Conditionally identically distributed species sampling sequences

Federico Bassetti[*], Irene Crimaldi[†], Fabrizio Leisen[‡]

June 9, 2008

Abstract

Conditional identity in distribution (Berti et al. (2004)) is a new type of dependence for random variables, which generalizes the well-known notion of exchangeability. In this paper, a class of random sequences, called *Generalized Species Sampling Sequences*, is defined and a condition to have conditional identity in distribution is given. Moreover, a class of generalized species sampling sequences that are conditionally identically distributed is introduced and studied: the *Generalized Ottawa sequences* (GOS). This class contains a "randomly reinforced" version of the Pólya urn and of the Blackwell-MacQueen urn scheme. For the empirical means and the predictive means of a GOS, we prove two convergence results toward suitable mixtures of Gaussian distributions. The first one is in the sense of *stable* convergence and the second one in the sense of *almost sure conditional* convergence. In the last part of the paper we study the length of the partition induced by a GOS at time n, i.e. the random number of distinct values of a GOS until time n. Under suitable conditions, we prove a strong law of large numbers and a central limit theorem in the sense of stable convergence. All the given results in the paper are accompanied by some examples.

Key-words: species sampling sequence, conditional identity in distribution, stable convergence, almost sure conditional convergence, generalized Pólya urn.

1 Introduction

A sequence $(X_n)_{n\geq 1}$ of random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ taking values in a Polish space, is said a *species sampling sequence* if (a version) of the regular conditional distribution of X_{n+1} given $X(n) := (X_1, \dots, X_n)$ is the transition kernel

$$K_{n+1}(\omega, \cdot) := \textstyle\sum_{k=1}^{n} \tilde{p}_{n,k}(\omega)\delta_{X_k(\omega)}(\cdot) + \tilde{r}_n(\omega)\mu(\cdot) \tag{1}$$

[*]Department of Mathematics, University of Pavia, via Ferrata 1, 27100 Pavia, Italy. federico.bassetti@unipv.it

[†]Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy. crimaldi@dm.unibo.it

[‡]Department of Economics, University of Pavia, Via San Felice, 5, 27100 Pavia, Italy. fabrizio.leisen@unimore.it

where $\tilde{p}_{n,k}(\cdot)$ and $\tilde{r}_n(\cdot)$ are real–valued measurable functions of $X(n)$ and μ is a probability measure. See Pitman (1996).

As explained in Hansen and Pitman (2000), a species sampling sequence $(X_n)_{n\geq 1}$ can be interpreted as the sequential random sampling of individuals' species from a possibly infinite population of individuals belonging to several species. If, for the sake of simplicity, we assume that μ is diffuse, then the interpretation is the following. The species of the first individual to be observed is assigned a random tag X_1, distributed according to μ. Given the tags $X_1, \dots X_n$ of the first n individuals observed, the species of the $(n+1)$-th individual is a new species with probability $\tilde{r}_n$ and it is equal to the observed species X_k with probability $\sum_{j=1}^n \tilde{p}_{n,j} I_{\{X_j = X_k\}}$.

The concept of species sampling sequence is naturally related to that of random partition induced by a sequence of observations. Given a random vector $X(n) = (X_1, \dots, X_n)$, we denote by L_n the (random) number of distinct values of $X(n)$ and by $X^*(n) = (X_1^*, \dots, X_{L_n}^*)$ the random vector of the distinct values of $X(n)$ in the order in which they appear. The *random partition induced by* $X(n)$ is the random partition of the set $\{1, \dots, n\}$ given by $\pi^{(n)} = [\pi_1^{(n)}, \dots, \pi_{L_n}^{(n)}]$ where

$$i \in \pi_k^{(n)} \Leftrightarrow X_i = X_k^*.$$

Two distinct indices i and j clearly belong to the same block $\pi_k^{(n)}$ for a suitable k if and only if $X_i = X_j$. It follows that the *prediction rule* (1) can be rewritten as

$$K_{n+1}(\omega, \cdot) = \sum_{k=1}^{L_n(\omega)} \tilde{p}_{n,k}^*(\omega) \delta_{X_k^*(\omega)}(\cdot) + \tilde{r}_n(\omega)\mu(\cdot) \tag{2}$$

where

$$\tilde{p}_{n,k}^* := \sum_{j \in \pi_k^{(n)}} \tilde{p}_{n,j}.$$

In Hansen and Pitman (2000) it is proved that if μ is diffuse and $(X_n)_{n\geq 1}$ is an exchangeable sequence, then the coefficients $\tilde{p}_{n,k}^*$ are almost surely equal to some function of $\pi^{(n)}$ and they must satisfy a suitable recurrence relation. Although there are only a few explicit prediction rules which give rise to exchangeable sequences, this kind of prediction rules are appealing for many reasons. Indeed, exchangeability is a very natural assumption in many statistical problems, in particular from the Bayesian viewpoint, as well for many stochastic models. Moreover, remarkable results are known for exchangeable sequences: among others, such sequences satisfy a strong law of large numbers and they can be completely characterized by the well–known de Finetti representation theorem. See, e.g., Aldous (1985). Further, for an exchangeable sequence the empirical mean $\sum_{k=1}^n f(X_k)/n$ and the predictive mean, i.e. $E[f(X_{n+1})|X_1, \dots, X_n]$, converge to the same limit as the number of observations goes to infinity. This fact can be invoked to justify the use of the empirical mean in the place of the predictive mean, which is usually harder to compute. Nevertheless, in some situations the assumption of exchangeability can be too restrictive. For instance, instead of a classical Pólya urn scheme, it may be useful to deal with the so called randomly reinforced Pólya urn scheme. See, for example, Crimaldi (2007), Crimaldi and Leisen (2008), Flournoy and May (2008) and May, Paganoni and Secchi (2005). Such a process fails to be exchangeable but it can be still described with a prediction rule which is not too far from (1), see Example 3.4 of the present paper. Our purpose is to introduce and study a class of *generalized species sampling sequences*, which are generally not exchangeable but which still have interesting mathematical properties.

We thus need to recall the notion of *conditional identity in distribution*, introduced and studied in Berti, Pratelli and Rigo (2004). Such form of dependence generalizes the notion of exchangeability preserving some of its nice predictive properties. One says that a sequence $(X_n)_{n\geq 1}$, defined on $(\Omega, \mathcal{A}, P)$ and taking values in a measurable space $(E, \mathcal{E})$, is *conditionally identically distributed* with respect to a filtration $\mathcal{G} = (\mathcal{G}_n)_{n\geq 0}$ (in the sequel, $\mathcal{G}$-CID for short), whenever $(X_n)_{n\geq 1}$ is $\mathcal{G}$–adapted and, for each $n \geq 0$, $j \geq 1$ and every measurable real–valued bounded function f on E,

$$E[f(X_{n+j}) \,|\, \mathcal{G}_n] = E[f(X_{n+1}) \,|\, \mathcal{G}_n].$$

This means that, for each $n \geq 0$, all the random variables X_{n+j}, with $j \geq 1$, are identically distributed conditionally on $\mathcal{G}_n$. It is clear that every exchangeable sequence is a CID sequence with respect to its natural filtration but a CID sequence is not necessarily exchangeable. Moreover, it is possible to show that a $\mathcal{G}$–adapted sequence $(X_n)_{n\geq 1}$ is $\mathcal{G}$-CID if and only if, for each measurable real–valued bounded function f on E,

$$V_n^f := \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n]$$

is a $\mathcal{G}$–martingale, see Berti, Pratelli and Rigo (2004). Hence, the sequence $(V_n^f)_{n\geq 0}$ converges almost surely and in L^1 to a random variable V_f. One of the most important features of CID sequences is the fact that this random variable V_f is also the almost sure limit of the empirical means. More precisely, CID sequences satisfy the following strong law of large numbers: for each real–valued bounded measurable function f on E, the sequence $(M_n^f)_{n\geq 1}$, defined by

$$M_n^f := \frac{1}{n} \textstyle\sum_{k=1}^n f(X_k), \tag{3}$$

converges almost surely and in L^1 to V_f. It follows that also the predictive mean $\mathrm{E}[f(X_{n+1})|X_1, \ldots, X_n]$ converges almost surely and in L^1 to V_f. In other words, CID sequences share with exchangeable sequences the remarkable fact that the predictive mean and the empirical mean merge when the number of observations diverges. Unfortunately, while, for an exchangeable sequence, we have $V_f = E[f(X_1)|\mathcal{T}] = \int f(x)m(\omega, \mathrm{d}x)$, where $\mathcal{T}$ is the tail–σ–field and m is the random directing measure of the sequence, it is difficult to characterize explicitly the limit random variable V_f for a CID sequence. Indeed no representation theorems are available for CID sequences. See, e.g., Aletti, May and Secchi (2007).

The paper is organized as follows. In Section 2 we state our definition of generalized species sampling sequence, we discuss some examples and we give a condition under which a generalized species sampling sequence is CID with respect to a suitable filtration $\mathcal{G}$. In Sections 3 and 4 we deal with a particular class of generalized species sampling sequences which are CID: the *generalized Ottawa sequences* (GOS for short). We prove that, for a GOS, under suitable conditions, the sequence $\sqrt{n}(M_n^f - V_n^f)$ converges in the sense of *stable* convergence to a mixture of Gaussian distributions. Moreover, we show that, under suitable conditions, also $\sqrt{n}(V_n^f - V_f)$ converges in the sense of *almost sure conditional convergence* to another mixture of Gaussian distributions. Both types of convergences are stronger than the convergence in distribution. These results are accompanied by two examples.

In Section 5 we study the length L_n of the random partition induced by a GOS at time n, i.e. the random number of the distinct values assumed by a GOS until time n. In particular, a strong law of large numbers and a stable central limit theorem are presented. This section is also enriched by some examples. The paper closes by a section devoted to proofs and by an appendix in which the reader can find some results used for the proofs.

2 Prediction rules which generate a CID sequence

The Blackwell–MacQueen urn scheme provides the most famous example of exchangeable prediction rule, that is

$$P\{X_{n+1} \in \cdot \,|X_1, \ldots, X_n\} = \textstyle\sum_{i=1}^n \frac{1}{\theta + n} \delta_{X_i}(\cdot) + \frac{\theta}{\theta + n}\mu(\cdot)$$

where θ is a strictly positive parameter and μ is a probability measure, see, e.g., Blackwell and MacQueen (1973) and Pitman (1996). This prediction rule determines an exchangeable sequence $(X_n)_{n\geq 1}$ whose directing random measure is a Dirichlet process with parameter $\theta\mu(\cdot)$, see Ferguson (1973). According to this prediction rule, if μ is diffuse, a new species is observed with probability $\theta/(\theta + n)$ and an old species X_j^* is observed with probability proportional to the cardinality of $\pi_j^{(n)}$, a sort of *preferential attachment principle*. This rule has its analogous in term of random partitions in the so–called *Chinese restaurant process*, see Pitman (2006) and the references therein.

A *randomly reinforced prediction rule* of the same kind could work as follows:

$$P\{X_{n+1} \in \cdot \,|X_1, \ldots, X_n, Y_1, \ldots, Y_n\} = \textstyle\sum_{i=1}^n \frac{Y_i}{\theta + \sum_{j=1}^n Y_j} \delta_{X_i}(\cdot) + \frac{\theta}{\theta + \sum_{j=1}^n Y_j}\mu(\cdot) \qquad (4)$$

where μ is a probability measure and $(Y_n)_{n\geq 1}$ is a sequence of independent positive random variables. If μ is diffuse, then we have the following interpretation: each individual has a random positive weight Y_i and, given the first n tags $X(n) = (X_1, \ldots, X_n)$ together with the weights $Y(n) = (Y_1, \ldots, Y_n)$, it is supposed that the species of the next individual is a new species with probability $\theta/(\theta + \sum_{j=1}^n Y_j)$ and one of the species observed so far, say X_l^*, with probability $\sum_{i\in\pi_l^{(n)}} Y_i/(\theta + \sum_{j=1}^n Y_j)$. Again a preferential attachment principle. Note that, in this case, instead of describing the law of $(X_n)_{n\geq 1}$ with the sequence of the conditional distributions of X_{n+1} given $X(n)$, we have a latent process $(Y_n)_{n\geq 1}$ and we characterize $(X_n)_{n\geq 1}$ with the sequence of the conditional distributions of X_{n+1} given $(X(n), Y(n))$.

Now that we have given an idea, let us formalize what we mean by *generalized species sampling sequence*. Let $(\Omega, \mathcal{A}, P)$ be a probability space and E and S be two Polish spaces, endowed with their Borel σ-fields $\mathcal{E}$ and $\mathcal{S}$, respectively. In the sequel, $\mathcal{F}^Z = (\mathcal{F}_n^Z)_{n\geq 0}$ will stand for the natural filtration associated with any sequence of random variables $(Z_n)_{n\geq 1}$ on $(\Omega, \mathcal{A}, P)$ and we set $\mathcal{F}_\infty^Z = \vee_{n\geq 0}\mathcal{F}_n^Z$. Finally, $\mathcal{P}_n$ will denote the set of all partitions of $\{1, \ldots, n\}$.

We shall say that a sequence $(X_n)_{n\geq 1}$ of random variables on $(\Omega, \mathcal{A}, P)$, with values in E, is a generalized species sampling sequence if:

- (h_1) X_1 has distribution μ.

- (h_2) There exists a sequence $(Y_n)_{n\geq 1}$ of random variables with values in $(S,\mathcal{S})$ such that, for each $n\geq 1$, a version of the regular conditional distribution of X_{n+1} given

$$\mathcal{F}_n := \mathcal{F}_n^X \vee \mathcal{F}_n^Y$$

is

$$K_{n+1}(\omega,\cdot) = \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}(\omega), Y(n)(\omega))\delta_{X_i(\omega)}(\cdot) + r_n(\pi^{(n)}(\omega), Y(n)(\omega))\mu(\cdot) \qquad (5)$$

with $p_{n,i}(\cdot,\cdot)$ and $r_n(\cdot,\cdot)$ suitable measurable functions defined on $\mathcal{P}_n \times S^n$ with values in $[0,1]$.

- (h_3) X_{n+1} and $(Y_{n+j})_{j\geq 1}$ are conditionally independent given $\mathcal{F}_n$.

Example 2.1. Let μ be a probability measure on E, $(\nu_n)_{n\geq 1}$ be a sequence of probability measures on S, $(r_n)_{n\geq 1}$ and $(p_{n,i})_{n\geq 1,\,1\leq i\leq n}$ be measurable functions such that

$$r_n : \mathcal{P}_n \times S^n \to [0,1], \qquad p_{n,i} : \mathcal{P}_n \times Z^n \to [0,1]$$

and

$$\textstyle\sum_{i=1}^n p_{n,i}(q_n, y_1,\dots,y_n) + r_n(q_n, y_1,\dots,y_n) = 1 \qquad (6)$$

for each $n\geq 1$ and each $(q_n, y_1,\dots,y_n)$ in $\mathcal{P}_n \times S^n$. By the Ionescu Tulcea Theorem, there are two sequences of random variables $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$, defined on a suitable probability space $(\Omega,\mathcal{A},P)$, taking values in E and S respectively, such that conditions (h_1), (h_2) and the following condition are satisfied:

- Y_{n+1} has distribution ν_{n+1} and it is independent of the σ-field

$$\mathcal{F}_n \vee \sigma(X_{n+1}) = \mathcal{F}_{n+1}^X \vee \mathcal{F}_n^Y.$$

This last condition implies that, for each n, $(Y_{n+j})_{j\geq 1}$ is independent of $\mathcal{F}_{n+1}^X \vee \mathcal{F}_n^Y$. It follows, in particular, that $(Y_n)_{n\geq 1}$ is a sequence of independent random variables. Therefore, also (h_3) holds true. Indeed, for each real–valued bounded $\mathcal{F}_n$-measurable random variable V, each bounded Borel function f on E, each $j\geq 1$ and each bounded Borel function h on S^j, we have

$$\begin{aligned}
\mathrm{E}[Vf(X_{n+1})h(Y_{n+1},\dots,Y_{n+j})] &= \mathrm{E}\big[\,Vf(X_{n+1})\mathrm{E}[h(Y_{n+1},\dots,Y_{n+j})\,|\,\mathcal{F}_n \vee \sigma(X_{n+1})]\,\big] \\
&= \mathrm{E}\big[\,Vf(X_{n+1})\textstyle\int h(y_{n+1},\dots,y_{n+j})\,\nu_{n+1}(\mathrm{d}y_{n+1})\dots(\mathrm{d}y_{n+1})\,\big] \\
&= \mathrm{E}\big[\,V\mathrm{E}[f(X_{n+1})\,|\,\mathcal{F}_n]\textstyle\int h(y_{n+1},\dots,y_{n+j})\,\nu_{n+1}(\mathrm{d}y_{n+1})\dots(\mathrm{d}y_{n+1})\,\big].
\end{aligned}$$

On the other hand, we have

$$\mathrm{E}[h(Y_{n+1},\dots,Y_{n+j})\,|\,\mathcal{F}_n] = \textstyle\int h(y_{n+1},\dots,y_{n+j})\,\nu_{n+1}(\mathrm{d}y_{n+1})\dots(\mathrm{d}y_{n+1})$$

hence

$$\mathrm{E}[f(X_{n+1})h(Y_{n+1},\dots,Y_{n+j})\,|\,\mathcal{F}_n] = \mathrm{E}[f(X_{n+1})\,|\,\mathcal{F}_n]\mathrm{E}[h(Y_{n+1},\dots,Y_{n+j})\,|\,\mathcal{F}_n].$$

This fact is sufficient in order to conclude that also assumption (h_3) is verified. $\Diamond$

In order to state our first result concerning generalized species sampling sequences, we need some further notation. Set

$$p^*_{n,j}(\pi^{(n)}) = p^*_{n,j}(\pi^{(n)}, Y(n)) := \textstyle\sum_{i\in\pi_j^{(n)}} p_{n,i}(\pi^{(n)}, Y(n)) \quad \text{for } j = 1,\dots,L_n$$

and

$$r_n := r_n(\pi^{(n)}, Y(n)).$$

Given a partition $\pi^{(n)}$, denote by $[\pi^{(n)}]_{j+}$ the partition of $\{1,\dots,n+1\}$ obtained by adding the element $(n+1)$ to the j-th block of $\pi^{(n)}$. Finally, denote by $[\pi^{(n)};(n+1)]$ the partition obtained by adding a block containing $(n+1)$ to $\pi^{(n)}$. For instance, if $\pi^{(3)}=[(1,3);(2)]$, then $[\pi^{(3)}]_{2+}=[(1,3);(2,4)]$ and $[\pi^{(3)};(4)]=[(1,3);(2);(4)]$.

Theorem 2.2. *A generalized species sampling sequence $(X_n)_{n\geq 1}$ with μ diffuse is a CID sequence with respect to the filtration $\mathcal{G}=(\mathcal{G}_n)_{n\geq 0}$ with $\mathcal{G}_n:=\mathcal{F}_n^X\vee\mathcal{F}_\infty^Y$ if and only if, for each n, the following condition holds P-almost surely:*

$$p^*_{n,j}(\pi^{(n)}) = r_n p^*_{n+1,j}([\pi^{(n)};\{n+1\}]) + \textstyle\sum_{l=1}^{L_n} p^*_{n+1,j}([\pi^{(n)}]_{l+}) p^*_{n,l}(\pi^{(n)}) \tag{7}$$

for $1\leq j\leq L_n$.

The next example generalizes the well-known two parameter Poisson-Dirichlet process.

Example 2.3. Let $\theta>0$ and $\alpha\geq 0$. Moreover, let μ be a probability measure on E and, $(\nu_n)_{n\geq 1}$ be a sequence of probability measures on $(\alpha,+\infty)$. Consider the following sequence of functions

$$p_{n,i}(q_n,y(n)) := \frac{y_i-\alpha/C_i(q_n)}{\theta+\sum_{j=1}^n y_j}$$
$$r_n(q_n,y(n)) := \frac{\theta+\alpha L(q_n)}{\theta+\sum_{j=1}^n y_j}$$

where $y(n)=(y_1,\dots,y_n)\in(\alpha,+\infty)^n$, $q_n\in\mathcal{P}_n$, $C_i(q_n)$ is the cardinality of the block in q_n which contains i and $L(q_n)$ is the number of blocks of q_n. It is easy to see that such functions satisfy (6). Hence, by Example 2.1, there exists a generalized species sampling sequence $(X_n)_{n\geq 1}$ for which

$$P\{X_{n+1}\in\cdot\,|X(n),Y(n)\} = \textstyle\sum_{l=1}^{L_n}\dfrac{\sum_{i\in\pi_l^{(n)}} Y_i-\alpha}{\theta+\sum_{j=1}^n Y_j}\delta_{X_l^*}(\cdot) + \dfrac{\theta+\alpha L_n}{\theta+\sum_{j=1}^n Y_j}\mu(\cdot). \tag{8}$$

where $(Y_n)_{n\geq 1}$ is a sequence of independent random variables such that each Y_n has law ν_n. If μ is diffuse, one can easily check that (7) of Theorem 2.2 holds and so $(X_n)_{n\geq 1}$ is a CID sequence with respect to $\mathcal{G}=(\mathcal{F}_n^X\vee\mathcal{F}_\infty^Y)_{n\geq 1}$.

It is worthwhile noting that if $Y_n=1$ for every $n\geq 1$ and α belongs to $[0,1]$, then we get an exchangeable sequence directed by the well-known two parameter Poisson-Dirichlet process: i.e. an exchangeable sequence described by the prediction rule

$$P\{X_{n+1}\in\cdot\,|X_1,\dots,X_n\} = \textstyle\sum_{l=1}^{L_n}\dfrac{|\pi_l^{(n)}|-\alpha}{\theta+n}\delta_{X_l^*}(\cdot)+\dfrac{\theta+\alpha L_n}{\theta+n}\mu(\cdot).$$

See, e.g., Pitman and Yor (1997) and Pitman (2006). $\diamond$

A special case of the previous example is the randomly reinforced Blackwell-McQueen urn scheme (4). However this prediction rule may be collocated in a more general class of generalized species sampling sequences, that are CID. In the next sections, we shall introduce and study this class, called "Generalized Ottawa Sequences".

3 Generalized Ottawa sequences

We shall say that a generalized species sampling sequence $(X_n)_{n\geq 1}$ is a *generalized Ottawa sequence* or, more briefly, a GOS, if for every $n\geq 1$

- The functions r_n and $p_{n,i}$ $(i = 1, \dots, n)$ do not depend on the partition, hence

$$K_{n+1}(\omega, \cdot) = \textstyle\sum_{i=1}^n p_{n,i}(Y(n)(\omega))\delta_{X_i(\omega)}(\cdot) + r_n(Y(n)(\omega))\mu(\cdot). \tag{9}$$

- The functions r_n are striclty positive and

$$r_n(Y_1, \dots, Y_n) \geq r_{n+1}(Y_1, \dots, Y_n, Y_{n+1}) \tag{10}$$

  almost surely.

- The functions $p_{n,i}$ satisfy

$$\begin{aligned} p_{n,i} &:= \frac{r_n}{r_{n-1}} p_{n-1,i} \qquad \text{for } i = 1, \dots, n-1 \\ p_{n,n} &:= 1 - \frac{r_n}{r_{n-1}} \end{aligned} \tag{11}$$

  with $r_0 = 1$.

For simplicity, from now on, we shall denote by r_n and $p_{n,i}$ the $\mathcal{F}_n^Y$-measurable random variables $r_n(Y(n))$ and $p_{n,i}(Y(n))$, that is $r_n := r_n(Y(n))$ and $p_{n,i} := p_{n,i}(Y(n))$.

First of all let us stress that any GOS is a CID sequence with respect to the filtration $\mathcal{G} = (\mathcal{F}_n^X \vee \mathcal{F}_\infty^Y)_{n \geq 0}$. Indeed, since $\mathcal{G}_n = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$, condition $(h3)$ implies that

$$\mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{F}_n] \tag{12}$$

for each bounded Borel function f on E and hence, by $(h2)$, one gets

$$V_n^f := \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] = \textstyle\sum_{i=1}^n p_{n,i} f(X_i) + r_n \mathrm{E}[f(X_1)].$$

Since the random variables $p_{n+1,i}$ are $\mathcal{G}_n$-measurable it follows that

$$\begin{aligned} \mathrm{E}[V_{n+1}^f \,|\, \mathcal{G}_n] &= \textstyle\sum_{i=1}^n p_{n+1,i} f(X_i) + p_{n+1,n+1} \mathrm{E}[f(X_{n+1}) \,|\, \mathcal{G}_n] + r_{n+1} \mathrm{E}[f(X_1)] \\ &= \frac{r_{n+1}}{r_n} \textstyle\sum_{i=1}^n p_{n,i} f(X_i) + V_n^f - \frac{r_{n+1}}{r_n} V_n^f + r_{n+1} \mathrm{E}[f(X_1)] \\ &= \frac{r_{n+1}}{r_n} V_n^f - r_{n+1} \mathrm{E}[f(X_1)] + V_n^f - \frac{r_{n+1}}{r_n} V_n^f + r_{n+1} \mathrm{E}[f(X_1)] = V_n^f. \end{aligned}$$

Some examples follow.

Example 3.1. Consider a GOS for which

$$Y_n = a_n$$

where $(a_n)_{n \geq 0}$ is a *decreasing numerical* sequence with $a_0 = 1$, $a_n > 0$ and $r_n(y_1, \dots, y_n) = y_n$. $\Diamond$

Example 3.2. Let $(Y_n)_{n \geq 1}$ be a Markov chain taking values in $(0, 1]$, with $Y_1 = 1$ and transition probability kernel given by

$$P\{Y_{n+1} \leq x | Y_n\} = \frac{x}{Y_n} I_{(0, Y_n)}(x) + I_{[Y_n, +\infty)}(x) \quad n \geq 1.$$

Then we have $Y_n \geq Y_{n+1}$ a.s. for all $n \geq 1$. Thus we can consider a GOS with $r_n(y_1, \dots, y_n) = y_n$. $\Diamond$

As we shall see in the next example, the randomly reinforced Blackwell–McQueen urn scheme gives rise to a GOS.

Example 3.3. Let μ be a probability measure on E, $(\nu_n)_{n\geq 1}$ be a sequence of probability measures on S and (r_n), $(p_{n,i})$ measurable functions as in (10) and (11). Following Example 2.1, there exist two sequences of random variables $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$, defined on a suitable probability space $(\Omega, \mathcal{A}, P)$, such that each Y_n has law ν_n and it is independent of $\mathcal{F}_n^X \vee \mathcal{F}_{n-1}^Y$ and $(X_n)_{n\geq 1}$ follows the prediction rule (9), i.e. it is a GOS.

As special case one can consider $S = \mathbb{R}_+$ and

$$r_n(y_1, \ldots, y_n) = \frac{\theta}{\theta + \sum_{j=1}^n y_j}$$

with $\theta > 0$. $\diamondsuit$

Particular case of the previous example is the following randomly reinforced Pólya urn.

Example 3.4 (A randomly reinforced Pólya urn)**.** An urn contains b black and r red balls, b and r being strictly positive integer numbers. Repeatedly (at each time $n \geq 1$), one ball is drawn at random from the urn and then replaced together with a positive random number Y_n of additional balls of the same color. For each n, the random number Y_n must be independent of the preceding numbers and of the drawings until time n. If we denote by X_n the indicator function of the event $\{$black ball at time $n\}$, then we clearly have $E = \{0, 1\}$,

$$\mu(0) = \frac{r}{b+r}, \qquad \mu(1) = \frac{b}{b+r},$$

and

$$P\{X_{n+1} \in \cdot \ |X(n), Y(n)\} = \frac{1}{b + r + \sum_{j=1}^n Y_j} \sum_{i=1}^n Y_i \delta_{X_i(\omega)}(\cdot) + \frac{b+r}{b + r + \sum_{j=1}^n Y_j} \mu(\cdot).$$

Note that the sequence $(X_n)_{n\geq 1}$ is generally not exchangeable. Indeed, it is straightforward to prove that, even if the random variables Y_n are identically distributed, the sequence $(X_n)_{n\geq 1}$ is not exchangeable (apart from particular cases). $\diamondsuit$

4 Convergence results for a GOS

In this section we prove some limit theorems for a GOS under stable convergence and almost sure conditional convergence.

Stable convergence has been introduced by Rényi (1963) and subsequently studied by various authors, see, for example, Aldous and Eagleson (1978), Jacod and Memin (1981), Hall and Heyde (1980). A detailed treatment, including some strengthened forms of stable convergence, can be found in Crimaldi, Letta and Pratelli (2007).

Given a probability space $(\Omega, \mathcal{A}, P)$ and a Polish space E (endowed with its Borel σ-field $\mathcal{E}$), a kernel K on E is a family $K = (K(\omega, \cdot))_{\omega\in\Omega}$ of probability measure on E such that, for each bounded Borel function g on E, the function

$$K(g)(\omega) = \int g(x) K(\omega, \mathrm{d}x)$$

is measurable with respect to $\mathcal{A}$. Given a sub-σ-field $\mathcal{H}$ of $\mathcal{A}$, we say that the kernel K is $\mathcal{H}$-measurable if, for each bounded Borel function g on E, the random variable $K(g)$ is measurable with respect to

$\mathcal{H}$. In the following, the symbol $\mathcal{N}$ will denote the sub-σ-field generated by the P-negligible events of $\mathcal{A}$. Given a sub-σ-field $\mathcal{H}$ of $\mathcal{A}$ and a $\mathcal{H} \vee \mathcal{N}$-measurable kernel K on E, a sequence $(Z_n)_{n\geq 1}$ of random variables on $(\Omega, \mathcal{A}, P)$ with values in E converges $\mathcal{H}$*-stably* to K if, for each bounded continuous function g on E and for each $\mathcal{H}$–measurable real–valued bounded random variable W

$$\mathrm{E}[g(Z_n)\, W] \longrightarrow \mathrm{E}[K(g)\, W].$$

If $(Z_n)_{n\geq 1}$ converges $\mathcal{H}$-stably to K then, for each $A \in \mathcal{H}$ with $P(A) \neq 0$, the sequence $(Z_n)_{n\geq 1}$ converges in distribution under the probability measure $P_A = P(\cdot|A)$ to the probability measure $P_A K$ on E given by

$$P_A K(B) = P(A)^{-1}\mathrm{E}[I_A K(\cdot, B)] = \int K(\omega, B)\, P_A(\mathrm{d}\omega) \qquad \text{for each } B \in \mathcal{E}. \tag{13}$$

In particular, if $(Z_n)_{n\geq 1}$ converges $\mathcal{H}$-stably to K, then $(Z_n)_{n\geq 1}$ converges in distribution to the probability measure PK on E given by

$$PK(B) = \mathrm{E}[K(\cdot, B)] = \int K(\omega, B)\, P(\mathrm{d}\omega) \qquad \text{for each } B \in \mathcal{E}. \tag{14}$$

Moreover, if all the random variables Z_n are $\mathcal{H}$-measurable, then the $\mathcal{H}$-stable convergence obviously implies the $\mathcal{A}$-stable convergence.

Given a filtration $\mathcal{G} = (\mathcal{G}_n)_{n\geq 0}$ and a kernel K on E, we shall say that, with respect to $\mathcal{G}$, the sequence $(Z_n)_{n\geq 1}$ converges to K in the sense of the almost sure conditional convergence if, for each bounded continuous function g, we have

$$\mathrm{E}[g(Z_n)\,|\,\mathcal{G}_n] \longrightarrow K(g) \qquad \text{almost surely.}$$

If $(Z_n)_{n\geq 1}$ converges to K in the sense of the almost sure conditional convergence with respect to a filtration $\mathcal{G}$, then $(Z_n)_{n\geq 1}$ also converges $\mathcal{G}_\infty$-stably to K, see Crimaldi (2007).

Throughout the paper, if U is a positive random variable, we shall call the Gaussian kernel associated with U the family

$$\mathcal{N}(0, U) = \big(\mathcal{N}(0, U(\omega))\big)_{\omega\in\Omega}$$

of Gaussian distributions with zero mean and variance equal to $U(\omega)$ (with $\mathcal{N}(0, 0) := \delta_0$). Note that, in this case, the probability measure defined in (13) and (14) is a mixture of Gaussian distributions.

It is worthwhile to recall that, if $(X_n)_{n\geq 1}$ is a GOS, then it is a CID sequence with respect to the filtration $\mathcal{G} = (\mathcal{F}_n^X \vee \mathcal{F}_\infty^Y)_{n\geq 0}$ (as shown in Section 3) and so the sequence V_n^f (defined in section 3) converges almost surely and in L^1 to a random variable V_f, whenever f is a bounded Borel function on E. Moreover, the random variable V_f is also the almost sure (and in L^1) limit of the empirical mean

$$M_n^f = \frac{1}{n}\sum\nolimits_{k=1}^{n} f(X_k).$$

We are ready to state the main theorems of this section.

Theorem 4.1. *Let $(X_n)_{n\geq 1}$ be a GOS. Using the above notation, for each bounded Borel function f and each $n \geq 1$, let us set*

$$S_n^f = \sqrt{n}(M_n^f - V_n^f)$$

and, for $1 \leq j \leq n$,

$$Z_{n,j}^f = \frac{1}{\sqrt{n}}\big[f(X_j) - jV_j^f + (j-1)V_{j-1}^f\big] = \frac{1}{\sqrt{n}}(1 + jp_{j,j})\big[f(X_j) - V_{j-1}^f\big].$$

Suppose that:

(a) $U_n^f := \sum_{j=1}^n (Z_{n,j}^f)^2 \xrightarrow{P} U_f$.

(b) $(Z_n^f)^* := \sup_{1\leq j\leq n} |Z_{n,j}^f| \xrightarrow{L^1} 0$.

Then the sequence $(S_n^f)_{n\geq 1}$ *converges* $\mathcal{A}$*-stably to the Gaussian kernel* $\mathcal{N}(0, U_f)$.

In particular, condition (a) and (b) are satisfied if the following conditions hold:

(a1) $U_n^f \xrightarrow{a.s.} U_f$.

(b1) $\sup_{n\geq 1} \mathrm{E}[(S_n^f)^2] < +\infty$.

Let us see an application of the previous theorem in the next example.

Example 4.2. Let us consider the setting of Example 3.3 with

$$r_k = \frac{\theta}{\theta + \sum_{i=1}^k Y_i}.$$

where $\theta > 0$ and the random variables Y_n are identically distributed with $Y_n \geq \gamma > 0$ and $\mathrm{E}[Y_n^4] < +\infty$. Given a bounded Borel function f on E we are going to prove that the sequence $(S_n^f)_{n\geq 1}$ (defined in Theorem 4.1) converges $\mathcal{A}$-stably to the Gaussian kernel

$$\mathcal{N}\big(0, \Delta(V_{f^2} - V_f^2)\big),$$

where $\Delta := \mathrm{Var}[Y_1]/\mathrm{E}[Y_1]^2$.

Without loss of generality, we may assume that f takes values in $[0,1]$. Let us observe that, after some calculations, we have

$$S_n^f = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^n f(X_i) - n\,\frac{\theta \mathrm{E}[f(X_1)] + \sum_{i=1}^n Y_i f(X_i)}{\theta + \sum_{j=1}^n Y_j}\right).$$

If we set $b := \theta\mathrm{E}[f(X_1)]$ and $\widehat{Y}_i := Y_i - \mathrm{E}[Y_i] = Y_i - m$, then we can write

$$S_n^f = \frac{1}{\sqrt{n}}\frac{1}{(\theta + \sum_{j=1}^n Y_j)}\big[(\theta + \textstyle\sum_{j=1}^n \widehat{Y_j})\sum_{i=1}^n f(X_i) - nb - n\sum_{i=1}^n \widehat{Y}_i f(X_i)\big].$$

Therefore, since $Y_n \geq \gamma$ and $0 \leq f(X_n) \leq 1$ for each n, we obtain

$$\begin{aligned}\mathrm{E}[(S_n^f)^2] &\leq \frac{2n}{(\theta+\gamma n)^2}\big(\mathrm{E}[\,(\theta + \textstyle\sum_{j=1}^n \widehat{Y_j})^2\,] + \mathrm{E}[\,(b + \sum_{i=1}^n \widehat{Y}_i f(X_i)\,)^2\,]\big)\\ &\leq \frac{2n}{(\theta+\gamma n)^2}\big(\theta^2 + b^2 + 2n\mathrm{Var}[Y_1]\big) \leq C\end{aligned}$$

where C is a suitable constant. Finally, let us observe that, after some calculations, we get

$$\begin{aligned}U_n^f &= \frac{1}{n}\textstyle\sum_{j=1}^n \big[f(X_j) - jV_j^f + (j-1)V_{j-1}^f\big]^2\\ &= \frac{1}{n}\textstyle\sum_{j=1}^n \big[f(X_j) + A_j^f Y_j - B_j Y_j f(X_j) - V_{j-1}^f\big]^2,\end{aligned}$$

where

$$\begin{aligned}A_j^f &= j\big[(\theta + \textstyle\sum_{i=1}^j Y_i)(\theta + \sum_{i=1}^{j-1} Y_i)\big]^{-1}\big[b - \theta f(X_j) + \sum_{i=1}^{j-1} Y_i f(X_i)\big],\\ B_j &= j\big[(\theta + \textstyle\sum_{i=1}^j Y_i)(\theta + \sum_{i=1}^{j-1} Y_i)\big]^{-1}\sum_{i=1}^{j-1} Y_i.\end{aligned}$$

Hence, we have

$$\begin{aligned} U_n^f &= \frac{1}{n}\sum_{j=1}^n \big[f^2(X_j) + (A_j^f)^2 Y_j^2 + B_j^2 Y_j^2 f^2(X_j) + (V_{j-1}^f)^2 - 2f(X_j)V_{j-1}^f\big] \\ &+ \frac{2}{n}\sum_{j=1}^n \big[A_j^f Y_j f(X_j) - B_j Y_j f^2(X_j) - A_j^f Y_j V_{j-1}^f - A_j^f B_j Y_j^2 f(X_j) + B_j Y_j f(X_j) V_{j-1}^f\big]. \end{aligned}$$

Recall that we have the following almost sure convergences:

$$f^q(X_n)/n \longrightarrow 0 \,(\text{for } q = 1, 2), \quad V_n^f \longrightarrow V_f$$

$$\frac{1}{n}\sum_{j=1}^n Y_j^r \longrightarrow \mathrm{E}[Y_1^r] \,(\text{for } r = 1, 2), \quad \frac{1}{n}\sum_{j=1}^n f(X_j) \longrightarrow V_f, \quad \frac{1}{n}\sum_{j=1}^n f^2(X_j) \longrightarrow V_{f^2}.$$

From the above relations, we get

$$B_j \xrightarrow{a.s.} 1/\mathrm{E}[Y_1].$$

In order to study the convergence of $\frac{1}{n}\sum_{j=1}^n Y_j^r f^q(X_j)$ for $r, q = 1, 2$, let us set

$$Z_n := \sum_{j=1}^n \frac{1}{j}\big(Y_j^r f^q(X_j) - \mathrm{E}[Y_j^r f^q(X_j) \,|\, \mathcal{F}_{j-1}]\big).$$

The sequence $(Z_n)_{n\geq 1}$ is a martingale with respect to $\mathcal{F} = (\mathcal{F}_n)_{n\geq 1}$ such that

$$\begin{aligned} \mathrm{E}[Z_n^2] &= \sum_{j=1}^n \frac{1}{j^2}\mathrm{E}\big[\,(Y_j^r f^q(X_j) - \mathrm{E}[Y_j^r f^q(X_j) \,|\, \mathcal{F}_{j-1}]\,)^2\big] \\ &\leq 2\mathrm{E}[Y_1^{2r}]\sum_{j=1}^\infty \frac{1}{j^2} < +\infty. \end{aligned}$$

Therefore, by Kronecher's lemma, we find that

$$\frac{1}{n}\sum_{j=1}^n \big(Y_j^r f^q(X_j) - \mathrm{E}[Y_j^r f^q(X_j) \,|\, \mathcal{F}_{j-1}]\big) \xrightarrow{a.s.} 0.$$

On the other hand, since Y_j is independent of $\mathcal{F}_j^X \vee \mathcal{F}_{j-1}^Y$ by assumption, we have

$$\mathrm{E}[Y_j^r f^q(X_j) \,|\, \mathcal{F}_{j-1}] = \mathrm{E}[Y_1^r]\mathrm{E}[f^q(X_j) \,|\, \mathcal{F}_{j-1}] = \mathrm{E}[Y_1^r]V_{j-1}^{f^q} \xrightarrow{a.s.} \mathrm{E}[Y_1^r]V_{f^q}.$$

Since $n^{-1}\sum_{j=1}^n a_j d_j \overset{a.s.}{\to} ad$ whenever

$$a_j \geq 0, \quad d_j \overset{a.s.}{\to} d, \quad n^{-1}\sum_{j=1}^n a_j \overset{a.s.}{\to} a, \tag{15}$$

we obatin that

$$\frac{1}{n}\sum_{j=1}^n \mathrm{E}[Y_j^r f^q(X_j) \,|\, \mathcal{F}_{j-1}] \xrightarrow{a.s.} \mathrm{E}[Y_1^r]V_{f^q}$$

and so

$$\frac{1}{n}\sum_{j=1}^n Y_j^r f^q(X_j) \xrightarrow{a.s.} \mathrm{E}[Y_1^r]V_{f^q}.$$

In particular, we get

$$A_j^f \xrightarrow{a.s.} \frac{V_f}{\mathrm{E}[Y_1]}.$$

Summing up, we have proved that U_n^f is a sum of terms of the type $n^{-1}\sum_{j=1}^n a_j d_j$, where (a_j) and (d_j) satisfy conditions (15) and so we finally get that U_n^f converges a.s. to $U_f = \Delta(V_{f^2} - V_f^2)$. By Theorem 4.1, we conclude that S_n^f converges $\mathcal{A}$-stably to the Gaussian kernel $\mathcal{N}\big(0, \Delta(V_{f^2} - V_f^2)\big)$. $\Diamond$

The second result of this section is contained in the following theorem.

Theorem 4.3. *Let $(X_n)_{n\geq 1}$ be a GOS and f be a bounded Borel function. Using the previous notation, for $n \geq 0$ set*

$$Q_n := p_{n+1,n+1} = 1 - \frac{r_{n+1}}{r_n} \qquad \text{and} \qquad W_n^f := \sqrt{n}(V_n^f - V_f).$$

Suppose that the following conditions are satisfied:

(i) $n \sum_{k\geq n} Q_k^2 \xrightarrow{a.s.} H$, *where H is a positive real random variable.*

(ii) $\sum_{k\geq 0} k^2 \, \mathrm{E}[Q_k^4] < \infty$.

Then the sequence $(W_n^f)_{n\geq 0}$ converges to the Gaussian kernel

$$\mathcal{N}\big(0, H(V_{f^2} - V_f^2)\big)$$

in the sense of the almost sure conditional convergence with respect to the filtrations $\mathcal{F} = (\mathcal{F}_n^X \vee \mathcal{F}_n^Y)_{n\geq 0}$ and $\mathcal{G} = (\mathcal{F}_n^X \vee \mathcal{F}_\infty^Y)_{n\geq 0}$.

In particular, we have

$$W_n^f \overset{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}\big(0, H(V_{f^2} - V_f^2)\big).$$

Corollary 4.4. *Using the notation of Theorem 4.3, let us set for $k \geq 0$*

$$\rho_k = \frac{1}{r_{k+1}} - \frac{1}{r_k}$$

and assume the following conditions:

(a) $r_k \leq c_k$ *a.s. with* $\sum_{k\geq 0} k^2 c_{k+1}^4 < \infty$ *and* $k r_k \overset{a.s.}{\to} \alpha$, *where c_k, α are strictly positive constants.*

(b) The random variable ρ_k are independent and identically distributed with $\mathrm{E}[\rho_k^4] < \infty$.

Finally, let us set $\beta := \mathrm{E}[\rho_k^2]$ and $h := \alpha^2 \beta$.

Then, the conclusion of Theorem 4.3 holds true with H equal to the constant h.

Example 4.5. Let us consider the setting of Example 3.3 with

$$r_k = \frac{\theta}{\theta + \sum_{i=1}^k Y_i}.$$

where $\theta > 0$ and the random variables Y_n are identically distributed with $Y_n \geq \gamma > 0$ and $\mathrm{E}[Y_n^4] < +\infty$. Let us set $\mathrm{E}[Y_1] = m$ and $\mathrm{E}[Y_1^2] = \delta$. We have $r_k \leq c_k = \theta/(\theta + \gamma k)$ and, by the strong law of large numbers, we have

$$k r_k = \frac{\theta k}{\theta + \sum_{i=1}^k Y_i} \xrightarrow{a.s.} \theta/m.$$

Furthermore we have

$$\rho_k = \frac{1}{r_{k+1}} - \frac{1}{r_k} = \frac{Y_{k+1}}{\theta}$$

and so $\beta = \mathrm{E}[\rho_k^2] = \delta/\theta^2$. Therefore the above corollary holds with $h = \delta/m^2$. $\Diamond$

The particular generalized Pòlya urn discussed in Crimaldi (2007) (Cor. 4.1) and in May, Paganoni and Secchi (2005) is included in the above example.

5 Random partition induced by a GOS

Exchangeable species sampling sequences are strictly connected with exchangeable random partitions. Random partitions have been studied extensively, see, for instance Pitman (2006) and the references theirin.

In this section we investigate some properties of the length L_n of the random partition induced by a GOS at time n, i.e. the random number of distinct values of GOS until time n.

Let $A_0 := E$ and $A_n(\omega) := E \setminus \{X_1(\omega), \dots, X_n(\omega)\} = \{y \in E : \ y \notin \{X_1(\omega), \dots, X_n(\omega)\}\}$ for $n \geq 1$ and define the following $\mathcal{F}_n$-measurable random variable:

$$s_n := r_n(Y(n))\mu(A_n) = r_n\mu(A_n).$$

Remark 5.1. Reconsidering the species interpretation, given $X(n) = (X_1, \dots X_n)$ and $Y(n) = (Y_1, \dots, Y_n)$, the species of the $(n+1)$-th individual is a new species with probability s_n and one of the species observed so far with probability $1 - s_n$. In particular one has

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = s_n = r_n\mu(A_n).$$

If the probability measure μ is diffuse, then $s_n = r_n$.

If μ is diffuse and the coefficients r_n are deterministic (such as in Example 3.1), then the sequence of the increments $(L_n - L_{n-1})_{n\geq 1}$ (with $L_0 := 0$) is a sequence of independent random variables such that, for each n, the distribution of $L_n - L_{n-1}$ is a Bernoulli distribution with parameter r_{n-1}, hence it is immediate to deduce, under suitable conditions, both a strong law of large numbers and a central limit theorem for $(L_n)_{n\geq 1}$.

In this section we prove a law of large numbers and a central limit theorem for a GOS. Moreover, some examples of GOS that satisfy the hypotheses of these results are given.

Theorem 5.2. *Let $(X_n)_{n\geq 1}$ be a generalized species sampling sequence. Suppose that there exists a sequence $(h_n)_{n\geq 1}$ of real numbers and a random variable L such that the following properties hold:*

$$h_n \geq 0, \quad h_n \uparrow +\infty, \quad \textstyle\sum_{j\geq 1} \dfrac{\mathrm{E}[s_j(1-s_j)]}{h_j^2} < +\infty, \quad \dfrac{1}{h_n}\textstyle\sum_{j=0}^n s_j \xrightarrow{a.s.} L.$$

Then we have $L_n/h_n \xrightarrow{a.s.} L$.

Remark 5.3. Let us note that, for each n, we have

$$\mathrm{E}[L_{n+1} \,|\, \mathcal{F}_n] = L_n + s_n \geq L_n.$$

Hence the sequence $(L_n)_{n\geq 0}$ is a positive submartingale with $\mathrm{E}[L_{n+1}] = \mathrm{E}[L_n] + \mathrm{E}[s_n]$. Therefore $(L_n)_{\geq 0}$ is bounded in L^1 if and only if we have $\sum_{k\geq 0} \mathrm{E}[s_k] < +\infty$ and, in this case, $(L_n)_{n\geq 0}$ converges almost surely to an integrable random variable. It follows that, for each sequence $(h_n)_{n\geq 0}$ with $h_n \to +\infty$, the ratio L_n/h_n goes almost surely to zero. An example of this situation is given by Example 3.2 with μ diffuse. Indeed, in this case, we have $\mathrm{E}[s_n] = \mathrm{E}[Y_n] = (1/2)^{n-1}$.

Theorem 5.4. *Let $(X_n)_{n\geq 1}$ be a GOS with μ diffuse and suppose there exists a sequence $(h_n)_{n\geq 1}$ of real numbers and a positive random variable σ^2 such that the following properties hold:*

$$h_n \geq 0, \quad h_n \uparrow +\infty, \quad \sigma_n^2 := \frac{\sum_{j=1}^n r_j(1-r_j)}{h_n} \xrightarrow{a.s.} \sigma^2.$$

Then, setting $R_n := \sum_{j=1}^n r_j$, we have

$$T_n := \frac{L_n - R_{n-1}}{\sqrt{h_n}} \overset{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}(0, \sigma^2).$$

Corollary 5.5. *Under the same assumptions of Theorem 5.4, if $P(\sigma^2 > 0) = 1$, then we have*

$$\frac{T_n}{\sigma_n} = \frac{(L_n - R_{n-1})}{\sqrt{\sum_{j=1}^n r_j(1 - r_j)}} \overset{\mathcal{A}-stably}{\longrightarrow} \mathcal{N}(0, 1).$$

Example 5.6. Let us consider Example 3.1 with μ diffuse and

$$a_n = \frac{\theta}{\theta + n^{1-\alpha}}$$

with $\theta > 0$ and $0 < \alpha < 1$. We have $s_n = r_n = a_n$ and, setting $h_n = n^\alpha$ and $L = \theta/\alpha$, the assumptions of Theorem 5.2 are satisfied. Indeed we have

$$\begin{aligned}\sum_j \frac{r_j(1 - r_j)}{j^{2\alpha}} &= \sum_j \frac{j^{1-\alpha}}{(\theta + j^{1-\alpha})^2 j^{2\alpha}} \\ &= \sum_j \left(\frac{j^{1-\alpha}}{\theta + j^{1-\alpha}}\right)^2 \frac{1}{j^{\alpha+1}} < +\infty.\end{aligned}$$

Moreover, since

$$\frac{1}{n^\alpha}\sum_{j=1}^n \frac{1}{j^{1-\alpha}} \longrightarrow \frac{1}{\alpha} \qquad \text{for } \alpha \in (0, 1), \tag{16}$$

we have

$$\frac{1}{n^\alpha} R_n = \frac{1}{n^\alpha}\sum_{j=1}^n \frac{\theta}{\theta + j^{1-\alpha}} \longrightarrow \frac{\theta}{\alpha}.$$

Thus we have $L_n/n^\alpha \overset{a.s.}{\longrightarrow} \theta/\alpha$. Finally, since

$$\frac{1}{h_n}\sum_{j=1}^n a_j b_j \to b, \tag{17}$$

provided that $a_j \geq 0$, $\sum_{j=1}^n a_j/h_n \to 1$ and $b_n \to b$ as $n \to +\infty$, it is easy to see that

$$\sigma_n^2 = \frac{\sum_{j=1}^n r_j(1 - r_j)}{n^\alpha} = \frac{\theta}{n^\alpha}\sum_{j=1}^n \frac{j^{1-\alpha}}{(\theta + j^{1-\alpha})^2} = \frac{\theta}{n^\alpha}\sum_{j=1}^n \left(\frac{j^{1-\alpha}}{\theta + j^{1-\alpha}}\right)^2 \frac{1}{j^{1-\alpha}} \to \theta/\alpha.$$

Therefore, by Theorem 5.4, we obtain

$$T_n = \frac{L_n - R_{n-1}}{n^{\alpha/2}} \overset{\mathcal{A}-\text{stably}}{\longrightarrow} \mathcal{N}(0, \theta).$$

$\Diamond$

Example 5.7. Let us consider the setting of Example 3.3 with μ diffuse and

$$r_n = \frac{\theta}{\theta + \sum_{i=1}^n Y_i}.$$

where $\theta > 0$ and the random variables Y_n are independent identically distributed positive random variable with $\mathrm{E}[Y_n] = m > 0$. Then $s_n = r_n$ and, setting $h_n = \log n$ and $L = c/m$, the assumptions of Theorem 5.2 are satisfied. Indeed

$$\sum_j \frac{\mathrm{E}\big[r_j(1 - r_j)\big]}{(\log j)^2} \leq \sum_j \frac{1}{(\log j)^2} < +\infty.$$

Moreover, by the strong law of large numbers, we have

$$\left(\frac{\theta}{j} + \frac{1}{j}\sum_{i=1}^j Y_i\right)^{-1} \overset{a.s.}{\longrightarrow} 1/m.$$

Therefore, since $\frac{1}{\log n}\sum_{j=1}^n \frac{1}{j} \to 1$, by (17), we can conclude that

$$\frac{1}{\log n} R_n = \frac{\theta}{\log n}\sum_{j=1}^n \frac{1}{\theta + \sum_{i=1}^j Y_i} = \frac{\theta}{\log n}\sum_{j=1}^n \frac{1}{j}\left(\frac{\theta}{j} + \frac{1}{j}\sum_{i=1}^j Y_i\right)^{-1} \xrightarrow{a.s.} \frac{\theta}{m}$$

and so $L_n/\log n \xrightarrow{a.s.} \theta/m$. Moreover, by (17) and the strong law of large numbers, we have

$$\begin{aligned}
\sigma_n^2 &= \frac{\sum_{j=1}^n r_j(1-r_j)}{\log n} = \frac{\theta}{\log n}\sum_{j=1}^n \frac{\sum_{i=1}^j Y_i}{(\theta + \sum_{i=1}^j Y_i)^2} \\
&= \frac{\theta}{\log n}\sum_{j=1}^n \left(\frac{\sum_{i=1}^j Y_i/j}{\theta/j + \sum_{i=1}^j Y_i/j}\right)^2 \frac{j}{\sum_{i=1}^j Y_i}\,\frac{1}{j} \to \theta/m.
\end{aligned}$$

Therefore, by Theorem 5.4, we obtain

$$T_n = \frac{L_n - R_{n-1}}{\sqrt{\log n}} \overset{\mathcal{A}\text{−stably}}{\longrightarrow} \mathcal{N}(0, \theta/m)$$

and so

$$\frac{L_n - R_{n-1}}{\sqrt{\frac{\theta}{m}\log n}} \overset{\mathcal{A}\text{−stably}}{\longrightarrow} \mathcal{N}(0, 1).$$

If we take $Y_i = 1$ for all i, we find the well known results for the asymptotic distribution of the length of the random partition obtained with the Blackwell–McQueen urn scheme. Indeed, since $\sum_{j=1}^n j^{-1} - \log n = \gamma + O(\frac{1}{n})$, one gets

$$\frac{L_n - \theta\log n}{\sqrt{\theta\log n}} \overset{\mathcal{A}\text{−stably}}{\longrightarrow} \mathcal{N}(0, 1).$$

See, for instance, pages 68-69 in Pitman (2006). $\diamondsuit$

6 Proofs.

This section contains all the proofs of the paper. Recall that

$$\mathcal{F}_n = \mathcal{F}_n^X \vee \mathcal{F}_n^Y \qquad \text{and} \qquad \mathcal{G}_n = \mathcal{F}_n^X \vee \mathcal{F}_\infty^Y = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$$

and so condition $(h3)$ of the definition of generalized species sampling sequence implies that

$$V_n^g := \mathrm{E}[g(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[g(X_{n+1}) \,|\, \mathcal{F}_n]$$

for each bounded Borel function g on E.

6.1 Proof of Theorem 2.2

We start with a useful lemma.

Lemma 6.1. *If $(X_n)_{n\geq1}$ is a generalized species sampling sequence, then we have*

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \sum_{j\in\pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n)) + r_n(\pi^{(n)}, Y(n))\mu(\{X_l^*\})$$

for each $l = 1, \ldots, L_n$. Moreover,

$$\mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n)) \int_{A_n} f(y)\,\mu(\mathrm{d}y).$$

holds true with $A_0 := E$ *and* A_n *the random "set" defined by*

$$A_n(\omega) := E \setminus \{X_1(\omega), \dots, X_n(\omega)\} = \{y \in E : \; y \notin \{X_1(\omega), \dots, X_n(\omega)\}\} \quad \textit{for } n \geq 1.$$

In particular, we have

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{G}_n] = P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n))\mu(A_n) := s_n(\pi^{(n)}, Y(n))$$

If μ *is diffuse, we have*

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = \mathrm{P}[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n))$$

for each $l = 1, \dots, L_n$ *and*

$$\mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] = \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n))\mathrm{E}[f(X_1)]$$

and

$$P[L_{n+1} = L_n + 1 \,|\, \mathcal{G}_n] = P[L_{n+1} = L_n + 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n)).$$

Proof. Since $\mathcal{G}_n = \mathcal{F}_n \vee \sigma(Y_{n+j} : j \geq 1)$, condition (h_3) implies that

$$P[n+1 \in \pi_l^{(n+1)} \,|\, \mathcal{G}_n] = P[X_{n+1} = X_l^* \,|\, \mathcal{G}_n] P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n].$$

Hence, by assumption $(h2)$, we have

$$\begin{aligned} P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n))\delta_{X_i}(X_l^*) + r_n(\pi^{(n)}, Y(n))\mu(\{X_l^*\}) \\ &= \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n)) + r_n(\pi^{(n)}, Y(n))\mu(\{X_l^*\}). \end{aligned}$$

for each $l = 1, \dots, L_n$. If μ is diffuse, we obtain

$$P[X_{n+1} = X_l^* \,|\, \mathcal{F}_n] = \textstyle\sum_{j \in \pi_l^{(n)}} p_{n,j}(\pi^{(n)}, Y(n))$$

for each $l = 1, \dots, L_n$.

Now, we observe that

$$I_{\{L_{n+1}=L_n+1\}} = I_{B_n}(X_1, \dots, X_n, X_{n+1})$$

where $B_n = \{(x_1, \dots, x_{n+1}) : \; x_{n+1} \notin \{x_1, \dots, x_n\}\}$. Thus, by (h_3) and (h_2), we have

$$\begin{aligned} \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] &= \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{F}_n] \\ &= \textstyle\int I_{B_n}(X_1, \dots, X_n, y) f(y) K_{n+1}(\cdot, \mathrm{d}y) \\ &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n)) \int_{A_n} f(y)\delta_{X_i}(\mathrm{d}y) + r_n(\pi^{(n)}, Y(n)) \int_{A_n} f(y)\mu(\mathrm{d}y) \\ &= r_n(\pi^{(n)}, Y(n)) \textstyle\int_{A_n} f(y)\mu(\mathrm{d}y). \end{aligned}$$

If we take $f = 1$, we get

$$P[U_{n+1} = 1 \,|\, \mathcal{G}_n] = P[U_{n+1} = 1 \,|\, \mathcal{F}_n] = r_n(\pi^{(n)}, Y(n))\mu(A_n).$$

Finally, if μ is diffuse, then $\mu(A_n(\omega)) = 1$ for each ω and so we have

$$\textstyle\int_{A_n} f(y)\mu(\mathrm{d}y) = \mathrm{E}[f(X_1)].$$

□

Proof of Theorem 2.2. Let us fix a bounded Borel function f on E. Using the given prediction rule, we have

$$\begin{aligned} V_n^f &= \textstyle\sum_{i=1}^n p_{n,i}(\pi^{(n)}, Y(n)) f(X_i) + r_n(\pi^{(n)}, Y(n)) \mathrm{E}[f(X_1)] \\ &= \textstyle\sum_{j=1}^{L_n} p_{n,j}^*(\pi^{(n)}) f(X_j^*) + r_n \mathrm{E}[f(X_1)]. \end{aligned}$$

The sequence (X_n) is $\mathcal{G}$-cid if and only if for each bounded Borel function f on E, the sequence $(V_n^f)_{n\geq 0}$ is a $\mathcal{G}$-martingale. We observe that we have (for the sake of simplicity we skip the dependence on $(Y_n)_{n\geq 1}$)

$$\begin{aligned} \mathrm{E}[V_{n+1}^f \,|\, \mathcal{G}_n] &= \textstyle\sum_{i=1}^n f(X_i) E_i + \mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] + \mathrm{E}[r_{n+1} \,|\, \mathcal{G}_n] \bar{f} \\ &= \textstyle\sum_{j=1}^{L_n} f(X_j^*) \sum_{i\in\pi_j^{(n)}} E_i + \mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] + \mathrm{E}[r_{n+1} \,|\, \mathcal{G}_n] \bar{f} \end{aligned}$$

where $E_i = \mathrm{E}[p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n]$ and $\bar{f} = \mathrm{E}[f(X_1)]$.
Now we are going to compute the various conditional expectations which appear in the second member of above equality. Since μ is diffuse, using Lemma 6.1, we have

$$\begin{aligned} E_i &= \mathrm{E}[p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] \\ &= \textstyle\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}} p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}} \,|\, \mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} \,|\, \mathcal{G}_n] p_{n+1,i}([\pi^{(n)}; n+1]) \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) \sum_{j\in\pi_l^{(n)}} p_{n,j}(\pi^{(n)}) + r_n p_{n+1,i}([\pi^{(n)}; n+1]) \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,i}([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) + r_n p_{n+1,i}([\pi^{(n)}; n+1]) \end{aligned}$$

and so

$$\begin{aligned} \sum_{i\in\pi_j^{(n)}} E_i &= \sum_{l=1,l\neq j}^{L_n} p_{n+1,j}^*([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) + \sum_{i\in\pi_j^{(n)}} p_{n+1,i}([\pi^{(n)}]_{j+}) p_{n,j}^*(\pi^{(n)}) + r_n p_{n+1,j}^*([\pi^{(n)}; n+1]) \\ &= \textstyle\sum_{l=1}^{L_n} p_{n+1,j}^*([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) - p_{n+1,n+1}([\pi^{(n)}]_{j+}) p_{n+1,j}^*(\pi^{(n)}) + r_n p_{n+1,j}^*([\pi^{(n)}; n+1]) \end{aligned}$$

Moreover, using Lemma 6.1 again, we have

$$\begin{aligned} &\mathrm{E}[p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] = \\ &\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}} p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} p_{n+1,n+1}(\pi^{(n+1)}) f(X_{n+1}) \,|\, \mathcal{G}_n] = \\ &\sum_{l=1}^{L_n} \mathrm{E}[I_{\{n+1\in\pi_l^{(n+1)}\}} \,|\, \mathcal{G}_n] p_{n+1,n+1}([\pi^{(n)}]_{l+}) f(X_l^*) + \mathrm{E}[I_{\{L_{n+1}=L_n+1\}} f(X_{n+1}) \,|\, \mathcal{G}_n] p_{n+1,n+1}([\pi^{(n)}]; n+1) = \\ &\textstyle\sum_{l=1}^{L_n} \left(\sum_{k\in\pi_l^{(n)}} p_{n,k}(\pi^{(n)})\right) p_{n+1,n+1}([\pi^{(n)}]_{l+}) f(X_l^*) + r_n p_{n+1,n+1}([\pi^{(n)}]; n+1) \bar{f} = \\ &\textstyle\sum_{l=1}^{L_n} p_{n,l}^*(\pi^{(n)}) p_{n+1,n+1}([\pi^{(n)}]_{l+}) f(X_l^*) + r_n p_{n+1,n+1}([\pi^{(n)}]; n+1) \bar{f}. \end{aligned}$$

Finally we have

$$\begin{aligned} \mathrm{E}[r_{n+1} \,|\, \mathcal{G}_n] &= 1 - \textstyle\sum_{i=1}^{n+1} \mathrm{E}[p_{n+1,i}(\pi^{(n+1)}) \,|\, \mathcal{G}_n] \\ &= 1 - \textstyle\sum_{i=1}^n E_i - E_{n+1} \\ &= 1 - \textstyle\sum_{i=1}^n E_i - \sum_{l=1}^{L_n} p_{n,l}^*(\pi^{(n)}) p_{n+1,n+1}([\pi^{(n)}]_{l+}) - r_n p_{n+1,n+1}([\pi^{(n)}]; n+1) \end{aligned}$$

Thus we get

$$\mathrm{E}[V_{n+1}^f \,|\, \mathcal{G}_n] = \sum_{j=1}^{L_n} c_{n,j} f(X_j^*) + (1 - \sum_{j=1}^{L_n} c_{n,j})\bar{f}$$

where

$$\begin{aligned} c_{n,j} &= \sum_{i \in \pi_j^{(n)}} E_i + p_{n+1,n+1}([\pi^{(n)}]_{j+}) p_{n,j}^*(\pi^{(n)}) \\ &= r_n p_{n+1,j}^*([\pi^{(n)}; n+1]) + \sum_{l=1}^{L_n} p_{n+1,j}^*([\pi^{(n)}]_{l+}) p_{n,l}^*(\pi^{(n)}) \end{aligned}$$

We can conclude that $(X_n)_{n\geq 1}$ is $\mathcal{G}$-cid if and only if we have, for each bounded Borel function f on E and each n

$$\sum_{j=1}^{L_n} p_{n,j}^* f(X_j^*) + r_n \bar{f} = \sum_{j=1}^{L_n} c_{n,j} f(X_j^*) + (1 - \sum_{j=1}^{L_n} c_{n,j})\bar{f} \qquad P\text{-almost surely.}$$

Since E is a Polish space, we may affirm that $(X_n)_{n\geq 1}$ is $\mathcal{G}$-cid if and only if, for each n, we have P-almost surely

$$\sum_{j=1}^{L_n} p_{n,j}^* \delta_{X_k^*}(\cdot) + r_n \mu(\cdot) = \sum_{j=1}^{L_n} c_{n,j} \delta_{X_k^*}(\cdot) + (1 - \sum_{j=1}^{L_n} c_{n,j})\mu(\cdot)$$

But this last equality holds if and only if, for each n, we have P-almost surely

$$p_{n,j}^* = c_{n,j} \qquad \text{for } 1 \leq j \leq L_n \; ;$$

that is

$$p_{n,j}^*(\pi^{(n)}) = r_n p_{n+1,j}^*([\pi^{(n)}; \{n+1\}]) + \sum_{l=1}^{L_n} p_{n+1,j}^*([\pi^{(n)}]_{l+},) p_{n,l}^*(\pi^{(n)})$$

This is exactly the condition in the statement of the Theorem 2.2.

□

6.2 Proofs of Section 4

Proof of Theorem 4.1. We will use Theorem A.2 in the Appendix. For each $n \geq 1$, let us set

$$D_n^f = \sqrt{n}(M_n^f - V_f) = \frac{1}{\sqrt{n}}\left[\sum_{k=1}^n f(X_k) - nV_f\right],$$

and, for $0 \leq j \leq n$,

$$L_{n,j}^f = \mathrm{E}[D_n^f \,|\, \mathcal{G}_j] \qquad \mathcal{F}_{n,j} = \mathcal{G}_j.$$

Then, for each $n \geq 1$, the sequence $(L_{n,j})_{0\leq j\leq n}$ is a martingale with respect to $(\mathcal{F}_{n,j})_{0\leq j\leq n}$ such that $L_{n,0} = \mathrm{E}[D_n^f|\mathcal{G}_0] = 0$ and, for $1 \leq j \leq n$,

$$L_{n,j}^f - L_{n,j-1}^f = \mathrm{E}[D_n^f \,|\, \mathcal{G}_j] - \mathrm{E}[D_n^f \,|\, \mathcal{G}_{j-1}] = Z_{n,j}^f.$$

Indeed, using (12) we have

$$\begin{aligned} &\mathrm{E}[D_n^f \,|\, \mathcal{G}_j] - \mathrm{E}[D_n^f \,|\, \mathcal{G}_{j-1}] \\ &= \frac{1}{\sqrt{n}}\left[\sum_{k=1}^j f(X_k) + (n-j)V_j^f - nV_j^f - \sum_{k=1}^{j-1} f(X_k) - (n-j+1)V_{j-1}^f + nV_{j-1}^f\right] \\ &= \frac{1}{\sqrt{n}}\left[f(X_j) - jV_j^f + (j-1)V_{j-1}^f\right]. \end{aligned}$$

Moreover, we have

$$S_n^f = \mathrm{E}[D_n^f \,|\, \mathcal{G}_n] = L_{n,n}^f = \sum_{j=1}^n Z_{n,j}^f.$$

Finally, we have

$$\mathcal{H}_j = \liminf_n \mathcal{F}_{n,j\wedge n} = \liminf_n \mathcal{G}_{j\wedge n} = \mathcal{G}_j$$

and, if we set

$$\mathcal{H} = \bigvee_{j\geq 0}\mathcal{H}_j = \bigvee_{j\geq 0}\mathcal{G}_j,$$

then the random variable U_f is measurable with respect to the σ-field $\mathcal{H}\vee\mathcal{N}$. At this point we can apply Theorem A.2 and the proof of the first assertion is concluded.

If conditions (a1) holds, then condition (a) is obviously verified. Moreover we have

$$Z^f_{n,j} = \frac{1}{\sqrt{n}}\, Z^f_j$$

where

$$Z^f_j = f(X_j) - jV^f_j + (j-1)V^f_{j-1}.$$

We can write

$$\frac{1}{n}(Z^f_n)^2 = (Z^f_{n,n})^2 = \sum_{j=1}^{n}(Z^f_{n,j})^2 - \frac{1}{n}\sum_{j=1}^{n-1}(Z^f_j)^2 = U^f_n - \frac{n-1}{n}U^f_{n-1} \xrightarrow{a.s.} 0,$$

This fact implies that

$$(Z^f_n)^* = \sup_{1\leq j\leq n}|Z^f_{n,j}| \xrightarrow{a.s.} 0,$$

Indeed,

$$\sup_{0\leq j\leq n}(Z^f_{n,j})^2 = \frac{1}{n}\sup_{0\leq j\leq n}(Z^f_j)^2 \xrightarrow{a.s.} 0.$$

Further, we have

$$\begin{aligned}\mathrm{E}\big[\big((X^f_n)^*\big)^2\big] &= \mathrm{E}[\sup_{1\leq j\leq n}(Z^f_{n,j})^2] \leq \sum_{j=1}^{n}\mathrm{E}[(Z^f_{n,j})^2] = \sum_{j=1}^{n}\mathrm{E}\big[\big(L^f_{n,j} - L^f_{n,j-1}\big)^2\big] \\ &= \sum_{j=1}^{n}\mathrm{E}\big[(L^f_{n,j})^2\big] - \mathrm{E}\big[(L^f_{n,j-1})^2\big] = \mathrm{E}\big[(L^f_{n,n})^2\big] = \mathrm{E}[(S^f_n)^2].\end{aligned}$$

From (b1) and the above relations, we obtain that the sequence $\big((Z^f_n)^*\big)_n$ is bounded in L^2 and so we get condition (b). □

Proof of Theorem 4.3. Without loss of generality, we may assume $|f|\leq 1$. It will be sufficient to prove that the sequence $(V^f_n)_{n\geq 0}$ satisfies conditions (a) and (b) of Theorem A.3, with $U = H(V_{f^2} - V^2_f)$. To this end, we observe firstly that, after some calculations, we have

$$V^f_k - V^f_{k+1} = \big[V^f_k - f(X_{k+1})\big]\, Q_k. \tag{18}$$

From this equality we get $|V^f_k - V^f_{k+1}| \leq Q_k$, and so, using assumption (ii), we find

$$\sup_k k^2\,|V^f_k - V^f_{k+1}|^4 \leq \sum_{k\geq 0} k^2 Q^4_k \in L^1.$$

Furthermore, by (18), we have

$$\sum_{k\geq n}(V^f_k - V^f_{k+1})^2 = \sum_{k\geq n}\big[V^f_k - f(X_{k+1})\big]^2 Q^2_k \qquad \text{for } n\to+\infty.$$

Therefore, in order to complete the proof, it suffices to prove, for $n\to+\infty$, the following convergence:

$$n\sum_{k\geq n}\big[V^f_k - f(X_{k+1})\big]^2 Q^2_k \xrightarrow{a.s.} H(V_{f^2} - V^2_f).$$

The above convergence can be rewritten as

$$n\sum_{k\geq n}\left[(V_k^f)^2+f^2(X_{k+1})-2V_k^f f(X_{k+1})\right]Q_k^2 \xrightarrow{a.s.} H(V_{f^2}-V_f^2). \tag{19}$$

Now, by assumption (i) and the almost sure convergence of $(V_k^f)_k$ to V_f and of $(V_k^{f^2})_k$ to V_{f^2}, we have

$$n\sum_{k\geq n}V_k^f\,Q_k^2 \xrightarrow{a.s.} V_f H \tag{20}$$

$$n\sum_{k\geq n}(V_k^f)^2\,Q_k^2 \xrightarrow{a.s.} (V_f)^2 H \tag{21}$$

$$n\sum_{k\geq n}V_k^{f^2}\,Q_k^2 \xrightarrow{a.s.} V_{f^2} H \tag{22}$$

Thus, it will be enough to prove the following convergence:

$$n\sum_{k\geq n}\left[g(X_{k+1})-V_k^g\right]Q_k^2 \xrightarrow{a.s.} 0 \tag{23}$$

where g is a bounded Borel function with $|g|\leq 1$. Indeed, from (23) with $g=f^2$ and (22), we obtain

$$n\sum_{k\geq n}f^2(X_{k+1})\,Q_k^2 \xrightarrow{a.s.} V_{f^2}H, \tag{24}$$

Moreover, from (23) with $g=f$ and (20), we obtain

$$n\sum_{k\geq n}f(X_{k+1})\,Q_k^2 \xrightarrow{a.s.} V_f H, \tag{25}$$

and so, by the almost sure convergence of $(V_k^f)_k$ to V_f, we get

$$n\sum_{k\geq n}V_k^f\,f(X_{k+1})\,Q_k^2 \xrightarrow{a.s.} (V_f)^2 H, \tag{26}$$

Then convergence relations (21), (24) and (26) lead us to the desired relation (19).

In order to prove (23), we consider the process $(Z_n)_{n\geq 0}$ defined by

$$Z_n := \sum_{k=0}^{n-1}k\left[g(X_{k+1})-V_k^g\right]Q_k^2.$$

It is a martingale with respect to the filtration $\mathcal{G}=(\mathcal{G}_n)_{n\geq 0}$. Moreover, by assumption (ii), we have

$$\mathrm{E}[Z_n^2]=\sum_{k=0}^{n-1}k^2\,\mathrm{E}\left[(g(X_{k+1})-V_k^g)^2\,Q_k^4\right]\leq\sum_{k\geq 0}k^2\,\mathrm{E}[Q_k^4]<\infty. \tag{27}$$

The martingale $(Z_n)_{n\geq 1}$ is thus bounded in L^2 and so it converges almost surely; that is, the series

$$\sum_{k\geq 0}k\left[g(X_{k+1})-V_k^g\right]Q_k^2$$

is almost surely convergent. On the other hand, by a well-known Abel's result, the convergence of a series $\sum_k a_k$, with $a_k\in\mathbb{R}$, implies the convergence of the series $\sum_k k^{-1}a_k$ and the relation $n\sum_{k\geq n}k^{-1}a_k\to 0$ for $n\to+\infty$. Applying this result, we find (23) and the proof is so concluded. □

Proof of Corollary 4.4. It will suffice to verify that condition (i) and (ii) of Theorem 4.3 hold with $H=h$. With regard to condition (ii), it is enough to observe that, by the obvious inequality $Q_k=r_{k+1}\rho_k\leq c_{k+1}\rho_k$ and the identity in distribution of the random variables ρ_k, we have

$$\sum_{k\geq 0}k^2\,\mathrm{E}[Q_k^4]\leq\sum_{k\geq 0}k^2c_{k+1}^4\,\mathrm{E}[\rho_k^4]=\mathrm{E}[\rho_0^4]\sum_{k\geq 0}k^2c_{k+1}^4<\infty.$$

In order to prove condition (i) of Theorem 5.4 (with $H = h$), we observe that the series

$$\sum_k k^{-1}\,(\rho_k^2 - \beta)$$

is almost surely convergent: indeed, the random variables $Z_k := k^{-1}\,(\rho_k^2 - \beta)$ are independent, centered and square-integrable, with $\mathrm{Var}[Z_k] = k^{-2}\,\mathrm{Var}[\rho_1^2]$. Therefore, by the above mentioned Abel's result, we obtain the almost sure convergence of the series

$$\sum_k k^{-2}\,(\rho_k^2 - \beta)$$

and the relation (for $n \to +\infty$)

$$n \sum_{k \geq n} k^{-2}\,(\rho_k^2 - \beta) \xrightarrow{a.s.} 0.$$

Since we have $n \sum_{k \geq n} k^{-2} \to 1$ for $n \to +\infty$, the above relation can be rewritten in the form

$$n \sum_{k \geq n} k^{-2} \rho_k^2 \xrightarrow{a.s.} \beta.$$

Now we observe that [1]

$$Q_k^2 = r_{k+1}^2 \rho_k^2 \overset{a.s.}{\sim} \alpha^2 k^{-2} \rho_k^2.$$

Hence, for $n \to +\infty$, we have

$$n \sum_{k \geq n} Q_k^2 \overset{a.s.}{\sim} \alpha^2 n \sum_{k \geq n} k^{-2}\,\rho_k^2 \xrightarrow{a.s.} \alpha^2 \beta = h.$$

Condition (i) of Theorem 4.3 (with $H = h$) is thus proved and the proof is concluded. $\square$

6.3 Proofs of Section 5

In order to study the asymptotic behavior of $(L_n)_{n \geq 1}$ it will be useful to introduce the sequence of the increments

$$U_1 := L_1 = 1 \qquad \text{and} \qquad U_n := L_n - L_{n-1} \quad \text{for } n \geq 2.$$

Clearly $(U_n)_{n \geq 1}$ is a sequence of random variables with values in $\{0, 1\}$ such that, for each $n \geq 1$, the random variable U_n is $\mathcal{F}_n^X$-measurable and $L_n = \sum_{i=1}^n U_n$.

Proof of Theorem 5.2. Without loss of generality, we can assume $h_n > 0$ for each n. Let us set

$$Z_0 := 0 \qquad Z_n := \sum_{j=1}^n (U_j - s_{j-1})/h_j.$$

Then $Z = (Z_n)_{n \geq 0}$ is a martingale with respect to the filtration $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$. Indeed, by Lemma 6.1, we have

$$\mathrm{E}[Z_{n+1} - Z_n \,|\, \mathcal{F}_n] = \mathrm{E}[U_{n+1} - s_n \,|\, \mathcal{F}_n] = \mathrm{E}[I_{\{L_{n+1} = L_n + 1\}} - s_n \,|\, \mathcal{F}_n] = 0.$$

Moreover, we have

$$\mathrm{E}[U_{n+1}] = P(L_{n+1} = L_n + 1) = \mathrm{E}[s_n]$$

and

$$\begin{aligned} \mathrm{E}\big[(U_{n+1} - s_n)^2\big] &= \mathrm{E}\big[\,\mathrm{E}[(U_{n+1} - s_n)^2 \,|\, \mathcal{F}_n]\,\big] \\ &= \mathrm{E}[(1 - s_n)^2 s_n + s_n^2 (1 - s_n)] = \mathrm{E}[s_n(1 - s_n)]. \end{aligned}$$

[1] Given two sequences (a_n), (b_n) of random variables, the notation $a_n \overset{a.s.}{\sim} b_n$ means that $\frac{a_n}{b_n} \overset{a.s.}{\to} 1$.

Therefore we obtain

$$\mathrm{E}[Z_n^2] = \sum_{j=1}^n \mathrm{E}\big[(U_j - s_{j-1})^2\big]/h_j^2 = \sum_{j=1}^n \mathrm{E}\big[s_{j-1}(1-s_{j-1})\big]/h_j^2$$

and so

$$\sup_n \mathrm{E}[Z_n^2] \leq \sum_{j\geq 1} \mathrm{E}\big[s_{j-1}(1-s_{j-1})\big]/h_j^2 < +\infty.$$

It follows that $(Z_n)_{n\geq 1}$ converges almost surely and, by Kronecker's lemma, we get

$$\frac{1}{h_n}\big(L_n - \sum_{j=1}^n s_{j-1}\big) = \frac{1}{h_n}\sum_{j=1}^n (U_j - s_{j-1}) \xrightarrow{a.s.} 0.$$

Therefore, since $\sum_{j=1}^n s_{j-1}/h_n = \sum_{j=0}^n s_j/h_n - s_n/h_n \xrightarrow{a.s.} L$, we obtain $L_n/h_n \xrightarrow{a.s.} L$. □

In order to prove Theorem 5.4 we need a preliminary lemma.

Lemma 6.2. *If $(X_n)_{n\geq 1}$ is a GOS with μ diffuse, then (with the previous notation), for each fixed k, a version of the conditional distribution of $(U_j)_{j\geq k+1}$ given $\mathcal{G}_k$ is the kernel Q_k so defined:*

$$Q_k(\omega,\cdot) := \bigotimes_{j=k+1}^{\infty} \mathcal{B}\big(1, r_{j-1}(\omega)\big)$$

where $\mathcal{B}\big(1, r_{j-1}(\omega)\big)$ denotes the Bernoulli distribution with parameter $r_{j-1}(\omega)$.

Proof. It is enough to verify that, for each $n \geq 1$, for each $\epsilon_{k+1}, \ldots, \epsilon_{k+n} \in \{0,1\}$ and for each $\mathcal{G}_k$-measurable real–valued bounded random variable Z, we have

$$\mathrm{E}\big[Z I_{\{U_{k+1}=\epsilon_{k+1},\ldots,U_{k+n}=\epsilon_{k+n}\}}\big] = \mathrm{E}\big[Z \prod_{j=k+1}^{k+n} r_{j-1}^{\epsilon_j}(1-r_{j-1})^{1-\epsilon_j}\big]. \tag{28}$$

We go on with the proof by induction on n. For $n = 1$, by Lemma 6.1, we have

$$\mathrm{E}\big[Z I_{\{U_{k+1}=\epsilon_{k+1}\}}\big] = \mathrm{E}\big[Z\mathrm{E}[I_{\{U_{k+1}=\epsilon_{k+1}\}} \,|\, \mathcal{G}_k]\big] = \mathrm{E}[Z r_k^{\epsilon_{k+1}}(1-r_k)^{1-\epsilon_{k+1}}].$$

Assume that (28) is true for $n-1$ and let us prove it for n. Let us fix an $\mathcal{G}_k$-measurable real–valued bounded random variable Z. By Lemma 6.1, we have

$$\begin{aligned}
&\mathrm{E}\big[Z I_{\{U_{k+1}=\epsilon_{k+1},\ldots,U_{k+n}=\epsilon_{k+n}\}}\big] = \mathrm{E}\big[Z I_{\{U_{k+1}=\epsilon_{k+1},\ldots,U_{k+n-1}=\epsilon_{k+n-1}\}} \mathrm{E}[U_{k+n} = \epsilon_{k+n} \,|\, \mathcal{G}_{k+n-1}]\big] \\
&= \mathrm{E}\big[Z r_{k+n-1}^{\epsilon_{k+n}}(1-r_{k+n-1})^{1-\epsilon_{k+n}} I_{\{U_{k+1}=\epsilon_{k+1},\ldots,U_{k+n-1}=\epsilon_{k+n-1}\}}\big].
\end{aligned}$$

We have done because also the random variable $Z r_{k+n-1}^{\epsilon_{k+n}}(1-r_{k+n-1})^{1-\epsilon_{k+n}}$ is $\mathcal{G}_k$-measurable and (28) is true for $n-1$. □

Proof of Theorem 5.4. Without loss of generality, we can assume $h_n > 0$ for each n. In order to prove the desidered $\mathcal{A}$-stable convergence, it is enough to prove the $\mathcal{F}_\infty^X \vee \mathcal{F}_\infty^Y$-stable convergence of (T_n) to $\mathcal{N}(0,\sigma^2)$. But, in order to prove this last convergence, since we have $\mathcal{F}_\infty^X \vee \mathcal{F}_\infty^Y = \bigvee_k \mathcal{G}_k$, it suffices to prove that, for each k and A in $\mathcal{G}_k$ with $P(A) \neq 0$, the sequence (T_n) converges in distribution under P_A to the probability measure $P_A\mathcal{N}(0,\sigma^2)$. In other words, it is sufficient to fix k and to verify that $(T_{k+n})_n$ (and so $(T_n)_n$) converges $\mathcal{G}_k$-stably to $\mathcal{N}(0,\sigma^2)$. (Note that the kernel $\mathcal{N}(0,\sigma^2)$ is $\mathcal{G}_k \vee \mathcal{N}$-measurable for each fixed k.) To this end, we observe that we have

$$T_{k+n} = \frac{\sum_{j=1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} = \frac{\sum_{j=1}^{k}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} + \frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}}.$$

Obviously, for $n \to +\infty$, we have

$$\frac{\sum_{j=1}^{k}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} \xrightarrow{a.s.} 0.$$

Therefore we have to prove

$$\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} \overset{\mathcal{G}_k-\text{stably}}{\longrightarrow} \mathcal{N}(0, \sigma^2). \tag{29}$$

From Lemma 6.2 we know that a version of the the conditional distribution of $(U_j)_{j \geq k+1}$ given $\mathcal{G}_k$ is the kernel Q_k so defined:

$$Q_k(\omega, \cdot) = \bigotimes_{j=k+1}^{\infty} \mathcal{B}\big(1, r_{j-1}(\omega)\big).$$

On the canonical space $\mathbb{R}^{\mathbb{N}^*}$ let us consider the canonical projections $(\xi_j)_{j \geq k+1}$. Then, for each $n \geq 1$, a version of the conditional distribution of

$$\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}}$$

given $\mathcal{G}_k$ is the kernel N_{k+n} so characterized: for each ω, the probability measure $N_{k+n}(\omega, \cdot)$ is the distribution, under the probability measure $Q_k(\omega, \cdot)$, of the random variable (which is defined on the canonical space)

$$\frac{\sum_{j=k+1}^{k+n}\big(\xi_j - r_{j-1}(\omega)\big)}{\sqrt{h_{k+n}}}.$$

On the other hand, for almost every ω, under $Q_k(\omega, \cdot)$, the random variables

$$Z_{n,i} := \frac{\xi_{k+i} - r_{k+i-1}(\omega)}{\sqrt{h_{k+n}}} \qquad \text{for } n \geq 1,\ 1 \leq i \leq n$$

form a triangular array which satisfies the assumptions of Theorem A.1 in the Appendix. Indeed, we have the row-independence property and

$$\mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}] = 0, \qquad \mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] = \frac{r_{k+i-1}(\omega)\big(1 - r_{k+i-1}(\omega)\big)}{h_{k+n}}.$$

Therefore, by assumption, for $n \to +\infty$, we have for almost every ω,

$$\sum_{i=1}^{n}\mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] = \frac{\sum_{i=1}^{n} r_{k+i-1}(\omega)\big(1 - r_{k+i-1}(\omega)\big)}{h_{k+n}} = \sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}} \longrightarrow \sigma^2(\omega).$$

Moreover, under $Q_k(\omega, \cdot)$, we have $Z_n^* := \sup_i Z_{n,i} \leq 2/\sqrt{h_{k+n}} \longrightarrow 0$. Finally, we observe that, setting $V_n := \sum_{i=1}^{n} Z_{n,i}^2$, we have

$$\mathrm{E}^{Q_k(\omega,\cdot)}[V_n^2] = \mathrm{Var}^{Q_k(\omega,\cdot)}[V_n] + \left(\sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}}\right)^2$$

with

$$\begin{aligned}\mathrm{Var}^{Q_k(\omega,\cdot)}[V_n] = \sum_{i=1}^{n}\mathrm{Var}^{Q_k(\omega,\cdot)}[Z_{n,i}^2] &\leq \sum_{i=1}^{n}\mathrm{E}^{Q_k(\omega,\cdot)}[Z_{n,i}^4] \\ &\leq 4\left(\sigma_{k+n}^2(\omega) - \frac{h_{k-1}\sigma_{k-1}^2(\omega)}{h_{k+n}}\right)\frac{1}{h_{k+n}}.\end{aligned}$$

Since, for almost every ω, the sequence $(\sigma_n^2(\omega))_n$ is bounded and $h_n \uparrow +\infty$, it follows that, for almost every ω, the sequence $(V_n)_n$ is bounded in L^2 under $Q_k(\omega, \cdot)$ and so uniformly integrable. Theorem A.1 assures that, for almost every ω, the sequence of probability measures

$$\big(N_{k+n}(\omega, \cdot)\big)_{n \geq 1}$$

weakly converges to the Gaussian distribution $\mathcal{N}\big(0,\sigma^2(\omega)\big)$. This fact implies that, for each bounded continuous function g, we have

$$\mathrm{E}\left[g\left(\frac{\sum_{j=k+1}^{k+n}(U_j - r_{j-1})}{\sqrt{h_{k+n}}} \right) \mid \mathcal{G}_k \right] \xrightarrow{a.s.} \mathcal{N}(0,\sigma^2)(g).$$

It obviously follows the $\mathcal{G}_k$-stable convergence (29). □

A Appendix

For the reader's convenience, we state some results used above.

Theorem A.1. *Let $(Z_{n,i})_{n\geq 1,\, 1\leq i\leq k_n}$ be a triangular array of square integrable centered random variables on a probability space $(\Omega,\mathcal{A},P)$. Suppose that, for each fixed n, $(Z_{n,i})_i$ is independent ("row-independence property"). Moreover, set*

$$\sigma_{n,i}^2 := \mathrm{E}[Z_{n,i}^2] = \mathrm{Var}[Z_{n,i}], \qquad \sigma_n^2 := \textstyle\sum_{i=1}^{k_n}\sigma_{n,i}^2,$$
$$V_n := \textstyle\sum_{i=1}^{k_n} Z_{n,i}^2, \qquad Z_n^* := \sup_{1\leq i\leq k_n}|Z_{n,i}|$$

and assume that $(V_n)_{n\geq 1}$ is uniformly integrable, $Z_n^* \xrightarrow{P} 0$ and $\sigma_n^2 \longrightarrow \sigma^2$.

Then $\sum_{i=1}^{k_n} Z_{n,i} \overset{in\ law}{\longrightarrow} \mathcal{N}(0,\sigma^2)$.

Proof. In Hall and Heyde (1980) (see pp. 53–54) it is proved that, under the uniform integrability of (V_n), the convergence in probability to zero of $(Z_n^*)_{n\geq 1}$ is equivalent to the Lindeberg condition. Hence, it is possible to apply Corollary 3.1 (pp. 58-59) in Hall and Heyde (1980) with $\mathcal{F}_{n,i} = \sigma(Z_{n,1},\dots,Z_{n,i})$.

Theorem A.2. *(See Th. 5 and Cor. 7 of sec. 7 in Crimaldi, Letta and Pratelli (2007))*
Let $(l_n)_{n\geq 1}$ be a sequence of strictly positive integers. On a probability space $(\Omega,\mathcal{A},P)$, for each $n\geq 1$, let $(\mathcal{F}_{n,j})_{0\leq j\leq l_n}$ be a filtration and $(L_{n,j})_{n\geq 1, 0\leq j\leq l_n}$ be a triangular array of real random variables such that, for each n, the family $(L_{n,j})_{0\leq j\leq l_n}$ is a martingale with respect to $(\mathcal{F}_{n,j})_{0\leq j\leq l_n}$ and $L_{n,0}=0$. For each pair (n,j), with $n\geq 1$, $1\leq j\leq l_n$, let us set $Z_{n,j} = L_{n,j} - L_{n,j-1}$ and

$$S_n = \textstyle\sum_{j=1}^{l_n} Z_{n,j} = L_{n,l_n}, \qquad U_n = \sum_{j=1}^{l_n} Z_{n,j}^2, \qquad Z_n^* = \sup_{1\leq j\leq l_n}|Z_{n,j}|.$$

Let us suppose that the sequence $(U_n)_{n\geq 1}$ converges in probability to a positive random variable U and the sequence $(Z_n^*)_{n\geq 1}$ converges in L^1 to zero. Finally, let $\mathcal{N}$ be the sub-σ-field generated by the P-negligible events of $\mathcal{A}$ and let us set

$$\mathcal{H}_j = \liminf\nolimits_n \mathcal{F}_{n,j\wedge l_n} \quad \text{for } j\geq 0, \qquad \mathcal{H} = \textstyle\bigvee_{j\geq 0}\mathcal{H}_j.$$

If U is measurable with respect to the σ-field $\mathcal{H}\vee\mathcal{N}$, then $(S_n)_{n\geq 1}$ converges $\mathcal{H}$-stably to the Gaussian kernel $\mathcal{N}(0,U)$.

Theorem A.3. *(see Crimaldi, 2007)*
On $(\Omega,\mathcal{A},P)$, let $(V_n)_{n\geq 0}$ be a real martingale with respect to a filtration $\mathcal{G}=(\mathcal{G}_n)_{n\geq 0}$. Suppose that $(V_n)_{n\geq 0}$ converges in L^1 to a random variable V. Moreover, setting

$$U_n := n\textstyle\sum_{k\geq n}(V_k - V_{k+1})^2, \qquad Z := \sup_k \sqrt{k}\,|V_k - V_{k+1}|, \tag{30}$$

assume that the following conditions hold:

(a) The random variable Z is integrable.

(b) The sequence $(U_n)_{n\geq 1}$ converges almost surely to a positive real random variable U.

Then, with respect to $\mathcal{G}$, the sequence $(W_n)_{n\geq 1}$ defined by

$$W_n := \sqrt{n}(V_n - V) \tag{31}$$

converges to the Gaussian kernel $\mathcal{N}(0, U)$ in the sense of the almost sure conditional convergence.

Obviously the previous almost sure conditional convergence also holds with respect to any filtration $\mathcal{F}'$ such that $\mathcal{F}_n^V \subset \mathcal{F}'_n \subset \mathcal{G}_n$.

References

Aldous D.J and Eagleson G.K. (1978). On mixing and stability of limit theorems. *Ann. Prob.* **6** 325–331.

Aldous D.J. (1985). Exchangeability and related topics. *Ecole d'Eté de Probabilités de Saint-Flour*, XIII—1983, 1–198, Lecture Notes in Math., 1117, Springer, Berlin.

Aletti G., May C., Secchi P. (2007). On the distribution of the limit proportion for a two-color, randomly reinforced urn with equal reinforcement distributions. *Adv. in Appl. Probab.* **39** no. 3 690–707.

Berti P., Pratelli L. and Rigo P. (2004). Limit Theorems for a Class of Identically Distributed Random Variables. *Ann. Probab.* **32** 2029–2052.

Blackwell D. and MacQueen J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355.

Crimaldi I. and Leisen F. (2008). Asymptotic results for a generalized Pólya urn with "multi-updating" and applications to clinical trials. *Communications in Statistics - Theory and Methods*, **37**(17), forthcoming.

Crimaldi I. (2007). An almost sure conditional convergence result and an application to a generalized Pólya urn. Submitted.
Currently available at: http://www.dm.unibo.it/~crimaldi/articoli-inviati.html"

Crimaldi I., Letta G. and Pratelli L. (2007). A strong form of stable convergence. *Séminaire de Probabilités XL*, LNM **1899** 203–225.

Ferguson T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 1047–1054.

Flournoy N. and May C. (2008). Asymptotics in response-adaptive designs generated by a two-colors, randomly reinforced urn. *Ann. Statist.*, forthcoming.

Hall P. and Heyde C.C. (1980). Martingale limit theory and its application. Academic press.

Hansen B. and Pitman J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.* **46** 251–256.

Jacod J. and Memin J. (1981). Sur un type de convegence en intermédiaire entre la convergence en loi et la convergence en probabilité. *Sem. de Prob. XV*, LNM **850** 529–546.

May C., Paganoni A. and Secchi P. (2005). On a two-color generalized Pólya urn. *Metron* **63** no. 1 115–134.

Pitman J. (1996). Some developments of the Blackwell-MacQueen urn scheme, In: Ferguson, T.S. et al. (Eds.), *Statistics, Probability and Game Theory; Papers in honor of David Blackwell,* Lecture Notes-Monograph Series, Institute of Mathematical Statistics, Hayward, California, **30** 245–267.

Pitman J. (2006) *Combinatorial Stochastic Processes.* Ecole d'Eté de Probabilités de Saint-Flour XXXII, LNM 1875, Springer.

Pitman J. and Yor M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900.

Rényi A. (1963). On Stable Sequences of Events, *Sankhya* Ser. A. **25** 293–302.