Limit Theorems for Multiplicative Processes

Quansheng Liu * Emmanuel Rio † Alain Rouault ‡

April 30, 2002

Abstract

Let W be a non-negative random variable with $EW = 1$, and let $\{W_{\mathbf{i}}\}$ be a family of independent copies of W, indexed by all the finite sequences $\mathbf{i} = i_1...i_n$ of positive integers. For fixed r and n the random multiplicative measure μ_r^n has, on each r-adic interval $A^r_{i_1...i_n}$ at n-th level, the density $W_{i_1}...W_{i_1...i_n}$ with respect to the Lebesgue measure on $[0,1]$. If $EW \log W < \log r$, the sequence $\{\mu_r^n\}_n$ converges a.s. weakly to the Mandelbrot measure μ_r^∞. For each fixed $1 \leq n \leq \infty$, we study asymptotic properties for the sequence of random measures $\{\mu_r^n\}_r$ as $r \to \infty$. We prove uniform laws of large numbers, functional central limit theorems, a functional law of iterated logarithm, and large deviation principles. The function-indexed processes $\{\mu_r^n(f), f \in \mathcal{G}\}$ is a natural extension to a tree-indexed process at n-th level of the usual smoothed partial-sum process corresponding to $n = 1$. The results extend the classical ones for $\{\mu_r^1\}_r$, and the recent ones for the masses of $\{\mu_r^\infty\}_r$ established in [26].

A.M.S. Classification: Primary 60G57; Secondary 60F17, 60F10
Key words and phrases: Self-similar cascades, Mandelbrot's martingales, random measures, laws of large numbers, functional central limit theorems, functional law of the iterated logarithm, large deviations.

1 Introduction

Let us recall the basic model of the Mandelbrot cascade [27, 28, 19]. Let $\mathbb{N}^\star = \{1, 2, \cdots\}$ be the set of positive integers, and let

$$\mathbb{U} = \{\emptyset\} \cup \cup_{k=1}^\infty (\mathbb{N}^\star)^k$$

be the set of all finite sequences $i_1...i_n$ containing the null sequence $\emptyset$. Let W be a non-negative random variable with $EW = 1$ and $P(W = 1) \neq 1$, and let $\{W_{\mathbf{i}} : \mathbf{i} \in \mathbb{U}\}$ be independent copies of W; all the random variables are defined on a probability space $(\Omega, \mathcal{F}, P)$. Let λ be the Lebesgue measure on $[0,1]$. Fix $r \geq 2$; for every $n \geq 1$, let μ_r^n be the random measure on $[0,1]$, having on each r-adic interval $A^r_{i_1...i_n} = [\sum_{k=1}^n (i_k - 1) r^{-k}, \sum_{k=1}^n (i_k - 1) r^{-k} + r^{-n})$ the density $W_{i_1}...W_{i_1...i_n}$ with respect to the Lebesgue measure. In other words,

$$\mu_r^n(f) = \int f d\mu_r^n = \sum_{1 \leq i_1, ..., i_n \leq r} W_{i_1}...W_{i_1...i_n} \int_{A^r_{i_1...i_n}} f d\lambda, \tag{1.1}$$

*LMAM, Bâtiment Y. Coppens, Campus de Tohannic, Université de Bretagne-Sud, BP573, 56017 Vannes, France. e-mail: Quansheng.Liu@univ-ubs.fr
†LAMA CNRS UMR8100, Bâtiment Fermat, Université de Versailles, 78035 Versailles, France. e-mail: rio@math.uvsq.fr
‡LAMA CNRS UMR8100, Bâtiment Fermat, Université de Versailles, 78035 Versailles, France. e-mail: rouault@math.uvsq.fr

for each $f \in \mathcal{L}^1([0,1],\lambda)$.

Let $\mathcal{F}_0$ be the trivial σ-algebra, and for $n > 1$, let $\mathcal{F}_{n-1}$ be the σ-algebra generated by $\{W_{i_1}, ..., W_{i_1...i_{n-1}} : 1 \le i_1 \le r, ..., 1 \le i_{n-1} \le r\}$.

For fixed $r \geq 2$, almost surely (a.s.) the sequence of random measures $\{\mu_r^n\}_n$ is weakly convergent, as $n \to \infty$. Let μ_r^∞ be the Borel extension of this weak limit. The random Borel measure μ_r^∞ on $[0,1]$ is called the Mandelbrot measure for multiplicative cascades; it is not a.s. the null measure if and only if $\log r > EW \log W$ (see Kahane and Peyrière [19]). This measure and its extensions have been studied by many authors, see for example ([19], [36] [3], [24]).

Here we are mainly interested in the asymptotic behavior of $(\mu_r^n)_r$ as $r \to \infty$.

For $n = 1$, $\{\mu_r^1\}_r$ is usually called the smoothed partial-sum process. Its weak limit is λ. Many classical results about random walks, such as uniform laws of large numbers, Donsker's and Strassen's theorems, and large deviation principles can be expressed in terms of this process, see for example [2], [4], [5], [16], [17], [20], [35], [40], [12].

For $n = \infty$, it is known from [19] that the Hausdorff dimension of the support of μ_r^∞ is a.s. $1 - (EW \log W / \log r)$. So it is natural to conjecture that in some sense, μ_r^∞ tends to λ.

For $1 \le n < \infty$, the process $(\mu_r^n)_r$ is a natural extension of $(\mu_r^1)_r$ to a tree-indexed process, and may be referred as a "tree-indexed smoothed partial-sum process at n-th level"; proving convergence theorems for such a process is not evident because the summands in the definition of μ_r^n are not independent of each other.

A fundamental recursive relation will make tractable the extensions of known (and unknown) results on $(\mu_r^1)_r$ to $(\mu_r^n)_r$ for fixed n (finite or infinite). Fix $1 \le k \le r$; if the weights $W_{i_1}, ..., W_{i_1...i_n}$ in (1.1) are replaced by $W_{ki_1}, ..., W_{ki_1...i_n}$, the corresponding measures will be denoted by $\mu_r^n \circ \Theta_k$ $(1 \le n < \infty)$, and its weak limit (as $n \to \infty$) by $\mu_r^\infty \circ \Theta_k$; notice that the measures μ_r^n and μ_r^∞ depend on the marked r-ary tree with marks $W_{i_1...i_n}$ associated to each node $i_1...i_n$, while $\mu_r^n \circ \Theta_k$ and $\mu_r^\infty \circ \Theta_k$ depend on its shift at k. Θ_k may be considered as the shift operator to the node k in the space of marked trees. For fixed r and f, the random variables $(\mu_r^\infty \circ \Theta_k)(f)$, $1 \le k \le r$, are independent of each other and have the same law as $\mu_r^\infty(f)$, and as a family they are independent of $(W_1, ..., W_r)$. For $k = 1, ..., r$, let τ_k^r be the operator acting on functions from $[0,1]$ to $\mathbb{R}$, defined by

$$\tau_k^r f(x) = f\Big(\frac{k-1+x}{r}\Big), \qquad x \in [0,1].$$

Since $t \in A_{i_1...i_n}^r$ if and only if $r(t - \frac{i_1-1}{r}) \in A_{i_2...i_n}^r$, we have, for $f \in \mathcal{L}^1([0,1],\lambda)$,

$$\mu_r^n(f) = \int f d\mu_r^n = \sum_{k=1}^r W_k \sum_{1 \le i_2,...,i_n \le r} W_{ki_2}....W_{ki_2...i_n} \int_{A_{i_2...i_n}^r} \frac{1}{r} f(\frac{s+k-1}{r}) ds, \tag{1.2}$$

so that for each $1 \le n < \infty$,

$$\mu_r^n(f) = \frac{1}{r} \sum_{k=1}^r W_k (\mu_r^{n-1} \circ \Theta_k)(\tau_k^r f), \tag{1.3}$$

where by convention $\mu_r^0 \circ \Theta_k = \lambda$. The mass of μ_r^n is $Z_r^n := \mu_r^n(1)$. Taking the limit as $n \to \infty$ in (1.3), we see that a.s. for every $f \in \mathcal{C}([0,1])$,

$$\mu_r^\infty(f) = \frac{1}{r} \sum_{k=1}^r W_k (\mu_r^\infty \circ \Theta_k)(\tau_k^r f)\,. \tag{1.4}$$

This equation and its version for masses $Z_r^\infty := \mu_r^\infty(1)$,

$$Z_r^\infty = \frac{1}{r} \sum_{k=1}^r W_k \left(Z_r^\infty \circ \Theta_k\right), \tag{1.5}$$

have been studied by many authors ([19], [10], [18], [23], [25]). Asymptotic properties of the masses Z_r^∞ as $r \to \infty$, have been studied in [26].

Let us explain briefly the main results of the paper.

In section 2, we study the convergence of $\{\mu_r^n\}_n$ for fixed r as $n \to \infty$. Let $f \in \mathcal{L}^1([0,1],\lambda)$ be fixed. The sequence $\{\mu_r^n(f)\}_n$ is a martingale, and converges a.s. as $n \to \infty$. Let $\mu_r(f)$ be its limit. We compare $\mu_r(f)$ with $\mu_r^\infty(f)$, establish the convergence in L^p $(p \geq 1)$ of $\{\mu_r^n(f)\}_n$ for $f \in \mathcal{L}^p([0,1],\lambda)$, and obtain the rate of convergence in the case where $p > 1$.

In section 3 we prove different laws of large numbers for the random measures $\{\mu_r^n\}_r$, as $r \to \infty$. It turns out that the limit is the Lebesgue measure λ. The main results are given in Subsection 3.2, where we need only the assumption $EW = 1$ for $n < \infty$, and $EW \log W < \infty$ for $n = \infty$. (Actually in the first case, only $EW < \infty$ is essential). In particular we prove that the masses Z_r^∞ converge to 1 a.s., improving Theorem 1.1 of [26]. In what follows, each random measure ν will be considered as a function-indexed process $\{\nu(f), f \in \mathcal{G}\}$ for some class $\mathcal{G}$ of functions. For $n < \infty$ the strong law of large numbers (SLLN) in $\ell^\infty(\mathcal{G})$ is established for the random measures $\{\mu_r^n - \lambda\}_r$ under the usual assumption of relative compactness of $\mathcal{G}$ in $\mathcal{L}^1([0,1],\lambda)$ (cf. [35], p. 231). For $n = \infty$, since we need the recursive relation (1.3), the SLLN for $\{\mu_r^\infty - \lambda\}_r$ holds under some strengthened bracketing condition. We mention that, for classes of sets and random measures constructed with the help of some ergodic transformation (the transformation $x \to rx - [rx]$ for multiplicative cascades), Talagrand's combinatorial criterion (cf. [34]) may fail to hold, as proven in [30].

Section 4 is devoted to functional central limit theorems and section 5 to the functional law of iterated logarithm. In these two sections we need $EW^2 < \infty$, and prove the results for classes $\mathcal{G}$ satisfying a universal entropy condition. Actually this condition is convenient to handle the differences $\{\mu_r^n - \mu_r^{n+1}\}_r$ conditionally on $\mathcal{F}_n$. The functional limit theorems for $\{\mu_r^\infty - \lambda\}_r$ are derived from the case $n < \infty$ under some additional regularity condition on $\mathcal{G}$.

In Section 6, we study large deviation principles (LDP) for $(\mu_r^n)_r$ with $n \leq \infty$ fixed. The LDP for the random measure $(\mu_r^1)_r$ in the weak topology is related to the Mogulskii theorem ([12] p.176 and p. 318) which holds under the assumption

$$E\exp(tW) < \infty \quad \text{for every } t > 0\,. \tag{A_1}$$

We will prove that the LDP holds in the strong (or τ) topology under the assumption (A_1) if $1 \leq n < \infty$, and under the assumption

$$\overline{w} := \text{ess sup } W < \infty\,, \tag{A_∞}$$

if $n = \infty$. (Notice that (A_∞) was the assumption used for the corresponding result for masses established in [26]). Moreover, for $n \leq \infty$, we give a uniform LDP in $\ell^\infty(\mathcal{G})$ for some classes $\mathcal{G}$.

2 The Mandelbrot measures

In this section $r \geq 2$ is fixed unless the contrary is mentionned.

Recall that for each fixed $f \in \mathcal{L}^1([0,1],\lambda)$, $\{\mu_r^n(f)\}_n$ is a martingale; by the martingale convergence theorem and by considering the positive and negative parts of f, we see that the limit

$$\mu_r(f) = \lim_n \mu_r^n(f) \tag{2.1}$$

exists a.s.. Let D be a countable dense subset of $\mathcal{C}([0,1])$ equipped with the sup norm $\|.\|_\infty$. Then a.s. (2.1) holds for all $f \in D$, and therefore for all $f \in \mathcal{C}([0,1])$ since $|\mu_r^n(f)| \leq \|f\|_\infty \mu_r^n(1)$ and $\mu_r(f) \leq \|f\|_\infty \mu_r(1)$. Hence

$$\text{a.s.} \quad \mu_r^\infty(f) = \mu_r(f) \quad \text{for all } f \in \mathcal{C}([0,1]) \tag{2.2}$$

(for any Borel measure μ and any integrable function f, we always write $\mu(f)$ for $\int f\, d\mu$; of course $\mu(A) = \mu(\mathbf{1}_A)$ for a Borel set A).

Kahane and Peyrière [19] proved that the positive martingale $\{\mu_r^n(1)\}_n$ is uniformly integrable if and only if $EW \log W < \log r$. In that case $\mu_r^n(1) \to \mu_r^\infty(1)$ a.s. and in L^1. Moreover a.s. $\mu_r^\infty(1) > 0$. For $p > 1$, the convergence of $\mu_r^n(1)$ towards $\mu_r^\infty(1)$ holds in L^p if (and only if) $EW^p < r^{p-1}$. In this section, we are interested in the extension of these results to the convergence of $\mu_r^n(f)$ to $\mu_r^\infty(f)$ for a fixed f in $\mathcal{L}^1([0,1],\lambda)$ or in $\mathcal{L}^p([0,1],\lambda)$.

Notice that we do not have that, a.s., for all bounded measurable functions f, $\mu_r^\infty(f) = \lim_n \mu_r^n(f)$. Indeed, from [19], for ω not in a negligible set N, the support of $\mu_r^\infty = \mu_{r,\omega}^\infty$ has Hausdorff dimension $1 - \frac{EW\log W}{\log r}$ and therefore has λ-measure 0. So considering $\omega_0 \notin N$, and $f = \mathbf{1}_{\text{supp}\mu_{r,\omega_0}^\infty}$, we have $\mu_{r,\omega_0}^\infty(f) = \mu_{r,\omega_0}^\infty(1) > 0$. But $\mu_{r,\omega_0}^n(f) = 0$ for all n, and thus the convergence does not hold for this function.

This is the reason why the following theorem involves fixed f.

Theorem 2.1 *a) If $EW \log W < \log r$, then for each fixed $f \in \mathcal{L}^1([0,1],\lambda)$, we have*

$$\mu_r^n(f) \to \mu_r^\infty(f) \ \text{ in } L^1 \text{ as } n \to \infty, \qquad \text{and} \quad \mu_r^\infty(f) = \mu_r(f) \text{ a.s.}. \tag{2.3}$$

b) If $EW^p < r^{p-1}$, then for each fixed $f \in \mathcal{L}^p([0,1],\lambda)$, we have

$$\mu_r^n(f) \to \mu_r^\infty(f) \ \text{ in } L^p \text{ as } n \to \infty. \tag{2.4}$$

To prove the L^1 convergence, we need two lemmas. In the first Lemma, we make clear the measurability of $\mu_r^\infty(f)$; in the second, we show that the expected value of $\mu_r^n(f)$ remains the same when taking the limit, so that the linear functionals $f \mapsto E\mu_r(f)$ and $f \mapsto E\mu_r^\infty(f)$ are continuous (from $\mathcal{L}^1([0,1],\lambda)$ to $\mathbb{R}$), uniformly in r.

Lemma 2.2 *If f is Borel and positive, then the application $\omega \mapsto \mu_r^\infty(f)$ is measurable.*

Lemma 2.3 *If $EW \log W < \log r$, then for each fixed $f \in \mathcal{L}^1([0,1],\lambda)$, we have*

$$E\mu_r(f) = E\mu_r^\infty(f) = \lambda(f). \tag{2.5}$$

For the L^p convergence, we have in fact the following more precise result, which gives rates of convergence. Recall that $\mu_r^0 = \lambda$.

Proposition 2.4 *Let $p > 1$, $f \in \mathcal{L}^p([0,1],\lambda)$, and assume $EW^p < \infty$. Set $\alpha(p) = 1 - 1/p$ if $p \le 2$, $\alpha(p) = 1/2$ if $p > 2$, $B_p = 2\min\{\sqrt{k} : k \in \mathbb{N}, k \ge p/2\}$ if $p \neq 2$, and $B_p = 1$ if $p = 2$.*

a) If $n \ge 1$ and $r \ge 2$, then

$$\left(E|\mu_r^n(f) - \mu_r^{n-1}(f)|^p\right)^{1/p} \le \left[\frac{B_p\|W\|_p}{r^{\alpha(p)}}\right]^n \frac{\|W-1\|_p}{\|W\|_p}\|f\|_p; \tag{2.6}$$

b) if $0 \le m < n \le \infty$ and $r \ge 2$ is large enough such that $B_p\|W\|_p < r^{\alpha(p)}$, then

$$(E|\mu_r^n(f) - \mu_r^m(f)|^p)^{1/p} \le \frac{\|W-1\|_p}{\|W\|_p}\left[1 - \frac{B_p\|W\|_p}{r^{\alpha(p)}}\right]^{-1}\left[\frac{B_p\|W\|_p}{r^{\alpha(p)}}\right]^{m+1}\|f\|_p. \tag{2.7}$$

In particular (taking $n = \infty$ or $m = 0$), if $r > 1$ is fixed large enough and $n \to \infty$, then $\mu_r^n(f)$ converges to $\mu_r^\infty(f)$ in L^p at a geometric rate; if $1 \le n \le \infty$ is fixed and $r \to \infty$, then $\mu_r^n(f)$ converges to $\lambda(f)$ in L^p at a polynomial rate (as $r^{-\alpha(p)}$).

For the case $p = 2$, additional results will be given in Section 4.

Proof of Lemma 2.2. For each fixed $f \in \mathcal{C}([0,1])$, $\mu_r^\infty(f)$ is measurable as the a.s. limit of the r.v.'s $\mu_r^n(f)$ $(n \to \infty)$. Therefore $\mu_r^\infty([a,b])$ is measurable for each fixed interval $[a,b] \subset [0,1]$, because we can find $f_k \in \mathcal{C}([0,1])$ with $0 \le f_k \le 1$ and $f_k \to \mathbf{1}_{[a,b]}$, so that $\mu_r^\infty([a,b]) = \lim_k \mu_r^\infty(f_k)$.

Let $\mathcal{B}$ be the Borel σ-field on $[0,1]$ and let $\mathcal{A}$ be the class of all $A \in \mathcal{B}$ such that the function $\omega \mapsto \mu_r^\infty(A)$ is measurable. Then $\mathcal{A}$ contains all the intervals $[a,b]$. By the σ-additivity of μ_r^∞, it is a Dynkin system and therefore $\mathcal{A} = \mathcal{B}$. Since each non-negative Borel function is the limit of a non-decreasing sequence of simple functions, this implies the measurability of $\mu_r^\infty(f)$ for f positive.

Proof of Lemma 2.3. (a) We first prove that $E\mu_r(f) = \lambda(f)$. Clearly, for each $1 \le n < \infty$,

$$E\mu_r^n(f) = \lambda(f). \tag{2.8}$$

We assume for the moment that $f \in \mathcal{L}^\infty([0,1],\lambda)$. Since $EW \log W < \log r$, $\mu_r^n(1) \to \mu_r(1)$ in L^1 (cf. [19]) and is therefore uniformly integrable. As $|\mu_r^n(f)| \le \|f\|_\infty \mu_r^n(1)$, this implies that $\{\mu_r^n(f)\}_n$ is also uniformly integrable, so that

$$\mu_r^n(f) \to \mu_r(f) \text{ in } L^1 . \tag{2.9}$$

Letting $n \to \infty$ in (2.8), we see that $E\mu_r(f) = \lambda(f)$.

If $f \in \mathcal{L}^1([0,1],\lambda)$. Fatou's lemma and (2.8) yield

$$E\mu_r(f) \le \lambda(f) \quad \text{ if } 0 \le f \in \mathcal{L}^1([0,1],\lambda).$$

Therefore the functional $f \mapsto E\mu_r(f)$ on $\mathcal{L}^1([0,1],\lambda)$ is 1-Lipschitz . On $\mathcal{L}^\infty([0,1],\lambda)$, it coincides with the continuous functional $f \mapsto \lambda(f)$. By density of $\mathcal{L}^\infty([0,1],\lambda)$ in $\mathcal{L}^1([0,1],\lambda)$, this implies that $E\mu_r(f) = \lambda(f)$ for all $f \in \mathcal{L}^1([0,1],\lambda)$.

(b) We then prove that $E\mu_r^\infty(f) = \lambda(f)$. Set $\bar\mu_r^\infty(A) = E\mu_r^\infty(A)$, for $A \in \mathcal{B}$ (recall that $\mathcal{B}$ is the Borel σ-field on $[0,1]$). The set function $\bar\mu_r^\infty$ is well defined by Lemma 2.2. The σ-additivity of μ_r^∞ implies that of $\bar\mu_r^\infty$. Therefore $\bar\mu_r^\infty$ is a Borel measure on $[0,1]$. For $f \in \mathcal{C}([0,1])$, we have

$$\bar\mu_r^\infty(f) = E\mu_r^\infty(f) = E\mu_r(f) = \lambda(f).$$

Therefore the measures $\bar\mu_r^\infty$ and λ coincide, so that $E\mu_r^\infty(f) = \lambda(f)$ for all $f \in \mathcal{L}^1([0,1],\lambda)$. □

Proof of Part a) of Theorem 2.1. Fix $f \in \mathcal{L}^1([0,1],\lambda)$. Let $\epsilon > 0$ be arbitrarily fixed, and take $g \in \mathcal{C}([0,1])$ such that $\lambda(|f-g|) < \epsilon$. By the triangle inequality and Lemma 2.3,

$$\begin{aligned} E|\mu_r^n(f) - \mu_r^\infty(f)| &\le E|\mu_r^n(f-g)| + E|\mu_r^n(g) - \mu_r^\infty(g)| + E|\mu_r^\infty(g-f)| \\ &\le 2\lambda(|f-g|) + E|\mu_r^n(g) - \mu_r^\infty(g)|. \end{aligned} \tag{2.10}$$

Because $g \in \mathcal{C}([0,1])$, we have $\mu_r^n(g) \to \mu_r^\infty(g) = \mu_r(g)$ in L^1 (cf. (2.9). Therefore letting $n \to \infty$ in (2.10), we see that

$$\limsup_n E|\mu_r^n(f) - \mu_r^\infty(f)| \le 2\epsilon,$$

so that $\mu_r^n(f) \to \mu_r^\infty(f)$ in L^1. Since $\mu_r^n(f) \to \mu_r(f)$ a.s., it follows that $\mu_r^\infty(f) = \mu_r(f)$ a.s. □

Proof of Proposition 2.4 and of Part b) of Theorem 2.1. We know that (see [8], p. 356) if $p > 1$ and $\{X_i : i \ge 1\}$ are independent real random variables with $E|X_i|^p < \infty$ and $EX_i = 0$, then for all $r \ge 1$,

$$E\,(|\sum_{i=1}^r X_i|^p) \le \begin{cases} (B_p)^p E(\sum_{i=1}^r |X_i|^p) & \text{if } 1 < p \le 2, \\ (B_p)^p E(\sum_{i=1}^r |X_i|^p)\; n^{(p/2)-1} & \text{if } p > 2. \end{cases} \tag{2.11}$$

Assume $p \leq 2$. From the decomposition (see (1.3))

$$(\mu_r^{n+1} - \mu_r^n)(f) = \frac{1}{r}\sum_{k=1}^{r} W_k(\mu_r^n - \mu_r^{n-1})(f) \circ \Theta_k(\tau_k^r f), \tag{2.12}$$

and (2.11), we get that

$$E|(\mu_r^{n+1} - \mu_r^n)(f)|^p \leq \frac{B_p^p}{r^p}\sum_{k=1}^{r} E(W^p)E|(\mu_r^n - \mu_r^{n-1})(\tau_k^r f)|^p.$$

For $n \geq 1$, let

$$C_n = \sup_{\|f\|_p = 1} E|(\mu_r^n - \mu_r^{n-1})(f)|^p.$$

Then by the previous inequality and the identity $\|f\|_p^p = r^{-1}\sum_{k=1}^r \|\tau_k^r f\|_p^p$, we obtain

$$C_{n+1} \leq B_p^p E(W^p) r^{1-p}\, C_n\,,$$

so that for all $n \geq 1$.

$$C_n \leq \left[B_p^p E(W^p) r^{1-p}\right]^{n-1} C_1.$$

From the decomposition

$$\mu_r^1(f) - \mu_r^0(f) = \sum_{k=1}^{r}(W_k - 1)\int_{A_k^r} f\, d\lambda$$

and (2.11), we get that

$$\begin{aligned} E|\mu_1^r(f) - \mu_r^0(f)|^p &\leq B_p^p E\left(|W-1|^p\right)\sum_{k=1}^{r}\int_{A_k^r}|f|^p\, d\lambda \\ &\leq B_p^p E\left(|W-1|^p\right)\, r^{1-p}\int_0^1 |f|^p d\lambda \end{aligned}$$

by Hölder's inequality. Consequently, for any $n \geq 1$,

$$C_n \leq \left[B_p^p E(W^p) r^{1-p}\right]^n \frac{E(|W-1|^p)}{E(W^p)}\,, \tag{2.13}$$

which implies (2.6).

Therefore, from the triangle inequality we obtain, for $0 \leq m < n < \infty$,

$$\left(E|\mu_r^n(f) - \mu_r^m(f)|^p\right)^{1/p} \leq \sum_{k=m+1}^{n}\left[B_p\|W\|_p r^{(1-p)/p}\right]^k \frac{\|W-1\|_p\,\|f\|_p}{\|W\|_p}. \tag{2.14}$$

When $B_p\|W\|_p < r^{(p-1)/p}$, this gives (2.7) for $n < \infty$.

In particular, $\{\mu_r^n(f)\}_n$ is a Cauchy sequence in L^p, so that it converges in L^p to $\mu_r^\infty(f)$ by Theorem 2.1 a). Letting $n \to \infty$ we see that (2.7) still holds for $n = \infty$.

If $p > 2$, the proof is similar: the only change is that, instead of (2.13), we have

$$C_n \leq \left[\frac{B_p^p E(|W|^p)}{r^{p/2}}\right]^n \frac{E(|W-1|^p)}{E(W^p)}, \qquad n \geq 1. \qquad \Box \tag{2.15}$$

3 Laws of Large Numbers

3.1 Useful Tools

We first give two tools needed for the proof of main results stated in the next subsection.

The first tool is an extension of the classical strong law of large numbers (SLLN) to a tree-indexed process, with or without weights.

Proposition 3.1 *Fix $n \geq 1$ and let $U^1, U^2, ..., U^n$ be independent and integrable r.v. Let $\left(U^n_{i_1...i_n}\right)$ be a family of independent r.v. indexed by $(n, i_1, .., i_n)$, such that for every n, $U^n_{i_1...i_n}$ has the same distribution as U^n.*

(a) For $r \geq 1$, set

$$S^n_r = \frac{1}{r^n} \sum_{1 \leq i_1,...,i_n \leq r} U^1_{i_1} ... U^n_{i_1...i_n}, \tag{3.1}$$

and let $\mathcal{H}^n_r$ be the σ-field generated by $\{S^n_k\ ,\ k \geq r\}$. Then $(S^n_r, \mathcal{H}^n_r)_r$ is a reverse martingale, and

$$\lim_{r\to\infty} S^n_r = \prod_{k=1}^{n} EU^k \qquad a.s.\ and\ in\ L^1.$$

(b) Assume additionally $EU^n = 0$. If $\mathbf{a} = \left(a^r_{i_1...i_n} : 1 \leq i_1...i_n \leq r; r \geq 1\right)$ is a family of real numbers such that $\|\mathbf{a}\|_\infty := \sup_r \max_{1 \leq i_1...i_n \leq r} |a^r_{i_1...i_n}| < \infty$, then as $r \to \infty$,

$$\Gamma_r(\mathbf{a}) \ := \ \frac{1}{r^n} \sum_{1 \leq i_1,..,i_n \leq r} U^1_{i_1} ... U^n_{i_1...i_n} a^r_{i_1...i_n} \ \to\ 0 \qquad a.s.\ and\ in\ L^1. \tag{3.2}$$

The proof of Part (a) is similar to that in [32] (Theorem 51.3 p. 150) for the classical SLLN, which corresponds to the case $n = 1$. The key point in the proof of Part (b) is the following truncating argument.

Lemma 3.2 *Assume that the conditions of Proposition 3.1 are satisfied. For $M > 0$, let $\bar{U}^k_{i_1...i_k} := (-M \vee U^k_{i_1...i_k}) \wedge M$ be the truncation of $U^k_{i_1...i_k}$ at level M, and set $\widetilde{U}^k_{i_1...i_k} := \bar{U}^k_{i_1...i_k} - E\bar{U}^k_{i_1...i_k}$ and*

$$\Gamma^M_r(\mathbf{a}) \ := \ \frac{1}{r^n} \sum_{1 \leq i_1,..,i_n \leq r} \bar{U}^1_{i_1} ... \bar{U}^{n-1}_{i_1...i_{n-1}} \widetilde{U}^n_{i_1...i_n} a^r_{i_1...i_n}.$$

Then

$$\lim_{M\to\infty} \limsup_{r} \sup_{\mathbf{a}:\|\mathbf{a}\|_\infty \leq 1} |\Gamma_r(\mathbf{a}) - \Gamma^M_r(\mathbf{a})| = 0 \quad a.s.. \tag{3.3}$$

The second tool is the L^1 law of large numbers for a triangular array of independent non necessarily identically distributed random variables. It can easily be derived from Feller's weak law of large numbers (cf. [31] chapter 9 p.258), since the sequence (U_n) in the following proposition is uniformly integrable.

Proposition 3.3 *Let $\{U_{n,k} : n \geq 1, 1 \leq k \leq r_n\}$ be a triangular array of row-wise independent, integrable and centered real random variables such that $\lim_{n\to\infty} r_n = \infty$. If the family $\{U_{n,k} : n \geq 1, 1 \leq k \leq r_n\}$ is uniformly integrable, then as $n \to \infty$,*

$$U_n := \frac{1}{r_n} \sum_{k=1}^{r_n} U_{n,k} \to 0 \ in\ L^1.$$

Proof of Proposition 3.1 a) For simplicity, we only give details for $n = 2$; the case $n \neq 2$ may be managed in the same way. By linearity and exchangeability,

$$E\left(S_r^2|\mathcal{H}_{r+1}^2\right) = E\left(U_1^1 U_{11}^2|\mathcal{H}_{r+1}^2\right) \quad (3.4)$$

$$= E\left(S_{r+1}^2|\mathcal{H}_{r+1}^2\right) = S_{r+1}^2 \quad (3.5)$$

Therefore $\{(S_r^2, \mathcal{H}_r^2)\}_r$ is a reverse martingale, so that the limit $L = \lim_r S_r^2$ exists a.s. and in L^1. We shall identify the limit L. For fixed $k > 0$, we have $S_{r+k}^2 = A_{r+k} + B_{r+k} + C_{r+k}$, where

$$A_{r+k} = \frac{1}{(r+k)^2} \sum_{i_1=1}^{k} U_{i_1}^1 \left(\sum_{i_2=1}^{r+k} U_{i_1,i_2}^2\right), \quad (3.6)$$

$$B_{r+k} = \frac{1}{(r+k)^2} \sum_{i_1=k+1}^{k+r} \left(U_{i_1}^1 \sum_{i_2=1}^{k} U_{i_1,i_2}^2\right), \quad (3.7)$$

$$C_{r+k} = \frac{1}{(r+k)^2} \sum_{i_1=k+1}^{k+r} \left(U_{i_1}^1 \sum_{i_2=k+1}^{k+r} U_{i_1,i_2}^2\right). \quad (3.8)$$

By the SLLN, for every $i_1 \leq k$, $\lim \frac{1}{(r+k)} \sum_{i_2=1}^{r+k} U_{i_1,i_2}^2 = E(U^2)$ a.s. so that $\lim_r A_{r+k} = 0$ a.s. Similarly, $\lim_r B_{r+k} = 0$ a.s. because $\{U_{i_1}^1 \sum_{i_2=1}^{k} U_{i_1,i_2}^2;\ i_1 \geq k+1\}$ is an i.i.d. family (with mean $kE(U^1)E(U^2)$). Hence

$$L = \limsup_r C_{r+k} \quad a.s.$$

But C_{n+k} is measurable with respect to $\mathcal{I}_k$, the σ-field generated by U_j^1 and $U_{ij}^2, i, j \geq k$; hence L is $\mathcal{I}_k$-measurable for every k. From the 0-1 law, it is a.s. constant. Therefore $L = EL = \lim ES_r^2 = \prod_{k=1}^{n}(EU^k)$. $\square$

Proof of Lemma 3.2 We estimate the cost of truncating (see Etemadi [14]). Changing $U_{i_1..i_n}^n$ into $\bar{U}_{i_1..i_n}^n$ looses no more than

$$\Pi_1^r = \frac{1}{r^n} \sum_{i_1...i_n} |U_{i_1}^1|...|U_{i_1...i_{n-1}}^{n-1}||U_{i_1..i_n}^n - \bar{U}_{i_1..i_n}^n|,$$

and from lemma 3.1 we have a.s.

$$\limsup_r \Pi_1^r = E|U^n - \bar{U}^n| \prod_{k=1}^{n-1} E|U^k| = E\left(|U^n| - M\right)^+ \prod_{k=1}^{n-1} E|U^k|. \quad (3.9)$$

Let us replace $\bar{U}^n$ by $\widetilde{U}^n = \bar{U}^n - E\bar{U}^n$ to get a centered r.v.. Since $EU^n = 0$, $E\bar{U}^n = E(M + U)^- - E(U - M)^+$ and $|E\bar{U}^n| \leq E(M + U^n)^- + E(U^n - M)^+ = E\left(|U^n| - M\right)^+$ and we loose no more than

$$\Pi_2^r = E\left(|U^n| - M\right)^+ \frac{1}{r^{n-1}} \sum_{i_1...i_{n-1}} |U_{i_1}^1|...|U_{i_1...i_{n-1}}^{n-1}|$$

which gives again

$$\limsup_r \Pi_2^r = E\left(|U^n| - M\right)^+ \prod_{k=1}^{n-1} E|U^k|. \quad (3.10)$$

Let us change sequentially every $U_{i_1..i_k}^k$ into $\bar{U}_{i_1..i_k}^k$ starting from $k = 1$ to $k = n-1$. At the k-th step we loose no more than

$$\Pi_{3,k}^r = \frac{1}{r^n} \sum_{i_1...i_n} |\bar{U}_{i_1}^1|...|\bar{U}_{i_1...i_{k-1}}^{k-1}| \left(|U_{i_1...i_k}^k - \bar{U}_{i_1...i_k}^k|\right) |U_{i_1...i_{k+1}}^{k+1}|....|U_{i_1..i_{n-1}}^{n-1}||\widetilde{U}_{i_1...i_n}^n|,$$

which gives again

$$\begin{aligned}\limsup_r \Pi^r_{3,k} &= \left(\prod_{j=1}^{k-1} E|\bar{U}^j|\right) E\left(|U^k| - M\right)^+ \left(\prod_{l=k+1}^{n-1} E|U^j|\right) E|\widetilde{U}_n| \\ &\leq 2\left(\prod_{j\neq k\leq n} E|U^j|\right) E\left(U^k - M\right)^+ .\end{aligned}$$

The sum of looses for fixed n has a.s. a limsup bounded by

$$\limsup_r \left(\Pi^r_1 + \Pi^r_2 + \sum_{k=1}^{n-1}\Pi^r_{3,k}\right) \leq 2\left(\prod_{j=1}^{n} E|U^j|\right)\left(\sum_{k=1}^{n}\frac{E\left(|U^k| - M\right)^+}{EU^k}\right) .$$

Since M is arbitrary, the proof is ended.

Proof of Proposition 3.1 b) We use the method of [7]. From Lemma 3.2, we can assume that the r.v.'s are bounded by M.

For the a.s. convergence, it is sufficient to prove that

$$\limsup_r \Gamma_r(a) \leq 0\,, \tag{3.11}$$

since after changing U^n into $-U^n$ we get the reverse inequality. By independence, for all $t \in \mathbb{R}$,

$$E\exp tr^n\Gamma_r(a) = E\prod_{i_1,..,i_n} E\left[\exp tU^1_{i_1}...U^{n-1}_{i_1...i_{n-1}}U^n a^r_{i_1...i_n}|\mathcal{F}_{n-1}\right] . \tag{3.12}$$

From Hoeffding's inequality, if $X\in[a,b]$,

$$\log Ee^{t(X-EX)} \leq \frac{1}{8}(b-a)^2t^2 . \tag{3.13}$$

Here, since each $|U^k|$ is bounded by M and each $|a^r_{\mathbf{i}}|$ by $\|a\|_\infty$, we get

$$\log E\left[\exp tU^1_{i_1}...U^{n-1}_{i_1...i_{n-1}}U^n a^r_{i_1...i_n}|\mathcal{F}_{n-1}\right] \leq \frac{1}{2}t^2\|a\|^2_\infty M^{2n} . \tag{3.14}$$

Hence by (3.12),

$$\log E\exp tr^n\Gamma_r(a) \leq \frac{1}{2}t^2\|a\|^2_\infty M^{2n} r^n\,, \tag{3.15}$$

so that

$$P\left(\Gamma_r(a) > \delta\right) \leq \exp\left(-\frac{\delta^2 r^n}{2M^{2n}\|a\|^2_\infty}\right) . \tag{3.16}$$

Therefore (3.11) follows from the Borel-Cantelli lemma.

For the L^1 convergence, it suffices to show the uniform integrability of $\{\Gamma_r(a)\}_r$. Actually $\Gamma_r(a)$ is bounded by

$$\bar{\Gamma}_r := \frac{\|a\|_\infty}{r^n}\sum_{1\leq i_1,\ldots,i_n\leq r}|U^1_{i_1}...U^n_{i_1...i_n}|\,,$$

which converges in L^1 by Proposition 3.1 a), and is therefore uniformly integrable. □

3.2 Main results

For $n \leq \infty$ and some subset $\mathcal{G}$ of $\mathcal{L}^1([0,1],\lambda)$, we shall study L^1 and a.s. convergence of

$$\|\mu_r^n - \lambda\|_{\mathcal{G}} := \sup_{f\in\mathcal{G}} |\mu_r^n(f) - \lambda(f)| \tag{3.17}$$

as $r \to \infty$. In order to obtain uniform convergence results for finite n, we need finiteness of metric entropy in $\mathcal{L}^1([0,1],\lambda)$.

Definition 3.4 *Let (V,d) be an arbitrary semi-metric space and T be a subset of V. The covering number $N(\epsilon,T,d)$ is the minimal number of balls of radius ϵ needed to cover T. The entropy number is $H(\epsilon,T,d) = \log N(\epsilon,T,d)$. The subset T is said to be totally bounded in (V,d) if $N(\epsilon,T,d)$ is finite for all $\epsilon > 0$.*

For $n = \infty$, we need a stronger assumption.

Definition 3.5 *For $f,g \in \mathcal{L}^1([0,1],\lambda)$ such that $f \leq g$, the bracket $[f,g]$ is the set of all $h \in \mathcal{L}^1([0,1],\lambda)$ such that $f \leq h \leq g$; it is called an ϵ-bracket if $\lambda(g-f) \leq \epsilon$. The class $\mathcal{G}$ is said to be totally bounded with brackets in $\mathcal{L}^1([0,1],\lambda)$ if it can be covered by a finite number of ϵ-brackets, for all $\epsilon > 0$.*

As in the setting of smoothed partial-sum processes, for each $n < \infty$ fixed, the process $\{\mu_r^n(f), f \in \mathcal{G}\}$ has $d_\lambda^{(1)}$ continuous sample paths ($d_\lambda^{(1)}(f,g) = \lambda(|f-g|)$ for $f,g \in \mathcal{L}^1([0,1],\lambda)$), so the uniform deviation defined in (3.17) is measurable and can be handled as the maximum on some countable dense subset.

Theorem 3.6 *Let $1 \leq n < \infty$ be fixed.*

(a) $\lim_r Z_r^n = 1$ a.s. and in L^1.

(b) For $f \in \mathcal{L}^1([0,1],\lambda)$

$$\lim_r \mu_r^n(f) = \lambda(f) \text{ in } L^1 . \tag{3.18}$$

(c) If $\mathcal{G}$ is a class of uniformly bounded functions, totally bounded in $\mathcal{L}^1([0,1],\lambda)$, then

$$\lim_r \|\mu_r^n - \lambda\|_{\mathcal{G}} = 0 \quad \text{a.s. and in } L^1 . \tag{3.19}$$

Remark : As can be seen in the proof, the uniform boundedness is not needed for the a.s. convergence if additionally $\operatorname{ess\,sup} W < \infty$.

Theorem 3.7 *Assume $EW \log^+ W < \infty$.*

(a) $\lim_r Z_r^\infty = 1$ a.s and in L^1.

(b) For $f \in \mathcal{L}^1([0,1],\lambda)$,

$$\lim_r \mu_r^\infty(f) = \lambda(f) \text{ in } L^1 . \tag{3.20}$$

(c) If $\mathcal{G}$ is a subset of $\mathcal{L}^1([0,1],\lambda)$ such that, for each $\epsilon > 0$, it can be covered by a finite number of ϵ-brackets $[f_i,g_i]$, with f_i and g_i measurable, bounded and λ-a.e. continuous, then

$$\lim_r E^* \|\mu_r^\infty - \lambda\|_{\mathcal{G}} = 0 \quad \text{and} \quad \lim_r \|\mu_r^\infty - \lambda\|_{\mathcal{G}} = 0 \;\; P^* \text{a.s.}, \tag{3.21}$$

where P^* and E^* denote the corresponding outer probability and outer expectation.

(d) *Furthermore if* $EW^{1+\delta} < \infty$ *for some* $\delta > 0$*, then the assumption of a.e. continuity of* f_i, g_i *can be removed in (c).*

Remarks

1) Among the classes satisfying assumptions of Theorem 3.7 c), let us mention:
   - any relatively compact subset of $\mathcal{C}([0,1])$;
   - the class $\{\mathbf{1}_{[0,t]},\ t \in [0,1]\}$;
   - more generally for $\eta : [0,1] \to \mathbb{R}^+$, increasing and continuous with $\eta(0) = 0$, the class

$$\mathcal{S}_\eta = \{\mathbf{1}_C \text{ with } C \subset [0,1] \ ; \ \lambda\left((\partial C)^\varepsilon\right) \le \eta(\varepsilon), \text{ for all } \varepsilon \in [0,1]\}, \tag{3.22}$$

   where ∂C denotes the boundary of C and $(\partial C)^\epsilon$ the set of x such that $d(x, \partial C) \le \epsilon$.

2) The L^1 convergence of Part (a) was announced in ([26], Corollary to Theorem 1.1); but there was an error in the proof of (2.8) of that paper.

Proof of Theorem 3.6 Part (a) is a direct consequence of Lemma 3.1(a). To prove Parts (b) and (c), we first remark that for each $f \in \mathcal{L}^\infty([0,1], \lambda)$ and $1 \le n < \infty$,

$$\mu_r^n(f) - \mu_r^{n-1}(f) \to 0 \quad \text{a.s. and in } L^1 \quad (r \to \infty), \tag{3.23}$$

by applying Proposition 3.1(b) to the decomposition

$$\mu_r^n(f) - \mu_r^{n-1}(f) = \sum_{1 \le i_1, \ldots, i_n \le r} W_{i_1} \ldots W_{i_1 \ldots i_{n-1}} (W_{i_1 \ldots i_n} - 1) \int_{A^r_{i_1 \ldots i_n}} f \, d\lambda \,. \tag{3.24}$$

Since $\mu_r^0 = \lambda$, (3.23) implies that, for each $f \in \mathcal{L}^\infty([0,1], \lambda)$ and $1 \le n < \infty$,

$$\mu_r^n(f) - \lambda(f) \to 0 \quad \text{a.s. and in } L^1 \quad (r \to \infty). \tag{3.25}$$

By the density of $\mathcal{L}^\infty([0,1], \lambda)$ in $\mathcal{L}^1([0,1], \lambda)$, the equation (2.8) and the inequality

$$|\mu_r^n(f) - \lambda(f)| \le \mu_r^n(|f - g|) + |\mu_r^n(g) - \lambda(g)| + \lambda(|g - f|), \tag{3.26}$$

we see that the L^1 convergence in (3.25) still holds for every $f \in \mathcal{L}^1([0,1], \lambda)$, which ends the proof of b).

For Part (c), we assume that $\mathcal{G}$ is uniformly bounded by 1 for the sake of simplicity. To prove the a.s. convergence, it is enough to show that for every $n < \infty$

$$\lim_r \|\mu_r^n - \mu_r^{n-1}\|_{\mathcal{G}} = 0 \ \text{ a.s.}. \tag{3.27}$$

From Lemma 3.2, it is sufficient to prove (3.27) when all the r.v. $W_{\mathbf{i}}$ are bounded by a constant $M \ge 1$. Since $\mathcal{G}$ is totally bounded, for every $\epsilon > 0$ one can find $f_1, ..f_N \in \mathcal{L}^1([0,1], \lambda)$ such that for every $f \in \mathcal{G}$ there is some f_i such that $\lambda(|f - f_i|) \le \epsilon$. Actually we can choose the functions f_i in $\mathcal{L}^\infty([0,1], \lambda)$ since it is dense in $\mathcal{L}^1([0,1], \lambda)$.

By definition of μ_r^n, we then have, for $g \in \mathcal{L}^1([0,1], \lambda)$ and $n \ge 0$

$$|\mu_r^n(g)| \le M^n \lambda(|g|).$$

Hence, for $f \in \mathcal{G}$ and $\lambda(|f - f_i|) \le \epsilon$,

$$\left| \left(\mu_r^n(f) - \mu_r^{n-1}(f)\right) - \left(\mu_r^n(f_i) - \mu_r^{n-1}(f_i)\right) \right| \le 2M^n \epsilon \,. \tag{3.28}$$

Now, from (3.25) a.s. for every $1 \le i \le N$,

$$\lim_r \left(\mu_r^n(f_i) - \mu_r^{n-1}(f_i)\right) = 0\,. \tag{3.29}$$

Jointly with (3.28) it yields a.s.

$$\limsup_r \|\mu_r^n - \mu_r^{n-1}\|_{\mathcal{G}} \le 2M^n\epsilon \tag{3.30}$$

for every ϵ. This gives the a.s. convergence of Part c).

To get the L^1 convergence, it is enough to prove, for every fixed $n < \infty$, the uniform integrability of $\left(\|\mu_r^n - \mu_r^{n-1}\|_{\mathcal{G}}\right)_r$. But this is indeed the case because $\|\mu_r^n - \mu_r^{n-1}\|_{\mathcal{G}}$ is bounded by

$$S_r^n := r^{-n} \sum_{i_1,\ldots,i_n} W_{i_1}\ldots W_{i_1\ldots i_{n-1}}|W_{i_1,\ldots,i_n} - 1| \tag{3.31}$$

which by Proposition 3.1 converges in L^1 and is therefore uniformly integrable. □

Proof of Theorem 3.7

a) For $n \le +\infty$, let $\mathcal{H}_r^n$ be the σ-field generated by Z_s^n, $s \ge r$. By Proposition 3.1 a), for each $n < \infty$, $\{Z_r^n, \mathcal{H}_r^n\}$ is a reverse martingale, which means that for every integer $p \ge 1$ and every bounded and continuous function $g : \mathbb{R}^p \to \mathbb{R}$, we have

$$E\left(Z_r^n g\left(Z_{r+1}^n, Z_{r+2}^n, \ldots, Z_{r+p}^n\right)\right) = E\left(Z_{r+1}^n\, g\left(Z_{r+1}^n, Z_{r+2}^n, \ldots, Z_{r+p}^n\right)\right). \tag{3.32}$$

Let $r_0 \ge 2$ be such that $EW\log W < \log r_0$. For each fixed $r \ge r_0$, as $n \to \infty$, $Z_r^n \to Z_r^\infty$ a.s. and in L^1. Therefore using uniform integrability, we may let $n \to \infty$ in (3.32) , showing that $\{Z_r^\infty, \mathcal{H}_r^\infty\}$ is also a reverse martingale. Therefore Z_r^∞ converges a.s. and is uniformly integrable. To identify the limit, we will see in b) below that $Z_r^\infty \to 1$ in L^1, so that the proof is finished.

b) We first prove that for each $f \in \mathcal{L}^\infty([0,1],\lambda)$,

$$\mu_r^\infty(f) \to \lambda(f) \text{ in } L^1. \tag{3.33}$$

By extension of (1.3) to the associated Borel measures we get the decomposition

$$\mu_r^\infty(f) - \lambda(f) = \frac{1}{r}\sum_{k=1}^{r} \left[W_k \mu_r^\infty \circ \Theta_k(\tau_k^r f) - \lambda(\tau_k^r f)\right].$$

Since $|\mu_r^\infty \circ \Theta_k(\tau_k^r f) - \lambda(\tau_k^r f)| \le Z_r^\infty \circ \Theta_k$, the family $\{W_k\mu_r^\infty \circ \Theta_k(\tau_k^r f) - \lambda(\tau_k^r f)\}_{k,r}$ is uniformly integrable, so that Proposition 3.3 gives (3.33). By the density of $\mathcal{L}^\infty([0,1],\lambda)$ in $\mathcal{L}^1([0,1],\lambda)$, using equation (2.5) and inequality (3.26) with $n = \infty$ we see that (3.33) holds for $f \in \mathcal{L}^1([0,1],\lambda)$. This ends the proof of b). Taking $f = 1$ in (3.33), we see that $Z_r^\infty \to 1$ in L^1, which ends also the proof of a).

c) Let us first reduce the problem to a simpler one involving only one function. Let $\epsilon > 0$, and let $\{[f_i, g_i] : 1 \le i \le N\}$ be a cover of $\mathcal{G}$ by ϵ-brackets, with f_i and g_i measurable, bounded and λ-a.s. continuous. If $f \in [f_i, g_i]$, then

$$\mu_r^\infty(f) - \lambda(f) \le \mu_r^\infty(g_i) - \lambda(f_i) = [\mu_r^\infty(g_i) - \lambda(g_i)] + [\lambda(g_i) - \lambda(f_i)]$$

and

$$\mu_r^\infty(f) - \lambda(f) \ge \mu_r^\infty(f_i) - \lambda(g_i) = [\mu_r^\infty(f_i) - \lambda(f_i)] + [\lambda(f_i) - \lambda(g_i)].$$

Therefore

$$\|\mu_r^\infty - \lambda\|_{\mathcal{G}} \le \max\{|\mu_r^\infty(g_i) - \lambda(g_i)|, |\mu_r^\infty(f_i) - \lambda(f_i)| : 1 \le i \le N\} + \epsilon. \tag{3.34}$$

c1) To prove the a.s. convergence, it is convenient to introduce the random measures $\tilde{\mu}_r^n$ defined by

$$\tilde{\mu}_r^n = \frac{1}{r}\sum_{k=1}^r W_k(Z_r^{n-1}\circ\Theta_k)\delta_{\frac{k}{r}}, \quad 1 \le n \le \infty, \tag{3.35}$$

(recall that by convention $Z_r^{n-1}\circ\Theta_k = 1$ if $n = 1$, and $= Z_r^\infty\circ\Theta_k$ if $n = \infty$), and to compare it with μ_r^n with the help of (1.3).

Let us first prove that a.s. for all $t \in [0, 1]$,

$$\lim_r \tilde{\mu}_r^\infty([0,t]) = t. \tag{3.36}$$

For fixed $t \in (0, 1]$ and $1 \le n < \infty$, set

$${}^tZ_r^n := \frac{r}{[rt]}\tilde{\mu}_r^n([0,t]) = \frac{1}{[rt]}\sum_{i_1=1}^{[rt]} W_{i_1}\sum_{1\le i_2,..i_n\le r}\frac{W_{i_1i_2}...W_{i_1...i_n}}{r^{n-1}}, \tag{3.37}$$

where $[x]$ is the integer part of x. By (3.37), if $EW\log W < \log r$, then as $n\to\infty$, ${}^tZ_r^n$ converges a.s. and in L^1 to

$${}^tZ_r^\infty := \frac{1}{[rt]}\sum_{k=1}^{[rt]} W_k Z_r^\infty\circ\Theta_k. \tag{3.38}$$

For $1 \le n \le \infty$, let ${}^t\mathcal{H}_r^n$ be the σ-field generated by $\{{}^tZ_k^n\ ,\ k \ge r\}$. Let $r_t \ge t^{-1}$ be such that $EW\log W < \log r$. Just like $\{Z_r^n\}_r$, for each fixed $1 \le n \le \infty$, the sequence $\{{}^tZ_r^n\}_{r\ge r_t}$ is a reverse martingale with respect to $\{{}^t\mathcal{H}_r^n\}_{r\ge t^{-1}}$ (the proof is similar), so that it converges a.s. and in L^1. To identify the limit of ${}^tZ_r^\infty$, we use

$${}^tZ_r^\infty - 1 = \frac{1}{[rt]}\sum_{k=1}^{[rt]}(W_kZ_r^\infty\circ\Theta_k - 1) \tag{3.39}$$

and Proposition 3.3 to conclude that ${}^tZ_r^\infty \to 1$ in L^1. Since $\tilde{\mu}_r^\infty([0,t]) = \frac{[rt]}{r}\,{}^tZ_r^\infty$, it follows that $\lim_r \tilde{\mu}_r^\infty([0,t]) = t$ a.s. (and in L^1). By a classical monotonicity argument, this implies (3.36), hence the a.s. weak convergence of $\tilde{\mu}_r^\infty$ to λ. To get the similar result for μ_r^∞, observe first that, from (1.4)

$$\mu_r^\infty(f) - \tilde{\mu}_r^\infty(f) = \frac{1}{r}\sum_{k=1}^r W_k\ \mu_r^\infty\circ\Theta_k(\tau_k^r f - f(k/r)).$$

Since , for $f \in \mathcal{C}$, $\sup_{x\in[0,1]}|\tau_k f(x) - f_{k,r}| \le \omega_f(r^{-1})$, where $\omega_f(h)$ is the maximal oscillation of f on intervals of size h, $h > 0$, we have

$$|\ \mu_r^\infty(f) - \tilde{\mu}_r^\infty(f)\ | \le \frac{\omega_f(r^{-1})}{r}\sum_{k=1}^r W_kZ_r^\infty\circ\Theta_k\ = \omega_f(r^{-1})Z_r^\infty, \tag{3.40}$$

where the last equality holds by (1.5). This yields the a.s. weak convergence of μ_r^∞ to λ. Therefore (cf. [6], p.163, Proposition 8.12) a.s. for all f measurable, bounded and λ-a.s. continuous,

$$\lim_r \mu_r^\infty(f) = \lambda(f). \tag{3.41}$$

Replacing f by f_i, g_i in the above display and using (3.34), we see that

$$P^* \text{ a.s. } \limsup_r \|\mu_r^\infty - \lambda\|_{\mathcal{G}} \le \epsilon\,.$$

The proof of the a.s. convergence is therefore finished.

c2) Taking E^* in (3.34) and using b) gives the L^1-convergence.

d) We notice that if $EW^{1+\delta} < \infty$, then by Theorem 2.4, for n sufficiently large and f in $\mathcal{L}^{1+\delta}([0,1],\lambda)$,

$$E\sum_r |\mu_r^\infty(f) - \mu_r^n(f)|^{1+\delta} < \infty,$$

so that a.s. $\lim_r |\mu_r^\infty(f) - \mu_r^n(f)| = 0$. Since a.s. $\lim_r |\mu_r^n(f) - \lambda(f)| = 0$ for $f \in \mathcal{L}^\infty([0,1],\lambda)$, it follows that a.s. (3.41) holds for $f \in \mathcal{L}^\infty([0,1],\lambda)$. Therefore $\lim_r \|\mu_r^\infty - \lambda\|_{\mathcal{G}} = 0$ a.s. provided that $\mathcal{G}$ can be covered by a finite number of ϵ-brackets $[f_i, g_i]$ with $f_i, g_i \in \mathcal{L}^\infty([0,1],\lambda)$. □

4 Functional Central Limit Theorems

In this section, assuming $EW^2 < \infty$, we establish functional central limit theorems for the process $\{\mu_r^n(f) : f \in \mathcal{G}\}$, where $\mathcal{G}$ is a subclass of $\mathcal{L}^2([0,1],\lambda)$.

Definition 4.1 *A class $\mathcal{G}$ is called regular if there exists r_0 such that*

$$a.s. \quad \forall r > r_0 \quad \forall f \in \mathcal{G} \quad \lim_n \mu_r^n(f) = \mu_r^\infty(f)\,. \tag{4.1}$$

Remark: 1) Since a.s. $\{\mu_r^n\}_n$ converges weakly to μ_r^∞ (which is a.s. atomless), the class $\mathcal{C}([0,1])$ of continuous functions and the class $\{1_{[0,t]}, t \in [0,1]\}$ of cells are regular classes. Consequently the class of functions with bounded variation is also regular.

2) For $\eta(x) = x^\delta$ and $\delta > 0$, the class $\mathcal{S}_\eta$ is regular (see the Hausdorff dimensions of the supports of μ_r^∞ and [6], p.163, Prop. 8.12). On the contrary, for $\eta(x) = (|\log x| + 1)^{-1}$, the class $\mathcal{S}_\eta$ (see (3.22)) is not regular, due to the remark before Theorem 2.1.

The following assumption prevails in this section (cf. [35], p.127 and 133).

Assumption 4.2 *$\mathcal{G}$ is a subclass of $\mathcal{L}^2([0,1],\lambda)$ consisting of uniformly bounded functions and satisfying the following universal entropy condition:*

$$I_{\mathcal{G}} := \int_0^\infty \sup_Q \sqrt{H\left(x, \mathcal{G}, L_2(Q)\right)}\, dx < \infty\,, \tag{4.2}$$

where the sup is taken over all probability measures Q on $[0,1]$ and H is the entropy number of Definition 3.4.

Set

$$\sigma = \sqrt{E[(W-1)^2]} \quad \text{and} \quad \sigma_n = \sqrt{(\sigma^2+1)^n \sigma^2}.$$

We start from a result of [40] (Section 7.5).

Theorem 4.3 (Ziegler) *If $\mathcal{G}$ satisfies Assumption 4.2, then as $r \to \infty$,*

$$\frac{\sqrt{r}}{\sigma}(\mu_r^1 - \lambda) \Rightarrow \mathbb{G} \;\; in\; \ell^\infty(\mathcal{G})\,, \tag{4.3}$$

where the limit $\mathbb{G}$ is a centered Gaussian process indexed by $\mathcal{G}$, with $d_\lambda^{(2)}$-continuous paths and covariance $E\left(\mathbb{G}(f)\mathbb{G}(g)\right) = \lambda(fg)$, and where $\Rightarrow$ stands for the weak convergence.

In the following theorems, we extend this result to all orders (Th. 4.4), and obtain other new limit results (Th. 4.5).

Theorem 4.4 *Let $\mathcal{G}$ satisfy Assumption 4.2. Then :*
a) for every $1 \leq n < \infty$ fixed,

$$\frac{\sqrt{r}}{\sigma}(\mu_r^n - \lambda) \Rightarrow \mathbb{G}, \ \ in\ \ell^\infty(\mathcal{G})\,; \tag{4.4}$$

b) for $n = \infty$, if $\mathcal{G}$ is regular (and satisfies Assumption 4.2), then (4.4) holds.

Theorem 4.5 *Let $\mathcal{G}$ satisfy Assumption 4.2. Then :*
a) for every $0 \leq n < \infty$, setting

$$\xi_r^n := \frac{r^{(n+1)/2}}{\sigma_n}\left(\mu_r^{n+1} - \mu_r^n\right)$$

the sequence of random processes $\{\xi_r^n\}_r$ converges weakly in $\ell^\infty(\mathcal{G})$ to $\mathbb{G}$;
b) for every $d \geq 0$, the multivariate random process $\vec{\xi}_r := \left(\xi_r^0, \ldots, \xi_r^d\right)$ converges weakly in $\ell^\infty\left(\mathcal{G}, \mathbb{R}^{d+1}\right)$ to $(\mathbb{G}_0, \ldots, \mathbb{G}_d)$ as $r \to \infty$, where $\mathbb{G}_0, \ldots, \mathbb{G}_d$ are independent Gaussian processes with the same distribution as $\mathbb{G}$.

Remark: Combining methods used in the proofs of these theorems, one can show that the last component of $\vec{\xi}_r$ may be replaced by

$$\zeta_r^{d,m} := \frac{r^{(d+1)/2}}{\sigma_d}\left(\mu_r^m - \mu_r^d\right)$$

for every $m > d$ (assuming regularity for $m = \infty$).

4.1 Proof of Theorem 4.4

The main ingredients are Ziegler's theorem and estimates of the increments $\|\mu_r^{n+1} - \mu_r^n\|_{\mathcal{G}}$ in quadratic mean.

Proposition 4.6 *Let $\mathcal{G}$ satisfy Assumption 4.2 and let*

$$\Delta_r^n := \|\mu_r^{n+1} - \mu_r^n\|_{\mathcal{G}} := \sup_{f\in\mathcal{G}} |\mu_r^{n+1}(f) - \mu_r^n(f)|, \qquad n \geq 1. \tag{4.5}$$

Then for some constant $K > 0$,

$$E\left(\Delta_r^n\right)^2 \leq K I_{\mathcal{G}} \sigma_n^2 r^{-(n+1)}. \tag{4.6}$$

<u>Proof.</u> From (1.1) we have

$$\mu_r^{n+1}(f) - \mu_r^n(f) = \sum_{1\leq i_1,\ldots,i_n\leq r} p_{i_1\ldots i_n}\left(\sum_{i_{n+1}=1}^{r} Y_{i_1\ldots i_{n+1}} \int_{A^r_{i_1\cdots i_{n+1}}} f\, d\lambda\right), \tag{4.7}$$

where $p_{i_1\ldots i_n} = W_{i_1}\ldots W_{i_1\ldots i_n}$ and $Y_{i_1\ldots i_{n+1}} = W_{i_1\ldots i_{n+1}} - 1$. Then, conditionally on $\mathcal{F}_n$, $\mu_r^{n+1} - \mu_r^n$ is a centered random measure. From the symmetrization lemma ([21] p.152 , [35] p.108-111),

$$E^{\mathcal{F}_n}\Phi\left(\Delta_r^n\right) \leq E_{n,\epsilon}\Phi\left(2\|\rho_r^n\|_{\mathcal{G}}\right), \tag{4.8}$$

where Φ is nondecreasing and convex,

$$\rho_r^n(f) = \sum_{1 \le i_1, \ldots, i_{n+1} \le r} p_{i_1 \ldots i_n} \epsilon_{i_1 \ldots i_{n+1}} Y_{i_1 \ldots i_{n+1}} \int_{A^r_{i_1 \cdots i_{n+1}}} f \, d\lambda \,, \tag{4.9}$$

ϵ are Rademacher variables independent of everything and $E_{n,\epsilon}$ denotes expectation with respect to variables ϵ and $Y_{i_1 \ldots i_{n+1}}$, conditionally on $\mathcal{F}_n$.

Choosing $\Phi(x) = x^2$ in (4.8) we get

$$E^{\mathcal{F}_n} \left(\Delta_r^n\right)^2 \le 4 E_{n,\epsilon} \left(\|\rho_r^n\|_{\mathcal{G}}^2 \right). \tag{4.10}$$

Let ν_r^{n+1} be the measure defined by

$$\nu_r^{n+1}(f) = \sum_{1 \le i_1, \ldots, i_{n+1} \le r} p^2_{i_1 \ldots i_n} Y^2_{i_1 \ldots i_{n+1}} r^{-(n+1)} \int_{A^r_{i_1 \cdots i_{n+1}}} f \, d\lambda \,,$$

and let $\mathbb{P}_r^{n+1}$ be the probability measure determined by

$$\mathbb{P}_r^{n+1}(f) = \frac{\nu_r^{n+1}(f)}{\nu_r^{n+1}(1)}. \tag{4.11}$$

Notice that

$$E \nu_r^{n+1}(1) = \sigma_{n+1}^2 \, r^{-(n+1)} \tag{4.12}$$

and that from Proposition 3.1,

$$\lim_r r^{(n+1)} \nu_r^{n+1}(1) = \sigma_{n+1}^2 \qquad \text{a.s.} \, . \tag{4.13}$$

From Hoeffding's inequality, the process $\left\{ \frac{\rho_r^n(f)}{\sqrt{\nu_r^{n+1}(1)}} \, , \; f \in \mathcal{G} \right\}$ is subgaussian for the semi-norm $L_2(\mathbb{P}_r^{n+1})$ ([35] p. 100-101). Hence the maximal inequality ([35], Corollary 2.2.5 p. 98) yields

$$E^{\mathcal{F}_{n+1}} \left(\|\rho_r^n\|_{\mathcal{G}}^2 \right) \;\; \le \;\; K \, \nu_r^{n+1}(1) \, I_r^n(\mathcal{G}), \tag{4.14}$$

where K is an universal constant, and

$$I_r^n(\mathcal{G}) \;\; = \;\; \int_0^\infty \sqrt{H\left(x, \mathcal{G}, L_2(\mathbb{P}_r^{n+1})\right)} \, dx \, . \tag{4.15}$$

Looking back at (4.10), we get

$$E \left(\Delta_r^n\right)^2 \le 4 K E \left(\nu_r^{n+1}(1) I_r^n(\mathcal{G}) \right). \tag{4.16}$$

We obtain the conclusion (4.6) from (4.16), (4.14), (4.12) and the universal entropy condition (4.2). □

<u>Proof of Theorem 4.4.</u> By Proposition 4.6, $\sqrt{r}\|\mu_r^n - \mu_r^1\|_{\mathcal{G}}$ tends to 0 in L^2 as $r \to \infty$, so that Part a) is a consequence of Ziegler's theorem.

To prove Part b) (the case $n = \infty$), we note that from Proposition 4.6, for $r > 4(1 + \sigma^2)$,

$$\sum_{n=1}^{\infty} \sqrt{E \left(\Delta_r^n\right)^2} \; < \; \frac{2\sigma_1 \sqrt{K I_{\mathcal{G}}}}{r}. \tag{4.17}$$

Since the class is regular,

$$\text{a.s.} \;\; \|\mu_r^\infty - \mu_r^1\|_{\mathcal{G}} \le \sum_{n=1}^{\infty} \Delta_r^n \, . \tag{4.18}$$

It follows that

$$\sqrt{E \left(\|\mu_r^\infty - \mu_r^1\|_{\mathcal{G}}^2 \right)} \le \frac{2\sigma_1 \sqrt{K I_{\mathcal{G}}}}{r}. \tag{4.19}$$

Consequently, the FCLT for $n = \infty$ is also derived from Ziegler's theorem. □

4.2 Proof of Theorem 4.5

By the Cramér-Wold device (cf. [40], p.248 for more details), the convergence of finite dimensional marginals is a by-product of the following theorem.

Theorem 4.7 *Assume $EW^2 < \infty$ and $f \in \mathcal{L}^2([0,1],\lambda)$. For each fixed $d \geq 0$, as $r \to \infty$,*

$$\vec{\xi}_r(f) \Rightarrow \mathcal{N}\left(0, \lambda(f^2) Id_{d+1}\right).$$

In its proof we shall need the following result.

Proposition 4.8 *If $f \in \mathcal{L}^2([0,1],\lambda)$ and $0 \leq m < n \leq \infty$, then:*

$$r^{m+1} E(\mu_r^{m+1}(f) - \mu_r^m(f))^2 \leq \sigma_m^2 \lambda(f^2); \tag{4.20}$$

$$\lim_r r^{m+1} E(\mu_r^n(f) - \mu_r^m(f))^2 = \sigma_m^2 \lambda(f^2). \tag{4.21}$$

Proof. (4.20) is just Proposition 2.4 a) applied to $p = 2$. Let us prove (4.21). From (4.7) we get

$$E[(\mu_r^{n+1}(f) - \mu_r^n(f))^2 \mid \mathcal{F}_n] = \sum_{1 \leq i_1,\dots,i_{n+1} \leq r} p_{i_1,\dots,i_n}^2 \sigma^2 \Big[\int_{A^r_{i_1\cdots i_{n+1}}} f\, d\lambda\Big]^2, \tag{4.22}$$

so that

$$E\left(\mu_r^{n+1}(f) - \mu_r^n(f)\right)^2 = \sigma_n^2 \sum_{1 \leq i_1,\dots,i_{n+1} \leq r} \Big[\int_{A^r_{i_1\cdots i_{n+1}}} f d\lambda\Big]^2. \tag{4.23}$$

Now,

$$\lim_r r^n \sum_{1 \leq i_1,\dots,i_n \leq r} \Big[\int_{A^r_{i_1\cdots i_n}} f d\lambda\Big]^2 = \lambda(f^2). \tag{4.24}$$

In fact, if f is continuous, the mean value theorem allows us to write the left hand side of (4.24) as a Riemann-sum of f^2, so that (4.24) holds; it extends to $f \in \mathcal{L}^2([0,1],\lambda)$ by a density argument. This yields

$$\lim_r r^{n+1} E(\mu_r^{n+1}(f) - \mu_r^n(f))^2 = \sigma_n^2 \lambda(f^2). \tag{4.25}$$

Since $\{\mu_r^n(f)\}_{n\geq 0}$ is a martingale, if $0 \leq m < n < \infty$, then

$$E(\mu_r^n(f) - \mu_r^m(f))^2 = \sum_{k=m}^{n-1} E(\mu_r^{k+1}(f) - \mu_r^k(f))^2. \tag{4.26}$$

Letting $n \to \infty$, we see that the equality still holds for $n = \infty$. By (4.20),

$$\sum_{k=m+1}^{\infty} E(\mu_r^{k+1}(f) - \mu_r^k(f))^2 = O(r^{-m-2}).$$

Therefore the claim (4.21) follows from (4.25) and (4.26). □

Proof of Theorem 4.7. The conclusion for $d = 0$ was proved in [16, Theorem 4.18], and can be easily checked by Lindeberg's theorem, using the decomposition

$$r(\mu_r^1(f) - \lambda(f)) = \sum_{k=1}^{r} (W_k - 1) \int_{A^r_k} f d\lambda. \tag{4.27}$$

For $d > 0$, as above with (4.27), we want to get benefit of the decomposition (2.12) to apply a multidimensional version of the Lyapunov's theorem as stated by Yurinskii [38].

From (2.12) and (4.27) we can write

$$\vec{\xi}_r(f) = \sum_{k=1}^{r} \vec{\xi}_{r,k}(f)\,,$$

where the $\vec{\xi}_{r,k}(f)$ $(k = 1, \dots, r)$ are independent, centered random vectors. For $l = 1, \dots, d$, the l-th component of $\vec{\xi}_{r,k}(f)$ is

$$\left[\vec{\xi}_{r,k}(f)\right]_l \;:=\; r^{-1/2}\, W_k\, \xi_r^l \circ \Theta_k(\tau_k^r f) = r^{(l-1)/2} W_k \left(\mu_r^l - \mu_r^{l-1}\right) \circ \Theta_k(\tau_k^r f), \tag{4.28}$$

$$\left[\vec{\xi}_{r,k}(f)\right]_0 \;:=\; r^{-1/2}\,(W_k - 1)\,\lambda(\tau_k^r f)\,. \tag{4.29}$$

From Proposition 4.8, for each $l \geq 0$,

$$\lim_r \operatorname{Var}\, \xi_r^l(f) = \lambda(f^2)\,. \tag{4.30}$$

From the martingale property, for $i \neq l$,

$$E\left(\mu_r^i(f) - \mu_r^{i-1}(f)\right)\left(\mu_r^l(f) - \mu_r^{l-1}(f)\right) = 0. \tag{4.31}$$

We have then

$$\lim_r \operatorname{Cov}\, \vec{\xi}_r = \lambda(f^2)\mathrm{Id}_{d+1}\,.$$

To apply Theorem 1 of [38], we first assume that there exists $M > 0$ such that

$$W \leq M \quad \text{and} \quad |f| \leq M\,. \tag{4.32}$$

We have only to check that

$$\lim_r \sum_{k=1}^{r} E\|\vec{\xi}_{r,k}(f)\|^3 = 0\,, \tag{4.33}$$

where $\|.\|$ is the usual Euclidean norm in $\mathbb{R}^{d+1}$. It is equivalent to prove that

$$\lim_r \sum_{l=0}^{d} \sum_{k=1}^{r} E\left|\left[\vec{\xi}_{r,k}(f)\right]_l\right|^3 = 0\,. \tag{4.34}$$

By the definition (4.28) and independence,

$$E\left|\left[\vec{\xi}_{r,k}(f)\right]_l\right|^3 = r^{3(l-1)/2}\,\frac{EW^3}{\sigma_l^3}\; E\left|\left(\mu_r^l - \mu_r^{l-1}\right) \circ (\tau_k^r f)\right|^3\,.$$

Applying Proposition 2.4 a) with $p = 3$, we get

$$E\left|\left[\vec{\xi}_{r,k}(f)\right]_l\right|^3 \leq C r^{-3/2} \lambda\left(|\tau_k^r(f)|^3\right)\,, \tag{4.35}$$

where C is a constant independent of r and k. Since

$$\lambda\left(|\tau_k^r(f)|^3\right) = \int_{A_r^k} |f^3|\; d\lambda\,,$$

adding (4.35) over l and k yields

$$\sum_{l=0}^{d} \sum_{k=1}^{r} E\left|\left[\vec{\xi}_{r,k}(f)\right]_l\right|^3 \leq C r^{-1/2}\,,$$

where C is another constant. This ends the proof of (4.34). We are in the conditions of Yurinskii's theorem and the convergence to the Gaussian distribution holds true.
If (4.32) is not true, let $M > 0$ be arbitrary and $\bar{f} = (-M) \vee f \wedge M$. From (4.20),

$$E\|\vec{\xi}_r(f) - \vec{\xi}_r(\bar{f})\|^2 \leq (d+1)\ \lambda\left(f^2 \mathbf{1}_{|f|>M}\right). \tag{4.36}$$

Now, for $\overline{W} = W \wedge M$ we apply again the technique of successive truncations used in the proof of lemma 3.2.

Let $G_{r,0} = \mu_r^{n+1}(f) - \mu_r^n(f)$ and for $k = 1, \ldots, n$, let

$$G_{r,k} := \sum_{1 \leq i_1, \ldots, i_{n+1} \leq r} \overline{W}_{i_1} \ldots \overline{W}_{i_1 \ldots i_k} W_{i_1 \ldots i_{k+1}} \ldots W_{i_1 \ldots i_n} \left(W_{i_1 \ldots i_{n+1}} - 1\right) \int_{A^r_{i_1 \cdots i_{n+1}}} f\, d\lambda.$$

At last, set

$$G_{r,n+1} := \sum_{1 \leq i_1, \ldots, i_{n+1} \leq r} \overline{W}_{i_1} \ldots \overline{W}_{i_1 \ldots i_n} \left(\overline{W}_{i_1 \ldots i_{n+1}} - E(\overline{W})\right) \int_{A^r_{i_1 \cdots i_{n+1}}} f\, d\lambda.$$

From the usual conditioning method we get

$$E\left[G_{r,k} - G_{r,k+1}\right]^2 = \left[E\left(\overline{W}\right)^2\right]^k E\left[(W-M)^+\right]^2 \sigma^2_{n-k-1} J_r^n, \qquad k < n, \tag{4.37}$$

$$E\left[G_{r,n} - G_{r,n+1}\right]^2 = \left[E\left(\overline{W}\right)^2\right]^n \text{Var}\ (W-M)^+ J_r^n, \tag{4.38}$$

with

$$J_r^n := \sum_{1 \leq i_1, \ldots, i_{n+1} \leq r} \left(\int_{A^r_{i_1 \cdots i_{n+1}}} f\, d\lambda\right)^2.$$

Now, $E(\overline{W})^2 \leq EW^2$ and by the Cauchy-Schwarz inequality

$$\left[\int_{A^r_{i_1 \cdots i_{n+1}}} f d\lambda\right]^2 \leq \lambda(A^r_{i_1 \ldots i_{n+1}}) \int_{A^r_{i_1 \cdots i_{n+1}}} f^2 d\lambda, \tag{4.39}$$

so that we obtain for each $k = 0, \ldots, n$,

$$E\left[G_{r,k} - G_{r,k+1}\right]^2 \leq r^{-(n+1)} \frac{\sigma_n^2}{\sigma^2} E\left[(W-M)^+\right]^2 \lambda(f^2).$$

By the triangle inequality we get

$$E\left[G_{r,0} - G_{r,n+1}\right]^2 \leq r^{-(n+1)}(n+1)^2\ \frac{\sigma_n^2}{\sigma^2}\ E\left[(W-M)^+\right]^2 \lambda(f^2). \tag{4.40}$$

Starting from $\vec{\xi}_r$, let us replace each r.v. $W_{\mathbf{i}}$ by the corresponding $\overline{W}_{\mathbf{i}}$ except $W_{\mathbf{i}} - 1$ which we replace by $\overline{W}_{\mathbf{i}} - E\overline{W}$. Let us denote by $\vec{\eta}_r$ the new random vector. Then the previous inequality (4.40) becomes

$$E\left[\xi_r^n(f) - \frac{\bar{\sigma}_n}{\sigma_n}\eta_r^n(f)\right]^2 \leq \frac{(n+1)^2}{\sigma^2}\ E\left[(W-M)^+\right]^2 \lambda(f^2), \tag{4.41}$$

where for $n = 0, \ldots d$ we have set $\bar{\sigma}_n^2 = \left[E\left(\overline{W}^2\right)\right]^n \text{Var}\ \overline{W}$. Since $\lim_M \bar{\sigma}_n^2 = \sigma_n^2$, from (4.41), (4.30), (4.36) and the triangle inequality, we get

$$\lim_{M \to \infty} \sup_r\ E\left[\|\vec{\xi}_r(f) - \vec{\eta}_r(\bar{f})\|^2\right] = 0. \tag{4.42}$$

By the Yurinskii's theorem $\eta_r(f)$ (for fixed M) converges to the $(d+1)$-dimensional normal distribution with covariance matrix $\lambda(f^2)Id_{d+1}$. This completes the proof of Theorem 4.7. □

Let us end the proof of Theorem 4.5 by establishing tightness.
a) Let $\mathcal{G}$ be a uniformly bounded class and $\mathcal{G}_\delta = \{f-g; f,g \in \mathcal{G}, |f-g|_2 < \delta\}$. The tightness in $\ell^\infty(\mathcal{G})$ will be a consequence of

$$\lim_{\delta\downarrow 0}\limsup_{r\uparrow\infty} r^{\frac{n+1}{2}} E\left(\sup_{f\in\mathcal{G}_\delta} |\mu_r^{n+1}(f) - \mu_r^n(f)|\right) = 0\,. \tag{4.43}$$

From a classical lemma ([35] problem 5 p.93), it is enough to show that for every sequence $\delta_r \downarrow 0$

$$\limsup_{r\uparrow\infty} r^{\frac{n+1}{2}} E\left(\sup_{f\in\mathcal{G}_{\delta_r}} |\mu_r^{n+1}(f) - \mu_r^n(f)|\right) = 0. \tag{4.44}$$

Let

$$\theta_r^n = \sqrt{\sup_{f\in\mathcal{G}_{\delta_r}} \mathbb{P}_r^{n+1}(f^2)}.$$

For $x \geq \theta_r^n$, the set $\mathcal{G}_{\delta_r}$ fits in a single ball of radius x and $H\left(x, \mathcal{G}_{\delta_r}, L_2\left(\mathbb{P}_r^{n+1}\right)\right) = 0$, so that (4.15) becomes:

$$I_r^n\left(\mathcal{G}_{\delta_r}\right) = \int_0^{\theta_r^n} \sqrt{H\left(x, \mathcal{G}_{\delta_r}, L_2\left(\mathbb{P}_r^{n+1}\right)\right)}\, dx \tag{4.45}$$

Furthermore, we know ([35] p. 128) that

$$N\left(x, \mathcal{G}_{\delta_r}, L_2(Q)\right) \leq \left[N\left(x/2, \mathcal{G}, L_2(Q)\right)\right]^2 .$$

for every measure Q, which implies

$$I_r^n\left(\mathcal{G}_{\delta_r}\right) \leq 2\sqrt{2}\int_0^{\theta_r^n} \sup_Q \sqrt{H\left(x, \mathcal{G}, L_2(Q)\right)}\, dx\,, \tag{4.46}$$

hence,

$$\begin{aligned} r^{\frac{n+1}{2}} E\left(\sup_{f\in\mathcal{G}_{\delta_r}} |\mu_r^{n+1}(f) - \mu_r^n(f)|\right) &\leq \sqrt{E\left(\sup_{f\in\mathcal{G}_{\delta_r}} |\mu_r^{n+1}(f) - \mu_r^n(f)|\right)^2} \\ &\leq K'\sqrt{E\left[r^{(n+1)}\nu_r^{n+1}(1)\int_0^{\theta_r^n} \sup_Q \sqrt{H(x, \mathcal{G}, L_2(Q))}\, dx\right]}, \end{aligned}$$

where K' is a universal constant. From Proposition 3.1, the sequence $\{r^{n+1}\nu_r^{n+1}(1)\}_r$ is uniformly integrable and its a.s. limit is σ_{n+1}^2. To prove (4.44) it is then enough to show that $\theta_r^n \xrightarrow{P} 0$. Owing to (4.13), it is enough to prove that

$$r^{n+1}\sup_{f\in\mathcal{G}_{\delta_r}} \nu_r^{n+1}\left(f^2\right) \xrightarrow{P} 0. \tag{4.47}$$

Actually Lemma 3.2 gives us the a.s. convergence.

b) The tightness of the multivariate random process $\vec{\xi}_r$ follows then from the tightness of each component ξ_r^k, just proved in (4.43). □

5 Law of the Iterated Logarithm

In this section we establish a functional law of the iterated logarithm (LIL) for $\{\mu_r^n\}_r$ in the sense of Strassen. Let $\mathcal{K}$ denote the compact subset of functions F from $\mathcal{G}$ into the set of real numbers such that for some function g in the unit ball of $L^2([0,1])$,

$$F(f) = \int_0^1 f(t)g(t)dt \quad \text{for any } f \in \mathcal{G}\,.$$

We say that a sequence $\{\nu_r\}_r$ of centered random measures on $\ell^\infty(\mathcal{G})$ satisfies the functional LIL, if $\{(2r/LLr)^{1/2}\nu_r(f) : f \in \mathcal{G}\}$ is a.s. relatively compact with respect to the uniform metric over $\mathcal{G}$ and the set of limit points is exactly $\mathcal{K}$ a.s.. Here $LLr := \log\log r$ for $r \geq 3$.

As usual we need an assumption of stability by contractions ([4], [20]). For f in $\mathcal{G}$, and any t in $(0,1]$, the function $l_t f$ defined by $l_t f(x) = f(x/t)$ for x in $[0,t]$ and $l_t f(x) = 0$ for $x > t$. Let us define

$$\bar{\mathcal{G}} := \{l_t f \;:\; f \in \mathcal{G}, t \in [1/2, 1]\} \tag{5.1}$$

Theorem 5.1 *Assume that $EW^2 < \infty$ and that $\bar{\mathcal{G}}$ satisfies the entropy condition (Assumption 4.2). Let n be any positive integer. Then the sequence $\{\sigma^{-1}(\mu_r^n - \lambda)\}_r$ satisfies the functional LIL. If furthermore $\mathcal{G}$ is regular, then $\{\mu_r^\infty - \lambda\}_r$ satisfies the functional LIL.*

The main task is to obtain the functional LIL under the same integrability assumptions as in the FCLT. An important part of the proof is a reduction of the problem to the case $n = 1$ (subsections 5.4 -5.6.). It uses a new method of symmetrization detailed in subsections 5.2 -5.3.

For classes satisfying bracketing or local modulus conditions, Lacey [20] has proven the functional LIL for μ_r^1 under weakened assumptions on the integrability of the entropy. However the LIL seems to be unknown for function-indexed partial sum processes under random entropy conditions. Here for sake of simplicity, we restrict our attention to classes of functions with a finite entropy integral and prove in subsection 5.7. the functional LIL for μ_r^1.

We start by a reduction of the general case to the case $n = 2$.

5.1 Almost sure approximation for $n \geq 2$

From Proposition 4.6

$$P\Big(\Delta_r^n \geq a_{rn}\Big) \leq (C/r)^{n+1} a_{rn}^2. \tag{5.2}$$

Now, take $a_{rn} = a_{r2} r^{-(n-2)/4}$. Then

$$\sum_{r\geq 4C^2}\sum_{n\geq 2} P\Big(\Delta_r^n \geq a_{rn}\Big) \leq \sum_{r\geq 4C^2} a_{r2}^{-2} \sum_{n\geq 2} C^{n+1} r^{-(n+4)/2} \leq 2 \sum_{r\geq 4C^2} a_{r2}^{-2}(C/r)^3. \tag{5.3}$$

Next choosing $a_{r2} = r^{-3/4}$ we get via the Borel-Cantelli lemma that, for any $n \geq 2$,

$$\|\mu_r^n - \mu_r^2\|_{\mathcal{G}} = O(r^{-3/4}) \text{ a.s.}\,. \tag{5.4}$$

If $\mathcal{G}$ is regular, then for $r > r_0$,

$$\text{a.s. } \|\mu_r^\infty - \mu_r^2\|_{\mathcal{G}} \leq \sum_{n=2}^{\infty} \Delta_r^n\,, \tag{5.5}$$

and consequently (5.4) still holds for $n = \infty$.

5.2 A symmetrization method

In this subsection, we introduce a symmetrization method, which will be used in the forthcoming subsections to improve inequality (5.2) for $n = 1$.

Let U be a random variable with uniform distribution over $[0, 1]$ and F be a distribution function on the real line. We denote by F^{-1} the generalized inverse function of F.

Theorem 5.2 *Let $X = F^{-1}(U)$ and set*

$$Y = (F^{-1}(U) + F^{-1}(1 - U))/2 .$$

Then $X - Y$ is a symmetric random variable. Furthermore $2 \text{ Var } Y \le \text{ Var } X$.

We conjecture that Y has the minimal variance among those r.v.'s Z such that $X - Z$ is symmetric. To prove Theorem 5.2, we first note that

$$X - Y = (F^{-1}(U) - F^{-1}(1 - U))/2 .$$

Then the symmetry follows from the fact that $1 - U$ has the uniform distribution over $[0, 1]$. Next

$$2 \text{ Var } Y = \text{ Var } X + \text{ Cov } (F^{-1}(U), F^{-1}(1 - U)) \le \text{Var} X,$$

since the above covariance is nonpositive (see Fréchet [15]), which completes the proof.

5.3 A symmetric decomposition for real random variables

Throughout the section W is a real-valued random variable with finite variance. Our aim is to decompose $W - E(W)$ into a series of symmetric r.v. converging at a geometric rate. Let $(\eta_l)_l$ be a sequence of i.i.d. random variables with uniform distribution over $[0, 1]$, independent of W.

Let $W^0 = W$ and F_0 denote the d.f. of W^0. At the first step, from Skorohod lemma, we know that

$$U_0 = F_0(W^0 - 0) + \eta_0(F_0(W^0) - F_0(W^0 - 0))$$

has the uniform distribution over $[0, 1]$ and satisfies $F_0^{-1}(U_0) = W^0$ a.s.. If

$$W^1 = (F_0^{-1}(U_0) + F_0^{-1}(1 - U_0))/2 ,$$

then $(W^0 - W^1)$ is symmetric by Theorem 5.2 and

$$\text{Var } W^1 \le \frac{1}{2} \text{ Var } W^0.$$

Define then by induction, for any integer l the r.v. W^{l+1} from W^l and η_l exactly in the same way as at the first step: denoting by F_l the d.f. of W^l we set

$$U_l = F_l(W^l - 0) + \eta_l(F_l(W^l) - F_l(W^l - 0)) \tag{5.6}$$

and define

$$W^{l+1} = (F_l^{-1}(U_l) + F_l^{-1}(1 - U_l))/2 . \tag{5.7}$$

Then the r.v. W^l satisfy

$$E(W^l) = E(W^0) \quad \text{and} \quad \text{Var } W^l \le 2^{-l} \text{ Var } W^0 . \tag{5.8}$$

Furthermore, since

$$W^l - W^{l+1} = (F_l^{-1}(U_l) - F_l^{-1}(1 - U_l))/2 ,$$

we get that

$$\text{Var }(W^l - W^{l+1}) \le \text{ Var } W^l \le 2^{-l} \text{ Var } W^0 . \tag{5.9}$$

Hence

$$W - E(W) = \sum_{l=0}^{\infty}(W^l - W^{l+1}), \tag{5.10}$$

the series being geometrically convergent in L^2 and consequently a.s. convergent.

5.4 Decomposition of $\mu_r^1 - \mu_r^2$.

In this section we provide a decomposition for $\mu_r^2 - \mu_r^1$. Recall

$$\mu_r^2(f) - \mu_r^1(f) = \sum_{i=1}^{r}\sum_{j=1}^{r} W_i(W_{ij} - 1)\int_{A_{ij}^r} f d\lambda.$$

Now let (η_{ij}^l) be double arrays of independent r.v. with the uniform distribution over $[0,1]$, independent of the initial sequences of the multiplicative process. We then define the r.v. W_{ij}^l from W_{ij} and the auxiliary r.v. $(\eta_{ij}^l)_l$ exactly in the same way as in the previous section. Then the so defined W_{ij}^l satisfy

$$E(W_{ij}^l) = 1 \quad \text{and} \quad \text{Var } W_{ij}^l \le 2^{-l} \text{ Var } W_{ij}^0 . \tag{5.11}$$

Now we are in position to decompose the difference into sums of conditionally symmetric processes. From the above variance bounds the equality

$$\mu_r^2(f) - \mu_r^1(f) = \sum_{l=1}^{\infty}\sum_{i=1}^{r}\sum_{j=1}^{r} W_i(W_{ij}^{l-1} - W_{ij}^l)\int_{A_{ij}^r} f d\lambda \tag{5.12}$$

holds true a.s. and in L^2. Furthermore the r.v. $W_{ij}^{l-1} - W_{ij}^l$ have a symmetric distribution. Next we give exponential bounds related to this decomposition.

5.5 Exponential bounds

We start by a symmetrization procedure. From the symmetry of $(W_{ij}^{l-1} - W_{ij}^l)$, if

$$\varepsilon_{ij}^l = \text{sign}(W_{ij}^{l-1} - W_{ij}^l)$$

(we use the convention $\text{sign}(0) = 0$), then ε_{ij}^l is symmetric with values in $\{-1, 0, 1\}$ conditionally to $|W_{ij}^l - W_{ij}^{l-1}|$. Consequently the random process

$$\sum_{i=1}^{r}\sum_{j=1}^{r} W_i(W_{ij}^{l-1} - W_{ij}^l)\int_{A_{ij}^r} f d\lambda$$

when f varies in $\mathcal{G}$, is a Rademacher process, conditionnally to the σ-field $\mathcal{B}_l$ generated by the r.v. W_i and $T_{ij}^l = |W_{ij}^l - W_{ij}^{l-1}|$. Next Theorem 4.7 in Ledoux-Talagrand [21] applies: if

$$X_{rl} = \sup_{f\in\mathcal{G}}\left|\sum_{i=1}^{r}\sum_{j=1}^{r} W_i(W_{ij}^{l-1} - W_{ij}^l)\int_{A_{ij}^r} f d\lambda\right|,$$

and if M_{rl} denotes a conditional median of X_{rl}, then

$$P(X_{rl} \ge M_{rl} + \sigma_{rl}t \mid \mathcal{B}_l) \le \exp(-t^2/8), \tag{5.13}$$

where

$$\sigma_{rl}^2 = \sup_{f\in\mathcal{G}} \sum_{i=1}^{r}\sum_{j=1}^{r} W_i^2 (T_{ij}^l)^2 \Big(\int_{A_{ij}^r} f\,d\lambda\Big)^2 .$$

Let

$$\theta_{rl} = r^{-2}\sqrt{\sum_{i=1}^{r}\sum_{j=1}^{r} W_i^2 (T_{ij}^l)^2}. \tag{5.14}$$

From a similar argument as in the proof of (4.14)

$$M_{rl} \le 2E(X_{rl} \mid \mathcal{B}_l) \le C\theta_{rl},$$

for some constant C depending on the universal entropy integral of $\mathcal{G}$. Note then that, since the functions of $\mathcal{G}$ take their values in $[-1,1]$,

$$\sigma_{rl} \le \theta_{rl} = r^{-2}\sqrt{\sum_{i=1}^{r}\sum_{j=1}^{r} W_i^2 (T_{ij}^l)^2}. \tag{5.15}$$

Hence, taking $t = t_{rl} = 4(1+\log r + \log l)$ in inequality (5.13), we get via the Borel-Cantelli Lemma that

$$P(X_{rl} - \theta_{rl}(C + 4 + 4\log r + 4\log l) > 0 \text{ i.o. } (r,l)\ ) = 0. \tag{5.16}$$

5.6 Almost sure approximation of μ_r^2 by μ_r^1

In this subsection we derive a.s. convergence rates for $\mu_r^2 - \mu_r^1$ from the exponential bounds of the above subsection. In fact, it remains to bound up the r.v. θ_{rl}. Let $r_k = 2^k$. Clearly

$$\sup_{r\in]r_{k-1},r_k]} \theta_{rl} \le 4\theta_{r_k l}. \tag{5.17}$$

Now, by Chebycheff's inequality,

$$P(\theta_{rl} \ge x) \le (r^2x)^{-2}E(W^2)\sum_{i=1}^{r}\sum_{j=1}^{r} E((T_{ij}^l)^2).$$

Note then that, since

$$E(T_{ij}^l)^2 \le 2^{1-l}E(W^2),$$

for $x = x_{kl} = 2^{-(l+3k)/4}$,

$$P(\theta_{r_k l} \ge x) \le C2^{-l}(r_k x)^{-2}.$$

Together with the Borel-Cantelli Lemma and the monotonicity properties of θ_{rl}, it implies that

$$P(\theta_{rl} > 4Cr^{-3/4}2^{-l/2} \text{ i.o. } (r,l)\ ) = 0 \tag{5.18}$$

Combining this bound with (5.16), we finally get:

$$P(X_{rl} \ge CL(r)r^{-3/4}2^{-l/2}(1+\log l) > 0 \text{ i.o. } (r,l)\ ) = 0. \tag{5.19}$$

Here $L(r) = \max(1,\log r)$. Now, summing on l, we derive from the above a.s. bound and (5.12) that

$$\|\mu_r^2 - \mu_r^1\|_{\mathcal{G}} = O(r^{-3/4}\log r) \text{ a.s.}. \tag{5.20}$$

5.7 Tightness of the normalized version of μ_r^1

Let $\mathcal{G}_\delta$ be the class of differences $f-g$ with f, g in $\mathcal{G}$ and $\|f-g\|_2 \leq \delta$. If $r_k = 2^k$, then

$$\sup_{r \in]r_{k-1}, r_k]} \sqrt{\frac{r}{LLr}} \|\mu_r^1\|_{\mathcal{G}_\delta} \leq 2\sqrt{\frac{r_k}{LLr_k}} \|\mu_{r_k}^1\|_{\bar{\mathcal{G}}_\delta},$$

Hence to prove the tightness, it is enough to prove that, for some $\varepsilon(\delta)$ converging to 0 as δ tends to 0,

$$\limsup_{k\to\infty} \sqrt{\frac{r_k}{LLr_k}} \|\mu_{r_k}^1\|_{\bar{\mathcal{G}}_\delta} \leq \varepsilon(\delta) \quad \text{a.s.} \tag{5.21}$$

As in the previous sections, we will decompose μ_r^1 into a sum of symmetric processes. Let (η_i^l) be arrays of independent r.v. uniformly distributed over $[0,1]$, independent of the initial sequences of the multiplicative process. We then define the r.v. W_i^l from W_i and the auxiliary variables $(\eta_i^l)_l$ by means of our symmetric decomposition. Then, for any g in $\bar{\mathcal{G}}_\delta$,

$$\mu_r^1(g) = \sum_{i=1}^r \sum_{l=1}^\infty (W_i^{l-1} - W_i^l) \int_{A_i^r} g d\lambda.$$

Let m be an arbitrary integer. We set

$$\bar{\mu}_r^1(g) = \sum_{i=1}^r \sum_{l=1}^m (W_i^{l-1} - W_i^l) \int_{A_i^r} g d\lambda.$$

The first step is to bound up $\|\mu_r^1 - \bar{\mu}_r^1\|_{\bar{\mathcal{G}}_\delta}$ again via Theorem 4.7 in [21]. Set

$$T_i^l = |W_i^l - W_i^{l-1}| \quad \text{and} \quad s_{rl} = r^{-1}\sqrt{\sum_{i=1}^r (T_i^l)^2}.$$

Let

$$Y_{rl} = \sup_{f\in\bar{\mathcal{G}}} \Big| \sum_{i=1}^r (W_i^{l-1} - W_i^l) \int_{A_i^r} f d\lambda \Big|.$$

With the same arguments as in the subsection on exponential bounds, we get that

$$P(Y_{rl} - s_{rl}(C+t) > 0) \leq \exp(-t^2/8), \tag{5.22}$$

where C is some constant depending only on the universal entropy integral of $\mathcal{G}$. Now

$$\|\mu_r^1 - \bar{\mu}_r^1\|_{\bar{\mathcal{G}}_\delta} \leq 2\sum_{l>m} Y_{rl}.$$

Hence taking $r = r_k$ and $t = t_{kl} = 4\sqrt{\log kl}$ in the above exponential bound, from the Borel-Cantelli Lemma

$$P\Big(\|\mu_{r_k}^1 - \bar{\mu}_{r_k}^1\|_{\bar{\mathcal{G}}_\delta} - 8\sum_{l>m} s_{r_k l}(C + \sqrt{\log kl}) > 0 \text{ i.o. } k \Big) = 0. \tag{5.23}$$

Now it remains to give an a.s. bound for the sum of the series

$$\sum_{l>m} s_{r_k l}(C + \sqrt{\log kl}).$$

By the Cauchy-Schwarz inequality,

$$\sum_{l>m} s_{r_k l}(C + \sqrt{\log kl}) \leq m^{-1/2} \sqrt{\sum_{l>m} s_{r_k l}^2 l(l-1)(C + \sqrt{\log kl})^2}.$$

Consequently, for $k \geq 3$ and $m \geq \exp(C^2)$,

$$\sum_{l>m} s_{r_k l}(C + \sqrt{\log kl}) \leq 2m^{-1/2}\sqrt{\log k}\sqrt{\sum_{l>m} s_{r_k l}^2 l(l-1)\log l}. \tag{5.24}$$

If $R_i^m = \sum_{l>m}(T_i^l)^2 l(l-1)\log l$, then $\sum_{l>m} s_{r_k l}^2 l(l-1)\log l = r^{-2}\sum_{i=1}^r R_i^m$.

From the above decomposition method into symmetric variables, the nonnegative variables $(R_i^m)_i$ are i.i.d.. Furthermore

$$E(R_i^m) \leq \sum_{l>m} 2^{-l} l(l-1)\log l \ E(W_i^2) \leq 16E(W^2).$$

and R_i^m is integrable. Hence, by the SLLN,

$$P(r^{-1}\sum_{i=1}^r R_i^m > 25E(W^2) \text{ i.o. } r) = 0$$

It follows that, for $k \geq 3$ and $m \geq \exp(C^2)$,

$$P(\sum_{l>m} s_{r_k l}(C + \sqrt{\log kl}) > 10\|W\|_2 m^{-1/2} 2^{-k/2}\sqrt{\log k} \text{ i.o. } k) = 0 \tag{5.25}$$

Combining (5.23) and (5.25), we finally obtain:

$$P\Big(\|\mu_{r_k}^1 - \bar{\mu}_{r_k}^1\|_{\bar{\mathcal{G}}_\delta} > 80\|W\|_2 m^{-1/2} 2^{-k/2}\sqrt{\log k} \text{ i.o. } k\Big) = 0. \tag{5.26}$$

Now we bound up the truncated version $\|\bar{\mu}_{r_k}^1\|_{\mathcal{G}_\delta}$. First, if

$$\varepsilon_i^l = \text{sign } (W_i^{l-1} - W_i^l),$$

then

$$\|\bar{\mu}_r^1\|_{\bar{\mathcal{G}}_\delta} \leq \sum_{l=1}^m \sup_{g \in \bar{\mathcal{G}}_\delta}\Big|\sum_{i=1}^r \varepsilon_i^l T_i^l \int_{A_i^r} g d\lambda\Big| =: \sum_{l=1}^m D_{rl}.$$

Again, by Theorem 4.7 in Ledoux-Talagrand [21], if $\mathcal{A}_l$ is the σ-field generated by the r.v. $(T_i^l)_i$, then

$$P(D_{rl} \geq 2M_{rl} + t\sigma_{rl} \mid \mathcal{A}_l) \leq \exp(-t^2/8), \tag{5.27}$$

where M_{rl} is a conditional median of Y_{rl} and

$$\sigma_{rl}^2 = \sup_{g \in \bar{\mathcal{G}}_\delta}\sum_{i=1}^r (T_i^l)^2\Big(\int_{A_i^r} g d\lambda\Big)^2.$$

Again from a similar argument as in the proof of (4.14)

$$M_{rl} \leq 2Cr^{-1}\sqrt{\sum_{i=1}^r (T_i^l)^2} = 2Cs_{rl},$$

for some constant C depending on the universal entropy integral of $\mathcal{G}$. Now we bound up σ_{rl}. Note that the elements of $\bar{\mathcal{G}}_\delta$ take their values in $[-2, 2]$ and have a L^2-norm less than δ. It follows that

$$\sigma_{rl}^2 \leq \max\{a_1(T_1^l)^2 + \cdots + a_r(T_r^l)^2 : 0 \leq a_i \leq 4/r, \sum a_i \leq \delta\}.$$

Clearly, in the above expression, we may replace the variables (T_i^l) by the order statistics $(T_{(i)}^l)$. Since the above expression is linear and increasing in the r.v. (a_i) the maximum is obtained when

$(a_1, a_2, \ldots, a_r)$ is some upper vertex of the polygon defined by the inequations $0 \leq a_i \leq 4/r$, $\sum a_i \leq \delta$. Then either $a_i = 0$ or $a_i = 4/r$ excepted for at most one point, and $\sum a_i = \delta$. Since we consider the order statistics, it is then easy to check that the maximum is achieved for $a_1 = 4/r, \ldots, a_p = 4/r$, $a_{p+1} = \delta - 4p/r$ and $a_i = 0$ for $i > p+1$, where $p = [\delta r/4]$. Hence we get that

$$\sigma_{rl}^2 \leq q_{rl}^2 = 4r^{-1} \int_0^{\delta/4} F_{rl}^{-1}(u)du$$

where F_{rl} is the empirical distribution function of the r-sample $(T_1^l)^2, \ldots, (T_r^l)^2$. Therefore, taking $r = r_k$ and $t = 4\sqrt{\log k}$ in the exponential inequality (5.27) we get via the Borel-Cantelli Lemma that

$$P(D_{r_k l} > 4Cs_{r_k l} + 4\sqrt{\log k}q_{r_k l} \text{ i.o. } k) = 0. \tag{5.28}$$

Now, by the SLLN for quantile processes in Van Zwet [39] (see also [1]), $\sqrt{r}q_{rl}$ converges a.s. to

$$\varepsilon_l(\delta) = \Big(4\int_0^{\delta/4} G_l^{-1}(u)du\Big)^{1/2},$$

where G_l^{-1} denotes the inverse distribution function of the r.v. T_i^l. Furthermore by the usual SLLN, rs_{rl}^2 converges to $E(T_1^l)$ a.s.. Hence

$$P(2^{k/2}D_{r_k l} > 4C\|W\|_2 2^{(1-l)/2} + 4\sqrt{\log k}\ \varepsilon_l(\delta) \text{ i.o. } k) = 0. \tag{5.29}$$

Now, summing on l in $[1, m]$, we get that

$$P\left(2^{k/2}\|\bar{\mu}_{r_k}^1\|_{\bar{\mathcal{G}}_\delta} > 16C\|W\|_2 + 4\sqrt{\log k}\left(\sum_{l=1}^{m}\varepsilon_l(\delta)\right) \text{ i.o. } k\right) = 0 \tag{5.30}$$

Finally the tightness property (5.21) follows easily from both (5.26) and (5.30).

5.8 Set of limit points

To obtain the set of limit points, we proceed as in [4], Sections 4 and 5. Throughout the subsection, let $M = \sup\{\|f\|_\infty : f \in \mathcal{G}\}$. For ε in $(0, 1/2)$, we define the sequence (q_k) of positive integers by $q_k = [\varepsilon^{-2}(1+\varepsilon)^k]$. Set $\varphi_k = (q_k/LLq_k)^{1/2}$. The first step is to give the limit points of the sequence of random functions $\{\varphi_k(\mu_{q_k}^1 - \lambda)(f) : f \in \mathcal{G}\}$.

Since $\mathcal{G}$ is relatively compact in $L^2([0,1])$, there exist some finite collection $\mathcal{A}(\delta)$ of elements of $\mathcal{G}$ and some measurable mapping Π_δ from $\mathcal{G}$ into $\mathcal{A}(\delta)$ such that $|g - \Pi_\delta g|_2 \leq \delta$ for any g in $\mathcal{G}$. Now by a density argument, for each g in $\mathcal{G}_\delta$ there exists some Lipschitzian Πg with values in $[-M, M]$ such that $\|g - \Pi g\|_2 \leq \delta$. Then, as in [4], section 4, from the Bernstein inequality,

$$\text{a.s. } \limsup_k \ \varphi_k \sup_{g \in \mathcal{A}(\delta)} |(\mu_{q_k}^1 - \lambda)(g - \Pi g)| \leq 2\delta.$$

The tightness property (5.21) gives then

$$\text{a.s. } \limsup_k \ \varphi_k \sup_{f \in \mathcal{G}} |(\mu_{q_k}^1 - \lambda)(f - \Pi(\Pi_\delta f))| \leq \varepsilon(\delta) + 2\delta. \tag{5.31}$$

Let M_δ denote the maximum of Lipschitz moduli of Πg for $g \in \mathcal{A}(\delta)$. By the Strassen invariance principle [33], there exists some Brownian motion $(B_t)_{t\geq 0}$ such that a.s.

$$\Delta(\mu_r^1, B) = \sup_{t \in [0,1]} |\mu_r^1([0,t] - t - r^{-1}B_{rt}| = o((LLr/r)^{1/2}).$$

Now, since the functions Πg are M_δ-Lipschitzian, integrating by parts, we have

$$\sup_{g\in\mathcal{A}(\delta)} |(\mu_r^1-\lambda)(\Pi g) - r^{-1}\int_0^1 \Pi g(t)dB_{rt}| \le M_\delta \Delta(\mu_r^1, B),$$

so that the strong invariance principle holds for the finite class of regular functions $\mathcal{L}(\delta) = \Pi(\mathcal{A}(\delta))$. Consequently the limit points of the processes $\{\sigma^{-1}\sqrt{r/LLr}(\mu_r^1-\lambda)(g) : g \in \mathcal{L}(\delta)\}$ are exactly Strassen's limit points.

Next, proceeding as in [4], Sections 4 and 5, we infer from (5.31) that the set of limit points of the sequence $(\sigma^{-1}\varphi_k(\mu_{q_k}^1 - \lambda)(f))_{f\in\mathcal{G}}$ is contained in $\mathcal{K}^{15\varepsilon}$ and is 30ε-dense in $\mathcal{K}^{15\varepsilon}$.

Now, for F in $\mathcal{K}$,

$$F(f) = \int_0^1 f(x)g(x)dx\,,$$

and t in $(0,1]$ let

$$F_t(f) = \int_0^1 f(x)tg(tx)dx\,.$$

To complete the proof, one can then use the trick in [4], pages 603-604 to ensure that the random functions $\sigma^{-1}\sqrt{r/LLr}(\mu_r^1-\lambda)$ for r in $(q_k, q_{k+1}]$, are near from the functions $F_{r/q_{k+1}}$ for F in $\mathcal{K}^{15\epsilon}$ near from the random function $(\sigma^{-1}\varphi_k(\mu_{q_k}^1-\lambda))$. □

6 Large Deviation Principles

We shall prove large deviation principles satisfied by $(\mu_r^n)_r$ with fixed $n \le \infty$. Each μ_r^n can be considered as an element of the set $\mathcal{M}$ of Borel measures acting on the class of continuous functions (weak topology) or on bounded measurable functions (strong topology).

As in the previous sections, let $\mathcal{B}$ be the Borel σ-field on $[0,1]$. For $A\in\mathcal{B}$, let p_A be the map $\mathcal{M}\to \mathbb{R}^+$ defined by $p_A(\nu) = \nu(A)$. The strong topology or $\tau-$topology on $\mathcal{M}$ is the coarsest topology which makes continuous the maps $p_A,\ A\in\mathcal{B}$. It turns $\mathcal{M}$ into a gauge space with the separating family $\mathcal{D} := \{d_A\}_{A\in B}$ of pseudo-metrics, where $d_A(\nu,\xi) :=\mid \nu(A)-\xi(A)\mid$ for $\nu,\xi\in\mathcal{M}$. We equip $\mathcal{M}$ with $\mathbf{B}$, the σ-field generated by $p_A, A\in\mathcal{B}$. We denote by bm the set of bounded measurable functions on $[0,1]$.

Definition 6.1 *We say that a sequence of probability measures (P_r) on a regular Hausdorff space $(\mathcal{X},\mathcal{B})$ ($\mathcal{B}$ is the Borel σ-field) satisfies a Large Deviation Principle (LDP) with a rate function $I : \mathcal{X}\to[0,+\infty]$ if I is a lower semicontinuous function (i.e. for all $a\in[0,\infty)$, the level set $\{x : I(x)\le a\}$ is a closed subset of X) such that:*

*(a) (**upper bound**) for any closed set F of $\mathcal{X}$,*

$$\limsup_{r\to\infty} \frac{1}{r}\log P_r(F) \le -\inf_{x\in F} I(x);$$

*(b) (**lower bound**) for any open set G of $\mathcal{X}$,*

$$-\inf_{x\in G} I(x) \le \liminf_{r\to\infty}\frac{1}{r}\log P_r(G).$$

I is called a good rate function if its level sets are compact subsets of $\mathcal{X}$.

We say that a sequence of random elements in $(\mathcal{X},\mathcal{B})$ satisfies a LDP if so does the sequence of their distributions.

Let Λ be the cumulant generating function (c.g.f) of W, defined for $t \in \mathbb{R}$ by

$$\Lambda(t) = \log E \exp(tW) \in (-\infty, +\infty],$$

and let Λ^* be its Fenchel-Legendre dual defined for $x \in \mathbb{R}$ by

$$\Lambda^*(x) = \sup_t \{tx - \Lambda(t)\} \in [0, +\infty].$$

As was mentioned at the end of Section 1, we shall need the assumption

$$E \exp(tW) < \infty \quad \text{for all } t \in \mathbb{R} \tag{A_1}$$

for $n < \infty$, and

$$\overline{w} := \text{ess sup } W < \infty \tag{A_∞}$$

for $n = \infty$. In [26] (Theorem 1.4) we proved that under (A_∞), the family $(Z_r^\infty)_{r\geq 2}$ satisfies the LDP in $\mathbb{R}$ with good rate function Λ^*. Here we establish LDP for the sequence of random measures $(\mu_r^n)_r$ for each fixed $n \leq \infty$.

6.1 The LDP for $(\mu_r^1)_r$ in the strong topology

The LDP for $(\mu_r^1)_r$ is related to the so-called Mogulskii-type theorems. In those results, both measures μ_r^1 and $\tilde{\mu}_r^1 = \frac{1}{r}\sum_{k=1}^r W_k \delta_{k/r}$ are studied via their distribution functions

$$M_r(t) = \mu_r^1([0,t]), \quad \tilde{M}_r(t) = \tilde{\mu}_r^1([0,t]) \in [0,1].$$

Under Assumption (A_1), it is known ([12] Section 5.2) that both $(M_r)_r$ and $(\tilde{M}_r)_r$ satisfy a LDP in $\mathcal{C}_0([0,1])$ equipped with the uniform topology, with the good rate function

$$I^{\uparrow}(\phi) = \begin{cases} \int_0^1 \Lambda^*(\dot{\phi}(t))\, dt & \text{if } \phi \in \mathcal{AC},\ \phi(0) = 0 \\ +\infty & \text{otherwise} \end{cases} \tag{6.1}$$

where $\mathcal{AC}$ denotes the space of absolute continuous functions. We deduce the following result for $(\mu_r^1)_r$.

Theorem 6.2 *If (A_1) holds, then the sequence $(\mu_r^1)_r$ satisfies the LDP on $\mathcal{M}$ equipped with the strong topology, with the convex good rate function I with the convex good rate function I :*

$$I(\nu) = \begin{cases} \int_0^1 \Lambda^*(\dot{\nu})\, d\lambda & \textit{if } \dot{\nu} = \frac{d\nu}{d\lambda} \textit{ exists} \\ +\infty & \textit{otherwise.} \end{cases} \tag{6.2}$$

Remark: Recently, Najim [29] proved that a LDP holds in the weak topology under the weaker assumption $E\exp(tW) < \infty$ for some $t > 0$, with a good rate function $\tilde{I}$ involving the singular part of ν.

Proof. Let $\mathcal{C}_0^{\uparrow}$ be the set of real continuous non decreasing functions of $[0,1]$ vanishing in 0. This is a closed subset of $\mathcal{C}$ and of course $P(M_r \in \mathcal{C}_0^{\uparrow}) = 1$. From the above result and Lemma 4.1.5 of [12], we see that $(M_r)_r$ satisfies the LDP in $\mathcal{C}_0^{\uparrow}$ equipped with the induced norm. Now, the natural application $\mathcal{C}_0^{\uparrow} \to \mathcal{M}$ which maps f to the measure ν such that $\nu((a,b]) = f(b) - f(a)$, is injective and continuous, so by the contraction principle, ([12] Theorem 4.2.1) we see that $\left(\mu_r^1\right)_r$ satisfies the LDP in $\mathcal{M}$ (with the weak topology) with the rate function I. To extend the result to the strong topology, it is enough to prove that $(\mu_r^1)_r$ is τ-exponentially tight. Since Λ is finite everywhere, then

$$\lim_{x\to+\infty} \frac{\Lambda^*(x)}{x} = \infty. \tag{6.3}$$

From [12] Lemma 6.2.16, this implies that I is good for the τ-topology, i.e. the sets $\Psi_I(\alpha) = \{\nu \in \mathcal{M} : I(\nu) \leq \alpha\}$ are compact. Actually

$$P(\mu_r^1 \in \Psi_I(\alpha)^c) = P(\sum_{k=1}^{r} \Lambda^*(W_k) > r\alpha).$$

From [12] Lemma 5.1.14, $E\exp(\gamma\Lambda^*(W)) < \infty$ for every $\gamma < 1$. By Markov's exponential inequality, this yields

$$\frac{1}{r}\log P\left(\sum_{k=1}^{r} \Lambda^*(W_k) > r\alpha\right) \leq -\frac{\alpha}{2} + \log Ee^{\Lambda^*(W)/2} \tag{6.4}$$

and then

$$\lim_{\alpha\to\infty} \limsup_r \frac{1}{r}\log P(\mu_r^1 \in \Psi_I(\alpha)^c) = -\infty\,, \tag{6.5}$$

and the exponential tightness is proved. □

6.2 The LDP for $(\mu_r^n)_r$, with $n \geq 2$, in the strong topology

Theorem 6.3 *a) For each fixed $2 \leq n < \infty$, if (A_1) holds, then the sequence $(\mu_r^n)_r$ satisfies the LDP on $\mathcal{M}$ equipped with the strong topology, with the good rate function I.*

b) If (A_∞) holds, then the sequence $(\mu_r^\infty)_r$ satisfies the LDP on $\mathcal{M}$ equipped with the strong topology, with the good rate function I.

Remark: For $n \geq 2$, note that the existence of the Laplace transform of the random variables $\mu_r^n(f)$ needs the much stronger assumption $E\exp(\theta W^n) < \infty$. Yet, the normalizing factors r^n in the multiplicative cascades at level n make it possible to get LDP without exponential moments.

To prove this theorem for $n > 1$, we will introduce a family of exponentially good approximations in the sense of [12] 4.2.2. The key point in our proof is that the approximations of the random measures (μ_r^n) will have the same LD rate function. This is the reason why the random variables $W_{\mathbf{i}}$ are not truncated at level 1, and the truncated variables $\widetilde{W}_{\mathbf{i}}$ below are normalized in such a way that $E\widetilde{W}_{\mathbf{i}} = 1$.

For $M > 0$, let $E_M = E(W \wedge M)$ and $\widetilde{W}_{\mathbf{i}}^M = (W_{\mathbf{i}} \wedge M)/E_M$ for $\mathbf{i} \in \mathbb{U}$. For $n \geq 2$, let

$$\mu_r^{n,M}(f) = \sum_{1\leq i_1,\ldots,i_n\leq r} W_{i_1}\widetilde{W}_{i_1 i_2}^M \ldots \widetilde{W}_{i_1\ldots i_n}^M \int_{A_{i_1\ldots i_n}^r} f\, d\lambda\,. \tag{6.6}$$

We denote by convention $\mu_r^{1,M} = \mu_r^1$.

Proposition 6.4 *Let $\delta > 0$ and $n \geq 2$.*

a) Under Assumption (A_1),

$$\limsup_{M\to\infty} \limsup_r \frac{1}{r}\log P\left(\sup_{||f||_\infty\leq 1} |\mu_r^n(f) - \mu_r^{n,M}(f)| > \delta\right) = -\infty\,. \tag{6.7}$$

b) Under Assumption (A_1), for every $f \in bm$ and $M > 0$,

$$\limsup_r \frac{1}{r}\log P(|\,\mu_r^{n,M}(f) - \mu_r^1(f)\,| > \delta) = -\infty\,. \tag{6.8}$$

Proof of Theorem 6.3

a) By Proposition 6.4, for every $f \in bm$ and every $\delta > 0$, we have

$$\limsup_r \frac{1}{r} \log P\left(\mid \mu_r^n(f) - \mu_r^1(f) \mid > \delta\right) = -\infty . \tag{6.9}$$

From [13] Theorem 1.6 and Corollary 1.10 p.236 (see also the first paragraph of p.240), this implies that $(\mu_r^n)_r$ and $(\mu_r^1)_r$ are exponentially equivalent in the strong topology. Since $(\mu_r^1)_r$ satisfies the LDP by Theorem 6.2, we conclude that so does $(\mu_r^n)_r$.

b) Since $\bar{w} \geq 1$, we have $|W-1| \leq \bar{w}$ and $\|f\|_\infty \leq 1$, whence, by (3.16),

$$P(|\mu_r^k(f) - \mu_r^{k-1}(f)| > \delta) \leq 2 \exp\left(-\frac{\delta^2 r^k}{2\bar{w}^{2k}}\right), \tag{6.10}$$

for each $1 \leq k < \infty$. This yields

$$\begin{aligned} P(|\mu_r^\infty(f) - \mu_r^1(f)| > \delta) &\leq P\left(\sum_{k=2}^{\infty} |\mu_r^k(f) - \mu_r^{k-1}(f)| > \delta\right) \\ &\leq \sum_{k=2}^{\infty} P\left(|\mu_r^k(f) - \mu_r^{k-1}(f)| > \delta 2^{-k+1}\right) \leq C \exp\left(-\frac{\delta^2 r^2}{8\bar{w}^4}\right) \end{aligned}$$

for r large enough, where $C > 0$ is a constant independent of r (one can choose $C = 2 + \delta^{-2}$). This gives

$$\limsup_r \frac{1}{r} \log P(|\mu_r^\infty(f) - \mu_r^1(f)| > \delta) = -\infty , \tag{6.11}$$

so that $(\mu_r^\infty)_r$ and $(\mu_r^1)_r$ are exponentially equivalent and we end the proof as in a). □

Using Proposition 6.4 b) and arguing as before we get also:

Corollary 6.5 *For every $n \geq 1$, the sequence $\left(\mu_r^{n,M}\right)_r$ satisfies the LDP with the good rate function I.*

Proof of Proposition 6.4

a) We start with

$$\sup_{\|f\|_\infty \leq 1} |\mu_r^n(f) - \mu_r^{n,M}(f)| \leq \tilde{D}_r^n(M), \tag{6.12}$$

where

$$\tilde{D}_r^n(M) = r^{-n} \sum_{1 \leq i_1, \ldots, i_n \leq r} W_{i_1} |W_{i_1 i_2} \ldots W_{i_1 \ldots i_n} - \widetilde{W}_{i_1 i_2}^M \ldots \widetilde{W}_{i_1 \ldots i_n}^M|, \qquad n \geq 2.$$

It is then enough to apply the following lemma whose proof is postponed.

Lemma 6.6 *Under Assumption (A_1), for $2 \leq n < \infty$ and any $\delta > 0$,*

$$\lim_{M \to \infty} \limsup_r \frac{1}{r} \log P\left(\tilde{D}_r^n(M) > \delta\right) = -\infty . \tag{6.13}$$

b) By the triangle inequality, it is enough to prove that for every $n \geq 2$ (δ and f as above),

$$\limsup_r \frac{1}{r} \log P(\mid \mu_r^{n,M}(f) - \mu_r^{n-1,M}(f) \mid > \delta) = -\infty . \tag{6.14}$$

Actually, for $n \geq 2$

$$\mu_r^{n,M}(f) - \mu_r^{n-1,M}(f) = \sum_{1 \leq i_1,\ldots,i_n \leq r} W_{i_1} \widetilde{W}^M_{i_1 i_2} \ldots \widetilde{W}^M_{i_1 \ldots i_{n-1}} \left(\widetilde{W}^M_{i_1 \ldots i_n} - 1 \right) \int_{A^r_{i_1 \ldots i_n}} f \, d\lambda. \qquad (6.15)$$

Denoting

$$Q_{i_1} = P\left(\mid r \sum_{1 \leq i_2,\ldots,i_n \leq r} \widetilde{W}^M_{i_1 i_2} \ldots \widetilde{W}^M_{i_1 \ldots i_{n-1}} \left(\widetilde{W}^M_{i_1 \ldots i_n} - 1 \right) \int_{A^r_{i_1 \ldots i_n}} f \, d\lambda \mid > \epsilon \right), \quad i_1 = 1, .., r, \quad (6.16)$$

the union of events bound gives

$$P\left(\mid \mu_r^{n,M}(f) - \mu_r^{n-1,M}(f) \mid > \delta \right) \leq P(Z_r^1 > \delta/\epsilon) + \sum_{i_1=1}^{r} Q_{i_1}. \qquad (6.17)$$

By the Chernov bound, we have for $\delta > \epsilon$,

$$P(Z_r^1 > \delta/\epsilon) \leq e^{-r\Lambda^*(\delta/\epsilon)}. \qquad (6.18)$$

For every $M > 0$, the variables $\widetilde{W}_{\mathbf{i}}$ and $\widetilde{W}_{\mathbf{i}} - 1$ are bounded by some $K_M > 0$, hence by (3.16),

$$Q_{i_1} \leq 2 \exp\left(-\frac{\epsilon^2 r^{n-1}}{2(K_M)^{2n-2}} \right) \qquad (6.19)$$

Inequalities (6.18) and (6.19) yield

$$\limsup_r \frac{1}{r} \log P(\mid \mu_r^{n,M}(f) - \mu_r^{n-1,M}(f) \mid > \delta) \leq -\Lambda^*(\delta/\epsilon) \qquad (6.20)$$

for every $\epsilon < \delta$. Letting $\epsilon \to 0$, we get (6.14) for $n \geq 3$.

For the case $n = 2$, we apply Hoeffding's inequality (3.13) to the variable $X = W \wedge M$ to get

$$\log E \exp\left[t\left(W \wedge M - E_M\right)\right] \leq \left(\frac{t^2 M^2}{8} \wedge |t| M \right). \qquad (6.21)$$

By conditioning and independence, we have for every real θ,

$$\frac{1}{r} \log E \exp\left[\theta r \sum_{1 \leq i_1, i_2 \leq r} W_{i_1} \left(W_{i_1 i_2} \wedge M - E_M \right) \int_{A^r_{i_1 i_2}} f \, d\lambda \right] \leq \log E \exp\left(\frac{\theta^2 M^2 W^2}{8r} \wedge |\theta| M W \right),$$

which, by the dominated convergence theorem, leads to

$$\limsup_r \frac{1}{r} \log E \exp\left[\theta r \sum_{1 \leq i_1, i_2 \leq r} W_{i_1} \left(W_{i_1 i_2} \wedge M - E_M \right) \int_{A^r_{i_1 i_2}} f \, d\lambda \right] \leq 0. \qquad (6.22)$$

Therefore (6.14) for $n = 2$ follows applying Markov's exponential inequality and letting $\theta \to +\infty$.
□

Proof of Lemma 6.6 In a first step we will prove that if

$$D_r^n(M) = r^{-n} \sum_{1 \leq i_1,\ldots,i_n \leq r} |W_{i_1} \ldots W_{i_1 \ldots i_n} - \widetilde{W}^M_{i_1} \ldots \widetilde{W}^M_{i_1 \ldots i_n}|, \qquad n \geq 1,$$

then for each $1 \le n < \infty$ and $\delta > 0$,

$$\lim_{M\to\infty} \limsup_r \frac{1}{r} \log P\left(D_r^n(M) > \delta\right) = -\infty\,. \tag{6.23}$$

Since $E_M \uparrow 1$ as $M \uparrow \infty$, we have $E_M \ge 1/2$ for $M > M_0$ (say). To prove (6.23) and (6.13) we will take $M > M_0$. We first consider the case $n = 1$. Let

$$c_M = E|W - \widetilde{W}^M| \ \text{ and } \ \Lambda_M(\theta) := \log E\left[\exp\left(\theta|W - \widetilde{W}^M|\right)\right].$$

By the Chernov bound, we have for $\delta > c_M$,

$$P(D_r^1(M) > \delta) \le e^{-r\Lambda_M^*(\delta)}\,, \tag{6.24}$$

where Λ_M^* is the Legendre dual of Λ_M.

Since $\lim_M \widetilde{W}^M = W$ a.s. and $|W - \widetilde{W}^M| \le W$ for $M > M_0$, we have, by (A_1) and the dominated convergence theorem,

$$\lim_M c_M = 0 \ \text{ and } \ \lim_M \Lambda_M(\theta) = 0 \ \text{ for every } \ \theta \in \mathbb{R}\,.$$

Then for fixed $x > c_M$ and every $\theta > 0$, we have $\Lambda_M^*(x) > \theta x - \Lambda_M(\theta)$, so that

$$\lim_M \Lambda_M^*(x) = +\infty \quad \text{for every } x > 0\,, \tag{6.25}$$

Together with (6.24), this gives (6.23) for $n = 1$.

We now proceed by induction on n. The induction hypothesis at order k is that (6.23) holds (with j instead of n) for every $1 \le j \le k$ and every $\delta > 0$. At level n, we start from

$$\begin{aligned} |W_{i_1}...W_{i_1...i_n} - \widetilde{W}_{i_1}^M...\widetilde{W}_{i_1...i_n}^M| \ &\le \ W_{i_1}|W_{i_1i_2}...W_{i_1...i_n} - \widetilde{W}_{i_1i_2}^M...\widetilde{W}_{i_1...i_n}^M| \\ &+ \ |W_{i_1} - \widetilde{W}_{i_1}^M|\, \widetilde{W}_{i_1i_2}^M...\widetilde{W}_{i_1...i_n}^M\,, \end{aligned}$$

so that, summing up, we get

$$D_r^n(M) \le \tilde{D}_r^n(M) + \bar{D}_r^n(M)\,, \tag{6.26}$$

where

$$\bar{D}_r^n(M) = r^{-n} \sum_{1\le i_1,\dots,i_{n-1}\le r} |W_{i_1} - \widetilde{W}_{i_1}^M|\, \widetilde{W}_{i_1i_2}^M...\widetilde{W}_{i_1...i_n}^M\,.$$

Write

$$\tilde{D}_r^n(M) = \frac{1}{r}\sum_{k=1}^r W_k\ D_r^{n-1}(M)\circ\Theta_k\,. \tag{6.27}$$

If $D_r^{n-1}(M)\circ\Theta_k \le \epsilon$ for all $1 \le k \le r$, then $\tilde{D}_r^n(M) \le \epsilon Z_r^1$. Therefore, for $\delta > 2\epsilon$,

$$\begin{aligned} P(\tilde{D}_r^n(M) > \frac{\delta}{2}) \ &\le \ P(Z_r^1 > \frac{\delta}{2\epsilon}) + rP(D_r^{n-1}(M) > \epsilon) && (6.28)\\ &\le \ e^{-r\Lambda^*(\delta/2\epsilon)} + rP(D_r^{n-1}(M) > \epsilon), && (6.29) \end{aligned}$$

where the last step holds by the Chernov bound. Similarly,

$$P(\bar{D}_r^n(M) > \frac{\delta}{2}) \le P(D_r^1(M) > \epsilon) + rP(Z_r^{n-1}(M) > \frac{\delta}{2\epsilon}), \tag{6.30}$$

where

$$Z_r^{n-1}(M) = \frac{1}{r^{n-1}} \sum_{i_1...i_{n-1}} \widetilde{W}_{i_1}^M...\widetilde{W}_{i_1...i_{n-1}}^M\,. \tag{6.31}$$

To manage this expression we use the following lemma.

Lemma 6.7 *With the notations of Proposition 3.1, assume that $U^1, U^2, .., U^n$ are positive and bounded, and have mean 1. Then $(S_r^n)_r$ satisfies the LDP with good rate function $I(x) = L_1^*(x)$, where L_1^* is the Legendre dual of $L_1(\theta) := \log E e^{\theta U^1}$.*

To apply Lemma 6.7 we use $L_1(\theta) = \log E e^{\theta \widetilde{W}^M}$. Since $L_1(\theta) \le \Lambda(\theta / E_M)$ for $\theta > 0$, we have $L_1^*(x) \ge \Lambda^*(x E_M)$ for $x E_M > 1$. If $M > M_0$, then $E_M \ge 1/2$, so that if $\delta > 4\epsilon$, then

$$\limsup_r \frac{1}{r} \log P\left(Z_r^{n-1}(M) > \frac{\delta}{2\epsilon} \right) \le -\Lambda^* \left(\frac{\delta}{4\epsilon} \right) . \tag{6.32}$$

Let $\ell_M^n(\delta) := \limsup_r \frac{1}{r} \log P\left(D_r^n(M) > \delta\right)$. Let $A > 0$ be arbitrary. Choose $\epsilon < \delta/4$ such that $\Lambda^*(\delta/4\epsilon) > A$. By (6.29), (6.30), (6.32), we see that

$$\ell_M^n(\delta) \le \max \left\{ -A, \ell_M^1(\epsilon), \ell_M^{n-1}(\epsilon) \right\} . \tag{6.33}$$

Hence, by the induction hypothesis, $\lim_{M \to \infty} \ell_M^n(\delta) \le -A$. This ends the proof of (6.23) since A is arbitrary. Now, collecting (6.23) and (6.29), we get (6.13). $\square$

Proof of Lemma 6.7 We apply the Gärtner-Ellis theorem. By independence,

$$\frac{1}{r} \log E \exp\left(r\theta S_r^n\right) = E \exp\left(\theta U^1 \tilde{S}_r^{n-1} \right) , \tag{6.34}$$

where

$$\tilde{S}_r^{n-1} =: \frac{1}{r^{n-1}} \sum_{1 \le i_1, \dots, i_{n-1} \le r} U_{i_1}^2 \dots U_{i_1 \dots i_{n-1}}^n , \tag{6.35}$$

is independent of U^1. Now, $\lim_{r \to \infty} \tilde{S}_r^{n-1} = 1$ a.s. by Proposition 3.1 a). Since the variables $U^1 \tilde{S}_r^{n-1}$ are uniformly bounded the dominated convergence theorem gives

$$\lim_r \frac{1}{r} \log E \exp\left(r\theta S_r^n\right) = L_1(\theta) , \tag{6.36}$$

and the LDP holds with rate function L_1^*. $\square$

6.3 Uniform LDP

Let $\mathcal{G}$ be a class of uniformly bounded functions. For every positive measure $\nu \in \mathcal{M}$ let $\nu^{\mathcal{G}}$ be the restriction of ν to $\mathcal{G}$. Let $\ell^\infty(\mathcal{G})$ be the set of all bounded functions from $\mathcal{G}$ to $\mathbb{R}$ provided with the norm $\|F\|_{\mathcal{G}} := \sup_{f \in \mathcal{G}} |F(f)|$. To our knowledge, the following result in the case $n = 1$ (smoothed partial-sum process) is not in the literature.

Theorem 6.8 *For $n < \infty$, under assumption (A_1), if $\mathcal{G}$ is a class of uniformly bounded and measurable functions, and is totally bounded in $\mathcal{L}^1([0,1], \lambda)$, then $(\mu_r^n)^{\mathcal{G}}$ $(r \ge 1)$ satisfies the LDP in $\ell^\infty(\mathcal{G})$ with the good rate function :*

$$I^{\mathcal{G}}(F) = \inf\{ I(\nu) \mid \nu \in \mathcal{M} \text{ and } \nu^{\mathcal{G}} = F \}. \tag{6.37}$$

Proof. From Proposition 6.4 a) we deduce that for n fixed, the family $\left(\mu_r^{n,M}\right)^{\mathcal{G}}$ is an exponential approximation of $(\mu_r^n)^{\mathcal{G}}$. It is then enough to prove that $\left(\mu_r^{n,M}\right)^{\mathcal{G}}$ satisfies a LDP in $\ell^\infty(\mathcal{G})$ with rate function $I^{\mathcal{G}}$. We use the method of [37] (see also [11]). For marginals (projection of $\mu_r^{n,M}$ on a finite family of functions of $\mathcal{G}$) the LDP is a consequence of Corollary 6.5. To get the uniform LDP it is then enough to prove :

a) if for $m \geq 1$, π_m is a collection of measurable projection on a finite m^{-1}-net of $\mathcal{G}$, and if $E_{m,\nu}(.) :=< \pi_m(.), \nu >\in \ell^\infty(\mathcal{G})$, then $\|\nu - E_{m,\nu}\|_{\mathcal{G}} \to 0$ as $m \to \infty$, uniformly over $\{\nu : I(\nu) \leq \alpha\}$ for all $\alpha > 0$;

b) the exponential uniform continuity: for every $\delta > 0$,

$$\lim_{\eta \downarrow 0} \limsup_{r} \frac{1}{r} \log P\left(\|\mu_r^{n,M}\|_{\mathcal{G}_\eta} > \delta\right) = -\infty\,, \tag{6.38}$$

where

$$\mathcal{G}_\eta = \{f - g \mid f, g \in \mathcal{G} \;\; \lambda(|f - g|) \leq \eta\}\,. \tag{6.39}$$

To prove a), take $\phi \in bm$ such that $\|\phi\|_\infty \leq 1$ and $\lambda(|\phi|) \leq m^{-1}$. If $I(\nu) < \infty$, then $d\nu = g d\lambda$ and $I(\nu) = \int \Lambda^*(g) d\lambda$. From (6.3) $\{g \in \mathcal{L}^1([0,1], \lambda)\;;\; \int \Lambda^*(g) d\lambda \leq \alpha\}$ is uniformly integrable ([12] p.266). Therefore, using

$$|\,\nu(\phi)\,| = |\int \phi \dot{\nu} d\lambda\,| \leq A m^{-1} + \int_{\dot{\nu} > A} \dot{\nu} d\lambda\,, \tag{6.40}$$

for every A, we see that

$$\lim_{m \to \infty} \sup\{|\,\nu(\phi)\,|\;;\; \lambda(|\,\phi\,|) \leq m^{-1}, \|\phi\|_\infty \leq 1,\; I(\nu) \leq \alpha\} = 0\,,$$

for every $\alpha > 0$, which implies a).

To prove b), we may assume that $\|f\|_\infty \leq 1/2$ and $\|g\|_\infty \leq 1/2$. Let $\mathcal{G}'_\eta = \{|h| : \; h \in \mathcal{G}\}$. Since

$$\|\mu_r^{n,M}\|_{\mathcal{G}_\eta} \leq \frac{M^{n-1}}{E_M^{n-1}} \|\mu_r^1\|_{\mathcal{G}'_\eta} \tag{6.41}$$

it is enough to prove

$$\lim_{\eta \downarrow 0} \limsup_{r} \frac{1}{r} \log P\left(\|\mu_r^1\|_{\mathcal{G}'_\eta} > \delta\right) = -\infty\,. \tag{6.42}$$

By the duality inequality, for every $x \in \mathbb{R}$,

$$x W_k \leq \Lambda(x) + \Lambda^*(W_k). \tag{6.43}$$

Let $h \in \mathcal{G}_\eta$. Since Λ is convex and $\Lambda(0) = 0$, we have

$$\Lambda(x) \leq \frac{x}{\theta} \Lambda(\theta) \text{ if } \; 0 \leq x \leq \theta\,,$$

so that for $\theta > 0$ and $x = \theta r \int_{A_k^r} |h|\, d\lambda$, (6.43) yields

$$\theta r W_k \int_{A_k^r} |h|\, d\lambda \leq \Lambda(\theta) r \int_{A_k^r} |h|\, d\lambda + \Lambda^*(W_k). \tag{6.44}$$

Adding up for $1 \leq k \leq r$, we get successively

$$\theta \|\mu_r^1\|_{\mathcal{G}'_\eta} \;\leq\; \eta \Lambda(\theta) + \frac{1}{r} \sum_k \Lambda^*(W_k)\,, \tag{6.45}$$

$$\frac{1}{r} \log P\left(\|\mu_r^1\|_{\mathcal{G}'_\eta} > \delta\right) \;\leq\; \frac{1}{r} \log P\left(\sum_{k=1}^{r} W_k > r\,(\theta\delta - \eta\Lambda(\theta))\right) \tag{6.46}$$

$$\leq\; -\frac{1}{2} (\theta\delta - \eta\Lambda(\theta)) + \log E \exp(\Lambda^*(W)/2) \tag{6.47}$$

(by (6.4)). Hence

$$\lim_{\eta \downarrow 0} \limsup_{r} \frac{1}{r} \log P\left(\|\mu_r^1\|_{\mathcal{G}'_\eta} > \delta\right) \leq -\frac{\theta\delta}{2} + \log E \exp(\Lambda^*(W)/2)\;. \tag{6.48}$$

Therefore (6.38) follows by letting $\theta \to \infty$.

Theorem 6.9 *Assume* (A_∞). *If* $\mathcal{G}$ *is a countable subset of* $\mathcal{L}^1([0,1],\lambda)$ *such that, for each* $\epsilon > 0$, *it can be covered by a finite number of* ϵ*-brackets* $[f_i, g_i]$, *with* f_i *and* g_i *measurable and bounded, then* $(\mu_r^\infty)^{\mathcal{G}}$ $(r \geq 1)$ *satisfies the LDP in* $\ell^\infty(\mathcal{G})$ *with the good rate function* $I^{\mathcal{G}}$ *(defined in (6.37))*.

Here we state the conclusion for a countable family to avoid measurability problem.

Proof. We adapt [11] p.202. For every integer m, let $[f_i^m, g_i^m]$ $(i = 1, ..., N_m)$ be a covering of $\mathcal{G}$ by m^{-1}-brackets. It is suficient to prove that

$$\lim_m \sup_{i=1,...N_m} \limsup_r \frac{1}{r} \log P\left(\sup\{\mid \mu_r^\infty(f) - \mu_r^\infty(f_i^m) \mid ; \ f \in [f_i^m, g_i^m]\} > \delta\right) = -\infty.$$

Actually, if $f \in [f_i^m, g_i^m]$ we have

$$\mid \mu_r^\infty(f) - \mu_r^\infty(f_i^m) \mid = \mu_r^\infty(f) - \mu_r^\infty(f_i^m) \leq \mu_r^\infty\left(g_i^m - f_i^m\right),$$

so that, by (6.11), it is enough to prove

$$\lim_m \sup_{\phi \geq 0; \ \lambda(\phi) \leq m^{-1}} \limsup_r \frac{1}{r} \log P(\mu_r^1(\phi) > \delta) = -\infty\,.$$

But this is obvious since $0 \leq \mu_r^1(\phi) \leq \bar{w}\lambda(\phi)$. □

References

[1] Aaronson, J. and Burton, R. and Dehling, H. and Gilat, D. and Hill, T. and Weiss, B. (1996): Strong laws for L- and U-statistics. *Trans. Amer. Math. Soc.* **7**, 2845–2866.

[2] Alexander, S.A. and Pyke, R. (1986): A uniform central limit theorem for set-indexed partial-sum processes with finite variance. *Ann. Probab.* **14**, 582-597.

[3] Barral, J. (1999): Moments, continuité, et analyse multifractale des martingales de Mandelbrot. *Probab. Theory Related Fields* **113** , no. 4, 535-569.

[4] Bass, R.F. (1985): Law of iterated logarithm for set-indexed partial sum processes with finite variance. *Z. Wahrsch. Verw. Gebiete*, **70**, 591-608.

[5] Bass, R.F. and Pyke, R. (1984): A strong law of large numbers for partial-sum processes indexed by sets. *Ann. Probab.* **12** 268-271.

[6] Breiman, L. (1992): Probability. Classics in Applied Mathematics 7, SIAM 1992.

[7] Choi, B.D. and Sung, S.H. (1987): Almost sure convergence theorems of weighted sums of random variables. *Stochastic Anal. Appl.* **5**, 365-377 .

[8] Chow, Y.S. and Teicher H. (1978): *Probability: Independence, Interchangeability and Martingales*, Springer-Verlag, New-York.

[9] Collet, P. and Koukiou, F. (1992): Large deviations for multiplicative chaos. *Comm. Math. Phys.* **147**, 329-342.

[10] Durrett, R. and Liggett, T. (1983): Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebeite* **64** 275-301.

[11] Dembo, A. and Zajic, T. (1995): Uniform large and moderate deviations for functional empirical processes *Stochastic Process. Appl.* **67**, 195-211.

[12] Dembo, A. and Zeitouni, O. (1998): *Large Deviations Techniques and Applications*, Springer, New York.

[13] Eichelsbacher, P. and Schmock, U. (1998): Exponential approximations in completely regular topological spaces and extensions of Sanov's theorem. *Stochastic Process. Appl.*, **77**, 233-251.

[14] Etemadi, N. (1981): An elementary proof of the strong law of large numbers. *Z. Wahrsch. Verw. Gebiete*, **55**, 119–122.

[15] Fréchet, M. (1957): Sur la distance de deux lois de probabilité. (French) *C. R. Acad. Sci. Paris* **244**, 689–692.

[16] Gaenssler, P. and Ziegler, K. (1994): On function-indexed partial-sum processes. In *Prob. Theory and Math.Stat.* (B. Grigelionis *et al.,* Eds.), 285-311. VSP/TEV Vilnius.

[17] Gaenssler, P. and Ziegler, K. (1994a): A uniform law of large-numbers for set-indexed processes with applications to empirical and partial-sum processes. In *Probability in Banach Spaces,* Vol. 9, pp. 385-400. Birkhaüser, Boston.

[18] Guivarc'h, Y. (1990): Sur une extension de la notion de loi semi-stable. *Ann. IHP* **26** 261-285.

[19] Kahane, J.P. and Peyrière, J. (1976): Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.* **22** 131-145.

[20] Lacey, M. (1989): Law of the iterated logarithm for partial sum processes indexed by functions. *J. Theoret. Probab.*, **2**, 377–398.

[21] Ledoux, M. and Talagrand, M. (1991): *Probability in Banach Spaces.* Springer Verlag, New York.

[22] Liu, Q. (1996): The growth of an entire characteristic function and the tail probabilities of the limit of a tree martingale. In *Trees.* Progress in Probability, vol.40, pp. 51-80, 1996, Birkhäuser Verlag, Basel. Eds.: B.Chauvin, S.Cohen, A.Rouault.

[23] Liu, Q. (1998): Fixed points of a generalized smoothing transformation and applications to branching random walks. *Adv. Appl. Prob.*, **30**, 85-112.

[24] Liu, Q. (2000) On generalized multiplicative cascades. Stoch. Proc. Appl. 86 (2000) 263-286.

[25] Liu, Q. (2001): Asymptotic properties and absolute continuity of laws stable by random weighted mean. Stoch. Proc. Appl. (2001), 83-107.

[26] Liu, Q. and Rouault, A. (2000): Limit theorems for Mandelbrot's multiplicative cascades. *Ann. Appl. Probab.* **10**, 218-231.

[27] Mandelbrot, B. (1974a): Multiplications aléatoires et distributions invariantes par moyenne pondérée aléatoire. *C.R.Acad. Sci. Paris Ser.I* **278**, 289-292 and 355-358.

[28] Mandelbrot, B. (1974b): Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62**, 331-333.

[29] Najim, J. (2001): A Cramér type theorem for weighted empirical means, *Elec. J. of Prob.* (to appear)

[30] Nobel, A.(1995): A counter example concerning ergodic theorems for a class of functions *Statist. Probab. Lett.* **17**, 169-172.

[31] Petrov, V.V. (1975): *Sums of Independent Random Variables*, Springer-Verlag, Berlin.

[32] Rogers, L. C. G.; Williams, D. (1994) *Diffusions, Markov processes, and martingales. Vol. 1.*, John Wiley and Sons, Chichester. Reprint of the second edition (1994), Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.

[33] Strassen, V. (1964): An invariance principle for the law of iterated logarithm *Z. Wahrsch. Verw. Gebiete* **3** 211-226.

[34] Talagrand, M. (1987): The Glivenko-Cantelli problem. *Ann. Probab.* **15**, 837-870.

[35] van der Vaart, A.W. and Wellner J.A. (1996): *Weak convergence and empirical processes*, Springer, New-York.

[36] Waymire, E.C.; Williams, D. (1996): A cascade ecomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* **348** 585-632.

[37] Wu, Liming (1994): Large Deviations, Moderate Deviations and LIL for Empirical Processes *Ann. Probab.* **22** 17-27.

[38] Yurinskii (1977): On the error of the Gaussian approximation for convolutions. *Theory Prob. Appl.* **XXII** 236-247.

[39] van Zwet, W.R. (1980): A strong law for linear functions of order statistics. *Ann. Probab.* **8**, 986-990.

[40] Ziegler, K. (1997): Functional Central Limit Theorems for Triangular Arrays of Function-Indexed Processes under Uniformly Integrated Entropy Conditions. *J. of Multivariate Analysis* **62** 233-272.