



HAL
open science

Développements récents sur l'estimation fonctionnelle par méthodes d'ondelettes

Christophe Chesneau

► **To cite this version:**

Christophe Chesneau. Développements récents sur l'estimation fonctionnelle par méthodes d'ondelettes. Statistiques [math.ST]. Université de Caen, 2014. tel-00975601

HAL Id: tel-00975601

<https://theses.hal.science/tel-00975601v1>

Submitted on 8 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Caen Basse-Normandie

Mémoire d'habilitation

Développements récents
sur l'estimation fonctionnelle
par méthodes d'ondelettes

Christophe Chesneau

Caen, le 30 mars 2014

Rapporteurs

Youming Liu, Université de Pékin,
B.L.S. Prakasa Rao, Université de Hyderabad,
Theofanis Sapatinas, Université de Chypre,

Examineurs

Fabienne Comte, Université Paris 5,
Arnak Dalalyan, ENSAE,
Djamal Louani, Université de Reims,
Mohamed Didi Biha, Université de Caen,
Jalal Fadili, ENSI Caen.

Date de soutenance

Jeudi 03 Avril 2014

Remerciements

Je tiens en premier lieu à remercier les membres du LMNO de l'Université de Caen Basse-Normandie pour le soutien qu'ils m'ont apporté au cours de ces dernières années.

Je remercie également mes collaborateurs pour le temps, le travail et le plaisir partagés : [ORDRE ALÉATOIRE] Thomas Willer, Mohammad Abbaszadeh, Hassan Doosti, Jalal Fadili, Gabriel Peyré, Charles Dossal, Fabienne Comte, Jean-Luc Starck, Isha Dewan, Maher Kachour, Esmaeil Shirazi, Yogendra Chaubey, Guillaume Lecué, Fabien Navarro, Mohamed Hebiri, Jan Bulla, Ingo Bulla, Mergner Sacha, André Sesboüé, Adrien Mallet, Micheline Guéguen, Nathalie Demasures, François Kauffmann, Nargess Hosseinioun, Taoufik Sassi, Bertrand Maillot, Dimitris Karlis.

Je veux témoigner ma profonde gratitude à B.L.S. Prakasa Rao, Theofanis Sapatinas et Youming Liu qui m'ont fait l'honneur d'accepter d'être les rapporteurs de cette HDR.

L'expression de ma reconnaissance et de ma considération s'adressent également à Fabienne Comte, Arnak Dalalyan, Djamal Louani, Mohamed Didi Biha et Jalal Fadili qui ont bien voulu accepter de faire partie du jury.

La participation de ces chercheurs prestigieux à mon HDR est un immense privilège, qu'ils en soient tous chaleureusement remerciés.

Mes derniers remerciements vont à mes amis et à ma famille pour leur soutien, leur affection et leur joie de vivre. Merci donc à Stéphanie, Annabelle, Baptiste, Christian, Marie et Émilie.

Table des matières

1	Introduction	6
1.1	Avant-propos	6
1.2	Cadre général	7
1.3	Notion de performance	8
2	Méthodes d'ondelettes	11
2.1	Base d'ondelettes	11
2.2	Méthode d'ondelettes et parcimonie	12
2.3	Estimateurs linéaires en ondelettes	13
2.4	Estimateurs de seuillage en ondelettes	14
3	Seuillage par blocs	18
3.1	Motivation	18
3.2	Résultat général I	19
3.3	Applications	22
3.3.1	Régression nonparamétrique	22
3.3.2	Densités pondérées	23
3.3.3	Déconvolution	24
3.4	Cadre multidimensionnel	25
3.5	Perspectives I	25
4	Double seuillage	26
4.1	Motivation	26
4.2	Résultat général II	27

4.3	Applications	29
4.3.1	Régression nonparamétrique	29
4.3.2	Régression nonparamétrique multicanal	30
4.3.3	Convolution de densités mélangées	32
4.4	Perspectives II	33
5	Estimation en ondelettes et dépendance	34
5.1	Motivation	34
5.2	Résultat général III	34
5.3	Applications	37
5.3.1	Déconvolution de densités	37
5.3.2	Modèle à bruit multiplicatif	39
5.3.3	Régression nonparamétrique	40
5.4	Perspectives III	43
6	Compléments	45

1 Introduction

1.1 Avant-propos

Mes premiers pas dans la recherche ont eu lieu en 2003 dans le LPMA à l'Université Paris 6, marquant le commencement d'une thèse de doctorat. Ils ont été guidés par Gérard Kerkyacharian sur le terrain de l'estimation fonctionnelle par méthodes d'ondelettes. J'ai alors contribué à la détermination de vitesse minimax pour des modèles statistiques émergents, la construction d'estimateurs utilisant des bases d'ondelettes déformées et le développement de l'estimateur de seuillage par blocs en ondelettes. La soutenance a eu lieu le 07 Décembre 2006 devant le jury composé de Anatoli Juditsky, Laurent Cavalier, Gérard Kerkyacharian, Alexandre Tsybakov et Stéphane Bouche-ron.

En 2007, le LMNO de l'Université de Caen Basse-Normandie m'a ouvert ses portes. Dès lors, j'ai pu développer ma thématique, élargir mon horizon mathématique et collaborer avec des chercheurs prestigieux. J'ai également eu l'opportunité et le plaisir de co-encadrer des étudiants motivés, notamment Maher Kachour (enseignant-chercheur à l'IDRAC de Lyon) avec Jalal Fadili, et Fabien Navarro (ATER au LMJL à Nantes) avec Jalal Fadili et Taoufik Sassi. L'envie d'encadrer et de diriger les recherches de mon propre chef est le principal moteur de cette HDR.

À ce jour, l'essentiel de mes travaux s'articule autour de l'estimation de fonctions inconnues émanant de phénomènes aléatoires. La richesse de ces phénomènes combinée avec l'amélioration constante des méthodes d'estima-

tion nourrissent mon intérêt pour le sujet. J'ai toutefois choisi de me spécialiser dans les méthodes d'ondelettes. La principale raison est qu'elles bénéficient d'une grande faculté d'adaptation à la complexité du problème posé, tout en ayant des performances d'estimation remarquables. Cela est présenté dans la première partie de ce rapport. Les trois autres parties concernent trois de mes résultats les plus significatifs. En outre, ils sont applicables à une multitude de modèles statistiques, ouvrant ainsi un large champ d'applications, et améliorent certains aspects de résultats existants.

1.2 Cadre général

Le point de départ est l'étude d'un phénomène aléatoire dont la modélisation fait intervenir une fonction inconnue f . Par exemple, f peut être la densité d'une v.a.r. représentant un caractère quantitatif continu (taille, superficie, teneur...) ou un signal perturbé par du bruit. On dispose alors de n données correspondantes aux réalisations de n éléments aléatoires issues de cette modélisation. L'objectif est d'estimer f du mieux possible à l'aide des données. Pour ce faire, partant des n éléments aléatoires associés, on cherche à construire un estimateur \hat{f} de f aussi performant que possible. Parmi les méthodes d'estimation existantes, il y a les méthodes à noyaux, les méthodes par polynômes locaux, les méthodes pénalisées et les méthodes d'ondelettes. Le sujet principal de ce rapport concerne ces dernières. Afin de motiver ce choix, la notion de "performance d'un estimateur" doit être clarifiée.

Dans ce qui suit, pour raison de simplicité, nous supposons que le support

de f est l'intervalle $[0, 1]$ et que

$$f \in \mathbb{L}^2([0, 1]) = \left\{ f : [0, 1] \rightarrow \mathbb{R}; \|f\|_2 = \left(\int_0^1 f^2(x) dx \right)^{1/2} \right\}.$$

1.3 Notion de performance

Pour évaluer la performance d'un estimateur \hat{f} de f , plusieurs outils mathématiques sont possibles. Parmi ceux-ci, nous considérons le risque \mathbb{L}^2 défini par :

$$R(\hat{f}, f) = \mathbb{E}(\|\hat{f} - f\|_2^2).$$

Ce risque mesure l'erreur quadratique intégrée moyenne que commet \hat{f} dans l'estimation de f . Naturellement, plus petite est cette erreur, meilleur sera l'estimateur, et plus on dispose d'informations sur f (i.e., plus n est grand) plus la précision de \hat{f} dans l'estimation de f doit être améliorée. Afin de mesurer cette précision, on étudie la suite $(\varphi_n)_n$ convergeant le plus rapidement possible vers 0 telle qu'il existe une constante $C > 0$ satisfaisant

$$R(\hat{f}, f) \leq C\varphi_n.$$

Une telle suite est appelée vitesse de convergence de \hat{f} .

Cependant, cette étude ne peut se faire sans des hypothèses de régularité sur f . On est donc contraint de supposer que f appartient à un espace fonctionnel \mathcal{S} évaluant cette régularité. Naturellement, f étant inconnue, plus riche est \mathcal{S} , plus réaliste est cette hypothèse. Cela motive notre intérêt pour

les boules de Besov, notée $B_{p,r}^s(M)$, contenant des fonctions de natures aussi diverses que variées. La définition mathématique précise de ces boules est donnée en appendice.

Quand cela est possible, nous adopterons l’approche minimax. L’enjeu de cette approche est de déterminer la vitesse de convergence atteinte par le “meilleur” estimateur pour la “pire” des fonctions à estimer appartenant à \mathcal{S} . Autrement écrit, on s’intéresse à la suite ρ_n définie par

$$\rho_n = \inf_{\hat{f}} \sup_{f \in \mathcal{S}} R(\hat{f}, f),$$

où l’infimum est pris sur l’ensemble de tous les estimateurs \hat{f} de f possible. La suite ρ_n est appelée vitesse optimale (au sens minimax). Plus la vitesse de convergence du risque d’un estimateur est proche de cette vitesse optimale, plus l’estimateur sera considéré performant.

La notion de performance se complète avec celle d’adaptativité. Un estimateur est dit adaptatif si sa construction ne dépend ni de f , ni de sa régularité inconnue. Il est donc réaliste et son implémentation pratique peut être envisagée. Les estimateurs non adaptatifs sont néanmoins intéressants pour poser les bases théoriques de l’approche minimax ; c’est généralement ce type d’estimateur que l’on construit pour exhiber une borne supérieure fine du risque, étape nécessaire à la détermination de la vitesse optimale.

Une description complète de ce cadre statistique est donnée dans le livre de référence [95].

Dans la section suivante, nous expliquons pourquoi et comment les mé-

thodes d'ondelettes nous permettent de construire des estimateurs \hat{f} à la fois adaptatifs et performants.

2 Méthodes d'ondelettes

2.1 Base d'ondelettes

L'origine des méthodes d'ondelettes repose sur la décomposition de f sur une base d'ondelettes. Pour fixer les notations, nous considérons la base d'ondelettes sur $[0, 1]$ introduite par [45] brièvement décrite ci-après.

Soient N un entier positif, ϕ et ψ les fonctions d'ondelettes initiales de la famille Daubechies $db2N$ et

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k).$$

Avec un traitement approprié aux bornes $\{0\}$ et $\{1\}$, il existe un entier τ tel que, pour tout $\ell \geq \tau$, la collection

$$\mathcal{W} = \{\phi_{\ell,k}, k \in \{0, \dots, 2^\ell - 1\}; \psi_{j,k}; j \in \{\ell, \dots, \infty\}, k \in \{0, \dots, 2^j - 1\}\}$$

forme une base orthonormée de $\mathbb{L}^2([0, 1])$.

Pour tout entier $\ell \geq \tau$, f peut être décomposée sur \mathcal{W} comme

$$f(x) = \sum_{k=0}^{2^\ell-1} \alpha_{\ell,k} \phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}(x), \quad x \in [0, 1], \quad (1)$$

où $\alpha_{j,k}$ et $\beta_{j,k}$ désignent les coefficients d'ondelettes de f définis par

$$\alpha_{j,k} = \int_0^1 f(x) \phi_{j,k}(x) dx, \quad \beta_{j,k} = \int_0^1 f(x) \psi_{j,k}(x) dx. \quad (2)$$

L'idée principale est d'écrire f comme la somme de deux termes complémentaires : un terme d'approximation correspondant à l'allure globale de f et un terme de détail correspondant à ses singularités (locales). Cette décomposition a la particularité d'être parcimonieuse : seul un petit nombre de $\beta_{j,k}$, caractérisés par leur grande magnitude, contiennent l'essentiel des détails de f . La puissance des méthodes d'ondelettes repose sur cette caractéristique.

Un autre intérêt des bases d'ondelettes est d'exprimer les espaces fonctionnels habituels, comme $B_{p,r}^s(M)$, par des coefficients d'ondelettes.

Ces aspects sont détaillés dans [83] [81].

2.2 Méthode d'ondelettes et parcimonie

L'estimation en ondelettes de f passe par l'estimation des coefficients inconnus $\alpha_{j,k}$ et $\beta_{j,k}$, généralement plus simple à mettre en œuvre. Cependant, comme il y a une infinité de $\beta_{j,k}$ et que chaque estimation engendre de l'erreur, une sélection de ces coefficients doit être faite. La nature parcimonieuse de la décomposition en ondelettes intervient à ce niveau : afin de limiter l'erreur d'estimation, le schéma classique est d'estimer tous les $\alpha_{j,k}$, lesquels sont en nombre fini, et uniquement les $\beta_{j,k}$ ayant de grandes magnitudes. Les autres sont réduits à zéro. Un estimateur en ondelettes de f s'obtient par la reconstruction des estimateurs sélectionnés sur la base d'ondelettes initiale. Une possibilité d'estimateur prenant en compte ces considérations est donc

de la forme :

$$\hat{f}(x) = \sum_{k=0}^{2^\tau-1} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{\infty} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \mathbf{1}_{\mathcal{D}_{j,k}} \psi_{j,k}(x),$$

où $\hat{\alpha}_{\tau,k}$ désigne un estimateur de $\alpha_{\tau,k}$, $\hat{\beta}_{j,k}$ un estimateur de $\beta_{j,k}$, $\mathbf{1}$ la fonction indicatrice et

$$\mathcal{D}_{j,k} = \{(j, k) \in \{\tau, \dots, \infty\} \times \{0, \dots, 2^j - 1\}; |\hat{\beta}_{j,k}| \text{ est suffisamment grand}\}.$$

2.3 Estimateurs linéaires en ondelettes

Les premiers estimateurs en ondelettes ont été introduits par [59] [72, 73] dans le cadre de l'estimation de densités. Leur spécificité repose sur une troncature de la somme du terme de détail estimé à un certain niveau j_1 . On parle alors d'estimateurs linéaires en ondelettes. Leur expression analytique est de la forme :

$$\hat{f}_*(x) = \sum_{k=0}^{2^\tau-1} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \psi_{j,k}(x),$$

où $\hat{\alpha}_{\tau,k}$ désigne un estimateur de $\alpha_{\tau,k}$ et $\hat{\beta}_{j,k}$ un estimateur de $\beta_{j,k}$.

L'idée sur laquelle repose cet estimateur est la suivante : plus f est régulière, plus $|\beta_{j,k}|$ décroît à mesure que j croît. Par exemple, sous l'hypothèse que $f \in B_{\infty,\infty}^s(M)$, il existe une constante $C > 0$ telle que

$$|\beta_{j,k}| \leq C 2^{-j(s+1/2)}.$$

La troncature permet donc de faire une sélection brutale des $\beta_{j,k}$ les plus intéressants à estimer.

Pour de nombreux modèles statistiques, si f est suffisamment régulière, il existe un niveau j_1 tel que \hat{f}_* soit performant. Typiquement, si $f \in B_{p,r}^s(M)$ avec $p \geq 2$, il existe un niveau j_1 tel que \hat{f}_* atteigne la vitesse optimale. Il est même possible d'estimer ce j_1 si celui-ci dépend de la régularité de f , rendant ainsi \hat{f}_* adaptatif. Cependant, peu importe le choix de j_1 , les bonnes performances de \hat{f}_* s'écroulent si f est irrégulière. C'est notamment le cas si $f \in B_{p,r}^s(M)$ avec $p \in [1, 2)$. Pour améliorer ce point, une sélection plus fine des $\hat{\beta}_{j,k}$ doit être faite. Des alternatives satisfaisantes sont proposées par les estimateurs de seuillage en ondelettes présentées ci-après.

Plus d'éléments sur les estimateurs linéaires en ondelettes sont consultables dans [69]. Pour l'étude de \hat{f}_* construit avec une estimation de j_1 , voir [1].

2.4 Estimateurs de seuillage en ondelettes

Les estimateurs de seuillage en ondelettes ont été introduits par [52, 53] dans le cadre du modèle de régression à pas équidistants, puis développés par [54] dans le contexte de l'estimation de densités. Leur originalité repose sur une sélection individuelle des $\hat{\beta}_{j,k}$, en plus d'une troncature large de la somme du terme de détail estimée. Cette sélection s'opère par seuillage : les estimateurs $\hat{\beta}_{j,k}$ dont la magnitude dépassent un certain seuil sont gardés, les autres sont supprimés. Le plus célèbre estimateur de seuillage en ondelettes

est celui utilisant un seuillage dur. Son expression analytique est de la forme :

$$\hat{f}_0(x) = \sum_{k=0}^{2^\tau-1} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \kappa \lambda_j\}} \psi_{j,k}(x), \quad (3)$$

où $\hat{\alpha}_{\tau,k}$ désigne un estimateur de $\alpha_{\tau,k}$, $\hat{\beta}_{j,k}$ un estimateur de $\beta_{j,k}$, $\mathbf{1}$ la fonction indicatrice, κ une constante et λ_j une quantité dépendante de n et j .

Pour de nombreux modèles statistiques, sous des hypothèses de régularité larges, on peut déterminer un niveau j_1 et un seuil $\kappa \lambda_j$ tels que \hat{f}_0 soit à la fois performant et adaptatif. Typiquement, si $f \in B_{p,r}^s(M)$ avec $p \geq 1$, \hat{f}_0 atteint la vitesse de convergence “presque” optimale.

Plus précisément, en combinant les résultats de [74] et [14], on est en mesure de montrer le résultat suivant :

Résultat général initial. *Soient f une fonction inconnue que l'on souhaite estimer via n éléments aléatoires, (1) la décomposition en ondelettes associée et \hat{f}_0 l'estimateur (3) défini avec*

$$\lambda_j = 2^{\delta j} \sqrt{\frac{\ln n}{n}}, \quad 2^{j_1} = \left[\left(\frac{n}{\ln n} \right)^{1/(2\delta+1)} \right].$$

On suppose que $\hat{\alpha}_{j,k}$, $\hat{\beta}_{j,k}$, δ et κ vérifient :

(A1) *Il existe une constante $C > 0$ telle que, pour tout $j \in \{\tau, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,*

$$\mathbb{E} (|\hat{\alpha}_{j,k} - \alpha_{j,k}|^2) \leq C 2^{2\delta j} \frac{1}{n}.$$

(A2) Il existe une constante $C > 0$ telle que, pour tout $j \in \{\tau, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\mathbb{E} \left(|\hat{\beta}_{j,k} - \beta_{j,k}|^4 \right) \leq C 2^{4\delta j} \frac{1}{n^2}.$$

(A3) Il existe une constante $C > 0$ telle que, pour tout $j \in \{\tau, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\mathbb{P} \left(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \frac{\kappa}{2} \lambda_j \right) \leq C \frac{1}{n^2}.$$

Alors, sous l'hypothèse que $f \in B_{p,r}^s(M)$ avec $M > 0$, $r \geq 1$ et $\{p \geq 2$ et $s \in (0, N)\}$ ou $\{p \in [1, 2)$ et $s \in ((2\delta + 1)/p, N)\}$, il existe une constante $C > 0$ telle que

$$R(\hat{f}_0, f) \leq C \left(\frac{\ln n}{n} \right)^{2s/(2s+2\delta+1)}.$$

Pour de nombreux modèles statistiques,

- les hypothèses **(A1)**, **(A2)** et **(A3)** sont satisfaites par des estimateurs naturels (sans biais ou convergents) de $\alpha_{j,k}$ et $\beta_{j,k}$. Les preuves peuvent faire appel à des inégalités probabilistes techniques, notamment l'inégalité de Rosenthal [92] pour **(A2)**, et l'inégalité de Bernstein [88] pour **(A3)**.
- $(\ln n/n)^{2s/(2s+2\delta+1)}$ correspond à la vitesse “presque” optimale. La nuance “presque” est dûe au facteur logarithmique, souvent de trop (et incontournable pour \hat{f}_0 [20]).

Depuis l'avènement de \hat{f}_0 , les estimateurs en ondelettes n'ont eu de cesse

d'évoluer. En outre, ils se sont enrichis de nouvelles règles de seuillage efficaces et/ou adaptés à de nombreux modèles statistiques émergents.

Mes contributions les plus significatives sur ce sujet s'inscrivent dans cette démarche. Les trois premières parties de ce rapport présentent une synthèse de celles-ci. La première partie est centrée autour de l'estimateur de seuillage par blocs en ondelettes, la seconde présente un estimateur en ondelettes construit à partir d'un "double seuillage" individuel et la troisième concerne l'estimation en ondelettes adaptée au traitement d'observations dépendantes.

3 Seuillage par blocs

3.1 Motivation

Le premier estimateur de seuillage par blocs en ondelettes a été élaboré par [65, 66] pour le modèle de densités. Sa particularité est d’opérer une troncature large de la somme du terme de détail estimée, combinée avec une sélection groupée de ses estimateurs. Ces groupes forment alors des “blocs” successifs, distincts et de même taille. La sélection s’opère par seuillage : les estimateurs d’un groupe dont la norme ℓ_2 dépasse un certain seuil sont gardés, les autres sont supprimés. L’expression analytique de l’estimateur de seuillage par blocs en ondelettes est de la forme :

$$\hat{f}_1(x) = \sum_{k=0}^{2^{j_0}-1} \hat{\alpha}_{j_0,k} \phi_{\tau,k}(x) + \sum_{j=j_0}^{j_1} \sum_{K \in A_j} \sum_{k \in U_{j,K}} \hat{\beta}_{j,k} \mathbf{1}_{\{|\hat{b}_{j,K}| \geq \kappa \lambda_j\}} \psi_{j,k}(x), \quad (4)$$

où $\hat{\alpha}_{j_0,k}$ désigne un estimateur de $\alpha_{j_0,k}$, $\hat{\beta}_{j,k}$ un estimateur de $\beta_{j,k}$,

$$\hat{b}_{j,K} = \left(\sum_{k \in U_{j,K}} |\hat{\beta}_{j,k}|^2 \right)^{1/2},$$

$2^{j_0} = L$, $A_j = \{1, \dots, 2^j L^{-1}\}$ et

$$U_{j,K} = \{k \in \{0, \dots, 2^j - 1\}; (K-1)L \leq k \leq KL - 1\}$$

(L étant un entier caractérisant la taille commune des blocs), κ une constante et λ_j une quantité dépendante de n et j .

Comme, pour tout $k \in U_{j,K}$,

$$\left\{ |\hat{\beta}_{j,k}| \geq \kappa \lambda_j \right\} \subseteq \left\{ |\hat{b}_{j,K}| \geq \kappa \lambda_j \right\},$$

\hat{f}_1 prend en compte plus de $\hat{\beta}_{j,k}$ que \hat{f}_0 . Il estime ainsi plus de détails de f , au risque de cumuler les erreurs d'estimation. Toutefois, avec une bonne calibration des paramètres, cette stratégie s'avère efficace. Cela a été mis en exergue sur le plan théorique et pratique par [4, 5] [6] pour le modèle de régression à pas équidistants.

3.2 Résultat général I

Dans [13, 14], j'ai établi des résultats généraux sur les performances de \hat{f}_1 . Ceux-ci annoncent les conditions que doivent vérifier le modèle statistique et les paramètres pour rendre \hat{f}_1 performant et compétitif avec \hat{f}_0 . La notion de performance est abordée via deux approches complémentaires : l'approche minimax sous le risque \mathbb{L}^p sur des boules de Besov dans [14] et l'approche maxiset sous le risque \mathbb{L}^p dans [13].

Le résultat principal de [14] sous le risque \mathbb{L}^2 est présenté ci-après.

Résultat général I. *Soient f une fonction inconnue que l'on souhaite estimer via n éléments aléatoires, (1) la décomposition en ondelettes associée*

et \hat{f}_1 l'estimateur (4) défini avec

$$\lambda_j = 2^{\delta j} \sqrt{\frac{\ln n}{n}}, \quad L = [\ln n], \quad 2^{j_1} = \left[\left(\frac{n}{\ln n} \right)^{1/(2\delta+1)} \right]. \quad (5)$$

On suppose que $\hat{\alpha}_{j,k}$, $\hat{\beta}_{j,k}$, δ et κ vérifient :

(B1) Il existe une constante $C > 0$ telle que, pour tout $j \in \{j_0, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\mathbb{E} (|\hat{\alpha}_{j,k} - \alpha_{j,k}|^2) \leq C 2^{2\delta j} \frac{1}{n}.$$

(B2) Il existe une constante $C > 0$ telle que, pour tout $j \in \{j_0, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\mathbb{E} (|\hat{\beta}_{j,k} - \beta_{j,k}|^4) \leq C 2^{4\delta j} \frac{1}{n^2}.$$

(B3) Il existe une constante $C > 0$ telle que, pour tout $j \in \{j_0, \dots, j_1\}$ et $K \in A_j$,

$$\mathbb{P} \left(\left(\sum_{k \in U_{j,K}} |\hat{\beta}_{j,k} - \beta_{j,k}|^2 \right)^{1/2} \geq \frac{\kappa}{2} \lambda_j \right) \leq C \frac{1}{n^2}.$$

Alors, sous l'hypothèse que $f \in B_{p,r}^s(M)$ avec $M > 0$, $s \in (0, N)$, $p \geq 1$ et $r \geq 1$, il existe une constante $C > 0$ telle que

$$R(\hat{f}_1, f) \leq C \varphi_n,$$

où

$$\varphi_n = \begin{cases} n^{-2s/(2s+2\delta+1)} & \text{si } p \geq 2 \text{ et } s \in (0, N), \\ \left(\frac{\ln n}{n}\right)^{2s/(2s+2\delta+1)} & \text{si } p \in [1, 2) \text{ et } s \in ((2\delta + 1)/p, N). \end{cases}$$

Pour de nombreux modèles statistiques, les hypothèses **(B1)**, **(B2)** et **(B3)** sont satisfaites par des estimateurs naturels de $\alpha_{j,k}$ et $\beta_{j,k}$. Les preuves peuvent faire appel à des outils probabilistes sophistiqués, notamment l'inégalité de Rosenthal [92] pour **(B2)**, et l'inégalité de Cirelson [44] et/ou l'inégalité de Talagrand [93] pour **(B3)**.

La contribution du résultat I réside dans l'obtention de la vitesse de convergence φ_n , meilleure que celle atteinte par \hat{f}_0 , combinée avec la souplesse des hypothèses formulées. Il peut alors s'appliquer à de nombreux modèles statistiques. En outre, c'est le cas du modèle de régression à plans d'expériences aléatoires [12], du modèle de déconvolution avec bruit blanc gaussien [14] [18] (pour l'estimation de dérivées de f), divers modèles de densités [16, 21] [85] et du modèle de déconvolution multicanal avec bruit blanc gaussien [84]. Pour chacun de ces modèles, nos résultats améliorent ceux relatifs à \hat{f}_0 . Les études numériques entreprises confirment cette amélioration.

Trois de ces résultats sont brièvement présentés ci-après.

3.3 Applications

3.3.1 Régression nonparamétrique

On considère le modèle de régression à plans d'expériences aléatoires : soient $(Y_1, X_1), \dots, (Y_n, X_n)$ n couples de v.a.r. i.i.d. telles que, pour tout $i \in \{1, \dots, n\}$,

$$Y_i = f(X_i) + \xi_i,$$

$f : [0, 1] \rightarrow \mathbb{R}$ est la fonction de régression inconnue, ξ_1, \dots, ξ_n sont n v.a.r. i.i.d. avec $\xi_1 \sim \mathcal{N}(0, 1)$, X_1, \dots, X_n sont i.i.d. avec $X_1 \sim \mathcal{U}(0, 1)$ et, pour tout $i \in \{1, \dots, n\}$, X_i et ξ_i sont indépendantes. On cherche à estimer f à partir de $(Y_1, X_1), \dots, (Y_n, X_n)$. Des applications de ce modèle sont décrites dans [68].

On suppose que f est bornée.

Soit \hat{f}_1 l'estimateur (4) défini avec

$$\hat{\alpha}_{j,k} = \frac{1}{n} \sum_{i=1}^n Y_i \phi_{j,k}(X_i), \quad \hat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^n Y_i \psi_{j,k}(X_i),$$

(5), $\delta = 0$ et κ suffisamment grande. Alors les hypothèses **(B1)**, **(B2)** et **(B3)** sont vérifiées et le résultat I s'applique [12].

3.3.2 Densités pondérées

On considère le modèle de densités pondérées : soit X_1, \dots, X_n n v.a.r. i.i.d. de densité commune

$$g(x) = \frac{f(x)w(x)}{\mu},$$

où $f : [0, 1] \rightarrow [0, \infty[$ est une densité inconnue, w est une fonction connue et μ est la constante de normalisation inconnue. On cherche à estimer f à partir de X_1, \dots, X_n .

On suppose que f est bornée et que w est minorée et majorée.

Soit \hat{f}_1 l'estimateur (4) défini avec

$$\hat{\alpha}_{j,k} = \frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{1}{w(X_i)} \phi_{j,k}(X_i), \quad \hat{\beta}_{j,k} = \frac{\hat{\mu}}{n} \sum_{i=1}^n \frac{1}{w(X_i)} \psi_{j,k}(X_i),$$

$$\hat{\mu} = \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{w(X_i)} \right)^{-1},$$

(5), $\delta = 0$ et κ suffisamment grande. Alors les hypothèses **(B1)**, **(B2)** et **(B3)** sont vérifiées et le résultat I s'applique [21], complétant ainsi [61, Theorem 2].

3.3.3 Déconvolution

On considère le modèle de convolution en bruit blanc gaussien : soit $\{Y_t; t \in [0, 1]\}$ un processus aléatoire tel que

$$dY_t = (f \star g)(t)dt + \frac{1}{\sqrt{n}}dW_t,$$

où $f : [0, 1] \rightarrow \mathbb{R}$ est une fonction 1-périodique inconnue, $\{W_t; t \in [0, 1]\}$ est un mouvement brownien standard et g est une fonction 1-périodique connue. On cherche à estimer f à partir de $\{Y_t; t \in [0, 1]\}$.

On suppose que f est bornée et que g vérifie : il existe deux constantes $c > 0$ et $\omega > 0$ telles que

$$|\mathcal{F}(g)(x)| \geq \frac{c}{(1+x^2)^{\omega/2}},$$

où $\mathcal{F}(g)(x)$ désigne la transformée de Fourier de g .

Soit \hat{f}_1 l'estimateur (4) défini avec la base d'ondelettes de Meyer 1-périodique [83] [70],

$$\hat{\alpha}_{j,k} = \sum_{\ell \in \mathbb{Z}} \frac{\overline{\mathcal{F}(\phi_{j,k})(\ell)}}{\mathcal{F}(g)(\ell)} \int_0^1 e^{-2\pi i \ell t} dY_t, \quad \hat{\beta}_{j,k} = \sum_{\ell \in \mathbb{Z}} \frac{\overline{\mathcal{F}(\psi_{j,k})(\ell)}}{\mathcal{F}(g)(\ell)} \int_0^1 e^{-2\pi i \ell t} dY_t,$$

(5), $\delta = \omega$ et κ suffisamment grande. Les hypothèses **(B1)**, **(B2)** et **(B3)** sont vérifiées, le résultat I s'applique, améliorant [70, Proposition 1] obtenu avec l'estimateur de seuillage dur en ondelettes. Cela est un cas particulier de l'étude entreprise dans [18].

3.4 Cadre multidimensionnel

Une autre contribution au sujet concerne l'adaptation de \hat{f}_1 à l'estimation des fonctions multidimensionnelles. Dans [19], nous proposons un tel estimateur pour un modèle de suites gaussiennes hétéroscédastiques. Avec un choix judicieux de la base d'ondelettes et des paramètres, nous montrons ses bonnes performances théoriques et pratiques. En outre, il s'avère efficace dans le débruitage d'images complexes, allant jusqu'à un temps de calcul six fois plus rapide que les estimateurs habituellement utilisés, dont BLS-GSM (Bayes Least-Squares estimate of Gaussian Scale Mixture).

Ce travail s'est étendu dans [20] au cas de la déconvolution de fonctions multidimensionnelles.

3.5 Perspectives I

Ces dernières années, des méthodes aléatoires ont été mise en place pour choisir une taille de blocs et un seuil améliorant les performances de \hat{f}_1 . Il y a notamment la méthode SURE proposée par [7] dans le cadre du modèle de régression à pas équidistants. Il serait intéressant d'étendre cette méthode à des modèles statistiques plus sophistiqués, comme ceux présentés précédemment. Également, l'adaptation de cet estimateur au cas multidimensionnel reste un challenge d'actualité. Une autre perspective serait la construction d'un estimateur de seuillage par blocs en ondelettes supportant les données dépendantes dans un cadre statistique large, complétant ainsi les travaux précurseurs de [94] [78, 79].

4 Double seuillage

4.1 Motivation

L'estimateur de "double seuillage" en ondelettes a été introduit par [49] (sous une forme non adaptative) dans le contexte du modèle de régression à plans d'expériences aléatoires. Celui-ci reprend la construction de l'estimateur de seuillage individuel en ondelettes en introduisant un nouveau "seuillage" dans la construction de $\hat{\beta}_{j,k}$.

En notant W_1, \dots, W_n les éléments aléatoires d'où émanent f , on propose dans [11] la définition générale suivante :

$$\hat{f}_2(x) = \sum_{k=0}^{2^\tau-1} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \kappa \lambda_j\}} \psi_{j,k}(x), \quad (6)$$

où

$$\hat{\alpha}_{j,k} = \frac{1}{v_n} \sum_{i=1}^n q_i(\phi_{j,k}, W_i), \quad \hat{\beta}_{j,k} = \frac{1}{v_n} \sum_{i=1}^n q_i(\psi_{j,k}, W_i) \mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| \leq \varsigma_j\}},$$

$q_i : \mathbb{L}^2([0, 1]) \times W_i(\Omega) \rightarrow \mathbb{R}$ est une certaine fonction, κ une constante, v_n , λ_j et ς_j sont des quantités dépendantes de n et j .

Le rôle de la fonction indicatrice présente dans $\hat{\beta}_{j,k}$ est de sélectionner les W_1, \dots, W_n ; si $|q_i(\psi_{j,k}, W_i)|$ dépasse un certain seuil, alors W_i est négligé, sinon W_i est conservé. Ce processus est néanmoins complexe : les éléments de la base d'ondelettes sont impliqués dans cette sélection, et les choix des seuils λ_j et ς_j rendant \hat{f}_2 performant sont liés.

L'apport de cette nouvelle sélection permet de relâcher des hypothèses sur le modèle statistique tout en gardant les bonnes performances de l'estimateur de seuillage en ondelettes. Par exemple, bien que non adaptatif, l'estimateur proposé par [49] s'avère performant avec seulement des moments d'ordre 3 pour la variable modélisant le bruit. En outre, la connaissance a priori de la loi de cette variable n'est pas nécessaire. On est donc dans un cadre plus large que l'hypothèse standard de normalité.

4.2 Résultat général II

Dans [11], nous établissons un résultat général sur les performances de \hat{f}_2 . Ce dernier explicite les conditions que doivent vérifier le modèle statistique et le choix des paramètres pour que \hat{f}_2 atteigne des vitesses de convergence "rapides" sous le risque \mathbb{L}^2 sur des boules de Besov. Il est présenté ci-dessous.

Résultat général II. *Soient f une fonction inconnue que l'on souhaite estimer via n éléments aléatoires indépendantes W_1, \dots, W_n , (1) la décomposition en ondelettes associée et \hat{f}_2 l'estimateur (6) défini avec*

$$\lambda_j = \theta_\psi 2^{\delta j} \sqrt{\frac{\ln \mu_n}{\mu_n}}, \quad \varsigma_j = \theta_\psi 2^{\delta j} \frac{v_n}{\sqrt{\mu_n \ln \mu_n}}, \quad 2^{j_1} = \lfloor \mu_n^{1/(2\delta+1)} \rfloor \quad (7)$$

et $\kappa \geq 2 + 8/3 + 2\sqrt{4 + 16/9}$.

On suppose que $q_1, \dots, q_n, v_n, \mu_n, \delta$ et θ_ψ vérifient :

(C1) Pour $\gamma \in \{\phi, \psi\}$, tout $j \in \{\tau, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$, on a

$$\mathbb{E} \left(\frac{1}{v_n} \sum_{i=1}^n q_i(\gamma_{j,k}, W_i) \right) = \int_0^1 f(x) \gamma_{j,k}(x) dx.$$

(C2) Pour $\gamma \in \{\phi, \psi\}$, tout $j \in \{\tau, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$, on a

$$\sum_{i=1}^n \mathbb{E} (|q_i(\gamma_{j,k}, W_i)|^2) \leq \theta_\gamma^2 2^{2\delta j} \frac{v_n^2}{\mu_n}.$$

Alors, sous l'hypothèse que $f \in B_{p,r}^s(M)$ avec $M > 0$, $r \geq 1$ et $\{p \geq 2$ et $s \in (0, N)\}$ ou $\{p \in [1, 2)$ et $s \in ((2\delta + 1)/p, N)\}$, il existe une constante $C > 0$ telle que

$$R(\hat{f}_2, f) \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\delta+1)}.$$

Pour de nombreux modèles statistiques, les hypothèses (C1) et (C2) sont vérifiables avec des outils probabilistes élémentaires. De plus, $(\ln \mu_n / \mu_n)^{2s/(2s+2\delta+1)}$ correspond souvent à la vitesse “presque” optimale au sens minimax. La nuance “presque” est due au facteur logarithmique $(\ln \mu_n)^{2s/(2s+2\delta+1)}$, souvent de trop.

L'expression générale du résultat II, dont la grande flexibilité des hypothèses à vérifier, est l'aboutissement de plusieurs travaux traitant de modèles statistiques variés et complexes. C'est le cas du modèle de régression à plans d'expériences aléatoires avec erreur [22] (pour l'estimation des dérivées de f), certains modèles où interviennent des poids ou des v.a.r. discrètes, notamment des modèles de densités mélangées [27] [11] et des modèles multicaux aléatoires [25] [28]. À ma connaissance, pour chacun d'entre eux, notre esti-

mateur est le seul à proposer une solution adaptative en ondelettes, qui plus est, performante. L'aspect numérique traité dans [28] va également dans ce sens. Une version particulière de cet estimateur a été développée pour le modèle de régression biaisée à plans d'expériences aléatoires dans [40], étant ainsi en marge de notre résultat général. De nouveau, des vitesses de convergences rapides y sont établies.

Trois de ces résultats sont brièvement présentés ci-après.

4.3 Applications

4.3.1 Régression nonparamétrique

On considère le modèle de régression à plans d'expériences aléatoires : soient $(Y_1, X_1), \dots, (Y_n, X_n)$ n couples de v.a.r. i.i.d. telles que, pour tout $i \in \{1, \dots, n\}$,

$$Y_i = f(X_i) + \xi_i,$$

$f : [0, 1] \rightarrow \mathbb{R}$ est la fonction de régression inconnue, ξ_1, \dots, ξ_n sont n v.a.r. i.i.d. avec $\mathbb{E}(\xi_1^2) < \infty$, X_1, \dots, X_n sont i.i.d. avec $X_1 \sim \mathcal{U}(0, 1)$ et, pour tout $i \in \{1, \dots, n\}$, X_i et ξ_i sont indépendantes. On cherche à estimer f à partir de $(Y_1, X_1), \dots, (Y_n, X_n)$. Des applications concrètes sont étudiées dans, par exemple, [68].

On suppose que f est bornée.

Soit \hat{f}_2 l'estimateur (6) défini avec

$$W_i = (Y_i, X_i), \quad q_i(\gamma, (x, x_*)) = x\gamma(x_*),$$

(7), $\delta = 0$, θ_ψ et κ suffisamment grandes, et $v_n = \mu_n = n$. Alors les hypothèses **(C1)** et **(C2)** sont vérifiées et le résultat II s'applique. Cela améliore [49, Theorem 4] dans le sens où notre estimateur est adaptatif et seuls des moments d'ordre 2 sur ξ_1 sont requis.

4.3.2 Régression nonparamétrique multicanal

On considère le modèle de régression à plans d'expériences aléatoires multicanal : soient $(Y_1, X_1), \dots, (Y_n, X_n)$ n couples de v.a.r. i.i.d. telles que, pour tout $i \in \{1, \dots, n\}$,

$$Y_i = f_{V_i}(X_i) + \xi_i,$$

V_1, \dots, V_n sont n v.a.r. discrètes indépendantes non-observées telles que $V_i(\Omega) = \{1, \dots, m\}$, $m \in \mathbb{N}^*$, pour tout $\nu \in \{1, \dots, m\}$, $f_\nu : [0, 1] \rightarrow \mathbb{R}$ sont des fonctions de régression inconnues, ξ_1, \dots, ξ_n sont n v.a.r. i.i.d. avec $\mathbb{E}(\xi_1^2) < \infty$, X_1, \dots, X_n sont i.i.d. avec $X_1 \sim \mathcal{U}(0, 1)$ et, pour tout $i \in \{1, \dots, n\}$, X_i , V_i et ξ_i sont indépendantes. Pour un ℓ fixé dans $\{1, \dots, m\}$, on cherche à estimer f_ℓ à partir de $(Y_1, X_1), \dots, (Y_n, X_n)$. Ce modèle est une version mélangée du modèle de régression introduit par [91].

On suppose que f_1, \dots, f_m sont bornées. On pose

$$w_k(i) = \mathbb{P}(V_i = k)$$

et on suppose que le déterminant de la matrice

$$\Gamma_n = \left(\frac{1}{n} \sum_{i=1}^n w_k(i) w_r(i) \right)_{(k,r) \in \{1, \dots, m\}^2}$$

est strictement positif.

Soit \hat{f}_2 l'estimateur (6) défini avec

$$W_i = (Y_i, X_i), \quad q_i(\gamma, (x, x_*)) = a_\ell(i) x \gamma(x_*),$$

où $(a_\ell(1), \dots, a_\ell(n))$ est l'unique vecteur réel vérifiant

$$\min_{b \in \mathbb{R}^n} \frac{1}{n} \sum_{i=1}^n b_i^2 \quad \text{tel que} \quad \frac{1}{n} \sum_{i=1}^n b_i w_\nu(i) = \delta_{\ell, \nu},$$

pour $\nu \in \{1, \dots, m\}$, ($\delta_{\ell, \nu}$ désignant le symbole de Kronecker), (7), $\delta = 0$, θ_ψ et κ suffisamment grandes, $\nu_n = n$ et $\mu_n = n/\rho_n$ avec $\rho_n = (1/n) \sum_{i=1}^n a_\ell^2(i)$.

Alors les hypothèses **(C1)** et **(C2)** sont vérifiées et le résultat II s'applique, offrant un premier élément théorique à ce problème. Les preuves ainsi qu'une étude numérique sont présentées dans [28].

Notons que, grâce au double seuillage, seul un moment d'ordre 2 sur le bruit a été supposé et qu'aucune hypothèse de bornitude particulière sur $a_\ell(i)$ n'a été formulée.

4.3.3 Convolution de densités mélangées

On considère le modèle de convolutions de densités mélangées : soient Y_1, \dots, Y_n n v.a.r. indépendantes telles que, pour tout $i \in \{1, \dots, n\}$,

$$Y_i = X_i + \epsilon_i, \quad (8)$$

où X_1, \dots, X_n sont n v.a.r. indépendantes, $\epsilon_1, \dots, \epsilon_n$ sont n v.a.r. i.i.d. de densité commune g et, pour tout $i \in \{1, \dots, n\}$, X_i et ϵ_i sont indépendantes. Pour tout $i \in \{1, \dots, n\}$, la densité de X_i est

$$h_i(x) = \sum_{\nu=1}^m w_\nu(i) f_\nu(x),$$

où $w_\nu(i)$ vérifie $w_\nu(i) > 0$ et $\sum_{\nu=1}^m w_\nu(i) = 1$, et, pour tout $\nu \in \{1, \dots, m\}$, $f_\nu : [0, 1] \rightarrow [0, \infty[$ est une densité inconnue. Pour un ℓ fixé dans $\{1, \dots, m\}$, on cherche à estimer f_ℓ à partir de Y_1, \dots, Y_n . Ce problème statistique étend le modèle de convolution de densités [8] [51] [62] [87] [63] [3] [46] [48] [76], et aussi le modèle de densités mélangées [80] [67] [89] [90].

On suppose que f_1, \dots, f_m sont bornées, le déterminant de la matrice

$$\Gamma_n = \left(\frac{1}{n} \sum_{i=1}^n w_k(i) w_r(i) \right)_{(k,r) \in \{1, \dots, m\}^2}$$

est strictement positif et g vérifie : il existe deux constantes $c > 0$ et $\omega > 0$ telles que

$$|\mathcal{F}(g)(x)| \geq \frac{c}{(1+x^2)^{\omega/2}},$$

où $\mathcal{F}(g)(x)$ désigne la transformée de Fourier de g .

Soit \hat{f}_2 l'estimateur (6) défini avec

$$W_i = Y_i, \quad q_i(\gamma, x) = a_\ell(i) \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma)(y)}}{\mathcal{F}(g)(y)} e^{-iyx} dy,$$

où $(a_\ell(1), \dots, a_\ell(n))$ est l'unique vecteur réel vérifiant

$$\min_{b \in \mathbb{R}^n} \frac{1}{n} \sum_{i=1}^n b_i^2 \quad \text{tel que} \quad \frac{1}{n} \sum_{i=1}^n b_i w_\nu(i) = \delta_{\ell, \nu},$$

pour $\nu \in \{1, \dots, m\}$, (7), $\delta = \omega$, θ_ψ et κ suffisamment grandes, $v_n = n$ et $\mu_n = n/\rho_n$ avec $\rho_n = (1/n) \sum_{i=1}^n a_\ell^2(i)$.

Alors les hypothèses **(C1)** et **(C2)** sont vérifiées et le résultat II s'applique [27]. Cela généralise les études entreprises par [63] [89], tout en améliorant les vitesses de convergence de [89] si on se place dans le cas non-bruité. Cela est dû à une meilleure gestion des poids $a_\ell(i)$ par notre estimateur.

4.4 Perspectives II

Afin d'améliorer la vitesse de convergence, il serait intéressant de combiner l'approche du double seuillage avec celle du seuillage par blocs en ondelettes. Ainsi, dans un premier temps, on ferait une sélection individuelle des estimateurs des coefficients d'ondelettes, puis une sélection groupée de ces estimateurs. L'adaptation de l'estimateur de double seuillage au cadre de données dépendantes est aussi une perspective à explorer. Un pas dans ce sens est proposé dans [41].

5 Estimation en ondelettes et dépendance

5.1 Motivation

Plusieurs phénomènes aléatoires font apparaître de la dépendance dans leur modélisation. La question de l'estimation d'une fonction inconnue f à partir de v.a.r. dépendantes s'est donc naturellement posée. Parmi les structures de dépendance souvent rencontrées se distingue l' α -mélangeance. Elle est vérifiée, entre autres, par de nombreux processus autorégressifs et certaines chaînes de Markov.

Dans ce contexte, quelques estimateurs en ondelettes ont été développés [77] [82] [96] [86] [55] [56] [57]. Cependant, à ma connaissance, aucun d'entre eux n'atteint le but de l'adaptativité dans un cadre statistique général et sous des conditions larges de régularité comme, par exemple, $f \in B_{p,r}^s(M)$, avec $p \geq 1$. Ce constat a motivé une partie de mes recherches.

5.2 Résultat général III

Sous l'hypothèse d' α -mélangeance, j'ai développé l'estimateur de seuillage dur en ondelettes (adaptatif) pour plusieurs modèles statistiques complexes, dont le modèle de densités pondérées [24], le modèle de déconvolution de densités [33], un modèle de densités de type censure multiplicatif [35] et le modèle de régression nonparamétrique à plans d'expériences aléatoires [29, 41]. Ses performances sont évaluées par la détermination de vitesses de convergence sous le risque \mathbb{L}^2 sur des boules de Besov.

Dans [43], ces résultats ont été généralisés et améliorés. En notant W_1, \dots, W_n les éléments aléatoires d'où émanent f , on considère l'estimateur de seuillage dur en ondelettes \hat{f}_3 défini par

$$\hat{f}_3(x) = \sum_{k=0}^{2^{j_0}-1} \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \kappa \lambda_j\}} \psi_{j,k}(x), \quad (9)$$

où

$$\hat{\alpha}_{j,k} = \frac{1}{n} \sum_{i=1}^n q(\phi_{j,k}, W_i), \quad \hat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^n q(\psi_{j,k}, W_i), \quad (10)$$

$q : \mathbb{L}^2([0, 1]) \times W_1(\Omega) \rightarrow \mathbb{R}$, κ une constante et λ_j est une quantité dépendante de n et j .

Le résultat principal de [43] annonce les conditions que doivent vérifier le modèle statistique, les éléments de la base d'ondelettes et le choix des paramètres pour que \hat{f}_3 atteigne des vitesses de convergence "rapides", toujours sous le risque \mathbb{L}^2 sur des boules de Besov.

Résultat général III. *Soient f une fonction inconnue que l'on souhaite estimer via n éléments aléatoires W_1, \dots, W_n issus d'un processus aléatoire $(W_t)_{t \in \mathbb{Z}}$ strictement stationnaire et α -mélangeant à décroissance exponentielle (décrit dans l'appendice), (1) la décomposition en ondelettes associée et \hat{f}_3 l'estimateur (9) défini avec*

$$\lambda_j = 2^{\delta j} \sqrt{\frac{\ln n}{n}}, \quad 2^{j_0} = \lceil \ln n \rceil, \quad 2^{j_1} = \left\lceil \left(\frac{n}{(\ln n)^3} \right)^{1/(2\delta+1)} \right\rceil \quad (11)$$

et κ suffisamment grande.

On suppose que q et δ vérifient :

(D1) Pour $\gamma \in \{\phi, \psi\}$, tout $j \in \{j_0, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\mathbb{E}(q(\gamma_{j,k}, W_1)) = \int_0^1 f(x)\gamma_{j,k}(x)dx,$$

(D2) Il existe une constante $C > 0$ telle que, pour $\gamma \in \{\phi, \psi\}$, tout $j \in \{j_0, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

(i) $\sup_{x \in W_1(\Omega)} |q(\gamma_{j,k}, x)| \leq C2^{\delta j}2^{j/2},$

(ii) $\mathbb{E}(|q(\gamma_{j,k}, W_1)|^2) \leq C2^{2\delta j},$

(iii) pour tout $m \geq 1,$

$$|\mathbb{C}_{ov}(q(\gamma_{j,k}, W_{m+1}), q(\gamma_{j,k}, W_1))| \leq C2^{2\delta j}2^{-j},$$

où \mathbb{C}_{ov} désigne la covariance.

Alors, sous l'hypothèse que $f \in B_{p,r}^s(M)$ avec $M > 0$, $r \geq 1$ et $\{p \geq 2$ et $s \in (0, N)\}$ ou $\{p \in [1, 2)$ et $s \in ((2\delta + 1)/p, N)\}$, il existe une constante $C > 0$ telle que

$$R(\hat{f}_3, f) \leq C \left(\frac{\ln n}{n} \right)^{2s/(2s+2\delta+1)}.$$

Les hypothèses **(D1)** et **(D2)** sont peu restrictives; on est en mesure de trouver q et δ les vérifiant pour de nombreux modèles statistiques. Il est intéressant de noter que **(D2)**-**(iii)** est vérifiée si les deux hypothèses de bornitudes **(D3)** et **(D4)** présentées ci-dessous le sont.

(D3) Soient u la densité de W_1 et $u_{(W_1, W_{m+1})}$ celle de (W_1, W_{m+1}) pour tout $m \in \mathbb{Z}$. On suppose qu'il existe une constante $C > 0$ telle que

$$\sup_{m \in \mathbb{Z}} \sup_{(x, y) \in W_1(\Omega) \times W_{m+1}(\Omega)} |u_{(W_1, W_{m+1})}(x, y) - u(x)u(y)| \leq C.$$

(D4) Il existe deux constantes $C > 0$ et $\delta \geq 0$ vérifiant, pour $\gamma \in \{\phi, \psi\}$, tout $j \in \{j_0, \dots, j_1\}$ et $k \in \{0, \dots, 2^j - 1\}$,

$$\int_{W_1(\Omega)} |q(\gamma_{j,k}, x)| dx \leq C 2^{\delta j} 2^{-j/2}.$$

Notons que $(\ln n/n)^{2s/(2s+2\delta+1)}$ correspond à la vitesse de convergence atteinte par l'analogue de \hat{f}_3 dans le cas indépendant, i.e., \hat{f}_0 (voir le résultat I), laquelle est souvent presque optimale au sens minimax. Cela améliore les vitesses de convergence présentées dans [33, 35].

Dans ce qui suit, la flexibilité de ce résultat est illustrée en considérant trois modèles statistiques sophistiqués.

5.3 Applications

5.3.1 Déconvolution de densités

On considère le modèle de déconvolution de densités : soit $(Y_t)_{t \in \mathbb{Z}}$ un processus strictement stationnaire tel que

$$Y_t = X_t + \epsilon_t, \tag{12}$$

où $(X_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire de densité inconnue $f : [0, 1] \rightarrow [0, \infty[$ et $(\epsilon_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire de densité connue $g : [0, 1] \rightarrow [0, \infty[$. On suppose que, pour tout $t \in \mathbb{Z}$, X_t et ϵ_t sont indépendantes et que $(Y_t)_{t \in \mathbb{Z}}$ est α -mélangeant à décroissance exponentielle. On cherche à estimer f à partir de Y_1, \dots, Y_n . Sous des conditions de dépendance, ce modèle a déjà été étudié via différentes approches dans [82] [75] [47] [96].

On suppose que f est bornée, la densité de Y_1 est bornée, la densité de (Y_1, Y_{m+1}) est bornée pour tout $m \in \mathbb{Z}$, et g vérifie : il existe deux constantes $c > 0$ et $\omega > 0$ telles que

$$|\mathcal{F}(g)(x)| \geq \frac{c}{(1+x^2)^{\omega/2}},$$

où $\mathcal{F}(g)(x)$ désigne la transformée de Fourier de g .

Soit \hat{f}_3 l'estimateur (9) défini avec

$$W_i = Y_i, \quad q(\gamma, x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma)}(y)}{\mathcal{F}(g)(y)} e^{-iyx} dy,$$

(11), $\delta = \omega$ et κ suffisamment grande. Alors les hypothèses **(D1)** et **(D2)** sont vérifiées et le résultat III s'applique [33, 43].

Cela étend le résultat principal de [63] au cas de l' α -mélangeance, améliore [96] quant à l'adaptativité et la vitesse de convergence, et aussi [33, Proposition 5.1] quant à la vitesse de convergence ; on gagne un terme logarithmique.

5.3.2 Modèle à bruit multiplicatif

On considère le modèle à bruit multiplicatif : soit $(Y_t)_{t \in \mathbb{Z}}$ un processus strictement stationnaire tel que

$$Y_t = \sigma_t^2 Z_t, \quad (13)$$

où $(\sigma_t^2)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire de densité inconnue $f : [0, 1] \rightarrow [0, \infty[$ et $(Z_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire de densité connue $g : [0, 1] \rightarrow [0, \infty[$. On suppose que, pour tout $t \in \mathbb{Z}$, σ_t^2 and Z_t sont indépendantes et que $(Y_t)_{t \in \mathbb{Z}}$ est α -mélangeant à décroissance exponentielle. On cherche à estimer f à partir de Y_1, \dots, Y_n . Ce problème inclut des modèles de type GARCH. Il a notamment été traité dans [47].

On suppose que f est bornée, la densité de Y_1 est bornée, la densité de (Y_1, Y_{m+1}) est bornée pour tout $m \in \mathbb{Z}$, et que g vérifie : il existe un entier ω tel que

$$g(x) = \frac{1}{(\omega - 1)!} (-\ln x)^{\omega-1}, \quad x \in [0, 1].$$

Remarquons que g correspond à la densité de la v.a.r. $\prod_{i=1}^{\omega} U_i$, où U_1, \dots, U_{ω} sont ω i.i.d. v.a.r. de la loi commune la loi $\mathcal{U}([0, 1])$.

Soit \hat{f}_3 l'estimateur (9) défini avec

$$W_i = Y_i, \quad q(\gamma, x) = T_{\omega}(\gamma)(x),$$

$T_\omega(\gamma)(x) = T(T_{\omega^{-1}}(\gamma))(x)$, avec

$$T(\gamma)(x) = (x\gamma(x))',$$

(11), $\delta = \omega$ et κ suffisamment grande. Alors les hypothèses **(D1)** et **(D2)** sont vérifiées et le résultat III s'applique [35, 43].

Cela améliore [35, Theorem 2] quant à la vitesse de convergence ; on gagne une puissance $1/2$. À ma connaissance, notre estimateur est le seul à apporter une solution adaptative directe au problème avec le MISE (“directe” car sans transformer le problème et sans utiliser le domaine fréquentiel via la transformée de Fourier).

5.3.3 Régression nonparamétrique

On considère le modèle de régression nonparamétrique : soit $(W_t)_{t \in \mathbb{Z}}$ un processus strictement stationnaire tel que, pour tout $t \in \mathbb{Z}$, $W_t = (Y_t, X_t)$,

$$Y_t = f(X_t) + \xi_t, \tag{14}$$

$f : [0, 1] \rightarrow \mathbb{R}$ est la fonction de régression inconnue, $(X_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire de densité inconnue $g : [0, 1] \rightarrow [0, \infty[$ et $(\xi_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire. On suppose que, pour tout $t \in \mathbb{Z}$, X_t and ξ_t sont indépendants et que $(W_t)_{t \in \mathbb{Z}}$ est α -mélangeant à décroissance exponentielle. On cherche à estimer f à partir de W_1, \dots, W_n .

Des applications de ce problème sont décrites dans [68]. Des méthodes en

ondelettes non adaptatives ont été proposées par [86] [55] [56] [57].

On suppose que f et g sont bornées, g est minorée par une constante positive notée c_* , la densité de W_1 est bornée, la densité de (W_1, W_{m+1}) est bornée et, pour tout $m \in \mathbb{Z}$, et ξ_1 est bornée.

On considère l'estimateur \hat{f} défini par

$$\hat{f}(x) = \frac{\hat{v}(x)}{\hat{g}(x)} \mathbf{1}_{\{|\hat{g}(x)| \geq c_*/2\}}, \quad (15)$$

où

– \hat{v} est défini comme (9) avec

$$W_i = (Y_i, X_i), \quad q(\gamma, (x, x_*)) = x\gamma(x_*)$$

(11) et $\delta = 0$, lequel est un estimateur de $v(x) = f(x)g(x)$,

– \hat{g} est défini par (9) avec

$$W_i = X_i, \quad q(\gamma, x) = \gamma(x)$$

(11) et $\delta = 0$, lequel est un estimateur de $g(x)$.

Alors les hypothèses **(D1)** et **(D2)** sont vérifiées; le résultat III s'applique séparément pour \hat{v} et \hat{g} , entraînant le résultat suivant.

Résultat. [43] *On reprend les notations et hypothèses ci-dessus. Soit \hat{f} l'estimateur (15). Sous l'hypothèse que $v = fg \in B_{p,r}^s(M)$ et $g \in B_{p,r}^s(M)$ avec $M > 0$, $r \geq 1$ et $\{p \geq 2$ et $s \in (0, N)\}$ ou $\{p \in [1, 2)$ et $s \in (1/p, N)\}$, il*

existe une constante $C > 0$ telle que

$$R(\hat{f}, f) \leq C \left(\frac{\ln n}{n} \right)^{2s/(2s+1)}.$$

Cela complète [86] sur la détermination de la vitesse de convergence pour le risque \mathbb{L}^2 sur des boules de Besov.

Extension

Principalement à cause de **(D2)-(i)**, le cas où ξ_1 est non bornée rend inapplicable notre résultat. Il a fait l'objet d'un traitement particulier dans [29], qui fut ensuite amélioré dans [41] (et étendu au cas multidimensionnel). Dans ce travail, je propose un estimateur de seuillage en ondelettes performant sous des conditions faibles sur le modèle. Les cas de la sous-gaussianité et la seule existence des moments d'ordre 2 pour ξ_1 sont traités, élargissant alors les hypothèses standards généralement posées.

Entre autre, dans [41], le résultat suivant est prouvé.

Résultat. *On reprend les notations et hypothèses ci-dessus avec $g(x) = 1$ (pour simplifier) et l'hypothèse sur ξ_1 : il existe deux constantes $v > 0$ et $\sigma > 0$ telles que, pour tout $\lambda > 0$,*

$$\mathbb{P}(|\xi_1| \geq \lambda) \leq 2ve^{-\lambda^2/(2\sigma^2)}.$$

Soit \hat{f}_3 l'estimateur (9) défini avec

$$2^{j_0} = \lceil \tau(\ln n)^2 \rceil, \quad 2^{j_1} = \left\lceil \frac{n}{(\ln n)^4} \right\rceil,$$

$\delta = 0$, κ suffisamment grande et

$$W_i = (Y_i, X_i), \quad q(\gamma, (x, x_*)) = x\gamma(x_*).$$

Alors, sous l'hypothèse que $f \in B_{p,r}^s(M)$ avec $M > 0$, $r \geq 1$ et $\{p \geq 2$ et $s \in (0, N)\}$ ou $\{p \in [1, 2)$ et $s \in (1/p, N)\}$, il existe une constante $C > 0$ telle que

$$R(\hat{f}_3, f) \leq C \left(\frac{\ln n}{n} \right)^{2s/(2s+1)}.$$

J'ai également travaillé sur l'élaboration d'estimateurs linéaires en ondelettes pour des modèles statistiques et des structures de dépendance jusqu'alors inexplorées avec cette méthode. Voir [26] pour le modèle de mélange de densités sous une dépendance quadrant, [31] pour un modèle de densités de type censure multiplicatif sous différentes structures de dépendance, puis [30] pour le modèle de densités pondérées stratifiées sous la dépendance quadrant et enfin, [32] pour le modèle de densités mélangées sous la dépendance quadrant (dans lequel on estime les dérivées de f).

5.4 Perspectives III

L'extension de ce travail au seuillage par blocs en ondelettes est une piste à explorer. Il est vraisemblable que l'étude de l'estimateur de seuillage global

dans le cadre de la β -mélangeance considéré dans [94] fournisse des éléments utiles à cela. Il serait aussi intéressant d'étendre notre résultat au processus α -mélangeant à décroissance polynomiale. Pour finir, le fait de supposer f à support compact exclut certaines modélisations dans ce cadre dépendant. Une solution à ce problème serait l'adaptation des outils développés par [71] pour des v.a.r. indépendantes.

6 Compléments

J'ai aussi contribué à d'autres thématiques statistiques, notamment

- l'agrégation d'estimateurs [17],
- l'estimation paramétrique en ondelettes [23] [37, 42],
- l'estimation de convolution de densités [36] [39],
- l'estimation de densités pour des v.a.r. fonctionnelles [38],
- les méthodes dérivées du LASSO [15] [58],
- les séries temporelles à valeurs entières [34].

Appendice

Boules de Besov. Soient $M > 0$, $s > 0$, $p \geq 1$ et $r \geq 1$. On pose, pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$ et $\epsilon \geq 0$,

$$\Delta_\epsilon(f)(x) = f(x + \epsilon) - f(x),$$

$\Delta_\epsilon^2(f)(x) = \Delta_\epsilon(\Delta_\epsilon(f))(x)$ et de même,

$$\Delta_\epsilon^N(f)(x) = \Delta_\epsilon^{N-1}(\Delta_\epsilon(f))(x), \quad N \in \mathbb{N}.$$

Soit

$$\rho^N(t, f, p) = \sup_{|\epsilon| \leq t} \left(\int_0^1 |\Delta_\epsilon^N(f)(x)|^p dx \right)^{1/p}.$$

Alors, pour $s \in [0, N)$, on définit la boule de Besov $B_{p,r}^s(M)$ de rayon $M > 0$ par

$$B_{p,r}^s(M) = \left\{ f \in \mathbb{L}^p([0, 1]); \left(\int_0^1 \left(\frac{\rho^N(t, f, p)}{t^s} \right)^r \frac{dt}{t} \right)^{1/r} \leq M \right\},$$

avec la modification usuelle quand $p = \infty$ ou $r = \infty$.

Le paramètre s est lié à la régularité de la fonction f , alors que p et r sont des paramètres de norme.

On a alors l'implication suivante ([69], Corollary 9.1) : si $f \in B_{p,r}^s(M)$ avec $s \in (0, N)$ alors il existe une constante $M^* > 0$ (dépendante de M) telle

que les coefficients d'ondelettes (2) vérifient

$$2^{\tau(1/2-1/p)} \left(\sum_{k=0}^{2^{\tau}-1} |\alpha_{\tau,k}|^p \right)^{1/p} + \left(\sum_{j=\tau}^{\infty} \left(2^{j(s+1/2-1/p)} \left(\sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq M^*,$$

avec la modification usuelle quand $p = \infty$ ou $r = \infty$.

Pour certains s , p et r , $B_{p,r}^s(M)$ contient des boules fonctionnelles standards, comme les boules de Hölder et les boules de Sobolev [50] [83] [69].

α -mélangeance à décroissance exponentielle. Soit $(V_t)_{t \in \mathbb{Z}}$ un processus aléatoire stationnaire. On dit que $(V_t)_{t \in \mathbb{Z}}$ est α -mélangeant à décroissance exponentielle si et seulement si il existe deux constantes $\gamma > 0$ et $\theta > 0$ telles que

$$\sup_{m \geq 1} (e^{\theta m} \alpha_m) \leq \gamma,$$

où α_m est le m -ième coefficient d' α -mélange de $(V_t)_{t \in \mathbb{Z}}$ défini par

$$\alpha_m = \sup_{(A,B) \in \mathcal{F}_{-\infty,0}^V \times \mathcal{F}_{m,\infty}^V} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

$\mathcal{F}_{-\infty,0}^V$ est la σ -algèbre engendrée par les variables aléatoires (ou vecteurs aléatoires) \dots, V_{-1}, V_0 et $\mathcal{F}_{m,\infty}^V$ est la σ -algèbre engendrée par les variables aléatoires (ou vecteurs aléatoires) V_m, V_{m+1}, \dots . L'importance de cette dépendance dans la modélisation de phénomènes aléatoires est décrite dans [2]

[9] [60] [64].

Références

- [1] Aubin, J.B. (2005). Estimation fonctionnelle par projection adaptative et applications, *Thèse de l'Université de Paris VI*.
- [2] Bradley, R.C. (2007). *Introduction to strong mixing conditions*, Vol. 1,2,3. Kendrick Press.
- [3] Butucea, C. and Matias, C. (2005). Minimax estimation of the noise level and of the signal density in a semiparametric convolution model, *Bernoulli*, 11, **2**, 309-340.
- [4] Cai, T. (1997). On adaptivity of blockshrink wavelet estimator over Besov spaces, *Technical report 97-05*, Purdue University.
- [5] Cai, T. (1999). Adaptive wavelet estimation : a block thresholding and oracle inequality approach, *Annals of Statistics*, 27, 898-924.
- [6] Cai, T. et Silverman, B.W. (2001). Incorporating information on neighboring coefficients into wavelet Estimation, *Sankhya*, 63, 127-148.
- [7] Cai, T. and Zhou, H. (2009). A data-driven block thresholding approach to wavelet estimation, *Annals of Statistics*, 37, 569-595.
- [8] Carroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density, *Journal of the American Statistical Association*, 83, 1184-1186.
- [9] Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models, *Econometric Theory*, 18, 17-39.

- [10] Chaubey, Y.P., Chesneau, C. and Shirazi, E. (2013). Wavelet-based estimation of regression function for dependent biased data under a given random design, *Journal of Nonparametric Statistics*, 25, 1, 53-71.
- [11] Chaubey, Y.P., Chesneau, C. and Doosti, H. (2013). Adaptive wavelet estimation of a density from mixtures under multiplicative censoring, preprint.
- [12] Chesneau, C. (2007). Wavelet block thresholding for samples with random design : a minimax approach under the \mathbb{L}^p risk, *Electronic Journal of Statistics*, 1, 331-346.
- [13] Chesneau, C. (2008). On the maxiset comparison between hard and block thresholding methods, *Statistics and Probability Letters*, 78, 6, 675-681.
- [14] Chesneau, C. (2008). Wavelet estimation via block thresholding : a minimax study under \mathbb{L}^p risk, *Statistica Sinica*, 18, 3, 1007-1024.
- [15] Chesneau, C. and Hebiri. M. (2008). Some theoretical results on the grouped variables Lasso, *Mathematical Methods of Statistics*, 17, 4, 317-326.
- [16] Chesneau, C. (2009). Block thresholding for a density estimation problem with a change-point, *Asian European Journal of Mathematics*, 2, 4, 545-555.
- [17] Chesneau, C. and Lecué, G. (2009). Adapting to unknown smoothness by aggregation of thresholded wavelet estimators, *Statistica Sinica*, special

- issue "Multiscale Methods and Statistics : a productive Marriage", 19, 1407-1417.
- [18] Chesneau, C. (2010). Wavelet estimation of the derivatives of an unknown function from a convolution model, *Current Development in Theory and Applications of Wavelets*, 4, 2, 131-151.
- [19] Chesneau, C., Fadili, J. and Starck, J.-L. (2010). Stein block thresholding for image denoising, *Applied and Computational Harmonic Analysis*, 28, 1, 67-88.
- [20] Chesneau, C., Fadili, J. and Starck, J.-L. (2010). Stein block thresholding for wavelet-based image deconvolution, *Electronic Journal of Statistics*, 4, 415-435.
- [21] Chesneau, C. (2010). Wavelet block thresholding for density estimation in the presence of bias, *Journal of the Korean Statistical Society*, 39, 1, 43-53.
- [22] Chesneau, C. (2010). On adaptive wavelet estimation of the regression function and its derivatives in an errors-in-variables model, *Current Development in Theory and Applications of Wavelets*, 4, 2, 185-208.
- [23] Chesneau, C. (2011). On adaptive wavelet estimation of a quadratic functional from a deconvolution problem, *Annals of the Institute of Statistical Mathematics*, 63, 2, 405-429.
- [24] Chesneau, C. (2011). Adaptive wavelet estimation of a biased density for strongly mixing sequences, *International Journal of Mathematics and Mathematical Sciences*, Volume 2011, Article ID 604150, 21 pages.

- [25] Chesneau, C. (2011). Adaptive wavelet estimator for a function and its derivatives in an indirect convolution model, *Journal of Statistical Theory and Practice*, 5, 2, 303-326.
- [26] Chesneau, C. (2011). A note on the wavelet deconvolution of a density from mixtures under quadrant dependence, *Current Development in Theory and Applications of Wavelets*, 5,1, 51-64.
- [27] Chesneau, C. (2011). Wavelet density estimators for the deconvolution of a component from a mixture, *Sankhya : The Indian Journal of Statistics*, 73-A, 2, 245-266.
- [28] Chesneau, C. and Fadili, J. (2012). Adaptive wavelet estimation of a function in an indirect regression model, *AStA Advances in Statistical Analysis*, 96, 1, 25-46.
- [29] Chesneau, C. (2012). Adaptive wavelet regression in random design and general errors with weakly dependent data, *Acta Universitatis Apulensis*, 29, 65-84.
- [30] Chesneau, C., Dewan, I. and Doosti, H. (2012). Wavelet linear density estimation for associated stratified size-biased sample, *Journal of Nonparametric Statistics*, 24, 2, 429-445.
- [31] Chesneau, C. and Doosti, H. (2012). Wavelet linear density estimation for a GARCH model under various dependence structures, *Journal of the Iranian Statistical Society*, 11, 1, 1-21.
- [32] Chesneau, C. (2012). Wavelet linear estimation of a density and its derivatives from observations of mixtures under quadrant dependence, *ProbS-*

- tat Forum*, 5, 38-46.
- [33] Chesneau, C. (2012). On the adaptive wavelet deconvolution of a density for strong mixing sequences, *Journal of the Korean Statistical Society*, 41, 423-436.
- [34] Chesneau, C. and Kachour, M. (2012). A parametric study for the first-order signed integer-valued autoregressive process, *Journal of Statistical Theory and Practice*, 6, 4, 760-782.
- [35] Chesneau, C. (2013). Wavelet estimation of a density in a GARCH-type model, *Communications in Statistics : Theory and Methods*, 42, 1, 98-117.
- [36] Chesneau, C., Comte, F. and Navarro, F. (2013). Fast nonparametric estimation for convolutions of densities, *Canadian Journal of Statistics*, (to appear).
- [37] Chesneau, C., Kachour, M. and Navarro, F. (2013). A note on the adaptive estimation of a quadratic functional from dependent observations, *IStatistik : Journal of the Turkish Statistical Association*, 6, 1, 10-26.
- [38] Chesneau, C., Kachour, M. and Maillot, B. (2013). Nonparametric estimation for functional data by wavelet thresholding, *REVSTAT - Statistical Journal*, 11, 2, 211-230.
- [39] Chesneau, C. and Navarro, F. (2013). On a plug-in wavelet estimator for convolutions of densities, *Journal of Statistical Theory and Practice*, (to appear).

- [40] Chesneau, C. and Shirazi, E. (2013). Nonparametric wavelet regression based on biased data, *Communications in Statistics : Theory and Methods*, (to appear).
- [41] Chesneau, C. (2013). On the adaptive wavelet estimation of a multidimensional regression function under α -mixing dependence : Beyond the standard assumptions on the noise, *Commentationes Mathematicae Universitatis Carolinae*, 4, 527-556.
- [42] Chesneau, C., Kachour, M. and Navarro, F. (2013). On the estimation of density-weighted average derivative by wavelet methods under various dependence structures, *Sankhya : The Indian Journal of Statistics*, (to appear).
- [43] Chesneau, C. (2014). A general result on the mean integrated squared error of the wavelet hard thresholding estimator under α -mixing dependence, *Journal of Probability and Statistics*, Volume 2014, Article ID 403764, 12 pages.
- [44] Cirelson, B.S., Ibragimov, I.A., and Sudakov, V.N. (1976). *Norm of Gaussian Sample Functions*, Springer Verlag, Berlin.
- [45] Cohen, A., Daubechies, I., Jawerth, B. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms, *Applied and Computational Harmonic Analysis*, 24, 1, 54-81.
- [46] Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006). Penalized contrast estimator for density deconvolution, *Canadian Journal of Statistics*, 34, 431-452.

- [47] Comte, F., Dedecker, J. and Taupin, M.-L. (2008). Adaptive density deconvolution for dependent inputs with measurement errors, *Mathematical Methods of Statistics*, 17, 2, 87-112.
- [48] Delaigle, A. and Gijbels, I. (2006). Estimation of boundary and discontinuity points in deconvolution problems, *Statistica Sinica*, 16, 773 -788.
- [49] Delyon, B. and Juditsky, A. (1996). On minimax wavelet estimators, *Applied Computational Harmonic Analysis*, 3, 215-228.
- [50] DeVore, R. and Popov, V. (1988). Interpolation of Besov spaces, *Transactions of the American Mathematical Society*, 305, 397-414.
- [51] Devroye, L. (1989). Consistent deconvolution in density estimation, *Canadian Journal of Statistics*, 17, 235-239.
- [52] Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, 81, 425-455.
- [53] Donoho, D.L. and Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinkage, *Journal of the American Statistical Association*, 90, 1200-1224.
- [54] Donoho, D.L., Johnstone, I.M., Kerkycharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding, *Annals of Statistics*, 24, 508-539.
- [55] Doosti, H., Afshari, M. and Niroumand, H.A. (2008). Wavelets for non-parametric stochastic regression with mixing stochastic process, *Communication in Statistics : Theory and Methods*, 37, 3, 373-385.

- [56] Doosti, H., Islam, M.S., Chaubey, Y.P. and Gora, P. (2010). Two dimensional wavelets for nonlinear autoregressive models with an application in dynamical system, *Italian Journal of Pure and Applied Mathematics*, 27, 39-62.
- [57] Doosti, H. and Niroumand, H.A. (2009). Multivariate stochastic regression estimation by wavelets for stationary time series, *Pakistan journal of Statistics*, 27, 1, 37-46.
- [58] Dossal, C., Kachour, M., Fadili, J., Peyré, G. and Chesneau, C. (2013). The degrees of freedom of the Lasso for general design matrix, *Statistica Sinica*, 23, 2, 809-828.
- [59] Doukhan, P. and Leon, J. (1990). Déviation quadratique d'estimateur de densité par projection orthogonale, *Comptes Rendus Acad. Sciences Paris (A)*, 310, 424-430.
- [60] Doukhan, P. (1994). *Mixing. Properties and Examples*, Lecture Notes in Statistics 85, Springer Verlag, New York.
- [61] Efromovich S. (2004). Density estimation for biased data, *Annals of Statistics*, 32, 1137-1161.
- [62] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem, *Annals of Statistics*, 19, 1257-1272.
- [63] Fan, J. and Koo, J.Y. (2002). Wavelet deconvolution, *IEEE transactions on information theory*, 48, 734-747.

- [64] Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000). Stochastic volatility models as hidden Markov models and statistical applications, *Bernoulli*, 6, 6, 1051-1079.
- [65] Hall, P., Kerkycharian, G. and Picard, D. (1998). Block threshold rules for curve estimation using kernel and wavelet methods, *Annals Statistics*, 26, 922-942.
- [66] Hall, P., Kerkycharian, G. and Picard, D. (1999). On the minimax optimality of block thresholded wavelet estimators, *Statistica Sinica*, 9, 33-49.
- [67] Hall, P. and Zhou, X.H. (2003). Nonparametric estimation of component distributions in a multivariate mixture, *Annals of Statistics*, 31, 1, 201-224.
- [68] Härdle, W. (1990). *Applied nonparametric regression*, Cambridge University Press.
- [69] Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelet, approximation and statistical applications*, Lectures Notes in Statistics, New York 129, *Springer Verlag*.
- [70] Johnstone, I., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting, *Journal of the Royal Statistical Society : Series B*, 66, 547-573.
- [71] Juditsky, A. and Lambert-Lacroix, S. (2004). On minimax density estimation on \mathbb{R} , *Bernoulli*, 10, 2, 187-220.
- [72] Kerkyacharian, G. and Picard, D. (1992). Density estimation in Besov spaces, *Statist. Probab. Lett.*, volume 13, 15-24.

- [73] Kerkyacharian, G. and Picard, D. (1993). Density estimation by kernel and wavelets methods : optimality of Besov spaces, *Statistics and Probability Letter*, 18, 327-336.
- [74] Kerkyacharian, G. and Picard, D. (2000). Thresholding algorithms, maxisets and well-concentrated bases, *Test*, 9, 2, 283-344.
- [75] Kulik, R. (2008). Nonparametric deconvolution problem for dependent sequences, *Electronic J. of Statist.*, 2, 722-740.
- [76] Lacour, C. (2006). Rates of convergence for nonparametric deconvolution, *Comptes Rendus de l'Académie des Sciences de Paris Série I Math.*, 342 (11), 877-882.
- [77] Leblanc, F. (1996). Wavelet linear density estimator for a discrete time stochastic process : \mathbb{L}^p -losses. *Statistics and Probability Letters*, 27, 71-84.
- [78] Li, L. and Xiao, Y. (2007). On the minimax optimality of block thresholded wavelet estimators with long memory data, *Journal of Statistical Planning and Inference*, 137, 2850-2869.
- [79] Li, L. and Xiao, Y. (2010). A Note on Block Thresholded Wavelet Estimators with Correlated Noise, *Communications in Statistics - Theory and Methods*, 39, 1111-1128.
- [80] Maiboroda, R. E. (1996). Estimators of components of a mixture with varying concentrations, *Ukrainian Mathematical Journal*, 48, 4, 562-566.
- [81] Mallat, S. (2009). *A wavelet tour of signal processing*, Elsevier/ Academic Press, Amsterdam, third edition, the sparse way, with contributions from Gabriel Peyré.

- [82] Masry, E. (1993). Strong consistency and rates for deconvolution of multivariate densities of stationary processes, *Stochastic Processes and their Applications*, 47, 53-74.
- [83] Meyer, Y. (1992). *Wavelets and Operators*, Cambridge University Press, Cambridge.
- [84] Navarro, F., Chesneau, C., Fadili, J. and Sassi, T. (2013). Block thresholding for wavelet-based estimation of function derivatives from a heteroscedastic multichannel convolution model, *Electronic Journal of Statistics*, 7, 0, 428-453.
- [85] Navarro, F., Chesneau, C. and Fadili, J. (2014). On adaptive wavelet estimation of a class of weighted densities, *Communications in Statistics - Simulation and Computation*, (to appear)
- [86] Patil, P.N. and Truong, Y.K. (2001). Asymptotics for wavelet based estimates of piecewise smooth regression for stationary time series, *Annals of the Institute of Statistical Mathematics*, 53, 1, 159-178.
- [87] Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution, *Annals of Statistics*, 27, 2033-2053.
- [88] Petrov, V. V. (1995). *Limit Theorems of Probability Theory : Sequences of Independent Random Variables*, Oxford : Clarendon Press.
- [89] Pokhyl'ko, D. (2005). Wavelet estimators of a density constructed from observations of a mixture, *Theory of Probability and Mathematical Statistics*, 70, 135-145.

- [90] Prakasa Rao, B.L.S. (2010). Wavelet linear estimation for derivatives of a density from observations of mixtures with varying mixing proportions, *Indian Journal of Pure and Applied Mathematics*, 41, 1, 275-291.
- [91] Racine J.S. and Li, Q. (2004). Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data, *Journal of Econometrics*, 119, 1, 99-130.
- [92] Rosenthal, H.P. (1970). On the subspaces of \mathbb{L}^p ($p \geq 2$) spanned by sequences of independent random variables, *Israel Journal of Mathematics*, 8, 273-303.
- [93] Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes, *Annals of Probability*, 22, 28-76.
- [94] Tribouley, K. and Viennet, G. (1998). \mathbb{L}^p adaptive density estimation in a β -mixing framework, *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 34, 2, 179-208.
- [95] Tsybakov, A.B. (2004). *Introduction à l'estimation non-paramétrique*, Springer, New York.
- [96] van Zanten, H. and Zareba, P. (2008). A note on wavelet density deconvolution for weakly dependent data, *Statistical Inference for Stochastic Processes*, 11, 207-219.