Applications statistiques de suites faiblement dépendantes et de systèmes dynamiques

Clémentine Prieur

Doctorat de Mathématiques de
l'Université de Cergy-Pontoise

APPLICATIONS STATISTIQUES DE SUITES FAIBLEMENT DEPENDANTES ET DE SYSTEMES DYNAMIQUES

par

Clémentine PRIEUR

Thèse soutenue le 13 décembre 2001 devant le jury composé de:

Pr. Michel BENAIM
Pr. Patrice BERTAIL
Pr. Jean DIEBOLT (Rapporteur)
Pr. Paul DOUKHAN (Directeur de thèse)
Pr. Guy FAYOLLE
Pr. Gérard GREGOIRE
Pr. Emmanuel RIO (Rapporteur)
Pr. Denis TALAY

Remerciements

L'heure est enfin venue aux remerciements! Je tiens d'abord à exprimer à Paul Doukhan ma profonde reconnaissance pour avoir dirigé mes travaux. Il a réussi à me communiquer son goût pour la recherche, je l'en remercie chaleureusement. Je lui suis également reconnaissante pour sa disponibilité et ses idées toujours judicieuses. Je tiens encore à préciser que son groupe de travail a été pour moi un lieu d'échange et d'enrichissement.

Je suis très touchée de l'honneur que m'ont fait Jean Diebolt et Emmanuel Rio en acceptant de rapporter cette thèse. Leurs questions et leurs remarques m'ont été très utiles.

Je remercie vivement Patrice Bertail, Michel Benaïm, Gérard Grégoire, Guy Fayolle et Denis Talay d'avoir accepté de faire partie du jury.
C'est à la suite d'un exposé de Gérard Grégoire à l'université de Cergy-Pontoise que je me suis intéressée aux problèmes de rupture. Je l'en remercie sincèrement.
Merci à Michel Benaïm pour les références bibliographiques qu'il a toujours accepté de me donner.
Merci à Denis Talay pour ses nombreux conseils.

Je remercie les probabilistes et les statisticiens de l'IRMAR qui m'ont accueillie au sein de leur groupe de travail durant l'année que j'ai passée à Rennes.

Merci à Patrick Ango Nze qui a accepté de travailler avec moi. Grâce à lui j'ai appris ce qu'était le travail en équipe.
Je n'oublie pas Sana Louhichi. Elle a toujours été à mon écoute et ses conseils m'ont été extrêmement utiles.

Un grand merci également à Viviane Baladi, Véronique Maume-Deschamps et Hans Rugh qui m'ont, à plusieurs reprises, donné de précieux conseils. Ils m'ont beaucoup appris sur les systèmes dynamiques et ont fait preuve à mon égard d'une très grande patience. Je tiens à leur exprimer ici ma sympathie et ma reconnaissance.
Je remercie Jonathan de Halleux qui, avec toute sa jovialité, a accompagné mes premiers pas en C.
Merci encore à Odile Brandière, Anne Broise, Jérôme Dedecker, ainsi qu'à tous ceux qui m'ont consacré de leur temps.

Je tiens à exprimer ma gratitude envers l'équipe de Cergy qui m'a accueillie pendant la durée de cette thèse. Merci au bureau 572, réputé pour sa convivialité et la saveur de ses thés. Merci à Christel, Mélissa, Christophe, Frédéric, Sylvain, ... Je n'oublie pas Annette Rasmussen qui a su, par sa disponibilité et sa gentillesse, alléger mes démarches administratives.

Merci à ma famille.
Merci aux copains de Cachan!

Enfin, je tiens à remercier celui qui, tant sur le plan affectif que sur le plan scientifique, a été mon soutien quotidien. Merci Christophe.

A mon mari, Christophe

Résumé : Cette thèse porte sur l'étude d'applications statistiques de suites dépendantes et stationnaires. Nous étudions deux classes de suites dépendantes. Nous nous intéressons d'une part à des suites faiblement dépendantes, où notre notion de dépendance faible est une variante de la notion introduite par Doukhan & Louhichi [21], d'autre part à certains systèmes dynamiques présentant une propriété de décroissance des corrélations. Nous traitons du comportement asymptotique du processus empirique, fondamental en statistiques. Nous étudions aussi un estimateur à noyau de la densité dans nos deux cadres de dépendance. Enfin, nous nous intéressons à un problème de rupture d'une fonction de régression en dépendance faible. A ces fins, nous développons des idées de Rio [53] pour montrer un théorème limite centrale en dépendance faible, ainsi que des nouvelles inégalités de moments qui étendent celles de Louhichi [40]. Enfin, nous illustrons certains de nos résultats par des simulations.

AMSC 2000 : 37 A 50, 37 D 20, 37 M 10, 60 E 15, 60 F 05, 60 F 15, 60 F 17, 60 G 07, 60 G 10, 60 G 99, 62 G 08, 62 J 05.

Mots clés : Dépendance faible, suites stationnaires, systèmes dynamiques, décroissance des corrélations, théorème limite centrale, théorème de Lindeberg, densité invariante, estimation non-paramétrique, estimation à noyau, inégalités de moments, processus empirique, loi forte des grands nombres de Marcinkiewicz-Zygmund, rupture, fonction de régression.

Abstract: This thesis deals with the study of some statical applications of dependent and stationary sequences. We study two classes of dependent sequences : weak dependent sequences in a sense which is a variation on the notion introduced by Doukhan & Louhichi [21], and some dynamical systems whose correlations decrease. We study the asymptotical behaviour of the empirical process, which is very important in statistics. We also study standard kernel density estimates in both frames. We finally investigate a change point estimator of some regression function in our frame of weak dependence. The main tool to study these applications is a variation on Rio's ideas [53] to prove a Central Limit Theorem for weakly dependent sequences as far as some new moment inequalities which extend Louhichi's inequalities [40]. We also make some numerical simulations concerning some of our results.

AMSC 2000: 37 A 50, 37 D 20, 37 M 10, 60 E 15, 60 F 05, 60 F 15, 60 F 17, 60 G 07, 60 G 10, 60 G 99, 62 G 08, 62 J 05.

Key words: Weak dependence, stationary sequences, dynamical systems, decay of correlations, central limit Theorem, Lindeberg Theorem, invariant density, non-parametric estimation, kernel estimates, moment inequalities, empirical process, Marcinkiewicz-Zygmund strong law for large numbers, change point estimation, regression function.

Table des matières

Travaux de l'auteur

[3] P. Ango Nze & C. Prieur, Estimation d'une rupture en dépendance faible. Note soumise aux *C.R.A.S. Paris* Série **1** (2001).

[16] C. Coulon-Prieur & P. Doukhan, A triangular central limit theorem under a new weak dependence condition. *Statistic & Probability Letters*, **47** (2000), 61-68.

[48] C. Prieur, Density Estimation of Dynamical Systems, dans *Proceedings of the* 21^{st} *International Seminar On Stability Problems for Stochastic Models*, Eger (Hungary) 2001. To appear in Kluwer-Plenum, *Journal of Math. Sciences.*

[49] C. Prieur, Estimation de la densité invariante de systèmes dynamiques en dimension 1. *C.R.A.S. Paris* Série **1**, **332** (2001), 761-764.

[50] C. Prieur, Density Estimation For One-Dimensional Dynamical Systems. *ESAIM, Probab. & Statist.* `www.emath.fr/ps` **5** (2001), 51-76.

[51] C. Prieur, *An Empirical Functional Central Limit Theorem For Weakly Dependent Sequences*, Préprint de l'Université de Cergy-Pontoise (octobre 2001). Article soumis.

Bibliographie

[1] A. Amroun, A., *Systèmes dynamiques perturbés. Sur une classe de fonctions zéta dynamiques,* Thèse de Doctorat de l'Université Paris 6, Spécialité Mathématique (1995).

[2] P. Ango Nze and P. Doukhan, Non-parametric Minimax estimation in a weakly dependent framework I: Quadratic properties. *Math. Meth. Statist.* **5** no. 4 (1996), 404-423.

[3] P. Ango Nze and C. Prieur, Estimation d'une rupture en dépendance faible. Note soumise aux *C.R.A.S. Paris* Série **1** (2001).

[4] V. Baladi, M. Benedicks and V. Maume-Deschamps, Almost sure rates of mixing for i.i.d. unimodal maps. Préprint (1999), à paraître dans *Ann. E.N.S.*

[5] A.D. Barbour, R.M. Gerrard and G. Reinert, Iterates of expanding maps. *Probab. Theory Related Fields* **116** no. 2 (2000), 151-180.

[6] L.E. Baum and M. Katz, Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120** (1965), 108-123.

[7] H. Berbee, Convergence Rates in the Strong Law of Bounded Mixing Sequences, *Probab. Th. Rel. Fields* **74** (1987), 255-270.

[8] D. Bosq and D. Guégan, Nonparametric estimation of the chaotic function and the invariant measure of a dynamical system. *Statistics & Probability Letters* **25** (1995), 201-212.

[9] D. Bosq and J.P. Lecoutre, *Théorie de l'estimation fonctionnelle.* Collection "Economie et statistiques avancées". Série: Ecole Nationale de la Statistique et de l'Administration Economique et Centre d'Etudes des Programmes Economiques". Economica (1987).

[10] J. Bretagnolle and C. Huber, Estimation des densités : risque minimax. *Z. Wahrsch. verw. Gebiete* **47** (1979), 119-137.

[11] A. Broise, F. Dal'bo and M. Peigné, Etudes spectrales d'opérateurs de transfert et applications. *Astérisque* **238** (1996). Société math. de France.

[12] K.C. Chanda and F.H. Ruymgart, General linear processes: a property of the empirical process applied to density and mode estimation. *J. Time Ser. Anal.* **11** no. 3 (1990), 185-199.

[13] Y.S. Chow and T.L. Lai, Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings. *Trans. Amer. Math. Soc.* **208** (1975), 51-72.

[14] P. Collet, Some ergodic properties of maps of the interval, dans *dynamical systems*, coécrit par R. Bamon, J.M. Gambaudo et S. Martinez. Hermann, Paris (1996).

[15] V. Couallier, *Inférence statistique pour des estimateurs de discontinuités dans un cadre non paramétrique,* Thèse de l'Université Paul Sabatier, Toulouse 3 (2000).

[16] C. Coulon-Prieur and P. Doukhan, A triangular central limit Theorem under a new weak dependence condition. *Statistics & Probability Letters* **47** (2000), 61-68.

[17] W. De Melo and S. Van Strien, *One-Dimensional Dynamics.* Springer-Verlag (1993).

[18] P. Doukhan, *Mixing: Properties and Examples.* Springer Verlag, *Lecture Notes in Statistics* **85** (1994).

[19] P. Doukhan, *Models, Inequalities and Limit Theorems for Stationary Sequences,* coécrit par P. Doukhan, G. Oppenheim and M. Taqqu. Birkhaüser (à paraître).

[20] P. Doukhan and G. Lang, *Rates in the empirical central limit theorem for stationary weakly dependent random fields*, Préprint no. 15/00 de l'université de Cergy-Pontoise (mai 2000).

[21] P. Doukhan and S. Louhichi, A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications* **84** (1999), 313-342.

[22] P. Doukhan and S. Louhichi, Functional estimation of a density under a new weak dependence condition. *Scandinavian Journal of Statistics* **28** no. 2 (2001), 325-342.

[23] P. Doukhan and F. Portal, Moments de variables aléatoires mélangeantes. *C.R.A.S. Paris* Série **1**, **297** (1983), 129-132.

[24] J. Esary, F. Proschan and D. Walkup, Association of random variables with applications. *Ann. Math. Statist.* **38** (1967), 1466-1476.

[25] J. Fan and I. Gibjels, Variable bandwidth and local linear regression smoothers. *Ann. Stat.* **20** (1992), 2008-2036.

[26] L. Giraitis, P. Kokoszka and R. Leipus, Stationary ARCH models: dependence structure and Central Limit Theorem. *Econometric Theory* **16** (01) (2001), 3-22.

[27] G. Grégoire and Z. Hamrouni, Change Point Estimation by Local Linear Smoothing. To appear in *Journal of Multivariate Analysis* (2001).

[28] Z. Hamrouni, *Inférence statistique par lissage linéaire local pour une fonction de régression présentant des discontinuités,* Thèse de l'Université Joseph Fourier, Grenoble 1 (1999).

[29] I.A. Ibragimov, Some limit theorems for stationary processes. *Theory Probab. Appl.* **7**, 349-382.

[30] B. Isha and R. Prakasa, Kernel-type density and failure rate estimation for associated sequences. *Ann. Inst. Statist. Math.* **47** (1995), 253-266.

[31] A.N. Kolmogorov and Y.A. Rozanov, On the strong mixing conditions for stationary Gaussian sequences. *Th. Probab. Appl.* **5**, 204-207.

[32] T.L. Lai, Convergence rates and r-quick versions of the strong law for stationary mixing sequences. *Ann. Probab.* **5** no. 5 (1977), 693-706.

[33] A. Lasota and M. Mackey, *Probabilistic properties of deterministic systems.* Cambridge University Press (1985).

[34] A. Lasota and J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186** (1973), 481-488.

[35] J. W. Lindeberg, Eine neue Herleitung des Exponentialgesetzes. *Wahr. Math. Zeitschrift* **15** (1922), 211-225.

[36] C. Liverani, Decay of correlations for piecewise expanding maps. *J. Statist. Phy.* **78** no. 3-4 (1995), 1111-1129.

[37] C. Liverani, *Central limit Theorem for deterministic systems* dans *Proceedings of the international Congress on Dynamical Systems,* Montevideo 95. Pittman, *Research Notes in Mathematics series* (1997).

[38] S. Louhichi, *Théorèmes limites pour des suites positivement ou faiblement dépendantes*, Thèse de l'Université Paris Sud (1998).

[39] S. Louhichi, Rates of convergence in the CLT for some weakly dependent random variables. To appear in *Theory Probab. Appl.* (1999).

[40] S. Louhichi, *Independence Via Uncorrelatedness, Examples and Moment inequalities*, Préprint de l'Université Paris 11 (2000).

[41] J.A. Mac Donald and A.B. Owen, Smoothing with split linear fits, *Technometrics* **28** no. 3 (1986), 195-208.

[42] J. Maës, *Statistique non paramétrique des processus dynamiques réels en temps discret.* Thèse de l'Université Paris 6 (1999).

[43] F.A. Moricz, R.J. Serfling and W.F. Stout, Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *Ann. Probab.* Vol. **10** no.4 (1982), 1032-1040.

[44] M. Peligrad and S. Utev, Central Limit Theorem for Linear Processes. *Ann. Probab.* **25** no. 1 (1997), 443-456.

[45] V.V. Petrov, *Limit theorems of probability theory : sequences of independent random variables.* Clarendon Press, Oxford (1995).

[46] D. Pollard, *Convergence of Stochastic Processes.* Springer Verlag, *Springer Series in Statistics* (1984).

[47] R. Prakasa, *Nonparametric functional estimation.* Academic Press, New York (1983).

[48] C. Prieur, Density Estimation of Dynamical Systems, dans *Proceedings of the* 21^{st} *International Seminar On Stability Problems for Stochastic Models*, Eger (Hungary) 2001. To appear in Kluwer-Plenum, *Journal of Math. Sciences.*

[49] C. Prieur, Estimation de la densité invariante de systèmes dynamiques en dimension 1. *C.R.A.S. Paris* Série **1**, **332** (2001), 761-764.

[50] C. Prieur, Density Estimation For One-Dimensional Dynamical Systems. *ESAIM, Probab. & Statist.* `www.emath.fr/ps` **5** (2001), 51-76.

[51] C. Prieur, *An Empirical Functional Central Limit Theorem For Weakly Dependent Sequences*, Préprint no. 18/01 de l'Université de Cergy-Pontoise (octobre 2001). Article soumis.

[52] E. Rio, Inégalités de moments pour les suites stationnaires et fortement mélangeantes. *C.R.A.S. Paris* Série **1**, **318** (1994), 355-360.

[53] E. Rio, About the Lindeberg method for strongly mixing sequences. *ESAIM, Probab. & Statist.* `www.emath.fr/ps` **1** (1995), 35-61.

[54] E. Rio, Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes. *Probab. Theory Relat. Fields* **104** (1996), 255-282.

[55] E. Rio, *Théorèmes Limites pour des suites faiblement dépendantes.* Springer, *SMAI Mathématiques et Applications* **31** (2000).

[56] P.M. Robinson, Non parametric estimators for time series. *J. Time Ser. Anal.* **4** no. 3 (1983), 185-207.

[57] M. Rosenblatt, A central limit Theorem and a strong mixing condition. *Proc. Nat. Ac. Sc. U.S.A.* **42** (1956), 43-47.

[58] M. Rosenblatt, Stochastic curve estimation, in *NSF-CBMS Regional Conference Series in Probability and Statistics*, **3** (1991).

[59] H.P. Rosenthal, On the subspaces of L^p $(p > 2)$, spanned by sequences of independent random variables. *Israël J. Math.* **8** (1970), 273-303.

[60] Y.A. Rozanov and V.A. Volkonski, Some limit theorems for random functions. Part. 1 *Theory Probab. Appl.* **4** (1959), 178-197. Part. 2 *Theory Probab. Appl.* **6** (1961), 186-197.

[61] W. Rudin, *Real and complex analysis.* McGraw-Hill Series in Higher Mathematics, Second Edition (1974).

[62] Q.M. Shao, Maximal inequalities for partial sums of $\rho-$mixing sequences. *Ann. Probab.* **23** no. 2 (1995), 948-965.

[63] Q.M. Shao and H. Yu, Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* **24** no. 4 (1996), 2098-2127.

[64] W.F. Stout, *Almost sure convergence.* Academis, New York (1974).

[65] M. Viana, *Stochastic dynamics of deterministic systems, Instituto de Matematica Pura e Aplicada.* IMPA, **21** (1997).

[66] C.S. Withers, Central limit Theorems for dependent variables. *Z. Wahrsch. Verw. Gebiete* **57** (1981), 509-534 (Corrigendum (1983), **63**, 555).

Introduction

Cette thèse porte sur l'étude d'applications statistiques de suites aléatoires $(X_n)_{n\in\mathbb{Z}}$ stationnaires et dépendantes. Nous étudions deux classes de suites dépendantes : des suites faiblement dépendantes au sens de Coulon-Prieur & Doukhan [16] (variante de la notion de dépendance introduite par Doukhan & Louhichi [21]) d'une part, et certains systèmes dynamiques ergodiques d'autre part.

Nous traitons du comportement asymptotique du processus empirique, processus fondamental en statistiques. Nous précisons aussi le comportement asymptotique d'un estimateur à noyau de la densité dans nos deux cas de dépendance. Enfin, nous étudions un problème de rupture d'une fonction de régression, également à l'aide d'un estimateur à fenêtre.

A ces fins, nous développons des idées de Rio [53] conduisant au théorème limite centrale pour nos suites dépendantes, ainsi qu'à des inégalités de moments étendant celles de Louhichi [40]. Ces inégalités de moments permettent de prouver la tension de la suite des fonctions de répartition empirique recentrées et renormalisées.

Dans cette introduction, nous décrivons en détails les travaux présentés dans cette thèse.

DEPENDANCE FAIBLE

Depuis quelques années déjà, l'étude de la dépendance de suites de variables aléatoires s'est beaucoup développée, car l'indépendance reflète souvent mal la réalité (par exemple en ferromagnétisme, le modèle d'Ising fait intervenir des suites de variables aléatoires associées). En statistiques, on sépare principalement la dépendance en deux catégories : la dépendance faible et la dépendance forte. Rappelons ici que l'un des principaux objectifs des statistiques est l'obtention de tests consistants, et qu'en termes probabilistes, un test est construit à l'aide d'un théorème limite. Partant de cette idée, nous pouvons distinguer les deux types de dépendance, faible et forte. En effet, en dépendance faible, les théorèmes limite associés aux statistiques de Donsker et de Kolmogorov se produisent avec une normalisation en $\sqrt{n}$ et la limite est un processus plutôt irrégulier, tandis qu'en dépendance forte, le processus limite est très régulier et on rencontre diverses normalisations.

Nous allons considérer ici uniquement le cas de la dépendance faible. Il existe de nombreuses notions de dépendance faible. Rappelons ici la notion de mélange qui s'exprime en termes de coefficients de mélange entre des tribus engendrées par le passé et le futur de la suite $(X_n)_{n\in\mathbb{Z}}$: pour $p \in \mathbb{Z}$, $\mathcal{F}_{-\infty}^{p} = \sigma(X_i,\ i \le p)$ et $\mathcal{F}_{p}^{+\infty} = \sigma(X_i,\ i \ge p)$. Si $(\Omega,\mathcal{A},\mathbb{P})$ désigne un espace probabilisé, on mesure la dépendance entre deux tribus $\mathcal{U}$ et $\mathcal{V}$ grâce aux coefficients qui suivent :

– $\alpha(\mathcal{U},\mathcal{V}) = \sup\{|P(U \cap V) - P(U)P(V)|;\ U \in \mathcal{U},\ V \in \mathcal{V}\}$ (mélange fort, Rosenblatt [57]);

– $\beta(\mathcal{U},\mathcal{V}) = \mathbf{E}\{\sup |P(V|\mathcal{U}) - P(V)|\,;\ V \in \mathcal{V}\}$ (Rozanov & Volkonski [60]);

– $\rho(\mathcal{U},\mathcal{V}) = \sup\{|Corr(u,v)|\,;\ u \in \mathbb{L}^2(\mathcal{U}),\ v \in \mathbb{L}^2(\mathcal{V})\}$ (Kolmogorov & Rozanov [31]);

– $\Phi(\mathcal{U},\mathcal{V}) = \sup\{|P(V|U) - P(V)|\,;\ U \in \mathcal{U},\ V \in \mathcal{V}\}$ (mélange uniforme, Ibragimov [29]).

Soit alors $(X) = (X_n)_{n\in\mathbb{Z}}$ une suite de variables aléatoires. On dit que que cette suite est $\beta-$mélangeante si la suite des coefficients de $\beta-$mélange $\beta(n) = \sup_{p\in\mathbb{Z}} \beta\left(\mathcal{F}^p_{-\infty},\mathcal{F}^{+\infty}_{p+n}\right)$ converge vers 0 lorsque n tend vers $+\infty$. Les notions de $\alpha-$mélange, de $\rho-$mélange ou de $\Phi-$mélange sont définies de façon analogue.

Nous avons les relations suivantes :

– $\Phi-$mélange $\Rightarrow$ $\rho-$mélange $\Rightarrow$ $\alpha-$mélange;

– $\Phi-$mélange $\Rightarrow$ $\beta-$mélange $\Rightarrow$ $\alpha-$mélange.

Beaucoup de travaux portent sur les propriétés et les applications des suites mélangeantes (voir Doukhan [18] ou encore Rio [55]). On sait montrer que certains processus sont mélangeants : par exemple, sous certaines conditions, les processus linéaires sont $\alpha-$mélangeants. Il en est de même des chaînes de Markov. On peut même alors donner un ordre de grandeur explicite pour les coefficients d'$\alpha-$mélange. Cependant, l'évaluation des coefficients de mélange reste souvent difficile. C'est pourquoi Doukhan & Louhichi [21] ont cherché à définir une notion de dépendance plus générale que le mélange, où l'écart mesuré n'est plus celui entre les tribus engendrées par le passé et le futur des variables aléatoires, mais l'écart entre des fonctions (appartenant à une classe "convenable") du passé et du futur de la suite.

Soit E un espace euclidien de la forme $\mathbb{R}^d$ muni de sa norme euclidienne $||.||$. On considère une suite de variables aléatoires $(X_n)_{n\in\mathbb{Z}}$ à valeurs dans E.

Définition 1 (Doukhan & Louhichi [21]) *Soient* $\mathcal{L} = \cup_{p=1}^{+\infty}\mathcal{L}^p$, *où* $\mathcal{L}^p$ = *classe de fonctions de* E^p *dans* $\mathbb{R}$, $\Psi : \mathcal{L}\times\mathcal{L}\times\mathbb{N}^{*2} \to \mathbb{R}_+$, *et* $\theta = (\theta_r)_{r\in\mathbb{N}}$ *où* $\theta_r \searrow 0$.

La suite $(X_n)_{n\in\mathbb{Z}}$ *est dite* $(\theta,\mathcal{L},\Psi)$ *faiblement dépendante si*
$\forall r \in \mathbb{N}$, $\forall u,v \in \mathbb{N}^*$, $\forall (h,k) \in \mathcal{L}^u \times \mathcal{L}^v$,
$\forall i_1 \le i_2 \le \cdots \le i_u < i_u + r \le j_1 \le \cdots \le j_v$,
$|\mathrm{Cov}\,(h(X_{i_1},\ldots,X_{i_u}),k(X_{j_1},\ldots,X_{j_v}))| \le \theta_r\ \Psi(h,k,u,v)$.

Remarquons que si une suite $(X_n)_{n\in\mathbb{Z}}$ est fortement mélangeante au sens de Rosenblatt [57], elle est $(\theta,\mathcal{L},\Psi)$ faiblement dépendante avec $\mathcal{L} = \mathbb{L}^\infty$, $(\theta_r) = (\alpha_r)$, $\Psi(h,k,u,v) = 4||h||_\infty||k||_\infty$.

Ce qui importe dans la définition 1 ci-dessus, c'est le choix de la classe $\mathcal{F}$ de fonctions du passé et du futur. En pratique, nous allons nous intéresser aux deux variantes suivantes de la définition 1 (définitions 2 et 3 ci-dessous).

Commençons par définir l'ensemble $\mathbb{L}^\infty$ comme l'ensemble des fonctions numériques mesurables et bornées sur un espace de la forme $\mathbb{R}^p$. On note $||.||_\infty$ la norme usuelle sur $\mathbb{L}^\infty$. Maintenant, si $q \in \mathbb{N}^*$, on munit l'ensemble $F = E^q$ de la norme $||(x_1,\ldots,x_q)||_F = ||x_1|| + \cdots + ||x_q||$. Soit alors $h : F = E^q \to \mathbb{R}$ une fonction numérique sur F, son module de Lipschitz est défini par

$$\mathrm{Lip}(h) = \sup_{x\neq y} \frac{|h(x) - h(y)|}{||x-y||_F}.$$

Nous sommes maintenant en mesure de donner les deux définitions qui suivent, variantes de la définition 1.

Définition 2 ($s-$dépendance) *La suite $(X_n)_{n\in\mathbb{Z}}$ est $s-$faiblement dépendante s'il existe une suite $\theta = (\theta_r)_{r\in\mathbb{N}}$, $\theta_r \searrow 0$, telle que*

$\forall r \in \mathbb{N}$, $\forall u \in \mathbb{N}^*$, $\forall h \in \mathbb{L}^\infty$, $||h||_\infty \leq 1$,

$\forall k \in \mathcal{L} := \cup_{p=1}^{+\infty}\{f : E^p \to \mathbb{R}, f \in \mathbb{L}^\infty, ||f||_\infty \leq 1, \mathrm{Lip}(f) < \infty\}$,

$\forall i_1 \leq \cdots \leq i_u < i_u + r \leq j$,

$|\mathrm{Cov}\,(h(X_{i_1},\ldots,X_{i_u}),k(X_j))| \leq \mathrm{Lip}(k)\theta_r$.

Définition 3 ($a-$dépendance) *La suite $(X_n)_{n\in\mathbb{Z}}$ est $a-$faiblement dépendante s'il existe une suite $\theta = (\theta_r)_{r\in\mathbb{N}}$, $\theta_r \searrow 0$, telle que*

$\forall r \in \mathbb{N}$, $\forall u \in \mathbb{N}^*$, $\forall h,k \in \mathcal{L} := \cup_{p=1}^{+\infty}\{f : E^p \to \mathbb{R}, f \in \mathbb{L}^\infty, ||f||_\infty \leq 1, \mathrm{Lip}(f) < \infty\}$,

$\forall i_1 \leq \cdots \leq i_u < i_u + r \leq j$,

$|\mathrm{Cov}\,(h(X_{i_1},\ldots,X_{i_u}),k(X_j))| \leq \mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r$.

Donnons maintenant quelques exemples de suites $s-$ ou $a-$dépendantes.

Suites $s-$dépendantes :

Nous commençons par donner la définition suivante :

Définition 4 (shift) *Soit $(\eta_n)_{n\in\mathbb{Z}}$ une suite stationnaire de va réelles, et soit F une fonction mesurable définie sur $\mathbb{R}^\mathbb{Z}$. La suite stationnaire $(X_n)_{n\in\mathbb{Z}}$ définie par $X_n = F(\eta_n,\eta_{n-1},\eta_{n-2},\ldots)$ est appelée un shift de Bernoulli causal.*

Si la suite des innovations $(\eta_n)_{n\in\mathbb{Z}}$ est identiquement distribuée indépendante (i.i.d.), si $(\theta_r)_{r\in\mathbb{N}}$ désigne une suite positive et décroissant vers 0 telle que

$$\mathbf{E}\,|F(\eta_0,\eta_{-1},\eta_{-2},\ldots) - F(\eta_0,\ldots,\eta_{-r},0,0,\ldots)| \leq \theta_r,$$

alors $(X_n)_{n\in\mathbb{Z}}$ est $s-$faiblement dépendante (avec la suite (θ_r)) (voir Doukhan & Louhichi [21]). D'où le s (comme shift) dans $s-$dépendance.

Donnons quelques exemples de telles situations :

- Modèle autorégressif fonctionnel à valeurs réelles
  Supposons qu'il existe $0 \leq k < 1$ et $T : \mathbb{R} \to \mathbb{R}$ tels que $|T(u) - T(u')| \leq k|u - u'|$ pour tous u, $u' \in \mathbb{R}$, et $(\eta_n)_{n\in\mathbb{Z}}$ i.i.d. telle que $\mathbf{E}|\eta_0| < \infty$,
  alors la suite $(X_n)_{n\in\mathbb{Z}}$ définie par $X_n = T(X_{n-1}) + \eta_n$, est $s-$faiblement dépendante avec $\theta_r = Ck^r$ où $C > 0$.

- La chaîne de Markov stationnaire, non mélangeante, à innovations binomiales i.i.d. $(\eta_n)_{n\in\mathbb{Z}}$, et définie par $\xi_n = (\xi_{n+1} + \eta_n)/2$ a pour loi marginale la loi uniforme sur $[0,1]$ et est $s-$dépendante avec $\theta_r = \mathcal{O}\,(2^{-r})$.

- Les processus linéaires $X_n = \sum_{k=0}^{\infty} a_k\eta_{n-k}$, où la suite des innovations $(\eta_n)_{n\in\mathbb{Z}}$ est i.i.d. et telle que $\mathbf{E}|\eta_0| < \infty$, sont $s-$faiblement dépendants avec $\theta_r = \mathbf{E}|\eta_0|\sum_{k>r}|a_k|$. Si de plus les innovations η_i sont centrées et dans $\mathbb{L}^2$, nous pouvons choisir $\theta_r = \sqrt{\mathbf{E}|\eta_0|^2\sum_{k>r}|a_k|^2}$.

- Le processus bilinéaire défini par la relation $X_n = aX_{n-1} + bX_{n-1}\eta_{n-1} + \eta_n$, où $(\eta_n)_{n\in\mathbb{Z}}$ est i.i.d., est également $s-$faiblement dépendant avec un coefficient de dépendance faible $\theta_r = \frac{c^r\,(r+1)}{c-1}$, où $c = \mathbf{E}\,|a + b\eta_0| < 1$, qui décroît géométriquement.

Nous renvoyons à Coulon-Prieur & Doukhan [16] et à Prieur [51] pour une liste d'exemples plus exhaustive.

Suites $a-$dépendantes :

Définition 5 (association) *La suite $(\xi_n)_{n\in\mathbb{Z}}$ est associée si pour toutes h,k réelles croissantes coordonnée par coordonnée sur $\mathbb{R}^A$ ($A \subset \mathbb{Z}$, A fini) telles que $\mathbf{E}\left(h^2(\xi_A) + k^2(\xi_A)\right) < \infty$,*
$\mathrm{Cov}\left(h(\xi_A),k(\xi_A)\right) \geq 0$.

Les suites associées sont $a-$faiblement dépendantes avec $\theta_r = \sup_{t\in\mathbb{Z}} \sum_{k\geq r} \mathrm{Cov}\left(\xi_t,\xi_{t+k}\right)$. D'où le a (comme association) de $a-$dépendance.
Les processus gaussiens satisfont aussi la $a-$dépendance faible, mais avec
$\theta_r = \sup_{t\in\mathbb{Z}} \sum_{k\geq r} \left|\mathrm{Cov}\left(\xi_t,\xi_{t+k}\right)\right|$.

Les trois premiers chapitres de cette thèse présentent certaines applications statistiques de suites $s-$ ou $a-$faiblement dépendantes. Le chapitre 4 porte sur l'estimation de la densité invariante de certains systèmes dynamiques (en dimension 1) ayant une propriété dite de décroissance des corrélations. Enfin, le dernier chapitre (chapitre 5) contient les résultats de simulations portant sur le chapitre 4. Nous allons maintenant donner un résumé de chacun des chapitres de cette thèse.

Chapitre 1 : Théorème Limite Centrale Triangulaire et Applications Statistiques (d'après Coulon-Prieur & Doukhan [16]).

Dans ce chapitre, nous établissons un théorème limite centrale pour des tableaux triangulaires de suites faiblement dépendantes au sens des définitions 2 et 3 précédentes.

La méthode jusqu'alors la plus usuelle pour montrer un théorème limite centrale pour des suites dépendantes est une décomposition en blocs de Bernstein suivie de la méthode de Lindeberg, standard pour les suites indépendantes (voir par exemple Withers [66], Doukhan & Louhichi [21]).

Dans ce chapitre, nous n'utilisons pas de blocs de Bernstein. Nous étendons la méthode de Lindeberg utilisée par Rio [53, 54] dans le cadre mélangeant (ou par exemple par Louhichi [39] dans le cadre des suites positivement dépendantes) à notre cadre de dépendance faible.

Nous montrons, pour une suite de variables aléatoires $(X_n)_{n\geq 1}$ faiblement dépendantes, un théorème limite centrale pour des tableaux triangulaires de la forme

$$\left(X_{n,k} = g_{n,k}(X_k)\right)_{n\geq 1, 1\leq k\leq k_n}, \tag{1}$$

où pour tout $n \geq 1$, $k_n \in \mathbb{N}^*$, avec $k_n \underset{n\infty}{\longrightarrow} +\infty$, et où les $g_{n,k}$ sont des fonctionnelles numériques lipschitziennes bornées (voir Coulon-Prieur & Doukhan [16]).

Théorème 1 *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires stationnaires, à valeurs dans un espace $\mathbb{R}^d$, et $s-$faiblement dépendantes. Soit le tableau triangulaire $(X_{n,k})_{n\geq 1, 1\leq k\leq k_n}$ défini par (1). On définit*

$$S_n := X_{n,1} + \cdots + X_{n,k_n}, \ \ et\ S_{k,n} = X_{n,1} + \cdots + X_{n,k}, \ 1 \leq k \leq k_n. \tag{2}$$

Nous faisons les hypothèses suivantes :

$$\begin{array}{rl} & \operatorname{Var}(S_n) \xrightarrow[n\,\infty]{} \sigma^2 > 0 \\ \exists\, n_0 \geq 1 \;/ & n \geq n_0 \Rightarrow v_{k,n} := \operatorname{Var}(S_{k,n}) - \operatorname{Var}(S_{k-1,n}) > 0, \forall\, 1 \leq k \leq k_n. \end{array} \tag{3}$$

On définit enfin les quantités suivantes :

$$\begin{array}{rcl} \lambda_n = \sup_{1\leq k\leq k_n} \operatorname{Lip}(g_{n,k}) & , & M_n = \sup_{1\leq k\leq k_n} \|g_{n,k}\|_\infty, \\ \delta_n = \sup_{1\leq k\leq k_n} \mathbf{E}|X_{n,k}| & , & \Delta_n = \sup_{1\leq k\neq l\leq k_n} \mathbf{E}|X_{n,k}X_{n,l}|. \end{array} \tag{4}$$

Alors, si

$$(k_nM_n + k_n^{2/3})M_n\delta_n \to 0,\; k_nM_n\sum_{p=1}^{k_n}\min(\lambda_n\theta_p,\Delta_n+\delta_n^2) \to 0,\; k_n\sum_{p=1}^{k_n}\min(M_n\lambda_n\theta_p,\Delta_n) \to 0,$$

lorsque n tend vers l'infini, nous avons :

$$S_n \xrightarrow[n\,\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2).$$

Nous avons aussi l'équivalent pour la $a-$dépendance.

Théorème 2 *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires stationnaires, à valeurs dans un espace $\mathbb{R}^d$, et $a-$faiblement dépendantes. Soit le tableau triangulaire $(X_{n,k})_{n\geq 1, 1\leq k\leq k_n}$ défini par (1). Nous reprenons les notations (2) et (4) du théorème 1.*

Alors, sous les hypothèses (3) du théorème 1, et si

$$(k_nM_n + k_n^{2/3})M_n\delta_n \to 0,\; k_nM_n\sum_{p=1}^{k_n}\min(\lambda_n^2\theta_p,\Delta_n+\delta_n^2) \to 0,\; k_n\sum_{p=1}^{k_n}\min(\lambda_n^2\theta_p,\Delta_n) \to 0,$$

lorsque n tend vers l'infini, nous avons :

$$S_n \xrightarrow[n\,\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2).$$

Dans Coulon-Prieur & Doukhan [16], nous donnons deux applications de ces théorèmes limite centrale triangulaires. La première porte sur des tableaux linéaires (les $g_{n,k}$ sont alors des fonctions linéaires de la forme $g_{n,k}(x) = a_{n,k} \times x$, voir aussi Peligrad & Utev [44]). La deuxième porte sur le problème d'estimation non paramétrique d'une densité. Nous développons ci-dessous le problème de l'estimation de la densité marginale d'une suite stationnaire de variables aléatoires faiblement dépendantes.

Estimation de la densité d'une suite stationnaire faiblement dépendante

Nous considérons une suite stationnaire $(X_n)_{n\geq 1}$ faiblement dépendante. Nous supposons que la loi marginale de cette suite $(X_n)_{n\geq 1}$ admet une densité notée f. Nous voulons estimer f par la méthode de noyau. L'estimateur de f que nous considérons est le suivant :

$$\hat{f}_{n,b_n}(x) := \frac{1}{nb_n}\sum_{k=1}^{n} K\left(\frac{x-X_k}{b_n}\right),$$

où $(b_n)_{n\geq 1} \in (\mathbb{R}_+)^{\mathbb{N}^*}$ désigne la fenêtre, $b_n \underset{n\infty}{\longrightarrow} 0$, et où K est une fonction à valeurs réelles appelée noyau (voir par exemple Bosq & Guégan [8]).

Nous supposons que la densité marginale f est continue et que les densités jointes des vecteurs (X_0,X_n) sont uniformément bornées. Alors, sous certaines hypothèses portant sur le noyau K (K lipschitzien, d'intégrale 1, à support compact tel que $\int K^2 > 0$ par exemple), nous obtenons un théorème limite centrale, avec la normalisation $\sqrt{nb_n}$, pour l'estimateur recentré $\hat{f}_{n,b_n}(x) - \mathbf{E}\hat{f}_{n,b_n}(x)$, à condition que $nb_n \underset{n\infty}{\longrightarrow} +\infty$, et que la suite des coefficients de dépendance faible de $(X_n)_{n\geq 1}$ satisfasse $\theta_r = \mathcal{O}\left(r^{-a}\right)$, $a > 3$ dans le cas de la $s-$dépendance, ou $\theta_r = \mathcal{O}\left(r^{-a}\right)$, $a > 4$ dans le cas de la $a-$dépendance. Plus précisément :

$$\mathbf{Si}\ (\mathbf{x_1},\ldots,\mathbf{x_l})/\ \forall\, \mathbf{1} \leq \mathbf{i} \leq \mathbf{l}, \mathbf{f(x_i)} \neq \mathbf{0},\ ,\ \mathbf{si}\ \mathbf{Y_n(x)} := \frac{\sqrt{\mathbf{nb_n}}\left(\mathbf{\hat{f}_{n,b_n}(x) - E\hat{f}_{n,b_n}(x)}\right)}{\sqrt{\mathbf{f(x)}\int \mathbf{K^2}}},\ \mathbf{alors}$$

$$(\mathbf{Y_n(x_1)},\ldots,\mathbf{Y_n(x_l)}) \underset{\mathbf{n\infty}}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\left(\mathbf{0},\mathbf{I_l}\right) \tag{5}$$

(voir Coulon-Prieur & Doukhan [16] pour une formulation plus précise).

Nous en déduisons, si $f > 0$, la convergence des répartitions fini-dimensionnelles de la suite de processus $(Y_n(x),\, x \in \mathbb{R})$ vers les répartitions fini-dimensionnelles du bruit blanc gaussien. La suite de processus $(Y_n(x),\, x \in \mathbb{R})$ n'est donc pas tendue dans $C(\mathbb{R})$.

Ce résultat est du même type que ceux obtenus dans les cas i.i.d. ou mélangeant (voir par exemple Rosenblatt [58], Robinson [56]). Dans le cadre de la dépendance faible, Doukhan & Louhichi [22] obtenait déjà un tel résultat pour l'estimation de la densité, mais sous des hypothèses plus restrictives : $\theta_r = \mathcal{O}\left(r^{-a}\right)$, $a > max\{9,\frac{3}{2}(1+\delta^{-1})\}$ si $b_n \sim n^{-\delta}$ pour la $s-$dépendance ou $\theta_r = \mathcal{O}\left(r^{-a}\right)$, $a > 12$ pour le cas de la $a-$dépendance. Ces auteurs utilisaient dans leur article des blocs de Bernstein.

L'étude du comportement du biais $\mathbf{E}\hat{f}_n(x) - f(x)$ est classique. Elle dépend de la régularité de f et non des propriétés de dépendance de la suite $(X_n)_{n\geq 1}$. Nous renvoyons à Prieur [50] pour une étude détaillée du terme de biais.

Donnons maintenant un schéma volontairement simplifié de la méthode que nous utilisons pour montrer un théorème limite centrale en dépendance faible, notamment les théorèmes 1 et 2 ci-dessus (voir Coulon-Prieur & Doukhan [16]), mais aussi le théorème limite centrale qui nous permet dans Prieur [51] d'obtenir la convergence des répartitions fini-dimensionnelles du processus empirique (voir chapitre 2), ou encore le théorème limite centrale pour notre estimateur de régression linéaire locale au chapitre 3.

Soit $(Z_n)_{n\geq 1}$ une suite stationnaire de variables aléatoires centrées et faiblement dépendantes.

Notons $S_n = Z_1 + \cdots + Z_n$, $S_0 = 0$ et $\sigma_n = \sqrt{\operatorname{Var} S_n}$. $\sigma = \lim_{n\infty} \sigma_n$ est supposée finie et strictement positive. Soit η une gaussienne centrée réduite, et pour $1 \leq k \leq n$, $v_k = \operatorname{Var} S_k - \operatorname{Var} S_{k-1}$. Supposons les v_k strictement positifs.

Introduisons alors une suite $(Y_n)_{n\geq 1}$ de gaussiennes centrées, de variance v_n, indépendantes. Choisissons la suite $(Y_n)_{n\geq 1}$ indépendante de la suite de départ $(Z_n)_{n\geq 1}$.

Posons $T_k = \sum_{i=k+1}^{n} Y_i$, $k \leq n-1$, $T_n = 0$.

Pour montrer la convergence de S_n vers une gaussienne centrée de variance σ^2, nous montrons pour une classe de fonctions f suffisamment large que la différence $\mathbf{E}\,\{f(S_n) - f(\sigma_n\,\eta)\}$ tend vers 0.

Méthode de Lindeberg

Nous commençons par utiliser la décomposition proposée par Lindeberg [35] dans le cas indépendant :

$$\begin{aligned}\mathbf{E}\,\{f(S_n) - f(\sigma_n\,\eta)\} &= \textstyle\sum_{k=1}^n \Delta_k(f) \\ &= \textstyle\sum_{k=1}^n \mathbf{E}\,\{f(S_{k-1} + Z_k + T_k) - f(S_{k-1} + Y_k + T_k)\}.\end{aligned}$$

Dans le cas indépendant, on conclut avec un développement de Taylor. Mais lorsque la suite $(X_n)_{n\geq 1}$ n'est pas indépendante, cette décomposition ne suffit pas. C'est pourquoi nous décomposons à nouveau chacun des termes $\Delta_k(f)$:

$$\Delta_k(f) = \Delta_k^{(1)}(f) - \Delta_k^{(2)}(f),$$

où

$$\Delta_k^{(1)}(f) = \mathbf{E}\ f(S_{k-1} + Z_k + T_k) - \mathbf{E}\ f(S_{k-1} + T_k) - \frac{v_k}{2}\mathbf{E}\ f''(S_{k-1} + T_k), \tag{6}$$

$$\Delta_k^{(2)}(f) = \mathbf{E}\ f(S_{k-1} + Y_k + T_k) - \mathbf{E}\ f(S_{k-1} + T_k) - \frac{v_k}{2}\mathbf{E}\ f''(S_{k-1} + T_k). \tag{7}$$

Il nous faut alors trouver de bonnes bornes de majoration pour les termes $\Delta_k^{(1)}(f)$ et $\Delta_k^{(2)}(f)$.

Pour $\Delta_k^{(2)}(f)$, les propriétés d'indépendance de la suite $(Y_n)_{n\geq 1}$ nous permettent de conclure à l'aide d'un développement de Taylor. La méthode utilisée ensuite est inspirée de Rio [53]. La principale étape dans la preuve de Rio est l'évaluation de

$$\Delta_k^{(1)}(f) = \mathbf{E}\ f(S_{k-1} + Z_k + T_k) - \mathbf{E}\ f(S_{k-1} + T_k) - \frac{v_k}{2}\mathbf{E}\ f''(S_{k-1} + T_k).$$

Notre objectif est donc de contrôler les termes $\Delta_k^{(1)}(f)$ pour des suites faiblement dépendantes au sens des définitions 2 et 3. Nous faisons apparaître des covariances que nous pouvons majorer, soit de façon directe en utilisant la décomposition

$$\mathrm{Cov}\,(g(Z_{i_1},\ldots,Z_{i_u}),k(Z_j)) = \mathbf{E}g(Z_{i_1},\ldots,Z_{i_u})k(Z_j) - \mathbf{E}g(Z_{i_1},\ldots,Z_{i_u})\ \mathbf{E}k(Z_j),$$

soit en nous appuyant sur l'une des inégalités de covariance intervenant dans les définitions 2 et 3 :

$$|\mathrm{Cov}\,(h(Z_{i_1},\ldots,Z_{i_u}),k(Z_j))| \leq \|h\|_\infty \mathrm{Lip}(k)\theta_r,$$

ou

$$|\mathrm{Cov}\,(h(Z_{i_1},\ldots,Z_{i_u}),k(Z_j))| \leq \mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r.$$

Le but est de trouver des majorations assez fines pour qu'en sommant les majorants de (6) et (7) sur les k allant de 1 à k_n, nous obtenions une quantité qui tende vers 0 lorsque n tend vers l'infini (voir Coulon-Prieur & Doukhan [16], Prieur [51]).

Cette méthode fonctionne encore si nous considérons un tableau triangulaire $(Z_{n,k})_{n\geq 1,\,1\leq k\leq k_n}$ à la place de la suite $(Z_n)_{n\geq 1}$. C'est le cas dans Coulon-Prieur & Doukhan

[16] pour montrer le théorème limite centrale triangulaire. Les sommes considérées sont alors les sommes $S_{n,n} = Z_{n,1} + \cdots + Z_{n,k_n}$ et $S_{k,n} = Z_{n,1} + \cdots + Z_{n,k}$.

Chapitre 2 : Théorème Limite Centrale Fonctionnel Empirique (d'après Prieur [51]).

L'objet de ce chapitre est l'étude du processus empirique construit à partir d'une suite de variables aléatoires $s-$faiblement dépendante au sens de la définition 2. Nous obtenons un théorème limite centrale pour la suite des fonctions de répartition empirique recentrées et renormalisées (théorème 3 ci-dessous).

Considérons une suite stationnaire $(X_n)_{n\geq 1}$ de variables aléatoires ayant pour fonction de répartition commune F supposée lipschitzienne. La suite $(X_n)_{n\geq 1}$ est supposée $s-$faiblement dépendante. Le processus empirique associé à la suite $(X_n)_{n\geq 1}$ est défini comme suit :

$$\forall\, t \in \mathbb{R},\, F_n(t) = \frac{1}{n}\sum_{k=0}^{n-1} 1_{X_k \leq t} \text{ et } U_n(t) = \sqrt{n}\left(F_n(t) - F(t)\right). \tag{8}$$

Théorème 3 (Théorème limite centrale fonctionnel empirique) *Soit $(X_n)_{n\geq 1}$ une suite stationnaire de variables aléatoires à valeurs réelles, ayant une fonction de répartition commune F supposée lipschitzienne. Supposons que $(X_n)_{n\geq 1}$ soit $s-$faiblement dépendante avec $\theta_r = \mathcal{O}\left((r+1)^{-2-2\sqrt{2}-\nu}\right)$, $\nu > 0$. Alors la suite de processus $\{U_n(t),\, t \in \mathbb{R}\}_{n\geq 1}$ définie par (8) converge en distribution, dans l'espace de Skohorod $\mathcal{D}(\mathbb{R})$, vers le processus gaussien centré indexé par $\mathbb{R}$ et de covariance définie par*

$$\Gamma(s,t) = \sum_{k=-\infty}^{+\infty} \operatorname{Cov}\left(1_{X_0 \leq s}, 1_{X_{|k|} \leq t}\right).$$

La preuve de ce résultat repose essentiellement sur des nouvelles inégalités de moments (lemme 1 ci-dessous) et sur un théorème limite centrale en dépendance faible (théorème 4 ci-dessous). Doukhan et Louhichi obtiennent aussi dans [21] un théorème limite centrale fonctionnel pour le processus empirique associé à une suite faiblement dépendante, mais la méthode employée ici conduit à des hypothèses moins restrictives aussi bien pour la tension de la suite de processus que pour la convergence de ses répartitions fini-dimensionnelles.

Inégalités de moments

Pour montrer nos inégalités de moments, nous adaptons la méthode proposée par Louhichi [40] pour calculer les moments des sommes partielles d'une suite de variables aléatoires $(X_n)_{n\geq 1}$ vérifiant une condition de dépendance faible que nous pouvons ainsi résumer : "la non-corrélation entraîne l'indépendance", décrite par l'inégalité de covariance suivante :

$$\left|\operatorname{Cov}\left(h(X_i,\, i \in A), k(X_j,\, j \in B)\right)\right| \leq \sum_{i\in A}\sum_{j\in B} \left\|\frac{\partial h}{\partial x_i}\right\|_\infty \left\|\frac{\partial k}{\partial x_j}\right\|_\infty \left|\operatorname{Cov}(X_i, X_j)\right|,$$

où A et B sont des sous-ensembles disjoints arbitraires de $\mathbb{N}^*$ et h, k des fonctions qui ont une différentielle bornée.

Rappelons dans un premier temps l'inégalité de Rosenthal en indépendant :

Soit $X = (X_1, \ldots, X_n)$ un vecteur centré de variables aléatoires réelles de variances finies. Soit $S_n = X_1 + \cdots + X_n$, $S_0 = 0$. Alors, si $r \in]2, +\infty[$, on a

$$E|S_n|^r \leq C_r\{(\mathrm{Var}\, S_n)^{\frac{r}{2}} + n\mathbf{E}|X_1|^r\}.$$

Cette inégalité a de nombreuses applications probabilistes et statistiques, et son principal avantage est de ramener le contrôle des moments d'ordres grands des sommes partielles à celui des moments d'ordre 2. Citons par exemple Rosenthal [59], Petrov [45] pour d'autres versions de cette inégalité, Stout [64], Bretagnolle & Huber [10] pour des applications variées.

Comme beaucoup de modèles statistiques ne font pas intervenir des suites indépendantes de variables aléatoires, il est intéressant de rechercher des inégalités de moments en dépendance faible. Beaucoup d'auteurs parmi lesquels Doukhan & Portal [23], Rio [52, 53], Shao [62], ... ont montré des versions de l'inégalité de Rosenthal dans le cadre du mélange. En dépendance faible, nous avons des résultats de Doukhan & Louhichi [21] concernant les moments d'ordres entiers pairs. Dans cette thèse nous montrons des inégalités de moments d'ordres $r \in]2, +\infty[$ pour des suites $s-$faiblement dépendantes.

Lemme 1 (Inégalités de moments) *Soit $r \in]2, +\infty[$. Soit $(X_n)_{n\geq 1}$ une suite stationnaire et centrée de variables aléatoires $s-$faiblement dépendantes. Supposons de plus que la suite $(X_n)_{n\geq 1}$ est uniformément bornée par 1. On définit, pour $n \geq 1$, $S_n = X_1 + \cdots + X_n$ et $S_0 = X_0 = 0$. Alors il existe un réel positif C_r dépendant seulement de r tel que*

$$\mathbf{E}|S_n|^r \leq C_r \left(M_{2,n}^{r/2} + M_{r,n}\right),$$

où $M_{r,n} := n\sum_{i=0}^{n-1}(i+1)^{r-2}\theta_i$.

Outre leur intérêt théorique, ces inégalités sont appliquées pour montrer la tension du processus empirique associé à une suite de variables $s-$faiblement dépendantes (voir Prieur [51]). Nous déduisons également de ces inégalités de moments une vitesse de convergence dans une loi forte de Marcinkiewicz-Zygmund ([51]).

Théorème Limite Centrale

Nous montrons aussi un théorème limite centrale pour des suites de variables aléatoires $s-$faiblement dépendantes.

Théorème 4 (Théorème limite centrale) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite stationnaire de variables aléatoires centrées, $s-$dépendantes avec $\theta_r = \mathcal{O}\left((r+1)^{-a}\right)$, $a > 3/2$. Soit $S_n = X_0 + \cdots + X_{n-1}$. Supposons que $\mathrm{Var}\,(S_n)/n \underset{n\infty}{\longrightarrow} \sigma^2 > 0$. Alors,*

$$\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1} X_i \xrightarrow[n\infty]{\mathcal{L}} \mathcal{N}\left(0, \sigma^2\right).$$

Pour montrer ce théorème, nous utilisons la méthode décrite dans le paragraphe **Chapitre 1** à la page 16 de cette introduction. Ce théorème permet de retrouver la convergence fini-dimensionnelle dans le théorème limite centrale fonctionnel empirique, avec en plus un affaiblissement des hypothèses.

Chapitre 3 : Problème de rupture (travail joint avec Patrick Ango Nze).

Lorsqu'un phénomène se produit sur une longue période, la question de sa stabilité se pose. Il peut effectivement se produire une altération du modèle à partir d'un certain instant τ. Cette altération peut se traduire, lorsqu'il s'agit d'une fonction de régression, par un saut de cette fonction, ou par une rupture de pente. De façon générale, une rupture est un changement instantané (à l'échelle de la période d'échantillonnage) dans les paramètres du système étudié. Les modèles de rupture en statistiques mathématiques connaissent un essor important. Les applications sont nombreuses. Parmi les domaines d'application, citons la finance, les données médicales, l'écologie, le traitement du signal ou encore l'étude des séismes. Une rupture peut avoir des conséquences plus ou moins importantes comme la contamination par un virus (rupture du taux de certains marqueurs dans le sang), le changement de tendances à la bourse, etc... Nous renvoyons aux thèses de Couallier [15] et de Hamrouni [28] pour des exemples plus détaillés.

Dans cette thèse, nous présentons un travail joint avec Patrick Ango Nze. Nous nous intéressons à des problèmes de discontinuités dans le cadre du modèle défini ci-dessous.

Soit $(X_1,Y_1),\ldots,(X_n,Y_n)$ un ensemble de vecteurs aléatoires, où la suite $(X_n)_{n\geq 1}$ est une suite stationnaire de variables aléatoires $s-$faiblement dépendantes au sens de la définition 2 de cette introduction. Nous supposons que cet ensemble de vecteurs est gouverné par le modèle

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i,$$

où la fonction $m(\cdot)$ est lisse sauf en un point τ, la fonction $\sigma(\cdot)$ est continuement différentiable et les ε_i sont des variables aléatoires i.i.d. centrées réduites qui sont indépendantes des X_i. Les fonctions $m(\cdot)$ et $\sigma(\cdot)$ sont inconnues. On peut supposer en toute généralité que les X_i sont à valeurs dans l'intervalle $[0,1]$.

Nous faisons maintenant les hypothèses suivantes sur le modèle :

(1) $\mathbf{E}\,\varepsilon = 0$, $\mathbf{E}\,\varepsilon^2 = 1$, $\mathbf{E}\,|\varepsilon|^3 < \infty$.

(2) La densité marginale f de la suite stationnaire $(X_n)_{n\geq 1}$ est strictement positive et continue.

(3) Les densités jointes des couples (X_0,X_n), $f_n(x,y)$, sont uniformément bornées : $\sup_{n\geq 1}\sup_{x,\,y}|f_n(x,y)| < \infty$.

(4) Les restrictions de la fonction de régression $m(x) = \mathbf{E}(Y|X=x)$ aux intervalles $[0,\tau]$ et $[\tau,1]$ ont toutes deux des dérivées secondes continues. On en déduit que la fonction de régression $m(\cdot)$ et ses dérivées première et seconde ont chacune une limite à droite et une limite à gauche au point τ.

(5) La variance conditionnelle $\sigma^2(x) = \mathrm{Var}\,(Y|X=x)$ est continuement différentiable.

(6) L'instant de saut τ de la fonction de régression est supposé appartenir à l'intervalle ouvert]0,1[.

Les hypothèses (2) et (5) peuvent être affaiblies (voir chapitre 3). L'hypothèse (3) est classique en dépendance faible. Les autres hypothèses sont des hypothèses techniques qui interviennent dans les démonstrations.

Dans la suite, pour $\nu = 0, 1, 2$ nous utilisons les notations suivantes :

(i) $m_-^{(\nu)}(\tau) = \lim_{t \nearrow \tau} m^{(\nu)}(t)$,

(ii) $m_+^{(\nu)}(\tau) = \lim_{t \searrow \tau} m^{(\nu)}(t)$,

(iii) $\gamma^{(\nu)}(\tau) = m_+^{(\nu)}(\tau) - m_-^{(\nu)}(\tau)$.

Le but est d'estimer l'amplitude du saut $\gamma(\tau) = m_+(\tau) - m_-(\tau)$.

Il existe différentes approches, aussi bien semi paramétriques que non paramétriques, pour étudier les sauts d'une fonction de régression. Nous avons sélectionné l'approche de Grégoire & Hamrouni [27]. Cette approche, non paramétrique, repose sur le lissage à droite et à gauche par régression locale. Avant de préciser cette méthode, notons que notre choix est en particulier motivé par les deux points qui suivent : d'une part la méthode de régression polynômiale locale présente une propriété d'absence d'effet de bord (la fonction de régression peut être estimée avec la même vitesse de convergence près des bords que dans la partie centrale), d'autre part cette méthode semble se comporter de façon satisfaisante pour des tailles d'échantillon relativement modestes (voir Grégoire & Hamrouni [27] pour plus de détails).

Montrons ci-dessous comment l'estimateur de régression linéaire locale est obtenu.

Régression Linéaire Locale

Soient K un noyau et h_n une fenêtre. L'estimateur de régression linéaire locale est obtenu en résolvant le problème de minimisation suivant :

$$\min_{\alpha_x,\, \beta_x} \sum_{i=1}^{n} (Y_i - \alpha_x - \beta_x(x - X_i))^2 K\left(\frac{x - X_i}{h_n}\right).$$

Sous certaines hypothèses, ce problème de minimisation (qui correspond à un problème des moindres carrés pondérés dans le cas indépendant) admet une unique solution $(\hat{\alpha_x}, \hat{\beta_x})$. L'estimateur de régression linéaire locale est alors donné par la définition 6 ci-dessous.

Définition 6 (Estimateur de régression linéaire locale) *L'estimateur de régression linéaire locale de $m(x)$ est défini par*

$$\hat{m}(x) = \hat{\alpha}_x = \frac{\sum_{i=1}^n \omega_i(x) Y_i}{\sum_{i=1}^n \omega_i(x)} ,$$

où, pour $i = 1, \ldots, n$,

$$\omega_i(x) = K\left(\frac{x - X_i}{h_n}\right) (S_2(x) - (x - X_i) S_1(x))$$

avec, pour $l = 0, 1, 2$,

$$S_l(x) = \sum_{i=1}^{n} (x - X_i)^l K\left(\frac{x - X_i}{h_n}\right) .$$

Nous étudions le cas où la rupture se produit en un seul instant τ connu. Nous considérons alors deux noyaux $K_+(\cdot)$ et $K_-(\cdot)$ tels que :

- $K_+ : [-1{,}0] \to \mathbb{R}_+$ continuement différentiable,
- $K_- : [0{,}1] \to \mathbb{R}_+$ continuement différentiable et $K_-(x) = K_+(-x)$.

Posons :

(i) $K_l^+ = \int_{-1}^0 x^l K_+(x)dx$, $l \in \mathbb{N}$,

(ii) $B_+ = (K_2^+)^2 - K_3^+ K_1^+$,

(iii) $V_+ = \int_{-1}^0 (K_2^+ - xK_1^+)^2 K_+^2(x)dx$.

Nous estimons le saut éventuel au point τ, $\gamma(\tau) = m_+(\tau) - m_-(\tau)$, en utilisant un estimateur à droite pour $m_+(\tau)$ et un autre à gauche pour $m_-(\tau)$. Ces estimateurs $\hat{m}_+(\tau)$ et $\hat{m}_-(\tau)$ sont construits respectivement à partir de K_+ et de K_- comme dans la définition 6. Parmi les premiers auteurs à proposer une méthode reposant sur ce principe, citons Mac Donald & Owen [41]. Nous obtenons dans un premier temps une expression pour l'erreur quadratique conditionnelle moyenne :

Lemme 2 (Erreur quadratique conditionnelle moyenne) *On suppose que la suite $(X_n)_{n\in\mathbb{N}}$ est $s-$faiblement dépendante. Si le paramètre de fenêtre vérifie $h_n \to 0$ et $nh_n \to +\infty$, si de plus $\theta_r = r^{-a}$ avec $a \geq 3$, alors l'erreur quadratique conditionnelle moyenne de $\widehat{\gamma}(\tau)$ est*

$$\mathbf{E}\left[(\widehat{\gamma}(\tau) - \gamma(\tau))^2 | X_1, \ldots, X_n\right] = \left[\frac{\gamma^{(2)}(\tau)}{2} h_n^2 B_+\right]^2 + \frac{2\sigma^2(\tau)V_+}{nh_n f(\tau)} + \ o_P\left(h_n^4 + \frac{1}{nh_n}\right).$$

Nous montrons ensuite, en utilisant la méthode que nous avons décrite ci-dessus (page 16), que $\hat{\gamma}(\tau) = \hat{m}_+(\tau) - \hat{m}_-(\tau)$ est asymptotiquement normal :

Théorème 5 (Théorème limite centrale) *On suppose que la suite $(X_n)_{n\in\mathbb{N}}$ est $s-$faiblement dépendante. Si $\theta_r = r^{-a}$ avec $a \geq 3$, et si $nh^5 \to 0$, $nh \to +\infty$ alors*

$$\sqrt{nh}\left(\widehat{m}_-(\tau) - m_-(\tau), \widehat{m}_+(\tau) - m_+(\tau)\right)'$$

converge en loi vers la loi normale en dimension 2

$$\mathcal{N}\left(0, \left(V_+\sigma^2(\tau)/f(\tau)\right) I_2\right)$$

où I_2 est la matrice identité en dimension 2.

Le cas où la rupture est localisée correspond à certains cas pratiques, par exemple au cas où un expérimentateur produit lui-même une discontinuité dans les paramètres (en médecine: une intervention chirurgicale, le démarrage d'un nouveau traitement, etc).

Le cas où la rupture n'est pas localisée est traité dans le cas indépendant par Grégoire & Hamrouni [27] par exemple, mais une adaptation de leurs résultats en dépendance faible présente de nombreuses difficultés. En particulier, il faut résoudre le problème de la convergence de suites de variables aléatoires faiblement dépendantes vers un processus de Poisson composé. Ce cas est encore à l'étude.

Chapitre 4 : Estimation de la densité invariante de systèmes dynamiques (d'après Prieur [49, 50]).

Dans ce chapitre, nous étudions des systèmes dynamiques en dimension 1 ayant une propriété de décroissance des corrélations. Avant de préciser ce que nous entendons par décroissance des corrélations, expliquons pourquoi nous nous intéressons aux propriétés statistiques des systèmes dynamiques.

Ce n'est pas parce qu'un système dynamique est simple et déterministe, qu'il ne peut pas présenter des comportements "imprévisibles". Il se peut en effet que, choisissant deux conditions initiales voisines, nous obtenions après un temps fini deux orbites ayant un comportement totalement différent. Cette observation motive l'étude des propriétés statistiques des systèmes dynamiques, puisqu'il n'est pas toujours possible de tirer une information intéressante de l'observation orbite par orbite (voir par exemple Collet [14], Viana [65]). L'étude qui va suivre porte sur le problème de l'estimation de la densité invariante de systèmes dynamiques. C'est-à-dire que nous allons étudier la mesure invariante de certains systèmes dynamiques, non pas directement, mais à travers sa densité, lorsqu'elle existe, par rapport à une mesure de référence. Les systèmes dynamiques étudiés dans cette thèse sont essentiellement uniformément dilatants, et leurs propriétés (en termes de covariances) se rapprochent de celles de suites faiblement dépendantes (propriété de décroissance des corrélations (12) définie page 24).

Dans la suite, T désigne une fonction d'un intervalle fermé $I \subset \mathbb{R}$ dans lui-même. Nous nous plaçons sur l'espace probabilisé $(I, \mathcal{B}(I), \lambda_I)$ où $\mathcal{B}(I)$ est la tribu des boréliens de I et λ_I la mesure de Lebesgue restreinte à I.

Définition 7 (Mesure invariante) *La mesure μ est $T-$invariante si, pour tout ensemble mesurable A, $\mu(T^{-1}A) = \mu(A)$.*

Ceci est encore équivalent à $\int f \circ T d\mu = \int f d\mu \ \forall f \in \mathbb{L}^1(d\mu)$.

Par exemple la mesure de Lebesgue est invariante par $T(x) = 2x \bmod 1$.

Dans cette thèse nous considérons une classe de fonctions T, notée $\mathcal{T}$, qui est décrite en détails dans Prieur [50]. Disons pour simplifier que la principale hypothèse est de supposer que T admet une unique mesure invariante μ_0 absolument continue par rapport à la mesure de Lebesgue. Notons f la densité de μ_0 par rapport à λ_I. Notre but est d'estimer f. Nous introduisons à cette fin la suite $(X_n)_{n\in\mathbb{N}}$ définie par :

$$X_0 \sim \mu_0 \text{ , et } X_n = T^n X_0 \text{ , } \forall\, n \geq 1. \tag{9}$$

Cette suite est stationnaire car X_0 suit la loi invariante μ_0.

Nous définissons maintenant un estimateur à noyau pour la densité f à partir de la suite $(X_n)_{n\in\mathbb{N}}$:

$$\hat{f}_{n,b_n}(x) = \frac{1}{nb_n} \sum_{k=0}^{n-1} K\left(\frac{x - X_k}{b_n}\right), \tag{10}$$

avec une suite $(b_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ et un noyau $K : \mathbb{R} \to \mathbb{R}$ à support compact D tels que :

$$b_n \underset{n\infty}{\longrightarrow} 0,\ \int_D K(t) d\lambda_I(t) = 1,\ 0 < \int_D K^2(t) d\lambda_I(t) < \infty. \tag{11}$$

Nous nous intéressons aux suites $(X_n)_{n\in\mathbb{N}}$ définies comme ci-dessus qui satisfont de plus une hypothèse de décroissance des corrélations :

$$\exists\ \kappa\ ,\ \exists\ (d_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}},\ \sum_{k=0}^{+\infty} d_k < \infty \text{ tels que } \forall\ n \geq 0, \forall\ h,\ k \in \mathcal{VB},$$
$$|\mathrm{Cov}\,(h(X_0),k(X_n))| \leq \kappa\ \|k\|_1 \|h\|_{\mathcal{VB}}\ d_n. \tag{12}$$

L'ensemble $\mathcal{VB}$ est l'ensemble des fonctions $h : I \to \mathbb{R}$, à variation bornée et de norme $\mathbb{L}^1$ finie. La norme standard sur $\mathbb{L}^1$ est notée $\|.\|_1$ et $\|h\|_{\mathcal{VB}} = \mathcal{V}(h) + \|h\|_1$, où $\mathcal{V}(h)$ désigne la variation de h. Nous renvoyons à Prieur [50] pour une définition plus précise de $\mathcal{VB}$. Dans la suite, nous prendrons $K \in \mathcal{VB}$.

Cette hypothèse de décroissance des corrélations nous rappelle vaguement la notion de dépendance faible. Le système oublie le passé lorsque le temps s'écoule. Certains parlent de mélangeance du système dynamique, à ne pas confondre avec la notion de mélange présentée page 12. En fait, $(X_n)_{n\in\mathbb{N}}$ ne peut être $\alpha-$mélangeante (et donc ne peut être $\beta-$, $\rho-$, ou $\Phi-$mélangeante en raison des implications de la page 12). En effet, comme $\sigma(X_k) \subset \sigma(X_{k-1})$ pour tout k dans $\mathbb{N}^*$, il est possible de montrer que

$$\alpha(n) \geq \sup_{k\geq 0} \alpha\left(\sigma(X_k),\, \sigma(X_k)\right) = \frac{1}{4}.$$

La suite $(X_n)_{n\in\mathbb{N}}$ ne vérifie pas non plus en général les conditions de dépendance faible données par les définitions 2 et 3 page 13. Nous ne pouvons donc pas utiliser directement les résultats de Coulon-Prieur & Doukhan [16] (chapitre 1). L'objet de ce chapitre est donc d'adapter la méthode utilisée au chapitre 1 pour estimer la densité marginale d'une suite stationnaire de variables aléatoires faiblement dépendantes.

Dans un premier temps, nous proposons une étude directe de la convergence en moyenne quadratique de notre estimateur (lemme 3 ci-dessous). L'ordre de grandeur que nous obtenons pour la vitesse est plus fin que celui donné par Bosq & Guégan [8] ou plus récemment par Maës [42]. Notre ordre nous permet ensuite d'obtenir un théorème limite centrale (théorème 6 ci-dessous) ainsi que des variantes de ce théorème limite centrale prenant en compte le biais de notre estimateur (voir Prieur [50]). Nous montrons ces théorèmes en adaptant la méthode qui nous a permis aux chapitres 1, 2 et 3 d'obtenir des théorèmes limites centrales en dépendance faible. Cette méthode est décrite dans cette introduction à la page 16.

Lemme 3 (Convergence en moyenne quadratique) *Soient $T \in \mathcal{T}$ et μ_0 la mesure $T-$invariante absolument continue par rapport à la mesure de Lebesgue. Soit S le support de μ_0. Pour presque tout $x \in S$, nous avons :*

$$\mathrm{Var}\,\left(\hat{f}_{n,b_n}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)ds + o(1)\right),$$

où $\hat{f}_{n,b_n}(x)$ est défini par (10) et (11).

Théorème 6 (TLC pour l'estimateur de la densité invariante) *Soient $T \in \mathcal{T}$ et μ_0 la mesure $T-$invariante absolument continue par rapport à la mesure de Lebesgue. Soit S le support de μ_0. On définit*

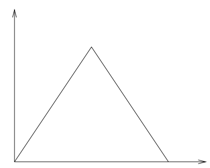
$$Y_n(x) := \sqrt{nb_n}\left(\hat{f}_{n,b_n}(x) - \mathbf{E}\hat{f}_{n,b_n}(x)\right),$$

où $\hat{f}_{n,b_n}(x)$ est défini par (10) et (11). Alors, si $nb_n \xrightarrow[n\,\infty]{} +\infty$, nous avons pour presque tout $(x_1,\ldots,x_l) \in S^l$:

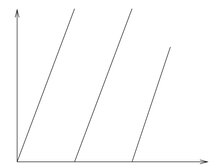
$$(Y_n(x_1),\ldots,Y_n(x_l)) \xrightarrow[n\,\infty]{\mathcal{L}} \mathcal{N}(0,\Sigma_l),$$

avec $\Sigma_l := \left(\int_D K^2(s)ds\right) \times diag\,(f(x_1),\ldots,f(x_l))$, où $diag\,(u_1,\ldots,u_l)$ désigne la matrice diagonale carrée de taille l ayant les u_i pour termes diagonaux.

Parmi les fonctions T satisfaisant nos hypothèses, citons les fonctions “tentes” et les fonctions $r-$adiques (voir figures ci-dessous) qui sont toutes deux des fonctions de Lasota-Yorke (définies dans Prieur [50] où les exemples sont également étudiés plus en détails).



Fonction "tente"



Fonction $r-$adique

Ces fonctions dilatantes par morceaux sont utiles lorsqu'on étudie certains systèmes lisses non-uniformément hyperboliques comme les fonctions unimodales par exemple.

Les résultats obtenus (lemme 3 et théorème 6) sont intéressants dans la mesure où ils nous amènent à rapprocher les notions de dépendance faible (définitions 2 et 3) et de systèmes dynamiques, tout en soulignant les différences entre ces deux notions. En effet, nous obtenons bien un théorème limite centrale du même type que le résultat de convergence (5) (voir page 16) pour l'estimateur de la densité marginale d'une suite stationnaire faiblement dépendante, mais pour celà, nous sommes amenés à restreindre la classe de systèmes dynamiques que nous considérons. Nous pouvons expliquer pourquoi notre méthode (qui est une variante de la méthode utilisée pour montrer le théorème limite centrale triangulaire dans Coulon-Prieur & Doukhan [16]) nous donne accès à une classe relativement restreinte de systèmes dynamiques. Notre hypothèse de décroissance des corrélations est assez forte. Elle porte en effet sur le contrôle des covariances $\mathrm{Cov}\,(h(X_0),k(X_n))$ où h et k sont dans l'ensemble $\mathcal{VB}$ des fonctions à variation bornée et de norme $\mathbb{L}^1$ finie. Pour élargir la classe de systèmes dynamiques considérés, il faudrait réussir à affaiblir notre hypothèse de décroissance de corrélation. Il faudrait pouvoir supposer :

$\exists\ \kappa\ ,\ \exists\ (d_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}},\ \sum_{k=0}^{+\infty} d_k < \infty$ tels que $\forall\ n \geq 0,\ \forall\ h,\ k \in \mathcal{E}$,

$$|\mathrm{Cov}\,(h(X_0),k(X_n))| \leq \kappa\ \|k\|_1 \|h\|_{\mathcal{VB}}\ d_n,$$

où $\mathcal{E}$ est un sous-ensemble de $\mathbb{L}^1$ plus petit que $\mathcal{VB}$.

En raison de l'absence de densités jointes pour les couples (X_0,X_n), nous ne pouvons considérer l'ensemble des fonctions lipschitziennes (considéré en dépendance faible, voir définitions 2 et 3) comme ensemble de référence $\mathcal{E}$ dans notre hypothèse de décroissance des corrélations. En effet, lorsque nous majorons les quantités $\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right),K\left(\frac{x-X_k}{b_n}\right)\right)\right|$ en utilisant la relation $\mathrm{Cov}(X,Y) = \mathbf{E}XY - \mathbf{E}X\mathbf{E}Y$, nous ne pouvons plus mettre en facteur b_n^2. Le facteur que nous mettons en évidence est de la forme $b_n\ \varepsilon(n,k)$ où $\varepsilon(n,k) \xrightarrow[n\ \infty]{} 0$, mais de façon non uniforme en k (voir Prieur [50]). Comme au chapitre 1, nous majorons aussi ces covariances en utilisant les propriétés de dépendance de notre suite. Pour compenser la perte " $b_n\ \varepsilon_{n,k}$ au lieu de b_n^2 ", nous devons obtenir cette fois-ci des majorants qui convergent vers 0 plus vite que les majorants obtenus au chapitre 1 dans le cadre de la dépendance faible. C'est pourquoi la classe des fonctions lipschitziennes qui convenait au chapitre 1 ne convient plus ici. Pour des raisons semblables, nous ne pouvons pas non plus prendre l'ensemble des fonctions $\alpha-$höldériennes à la place de l'ensemble $\mathcal{VB}$.

Rappelons que la méthode que nous utilisons est une variante de la méthode utilisée en dépendance faible (Coulon-Prieur & Doukhan [16]). Nous cherchons donc un majorant assez fin des termes

$$\left|\Delta_k^{(1)}(f)\right| = \left|\mathbf{E}f(S_{k-1}+Z_k+T_k) - \mathbf{E}f(S_{k-1}+T_k) - \frac{v_k}{2}\mathbf{E}f''(S_{k-1}+T_k)\right|$$

où les Z_k dépendent ici de n (comme dans Coulon-Prieur & Doukhan [16], voir chapitre 1) :

$$Z_k = Z_{k,n} = \frac{1}{\sqrt{nb_n}}K\left(\frac{x-X_k}{b_n}\right).$$

Le point essentiel dans les calculs est alors la majoration des termes de la forme :

$$\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right|$$

Nous avons vu ci-dessus que

$$\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \leq C_1\ b_n\ \varepsilon(n,k), \tag{13}$$

où C_1 est une constante et où pour chaque k, $\varepsilon(n,k) \xrightarrow[n\ \infty]{} 0$.

Avec l'hypothèse de décroissance des corrélations (12) nous avons de plus :

$$\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \leq C_2\ b_n\ d_k, \tag{14}$$

où C_2 est une constante.

Nous montrons dans Prieur [49, 50] (voir chapitre 4) que les majorations (13) et (14) sont suffisantes pour obtenir le théorème limite centrale 6 ci-dessus.

Nous avons obtenu un théorème de convergence pour un estimateur $\hat{f}_{n,b_n}(x)$ que nous ne pouvons pas simuler sans connaître la mesure invariante μ_0. Pour les simulations, il va donc falloir considérer un nouvel estimateur de f qui se construit cette fois-ci à partir d'une nouvelle suite $(X'_n)_{n\in\mathbb{N}}$ non stationnaire (c'est-à-dire que X'_0 doit pouvoir suivre une loi quelconque qui peut être différente de la loi invariante μ_0). Nous introduisons dans un premier temps un opérateur très utile appelé opérateur de transfert (ou encore de Perron-Frobenius) associé à T. Cet opérateur est donné dans notre cadre par la formule explicite

$$\mathcal{L}_T w(x) = \sum_{y,\ Ty=x} \frac{w(y)}{|T'(y)|}\ ,\ \forall\ w \in \mathcal{VB}.$$

Au chapitre 4, le théorème 6.1. décrit les propriétés de $\mathcal{L}_T$ dans le cas où T est une fonction de type Lasota-Yorke (voir Liverani [36], Collet [14] ou Viana [65] par exemple). En résumé, si $p \in \mathcal{VB}$, $\mathcal{L}_T^n p$ se rapproche de la densité invariante f exponentiellement vite. Ceci nous permet d'estimer f à partir, non plus de $(X_n)_{n\in\mathbb{N}}$, mais de $(X_n^{'})$ définie par :

$$X_0^{'} \sim p(t)dt,\ X_n^{'} = T^n X_0^{'}\ \forall\ n \geq 1$$

où $p \in \mathcal{VB}$ est une fonction densité quelconque.

Nous obtenons alors un théorème limite centrale ainsi que des variantes de ce théorème limite centrale (théorèmes 6.2., 6.3. et 6.4. dans Prieur [50]) pour le nouvel estimateur

$$\hat{p}_{n,b_n}(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x-X_k^{'}}{b_n}\right). \tag{15}$$

Cet estimateur ne dépend plus que de paramètres connus. $X_0^{'}$ peut être simulée à partir de n'importe quelle fonction densité p à variation bornée. Ensuite, les $X_n^{'}$ sont simulés en appliquant à $X_0^{'}$ les itérées de la fonction T connue. En travaillant avec un grand nombre d'échantillons, nous pouvons visualiser avec des simulations la convergence de notre estimateur de la densité.

Chapitre 5 : Simulations numériques.

Dans ce chapitre, nous réalisons des simulations concernant les résultats du chapitre 4. Les notations sont les mêmes que celles du paragraphe précédent. Rappelons juste que l'objet du chapitre 4 est d'estimer la densité invariante, notée f, d'une dynamique T ayant de bonnes propriétés, en particulier celle de ne posséder qu'une seule mesure invariante μ_0 absolument continue par rapport à la mesure de Lebesgue. L'estimateur utilisé est l'estimateur à noyau $\hat{p}_{n,b_n}(x)$ décrit par l'expression (15) ci-dessus. Au chapitre 5, nous traçons des courbes approchées de la courbe théorique $y = f(x)$ en utilisant d'une part l'estimateur $\hat{p}_{n,b_n}(x)$, d'autre part le résultat de convergence en moyenne quadratique de $\hat{p}_{n,b_n}(x)$ suivant :

Lemme 4 (Convergence en moyenne quadratique de $\hat{p}_{n,b_n}(x)$) *Soient $T \in \mathcal{T}$ et μ_0 la mesure $T-$invariante absolument continue par rapport à la mesure de Lebesgue.*
Soit S le support de μ_0. Pour presque tout $x \in S$, nous avons :

$$\text{Var }(\hat{p}_{n,b_n}(x)) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)ds + o(1)\right). \tag{16}$$

Nous avons trouvé intéressant d'illustrer ce lemme par des simulations, car c'est lui qui nous a permis d'adapter la méthode de Lindeberg-Rio [54] pour montrer le résultat de convergence en loi pour notre estimateur $\hat{p}_{n,b_n}(x)$.

Expliquons dans un premier temps comment nous traçons les courbes approchées.

Nous écrivons un programme C qui permet de calculer, à partir d'un $m-$échantillon de loi $p(t)\,dt$, la variance empirique d'un vecteur

$$P_m^n(x) := \sqrt{nb_n}\left(\hat{p}_{n,b_n}^{(1)}(x), \ldots, \hat{p}_{n,b_n}^{(m)}(x)\right),$$

constitué de m copies indépendantes de $\sqrt{nb_n}\,\hat{p}_{n,b_n}(x)$.

Notons $\mathrm{Var}_{emp}\,(P^n_m(x))$ cette variance empirique. Nous savons, par la loi forte des grands nombres, qu'à n fixé,

$$\mathrm{Var}_{emp}\,(P^n_m(x)) \xrightarrow[m\infty]{p.s.} (nb_n)\,\mathrm{Var}\,(\hat{p}_{n,b_n}(x))\,. \tag{17}$$

Nous écrivons ensuite un programme, en MATLAB, nous permettant de tracer les courbes $y = \dfrac{1}{\int K^2(s)\,ds}\mathrm{Var}_{emp}\,(P^n_m(x))$. Grâce à (16) et (17), nous espérons que pour de grandes valeurs de m et de n, les courbes empiriques vont se rapprocher de la coube théorique $y = f(x)$.

Nos simulations ont été faites pour la fonction T unimodale : $T(x) = 4x(1-x)$, et pour le noyau $K_1(x) = \mathbf{1}_{[-1/2,1/2]}(x)$. Comme loi initiale nous avons pris la loi uniforme sur [0,1] dans un premier temps, puis la loi de densité $p_2(x) = \dfrac{2\,e^{-2x}}{1-e^{-2}}\,\mathbf{1}_{[0,1]}(x)$.

Nous avons considéré des fenêtres de la forme $b_n = n^{-a}$, $0 < a < 1$. Nos courbes convergent effectivement, mais la vitesse de convergence semble limitée par la vitesse de convergence de la fenêtre b_n vers 0. Plus $a = \dfrac{-\ln(b_n)}{\ln(n)}$ est proche de 1, plus la convergence est rapide.

Remarquons maintenant que pour $1/2 \leq a < 1$, nb_n tend vers l'infini, mais pas nb_n^2. Pourtant, pour ces valeurs de a, nous observons bien la convergence. Cette observation confirme que la vitesse de convergence en moyenne quadratique de $\hat{p}_{n,b_n}(x)$, en $\mathcal{O}\left(\dfrac{1}{nb_n^2}\right)$, obtenue par exemple dans [8], n'était pas optimale, et que, comme nous l'énonçons au lemme 4, $\mathrm{Var}\,(\hat{p}_{n,b_n}(x)) = \mathcal{O}\left(\dfrac{1}{nb_n}\right)$.

Notons enfin que nous avions initialement écrit le programme C, qui calcule $\mathrm{Var}_{emp}\,(P^n_m(x))$, et dont nous donnons la structure au chapitre 5, en MATLAB. Le programme mettait alors un temps très long (plusieurs heures) à faire les calculs pour des valeurs de m et de n dépassant 1000. C'est pourquoi nous avons choisi de programmer en C. Nous avons écrit le programme permettant de tracer les courbes à l'aide de MATLAB, qui est agréable à utiliser en ce qui concerne les graphes.

Dans cette thèse nous donnons donc dans un premier temps diverses applications statistiques de suites faiblement dépendantes (chapitres 1 à 3). Parmi les techniques que nous utilisons figurent plusieurs variantes d'une méthode, inspirée par Rio [53], dont nous avons présenté un schéma simplifié dans cette introduction. Dans un second temps, nous montrons que nous pouvons étendre l'utilisation de telles techniques au cas de certains systèmes dynamiques (chapitre 4). Enfin, au chapitre 5, nous donnons quelques simulations numériques relatives aux résultats théoriques du chapitre 4.

Chapitre 1

Théorème Limite Centrale Triangulaire et Applications Statistiques

Dans ce chapitre, nous présentons un article publié dans *Statistic & Probability Letters*, (Coulon-Prieur & Doukhan, 2000 [16]). Nous proposons un théorème de limite centrale triangulaire dans le cadre de la dépendance faible. Nous appliquons ensuite ce théorème à certains tableaux linéaires ainsi qu'au problème de l'estimation d'une densité par la méthode de noyau.

ELSEVIER

Statistics & Probability Letters 47 (2000) 61–68

STATISTICS & PROBABILITY LETTERS

www.elsevier.nl/locate/stapro

A triangular central limit theorem under a new weak dependence condition

Clémentine Coulon-Prieur, Paul Doukhan *

University Cergy Pontoise, UPRES A8088 CNRS, Mathématiques, Department of Economy, 33 Boulevard du Port, 95011 Cergy-Pontoise Cedex, France

Received November 1998; received in revised form May 1999

Abstract

We use a new weak dependence condition from Doukhan and Louhichi (Stoch. Process. Appl. 1999, 84, 313–342) to provide a central limit theorem for triangular arrays; this result applies for linear arrays (as in Peligrad and Utev, Ann. Probab. 1997, 25(1), 443–456) and standard kernel density estimates under weak dependence. This extends on strong mixing and includes non-mixing Markov processes and associated or Gaussian sequences. We use Lindeberg method in Rio (Probab. Theory Related Fields 1996, 104, 255–282).

MSC: 60E15; 60F05; 60F17; 60G10; 60G99

Keywords: Stationary sequences; Lindeberg theorem; Central limit theorem; Non-parametric estimation; *s*- and *a*-weakly dependent

1. Introduction

We use here a new weak dependence condition from Doukhan and Louhichi (1997,1999) in order to provide a central limit theorem (CLT in short) for triangular arrays; this result is used both for standard kernel density estimates and for general linear combinations of weakly dependent sequences, analogously to Peligrad and Utev (1997).

We work here under a fundamental *causality* assumption in order to be in position to use the Lindeberg method for dependent sequences developed by Rio (1995,1996).

Contrarily to previous authors, Rio does not use Bernstein blocks to prove a central limit theorem. The standard ways of proving such limit theorems for dependent random sequences are, after a decomposition into Bernstein blocks, to make use of the standard techniques for i.i.d. sequences; Lindeberg and Stein are the two techniques usually used (see references e.g. in Rio, 1996).

* Corresponding author.
E-mail address: doukhan@u-cergy.fr (P. Doukhan)

PII: S0167-7152(99)00138-8

For example, Bernstein blocks are used with Lindeberg method in Doukhan and Louhichi (1999): the results presented in this note clearly improve on CLTs stated in this basic paper under a more general non-causal frame.

A CLT is obtained for the kernel density estimates. The result is analogue to the one obtained for i.i.d. or mixing samples (see Rosenblatt, 1991; Robinson, 1983). Under our frame, it has to be noticed that we require the same assumptions here for such a CLT than to obtain minimax second order properties of the kernel estimates (see Doukhan and Louhichi, 1997).

Finally, the linear triangular CLT requires the same assumption as to prove the standard $\sqrt{n}$-CLT.

We now introduce our dependence frame; it is a variation on the definition in Doukhan and Louhichi (1999). Assume that, for convenient functions h and k,

$$\mathrm{Cov}(h(\text{'past'}), k(\text{'future'})),$$

converge to 0 as the distance between the 'past' and the 'future' converges to infinity. Here 'past' and 'future' refer to the values of some time series of interest. Asymptotically, this means that independence holds if we use a *determining* function class.

More precisely, E being some Euclidean space $\mathbb{R}^d$ endowed with its Euclidean norm $\|\,.\,\|$, we shall consider a sequence of E-valued random variables $(\xi_n)_{n\in\mathbb{N}}$. We define $\mathbb{L}^\infty$ as the set of measurable and bounded numerical functions on some space $\mathbb{R}^u$ and its norm is classically written $\|\bullet\|_\infty$. Moreover, let $u\in\mathbb{N}^*$ be a positive integer we endow the set $F=E^u$ with the norm $\|(x_1,\ldots,x_u)\|_F=\|x_1\|+\cdots+\|x_u\|$. Let now $h:F=E^u\to\mathbb{R}$ be a numerical function on F, we set

$$\mathrm{Lip}(h)=\sup_{x\neq y}\frac{|h(x)-h(y)|}{\|x-y\|_F},$$

the Lipschitz modulus of h. Define

$$\mathscr{L}=\bigcup_{u=1}^{\infty}\{h\in\mathbb{L}^\infty(\mathbb{R}^u,\mathbb{R});\ \|h\|_\infty\leqslant 1,\ \mathrm{Lip}(h)<\infty\}. \tag{1}$$

Definition 1. The sequence $(\xi_n)_{n\in\mathbb{N}}$ is s-weakly (resp. a-weakly) dependent, if for some sequence $\theta=(\theta_r)_{r\in\mathbb{N}}$ decreasing to zero at infinity and any $(u+2)$-tuple $(i_1,\ldots,i_u,j_1,j_2)$ with $i_1\leqslant\cdots\leqslant i_u<i_u+r\leqslant j_1\leqslant j_2$, and $h\in\mathbb{L}^\infty$ satisfies $\|h\|_\infty\leqslant 1$ and $k\in\mathscr{L}$,

$$|\mathrm{Cov}(h(\xi_{i_1},\ldots,\xi_{i_u}),k(\xi_{j_1},\xi_{j_2}))|\leqslant\mathrm{Lip}(k)\theta_r \tag{2}$$

and, respectively, for $h,k\in\mathscr{L}$

$$|\mathrm{Cov}(h(\xi_{i_1},\ldots,\xi_{i_u}),k(\xi_{j_1},\xi_{j_2}))|\leqslant\mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r. \tag{3}$$

Weak dependence conditions are shown to hold in, either causal or noncausal frames in Doukhan and Louhichi (1997a). For this, consider also v-tuples $(j_1,\ldots,j_v)$ with $i_1\leqslant\cdots\leqslant i_u<i_u+r\leqslant j_1\leqslant\cdots\leqslant j_v$, then such weak dependence conditions are defined for functions h and k defined on E^u and E^v, respectively, through inequalities

$$|\mathrm{Cov}(h(\xi_{i_1},\ldots,\xi_{i_u}),k(\xi_{j_1},\ldots,\xi_{j_v}))|\leqslant v\mathrm{Lip}(k)\theta_r \tag{4}$$

if $k\in\mathscr{L}$ and $\|h\|_\infty\leqslant 1$ or

$$|\mathrm{Cov}(h(\xi_{i_1},\ldots,\xi_{i_u}),k(\xi_{j_1},\ldots,\xi_{j_v}))|\leqslant\min(u,v)\mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r \tag{5}$$

if $h,k\in\mathscr{L}$. Strong mixing defined by Rosenblatt (see e.g., Doukhan, 1994) is a variation of such definitions (see Doukhan and Louhichi, 1999); however mixing refers to σ-algebras rather than to random variables. For completeness, we now recall some examples adapted from Doukhan and Louhichi (1999), where noncausal cases are also considered:

Definition 2. Let $(\eta_n)_{n\in\mathbb{Z}}$ be a stationary sequence of real valued r.v's and F be a measurable function defined on $\mathbb{R}^{\mathbb{N}}$. The stationary sequence $(\xi_n)_{n\in\mathbb{Z}}$ defined by $\xi_n = F(\eta_n, \eta_{n-1}, \eta_{n-2}, \ldots)$ is called a causal Bernoulli shift. We denote $(\delta_k)_{k\in\mathbb{N}}$ any nonnegative sequence such that

$$\mathbb{E}|F(\eta_0, \eta_{-1}, \eta_{-2}, \ldots) - F(\eta_0, \ldots, \eta_{-r}, 0, 0, \ldots)| \leqslant \delta_r.$$

Causal shifts with i.i.d. innovations $(\eta_k)_{k\in\mathbb{Z}}$ satisfy (4) with $\theta_r = 2\delta_r$ (see Rio, 1996). Examples of such situations follow:

- The example of the nonmixing stationary Markov chain with i.i.d. Binomial innovations $(\eta_t)_{t\in\mathbb{Z}}$, $\xi_n = (\xi_{n-1} + \eta_n)/2$ satisfies $\delta_r = \mathcal{O}(2^{-r})$; its marginal distribution is uniform on $[0, 1]$.
- The real-valued functional autoregressive model $\xi_t = r(\xi_{t-1}) + \eta_t$, $r : \mathbb{R} \to \mathbb{R}$. If $|r(u) - r(u')| \leqslant c|u - u'|$ for some $0 \leqslant c < 1$ and for all $u, u' \in \mathbb{R}$, and if the i.i.d. innovation process $(\eta_t)_{t\in\mathbb{Z}}$ satisfies $\mathbb{E}\|\xi_0\| < \infty$, then s-dependence holds with $\theta_r = \delta_r = Cc^r$ for some constant $C > 0$.
- Chaotic expansions associated with the discrete chaos generated by the sequence $(\eta_t)_{t\in\mathbb{Z}}$. In a condensed formulation we write, $F(x) = \sum_{k=0}^{\infty} F_k(x)$, $x \in \mathbb{R}^{\mathbb{N}}$ for

$$F_k(x) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_k=0}^{\infty} a^{(k)}_{j_1,\ldots,j_k} x_{j_1} x_{j_2} \ldots x_{j_k}, \quad k \geqslant 1,$$

  where $F_k(x)$ denotes the kth-order chaos contribution and $F_0(x) = a_0^{(0)}$ is only a centering constant. In short we write in the vectorial notation, $F_k(x) = \sum_{j\in\mathbb{N}^k} a_j^{(k)} x_j$. Processes associated with a finite number of chaos (i.e. $F_k \equiv 0$ if $k > k_0$ for some $k_0 \in \mathbb{N}$) are also called *Volterra processes*. A simple and general condition for $\mathbb{L}^1$-convergence of this expansion, still written in a condensed notation, is $\sum_{k=0}^{\infty}\{\sum_{j\in\mathbb{N}^k} |a_j^{(k)}| \mathbb{E}|\xi_0|^k\} < \infty$. This condition allows to define the distribution of such shift processes. A suitable bound for δ_r is then

$$\delta_r = \sum_{k=0}^{\infty}\left\{\sum_{j\in\mathbb{N}^k;\, \|j\|_\infty > r} \left|a_j^{(k)}\right| \mathbb{E}|\xi_0|^k\right\} < \infty.$$

- For example, linear processes $\xi_n = \sum_{k=0}^{\infty} a_k \eta_{n-k}$ which include ARMA models are those with $F_k(x) \equiv 0$ for all $k > 1$. A first choice is $\delta_r = \mathbb{E}|\eta_0| \sum_{k>r} |a_k|$ for the linear process with i.i.d. innovations such that $\mathbb{E}|\eta_0| < \infty$.
  For centered and $\mathbb{L}^2$ innovations, another choice is $\delta_r = \sqrt{\mathbb{E}|\eta_0|^2 \sum_{k>r} |a_k|^2}$.
- The simple bilinear process with the recurrence equation $\xi_t = a\xi_{t-1} + b\xi_{t-1}\eta_{t-1} + \eta_t$. Such processes are associated with the chaotic representation in

$$F(x) = \sum_{j=1}^{\infty} x_j \prod_{s=0}^{j-1} (a + bx_s), \quad x \in \mathbb{R}^{\mathbb{Z}}.$$

  If $c = \mathbb{E}|a + b\xi_0| < 1$ then $\delta_r = \theta_r = c^r(r + 1)/(c - 1)$ has a geometric decay rate.

Definition 3 (Esary et al., 1967)**.** The sequence $(\xi_n)_{n\in\mathbb{Z}}$ is associated if for all coordinatewise increasing real-valued functions h and k on $\mathbb{R}^A$ $(A \subset \mathbb{Z})$, $\mathrm{Cov}(h(\xi_A), k(\xi_A)) \geqslant 0$, if $E[h^2(\xi_A) + k^2(\xi_A)] < \infty$.

Associated sequences satisfy (5) with $\theta_r = \sup_{t\in\mathbb{Z}} \sum_{k\geqslant r} |\mathrm{Cov}(\xi_t, \xi_{t+k})|$. Absolute values are useless here, but Gaussian sequences also satisfy this condition with them. Hence combinations of such independent processes yield examples of weak dependent sequences which are neither Gaussian nor associated.

The paper is organized as follows. Our main results are stated with their applications in Section 2, and Section 3 is devoted to prove the main results.

2. Main results

This paper is concerned with triangular arrays $(X_{n,k})_{k=1,\dots,k_n}$ for $n=1,2,\dots$ defined through an E-valued weakly dependent sequence $(\xi_n)_{n\in\mathbb{N}}$ by numerical Lipschitz functions $g_{n,k}$ defined on E for $k=1,\dots,k_n$ and $n=1,2,\dots$, we assume that the sequence of integers k_n increases to infinity with n. We set

$$X_{n,k}=g_{n,k}(\xi_k) \quad \text{and} \quad S_n=X_{n,1}+\cdots+X_{n,k_n}.$$

We assume in the following that $\mathbb{E}X_{n,k}=0$. Now let $S_{k,n}=X_{n,1}+\cdots+X_{n,k}$ for $1\leqslant k\leqslant k_n$, we also suppose that the exist constants $\sigma,\alpha>0$ such that

$$\lim_{n\to\infty}\operatorname{Var}S_n=\sigma^2>0 \quad \text{and} \quad v_{k,n}=\operatorname{Var}S_{k,n}-\operatorname{Var}S_{k-1,n}\geqslant\frac{\alpha}{n}, \tag{6}$$

for each $k\in\{1,\dots,k_n\}$ and for any integer n.

We shall also set

$$\begin{aligned}\lambda_n&=\sup_{1\leqslant k\leqslant k_n}\operatorname{Lip}(g_{k,n}), \quad M_n=\sup_{1\leqslant k\leqslant k_n}\|g_{k,n}\|_\infty,\\ \delta_n&=\sup_{1\leqslant k\leqslant k_n}\mathbb{E}|X_{k,n}| \quad \text{and} \quad \Delta_n=\sup_{1\leqslant k\neq l\leqslant k_n}\mathbb{E}|X_{k,n}X_{l,n}|.\end{aligned} \tag{7}$$

We are now in a position to state the two main results of this paper.

Theorem 1. *Assume that the E-valued sequence satisfies the s-weak dependence condition and the triangular array $(X_{n,k})_{1\leqslant k\leqslant k_n}$ defined as before satisfies assumption* (6), *then if*

$$(k_nM_n+k_n^{2/3})M_n\delta_n\to 0, \quad k_nM_n\sum_{p=1}^{k_n}\min(\lambda_n\theta_p,\Delta_n)\to 0, \quad k_n\sum_{p=1}^{k_n}\min(M_n\lambda_n\theta_p,\Delta_n)\to 0,$$

as $n\to\infty$, we obtain

$$S_n\to_{n\to\infty}\mathcal{N}(0,\sigma^2) \quad \textit{in distribution.}$$

Theorem 2. *Assume that the E-valued sequence satisfies the a-weak dependence condition and the triangular array $(X_{n,k})_{1\leqslant k\leqslant k_n}$ defined as before satisfies assumption* (6), *then if*

$$(k_nM_n+k_n^{2/3})M_n\delta_n\to 0, \quad k_nM_n\sum_{p=1}^{k_n}\min(\lambda_n^2\theta_p,\Delta_n)\to 0, \quad k_n\sum_{p=1}^{k_n}\min(\lambda_n^2\theta_p,\Delta_n)\to 0,$$

as $n\to\infty$, we obtain

$$S_n\to_{n\to\infty}\mathcal{N}(0,\sigma^2) \quad \textit{in distribution.}$$

Applications of those results given in Section 3 will prove that they lead to real improvement of previous results. They also lead to completely new results extending e.g. in Peligrad and Utev (1997).

2.1. Linear triangular arrays

The random variables ξ_k are supposed to be uniformly bounded, real valued and centered at expectation and here

$$g_{n,k}(x)=a_{n,k}x,$$

hence setting $b_n=\sup_{1\leqslant k\leqslant k_n}|a_{n,k}|$ we obtain, $\lambda_n\leqslant Cb_n$, $M_n\leqslant Cb_n$, $\delta_n\leqslant Cb_n$, $\Delta_n\leqslant Cb_n^2$, for some constant $C>0$. We deduce the following result:

Corollary 1. *Assume that the $\mathbb{R}$-valued sequence satisfies either the s-weak dependence condition or the a-weak dependence condition and assumptions* (6). *If moreover,* $k_n b_n^2 \to 0$, *as* $n \to \infty$ *and* $\sum_{p=1}^{\infty} \theta_p < \infty$ *then*

$$S_n \to \mathcal{N}(0, \sigma^2) \quad \textit{in distribution, as } n \to \infty.$$

Notice first that as in Peligrad and Utev (1997), we need $b_n \to 0$. Moreover, those authors also assume the first condition in (6); the second one is however not always satisfied (as for the example of regression with a fixed design described without proof in the latter paper).

The proof also yields the (standard) CLT $n^{-1/2} \sum_1^n \xi_k \to \mathcal{N}(0, \sigma^2)$ for a stationary and bounded weakly dependent sequence with $\sigma^2 > 0$ and $\sum_{p=0}^{\infty} \theta_p < \infty$.

2.2. Density estimation

Let u be some numerical function with integral 1, Lipschitzian and rapidly convergent to 0 at infinity, for simplicity, we assume here that it is compactly supported; one classically (see e.g. Rosenblatt, 1991) defines kernel density estimates for the marginal density of the process (ξ_n) by setting for some fixed sequence (h_n) of positive real numbers such that $h \to 0$ (for clarity we write h instead h_n) and $nh \to 0$ as $n \to \infty$,

$$\hat{f}(x) = f_{n,h_n}(x) \quad \text{with} \quad f_{n,h}(x) = \frac{1}{nh} \sum_{k=1}^{n} u\left(\frac{x - \xi_k}{h}\right).$$

We assume from now on, that the marginal density of X_n exists and we denote it f. The study of the bias of this estimate is purely analytical and does not depend on the dependence properties of the sequence (ξ_n) as it is noticed in Rosenblatt (1991). We thus restrict our attention to the centered estimation process, written as

$$Z_n(x) := Z_{n,h_n}(x) = \sqrt{nh_n}(\hat{f}(x) - \mathbb{E}\hat{f}(x)).$$

Let $x_1, \ldots, x_l \in \mathbb{R}$, if one wants to know the asymptotic behaviour in distribution of the vector $(Z_n(x_1), \ldots, Z_n(x_l))$, it is sufficient to use the previous theorems with $k_n = n$ and, $X_{n,k} = (1/\sqrt{nh}) \sum_{j=1}^{l} s_j u(\xi_k - x_j/h)$, for arbitrary numbers $s_1, \ldots, s_l \in \mathbb{R}$, to see that this example enters the frame of a general triangular array.

Corollary 2. *Assume that the previous s-weak dependence* (*resp. a-*) *condition holds for the stationary sequence* $(\xi_n)_{n \in \mathbb{N}}$ *with for some positive* $a < \frac{1}{3}$ (*resp.* $a < \frac{1}{4}$) $\sum_{p=1}^{\infty} \theta_p^a < \infty$, *then the finite-dimensional marginals* $(Y_n(x_1), \ldots, Y_n(x_l))$, *of the process* $Y_n(x) \equiv Z_n(x)/\sqrt{f(x) \int_{-\infty}^{\infty} u^2(t)\,\mathrm{d}t}$ *converge in distribution to an* $\mathcal{N}(0, I_l)$ *random variable if we assume moreover that* $f(x_1) \neq 0, \ldots, f(x_l) \neq 0$, *that* ξ_0*'s marginal admits a continuous marginal density* f *and the marginal densities* $f_k(x, y)$ *of the bivariate random variables* (ξ_0, ξ_k) *exist for any* $k > 0$ *and satisfy* $\sup_{k>0} \sup_{(x,y) \in \mathbb{R}^2} f_k(x, y) < \infty$.

Remarks

- These conditions hold, respectively, if $\theta_r = \mathcal{O}(r^{-a})$ for some $a > 3$ (resp. $a > 4$).
- This result improves on a previous result in Doukhan and Louhichi (1997) (see also Isha and Prakasa, 1995 or Chanda and Ruymgart, 1990), e.g. under association we need $\mathrm{Cov}(\xi_0, \xi_r) = \mathcal{O}(r^{-a})$ for $a > 5$ while the previous result was obtained assuming $a > 12$ and for causal shifts it was needed that $\theta_r = \mathcal{O}(r^{-a})$ for some $a > \max\{9, \frac{3}{2}(1 + \delta^{-1})\}$ if $h \sim n^{-\delta}$. In both cases, the result seems to be new.
- For strongly mixing sequences, the condition $\alpha_n = \mathcal{O}(n^{-a})$ for $a > 1$ ensures this CLT as proved by Robinson (1983) (and also Ango Nze and Doukhan, 1996); this assumptions is of a different nature, e.g. linear processes satisfy the mixing conditions (under additional regularity conditions; see Doukhan, 1994, Chapter 2.3) the decay rate of the coefficients are there more restrictive.

Up to constants, we obtain $\lambda_n = 1/(h\sqrt{nh})$, $M_n = \lambda_n h$, $\Delta_n = h/n$, and $\delta_n = h/\sqrt{nh}$. Conditions (6) follow by standard arguments from assumptions on marginal densities and from $\sum_{p=0}^{\infty} \theta_p^a < \infty$ for $3a < 1$ (resp. for $4a < 1$) (see e.g. Doukhan and Louhichi, 1997 or Ango Nze and Doukhan, 1996). We now split the proof in the cases considered.

2.2.1. s-dependence

Now $nM_n^2\delta_n + n^{2/3}M_n\delta_n = 1/(\sqrt{nh}) + (4h^{1/3}/(nh)^{1/3})$, so that we just need $nh \to_{n\to\infty} \infty$. Let $a \in]0,1]$,

$$nM_n \sum_{p=1}^{\infty} \min(\lambda_n\theta_p, \Delta_n) \leqslant \sum_{p=1}^{\infty} \left(\frac{\theta_p}{h^2}\right)^a \left(\sqrt{\frac{h}{n}}\right)^{1-a} = \frac{1}{(nh)^{(1-a)/2}} h^{1-3a} \sum_{p=1}^{\infty} \theta_p^a.$$

We need that for some $a \leqslant \frac{1}{3}$, $\sum_{p=1}^{\infty} \theta_p^a < \infty$.

$$n \sum \min(\Delta_n, M_n\lambda_n\theta_p) \leqslant \sum_{p=1}^{\infty} \min\left(h, \frac{n\theta_p}{nh^2}\right) \leqslant \sum_{p=1}^{\infty} \theta_p^a h^{-2a} h^{1-3a}.$$

If for some $0 < a < \frac{1}{3}$, $\sum_{p=1}^{\infty} \theta_p^a < \infty$, the previous expression tends to 0 when $n \to \infty$.

2.2.2. a-dependence

Now, $nM_n^2\delta_n = 1/\sqrt{nh} \to 0$ as $n \to \infty$. We also consider

$$nM_n \sum_{p=1}^{\infty} \min(\lambda_n^2\theta_p, \Delta_n) = \sum_{p=1}^{\infty} \min\left(\sqrt{\frac{h}{n}}, \frac{\theta_p}{n^{1/2}h^{7/2}}\right) \leqslant \frac{1}{\sqrt{nh}} h^{1-4a} \sum_{p=1}^{\infty} \theta_p^a \quad \text{for } a \in]0,1].$$

So we need that there exists $a \leqslant \frac{1}{4}$ with $\sum_{p=1}^{\infty} \theta_p^a < \infty$.

$$n \sum_{p=1}^{\infty} \min(\lambda_n^2\theta_p, \Delta_n) = \sqrt{nh} \sum_{p=1}^{\infty} \min\left(\sqrt{\frac{h}{n}}, \frac{\theta_p}{n^{1/2}h^{7/2}}\right) \leqslant h^{1-4a} \sum_{p=1}^{\infty} \theta_p^a.$$

If there is some positive $a < \frac{1}{4}$ with $\sum_{p=1}^{\infty} \theta_p^a < \infty$, we conclude the proof. □

3. Proofs

3.1. *Proof of Theorem* 1

The proof is a variation of Lindeberg method after Rio (1995). Consider a bounded thrice differentiable function $h : \mathbb{R} \to \mathbb{R}$ with continuous and bounded derivatives. Set $C_j = \|h^{(j)}\|_\infty$, for $j = 0, 1, 2, 3$. Also consider $\sigma_n^2 = \operatorname{Var} S_n$. Set, for some standard Gaussian r.v. η, $\Delta_n(h) = \mathbb{E}(h(S_n) - h(\sigma_n\eta))$. The theorem will follow from assumptions (6), if we prove that

$$\lim_{n\to\infty} \Delta_n(h) = 0.$$

Recall that $v_{k,n} > 0$ for each k and set $Y_{n,k} \sim \mathcal{N}(0, v_{n,k})$. The sequence $(Y_{n,k})_{1\leqslant k\leqslant k_n, n\geqslant 1}$ is assumed to be independent and independent of the sequence $(\xi_k)_{k\in\mathbb{N}}$, and set, if $1 \leqslant k \leqslant k_n$, $T_{k,n} = \sum_{j=k+1}^{k_n} Y_{n,j}$, empty sums are, as usual, set equal to 0. We are in position to use Rio's decomposition

$$\Delta_n(h) = \sum_{k=2}^{k_n} \Delta_{k,n}(h) \tag{8}$$

with $\Delta_{k,n}(h) = \mathbb{E}(h(S_{k-1,n} + X_{n,k} + T_{k,n}) - h(S_{k-1,n} + Y_{n,k} + T_{k,n}))$.

The function $x \to h_{k,n}(x) = \mathbb{E}\, h(x+T_{k,n})$ has the same derivability properties as h, e.g. for $0 \leqslant j \leqslant 3$, $\|h_{k,n}^{(j)}\| \leqslant C_j$; now, from independence of the Gaussian r.v. $T_{k,n}$ and the process $(\xi_n)_{n\in\mathbb{N}}$, we write $\Delta_{k,n}(h) = \Delta_{k,n}^{(1)}(h) - \Delta_{k,n}^{(2)}(h)$, with

$$\Delta_{k,n}^{(1)}(h) = \mathbb{E}\, h_{k,n}(S_{k-1,n} + X_{n,k}) - \mathbb{E}\, h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\, h''_{k,n}(S_{k-1,n}),$$

$$\Delta_{k,n}^{(2)}(h) = \mathbb{E}\, h_{k,n}(S_{k-1,n} + Y_{n,k}) - \mathbb{E}\, h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\, h''_{k,n}(S_{k-1,n}).$$

Bound of $\Delta_{k,n}^{(2)}(h)$. Using Taylor expansion yields for some random variable valued $\rho_{n,k} \in (0,1)$: $\Delta_{k,n}^{(2)}(h) = \mathbb{E}\, h'_{k,n}(S_{k-1,n})Y_{n,k} + \frac{1}{2}\mathbb{E}\, h''_{k,n}(S_{k-1,n})(Y_{n,k}^2 - v_{n,k}) + \frac{1}{6}\mathbb{E}\, h_{k,n}^{(3)}(S_{k-1,n} + \rho_{n,k}Y_{n,k})Y_{n,k}^3$. From the independence of the Gaussian sequence, $|\Delta_{k,n}^{(2)}(h)| \leqslant (C_3/6)\mathbb{E}|Y_{n,k}|^3$, hence $|\Delta_{k,n}^{(2)}(h)| \leqslant (2C_3 v_{k,n}^{3/2}/3\sqrt{2\pi})$. Now $v_{k,n} = \operatorname{Var} X_{n,k} + 2\sum_{j=1}^{k-1} \operatorname{Cov}(X_{n,j}, X_{n,k})$, hence $v_{k,n} \leqslant M_n\delta_n + 2\sum_{j=1}^{k_n-1} \min(\lambda_n M_n \theta_j, \Delta_n)$. We thus need

$$k_n^{2/3}\left[M_n\delta_n + 2\sum_{j=1}^{k-1}\min(\lambda_n M_n\theta_j, \Delta_n)\right] \to_{n\to\infty} 0. \tag{9}$$

Bound of $\Delta_{k,n}^{(1)}(h)$. Set $\Delta_{k,n}^{(1)}(h) = \mathbb{E}\delta_{k,n}^{(1)}(h)$ we write, again with some random $\tau_{k,n} \in (0,1)$, $\delta_{k,n}^{(1)}(h) = h'_{k,n}(S_{k-1,n})X_{n,k} + \frac{1}{2}h''_{k,n}(S_{k-1,n})(X_{n,k}^2 - v_{n,k}) + \frac{1}{6}(h_{k,n}^{(3)}(S_{k-1,n} + \tau_{n,k}X_{n,k})X_{n,k}^3)$. We analyze separately the term in the previous expression

$$\frac{1}{6}|\mathbb{E}\, h_{k,n}^{(3)}(S_{k-1,n} + \tau_{n,k}X_{n,k})X_{n,k}^3| \leqslant \frac{C_3}{6}M_n^2\delta_n. \tag{10}$$

To estimate the median term, we write $\operatorname{Cov}(h''_{k,n}(S_{k-1,n}), X_{n,k}^2) = \sum_{j=1}^{k-1}\operatorname{Cov}((h''_{k,n}(S_{j,n}) - h''_{k,n}(S_{j-1,n}), X_{n,k}^2)$, hence

$$|\operatorname{Cov}(h''_{k,n}(S_{k-1,n}), X_{n,k}^2)| \leqslant \max(2C_2, C_3)M_n\sum_{j=1}^{k-1}\min(\lambda_n\theta_j, \Delta_n). \tag{11}$$

$$|\operatorname{Cov}(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}), X_{n,k})| \leqslant C\min(M_n\lambda_n\theta_{k-i}, \Delta_n) \tag{12}$$

and

$$|Eh''_{k,n}(S_{k-1,n})EX_{n,i}X_{n,k}| \leqslant C\min(M_n\lambda_n\theta_{k-i}, \Delta_n). \tag{13}$$

Adding (12) and (13) and summing up the expression for all i we get

$$\left|\mathbb{E}\, h'_{k,n}(S_{k-1,n})X_{n,k} - \mathbb{E}\, h''_{k,n}(S_{k-1,n})\sum_{i=1}^{k-1}\mathbb{E}X_{n,i}X_{n,k}\right| \leqslant C\sum_{p=1}^{k-1}\min(M_n\lambda_n\theta_p, \Delta_n). \tag{14}$$

We add Eqs. (10), (11) and (14) to obtain

$$|\Delta_{k,n}^{(1)}(h)| \leqslant \max(2C_2, C_3)\left[M_n^2\delta_n + M_n\sum_{p=1}^{k-1}\min(\lambda_n\theta_p, \Delta_n) + \sum\min(M_n\lambda_n\theta_p, \Delta_n)\right]. \tag{15}$$

Now we sum for all k to conclude

$$\left|\sum_{k=2}^{k_n}\Delta_{k,n}^{(1)}(h)\right| \leqslant C\left[k_nM_n^2\delta_n + k_nM_n\sum_{p=1}^{\infty}\min(\lambda_n\theta_p, \Delta_n) + k_n\sum_{p=1}^{\infty}\min(M_n\lambda_n\theta_p, \Delta_n)\right]. \tag{16}$$

3.2. Proof of Theorem 2

Only replace, respectively, Eqs. (11)–(14) by

$$|\mathrm{Cov}(h''_{k,n}(S_{k-1,n}), X^2_{n,k})| \leqslant CM_n \sum_{j=1}^{k-1} \min(\lambda_n^2\theta_j, \Delta_n),$$

$$|\mathrm{Cov}(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}), X_{n,k})| \leqslant C \min(\Delta_n, \lambda_n^2\theta_{k-i}),$$

$$|\mathbb{E}\, h''_{k,n}(S_{k-1,n})\mathbb{E}\, X_{n,i}X_{n,k}| \leqslant C \min(\Delta_n, \lambda_n^2\theta_{k-i})$$

and

$$\left|\mathbb{E}\, h'_{k,n}(S_{k-1,n})X_{n,k} - \mathbb{E}\, h''_{k,n}(S_{k-1,n}) \sum_{i=1}^{k-1} \mathbb{E}\, X_{n,i}X_{n,k}\right| \leqslant C \sum_{p=1}^{k-1} \min(\lambda_n^2\theta_{k-i}, \Delta_n).$$

Then we replace (15) by

$$|\Delta^{(1)}_{k,n}(h)| \leqslant C\left[M_n^2\delta_n + M_n \sum_{p=1}^{k-1} \min(\lambda_n^2\theta_p, \Delta_n) + \sum_{p=1}^{k-1} \min(\lambda_n^2\theta_p, \Delta_n)\right].$$

Finally we replace (16) by

$$\left|\sum_{k=2}^{k_n} \Delta^{(1)}_{k,n}(h)\right| \leqslant Ck_nM_n^2\delta_n + k_nM_n \sum_{p=1}^{\infty} \min(\lambda_n^2\theta_p, \Delta_n) + \sum_{p=1}^{\infty} \min(\lambda_n^2\theta_p, \Delta_n).$$

References

Ango Nze, P., Doukhan, P., 1996. Non-parametric Minimax estimation in a weakly dependent framework Is: Quadratic properties. Math. Meth. Statist. 5 (4), 404–423.

Chanda, K.C., Ruymgart, F.H., 1990. General linear processes: a property of the empirical process applied to density and mode estimation. J. Time Ser. Anal. 11 (3), 185–199.

Doukhan, P., 1994. Mixing: Properties and Examples, Lecture Notes in Statistics, Vol. 85. Springer, Berlin.

Doukhan, P., Louhichi, S., 1997. Functional estimation of a density under a new weak dependence condition. Université de Cergy Pontoise, Prépublication 10.97 (a short version of this appeared in C.R.A.S. Paris (1998) 327, 989–992).

Doukhan, P., Louhichi, S., 1999. A new weak dependence condition and applications to moment inequalities. Stoch. Process. Appl. 84, 313–342.

Esary, J., Proschan, F., Walkup, D., 1967. Association of random variables with applications. Ann. Math. Statist. 38, 1466–1476.

Isha, B., Prakasa, R., 1995. Kernel-type density and failure rate estimation for associated sequences. Ann. Inst. Statist. Math. 47, 253–266.

Peligrad, M., Utev, S., 1997. Central limit theorem for linear processes. Ann. Probab. 25 (1), 443–456.

Rio, E., 1995. About the Lindeberg method for strongly mixing sequences. ESAIM, Probab. Statist. 1, 35–61.

Rio, E., 1996. Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes. Probab. Theory Related Fields 104, 255–282.

Robinson, P.M., 1983. Non parametric estimators for time series. J. Time Ser. Anal. 4 (3), 185–207.

Rosenblatt, M., 1991. Stochastic curve estimation. NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 3.

Chapitre 2

Théorème Limite Centrale Fonctionnel Empirique

Dans ce chapitre, nous établissons un théorème limite centrale fonctionnel pour un processus empirique construit à partir d'une suite de variables aléatoires faiblement dépendantes. Les principales étapes dans la preuve de ce théorème sont l'établissement de nouvelles inégalités de moments et un théorème limite centrale pour les sommes partielles d'une suite de variables aléatoires faiblement dépendantes.

An Empirical Functional Central Limit Theorem For Weakly Dependent Sequences

C. Prieur

Préprint de l'Université de Cergy-Pontoise.

Abstract : In this paper we obtain a Functional Central Limit Theorem for the empirical process of a stationary sequence under a new weak dependence condition introduced by Doukhan & Louhichi [21]. This result improves on the Empirical Functional Central Limit Theorem in Doukhan & Louhichi [21]. Our proof relies on new moment inequalities and on a Central Limit Theorem. Techniques of proofs come respectively from Louhichi [40] and Rio [54]. We also deduce a rate of convergence in a Marcinkiewicz-Zygmund Strong Law.

AMS Mathematics Subject Classifications (2000) : 60E15; 60F05; 60F15; 60F17; 60G10.

Key words: stationary sequences, Rosenthal inequality, moment inequalities, Functional Central Limit Theorem, empirical process, Marcinkiewicz-Zygmund Strong Law, weakly dependent sequences, Lindeberg Theorem.

1 Introduction

In this paper we essentially give a Functional Central Limit Theorem (Theorem 1) for the empirical process of stationary weakly dependent sequences which improves on results in Doukhan & Louhichi [21]. Our dependence frame is described in section 1.1. The main steps to obtain our theorem are a new moment inequality (Lemma 3) and a Central Limit Theorem (CLT) (Theorem 2).

Moment inequalities under independence have already been studied. We recall here Rosenthal inequality:

$$\mathbf{E}|S_n|^r \leq C_r \left\{ (\text{Var } S_n)^{r/2} + n\mathbf{E}|X_1|^r \right\}, \tag{1}$$

where $X = (X_1, \ldots, X_n)$ is a centered vector of independent and identically distributed real valued random variables with finite variance, $S_n = X_1 + \cdots + X_n$, $S_0 = 0$, and $r \in]2, +\infty[$. Doukhan & Louhichi [21] obtain Rosenthal inequalities under weak dependence, but exponents are restricted to be even integers. Louhichi [40] get moment inequalities of order $r \in]2, +\infty[$ for a class of sequences satisfying the property "*the non correlation yields the independence.*" This property is called the $(AG)-$property and is satisfied by associated and Gaussian processes. In this paper, we extend Louhichi's moment inequalities to weakly dependent sequences (Lemma 3 stated in section 3). Weak dependence used here is precised by the definition 1 below. Examples of such sequences are described in section 1.2. The tightness of the empirical process is deduced from these moment inequalities. Another application of our moment inequalities is a Marcinkiewicz-Zygmund Strong Law (MZSL) for partial sums of bounded dependent random variables (Corollary 1).

We prove then a Central Limit Theorem from which we deduce the fi-di convergence. Its proof relies on a variation on the Lindeberg-Rio method (Rio, 1996) and not on Bernstein's blocks (used e.g. by Doukhan & Louhichi, 1999, to prove the fi-di convergence).

We relax assumptions of previous authors for both tightness and fi-di convergence.

The paper is organized as follows. Our main result is stated in Section 2. In Section 3 we state our new moment inequalities. We also write a corollary concerning rate of convergence for a Marcinkiewicz-Zygmund Strong Law (MZSL) for partial sums in our dependence frame. Finally sections 4, 5 and 6 are respectively devoted to the proofs of the main result, of the corollary concerning a MZSL and of our moment inequalities. We defer the proof of the Central Limit Theorem and of some technical lemmas respectively to Appendix A and Appendix B.

1.1 Weak dependence

Our dependence frame is a variation on the one in Doukhan & Louhichi [21]. We work here under a causality assumption which is fundamental in the proof of our moment inequalities (see section 6). More precisely, E being some Euclidean space $\mathbb{R}^d$ endowed with its Euclidean norm $\|.\|$, we shall consider a sequence of E-valued random variables $(X_n)_{n\in\mathbb{N}}$. We define $\mathbb{L}^\infty$ as the set of measurable and bounded numerical functions on some space $\mathbb{R}^m$, $m \in \mathbb{N}^*$ and its norm is classically written $\|.\|_\infty$. Moreover, let $u \in \mathbb{N}^*$ be a positive integer. We endow the set $F = E^u$ with the norm $\|(x_1,\dots,x_u)\|_F = \|x_1\| + \dots + \|x_u\|$. Let now $h : F = E^u \to \mathbb{R}$ be a numerical function on F, we set

$$\mathrm{Lip}(h) = \sup_{x\neq y} \frac{|h(x) - h(y)|}{\|x-y\|_F}$$

the Lipschitz modulus of h. Define $\mathcal{L} = \bigcup_{u=1}^\infty \{h \in \mathbb{L}^\infty(E^u,\mathbb{R});\ \mathrm{Lip}(h) < \infty\}$.

Definition 1

The sequence $(X_n)_{n\in\mathbb{N}}$ is $s-$weakly dependent (SWD) if for some sequence $\theta = (\theta_r)_{r\in\mathbb{N}}$ decreasing to zero at infinity, for any u-tuple $(i_1,\dots,i_u)$, $u \in \mathbb{N}^*$, and for any v-tuple $(j_1,\dots,j_v)$, $v \in \mathbb{N}^*$, with $i_1 \le \dots \le i_u \le i_u + r \le j_1 \le \dots \le j_v$, and $h \in \mathbb{L}^\infty$, $k \in \mathcal{L}$,

$$|\mathrm{Cov}\,(h(X_{i_1},\dots,X_{i_u}),k(X_{j_1},\dots,X_{j_v}))\,| \le v\,\|h\|_\infty\ \mathrm{Lip}(k)\ \theta_r. \tag{2}$$

1.2 Examples

Before stating our results, we give examples of (SWD) sequences in this section.

Definition 2

Let $(\eta_n)_{n\in\mathbb{Z}}$ be a stationary sequence of real valued random variables and F be a measurable function defined on $\mathbb{R}^{\mathbb{N}}$. The stationary sequence $(X_n)_{n\in\mathbb{Z}}$ defined by $X_n = F(\eta_n,\eta_{n-1},\eta_{n-2},\dots)$ is called a causal Bernoulli shift.

We denote $(\theta_r)_{r\in\mathbb{N}}$ any nonnegative and non increasing sequence such that

$$\mathbf{E}|F(\eta_0,\eta_{-1},\eta_{-2},\dots) - F(\eta_0,\dots,\eta_{-r},0,0,\dots)| \le \theta_r.$$

Causal shifts with independent identically distributed (i.i.d.) innovations $(\eta_k)_{k\in\mathbb{Z}}$ satisfy (2) with θ_r (see Doukhan & Louhichi, 1999).

Examples of such situations follow:

- The real-valued functional autoregressive model:
  If $T : \mathbb{R} \to \mathbb{R}$ is such that $|T(u) - T(u')| \le c|u-u'|$ for some $0 \le c < 1$ and for all $u,\ u' \in \mathbb{R}$, and if $(\eta_n)_{n\in\mathbb{Z}}$ is some i.i.d innovation process satisfying $\mathbf{E}|\eta_0| < \infty$, $(X_n)_{n\in\mathbb{N}}$ defined by

$$X_n = T(X_{n-1}) + \eta_n \tag{3}$$

  is $s-$weakly dependent with $\theta_r = Cc^r$ for some constant $C > 0$.

- The non mixing stationary Markov chain with i.i.d. Bernoulli innovations ($P(\eta_0 = 0) = P(\eta_0 = 1) = 1/2$) $X_n = (X_{n-1} + \eta_n)/2$ is $s-$weakly dependent with $\theta_r = \mathcal{O}(2^{-r})$; its marginal distribution is uniform on [0,1].
- Chaotic expansions associated with the discrete chaos generated by the sequence $(\eta_t)_{t\in\mathbb{Z}}$: In a condensed formulation we write, $F(x) = \sum_{k=0}^{\infty} F_k(x)$, $x \in \mathbb{R}^{\mathbb{N}}$ for

$$F_k(x) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_k=0}^{\infty} a^{(k)}_{j_1,\ldots,j_k} x_{j_1}x_{j_2}\cdots x_{j_k},\ k \geq 1$$

where $F_k(x)$ denotes the k-th order chaos contribution and $F_0(x) = a_0^{(0)}$ is only a centering constant. In short we write in the vectorial notation, $F_k(x) = \sum_{j\in\mathbb{N}^k} a_j^{(k)} x_j$. Processes associated with a finite number of chaos (i.e. $F_k \equiv 0$ if $k > k_0$ for some $k_0 \in \mathbb{N}$) are also called *Volterra processes*. A simple and general condition for $\mathbb{L}^1$-convergence of this expansion, still written in a condensed notation, is $\sum_{k=0}^{\infty}\left\{\sum_{j\in\mathbb{N}^k}\left|a_j^{(k)}\right|\mathbf{E}|\eta_0|^k\right\} < \infty$. This condition allows to define the distribution of such shift processes. A suitable bound for θ_r is then

$$\theta_r = \sum_{k=0}^{\infty}\left\{\sum_{\{j\in\mathbb{N}^k;\|j\|_\infty > r\}}\left|a_j^{(k)}\right|\mathbf{E}|\eta_0|^k\right\} < \infty.$$

  - If $\mathbf{E}|\eta_0| = 0$ and

$$F_k(x) = \sum_{0\leq j_1<\cdots<j_k} a^{(k)}_{j_1,\ldots,j_k} x_{j_1}x_{j_2}\cdots x_{j_k},\ k \geq 1,$$

    then $\theta_r = \left(\sum_{0\leq j_1<\cdots<j_k, j_k>r}(a^{(k)}_{j_1,\ldots,j_k})^2(\mathbf{E}\eta_0^2)^k\right)^{1/2}$.
  - Linear processes $X_n = \sum_{k=0}^{\infty} a_k\eta_{n-k}$ which include ARMA models are those with $F_k(x) \equiv 0$ for all $k > 1$. A first choice is $\theta_r = \mathbf{E}|\eta_0|\sum_{k>r}|a_k|$ for the linear process with i.i.d. innovations such that $\mathbf{E}|\eta_0| < \infty$. For centered and $\mathbb{L}^2$ innovations, another choice is thus $\theta_r = \sqrt{\mathbf{E}|\eta_0|^2\sum_{k>r}|a_k|^2}$.
  - The simple bilinear process with the recurrence equation $X_t = aX_{t-1} + bX_{t-1}\eta_{t-1} + \eta_t$. Such processes are associated with the chaotic representation in

$$F(x) = \sum_{j=1}^{\infty} x_j \prod_{s=0}^{j-1}(a + bx_s),\ x \in \mathbb{R}^{\mathbb{Z}}.$$

    If $c = \mathbf{E}|a + b\xi_0| < 1$ then $\theta_r = \frac{c^r(r+1)}{c-1}$ has a geometric decay rate.
  - ARCH(∞): The recurrence equation $X_t = (a + \sum_{j=1}^{\infty} b_j X_{t-j})\eta_t$ has the stationary solution:

$$X_t = a\sum_{l=0}^{\infty}\sum_{j_1\cdots j_l=1}^{\infty} b_{j_1}\cdots b_{j_l}\eta_t\eta_{t-j_1}\cdots\eta_{t-j_1-\cdots-j_l},$$

    with $a \geq 0$, $b_j \geq 0$ for all j, $\sum_{j=1}^{\infty} b_j < \infty$ and $(\eta_k)_{k\in\mathbb{Z}}$ a sequence of i.i.d. non-negative random variables (see e.g. Giraitis, Kokoszka and Leipus, 1998). Then we can prove that the process $(X_t)_{t\in\mathbb{Z}}$ is $s-$weakly dependent (see Doukhan, 2001).

2 The Empirical Functional Central Limit Theorem

Let $(X_n)_{n\in\mathbb{N}}$ denote a stationary sequence of real valued random variables. In this section we investigate some properties of the empirical process constructed from the stationary sequence $(X_n)_{n\in\mathbb{N}}$.

We get a Functional Limit Theorem for the empirical process under the (SWD)-weak dependence condition. We consider a stationary sequence $(X_n)_{n\in\mathbb{N}}$ of random variables with common distribution function F. For $t \in \mathbb{R}$, we consider the following processes:

$F_n(t) := \frac{1}{n}\sum_{k=0}^{n-1} 1_{X_k\le t}$ and $U_n(t) := \sqrt{n}\left(F_n(t) - F(t)\right)$.

We have the following result:

Theorem 1 *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of real valued random variables with common repartition function F supposed to be Lipschitz. Suppose that (X_n) satisfies the $(SWD)-$dependence condition, with $\theta_r = \mathcal{O}\left((r+1)^{-2-2\sqrt{2}-\nu}\right)$, for some $\nu > 0$. Then the sequence of processes*

$\{U_n(t);\ t\in\mathbb{R}\}_{n>0}$ converges in distribution in the Skohorod space $\mathcal{D}(\mathbb{R})$ to the centered Gaussian process indexed by $\mathbb{R}$ with covariance defined by $\Gamma(s,t) = \sum_{k=-\infty}^{+\infty} Cov(1_{X_0\le s}, 1_{X_{|k|}\le t})$.

In fact, Theorem 1 above can be decomposed in two parts: the tightness and the fi-di convergence.

Lemma 1 (Tightness) *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of real valued random variables with common distribution function F supposed to be Lipschitz. We assume that $(X_n)_{n\in\mathbb{N}}$ fulfils the $(SWD)-$dependence condition with $\theta_r = \mathcal{O}\left((r+1)^{-2-2\sqrt{2}-\nu}\right)$, for some $\nu > 0$.*

Then the sequence of processes $\{U_n(t);\ t\in\mathbb{R}\}_{n>0}$ is tight in the Skohorod space $\mathcal{D}(\mathbb{R})$.

Lemma 1 above is proved in section 4. It clearly improves on Doukhan & Louhichi [21] who assume $\theta_r = \mathcal{O}(r^{-(5+\nu)})$, $\nu > 0$. Indeed, in order to obtain tightness, those authors calculate the moment of order 4 of the partial sums. Here we just have to calculate some moment of order $2+\sqrt{2}$ as is shown in the proof of Lemma 1. Lemma 3 in the next section allows us indeed to calculate moment with an order r which is not necessarily an integer.

The fi-di convergence is deduced from the following Central Limit Theorem:

Theorem 2 (Central Limit Theorem) *Let $(X_n)_{n\in\mathbb{N}}$ de a stationary sequence of centered $(SWD)-$weakly dependent random variables with $\theta_r = \mathcal{O}\left((r+1)^{-a}\right)$, for some $a > 3/2$. We assume that $(X_n)_{n\in\mathbb{N}}$ is uniformly bounded. If $S_n = X_0 + \cdots X_{n-1}$, we assume that $\frac{\mathrm{Var}(S_n)}{n} \xrightarrow[n\infty]{} \sigma^2 > 0$. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1} X_i \xrightarrow[n\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2).$$

The method to prove Theorem 2 is a variation on Lindeberg-Rio method (Rio, 1996). We deduce from this CLT the following lemma:

Lemma 2 (Fi-di convergence) *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of real valued random variables with common distribution function F supposed to be Lipschitz. Suppose that $(X_n)_{n\in\mathbb{N}}$ satisfies the $(SWD)-$dependence condition, with a sequence $(\theta_r)_{r\in\mathbb{N}} := ((r+1)^{-a})_{r\in\mathbb{N}}$. Assume that $a > 3$. Then the finidimensional marginals of the process $\{U_n(t);\ t\in\mathbb{R}\}_{n>0}$ converge in distribution to the finidimensional marginals of the centered Gaussian process indexed by $\mathbb{R}$ with covariance defined by $\Gamma(s,t) = \sum_{k=-\infty}^{+\infty} \mathrm{Cov}\left(1_{X_0\le s}, 1_{X_{|k|}\le t}\right)$.*

The proofs of Theorem 2 and Lemma 2 are deferred to Appendix A. The condition $\theta_r = \mathcal{O}\left((r+1)^{-3-\delta}\right)$, for some $\delta > 0$ in Lemma 2 improves the condition $\theta_r = \mathcal{O}\left(r^{-4}\right)$ obtained by Doukhan & Louhichi [21] for fi-di convergence.

Now both tightness result and fi-di convergence result yield Theorem 1.

3 Moment inequalities

In the statements of the main results in section 2, we consider a (SWD) sequence $(X_n)_{n\in\mathbb{N}}$. To prove the tightness in Lemma 1, we need moment inequalities for the partial sums of a sequence $(Y_n) = (\varphi(X_n))$. We prove in section 4 that $(Y_n)_{n\in\mathbb{N}}$ is also $s-$weakly dependent. Therefore the goal of the following Lemma is to give moment inequalities for (SWD) sequences. Doukhan & Louhichi [21] prove moment inequalities for weakly dependent sequences. But the order of these inequalities is an integer not less than 2. Recently, Louhichi [40] proves moment inequalities of order $r \in]2, +\infty[$ but for sequences satisfying the $(AG)-$property. The following variation on Louhichi's Lemma [40] entails moment bounds with order $r \in]2, +\infty[$ for $(SWD)-$sequences:

Lemma 3 *Let r be a fixed real number > 2. Let (X_n) be a stationary sequence of centered and (SWD) random variables. Suppose moreover that this sequence is bounded by 1. Let $S_n := X_1 + X_2 + \cdots + X_n$, for $n \geq 1$ and $S_0 = X_0 = 0$. Then there exists a positive constant C_r depending only on r, such that*

$$\mathbf{E}|S_n|^r \leq C_r \left[s_n^r + M_{r,n}\right], \tag{4}$$

where $M_{r,n} := n\sum_{i=0}^{n-1}(i+1)^{r-2}\theta_i$, and $s_n^2 := M_{2,n} = n\sum_{i=0}^{n-1}\theta_i$.

We prove Lemma 3 in Section 6.

These moment inequalities also allow us to study the rate of convergence for a Marcinkiewicz-Zygmund Strong Law for partial sums of bounded dependent random variables. Such results appear for example in Lai [32] and in Berbee [7] for the mixing case and in Louhichi [38] for the associated case.

Let us first introduce the notion of $r-$quick convergence as in Lai [32].

Definition 3 *A sequence $(Z_n)_{n\in\mathbb{N}}$ of random variables converges to 0 $r-$quickly $(r > 0)$ if*

$$\mathbf{E}\left(N_\epsilon\right)^r < \infty \ \forall \varepsilon > 0,$$

where $N_\epsilon := \sup\{n \geq 1 : \ |Z_n| \geq \varepsilon\}$.

Note that $Z_n \longrightarrow 0$ $r-$quickly for some $r > 0$ implies $Z_n \longrightarrow 0$ a.s.

The following corollary is a convergence theorem for $s-$weakly dependent variables.

Corollary 1 *Let $(X_n)_{n\geq 1}$ be a stationary sequence of centered and (SWD) random variables. Let r be a fixed real number > 2. Suppose moreover that this sequence is bounded by some positive constant M. Assume that the coefficient of (SWD) satisfies $\theta_q = \mathcal{O}\left((q+1)^{-D}\right)$ with $D > r-1$. Then,*

– *for all $\frac{1}{2} < \alpha \leq 1$, for all $k < (\alpha - \frac{1}{2})r - 1$, and for all $\varepsilon > 0$, we have:*

$$\sum_{n\geq 1} n^k P(\max_{j\leq n}|S_j| \geq \varepsilon n^\alpha) < \infty; \tag{5}$$

– *for all $1 \geq \alpha > \frac{1}{2}$, for all $1 < p\alpha < (\alpha - \frac{1}{2})r + 1$, we have the four following assertions:*

1. $\sum_{n\geq 1} n^{p\alpha-2}P(\max_{j\leq n}|S_j| \geq \varepsilon n^\alpha) < \infty \ \forall \varepsilon > 0$,
2. $\mathbf{E}\{\sup_{n\geq 0}(|S_n| - \varepsilon n^\alpha)\}^{\frac{p\alpha-1}{\alpha}} < \infty \ \forall \varepsilon > 0$,
3. $\sum_{n\geq 1} n^{p\alpha-2}P(\sup_{j\geq n} j^{-\alpha}|S_j| \geq \varepsilon) < \infty \ \forall \varepsilon > 0$,
4. $n^{-\alpha}S_n \longrightarrow 0$ $(p\alpha - 1)$ – *quickly.*

The proof of Corollary 1 is written in section 5.

4 Proof of Theorem 1

This section is devoted to the proof of the main result (Theorem 1) stated in section 2. We prove the Functional Limit Theorem (Theorem 1) in 2 steps: the tightness and the fi-di convergence.

Proof of Lemma 1

In the following, C will denote some arbitrary constant which may vary from line to line. Let $s < t$ be two real numbers. We want to apply moment inequalities of Lemma 3 to the sequence $\left(\frac{1_{s<X_n\le t} - (F(t)-F(s))}{2}\right)_{n\in\mathbb{N}}$.

So we have to prove that this sequence which we call $(Y_n)_{n\in\mathbb{N}}$ satisfies a (SWD) dependence condition. For this we have to bound $C_{h,u,v} := \mathrm{Cov}\left(h(Y_{i_1},\dots,Y_{i_u}),k\left(Y_{j_1},\dots,Y_{j_v}\right)\right)$ for all $i_1 \le \cdots \le i_u \le i_u + r \le j_1 \le \cdots \le j_v$ and for all $h \in \mathbb{L}^\infty$, $k \in \mathcal{L}$. Let $\epsilon > 0$ such that $s+\varepsilon < t-\varepsilon$. Let us write $Y_n = \varphi(X_n)$. We want to smooth the function φ which is not Lipschitz. For this we consider the following Lipschitz function φ^ϵ smoothing φ :

– φ_ϵ is equal to φ on $]-\infty,s-\epsilon]\cup]s+\epsilon,t-\varepsilon]\cup]t+\varepsilon,+\infty[$;

– for $s-\varepsilon < x \le s+\varepsilon$, $\varphi_\varepsilon(x) = \frac{-1}{8\varepsilon^3}\left(x^3 - 3sx^2 + 3(s^2-\varepsilon^2)x - s^3 + 3s\varepsilon^2 - 2\varepsilon^3\right) - \frac{F(t)-F(s)}{2}$;

– for $t-\varepsilon < x \le t+\varepsilon$, $\varphi_\varepsilon(x) = \frac{-1}{8\varepsilon^3}\left(x^3 - 3tx^2 + 3(t^2-\varepsilon^2)x - t^3 + 3t\varepsilon^2 - 2\varepsilon^3\right) - \frac{F(t)-F(s)}{2}$.

We then have $||\varphi^\epsilon||_\infty \le ||\varphi||_\infty \le \frac{1}{2}$ and $\mathrm{Lip}(\varphi^\epsilon) \le \frac{3}{8\,\epsilon}$.

Then,

$$
\begin{aligned}
|C_{h,u,v}| &\le \left|\mathrm{Cov}\left(h(Y_{i_1},\cdots,Y_{i_u}),k\left(\varphi(X_{j_1}),\dots,\varphi(X_{j_v})\right) - k\left(\varphi^\epsilon(X_{j_1}),\dots,\varphi^\epsilon(X_{j_v})\right)\right)\right| \\
&\quad + \left|\mathrm{Cov}\left(h(Y_{i_1},\dots,Y_{i_u}),k\left(\varphi^\epsilon(X_{j_1}),\dots,\varphi^\epsilon(X_{j_v})\right)\right)\right| \\
&\le 2\,||h||_\infty \mathrm{Lip}(k) \textstyle\sum_{i=1}^v \mathbf{E}|\varphi(X_{j_i}) - \varphi^\epsilon(X_{j_i})| + ||h||_\infty\, \mathrm{Lip}(k)\ \mathrm{Lip}(\varphi^\varepsilon)\ \theta_r\ v \\
&\le 2||h||_\infty \mathrm{Lip}(k) v\ 2\ ||\varphi||_\infty P(X_0 \in]s-\varepsilon,s+\varepsilon]\cup]t-\varepsilon,t+\varepsilon]) + ||h||_\infty \mathrm{Lip}(k)\mathrm{Lip}(\varphi^\varepsilon)\ \theta_r\ v \\
&\le 2||h||_\infty \mathrm{Lip}(k) v\ 2\ ||\varphi||_\infty\ 4\varepsilon\ \mathrm{Lip}(F) + ||h||_\infty \mathrm{Lip}(k)\ \tfrac{3}{8\varepsilon}\ \theta_r\ v \\
&\le ||h||_\infty \mathrm{Lip}(k) v\ \left(8\ \varepsilon\ \mathrm{Lip}(F) + \tfrac{3\,\theta_r}{8\epsilon}\right).
\end{aligned}
$$

In the following, C denotes some positive constant which may vary from line to line.

So we have

$$|C_{h,u,v}| \le C||h||_\infty\ \mathrm{Lip}(k)\ v\ \left(\varepsilon + \frac{\theta_r}{\varepsilon}\right). \tag{6}$$

Then if $\sqrt{\theta_r} < \frac{t-s}{2}$, we take $\epsilon = \sqrt{\theta_r}$, and we get:

$$|C_{h,u,v}| \le C\ ||h||_\infty\ \mathrm{Lip}(k)\ v\ \sqrt{\theta_r}. \tag{7}$$

Moreover $|C_{h,u,v}| = \left|\mathrm{Cov}\left(h(Y_{i_1},\dots,Y_{i_u}),k\left(Y_{j_1},\dots,Y_{j_v}\right) - k\left({Y'}_{j_1},\dots,{Y'}_{j_v}\right)\right)\right|$, where we consider $Y'_n = \varphi(X'_n)$ with $(X'_n)_{n\in\mathbb{N}}$ independent of $(X_n)_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$, ${X'}_n \sim X_n$.

Then, $|C_{h,u,v}| \le 2||h||_\infty\ \mathrm{Lip}(k)\ \mathbf{E}\sum_{i=1}^v |\varphi(X_{j_i}) - \varphi(X'_{j_i})| \le 2\ ||h||_\infty\ \mathrm{Lip}(k)\ v\ \frac{1}{2}\ 2\mathbf{E}|1_{s<X_0\le t}|$.

Hence,

$$|C_{h,u,v}| \le C\ ||h||_\infty\ \mathrm{Lip}(k)\ v\ \left(F(t) - F(s)\right). \tag{8}$$

From (8) and as F is Lipschitz, we deduce that

$$|C_{h,u,v}| \le C\ ||h||_\infty\ \mathrm{Lip}(k)\ v\ |t-s|. \tag{9}$$

Hence, if $\sqrt{\theta_r} \geq \frac{t-s}{2}$, we get:

$$|C_{h,u,v}| \leq C \; ||h||_\infty \; \mathrm{Lip}(k) \; v \; \sqrt{\theta_r}.$$

Hence, for all real numbers $s < t$, inequality (7) holds. We conclude with inequalities (7) and (8),

$$|C_{h,u,v}| \leq C \; ||h||_\infty \; \mathrm{Lip}(k) \; v \; ((F(t) - F(s)) \wedge \sqrt{\theta_r}). \tag{10}$$

So $(Y_n)_{n \in \mathbb{N}}$ is (SWD) with $\theta_r' \leq C \; ((F(t) - F(s)) \wedge \sqrt{\theta_r})$. We prove the tightness by applying Lemma 3 to $(Y_n)_{n \in \mathbb{N}}$. In the following C_r denotes some arbitrary constant depending only on r and which may vary from line to line. Notice that $\frac{2}{\sqrt{n}} \sum_{k=0}^{n-1} Y_k = U_n(t) - U_n(s)$.

If we write $\theta_r = r^{-a}$, $a > 0$, we get

$$\begin{aligned} E\left|\tfrac{2}{\sqrt{n}} \textstyle\sum_{k=0}^{n-1} Y_k\right|^r \quad &\leq \quad C_r \left\{ \left(\textstyle\sum_{k=1}^{n} k^{\frac{-a}{2}} \wedge (F(t) - F(s)) \right)^{\frac{r}{2}} \right\} \\ &\quad + C_r \left\{ n^{\frac{2-r}{2}} \textstyle\sum_{k=1}^{n} (k+1)^{r-2} [k^{\frac{-a}{2}} \wedge (F(t) - F(s))] \right\} \end{aligned}$$

Hence if $a > 2$ and if $a > 2(r-1)$,

$$E\left|\frac{2}{\sqrt{n}} \sum_{k=0}^{n-1} Y_k\right|^r \leq C_r \{ (F(t) - F(s))^{\frac{r(a-2)}{2a}} + n^{\frac{2-r}{2}} \, (F(t) - F(s))^{\frac{2+a-2r}{a}} \}.$$

Now, if we chose $r = 2 + \sqrt{2}$, as F is continuous, it follows from Theorem 2.1 in Shao & Yu [63] that the sequence $\{U_n(t), t \in \mathbb{R}\}$ is tight in the Skohorod space $\mathcal{D}(\mathbb{R})$ as soon as $a > 2 + 2\sqrt{2}$. The choice $r = 2 + \sqrt{2}$ minimizes the condition on the dependence coefficient a. This concludes the proof of the tightness. □

To conclude the proof of Theorem 1 we now have to prove Lemma 2 of fi-di convergence. To prove it, we use a Central Limit Theorem (Theorem 2) whose proof is written in Appendix A.

Now Lemmas 1 and 2 together yield both the tightness and the fi-di convergence as soon as $a > 2 + 2\sqrt{2}$. This concludes the proof of Theorem 1.

5 Proof of Corollary 1

In this section we prove Corollary 1 stated in section 3. C_r will denote some constant depending only on r and which may vary from line to line.

We apply the moment inequalities to the partial sums of $(\frac{X_n}{M})$. We get as $D > r - 1$ and $r > 2$, for all n large enough:

$$\mathbf{E}\left|\frac{S_n}{M}\right|^r \leq C_r \; n^{\frac{r}{2}} \; \left(\sum_{i=0}^{+\infty} \theta_i \right)^{\frac{r}{2}}.$$

We then apply maximal inequalities in Moricz *et al* [43] and Bienaymé-Tchebycheff inequality to obtain: for $\frac{1}{2} < \alpha \leq 1$, for $k < (\alpha - \frac{1}{2})r - 1$ and for all $\varepsilon > 0$

$$P(\max_{1 \leq k \leq n} |S_k| \geq \varepsilon n^\alpha) \leq C_{r,\varepsilon} n^{\frac{r}{2} - \alpha r}, \tag{11}$$

where $C_{r,\varepsilon}$ only depends on r and ε.

From (11) we deduce (5) in Corollary 1. Now, Lemma 2 in Chow & Lai [13] together with (5) yield the four assertions *1.*, *2.*, *3.* and *4.* of Corollary 1.

Next section is devoted to the proof of moment inequalities stated in Lemma 3 of section 3.

6 Proof of Lemma 3

The proof is a variation on Louhichi's method [40] under our dependence frame. Let $p \geq 2$ be a fixed integer. We define the function $g_p : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+$ as in Louhichi [40]:

$$g_p(t,x) := \frac{1}{(p+1)!}\left[x^{p+1}1_{0\leq x\leq t} + (x^{p+1} - (x-t)^{p+1})1_{t<x}\right], \tag{12}$$

for any $x \geq 0$ and $g_p(t,x) = g_p(t,-x)$. Then we decompose the proof in several tool steps.

6.1 Step 1: Main terms

Let $p \geq 2$ be a fixed integer and $\mathcal{C}_p$ be the class of real-valued, p times continuously differentiable functions f such that $f(0) = \cdots = f^{(p)}(0) = 0$.
Let $\mathcal{F}_p(b_p,b_{p+1})$ be the subclass of $\mathcal{C}_{p+1}$ such that $\|f^{(p)}\|_\infty \leq b_p$ and $\|f^{(p+1)}\|_\infty \leq b_{p+1}$, where $\|f^{(i)}\|_\infty = \sup_{x\in\mathbb{R}} |f^{(i)}(x)|$ and $f^{(i)}$ denotes the differential of order i of the function f.
We recall a result in Louhichi [40] which is a generalization of equation (4.3) in Rio [53].
Lemma 4 *Let p be a fixed integer, $p \geq 2$. Let $\phi_p \in \Phi_p$, where*

$$\Phi_p := \{\phi : \mathbb{R}^+ \to \mathbb{R}^+; \phi \textit{ convex}, \phi(0) = \phi'(0) = \cdots = \phi^{(p)}(0) = 0,\ \phi^{(p)} \textit{ nondecreasing, concave}\}.$$

Suppose that $\lim_{x\to\infty} \phi_p^{(p+1)}(x) = 0$. Then

$$\phi_p(x) = \int_0^{+\infty} g_p(t,x)\nu_p(dt),$$

where ν_p is the Stieltjes measure of $-\phi_p^{(p+1)}$ defined by $\nu_p(dt) = -d\phi_p^{(p+1)}(t)$.
Lemma 4 above together with Fubini's Theorem yields

$$\mathbf{E}\phi_p(|S_n|) = \int_0^{+\infty} \mathbf{E}g_p(t,S_n)\nu_p(dt). \tag{13}$$

So we deduce the estimation of $\mathbf{E}\phi_p(|S_n|)$ from that of $\mathbf{E}g_p(t,S_n)$. Hence the goal of this step is to bound $\mathbf{E}f(S_n)$ for a "good" set of real valued functions f containing the functions $x \longrightarrow g_p(t,x)$, $t \geq 0$.

We notice that the function $x \longrightarrow g_p(t,x)$ as defined by (12) belongs to the set $\mathcal{F}_p(b_p,b_{p+1})$, with $b_p = t$ and $b_{p+1} = 1$. Hence we give in this step an estimation of $\mathbf{E}f(S_n)$, for $f \in \mathcal{F}_p(b_p,b_{p+1})$.
Let us first exhibit the main terms which appear in the proof.

We introduce the following notations:

$$\mathbf{E}_{p-1,k} = \sum_{0=:i_0\leq i_1\leq\ldots\leq i_{p-1}\leq k-1} \left|\mathbf{E}X_{k-i_0}X_{k-i_1}\cdots X_{k-i_{p-1}}\right|;$$

$$\mathbf{E}_{p-2,k}(\Delta f) = \sup_{0\leq u\leq 1} \sum_{0=:i_0\leq i_1\leq\ldots\leq i_{p-2}\leq k-1} \left|\mathbf{E}X_{k-i_0}X_{k-i_1}\cdots X_{k-i_{p-2}}\Delta_{p-2,k}(f)\right|,$$

where

$$\begin{aligned}\Delta_{p-2,k}(f) = \Delta_{p-2,k}(f,u) &= \left[f(S_{k-i_{p-2}-1} + uX_{k-i_{p-2}}) - f(S_{k-i_{p-2}-1})\right]\\ &= uX_{k-i_{p-2}}\int_0^1 f'\left(S_{k-i_{p-2}-1} + uvX_{k-i_{p-2}}\right)dv;\end{aligned}$$

and

$$\mathbf{E}_{p-2,k}(f) \quad = \sum_{0=:i_0\le i_1\le\ldots\le i_{p-2}\le k-1} |\mathbf{E}X_{k-i_0}X_{k-i_1}\cdots X_{k-i_{p-2}}f(S_{k-i_{p-2}-1})|.$$

Then we denote by $A_{p,n}$(respectively by $A_{p,n}(f)$, $A_{p,n}(\Delta f)$) the sum $\sum_{k=1}^n \mathbf{E}_{p,k}$ (respectively $\sum_{k=1}^n \mathbf{E}_{p,k}(f)$, $\sum_{k=1}^n \mathbf{E}_{p,k}(\Delta f)$).

For a real valued function f that belongs to the set $\mathcal{F}_p(b_p,b_{p+1})$, the quantity $|\mathbf{E}(f(S_n))|$ is evaluated by means of the main terms $\mathbf{E}_{p-2,k}(f^{(p-1)})$ and $\mathbf{E}_{p-2,k}(\Delta f^{(p-1)})$ as shows the following lemma.

Lemma 5 *Let p be a fixed integer, $p\ge 2$. Let (X_n) be a sequence of (SWD) random variables, centered and bounded by 1. There exists a positive constant C_p depending only on p, such that for any $f\in\mathcal{F}_p(b_p,b_{p+1})$*

$$\begin{aligned}|\mathbf{E}(f(S_n))| \quad &\le \quad C_p\left\{s_n^p(b_p\wedge b_{p+1}s_n)+(b_p\wedge b_{p+1})\sum_{i=0}^{n-1}(n-i)\theta_i\right.\\ &\left. + \sum_{k=1}^n \mathbf{E}_{p-2,k}(f^{(p-1)})+\sum_{k=1}^n \mathbf{E}_{p-2,k}(\Delta f^{(p-1)})\right\}.\end{aligned}$$

The covariance terms in Louhichi [40] are replaced here by bounds depending on the sequence $(\theta_r)_{r\ge 0}$.

Lemma 5 is proved in Appendix B.

6.2 Step 2: Evaluation of the main terms $\mathbf{E}_{p-2,k}(f)$ and $\mathbf{E}_{p-2,k}(\Delta f)$

The purpose of the second step is to evaluate the main terms $\mathbf{E}_{p-2,k}(f)$ and $\mathbf{E}_{p-2,k}(\Delta f)$ of Lemma 5. We need first the following preparatory lemmas.

Preparatory lemmas:

Let us recall the following notations: $M_{r,n} := n\sum_{i=0}^{n-1}(i+1)^{r-2}\theta_i$, $s_n^2 := M_{2,n} = n\sum_{i=0}^{n-1}\theta_i$.

Lemma 6 *[Doukhan and Louhichi [21]]*
Let (X_n) be a centered sequence of (SWD) random variables. Suppose that (X_n) is uniformly bounded by 1. Then, for any integer $p\ge 2$, there exists a positive constant C_p for which

$$\sum_{k=1}^n \mathbf{E}_{p-1,k}\le C_p\left\{s_n^p+M_{p,n}\right\}.$$

Lemma 6 is established by Doukhan and Louhichi [21], in order to give moment inequalities with integer order p.

The following lemma will often be used in the sequel and is proved in Appendix B.

Lemma 7 *Hölder's inequalities*
For all $p\ge 4$ and $m\in\{3,\ldots,p-1\}$, we have

$$M_{m,n}M_{p-m,n}\le s_n^{2p/(p-2)}M_{p,n}^{(p-4)/(p-2)}\le\left\{s_n^p+M_{p,n}\right\}. \tag{14}$$

Now define

$$M_{m,n}(b_1,b_2) := n\sum_{i=0}^{n-1}(b_1\wedge b_2(i+1))(i+1)^{m-2}\theta_i,$$

then

$$s_n^{p-m}(b_1 \wedge b_2 s_n) M_{m,n} \le s_n^p(b_1 \wedge b_2 s_n) + M_{p,n}(b_1,b_2). \tag{15}$$

Also

$$s_n^{p-m} M_{m,n}(b_1,b_2) \le s_n^p(b_1 \wedge b_2 s_n) + M_{p,n}(b_1,b_2). \tag{16}$$

The basic lemma:

The following lemma is the basic technical lemma of this section.

Lemma 8 *Let f be a real valued function of the set $\mathcal{F}_1(b_1,b_2)$. Let (X_n) be a centered sequence of (SWD) random variables. Suppose that (X_n) is uniformly bounded by 1. Then, for any integer $p \ge 2$, there exists a positive constant C_p depending only on p, for which*

$$\begin{aligned}\sum_{k=1}^n \mathbf{E}_{p-2,k}(\Delta f) &+ \sum_{k=1}^n \mathbf{E}_{p-2,k}(f)\\ &\le C_p \left\{ s_n^p(b_1 \wedge b_2 s_n) + \sum_{m=2}^{p-2} M_{m,n} M_{p-m,n}(b_1,b_2) + M_{p,n}(b_1,b_2) \right\}\end{aligned} \tag{17}$$

The right hand side term of (17) is similar to the one obtained in Lemma 5 of Louhichi [40]. However, details of the proof are slightly different due to the kind of dependence assumed.

Proof of Lemma 8

Using induction on $p \ge 2$, we will prove that each of the terms $\sum_{k=1}^n \mathbf{E}_{p-2,k}(\Delta f)$ and $\sum_{k=1}^n \mathbf{E}_{p-2,k}(f)$ is bounded by the right hand side (RHS) of (17).

For $p = 2$, we refer to Louhichi [38]. We can also deduce the calculations for the case $p = 2$ from the general case that we state just below.

Suppose now that (17) holds for the order $(p-1)$, we will prove it for p.

Our purpose is then to evaluate the following sums:

$$\sum_{k=1}^n \sum_{0=:i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1} \left| \mathbf{E} X_{k-i_0} X_{k-i_1} \cdots X_{k-i_{p-2}} f(S_{k-i_{p-2}-1}) \right|, \tag{18}$$

and

$$\sum_{k=1}^n \sum_{0=:i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1} \left| \mathbf{E} X_{k-i_0} X_{k-i_1} \cdots X_{k-i_{p-2}} \Delta_{p-2,k} f \right|, \tag{19}$$

We argue as Doukhan & Portal [23] : let $0 =: i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1$ be a fixed sequence of increasing integers, let m be the smallest integer for which $r_m := i_{m+1} - i_m = \max_{1 \le q \le p-2}(i_q - i_{q-1})$. Finally let $\sum^{(m)}_{0=:i_0 \le i_1 \le \ldots \le i_{p-2}}$ denote the sums over the subdivisions $i_1 \le \ldots \le i_{p-2} \le k-1$ for which the big lag $\max_{1 \le q \le p-2}(i_q - i_{q-1})$ is reached at the index m.

We also set $\sum^{(m),r_m}_{0=:i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1}$ for sums over the subdivisions $i_1 \le \ldots \le i_{p-2}$ such that the big lag $i_{m+1} - i_m = \max_{1 \le q \le p-2}(i_q - i_{q-1})$ equals r_m.

Hence, if B denotes a subset of $\mathbb{N}^{p-2}$ and g is a positive function on $\mathbb{N}^{p-2}$, then

$$\sum_{(0=:i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1) \in B} g(i_1, \ldots, i_{p-2}) \le \sum_{m=0}^{p-3} \sum_{r_m=0}^{k-1} \sum^{(m),r_m}_{(0=:i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1) \in B} g(i_1, \cdots, i_{p-2}). \tag{20}$$

We mean by the notation $(0 =: i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1) \in B$ that $(i_1, \ldots, i_{p-2}) \in B$ and that $0 =: i_0 \le i_1 \le \ldots \le i_{p-2} \le k-1$.

Study of the bound of (18):
We note the following decomposition: for g continuous and differentiable, if $g(0)=0$,

$$g(S_n)=\sum_{k=1}^{n}[g(S_k)-g(S_{k-1})]=\sum_{k=1}^{n}X_k\, g^{'}(S_{k-1}+u_kX_k),\text{ where for all } 1\le k\le n,\ 0<u_k<1. \quad (21)$$

Now we use calculations on the function f' instead of f, whose expression appears thanks to Taylor's formula (cf. equality (21)). Let us go into further details.
First, we write

$$f(S_{k-i_{p-2}-1})=[f(S_{k-i_{p-2}-1})-f(S_{k-i_{p-2}-r_m-1})]+f(S_{k-i_{p-2}-r_m-1}).$$

The last decomposition yields :

$$\begin{aligned} & |\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}}f(S_{k-i_{p-2}-1}))| && (22)\\ \le\ & |\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}}[f(S_{k-i_{p-2}-1})-f(S_{k-i_{p-2}-r_m-1})])|\\ & +|\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}}f(S_{k-i_{p-2}-r_m-1}))|\\ =:\ & J_1(i_1,\ldots,i_{p-2})+J_2(i_1,\ldots,i_{p-2}). \end{aligned}$$

Evaluation of $\sum_{i_1\le\ldots\le i_{p-2}}J_2(i_1,\ldots,i_{p-2})$:

The relation $\mathbf{E}(XY)=\mathrm{Cov}(X,Y)+\mathbf{E}(X)\mathbf{E}(Y)$ writes

$$\begin{aligned}\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}}f(S_{k-i_{p-2}-r_m-1})) &= \mathrm{Cov}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}},f(S_{k-i_{p-2}-r_m-1}))\\ &\quad+\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}})\mathbf{E}(f(S_{k-i_{p-2}-r_m-1})).\end{aligned} \quad (23)$$

The function f belongs to the set $\mathcal{F}_1(b_1,b_2)$, hence using Taylor's formula (recall that $f(0)=f'(0)=0$), we obtain

$$|f(x)|\le b_1|x|,\quad |f(x)|\le\frac{b_2x^2}{2},$$

this yields

$$|\mathbf{E}(f(S_{k-i_{p-2}-r_m-1}))|\le s_n(b_1\wedge b_2s_n). \quad (24)$$

We deduce from the last bound, Lemmas 6 and 7, that

$$\begin{aligned}&\sum_{k=1}^{n}\sum_{0\le i_1\le\ldots\le i_{p-2}\le k-1}|\mathbf{E}(X_kX_{k-i_1}\cdots X_{k-i_{p-2}})\mathbf{E}(f(S_{k-i_{p-2}-r_m-1}))|\\ &\quad\le C_{p-1}\left(M_{p-1,n}+s_n^{p-1}\right)s_n(b_1\wedge b_2s_n)\le C_p\left(M_{p,n}(b_1,b_2)+s_n^p(b_1\wedge b_2s_n)\right). \end{aligned} \quad (25)$$

In order to evaluate the first term of the RHS of (23), we use the following lemma whose proof is deferred to Appendix B.

Lemma 9 *Let (X_n) be a centered sequence of (SWD) random variables. Suppose that (X_n) is uniformly bounded by 1. Then, for any integer $p\ge 2$, there exists a positive constant C_p such that for any $f\in\mathcal{F}_1(b_1,b_2)$:*

$$\sum_{k=1}^{n}\sum_{m=0}^{p-2}\sum_{r_m=0}^{k-1}\sum_{0=:i_0\le i_1\le\ldots\le i_{p-1}}^{(m),r_m}\sup_{u\in[0,1]}|\mathrm{Cov}(X_k\cdots X_{k-i_m},X_{k-i_{m+1}}\cdots X_{k-i_{p-1}}f'(S_{k-i_{p-1}-1}+uX_{k-i_{p-1}}))|$$

$$\le C_p\{M_{p,n}(b_1,b_2)+\sum_{m=2}^{p-2}M_{m,n}M_{p-m,n}(b_1,b_2)+s_n^p(b_1\wedge b_2s_n)\}.$$

Now, the first equality in (21) and Taylor's formula yield :

$$\sum_{0\le i_1\le\ldots\le i_{p-2}\le k-1}^{(m),r_m} |\mathrm{Cov}(X_k X_{k-i_1}\cdots X_{k-i_{p-2}}, f(S_{k-i_{p-2}-r_m-1}))|$$
$$\le \sum_{(i)}^{(m),r_m} \sum_{i_{p-1}=i_{p-2}+r_m+1}^{k-1} \sup_{u\in[0,1]} |\mathrm{Cov}(X_k\cdots X_{k-i_{p-2}}, X_{k-i_{p-1}} f'(S_{k-i_{p-1}-1}+uX_{k-i_{p-1}}))|, \tag{26}$$

where (i) means $0\le i_1\le\ldots\le i_{p-2}$.

Remark 4 The two sums $\sum_{0\le i_1\le\ldots\le i_{p-2}}^{(m),r_m}\sum_{i_{p-1}=i_{p-2}+r_m+1}^{k-1}$ are taken over the subdivisions $0\le i_1\le\cdots\le i_{p-1}\le k-1$ such that $r_m := i_{m+1}-i_m = \max_{1\le q\le p-2}(i_q-i_{q-1})$ and that $i_{p-1}-i_{p-2}\ge r_m+1$, hence $i_{p-1}-i_{p-2}$ is the big lag of the subdivision $0\le i_1\le\cdots\le i_{p-1}\le k-1$. So if we sum some positive quantities, $\sum_{m=0}^{p-3}\sum_{r_m=0}^{k-1}\sum_{0\le i_1\le\ldots\le i_{p-2}}^{(m),r_m}\sum_{i_{p-1}=i_{p-2}+r_m+1}^{k-1}\le\sum_{m=0}^{p-2}\sum_{r_m=0}^{k-1}\sum_{0\le i_1\le\ldots\le i_{p-2}\le i_{p-1}}^{(m),r_m}$. $\diamond$

We take the sums over $m:\ 0\le m\le(p-3)$, $r_m:\ 0\le r_m\le k-1$ and over $k:\ 1\le k\le n$ in (26), and we use Remark 4 and Lemma 9 to obtain :

$$\sum_{k=1}^{n}\sum_{m=0}^{p-3}\sum_{r_m=0}^{k-1}\sum_{0\le i_1\le\ldots i_{p-2}\le k-1}^{(m),r_m} |\mathrm{Cov}(X_k X_{k-i_1}\cdots X_{k-i_{p-2}}, f(S_{k-i_{p-2}-r_m-1}))|$$

$$\le C_p\{M_{p,n}(b_1,b_2)+\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(b_1,b_2)+s_n^p(b_1\wedge b_2 s_n)\} \tag{27}$$

We deduce from (23), (25) and (27) that

$$\sum_{k=1}^{n}\sum_{m=0}^{p-3}\sum_{r_m=0}^{k-1}\sum_{0\le i_1\le\ldots i_{p-2}\le k-1}^{(m),r_m} |J_2(i_1,\ldots,i_{p-2})|$$

$$\le C_p\{M_{p,n}(b_1,b_2)+\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(b_1,b_2)+s_n^p(b_1\wedge b_2 s_n)\}. \tag{28}$$

Evaluation of $\sum_{i_1\le\ldots\le i_{p-2}} J_1(i_1,\ldots,i_{p-2})$:

Let us note that

$$|\mathbf{E}XY|\le|\mathrm{Cov}(X,Y)|+|\mathbf{E}X||\mathbf{E}Y|.$$

Using decomposition (21) and Remark 4, we then deduce

$\sum_{0\le i_1\le\ldots\le i_{p-2}\le k-1}^{(m),r_m} |\mathbf{E}(X_k X_{k-i_1}\cdots X_{k-i_{p-2}}[f(S_{k-i_{p-2}-1})-f(S_{k-i_{p-2}-r_m-1})])|$

$$\begin{aligned}
\leq\ & \textstyle\sum_{0\leq i_1\leq\ldots\leq i_{p-2}\leq k-1}^{(m),r_m} |\mathrm{Cov}(X_k\cdots X_{k-i_m},X_{k-i_{m+1}}\cdots X_{k-i_{p-2}}[f(S_{k-i_{p-2}-1})-f(S_{k-i_{p-2}-r_m-1})])| \\
& +\textstyle\sum_{i_1\leq\ldots\leq i_{p-2}}^{(m),r_m} |\mathbf{E}(X_k\cdots X_{k-i_m})\mathbf{E}(X_{k-i_{m+1}}\cdots X_{k-i_{p-2}}[f(S_{k-i_{p-2}-1})-f(S_{k-i_{p-2}-r_m-1})])| \\
\leq\ & \textstyle\sum_{i_1\leq\ldots\leq i_{p-1}}^{(m),r_m} \sup_{u\in[0,1]} |\mathrm{Cov}(X_k\ldots X_{k-i_m},X_{k-i_{m+1}}\ldots X_{k-i_{p-1}}f'(S_{k-i_{p-1}-1}+uX_{k-i_{p-1}})| \\
& +\textstyle\sum_{i_1\leq\ldots\leq i_{p-2}}^{(m),r_m} [|\mathbf{E}(X_kX_{k-i_1}\ldots X_{k-i_m})\mathbf{E}(X_{k-i_{m+1}}\ldots X_{k-i_{p-2}}f(S_{k-i_{p-2}-1})| \\
& +|\mathbf{E}(X_k\ldots X_{k-i_m})\mathrm{Cov}(X_{k-i_{m+1}}\ldots X_{k-i_{p-2}},f(S_{k-i_{p-2}-r_m-1}))| \\
& +|\mathbf{E}(X_k\ldots X_{k-i_m})\mathbf{E}(X_{k-i_{m+1}}\ldots X_{k-i_{p-2}})\mathbf{E}(f(S_{k-i_{p-2}-r_m-1}))|].
\end{aligned} \tag{29}$$

To give a bound for the sums over k, m and r_m of the RHS of (29), we use Lemma 9 in order to evaluate the first term, and for the other terms we use lemmas 6 and 7 and, respectively the inductive assumption, inequality (27) and inequality (24). We thus obtain :

$$\begin{aligned}
&\sum_{k=1}^{n}\sum_{m=0}^{p-3}\sum_{r_m=0}^{k-1}\sum_{0\leq i_1\leq\ldots\leq i_{p-2}\leq k-1}^{(m),r_m} |J_1(i_1,\ldots,i_{p-2})| \\
&\leq C_p\{\sum_{m=2}^{p-2} M_{m,n}\ M_{p-m,n}(b_1,b_2)+M_{p,n}(b_1,b_2)+s_n^p(b_1\wedge b_2s_n)\}.
\end{aligned} \tag{30}$$

Inequalities (22), (28), (30) entail

$$\begin{aligned}
&\sum_{k=1}^{n}\sum_{0\leq i_1\leq\ldots\leq i_{p-2}\leq k-1} \left|\mathbf{E}X_kX_{k-i_1}\cdots X_{k-i_{p-2}}f(S_{k-i_{p-2}-1})\right| \\
&\leq C_p\left\{s_n^p(b_1\wedge b_2s_n)+\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(b_1,b_2)+M_{p,n}(b_1,b_2)\right\}.
\end{aligned}$$

The last inequality proves that the quantity in (18) is bounded by the RHS of (17).
Let us now give a bound for (19).

Study of the bound of (19):

We obtain, using again the fact $|\mathbf{E}(XY)|\leq|\mathrm{Cov}(X,Y)|+|\mathbf{E}X||\mathbf{E}Y|$, a Taylor expansion, the inductive assumption and Lemma 9 :

$$\begin{aligned}
&\sum_{k=1}^{n}\sup_{0\leq u\leq 1}\sum_{i_1\leq\ldots\leq i_{p-2}\leq k-1} \left|\mathbf{E}X_kX_{k-i_1}\cdots X_{k-i_{p-2}}\Delta_{p-2,k}f\right| \\
&\leq C_p\left\{s_n^p(b_1\wedge b_2s_n)+\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(b_1,b_2)+M_{p,n}(b_1,b_2)\right\}.
\end{aligned}$$

This last bound proves that the quantity in (19) is bounded by the RHS of (17), for all integer $p\geq 2$.
□

6.3 Step 3 : End of the proof of Lemma 3

We are now in position to conclude. Let us first recall that the function $x \to g_p(t,x)$ belongs to the set $\mathcal{F}_p(b_p,b_{p+1})$, with $b_p = t$ and $b_{p+1} = 1$.
We note also that if f belongs to the set $\mathcal{F}_p(b_p,b_{p+1})$, then $f^{(p-1)} \in \mathcal{F}_1(b_1,b_2)$ with $b_1 = b_p$ and $b_2 = b_{p+1}$.

As $p \geq 2$, $(t \wedge 1)\sum_{i=0}^{n-1}(n-i)\theta_i \leq n\sum_{i=0}^{n-1}(t \wedge (i+1))(i+1)^{p-2}\theta_i = M_{p,n}(t,1)$.

Hence we deduce combining Lemmas 5 and 8

$$\mathbf{E}g_p(t,S_n) \leq C_p\left\{\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(t,1) + M_{p,n}(t,1) + s_n^p(t \wedge s_n)\right\}. \tag{31}$$

Let $\Phi_p \in \mathbf{\Phi_p}$. Taking into account Lemma 4 and the fact $g^{(p)}(x) = x \wedge t$, we deduce that

$$\phi_p^{(p)}(x) = \int_0^{+\infty}(t \wedge x)\nu_p(dt). \tag{32}$$

Inequalities (13), (31) and (32) yield :

$$\begin{aligned}\mathbf{E}\phi_p(|S_n|) \quad &\leq \quad C_p\left\{\sum_{m=2}^{p-2} M_{m,n}\left(n\sum_{i=0}^{n-1}\phi_p^{(p)}(i+1)(i+1)^{p-m-2}\theta_i\right)\right. \\ &+ \quad \left. n\sum_{i=0}^{n-1}\phi_p^{(p)}(i+1)(i+1)^{p-2}\theta_i + s_n^p\phi_p^{(p)}(s_n)\right\}.\end{aligned} \tag{33}$$

Now, using the concavity of $\phi_p^{(p)}$ yields :

$$\phi_p(x) = \frac{x^p}{(p-1)!}\int_0^1(1-t)^{p-1}\phi_p^{(p)}(tx)dt \geq x^p\phi_p^{(p)}(x)\int_0^1\frac{t(1-t)^{p-1}}{(p-1)!}dt.$$

Hence we deduce that

$$x^p\phi_p^{(p)}(x) \leq C_p\phi_p(x). \tag{34}$$

We conclude, combining inequalities (33) and (34) :

$$\begin{aligned}\mathbf{E}\phi_p(|S_n|) \quad &\leq \quad C_p\left\{\sum_{m=2}^{p-2} M_{m,n}\left(n\sum_{i=0}^{n-1}\phi_p(i+1)(i+1)^{-m-2}\theta_i\right)\right. \\ &\left. + \, n\sum_{i=0}^{n-1}\phi_p(i+1)(i+1)^{-2}\theta_i + \phi_p(s_n)\right\}\end{aligned}$$

Finally, for the suitable choice of the function ϕ_p

$$\phi_p(x) = x^r, \quad \text{for} \quad r \in]p,p+1],$$

and using inequality (14) in Lemma 7, we finish the proof of Lemma 3.

7 Appendix A

This appendix is devoted to the proofs of Theorem 2 and Lemma 2.

Theorem 6 in Doukhan & Louhichi [21] proves the fi-di convergence in another weak dependence frame. Following their approach but replacing their dependence conditions by (SWD) sequences, we have the following result: fi-di convergence holds if (X_i) is (SWD) with $\theta_r = \mathcal{O}(r^{-a})$ with $a \geq 4$, which is enough to yield the Empirical Functional Central Limit Theorem (Theorem 1). However, we prove here a general Central Limit Theorem (Theorem 2 stated in section 2) under weak dependence and we apply it to prove the fi-di convergence. It allows us to weaken assumption required for the fi-di convergence to $\theta_r = \mathcal{O}\left(r^{-a}\right)$ with $a > 3$ (Lemma 2). To prove Theorem 2 we do not use Bernstein's blocks as in Doukhan & Louhichi [21] but a variation on the Lindeberg-Rio method (Rio, 1996).

Proof of Theorem 2: In the following, if $x \in \mathbb{R}_+$, $[x]$ denotes the integer part of x.

We define $S_n = X_0 + \cdots + X_{n-1}$, and $S_k = X_0 + \cdots + X_{k-1}$. Let $M := \sup_{n\in\mathbb{N}} ||X_n||_\infty$.

Let $v_k = \mathrm{Var}\left(S_k\right) - \mathrm{Var}\left(S_{k-1}\right)$. $(S_0 = 0)$.

By assumption, $\sigma_n^2 = \frac{\mathrm{Var}\left(S_n\right)}{n} \xrightarrow[n\infty]{} \sigma^2 > 0$.

We have $v_k = \sum_{i=-(k-1)}^{+(k-1)} \mathrm{Cov}\left(X_0, X_{|i|}\right)$.

As $(X_n)_{n\in\mathbb{N}}$ is (SWD) and uniformly bounded by M, we have $\left|\mathrm{Cov}\left(X_0, X_{|i|}\right)\right| \leq M\, \theta_{|i|}$.

By assumption, $\theta_r = \mathcal{O}\left((r+1)^{-a}\right)$ with $a > 3/2$. Hence $\sum_{i\geq 0} \theta_i < \infty$ and then $v_k \xrightarrow[k\infty]{} \sum_{i=-\infty}^{+\infty} \mathrm{Cov}\left(X_0, X_{|i|}\right) < \infty$.

Césaro's Theorem yields $\sigma^2 := \lim_{n\infty} \frac{\mathrm{Var}\left(S_n\right)}{n} = \lim_{n\infty} \frac{\sum_{k=1}^n v_k}{n} = \lim_{k\infty} v_k$.

Hence there exists some positive integer k_0 such that

$$k \geq k_0 \Rightarrow v_k \geq \frac{\sigma^2}{2} > 0. \tag{35}$$

Then set $Y_k \sim \mathcal{N}(0, v_{k+1})$, $k \geq k_0 - 1$. The sequence $(Y_k)_{k\geq k_0-1}$ is assumed to be independent and independent of the sequence $(X_k)_{k\in\mathbb{N}}$. We also set, for $k \geq k_0 - 2$, $T_{k,n} = \sum_{j=k+1}^{n-1} Y_j$; empty sums are, as usual, set equal to 0.

Let $\eta \sim \mathcal{N}\left(0,1\right)$, and $h_{k,n}(x) = \mathbf{E}h(x + T_{k,n}) := \mathbf{E}1_{y\geq x+T_{k,n}}$, where $y \in \mathbb{R}$. In the following, C will denote some arbitrary constant which may vary from line to line but which is independent from y,k,n. We are in position to use Rio's decomposition. We define

$$\Delta_n(h_{k,n}) = \sum_{k=k_0-1}^{\left[\frac{n}{3}\right]} \Delta_{k,n}(h_{k,n}), \tag{36}$$

with $\Delta_{k,n}(h_{k,n}) = \mathbf{E}\left(h_{k,n}(S_k + X_k) - h_{k,n}(S_k + Y_k)\right)$.

Then

$$\Delta_n(h_{k,n}) = \mathbf{E}\left(h\left(S_{\left[\frac{n}{3}\right]+1} + T_{\left[\frac{n}{3}\right],n}\right) - h\left(S_{k_0-1} + T_{k_0-2,n}\right)\right), \text{ where } h(x) := 1_{y\geq x}. \tag{37}$$

Assume that there exists $\alpha_n \xrightarrow[n\to\infty]{} 0$ such that for all y,

$$|\Delta_n(h_{k,n})| \leq \alpha_n. \tag{38}$$

Setting $y = \sqrt{n}\, z$, we get for all z,

$$\left| P\left(\frac{S_{\left[\frac{n}{3}\right]+1} + T_{\left[\frac{n}{3}\right],n}}{\sqrt{n}} \leq z\right) - P\left(\frac{S_{k_0-1} + T_{k_0-2,n}}{\sqrt{n}} \leq z\right)\right| \leq \alpha_n. \tag{39}$$

Then, as $\frac{S_{k_0-1}+T_{k_0-2,n}}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right)$,

$$\frac{S_{\left[\frac{n}{3}\right]+1}}{\sqrt{n}} + \frac{Y_{\left[\frac{n}{3}\right]+1} + \cdots + Y_{n-1}}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right). \tag{40}$$

Moreover

$$\frac{Y_{\left[\frac{n}{3}\right]+1} + \cdots + Y_{n-1}}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \frac{2}{3}\sigma^2\right). \tag{41}$$

(40) and (41) yield

$$\frac{S_{[\frac{n}{3}]+1}}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\frac{\sigma^2}{3}\right). \tag{42}$$

Finally (42) yields

$$\frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right). \tag{43}$$

Let us now prove (38).

To bound the terms $\Delta_{k,n}(h_{k,n})$ one writes

$\Delta_{k,n}(h_{k,n}) = \Delta^{(1)}_{k,n}(h_{k,n}) - \Delta^{(2)}_{k,n}(h_{k,n})$, with

$$\Delta^{(1)}_{k,n}(h) = \mathbf{E}h(S_{k+1}+T_{k,n}) - \mathbf{E}h(S_k+T_{k,n}) - \frac{v_{k+1}}{2}\mathbf{E}h"(S_k+T_{k,n}), \tag{44}$$

$$\Delta^{(2)}_{k,n}(h) = \mathbf{E}h(S_k+Y_k+T_{k,n}) - \mathbf{E}h(S_k+T_{k,n}) - \frac{v_{k+1}}{2}\mathbf{E}h"(S_k+T_{k,n}). \tag{45}$$

Study of $\Delta^{(2)}_{k,n}(h)$:

Using a Taylor's decomposition we get

$$\begin{aligned}\Delta^{(2)}_{k,n}(h_{k,n}) &= \mathbf{E}\left\{h'_{k,n}(S_k)\,Y_k\right\}\\ &\quad+\mathbf{E}\left\{\tfrac{1}{2}h''_{k,n}(S_k)\left(Y_k^2-v_{k+1}\right)+\tfrac{1}{6}h^{(3)}_{k,n}(S_k+\rho_{k,n}Y_k)\,Y_k^3\right\},\end{aligned}$$

where $0<\rho_{k,n}<1$. Using the independence of the process $(Y_n)_{n\in\mathbb{N}}$ and the process $(X_n)_{n\in\mathbb{N}}$, one deduces that,

$$\Delta^{(2)}_{k,n}(h_{k,n}) = \frac{1}{6}\mathbf{E}\left\{h^{(3)}_{k,n}(S_k+\rho_{k,n}Y_k)\,Y_k^3\right\}.$$

Hence

$$\left|\Delta^{(2)}_{k,n}(h_{k,n})\right| \le ||h^{(3)}_{k,n}||_\infty\,\frac{2v_{k+1}^{3/2}}{3\sqrt{2\pi}} \le ||h^{(3)}_{k,n}||_\infty\; C\;\left(\sum_{i=0}^{k}\theta_i\right)^{\frac{3}{2}}. \tag{46}$$

Study of $\Delta^{(1)}_{k,n}(h)$:

Set $\Delta^{(1)}_{k,n}(h_{k,n}) = \mathbf{E}\delta^{(1)}_{k,n}(h_{k,n})$, then, from Taylor's formula again (with some random $\tau_{k,n}\in]0,1[$), we write

$\delta^{(1)}_{k,n}(h_{k,n}) = X_k h'_{k,n}(S_k) + \frac{1}{2}h"_{k,n}(S_k)(X_k^2-v_{k+1}) + \frac{1}{6}\left(h^{(3)}_{k,n}(S_k+\tau_{k,n}X_k)X_k^3\right)$.

We analyze separatedely the terms in the previous expression

$$\frac{1}{6}|\mathbf{E}h^{(3)}_{k,n}(S_k+\tau_{k,n}X_k)X_k^3| \le M^3\;\frac{||h^{(3)}_{k,n}||_\infty}{6} \le C\;||h^{(3)}_{k,n}||_\infty. \tag{47}$$

We then write

$$\mathbf{E}\left\{X_k h'_{k,n}(S_k) + \frac{1}{2}h"_{k,n}(S_k)(X_k^2-v_{k+1})\right\}$$

$$= \mathbf{E}\left\{X_k h'_{k,n}(S_k)\right\} + \frac{1}{2}\mathrm{Cov}\left(h''_{k,n}(S_k),X_k^2\right) - \left(\mathbf{E}h''_{k,n}(S_k)\right)\sum_{i=1}^{k}\mathbf{E}\left(X_0X_i\right). \tag{48}$$

We set $X_p = 0$ for all $p<0$ and $S_0 = 0$.

We then have

$$g(S_i) - g(0) = \sum_{j=1}^{i} \left(g(S_j) - g(S_{j-1}) \right). \tag{49}$$

Hence, using (49), a Taylor's decomposition and the independence properties of the sequence $(Y_k)_{k \geq k_0 - 1}$ we get

$$\left| \mathrm{Cov}\left(h''_{k,n}\left(S_k\right), X_k^2 \right) \right| \leq C \, ||h^{(3)}_{k,n}||_\infty \sum_{i=1}^{k} \theta_i \tag{50}$$

and

$$X_k h'_{k,n}(S_k) = X_k \sum_{i=1}^{k} \left\{ X_{i-1} h''_{k,n}(S_{i-1}) + \frac{X_{i-1}^2}{2} h^{(3)}_{k,n}(S_{i-1} + \nu_{k,n,i} X_{i-1}) \right\}, \tag{51}$$

where $0 < \nu_{k,n,i} < 1$.

Moreover using the weak dependence of $(X_n)_{n \in \mathbb{N}}$,

$$\left| \mathbf{E} X_k \frac{X_{i-1}^2}{2} h^{(3)}_{k,n}\left(S_{i-1} + \nu_{k,n} X_{i-1} \right) \right| \leq C \, ||h^{(3)}_{k,n}||_\infty \theta_{k-i+1}. \tag{52}$$

Let now $j := \sup(0, 2i - k)$.

We write

$$\begin{aligned} \mathbf{E} X_k X_{i-1} h''_{k,n}(S_{i-1}) \quad = \quad & \mathbf{E} \left\{ h''_{k,n}(S_{j-1}) X_k X_{i-1} \right\} \\ & + \mathbf{E} \left\{ \left(h''_{k,n}(S_{i-1}) - h''_{k,n}(S_{j-1}) \right) X_k X_{i-1} \right\}. \end{aligned}$$

We also have, using a Taylor's decomposition

$$|h''_{k,n}(S_{i-1}) - h''_{k,n}(S_{j-1})| \leq C \, ||h^{(3)}_{k,n}||_\infty (k - i).$$

Hence,

$$\begin{aligned} & \left| \mathbf{E} \left\{ \left(h''_{k,n}(S_{i-1}) - h''_{k,n}(S_{j-1}) \right) X_k X_{i-1} \right\} \right| \\ & \leq C \quad ||h^{(3)}_{k,n}||_\infty \ \left(\ k - i \ \right) \theta_{k-i+1} \\ & \leq C \, ||h^{(3)}_{k,n}||_\infty (k - i + 1) \theta_{k-i+1}. \end{aligned} \tag{53}$$

Then, still using a Taylor's decomposition

$$\left| \mathrm{Cov}\left(h''_{k,n}(S_{j-1}), X_k X_{i-1} \right) \right| \leq C ||h^{(3)}_{k,n}||_\infty \sum_{l=1}^{j-1} \theta_{i-l}. \tag{54}$$

We can verify that

$$\sum_{i=1}^{k} \sum_{l=1}^{j-1} \theta_{i-l} = \sum_{i=1}^{\left[\frac{k}{2}\right]} i\theta_i + \sum_{i=\left[\frac{k}{2}\right]+1}^{k-1} (k - i)\theta_i \leq \sum_{i=1}^{k-1} i\theta_i. \tag{55}$$

Then, using inequalities (51), (52), (53), (54) and (55),

$$\left| \mathbf{E}\left(X_k h'_{k,n}(S_k) \right) - \sum_{i=1}^{k} \mathbf{E}\left(h''_{k,n}(S_{j-1}) \, \mathbf{E}\left(X_k X_{i-1} \right) \right| \leq C ||h^{(3)}_{k,n}||_\infty \sum_{p=1}^{k} p \, \theta_p. \tag{56}$$

It remains to bound

$$\left\{\sum_{i=1}^{k}\mathbf{E}h''_{k,n}(S_{j-1})\mathbf{E}\left(X_kX_{i-1}\right)\right\}-\left\{\mathbf{E}h''_{k,n}(S_k)\sum_{i=1}^{k}\mathbf{E}\left(X_0X_i\right)\right\}.$$

We can rewrite it

$$\sum_{i=1}^{k}\mathbf{E}\left(h''_{k,n}(S_{j-1})-h''_{k,n}(S_k)\right)\mathbf{E}\left(X_kX_{i-1}\right).$$

Using a Taylor's decomposition and as $(k-j+1)\leq 2\,(k-i+1)$,

$$\left|\sum_{i=1}^{k}\mathbf{E}\left(h''_{k,n}(S_{j-1})-h''_{k,n}(S_k)\right)\mathbf{E}\left(X_kX_{i-1}\right)\right|\leq C||h^{(3)}_{k,n}||_\infty\sum_{i=1}^{k}i\,\theta_i. \tag{57}$$

Then, summing inequalities (47), (50), (56) and (57),

$$\left|\Delta^{(1)}_{k,n}(h_{k,n})\right|\leq C\;||h^{(3)}_{k,n}||_\infty\left(1+\sum_{i=1}^{k}\theta_i+\sum_{i=1}^{k}i\;\theta_i\right)\leq C\;||h^{(3)}_{k,n}||_\infty\left(1+\sum_{i=1}^{k}i\;\theta_i\right). \tag{58}$$

Hence using inequalities (46) and (58),

$$\left|\sum_{k=k_0-1}^{\left[\frac{n}{3}\right]}\Delta_{k,n}(h_{k,n})\right|\leq C\;\left(\left[\frac{n}{3}\right]-k_0+2\right)\sup_{k_0-1\leq k\leq\left[\frac{n}{3}\right]}||h^{(3)}_{k,n}||_\infty\left(1+\sum_{i=1}^{\left[\frac{n}{3}\right]}i\;\theta_i\right). \tag{59}$$

We now have to bound $\sup_{k_0-1\leq k\leq\left[\frac{n}{3}\right]}||h^{(3)}_{k,n}||_\infty$.

If $\Phi_1(z):=\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$, we can prove with (35) that

$$\sup_{k_0-1\leq k\leq\left[\frac{n}{3}\right]}||h^{(3)}_{k,n}||_\infty\leq C\;n^{\frac{-3}{2}}\sigma^{-3}||\Phi''_1||_\infty, \tag{60}$$

where C is independent of y (see also Rio, 1996).

Hence

$$\left|\sum_{k=k_0-1}^{\left[\frac{n}{3}\right]}\Delta_{k,n}(h_{k,n})\right|\leq C\;n^{\frac{-1}{2}}\sigma^{-3}||\Phi''_1||_\infty\left(1+\sum_{i=1}^{n}i\;\theta_i\right). \tag{61}$$

Let denote α_n the right hand side of (61). α_n does not depend on y and it tends to 0 as soon as $a>3/2$. Hence, by (43), this concludes the proof of Theorem 2. □

We are now in position to prove the fi-di convergence. We prove Lemma 2 for one-dimensional marginals. If one wants to prove it for any finidimensional marginals, it is sufficient to use the following proof with

$$Z_k(t_1,\ldots,t_r)=\sum_{j=1}^{r}\alpha_j\left(1_{X_k\leq t_j}-F(t_j)\right),$$

where $(t_1,\ldots,t_r)\in\mathbb{R}^r$ and for arbitrary numbers $\alpha_1,\ldots,\alpha_r$. To prove the convergence of

$(Z_n(t))_{n\in\mathbb{N}}:=\left(1_{X_n\leq t}-F(t)\right)_{n\in\mathbb{N}}$ for some $t\in\mathbb{R}$, we apply Theorem 2 to the sequence $(Z_n(t))_{n\in\mathbb{N}}$. We first need the following proposition:

Proposition 1 *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence with common distribution function F supposed to be Lipschitz. Suppose that $(X_n)_{n\in\mathbb{N}}$ satisfies the $(SWD)-$weak dependence condition, with a sequence $(\theta_r)_{r\in\mathbb{N}}$. Then the sequence $(Z_n(t))_{n\in\mathbb{N}} := (1_{X_n\le t} - F(t))_{n\in\mathbb{N}}$, which is uniformly bounded, satisfies the $(SWD)-$weak dependence condition, with a sequence $(\theta'_r)_{r\in\mathbb{N}}$ such that there exists some positive constant D satisfying $\theta'_r \le D\sqrt{\theta_r}$.*

Proof of Proposition 1:

In the following, C will denote some arbitrary constant which may vary from line to line but which is independent from u,v,r. Let $t\in\mathbb{R}$. We study the sequence $(Z_n(t))_{n\in\mathbb{N}}$. We want to prove that this sequence satisfies a $(SWD)-$dependence condition. For this we have to bound $C_{h,u,v} := \mathrm{Cov}\left(h(Z_{i_1},\ldots,Z_{i_u}),k\left(Z_{j_1},\ldots,Z_{j_v}\right)\right)$ for all $i_1\le\cdots\le i_u\le i_u+r\le j_1\le\cdots\le j_v$ and for all $h\in\mathbb{L}^\infty$, $k\in\mathcal{L}$. We write Z_i for $Z_i(t)$. Let us write $Z_n=\varphi(X_n)$. We want to smooth the function φ which is not Lipschitz. Let $\varepsilon>0$. We consider the following Lipschitz function φ^ϵ smoothing φ :

φ_ε is equal to φ on $[0,t-\varepsilon]\cup]t+\varepsilon,1]$ and for $t-\varepsilon<x\le t+\varepsilon$,

$\varphi_\varepsilon(x)=\frac{1}{4\varepsilon^3}\left(x^3-3tx^2+3(t^2-\varepsilon^2)x-t^3+3t\varepsilon^2+2\varepsilon^3\right)-F(t)$.

We then have $||\varphi^\varepsilon||_\infty\le||\varphi||_\infty\le 1$ and $\mathrm{Lip}(\varphi^\varepsilon)\le\frac{3}{4\,\varepsilon}$.

Hence,

$$\begin{aligned}
|C_{h,u,v}| &\le |\mathrm{Cov}\left(h(Z_{i_1},\ldots,Z_{i_u}),k\left(\varphi(X_{j_1}),\ldots,\varphi(X_{j_v})\right)-k\left(\varphi^\varepsilon(X_{j_1}),\ldots,\varphi^\varepsilon(X_{j_v})\right)\right)| \\
&\quad +|\mathrm{Cov}\left(h(Z_{i_1},\ldots,Z_{i_u}),k\left(\varphi^\varepsilon(X_{j_1}),\ldots,\varphi^\varepsilon(X_{j_v})\right)\right)| \\
&\le 2\,||h||_\infty\mathrm{Lip}(k)\mathbf{E}\textstyle\sum_{i=1}^v|\varphi(X_{j_i})-\varphi^\varepsilon(X_{j_i})|+||h||_\infty\mathrm{Lip}(k)\mathrm{Lip}(\varphi^\varepsilon)\;\theta_r\;v \\
&\le 2||h||_\infty\mathrm{Lip}(k)\;v\;2||\varphi||_\infty P\left(X_0\in]t-\varepsilon,t+\varepsilon]\right)+||h||_\infty\mathrm{Lip}(k)\mathrm{Lip}(\varphi^\varepsilon)\;\theta_r v \\
&\le ||h||_\infty\mathrm{Lip}(k)v\left(8\varepsilon\mathrm{Lip}(F)+\tfrac{3}{4\varepsilon}\theta_r\right)\le C\;||h||_\infty\mathrm{Lip}(k)v\left(\varepsilon+\tfrac{\theta_r}{\varepsilon}\right).
\end{aligned}$$

Then if we take $\varepsilon=\sqrt{\theta_r}$,

$$|C_{h,u,v}|\le C\;||h||_\infty\mathrm{Lip}(k)\;v\;\sqrt{\theta_r}. \tag{62}$$

So $(Y_n)_{n\in\mathbb{N}}$ is (SWD) with $\theta'_r\le C\;\sqrt{\theta_r}$. This concludes the proof of Proposition 1. □

Now, let $S_k(t)=\sum_{i=0}^{k-1}Z_i(t)$ for $1\le i\le n$, $S_0(t)=0$.

Let us prove that $\frac{\mathrm{Var}\,(S_n(t))}{n}\xrightarrow[n\infty]{}\sigma^2\ge 0$.

Let $v_k(t)=\mathrm{Var}\;(S_k(t))-\mathrm{Var}\;(S_{k-1}(t))$.

We have $v_k(t)\xrightarrow[k\infty]{}\sigma^2(t)=\sum_{i=-\infty}^{+\infty}\mathrm{Cov}\left(Z_0(t),Z_{|i|}(t)\right)<\infty$ as soon as $\theta_r=\mathcal{O}\left((r+1)^{-2-\delta}\right)$ with $\delta>0$. Césaro's Theorem yields then $\frac{\mathrm{Var}\,(S_n(t))}{n}=\frac{\sum_{k=1}^n v_k(t)}{n}\xrightarrow[n\infty]{}\sigma^2(t)$.

If $\sigma^2(t)=0$, the convergence of $\frac{S_n(t)}{n}$ is a simple fact. If $\sigma^2(t)>0$, thanks to Proposition 1, we are in position to apply Theorem 2 to $(Z_n(t))_{n\in\mathbb{N}}$ to obtain the fi-di convergence as soon as $\theta_r=\mathcal{O}\left((r+1)^{-a}\right)$ with $a>3$. This concludes the proof of Lemma 2.

8 Appendix B

Proof of Lemma 5 By induction on $p\ge 2$, and using the following decomposition

$$\text{for}\;\; g(0)=0,\;\; g(S_n)=\sum_{k=1}^n[g(S_k)-g(S_{k-1})]=\sum_{k=1}^n X_k g'(S_{k-1}+uX_k), \tag{63}$$

for some $0 < u < 1$, we show that, for any $f \in \mathcal{C}_{p-1}$,

$$|\mathbf{E}(f(S_n))| \leq C_p \sum_{k=1}^{n} \left[\mathbf{E}_{p-2,k}(f^{(p-1)}) + \mathbf{E}_{p-2,k}(\Delta f^{(p-1)}) \right] + \sum_{k=1}^{n} \mathbb{R}_{p,k}(f), \tag{64}$$

where

$$\mathbb{R}_{p,k}(f) := \left| \mathbf{E}\left[f(S_k) - f(S_{k-1}) - \cdots - \frac{{X_k}^{p-1}}{(p-1)!} f^{(p-1)}(S_{k-1}) \right] \right|.$$

Let now f be a fixed function of the set $\mathcal{F}_p(b_p, b_{p+1})$.
In view of (64), Lemma 5 is proved if we suitably evaluate $\mathbb{R}_{p,k}(f)$. Clearly

$$\mathbb{R}_{p,k}(f) \leq \mathbb{R}_{p+1,k}(f) + |\mathrm{Cov}(X_k^p, f^{(p)}(S_{k-1}))| + \mathbf{E}(|X_k^p|)\mathbf{E}|f^{(p)}|(S_{k-1}) =: I_{1,k} + I_{2,k} + I_{3,k}.$$

Taylor's formula and the (SWD) property yield :

$$I_{1,k} \leq C_p \mathbf{E}\left(|X_k|^p (b_p \wedge b_{p+1}|X_k|)\right),$$

$$I_{2,k} \leq C_p \left[b_p \mathbf{E}|X_k|^2 \wedge b_{p+1} \sum_{i=0}^{k-1} \theta_i \right] \leq C_p (b_p \wedge b_{p+1}) \sum_{i=0}^{k-1} \theta_i.$$

Finally

$$\begin{aligned} I_{3,k} &\leq C_p \ \mathbf{E}|X_k^p|(b_p \wedge b_{p+1} s_n) && (65)\\ &\leq C_p \ \mathbf{E}|X_k^p|(b_p \wedge b_{p+1} s_n) 1_{|X_k| \leq s_n} + C_p \ \mathbf{E}|X_k^p|(b_p \wedge b_{p+1} s_n) 1_{|X_k| > s_n} && (66)\\ &\leq C_p \ s_n^{p-2}(b_p \wedge b_{p+1} s_n)\mathbf{E}(X_k^2) + C_p \ \mathbf{E}\left(|X_k|^p (b_p \wedge b_{p+1}|X_k|)\right). && (67) \end{aligned}$$

The proof of Lemma 5 is complete by noting that, as $(X_n)_{n\in\mathbb{N}}$ is bounded by 1,

$$\sum_{k=1}^{n} \mathbf{E}\left(|X_k|^p (b_p \wedge b_{p+1}|X_k|)\right) \leq (b_p \wedge b_{p+1}) n \theta_0,$$

and $\sum_{k=1}^{n} \mathbf{E}(X_k^2) \leq s_n^2$. $\square$

Proof of Lemma 7 For all positive integer m, and all positive real numbers α and β that are conjugate (i.e $1/\alpha + 1/\beta = 1$) there holds, thanks to Hölder inequality

$$\sum_{i=0}^{n-1} (i+1)^{m-2} \theta_i \leq \left(\sum_{i=0}^{n-1} \theta_i \right)^{1/\alpha} \left(\sum_{i=0}^{n-1} (i+1)^{(m-2)\beta} \theta_i \right)^{1/\beta}.$$

We obtain $M_{m,n} \leq s_n^{2/\alpha} \left(M_{p,n}\right)^{1/\beta}$,
hence, for $\alpha = \frac{p-2}{p-m}$ and $\beta = \frac{p-2}{m-2}$, we obtain $M_{m,n} \leq s_n^{2(p-m)/(p-2)} M_{p,n}^{(m-2)/(p-2)}$.

There holds also $M_{p-m,n} \leq s_n^{2m/(p-2)} M_{p,n}^{(p-m-2)/(p-2)}$.

The two last inequalities prove the first inequality in (14). Then, bounding $M_{p,n}$ and s_n^p by the sum $M_{p,n} + s_n^p$, we get the second inequality of (14).
Now, we only prove (15) ((16) will hold with the same method).
Clearly

$$s_n^{p-m}(b_1 \wedge b_2 s_n) \left(n \sum_{(i+1) \leq (s_n \wedge n)} (i+1)^{m-2} \theta_i \right)$$

$$\leq s_n^{p-m}(b_1 \wedge b_2 s_n) s_n^{m-2}\left(n \sum_{i=0}^{n-1} \theta_i\right) \leq s_n^p(b_1 \wedge b_2 s_n) \tag{68}$$

and

$$n \sum_{i:\ (i+1)\geq s_n}^{n-1} \left(s_n^{p-m}(b_1 \wedge b_2 s_n)(i+1)^{m-2}\theta_i\right) \leq M_{p,n}. \tag{69}$$

Collecting inequalities (68) and (69) we obtain (15). □

Proof of Lemma 9 We write

$$f'(S_{k-i_{p-1}-1} + uX_{k-i_{p-1}}) = [f'(S_{k-i_{p-1}-1} + uX_{k-i_{p-1}}) - f'(S_{k-i_{p-1}-1})] + f'(S_{k-i_{p-1}-1})$$

so the covariance quantity of Lemma 9 is decomposed into two terms :

$$|\mathrm{Cov}(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}} f'(S_{k-i_{p-1}-1}))| \tag{70}$$

and

$$|\mathrm{Cov}(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}}[f'(S_{k-i_{p-1}-1} + uX_{k-i_{p-1}}) - f'(S_{k-i_{p-1}-1})])|. \tag{71}$$

- Study of the terms of type (70):
  We use the following decomposition:

$$|\mathrm{Cov}(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}} f'(S_{k-i_{p-1}-1}))| \leq$$

$$\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}}[f'(S_{k-i_{p-1}-1}) - f'(S_{k-i_{p-1}-r_m-1})]\right)\right|$$

$$+\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}} f'(S_{k-i_{p-1}-r_m-1})\right)\right|.$$

  - The (SWD) property yields

$$|J(k,r_m,i_1,\ldots,i_{p-1})| :=$$

$$\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}}[f'(S_{k-i_{p-1}-1}) - f'(S_{k-i_{p-1}-r_m-1})]\right)\right|$$

$$\leq C_p\ (b_1 \wedge b_2 r_m)\ \theta_{r_m},$$

  where $r_m = i_{m+1} - i_m = \max_{1\leq q\leq p-2}(i_q - i_{q-1})$. Thus

$$I_{r_m} := \sum_{0=:i_0\leq i_1\leq\ldots\leq i_{p-1}}^{(m),r_m} C_p\ (b_1 \wedge b_2 r_m)\ \theta_{r_m} \leq C_p {r_m}^{p-2}\ (b_1 \wedge b_2 r_m)\ \theta_{r_m}.$$

  Taking the sum over r_m in the last inequality, we obtain

$$\sum_{r_m=0}^{k-1} I_{r_m} \leq \sum_{l=0}^{k-1} C_p\ (b_1 \wedge b_2 l)\ l^{p-2}\ \theta_l.$$

  Summing up over $m \in \{0,\ldots,p-2\}$, $k \in \{0,\ldots,n\}$ in the last inequality yields

$$\sum_{k=1}^{n} \sum_{m=0}^{p-2} \sum_{r_m=0}^{k-1} \sum_{0\leq(i)\leq k-1}^{(m),r_m} |J(k,r_m,i_1,\ldots,i_{p-1})|$$

$$\leq \sum_{k=1}^{n} \sum_{l=0}^{k-1} C_p \, l^{p-2} \, (b_1 \wedge b_2(l+1)) \, \theta_l \leq C_p \, M_{p,n}(b_1,b_2), \tag{72}$$

where (i) denotes $0 =: i_0 \leq i_1 \leq \ldots \leq i_{p-1} \leq k-1$.

– We now use the fact

$$\mathrm{Cov}(X,YZ) = \mathrm{Cov}(XY,Z) + \mathbf{E}(XY)\mathbf{E}(Z) - \mathbf{E}(X)\mathrm{Cov}(Y,Z) - \mathbf{E}(X)\mathbf{E}(Y)\mathbf{E}(Z),$$

to get

$$\begin{aligned}
&\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_m}, X_{k-i_{m+1}} \cdots X_{k-i_{p-1}} f'(S_{k-i_{p-1}-r_m-1})\right)\right| \\
&\leq \left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| \\
&\quad + \left|\mathbf{E}(X_k \cdots X_{k-i_{p-1}})\mathbf{E}(f'(S_{k-i_{p-1}-r_m-1}))\right| \\
&\quad + \left|\mathbf{E}(X_k \cdots X_{k-i_m})\mathrm{Cov}(X_{k-i_{m+1}} \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1}))\right| \\
&\quad + \left|\mathbf{E}(X_k \cdots X_{k-i_m})\mathbf{E}(X_{k-i_{m+1}} \cdots X_{k-i_{p-1}})\mathbf{E}(f'(S_{k-i_{p-1}-r_m-1}))\right| \\
&=: I_1(i_1, \ldots, i_{p-1}) + I_2(i_1, \ldots, i_{p-1}) + I_3(i_1, \ldots, i_{p-1}) + I_4(i_1, \ldots, i_{p-1}).
\end{aligned}$$

Using Lemmas 6, 7 and arguing as for inequality (24) yield

$$\begin{aligned}
&\sum_{k=1}^{n} \sum_{0=:i_0 \leq i_1 \leq \ldots \leq i_{p-1} \leq k-1} I_2(i_1, \ldots, i_{p-1}) + \sum_{k=1}^{n} \sum_{0=:i_0 \leq i_1 \leq \ldots \leq i_{p-1} \leq k-1} I_4(i_1, \ldots, i_{p-1}) \\
&\quad \leq C_p\{s_n^p(b_1 \wedge b_2 s_n) + M_{p,n}(b_1,b_2) + \sum_{m=2}^{p-2} M_{m,n} M_{p-m,n}(b_1,b_2)\}.
\end{aligned} \tag{73}$$

Let us now study $I_1(i_1, \ldots, i_{p-1})$.
On one hand we have, using the SWD-property:

$$\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| \leq C_p \, b_1 \, \theta_{r_m}. \tag{74}$$

On the other hand, using Taylor's decomposition, we get:

$$\begin{aligned}
&\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| \qquad (75)\\
&= \left|\textstyle\sum_{i_p=i_{p-1}+r_m+1}^{k-1} \mathrm{Cov}(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_p}) - f'(S_{k-i_p-1}))\right| \\
&\leq \textstyle\sum_{i_p=i_{p-1}+r_m+1}^{k-1} C_p \, \theta_{i_p-i_{p-1}} \, b_2 \leq \sum_{l=r_m+1}^{k-1} C_p \, \theta_l \, b_2.
\end{aligned}$$

Moreover we write:

$$\begin{aligned}
&\left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| \\
&= \left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| 1_{b_1 \leq r_m b_2} \\
&\quad + \left|\mathrm{Cov}\left(X_k \cdots X_{k-i_{p-1}}, f'(S_{k-i_{p-1}-r_m-1})\right)\right| 1_{b_1 > r_m b_2} \\
&=: C_1(i_1, \cdots, i_{p-1}, r_m) + C_2(i_1, \cdots, i_{p-1}, r_m).
\end{aligned}$$

Inequalities (74) and (75) yield

$$\sum_{k=1}^{n}\sum_{m=0}^{p-2}\sum_{r_m=0}^{k-1}\sum_{0=:i_0\le i_1\le\cdots\le i_{p-1}\le k-1}^{(m),r_m} C_1(i_1,\cdots,i_{p-1},r_m)\le C_p\ M_{p,n}(b_1,b_2).$$

If now we note that:

$$\sum_{r_m=0,r_m<\frac{b_1}{b_2}}^{k-1}\ \sum_{l=r_m+1}^{k-1} l^{p-2}\,\theta_l\le\sum_{l=1}^{k-1}\ \sum_{r_m=0}^{l\wedge(b_1/b_2)} l^{p-2}\,\theta_l,$$

using once more inequality (75) we also get

$$\sum_{k=1}^{n}\sum_{m=0}^{p-2}\sum_{r_m=0}^{k-1}\sum_{0=:i_0\le i_1\le\cdots\le i_{p-1}\le k-1}^{(m),r_m} C_2(i_1,\cdots,i_{p-1},r_m)\le C_p\ M_{p,n}(b_1,b_2).$$

So,

$$\sum_{k=1}^{n}\ \sum_{i_1\le\ldots\le i_{p-1}\le k-1} I_1(i_1,\ldots,i_{p-1})\le C_p\ M_{p,n}(b_1,b_2). \tag{76}$$

We still have to bound $I_3(i_1,\ldots,i_{p-1})$.
For this we use the (SWD) property, as for the bound of $I_1(i_1,\ldots,i_{p-1})$, and Lemmas 6 and 7. We obtain

$$\sum_{k=1}^{n}\sum_{(i)} I_3(i_1,\ldots,i_{p-1})\le C_p\{M_{p,n}(b_1,b_2)+s_n^p(b_1\wedge b_2 s_n)+\sum_{m=2}^{p-2} M_{m,n}M_{p-m,n}(b_1,b_2)\},$$

which provides a bound for the sum of the terms of type (70).

- Study of the terms of type (71): (i) still denotes $i_1\le\ldots\le i_{p-1}\le k-1$.
  Let us now give a bound for (71). With Taylor,

$$\begin{aligned}&|\mathrm{Cov}\left(X_k\cdots X_{k-i_m},X_{k-i_{m+1}}\cdots X_{k-i_{p-1}}[f'(S_{k-i_{p-1}-1}+uX_{k-i_{p-1}})-f'(S_{k-i_{p-1}-1})]\right)|\\&\qquad\le C_p(b_1\wedge b_2)\theta_{r_m}.\end{aligned}$$

So,

$$\begin{aligned}&\sum_{(k),(i)}|\mathrm{Cov}(X_k\cdots X_{k-i_m},X_{k-i_{m+1}}\cdots X_{k-i_{p-1}}[f'(S_{k-i_{p-1}-1}+uX_{k-i_{p-1}})-f'(S_{k-i_{p-1}-1})])|\\&\qquad\le n\ C_p\sum_{i=0}^{n-1} i^{p-2}(b_1\wedge b_2(i+1))\theta_i\le C_p\ M_{p,n}(b_1\wedge b_2),\end{aligned} \tag{77}$$

where (k) denotes $1\le k\le n$ and (i) denotes $i_1\le\ldots\le i_{p-1}\le k-1$.
Inequalities (72), (73), (76) and (77) yield the proof of Lemma 9. □

Acknowledgements: The author is grateful to Sana Louhichi for fruitful discussions and helpful suggestions on this work.

Chapitre 3

Problème de rupture

Dans ce chapitre, nous présentons un travail réalisé en collaboration avec Patrick Ango Nze. Nous nous intéressons à un problème d'estimation de rupture pour une fonction de régression dans un cadre de dépendance faible.

Estimation d'une rupture en dépendance faible

Patrick Ango Nze[1] et Clémentine Prieur[2]

Résumé – On se place dans un cadre de dépendance faible proposé par Doukhan et Louhichi dans [21]. On estime le saut au point de rupture, connu, de la fonction de régression du modèle $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$. La suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est indépendante et identiquement distribuée (i.i.d.) et indépendante de la suite faiblement dépendante $(X_n)_{n\in\mathbb{N}}$. On énonce un théorème de limite centrale du saut.

Change point estimation for a weakly dependent sequence

Abstract – Let $(X_n,Y_n)_{n\in\mathbb{N}}$ be a stationary sequence governed by the model $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ where $(\varepsilon_n)_{n\in\mathbb{N}}$ is i.i.d. and independent from $(X_n)_{n\in\mathbb{N}}$. The latter sequence satisfy a weak dependence condition proposed by Doukhan and Louhichi in [21]. We provide a Central Limit Theorem for jumps in the regression function. Our method deals with linear local regression described in [27]. We use a variation on Lindeberg-Rio [55].

Key words and phrases. Stationary sequences, Lindeberg Theorem, Central Limit Theorem, nonparametric regression, weakly dependent sequences, local linear regression, change point.

MSC (2000) 60 E 15, 60 F 05, 60 G 10, 60 G 99, 62 G 08, 62 J 05.

3.1 Introduction, notations

Dans cette note, nous nous intéressons au problème d'estimation d'une rupture dans un modèle de régression : $Y_n = m(X_n) + \sigma(X_n)\varepsilon_n$ où la suite $(\varepsilon_n)_{n\in\mathbb{N}}$ est i.i.d. et indépendante de la suite $(X_n)_{n\in\mathbb{N}}$. On reprend l'étude de ce modèle par Grégoire & Hamrouni [27] dans le cadre indépendant, mais dans un contexte de dépendance faible : la suite $(X_n)_{n\in\mathbb{N}}$ est supposée faiblement dépendante au sens de la définition 8. La première partie est consacrée à la présentation du modèle étudié. On suppose que la fonction de régression présente un unique point de discontinuité τ connu. On cherche à estimer l'amplitude du saut. La méthode d'estimation retenue est une méthode non paramétrique. On estime le saut au point τ, $\gamma(\tau) = m_+(\tau) - m_-(\tau)$, en utilisant un estimateur à droite pour $m_+(\tau)$ et un autre à gauche pour $m_-(\tau)$. Les estimateurs à droite et à gauche choisis sont des estimateurs de régression linéaire locale. L'article [27] motive l'utilisation de cette procédure en vue de résoudre des problèmes de rupture. Disons pour résumer que ce choix permet de gérer les effets de bords. En outre, sur le plan pratique, ce type d'estimateurs se comporte relativement bien pour des tailles d'échantillon modestes.

1. Université Lille 3, UFR AES, BP 149, 59653, Villeneuve d'Ascq Cedex, France. Courriel : angonze@univ-lille3.fr

2. Université de Cergy-Pontoise, Laboratoire de mathématiques, Bâtiment A4, Site Saint Martin, 95011, Cergy-Pontoise Cedex, France. Courriel : prieur@math.u-cergy.fr

Nous renvoyons à [27] pour des références complémentaires. Les résultats de convergence pour l'estimateur de l'amplitude du saut sont présentés dans la deuxième partie. On montre en particulier un théorème de limite centrale (théorème 9 et corollaire 1).

3.1.1 Hypothèses de dépendance faible

Commençons par introduire la notion de dépendance faible utilisée dans cette note. Pour un espace mesurable et normé $(E,\|.\|)$ on définit la classe $\mathbb{L}^\infty$ réunion de la famille, $\mathbb{L}^\infty(E^u)$ $(u \in \mathbb{N}^*)$, des fonctions numériques mesurables et bornées sur l'ensemble E^u. Ce dernier est muni de la norme l^1 :

$$\|x\|_1 = \|x_1\| + \cdots + \|x_u\|, \qquad x = (x_1,\ldots,x_u) \in E^u.$$

Le coefficient de Lipschitz d'une fonction $f : E^u \longrightarrow \mathbb{R}$ étant défini par

$$\mathrm{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1},$$

on s'attache aux fonctions appartenant à la classe

$$\mathcal{L} = \bigcup_{u=1}^\infty \{f \in \mathbb{L}^\infty(E^u) : \ \mathrm{Lip}(f) < \infty, \|f\|_\infty \leq 1\}.$$

La définition de dépendance faible qui suit reprend l'article [21].

Définition 8 *Une suite $(Z_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $E = \mathbb{R}^D$ sera dite $(\theta,\mathcal{L},\psi)$-faiblement dépendante s'il existe une suite $\theta = (\theta_r)_{r\in\mathbb{N}}$ décroissant vers zéro, une fonction numérique ψ définie pour les quadruplets $(f,g,u,v) \in \mathcal{L}^2 \times \mathbb{N}^{*2}$, telles que pour tout triplet $(u,v,r) \in (\mathbb{N}^*)^2 \times \mathbb{N}$*

$$|\mathrm{Cov}(f(Z_{i_1},\ldots,Z_{i_u}),g(Z_{j_1},\ldots,Z_{j_v}))| \leq \psi(f,g,u,v)\theta_r$$

soit vérifiée pour tout u-uplet $(i_1,\ldots,i_u)$ et tout v-uplet $(j_1,\ldots,j_v)$ dès lors que $i_1 \leq \ldots \leq i_u < i_u + r \leq j_1 \leq \ldots \leq j_v$.

Cette définition présente une alternative intéressante à des conditions de dépendance plus classiques (le mélange, par exemple). Parmi les choix les plus courants de $\psi(f,g,u,v)$, on retient

$$\psi_1(f,g,u,v) = u\,\mathrm{Lip}(f) + v\,\mathrm{Lip}(g), \ \psi_1'(f,g,u,v) = v\,\mathrm{Lip}(g).$$

La ψ_1 dépendance faible s'applique à des modèles markoviens généraux ainsi qu'aux schémas de Bernoulli. La fonction ψ_1' renvoie à l'analogue causal de ψ_1. L'article [21] propose une revue plus détaillée de tels modèles applicatifs. Le lecteur intéressé pourra le consulter avec profit.

3.1.2 Hypothèses sur le modèle

Les observations forment une suite $(X_n,Y_n)_{n\in\mathbb{N}}$ de v.a. à valeurs dans $\mathbb{R}^2$. Elles obéissent au modèle

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i$$

pour une fonction $m(\cdot)$ régulière partout sauf au point connu τ, une fonction $\sigma(\cdot)$ continûment dérivable, et une suite de v.a. à valeurs réelles $(\varepsilon_n)_{n\in\mathbb{N}}$ i.i.d. centrée réduite. On suppose de plus que les suites $(X_n)_{n\in\mathbb{N}}$ et $(\varepsilon_n)_{n\in\mathbb{N}}$ sont mutuellement indépendantes.

Les hypothèses suivantes seront utilisées par la suite.

1. $\mathbb{E}\,\varepsilon_0 = 0, \mathbb{E}\,(\varepsilon_0^2) = 1, \mathbb{E}\,|\varepsilon_0|^3 < \infty.$

2. La loi de X_0 est concentée sur l'intervalle [0,1]. Elle possède une densité f strictement positive et continue.
3. Les lois des couples (X_0,X_k) possèdent des densités $f_k(x,y)$ uniformément bornées par rapport à k :
$$\sup_{k\geq 1}\sup_{x,\,y}|f_k(x,y)| < \infty.$$
4. La fonction de régression $m(\cdot) = \mathbb{E}(Y_0|X_0 = \cdot)$ est deux fois continûment dérivable dans chaque intervalle $[0,\tau]$ et $[\tau,1]$. En particulier m et ses deux dérivées successives possèdent des limites à gauche et à droite au point τ.
5. La variance conditionnelle $\sigma^2(\cdot) = \mathrm{Var}\,(Y_0|X_0 = \cdot)$ est continûment dérivable.
6. Le point de saut τ de la fonction $m(\cdot)$ est dans l'intervalle $]0,1[$.

On définit, pour $\nu = 0,1$ ou 2,

(i) $m_-^{(\nu)}(\tau) = \lim_{t\nearrow\tau} m^{(\nu)}(t)$,

(ii) $m_+^{(\nu)}(\tau) = \lim_{t\searrow\tau} m^{(\nu)}(t)$,

(iii) $\gamma^{(\nu)}(\tau) = m_+^{(\nu)}(\tau) - m_-^{(\nu)}(\tau)$.

3.1.3 Estimateurs linéaires locaux

Les fonctions m et σ sont inconnues. On utilise des estimateurs à noyau pour les estimer. Les définitions qui suivent reprennent [25]. Soit K un noyau, et $h_n = h$ une fenêtre. Les paramètres réels $\hat{\alpha}_x$ et $\hat{\beta}_x$ sont obtenus en résolvant le problème de minimisation de

$$\sum_{i=1}^{n}(Y_i - \alpha_x - \beta_x(x - X_i))^2 K_i\,(x)\,, \tag{1}$$

avec $K_i\,(x) = K\left(\dfrac{x - X_i}{h}\right)$.

Définition 9 *L'estimateur linéaire local de $m(\cdot)$ est*

$$\widehat{m}(x) = \hat{\alpha}_x = \frac{\sum_{i=1}^{n}\omega_i(x)Y_i}{\sum_{i=1}^{n}\omega_i(x)}\;, \tag{2}$$

avec

$$\omega_i(x) = K_i\,(x)\,(S_2(x) - (x - X_i)S_1(x))$$

où, pour $l \in \mathbb{N}$,

$$S_l(x) = \sum_{i=1}^{n}(x - X_i)^l K_i\,(x)\,.$$

3.1.4 Estimateurs à droite et à gauche

On reprend ici des hypothèses de [27] pour faciliter la lecture du texte. On considère un noyau $K_+ : [-1,0] \to \mathbb{R}_+$ continûment différentiable. On pose alors

(i) $K_l^+ = \int_{-1}^{0} x^l K_+(x)dx$, et $L_l^+ = \int_{-1}^{0} x^l K_+^2(x)dx$, $l \in \mathbb{N}$,

(ii) $B_+ = (K_2^+)^2 - K_3^+ K_1^+$,

(iii) $V_+ = \int_{-1}^{0} (K_2^+ - xK_1^+)^2 K_+^2(x)dx$.

On définit $K_-(x) = K_+(-x)$, puis les quantités K_l^-,L_l^-,B_- et V_- relativement au noyau K_-. On a alors $V_+ = V_-$, $B_+ = B_-$ et

$$V_+ = (K_2^+)^2 L_0^+ + (K_1^+)^2 L_2^+ - 2K_1^+ K_2^+ L_1^+.$$

D'autre part, il n'y a aucune perte de généralité à imposer la normalisation $K_2 K_0 - K_1^2 = 1$. En effet, si K est un noyau quelconque, $K/\sqrt{K_2K_0 - K_1^2}$ vérifie la contrainte de normalisation. Les noyaux K_+ et K_- conduisent alors à des estimateurs linéaires locaux $\widehat{m}_+$ et $\widehat{m}_-$ définis par la formule (2). L'article [27] motive l'utilisation de l'estimateur $\widehat{\gamma} = \widehat{m}_+ - \widehat{m}_-$ en vue de résoudre des problèmes de rupture.

3.2 Résultats

On suppose vérifiées les hypothèses du paragraphe 3.1.2.

Théorème 7 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1)$-faiblement dépendant. Si $\theta_r = r^{-a}$ avec $a \geq 3$, si $h_n \to 0$, et $nh_n \to +\infty$, alors, pour tout $l \in \mathbb{N}$, et pour tout $x \in]0,1[$*

$$\begin{aligned}
S_l\left(x\right) &= nh^{l+1} f\left(x\right) K_l \left(1 + o_P(1)\right), \\
R_l\left(x\right) &= \sum_{i=1}^{n} \left(x - X_i\right)^l K_i\left(x\right) \Big(m\left(X_i\right) - m\left(x\right) - \left(X_i - x\right) m'(x)\Big) \\
&= \frac{1}{2} nh^{l+3} m''\left(x\right) f\left(x\right) K_{l+2}\left(1 + o_P(1)\right), \qquad x \neq \tau \\
T_l\left(x\right) &= \sum_{i=1}^{n} \left(x - X_i\right)^l K_i^2\left(x\right) \sigma^2\left(X_i\right) \\
&= nh^{l+1} \sigma^2\left(x\right) f\left(x\right) L_l\left(1 + o_P(1)\right), \\
\xi_l(x) &= \sum_{i=1}^{n} (x - X_i)^l K_i\left(x\right) \sigma(X_i) \varepsilon_i \\
&= O_P\left(n^{1/2} h^{l+(1/2)}\right),
\end{aligned}$$

$$\sum_{i=1}^{n} \omega_i(x) = n^2 h^4 f^2(x)(1 + o_P(1)). \tag{1}$$

Démonstration. – On commence par poser : $S_l\left(x\right) = \sum_{i=1}^{n} Z_i(x)$. Alors, d'après l'inégalité de Bienaymé-Tchebycheff,

$$S_l\left(x\right) = n\,\mathbb{E}\left(Z_1(x)\right) + \mathcal{O}_P\left(\sqrt{\mathrm{Var}\ \left(S_l\left(x\right)\right)}\right). \tag{2}$$

On a

$$\begin{aligned}
\mathrm{Var}\,(S_l\,(x)) &= \mathrm{Var}\left[\sum_{i=1}^{n}(x-X_i)^l\,K_i(x)\right]\\
&= n\mathrm{Var}\,(x-X_0)^l\,K_0(x)+2\sum_{i=1}^{n-1}(n-i)\mathrm{Cov}\left((x-X_0)^l\,K_0(x),(x-X_i)^l\,K_i(x)\right)\\
&= nh^{2l+1}f\,(x)\int u^{2l}K(u)du\,(1+\mathrm{o}(1))+2\sum_{i=1}^{n-1}(n-i)A_i(x)
\end{aligned}$$

où $A_i(x)=\mathrm{Cov}\,(Z_1(x),Z_{i+1}(x))\,.$

Grâce à l'hypothèse d'existence et de bornitude des densités jointes f_j, il vient

$$|A_i\,(x)|\leq \|f_i\|_\infty h^{2l+2}\left\{\int |u^lK(u)|du\right\}^2=\mathcal{O}\left(h^{2l+2}\right).$$

L'hypothèse de $(\theta,\mathcal{L},\psi_1)$ dépendance faible conduit d'autre part à

$$|A_i\,(x)|\leq Ch^{2l-1}\theta_i. \tag{3}$$

Un argument de troncature des indices dû a Tran (voir [21]) conduit alors à la majoration

$$\left|\sum_{i=1}^{n-1}(n-i)A_i(x)\right|\leq C\,n\,h^{2l+2-(3/a)}.$$

En équilibrant les termes de covariance avec le terme de variance, on obtient pour $a\geq 3$,

$$\mathrm{Var}\,(S_l\,(x))\leq Cnh^{2l+1}.$$

Combinée avec la relation (2) cette majoration conduit au résultat annoncé pour S_l. Les termes R_l, T_l et ξ_l se traitant de manière analogue, le détail de la preuve est omis ici. Enfin, on déduit (1) de la relation $\sum_{i=1}^{n}\omega_i(x)=S_2(x)S_0(x)-S_1^2(x)$.

Remarque. – Dans le cas où $h=n^{-\delta}$ ($\delta\in]0,1[$), la condition sur le taux de dépendance s'énonce $\theta_r=r^{-a}$ avec $a>3\,\delta$.

Théorème 8 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1)$ faiblement dépendant. Si le paramètre de fenêtre vérifie $h_n\to 0$ et $nh_n\to+\infty$, si de plus $\theta_r=r^{-a}$ avec $a\geq 3$, alors l'erreur quadratique conditionnelle moyenne de $\widehat{\gamma}(\tau)$ est*

$$\mathbb{E}\left[(\widehat{\gamma}(\tau)-\gamma(\tau))^2|X_1,\ldots,X_n\right]=\left[\frac{\gamma^{(2)}(\tau)}{2}h_n^2B_+\right]^2+\frac{2\sigma^2(\tau)V_+}{nh_nf(\tau)}+\;o_P\left(h_n^4+\frac{1}{nh_n}\right).$$

Démonstration. – La décomposition de l'erreur quadratique conditionnelle moyenne fait ressortir les termes suivants :

$\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))^2|X_1,\ldots,X_n\right]$, $\mathbb{E}\left[(\widehat{m}_-(\tau)-m_-(\tau))^2|X_1,\ldots,X_n\right]$, et
$\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))(\widehat{m}_-(\tau)-m_-(\tau))|X_1,\ldots,X_n\right].$

Les deux premiers termes se traitent comme dans le cadre indépendant. On suit les décompositions de [27] :

$$\begin{aligned}\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))^2|X_1,\ldots,X_n\right] &= \left(\frac{\sum_{i=1}^n \omega_i^+(\tau)(m_+(X_i)-m_+(\tau))}{\sum_{i=1}^n \omega_i^+(\tau)}\right)^2 \\ &\quad+\frac{\sum_{i=1}^n(\omega_i^+(\tau))^2\sigma^2(X_i)}{\left(\sum_{i=1}^n \omega_i^+(\tau)\right)^2} \\ &= (B_n^+)^2+V_n^+.\end{aligned}$$

On établit une formule analogue avec m_-. Par contre, on n'a plus ici l'indépendance de $\widehat{m}_+(\tau)-m_+(\tau)$ et de $\widehat{m}_-(\tau)-m_-(\tau)$ conditionnellement à $X_1,\ldots,X_n$. Cependant, remarquons que conditionnellement à $X_1,\ldots,X_n$, $\omega_i^+(\tau)\omega_i^-(\tau)=0$ presque sûrement : c'est une conséquence du fait que les supports de K_+ et de K_- sont respectivement $[-1,0]$ et $[0,1]$. On a alors, grâce aux propriétés d'indépendance de la suite $(\varepsilon_i)_{i\in\mathbb{N}}$,

$$\mathbb{E}\left[(\widehat{m}_+(\tau)-m_+(\tau))(\widehat{m}_-(\tau)-m_-(\tau))|X_1,\ldots,X_n\right]=B_n^+\,B_n^-.$$

On retrouve alors le même résultat qu'en indépendant [27].

Théorème 9 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1')$ faiblement dépendant. Si $\theta_r=r^{-a}$ avec $a\geq 3$, et si $nh^5\to 0$, $nh\to+\infty$ alors*

$$\sqrt{nh}\left(\widehat{m}_-(\tau)-m_-(\tau)\,,\widehat{m}_+(\tau)-m_+(\tau)\right)'$$

converge vers la loi normale en dimension 2

$$\mathcal{N}\left(0,\left(V_+\sigma^2(\tau)/f(\tau)\right)I_2\right)$$

où I_2 est la matrice identité en dimension 2.

Corollaire 1 *On suppose que le processus $\boldsymbol{X}$ est $(\theta,\mathcal{L},\psi_1')$ faiblement dépendant. Si $\theta_r=r^{-a}$ avec $a\geq 3$, et si $nh^5\to 0$, $nh\to+\infty$ alors*

$$\sqrt{nh}\left(\widehat{\gamma}(\tau)-\gamma(\tau)\right)\xrightarrow{Loi}\mathcal{N}\left(0,2\tfrac{\sigma^2(\tau)}{f(\tau)}V_+\right).$$

Remarque. – Pour montrer le théorème 9, on utilise une technique semblable à celle développée dans [16]. L'hypothèse de causalité est essentielle en raison des termes de covariances, non symétriques, considérés. Le corollaire 1 découle immédiatement du théorème 9.

Démonstration du théorème 9. – Le contrôle du terme de biais suit la ligne de démonstration de [27].

$$\begin{aligned}\sqrt{nh}\left(\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\ldots,X_n\right)-m_+(\tau)\right) &= \sqrt{nh}\frac{\sum_{i=1}^n \omega_i^+(\tau)\left(m_+(X_i)-m_+(\tau)\right)}{\sum_{i=1}^n \omega_i^+(\tau)} \\ &= \sqrt{nh}\frac{h^2B_+m_+^{(2)}(\tau)}{2}\left(1+\mathrm{o}_P(1)\right) \\ &= \mathcal{O}_P\left(\sqrt{nh^5}\right).\end{aligned}$$

On obtient une borne analogue pour $\sqrt{nh}\left(\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)-m_-(\tau)\right)$. On déduit de la décomposition

$$\widehat{\gamma}(\tau)-\gamma(\tau)=\left(\widehat{m}_+(\tau)-m_+(\tau)\right)-\left(\widehat{m}_-(\tau)-m_-(\tau)\right),$$

que si $nh^5\to 0$, alors

$$\sqrt{nh}\left(\mathbb{E}\left(\widehat{\gamma}(\tau)|X_1,\ldots,X_n\right)-\gamma(\tau)\right)=\mathrm{o}_P(1). \tag{4}$$

On se propose de prouver le comportement asymptotiquement gaussien de

$$\left(\widehat{m}_+(\tau)-\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\cdots,X_n\right),\widehat{m}_-(\tau)-\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)\right)$$

$$=\left(\frac{\sum_{i=1}^n\omega_i^+(\tau)\sigma\left(X_i\right)\varepsilon_i}{\sum_{i=1}^n\omega_i^+(\tau)},\frac{\sum_{i=1}^n\omega_i^-(\tau)\sigma\left(X_i\right)\varepsilon_i}{\sum_{i=1}^n\omega_i^-(\tau)}\right)$$

en étudiant la v.a.

$$A_{a,b}=a\left(\widehat{m}_+(\tau)-\mathbb{E}\left(\widehat{m}_+(\tau)|X_1,\ldots,X_n\right)\right)+b\left(\widehat{m}_-(\tau)-\mathbb{E}\left(\widehat{m}_-(\tau)|X_1,\ldots,X_n\right)\right),$$

avec $a,b\in\mathbb{R}$. D'après le théorème 7, comme $a\geq 3$, il vient que

$$A_{a,b}=\frac{1}{n^2h^4f^2(\tau)}\left(\sum_{i=1}^n\left(a\omega_i^+(\tau)+b\omega_i^-(\tau)\right)\sigma\left(X_i\right)\varepsilon_i\right)\left(1+\mathrm{o}_P(1)\right).$$

D'autre part, on a la relation

$$\sum_{i=1}^n\omega_i^+(\tau)\sigma\left(X_i\right)\varepsilon_i=S_2^+(\tau)\xi_0^+(\tau)-S_1^+(\tau)\xi_1^+(\tau).$$

On pose

$$\Omega_i^+=\frac{\left(hK_2^+-\left(\tau-X_i\right)K_1^+\right)K_+\left(\frac{\tau-X_i}{h}\right)\sigma\left(X_i\right)\varepsilon_i}{f(\tau)}.$$

On définit la v.a. Ω_i^- par analogie immédiate. Enfin $\Omega_i=a\Omega_i^++b\Omega_i^-$. D'après le théorème 7, il vient alors

$$A_{a,b}=\frac{1}{nh^2}\sum_{i=1}^n\Omega_i+\mathrm{o}_P(n^{-1/2}h^{-1/2}). \tag{5}$$

On applique la variante de la méthode de Lindeberg-Rio décrite dans [16]. Soit une fonction $\varphi:\mathbb{R}\longrightarrow\mathbb{R}$ possédant des dérivées continues et bornées jusqu'à l'ordre trois. Soit η une v.a. réelle gaussienne centrée et réduite. On veut établir

$$\Delta_n\left(\varphi\right)=\mathbb{E}\left(\varphi\left(S_n\right)-\varphi\left(\sigma_n\eta\right)\right)\to 0 \tag{6}$$

avec $S_n=\frac{\sum_{i=1}^n\Omega_i}{\sqrt{nh^3}}$ et $\sigma_n^2=\mathrm{Var}\left(S_n\right)$. Les v.a. Ω_i sont centrées et décorrélées. Alors

$$\begin{aligned}\sigma_n^2 &= \frac{1}{nh^3}\left(\sum_{i=1}^n\mathrm{Var}\left(\Omega_i\right)+2\sum_{i=2}^n(n-i+1)\mathrm{Cov}\left(\Omega_1,\Omega_i\right)\right) && (7)\\ &= \left(a^2+b^2\right)\sigma^2(\tau)\frac{V_+}{f(\tau)}(1+\mathrm{o}(1)). && (8)\end{aligned}$$

On définit $S_{k,n}=\frac{\sum_{i=1}^k\Omega_i}{\sqrt{nh^3}}$, avec $1\leq k\leq n$ et on prend par convention $S_{0,n}=0$.

La propriété de décorrélation décrite plus haut conduit alors à vérifier que

$$\begin{aligned} v_{k,n} &= \text{Var}\,(S_{k,n}) - \text{Var}\,(S_{k-1,n}) \\ &= \frac{1}{nh^3}\text{Var}\,(\Omega_k) + 2\frac{1}{nh^3}\sum_{i=1}^{k-1}\text{Cov}\,(\Omega_i,\Omega_k) \\ &= \frac{(a^2+b^2)\,\sigma^2(\tau)V_+}{n\,f(\tau)} + \text{o}(\frac{1}{n}) \end{aligned}$$

est strictement positif pour n suffisamment grand.

Soit une suite $(\eta_{k,n})_{n\in\mathbb{N}^*,1\leq k\leq n}$ de v.a. gaussiennes indépendantes où la loi de $\eta_{k,n}$ est centrée et de variance $v_{k,n}$. On définit enfin $\varphi_{k,n}(x) = \mathbb{E}\,\varphi\,(x+T_{k,n})$ avec $T_{k,n} = \sum_{i=k+1}^{n}\eta_{i,n}$. Les fonctions $\varphi_{k,n}$ ont les mêmes propriétés de dérivabilité que φ : si pour $0\leq j\leq 3$, $\|\varphi^{(j)}\|_\infty \leq C_j$, alors $\|\varphi_{k,n}^{(j)}\|_\infty \leq C_j$.

Alors il existe $0 < \iota_{k,n} < 1$ tel que

$$\begin{aligned} |\Delta_n\,(\varphi)| &\leq \sum_{k=1}^{n}\left|\Delta_{k,n}^1\,(\varphi)\right| + \sum_{k=1}^{n}\left|\Delta_{k,n}^2\,(\varphi)\right| \\ \Delta_{k,n}^1\,(\varphi) &= \mathbb{E}\varphi_{k,n}\,(S_{k-1,n}+\eta_{k,n}) - \mathbb{E}\varphi_{k,n}\,(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\varphi_{k,n}^{(2)}\,(S_{k-1,n})\,, \\ \Delta_{k,n}^2\,(\varphi) &= \mathbb{E}\varphi_{k,n}\,(S_{k,n}) - \mathbb{E}\varphi_{k,n}\,(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbb{E}\varphi_{k,n}^{(2)}\,(S_{k-1,n}) \\ &= \mathbb{E}\varphi_{k,n}^{'}\,(S_{k-1,n})\,\frac{\Omega_k}{\sqrt{nh^3}} + \frac{1}{2}\mathbb{E}\varphi_{k,n}^{(2)}\,(S_{k-1,n})\left(\frac{\Omega_k^2}{nh^3} - v_{k,n}\right) \\ &\quad + \frac{1}{6}\mathbb{E}\varphi_{k,n}^{(3)}\left(S_{k-1,n} + \iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3. \end{aligned} \tag{9}$$

On borne $\Delta_{k,n}^1\,(\varphi)$ par

$$\begin{aligned} \left|\frac{1}{6}\mathbb{E}\varphi_{k,n}^{(3)}\,\left((S_{k-1,n}+\iota_{k,n}\eta_{k,n})\,\eta_{k,n}^3\right)\right| &\leq \frac{\|\varphi_{k,n}^{(3)}\|_\infty}{6}\mathbb{E}\left|\eta_{k,n}^3\right| \\ &\leq \frac{2\|\varphi_{k,n}^{(3)}\|_\infty v_{k,n}^{3/2}}{3\sqrt{2\pi}} = \mathcal{O}\left(n^{-3/2}\right). \end{aligned} \tag{10}$$

D'où

$$\sum_{i=1}^{n}\left|\Delta_{k,n}^1\,(\varphi)\right| = \mathcal{O}\left(n^{-1/2}\right) \to 0. \tag{11}$$

Pour étudier $\Delta_{k,n}^2$ on utilise le développement de Taylor (9).

En utilisant la propriété d'indépendance de la suite centrée $(\varepsilon_i)_{i\in\mathbb{N}}$, il vient

$$\mathbb{E}\varphi_{k,n}^{'}\,(S_{k-1,n})\,\frac{\Omega_k}{\sqrt{nh^3}} = 0. \tag{12}$$

D'autre part, comme les Ω_i sont décorrélées, et en notant $\Omega_i = \Psi_n(X_i)\,\varepsilon_i$,

$$\begin{aligned}
\mathbb{E}\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right)\left(\frac{\Omega_k^2}{nh^3}-v_{k,n}\right) &= \operatorname{Cov}\left(\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right),\frac{\Omega_k^2}{nh^3}\right)\\
&= \sum_{i=1}^{k-1}\mathbb{E}\left(\varphi_{k,n}^{(3)}\left(S_{i-1,n}^{*}\right)\frac{\Omega_i}{\sqrt{nh^3}}\frac{\Psi_n^2(X_k)}{nh^3}\right)\mathbb{E}(\varepsilon_k^2)\\
&\quad-\sum_{i=1}^{k-1}\left(\mathbb{E}\varphi_{k,n}^{(3)}\left(S_{i-1,n}^{*}\right)\frac{\Omega_i}{\sqrt{nh^3}}\right)\left(\mathbb{E}\frac{\Psi_n^2(X_k)}{nh^3}\mathbb{E}\left(\varepsilon_k^2\right)\right),
\end{aligned}$$

avec $S_{i-1,n}^{*}=S_{i-1,n}+\iota_i\frac{\Omega_i}{\sqrt{nh^3}}$, $0<\iota_i<1$. D'où

$$\begin{aligned}
&\left|\mathbb{E}\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right)\left(\tfrac{\Omega_k^2}{nh^3}-v_{k,n}\right)\right|\\
&\qquad\le \tfrac{\|\varphi_{k,n}^{(3)}\|_\infty}{\sqrt{n^3h^9}}\textstyle\sum_{i=1}^{k-1}\left(\mathbb{E}\left|\Psi_n(X_i)\varepsilon_i\Psi_n^2(X_k)\right|+\mathbb{E}\left|\Psi_n(X_i)\varepsilon_i\right|\mathbb{E}\left(\Psi_n^2(X_k)\right)\right)\\
&\qquad\le C\textstyle\sum_{i=1}^{k-1}\dfrac{h^{1/2}}{n^{3/2}}.
\end{aligned}$$

De plus, puisque d'une part ε_k est indépendante de $(\varepsilon_1,\ldots,\varepsilon_{k-1},\Psi_n(X_1),\ldots,\Psi_n(X_k))$, et d'autre part X_k est indépendante de $(\varepsilon_1,\ldots,\varepsilon_{k-1})$,

$$\begin{aligned}
\operatorname{Cov}\left(\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right),\frac{\Omega_k^2}{nh^3}\right) &= \sum_{i=1}^{k-1}\operatorname{Cov}\left(\varphi_{k,n}^{(3)}\left(S_{i-1,n}^{*}\right)\frac{\Omega_i}{\sqrt{nh^3}},\frac{\Psi_n^2(X_k)\varepsilon_k^2}{nh^3}\right)\\
&= \sum_{i=1}^{k-1}\mathbb{E}\left(\operatorname{Cov}\left(\varphi_{k,n}^{(3)}\left(S_{i-1,n}^{*}\right)\frac{\Omega_i}{\sqrt{nh^3}},\frac{\Psi_n^2(X_k)}{nh^3}|\varepsilon_1,\cdots,\varepsilon_{k-1}\right)\right).
\end{aligned}$$

En utilisant la dépendance des X_i on obtient :

$$\begin{aligned}
\left|\operatorname{Cov}\left(\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right),\frac{\Omega_k^2}{nh^3}\right)\right| &\le \sum_{i=1}^{k-1}\frac{\|\varphi_{k,n}^{(3)}\|_\infty\|\Psi_n\|_\infty\mathbb{E}|\varepsilon_i|\operatorname{Lip}\left(\Psi_n^2\right)}{\sqrt{nh^3}\,nh^3}\theta_{k-i}\\
&\le \sum_{i=1}^{k-1}\frac{\theta_i}{n^{3/2}h^{5/2}}. \qquad (13)
\end{aligned}$$

On en tire que

$$\begin{aligned}
\sum_{k=1}^{n}\left|\frac{1}{2}\mathbb{E}\varphi_{k,n}^{(2)}\left(S_{k-1,n}\right)\left(\frac{\Omega_k^2}{nh^3}-v_{k,n}\right)\right| &\le C\sum_{k=1}^{n}\sum_{i=1}^{k-1}\min\left(\frac{h^{1/2}}{n^{3/2}},\frac{\theta_i}{n^{3/2}h^{5/2}}\right) \qquad (14)\\
&\le C\frac{1}{\sqrt{nh}}\sum_{i=1}^{n-1}\min\left(h,\frac{\theta_i}{h^2}\right)=\mathcal{O}\left(\frac{h^{1-(3/a)}}{\sqrt{nh}}\right). \qquad (15)
\end{aligned}$$

Comme $nh\to\infty$ et $a\ge 3$, cette borne tend vers 0.

Enfin,

$$\frac{1}{6}\mathbb{E}\left|\varphi_{k,n}^{(3)}\left(S_{k-1,n}+\iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3\right| \quad \leq \quad C\,\mathbb{E}\left|\frac{\Psi_n^3(X_k)}{\sqrt{n^3h^9}}\right|\mathbb{E}\left|\varepsilon_k^3\right|$$

D'où

$$\sum_{k=1}^{n}\left|\frac{1}{6}\mathbb{E}\varphi_{k,n}^{(3)}\left(S_{k-1,n}+\iota_{k,n}\frac{\Omega_k}{\sqrt{nh^3}}\right)\left(\frac{\Omega_k}{\sqrt{nh^3}}\right)^3\right| \quad = \quad \mathcal{O}\left(\frac{1}{\sqrt{nh}}\right)\to 0. \qquad (16)$$

Ceci achève la démonstration.

Remarque. – Si le paramètre de fenêtre vérifie la condition $h = n^{-\delta}$ avec $1/5 < \delta < 1$, alors le théorème de limite centrale 9 est valable pour $a > 6\delta/(1+\delta)$.

Chapitre 4

Estimation de la densité invariante de systèmes dynamiques

Dans ce chapitre, nous présentons des résultats publiés aux *Comptes Rendus de l'Académie des Sciences*, (Prieur, 2001 [49]), et à *ESAIM P & S*, (Prieur, 2001 [50]). Ce sont des résultats de convergence pour l'estimateur à noyau de la densité invariante de certains systèmes dynamiques.

C. R. Acad. Sci. Paris, t. 332, Série I, p. 761–764, 2001
Statistique/*Statistics*
(Systèmes dynamiques/*Dynamical Systems*)

Estimation de la densité invariante de systèmes dynamiques en dimension 1

Clémentine PRIEUR

Mathématiques, Université de Cergy-Pontoise, site Saint-Martin, 2, avenue Adolphe-Chauvin, 95302 Cergy-Pontoise cedex, France
Courriel : prieur@math.u-cergy.fr

(Reçu le 22 décembre 2000, accepté le 9 février 2001)

Résumé. Dans cette Note nous établissons un théorème limite central pour les estimateurs à noyau de la densité invariante de certains systèmes dynamiques en dimension 1. La preuve de ce théorème se fait en deux étapes essentielles : l' étude de la vitesse de convergence de la variance de l' estimateur puis l' obtention du théorème limite central par une variante de la méthode de Lindeberg–Rio [6]. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Density estimation for one-dimensional dynamical systems

Abstract. *In this Note we obtain a central limit theorem for standard kernel invariant density estimates of one-dimensional dynamical systems. The two main steps in the proof of this theorem are the following: the study of speed of convergence for the variance of the estimator and then a variation on the Lindeberg–Rio method* [6]. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Un problème classique est celui de l' estimation de la densité marginale f d' une suite stationnaire $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires dépendantes. On sait que si la suite $(X_n)_{n\in\mathbb{N}}$ satisfait des conditions de mélange [7] ou de dépendance faible [3], on a le résultat suivant :

$$\sqrt{nb_n}\,\big[\widehat{f}_n(x)-\mathbb{E}\widehat{f}_n(x)\big]\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}\left(0,f(x)\int_{-\infty}^{+\infty}K^2(t)\,\mathrm{d}t\right),\tag{1}$$

où $\widehat{f}_n(x)$ est un estimateur classique à noyau défini de la façon suivante :

$$\widehat{f}_n(x)=\frac{1}{nb_n}\sum_{k=0}^{n-1}K\left(\frac{x-X_k}{b_n}\right),\tag{2}$$

Note présentée par Paul DEHEUVELS.

S0764-4442(01)01909-7/FLA

avec une suite $(b_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ et un noyau $K : \mathbb{R} \to \mathbb{R}_+$ à support compact D tels que :

$$b_n \xrightarrow[n\to+\infty]{} 0, \quad nb_n \xrightarrow[n\to+\infty]{} +\infty, \quad \text{et} \quad \int_D K(t)\,\mathrm{d}t = 1, \quad \int_D K^2(t)\,\mathrm{d}t > 0.$$

Dans cette Note, nous nous restreignons à la classe de systèmes dynamiques :

$$\forall\, n \geqslant 1, \quad X_n = T^n X_0, \tag{3}$$

où T est une fonction d'un intervalle $I \subset \mathbb{R}$ dans lui-même admettant une mesure de probabilité invariante μ_0 absolument continue par rapport à la mesure de Lebesgue, et X_0 une variable aléatoire de loi μ_0. Notons f la densité de μ_0 par rapport à la mesure de Lebesgue. Le but est d'estimer f. Nous travaillons sous une hypothèse de contrôle des corrélations. Avant de l'écrire, définissons l'ensemble de fonctions suivant :

DÉFINITION 1.1. – $\mathcal{VB}$ est l'ensemble des fonctions h de I dans $\mathbb{R}$ à variation bornée et de norme $\mathbb{L}^1$ finie (*voir* [8] par exemple). Si $\|h\|_{\mathcal{VB}} := \mathcal{V}(h) + \|h\|_1$, où $\mathcal{V}(h)$ est la variation de h et $\|\cdot\|_1$ la norme standard sur $\mathbb{L}^1$, alors $\|\cdot\|_{\mathcal{VB}}$ est une norme et $\mathcal{VB}$ muni de cette norme est un espace de Banach.

Nous pouvons maintenant supposer :

$$\exists\,\kappa > 0,\ \exists (d_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}},\ \sum_{k=0}^{+\infty} d_k < \infty \text{ tels que } \forall\, n \geqslant 0,\ \forall h,\ k \in \mathcal{VB},$$
$$\bigl|\operatorname{Cov}\bigl(h(X_0), k(X_n)\bigr)\bigr| \leqslant \kappa\, \|k\|_1\, \|h\|_{\mathcal{VB}}\, d_n, \tag{4}$$

où Cov désigne la covariance par rapport à la mesure invariante μ_0. Ce contrôle de covariances permet de conclure à l'existence d'un théorème de type (1). C'est l'objet du théorème 3.1 énoncé dans le paragraphe 3.

2. Hypothèses sur le modèle et exemples

Dans cette partie nous donnons les hypothèses techniques requises pour montrer le théorème limite central. Prenons $K \in \mathcal{VB}$. Nous considérons ensuite un intervalle $I \subset \mathbb{R}$, et T une application de I dans lui-même. Notons λ la mesure de Lebesgue, et $\operatorname{int}(I)$ l'intérieur de I. Nous supposons :

- pour tout k dans $\mathbb{N}$, pour tout x dans $\operatorname{int}(I)$, $\lim_{t\to0^-} T^k(x-t) =: T^k(x^+)$ et $\lim_{t\to0^+} T^k(x-t) =: T^k(x^-)$ existent ;
- pour tout k dans $\mathbb{N}^*$, posons $D_-^k := \{x \in \operatorname{int}(I),\ T^k(x^-) = x\}$ et $D_+^k := \{x \in \operatorname{int}(I),\ T^k(x^+) = x\}$. Soit $I_1 := \operatorname{int}(I) \setminus \bigcup_{k\in\mathbb{N}^*}(D_-^k \cup D_+^k) =: \operatorname{int}(I) \setminus B$. Alors $\lambda(B) = 0$;
- T admet au moins une mesure de probabilité invariante μ_0 qui est absolument continue par rapport à la mesure de Lebesgue : $\mathrm{d}\mu_0 = f\,\mathrm{d}\lambda$;
- soit $S := \operatorname{supp}(\mu_0)$ le support de μ_0. Si $I_2 := \{x \in S \subset I,\ f \text{ est continue en } x \text{ et } f(x) > 0\}$, alors $\lambda(S \setminus I_2) = 0$.

X_0 est une variable aléatoire de loi μ_0. Nous rappelons que $(X_n)_{n\in\mathbb{N}}$ est définie par (3). Enfin, nous faisons l'hypothèse de contrôle des corrélations décrite par l'inégalité (4) de l'introduction.

Dans le cas où la décroissance dans (4) est exponentielle, il existe des exemples simples (qui peuvent s'étendre par conjugaison) :

- *les fonctions de type Lasota–Yorke*. – Pour une définition plus précise de ces fonctions, nous renvoyons au chapitre 3 de Viana [8] ou à [5].
  - Les fonctions « tentes » (*voir* figure 1, où le sommet est le point $[0.5, 0.75]$).
  - Les fonctions r-adiques (*voir* figure 2, où $r = 8/3$).
- *Les fonctions à nombre infini dénombrable d' intervalles de monotonicité* – La fonction de Gauss par exemple. Elle est définie sur $[0, 1]$ par $T(x) = \frac{1}{x} - \left[\frac{1}{x}\right]$ si $x \neq 0$ et $T(0) = 0$.

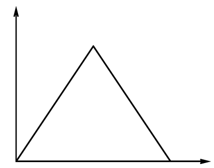


Figure 1. – Fonction « tente ».

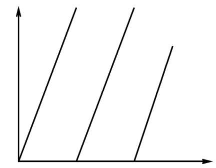


Figure 2. – Fonction r-adique

3. Énoncé du théorème limite central et éléments de preuve

Dans un premier temps, énonçons le résultat qui fait l'objet de cette Note. C'est un théorème limite central portant sur le processus d'estimation centré : $Y_n(x) := \sqrt{nb_n}\big(\widehat{f}_n(x) - \mathbf{E}\,\widehat{f}_n(x)\big)$, où $\widehat{f}_n(x)$ est défini à partir de $(X_n)_{n\in\mathbb{N}}$ comme dans (2).

THÉORÈME 3.1. – *Supposons que T satisfait les hypothèses du paragraphe 2, et que pour tout $1 \leqslant i \leqslant \ell$, $x_i \in I_1 \cap I_2$. Alors les répartitions fini-dimensionnelles $(\overline{Y_n}(x_1), \ldots, \overline{Y_n}(x_\ell))$, du processus*

$$\overline{Y_n}(x) \equiv \frac{Y_n(x)}{\sqrt{f(x)\int_{-\infty}^{+\infty} K^2(t)\,\mathrm{d}t}}$$

convergent en loi, quand n tend vers $+\infty$, vers $\mathcal{N}(0, I_\ell)$.

Remarque 3.1. – Pour $\ell = 1$ ce théorème est un théorème limite central. Si l arbitraire, si J est un intervalle compact inclus dans I, ce résultat montre que la suite des processus $\{\overline{Y_n}(x);\ x \in J\}_{n\geqslant 1}$ n'est pas tendue dans $\mathcal{C}(J)$; ses répartitions fi ni-dimensionnelles convergent en effet vers celles du bruit blanc gaussien $\dot{W}$.

Pour montrer ce théorème nous commençons par montrer un résultat de convergence en moyenne quadratique pour cet estimateur.

LEMME 3.1. – *Supposons que T satisfait les hypothèses du paragraphe 2. Soit $x \in I_1 \cap I_2$. Alors il existe une suite $(\varepsilon_n)_{n\in\mathbb{N}} \in (\mathbb{R})^{\mathbb{N}}$ qui tend vers 0 quand n tend vers l' infini telle que:*

$$\mathrm{Var}\big(\widehat{f}_n(x)\big) = \frac{f(x)\int_D K^2(s)\,\mathrm{d}s}{nb_n}\,[1 + \varepsilon_n].$$

Remarque 3.2. – Une première évaluation de la vitesse de convergence de $\widehat{f}_n(x)$ (en $\mathcal{O}(1/(nb_n^2))$) est donnée dans [2] (ou plus récemment dans [4]). Notre résultat donne lieu au théorème limite central, il fournit donc « la » vitesse de convergence.

Pour montrer un tel résultat dans un cadre de dépendance faible, il est tout à fait raisonnable de supposer l'existence de densités jointes régulières pour les couples $(X_0, X_k)_{k\in\mathbb{N}^*}$. Ici, les lois jointes sont singulières et cette absence de régularité dissocie l'étude de la dépendance faible de celle des systèmes dynamiques. Pour montrer le lemme 3.1, nous utilisons ici les deux lemmes ci-dessous :

LEMME 3.2. – *Supposons que T satisfait les hypothèses du paragraphe 2. Soit $x \in I_1 \cap I_2$. Alors pour chaque k dans $\mathbb{N}^*$, il existe une suite $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ telle que*

$$\mathrm{Cov}\left(K\left(\frac{x - X_0}{b_n}\right), K\left(\frac{x - X_k}{b_n}\right)\right) = b_n\,\varepsilon(n,k).$$

LEMME 3.3. – *Supposons que T satisfait les hypothèses du paragraphe* 2. *Soit* $x \in I_1 \cap I_2$. *Alors*

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\operatorname{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) \xrightarrow[n\to\infty]{} 0.$$

Le point essentiel dans la preuve du lemme 3.2 est l'étude, pour $k \geqslant 1$ de $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right]$. Pour n grand, pour $k \geqslant 1$ fixé nous avons :

$$\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right] \leqslant b_n \int_D K(t)f(x-tb_n)K\left(\frac{x-T^k(x-tb_n)}{b_n}\right)\mathrm{d}t.$$

Comme K est à support D compact, nous avons

$$x-T^k(x-tb_n) \xrightarrow[n\to+\infty]{} x-T^k(x\mp) \neq 0, \quad \text{uniformément en } t \in D,\ t \geqslant 0 \text{ ou } t \in D,\ t \leqslant 0.$$

Donc, pour k fixé, $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right] = \mathrm{o}(b_n)$, les termes $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)\right]\mathbb{E}\left[K\left(\frac{x-X_k}{b_n}\right)\right]$ se traitent de façon classique. Nous utilisons ensuite le lemme 3.2 ainsi que le contrôle des corrélations (4) pour montrer le lemme 3.3. La fin de la preuve du théorème 3.1 (*cf.* [5]) repose (comme dans [3]) sur une variation de la méthode de Lindeberg–Rio.

Remarque 3.3. – L'étude du biais est classique, elle fait appel à la régularité de f qui dans certains cas [1] se déduit de celle de T. D'autre part, nous pouvons étendre nos résultats au cas où X_0 ne suit plus forcément la loi invariante. $(X_n)_{n\in\mathbb{N}}$ n'est alors plus stationnaire. L'étude du cas non stationnaire repose essentiellement sur les propriétés de l'opérateur de Perron–Frobenius associé à T défini de la façon suivante :

$$\mathcal{L}: \begin{cases} \mathcal{BV} \longrightarrow \mathcal{BV}, \\ \omega \longmapsto \mathcal{L}\omega(x) = \displaystyle\sum_{Ty=x}\frac{\omega(y)}{|T'(y)|}. \end{cases}$$

Références bibliographiques

[1] Amroun A., Systèmes dynamiques perturbés. Sur une classe de fonctions ζ dynamiques, Thèse de doctorat de l'Université Paris-6, 1995.
[2] Bosq D., Guégan D., Nonparametric estimation of the chaotic function and the invariant measure of a dynamical system, Statis. & Probab. Let. 25 (1995) 201–212.
[3] Coulon-Prieur C., Doukhan P., A triangular central limit theorem under a new weak dependence condition, Statis. & Probab. Let. 47 (2000) 61–68.
[4] Maës J., Statistique non paramétrique des processus dynamiques réels en temps discret, Thèse de doctorat de l'Université Paris-6, 1999.
[5] Prieur C., Functional estimation of dynamical systems, Preprint de l'Université de Cergy-Pontoise, n° 7/00, 2000.
[6] Rio E., Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes, Probab. Theory Relat. Fields 104 (1996) 255–282.
[7] Robinson P.M., Nonparametric estimators for time series, J. Time Ser. Anal. 4–3 (1983) 185–207.
[8] Viana M., Stochastic dynamics of deterministic systems, Instituto de Matematica Pura e Aplicada, IMPA 21 (1997).

Résumé d'un texte qui sera conservé pendant cinq ans dans les Archives de l'Académie et dont copie peut être obtenue.

ESAIM: Probability and Statistics September 2001, Vol. 5, 51–76
URL: `http://www.emath.fr/ps/`

DENSITY ESTIMATION FOR ONE-DIMENSIONAL DYNAMICAL SYSTEMS

Clémentine Prieur[1]

Abstract. In this paper we prove a Central Limit Theorem for standard kernel estimates of the invariant density of one-dimensional dynamical systems. The two main steps of the proof of this theorem are the following: the study of rate of convergence for the variance of the estimator and a variation on the Lindeberg–Rio method. We also give an extension in the case of weakly dependent sequences in a sense introduced by Doukhan and Louhichi.

Mathematics Subject Classification. 37D20, 37M10, 37A50, 60G07, 60G10.

Received January 19, 2001. Revised March 9, June 12 and July 3, 2001.

1. Introduction

This paper considers estimation of the marginal density f of a stationary sequence $(X_n)_{n\in\mathbb{N}}$ of dependent random variables. If $(X_n)_{n\in\mathbb{N}}$ satisfies mixing conditions, Robinson [24] obtains the following result:

$$\sqrt{nb_n}\left[\hat{f}_n(x)-\mathbf{E}\hat{f}_n(x)\right]\xrightarrow[n\to+\infty]{\mathcal{D}}\mathcal{N}\left(0,f(x)\int_{-\infty}^{+\infty}K^2(t)\mathrm{d}t\right),\tag{1.1}$$

where $\hat{f}_n(x)$ is a standard kernel density estimate defined as follows (see Rosenblatt [25]):

$$\hat{f}(x)=\hat{f}_n(x)=\frac{1}{nb_n}\sum_{k=0}^{n-1}K\left(\frac{x-X_k}{b_n}\right),\tag{1.2}$$

with a sequence $(b_n)_{n\in\mathbb{N}}\in(\mathbb{R}_+)^{\mathbb{N}}$ and a compact supported kernel $K:\mathbb{R}\to\mathbb{R}$ (we note D its support) satisfying:

$$b_n\xrightarrow[n\to+\infty]{}0,\ \text{and}\ \int_D K(t)\mathrm{d}t=1,\ 0<\int_D K^2(t)\mathrm{d}t<\infty.\tag{1.3}$$

Quote that the last assumption, $\int_D K^2(t)\mathrm{d}t<\infty$, holds for K measurable and bounded.

Keywords and phrases: Dynamical systems, decay of correlations, invariant probability, stationary sequences, Lindeberg theorem, Central Limit Theorem, bias, nonparametric estimation, s-weakly and a-weakly dependent.

[1] Université de Cergy-Pontoise, Laboratoire de Mathématiques, bâtiment A4, Site Saint-Martin, 95011 Cergy-Pontoise Cedex, France; e-mail: `prieur@math.u-cergy.fr`

The purpose of this paper is to prove such a result for a certain class of one-dimensional dynamical systems. Let us introduce the class $(\mathcal{T})$ of dynamical systems:

$$\forall n \geq 1,\ X_n = T^n X_0 \tag{1.4}$$

such that:

- $T : I \to \mathbb{R}$ is a function defined on a closed interval $I \subset \mathbb{R}$;
- T admits an invariant probability measure μ_0 absolutely continuous with respect to the Lebesgue measure;
- the random variable X_0 has the distribution μ_0.

Therefore $(X_n)_{n\in\mathbb{N}}$ is a stationary sequence. Let f denote the density of μ_0 with respect to Lebesgue.

We also assume a control of the correlations. Before stating this control, let us define the set of functions $\mathcal{BV}$. We first define the variation of a function $\varphi : I = [L, R] \to \mathbb{R}$ (L and R can be respectively set equal to $-\infty$ and $+\infty$) by

$$\mathcal{V}(\varphi) = \sup_J \ \sup \sum_{i=1}^{n} |\varphi(x_{i-1}) - \varphi(x_i)|$$

where the first supremum is taken over all compact subsets $J = [L_J, R_J]$ of $I = [L, R]$, and where the second supremum is taken over all finite partitions $L_J = x_0 < x_1 < \cdots < x_n = R_J$, $n \geq 1$, of J (see [26]). Now $\mathcal{BV}$ denotes the set of functions $h : I \to \mathbb{R}$ with bounded variation and whose $\mathbb{L}^1-$norm is finite (see *e.g.* [27]). If $||h||_{\mathcal{BV}} := \mathcal{V}(h) + ||h||_1$ where $\mathcal{V}(h)$ is the variation of h and $||\cdot||_1$ is the standard norm on $\mathbb{L}^1$, then $||\cdot||_{\mathcal{BV}}$ is a norm and $\mathcal{BV}$ endowed with this norm is a Banach space.

We can now state the assumption on the correlations.

There exist a positive constant κ and a sequence of non-negative numbers $(d_n)_{n\in\mathbb{N}}$ satisfying $\sum_{k=0}^{+\infty} d_k < \infty$ such that

$$\forall n \geq 0,\ \forall\, h,\ k \in \mathcal{BV}, \quad |\mathrm{Cov}\,(h(X_0), k(X_n))\,| \leq \kappa \|k\|_1 \|h\|_{\mathcal{BV}}\, d_n, \tag{1.5}$$

where Cov denotes the covariance with respect to the invariant measure μ_0. This control yields the Theorem 1.1 (see Sect. 4).

The paper is organized as follows. In Section 2, we precise the assumptions on the class $\mathcal{T}$ of dynamical systems. In Section 3 we study the convergence in mean squares of our density estimates. This is the purpose of Lemma 3.1 which constitutes the tool step of this paper and improves existing results (see Bosq and Guégan [5], Maës [19]). Section 4 is devoted to the statement and the proof of the main result: the Central Limit Theorem (CLT) of type (1.1). Lemma 3.1, together with a variation on the Lindeberg–Rio method [22, 23], yields the central limit Theorem 4.3. Hence the rate of convergence obtained in Lemma 3.1 is "the good one". Theorem 4.3 does not involve the bias term. However the analysis of the bias is standard (see [2, 25]) and is not linked with the dependent structure of the subjacent sequence $(X_n)_{n\in\mathbb{N}}$ but only on its marginal distribution; this yields Theorems 4.1 and 4.2. Some examples follow in Section 5.

In Section 6 we extend the results of Sections 3 and 4 to the case where X_0 does not follow the invariant law μ_0. Properties of the Perron–Frobenius operator allow to conclude for Lasota–Yorke transformations T (Lem. 6.1 and Ths. 6.2, 6.3 and 6.4). Thanks to the results of this section, we can choose a density function p and construct a (non-stationary) sequence $(X_n^{'})_{n\in\mathbb{N}}$ as follows:

$$X_0^{'} \text{ has the distribution } p(t)\mathrm{d}t,\ \ X_n^{'} = T^n X_0^{'},\ n \geq 1, \tag{1.6}$$

and we then estimate the invariant density f with

$$\hat{p}(x) = \hat{p}_n(x) = \frac{1}{nb_n} \sum_{k=0}^{n-1} K\left(\frac{x - X_k^{'}}{b_n}\right) \tag{1.7}$$

where b_n and K are defined as in (1.3).

Finally, the appendix is devoted to an extension to the case of weakly dependent sequences in a sense introduced by Doukhan and Louhichi [13] (Th. A.1).

2. Definition of the class $\mathcal{T}$ of transformations

In this part we detail the technical assumptions required for the Central Limit Theorems given in Section 4. We first assume that the kernel K is in the set $\mathcal{BV}$. We consider a closed interval $I := [L, R] \subset \mathbb{R}$, and T a function from I into itself. We denote λ the Lebesgue measure on I and $\mathrm{int}(I)$ the interior of I. We assume:

- for all k in $\mathbb{N}$, for all x in $\mathrm{int}(I)$, $\lim_{t \to 0^+} T^k(x+t) =: T^k(x^+)$ and $\lim_{t \to 0^-} T^k(x+t) =: T^k(x^-)$ exist;
- for all k in $\mathbb{N}^*$, denote $D_-^k := \{x \in \mathrm{int}(I),\ T^k(x^-) = x\}$ and $D_+^k := \{x \in \mathrm{int}(I),\ T^k(x^+) = x\}$. Let $I_1 := \mathrm{int}(I) \setminus B$, where $B := \bigcup_{k \in \mathbb{N}^*} (D_-^k \cup D_+^k)$. We assume $\lambda(B) = 0$;
- T admits at least an invariant probability measure μ_0 which is absolutely continuous with respect to Lebesgue measure: $d\mu_0 = f d\lambda$;
- let $S := \mathrm{supp}(\mu_0)$ be the support of μ_0. If $I_2 := S \cap C(f)$ where $C(f)$ denotes the continuity set of f, then $\lambda(S \setminus I_2) = 0$.

Recall that μ_0 is the distribution of X_0 and that $(X_n)_{n \in \mathbb{N}}$ is defined by (1.4). Finally we assume the control of correlations described by inequality (1.5) in the introduction (note that the absolutely continuous invariant probability measure μ_0 is unique because of the control of correlations (1.5)).

The following "tent-map" (see Fig. 1 in Sect. 5) belongs to $\mathcal{T}$ (hence $\mathcal{T} \neq \varnothing$):

$$T(x) = \begin{cases} \frac{3}{2}x & \text{if } 0 \le x \le \frac{1}{2} \\ \frac{-3}{2}x + \frac{3}{2} & \text{if } \frac{1}{2} < x \le 1. \end{cases}$$

In the next sections, $\lambda(D)$ denotes the Lebesgue measure of the compact set D.

3. Convergence in mean squares in the stationary case

In this section, we provide the mean squares convergence of the invariant density estimates for dynamical systems in the class $\mathcal{T}$. Let T be in $\mathcal{T}$, then the series of correlations are summable (1.5).

Remark 3.1. In many examples (see Sect. 5) we have a stronger property: the exponential decay of correlations. A classical way to prove it is by using the theory of transfer operators (see Collet [8]).

Recall that μ_0 is the absolute continuous invariant probability measure for T, that f is its density with respect to Lebesgue measure on I, and that $X_0 \sim \mu_0$. $(X_k)_{k \in \mathbb{N}}$ is then a stationary process with marginal density f. We estimate f by standard kernel estimates of the invariant density. We get the following mean squares convergence result:

Lemma 3.1. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n \to \infty]{} 0$. Then if $x \in I_1 \cap I_2$, we have*

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right).$$

Remark 3.2. A first evaluation, $\mathcal{O}\left(\frac{1}{nb_n^2}\right)$, of the rate of convergence of $\hat{f}_n(x)$ is given in [5] (or more recently in [19]). Our result provides $\lim_{n\to+\infty}(nb_n)\mathrm{Var}\left(\hat{f}(x)\right) = C,\ C \geq 0$. This accurate rate of convergence is necessary in order to obtain the forthcoming Central Limit Theorem (Th. 4.3) from Lindeberg–Rio technique.

To prove such a result in a mixing frame, the authors (*e.g.* [13]) usually assume that the couples $\{(X_0, X_k)\}_{k\in\mathbb{N}^*}$ have regular joint densities. Here these distributions are singular and therefore our study is quite different. We first need the two following lemmas:

Lemma 3.2. *Assume that T is in the class $\mathcal{T}$ and let $x \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then for each $k \in \mathbb{N}^*$, there exists a sequence $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ such that*

$$\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n\,\varepsilon(n,k).$$

Lemma 3.3. *Assume that T is in the class $\mathcal{T}$ and let $x \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then*

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) \xrightarrow[n\to\infty]{} 0.$$

Proof of Lemma 3.2. Recall that D denotes the support of K. Let k fixed in $\mathbb{N}^*$, then

$$\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = \int_L^R K\left(\frac{x-T^k s}{b_n}\right) K\left(\frac{x-s}{b_n}\right) f(s)\mathrm{d}s - \left(\int_L^R K\left(\frac{x-s}{b_n}\right) f(s)\mathrm{d}s\right)^2.$$

- <u>Study of $Q_n := \left(\int_L^R K\left(\frac{x-s}{b_n}\right) f(s)\mathrm{d}s\right)^2$</u>

$$0 \leq Q_n = b_n^2 \left(\int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K(t)f(x-tb_n)\mathrm{d}t\right)^2 \leq b_n^2 \left(\int_D |K(t)|\, f(x-tb_n)\mathrm{d}t\right)^2.$$

Let $l_n(t) := |K(t)|\, f(x-tb_n)$. As f is continuous at point x ($x \in I_2$ implies $x \in C(f)$), we get for all $t \in D$: $l_n(t) \xrightarrow[n\to\infty]{} |K(t)|\, f(x)$. Let $\varepsilon > 0$. As D is compact, the convergence of $f(x-tb_n)$ to $f(x)$ is uniform in $t \in D$. Hence there exists a positive integer N such that for $n \geq N$, $|l_n(t)| \leq |K(t)|\ (f(x)+\varepsilon)\ \forall t \in D$.

Then, as $\int_D |K(t)|\ (f(x)+\varepsilon)\,\mathrm{d}t \leq (f(x)+\varepsilon)\ \sqrt{\lambda(D)}\ \sqrt{\int_D K^2(t)\mathrm{d}t} < \infty$, the dominated convergence theorem yields ([26], p. 27)

$$\int_D |K(t)|\, f(x-tb_n)\mathrm{d}t \xrightarrow[n\to\infty]{} f(x) \int_D |K(t)|\,\mathrm{d}t < \infty.$$

Hence $b_n \left(\int_D |K(t)|\, f(x-tb_n)\mathrm{d}t\right)^2 \xrightarrow[n\to\infty]{} 0$ and there exists $\varepsilon_1(n) \xrightarrow[n\to\infty]{} 0$ such that $Q_n = b_n\,\varepsilon_1(n)$.

- Study of $A_{n,k} := \int_L^R K\left(\frac{x-T^k s}{b_n}\right) K\left(\frac{x-s}{b_n}\right) f(s)\mathrm{d}s$

$$\begin{aligned} 0 \leq |A_{n,k}| &= b_n \left| \int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) K(t) f(x-tb_n)\mathrm{d}t \right| \\ &\leq b_n \int_D \left| K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) K(t) \right| f(x-tb_n)\mathrm{d}t. \end{aligned}$$

As D is compact, $tb_n \to 0$ uniformly in $t \in D$ as n tends to infinity. Hence the existence of onesided limits at point x implies

$x - T^k(x-tb_n) \longrightarrow x - T^k(x^-)$, uniformly with respect to $t \in D_+^* := D \cap \mathbb{R}_+^*$;

$x - T^k(x-tb_n) \longrightarrow x - T^k(x^+)$, uniformly with respect to $t \in D_-^* := D \cap \mathbb{R}_-^*$.

Moreover, by assumption, $x - T^k(x^-) \neq 0$ and $x - T^k(x^+) \neq 0$. Let $D^* := D \setminus \{0\}$.

From the compactness of D we exhibit some $n_0 \in \mathbb{N}^*$ such that

$n \geq n_0 \Longrightarrow \left[\forall\, t \in D^* :\ K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) = 0\right].$

Then $n \geq n_0 \Longrightarrow \frac{A_{n,k}}{b_n} = 0$ and we can write $A_{n,k} = b_n\varepsilon_2(n,k)$, with $\varepsilon_2(n,k) = 0$ for n large enough.

Then $\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n[\varepsilon_2(n,k) - \varepsilon_1(n)]$. Hence for each $k \in \mathbb{N}^*$, there exists $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ such that $\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n\, \varepsilon(n,k)$, which concludes the proof. □

Proof of Lemma 3.3. We have

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \leq \frac{1}{b_n}\sum_{k=1}^{n-1}\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right|\cdot \quad (3.1)$$

By the control (1.5) of correlations and from Lemma 3.2 and inequality (3.1), there exists a constant $M > 0$ such that

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \leq M\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right).$$

Let $\varepsilon > 0$. Since $\sum_{k=0}^{\infty} d_k < \infty$, there exists $k(\varepsilon) \in \mathbb{N}^*$ such that $\sum_{k=k(\varepsilon)}^{\infty} d_k < \frac{\varepsilon}{2}$. So for $n \geq k(\varepsilon)$,

$$\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right) < \sum_{k=1}^{k(\varepsilon)-1}\varepsilon(n,k) + \frac{\varepsilon}{2}\cdot$$

Hence there exists some $n_0 \geq k(\varepsilon)$ such that $n \geq n_0$ implies

$$\sum_{k=1}^{k(\varepsilon)-1}\varepsilon(n,k) + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This entails $\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right) \xrightarrow[n\to\infty]{} 0$ and concludes the proof. □

Remark 3.3. Using similar arguments, we can prove that for all $0 \leq i < j \leq n-1$, and for any bounded function φ,

$$\mathrm{Cov}\left(\varphi(X_0)\left[K\left(\frac{x-T^jX_0}{b_n}\right)-\mathbf{E}K(\frac{x-T^jX_0}{b_n})\right], K\left(\frac{x-T^iX_0}{b_n}\right)-\mathbf{E}K\left(\frac{x-T^iX_0}{b_n}\right)\right) \leq b_n \times \varepsilon(n, j-i),$$

with $\varepsilon(n, j-i) \xrightarrow[n\to\infty]{} 0$. We make use of this remark in the next section to prove the central limit Theorem 4.3.

We are now in position to prove Lemma 3.1.

Proof Lemma 3.1. Let $\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x)$ denote the variance of $\frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x-\widetilde{X_k}}{b_n}\right)$ where the $\widetilde{X_k}$'s are independent copies of the X_k's. We have:

$$\mathrm{Var}\,(\hat{f}(x)) = \mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) + \frac{2}{n^2b_n^2}\sum_{0\leq i<j\leq n-1} \mathrm{Cov}\left(K\left(\frac{x-X_i}{b_n}\right), K\left(\frac{x-X_j}{b_n}\right)\right)\cdot \tag{3.2}$$

As $x \in I_1 \subset \mathrm{int}(I)$, we have for n large enough

$$\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) = \frac{1}{nb_n}\int_D K^2(s)\ f(x-sb_n)\mathrm{d}s - \frac{1}{n}\left(\int_D K(s)\ f(x-sb_n)\mathrm{d}s\right)^2.$$

As f is continuous at point x, and as K is compactly supported and satisfies (1.3), apply twice the dominated convergence theorem (see *e.g.* [26], p. 27) to obtain

$$\begin{aligned}
\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) &= \frac{1}{nb_n}\int_D K^2(s)\ f(x-sb_n)\mathrm{d}s - \frac{1}{n}\left(\int_D K(s)\ f(x-sb_n)\mathrm{d}s\right)^2\\
&= \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{1}{nb_n}\ \left\{b_n\left(\left(\int_D K(s)f(x)\mathrm{d}s\right)^2 + o(1)\right)\right\}\\
&= \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right).
\end{aligned}$$

Remark 3.4. Quote that Bosq and Lecoutre ([6], p. 76) ask an additional differentiability assumption on f in order to obtain an equivalent of the bias together with this result.

Hence

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{2}{n^2b_n^2}\sum_{k=1}^{n-1}(n-k)\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\cdot$$

Then

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{2}{nb_n}\left(\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right)\cdot$$

Hence by Lemma 3.3,

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right),$$

which concludes the proof. □

Remark 3.5. We may also study the MISE defined as follows

$$\text{MISE}\,(\hat{f}, f) := \int \mathbf{E}(\hat{f} - f)^2 \mathrm{d}x.$$

As usual (see [2]) the rate of convergence of the MISE to 0 depends on the regularity of f. If T is Lasota–Yorke and Markov, we can deduce the regularity of the density f from the one of T (see [1]).

4. A Central Limit Theorem in the stationary case

Many versions of a Central Limit Theorem for the partial sums of dynamical systems $\frac{1}{\sqrt{n}} \sum_{j=1}^{[ns]} \phi(T^j(x))$ have been proved in the literature. For example Liverani [18], Viana [27] prove a CLT in the case where $s = 1$ for some piecewise expanding dynamical systems. Barbour *et al.* [4] prove a functional CLT with respect to s in the case where T is some expanding map of the unit interval into itself. They use first a coupling method: they prove that the iterates of T can be closely tied to an $m-$dependent process. Then they use techniques which are derived using Stein's method, so they obtain bounds on the rate of convergence. Here we prove a CLT for the density estimates.

We study the following process

$$U_n(x) := \sqrt{nb_n}(\hat{f}(x) - f(x)). \tag{4.1}$$

We do not use a decomposition in Bernstein blocks. Here the idea is to adapt the Lindeberg method after Rio [23]. To be in position to use such a method, we need the mean squares convergence result stated in Section 3 (Lem. 3.1). We then study the bias term by a Taylor's decomposition. Let us now state the central limit results for the invariant density estimates in the case of dynamical systems in the class $\mathcal{T}$.

We first precise the following notations. If $l \in \mathbb{N}^*$ and if $a,\ y_1, \dots, y_l \in \mathbb{R}$, $a * (y_i)_{1 \le i \le l}$ denotes the vector of $\mathbb{R}^l$ whose coordinates are $a * y_1, \dots, a * y_l$, and $a * \text{diag}\,(y_1, \dots, y_l)$ denotes the diagonal matrix whose diagonal terms are equal to $a * y_1, \dots, a * y_l$. We also define

$$\Sigma_l := \left(\int_D K^2(s)\mathrm{d}s \right) * \text{diag}\,(f(x_1), \dots, f(x_l)). \tag{4.2}$$

Theorem 4.1. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let m be a positive integer. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f is m-times continuously differentiable on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \le j \le m-1$ and that nb_n^{2m+1} converges to some non-negative constant ρ_m as n tends to infinity. Then*

$$(U_n(x_1),\ \dots,\ U_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(\frac{(-1)^m\ \sqrt{\rho_m}}{m!} \int_D s^m\ K(s)\mathrm{d}s * \left(f^{(m)}(x_i)\right)_{1\le i\le l},\ \Sigma_l \right),$$

where Σ_l is defined by (4.2).

Remark 4.1. For example, when $m = 1$ Theorem 4.1 yields

$$(U_n(x_1),\ \dots,\ U_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(-\sqrt{\rho_1} \int_D s\ K(s)\mathrm{d}s * (f'(x_i))_{1\le i\le l},\ \Sigma_l \right).$$

We can also write such a theorem if the regularity of the invariant density f in terms of Hölder spaces is not necessarily an integer. Let ν denote the regularity of the function f, this means that setting $\nu = \alpha + \beta$ with $\alpha \in \mathbb{N}$ and $0 \le \beta < 1$ there exists a constant $A > 0$ such that f is α-times continuously differentiable with $|f^{(\alpha)}(x) - f^{(\alpha)}(y)| \le A|x - y|^\beta$ for $x,\ y$ belonging to an arbitrary compact interval. We get the following result:

Theorem 4.2. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let $\nu = \alpha + \beta$ with $\alpha \in \mathbb{N}$ and $0 \le \beta < 1$. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f has the regularity ν on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \le j \le \alpha$ and that $nb_n^{2\nu+1} \xrightarrow[n\to+\infty]{} 0$. Then*

$$(U_n(x_1),\ \ldots\ ,\ U_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

Example 4.1. *In many cases, the regularity of the invariant density f can be deduced from the one of the dynamic T. Let us give some examples and references.*

We consider a transformation T of the interval $I = [0,1]$. We assume that T is Lasota-Yorke in a sense defined in Section 5. Then there exists a countable partition (finite or not) of I, $\{a_j\}_{j\in J}$, such that for all $j \in J$, on $]a_j, a_{j+1}[$, T is expansive (see Assumpt. 5.2). For all $j \in J$, we assume that T admits a continuous extension to $I_j := [a_j, a_{j+1}]$. Let us denote $\overline{T}(I_j)$ the image of I_j by the extension of T on I_j. We assume moreover that T is of Markov type, i.e.

1. *if the partition $\{a_j\}_{j\in J}$ is finite, we assume that for all $j \in J$, there exists $K_j \subset J$ such that $\overline{T}(I_j) = \bigcup_{k\in K_j} I_k$. Now if A is a subset of I, let clos(A) denotes the closure of A. We assume in that case that there exists a positive integer p such that for all $j \in J$, clos$\left(T^p\left(]a_j, a_{j+1}[\right)\right) = I$;*
2. *if the partition $\{a_j\}_{j\in J}$ is infinitely countable, we assume that for all $j \in J$, $\overline{T}(I_j) = I$.*

We now give easy examples of such transformations T.

1. *the r-adic maps*
$$T(x) = rx\ [1],\ \text{where } r \in \mathbb{N},\ r \ge 2.$$
We have $0 = a_0 < \frac{1}{r} < \cdots < \frac{r-1}{r} < a_r = 1$;
2. *generalization of the r-adic maps*
$$T(x) = rx + c\ [1],\ \text{where } r > 1,\ 0 \le c < 1,\ \text{and } r + c \in \mathbb{N},\ c(r+1) \in \mathbb{N}^*.$$
For these maps we can take $0 = a_0 < \frac{1-c}{r} < \cdots < \frac{(r+c-1)-c}{r} < a_{r+c} = 1$;
3. *some piecewise linear transformations.*
There exists a countable partition (finite or not), $\{a_j\}_{j\in J}$, of I such that for all $j \in J$, $\overline{T}(I_j) = I$, T is linear and continuously differentiable on $]a_j, a_{j+1}[$ and $|T'(x)| \ge 1 + \varepsilon$ where $\varepsilon > 0$;
4. *the Gauss map*
$$T(x) = \frac{1}{x} - \left[\frac{1}{x}\right]\ \text{for } x \ne 0 \text{ and } T(0) = 0.$$
Then for all $j \in J = \mathbb{N}^*$, $I_j = [a_j, a_{j+1}] = \left[\frac{1}{j+1}, \frac{1}{j}\right]$.
For this map we know the exact form of the invariant density, $f(x) = \frac{1}{\log(2)}(1+x)^{-1}$, which is infinitely continuously differentiable.

For these transformations T, we know (e.g. [1, 7, 15]) that the regularity of the invariant density f depends on the one of T. If the partition is finite and if T is piecewise two-times continuously differentiable, then f is piecewise continuously differentiable. If moreover T is onto on each interval of the partition, then f is continuously differentiable. In the case where the partition is infinitely countable, a further assumption on the Schwarzian derivative of T is needed (see e.g. [1, 10]) to conclude that f is Hölder continuous.

If the regularity of f is not an integer, no equivalent of the bias seems to be known. The optimal rate is reached but we do not get an explicit equivalent of the bias. Hence we sometimes prefer not to consider the bias term but rather to restrict our attention to the centered estimation process

$$Y_n(x) := \sqrt{nb_n}(\hat{f}(x) - \mathbf{E}\hat{f}(x)). \tag{4.3}$$

We get the following result (see also Applications 6.1 and 6.2 in Sect. 6 for some use of this result):

Theorem 4.3. *Let T be in the class $\mathcal{T}$, and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Then if for all $1 \le i \le l$, $x_i \in I_1 \cap I_2$,*

$$(Y_n(x_1),\ \dots,\ Y_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\ \Sigma_l),$$

where Σ_l is defined by (4.2).

Remark 4.2. For $l = 1$ Theorem 4.3 is a CLT. Let $J \subset I_1 \cap I_2$ be a compact subinterval of $I_1 \cap I_2$. Working with arbitrary l and with some $f > 0$ implies that the sequence of estimation processes $\left(\frac{Y_n(x)}{\sqrt{f(x)}},\ x \in J\right)_{n\in\mathbb{N}^*}$ is not tight in $C(J)$; its limit is indeed $\sqrt{\int_D K^2(s)\mathrm{d}s}\ \dot{W}$, where $\dot{W}$ is the Gaussian white noise. Now for sake of simplicity we develop the proof for $l = 1$. The general case is similar. If one wants to know the asymptotic behaviour in distribution of the vector $(Y_n(x_1),\ \dots,\ Y_n(x_l))$, it is sufficient to use the following proof of Theorem 4.3 (in the case $l = 1$) with $\frac{1}{\sqrt{nb_n}}\sum_{j=1}^{l} s_j K(\frac{x_j - X_k}{b_n})$, for arbitrary numbers $s_1,\ \dots,\ s_l \in \mathbb{R}$, instead of $\frac{1}{\sqrt{nb_n}}K(\frac{x-X_k}{b_n})$.

We first prove Theorem 4.3 and then deduce Theorems 4.1 and 4.2 by studying the bias term defined by $\mathcal{BIAS}_n(x) = \mathbf{E}\hat{f}_n(x) - f(x)$.

Proof of Theorem 4.3 with $l = 1$. Let us first notice that if $f(x) = 0$, then $Y_n(x)$ tends to zero in mean squares (Lem. 3.1), so it also converges to zero in law.

From now on, we suppose that $f(x) > 0$. Let $g_n(t) = \frac{1}{\sqrt{nb_n}}K(\frac{x-t}{b_n})$, $M_n = \|g_n\|_\infty$, $l_n = \|g_n\|_{\mathcal{BV}}$ and $\delta_n = \|g_n\|_1$.

In the following c will denote some constant independent of k and n, which may vary from line to line. We have $M_n \le \frac{c}{\sqrt{nb_n}}$, $l_n \le \frac{c}{\sqrt{nb_n}}$ and $\delta_n \le \frac{c\,b_n}{\sqrt{nb_n}}$, where c is positive. We recall that for all $h,\ k$ in $\mathcal{BV}$

$$||h \cdot k||_{\mathcal{BV}} \le \|h\|_\infty ||k||_{\mathcal{BV}} + \|k\|_\infty ||h||_{\mathcal{BV}}. \tag{4.4}$$

We set, for $k = 0,\ \dots,\ n-1$ and $n = 1,\ 2,\ \dots,$

$$Z_{n,k} = g_n(X_{n-k-1}) - \mathbf{E}\left(g_n(X_{n-k-1})\right),\ \text{and}\ S_n = Z_{n,0} + \cdots + Z_{n,n-1}.$$

Now let $S_{k,n} = Z_{n,0} + \cdots + Z_{n,k}$ for $0 \le k \le n-1$. Empty sums are, as usual, set equal to 0.

Hence

$$\lim_{n\to\infty} \operatorname{Var} S_n = f(x)\int_D K^2(s)\mathrm{d}s > 0. \tag{4.5}$$

Consider now a bounded thrice differentiable function $h : \mathbb{R} \to \mathbb{R}$ with continuous and bounded derivatives. Set $C_j = \|h^{(j)}\|_\infty$, for $j = 0,\ 1,\ 2,\ 3$. Set $\sigma_n^2 = \operatorname{Var} S_n$. For some standard Gaussian random variable η, write $\Delta_n(h) = \mathbf{E}(h(S_n) - h(\sigma_n\eta))$. The theorem will follow from (4.5), if we prove that $\lim_{n\to\infty}\Delta_n(h) = 0$. Let $v_{k,n} := \operatorname{Var} S_{k,n} - \operatorname{Var} S_{k-1,n}$, for $1 \le k \le n-1$.

$$v_{k,n} = 2\sum_{l=0}^{k-1} \operatorname{Cov}\left(Z_{n,k}, Z_{n,l}\right) + \mathbf{E}Z_{n,k}^2 =: a_{n,k}^1 + a_n^2,$$

with

$$a_n^2 \sim \frac{f(x)}{n}\int_D K^2(t)\mathrm{d}t,$$

a_n^2 independent of k, and

$$0 \leq a_{n,k}^1 \leq \frac{2}{n}\sum_{l=1}^{k}\min\left(\varepsilon(n,l),d_l\right).$$

So, as in Lemma 3.3, we show that $\sup_{0\leq k\leq n-1}(na_{n,k}^1) \longrightarrow 0$ as $n \longrightarrow \infty$. Then there exists $n_o \in \mathbb{N}^*$ such that for all $n \geq n_o$ and for all $k \in \mathbb{N}^*$ such that $0 \leq k \leq n-1$ we have $v_{k,n} > 0$. Therefore we can consider $Y_{n,k} \sim \mathcal{N}(0, v_{k,n})$ for sufficiently large n and for $0 \leq k \leq n-1$.

Let us assume that the array $\{Y_{n,k};\ 0 \leq k \leq n-1,\ n \geq n_0\}$ is independent and is independent of the sequence $(X_k)_{k\in\mathbb{N}}$. If $0 \leq k \leq n-1$, set $T_{n,k} = \sum_{j=k+1}^{n-1} Y_{n,j}$, still with empty sums set equal to 0. We can now write Rio's decomposition

$$\Delta_n(h) = \sum_{k=0}^{n-1}\Delta_{k,n}(h),$$

with $\Delta_{k,n}(h) = \mathbf{E}\left[h(S_{k-1,n}+Z_{n,k}+T_{n,k}) - h(S_{k-1,n}+Y_{n,k}+T_{n,k})\right]$.

The function $x \mapsto h_{k,n}(x) = \mathbf{E}h(x+T_{n,k})$ has the same derivability properties as h, *e.g.* for $0 \leq j \leq 3$, $\|h_{k,n}^{(j)}\| \leq C_j$; now we write $\Delta_{k,n}(h) = \Delta_{k,n}^{(1)}(h) - \Delta_{k,n}^{(2)}(h)$, with

$$\begin{aligned}
\Delta_{k,n}^{(1)}(h) &= \mathbf{E}h_{k,n}(S_{k-1,n}+Z_{n,k}) - \mathbf{E}h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbf{E}h_{k,n}''(S_{k-1,n}),\\
\Delta_{k,n}^{(2)}(h) &= \mathbf{E}h_{k,n}(S_{k-1,n}+Y_{n,k}) - \mathbf{E}h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbf{E}h_{k,n}''(S_{k-1,n}).
\end{aligned}$$

- Bound of $\Delta_{k,n}^{(2)}(h)$.

  Using Taylor expansion yields for some (random) $\rho_{n,k} \in (0,\ 1)$:

$$\Delta_{k,n}^{(2)}(h) = \mathbf{E}h_{k,n}'(S_{k-1,n})Y_{n,k} + \frac{1}{2}\mathbf{E}h_{k,n}''(S_{k-1,n})(Y_{n,k}^2 - v_{k,n}) + \frac{1}{6}\mathbf{E}h_{k,n}^{(3)}(S_{k-1,n}+\rho_{n,k}Y_{n,k})Y_{n,k}^3.$$

From the independence of the Gaussian sequence $(Y_{n,k})_{n\in\mathbb{N},\,0\leq k\leq n-1}$ and the process $(X_n)_{n\in\mathbb{N}}$,

$$|\Delta_{k,n}^{(2)}(h)| \leq \frac{C_3}{6}\mathbf{E}|Y_{n,k}|^3\ ,$$

hence

$$|\Delta_{k,n}^{(2)}(h)| \leq \frac{2C_3 v_{k,n}^{\frac{3}{2}}}{3\sqrt{2\pi}}.$$

Now

$$v_{k,n} = \mathrm{Var}\, Z_{n,k} + 2\sum_{j=0}^{k-1}\mathrm{Cov}(Z_{n,j},\ Z_{n,k}),$$

hence

$$v_{k,n} \leq 4M_n\delta_n + 2\sum_{j=1}^{k}2\delta_n l_n d_j. \tag{4.6}$$

We thus need

$$n^{\frac{2}{3}}\left(\frac{1}{n}+\sum_{j=1}^{k}\frac{d_j}{n}\right)\xrightarrow[n\to\infty]{}0.$$

As $\sum_{k=0}^{\infty} d_k<\infty$, the last assertion is always true.

- Bound of $\Delta^{(1)}_{k,n}(h)$.

Set $\Delta^{(1)}_{k,n}(h)=\mathbf{E}\delta^{(1)}_{k,n}(h)$. We write, again with some random $\tau_{n,k}\in(0,\ 1)$,

$$\delta^{(1)}_{k,n}(h)=h^{'}_{k,n}(S_{k-1,n})Z_{n,k}+\frac{1}{2}h^{''}_{k,n}(S_{k-1,n})(Z^2_{n,k}-v_{k,n})+\frac{1}{6}\left(h^{(3)}_{k,n}(S_{k-1,n}+\tau_{n,k}Z_{n,k})Z^3_{n,k}\right).$$

We analyze separately the terms in the previous expression. We have

$$\frac{1}{6}\left|\mathbf{E}h^{(3)}_{k,n}(S_{k-1,n}+\tau_{n,k}Z_{n,k})Z^3_{n,k}\right|\leq\frac{C_3}{6}(4M_n^2)(2\delta_n). \tag{4.7}$$

To estimate the middle term, write (with Rio)

$$\mathrm{Cov}\left(h^{''}_{k,n}(S_{k-1,n}),\ Z^2_{n,k}\right)=\sum_{j=0}^{k-1}\mathrm{Cov}\left(h^{''}_{k,n}(S_{j,n})-h^{''}_{k,n}(S_{j-1,n}),\ Z^2_{n,k}\right).$$

By Taylor,

$$\mathrm{Cov}\left(h^{''}_{k,n}(S_{j,n})-h^{''}_{k,n}(S_{j-1,n}),\ Z^2_{n,k}\right)=\mathrm{Cov}\left(Z_{n,j}\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),Z^2_{n,k}\right),$$

for some $0<u<1$.

We can also write

$$Z^2_{n,k}=g_n^2(X_{n-k-1})-2\mathbf{E}\left(g_n(X_{n-k-1})\right)g_n(X_{n-k-1})+\left[\mathbf{E}\left(g_n(X_{n-k-1})\right)\right]^2,$$

and

$$Z_{n,j}=g_n(X_{n-j-1})-\mathbf{E}\left(g_n(X_{n-j-1})\right).$$

From those decompositions, the dominant term of

$$\mathrm{Cov}\left(Z_{n,j}\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),Z^2_{n,k}\right)$$

is

$$\mathrm{Cov}\left(g_n(X_{n-j-1})\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),g_n^2(X_{n-k-1})\right),$$

by replacing twice $Z_{n,l}$ by $g(X_{n-l-1})$. Hence using (4.4) and Remark 3.3 we obtain:

$$\mathrm{Cov}\left(g_n(X_{n-j-1})\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),g_n^2(X_{n-k-1})\right)\leq C_3\ \delta_n\ 2M_nl_nd_{k-j}. \tag{4.8}$$

So, as it is the dominant term, we get the same bound for the other terms. Summing up yields:

$$\left|\mathrm{Cov}\left(h^{''}_{k,n}(S_{k-1,n}),\ Z^2_{n,k}\right)\right|\leq 4\sum_{j=0}^{k-1}C_3\ \delta_n\ 2M_nl_nd_j. \tag{4.9}$$

Proceeding as for (4.9) implies:

$$\left|\mathrm{Cov}\left(h^{'}_{k,n}(S_{i,n})-h^{'}_{k,n}(S_{i-1,n}),\ Z_{n,k}\right)\right|\leq 2C_2\delta_nl_nd_{k-i}. \tag{4.10}$$

Hence from (4.10) and from Remark 3.3:

$$\left|\mathrm{Cov}\left(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}),\ Z_{n,k}\right)\right| \leq c \min\left(\delta_n l_n d_{k-i},\ \frac{\varepsilon(n,k-i)}{n}\right). \tag{4.11}$$

We also have

$$\left|\mathbf{E}h''_{k,n}(S_{k-1,n})\mathbf{E}Z_{n,i}Z_{n,k}\right| \leq c \min\left(\delta_n l_n d_{k-i},\ \frac{\varepsilon(n,k-i)}{n}\right). \tag{4.12}$$

Adding (4.11) and (4.12) and summing up the expression for all i yields:

$$\left|\mathbf{E}h'_{k,n}(S_{k-1,n})Z_{n,k} - \mathbf{E}h''_{k,n}(S_{k-1,n})\sum_{i=0}^{k-1}\mathbf{E}Z_{n,i}Z_{n,k}\right| \leq c\sum_{p=1}^{k}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right). \tag{4.13}$$

We add equations (4.7), $\frac{1}{2}$(4.9) and (4.13) to obtain:

$$\left|\Delta^{(1)}_{k,n}(h)\right| \leq c\left(M_n^2\delta_n + \delta_n M_n l_n \sum_{p=0}^{k-1} d_p + \sum_{p=1}^{k}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right). \tag{4.14}$$

We sum (4.14) for all k to conclude:

$$\left|\sum_{k=0}^{n-1}\Delta^{(1)}_{k,n}(h)\right| \leq c\times n\left(M_n^2\delta_n + \delta_n M_n l_n \sum_{p=0}^{\infty} d_p + \sum_{p=1}^{\infty}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right). \tag{4.15}$$

With the techniques used to prove Lemma 3.3 we can prove

$$\lim_{n\to\infty} n\left(\sum_{p=1}^{\infty}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right) = 0.$$

Replacing $M_n,\ \delta_n,\ l_n$ by their upper bounds, we easily see on (4.15) that

$$\sum_{k=0}^{n-1}\Delta^{(1)}_{k,n}(h) \xrightarrow[n\to\infty]{} 0.$$

This concludes the proof of Theorem 4.3. $\square$

Remark 4.3. The proof of Theorem 4.3 extends immediately to the case $g_n(t) = \frac{1}{\sqrt{nb_n}}K(\frac{x-\psi(t)}{b_n})$, where $\psi : I \to \mathbb{R}$ is some monotone function.

Instead of $\mathcal{BV}$ we can also take a Banach space $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ whose norm satisfies:

- there exists $M > 0$ such that for all n in $\mathbb{N}$, $\| K\left(\frac{x-\cdot}{b_n}\right)\|_{\mathcal{B}} \leq M$;
- $\sqrt{nb_n}\ \| K^2\left(\frac{x-\cdot}{b_n}\right)\|_{\mathcal{B}} \xrightarrow[n\to\infty]{} 0$.

Equation (4.9) and Lemma 3.2 still hold, thus we can prove Theorem 4.3.

Lipschitz norm does not yield the second point above. Hence we cannot replace the norm $||.||_{\mathcal{BV}}$ by the Lipschitz norm. Therefore the class of examples is not really large. It is a real problem which is due to the kernel density estimates whose Lipschitz norm has order $\mathcal{O}\left(\frac{1}{b_n}\right) \xrightarrow[n\to+\infty]{} +\infty$.

Proof of Theorems 4.1 and 4.2. For sake of simplicity, we write the proof for $l = 1$. We first write

$$\hat{f}_n(x) - f(x) = \left[\hat{f}_n(x) - \mathbf{E}\hat{f}_n(x)\right] + \left[\mathbf{E}\hat{f}_n(x) - f(x)\right].$$

The behaviour of $\hat{f}_n(x) - \mathbf{E}\hat{f}_n(x)$ is given by Theorem 4.3. Hence we restrict our attention to $\mathcal{BIAS}_n(x) = \mathbf{E}\hat{f}_n(x) - f(x)$.

The sequence $(X_n)_{n\in\mathbb{N}}$ is stationary because X_0 follows the invariant law μ_0.

Hence as $\int_D K(s)\mathrm{d}s = 1$,

$$\mathcal{BIAS}_n(x) = \int_D K(s)[f(x - sb_n) - f(x)]\mathrm{d}s. \tag{4.16}$$

Let ν denote the regularity of f on V.

Case $\nu \in \mathbb{N}^*$ (Th. 4.1):

As D is compact and as $b_n \xrightarrow[n\to\infty]{} 0$, there exists some $n_0 \in \mathbb{N}^*$ such that for $n \geq n_0$ the interval $J_{s,n} := [\min(x - sb_n, x),\ \sup(x - sb_n, x)]$ is included in V for all $s \in D$. Hence using (4.16) and Taylor's decomposition on each $J_{s,n}$ we get for $n \geq n_0$

$$\mathcal{BIAS}_n(x) = \int_D K(s)\left\{\sum_{j=1}^{m-1} \frac{(-sb_n)^j}{j\,!}\ f^{(j)}(x) + \frac{(-sb_n)^m}{m!} f^{(m)}(x - st_{s,n}b_n)\right\}\mathrm{d}s,$$

where for all $n \geq n_0$ and for all $s \in D$, $t_{s,n}$ is some real satisfying $0 < t_{s,n} < 1$.

Hence, as $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \leq j \leq m-1$,

$$\mathcal{BIAS}_n(x) = \int_D K(s)\frac{(-sb_n)^m}{m!}\ f^{(m)}(x - st_{s,n}b_n)\mathrm{d}s.$$

Then as $f^{(m)}$ is continuous in x and as $0 < t_{s,n} < 1$ for all $s \in D$ and for all $n \geq n_0$, we have for each fixed $s \in D : t_{s,n}sb_n \xrightarrow[n\to\infty]{} 0$. Hence, proceeding as in the proof of Lemma 3.2 (study of Q_n), we get, by the dominated convergence theorem,

$$\int_D \frac{(-s)^m}{m!}K(s)f^{(m)}(x - st_{s,n}b_n)\mathrm{d}s \xrightarrow[n\to\infty]{} f^{(m)}(x)\int_D \frac{(-s)^m}{m!}K(s)\mathrm{d}s.$$

Hence, as soon as there exists $\rho_m \in \mathbb{R}^+$ such that $nb_n^{2m+1} \xrightarrow[n\to\infty]{} \rho_m$,

$$\sqrt{nb_n}\ \mathcal{BIAS}_n(x) \xrightarrow[n\to\infty]{} \frac{(-1)^m\sqrt{\rho_m}}{m!}\ f^{(m)}(x)\int_D s^m K(s)\mathrm{d}s.$$

This concludes the study of Theorem 4.1.

General case (Th. 4.2):

Here we use the integral form of Taylor's decomposition. Let $n_0 \in \mathbb{N}^*$ be such that for $n \geq n_0$ the interval $J_s := [\min(x - sb_n, x),\ \sup(x - sb_n, x)]$ is included in V for all $s \in D$. Recall that $\nu = \alpha + \beta$ where $\alpha \in \mathbb{N}$ and $0 \leq \beta < 1$.

Empty sums are set equal to 0.

For $n \geq n_0$,

$$\mathcal{BIAS}_n(x) = \int_D \left\{\sum_{j=1}^{\alpha-1} \frac{(-sb_n)^j}{j\,!} f^{(j)}(x) + \int_0^1 \frac{(-sb_n)^\alpha}{(\alpha-1)!}(1-t)^{\alpha-1} f^{(\alpha)}(x - stb_n)\mathrm{d}t\right\} K(s)\mathrm{d}s.$$

As $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \le j \le \alpha$ we deduce:

$$\mathcal{BIAS}_n(x) = \int_D \int_0^1 \frac{(1-t)^{\alpha-1}}{(\alpha-1)!}(-b_n)^\alpha s^\alpha [f^{(\alpha)}(x - stb_n) - f^{(\alpha)}(x)]\mathrm{d}tK(s)\mathrm{d}s.$$

As f has the regularity $\nu = \alpha + \beta$ (in terms of Hölder spaces) on V, there exists some constant A independent of n such that:

$$|\mathcal{BIAS}_n(x)| \le \int_D \frac{1}{(\alpha-1)!}\, b_n^\alpha s^\alpha \; A\, b_n^\beta s^\beta \; |K(s)|\,\mathrm{d}s = \frac{A\, b_n^\nu}{(\alpha-1)!} \int_D s^\nu \,|K(s)|\,\mathrm{d}s.$$

As D is compact and as $\int_D K^2(s)\mathrm{d}s < \infty$, we have $\int_D s^\nu\,|K(s)|\,\mathrm{d}s < \infty$. Hence there exists some non-negative constant C independent of n such that:

$$|\mathcal{BIAS}_n(x)| \le C\, b_n^\nu.$$

So $\sqrt{nb_n}\; \mathcal{BIAS}_n(x) \xrightarrow[n\to\infty]{} 0$ as soon as $nb_n^{2\nu+1} \xrightarrow[n\to\infty]{} 0$. This concludes the proof of Theorem 4.2. □

5. Examples of dynamical systems in the class $\mathcal{T}$

Without being exhaustive we will now give some examples of dynamics T which satisfy all the previous assumptions.

- Lasota–Yorke functions

Let T be some piecewise smooth expanding map of the interval $[0,1]$. Following Viana ([27], Chap. 3), we introduce the following set of assumptions for T.

Assumption 5.1. *(regularity). There exists* $0 = a_0 < a_1 < \cdots < a_l = 1$ *such that the restriction of* T *to each* $\eta_i = (a_{i-1}, a_i)$ *is of class* C^1*, with* $|T'(x)| > 0$ *for all* $x \in \eta_i$ *and* $i = 1, \ldots, l$.
Moreover, the function $g_{\eta_i} = \frac{1}{|T'_{\eta_i}|}$ *has bounded variation for* $i = 1, \ldots, l$.

If h is some function on I and if $J \subset I$, h_J denotes the restriction of h to J.
Using this notation: (T_{η_i}) and g_{η_i} admit continuous extensions to $\bar{\eta}_i = [a_{i-1}, a_i]$ for each $i = 1, \ldots, l$. Since modifying the values of a map over a finite set of points does not change its statistical properties, we may assume that T is either left-continuous or right-continuous (or both) at a_i, for each $i = 1, \ldots, l$.
Then let $P^{(1)}$ be some partition of I into intervals η such that $\eta_i \subset \eta \subset \bar{\eta}_i$ for some i and such that (T_η) is continuous.
For $n \ge 1$, $P^{(n)}$ is the Markov partition of I: $P^{(n)}(x) = P^{(n)}(y)$ if and only if $P^{(1)}(T^j(x)) = P^{(1)}(T^j(y))$ for all $0 \le j < n$. ($P^{(n)}$ is the largest partition on which T^n is monotone.)
Given $\eta \in P^{(n)}$, denote $g_\eta^{(n)} = \frac{1}{|(T_\eta^n)'|}$.

Assumption 5.2. *(expansivity). There exist* $C_1 > 0$ *and* $\lambda_1 < 1$ *such that* $\sup_{t\in\eta} g_\eta^{(n)}(t) \le C_1\lambda_1^n$ *for all* $\eta \in P^{(n)}$ *and all* $n \ge 1$.

Assumption 5.3. *(topological mixing). There is an interval* $I_* \subset I = [0,1]$ *such that* $T(I_*) = I_*$*, every orbit* $T^n(x)$*,* $x \in (0,1)$*, eventually enters* I_**, and* T_{I_*} *is topologically mixing: for each interval* $J \subset I_*$ *there is* $n \ge 1$ *such that* $T^n(J) = I_*$.

Lasota and Yorke [16], Liverani [17], Viana [27] and others study such functions. It may be shown (Viana [27]) that T admits a unique absolutely continuous invariant probability measure μ_0 ($\mathrm{d}\mu_0 = f\ \mathrm{d}t$ where $\mathrm{d}t$ is the Lebesgue measure on $[0,1]$). In addition, μ_0 is ergodic and its support coincides with I_*.

We also have that f has bounded variation on $[0,1]$. This implies that f is continuous on $[0,1]$ except at most for countably many points. So our assumptions in Section 1 are satisfied.

The "tent-maps" having a large enough slope and the r-adic transformations, for $r > 1$, of the interval are two examples of Lasota–Yorke functions.

- **"tent-map"**
  It has $T^{'}$ constant and strictly larger than 1 in absolute value, in each of the monotonicity intervals $[0,\ c)$ and $(c,\ 1]$.
  Moreover if $c = \frac{1}{2}$ and $|T'(x)| = \sigma > \sqrt{2}$ for all $x \neq c$, we have $I_* = [T^2(c),\ T(c)]$.

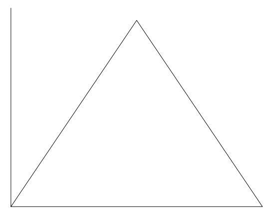


FIGURE 1. The "tent map" for $c = 0.5$ and $T(c) = 0.75$.

- **r-adic transformations** $(r > 1)$ of the interval $[0,\ 1]$. We have $I_* = I$.

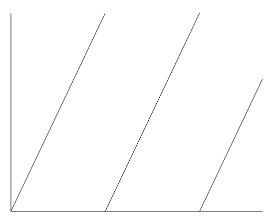


FIGURE 2. The r-adic map with $r = 8/3$.

- Functions with infinitely many monotonicity intervals. The precedent case extends, under appropriate conditions, to piecewise expanding maps with countably many domains of smoothness and monotonicity (see *e.g.* Broise [7], Viana [27]). For example if we consider the Gauss-map, that is the map T defined by $T(x) = \frac{1}{x} - [\frac{1}{x}]$ for $x \neq 0$ and $T(0) = 0$, we have summable decay of correlations and an invariant probability measure μ_0 absolutely continuous with respect to Lebesgue on $[0,1]$ and whose density has bounded variation. We have, keeping the former notation, $I_* = I$. Furthermore the Gauss-map satisfies the assumptions in Section 1. Therefore we have the result for the invariant density estimates in that case.

Remark 5.1. Let $T : [0,1] \to [0,1]$ defined by $T(x) = 4x(1-x)$, then Theorem 4.3 still holds because this map is obtained from a "tent map" by conjugation.

6. A non-stationary case

Lasota–Yorke functions T, introduced in Section 5, have additional properties which allow us to extend the previous results to the non-stationary case. We use the same definitions as before (*e.g.* (X_n) is the stationary dynamical system). Now let p be any density function on $I = [0,1]$ with bounded variation. We set $p(t) = 0$

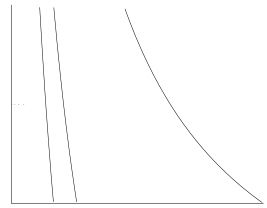


FIGURE 3. The Gauss-map.

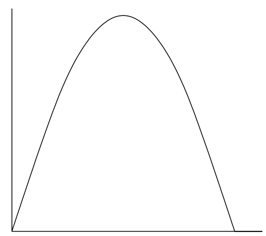


FIGURE 4. $T(x) = 4x(1-x)$.

for $t \notin I$. We define a random variable $X_0^{'}$ with distribution $p(t)dt$ and the (non-stationary) dynamical system $X_n^{'} = T^n X_0^{'}$, $n \geq 1$ (as in (1.6)).

6.1. Classical results

In the case of Lasota–Yorke function T, the invariant density f of T has bounded variation. We define the Perron–Frobenius operator $\mathcal{L}$ (for sake of simplicity we write $\mathcal{L}$ for $\mathcal{L}_T$) as follows:

$$\mathcal{L} : \begin{cases} \mathcal{BV} & \rightarrow & \mathcal{BV} \\ \omega & \mapsto & \mathcal{L}\omega(x) = \sum_{Ty=x} \dfrac{\omega(y)}{|T'(y)|} \end{cases}.$$

The following theorem collects properties of both the Perron–Frobenius operator $\mathcal{L}$ and the associated invariant density f of the Lasota–Yorke function T.

Theorem 6.1. (Liverani [17], Collet [8], Viana [27])

- $$h \in \mathcal{BV} \Longrightarrow \mathcal{L}^n h \in \mathcal{BV}\ \forall n \in \mathbb{N}. \textit{ Moreover } \sup_{n\in\mathbb{N}} ||\mathcal{L}^n h||_{\mathcal{BV}} < \infty. \tag{6.1}$$
- *If f denotes the invariant density, then there exists some $\gamma > 0$ such that:*
$$\frac{1}{\gamma} \leq f(t) \leq \gamma \textit{ for all } t \in I. \tag{6.2}$$
Then $\frac{1}{f}$ has also bounded variation.
- $$\exists\, R > 0,\ \exists\, 0 \leq \lambda < 1,\ \forall j \in \mathbb{N} : ||\mathcal{L}^j p - f||_\infty \leq R\, \lambda^j. \tag{6.3}$$
- *The correlations decrease exponentially fast. Hence there exists $\kappa > 0$ such that for any $h,\ k \in \mathcal{BV}$,*
$$|\mathrm{Cov}\,(h(X_0), k(X_n))\,| \leq \kappa\ \|k\|_1\ \|h\|_{\mathcal{BV}}\ \lambda^n\ \ \forall n \geq 0, \tag{6.4}$$

where Cov *denotes the covariance with respect to the invariant probability measure* μ_0, *and* λ *is the same as in* (6.3).

Remark 6.1. The second assertion of Theorem 6.1 yields that if T is Lasota–Yorke, we can not have $f(x) = 0$.

6.2. Convergence in mean squares

Let us first recall the definition (1.7) of $\hat{p}(x)$:

$$\hat{p}(x) = \hat{p}_n(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x - X_k^{'}}{b_n}\right),$$

where b_n and K are defined in the introduction by (1.3). The following result extends Lemma 3.1 to non-stationary dynamical systems, I_1 and I_2 being defined as in Section 2.

$\lambda(D) < \infty$ still denotes the Lebesgue measure of the compact D.

Lemma 6.1. *Let* T *be a Lasota–Yorke function and* $\hat{p}_n(x)$ *be defined by (1.3, 1.6) and (1.7). Assume that* $b_n \xrightarrow[n\to\infty]{} 0$. *Then for* $x \in I_1 \cap I_2$ *we have*

$$\mathrm{Var}\left(\hat{p}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right).$$

Proof of Lemma 6.1. Write

$$(nb_n)\mathrm{Var}\left(\hat{p}(x)\right) = V_n + \frac{2}{nb_n}\sum_{1\le i<j\le n} C_{i,j} \tag{6.5}$$

with

$$V_n := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\mathbf{E}K^2\left(\frac{x - X_k^{'}}{b_n}\right) - \sum_{k=0}^{n-1}\left(\mathbf{E}K\left(\frac{x - X_k^{'}}{b_n}\right)\right)^2\right) \tag{6.6}$$

and

$$C_{i,j} := \mathrm{Cov}\left(K\left(\frac{x - X_i'}{b_n}\right), K\left(\frac{x - X_j'}{b_n}\right)\right). \tag{6.7}$$

Note that $C_{i,j}$ depends on n.

- **Study of** $\frac{2}{nb_n}\sum_{0\le i<j\le n-1} C_{i,j}$**:**

$$\begin{aligned}
|C_{i,j}| &= \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)K\left(\frac{x - T^{j-i}t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^j p(t)\mathrm{d}t\right| \\
&\le \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)K\left(\frac{x - T^{j-i}t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)f(t)\mathrm{d}t\right| \\
&\quad + \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^j p(t) - f(t)\right)\mathrm{d}t\right| =: A_{n,i,j} + B_{n,i,j}.
\end{aligned}$$

By inequality (6.2) of Theorem 6.1, f is bounded below by $\frac{1}{\gamma} > 0$, so we can write:

$$A_{n,i,j} = \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) K\left(\frac{x-T^{j-i}t}{b_n}\right) \frac{\mathcal{L}^i p(t)}{f(t)} f(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right) \frac{\mathcal{L}^i p(t)}{f(t)} f(t)\mathrm{d}t \int_0^1 K\left(\frac{x-t}{b_n}\right) f(t)\mathrm{d}t \right|.$$

We note that $A_{n,i,j} = |\mathrm{Cov}\,(h(X_0), k(X_{j-i}))\,|$, where $h : t \mapsto K(\frac{x-t}{b_n})\frac{\mathcal{L}^i p(t)}{f(t)}$ and $k : t \mapsto K(\frac{x-t}{b_n})$.

In the following c will denote some constant independent of n, i, and j which may vary from line to line.

Quote that there exists a constant c independent of n such that $||\mathcal{L}^n p||_\infty \le c$. Indeed, $||\mathcal{L}^n p||_\infty \le ||\mathcal{L}^n p||_{\mathcal{BV}}$ and by assertion (6.1) in Theorem 6.1, $\sup_{n\in\mathbb{N}} ||\mathcal{L}^n p||_{\mathcal{BV}} < \infty$.

We obtain $A_{n,i,j} \le c\, b_n \lambda^{j-i}$ by using Theorem 6.1.

Analogously, for $B_{n,i,j}$, we have

$$B_{n,i,j} \le \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t \right| \int_0^1 \left| K\left(\frac{x-t}{b_n}\right) \left(\mathcal{L}^j p(t) - f(t)\right) \right| \mathrm{d}t \le b_n^2\, c\, \lambda^j.$$

Using the bound $\lambda < 1$ now yields $A_{n,i,j} + B_{n,i,j} \le c\, b_n \lambda^{j-i}$.

Remark 6.2. Using similar arguments, we can prove that given any $h \in \mathcal{BV}$ and $k \in \mathbb{L}^1(dt)$ we have for all $0 \le i < j \le n-1$

$$|\mathrm{Cov}\left(h(X_i'), k(X_j')\right)| \le c\, ||h||_{\mathcal{BV}} ||k||_1 \lambda^{j-i}. \tag{6.8}$$

Using Theorem 6.1 and the proof of Lemma 3.2, we get

$$\begin{aligned} |C_{i,j}| &= \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) K\left(\frac{x-T^{j-i}t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^j p(t)\mathrm{d}t \right| \\ &\le b_n\, \varepsilon(n, j-i), \end{aligned}$$

where for $j-i$ fixed in $\mathbb{N}^*$, $\varepsilon(n, j-i) \xrightarrow[n\longrightarrow\infty]{} 0$.

Therefore

$$\left| \frac{2}{nb_n} \sum_{0\le i<j\le n-1} C_{i,j} \right| \le \frac{2\,c}{n} \sum_{0\le i<j\le n-1} \min\left(\lambda^{j-i}, \varepsilon(n, j-i)\right) \le c \sum_{k=1}^{n-1} \min\left(\lambda^k, \varepsilon(n,k)\right) \cdot \tag{6.9}$$

The right hand side of this inequality tends to 0 as n tends to infinity.

Remark 6.3. Using similar arguments, we can prove that for all $0 \le i < j \le n-1$, and for any bounded function φ,

$$\mathrm{Cov}\left(\varphi(X_0)\left(K\left(\frac{x-T^jX_0'}{b_n}\right) - \mathbf{E}K\left(\frac{x-T^jX_0'}{b_n}\right)\right), K\left(\frac{x-T^iX_0'}{b_n}\right) - \mathbf{E}K\left(\frac{x-T^iX_0'}{b_n}\right)\right) \le b_n\, \varepsilon(n, j-i),$$

with $\varepsilon(n, j-i) \xrightarrow[n\to\infty]{} 0$.

• **Study of V_n:**

Using $\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x)$ introduced in formula (3.2), it is worth decomposing V_n as follows:

$$V_n = (nb_n)\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) + (s_n + s_n^{'}) := (nb_n)\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) + \Big(V_n - (nb_n)\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x)\Big), \tag{6.10}$$

where

$$s_n := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\mathbf{E}K^2\left(\frac{x-X_k^{'}}{b_n}\right) - \mathbf{E}K^2\left(\frac{x-X_k}{b_n}\right)\right)$$

and

$$s_n^{'} := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\left(\mathbf{E}K\left(\frac{x-X_k^{'}}{b_n}\right)\right)^2 - \left(\mathbf{E}K\left(\frac{x-X_k}{b_n}\right)\right)^2\right).$$

We have

$$|s_n| \;=\; \left|\tfrac{1}{nb_n}\left(\sum_{k=0}^{n-1}\int_0^1 K^2\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) - f(t)\right)\mathrm{d}t\right)\right|.$$

Hence, using (6.3) in Theorem 6.1 we get

$$|s_n| \le \frac{1}{n}\int_D K^2(s)\sum_{k=0}^{n-1}\left|\left(\mathcal{L}^k p(x-sb_n) - f(x-sb_n)\right)\right|\mathrm{d}s \le \frac{R\,\int_D K^2(s)\mathrm{d}s}{n}\sum_{k=0}^{n-1}\lambda^k. \tag{6.11}$$

The right hand side of this inequality tends to 0 as n tends to infinity.

Let $K_n' = K\left(\frac{x-X_n^{'}}{b_n}\right)$, $K_n = K\left(\frac{x-X_n}{b_n}\right)$. Then using the following identity

$$\left(\mathbf{E}K_n'\right)^2 - \left(\mathbf{E}K_n\right)^2 = \left(\mathbf{E}\left(K_n'-K_n\right)\right)\left(\mathbf{E}\left(K_n'+K_n\right)\right),$$

we have by (6.1, 6.2) and (6.3) in Theorem 6.1:

$$\begin{aligned}|s_n^{'}| \;&=\; \left|\frac{1}{nb_n}\sum\nolimits_{k=0}^{n-1}\left(\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) - f(t)\right)\mathrm{d}t\right)\left(\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) + f(t)\right)\mathrm{d}t\right)\right| \\ &\le\; \frac{R}{n}\left(\sum\nolimits_{k=0}^{n-1}\lambda^k\right)\; b_n\;\left(\sup_n ||\mathcal{L}_n p||_\infty + \gamma\right)\;\lambda(D)\;\int_D K^2(s)\mathrm{d}s.\end{aligned}$$

As $0 \le \lambda < 1$ and as $\int_D K^2(s)\mathrm{d}s < \infty$, the right hand side of this inequality tends also to 0 as n tends to infinity.

Hence by (6.10, 6.11) and (6.2),

$$V_n \sim (nb_n)\;\mathrm{Var}_{\,\mathrm{ind}}\hat{f}(x) \sim f(x)\int_D K^2(t)\mathrm{d}t > 0 \;\text{ as } n \text{ tends to infinity.} \tag{6.12}$$

Collecting (6.5) and bounds (6.9) and (6.12) yields the result:

$$(nb_n)\mathrm{Var}\left(\hat{p}(x)\right) \xrightarrow[n\to+\infty]{} f(x)\int_D K^2(t)\mathrm{d}t.$$

□

6.3. **Central Limit Theorem**

Theorems 6.2, 6.3 and 6.4 below deal with the non-stationary case and are analogous to Theorems 4.1, 4.2 and 4.3 of Section 4. The setting is not really different from the stationary one. We first study the following process

$$U_n^{'}(x) := \sqrt{nb_n}(\hat{p}(x) - f(x)). \tag{6.13}$$

Theorem 6.2. *Let T be a Lasota–Yorke function and $\hat{p}_n(x)$ be defined by (1.3, 1.6) and (1.7). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \leq i \leq l$, let $x_i \in I_1 \cap I_2$. Let m be a positive integer. Assume that for each i, $1 \leq i \leq l$, there exists a neighbourhood V_i of x_i such that the invariant density f is $m-$times continuously differentiable on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \leq j \leq m-1$ and that nb_n^{2m+1} converges to some non-negative constant ρ_m as n tends to infinity. Then*

$$(U_n'(x_1),\ \ldots\ ,\ U_n'(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(\frac{(-1)^m\ \sqrt{\rho_m}}{m!}\ \int_D s^m\ K(s)\mathrm{d}s * \left(f^{(m)}(x_i)\right)_{1\leq i\leq l},\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

Let us now consider the case where the invariant density f has a regularity $\nu = \alpha + \beta$ with $0 \leq \beta < 1$ and $\alpha \in \mathbb{N}$.

Theorem 6.3. *Let T be a Lasota–Yorke function and $\hat{p}(x)$ be defined by (1.3, 1.6) and (1.7). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \leq i \leq l$, let $x_i \in I_1 \cap I_2$. Let $\nu = \alpha + \beta$ with $0 \leq \beta < 1$ and $\alpha \in \mathbb{N}$. Assume that for each i, $1 \leq i \leq l$, there exists a neighbourhood V_i of x_i such that the invariant density f has the regularity ν on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \leq j \leq \alpha$ and that $nb_n^{2\nu+1} \xrightarrow[n\to+\infty]{} 0$. Then*

$$(U_n'(x_1),\ \ldots\ ,\ U_n'(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

As in the stationary case, we sometimes prefer to study the centered estimation process

$$Y_n^{'}(x) := \sqrt{nb_n}(\hat{p}(x) - \mathbf{E}\hat{p}(x)). \tag{6.14}$$

We then get:

Theorem 6.4. *Let T be a Lasota–Yorke function and $\hat{p}(x)$ be defined by (1.3, 1.6) and (1.7). For all $1 \leq i \leq l$, let $x_i \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$, $nb_n \xrightarrow[n\to\infty]{} \infty$. Then the finite dimensional marginals $(\overline{Y_n'}(x_1),\ \ldots\ ,\ \overline{Y_n'}(x_l))$ of the process*

$$\overline{Y_n'}(x) \equiv \frac{Y_n^{'}(x)}{\sqrt{f(x)\int_{-\infty}^{\infty} K^2(t)\mathrm{d}t}}$$

converge in distribution to a standard $\mathcal{N}(0,\ I_l)$ random variable.

Remark 6.4. Here we have normalized the process $Y_n^{'}(x)$. It is possible as $f(x) > 0$ for Lasota–Yorke functions T (see Th. 6.1).

As in the stationary case, we first prove Theorem 6.4 and then deduce Theorems 6.2 and 6.3 by studying the bias term. For the proof of Theorem 6.4 we develop the proof for $l = 1$ for sake of simplicity. If one wants to know the asymptotic behaviour in distribution of the vector $(Y_n^{'}(x_1),\ \ldots\ ,\ Y_n^{'}(x_l))$, it is sufficient to use the following proof (in the case $l=1$) with $\frac{1}{\sqrt{nb_n}}\sum_{j=1}^{l} s_j K(\frac{x_j-X_k^{'}}{b_n})$, for arbitrary numbers $s_1,\ \dots,\ s_l\in\mathbb{R}$, instead of $\frac{1}{\sqrt{nb_n}}K(\frac{x-X_k^{'}}{b_n})$.

Let us first give two applications of Theorem 6.4.

Application 6.1 Let r be a positive integer. For all $1\le i\le r$ we define

$$\hat{p}_n^{(i)}(x)=\frac{1}{nb_n}\sum_{k=0}^{n-1}K\left(\frac{x-X_k^{'\,(i)}}{b_n}\right)$$

where for all n, $\left(X_0^{'\,(i)},\dots,X_{n-1}^{'\,(i)}\right)\sim\left(X_0^{'},\dots,X_{n-1}^{'}\right)$. We assume that the sequences $X^{(i)}:=\left(X_k^{'\,(i)}\right)_{k\in\mathbb{N}}$, for $i\in\mathbb{N}$, are independent of each other. Hence $\hat{p}_n^{(0)}(x),\hat{p}_n^{(1)}(x),\dots,\hat{p}_n^{(r)}(x)$ are $r+1$ independent copies of $\hat{p}_n(x)$, and we can consider

$$\hat{Y}_n^{'}(x):=\sqrt{nb_n}\left\{\hat{p}_n^{(0)}(x)-\frac{\hat{p}_n^{(1)}(x)+\dots+\hat{p}_n^{(r)}(x)}{r}\right\}.$$

Let assume that $b_n\xrightarrow[n\to\infty]{}0$, $nb_n\xrightarrow[n\to\infty]{}\infty$. We also assume that r depends on n, $r=r(n)$, with $\frac{r(n)}{n}\le C$ where C is some positive constant and $r(n)b_n\xrightarrow[n\to\infty]{}\infty$. For example take $r(n)=n$. For all i and for large n,

$$\left|\hat{p}_n^{(i)}(x)-\mathbf{E}\hat{p}_n(x)\right|\le\frac{2\,||K||_\infty}{b_n}$$

and

$$\begin{aligned}\textstyle\sum_{i=1}^{r}\text{Var}\left(\hat{p}_n^{(i)}(x)\right)&=r\ \text{Var}\left(\hat{p}_n(x)\right)=\frac{r(n)}{nb_n}\left(f(x)\int_D K(s)\mathrm{d}s+\ o_n(1)\right)\\&\le\tfrac{C}{b_n}\left(f(x)\int_D K(s)\mathrm{d}s+\ 1\right).\end{aligned}$$

Hence using Bernstein's inequality in Pollard ([20], pp. 192, 193) we get for all $\eta>0$

$$P\left(\left|\tfrac{\hat{p}_n^{(1)}(x)+\dots+\hat{p}_n^{(r)}(x)}{r}-\mathbf{E}\hat{p}_n(x)\right|>\eta\right)\quad\le\quad 2\exp\left(\frac{-\eta^2\ r^2\ b_n}{2\ C\left(f(x)\ \int_D K^2(s)\mathrm{d}s+1\right)+\frac{4}{3}\ \eta\ r\ ||K||_\infty}\right).$$

The exponential term above tends to 0 as n tends to infinity. Hence, we can approach $Y_n^{'}(x)$ by the empirical quantity $\hat{Y'}_n(x)$. The advantage of $\hat{Y'}_n(x)$ is that it can be simulated. Indeed it does not involve the knowledge of $f(x)$.

Application 6.2 Now let $\hat{p}_n^{(1)}(x)$ and $\hat{p}_n^{(2)}(x)$ be two independent copies of $\hat{p}_n(x)$. The difference

$$\Phi_n(x)=\sqrt{nb_n}\left(\{\hat{p}_n^{(1)}(x)-\mathbf{E}\hat{p}_n^{(1)}(x)\}-\{\hat{p}_n^{(2)}(x)-\mathbf{E}\hat{p}_n^{(2)}(x)\}\right)$$

does not depend on $\mathbf{E}\hat{p}_n(x)$. Indeed $\Phi_n(x)=\sqrt{nb_n}\left(\hat{p}_n^{(1)}(x)-\hat{p}_n^{(2)}(x)\right)$.

Moreover, $\Phi_n(x)$ converges in distribution to a $\mathcal{N}(0,2\ f(x)\int_D K^2(s)\mathrm{d}s)$ random variable as n tends to infinity as soon as $b_n\xrightarrow[n\to\infty]{}0$, $nb_n\xrightarrow[n\to\infty]{}\infty$. To estimate $f(x)$ it can be useful to work with $\Phi_n(x)$ instead of $U_n^{'}(x)$

as $f(x)$ only appears in the variance term of the limit of $\Phi_n(x)$ and not in the quantity $\Phi_n(x)$ itself. Hence to simulate $\Phi_n(x)$ we have neither to approach $\mathbf{E}\hat{p}_n(x)$ using exponential inequalities as in Application 6.1 nor to know $f(x)$ (as $X_0^{'}$ has the distribution $p(t)\,\mathrm{d}t$ where p is known).

Proof of Theorem 6.4 with $l=1$. We use notations $g_n,\ M_n,\ l_n,$ and δ_n of Theorem 4.3. As in the proof of Theorem 4.3 we obtain $M_n \le \frac{c}{\sqrt{nb_n}}$, $l_n \le \frac{c}{\sqrt{nb_n}}$ and $\delta_n \le \frac{c\,b_n}{\sqrt{nb_n}}$, for some positive constant c.

We set for $k=0,\ \dots,\ n-1$ and $n=1,\ 2,\ \dots,$

$$Z_{n,k}^{'} = g_n(X_{n-k-1}^{'}) - \mathbf{E}\left(g_n(X_{n-k-1}^{'})\right), \text{ and } S_n^{'} = Z_{n,0}^{'} + \cdots + Z_{n,n-1}^{'}.$$

Now let $S_{k,n}^{'} = Z_{n,0}^{'} + \cdots + Z_{n,k}^{'}$ for $0 \le k \le n-1$. Empty sums are, as usual, set equal to 0.

Recall that

$$\lim_{n\to\infty} \operatorname{Var} S_n = f(x)\int_D K^2(s)\mathrm{d}s > 0. \tag{6.15}$$

We still consider a bounded thrice differentiable function $h:\mathbb{R}\to\mathbb{R}$ with continuous and bounded derivatives, with $C_j = \|h^{(j)}\|_\infty$, for $j=0,\ 1,\ 2,\ 3$. As in the proof of Theorem 4.3, also consider $\sigma_n^2 = \operatorname{Var} S_n$, and set in that case for some standard Gaussian r.v. η, $\Delta_n(h) = \mathbf{E}(h(S_n^{'}) - h(\sigma_n\eta))$. As in the proof of Theorem 4.3, the theorem will follow if we prove that $\lim_{n\to\infty}\Delta_n(h)=0$.

Let us assume $\{Y_{n,k};\ 0\le k\le n-1,\ n\ge n_0\}$ to be defined as in Section 3. $T_{n,k}$ is also defined as before. We are now in position to use Rio's decomposition

$$\Delta_n(h) = \sum_{k=0}^{n-1}\Delta_{k,n}(h),$$

with $\Delta_{k,n}(h) = \mathbf{E}(h(S_{k-1,n}^{'} + Z_{n,k}^{'} + T_{n,k}) - h(S_{k-1,n}^{'} + Y_{n,k} + T_{n,k}))$.

We still use the function $x \to h_{k,n}(x) = \mathbf{E}h(x+T_{n,k})$, which has the same derivability properties as the function h.

We proceed as in Section 4.

Inequality (6.8) replaces inequality (1.5).

For example inequality (4.12) in the proof of Theorem 4.3 is replaced by

$$|\mathbf{E}h_{k,n}^{''}(S_{k-1,n}^{'})\mathbf{E}Z_{n,i}^{'}Z_{n,k}^{'}| \le c\min\left(\delta_n l_n\lambda^{k-i},\ \frac{\varepsilon(n,k-i)}{n}\right).$$

Hence doing this with each inequality of the proof of Theorem 4.3 we conclude the proof of Theorem 6.4. $\square$

Let us now prove Theorems 6.2 and 6.3.

Proof of Theorems 6.2 and 6.3. We have the following decomposition:

$$\hat{p}_n(x) - f(x) = \hat{p}_n(x) - \mathbf{E}\hat{p}_n(x) + \mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x) + \mathbf{E}\hat{f}_n(x) - f(x).$$

The term $\hat{p}_n(x) - \mathbf{E}\hat{p}_n(x)$ is studied in Theorem 6.4. The term $\mathbf{E}\hat{f}_n(x) - f(x) = \mathcal{BIAS}_n(x)$ is studied in the proof of Theorems 4.1 and 4.2. It does not depend on the density p. So we just have to study the term $\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)$. We have

$$\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x) = \frac{1}{n}\int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K(s)\sum_{k=0}^{n-1}\left[\mathcal{L}^k p(x-sb_n) - f(x-sb_n)\right]\mathrm{d}s. \tag{6.16}$$

Using inequality (6.3) in Theorem 6.1, equality (6.16) and $\int_D K^2(s)\mathrm{d}s < \infty$ we get

$$\left|\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)\right| \leq \frac{R}{n}\sum_{k=0}^{n-1}\lambda^k \sqrt{\int_D K^2(s)\mathrm{d}s}\,\sqrt{\lambda(D)} \leq \frac{R}{n}\,\frac{1}{1-\lambda}\sqrt{\int_D K^2(s)\mathrm{d}s}\,\sqrt{\lambda(D)} \tag{6.17}$$

where $0 \leq \lambda < 1$. As $b_n \xrightarrow[n\to\infty]{} 0$ and as $\int_D K^2(s)\mathrm{d}s < \infty$, inequality (6.17) yields $\sqrt{nb_n}\;\left|\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)\right| \xrightarrow[n\to\infty]{} 0$. It concludes the proof of Theorems 6.2 and 6.3. ☐

Remark 6.5. Notice that $\sup_{t\in I}|\mathcal{L}^n 1(t) - f(t)|$ tends to 0 exponentially fast, hence $\mathcal{L}^n 1$ also appears to be a good evaluation of f. Unfortunately explicit computations of $\mathcal{L}^n 1$ involve a complete knowledge of iterated preimages with respect to T. Example given for $r \in \mathbb{N}^*$ the $r-$adic transformation involves r^n such preimages.

Appendix A. Appendix

Extension to weak dependent sequences

We extend here the results of this paper to weak dependent sequences. Let us first introduce our dependence frame which is a variation on the definition in Doukhan and Louhichi [13]. Assume that, for convenient functions h and k,

$$\mathrm{Cov}\,(h(\text{“past”}), k(\text{“future”}))$$

converges to 0 as the distance between the “past” and the “future” converges to infinity. Here “past” and “future” refer to the values of some time series of interest. Asymptotically, this means that independence holds if we use a *determining* function class.

More precisely, E being some Euclidean space $\mathbb{R}^d$ endowed with its Euclidean norm $\|.\|$, we shall consider a sequence of E-valued random variables $(\xi_n)_{n\in\mathbb{N}}$. We define $\mathbb{L}^\infty$ as the set of measurable and bounded numerical functions on some space $\mathbb{R}^k$, $k \in \mathbb{N}^*$ and its norm is classically written $\| \bullet \|_\infty$.

Moreover, let $u \in \mathbb{N}^*$ be a positive integer. We endow the set $F = E^u$ with the norm

$$\|(x_1, \ldots, x_u)\|_F = \|x_1\| + \cdots + \|x_u\|.$$

Let now $h : F = E^u \to \mathbb{R}$ be a numerical function on F, we set

$$\mathrm{Lip}(h) = \sup_{x\neq y}\frac{|h(x) - h(y)|}{\|x - y\|_F}$$

the Lipschitz modulus of h. Define

$$\mathcal{L} = \bigcup_{u=1}^{\infty}\left\{h \in \mathbb{L}^\infty(E^u, \mathbb{R});\ \|h\|_\infty \leq 1, \mathrm{Lip}(h) < \infty\right\}\cdot \tag{A.1}$$

Definition A.1. The sequence $(\xi_n)_{n\in\mathbb{N}}$ is s-weakly (resp. a-weakly) dependent, if for some sequence $\theta = (\theta_r)_{r\in\mathbb{N}}$ decreasing to zero at infinity and any (u+1)-tuple $(i_1, \ldots, i_u, j_1)$ with $i_1 \leq \ldots \leq i_u < i_u + r \leq j_1$, for $h \in \mathbb{L}^\infty$ satisfying $\|h\|_\infty \leq 1$ and for $k \in \mathcal{L}$,

$$|\mathrm{Cov}\,(h(\xi_{i_1}, \ldots, \xi_{i_u}), k(\xi_{j_1}))\,| \leq \mathrm{Lip}(k)\theta_r, \tag{A.2}$$

and respectively for $h, k \in \mathcal{L}$

$$|\mathrm{Cov}\,(h(\xi_{i_1}, \ldots, \xi_{i_u}), k(\xi_{j_1}))\,| \leq \mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r. \tag{A.3}$$

The results presented in this appendix improve CLTs stated by Doukhan and Louhichi in a more general non-causal frame (see [13]). We work here indeed under a fundamental causality assumption. Contrarly to Doukhan and Louhichi [13], we do not use Bernstein blocks but a variation on the Lindeberg–Rio method. We also relax assumptions in Coulon–Prieur and Doukhan [9]. Indeed, in [9], the authors need two points in the future. Here we just consider one point in the future ξ_{j_1}.

Note that the notions of weak dependence and dynamical systems are not that much different. For example let us define the autoregressive model by:

$$\xi_n = T(\xi_{n-1}) + \eta_n, \tag{A.4}$$

with $T : \mathbb{R} \to \mathbb{R}$ such that $|T(u) - T(u')| \le c|u - u'|$ for some $0 \le c < 1$ and for all $u,\ u' \in \mathbb{R}$, and with $(\eta_n)_{n\in\mathbb{Z}}$ some real valued i.i.d inovation process satisfying $\mathbf{E}|\eta_0| < \infty$. This model is s-weakly dependent. A generalization of this model is given by:

$$X_{n+1} = F(X_n, \varepsilon_{n+1}),$$

with $(\varepsilon_i)_{i\in\mathbb{N}}$ a sequence of independent random variables (r.v.s) and with F a measurable function. Such Markov chains are actually noisy dynamical systems (see Baladi *et al.* [3]).

We refer to [9] for further examples of weak dependent sequences.

Density estimation in the case of weak dependence

We are now going to extend Theorem 4.3.

$Y_n(x)$ is defined as in Section 4, that is

$Y_n(x) := \sqrt{nb_n}\left(\hat{f}(x) - \mathbf{E}\hat{f}(x)\right)$, where $\hat{f}_n(x) = \frac{1}{nb_n}\sum_{k=1}^{n} K\left(\frac{x-\xi_k}{b_n}\right)$, and K is supposed to be Lipschitz.

Theorem A.1. *Assume that the previous $s-$weak dependence (resp. $a-$) condition holds for the stationary real valued sequence $(\xi_n)_{n\in\mathbb{N}}$ with for some positive $a < \frac{1}{3}$ (resp. $a < \frac{1}{4}$) $\sum_{p=1}^{\infty} \theta_p^a < \infty$, then the finite dimensional marginals $(\overline{Y_n}(x_1), \ldots, \overline{Y_n}(x_l))$, of the process $\overline{Y_n}(x) \equiv Y_n(x)/\sqrt{f(x)\int_{-\infty}^{\infty} u^2(t)\mathrm{d}t}$ converge in distribution to an $\mathcal{N}(0, I_l)$ random variable if we assume moreover that $f(x_1) \neq 0, \ldots, f(x_l) \neq 0$, that ξ_0's marginal admits a continuous marginal density f and the marginal densities $f_k(x, y)$ of the bivariate random variables (ξ_0, ξ_k) exist for any $k > 0$ and satisfy $\sup\limits_{k>0}\ \sup\limits_{(x,y)\in\mathbb{R}^2} f_k(x, y) < \infty$.*

Remarks.

- Here we need the existence of marginal densities $f_k(x, y)$ of the bivariate random variables (ξ_0, ξ_k). It is a classical assumption in that frame. We recall that in the case of dynamical systems, such densities are singular. The strong estimate

$$\mathrm{Cov}\left(K\left(\frac{x-\xi_j}{b_n}\right),\ K\left(\frac{x-\xi_i}{b_n}\right)\right) \le C\ b_n^2$$

is standard under this condition while in the dynamical case, we just use that

$$\mathrm{Cov}\left(K\left(\frac{x-T^jX_0}{b_n}\right),\ K\left(\frac{x-T^iX_0}{b_n}\right)\right) \le C\ b_n\ \varepsilon(n, j-i)$$

for a sequence $\varepsilon(n, k) \xrightarrow[n\to\infty]{} 0$ for any k. We also replace the summable decay of correlations (see (1.5)) of dynamical systems in the class $\mathcal{T}$ by a weak dependence condition.

Furthermore, in the case of stationary dynamical systems, we do not have any reason to suppose that

$$\sup_{1\le k\le n} \frac{1}{b_n}\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right),\ K\left(\frac{x-X_k}{b_n}\right)\right)$$

tends to 0 as n tends to infinity, so using our estimates it appears to be hopeless to consider the Lipschitz or an Hölder norms as in the case of weak dependence instead of the norm $|| \bullet ||_{\mathcal{BV}}$ without adding assumptions on the sequence $(b_n)_{n\in\mathbb{N}}$.

- The conditions hold respectively if $\theta_r = \mathcal{O}(r^{-a})$ for some $a > 3$ (resp. $a > 4$).
- This result improves on a previous result in Doukhan and Louhichi [14], *e.g.* under association we need $\mathrm{Cov}(\xi_0, \xi_r) = \mathcal{O}(r^{-a})$ for $a > 5$ while the previous result was obtained assuming $a > 12$ and for causal shifts it was needed that $\theta_r = \mathcal{O}(r^{-a})$ for some $a > \max\{9, \frac{3}{2}(1+\delta^{-1})\}$ if $b_n \sim n^{-\delta}$.
- For strongly mixing sequences, the condition $\alpha_n = \mathcal{O}(n^{-a})$ for $a > 1$ ensures this CLT as proved by Robinson [24] (and also Ango–Nze and Doukhan [2]); this asumption is of a different nature, *e.g.* linear processes satisfy mixing conditions (under additional regularity conditions, see Doukhan ([11], Chap. 2.3). The decay rate of the coefficients are there more restrictive.

The proof of Theorem A.1 is a variation on the proof of Theorem 4.3. We refer to Coulon–Prieur and Doukhan [9], where the proof is written under stronger assumptions (in terms of dimension of the "future"). The same techniques provide some general limit theorem for triangular arrays under weak dependence (see [9]).

The author is grateful to the anonymous referees for constructive comments.

References

[1] A. Amroun, *Systèmes dynamiques perturbés. Sur une classe de fonctions zéta dynamiques*, Thèse de Doctorat de l'Université Paris 6, Spécialité Mathématique (1995).

[2] P. Ango Nze and P. Doukhan, Non-parametric Minimax estimation in a weakly dependent framework I: Quadratic properties. *Math. Methods Statist.* **5-4** (1996) 404-423.

[3] V. Baladi, M. Benedicks and V. Maume–Deschamps, Almost sure rates of mixing for i.i.d. unimodal maps. *Ann. E.N.S.* (to appear).

[4] A.D. Barbour, R.M. Gerrard and G. Reinert, Iterates of expanding maps. *Probab. Theory Related Fields* **116** (2000) 151-180.

[5] D. Bosq and D. Guégan, Nonparametric estimation of the chaotic function and the invariant measure of a dynamical system. *Statist. Probab. Lett.* **25** (1995) 201-212.

[6] D. Bosq and J.P. Lecoutre, *Théorie de l'estimation fonctionnelle.* Collection "Économie et statistiques avancées". Série : École Nationale de la Statistique et de l'Administration Économique et Centre d'Études des Programmes Economiques". Economica (1987).

[7] A. Broise, F. Dal'bo and M. Peigné, Études spectrales d'opérateurs de transfert et applications. *Astérisque* **238** (1996) Société Math. de France.

[8] P. Collet, Some ergodic properties of maps of the interval, in *dynamical systems*, edited by R. Bamon, J.M. Gambaudo and S. Martinez. Hermann, Paris (1996).

[9] C. Coulon–Prieur and P. Doukhan, A triangular central limit Theorem under a new weak dependence condition. *Statist. Probab. Lett.* **47** (2000) 61-68.

[10] W. De Melo and S. Van Strien, *One-Dimensional Dynamics.* Springer-Verlag (1993).

[11] P. Doukhan, *Mixing: Properties and Examples.* Springer Verlag, *Lecture Notes in Statist.* **85** (1994).

[12] P. Doukhan, *Models, Inequalities and Limit Theorems for Stationary Sequences*, edited by P. Doukhan, G. Oppenheim and M. Taqqu. Birkhaüser (to appear).

[13] P. Doukhan and S. Louhichi, A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** (1999) 313-342.

[14] P. Doukhan and S. Louhichi, Functional estimation of a density under a new weak dependence condition. *Scand. J. Statist.* **28** (2001) 325-342.

[15] A. Lasota and M. Mackey, *Probabilistic properties of deterministic systems.* Cambridge University Press (1985).

[16] A. Lasota and J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186** (1973) 481-488.

[17] C. Liverani, Decay of correlations for piecewise expanding maps. *J. Statist. Phys.* **78** (1995) 1111-1129.

[18] C. Liverani, Central limit Theorem for deterministic systems, in *Proc. of the international Congress on Dynamical Systems*, Montevideo 95. Pittman, *Res. Notes Math.* (1997).

[19] J. Maës, *Statistique non paramétrique des processus dynamiques réels en temps discret.* Thèse de l'Université Paris 6 (1999).

[20] D. Pollard, *Convergence of Stochastic Processes.* Springer Verlag, *Springer Ser. Statist.* (1984).

[21] R. Prakasa, *Nonparametric functional estimation.* Academic Press, New York (1983).

[22] E. Rio, About the Lindeberg method for strongly mixing sequences. *ESAIM: PS* **1** (1995) 35-61. `www.emath.fr/ps`

[23] E. Rio, Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes. *Probab. Theory Related Fields* **104** (1996) 255-282.
[24] P.M. Robinson, Non parametric estimators for time series. *J. Time Ser. Anal.* **4-3** (1983) 185-207.
[25] M. Rosenblatt, Stochastic curve estimation, in *NSF-CBMS Regional Conference Series in Probability and Statistics*, Vol. 3 (1991).
[26] W. Rudin, *Real and complex analysis.* McGraw-Hill Series in Higher Mathematics, Second Edition (1974).
[27] M. Viana, *Stochastic dynamics of deterministic systems, Instituto de Matematica Pura e Aplicada.* IMPA, Vol. 21 (1997).

Quelques remarques :

Nous donnons ici quelques précisions sur l'article
"*Density Estimation For One-Dimensional Dynamical Systems*"
présenté aux pages 83-108 de cette thèse.

– A la remarque 3.3. de la page 88, on peut remplacer $\varphi(X_0)$ par $\varphi(T^j X_0)$.
Posons $K_n(t) = K\left(\dfrac{x-t}{b_n}\right) - \mathbf{E}\, K\left(\dfrac{x-X_0}{b_n}\right)$. La remarque 3.3. s'écrit alors :

$$\forall\, 0 \leq i < j \leq n-1\,,\ \forall\, \varphi \text{ bornée,}$$
$$\left|\mathrm{Cov}\left(\varphi(T^j X_0) K_n(X_j), K_n(X_i)\right)\right| \leq \|\varphi\|_\infty\, b_n\, \varepsilon(n, j-i)\,, \tag{1}$$

avec $\varepsilon(n,j-i) \underset{n\infty}{\longrightarrow} 0$.
Notons que $\varepsilon(n,j-i)$ ne dépend pas de φ. La suite $\varepsilon(n,j-i)$ dépend uniquement de K, T, x, n et $j-i$.

– Précisons maintenant comment on applique la variante de la remarque 3.3. (inégalité (1) ci-dessus) pour obtenir les inégalités (4.11) et (4.12) de la page 94.
Grâce à un développement de Taylor, on commence par écrire :

$$h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}) = h''_{k,n}(S_{i-1,n} + \rho_{i,n} Z_{n,i})\, Z_{n,i}\,,$$

où $\rho_{i,n}$ est un réel strictement compris entre 0 et 1. On a alors :

$$\mathrm{Cov}\left(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}), Z_{n,k}\right) = \frac{\mathrm{Cov}\left(\varphi(T^{n-i-1}X_0)\, K_n(X_{n-i-1}), K_n(X_{n-k-1})\right)}{nb_n}\,,$$

avec $\varphi(t) = h''_{k,n}\left(\dfrac{K_n(T^i t) + \cdots + K_n(Tt) + \rho_{i,n} K_n(t)}{\sqrt{nb_n}}\right)$.

Comme $\|\varphi\|_\infty \leq \|h''_{k,n}\|_\infty \leq C_2$, on obtient alors avec (1) :

$$\begin{aligned} \left|\mathrm{Cov}\left(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}), Z_{n,k}\right)\right| &\leq \frac{1}{nb_n}\|\varphi\|_\infty\, b_n\, \varepsilon(n,k-i) \\ &\leq \frac{C_2}{n}\, \varepsilon(n,k-i). \end{aligned} \tag{2}$$

A l'aide des inégalités (4.10) de la page 93 et (2) ci-dessus nous déduisons l'inégalité (4.11) de la page 94.

Notons maintenant que

$$\left|\mathbf{E}\, h''_{k,n}(S_{k-1,n})\mathbf{E}\, Z_{n,i} Z_{n,k}\right| \;\leq\; \frac{C_2}{nb_n}\left|\mathrm{Cov}\left(K_n(X_{n-i-1}), K_n(X_{n-k-1})\right)\right|.$$

En appliquant alors (1) avec la fonction φ constante égale à 1, nous en déduisons :

$$\left|\mathbf{E}\, h''_{k,n}(S_{k-1,n})\mathbf{E}\, Z_{n,i}Z_{n,k}\right| \quad \leq \quad \frac{C_2}{n}\,\varepsilon(n,k-i). \tag{3}$$

Des inégalités (1.5) de la page 84 et (3) ci-dessus, nous déduisons l'inégalité (4.12) de la page 94.

Chapitre 5

Simulations numériques

Dans ce chapitre, nous illustrons certains résultats du chapitre 4 par des simulations numériques. Les programmes sont écrits en C et les figures tracées à l'aide de MATLAB.

Simulations numériques

1 Introduction

Dans ce chapitre, nous réalisons des simulations numériques relatives aux résultats du chapitre 4. Rappelons que l'objet du chapitre 4 est l'estimation de la densité invariante d'applications T d'un intervalle $I \subset \mathbb{R}$, fermé, dans lui-même. Les applications T considérées sont dans une certaine classe $\mathcal{T}$. Elles ont une unique mesure invariante, μ_0, absolument continue par rapport à la mesure de Lebesgue. Nous gardons les notations du chapitre 4. La densité de μ_0 par rapport à la mesure de Lebesgue est donc notée f. Nous estimons f à l'aide d'un estimateur à noyau construit de la façon suivante :

$$X_0^{'} \sim p(t)dt \text{ , où } p \text{ est une fonction densité à variation bornée,}$$

$$X_n^{'} = T^n X_0^{'} \text{ , } \forall\, n \geq 1 \text{ ,}$$

$$\hat{p}_{n,b_n}(x) = \frac{1}{nb_n} \sum_{k=0}^{n-1} K\left(\frac{x - X_k^{'}}{b_n}\right),$$

avec une suite $(b_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ et un noyau $K : \mathbb{R} \to \mathbb{R}$ à support compact $D \subset I$ tels que :

$$b_n \underset{n\infty}{\longrightarrow} 0,\ \int_D K(t)d\lambda_I(t) = 1,\ 0 < \int_D K^2(t)d\lambda_I(t) < \infty.$$

Pour un tel estimateur, nous obtenons en particulier le résultat de convergence en moyenne quadratique suivant (lemme 6.1. page 99) :

Lemme 1 (Convergence en moyenne quadratique) *Soient $T \in \mathcal{T}$ et μ_0 la mesure $T-$invariante absolument continue par rapport à la mesure de Lebesgue. Soit S le support de μ_0. Pour presque tout $x \in S$, nous avons :*

$$\text{Var }(\hat{p}_{n,b_n}(x)) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)ds + o(1)\right).$$

Nous renvoyons au chapitre 4 pour plus de détails sur ces résultats théoriques.

Pour vérifier ce résultat numériquement, nous avons écrit un programme en langage C. Dans un premier temps, nous calculons, pour différentes valeurs de x, et pour n et m fixés,

$$P_m^n(x) := \left(\hat{p}_n^{(1)}(x), \ldots, \hat{p}_n^{(m)}(x)\right) \cdot$$

Ici les $\hat{p}_n^{(i)}(x)$, $1 \leq i \leq m$ sont m copies indépendantes de $\hat{p}_n(x)$.

Nous calculons ensuite pour chaque x la variance empirique du vecteur $P_m^n(x)$:

$$\text{Var}_{emp}\,(P(x)) = \frac{1}{m}\sum_{i=1}^{m}\left\{\hat{p}_n^{(i)}(x) - \frac{1}{m}\sum_{j=1}^{m}\hat{p}_n^{(j)}(x)\right\}^2.$$

Par la loi forte des grands nombres, nous savons qu'à n fixé, $\text{Var}_{emp}\,(P_m^n(x))$ converge presque sûrement vers $\text{Var}\,(\hat{p}_n(x))$ lorsque m tend vers l'infini. Plus précisément, $\forall\, \alpha \in [0{,}1/2[$,

$m^\alpha\ \{\text{Var}_{emp}\,(P_m^n(x)) - \text{Var}\,(\hat{p}_n(x))\}$ tend vers 0 lorsque m tend vers l'infini (voir Baum & Katz [6]).

Enfin, nous interpolons la courbe $(nb_n)\,\text{Var}_{emp}\,(P_m^n(x))$ en fonction de x. Nous vérifions sur les figures que, pour de grandes valeurs de m et de n, cette courbe se rapproche de la courbe théorique $y = f(x) \int K^2(s)\, ds$.

Nous avons d'abord écrit ce programme à l'aide de MATLAB. Pour les grandes valeurs de m ou de n, le temps de calcul était alors très long, de l'ordre de plusieurs heures. En effet, le programme contient des boucles emboîtées qui conviennent mal à ce logiciel. Nous avons donc réécrit le programme en C pour pouvoir envisager de grandes valeurs de m et de n. La vitesse de convergence de $\text{Var}_{emp}\,(P_m^n(x))$ vers $\text{Var}\,(\hat{p}_n(x))$ étant assez lente ($m^{-\alpha}$, $\alpha < 1/2$) nous savons déjà que le nombre m doit être choisi plutôt grand (1000, 5000, 10000, ...).

2 Structure du programme

Nous avons choisi de faire les simulations pour la fonction unimodale définie sur [0,1] par : $T(x) = 4x(1-x)$. La densité invariante est alors connue et s'écrit : $f(x) = \dfrac{1}{\pi\sqrt{x(1-x)}}$ (voir par exemple Collet [14]).

Nous avons travaillé avec le noyau suivant : $K_1(x) = \mathbf{1}_{[-1/2,1/2]}(x)$.

Pour la loi initiale nous avons pris l'une des deux lois suivantes :

- la loi uniforme sur [0,1],
- la loi de densité $p_2(x) = \dfrac{2\,e^{-2x}}{1-e^{-2}}\,\mathbf{1}_{[0,1]}(x)$.

Ecrivons maintenant la structure du programme C pour K_1 et pour la loi uniforme. Le programme précis peut être obtenu sur simple demande à l'auteur.

STRUCTURE :

```
/* Calcul du noyau K1

Argument :
x, réel

Renvoi :
```

$$K_1(x) = \begin{cases} 1 \text{ si } |x| \leq 1/2 \\ 0 \text{ sinon} \end{cases}$$

/* Calcul de la fonction unimodale T

Argument :
x, réel dans $[0,1]$

Renvoi :
$T(x) = 4x(1-x)$

/* Calcul de la variance empirique d'un vecteur

Arguments :
m, taille du vecteur et U, vecteur de $\mathbb{R}^m$

Renvoi :
$V(m,U) = \frac{1}{m}\sum_{i=1}^{m}\left\{U[i] - \frac{1}{m}\sum_{j=1}^{m}U[j]\right\}^2$

/* Calcul de nb_n multiplié par la variance empirique d'un vecteur constitué par m copies indépendantes de $\hat{p}_n(x)$

Arguments :
x, réel dans $[0,1]$
n, nombre d'itérations de la fonction T
m, taille de l'échantillon $\left(\hat{p}_n^{(1)}(x), \ldots, \hat{p}_n^{(m)}(x)\right)$
a, réel dans $]0,1[$ ($b_n = n^{-a}$)

Renvoi : $est(x,n,m,a)$
Au départ, U est le vecteur nul de $\mathbb{R}^m$.
Pour j variant de 0 à $m-1$,
- le programme tire au hasard une valeur X dans $[0,1]$ (à l'aide de la fonction RAND bien initialisée) ;
- calcule $K_1\left(\dfrac{x-X}{b_n}\right)$ puis le met dans $U[j]$;
- Puis de $i=1$ à $n-1$,
    - met $T(X)$ dans X;
    - calcule $Y = K_1\left(\dfrac{x-X}{b_n}\right)$;

- met $Y+U[j]$ dans $U[j]$;

A ce stade du programme, le vecteur U contient un $m-$échantillon $(nb_n)\left(\hat{p}_n^{(1)}(x),\ldots,\hat{p}_n^{(m)}(x)\right)$.

Le programme sort alors $\dfrac{1}{nb_n}V(m,U)$, c'est-à-dire $(nb_n)\mathrm{Var}_{emp}\left(P_m^n(x)\right)$.

/* Corps du programme

Arguments :

s, entier ≥ 1 nombre de points x pour lesquels on fait les calculs

n, nombre d'itérations de la fonction T

m, taille de l'échantillon $\left(\hat{p}_n^{(1)}(x),\ldots,\hat{p}_n^{(m)}(x)\right)$

a, réel dans $]0{,}1[$ ($b_n=n^{-a}$)

Renvoi : $est(s,n,m,a)$

Pour s valeurs de x réparties sur $[0{,}1]$, le programme calcule $est(x,n,m,a)$.

Pour faire les graphes, nous écrivons un programme, à l'aide de MATLAB, qui utilise en données les valeurs calculées par le programme précédent écrit en C. Ce programme nous permet d'interpoler la courbe $est(x,n,m,a)$ en fonction de x.

Nous lançons ce programme pour différentes valeurs de n, m, et a. Etudions maintenant le résultat de ces simulations.

3 Simulations

Commençons par donner le graphe de la densité invariante par rapport à T : f.

Rappelons que $f(x) = \dfrac{1}{\pi\sqrt{x(1-x)}}$ et que f atteint son minimum $2/\pi$ en $x = 1/2$.

La ligne tracée en pointillés à la figure 1 est la droite $y = 2/\pi$.

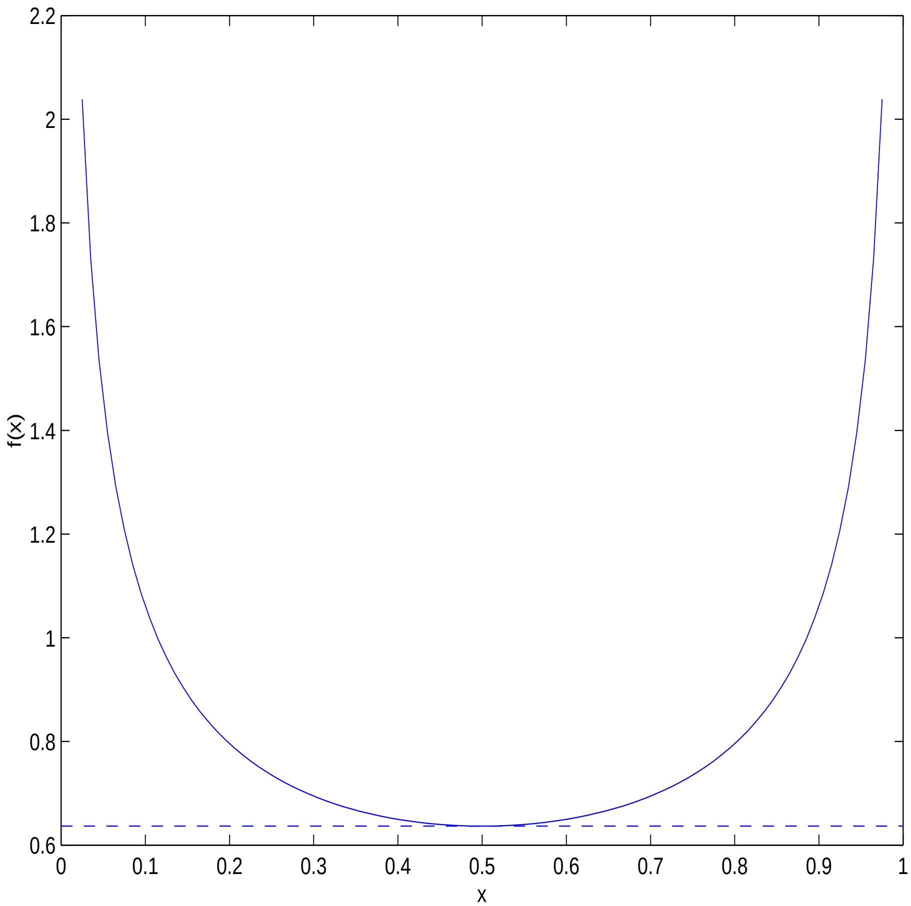


FIG. 1 – *Densité invariante f*

3.1 Noyau K_1 et loi initiale uniforme

Lançons maintenant la simulation pour $m = 5000$ et $n = 500$.

Sur la figure 2 ci-dessous, la courbe empirique est tracée en trait plein et la courbe $y = f(x)$ est tracée en pointillés.

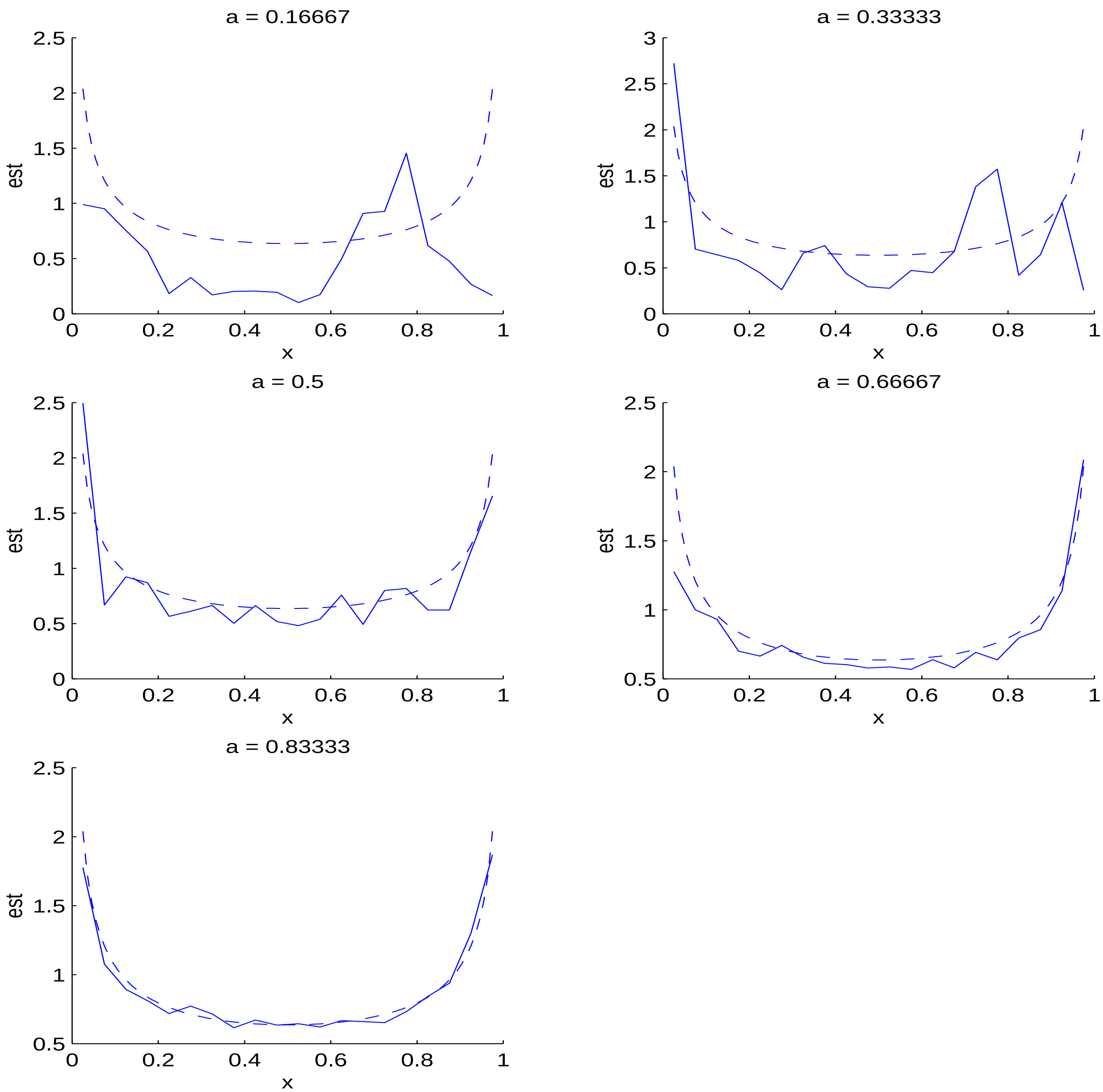


FIG. 2 – *Convergence quadratique,* $m = 5000$, $n = 500$, $s = 20$

Sur la figure 2 nous voyons que pour des suites b_n convergeant assez vite vers 0 ($a = -\ln(b_n)/\ln(n) \geq 0.5$), la courbe empirique se rapproche de la courbe théorique (figure 1). Ici la courbe théorique est celle de f car $\int K_1^2(s)\ ds = 1$. Lors de la preuve du lemme 6.1. au chapitre 4, nous avons vu qu'effectivement la fenêtre b_n jouait un rôle important. En effet, nous montrons au chapitre 4 que le point clé dans la preuve du lemme 6.1. du chapitre 4 (lemme 1 ci-dessus) est la vitesse à laquelle l'expression $\left|\mathrm{Cov}\left(K_1\left(\frac{x-X_0'}{b_n}\right), K_1\left(\frac{x-X_k'}{b_n}\right)\right)\right|$ converge vers 0. Or

$$\left|\mathrm{Cov}\left(K_1\left(\frac{x-X_0'}{b_n}\right), K_1\left(\frac{x-X_k'}{b_n}\right)\right)\right| \leq C\ b_n\ (\varepsilon(n,k) \wedge d_k)\,, \tag{1}$$

où pour k fixé, la suite $\varepsilon(n,k)$ converge vers 0 lorsque n tend vers l'infini. Rappelons que la suite $(d_n)_{n\in\mathbb{N}}$, définie par la relation (1.5) à la page 84 du chapitre 4, contrôle la décroissance des corrélations du système dynamique généré par T.

En s'intéressant aux détails de la démonstration du lemme 1 (lemme 6.1. page 99 du chapitre 4), nous voyons qu'à k fixé, la vitesse de décroissance de la suite $(\varepsilon(n,k))_{n\in\mathbb{N}}$ est limitée par la vitesse de convergence de la suite $b_n = n^{-a}$ vers 0.

Pour la fonction unimodale T choisie ici, la décroissance de d_n est exponentielle (voir [65]). Comme la suite b_n est choisie de la forme n^{-a}, c'est donc bien la valeur de a qui va limiter la vitesse de convergence vers 0 de la borne supérieure de l'expression (1) ci-dessus. Ceci explique que pour des petites valeurs de a, il faut prendre des valeurs de n beaucoup plus grande pour observer la convergence. C'est ce que nous remarquons sur la figure 4.

Notons enfin que pour $1/2 \leq a < 1$, nb_n^2 ne tend pas vers l'infini, et pourtant les courbes convergent. Nous vérifions ainsi numériquement que la convergence en moyenne quadratique de l'estimateur $\hat{p}_{n,b_n}(x)$ requiert simplement l'hypothèse “$nb_n \underset{n\infty}{\longrightarrow} +\infty$”, et non “$nb_n^2 \underset{n\infty}{\longrightarrow} +\infty$”.

Sur la figure 3, nous présentons un graphe en dimension trois représentant l'évolution des courbes empiriques $y = est(x,n,m,a)$ en fonction de la valeur de a pour $m = 5000$ et $n = 500$.

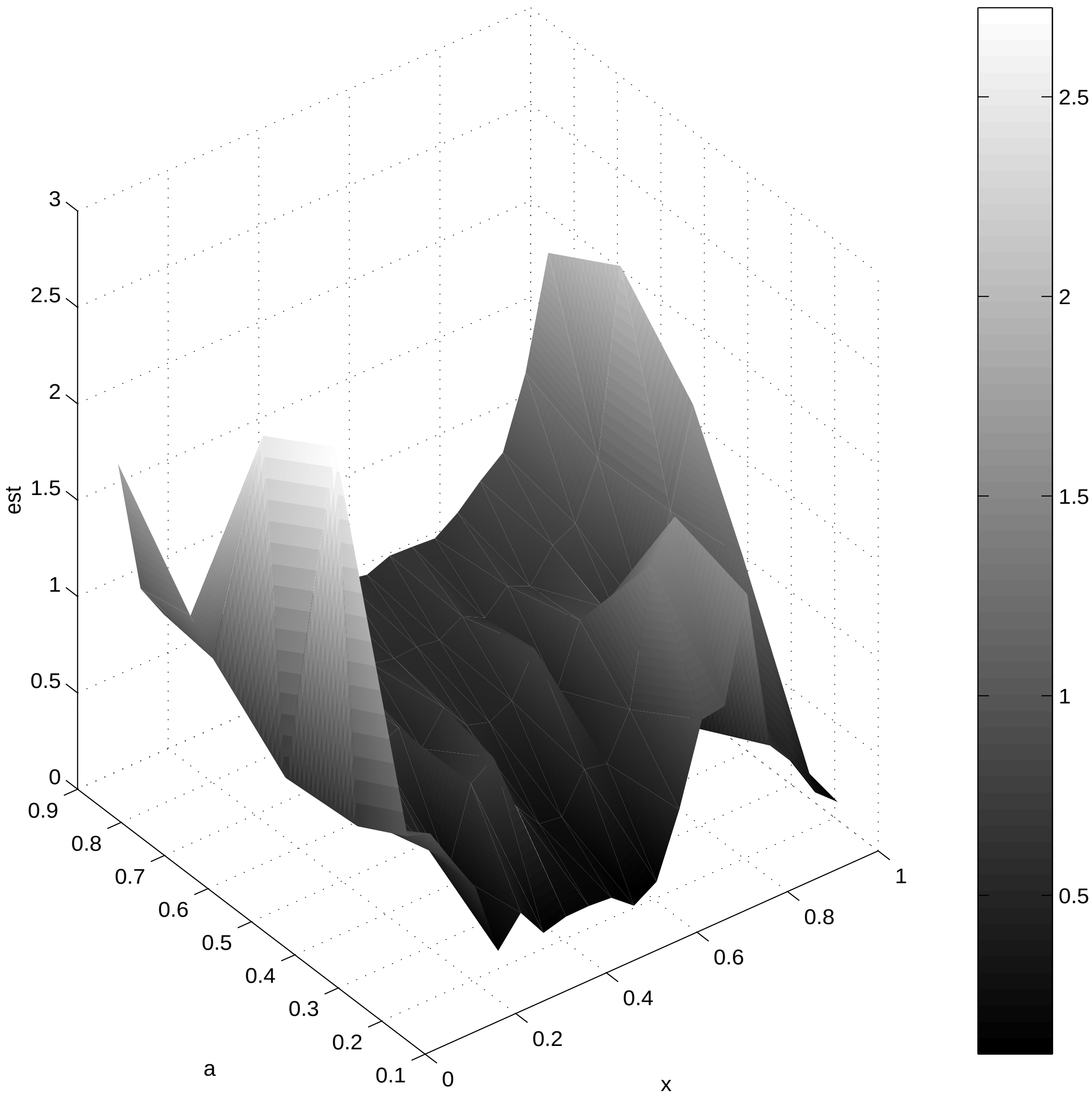


FIG. 3 – *Evolution de la convergence en fonction de la valeur* $a = -\ln(b_n)/\ln(n)$

Ce graphe confirme l'évolution des courbes de la figure 2 en fonction de la valeur de $a = -\ln(b_n)/\ln(n)$.

Sur la figure 4, nous prenons $a = 0.33$, $m = 5000$ et $n = 10000$. Nous observons alors effectivement que la courbe empirique (en trait plein) est plus proche de la coube théorique (en pointillés) que pour $a = 0.33$, $m = 5000$, $n = 500$ (figure 2). C'est donc bien qu'en prenant une valeur différente pour a, on change la vitesse de convergence.

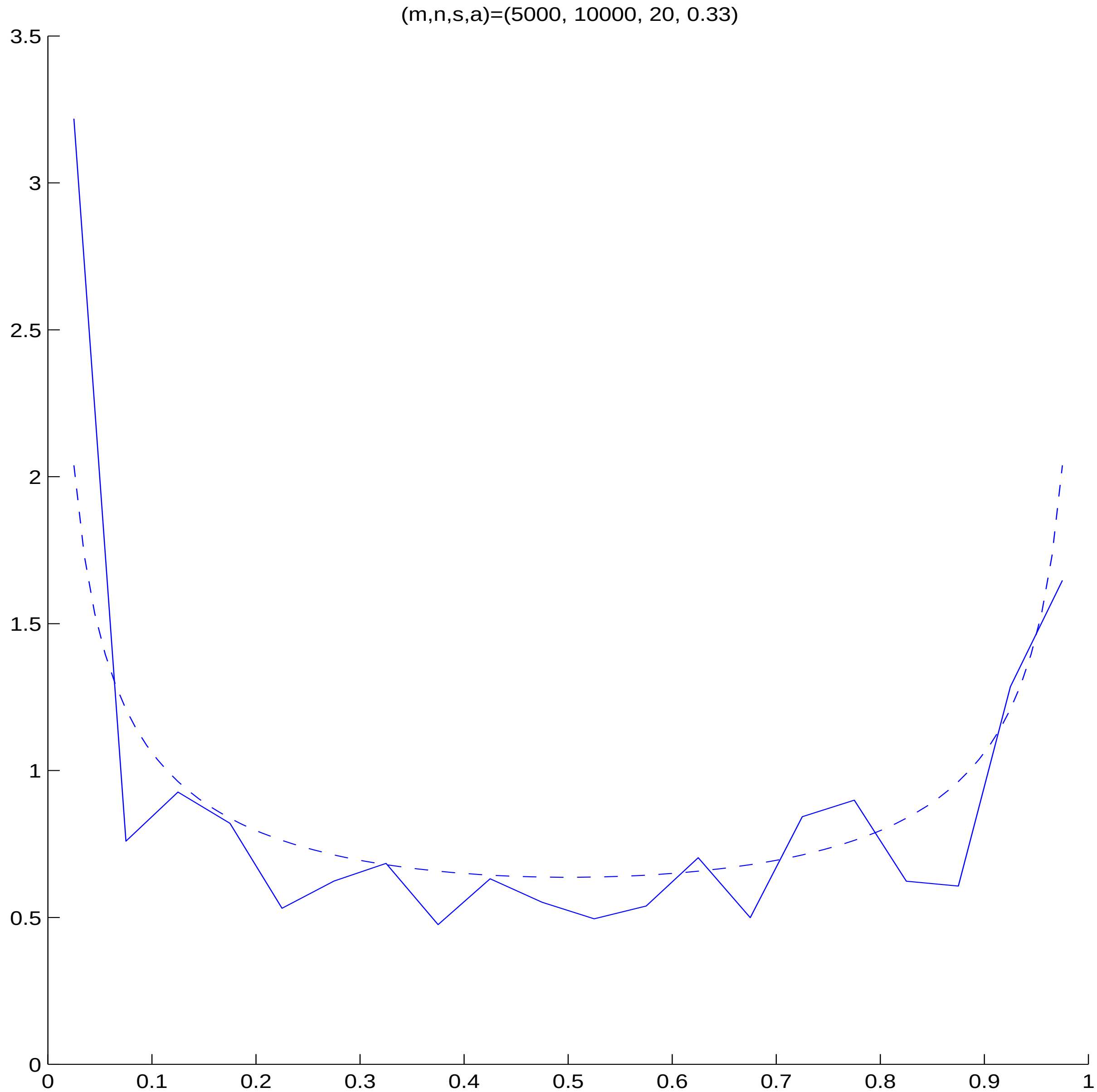


FIG. 4 – *Convergence quadratique*, $m = 5000$, $n = 10000$, $s = 20$, $a = 0.33$

3.2 Noyau K_1 et loi initiale p_2

Enfin, dans ce paragraphe, nous prenons le noyau K_1 mais la densité initiale $p_2(x)$. Là encore, nous observons des comportements similaires. Commençons par lancer la simulation pour $m = 5000$ et $n = 500$.

Sur la figure 5 ci-dessous, la courbe empirique est tracée en trait plein et la courbe $y = f(x)$ est tracée en pointillés.

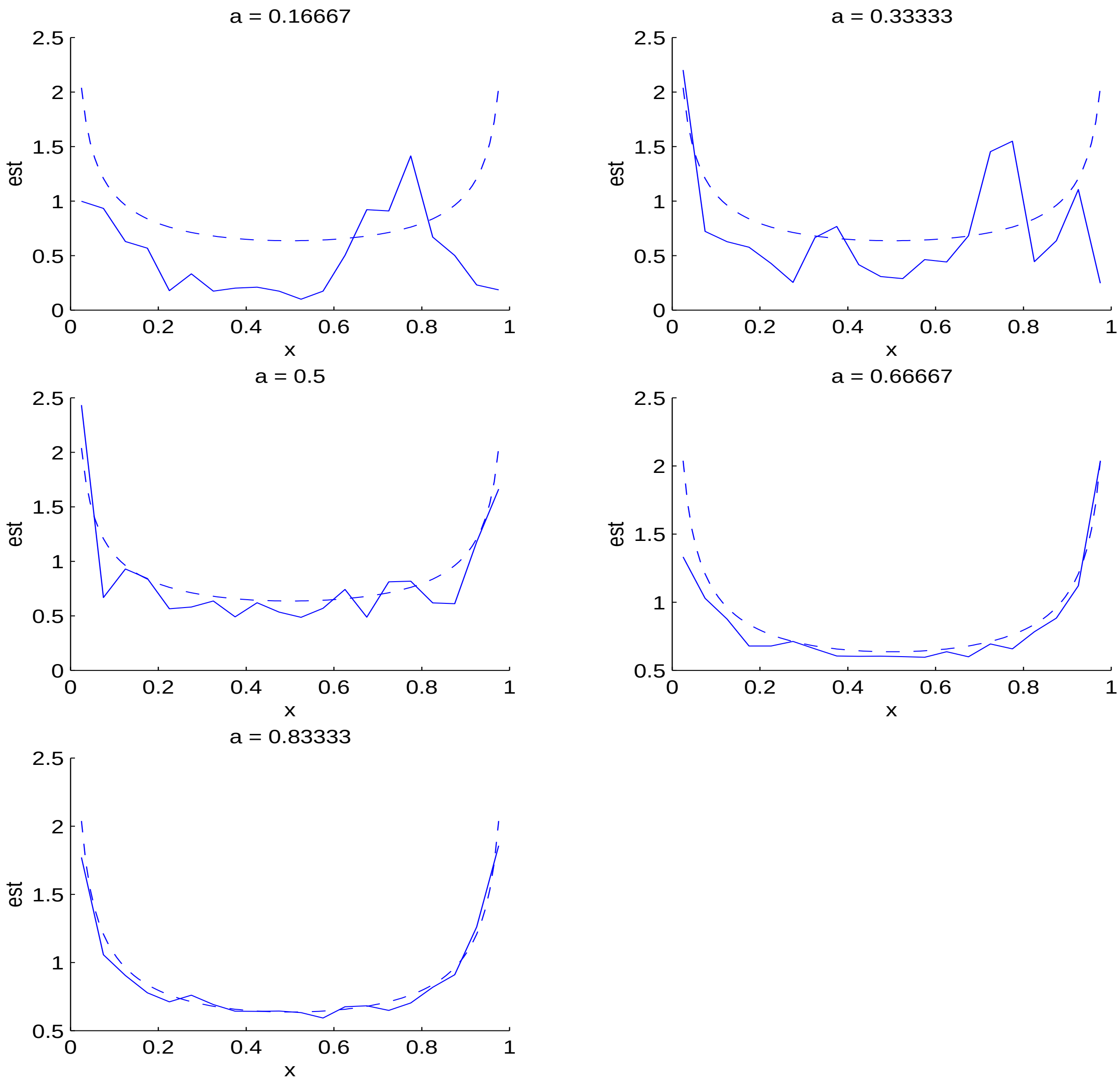


FIG. 5 – *Convergence quadratique,* $m = 5000$, $n = 500$, $s = 20$

Sur la figure 6, nous présentons un graphe en dimension trois représentant la variation des courbes empiriques $y = est(x,n,m,a)$ en fonction de la valeur de a pour $m = 5000$ et $n = 500$.

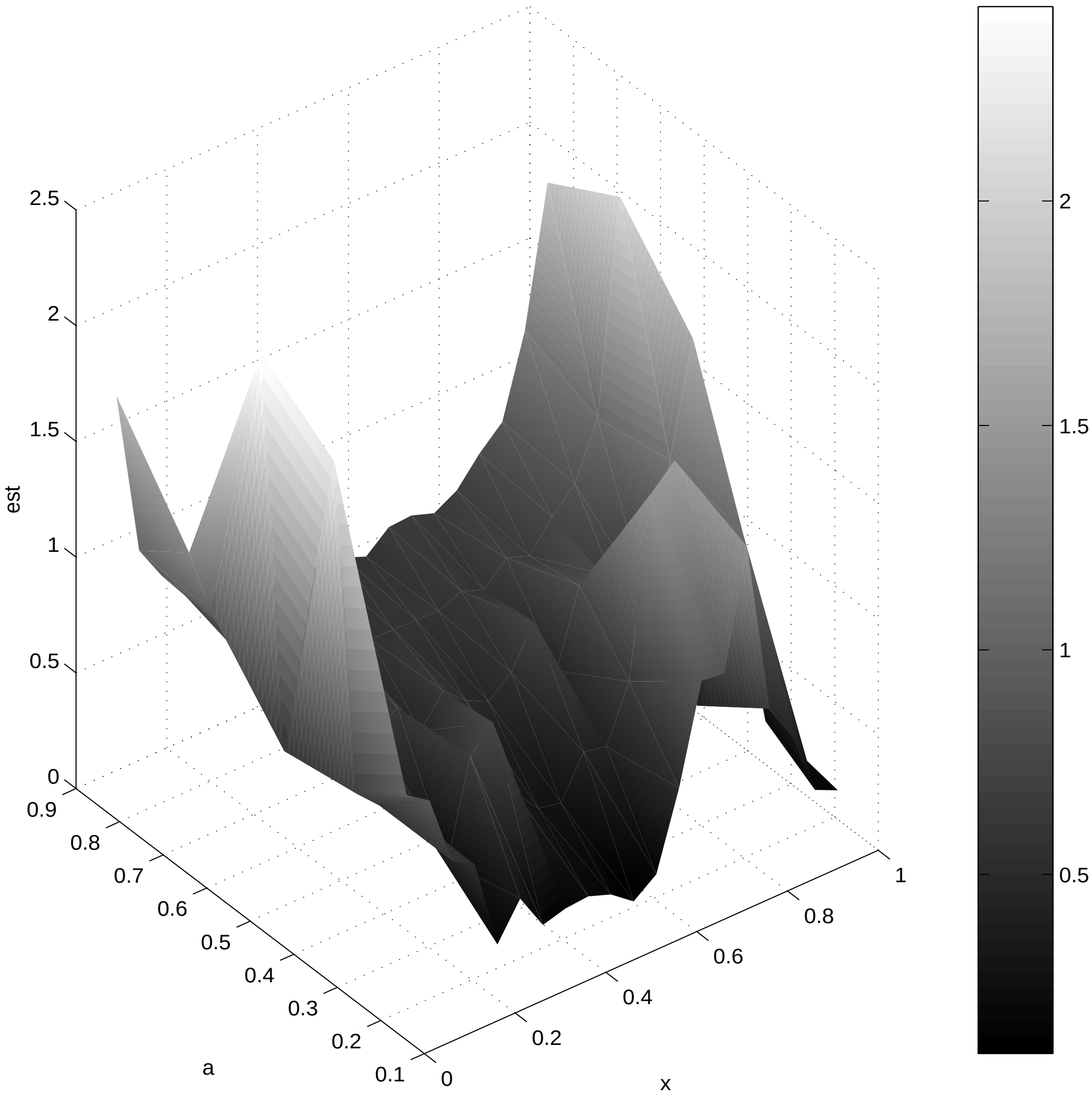


FIG. 6 – *Evolution de la convergence en fonction de la valeur* $a = -\ln(b_n)/\ln(n)$

Sur la figure 7, nous prenons $a = 0.33$, $m = 5000$ et $n = 10000$. Nous observons encore que la courbe empirique (en trait plein) est plus proche de la coube théorique (en pointillés) que pour $a = 0.33$, $m = 5000$ et $n = 500$ (figure 5).

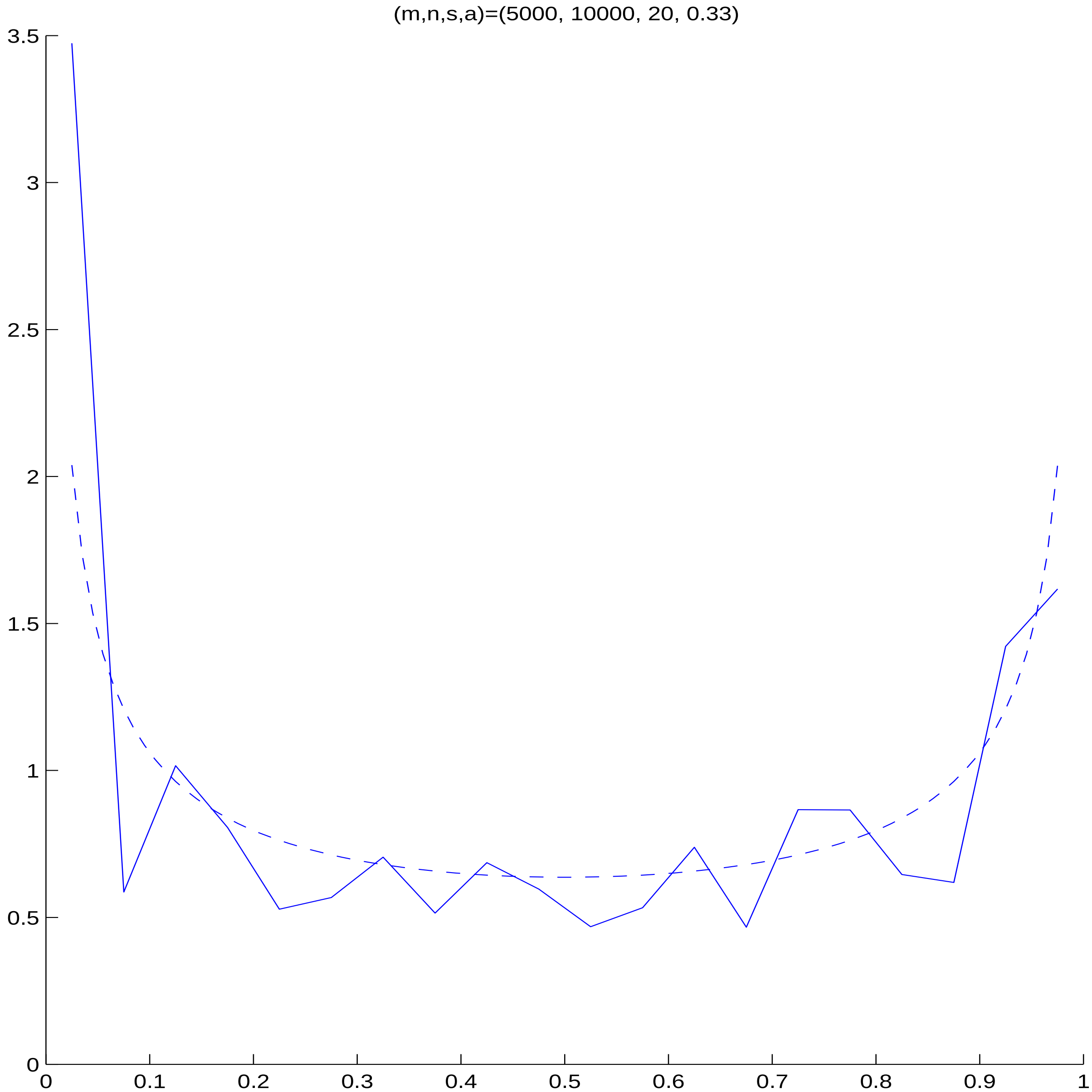


FIG. 7 – *Convergence quadratique,* $m = 5000$*,* $n = 10000$*,* $s = 20$*,* $a = 0.33$

4 Conclusion

Dans ce chapitre, nous avons choisi de ne présenter que les simulations relatives au résultat (lemme 1) de convergence en moyenne quadratique de l'estimateur de la densité étudié au chapitre 4.

Nous aurions également pu présenter des simulations concernant les résultats de convergence en loi. Nous aurions alors tracé la fonction de répartirion empirique de

$$\sqrt{nb_n}\,\{\hat{p}_n(x) - \mathbf{E}\hat{p}_n(x)\}$$

et vérifié que la courbe obtenue pour x fixé se rapprochait bien d'une répartition gaussienne.

Nous avons préféré mettre en valeur l'ordre de grandeur de la vitesse de convergence en moyenne quadratique de l'estimateur $\hat{p}_n(x)$, en $\mathcal{O}\left(\frac{1}{nb_n}\right)$, car c'est lui qui nous a permis au chapitre 4 d'utiliser une variation de la méthode de Lindeberg-Rio [54] pour montrer le théorème limite centrale (théorème 4.3. page 91) pour $\hat{p}_n(x)$.