\right)$$

Et par définition de $\widehat{a}$, on a $\widehat{a} = b^*$
On voit ici clairement apparaître le sens de l'opportunité de la plus petite valeur propre $\widehat{\lambda}_0$: elle signifie que l'on a chercher les p points sur lesquels se concentre en moyenne la mesure du processus $\widehat{Z}_t(\widehat{\theta};\widehat{\sigma})$. Celle-ci convergeant vers la mesure de X_0 qui a p points de support $a = (a_1, \ldots, a_p)$. L'estimateur $\widehat{a}$ convergera donc par un simple un jeu algébrique vers a à la vitesse de convergence de $\widehat{Z}_t(\widehat{\theta};\widehat{\sigma})$ vers X_0.

Cette méthode d'estimation est aisée à implémenter car il ne s'agit que de la recherche de la plus petite valeur propre d'une matrice bien déterminée et de recherche de zéros d'un polynôme d'une part et d'une inversion d'un système linéaire d'autre part.
De plus l'ensemble des objets définis dans l'expression des lois asymptotiques sont explicites et implémentables. En effet, la covariance Γ_a est une fonction de Γ qui dépend de la structure de dépendance des (X_t) et de $(\theta \star \epsilon)$ permettant d'obtenir un théorème central limite de matrice de covariance Γ. Elle dépend également des objets $A, B, C, D_\sigma^1 H$ et v_p^*. La dernière coordonnées de v^* se calcule aisémment pour tout p, voir par exemple Ralston et Rabinowitz [RR90], en fonction de p de a et de φ_1. Les expressions C et $D_\sigma^1 H$ se calculent avec l'aide des résultats donnés par le lemme 3.2 après avoir établi le type de fonction de contraste utilisée et la structure des (X_k). Enfin A et B sont directement calculables en fonction seulement de a , π et de la fonction φ_1, sans considération sur le type de dépendance du processus (X_k) ou sur la nature de la filtration $(u_k)_{k\in\mathbb{Z}}$.
La matrice de covariance Γ_π est définie par rapport aux mêmes objets que Γ_a avec de plus les matrices A_{φ_1} et $A_{\varphi_1'}$ qui dépendent des a_i, π_i, p et φ_1 d'une part et de P qui dépend de p d'autre part. Tous ces objets peuvent donc être estimés ce qui permet de construire des intervalles de confiance pour a et π.
Outre l'intérêt de ce résultat afin de permettre une restitution du signal X par la voie bayésienne, l'analyse de la distribution asymptotique obtenue est également intéressante en tant que telle. En effet ce résultat exprime quantitativement en fonction de quelles données du modèle la variabilité des estimateurs des points de support évolue.

4 Preuves.

Preuve du lemme 3.1 :
La matrice $D(d_X)$ est hermitienne car $(d_X)_{k,j} = (\overline{d_X})_{j,k}$. Elle est donc diagonalisable dans $\mathbb{C}$. On

note $(\lambda_i)_{0\le i\le p}$ ses $p+1$ valeurs propres complexes, de vecteurs propres associés $v^{(i)} = (v_j^{(i)})_{j=0,\dots,p}$. On a

$$\forall u \in \mathbb{C}^{p+1}, \quad {}^t\bar{u}\; D(d_X)\; u = \sum_{k=0}^{p}\sum_{j=0}^{p} \bar{u}_k u_j E(\varphi_1(X_0)^k \overline{\varphi_1(X_0)}^j)$$

Donc

$$\forall u \in \mathbb{C}^{p+1}, \quad {}^t\bar{u}\; D(d_X)\; u = E\left|\sum_{k=0}^{p} u_k \varphi_1(X_0)^k\right|^2 \tag{7}$$

où $|.|$ désigne le module d'un complexe.
Ainsi, toutes les valeurs propres de $D(d_X)$ sont réelles et positives. En les considérant ordonnées, on a :

$$0 \le \lambda_0 \le \lambda_1 \le \dots \le \lambda_p$$

Soit le polynôme P de $\varphi_1(x)$, de degré p, défini par la forme factorisée suivante

$$\forall x \in \mathbb{C}, \quad P(\varphi_1(x)) = \prod_{i=1}^{p} (\varphi_1(x) - \varphi_1(a_i))$$

Il existe une suite $(w_j)_{j=0,\dots,p}$ de complexes, telle que P s'écrive sous la forme linéarisée suivante

$$\forall x \in \mathbb{C}, \quad P(\varphi_1(x)) = \sum_{k=0}^{p} w_j \varphi_1(x)^k \tag{8}$$

et on a $P(\varphi_1(X_1)) = 0$ presque sûrement. En particulier avec l'égalité (7) et la forme linéarisée (8), on obtient

$${}^t\bar{w}\; D(d_X)\; w = E\left|P(\varphi_1(X_0))\right|^2 = 0 \Longrightarrow \quad \lambda_0 = 0, \;\text{ et }\; \exists\beta, \quad v^* = \beta w$$

Il reste à montrer que cette valeur propre est simple. Autrement dit que ${}^t\bar{u}\; D(d_X)\; u = 0$ si et seulement si u est colinéaire à v^*. Mais ${}^t\bar{u}\; D(d_X)\; u = 0$ si et seulement si le polynôme $P_u(\varphi_1(X_0)) = \sum_{k=0}^{p} u_k(\varphi_1(X_0))^k$ en $\varphi_1(X_0)$ est nul presque sûrement. Or $P_u = 0$ presque sûrement si et seulement si $\varphi_1(X_0)$ prend ses valeurs dans l'ensemble des racines de P_u. Comme les racines de P_u sont au nombre de p, et que l'ensemble des valeurs de $\varphi_1(X_0)$ est de cardinal p, par l'écriture factorisée, il vient que P_u est un multiple de P, et ainsi que u est colinéaire à v^*.

Preuve du théorème 3.1 :

a) Le lemme (3.1) nous assure que le vecteur propre v^* associé à la plus petite valeur propre λ_0 de $D(d_X)$ est bien défini, de manière unique, en fonction des moments de $\varphi_1(X_0)$, à un facteur près. Ainsi P_{v^*} est également bien défini de manière unique, à un facteur près. En tant que polynôme de degré p, il admet p racines complexes. Or, on a $a_1, \dots, a_p$ qui sont racines de P_{v^*} car $P_{v^*}(\varphi_1(X_0)) = 0$ presque sûrement.

b) Le système

$$\forall k = 0, \dots, p-1, \quad E(\varphi_1(X_0)^k) = \sum_{i=1}^{p} x_i \varphi_1(a_i)^k$$

admet un unique vecteur x solution dans $\mathbb{R}^p$. En effet, le déterminant de ce système est un déterminant de Van der Monde égal à

$$\prod_{1\le j<i\le p} (\varphi_1(a_i) - \varphi_1(a_j))$$

φ_1 étant injective et les points du support étant tous distincts **(M1)**, le déterminant est non nul. De plus, on a

$$\forall k = 0, \ldots, p-1, \quad E(\varphi_1(X_0)^k) = \sum_{i=1}^{p} \pi_i \varphi_1(a_i)^k$$

Ainsi, $\forall 1 \leq i \leq p, \quad \pi_i = x_i$.

Preuve du théorème 3.2 :

a) On considère le système triangulaire suivant:

$$\widehat{d}_W(\widehat{\theta}) = \Delta_{\varphi_2}(\widehat{\sigma}^2 \|\widehat{\theta}\|_2^2)\widehat{d}(\widehat{\theta}, \widehat{\sigma}) \tag{9}$$

Et, sous les hypothèses **(M1)** à **(M7)** le théorème 2.2 donne la convergence presque sûre de $\widehat{\theta}$ vers θ dans $l_1(\mathbb{Z})$ et de $\widehat{\sigma}$ vers σ_0 avec n. Par ergodicité du processus $(X_k)_{k\in\mathbb{Z}}$, le vecteur composé des moments conjugués empiriques calculé en $\widehat{W}(\widehat{\theta})$ converge presque sûrement vers le vecteur des moments conjugués calculé en $W_0(\theta)$. En inversant le système (9) et en utilisant la continuité de $\Delta_{\varphi_2}^{-1}$ en s et σ, on a, presque sûrement,

$$\lim_{n\to+\infty} D(\widehat{d}(\widehat{\theta}, \widehat{\sigma})) = D(d(\theta, \sigma_0)) \quad \text{p.s.} \tag{10}$$

Par construction $D(\widehat{d}(\widehat{\theta}, \widehat{\sigma}))$ est une matrice hermitienne. Elle est donc diagonalisable dans $\mathbb{C}$. On note $\widehat{\lambda}_i$ ses $p+1$ valeurs propres. On pourra vérifier par un raisonnement analogue à celui effectué dans la preuve du lemme 3.1, que celles-ci sont réelles. Comme $D(\widehat{d}(\widehat{\theta}, \widehat{\sigma}))$ converge vers $D(d(\theta, \sigma_0))$ presque sûrement, on obtient la convergence des valeurs propres :

$$\lim_{n\to+\infty} (\widehat{\lambda}_i)_{1\leq i\leq p} = (\lambda_i)_{1\leq i\leq p} \quad \text{presque sûrement}$$

En particulier, en tout n, on s'intéresse à la plus petite valeur propre de $D(\widehat{d}(\widehat{\theta}, \widehat{\sigma}))$, c'est-à-dire $\widehat{\lambda}_0 = \min_{k=1,\ldots,p}\{\widehat{\lambda}_k\}$ qui converge presque sûrement vers $\lambda_0 = 0$. Autrement dit, avec la convergence (10)

$$0 = \min_{i=0,\ldots,p} \{ \sum_{k,j=0}^{p} d(\theta, \sigma)_{k,j} \lim_{n\to\infty} \bar{\widehat{v}}_k^i \widehat{v}_j^i \}$$

où $\widehat{v}^{(k)}$ est le vecteur propre associé à l'instant n à la valeur propre $\widehat{\lambda}_k$. On note néanmoins $\widehat{v}^{(0)} = \widehat{v}^*$ et $v^{(0)} = v^*$. Ainsi, en particulier, $\widehat{v}^*$ converge vers v^*, et le polynôme $P_{\widehat{v^*}}$ défini par

$$P_{\widehat{v^*}}(\varphi_1(x)) = \sum_{i=1}^{p} \widehat{v}_i^* \varphi_1(x)^i$$

converge presque sûrement, pour tout x dans $\mathbb{R}$, vers $P_{v^*}(\varphi_1(x))$. Or $P_{v^*}(\varphi_1(x))$possèdant p racines distinctes $a_1, ..., a_p$, pour n suffisament grand, $P_{\widehat{v^*}}(\varphi_1(x))$ admet p racines distinctes $\{\widehat{\widehat{a}}_1, ..., \widehat{\widehat{a}}_p\}$ convergeant vers $\{a_1, \ldots, a_p\}$.
En tout n, on ordonne ces racines que l'on note $\{\widehat{a}_1, ..., \widehat{a}_p\}$. On a alors

$$\forall 1 \leq i \leq p, \quad \lim_{n\to\infty} \widehat{a}_i = a_i \quad p.s.$$

b) Le système

$$\forall k = 0, ..., p-1, \quad \widehat{d}_{k,0}(\widehat{\theta}, \widehat{\sigma}) = \sum_{i=1}^{p} \widehat{\pi}_i \varphi_1(\widehat{a}_i)^k$$

peut s'écrire matriciellement $\widehat{C} = {}^t\widehat{A}\,\widehat{\Pi}$ avec

$$\begin{aligned}
\widehat{C} &= \left(\widehat{d}_{k,0}(\widehat{\theta},\widehat{\sigma})\right)_{k=0,\dots,p-1} \\
\widehat{\Pi} &= (\widehat{\pi}_i)_{i=1,\dots,p} \\
\widehat{A} &= \begin{pmatrix} 1 & 1 & \cdot\ \cdot\ \cdot & 1 \\ \varphi_1(\widehat{a}_1) & \varphi_1(\widehat{a}_2) & \cdot\ \cdot\ \cdot & \varphi_1(\widehat{a}_p) \\ \varphi_1(\widehat{a}_1)^2 & \varphi_1(\widehat{a}_2)^2 & \cdot\ \cdot\ \cdot & \varphi_1(\widehat{a}_p)^2 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ \varphi_1(\widehat{a}_1)^{p-1} & \varphi_1(\widehat{a}_2)^{p-1} & \cdot\ \cdot\ \cdot & \varphi_1(\widehat{a}_p)^{p-1} \end{pmatrix}
\end{aligned}$$

Le déterminant de $\widehat{A}$ est le determinant de Van der Monde égal à :

$$\prod_{1 \le i < j \le p} (\varphi_1(\widehat{a}_i) - \varphi_1(\widehat{a}_j))$$

Or, comme, pour tout indice i $\lim_{n\to\infty} \widehat{a}_i = a_i$ *p.s.* , $a_i \neq a_j$ pour tout $j \neq i$, et φ_1 est injective, pour n suffisamment grand, on a $\varphi_1(\widehat{a}_i) \neq \varphi_1(\widehat{a}_j) \quad \forall j \neq i$. Le système est inversible et on achève la preuve avec la convergence presque sûre de $\widehat{C}$ vers C, $\widehat{A}$ vers A et $\widehat{\Pi}$ vers Π.

Preuve du Lemme 3.2 :
On utilise ici pour une part des arguments similaires à ceux introduits dans l'article ([GG99]) pour la preuve du lemme (4.1).
Lorsque $\sigma \le \sigma_0$, $\ H(\theta(\xi),\sigma)$ peut s'exprimer comme fonction des moments d'une variable aléatoire. Plus précisément, on a

$$\forall \sigma \le \sigma_0, \forall s \in \mathbb{Z}, \quad H(s,\sigma) = \frac{1}{(p+1)!} E\left[\prod_{i<j}\left|\varphi_1\left(Z_0^{(i)}(s)\right) - \varphi_1\left(Z_0^{(j)}(s)\right)\right|^2\right]$$

où $Z_0^{(i)}(s) = (s \star u \star X^{(i)})_0 + \sqrt{\sigma^2 - \sigma_0^2}(s \star \epsilon^{(i)})_0$ avec pour tout $i = 0,\dots,p$ et pour tout k entier relatif, les $X_k^{(i)}$ et les $\epsilon_k^{(i)}$ qui sont des copies indépendantes de X_k et ϵ_k respectivement.
Pour tout σ, H est une fonction infiniment dérivable en s et σ, et en particulier en $s = \theta$ et $\sigma = \sigma_0$. Pour tout $\sigma < \sigma_0$, on écrit la différentielle seconde par rapport à s_l et σ en ne gardant que les termes donnant des termes de puissance nulle en $\sqrt{\sigma_0^2 - \sigma^2}$. En effet, H étant infiniment différentiable en σ et s, elle l'est en particulier pour $\sigma_0 = \sigma$. Pour se faire, on développe au premier ordre les fonctions de $Z^{(i)}(s)$ au voisinage de σ_0. Par calcul direct, la différentielle partielle seconde $\frac{\partial^2 H(s,\sigma)}{\partial s_l \partial\sigma}$ est égale à

$$\begin{aligned}
&\frac{2}{(p+1)!}\sum_{i<j} E\Bigg[\mathcal{R}e\Bigg\{\frac{-\sigma}{\sqrt{\sigma_0^2-\sigma^2}}\Big(\overline{\varphi_{ij}}(\varphi"\epsilon\ (u\star X)_{-l})_{ij} + \overline{\varphi_{ij}}(\varphi'(\epsilon_{-l}))_{ij} + (\varphi'(u\star X)_{-l})_{ij}\overline{(\varphi'\epsilon)}_{ij}\Big) \\
&-\sigma\Big(\overline{\varphi_{ij}}(\varphi'''\epsilon^2(u\star X)_{-l})_{ij} + \overline{(\varphi'\epsilon)}_{ij}(\varphi"\epsilon\ (u\star X)_{-l})_{ij} + \overline{\varphi}_{ij}(\varphi''\epsilon\ \epsilon_{-l})_{ij} + \overline{(\varphi'\epsilon)}_{ij}(\varphi'\epsilon_{-l})_{ij} \\
+\quad &(\varphi'(u\star X)_{-l})_{ij}\overline{(\varphi''\epsilon^2)}_{ij} + \varphi''\epsilon(u\star X)_{-l})_{ij}\overline{(\varphi'\epsilon)}_{ij} + \overline{\varphi}_{ij}(\varphi''\epsilon\epsilon_{-l})_{ij} + (\varphi'\epsilon_{-l})_{ij}\overline{(\varphi'\epsilon)}_{ij}\Big)\Bigg\} \\
&\Bigg(\prod_{i'<j'\neq(i,j)}|\varphi_{i'j'}|^2 + \sum_{i',j'} 2\sqrt{\sigma_0^2-\sigma^2}\,\mathcal{R}e\left\{\overline{\varphi_{i'j'}}(\varphi'\epsilon)_{i'j'}\right\}\prod_{i"<j"\neq(i',j'),\neq(i,j)}|\varphi_{i"j"}|^2\Bigg)\Bigg]
\end{aligned}$$

$$+ \quad \frac{4}{(p+1)!}\sum_{i<j}\sum_{i'<j'\neq(i,j)} E\left[\left\{\frac{-\sigma}{\sqrt{\sigma_0^2-\sigma^2}}\mathcal{R}e\left\{\overline{\varphi_{ij}}(\varphi'(u\star X)_{-l})_{ij}\right\}\mathcal{R}e\left\{\overline{\varphi_{i'j'}}(\varphi'\epsilon)_{i'j'}\right\}\right.\right.$$

$$-\sigma\mathcal{R}e\left\{\overline{(\varphi'\epsilon)}_{ij}(\varphi'(u\star X)_{-l})_{ij}+\overline{\varphi_{ij}}(\varphi"\epsilon\,(u\star X)_{-l})_{ij}+\overline{\varphi_{ij}}(\varphi'\epsilon_{-l})_{ij}+\overline{\varphi_{ij}}(\varphi'(u\star X)_{-l})_{ij}\right\}$$

$$\mathcal{R}e\left\{\overline{\varphi_{i'j'}}(\varphi'\epsilon)_{i'j'}+\overline{\varphi_{i'j'}}(\varphi''\epsilon^2)_{i'j'}+(\varphi'\epsilon)_{i'j'}\overline{((\varphi'\epsilon)}_{i'j'}\right\}$$

$$\left.\left.\prod_{i''<j'',\neq(i,j)(i',j')}\left[|\varphi_{i''j''}|^2+2\sqrt{\sigma_0^2-\sigma^2}\mathcal{R}e\left\{\overline{\varphi_{i''j''}}(\varphi'\epsilon)_{i''j''}\right\}\right]\right\}\right]$$

où $\varphi_{ij}=\varphi_1\left(Z^{(i)}(s)\right)-\varphi_1\left(Z^{(j)}(s)\right)$, $\epsilon_{ij}=(s\star\epsilon^{(i)})_0-(s\star\epsilon^{(i)})_0$ et $(\epsilon_{-l})_{ij}=\epsilon_{-l}^{(i)}-\epsilon_{-l}^{(j)}$.
Enfin pour toute quantité g,h,l, on note $(ghl)_{ij}=g^{(i)}h^{(i)}l^{(i)}-g^{(j)}h^{(j)}l^{(j)}$.
On calcule maintenant la différentielle au point $(\theta(\xi^*),\sigma_0)$. Grâce à la p-concentration des $X_0^{(i)}$, on a

$$\forall(g_{ij})_{i<j},\quad \prod_{i<j}\left(g_{ij}\left(X^{(i)}\right)-g_{ij}\left(X^{(j)}\right)\right)=0\quad p.s.$$

On obtient alors

$$(p+1)!\frac{\partial^2 H(\theta(\xi^*),\sigma_0)}{\partial s_l\partial\sigma} = -4\sigma_0\theta_l\sum_{i<j}E\left(|\varphi_j'|^2|(\epsilon_{-l})_{ij}|^2\prod_{i'<j'}|\varphi_{i'j'}|^2\right)$$

$$- \quad 4\sigma_0\sum_{i<j}\sum_{i'<j'\neq(i,j)}E\left(\left(\mathcal{R}e\left\{\overline{\varphi_{i'j'}}(\varphi'\epsilon)_{i'j'}\right\}\mathcal{R}e\left\{(\varphi'(u\star X)_{-l})_{ij}\overline{(\varphi'\epsilon)}_{ij}\right\}\right.\right.$$

$$+ \quad \mathcal{R}e\left\{\overline{(\varphi'\epsilon)}_{ij}(\varphi'(u\star X)_{-l})_{ij}\right\}\mathcal{R}e\left\{\overline{\varphi_{i'j'}}(\varphi'\epsilon)_{i'j'}+\overline{\varphi_{i'j'}}(\varphi"\epsilon^2)_{i'j'}+|\varphi'\epsilon|^2_{i'j'}\right\}$$

$$+ \quad \mathcal{R}e\left\{\overline{\varphi}_{ij}(\varphi''\epsilon(u\star X)_{-l})_{ij}\right\}|\varphi'\epsilon|^2_{i'j'}+\mathcal{R}e\left\{\overline{(\varphi}_{ij}(\varphi'\epsilon_{-l})_{ij}\right\}|\varphi'\epsilon|^2_{i'j'}$$

$$+ \quad \left.\left.\mathcal{R}e\left\{(\varphi'(u\star X)_{-l})_{ij}\overline{(\varphi)}_{ij}\right\}|\varphi'\epsilon|^2_{i'j'}\right)\prod_{i"<j",\neq(i',j')\neq(i,j)}|\varphi_{i"j"}|^2\right)$$

où maintenant $\varphi_{ij}=\varphi_1(X_0^{(i)})-\varphi_1(X_0^{(j)})$, $\epsilon_{ij}=(\theta(\xi^*)\star\epsilon^{(i)})_0-(\theta(\xi^*)\star\epsilon^{(j)})_0$. On utilise la p-concentration des $X^{(i)}$, l'indépendance des variables ϵ et X, le fait que $E\left(|\epsilon_{-l}|_{ij}\right)=4$ et $E\left(|\epsilon_{ij}|^2\right)=4\|\theta(\xi^*)\|_2^2$ et $E\left(\epsilon_{i'}\overline{\epsilon_{ij}}\right)=2\|\theta(\xi^*)\|_2^2$ pour $i'=i$ et $=-2\|\theta(\xi^*)\|_2^2$ pour $i'=j$, et on obtient finalement

$$(p+1)!\frac{\partial^2 H(\theta(\xi^*),\sigma_0)}{\partial s_l\partial\sigma} = -16\sigma_0\theta_l\sum_{i<j}E\left(|\varphi_j'|^2\prod_{i'<j'}|\varphi_{i'j'}|^2\right)$$

$$- \quad 8\sigma_0\|\theta\|_2^2\sum_{i<j}E\left(\mathcal{R}e\left(\overline{\varphi'}_j\varphi"_j(u\star(X^{(i)}+X^{(j)}))_{-l}\right)\prod_{i'<j'}|\varphi_{i'j'}|^2\right)$$

$$- \quad 2\sigma_0\sum_{i<j}\sum_{i'<j'\neq(i,j)}E\left(\mathcal{R}e\left(\left(\overline{\varphi}_{i'j'}\varphi'_{i'j'}+|\varphi'_{i'j'}|^2\right)E\left(\epsilon_{i'}\overline{\epsilon}_{ij}\right)+2\overline{\varphi}_{i'j'}\varphi'_{j'}E\left(\epsilon_{i'j'}\overline{\epsilon}_{ij}\right)\right.\right.$$

$$+ \quad \left.8\overline{\varphi}_{ij}\varphi_j'\|\theta\|_2^2\right)(u\star X)_{-l})_{ij}|\varphi_j'|^2\prod_{i"<j",\neq(i',j')\neq(i,j)}|\varphi_{i"j"}|^2\Bigg)$$

$$- \quad 16\sigma_0\|\theta\|_2^2\sum_{i<j}\sum_{i'<j'\neq(i,j)}E\left(\mathcal{R}e\left(\overline{\varphi}_{ij}\varphi'_{ij}|\varphi'_{j'}|^2(u\star X)^{(i)}_{-l}\right)\prod_{i"<j"\neq(i,j)\neq(i',j')}|\varphi_{i"j"}|^2\right)$$

où encore, à un facteur $(p+1)!$ près

$$\frac{\partial^2 H(\theta(\xi^*),\sigma_0)}{\partial s_l \partial\sigma} = -16\sigma_0\theta_l \sum_{i<j} E\left(|\varphi'_j|^2 \prod_{i'<j'} |\varphi_{i'j'}|^2\right)$$

$$- \quad 8\sigma_0\|\theta\|_2^2 \sum_{i<j} E\left(\mathcal{R}e\left(\overline{\varphi'}_j\varphi"_j(u\star(X^{(i)}+X^{(j)}))_{-l}\right)\prod_{i'<j'}|\varphi_{i'j'}|^2\right)$$

$$- \quad 4\sigma_0\|\theta\|_2^2 \sum_{i<j} E\left(\mathcal{R}e\left((u\star X)^{(ij)}_{-l}|\varphi'_j|^2 \sum_{k=i+1,\neq j}^{p}\left(\overline{\varphi}_{ik}\varphi'_i + |\varphi'_{ik}|^2 + \overline{\varphi}_{ik}\varphi'_k\right) \prod_{i'<j'\neq(i,j)\neq(i,k)} |\varphi_{i'j'}|^2\right)\right.$$

$$- \quad \mathcal{R}e\left((u\star X)^{(ij)}_{-l}|\varphi'_j|^2 \sum_{k=j+1}^{p}\left(\overline{\varphi}_{jk}\varphi'_j + |\varphi'_{jk}|^2 + \overline{\varphi}_{jk}\varphi'_k\right) \prod_{i'<j'\neq(i,j)\neq(j,k)} |\varphi_{i'j'}|^2\right)$$

$$- \quad 2\mathcal{R}e\left((u\star X)^{(ij)}_{-l}|\varphi'_j|^2 \sum_{k=0}^{i-1}\left(\overline{\varphi}_{ki}\varphi'_i\right) \prod_{i'<j'\neq(i,j)\neq(k,i)} |\varphi_{i'j'}|^2\right)$$

$$+ \quad 2\mathcal{R}e\left((u\star X)^{(ij)}_{-l}|\varphi'_j|^2 \sum_{k=0,\neq i}^{j-1}\left(\overline{\varphi}_{kj}\varphi'_j\right) \prod_{i'<j'\neq(i,j)\neq(k,j)} |\varphi_{i'j'}|^2\right)\Bigg)$$

$$- \quad 16\sigma_0\|\theta\|_2^2 \sum_{i<j}\sum_{i'<j'\neq(i,j)} E\left(\mathcal{R}e\left((u\star X^{(ij)})_{-l}|\varphi'_j|^2\varphi'_j + (u\star X^{(i)})_{-l}|\varphi'_{j'}|^2\varphi'_{ij}\right)\overline{\varphi}_{ij}\prod_{i"<j"}|\varphi_{i"j"}|^2\right)$$

On procède de manière identique pour obtenir l'expression de la dérivée partielle en σ de H. Afin d'obtenir une condition suffisante garantissant que la différentielle seconde en s et σ est non nulle, on convole la différentielle seconde avec le filtre θ, et à $(p+1)!$ près

$$\left(D^2_{s,\sigma}H(\theta,\sigma_0)\star\theta\right)_0 = -16\sigma_0\|\theta\|_2^2 \sum_{i<j} E\left(|\varphi'_j|^2 \prod_{i'<j',\neq(ij)} |\varphi_{i'j'}|^2\right)$$

$$- \quad 8\sigma_0\|\theta\|_2^2 \sum_{i<j} E\left(\mathcal{R}e\left(\overline{\varphi'}_j\varphi_j"(X_0^{(i)}+X_0^{(j)})\prod_{i'<j',\neq(i,j)}|\varphi_{i'j'}|^2\right)\right)$$

$$- \quad 16\sigma_0\|\theta\|_2^2 \sum_{i<j}\sum_{i'<j'} E\left(\mathcal{R}e\left(\overline{\varphi}_{ij}\left(X_0^{(ij)}|\varphi'_j|^2\varphi'_j + X_0^{(i)}|\varphi'_{j'}|^2\varphi'_{ij}\right)\right)\prod_{i"<j",\neq(i',j'),\neq(i,j)}|\varphi_{i",j"}|^2\right)$$

Et comme $\sigma_0 \neq 0$ et $\|\theta\|_2$ non nulle on obtient l'hypothèse **(M11)** sous laquelle la différentielle seconde est non nulle.

Preuve du Lemme 3.3 :

On a pour tout n,

$$\sqrt{n}\ \left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma}) - d(\theta(\xi^*),\sigma_0)\right) = \sqrt{n}\ \left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma}) - \widehat{d}(\theta(\xi^*),\sigma_0)\right) + \sqrt{n}\ \left(\widehat{d}(\theta(\xi^*),\sigma_0) - d(\theta(\xi^*),\sigma_0)\right) \tag{11}$$

On procède à l'aide de développements des pseudo-moments conjugués empiriques autour de $\theta(\xi^*)$ et σ_0. On exprime le second terme de (11) en fonction des moments conjugués empiriques par inversion de la matrice $\Delta_{\varphi_2}(\sigma_0^2\|\theta(\xi^*)\|_2^2)$. L'ergodicité du processus $(X_k)_{k\in\mathbb{Z}}$ et la vitesse du terme de troncature $m(n)$ **(M9)**, permettent d'écrire le second terme de (11) en fonction de $\widehat{c}$ et c

$$\sqrt{n}\left(\widehat{d}(\theta(\xi^*),\sigma_0)-d(\theta(\xi^*),\sigma_0)\right)=\sqrt{n}\;\Delta_{\varphi_2}^{-1}(\sigma_0^2\|\theta\|_2^2)\;\left(\widehat{c}(\theta(\xi^*))-c(\theta(\xi^*))+o(\frac{1}{\sqrt{n}})\right) \tag{12}$$

On peut montrer que le premier terme de l'équation (11) s'exprime en fonction de $H_n(\theta(\xi^*);\sigma_0)$ de la manière suivante

$$\sqrt{n}\left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma})-\widehat{d}(\theta(\xi^*),\sigma_0)\right)=A\sqrt{n}H_n(\theta(\xi^*),\sigma_0)(1+o(1)) \tag{13}$$

où A est défini par

$$A=-\frac{D^1_{\xi,\sigma}d(\theta(\xi^*),\sigma_0)}{D^1_\sigma H(\theta(\xi^*),\sigma_0)}\begin{pmatrix} {}^tD^2_{\xi,\sigma}H(\theta(\xi^*),\sigma_0)\\ 1\end{pmatrix} \tag{14}$$

Avec la relation

$$H_n(\theta(\xi^*),\sigma_0)=D^1_d h(d(\theta(\xi^*),\sigma_0))\left(\widehat{d}(\theta(\xi^*),\sigma_0)-d(\theta(\xi^*),\sigma_0)\right)\;(1+o(1)) \tag{15}$$

Les équations (11), (12), (13) et (15) donnent

$$\sqrt{n}\left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma})-d(\theta(\xi^*),\sigma_0)\right)=\left(AD^1_dh+I_{(p+1)^2}\right)\Delta_{\varphi_2}^{-1}(\sigma_0^2\|\theta(\xi^*)\|_2^2)\sqrt{n}\left(\widehat{c}(\theta(\xi^*))-c(\theta(\xi^*))\right)(1+o(1))$$

où $D^1_dh=D^1_dh(d(\theta(\xi^*),\sigma_0))$ et $I_{(p+1)}$ est la matrice carré identité de taille $p+1$.
La distribution asymptotique de $\sqrt{n}\;(\widehat{c}(\theta(\xi^*))-c(\theta(\xi^*)))$ est donnée par l'hypothèse **(M10)**.
Pour conclure, il reste à montrer l'équation (13) et l'expression (14).

On établit au préalable le comportement asymptotique de l'estimateur du niveau de bruit $\widehat{\sigma}$. Pour ce faire, on utilise le développement à l'ordre 1 du gradient de J_n pour établir le comportement asymptotique de l'estimateur du niveau de bruit $\widehat{\sigma}$.

$$D^1J_n(\widehat{\theta},\widehat{\sigma})=D^1J_n(\theta,\sigma_0)+D^2J_n(\theta,\sigma_0)(\widehat{\xi}-\xi^*,\widehat{\sigma}-\sigma)(1+o(1))$$

$$\begin{aligned} 0 &= \begin{pmatrix} H_n(\theta,\sigma_0)\;{}^tD^1_\xi H_n(\theta,\sigma_0)\\ H_n(\theta,\sigma_0)D^1_\sigma H_n(\theta,\sigma_0)+\frac{\delta^2(n)}{2}\end{pmatrix}\\ &+ \left(\;H_n(\theta,\sigma_0)D^2H_n(\theta,\sigma_0)+\;{}^tD^1H_n(\theta,\sigma_0)D^1H_n(\theta,\sigma_0))\;\right)\begin{pmatrix}\widehat{\xi}-\xi^*\\ \widehat{\sigma}-\sigma_0\end{pmatrix}(1+o(1))\end{aligned}$$

Le système n'étant pas globalement inversible, on s'intéresse en particulier à la dernière ligne

$$\begin{aligned} 0 &= H_n(\theta,\sigma_0)D^1_\sigma H_n(\theta,\sigma_0)+\frac{\delta^2(n)}{2}\\ &+ \left(H_n(\theta,\sigma_0)D^2_{\xi,\sigma}H_n(\theta,\sigma_0)+D^1_\xi H_n(\theta,\sigma_0)D^1_\sigma H_n(\theta,\sigma_0)\;\right)(\widehat{\xi}-\xi^*)(1+o(1))\\ &+ \left(H_n(\theta,\sigma_0)D^2_\sigma H_n(\theta,\sigma_0)+(D^1_\sigma H_n(\theta,\sigma_0))^2\;\right)\;(\widehat{\sigma}-\sigma_0)(1+o(1))\end{aligned}$$

Maintenant, en introduisant l'équation (6), en multipliant par $\sqrt{n}$, grâce à la vitesse du terme de pénalisation $\sqrt{n}\delta^2(n)=o(1)$ **(M9)**, à la convergence presque sûre de H_n et $D^1_\xi H_n(\theta(\xi^*),\sigma_0)$ vers 0, garantie par l'ergodicité de $(X_k)_{k\in\mathbb{Z}}$ et le fait que $H(\theta(\xi^*),\sigma_0)=0$ et $D^1_\xi H(\theta(\xi^*),\sigma_0)=0$, on obtient

$$\sqrt{n}(\widehat{\sigma}-\sigma_0)=\frac{-1}{D^1_\sigma H(\theta(\xi^*),\sigma_0)}\sqrt{n}H_n(\theta(\xi^*),\sigma_0)(1+o(1)) \tag{16}$$

L'équation (13) s'obtient en injectant les développements au premier ordre de $\widehat{\xi}$ donné par (6) et de $\widehat{\sigma}$ donné par (16). La convergence presque sûre de $\widehat{\xi}$ et $\widehat{\sigma}$ vers ξ^* et σ^* respectivement, permet d'écrire le développement à l'ordre 1 du pseudo-moment conjugué empirique $\widehat{d}$ en ξ^* et σ_0

$$\sqrt{n}\ \left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma})-\widehat{d}(\theta(\xi^*),\sigma_0)\ \right)=\sqrt{n}\ D^1_{\xi,\sigma}\widehat{d}(\theta(\xi^*),\sigma_0)\left(\begin{array}{c}\widehat{\xi}-\xi^*\\ \widehat{\sigma}-\sigma_0\end{array}\right)(1+o(1))$$

$$\sqrt{n}\ \left(\widehat{d}(\theta(\widehat{\xi}),\widehat{\sigma})-\widehat{d}(\theta(\xi^*),\sigma_0)\ \right)=-\frac{D^1_{\xi,\sigma}\widehat{d}(\theta(\xi^*),\sigma_0)}{D^1_\sigma H_n(\theta(\xi^*),\sigma_0)}\left(\begin{array}{c}{}^tD^2_{\xi,\sigma}H_n(\theta(\xi^*),\sigma_0)\\ 1\end{array}\right)\sqrt{n}H_n(\theta(\xi^*),\sigma_0)(1+o(1))$$

Ce résultat couplé avec l'ergodicité de $(X_k)_{k\in\mathbb{Z}}$ assurant la convergence presque sûre de $D^1_{\xi,\sigma}\widehat{d}(\theta(\xi^*),\sigma_0)$ vers $D^1_{\xi,\sigma}d(\theta(\xi^*),\sigma_0)$ et de $D^1_\sigma H_n(\theta(\xi^*),\sigma_0)$ vers $D^1_\sigma H(\theta(\xi^*),\sigma_0)$ ainsi que $D^2_{\xi,\sigma}H_n(\theta(\xi^*),\sigma_0)$ vers $D^2_{\xi,\sigma}H(\theta(\xi^*),\sigma_0)$ et le lemme (3.2), constitue l'équation (14).
Enfin la matrice $B=AD^1_dh+I_{(p+1)^2}$ est non nulle car, en omettant d'écrire les variables en lesquelles sont calculées l'ensemble des fonctionnelles suivantes, on a

$$\begin{aligned} B &= -\frac{D^1_\xi d\ {}^tD^2_{\xi\sigma}HD^1_dh}{D^1_\sigma H}-\frac{D^1_\sigma dD^1_dh}{D^1_\sigma H}+I_{(p+1)^2}\\ &= \frac{-1}{D^1_\sigma H}\left(\left(D^1_\xi d\quad D^1_\sigma d\right)\left(\begin{array}{c}{}^tD^2_{\xi,\sigma}H\\ 1\end{array}\right)D^1_dh-D^1_\sigma HI_{(p+1)^2}\right)\end{aligned}$$

Or la trace de la matrice consituant le premier terme du membre de droite vaut

$$\begin{aligned}\text{trace } D^1d\left(\begin{array}{c}{}^tD^2_{\xi,\sigma}H\\ 1\end{array}\right)D^1_dh &= \sum_{i,j=0}^p\left(\sum_{l=1}^q\frac{\partial d_{ij}}{\partial\xi_l}\frac{\partial h}{\partial d_{ij}}\frac{\partial^2H}{\partial\xi_l\partial\sigma}\right)+\sum_{i,j=1}^p\frac{\partial d_{ij}}{\partial\sigma}\frac{\partial h}{\partial d_{ij}}\\ &= \sum_{l=1}^q\frac{\partial H}{\partial\xi_l}\frac{\partial^2H}{\partial\xi_l\partial\sigma}+D^1_\sigma H\\ &= D^1_\sigma H\end{aligned}$$

car $D^1_\xi H=0$. Et la trace de $D^1_\sigma HI_{(p+1)^2}$ vaut $(p+1)^2D^1_\sigma H$, ainsi comme $p\geq 2$ **(M1)** , B est non nulle. On remarque que la matrice B est non inversible car ${}^tD^1_dh$ est vecteur propre à gauche de la matrice $\left(D^1_\xi d\quad D^1_\sigma d\right)\left(\begin{array}{c}{}^tD^2_{\xi,\sigma}H\\ 1\end{array}\right)D^1_dh$ associé à la valeur propre $D^1_\sigma H(\theta(\xi^*),\sigma_0)$.

Preuve du théorème 3.3 :
A tout vecteur ${}^tc=(c_1,\ldots,c_p)$ de $\mathbb{C}^p$, on associe un vecteur ${}^tv(c)=(v_0(c),\ldots,v_p(c))\in\mathbb{C}^{p+1}$, normé, défini par

$$\forall y\in\mathbb{C},\quad \sum_{k=0}^p v_k(c)y^k=v_p(c)\prod_{i=1}^p(y-c_i) \tag{17}$$

En d'autres termes, si c représente les racines d'un polynôme de degré p, $v(c)$ représente le vecteur normé dont les coordonnées sont les coefficients de ce polynôme.
Cette application est bijective de $\mathbb{C}^p$ dans $\mathbb{C}^{p+1}\cap\{\|\ .\ \|_2=1\}$. Le vecteur $\widehat{a}=(\widehat{a}_1,\ldots,\widehat{a}_p)$ est, par définition, le zéro du polynôme en $\varphi_1(y)$ de degré p construit sur $\widehat{v}^*$ le vecteur propre normé associé à la plus petite valeur propre $\widehat{\lambda}_0$ de $D(\widehat{d}(\widehat{\theta};\widehat{\sigma})$. En d'autres termes, $\widehat{a}$ est défini par

$$\forall x\in\mathbb{R},\quad \sum_{k=0}^p\widehat{v}^*_k\varphi_1(x)^k=\widehat{v}^*_p\prod_{i=1}^p(\varphi_1(x)-\varphi_1(\widehat{a}_i))$$

Le vecteur $\widehat{a}$ est consistant, donc pour n suffisamment grand, $\widehat{a}$ a toutes ses coordonnées réelles. En notant ${}^t\underline{\varphi}_1(b)=(\varphi_1(b_1),\ldots,\varphi_1(b_p))$, on a

$$\widehat{v}^*=v(\underline{\varphi}_1(\widehat{a}))$$

où v est le vecteur défini par (17).
On définit maintenant

$$\begin{array}{rcl} \forall b \in \mathbb{R}^p, \quad \widehat{g}(b) & = & {}^t\overline{v(\underline{\varphi_1}(b))}\, D(\widehat{d}(\widehat{\theta};\widehat{\sigma}))\, v(\underline{\varphi_1}(b)) \\ & \text{et} & \\ g(b) & = & \overline{v(\underline{\varphi_1}(b))}\, D(d(\theta;\sigma_0))\, v(\underline{\varphi_1}(b)) \end{array}$$

Les applications $\widehat{g}$ et g atteignent leur minimum respectif en $\widehat{a}$ et a. On procéde classiquement par un développement de Taylor du gradient de $\widehat{g}$ au premier ordre en a pour obtenir la loi asymptotique de $\widehat{a}$. L'estimateur $\widehat{a}$ étant consistant, $\tilde{\widehat{a}}$ l'est également. On a, pour tout n,

$$0 = D_b^1\widehat{g}(a) + {}^t(\widehat{a}-a)D_b^2\widehat{g}(a)(1+o(1)) \tag{18}$$

Le terme de second ordre $D_b^2\widehat{g}(a)$ tend, en probabilité, vers $D_b^2 g(a)$ grâce à l'ergodicité du processus X. Par construction, le pseudo-moment conjugué $d(s,\sigma)$ est un vrai moment conjugué de $\varphi_1(X_0)$ lorsque $s=\theta$ et $\sigma=\sigma_0$. Ainsi, on peut également écrire l'application g sous la forme suivante

$$\forall b, \quad g(b) = E\left(\left|v_p(\underline{\varphi_1}(b))\right|^2 \prod_{i=1}^{p} \left|\varphi_1(X_0)-\varphi_1(b_i)\right|^2\right) \tag{19}$$

La différentielle seconde de g en a s'obtient facilement à partir de l'expression (19), en utilisant la propriété de concentration de la distribution de X_0 sur les points $a_1,\dots,a_p$. On obtient

$$D_b^2 g(a) = 2|v_p^*|^2\, A$$

où A est une matrice diagonale définie par

$$A = \begin{pmatrix} |\varphi_1'(a_1)|^2 K_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & |\varphi_1'(a_p)|^2 K_p \end{pmatrix} \quad \text{où} \quad K_j = E\left(\prod_{i=1,\neq j}^{p} |\varphi_1(X_0)-\varphi_1(a_i)|^2\right)$$

En utilisant la concentration de X_0 sur les p points de support $(a_1,\dots,a_p)$, l'injectivité de φ_1 et le fait que $v_p^* \neq 0$ car sinon nous aurions moins de p points de support, on obtient que $D_b^2 g(\tilde{a})$ est définie positive. On écrit le gradient de $\widehat{g}$ et celui de g à partir de leur définition initiale en utilisant la forme hermitienne de $D(d)$ pour tout d. Ainsi le développement (18) devient

$$\begin{array}{rcl} (\widehat{a}-a) & = & -\dfrac{A^{-1}}{2|v_p^*|^2}\left({}^t v^* \overline{\left(D(\widehat{d}(\widehat{\theta},\widehat{\sigma})) - D(d(\theta,\sigma_0))\right)} \overline{D_b^1 v(\underline{\varphi_1}(a))} \right. \\ & + & \left. {}^t\underline{v^*}\left(D(\widehat{d}(\widehat{\theta},\widehat{\sigma})) - D(d(\theta,\sigma_0))\right) \overline{D_b^1 v(\underline{\varphi_1}(a))}\right) \end{array}$$

ainsi

$$\sqrt{n}(\widehat{a}-a) = -\frac{A^{-1}}{2|v_p^*|^2} B\left(\widehat{d}(\widehat{\theta},\widehat{\sigma}) - d(\theta,\sigma_0)\right)(1+o(1)) \tag{20}$$

où $B \in \mathcal{M}_{p*(p+1)^2}$ est défini par colonne. On indexe chacune des $(p+1)^2$ colonnes par le couple $(i,j) \in \{0,\dots,p\}$ et

$$B_{.,(i,j)} = {}^t\overline{D_b^1 v_i(\underline{\varphi_1}(a))} v_j^* + {}^t D_b^1 v_j(\underline{\varphi_1}(a))\overline{v_i^*}$$

Et d'après le lemme 3.3, le vecteur des pseudo-moments conjugués empirique $\widehat{d}(\widehat{\theta};\widehat{\sigma})$ converge en loi, à la vitesse racine de n, vers une loi gaussienne centrée de matrice de covariance Γ_d. Et comme A est une matrice symétrique et réelle, on a finalement

$$\lim_{n\to+\infty} \sqrt{n}(\widehat{a}-a) = \mathcal{N}\left(0, \frac{1}{4|v_p^*|^4} A^{-1} B \Gamma_d\, {}^t\overline{B} A^{-1}\right) \quad \text{en loi}$$

On s'attache maintenant à établir la distribution asymptotique de $\widehat{\pi}$ qui se déduit aisément de celle des $\widehat{a}$. Soit

$$\begin{array}{rcl} f : [0,1]^p \otimes \mathbb{R}^p & \longrightarrow & \mathbb{C}^p \\ (\beta, b) & \longmapsto & a(\ldots, \sum_{j=1}^{p} \beta_j \varphi_1(b_j)^k, \ldots)_{k=0,\ldots,p-1} \end{array}$$

Par définition de $\widehat{\pi}$, on a

$$\begin{array}{rcl} f(\widehat{\pi}, \widehat{a}) & = & \left(\widehat{d}_{i,0}(\widehat{\theta}, \widehat{\sigma})\right)_{i=0,\ldots,p-1} = P\widehat{d}(\widehat{\theta}, \widehat{\sigma}) \\ \text{de plus,} & & \\ f(\pi, a) & = & (d_{i,0}(\theta, \sigma))_{i=0,\ldots,p-1} = Pd(\theta, \sigma) \end{array}$$

où P est la matrice projection de $\mathbb{C}^{(p+1)^2}$ dans $\mathbb{C}^p$ telle que $\forall w \in \mathbb{C}^{(p+1)^2}, Pw = v$ avec $v = (w_1, w_{(p+2)}, \ldots, w_{k(p+1)+1}, \ldots, w_{p^2})_{k=0,\ldots,(p-1)}$.
L'application f est continuement différentiable en tout point (β, b), et

$$\forall \beta \in [0,1]^p, \quad \lim_{n\to+\infty} D^1_\beta f(\beta, \widehat{a}) = D^1_\beta f(\beta, a) \quad p.s.$$

avec

$$D^1_\beta f(\pi, a) = \begin{pmatrix} 1 & \ldots & 1 \\ \varphi_1(a_1) & \ldots & \varphi_1(a_p) \\ \vdots & & \vdots \\ \varphi_1(a_1)^{p-1} & \ldots & \varphi_1(a_p)^{p-1} \end{pmatrix}$$

$$D^1_b f(\pi, a) = \begin{pmatrix} 0 & \ldots & 0 \\ & \left(\pi_j \varphi_1'(a_j) i \varphi_1(a_j)^{i-1}\right) & \\ & & \end{pmatrix}_{i=1,\ldots,p-1; j=1,\ldots,p}$$

On note ces matrice A_{φ_1} et $A_{\varphi_1'}$ respectivement.
Maintenant, en utilisant la consistance de $\widehat{\pi}$ et de $\widehat{a}$, on effectue un développement de f à l'ordre 1 en (β, b) autour de (π, a). On obtient

$$f(\widehat{\pi}, \widehat{a}) = f(\pi, a) + \left(D^1_b f(\pi, a)(\widehat{a} - a) + D^1_\beta f(\pi, a)(\widehat{\pi} - \pi)\right)(1 + o(1))$$

Enfin, A_{φ_1} est inversible, et l'équation (20) donne l'expression asymptotique de $\sqrt{n}(\widehat{a} - a)$ en fonction des pseudo-moments conjugués. Ainsi,

$$\sqrt{n}(\widehat{\pi} - \pi) = A^{-1}_{\varphi_1} \left(A_{\varphi_1'} \frac{A^{-1}}{2|v_p^*|^2} B - P\right) \sqrt{n}(\widehat{d} - d)(1 + o(1))$$

ce qui avec le lemme (3.3) achève la preuve.

References

[CL95] R. Chen and T. H. Li. Blind restoration of linearly degraded discrete signals by gibbs sampler. *IEEE Transactions on Signal Processing*, 43:2410–2413, 1995.

[D.H92] D.Huang. Symmetric solutions and eigenvalue problems of toeplitz systems. *IEEE Trans. Acous. Speech Signal Processing*, 40:3069–3074, 1992.

[Dou94] P. Doukhan. Mixing properties and examples. In *Lecture Notes in Statistics*, volume 85. Springer, springer verlag edition, 1994.

[Gau97] E. Gautherat. *Déconvolution aveugle des systèmes linéaires aléatoires discrets bruités ou non.* Thèse, Université Evry-Val d'Essonne, 1997.

[GG96] F. Gamboa and E. Gassiat. Blind deconvolution of discrete linear systems. *Annals of Statistics*, 24(5):1964–1981, 1996.

[GG98] E. Gassiat and E. Gautherat. Identification of noisy linear systems with discrete random input. *IEEE Transaction on Information Theory*, 44:1941–1952, 1998.

[GG99] E. Gassiat and E. Gautherat. Speed of convergence for the blind deconvolution of a linear systems with discrete random input. *Annals of Stat.*, 27:1684–1705, 1999.

[HH80] P. Hall and C. C. Heyde. *Martingale limit theory and its applications.* Academic press, 1980.

[LC95] J.S. Liu and R. Chen. Blind deconvolution via sequential imputations. *Journal of Amerrican Statistical Association*, 90:567–576, 1995.

[Li95] T. H. Li. Blind deconvolution of linear systems with multilevel nonstationary inputs. *Annals of Statistics*, 23:690–704, 1995.

[LS00] L. Li and T. Speed. Parametric deconvolution of positive spike trains. *Annals of Statistics*, 28(5):1279–1301, 2000.

[LS01] K. Li and K. Shedden. Monte carlo deconvolution of digital signals guided by the inverse filter. *Journal of the American Statistical Association*, 96(455):1014–1021, 2001.

[Mak81] J. Makhoul. On the eigenvect of symmetric toeplitz matrices. *IEEE Acoust Speech Signal Processing*, 29:868–872, 1981.

[Pet75] V. Petrov. *Sums of independent variables.* Springer Verlag, 1975.

[RR90] A. Ralston and P. Rabinowitz. *A First Course in Numerical Analysis*, volume 36. McGraw-Hill, 1990. New York.