Université d'Evry-val d'Essonne
UFR SFA, Département de Mathématiques

Document de synthèse pour l'habilitation à diriger des recherches

En Sciences
Spécialité : **Mathématiques**

Présenté par Dasha LOUKIANOVA

Théorèmes limites dans le cadre markovien. Application aux statistiques.

Soutenue le 6 juillet 2010 devant le jury composé de

Monsieur Laurent DENIS,
Madame Valentine GENON-CATALOT,
Monsieur Arnaud GUILLIN, rapporteur,
Monsieur Reinhard HÖPFNER, rapporteur
Monsieur Jean JACOD, rapporteur,
Madame Monique JEANBLANC.

Document de synthèse d'activité scientifique.

Dasha Loukianova *

26 mars 2011

*Département de Mathématiques, Université d'Evry-Val d'Essonne, Bd François Mitterrand, 91025 Evry Cedex, France. E-mail : dasha.loukianova@univ-evry.fr

Table des matières

1 Remerciements.

Je tiens à remercier tout d'abord Arnaud Guillin, Reinhard Höpfner et Jean Jacod pour avoir consacré de leur temps à lire ce mémoire d'habilitation et pour leur soutien. Merci à eux et à Valentine Genon-Catalot, Laurent Denis qui me font l'honneur de participer au jury.
Une partie des travaux présentés a été écrite pendant mon semestre de délégation au CNRS à MAP5, où j'ai été accueillie par Valentine Genon-Catalot et Fabienne Comte. Je voudrais les remercier chaleureusement pour cet accueil et pour leur grande disponibilité.
J'ai une pensée à Francis Comets, directeur de ma thèse de doctorat. Merci de m'avoir initiée à la thématique des théorèmes limites en probabilité, pour m'avoir donné l'exemple d'une grande rigueur et honnêteté en Mathématique.
Merci à mes co-auteurs : Eva Löcherbach, Oleg Loukianov et Shiqi Song : ce travail est aussi le leur. Dans ma vie scientifique j'ai eu occasion de travailler avec de nombreuses personnes, toutes ne sont pas devenues co-auteurs, mais le travail en commun a toujours été un plaisir. Merci à tous mes collaborateurs.
Apres ma thèse de doctorat j'ai eu une longue et pénible période sans recherche : tout mathématicien étant passé par là sait qu'en ces moments on pense avoir touché le fond. En 2002 je suis "remontée à la surface", mais cette période d'isolement s'est réellement terminée quand j'ai rencontré Eva. De tout mon coeur merci à elle.
Merci à Denis Feyel pour avoir lu ce manuscrit et avoir corrigé les principales fautes de français. Merci pour son amitié et son soutien indéfectible.
J'ai la chance de faire partie du laboratoire d'Analyse et Probabiltés de l'Université d'Evry, où en plus du très bon niveau mathématique on profite aussi d'un très bon climat humain. Je tiens à remercier tous les membres de ce laboratoire pour avoir contribué, chacun à sa manière, à ce précieux climat.

2 Introduction.

Dans ces notes je décris brièvement l' ensemble de mes travaux de recherche, depuis la thèse de doctorat et jusqu'aux derniers développements. Ces travaux se partage en 4 thèmes :

- Mécanique statistique. Etude du modèle de Hopfield de mémoire associative par les méthodes des grandes déviations.
- Théorèmes limites uniformes pour les familles de martingales dépendant d'un paramètre. Application à l'étude de l'estimateur de maximum de vraisemblance pour le paramètre de drift d'une diffusion.
- Théorèmes limites pour les fonctionnelles additives d'un processus de Markov récurrent. Application à l'estimation des diffusions, l'estimation par noyau, l'estimation pénalisée.
- Inégalités fonctionnelles.

Les quatre thèmes ne sont pas réellement liés, mais ont en commun le fait de toucher de prés ou de loin aux théorèmes limites dans le cadre Markovien. Ces travaux de recherche correspondent à 12 publication et une pré-publication et sont présentés dans ces notes par ordre chronologique. Ils sont aussi regroupés par sous-thèmes. Chaque section traite un sous-thème et chaque sous-section (sauf une ou deux exceptions) traite une publication. Les références générales sont données par un nom d'auteur et un nombre ; ceux où le nombre précède par la lettre L (et sans nom) corréspondent à la bibliographie personnelle.

3 Mécanique statistique. Etude du modèle de Hopfield de mémoire associative par les méthodes des grandes déviations $([L1],[L2])$.

J'ai commencé mes activités de recherche en Doctorat à Paris 7, sous la direction de Francis Comets. Mes travaux de thèse se situaient dans le domaine de la mécanique statistique et portaient sur un modèle de réseaux de neurones : modèle de Hopfield de mémoire associative. Dans ma thèse j'ai obtenu deux résultats. Le premier concerne la capacité de mémoire de ce modèle. Présenté dans ma thèse et dans la note de *CRAS* [L1], ce résultat a été très rapidement amélioré par M.Talagrand [Tal] et n'a pas été présenté dans une autre publication que [L1]. Le second concerne le seuil d'erreur de ce modèle. Il est publié dans *Probability Theory and Related Fields* [L2].

3.1 Description du modèle.

Le modèle de Hopfield est un réseau composé d'un grand nombre N de neurones pouvant être en deux états : "au repos" ou "excité". L'état du i-ème neurone est décrit par son spin $\sigma_i \in \mathcal{S} := \{-1, +1\}$ et l'état du réseau entier par une configuration de spins $\sigma \in S^N$. Ce réseau est destiné à mémoriser m configurations de spins particulières $X^1, \ldots, X^n$, appelées prototypes dans le sens suivant : il doit être capable d'associer un prototype à toute configuration initiale par le moyen d'un certain processus dynamique en temps discret et à valeurs dans l'espace de configurations $\mathcal{S}^N$. En réalité il existe plusieurs dynamiques possibles, mais chacune est faite en sorte à minimiser la fonctionnelle d'énergie

$$H_N(\sigma) = -\sum_{s=1}^{m} < \sigma, X^s >,$$

où $<,>$ est le produit scalaire dans $\mathbb{R}^{\mathbf{N}}$. L'idée conduisant à cette fonction d'énergie est que dans certains cas les prototypes réalisent les minima de H_N (par exemple quand ils sont orthogonaux.) Ainsi, du point de vue "statique", les questions liées au fonctionnement de la mémoire reviennent à l'étude du "relief énergie" de H_N. C'est le point de vue qu'on adopte ici. Pour comprendre comment le réseau se comporte dans une situation typique, on choisit le prototype au hasard, notamment on suppose que les variables X_i^s sont i.i.d. uniformes sur $\{-1, +1\}$. Puisqu'on s'intéresse aux réseaux de grandes tailles on suppose que $N \to \infty$. où $<,>$ dénote le produit scalaire dans $\mathbb{R}^{\mathbf{N}}$.

3.2 Capacité de mémoire ($[L1]$).

Une question importante posée par le modèle de Hopfield est sa capacité de mémoire, i.e. la relation d'échelle, quand $N \to \infty$ entre la taille du réseau N et le nombre de prototypes m pouvant être réstitués par le réseau exactement ou avec une petite erreur. L'écart entre deux configurations se mesure en termes de la distance de Hamming normalisée par N, cette dernière distance étant égale aux nombre de coordonnées différentes entre les deux configurations. Il a été montré rigoureusement par McEliece, Posner, Rodemich et Venkatesh [47] que si $m < \frac{N}{4 \log N}$, la restitution des prototypes est exacte, ce résultat a été précisé par Komlos et Paturi[41], qui ont montré que dans ce régime chaque prototype a un bassin d'attraction non-trivial. Si on tolère une petite erreur dans la restitution des prototypes, m peut croître plus rapidement. Le régime intéressant pour ce modèle est $m = \alpha N$ (" many patterns",) où de nombreuses études numériques ont montré l'existence d'un seuil critique α^*. Par des méthodes non-rigoureses Amit, Gutfreund et Sompolinsky [AGF] ont suggéré la valeur $\alpha_c = 0,14$. Cependant il n'existe à ce jour aucun résultat mathématique rigoureux confirmant cette valeur du α critique. Le premier résultat dans cette direction est dû à Ch. M. Newman [New], qui donne la borne inférieure pour le α_c critique, égale à $\alpha_c \gtrsim 0,056$. Dans $[L1]$ on a amélioré la borne inférieure de Newman en trouvant $\alpha_c \gtrsim 0,071$. Les méthodes utilisées sont celles des grandes déviations. Le résultat de $[L1]$ a été amélioré par M. Talagrand, qui commençait alors à travailler intensivement sur les modèles de la mécanique statistique. Le résultat de Talagrand donne $\alpha_c \gtrsim 0,84$. [Tal]. Pour formuler rigoureusement notre résultat nous avons besoin de notations supplémentaires :

Pour toute $\sigma \in \mathcal{S}^N$ et $r > 0$, notons $Sp(\sigma, r)$ la sphère de centre σ et de rayon de Hamming r, notons $B(\sigma, r)$ la boule correspondante. Pour $\delta \in]0, 1/2[$ et $\varepsilon > 0$ notons $A(N, m, \delta, \varepsilon)$ l'événement tel que chaque prototype soit entouré d'une barrière d'énergie de hauteur εN^2 sur la sphère de rayon $[\delta N]$:

$$A(N, m, \delta, \varepsilon) = \{\forall s \leq m; \quad \min_{\{\sigma' \in Sp(\sigma, r)\}} H_N(\sigma') > H_N(X^s) + \varepsilon N^2\}.$$

Ici $\delta < 1/2$, parce que ces barrières doivent séparer les prototypes différents, et que par la loi des grands nombres $\frac{d(X_s, X_l)}{N} \to \frac{1}{2}$ pour $s \neq l$ quand $N \to \infty$. Cet événement signifie que la dynamique issue d'un prototype converge vers un minimum local à l'intérieur de la boule $B(X^s, [\delta N])$. Ainsi l'erreur commise sur un prototype n'excédera pas $[\delta N]$. On comprend la valeur critique de la capacité de mémoire α^* dans le sens que dans tout régime plus lent les barrières autour des prototypes subsistent. Dans ce sens nous avons le théorème suivant

Théorème 3.1 *Soient $\alpha < \alpha^*$ et $m = m(N)$ une suite d'entiers positifs telle que $\lim_{N\to\infty} \frac{m(N)}{N} = \alpha$. Alors il existe $\delta \in]0, 1/2[$ et $\varepsilon > 0$ (tous deux dépendant seulement de α) tels que*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbf{P}[A^c(N, m, \delta, \varepsilon)] \leq -K(\alpha, \delta, \varepsilon).$$

où $K(\alpha, \delta, \varepsilon) > 0$. Numériquement on trouve $\alpha^* \approx 0,071 := \alpha_l$.

Notons que la même définition de la capacité critique α_c a été utilisée par C. Newman et M. Talagrand.

3.3 Seuil d'erreur dans la restitution des prototypes ($[L2]$).

Dans la section précédente nous avons parlé de la borne inférieure pour la valeur critique α_c de la capacité de mémoire. Il n'existe pas de résultats donnant la borne supérieure pour cette valeur. Dans le $[L2]$ on montre d'une autre façon qu'il existe bien une rupture dans le fonctionnement du modèle quand le nombre de prototypes vérifie $m = \alpha N$ avec α grand : on montre que pour tout $\alpha > 0$ il n'y a pas de minimums locaux du hamiltonien d'énergie H_N dans un certain voisinage des prototypes, de taille dépendant de α. Le résultat du $[L2]$ est le premier résultat "négatif" pour le modèle de Hopfield dans le régime "many patterns" $m = \alpha N$, il montre que l'erreur de restitution dans ce régime est inévitable. Notons $C(N, m, \delta)$ l'événement suivant :

$$C(N, m, \delta) = \{\exists s \leq m, \quad \exists \sigma \in B(X^s, [\delta N]), \quad \sigma \text{est un minimum local de } H_N\}.$$

Théorème 3.2 *Soit $\alpha > 0$ et $m = m(N)$ une suite d'entiers positifs dépendant de N telle que $\lim_{N\to\infty} \frac{m(N)}{N} = \alpha$. Il existe un $\delta > 0$ dépendant de α seulement tel que*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbf{P}[C(N, m, \delta)] \leq -K(\alpha, \delta)$$

où $K(\alpha, \delta) > 0$. De plus, si $\delta(\alpha)$ est la borne supérieure des δ pour lesquels ce résultat est vrai, alors

$$\liminf_{\alpha\to\infty} \delta(\alpha) \geq \delta_* \approx 0,05.$$

Les méthodes utilisées sont celles des grandes déviations et de l'association négative des variables. L'association négative permet de majorer les probabilités d'intersection d'événements "négativement associés" par des produits des probabilités.

4 Estimateur du maximum de vraisemblance dans le régime de convergence presque sûre ($[L3], [L4], [L5]$).

Dans ce chapitre nous présentons trois articles portant sur l'étude du MVS d'un paramètre de drift pour une diffusion à valeurs dans $\mathbb{R}^n$. Le cadre général (qui sera précisé dans chaque cas) est le suivant. Soit (Θ, d) un espace métrique compact. Considérons une famille de diffusions à valeurs dans $\mathbb{R}^n$ donnée par

$$dX_t^\theta = \sigma(X_t^\theta)dW_t + b(X_t^\theta, \theta)dt, \quad X_0^\theta = x, \quad \theta \in \Theta, \tag{4.1}$$

où W est un mouvement brownien à valeurs dans $\mathbb{R}^k$. Nous supposons que $a = \sigma\sigma^T$ est définie positive et que pour chaque $\theta \in \Theta$ il existe une solution de (4.1).

L'estimateur de maximum de vraisemblance du paramètre θ du drift est défini comme le processus $(\hat{\theta}_t), t \geq 0$ maximisant p.s. et pour chaque $t \geq 0$ la fonction de vraisemblance

$$L_t(\theta) = \exp[\int_0^t b_\theta^T a^{-1}\sigma(X_s)dW_s - \frac{1}{2}\int_0^t b_\theta^T a^{-1} b_\theta(X_s)ds], \tag{4.2}$$

qu'on écrira souvent sous la forme

$$L_t(\theta) = M_t(\theta) - \frac{1}{2}A_t(\theta). \tag{4.3}$$

Dans les trois articles résumés dans ce chapitre nous étudions des questions d'existence, de consistance et de vitesse de convergence pour cet estimateur. D'habitude ce genre d'étude est effectué du point de vue de la convergence en probabilité et dans le cadre de la récurrence positive, voir [43] pour un survey extensif. Dans ce chapitre nous nous intéressons systématiquement à la convergence presque sûre. L'intérêt du point de vue "presque sûr" est que la vitesse de convergence est aussi une fonction de l'observation. Dans $[L3]$ on considère le processus d'Ornstein-Uhlenbeck généralisé, récurrent positif, dans $[L4]$ la diffusion peut être récurrente positive où nul-récurrente, et dans $[L5]$ elle peut être récurrente où transiente, aussi Θ peut-il être infini-dimmensionnel.

4.1 Remarque sur les techniques de semi-groupes et la consistance de MLE [L3.].

Dans cette note de *Statistics & Probability Letters* on considère le processus d'Ornstein-Uhlenbeck multidimensionnel

$$dY_t = dW_t - \nabla U(\theta, Y_t)dt,$$

avec le drift donné par le gradient d'une fonction convexe U. La question de la convergence presque sûre de MLE est équivalente à la question de la convergence presque sûre vers zéro de

$$\sup_{\theta \in K} \frac{L_t(\theta)}{L_t(\theta_0)},$$

pour tout compact $K \subset \Theta$ ne contenant pas θ_0. Or pour certains f et pour $V(\theta, x) = \Delta U(\theta, x) - |\nabla U|^2(\theta, x)$ on a une majoration de la forme

$$\sup_{\theta \in K} \frac{L_t(\theta)}{L_t(\theta_0)} \leq \exp t \left\{ \sup_K \frac{f(X_t)}{t} + \sup_K \frac{1}{t} \int_0^t [V(\theta, X_s) - V(\theta_0, X_s)] ds \right\}. \tag{4.4}$$

Par la formule d'Ito et le théorème ergodique pour les fonctionnelles additives on a

$$\frac{f(X_t)}{t} = \frac{f(x)}{t} + \frac{1}{t} \int_0^t \nabla f(X_s) dW_s + \frac{1}{t} \int_0^t L^{\theta_0} f(X_s) ds \to \mu(L^{\theta_0} f),$$

où $L^\theta f(x) = \frac{1}{2} \Delta f(x) - \langle \nabla U(\theta, x), \nabla f(x) \rangle$ est le générateur du semi-groupe de X^θ. On remarque aussi que

$$[V(\theta, X_s) - V(\theta_0, X_s)] \leq L^{\theta_0} [U(\theta, x) - U(\theta_0, x)].$$

La question de la convergence du MLE ici est ramenée à la question du signe de la limite de l'expression entre les accolades dans (4.4). Cette limite est majorée par $\mu(L^{\theta_0} g)$, elle est donc négative car $\mu(Lg) = 0$. En raisonnant uniformément en θ on obtient la consistance du MVS.

4.2 Loi des grandes nombres uniforme pour une suite de martingales dépendant d'un paramètre. Application à la consistance p.s. de MVS ([L4]).

Dans l'article de *Statistics & Probability Letters* en collaboration avec Oleg Loukianov nous nous posons la question de la convergence *p.s.* de l'estimateur de MVS pour les diffusions d'une manière plus générale. La diffusion X est cependant supposée récurrente (positive ou nulle). Rappelons que la fonction de vraisemblance s'écrit

$$L_t(\theta) = M_t(\theta) - \frac{1}{2} A_t(\theta),$$

où $M_t(\theta)$ est une martingale locale continue et $A_t(\theta)$ sa variation quadratique,(4.2). On a aussi $L_t(\theta_0) = 0$. Ainsi pour établir la consistance du MVS il suffit de prouver pour tout t suffisamment grand et tout compact K ne contenant pas θ_0

$$\sup_{\theta \in K} L_t(\theta) < 0.$$

Pour montrer ceci nous considérons une famille $\{M_t(\theta);\ t \geq 0,\ \theta \in \Theta\}$ de martingales locales continues dépendant d'un paramètre appartenant à un espace métrique borné (Θ, d). La loi forte des grands nombres pour une martingale localement de carré intégrable normalisée par sa variation quadratique est un résultat classique et bien connu. Ici nous montrons une version uniforme de cette loi sous des conditions adaptées au problème de l'estimation par le MVS. Nous supposons que le processus de variation quadratique $(A_t(\theta, \theta'))$ de la martingale $(M_t(\theta) - M_t(\theta'))$ est hölderien par rapport au paramètre dans le sens où pour chaque $t \geq 0$ on a

$$\forall (\theta, \theta') \in \Theta^2, \quad A_t(\theta, \theta') \leq d^{2\delta}(\theta, \theta') V_t \tag{4.5}$$

avec $0 < \delta \leq 1$ et (V_t) processus positif croissant vers l'infini. C'est une condition naturelle satisfaite dès que la fonction du drift elle-même est höldérienne. Nous prouvons ensuite la loi des grands nombres uniforme sur les compacts K de Θ :

Théorème 4.1 *Supposons que la famille* $\{M_t(\theta)\}$ *satisfasse* (4.5), *alors*

$$\lim_{t\to\infty} \sup_{\theta\in K} \frac{|M_t(\theta)|}{A_t(\theta)} = 0 \quad p.s.$$

pour tout compact $K \in \Theta$ *satisfaisant*

$$\forall\theta \in K, \quad \liminf_{t\to\infty} \frac{A_t(\theta)}{V_t} > 0 \quad p.s.$$

La loi de grandes nombres uniforme qu'on obtient ainsi permet de montrer la convergence presque sûre de MLE pour les diffusions récurrentes (positives ou nulles) multidimensionnelles en observation continues. Ce résultat améliore les résultats de Van Zanten ([64] et [65]) où pour une diffusion unidimensionnelle et ergodique (=récurrente positive) on prouve la consistance en probabilité de MLE.

4.3 Loi du logarithme itéré uniforme pour les familles des martingales. Vitesse de convergence presque sûre de MVS ([$L5$]).

Rappelons que que le processus $(r_t^{-1}, t \geq 0)$ est une vitesse de convergence supérieure de l'estimateur $\hat{\theta}_t$ si $r_t d(\hat{\theta}_t, \theta_0)$ est borné dans un certain sens. On peut considérer deux notions importantes de bornitude :

$$\mathbf{P}\left(\limsup_{t\to\infty} r_t d(\hat{\theta}, \theta_0) = \infty\right) = 0 \quad \text{(vitesse p.s.)} \tag{4.6}$$

$$\lim_{M\to\infty} \limsup_{t\to\infty} \mathbf{P}\left(r_t d(\hat{\theta}_t, \theta_0) > M\right) = 0 \quad \text{(vitesse en probabilité)} \tag{4.7}$$

Pour la fonction de vraisemblance, reprenons suivant (4.2) et (4.3) la formule

$$L_t(\theta) = M_t(\theta) - \frac{1}{2} A_t(\theta).$$

Une façon d'obtenir une borne inférieure pour (r_t) (voir de A.W. van der Vaart et J.A. Wellner [63] ou S. van de Geer [60]) dans le cadre de M-estimation est la suivante :

Puisque $L_t(\theta_0) = 0$, chaque MVS $\hat{\theta}_t$ doit satisfaire $2M_t(\hat{\theta}_t) \geq A_t(\hat{\theta}_t)$. En majorant $(M_t(\theta))$ et en minorant $(A_t(\theta))$ par des fonctions de $d(\hat{\theta}_t, \theta_0)$ et en choisissant une normalisation appropriée pour que les bornes restent valables quand t tends vers l'infini, on obtient une estimation pour (r_t).

Pour mettre en oeuvre cette idée nous avons besoin d'hypothèses sur la régularité de $< M(\theta) - M(\theta') >_t$. Nous nous limitons ici aux cadre höldérien, car toutes les applications qu'on voudrait traiter s'y inscrivent.

Hypothèse 4.2 *Supposons qu'il existe deux processus continues U_t et V_t croissants vers l'infini de façon que $\liminf_{t\to\infty} U_t/V_t > 0$ et tels que l'on ait pour t suffisamment grand*

$$\forall \theta \in \Theta, \qquad d^{2k}(\theta, \theta_0) U_t \leq \langle M(\theta) - M(\theta_0) \rangle_t$$

et

$$\forall (\theta, \theta') \in \theta^2, \quad \langle M(\theta) - M(\theta') \rangle_t \leq d^{2\delta}(\theta, \theta') V_t,$$

avec $\delta \leq 1$ et $k \geq \delta$.

Ces propriétés sont assez naturelles et la première donne une minoration voulue pour $A_t(\theta)$. Ensuite il faut en déduire une majoration (module de continuité) pour $M(\theta)$. Cette majoration doit être uniforme en θ , ce qui implique la dépendance de cette borne de l'entropie métrique de Θ. Si on s'intéresse à la convergence en probabilité, l'inégalité maximale de Nishiyama [53] pour une famille de martingales dépendant d'un paramètre, développée ensuite par van Zanten et van der Vaart ([66], [62]) permet d'obtenir l' estimation suivante, utile dans ce contexte

$$\mathbf{E}\left(\frac{|M_t(\theta) - M_t(\theta')|}{\sqrt{V_t}} \right) \sim d^{\delta}(\theta, \theta').$$

Pour le contexte "presque sûr", auquel on s'intéresse, en se servant de la méthode de chaînage on obtient une sorte de loi du logarithme itéré uniforme : on a presque sûrement pour t suffisamment grand

$$\frac{|M_t(\theta) - M_t(\theta')|}{\sqrt{V_t \ln\ln V_t}} \leq C\phi(d(\theta, \theta')).$$

La fonction $\phi = x^{\delta} \times l(x)$, est décroissante et dépend de l'entropie métrique N de l'espace Θ. Notamment l doit satisfaire l'inégalité

$$\sum_{m\geq 1} m^2 N^2 (2^{-m} \exp(-l^2(2^{-m})) < \infty. \tag{4.8}$$

Dans le cas $\Theta \subset \mathbb{R}^d$, on a $N(\varepsilon) \sim \varepsilon^{-d}$, et on peut prendre $\phi(x) = x^{\delta}\sqrt{\ln(c/x)}$. Dans le cas $N(\varepsilon) \sim \exp(c\varepsilon^{-\alpha})$ avec $0 < \alpha < 2\delta$ on peut prendre $\phi(x) = x^{\delta - \frac{\alpha}{2}}$. En appliquant ce résultat à l'estimateur du MVS on obtient le théorème suivant :

Théorème 4.3 *Considérons une famille de diffusions donnée par* (4.1) *et vérifiant l'hypothèse* (4.2). *Si $r_t \nearrow +\infty$ satisfait*

$$r_t^{2k}\phi(\frac{1}{r_t}) \leq \sqrt{\frac{V_t}{\ln\ln V_t}},$$

alors

$$\mathbf{P}_{\theta_0}\left(\limsup_{t\to\infty} r_t d(\hat{\theta}_t, \theta_0) = \infty \right) = 0,$$

et donc $1/r_t$ est une vitesse supérieure de convergence du MVS.

Notons que dans des contextes différentes la loi du logarithme itéré uniforme pour certaines familles des martingales a déjà été utilisée en estimation ; voir R. Senoussi [57] pour les chaînes de Markov, S. van de Geer et L. Stougie [61].

Comme exemple type on peut considérer une diffusion dont le drift vérifie la propriété de Hölder

Exemple 4.4

$$\forall(\theta,\theta',x)\in\Theta^2\times\mathbb{R}^n,\quad |b(\theta,x)-b(\theta',x)|\le C(x)d^\delta(\theta,\theta'),$$

ainsi qu'une condition d'identifiabilité :

$$\forall(\theta,x)\in\Theta\times\mathbb{R}^n, |b(\theta,x)-b(\theta_0,x)|\ge K(x)d^k(\theta,\theta_0);$$

où $C(x)>0$ et $K(x)>0$. Alors on peut prendre

$$V_t=\int_0^t C^2|a^{-1}|(X_s)ds\quad et\quad U_t=\int_0^t K^2|a^{-1}|(X_s)ds.$$

Dans ce cas nous trouvons (c constante positive)

$$r_t^{2k-\delta}\sqrt{\ln(cr_t)}=\sqrt{\frac{V_t}{\ln\ln V_t}}.$$

Remarquons que V_t et U_t sont des fonctionnelles additives de (X). Dans le cas où (X) est récurrent positif, V_t peut être remplacé par t ce qui donne pour drift lipschitzien $(\delta=k=1)$ la vitesse $r_t\sqrt{\ln r_t}\sim\sqrt{t/\ln\ln t}$. C'est un peu moins bien que la vitesse attendue $r_t=\sqrt{t/\ln\ln t}$ valable dans le cas de la dépendance linéaire en θ.

Mis à part cet exemple naturel, on donne aussi dans [$L5$] des exemples de diffusions nul-récurrentes, transientes, et avec drift discontinu, pour lesquelles la vitesse p.s. de convergence du MVS peut être donnée par le théorème 4.3, ainsi qu'un exemple où l'espace des paramètres Θ est infini-dimmensionnel. Arrêtons-nous plus en détails sur ce cas.

Exemple 4.5 *Considérons une famille de diffusions*

$$dX_t=\theta(X_t)dt+dW_t,$$

où la fonction inconnue θ est à support dans $[0,1]$ et $\alpha\in]1/2,1]$ tel que

$$\sup_{0\le s\le 1}|\theta(s)|+\sup_{0\le s\le t\le 1}\frac{|\theta(t)-\theta(s)|}{|t-s|^\alpha}\le 1.$$

Alors X_t est nul-récurrent et sa mesure invariante μ_θ satisfait pour des constantes m_θ, M_θ

$$0<m_\theta\le d\mu_\theta/dx\le M_\theta.$$

Soit μ la mesure qui correspond à la vraie valeur du paramètre. On peut montrer que Θ est un compact dans $L^2(\mu)$ avec l'entropie $N(\varepsilon)\sim\exp(A\varepsilon^{-1/\alpha})$, de sorte que la condition d'entropie (4.8) est satisfaite avec $l(x)=cx^{-1/(2\alpha)}$.

Ici

$$A_t(\theta,\theta_1)=\int_0^t(\theta-\theta^1)(X_s)ds.$$

Posons

$$Z_t = \int_0^t \mathbf{1}_{[0,1]}(X_s)ds.$$

Par le théorème de la limite-quotient nous avons

$$\forall(\theta,\theta_1,) \quad \frac{A_t(\theta,\theta_1)}{Z_t} \to \frac{d^2(\theta,\theta_1)}{\mu([0,1])}.$$

On peut montrer que cette convergence est uniforme, ce qui nous donne à partir d'un certain temps (aléatoire)

$$\forall(\theta,\theta_1,) \quad \frac{d^2(\theta,\theta_1)}{2\mu([0,1])} \le \frac{A_t(\theta,\theta_1)}{Z_t} \le \frac{2d^2(\theta,\theta_1)}{\mu([0,1])}.$$

En choisissant $U_t = Z_t/2\mu([0,1])$ *et* $V_t = 2Z_t/\mu([0,1])$ *on satisfait toutes les hypothèses du théorème et l'on obtient*

$$r_t \sim \left(\frac{Z_t}{\ln\ln Z_t}\right)^{\frac{1}{2+1/\alpha}}.$$

Notons que ce taux de convergence, obtenu par des méthodes purement paramétriques, correspond aux taux non-paramétrique et minimax $r_t = t^{\alpha/(2\alpha+1)}$ *qu'on trouve dans le cas récurrent-positif et au sens de la convergence en probabilité, voir Delattre , Hoffmann et Kessler [23], et aussi Delattre et Hoffmann [22].*

5 Equivalent déterministe pour les fonctionnelles additives des diffusions récurrentes dans $\mathbb{R}$. Application aux statistiques ([L6], [L7]).

Considérons un processus de diffusion donné par une EDS

$$dX_t = \sigma(X_t)\,dW_t + b(X_t)\,dt, \tag{5.9}$$

où W_t est un mouvement brownien dans $\mathbb{R}$. Nous supposons dans toute cette section que l'espace d'état de $X = (X_t)_{t\ge 0}$, noté E est une intervalle fini ou infini de $\mathbb{R}$ et que X est récurrent au sens de Harris. Rappelons qu'un processus de Markov fort (X) à valeurs dans un espace polonais mesurable $(E, \mathcal{E})$ est récurrent-Harris s'il existe une mesure μ sigma-finie telle que pour chaque $A \in \mathcal{E}$ on ait

$$\mu(A) > 0 \Longrightarrow \forall x \in E : \quad P_x\left(\int_0^\infty 1_A(X_s)ds = \infty\right) = 1 \tag{5.10}$$

La mesure μ est alors invariante pour X. En fonction de la masse totale de E les processus de Harris se divisent en deux classes : si $\mu(E) < \infty$, X est dit récurrent positif (ou ergodique) et si $\mu(E) = \infty$, X est dit récurrent nul. Par exemple, le processus d' Ornstein-Uhlenbeck $dX_t = dW_t - X_t\,dt;\ E = \mathbb{R}$ est ergodique, μ étant la mesure gaussienne dans ce cas. Le mouvement brownien dans $\mathbb{R}^n$, $n \leq 2$ est nul-récurrent, admettant la mesure de Lebesgue comme mesure invariante. Un autre exemple de processus nul-récurrent est le carré de Bessel de paramètre $0 < \delta \leq 2$. Enfin, pour les diffusions à valeurs dans $\mathbb{R}$ cette notion coïncide avec $\forall x, y; \quad \mathbf{P}_x(T_y < \infty) = 1$, T_y étant le temps d'atteinte de x. En s'intéressant à des problèmes de statistique asymptotique pour les processus de diffusion nous sommes amenés à étudier le comportement limite des fonctionnelles additives de la forme $A_t = \int_0^t f(X_s)\, ds$ et des martingales de la forme $M_t = \int_0^t f(X_s)\, dW_s$. Dans le cas des diffusions ergodiques c'est un sujet bien étudié. On a la loi des grands nombres : si $\mu(f) < \infty$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(X_s)\, ds = \mu(f) \quad P_x - a.s.\ \forall x$$

et si $\mu(f^2) < \infty$, on a le théorème central limite

$$\frac{1}{\sqrt{t}} \int_0^t f(X_s)\, dW_s \longrightarrow \mathcal{N}(0, \mu(f^2)).$$

Pour les diffusions nul-récurrentes il n'existe pas de loi des grands nombres avec normalisation déterministe, et la convergence en loi a lieu uniquement pour la sous-classe des diffusions de variation régulière, voir Darling et Kac,[15], Touati [59] et Höpfner and Löcherbach [36]. Le seul théorème s'appliquant à toute la classe des diffusions récurrentes est le fameux théorème de Chacon-Ornstein, qui dit que deux fonctionnelles additives intégrables de X sont asymptotiquement equivalentes : $\forall f \geq 0$, $\mu(f) < \infty$; $\forall g \geq 0$, $0 < \mu(g) < \infty$;

$$\lim_{t\to\infty} \frac{\int_0^t f(X_s)\, ds}{\int_0^t g(X_s)\, ds} = \frac{\mu(f)}{\mu(g)} \quad P_x - a.s.\ \forall x \tag{5.11}$$

La version intégrale de ce théorème dit que $\mu - a.s.$

$$\lim_{t\to\infty} \frac{\mathbf{E}_x \int_0^t f(X_s)\, ds}{\mathbf{E}_x \int_0^t g(X_s)\, ds} = \frac{\mu(f)}{\mu(g)}; \tag{5.12}$$

où l'ensemble exceptionnel dépend de f et de g.

La littérature sur la statistique asymptotique pour les diffusions ergodiques en général et l'estimation du drift en particulier, est abondante (voir les travaux de Basawa et Prakasa Rao [4], Prakasa Rao [55], Fegin [26], Kutoyants [44], Dalalyan [17], Dalalyan et Kutoyants [**?**], Galtchouk et Pergamentchikov [30], Lanska [45], van Zanten [65, 66], Yoshida [68] ; on trouvera un survey extensif des résultats existant dans le livre [44] de Kutoyants). La situation est différente pour les diffusions nul-récurrentes. Comme nous l'avons mentionné plus haut, les diffusions à *variation régulière* forment une classe importante de diffusions nul-récurrentes. Pour ces diffusions les couples $(M_t/\sqrt{v_t}, A_t/v_t)$, convergent en loi quand $t \to \infty$, où $v_t = t^\alpha l(t)$ pour un $0 < \alpha \leq 1$ et où $l(t)$ est une fonction à variation lente. En utilisant cette convergence, Höpfner et Kutoyants [35] ont donné l'un des premiers exemples concernant la vitesse de convergence et la distribution limite du MVS d'un paramètre de drift pour une diffusion nul-récurrente.

Dans ce chapitre nous présentons un outil qui permet de traiter d'une manière unifiée le cas récurrent-positif et récurrent-nul dans certaines questions concernantes la vitesse de convergence des estimateurs. L'idée provient de l'observation suivante : la vitesse de convergence d'un estimateur est une notion de tension. Alors il se pose une question naturelle : Y-a t'il un résultat de tension pour les fonctionnelles additives avec normalisation déterministe, valable pour toute la classe des diffusions récurrentes ? Avec Oleg Loukianov dans [$L6$] nous avons donné une réponse affirmative à cette question : Pour chaque diffusion récurrente X il existe une fonction déterministe v_t, appelé dans la suite *équivalent déterministe de* X. Avec cette fonction explicitée plus bas on a le résultat suivant

$$\lim_{M\to\infty} \liminf_{t\to\infty} P(1/M < A_t/v_t < M) = 1. \tag{5.13}$$

Cette propriété a d'abord été prouvée par Chen [9] pour les chaînes de Markov Harris-récurrentes, sans rapport avec les statistiques. Dans [$L6$], en utilisant la méthode de décomposition en cycles, nous re-écrivons cette propriété pour les diffusions unidimensionnelles. Notons aussi que ultérieurement avec Eva Löcherbach ([$L8$]) nous avons étendu cette propriété à toute la classe des processus récurrents de Harris en temps continu. Pour définir l'équivalent déterministe v_t nous avons besoin d'une notion de fonction spéciale. Une fonction $f : \mathbb{R} \to \mathbb{R}^+$ mesurable et bornée est dite spéciale (J. Neveu [51]), si pour toute $h : \mathbb{R} \to \mathbb{R}^+$, mesurable bornée et telle que $\mu(h) > 0$, la h-résolvante $U^h f$ est bornée :

$$U^h f(x) = \mathbf{E}_x \int_0^\infty \exp\left[-\int_0^t h(X_s)ds\right] f(X_t)dt.$$

Cette notion peut être étendue aux fonctionnelles additives (M. Brancovan [7]). Une fonctionnelle additive A_t est spéciale si pour toute $h : R \to \mathbb{R}^+$, mesurable bornée et telle que $\mu(h) > 0$, la résolvante

$$U_A^h 1(x) = \mathbf{E}_x \int_0^\infty \exp\left[-\int_0^t h(X_s)ds\right] dA_t$$

est bornée.

Le temps local L_t^x est une *fonctionnelle additive spéciale* (SAF), d'autre part, toute fonction bornée à support compact est spéciale pour une diffusion dans $\mathbb{R}^n$. Plus précisément, si (P_t) est fortement fellerien, chaque fonction bornée à support compact est spéciale pour X, (Revuz [56]).

La propriété importante des fonctions et fonctionnelles spéciales est que pour elles le théorème de Chacon-Ornstein 5.12 est vérifié pour tout x, et que la version "renforcée" suivante a lieu :

si A, B sont deux fonctionnelles additives spéciales telles que $\|B\| := \mathbf{E}_\mu B_1 > 0$, pour tout couple (π_1, π_2) de mesures de probabilité on a (voir [7] pour la preuve)

$$\lim_{t\to\infty} \frac{\mathbf{E}_{\pi_1} A_t}{\mathbf{E}_{\pi_2} B_t} = \frac{\|A\|}{\|B\|}. \tag{5.14}$$

Maintenant pour normaliser une fonctionnelle additive nous choisissons une fonction spéciale g, $\mu(g) = 1$, et posons

$$v_t = \mathbf{E}_\pi \int_0^t g(X_s)ds,$$

où π est une probabilité sur E. Il est clair que v_t est ≥ 0 et croissante. En vertu de (5.14), pour tout autre choix v'_t d'un équivalent déterministe on aura $v_t/v'_t \to 1$, quand $t \to \infty$. Si X est une diffusion récurrente positive, $v_t \sim t$. Si X est nulle-récurrente à variation régulière d'indice α, $v_t \sim t^\alpha l(t)$ où l est une fonction à variation lente. Si X est un mouvement brownien dans $\mathbb{R}$, $v_t \sim \sqrt{t}$. Certains exemples sont moins explicites : si X est un mouvement brownien avec un drift à support compact, alors X est nulle- récurrente et $v_t \sim \sqrt{t}$. Mais dans le cas général on a besoin de connaître le semi-groupe de X pour évaluer v_t.

5.1 Application de l'équivalent déterministe à l'estimation paramétrique ([$L6$]).

Comme application de l'équivalent détérministe on calcule dans [$L6$] la vitesse de convergence de MVS pour le paramètre de drift d'une diffusion récurrente, avec drift höldérien par rapport au paramètre. Les conditions qu'on impose sur la fonction de drift sont standards et essentiellement les mêmes que dans [$L5$] : il faut que la martingale dans la définition de l'estimateur

$$\hat{\theta}_t = Argsup_{\theta \in \Theta} \left(M_t(\theta) - \frac{1}{2} A_t(\theta) \right), \tag{5.15}$$

vérifie pour t suffisamment grand une condition d'identifiabilité

$$\forall \theta \in \Theta, \qquad d^{2k}(\theta, \theta_0) U_t \leq \langle M(\theta) - M(\theta_0) \rangle_t \tag{5.16}$$

et une condition de Hölder

$$\forall (\theta, \theta') \in \theta^2, \quad \langle M(\theta) - M(\theta') \rangle_t \leq d^{2\delta}(\theta, \theta') V_t, \tag{5.17}$$

avec $\delta \leq 1$ et $k \geq \delta$, où U_t et V_t sont deux processus continus croissants vers l'infini de façon que $\liminf_{t \to \infty} U_t / V_t > 0$. La notion de vitesse qu'on considère ici est celle de la convergence en probabilité. Rappelons que $N= N(\Theta, d, \varepsilon)$ désigne l'entropie métrique de (Θ, d) c'est à dire le plus petit nombre de boules de d- rayon ε, nécessaires pour recouvrir Θ. Supposons que pour une certaine fonction $\phi : R^+ \to \mathbb{R}^+$ telle que $\phi(x)/x^p$ soit décroissante pour un $p < 2$, l'inégalité suivante soit satisfaite

$$\int_0^\eta \sqrt{\log(1 + N(\theta_\eta, d^\delta, \varepsilon))} d\varepsilon \leq \phi(\eta).$$

On a noté $\Theta_\eta = \{\theta \in \Theta : d^\delta(\theta, \theta_0) \leq \eta\}$. Par exemple nous pouvons prendre $\phi(\eta) = \eta$ si $\Theta \in \mathbb{R}^m$, mais certains compacts infini-dimensionnels satisfont aussi cette condition. Soit (v_t) l'équivalent déterministe de (X_t). Nous montrons que si le processus r_t satisfaisant

$$r_t^{2k} \phi(1/r_t^\delta) \leq \sqrt{v_t}$$

alors r_t^{-1} est une vitesse de convergence supérieure de MVS, c'est à dire que r_t vérifie

$$\lim_{M \to \infty} \limsup_{t \to \infty} \mathbf{P} \left(r_t d(\hat{\theta}_t, \theta_0) > M \right) = 0.$$

Pour les compacts dans $\mathbb{R}^n$ nous trouvons ainsi la vitesse $r_t = \sqrt{v_t^{1/(2k-\delta)}}$.

Les conditions (5.16) et (5.17) exprimées à l'aide de la fonction du drift deviennent

$$K(x) d^k(\theta, \theta') \leq |b(\theta, x) - b(\theta', x)| \leq C(x) d^\delta(\theta, \theta'),$$

nous pouvons donc prendre $U_t = \int_0^t \frac{K^2}{\sigma^2}(X_s)ds$ et $V_t = \int_0^t \frac{C^2}{\sigma^2}(X_s)ds$. La récurrence de X implique

$$\lim_{t\to\infty} V_t = \infty \quad \text{et} \quad \liminf_{t\to\infty} \frac{U_t}{V_t} \to \frac{\mu(K^2/\sigma^2)}{\mu(C^2/\sigma^2)} > 0,$$

dès que $0 < \mu(C^2/\sigma^2), \mu(K^2/\sigma^2) < \infty$. Ansi notre résultat généralise la vitesse bien connue $r_t = \sqrt{t^{1/(2k-\delta)}}$, trouvée pour les diffusions ergodiques par van Zanten [65],[66].

5.2 Application de l'équivalent déterministe à l'estimation non-paramétrique ([$L7$]).

Dans le cas paramétrique l'équivalent déterministe s'est avéré un bon outil, on voudrait le tester dans l'estimation non-parametrique. On considère une diffusion récurrente à valeurs dans $\mathbb{R}$ vérifiant des hypothèses standard pour l'existence d'une solution forte. On suppose que la fonction de drift b est α-Höldérienne au voisinage d'un certain point x_0. L'estimateur de Nadaraya-Watson est un estimateur par noyau de la fonction de drift au voisinage de x_0, cet estimateur s'écrit

$$\hat{b}^h_{x_0,t} = \frac{\int_0^t \phi(\frac{X_s-x_0}{h})dX_t}{\int_0^t \phi(\frac{X_s-x_0}{h})dt},$$

où ϕ est une fonction positive régulière à support compact. Le paramètre h représente la largeur d'une "bande" autour de x_0, c'est à dire la largeur de la fenêtre d'observation. Dans le cas récurrent positif h est une fonction de t qui peut être précisée par une procédé standard d'optimisation. Dans notre cas récurrent général, il se trouve qu'on peut faire une optimisation analogue en remplaçant t par l'équivalent déterministe v_t. Mais pour pouvoir utiliser cette démarche simple on a besoin de contrôler des quantités comme

$$A_t^{h_t} = \int_0^t \phi(\frac{X_s - x_0}{h_t})dt \quad \text{et} \quad M_t^{h_t} = \int_0^t \phi(\frac{X_s - x_0}{h_t})dW_t.$$

Notons qu'à cause de h_t qui dépend de t, $A_t^{h_t}$ n'est plus une fonctionnelle additive et $M_t^{h_t}$ n'est plus une martingale, il est donc difficile d'obtenir le contrôle de ces quantités par les méthodes classiques du calcul stochastique. Finalement pour obtenir ce contrôle, on démontre d'abord deux résultats de tension "uniforme" par rapport à $h \in [0, \delta]$ pour A_t^h et M_t^h, et ensuite on montre que les résultats de tension restent valables pour des suites "diagonales" $A_t^{h_t}$ et $M_t^{h_t}$. La vitesse de convergence de Nadaraya-Watson est alors $r_t = v_t^{\frac{\alpha}{2\alpha+1}}$, spécifiée par la largeur de la fenêtre $h_t = v_t^{\frac{-1}{2\alpha+1}}$. L'équivalent déterministe v_t dépend de la diffusion X dont on observe une trajectoire, mais qu'on ne connaît pas. Il est important de pouvoir exprimer la vitesse et la largeur de la fenêtre en fonction des quantités observables. Pour cela on montre finalement que le résultat reste valable quand on remplace v_t par une fonctionnelle additive intégrable V_t qui elle, est une fonction d'observations. Exprimée en termes de l'équivalent déterministe, dans le cas récurrent positif notre vitesse coïncide avec la vitesse bien connue $r_t = t^{\alpha/2\alpha+1}$ de même qu'avec la vitesse minimax $r_t = t^{\alpha/(4\alpha+2)}$ trouvée par Delattre and Hoffmann [22] pour le modèle du mouvement brownien avec drift à support compact. Rappelons que cette dernière diffusion est nulle-récurrente avec $v_t = \sqrt{t}$. Sous sa forme aléatoire notre vitesse coïncide avec celle de Delattre, Hoffmann and Kessler [23], où l'estimateur de Nadaraya-Watson a été étudié dans le contexte minimax . Notons finalement qu'on a obtenu ultérieurement avec Eva Löcherbach dans $[L8]$ un résultat similaire pour les diffusions multidimensionnelles, mais uniquement avec la vitesse et la fenêtre déterministes.

6 Scission de Nummelin en temps continu et applications ($[L8], [L9]$).

Les résultats de la section précédente portent sur les diffusions unidimensionnelles. La question de la généralisation de la notion d'équivalent déterministe aux cas des diffusions multidimensionnelles est l'objet de toute l'étude décrite dans cette section. Dans le cas unidimensionnel la méthode principale pour démontrer des théorèmes limites pour les fonctionnelles additives est la décomposition de la trajectoire d'une diffusion en cycles i.i.d. Pour les diffusions dans $\mathbb{R}$ la décomposition peut être faite par rapport à des visites successives d'un point donné de l'espace d'états, car la loi du moment de sortie est une masse de Dirac en ce point et "le futur" ne dépend du "passé" que par cette loi, ce qui donne l'indépendance et l'équidistribution des cycles successifs. Pour reproduire cette construction pour une diffusion à valeurs dans $\mathbb{R}^n$ il faut trouver des ensembles $A \in \mathcal{B}(\mathbb{R}^n)$ visités par le processus X_t pour t arbitrairement grand et tels que pour tout $x \in A$, on ait $P_t(X_{T_{A^c}}, dy) = \nu(dy)$, où ν est une probabilité. La question n'est pas évidente en dimension supérieure, car les points sont polaires, donc ne sont pas visités infiniment souvent, et si on essaie de décomposer la trajectoire suivant ses visites dans un ensemble de mesure invariante positive, la loi du moment de sortie dépend du point de sortie. Rappelons que pour les chaînes de Markov le même problème a été résolu en 1978 simultanément par Athreya, Ney et par Nummelin. Les ensembles A tels que la transition de la chaîne $P(x, dy)$ ne dépende pas de $x \in A$ s'appellent des *atomes*. On décompose la trajectoire de la chîne en cycles i.i.d. suivant les visites successives d'un atome. Quand l'ensemble d'états n'est pas dénombrable, les atomes n'existent pas toujours, mais on peut les introduire de manière artificielle. Cette construction, appelée "splitting" de Nummelin, a été introduite en 1978 par Athreya, Ney [2] et par Nummelin [54]. En voici un bref résumé.

Supposons que la transition P de la chaîne $(\bar{X}_n)_{n\in\mathbf{N}}$ à valeurs dans E pour un $C \in \mathcal{B}(E)$, $0 < \alpha \leq 1$, satisfasse la condition de minoration suivante

$$\mathbf{P}(x, dy) \geq \alpha \mathbf{1}_C(x)\nu(dy), \tag{6.18}$$

où ν est une probabilité portée par C, et où la mesure invariante vérifie $\mu(C) > 0$. Alors on définit une chaîne $(Z)_n = (Z_n^1, Z_n^2)_n$ à valeurs dans $E \times [0, 1]$ par la transition

$$Q((x, u), dy \times dv) = \begin{cases} \nu(dy) \times dv & \text{si} \quad (x, u) \in C \times [0, \alpha] \\ \frac{1}{1-\alpha}(P(x, dy) - \alpha\nu(dy)) \times dv & \text{si} \quad (x, u) \in C \times]\alpha, 1] \\ P(x, dy) \times dv & \text{si} \quad x \notin C. \end{cases} \tag{6.19}$$

La chaîne Z est récurrente et admet un atome $C \times [0, \alpha]$. Sa première marginale Z^1 est égale en loi à la chaîne initiale $(\bar{X})$, sa deuxième marginale est une suite de v.a. i.i.d. uniformes sur $[0, 1]$. La mesure invariante de la chaîne Z a comme première marginale la mesure invariante de la chaîne initiale. Les fonctionnelles additives de Z se décomposent en blocs i.i.d. En démontrant les théorèmes limites pour ces fonctionnelles on obtient automatiquement les théorèmes correspondant pour les fonctionnelles additives de $(\bar{X})$. Cependant cette méthode est très liée à la structure discrète de la chaîne, notament la notion de "pas suivant" et de "splitting" en temps continu restait une question ouverte.

6.1 Scission de Nummelin ([$L8$]).

Dans [$L8$] avec Eva Löcherbach, pour un processus (X_t) de Markov fort, en temps continu, récurrent, à valeurs dans un espace polonais E, de transition (P_t), on donne une construction analogue au splitting de Nummelin pour les chaînes. Cette construction permet d'introduire une décomposition en cycles presque indépendants pour les fonctionnelles additives de (X_t), en représentant (X_t) comme la première coordonnée d'un processus $(Z_t = (Z_t^1, Z_t^2, Z_t^3))$ à valeurs dans $E \times [0,1] \times E$, markovien. Pour ce dernier nous introduisons une suite de temps d'arrêts $(R_n, n \geq 1)$, par rapport à laquelle on définit une décomposition en cycles. Pour donner la construction précise nous avons besoin d'hypothèses et de considérations supplémentaires.

Hypothèse 6.1 *On suppose que pour tout $t > 0$ la transition P_t est absolument continue par rapport à la même mesure de référence Λ :*

$$P_t(x, dy) = p_t(x, y)\Lambda(dy).$$

Soit à présent $(\sigma_n)_{n\geq 1}$ une suite i.i.d. de v.a. exponentielles, indépendantes de X. On sait alors que la chaîne $\bar{X}_n := X_{T_n}$ est Harris-récurrente et que sa transition

$$U^1(x, dy) = \int_0^\infty e^{-t} P_t(x, dy)$$

satisfait la condition de minoration (6.18) (voir [36] proposition 6.7) pour un $0 < \alpha \leq 1$. Notons $Q((x,u), y)$ la transition de $E \times [0,1]$ dans E donnée par

$$Q((x,u), dy) = \begin{cases} \nu(dy) & \text{si} \quad (x,u) \in C \times [0, \alpha] \\ \frac{1}{1-\alpha}(U^1(x, dy) - \alpha\nu(dy)) & \text{si} \quad (x,u) \in C \times]\alpha, 1] \\ U^1(x, dy) & \text{si} \quad x \notin C. \end{cases} \tag{6.20}$$

Notons aussi que sous l'hypothèse (6.1) la transition U^1 est absolument continue par rapport à Λ, de densité $u^1(x,y) = \int_0^\infty e^{-t} p_t(x, x') = u^1(x, x')\Lambda(dx')$. Notons enfin que

$$\int_0^1 Q((x,u), dx')du = u^1(x, x')\Lambda(dx').$$

On construit alors le processus Z par récurrence

1. Posons $T_0 = 0$.
2. Posons $Z_0^1 = X_0 = x$. Choisissons Z_0^2 suivant la loi uniforme sur $[0,1]$, indépendamment de Z_0^1. Sachant $\{(Z_0^1, Z_0^2) = (x,u)\}$ choisissons Z_0^3 suivant $Q((x,u), dx')$.

3. Sachant $\{Z_0 = (x, u, x')\}$ choisissons

$$\sigma_1 \sim e^{-t} \frac{p_t(x, x')}{u^1(x, x')} dt \quad \text{sur} \quad \mathbb{R}^+.$$

4. Sachant $\{Z_0 = (x, u, x'), \sigma_1 = t\}$ posons $Z_s^2 = u$, $Z_s^3 = x'$ pour tout $0 \leq s < t$,
5. pour tout $0 \leq s < t$ choisissons

$$Z_s \sim \frac{p_s(x, y) p_{t-s}(y, x')}{p_t(x, x')} \Lambda(dy)$$

et sachant en plus que $Z_s^1 = y$, pour $s + u < t$, choisissons

$$Z_{s+u}^1 \sim \frac{p_u(y, y') p_{t-s-u}(y', x')}{p_{t-s}(y, x')} \Lambda(dy').$$

A l'étape 2 on pose $T_1 = \sigma_1 = t$, $Z_{T_1}^1 = x'$ et on construit le processus sur $[T_1; T_2 = T_1 + \sigma_2[$ suivant le schéma ci-dessus. Ainsi à l'étape $n \in \mathbf{N}$ on construit le processus entre $[T_n, T_{n+1} = T_n + \sigma_n[$.

On note $(\mathcal{F}_t)_{t \geq 0}$ la filtration engendrée par Z. Le processus Z ainsi construit vérifie les propriétés suivantes :
- La suite $(T_n - T_{n-1})_{n \geq 1}$ est une suite de v.a. i.i.d. $\exp(1)$, indépendante de la première coordonnée $(Z_t^1)_{t \geq 0}$ et indépendante de $\mathcal{F}_{T_n^-}$.
- Z est un processus de Markov.
- La première coordonnée Z^1 est égale en loi à X.
- $A = C \times [0, \alpha] \times E$ est un atome récurrent pour Z au sens suivant :
  Soit $R = inf\{n; Z_{T_n} \in A\}$. Alors la loi conditionnelle $\mathcal{L}(Z_{T_R}^3 | Z_{T_R}^1, Z_{T_R}^2) = \nu$, par conséquent la loi de $(Z_{T_{R+1}+s}); s \geq 0$ est égale à la loi de $(Z_s), s \geq 0$ partant de ν.

Par rapport aux visites de A par le processus Z on définit une décompositions en cycles comme suit : Soit

$$S_0 = 0, \quad , R_0 = 0, \quad S_{n+1} = \inf\{T_m > R_n; \quad Z_{T_m} \in A\},$$
$$R_{n+1} = \inf\{T_m > S_{n+1}, \quad n \geq 0\}.$$

La suite (R_n) a les propriétés :
1. pour tout $n \in \mathbf{N}$ $R_n < \infty$ p.s. et $R_{n+1} = R_n + R_1 \circ \theta_{R_n}$,
2. une fonctionnelle additive de Z $(A_t)_{t \geq 0}$ peut être décomposée en blocs $\xi_n := (A_{R_{n+1}} - A_{R_n})_{n \geq 0}$ qui sont deux-dépendants, c'est à dire $\forall n \in \mathbf{N}, \quad \xi_n$ est indépendant de $\{\xi_k, k \in \mathbf{N}, \quad k \neq n \pm 1.\}$

La continuité empêche d'obtenir la vraie indépendance des blocs, mais seulement la deux-dépendance décrite ci-dessus. Dans $[L8]$ on applique finalement la technique du splitting en temps continu pour démontrer l'existence d'un équivalent déterministe pour X. On obtiens le théorème

Théorème 6.2 *Soit g une fonction spéciale de X, soit m une probabilité sur E. Posons*

$$v_t = \mathbf{E}_m \int_0^t g(X_s) ds.$$

Alors pour toute fonctionnelle additive intégrable de X on a

$$\lim_{M \to \infty} \liminf_{t \to \infty} \mathbf{P}_\pi \left(1/M \leq A_t / v_t \leq M\right) = 1.$$

6.2 Application à l'estimation par noyau pour les diffusions multidimensionnelles ([$L8$]).

Nous utilisons l'équivalent déterministe qu'on a obtenu dans [$L8$] pour traiter le problème de l'estimation à noyau de la dérive d'une diffusion récurrente en dimension $d \geq 2$ – problème qui jusqu'alors n'a pu être traité que dans le cas $d = 1$, par des méthodes strictement unidimensionnelles, car utilisant le temps local, voir Delattre Hoffmann et Kessler [23], Kutoyants [44]. Le problème qu'on expose ici est une généralisation aux dimensions supérieures du problème qu'on a traité dans [$L7$].

Nous considérons la solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,$$

où $b : R^d \to \mathbb{R}^d$, $\sigma : R^d \to \mathbb{R}^{d\times m}$, avec b et σ continus et tels qu'une solution forte existe. Pour satisfaire les hypothèses du splitting nous supposerons aussi que le semi-groupe de X admet une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Enfin nous supposons que X est récurrent de Harris, avec la mesure invariante $\mu(dx) = p(x)dx$ où p est continue. A partir des observations continues d'une trajectoire de X nous nous proposons d'estimer la fonction de drift b au voisinage d'un point donné $x_0 \in \mathbb{R}^d$. On impose une condition de régularité sur b au voisinage de x_0 : Pour certains $\delta, \gamma > 0$ et $\alpha \in]0, 1[$ on suppose la condition de Hölder suivante :

$$\sup_{x\in[x_0-\delta,x_0+\delta]} |b(x) - b(x_0)| \times |x - x_0|^{-\alpha} \leq \gamma$$

L'estimation du drift se fait ici à l'aide de l'estimateur de Nadaraya-Watson, construit comme suit : Soit $\phi : \mathbb{R}^d \to \mathbb{R}^+$ continue à support compact, telle que $\int_{\mathbb{R}^d} \phi(x)dx) = 1$. Pour tout $h > 0$ posons $\phi_h(y) = \phi(\frac{y-x_0}{h})/h^d$. Le paramètre h joue le rôle de largeur de la fenêtre d'observation. Soit v_t l'équivalent déterministe de X et posons $h_t = v_t^{-\frac{1}{2\alpha+d}} \wedge \delta$. Définissons

$$\hat{b}_t = \frac{\int_0^t \phi_{h_t}(X_s)dX_s}{\int_0^t \phi_{h_t}(X_s)ds}.$$

Nous montrons que l'estimateur à noyau de Nadaraya-Watson associé à la largeur de la fenêtre $h_t = v_t^{-\frac{1}{2\alpha+d}}$ atteint la vitesse de convergence $r_t = v_t^{\alpha/(2\alpha+d)}$ au sens suivant :

$$\lim_{M\to\infty} \limsup_{t\to\infty} \mathbf{P}(r_t|\hat{b}_t - b(x_0)| > M) = 0.$$

6.3 Loi du logarithme itéré pour des fonctionnelles additives de processus de Markov récurrents ([$L9$]).

Comme autre application de la méthode du splitting en temps continu nous avons démontré avec Eva Löcherbach la loi du logarithme itéré pour des fonctionnelles additives A_t intégrables positives du processus de Markov $(X_t)_{t\geq 0}$ Harris-récurrent. Notons que ce sujet avait déjà été très étudié dans nombreux cas particuliers, par exemple pour le mouvement brownien en dimension un et deux par Csáki, Fôldes et Hu [11] , Földes et Hu [14], et aussi pour les chaînes Harris-récurrentes, voir les travaux de Chen [10],[11]. Pour tout $\lambda \geq 0$ notons $L_2(\lambda) = (\log\log\lambda)\wedge 1$. Soit A une fonctionnelle additive intégrable de X et (v_t) l'équivalent déterministe de X.

Théorème 6.3 *Il existe une constante $0 < c < \infty$, dépendant seulement du processus, telle que*

$$\limsup_{t\to\infty} \frac{A_t}{v(\frac{t}{L_2(v(t))})L_2(v(t))} = c\int_E \mu(dx)E_x(A_1) \tag{6.21}$$

Soit à présent une fonction f telle que
– $\mu(f) = 0$
– $|f|$ est bornée spéciale.
On dit en ce cas que f est une *charge*. Alors pour toute fonctionnelle additive integrable nous avons

Théorème 6.4 *Il existe une constante $\lambda_f \geq 0$ telle que*

$$\limsup_{t\to\infty} \frac{\int_0^t f(X_s)ds}{\sqrt{2A_tL_2(A_t)}} = (\mathbf{E}_\mu(A_1))^{-1/2}\lambda_f a.s.$$

Remarque 6.5 *Si en plus $\int_0^\infty f(x)P_tf(x)dt \in L^1(\mu)$, alors*

$$(\lambda_f)^2 = 2\int_E\int_0^\infty f(x)P_tf(x)dt\mu(dx).$$

Ces résultats généralisent ceux de Chen [10] pour les chaînes de Markov.

Exemple 6.6 *Comme exemple d'application de ces résultats considérons l'estima- teur du MVS pour le paramètre θ dans le modèle*

$$dX_t = \theta b(X_t)dt + \sigma dW_t, \quad X_0 = 0,$$

où $\sigma > 0$ est connu, b,σ sont localement lipschitziens et tels que X soit récurrent. L'estimateur du MVS s'écrit alors

$$\hat{\theta}_t = \int_0^t b(X_s)dX_s/\int_0^t b^2(X_s)ds.$$

A l'aide du théorème 6.4 nous obtenons une borne pour la vitesse de convergence p.s. de $\hat{\theta}_t$:

$$\limsup_{t\to\infty} \frac{\sqrt{I_t}}{\sqrt{\log\log I_t}}(\hat{\theta}_t - \theta_0) < \infty,$$

avec $I_t = \int_0^t b^2(X_s)ds$.

7 Estimation pénalisée du drift ([$L10$]).

Pendant mon semestre de délégation CNRS au laboratoire MAP5 (février-juillet 2008) je me suis intéressée au sujet de sélection de modèles en statistique, sujet très développé dans cette équipe notamment par Fabienne Comte et Valentine Genon-Catalot. Ce séjour a donné lieu à deux articles, en collaboration avec Eva Löcherbach et Oleg Loukianov, dont l'un est [$L10$]. Dans ce travail nous nous posons le problème de l'estimation de la fonction de drift sur un intervalle donné pour une diffusion unidimensionnelle récurrente positive X, solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x. \tag{7.22}$$

Nous supposons que la trajectoire $(X_s, 0 \le s \le t)$ est observée de façon continue, et que les coefficients sont sujets aux deux hypothèses suivantes :

Hypothèse 7.1 *1. b et σ sont localement lipschitziens et b est à croissance au plus linéaire à l'infini.*

2. Il existe $0 < \sigma_0^2 \le \sigma_1^2 < \infty$ tels que pour tout $x \in \mathbb{R}$, $\sigma_0^2 \le \sigma^2(x) \le \sigma_1^2$.

Hypothèse 7.2 *1. Il existe deux constantes connu M_0 et b_0 telles que $K \subset [-M_0, M_0]$ et pour tout x vérifiant $|x| \le M_0$, $|b(x)| \le b_0$.*

2. Il existe une constante positive γ telle que pour tout x vérifiant $|x| \ge M_0$,

$$xb(x) \le -\gamma.$$

3. La constante γ satisfait la condition $2\gamma > 31\sigma_1^2$.

La première hypothèse ne sert qu'à assurer l'existence et l'unicité d'une solution de (7.22) quand à la deuxième, elle implique une certaine vitesse de l'ergodicité de X, notamment elle assure que la densité invariante de X décroît à l'infini au moins comme $|x|^{-31}$. Le sens de 31 sera précisé plus bas. On ne fait aucune hypothèse de régularité de drift, on explore par conséquent la question de l'estimation adaptative de b.

On se donne la fonction de risque

$$Risk(\hat{b}, b) = \mathbf{E}_{x,b}[\frac{1}{t}\int_0^t (\hat{b} - b)(X_s)\mathbf{1}_K(X_s)ds]$$

et la fonction de contraste

$$\gamma_t(h) = \frac{1}{t}\int_0^t h^2(X_s)ds - \frac{2}{t}\int_0^t h(X_s)dX_s.$$

Notre estimateur $\hat{b}$ est obtenu par minimisation de γ_t sur une sous espace de dimension finie de $\mathbb{L}^2([0,1])$ noté $\mathcal{S}_m$ (un modèle), où $m \in \mathcal{M}_t$ et $\mathcal{M}_t$ est un certain ensemble d'indices, dont la taille et bornée par t. On défini donc d'abord

$$\hat{b}_m = argmin_{h \in S_m} \gamma_t(h).$$

Pour adapter l'estimateur à la régularité inconnue de la fonction b, on utilise la procedé de la "sélection des modèles", dans l'esprit de [13]. Pour une collection finie $\{\mathcal{S}_m : m \in \mathcal{M}_t\}$ de sous-espaces linéaires de $\mathrm{Ł}^2(K, dx)$ on définit

$$\hat{b}_t = \hat{b}_{S_{\hat{m}}}, \quad \text{où} \quad \hat{m} = argmin_{m\in\mathcal{M}_t}\{\gamma_t(\hat{b}_{S_m}) + pen(m)\}.$$

Le terme de la pénalisation $pen(m)$ est proportionnelle à la dimension de $\mathcal{S}_m$. On adopte cette définition quand la forme quadratique associé au γ_t est non-dégénérée. Sinon on pose $\hat{b} = 0$. Dans le $[L11]$, sous des conditions habituelles et naturelles sur l'ensemble de modèles nous prouvons que

$$Risk(\hat{b}, b) \leq C_1 min_{m\in\mathcal{M}_t}(\|b - Proj_{S_m} b\|^2_{L^2(K,dx)} + t^{-1}dim(S_m)) + C_2 t^{-1}. \tag{7.23}$$

La borne sur le risque est non-asymptotique et atteint la vitesse minimax optimale quand $t \to \infty$. L'intérêt principal de notre travail est que nous ne supposons pas que le processus se trouve dans le régime stationnaire, ni que la diffusion ait la propriété d'érgodicité géométrique (hypothèses standard dans ce contexte). En particulier (mais pas seulement) l'hypothèse d'érgodicité géométrique permet d'estimer la probabilité de l'ensemble A_t^c où $\hat{b}$ n'est pas défini et posé égal à zéro. L'estimation obtenue est en effet une surestimation, car elle donne l'ordre e^{-t}. En remplaçant l'hypothèse de l'érgodicité géometrique par l'hypothèse (7.2) on peut quand même estimer A_t^c, ainsi que donner d'autres estimations nécessaires analogues, mais à l'ordre exact $1/t$, dicté par le terme de la variance dans la borne (7.23). Pour mettre au point la preuve sans les deux hypothèses standard il nous a fallu démontrer une nouvelle inégalité de déviation dans le théorème ergodique $[L11]$. L'inégalité en question donne une borne polynomiale. L' hypothèse (7.2) calibre le drift de manière à avoir exactement la vitesse $1/t$ pour tous les termes complémentaires dans la borne (7.23), voir exactement le corollaire (8.12) du chapitre suivant. L'inégalité polynômiale est le sujet de l'article $[L11]$, exposé dans la section suivante.

8 Déviations polynomiales dans le théorème érgodique. ($[L11]$)

Soit X une diffusion uni-dimensionnelle avec données initiales arbitraires

$$dX_t = \beta(X_t)\,dt + \sigma(X_t)\,dW_t.$$

Dans toute cette section nous supposons une condition standard pour assurer l'existence et l'unicité d'une solution forte de cette équation :

Hypothèse 8.1 *1. Pour tout x, $\sigma^2(x) > 0$.*

2. β et σ sont localement lipschitziens et $|\sigma(x)| + |\beta(x)| \leq C(1 + |x|)$, pour certain $C > 0$.

Nous supposons aussi dans toute cette section que

Hypothèse 8.2 *X est récurrente positive de probabilité invariante μ.*

Si X est une diffusion récurrente positive le Théorème Ergodique dit que pour tout $x \in \mathbb{R}$, $f \in \mathbb{L}^1(\mu)$ and $\varepsilon > 0$

$$\lim_{t\to\infty} \mathbf{P}_x \left(\left| \frac{1}{t} \int_0^t f(X_s)ds - \mu(f) \right| \geq \varepsilon \right) \to 0. \tag{8.24}$$

Il est pourtant très utile dans de nombreuse application d'avoir une borne non-asymptotique pour la probabilité (8.24). On s'en est rendu compte pendant le travail sur l'estimation pénalisée du drift. On a aussi observé qu'il y a nécessité d'une telle borne dans d'autres situations liées à l'estimation non-asymptotique (voir [13], [30]). D'une manière très "sommaire" une telle borne est utile chaque fois qu'on veut remplacer la quantité aléatoire $\frac{1}{t}\int_0^t f(X_s)ds$ par la quantité déterministe $\mu(f)$, sur tout l'espace à l'exception d'un "petit" ensemble. "Petit" habituellement signifie que sa probabilité décroît exponentiellement vite quand $t \to \infty$, et ce cas a déjà été largement étudié, voir [12], [46]. Nous n'avons pas trouvé dans la littérature de résultats connus à ce jour, concernant d'autres taux. Et pourtant considérer des taux plus lents suffit dans nombreux cas et permet d'alléger les hypothèses sur le modèle. Dans [$L11$] nous obtenons une borne polynômiale pour la probabilité (8.24). La preuve est basée sur la méthode régénérative. Rappelons qu'une des façons d'introduire les instants de régénération pour une diffusion dans $\mathbb{R}$ est la suivante : fixons deux points $a < b$, $a, b \in \mathbb{R}$. Définissons une suite de temps d'arrêt $(S_n)_n$, $(R_n)_n$ par : $S_0 = 0$, $R_0 = 0$,

$$S_1 := \inf\{t \geq 0 : X_t = b\}, \quad R_1 := \inf\{t \geq S_1 : X_t = a\},$$

et pour $n \geq 1$,

$$S_{n+1} := \inf\{t > R_n : X_t = b\}, \quad R_{n+1} := \inf\{t \geq S_{n+1} : X_t = a\}.$$

Il est bien connu que la suite $(R_n)_n$ "découpe" le processus en des blocs i.i.d. dans le sens suivant : posons

$$\xi_n = \int_{R_n}^{R_{n+1}} f(X_s)ds, \quad n \geq 0,$$

où $f : R \to \mathbb{R}$ est mesurable bornée. Alors nous avons la proposition suivante :

Proposition 8.3 *Pour toute distribution initiale ν, la suite $(\xi_n)_{n\geq 1}$ est une suite i.i.d. sous $\mathbf{P}_\nu$. Pour tout $n \geq 1$, la loi de ξ_n sous $\mathbf{P}_\nu$ est la même que la loi de ξ_0 sous $\mathbf{P}_a$.*

Alors sous une condition naturelle d'intégrabilité polynômiale des cycles de vie délimités par deux temps de régénération succéssifs on peut formuler les deux théorèmes suivants :

Théorème 8.4 *Supposons les conditions 8.1 et 8.2. Soit $f \in L^1(\mu)$. Supposons que $\|f\|_\infty < \infty$. Soit ν une distribution initiale et $0 < \varepsilon < \|f\|_\infty$. Supposons qu'il existe un $p > 1$ tel que $\mathbf{E}_\nu(R_1)^{p/2} < \infty$ et $\mathbf{E}_\nu(R_2 - R_1)^p < \infty$. Alors pour tout $t \geq 1$ on a l'inégalité :*

$$\mathbf{P}_\nu \left(\left| \frac{1}{t} \int_0^t f(X_s)ds - \mu(f) \right| > \varepsilon \right) \leq \left\{ \begin{array}{ll} K(l,p,\nu,X)\frac{1}{\varepsilon^p}\|f\|_\infty^p\, t^{-p/2} & \text{if } p \geq 2 \\ K(l,p,\nu,X)\frac{1}{\varepsilon^p}\|f\|_\infty^p\, t^{-\frac{p-1}{2}} & \text{if } 1 < p < 2 \end{array} \right\}.$$

Ici $K(l,p,\nu,X)$ est une constante positive, différente dans les deux cas, qui dépend de l, p, ν et X (par la décomposition en cycles indépendantes), mais qui ne dépend pas de f, t, ε.

Dans le cas où f est en plus supposée bornée à support compact, nous pouvons formuler l'inégalité de déviations où la borne dépend de f à travers sa $\mathbb{L}^1(\mu)$- norme et non sa sup-norme. Dans certaines situations pratiques ceci est d'une importance majeure.

Théorème 8.5 *Supposons les conditions 8.1 et 8.2. Soient $f : \mathbb{R} \to \mathbb{R}$ bornée à support compact et ν une distribution initiale. Supposons qu'il existe un $p \in \mathbf{N}$, $p > 1$, tel que $\mathbf{E}_\nu(R_1)^p < \infty$ et $\mathbf{E}_\nu(R_2 - R_1)^p < \infty$. Alors pour tout $t \geq 1$, pour tout $0 < \varepsilon < \mu(|f|)$ on a l'inégalité :*

$$\mathbf{P}_\nu \left(\left| \frac{1}{t} \int_0^t f(X_s)ds - \mu(f) \right| > \varepsilon \right) \leq K(l,p,X) \frac{1}{\varepsilon^p} \mu(|f|)^p \, t^{-p/2}.$$

Ici $K(l,p,X)$ est une constate positive, dépendant de l, p, X, et indépendante de f, t, ε.

Pour $a \in \mathbb{R}$ notons T_a le temps d'atteinte de a par X ; $T_a = \inf\{t \geq 0; \quad X_t = a\}$.

Remarque 8.6 *Puisque $\mathbf{E}_\nu(R_1)^p \leq 2^{p-1}(\mathbf{E}_\nu T_b^p + \mathbf{E}_b T_a^p)$, nous pouvons voir que les hypothèses des théorèmes 8.4 et 8.5 sont satisfaites si $\mathbf{E}_a T_b^p < \infty$, $\mathbf{E}_b T_a^p < \infty$ et $\mathbf{E}_\nu T_b^{p/2} < \infty$.*

La question naturelle est alors celle des conditions sur les coefficients de X, pour que les moments des temps d'atteinte d'ordre p, $p > 0$, soient finis. Cette question a été déjà étudiée par plusieurs auteurs, voir par exemple [8], [25], [67].

Dans ([$L11$]) nous montrons la propriété de "tout ou rien" suivante : le moment $\mathbf{E}_x T_y^p$, $p \geq 1$, existe (ou non) simultanément pour tous les couples $x < y$, (resp. $x > y$) et nous donnons une étude fine des moments d'ordre $p > 1$ des temps d'atteinte dans le cas d'une diffusion dans $\mathbb{R}$. Pour formuler nos résultats nous avons besoin de deux hypothèses. La première va garantir la finitude des moments jusqu'à un certain ordre :

Hypothèse 8.7 *Ils existent $M_0 > 0$, $\sigma_0 > 0$, $-\infty < \gamma < 1$ et $r > 0$ tels que*

$$\sigma_0 |x|^\gamma \leq |\sigma(x)| \quad \textit{and} \quad -\frac{x\beta(x)}{\sigma^2(x)} \geq r \quad \textit{pour } |x| > M_0.$$

La deuxième hypothèse assure que à partir d'un certain ordre les moments des temps d'atteinte seront infinis.

Hypothèse 8.8 *Ils existent $M_0 > 0$, $\sigma_1 > 0$, $-\infty < \delta < 1$ et $R > 0$ tels que*

$$0 < |\sigma(x)| \leq \sigma_1 |x|^\delta \quad \textit{and} \quad 0 < -\frac{x\beta(x)}{\sigma^2(x)} \leq R \quad \textit{pour } |x| > M_0.$$

Soit

$$p^* = \sup\{p > 0 \ : \ \int_x^\infty \frac{\xi^p}{\sigma^2(\xi)s(\xi)} d\xi < \infty\}.$$

où

$$s(t) = \exp\left(-2 \int_0^t \frac{\beta(u)}{\sigma^2(u)} du \right).$$

Théorème 8.9 *Let $M_0 < a < x$ or $x < a < -M_0$.*

1. *Supposons que l'hypothèse 8.7 est vérifié avec $2r + 2\gamma > 1$. Pour chaque $m \in \mathbb{R}$ vérifiant $1 \leq m < (2r+1)(1-\gamma)^{-1}/2$ posons $\alpha = m - [m]$. Alors*

$$\mathbf{E}_x T_a^m \leq \frac{x^{2m(1-\gamma)}}{r_m \sigma_0^{2m}(1-\gamma)^m},$$

   où $r_m = (2r + 2\gamma - 1)^\alpha \prod_{k=1}^{[m]}(2r - 2(k+\alpha)(1-\gamma) + 1)$.

2. *Sous l'hypothèse 8.8, pour tout entier $n \geq 1$:*
   - *si $n \leq p^*(1-\delta)^{-1}/2 + 1$ alors*

$$\mathbf{E}_x T_a^n \geq \frac{(x-a)^{2n(1-\delta)}}{R_n \sigma_1^{2n} \kappa^n},$$

     où $R_n = \prod_{k=1}^{n}(2R - 2k(1-\delta) + 1)$ et $\kappa = 1 \vee (1-\delta)$.
   - *si $n > p^*(1-\delta)^{-1}/2 + 1$, en particulier si $n > (2R+1)(1-\delta)^{-1}/2$, alors $\mathbf{E}_x T_a^n = \infty$.*

La base de cette étude est une version généralisée de la formule de Kac, également démontrée dans [$L11$]. Rappelons que la formule de Kac classique relie le moment d'ordre n d'un temps d'atteinte (et plus généralement d'une fonctionnelle additive) avec le moment d'ordre $n-1$, (voir Fitzsimmons et Pitman [27].) Pour formuler la formule de Kac généralisée nous avons besoin de notations supplémentaires. Pour $a \leq x \leq b$ soit

$$T_{a,b} = \inf\{t \geq 0;\ X_t \notin]a,b[\}.$$

Soit G une fonction de Green associée au temps de sortie $T_{a,b}$.

$$G(a,b,x,\xi) = \begin{cases} \frac{(S(x)-S(a))(S(b)-S(\xi))}{S(b)-S(a)} & a \leq x \leq \xi \leq b \\ \frac{(S(b)-S(x))(S(\xi)-S(a))}{S(b)-S(a)} & a \leq \xi \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

Théorème 8.10 *(Formule de Kac généralisée.)*

Soit $f : \mathbb{R} \to \mathbb{R}$ telle que la fonction $x \to \mathbf{E}_x f'(T_{a,b})$ soit continue sur $[a,b]$. Alors

$$\mathbf{E}_x f(T_{a,b}) = f(0) + \int_{-\infty}^{+\infty} G(a,b,x,\xi)\mathbf{E}_\xi f'(T_{a,b}) m(\xi) d\xi. \tag{8.25}$$

En utlisant les conditions sur les coefficients données dans le théorème 8.9 nous donnons une versions "pratique " pour chacun des théorèmes 8.5 et 8.4.

Corollaire 8.11 *Supposons que X satisfasse l'hypothèse 8.1 et que l'hypothèse 8.7 soit vérifié avec $2r + 2\gamma > 1$. Soient $1 < p < (2r+1)(1-\gamma)^{-1}/2$ et $f \in L^1(\mu)$, telle que $\|f\|_\infty < \infty$. Alors pour tout loi initiale ν vérifiant $\int_{\mathbb{R}} |x|^{p(1-\gamma)} d\nu(x) < \infty$, pour tout $0 < \varepsilon < \|f\|_\infty$ et $t \geq 1$, on a l'inégalité*

$$\mathbf{P}_\nu\left(\left|\frac{1}{t}\int_0^t f(X_s)ds - \mu(f)\right| > \varepsilon\right) \leq \left\{\begin{array}{ll} K(l,p,\nu,X)\frac{1}{\varepsilon^p}\|f\|_\infty^p\, t^{-p/2} & \text{if } p \geq 2 \\ K(l,p,\nu,X)\frac{1}{\varepsilon^p}\|f\|_\infty^p\, t^{-\frac{p-1}{2}} & \text{if } 1 < p < 2 \end{array}\right\}. \tag{8.26}$$

Ici $K(l,p,\nu,X)$ est une constante positive différente dans chaque cas, qui ne dépend pas de f, t, ε. En particulier, (8.26) est valable sous $\mathbf{P}_x$ pour tout $x \in \mathbb{R}$.

Corollaire 8.12 *Supposons que X satisfasse l'hypothèse 8.1 et que l'hypothèse 8.7 soit vérifié avec $2r+4\gamma > 3$. Soit $f : \mathbb{R} \to \mathbb{R}$ bornée à support compact. Soit ν une loi initiale telle que $\int_{\mathbb{R}} |x|^{p(1-\gamma)} d\nu(x) < \infty$. Alors pour tout $p \in \mathbf{N}$, $2 \leq p < (2r+1)(1-\gamma)^{-1}/2$, pour tout $0 < \varepsilon < \mu(|f|)$ et $t \geq 1$, on a l'inégalité*

$$\mathbf{P}_\nu \left(\left| \frac{1}{t} \int_0^t f(X_s) ds - \mu(f) \right| > \varepsilon \right) \leq K(l,p,X) \frac{1}{\varepsilon^p} \mu(|f|)^p \, t^{-p/2}. \qquad (8.27)$$

Ici $K(l,p,X)$ est une constante positive qui ne dépend pas de f, t, ε.

9 Trou spectral et moments exponentiels des temps d'atteinte d'une diffusion unidimensionnelle ([$L12$]).

On a observé tout le long de la section précédente l'importance des moments polynômiaux des temps d'atteinte. La formule de Kac (8.25) donne une CNS pour l'existence de tels moments : le moment d'ordre $n+1$ existe si et seulement si le moment d'ordre n est intégrable par rapport à la mesure invariante. Il se pose alors la question naturelle de trouver une CNS d'existence des moments exponentiels. C'était une des motivations de l'article [$L12$], en collaboration avec Oleg Loukianov et Shiqi Song, article dont on donne le résumé dans ce chapitre.

Soit $(X_t \; ; \; t \geq 0)$ un processus de Markov régulier, continu et à valeurs dans $\mathbb{R}$. Nous supposons que X est récurrent positif. Soient $S(x)$ sa fonction d'échelle et $m(dx)$ la mesure de vitesse associée à S. Soit $a \in \mathbb{R}$ et $T_a = \inf\{t \geq 0 \;\; : X_t = a\}$ le temps d'atteinte de a par X. Définissons les quantités

$$B_a^+ = \sup_{x \geq a} m(]x, +\infty[)(S(x) - S(a)) \quad B_a^- = \sup_{x \leq a} m(]-\infty; x[)(S(a) - S(x)) \qquad (9.28)$$

Dans [$L12.$] nous étudions quelques relations caractéristiques entre les moments exponentiels des temps d'atteinte de X, la finitude des quantités B_a^+, B_a^- données par (9.28), les inégalités de Hardy et de Poincaré pour la forme de Dirichlet associée à X. Comme conséquence nous donnons une chaîne d'équivalences entre tous ces objets.

La première question que nous nous posons dans ce papier est l'existence des moments exponentiels $\mathbf{E}_x[e^{\lambda T_a}]$, $x \in \mathbb{R}$, $\lambda > 0$. Dans certains cas particuliers ces moments ont été bien étudiés. Mentionnons, par exemple, Ditlevsen [24] pour le processus d'Ornstein-Uhlenbeck , Giorno, Nobile, Riccardi, Sacredote [32] pour Ornstein-Uhlenbeck et Bessel, Deaconu et Wantz [19] pour une diffusion avec une forte dérive, Darling et Siegert [16] et le livre de Borodin Salminen [6] pour un overview des formules connues. Mais nous n'avons pas pu trouver dans la littérature un critère général simple d'existence des moments exponentiels en termes de la fonction d'échelle $S(x)$ et de la mesure de vitesse $m(dx)$. Dans [$L12.$] nous donnons un tel critère en termes de $B_a^\pm$. Enonçons d'abord la proposition suivante, également démontrée dans [$L12.$]. Cette proposition sera référée dans la suite comme la "propriété de tout-ou rien" des moments exponentiels.

Proposition 9.1 (tout-ou-rien) *Soit $a \in \mathbb{R}$ et $\lambda > 0$. Les propriétés suivantes sont équivalentes :*

– *pour un $x > a$,* $\mathbf{E}_x \exp(\lambda T_a) < \infty$
– *pour tout $x > a$,* $\mathbf{E}_x \exp(\lambda T_a) < \infty$
– $\int_a^{+\infty} \mathbf{E}_\xi \exp(\lambda T_a) dm(\xi) < \infty$
Des propriétés analogues sont vraies pour $x < a$.

A présent soit λ_a^+ le supremum des $\lambda > 0$ tels que $\mathbf{E}_x e^{\lambda T_a} < \infty$ pour un $x > a$ (par conséquent pour tout $x > a$). Soit λ_a^- le supremum de $\lambda > 0$ tels que $\mathbf{E}_x e^{\lambda T_a} < \infty$ pour un (et donc pour tout) $x < a$. Notre premier résultat dit que

$$\frac{1}{4B_a^+} \le \lambda_a^+ \le \frac{1}{B_a^+}$$
$$\frac{1}{4B_a^-} \le \lambda_a^- \le \frac{1}{B_a^-}$$

où B_a^+ et B_a^- peuvent eventuellement être infinis. La positivité de $\lambda_a^\pm$ est alors équivalente à la finitude de $B_a^\pm$. De plus, si B_a^+ (resp. B_a^-) est fini pour un $a \in \mathbb{R}$, il est de même pour tout $a \in \mathbb{R}$. En réalité les quantités B_a^+, B_a^- sont déjà parues il y a longtemps dans d'autres contextes. En utilisant la méthode de Krein, Kac et Krein [39] et Kotani, Watanabe [42] ont montré l'équivalence entre la finitude de $B_0^+ \vee B_0^-$ et l'existence du trou spéctral γ pour le générateur de X dans le cas $S(x) = x$ (équivalence de Kac-Krein). Ils ont notamment prouvé (théorème 4 de l'appendice de [42]) que

$$B_0^+ \vee B_0^- \le \gamma^{-1} \le 4(B_0^+ \vee B_0^-). \tag{9.29}$$

Plus tard M. Artola, G. Talenti and G. Tomaselli [50], Bobkov and Götze [5] et Malrieu et Roberto [48] ont utilisé des quantités semblables pour caractériser un couple de probabilités μ et ν sur $\mathbb{R}$ satisfaisant des inégalités du type Hardy, voir aussi Miclo [49]. Dans le cas où μ est absolument continue par rapport à la mesure de Lebesgue, Malrieu and Roberto considèrent les quantités

$$\tilde{B}_a^+ = \sup_{x \ge a} \int_x^\infty d\mu(t) \int_a^x \left(\frac{d\mu}{dt}\right)^{-1} dt,$$
$$\tilde{B}_a^- = \sup_{x \le a} \int_{-\infty}^x d\mu(t) \int_x^a \left(\frac{d\mu}{dt}\right)^{-1} dt.$$

Ils montrent [48, theorem 6.6.2] que la constante optimale C_p dans l'inégalité de Poincaré

$$Var_\mu(f) \le C \int_{-\infty}^{+\infty} \left(\frac{df}{dx}\right)^2 (x) d\mu(x),$$

vérifie

$$\frac{1}{2} \tilde{B}_{me}^+ \vee \tilde{B}_{me}^- \le C_p \le 4(\tilde{B}_{me}^+ \vee \tilde{B}_{me}^-). \tag{9.30}$$

où me est la médiane de μ. Notons que les quantités $\tilde{B}^\pm$ correspondent aux quantités $B^\pm$ quand $S(x) = x$, $m(dx) = \mu(dx)$; néanmoins nous ne supposons pas que m soit absolument continue par rapport à la mesure de Lebesgue.

En suivant ces idées nous relions dans [$L12$] B_a^+ et B_a^- à l'inégalité de Hardy et Poincaré associées avec la forme de Dirichlet de X $\mathcal{E}(\mathcal{F}) = \int \left(\frac{dF}{dS}\right)^2 dS$ et par conséquent au trou spéctral du générateur de X. Nous montrons que les meilleures constantes dans les inégalités de Hardy A_a^+, A_a^-

$$\int_a^\infty (F(x)-F(a))^2 dm(x) \leq A_a^+ \int_a^\infty \left(\frac{dF}{dS}\right)^2 (t) dS(t)$$
$$\int_{-\infty}^a (F(x)-F(a))^2 dm(x) \leq A_a^- \int_{-\infty}^a \left(\frac{dF}{dS}\right)^2 (t) dS(t)$$

sur un espace de fonctions approprié $\mathcal{F}$, satisfont

$$B_a^+ \leq A_a^+ \leq 4B_a^+$$
$$B_a^- \leq A_a^- \leq 4B_a^-$$

De plus, si c_P est la meilleure constante c de l'inégalité de Poincaré

$$\int_{-\infty}^{+\infty} \left(F(x) - \frac{m(F)}{m(\mathbb{R})}\right)^2 dm(x) \leq c \int_{-\infty}^{+\infty} \left(\frac{dF}{dS}\right)^2 (x) dS(x),$$

alors c_P satisfait

$$\sup_a (A_a^+ \wedge A_a^-) \leq c_P \leq \inf_a (A_a^+ \vee A_a^-). \tag{9.31}$$

La partie droite $\mathcal{E}(\mathcal{F}) = \int_{-\infty}^{+\infty} \left(\frac{dF}{dS}\right)^2 (x) dS(x)$ de l'inégalité de Poincaré est la forme de Dirichlet associée à X L'inégalité de Poincaré donne alors de manière usuelle la borne pour le trou spectral $\gamma = 1/c_P$ de générateur de X. Ainsi (9.31) donne une autre preuve de l'équivalence de Kac-Krein.

A ce stade citons le théorème de Carmona et Klein [8] affirmant que si le générateur d'un processus de Markov admet un trou spectral, alors les temps d'atteinte du processus possèdent des moments exponentiels. Notre résultat montre que dans nos conditions ces propriétés sont toutes les deux équivalentes à la finitude de B_a^+ et B_a^- pour un (et donc pour tout) $a \in \mathbb{R}$. Finalement, nous précisons cette équivalence en reliant de façon vraiment directe les moments exponentiels des temps d'atteinte au trou spectral associé à X. Pour cela nous remarquons d'abord que l'inégalité de Hardy est l'inégalité de Poincaré pour le processus tué en T_a. Ainsi les constantes de Hardy $A_a^\pm$ sont liées de la même façon au trous spectraux $\gamma_a^\pm$ de générateur de X tué en T_a. Notamment

$$\frac{1}{A_a^+} = \gamma_a^+ \quad \text{et} \quad \frac{1}{A_a^-} = \gamma_a^-.$$

Dans la dernière partie de $[L12]$ nous montrons que pour tout $a \in \mathbb{R}$,

$$\gamma_a^+ = \lambda_a^+ \quad \text{and} \quad \gamma_a^- = \lambda_a^-.$$

Une identité semblable pour les temps de sortie d'un domaine borné D est connue depuis les travaux de Khasminskii [40] et Friedman [28]. Plus exactement, si τ est le temps de sortie de D et X^D est le processus tué en τ, alors on a l'égalité entre la largeur du trou spectral de générateur de X^D et le supremum des $\lambda > 0$ tels que $\mathbf{E}_x e^{\lambda\tau} < \infty$ pour tout $x \in D$.

Dans notre article [$L12$] nous prouvons cette inégalité pour les demi-droites $]a;+\infty[$ et $]-\infty;a[$. Notons qu'on ne peut pas appliquer directement la méthode des EDP de [28] qui nécessite que τ soit borné. Nous nous servons de l'analyse spectrale, utilisable dans ce contexte car le générateur est automatiquement symétrique en dimension 1. Notons aussi que nous ne pouvons pas établir ce genre d'inégalités directement sur $\mathbb{R}$, car notre processus est conservatif et le temps de sortie de $\mathbb{R}$ est identiquement infini. Mais en utilisant la borne sur la meilleure constante de Poncaré c_P ci-dessus, nous donnons une relation entre le trou spectral global $\gamma = 1/c_P$ et $\lambda_a^\pm$

$$\sup_a \left(\lambda_a^+ \wedge \lambda_a^-\right) \leq \gamma \leq \inf_a \left(\lambda_a^+ \vee \lambda_a^-\right).$$

10 Projets de recherche.

Nous essayons ci-dessous de dégager quelques pistes de recherche.

10.1 Inégalité du type "déviations modérées " pour un processus de Harris récurrent ([$L13$]).

Soit $(X)_t; t \geq 0$, un processus fortement markovien, récurrent, à valeurs dans un espace polonais, de mesure invariante μ. En utilisant la méthode du splitting en temps continu, avec Eva Löcherbach nous obtenons dans [$L13$] des inégalités de déviation pour des fonctionnelles centrées de X : soit f une fonction bornée, à support compact, telle que $\mu(f) = 0$. Nous obtenons l'inégalité suivante : pour tout $\eta \in]0, \frac{1}{2}[$,

$$P_\pi \left(| \int_0^t f(X_s)ds| \geq v(t)^{\frac{1}{2}+\eta} x \right) \leq C_1 \exp\left(-C_2 v(t)^\eta \left(x^2 \wedge x\right)\right) + R_t.$$

Ici, R_t est un terme de reste qui est de l'ordre de $O(\exp\left(-\sqrt{v(t)}\right))$, où v_t est l'équivalent déterministe de X et les constantes C_1 et C_2 s'expriment en termes du processus et de sa mesure invariante. Mis à part le cadre non-asymptotique, l'intérêt principal de cette inégalité est que le résultat est valable dans les deux cas : récurrent positif ou récurrent nul. En utilisant cette inégalité comme le point de départ, on voudrait obtenir le principe de déviations modérées pour le processus en temps continu, question déjà bien étudié dans le cas des chaînes de Markov (voir Guillin [34], Djellout et Guillin [20]).

10.2 Splitting en temps continu.

Tout en continuant à développer des applications de la méthode de splitting en temps continu, on voudrait écrire un ouvrage pédagogique, illustré par des applications classiques qu'on pourrait trouver à cette méthode, comme par exemple la démonstration à l'aide de la décomposition en cycles du théorème ergodique pour les fonctionnelles additives des diffusions multidimensionnelles.

10.3 Trou spectral et moments des temps d'atteinte.

L'équivalence obtenue dans [$L12$] est valable pour une diffusion à valeur dans $\mathbb{R}$. Or l'implication trou spectral $\Longrightarrow$ l'existence des moments exponentielles des temps d'atteinte des ensembles de mesure invariante positive est vrai pour un processus de Markov beaucoup plus général (voir Carmona et Klein [8]). On voudrait donc éteindre cette équivalence.

10.4 Statistiques pour les diffusions multidimensionnelles.

On voudrait éteindre nos inégalités de déviation au cas multidimensionnel. Ceci pourrait être exploité en statistiques, notamment dans l'estimation adaptative du drift, où, à ma connaissance, il n'existe pratiquement pas de résultats en dimmension supérieur à 1.

10.5 Alignement des sequences.

Avec Antoine Chambaz (Paris 5), Catherine Mattias (Evry, laboratoire de Statistiques et génome) et Oleg Loukianov (Paris 12) nous avons entamé un travail sur l'alignement de séquences. Suite à la difficultés du sujet, mais aussi des nombreuses contraintes professionnelles et familiales nous avons abandonné ce travail à mi-chemin, mais c'est un domaine qui m'intéresse toujours et j'espère pouvoir contribuer dans son développement.

Le problème principal dans l'étude mathématique des séquences génomiques est d'établir la distribution du score (local et global) d'alignement. La réponse partielle à cette question a été donnée dans l'article pionnier de Dembo, Karlin et Zeitouni [21] (1994) concernant l'alignement de deux séquences de lettres i.i.d. telle que les délétions (c'est à dire alignement d'une lettre avec un vide) ne sont pas autorisés (ungapped case). Notons que dans ce cas le score global est simplement une somme de v.a. i.i.d. Dans le cas "avec gaps" le score global est un objet beaucoup plus compliqué mais cependant il a une structure super additive. En utilisant cette propriété Grossmann et Yakir en 2004[33] ont généralisé certains résultats de (DKZ 1994) au cas "avec gaps", mais toujours i.i.d. Le résultat principal de Grossmann et Yakir est une majoration pour la probabilité de déviation du maximum sur n du score global. Avec Antoine Chambaz, Catherine Mattias et Oleg Loukianov nous avons amélioré les résultats de Grossmann et Yakir sur deux plans : premièrement nous avons obtenu une estimation plus fine de la probabilité de déviation du maximum sur n du score global, deuxièmement nos résultats s'appliquent à des séquences qui sont des chaînes de Markov. On voudrais maintenant démontrer une minoration du même type pour cette probabilité.

11 Références des travaux présentés.

11.1 Travaux acceptés ou publiés.

[L1] Loukianova, D. Capacité de mémoire dans le modèle de Hopfield. CRAS **t.318**, Série 1, 157-160 (1994)

[L2] Loukianova, D. Lower bound on the restitution error in the Hopfield Model. Probab. Theory Rel. Fields. 107, 161-176 (1997).

[L3] Loukianova, D. Remark on semigroup technics and the MLE estimations, Statist. Probab. Letters **62**, 111-115 (2003).

[L4] Loukianova,D. ; Loukianov,O. Uniforme law of large numbers and consistency of estimators for Harris diffusions, Statist. Probab. Letters **74**, 347-355 (2005).

[L5] Loukianova,D. ; Loukianov,O. Almost sure rate of convergence of MLE for multidimensional diffusions, Dependence in Probability and Statistics, Lecture Notes in Statistics , **187** , 329-349 (2006)

[L6] Loukianova,D. ; Loukianov,O. Deterministic equivalents of additive functionals of recurrent diffusions and drift estimation, Stat. Inference Stoch. Process 11, Issue 2, 107-121 (2008).

[L7] Loukianova,D. ; Loukianov,O. Uniform deterministic equivalent of of recurrent diffusions and non-parametric drift estimation, Ann. I.H.Poincaré PS , Vol. 44, No4, 771-186 (2008).

[L8] Löcherbach, E., Loukianova, D. On Nummelin Splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multidimensional diffusions, Stoch. Proc. Appl., 118, 1301-1321, (2008).

[L9] Löcherbach, E., Loukianova, D. The law of iterated logarithm for additive functionals and martingale additive functionals of Harris recurrent Markov processes, Stoch. Proc. Appl., vol. 119, no7, 2312-2335 (2009).

[L10] Löcherbach, E., Loukianova, D, Loukianov,O. Penalized nonparametric drift estimation for ergodic diffusion, accepté pour la publication à ESAIM PS.

[L11] Löcherbach, E., Loukianova, D, Loukianov,O. Polynomial bounds in the Ergodic theorem for poitive recurrent one-dimensional diffusions and integrability of hitting times, accepté pour la publication à Ann. I.H.Poincaré PS.

[L12] Loukianova, D, Loukianov,O. Song, S. Poincaré inequality and exponential integrability of hitting times for a linear diffusion, accepté pour la publication à Ann. I.H.Poincaré PS.

11.2 Travaux en cours ou soumis.

Löcherbach, E., Loukianova, D.(2009) Moderate deviations for centered additive functionals of recurrent Harris processes.

Chambaz, A., Mattias,C., Loukianov, O., Loukianova, D., Strong large deviations for Markov superadditive processes and applications to sequence alignments.

Löcherbach, E., Loukianova, D., Polynomial regeneration time moments for recurrent Harris processes.

References

[1] Amit, D.J., Gutfreund, H., Sompolinsky, H. Spin glass models of neural networks. Phys. Rev. A32, 1007-101 (1985)

[2] Athreya, K.B., Ney, P., A new approach to the limit theory of recurrent Markov chains, Trans.Amer.Math.Soc.245 (1978) 493-501

[3] Balaji, S., Ramasubramanian, S., Passage time Moments for Multidimensional Diffusions, Journal Of Applied Probability 37, No. 1, 246–251, 2000.

[4] Basawa, I.V., Prakasa Rao B.L.S (1980) Statistical inference for stochastic processes. Academic Press, London.

[5] Bobkov, S., Göetze, F., Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, J. Funct. Anal. 163, 1 (1999), 1–28.

[6] Borodin, A., Salminen, P., Handbook of Brownian Motion-Facts and Formulae, Second Edition, Birkhäuser Verlag, 2002.

[7] Brancovan, M., Fonctionnelles additives spéciales des processus récurrents au sens de Harris. Wahrsch. Verw.Gebiete 47, (1979) 163-194.

[8] Carmona, R., Klein, A., Exponential moments for hitting times of uniformly ergodic markov processes, The Annals of Probability, 1983, Vol. 11, No. 3, 648-665.

[9] Chen, X. How often does a Harris Markov chain recur ?, Ann. Prob. 3, 1324-1346, (1999).

[10] Chen, X. The law of the iterated logarithm for functionals of Harris recurrent Markov chains, : Self normalization. J. of Theoretical Probab. 12, No2 (1999), 421-445.

[11] Chen, X. On the limit law of the second orde for additive functionals of Harris recurrent Markov chains. Probab. Theory Relat.Fields 134, N° 2(2006), 248-282.

[12] Cattiaux, P., Guillin, A. , Deviation bounds for additive functionals of Markov process, ESAIM PS, Vol 12, p 12-29,2008

[13] Comte, F., Genon-Catalot, V., Rozenholc, Y., Penalized nonparametric mean square estimation of the coefficients of diffusion processes. Bernoulli 13, No. 2, 514–543, 2007.

[14] Csáki,E.., Földes, A., Hu,Y. ; Strong approximations of additive functionnals of a planar Brownian motion. Stochastic Processes Appl. 109, No.2 (2004), 263-293.

[15] Darling, D.A., Kac, M. On occupation times for Markov processes, Trans.Amer. Math.Soc.84, 444-458 (1957)

[16] Darling, D., A., Siegert, A., J., F., The first passage problem for a continuous markov process,

[17] Dalalyan, A. Sharp adaptive estimation of the drift function for ergodic diffusion, Ann Statist. 33(2005) 2507-2528

[18] Dalalyan, A. Kutoyants, Y. On second order minimax estimation of the invariant density for ergodic diffusions. Statist. Decisions 22 (2004) 17-43

[19] Deaconu, M., Wantz, S., Comportement des temps d'atteinte d'une diffusion fortement rentrante, Seminaire de probabilités (Strasbourg), Tome 31(1997), p.168-175.

[20] Djellout, A. Guillin, A. Moderate deviations of Markov Chains with atom. Stochastic Processes and their Applications, (2001), 237-251.

[21] Dembo, A. Karlin, T., Zeitouni, O., Critical Phenomena for sequence matching with scoring, Annals of Probability, vol 22, $n°4$

[22] Delattre, S. , Hoffmann, M. Asymptotic equivalence for a null recurrent diffusion model. Bernoulli 8 (2002) 139–174.

[23] Delattre,S., Hoffmann, M., Kessler, M.. Dynamics adaptive estimation of a scalar diffusion. Prépublication PMA-762,(2001) Univ. Paris 6. Available at www.proba.jussieu.fr/mathdoc/preprints/.

[24] Ditlevsen, S. A result on the first-passage time of an Ornstein-Uhlenbeck process, Statistics and Probability Letters, 77 : 1744-1749, 2007.

[25] Douc, R., Guillin, A., Moulines, E., Bounds on Regeneration times and limit theorems for Subgeometric Markov Chains, Ann. Inst. H. Poincaré Prob Statist. 44, No. 2, 239-257, 2008.

[26] Fegin, P.D. (1976) Maximum likelihood estimation for continuous-time stochastic processes. Adv. Appl. Prob. 8, 712-736

[27] Fitzsimmons,P.J., Pitman, J. Kac's moment formula and the Feyman-Kac formula for additive functionals of a Markov process. Stoch. Proc. Appl. 79, (1999), 117-134.

[28] Friedman, A., The asymptotic behavior of the first real eigenvalue of a second Order Elliptic Operator with a small parameter in the highest derivatives, Indiana University Mathematics Journal, Vol. 22, No. 10 (1973)

[29] Fukushima, M., Oshima, Y., Takeda, M., *Dirichlet Forms and Symmetric Markov Processes*, de Gruyter Studies in Mathematics 19 (1994)

[30] Galtchouk, L., Pergamenshchikov, S., Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. Part 1 : Sharp non-asymptotic oracle inequalities. hal-00177875.

[31] Genon-Catalot, V., Comte, F., Rozenholc, Y. Penalized nonparametric mean square estimation of the coefficients of diffusion processes. Bernoulli 13, No 2, 514-543 (2007)

[32] Giorno, V., Nobile, A.G., Riccardi, L., Sacredote, L., Some remarks on the Raleigh process, J. Appl., Prob. 23, 398-408 (1986).

[33] Grossman, S. and Yakir, B. Large deviations for global maxima of independent su- peradditive processes with negative drift and an application to optimal sequence alignment. (2004) Bernoulli 10, 829-845.

[34] Guillin, A. Uniform moderate deviations of functional empirical processes of Markov chains. Probability and Mathematical Statistics, 20 (2), (2001), 237-260.

[35] Höpfner, R., Kutoyants Yu. On a problem of statistical inference in null recurrent diffusions. Stat.Inference Stoch. Process 6(1) :25-42

[36] Höpfner, R. Löcherbach E. Limit Theorems for null recurrent Markov processes, vol 768.Memoirs of the AMS. Providence, Rhode Island

[37] Kac, M., On the distribution of certain Wiener functionals. Trans. Amer. Math. Soc. 65, 1-13., 1949.

[38] Kac, M., On some connections between probability theory and differential and integral equations. In : Neyman, J. (Ed.) , Proc. 2nd Berkley Symp. Math. Stat. Prob., Univ. of California Press, Berkley, CA, pp. 189-215, 1951.

[39] Kac, I. S. ; Krein, M. G. Criteria for the discreteness of the spectrum of a singular string. (Russian) Izv. Vysˇs. Uˇcebn. Zaved. Matematika 1958 no. 2 (3), 136–153.

[40] Khasminskii, R., Z., On positive solutions of the equation $Ru+Vu=0$, Theor.Probability Appl. 4 (1959), 309-318

[41] Komlos, J., Paturi, R. Convergence results in an autoassociative memory model, Neural Networks 1, 239-250 (1988)

[42] Kotani, S. ; Watanabe, S. Krein's spectral theory of strings and generalized diffusion processes. Functional analysis in Markov processes (Katata/Kyoto, 1981), pp. 235–259, Lecture Notes in Math., 923, Springer, Berlin-New York, 1982.

[43] Kutoyants, Y.A. Statistical inference for Ergodic Diffusion Processes, Springer, Berlin, New York.

[44] Kutoyants, Y.A., Efficient density estimation for ergodic diffusion processes. Stat.Inference Stoch.Process.1, No.2(1998), 131-155.

[45] Lanska, V. Minimum contrast estimation in diffusion process J.Appl. Probab. 16, 65-75 (1979)

[46] Lezaud, P, Chernoff and Berry-Esséen inequalities for Markov processes, ESAIM P & S 5, 183–201, 2001.

[47] MacEliece, R.J., Posner, E.C., Rodemich, E., Venkatesh, S.S., The capacity of the Hopfield associative memory, IEEE Trans.Inform.Theory 33, 461-482 (1987)

[48] Malrieu, F., Roberto, C. Les inégalités de Sobolev logarithmiques et le trou spectral sur la droite reelle, *Sur les inégalités de Sobolev logarithmiques*, collection Panoramas et Synthèèses de la SMF

[49] Miclo, L, Quand est-ce que des bornes de Hardy permettent de calculer une constante de Poincar exacte sur la droite ? Ann. Fac. Sci. Toulouse Math. (6) 17 (2008), no. 1, 121–192

[50] Muckenhoupt, B. , Hardy's inequality with weights. Studia Math. XLIV (1972), 31-38.

[51] J. Neveu, Potentiel markovien récurrent des chaînes de Harris. Ann. Inst. Fourier 22, 85-130 (1972)

[52] Newman , Ch.M. Memory capacity and neural network models : Rigourous lower bound. Neutral Networks 1 : 223-238 (1988)

[53] Nishiyama,Y. A maximal inequality for continuous martingales and M-estimation in gaussian white noise model, The Annals of Statistics 27 (2), 675-696 (1999)

[54] Nummelin, E. A splitting technique for Harris recurrent Markov Chains, Z. Warscheinlichkeitstheor, Verw.Geb. 43 (1978) 309-318.

[55] Prakasa Rao BLS (1990) Statistical Inference for diffusion type proceses. Arnold London.

[56] Revuz, D. Markov chains, Revised ed. North Holland, Amsterdam, 1984.

[57] Senoussi, R., Uniform iterated logarithm laws for martingales and their application to functional estimation in controlled Markov chains, Stochastic Process.Appl. 89, 193-211 (2000)

[58] Talagrand, M. Rigourous results for the Hopfield model with many patterns, PTRF 110, n˚2, 177-276 (1998)

[59] Touati, A. Théorèmes limites pour les processus de Markov récurrent, C.R.Acad. Sci.Paris Sér.I Math.305 841-844.

[60] S.A. van de Geer, Empirical processes in M-estimation, Cambridg University Press, Cambridge (2000).

[61] S.A. van de Geer, L. Stougie, On rates of convergence and asymptotic normality in the multiknapsack problem, Mathematical Programming 51, 349-358 (1991).

[62] van der Vaart, A., van Zanten, H. Donsker theorems for diffusions : necessary and sufficient conditions, The Annals of Probability , Vol.33, N°4, 1422-1451, (2005)

[63] van der Vaart, A. Welhner, J.A. Weak Convergence and Empirical Processes, Springer-Verlag (1996)

[64] van Zanten J.H., 2001. A note on consistent estimation of multivariate parameters in ergodic diffusion models. Scand. J. Statist. 28 (4), 617-623.

[65] van Zanten J. H., 2003. On uniform laws of large numbers for ergodic diffusions and consistency of estimators. Statistic Inference for Stochastic Process. 6 (2), 199-213.

[66] van Zanten, H. On the rate of convergence of the maximal likelihood estimator in Brownian semimartingale models, Bernoulli 11 (4), 643-664 (2005).

[67] Veretennikov, A.Yu., On polynomial mixing bounds for stochastic differential equations, Stoch. Proc. Appl. 70, 115–127, 1997.

[68] Yoshida, N. Asymptotic behavior of M-estimators and related random field for diffusion process. Ann Inst Stat Math 42 (2) 221-251.