

Almost sure conditional convergence for a generalized Pólya urn

Irene Crimaldi*

Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5,
40126 Bologna, Italy.

September 23, 2007

Abstract: For a two-colors, randomly reinforced urn, we prove a result of almost sure conditional convergence toward a Gaussian kernel.

Keywords: Convergence of conditional distributions/expectations, Exchangeable sequences, Generalized Pólya urns.

1 Introduction

Urn models are a very popular topic because of their applications in various fields: sequential clinical trials, biology, industry and finance. A large number of “replacement policies” has been considered and studied by many authors, from different points of view and by means of different methods: see, for instance, Eggenberger-Pólya (1923), Pólya (1931), Pemantle (1990, 1991), Gouet (1993), Kotz et al. (2000), Dirienzo (2000), Moler et al. (2002), Paganoni-Secchi (2004), Amerio et al. (2004), Berti et al. (2004), Janson (2004, 2006), May et al. (2005), Muliere et al. (2006).

In the present paper we give a limit theorem for a two-colors, randomly reinforced urn, which is a generalization of the urn model considered in May et al. (2005). More precisely, we deal with the following experiment. An urn contains $b \in \mathbb{N}^*$ black and $r \in \mathbb{N}^*$ red balls. At each time $n \geq 1$, a ball is drawn from the urn and then it is put again in the urn together with other N_n balls of the same color. The numbers N_n are *randomly* chosen in \mathbb{N}^* . For each n , the way in which the number N_n is chosen may depend on n but it is independent of the results of the choices for the preceding numbers and of the preceding drawings. If we denote by Y_n the indicator function of the event {black ball at time n }, then, by some results in Berti et al. (2004), the sequence $(M_n)_{n \geq 1}$ defined by

$$M_n := \frac{1}{n} \sum_{i=1}^n Y_i \tag{1}$$

converges in L^1 and almost surely to a random variable V . Moreover, the random variable

$$V_n := (b + \sum_{i=1}^n Y_i N_i) (b + r + \sum_{i=1}^n N_i)^{-1} \tag{2}$$

represents the proportion of black ball in the urn at time n and the sequence $(V_n)_{n \geq 0}$ also converges in L^1 and almost surely to V . We shall prove the following limit theorem.

*crimaldi@dm.unibo.it

Theorem 1.1. *Using the previous notation, let*

$$Q_n := (\sum_{i=1}^{n+1} N_i)^{-1} N_{n+1} \quad \text{and} \quad W_n := \sqrt{n}(V_n - V).$$

Further, let us set $\mathcal{G}_n := \sigma(Y_1, N_1, \dots, Y_n, N_n, N_{n+1})$ and denote by K_n a version of the conditional distribution of W_n given \mathcal{G}_n .

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) $\sum_{k \geq 0} k^2 \mathbb{E}[Q_k^4] < \infty$.*

Then, for almost every ω in Ω , the sequence $(K_n(\omega, \cdot))_n$ of probability measures converges weakly to the Gaussian distribution

$$\mathcal{N}(0, H(\omega)(V(\omega) - V^2(\omega))).$$

More briefly, the statement of the above theorem can be so reformulated: with respect to the conditioning system $\mathcal{G} = (\mathcal{G}_n)_n$, the sequence $(W_n)_n$ converges to the Gaussian kernel

$$\mathcal{N}(0, H(V - V^2))$$

in the sense of the *almost sure conditional convergence* (see Sec. 2).

It may be worthwhile to underline two features of Theorem 1.1: first of all, we allow that the distribution of N_n arbitrarily depends on n and, secondly, the obtained convergence is quite strong (indeed, it implies not only stable convergence, but also stable convergence in strong sense).

The paper is so structured. Section 2 contains some general asymptotic results for martingales and sequences of conditionally identically distributed random variables (in the sense of Berti et al. (2004)). They follow from Proposition A.1 in appendix, which is a result of almost sure conditional convergence toward a Gaussian kernel for a family of martingales. This proposition is closely related in many aspects to known results about stable convergence (see Crimaldi-Pratelli (2005) and Hall-Heyde (1980)) and strong stable convergence (see Crimaldi et al. (2007)), but the novelty lies in the fact that the convergence of the conditional distributions/expectations holds almost surely. In section 3, we prove Theorem 1.1 and, finally, in section 4, we consider some special cases.

2 Preliminary results

In the sequel (Ω, \mathcal{A}, P) will denote a probability space and we shall briefly call a *kernel* a family $K = (K(\omega, \cdot))_{\omega \in \Omega}$ of probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that, for each bounded Borel function f on \mathbb{R} , the function Kf defined on Ω by

$$Kf(\omega) := \int f(x) K(\omega, dx)$$

is measurable with respect to \mathcal{A} .

In particular, given on (Ω, \mathcal{A}, P) a real random variable M and a positive real random variable V , the family $(\mathcal{N}(M(\omega), V(\omega)))_{\omega \in \Omega}$, where $\mathcal{N}(M(\omega), V(\omega))$ denotes the Gaussian distribution with mean $M(\omega)$ and variance $V(\omega)$, is a kernel, which will be said *Gaussian* and

denoted by $\mathcal{N}(M, V)$. Moreover, if X is a real random variable on (Ω, \mathcal{A}, P) and \mathcal{U} is a sub- σ -field of \mathcal{A} , a version of the *conditional distribution* of X given \mathcal{U} is a kernel K such that, for each bounded Borel function f on \mathbb{R} , the random variable Kf is a version of the conditional expectation $E[f(X) | \mathcal{U}]$.

Further, we shall call a *conditioning system* a sequence $\mathcal{G} = (\mathcal{G}_n)_n$ of sub- σ -fields of \mathcal{A} . In particular, a filtration is an increasing conditioning system.

Using the above terminology, we can give the following definition.

Definition 2.1. Given a sequence $(X_n)_n$ of real random variables on (Ω, \mathcal{A}, P) and a conditioning system \mathcal{G} , let us denote by K_n a version of the conditional distribution of X_n given \mathcal{G}_n . If K is a kernel such that, for almost every ω in Ω , the sequence $(K_n(\omega, \cdot))_n$ of probability measures on \mathbb{R} converges weakly to the probability measure $K(\omega, \cdot)$, then we shall say that, with respect to the conditioning system \mathcal{G} , the sequence $(X_n)_n$ converges to K in the sense of the *almost sure convergence of the conditional distributions* or, more briefly, in the sense of the *almost sure conditional convergence*.

If a sequence $(X_n)_n$ converges to a kernel K in the sense of the almost sure conditional convergence with respect to a conditioning system \mathcal{G} , then the conditional expectation $E[f(X_n) | \mathcal{G}_n]$ converges almost surely to the random variable Kf , for each bounded continuous function f on \mathbb{R} . Therefore, the sequence $(X_n)_n$ converges to K \mathcal{G} -stably in the strong sense (see Def. 4, sec. 4 in Crimaldi et al. (2007)). In particular, $(X_n)_n$ converges in distribution to the distribution PK defined by $PK(B) = \int K(\omega, B) P(d\omega)$, for each Borel set B of \mathbb{R} .

By Proposition A.1 in appendix, we obtain some general results of almost sure conditional convergence for martingales and conditionally identically distributed sequences of random variables.

Proposition 2.2. On (Ω, \mathcal{A}, P) , let $(V_n)_{n \in \mathbb{N}}$ be a real martingale with respect to a filtration $\mathcal{G} = (\mathcal{G}_n)_{n \in \mathbb{N}}$. Suppose that $(V_n)_n$ converges in L^1 to a random variable V . Moreover, setting

$$U_n := n \sum_{k \geq n} (V_k - V_{k+1})^2, \quad Y := \sup_k \sqrt{k} |V_k - V_{k+1}|, \quad (3)$$

assume that the following conditions hold:

- (a) The random variable Y 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

$$W_n := \sqrt{n}(V_n - V) \quad (4)$$

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

Proof. In order to apply Proposition A.1, let us consider, for each $n \geq 1$, the filtration $(\mathcal{F}_{n,h})_{h \in \mathbb{N}}$ and the process $(M_{n,h})_{h \in \mathbb{N}}$ defined by

$$\mathcal{F}_{n,0} = \mathcal{F}_{n,1} := \mathcal{G}_n, \quad M_{n,0} = M_{n,1} := 0$$

and, for $h \geq 2$,

$$\mathcal{F}_{n,h} := \mathcal{G}_{n+h-1}, \quad M_{n,h} := \sqrt{n}(V_n - V_{n+h-1}).$$

It is easy to verify that, with respect to $(\mathcal{F}_{n,h})_{h \geq 0}$, the process $(M_{n,h})_{h \geq 0}$ is a martingale which converges in L^1 to the random variable $M_{n,\infty} := W_n$. In addition, the increment $X_{n,j} := M_{n,j} - M_{n,j-1}$ is equal to zero for $j = 1$ and, for $j \geq 2$, it coincides with a random variable of the form $\sqrt{n}(V_k - V_{k+1})$ with $k \geq n$. Therefore, we have

$$\sum_{j \geq 1} X_{n,j}^2 = U_n$$

where U_n is the random variable defined in (3), and

$$X_n^* := \sup_{j \geq 1} |X_{n,j}| = \sqrt{n} \sup_{k \geq n} |V_k - V_{k+1}| \leq \sup_{k \geq n} \sqrt{k} |V_k - V_{k+1}| \leq Y. \quad (5)$$

Moreover, the relation

$$n(V_n - V_{n+1})^2 = U_n - \frac{n}{n+1} U_{n+1} \xrightarrow{a.s.} 0$$

easily implies the convergence $\sup_{k \geq n} \sqrt{k} |V_k - V_{k+1}| \xrightarrow{a.s.} 0$ and so, by (5), it also implies the convergence $X_n^* \xrightarrow{a.s.} 0$. Hence, if we take $k_n = 1$ for each n and \mathcal{U} equal to the completion (in \mathcal{A}) of the σ -field $\bigvee_n \mathcal{G}_n$, then the conditioning system $(\mathcal{F}_{n,k_n})_n$ coincides with the filtration \mathcal{G} and the assumptions of Proposition A.1 are satisfied. The proof is thus concluded. \square

In order to state the next result, we need the following notion, which was introduced in Berti et al. (2004) as an extension of the classical notion of exchangeability.

Definition 2.3. Given a filtration $\mathcal{G} = (\mathcal{G}_n)_{n \in \mathbb{N}}$ on (Ω, \mathcal{A}, P) , a sequence $(Y_n)_{n \geq 1}$ of real random variables on (Ω, \mathcal{A}, P) is said to be *conditionally identically distributed with respect to \mathcal{G}* or, more briefly, *\mathcal{G} -conditionally identically distributed* if it is adapted to \mathcal{G} and such that, for each fixed $n \geq 0$, all the random variables of the form Y_{n+j} with $j \geq 1$ have the same conditional distribution given \mathcal{G}_n .

It is obvious that exchangeable sequences are conditionally identically distributed with respect to their natural filtration. If $(Y_n)_{n \geq 1}$ is conditionally identically distributed with respect to a filtration \mathcal{G} and each random variable Y_n is integrable, then (see Berti et al. (2004) for details) the sequence $(M_n)_{n \geq 1}$ of the empirical means defined by (1) converges in L^1 and almost surely to a random variable V (which, in the particular case of an exchangeable sequence, coincides with a version of the conditional expectation of Y_n given the tail σ -field of $(Y_n)_n$). Further, the sequence $(V_n)_{n \in \mathbb{N}}$ defined by

$$V_n := \mathbb{E}[Y_{n+1} | \mathcal{G}_n] \quad (6)$$

coincides with the martingale which is closed by V ; that is, for each n , we have the equality

$$V_n = \mathbb{E}[V | \mathcal{G}_n]. \quad (7)$$

From Proposition 2.2, we get the following corollaries.

Corollary 2.4. *Let \mathcal{G} be a filtration and $(Y_n)_{n \geq 1}$ a \mathcal{G} -conditionally identically distributed sequence of integrable random variables on (Ω, \mathcal{A}, P) . Let $(V_n)_{n \in \mathbb{N}}$ be the martingale defined by (6). Using notation (3), assume that conditions (a) and (b) of Proposition 2.2 hold.*

Then, with respect to \mathcal{G} , the sequence $(W_n)_n$ of random variables defined by (4) converges to the Gaussian kernel $\mathcal{N}(0, U)$ in the sense of the almost sure conditional convergence.

Proof. It is sufficient to apply Proposition 2.2 to the martingale $(V_n)_n$.

Corollary 2.5. *With the same notation and assumptions as in Corollary 2.4 and using notation (1), set*

$$X_n := \sqrt{n}(M_n - V), \quad Z_n := \sqrt{n}(M_n - V_n).$$

Suppose that the sequence $(Z_n)_n$ converges almost surely to a real random variable Z .

Then, with respect to \mathcal{G} , the sequence $(X_n)_n$ converges to the Gaussian kernel $\mathcal{N}(Z, U)$ in the sense of the almost sure conditional convergence.

Proof. By Lemma A.3 in the appendix, it will suffice to verify that, for each fixed t in \mathbb{R} , we have

$$\mathbb{E}[\exp(itX_n) | \mathcal{G}_n] \xrightarrow{a.s.} \exp(itZ - \frac{1}{2}t^2U).$$

To this end, we observe that we can write $X_n = Z_n + W_n$ where W_n is the random variable defined by (4). Thus, since Z_n is \mathcal{G}_n -measurable, we have

$$\mathbb{E}[\exp(itX_n) | \mathcal{G}_n] = \exp(itZ_n) \mathbb{E}[\exp(itW_n) | \mathcal{G}_n].$$

In order to conclude, it is enough to use the assumption of almost sure convergence of Z_n to Z and Corollary 2.4. \square

Remark 2.6. Under the assumptions of Corollary 2.4 (respectively, Corollary 2.5), since \mathcal{G} is increasing, by Cor. 5, sec. 5 in Crimaldi et al. (2007), the sequence (W_n) (respectively, (X_n)) also converges to the kernel $\mathcal{N}(0, U)$ (respectively, $\mathcal{N}(Z, U)$) \mathcal{G}_∞ -stably, and so \mathcal{A} -stably (but, in general, this stable convergence is not in the strong sense. See Cor. 4, sec. 4 in Crimaldi et al. (2007)).

3 Proof of Theorem 1.1

Let us start with the mathematical formalization of the experiment described in section 1.

Given a sequence $(\mu_n)_{n \geq 1}$ of probability measures on \mathbb{N}^* , using the Ionescu Tulcea theorem, it is possible to build a probability space (Ω, \mathcal{A}, P) and, on it, a sequence $(Y_n)_{n \geq 1}$ of random variables with values in $\{0, 1\}$ and a sequence $(N_n)_{n \geq 1}$ of random variables with values in \mathbb{N}^* such that the following conditions are satisfied for each $n \geq 0$:

(c1) A version of the conditional distribution of Y_{n+1} given the σ -field

$$\mathcal{F}_n := \sigma(Y_1, N_1, \dots, Y_n, N_n) \quad (\text{where } \mathcal{F}_0 := \{\emptyset, \Omega\})$$

is the kernel $(\mathcal{B}(1, V_n(\omega)))_{\omega \in \Omega}$, where $\mathcal{B}(1, p)$ denotes the Bernoulli distribution with parameter p and V_n is the random variable defined by (2) (with $V_0 = b(b+r)^{-1}$).

(c2) The distribution of the random variable N_{n+1} is μ_{n+1} and N_{n+1} is independent of the random vector $[Y_1, N_1, \dots, Y_n, N_n, Y_{n+1}]$.

By condition (c1), we have $\mathbb{E}[Y_{n+1} | \mathcal{F}_n] = V_n$ for each $n \geq 0$. Moreover, by this equality and condition (c2), we also have $\mathbb{E}[Y_{n+1} | \mathcal{G}_n] = V_n$ (where \mathcal{G}_n is the sub- σ -field defined in the statement of Theorem 1.1). Now, we observe that, if we set

$$S_n := b + r + \sum_{i=1}^n N_i, \tag{8}$$

we can write $V_{n+1} = S_{n+1}^{-1}(V_n S_n + Y_{n+1} N_{n+1})$ and we have

$$\mathbb{E}[V_{n+1} | \mathcal{G}_n] = S_{n+1}^{-1}(V_n S_n + V_n N_{n+1}) = V_n.$$

In other words, the sequence $(V_n)_{n \in \mathbb{N}}$ is a martingale with respect to the filtration $\mathcal{G} = (\mathcal{G}_n)_{n \in \mathbb{N}}$. This fact, since each random variable Y_n takes values in $\{0, 1\}$, implies (see Berti et al. 2004) that the sequence $(Y_n)_{n \geq 1}$ is \mathcal{G} -conditionally identically distributed. Therefore, the sequence $(M_n)_n$ of random variables defined by (1) converges in L^1 and almost surely to a random variable V such that (7) holds for each n .

We are now ready to prove our main theorem.

Proof of Theorem 1.1. Assuming notation (3), it will be sufficient to prove that the sequence $(V_n)_n$ satisfies conditions (a) and (b) of Proposition 2.2, with $U = H(V - V^2)$. To this end, we observe firstly that, after some calculations, we have

$$V_k - V_{k+1} = (V_k - Y_{k+1}) N_{k+1} (b + r + \sum_{i=1}^{k+1} N_i)^{-1}. \quad (9)$$

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

$$\sup_k k^2 |V_k - V_{k+1}|^4 \leq \sum_{k \geq 0} k^2 Q_k^4 \in L^1.$$

Furthermore, we have

$$N_{k+1} (b + r + \sum_{i=1}^{k+1} N_i)^{-1} \sim Q_k \quad \text{for } k \rightarrow +\infty,$$

and hence, by (9),

$$\sum_{k \geq n} (V_k - V_{k+1})^2 \sim \sum_{k \geq n} (V_k - Y_{k+1})^2 Q_k^2 \quad \text{for } n \rightarrow +\infty.$$

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

$$n \sum_{k \geq n} (V_k - Y_{k+1})^2 Q_k^2 \xrightarrow{a.s.} H(V - V^2).$$

Since we have $Y_{k+1}^2 = Y_{k+1}$, the above convergence can be rewritten as

$$n \sum_{k \geq n} (V_k^2 + Y_{k+1} - 2V_k Y_{k+1}) Q_k^2 \xrightarrow{a.s.} H(V - V^2). \quad (10)$$

Now, by assumption (i) and the almost sure convergence of $(V_k)_k$ to V , we have

$$n \sum_{k \geq n} V_k Q_k^2 \xrightarrow{a.s.} V H \quad (11)$$

$$n \sum_{k \geq n} V_k^2 Q_k^2 \xrightarrow{a.s.} V^2 H. \quad (12)$$

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

$$n \sum_{k \geq n} (Y_{k+1} - V_k) Q_k^2 \xrightarrow{a.s.} 0. \quad (13)$$

Indeed, from this and (11), we obtain

$$n \sum_{k \geq n} Y_{k+1} Q_k^2 \xrightarrow{a.s.} V H, \quad (14)$$

and so

$$n \sum_{k \geq n} V_k Y_{k+1} Q_k^2 \xrightarrow{a.s.} V^2 H. \quad (15)$$

Then convergence relations (12), (14) and (15) lead us to the desired relation (10).

In order to prove (13), we consider the process $(Z_n)_{n \in \mathbb{N}}$ defined by

$$Z_n := \sum_{k=0}^{n-1} k (Y_{k+1} - V_k) Q_k^2.$$

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

$$\mathbb{E}[Z_n^2] = \sum_{k=0}^{n-1} k^2 \mathbb{E}[(Y_{k+1} - V_k)^2 Q_k^4] \leq \sum_{k \geq 0} k^2 \mathbb{E}[Q_k^4] < \infty. \quad (16)$$

The martingale $(Z_n)_n$ is thus bounded in L^2 and so it converges almost surely; that is, the series

$$\sum_{k \geq 0} k (Y_{k+1} - V_k) Q_k^2$$

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 $n \sum_{k \geq n} k^{-1} a_k \rightarrow 0$ for $n \rightarrow +\infty$. Applying this result, we find (13) and the proof is so concluded. \square

4 Some special cases

In this section we describe some special cases in which the assumptions of Theorem 1.1 hold. First of all we may remark that, in Theorem 1.1, condition (ii) is obviously satisfied if the sequence $(N_n)_n$ is uniformly bounded by a random variable C with $\mathbb{E}[C^4] < \infty$. Indeed, we have $Q_k \leq C(k+1)^{-1}$. Furthermore, we have the following results.

Corollary 4.1. (*case i.i.d.*)

Using the notation of Theorem 1.1, suppose that the random variables N_n are identically distributed and $\mathbb{E}[N_1^4] < +\infty$. Set

$$m := \mathbb{E}[N_1], \quad \delta := \mathbb{E}[N_1^2], \quad h := \delta/m^2.$$

Then, with respect to \mathcal{G} , the sequence $(W_n)_n$ converges to the Gaussian kernel

$$\mathcal{N}(0, h(V - V^2))$$

in the sense of the almost sure conditional convergence.

Proof. It will suffice to verify that condition (i) and (ii) of Theorem 1.1 hold with $H = h$. With regard to condition (ii), it is enough to observe that, by the obvious inequality $Q_k \leq N_{k+1}/(k+1)$ and the identity in distribution of the random variables N_k , we have

$$\sum_{k \geq 0} k^2 \mathbb{E}[Q_k^4] \leq \sum_{k \geq 0} k^2 \mathbb{E}[N_{k+1}^4 / (k+1)^4] \leq \mathbb{E}[N_1^4] \sum_{k \geq 0} (k+1)^{-2} < \infty.$$

In order to prove condition (i) of Theorem 1.1 (with $H = h$), we observe that the series

$$\sum_k k^{-1} (N_{k+1}^2 - \delta)$$

is almost surely convergent: indeed, the random variables $X_k := k^{-1} (N_{k+1}^2 - \delta)$ are independent, centered and square-integrable, with $\text{Var}[X_k] = k^{-2} \text{Var}[N_1^2]$. Therefore, by the above mentioned Abel's result, we obtain the almost sure convergence of the series

$$\sum_k k^{-2} (N_{k+1}^2 - \delta)$$

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

$$n \sum_{k \geq n} k^{-2} (N_{k+1}^2 - \delta) \xrightarrow{a.s.} 0.$$

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

$$n \sum_{k \geq n} k^{-2} N_{k+1}^2 \xrightarrow{a.s.} \delta.$$

Now, we observe that, by the strong law of large numbers, we have for $k \rightarrow +\infty$

$$\sum_{i=1}^{k+1} N_i \stackrel{a.s.}{\sim} m(k+1) \sim mk,$$

and so

$$Q_k^2 \stackrel{a.s.}{\sim} m^{-2} k^{-2} N_{k+1}^2.$$

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

$$n \sum_{k \geq n} Q_k^2 \stackrel{a.s.}{\sim} m^{-2} n \sum_{k \geq n} k^{-2} N_{k+1}^2 \xrightarrow{a.s.} m^{-2} \delta = h.$$

Condition (i) of Theorem 1.1 (with $H = h$) is thus proved and the proof is concluded. \square

Corollary 4.2. (*Classical Pólya urn*)

Using the notation of Corollary 4.1, suppose that the random variables N_n are all equal to a constant $c \in \mathbb{N}^*$. Set $X_n := \sqrt{n}(M_n - V)$, where $(M_n)_n$ is the sequence defined by (1).

Then, with respect to \mathcal{G} , each of the two sequences $(W_n)_n, (X_n)_n$ converges to the Gaussian kernel $\mathcal{N}(0, V - V^2)$ in the sense of the almost sure conditional convergence.

Proof. The assumptions of Corollary 4.1 are obviously fulfilled with $h = 1$. It follows the desired convergence for $(W_n)_n$. Finally, after some calculations, we get

$$\sqrt{n}|M_n - V_n| = \sqrt{n}(b+r+nc)^{-1}|(b+r)M_n - b| \leq (c\sqrt{n})^{-1}(2b+r) \rightarrow 0 \quad \text{for } n \rightarrow +\infty.$$

Thus, by Corollary 2.5, we also obtain the desired convergence for $(X_n)_n$. \square

A Appendix

Given a conditioning system $\mathcal{G} = (\mathcal{G}_n)_n$, if \mathcal{U} is a sub- σ -field of \mathcal{A} such that, for each real integrable random variable Y , the conditional expectation $E[Y | \mathcal{G}_n]$ converges almost surely to the conditional expectation $E[Y | \mathcal{U}]$, then we shall briefly say that \mathcal{U} is an *asymptotic σ -field* for \mathcal{G} . In order that there exists an asymptotic σ -field \mathcal{U} for a given conditioning system \mathcal{G} , it is obviously sufficient that the sequence $(\mathcal{G}_n)_n$ is increasing or decreasing. (Indeed we can take $\mathcal{U} = \bigvee_n \mathcal{G}_n$ in the first case and $\mathcal{U} = \bigcap_n \mathcal{G}_n$ in the second one.)

We are going to prove the following general result.

Proposition A.1. *On (Ω, \mathcal{A}, P) , for each $n \geq 1$, let $(\mathcal{F}_{n,h})_{h \in \mathbb{N}}$ be a filtration and $(M_{n,h})_{h \in \mathbb{N}}$ a real martingale with respect to $(\mathcal{F}_{n,h})_{h \in \mathbb{N}}$, with $M_{n,0} = 0$, which converges in L^1 to a random variable $M_{n,\infty}$. Set*

$$X_{n,j} := M_{n,j} - M_{n,j-1} \quad \text{for } j \geq 1, \quad U_n := \sum_{j \geq 1} X_{n,j}^2, \quad X_n^* := \sup_{j \geq 1} |X_{n,j}|.$$

Further, let $(k_n)_{n \geq 1}$ be a sequence of strictly positive integers such that $k_n X_n^* \xrightarrow{a.s.} 0$ and let \mathcal{U} be a sub- σ -field which is asymptotic for the conditioning system \mathcal{G} defined by $\mathcal{G}_n := \mathcal{F}_{n,k_n}$. Assume that the sequence $(X_n^*)_n$ is dominated in L^1 and that the sequence $(U_n)_n$ converges almost surely to a positive real random variable U which is measurable with respect to \mathcal{U} .

Then, with respect to the conditioning system \mathcal{G} , the sequence $(M_{n,\infty})_n$ converges to the Gaussian kernel $\mathcal{N}(0, U)$ in the sense of the almost sure conditional convergence.

For the proof of the above proposition, we need some lemmas. The first one is an immediate extension of Theorem 2 in Blackwell-Dubins (1962).

Lemma A.2. *Let \mathcal{G} be a conditioning system and \mathcal{U} a sub- σ -field of \mathcal{A} . Then the following conditions are equivalent:*

(a) *The sub- σ -field \mathcal{U} is an asymptotic σ -field for \mathcal{G} .*

(b) *For each sequence $(Y_n)_n$ of integrable real random variables such that there exists an integrable real random variable Z with $Y_n \leq Z$ for each n and such that the random variable $\limsup_n Y_n$ is integrable, we have*

$$\limsup_n \mathbb{E}[Y_n | \mathcal{G}_n] \leq \mathbb{E}[\limsup_n Y_n | \mathcal{U}] \quad \text{a.s.}$$

(c) *For each sequence $(Y_n)_n$ of integrable real random variables such that there exists an integrable real random variable Z with $Y_n \geq Z$ for each n and such that the random variable $\liminf_n Y_n$ is integrable, we have*

$$\mathbb{E}[\liminf_n Y_n | \mathcal{U}] \leq \liminf_n \mathbb{E}[Y_n | \mathcal{G}_n] \quad \text{a.s.}$$

(d) *For each sequence $(Y_n)_n$ of integrable complex random variables, which is dominated in L^1 and which converges almost surely to a complex random variable Y , the conditional expectation $\mathbb{E}[Y_n | \mathcal{G}_n]$ converges almost surely to the conditional expectation $\mathbb{E}[Y | \mathcal{U}]$.*

If K is a kernel, then, for each ω in Ω , we shall denote by $\widehat{K}(\omega, \cdot)$ the characteristic function of the probability measure $K(\omega, \cdot)$; that is, for each real number t , we shall set

$$\widehat{K}(\omega, t) := \int \exp(itx) K(\omega, dx).$$

Since the complex function $(\omega, t) \mapsto \widehat{K}(\omega, t)$ is measurable with respect to ω and continuous with respect to t , it is measurable on $(\Omega \times \mathbb{R}, \mathcal{A} \otimes \mathcal{B}(\mathbb{R}))$. With this notation, we can now state the second lemma.

Lemma A.3. *Let K be a kernel and $(K_n)_n$ a sequence of kernels. Then the following conditions are equivalent:*

(a) *For almost every ω in Ω , the sequence of probability measures $(K_n(\omega, \cdot))_n$ converges weakly to the probability measure $K(\omega, \cdot)$.*

(b) *For each fixed real number t , the sequence of complex random variables $(\widehat{K}_n(\cdot, t))_n$ converges almost surely to $\widehat{K}(\cdot, t)$.*

Proof. Implication (a) \Rightarrow (b) is obvious. In order to prove implication (b) \Rightarrow (a), we assume condition (b) and we set

$$A := \{(\omega, t) \in \Omega \times \mathbb{R} : \limsup_n |\widehat{K}_n(\omega, t) - \widehat{K}(\omega, t)| \neq 0\}.$$

Then, for each fixed real number t , the section $A^{-1}(t) = \{\omega : (\omega, t) \in A\}$ is negligible under P . Therefore the set A is negligible under the product measure $P \otimes \lambda$, where λ denotes the Lebesgue measure on \mathbb{R} . It follows that, for almost every ω in Ω , the section $A(\omega) = \{t : (\omega, t) \in A\}$ is negligible under λ . For such an ω , we have $\widehat{K}_n(\omega, \cdot) \rightarrow \widehat{K}(\omega, \cdot)$ almost everywhere with respect to λ . It is well-known that this fact suffices to assure the weak convergence of $(K_n(\omega, \cdot))_n$ to $K(\omega, \cdot)$. \square

Finally, the following lemma is proved in Crimaldi et al. (2007, Lemma 1, sec. 7).

Lemma A.4. Given a finite family $(X_j)_j$ of real random variables on (Ω, \mathcal{A}, P) , let

$$S := \sum_j X_j, \quad U := \sum_j X_j^2, \quad X^* := \sup_j |X_j|.$$

Further, given two real numbers b and t , with $b > 0$, and a random variable V with values in $[0, b]$, set

$$L := \prod_j (1 + itX_j), \quad D := \exp(itS) - L \exp(-\frac{1}{2}t^2V), \quad B := \{|t|X^* \leq 1, U \leq b\}.$$

Then, on the set B , we have

$$|D| \leq \kappa(b, t)(|U - V| + 2b|t|X^*), \quad \text{with} \quad \kappa(b, t) := \frac{1}{2}t^2 \exp(\frac{7}{2}bt^2).$$

We are now in a position to prove Proposition A.1. The following proof is substantially an adaptation of the proof of Theorem 5 in Crimaldi et al. (2007).

Proof of Proposition A.1. In view of Lemma A.3 and of the assumed almost sure convergence of U_n to U , it is sufficient to prove that, for each real number t , we have

$$\mathbb{E}[\exp(itM_{n,\infty}) | \mathcal{G}_n] - \exp(-\frac{1}{2}t^2U_n) \xrightarrow{a.s.} 0. \quad (17)$$

For $t = 0$, there is nothing to prove and so we shall assume $t \neq 0$. We note that, for each fixed $n \geq 1$, the sequence $(M_{n,h})_{h \geq 0}$ converges almost surely to $M_{n,\infty}$ and the (increasing) sequence $(U_{n,h})_{h \geq 0}$ defined by

$$U_{n,h} := \sum_{j=1}^h X_{n,j}^2$$

converges everywhere to U_n . Therefore, it is possible to choose, for each fixed n , an integer $l_n \geq k_n$ in such a way that, setting

$$A_n := \{ |M_{n,\infty} - M_{n,l_n}| \vee |U_n - U_{n,l_n}| > 1/n \},$$

we have $P(A_n) < 2^{-n}$. By Borel-Cantelli lemma, we have $P(\limsup_n A_n) = 0$ and so

$$M_{n,\infty} - M_{n,l_n} \xrightarrow{a.s.} 0, \quad U_n - U_{n,l_n} \xrightarrow{a.s.} 0.$$

Since the complex function $x \mapsto \exp(itx)$ is Lipschitz on \mathbb{R} and the real function $x \mapsto \exp(-\frac{1}{2}t^2x)$ is Lipschitz on \mathbb{R}_+ , we also have

$$\exp(itM_{n,\infty}) - \exp(itM_{n,l_n}) \xrightarrow{a.s.} 0, \quad (18)$$

$$\exp(-\frac{1}{2}t^2U_n) - \exp(-\frac{1}{2}t^2U_{n,l_n}) \xrightarrow{a.s.} 0. \quad (19)$$

Furthermore, since the conditioning system \mathcal{G} has an asymptotic σ -field, the relation (18) implies

$$\mathbb{E}[\exp(itM_{n,\infty}) - \exp(itM_{n,l_n}) | \mathcal{G}_n] \xrightarrow{a.s.} 0.$$

Thus, by this last fact and (19), if we set

$$Y_n := \mathbb{E}[\exp(itM_{n,l_n}) | \mathcal{G}_n] - \exp(-\frac{1}{2}t^2U_{n,l_n}),$$

the desired relation (17) is equivalent to the following one:

$$Y_n \xrightarrow{a.s.} 0. \quad (20)$$

In order to prove this last convergence, fixing a positive number a , let us define, for each $n \geq 1$, the stopping time J_n (with respect to the filtration $(\mathcal{F}_{n,h})_{0 \leq h \leq l_n}$) in the following way:

$$J_n(\omega) := l_n \wedge \inf\{h \in \mathbb{N} : h \leq l_n, U_{n,h}(\omega) \geq a\}.$$

We observe that the absolute values of the two differences

$$\exp(itM_{n,l_n}) - \exp(itM_{n,J_n}), \quad \exp(-\frac{1}{2}t^2 U_{n,l_n}) - \exp(-\frac{1}{2}t^2 (U_{n,l_n} \wedge a))$$

are bounded by constants 2 and 1 respectively and, by the definition of J_n , they vanish on the event $\{U_{n,l_n} < a\}$. Hence, if we set

$$Z_n := \mathbb{E}[\exp(itM_{n,J_n}) | \mathcal{G}_n] - \exp(-\frac{1}{2}t^2 (U_{n,l_n} \wedge a)), \quad (21)$$

we can write

$$|Y_n - Z_n| \leq 2\mathbb{E}[I_{\{U_{n,l_n} \geq a\}} | \mathcal{G}_n] + I_{\{U_{n,l_n} \geq a\}}.$$

Moreover, since $U_{n,l_n} \xrightarrow{a.s.} U$, we have

$$\limsup_n I_{\{U_{n,l_n} \geq a\}} \leq I_{\{U \geq a\}} \quad \text{a.s.},$$

and, since \mathcal{U} is asymptotic for \mathcal{G} and U is measurable with respect to \mathcal{U} , we find (using Lemma A.2)

$$\limsup_n |Y_n - Z_n| \leq 2\mathbb{E}[I_{\{U \geq a\}} | \mathcal{U}] + I_{\{U \geq a\}} = 3I_{\{U \geq a\}} \xrightarrow{a.s.} 0 \quad \text{for } a \rightarrow +\infty.$$

Summing up, it will suffice to prove that

$$Z_n \xrightarrow{a.s.} 0 \quad (22)$$

instead of (20). To this end, for each n , let us introduce the complex martingale $(L_{n,h})_{0 \leq h \leq l_n}$ (with respect to the filtration $(\mathcal{F}_{n,h})_{0 \leq h \leq l_n}$) defined as follows:

$$L_{n,0} := 1, \quad L_{n,h} := \prod_{j=1}^h (1 + itX_{n,j} I_{\{j \leq J_n\}}) \quad \text{for } 1 \leq h \leq l_n.$$

Since, for each $x \in \mathbb{R}$, we have $|1 + ix|^2 = 1 + x^2 \leq \exp(x^2)$, it follows that

$$1 \leq |L_{n,k_n}| \leq \exp[\frac{1}{2}t^2 (k_n X_n^*)^2] \xrightarrow{a.s.} 1, \quad (23)$$

$$|L_{n,J_n}| \leq \exp(\frac{1}{2}t^2 a)(1 + |tX_n^*|. \quad (24)$$

Moreover, we have

$$|\text{Arg}(L_{n,k_n})| \leq \sum_{j=1}^{k_n \wedge J_n} |\text{Arg}(1 + itX_{n,j})| = \sum_{j=1}^{k_n \wedge J_n} |\arctan(tX_{n,j})| \leq |t|k_n X_n^* \xrightarrow{a.s.} 0. \quad (25)$$

Since the martingale $(L_{n,h})_{0 \leq h \leq l_n}$ is stopped at J_n , we have $L_{n,J_n} = L_{n,l_n}$ and so

$$\mathbb{E}[L_{n,J_n} | \mathcal{G}_n] = \mathbb{E}[L_{n,l_n} | \mathcal{F}_{n,k_n}] = L_{n,k_n} \xrightarrow{a.s.} 1, \quad (26)$$

where the almost sure convergence to the constant 1 is a consequence of (23) and (25). Now, fix a positive number b with $b > a$ and set $V_n := \mathbb{E}[U \wedge b | \mathcal{G}_n]$. Then, since \mathcal{U} is asymptotic for \mathcal{G} and U is measurable with respect to \mathcal{U} , we have:

$$V_n \xrightarrow{a.s.} \mathbb{E}[U \wedge b | \mathcal{U}] = U \wedge b, \quad (27)$$

and so

$$V_n - (U_{n,l_n} \wedge b) \xrightarrow{a.s.} 0. \quad (28)$$

Further, set

$$B_n := \{|t|X_n^* \leq 1, X_n^* \leq \sqrt{b-a}\}, \quad D_n := \exp(itM_{n,J_n}) - L_{n,J_n} \exp(-\frac{1}{2}t^2V_n).$$

Since the assumption $k_n X_n^* \xrightarrow{a.s.} 0$ implies $X_n^* \xrightarrow{a.s.} 0$, we have $I_{B_n^c} \xrightarrow{a.s.} 0$. Moreover, by the definition of D_n and relation (24), we get

$$|D_n| \leq 1 + |L_{n,J_n}| \leq 1 + \exp(\frac{1}{2}t^2a)(1 + |t|X_n^*). \quad (29)$$

Therefore, since $(X_n^*)_n$ is dominated in L^1 and \mathcal{G} has an asymptotic σ -field, by Lemma A.2, we get

$$\mathbb{E}[|D_n| I_{B_n^c} | \mathcal{G}_n] \xrightarrow{a.s.} 0. \quad (30)$$

Moreover, since we have $B_n \subset \{|t|X_n^* \leq 1, U_{n,J_n} \leq b\}$, applying Lemma A.4 to the finite family $(X_{n,j} I_{\{j \leq J_n\}})_{1 \leq j \leq l_n}$, we find

$$|D_n| I_{B_n} \leq \kappa(b, t) (|(U_{n,J_n} \wedge b) - V_n| + 2b|t|X_n^*). \quad (31)$$

The positive random variable $(U_{n,l_n} \wedge b) - (U_{n,J_n} \wedge b)$ vanishes on $\{J_n = l_n\}$ and, at each point of $\{J_n \neq l_n\}$, it coincides with the difference of two elements of $[a, b]$. Thus, it is bounded by the constant $b - a$ and so, from (31), we get

$$|D_n| I_{B_n} \leq \kappa(b, t) (b - a + |(U_{n,l_n} \wedge b) - V_n| + 2b|t|X_n^*). \quad (32)$$

Then, from (28), (30) and Lemma A.2, we have a.s.

$$\limsup_n \mathbb{E}[|D_n| | \mathcal{G}_n] = \limsup_n \mathbb{E}[|D_n| I_{B_n} | \mathcal{G}_n] \leq \kappa(b, t) (b - a). \quad (33)$$

We now observe that, by equalities (26) and the measurability of V_n with respect to \mathcal{G}_n , we have

$$\mathbb{E}[D_n | \mathcal{G}_n] = \mathbb{E}[\exp(itM_{n,J_n}) | \mathcal{G}_n] - L_{n,k_n} \exp(-\frac{1}{2}t^2V_n),$$

and so the random variable Z_n defined by (21) can be expressed in the following way:

$$Z_n = \mathbb{E}[D_n | \mathcal{G}_n] + L_{n,k_n} \exp(-\frac{1}{2}t^2V_n) - \exp(-\frac{1}{2}t^2(U_{n,l_n} \wedge a)).$$

From this equality, letting n go to $+\infty$ and using (26), (27) and (33), we obtain

$$\limsup_n |Z_n| \leq \kappa(b, t) (b - a) + |\exp(-\frac{1}{2}t^2(U \wedge b)) - \exp(-\frac{1}{2}t^2(U \wedge a))|.$$

Finally, it suffices to let b decrease to a in order to obtain the desired relation (22). \square

Remark A.5. In the previous proof, the assumption that $(X_n^*)_n$ is dominated in L^1 is used only to get relations (30) and (33). Actually, it can be replaced by the assumption that $(X_n^*)_n$ is a sequence of integrable random variables such that $\mathbb{E}[X_n^* | \mathcal{G}_n] \xrightarrow{a.s.} 0$. This follows immediately from (29) and (32).

Remark A.6. Since Lemma A.3 can be extended to the multidimensional case, it is possible, by the same small trick employed for Theorem 5 in Crimaldi et al. (2007), to give a multidimensional version of Proposition A.1.

References

- Amerio, E., Muliere, P. and Secchi, P. (2004). Reinforced urn processes for modeling credit default distributions. *International Journal of Theoretical and Applied Finance*, vol. 7, No. 4, pp. 407-423.
- Blackwell, D. and Dubins, L. (1962). Merging of Opinions with Increasing Information. *Ann. Math. Stat.* 33(3), pp. 882-886.
- Berti, P., Pratelli, L. and Rigo, P. (2004). Limit Theorems for a Class of Identically Distributed Random Variables. *Ann. Probab.* 32(3A), pp. 2029-2052.
- Crimaldi, I., Letta, G. and Pratelli, L. (2007). A strong form of stable convergence. *Séminaire de Probabilités XL, Lecture Notes in Math.* vol. 1899, pp. 203-225.
- Crimaldi, I. and Pratelli, L. (2005). Convergence Results for Multivariate Martingales. *Stochastic Processes Appl.* 115(4), pp. 571-577.
- Dirienzo, A.G. (2000). Using urn models for the design of clinical trials. *Sankhya Series B* 62(1), pp. 43-69.
- EGGENBERGER, F. and PÓLYA, G. (1923). Über die Statistik verketteter Vorgänge. *Zeitschrift Angew. Math. Mech.* 3, pp. 279-289.
- Gouet, R. (1993). Martingale Functional Central Limit Theorems for a Generalized Pólya Urn. *Ann. Probab.* 21, No. 3, pp. 1624-1639.
- Hall, P. and Heyde, C.C. (1980). *Martingale limit theory and its application*. Academic Press.
- Janson, S. (2006). Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* 134, no. 3, pp. 417-452.
- Janson, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes Appl.* 110, pp. 177-245.
- Kotz, S., Mahmoud, H. and Robert, P. (2000). On generalized Pólya urn models. *Statistics & Probability Letters* 49, pp. 163-173.
- May, C., Paganoni, A.M., and Secchi, P. (2005). On a two-color generalized Pólya urn. *Metron* 63, No. 1, pp. 115-134.
- Moler, A., Plo, F. and San Miguel, M. (2002). Almost sure convergence of urn models in a random environment. *Journal of Mathematical Sciences*, vol. 111, No. 3, pp. 3566-3571.
- Muliere, P., Paganoni, A.M. and Secchi, P. (2006). A randomly reinforced urn. *Journal of Statistical Planning and Inference* 136, pp. 1853-1874.
- Paganoni, A.M. and Secchi, P. (2004). Interacting reinforced urn system. *Adv. Appl. Probab.* 36(3), pp. 791-804.
- Pemantle, R. (1991). When are Touchpoints Limits for Generalized Pólya URNS?. *Proceedings of American Mathematical Society*, vol. 113, No. 1, pp. 235-243.
- Pemantle, R. (1990). A time-dependent version of Pólya's urn. *Jour. Theor. Prob.* 3, pp. 627-637.
- Pólya, G. (1931). Sur quelques points de la théorie des probabilités. *Ann. Inst. Poincaré* 1, pp. 117-161.