



CHAPTER 18

**General Central Limit Theorems for Associated Sequences, by H. Sangaré and G.S. Lo**

**Harouna Sangaré.** Email : harounasangare@fast-usttb.edu.ml.

**Gane Samb Lo.** Email : gane-samb.lo@edu.ugb.sn, gslo@aust.edu.ng, ganesamblo@ganesamblo.net.

LERSTAD, Gaston Berger University, Saint-Louis, Sénégal (main affiliation), AUST - African University of Sciences and Technology, Abuja, Nigeria. Permanent address : 1178 Evanston Dr NW T3P 0J9, Calgary, Alberta, Canada.

**Abstract.** We provide general central limit theorems (*CLT*'s) for associated random variables (*rv*'s) following the approaches used by Newman (1980) and Olivera et al.(2012). Given some assumptions, a Lynderberg-Feller-Levy type theorem is stated. We next specify different particular *CLT* versions of associated sequences based on moment conditions. Comparisons and application are addressed too.

**Keywords.** Central limit theorems; weak convergence; Associated sequences; asymptotic Statistics Stationary sequences.

**AMS 2010 Mathematics Subject Classification.** 60FXX; 60F05; 60G11

*Cite the chapter as :*

Sangaré H. and Lo, G.S. (2018). General Central Limit Theorems for Associated Sequences. In *A Collection of Papers in Mathematics and Related Sciences, a festschrift in honour of the late Galaye Dia* (Editors : Seydi H., Lo G.S. and Diakhaby A.). Spas Editions, Euclid Series Book, pp. 321 – 342. Doi : 10.16929/sbs/2018.100-04-01

©Spas Editions, Saint-Louis - Calgary 2018. H. Seydi *et al.* (Eds.) *A Collection of Papers in Mathematics and Related Sciences, a festschrift in honour of the late Galaye Dia.* Doi : 10.16929/sbs/2018.100

## 1. Introduction

We consider the problem of the central limit theorem for associated sequences. This problem goes back to [Newman \(1980\)](#). Since then, a number of *CLT*'s, and strong laws of large numbers (*SLLN*'s) or weak laws of large numbers (*WLLN*'s), and invariance principles and laws of the iterated logarithm (*LIL*'s) have been provided in the recent literature by different authors. [Burton \*et al.\* \(1986\)](#) and [Dabrowski and Dehling \(1988\)](#) considered weakly associated random variables to establish invariance principles in the lines of [Newman and Wright \(1981\)](#), as well as Berry-Essen-type results and functional *LIL*'s. The weak convergence for empirical processes using associated sequences has been discussed by [Louhichi \(2000\)](#) and [Yu \(1993\)](#).

The most general *CLT* seems to be the one provided by [Cox and Grimmett \(1984\)](#) for arbitrary associated *rv*'s fulfilling a number of moment conditions and those given by [Oliveira \(2012\)](#).

This question arises in the active research field on the concept of association and its application in many sciences, especially in percolation theory in Physics and in Reliability. The books by [Rao \(2012\)](#) and the monograph by [Oliveira \(2012\)](#) present a review of these researches. The book of [Bulinski and Shashkin \(2007\)](#) treats random associated sequences and intensively uses properties of regularly varying functions and provides *CTL*'s, *LLN*'s, *LIL*'s and invariance principles.

Although many results concerning the *CLT* problem are available for such sequences, there are still a number of open problems, especially regarding nonstationary sequences.

Here, we intend to provide more general *CLT*'s for arbitrary associated sequences. Precisely, here, we want to use all the power of the Newman method and express the conditions in the most general frame based on moment conditions so that any other result might be derived from them. In such a way, a Lyapounov-Feller-Levy type Theorem will be possible to be stated given some general assumptions. From this approach, more general *CLT*'s may be conceived only by turning back the Newman's method.

The paper is organized as follows. Since association is the central notion used here, we first make a quick reminder of it in Section 2. In Section

3, we make a round up of CLT's available in the literature with the aim of comparing them to our findings. In Section 4, we state our general CLT versions for arbitrary associated  $rv$ 's and give comparison study and concluding remarks.

## 2. A brief Reminder of Association

The concept of association has been introduced by Lehman (1966) in the bivariate case, and extended later in the multivariate case by Esary *et al.* (1967). The concept of association for random variables generalizes that of independence and seems to model a great variety of stochastic models.

This property also arises in Physics, and is quoted under the name of FKG property Fortuin *et al.* (1971)), in percolation theory and even in Finance (see Jiazhu (2002)).

The definite definition is given by Esary *et al.* (1967) as follows.

DEFINITION 5. A finite sequence of  $rv$ 's  $(X_1, \dots, X_n)$  is associated if, for any couple of real and coordinate-wise non-decreasing functions  $h$  and  $g$  defined on  $\mathbb{R}^n$ , we have

$$(2.1) \quad Cov(h(X_1, \dots, X_n), g(X_1, \dots, X_n)) \geq 0,$$

whenever the covariance exists. An infinite sequence of  $rv$ 's are associated whenever all its finite sub-sequences are associated.

We have a few number of interesting properties to be found in (Rao (2012)) :

**(P1)** A sequence of independent  $rv$ 's is associated.

**(P2)** Partial sums of associated  $rv$ 's are associated.

**(P3)** Order statistics of independent  $rv$ 's are associated.

**(P4)** Non-decreasing functions and non-increasing functions of associated variables are associated.

**(P5)** Let the sequence  $Z_1, Z_2, \dots, Z_n$  be associated and let  $(a_i)_{1 \leq i \leq n}$  be positive numbers and  $(b_i)_{1 \leq i \leq n}$  real numbers. Then the  $rv$ 's  $a_i(Z_i - b_i)$  are associated.

As immediate other examples of associated sequences, we may cite Gaussian random vectors with non-negatively correlated components (see [Pitt \(1982\)](#)) and a homogeneous Markov chain (see [Daley \(1968\)](#)).

Demimartingales are set from associated centered variables exactly as martingales are derived from partial sums of centered independent random variables. We have

**DEFINITION 6.** *A sequence of rv's  $\{S_n, n \geq 1\}$  in  $L^1(\Omega, \mathcal{A}, \mathbb{P})$  is a demimartingale when for any  $j \geq 1$ , for any coordinate-wise non-decreasing function  $g$  defined on  $\mathbb{R}^j$ , we have*

$$(2.2) \quad \mathbb{E}((S_{j+1} - S_j) g(S_1, \dots, S_j)) \geq 0, \quad j \geq 1.$$

Two particular cases should be highlighted. First any martingale is a demimartingale. Secondly, partial sums  $S_0 = 0, S_n = X_1 + \dots + X_n, n \geq 1$ , of associated and centered random variables  $X_1, X_2, \dots$  are demimartingales. In this case, (2.2) becomes :

$$\mathbb{E}((S_{j+1} - S_j) g(S_1, \dots, S_j)) = \mathbb{E}(X_{j+1} g(S_1, \dots, S_j)) = Cov(X_{j+1}, g(S_1, \dots, S_j)),$$

since  $\mathbb{E}X_{j+1} = 0$ . Since  $(x_1, \dots, x_{j+1}) \mapsto x_{j+1}$  and  $(x_1, \dots, x_{j+1}) \mapsto g(x_1, \dots, x_j)$  are coordinate-wise non-decreasing functions and since the  $X_1, X_2, \dots$  are associated, we get

$$\mathbb{E}((S_{j+1} - S_j) g(S_1, \dots, S_j)) = Cov(X_{j+1} g(S_1, \dots, S_j)) \geq 0.$$

Finally, we present the following key results for associated sequences that one can find in almost any paper on that topic and that we need for our proofs. A detailed review on these results is given in [Sangharé and Lo \(2016\)](#).

**LEMMA 29** (Hoeffding (1940) (see [Rao \(2012\)](#))). *Let  $(X, Y)$  be a bivariate random vector such that  $\mathbb{E}(X^2) < +\infty$  and  $\mathbb{E}(Y^2) < +\infty$ . If  $(X_1, Y_1)$  and  $(X_2, Y_2)$  are two independent copies of  $(X, Y)$ , then we have*

$$2Cov(X, Y) = \mathbb{E}(X_1 - X_2)(Y_1 - Y_2).$$

We also have

$$Cov(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(x, y) dx dy,$$

where

$$H(x, y) = \mathbb{P}(X > x, Y > y) - \mathbb{P}(X > x)\mathbb{P}(Y > y).$$

LEMMA 30 (Newman (1980)). Suppose that  $X, Y$  are two random variables with finite variance and,  $f$  and  $g$  are  $\mathbb{C}^1$  complex valued functions on  $\mathbb{R}^1$  with bounded derivatives  $f'$  and  $g'$ . Then

$$|\text{Cov}(f(X), g(Y))| \leq \|f'\|_\infty \|g'\|_\infty \text{Cov}(X, Y).$$

The following lemma is the most used tool in this field.

LEMMA 31 (Newman and Wright (1981) Theorem, see Newman and Wright (1981)). Let  $X_1, X_2, \dots, X_n$  be associated, then we have for all  $t = (t_1, \dots, t_n) \in \mathbb{R}^k$ ,

$$(2.3) \quad \left| \psi_{(X_1, X_2, \dots, X_n)}(t) - \prod_{i=1}^n \psi_{X_i}(t_i) \right| \leq \frac{1}{2} \sum_{1 \leq i \neq j \leq n} |t_i t_j| \text{Cov}(X_i, X_j).$$

Before we proceed any further, let us make a round up of CLT's for associated sequences in stationary and non-stationary cases.

### 3. Central limit theorem for associated sequences

Let  $X_1, X_2, \dots, X_n$  be an associated sequence of mean-zero random variables defined on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Define for each  $n \geq 1$ ,  $S_n = X_1 + \dots + X_n$ . The CLT question for stationary associated sequence turns around Newman (1980) results in which is proved that  $S_n/\sqrt{n}$  converges to a normal random variable  $\mathcal{N}(0, \sigma^2)$  when

$$\sigma^2 = \text{Var}(X_1) + 2 \sum_{j=2}^{\infty} \text{Cov}(X_1, X_j) < +\infty.$$

And in such a situation,

$$s_n^2 = \text{Var}(S_n/\sqrt{n}) \rightarrow \sigma^2 \text{ as } n \rightarrow +\infty.$$

A number of invariance principles and other CLT's are available but they are generally adaptations of this Newman result. As to the general case, (see Cox and Grimmett (1984)), did not consider stationarity in their results which used triangular sequences. Formulated for simple sequences,

their result is that  $S_n/(s_n\sqrt{n})$  weakly converges to a normal random variable  $\mathcal{N}(0, 1)$  if  $\text{Var}(X_n)$  is asymptotically bounded against zero, and the sequence of the third moments  $\mathbb{E}|X_n|^3$  is bounded and there exists a function  $u(r)$ ,  $r \in \{0, 1, \dots\}$  such that  $u(r) \rightarrow 0$  as  $r \rightarrow +\infty$  and such for all  $k \geq 1$ , and all  $r \geq 0$

$$\sum_{j:|k-j|\geq r} \text{Cov}(X_j, X_k) \leq u(r).$$

Let us recall their *CLT* as follows

**THEOREM 48.** *Let  $X_1, X_2, \dots, X_n$  be an associated sequence of mean-zero random variables defined on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Suppose there exist positive and finite constants  $c_1$  and  $c_2$  such that*

$$(3.1) \quad \text{Var}(X_j) \geq c_1 \text{ and } \mathbb{E}|X_j|^3 \leq c_2 \text{ for all } j \geq 1,$$

*and there is a function  $u(r)$  of  $r \in \mathbb{N}$  such that  $u(r) \rightarrow 0$  as  $r \rightarrow +\infty$  and for any  $r \geq 1$ ,*

$$(3.2) \quad \sup_{j \geq 1} \sum_{i:|j-i|\geq r} \text{cov}(X_i, X_j) \leq u(r).$$

Then

$$S_n/s_n \rightsquigarrow N(0, 1) \text{ as } n \rightarrow +\infty.$$

*where throughout the text, the symbol  $\rightsquigarrow$  stands for the weak convergence .*

**Oliveira (2012)** has proved general *CLT*'s, still using the Newman approach.

First, they obtained

**THEOREM 49** (see **Oliveira (2012)**, page 105, Theorem 4.4). *Let  $X_n$ ,  $n \in \mathbb{N}$ , be centered, square-integrable and associated random variables. For each  $n \in \mathbb{N}$ , let  $\ell_n \in \mathbb{N}$  and  $m_n = \left\lceil \frac{n}{\ell_n} \right\rceil$ . Define, for  $j = 1, \dots, m_n$ ,  $Y_{j,\ell_n} = \sum_{i=(j-1)\ell_n+1}^{j\ell_n} X_i$  and  $Y_{m_n+1,\ell_n} = \sum_{i=m_n\ell_n+1}^n X_i$ . Assume that  $m_n \rightarrow +\infty$ , and that*

$$(3.3) \quad \frac{1}{s_n^2} \sum_{j=1}^{m_n} \text{Var} (Y_{j,\ell_n}) \rightarrow 1,$$

$$(3.4) \quad \left| \mathbb{E} \exp \left( \frac{it}{s_n} S_n \right) - \prod_{j=1}^{m_n} \mathbb{E} \exp \left( \frac{it}{s_n} Y_{j,\ell_n} \right) \right| \rightarrow 0, \quad t \in \mathbb{R},$$

and

$$(3.5) \quad \forall \varepsilon > 0, \quad \frac{1}{s_n^2} \sum_{j=1}^{m_n} \int_{\{|Y_{j,\ell_n}| \geq \varepsilon s_n\}} Y_{j,\ell_n} d\mathbb{P} \rightarrow 0.$$

Then

$$\frac{1}{s_n} S_n \rightsquigarrow \mathcal{N}(0, 1),$$

Next, he obtained the following result using a Feller-Levy condition.

**THEOREM 50** (see [Oliveira \(2012\)](#), page 108, Theorem 4.8). *Let  $X_n, n \in \mathbb{N}$ , be centered, square-integrable and associated random variables. Assume that*

$$(3.6) \quad u(n) \rightarrow 0, \quad u(1) < +\infty,$$

$$(3.7) \quad \inf_{n \in \mathbb{N}} \frac{1}{n} s_n^2 > 0,$$

$$(3.8) \quad \forall \varepsilon > 0, \quad \frac{1}{s_n^2} \sum_{j=1}^{m_n} \int_{\{|X_j| \geq \varepsilon s_n\}} X_j^2 d\mathbb{P} \rightarrow 0.$$

Then

$$\frac{1}{s_n} S_n \rightsquigarrow \mathcal{N}(0, 1).$$

**Remark on whether the assumptions of the theorem are enough to get the CLT.** It seems to us that the conditions given by this theorem are not enough, as we tried to show it in Subsection 4.2 below. We think that the following assumption, denoted *(Hab)* below,

$$\frac{1}{s_n^2} \text{Var} \left( \sum_{j=m(n)\ell(n)+1}^{r(n)} X_j \right) \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

should be added. This assumption is implied by this simpler one, denoted (*HNab*) below :

$$\frac{1}{s_n^2} \sum_{i=t_n}^{u_n} \text{var}(X_i) \rightarrow 0 \text{ as } n \rightarrow \infty,$$

for  $0 \leq t_n \leq u_n \leq n, u_n - t_n \leq \ell(n), (u_n - t_n)/n \rightarrow 0$  as  $n \rightarrow \infty$ .

For a stationary case, this assumption is immediate. The foundation of our remark is given in Point (1) in Subsubsection 4.2.3 of Subsection 4.2 in Section 4.

Our objective in this paper is to express *CLT*'s in the most general setting, still using the Newman approach and to derive the former results as particular cases. With respect to the former results described above, we simplify the approach and get the best we can do by formulating a Lyapounov-Feller-Levy type of *CLT*. The general conditions are next expressed on moment conditions stated also in a general setting. Existing versions are all included in our statements. And we conclude that more general *CLT*'s cannot be obtained without getting out the Newman approach.

Let us begin by introduce the following assumptions

There exists a sequence  $\ell(n)$  of positive integers such that  $n = m(n)\ell(n) + r(n)$ , with  $0 \leq r(n) < \ell(n)$ ,  $0 \leq m(n) \rightarrow +\infty$  and

$$(L) \quad (\ell(n)/n, r(n)/n) \rightarrow (0, 0) \text{ as } n \rightarrow +\infty.$$

We want to stress that the integers  $m = m(n)$ ,  $\ell = \ell(n)$  and  $r = r(n)$  depend of  $n$  throughout the text even though we may and do drop the label  $n$  in many situations for simplicity's sake.

On top of this general assumption, we may require the following ones.

$$(H0) \quad \frac{\ell(n)}{s_n^2} \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

$$(Ha) \quad \frac{\ell(n)}{s_n^2} \sum_{j=1}^{m(n)} \text{Var} \left( \frac{S_{j\ell(n)} - S_{(j-1)\ell(n)}}{\sqrt{\ell(n)}} \right) \rightarrow 1 \text{ as } n \rightarrow +\infty.$$

$$(Hab) \quad \frac{1}{s_n^2} \mathbb{V}ar \left( \sum_{i=m(n)\ell(n)+1}^{r(n)} X_j \right) \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

$$(Hb) \quad \sup_{1 \leq j \leq m(n)+1} \frac{\ell(n)}{s_n^2} \mathbb{V}ar \left( \frac{S_{j\ell(n)} - S_{(j-1)\ell(n)}}{\sqrt{\ell(n)}} \right) = C_1(n) \rightarrow 0 \text{ as } j \rightarrow +\infty.$$

We have for some  $\delta > 0$ ,  $\mathbb{E} |X_j|^{2+\delta} < +\infty, j \geq 1$  and the Lyapounov Condition holds

$$(Hc) \quad \frac{\ell^{3/2}(n)}{s_n^{2+\delta}} \sum_{j=1}^m \mathbb{E} \left| \frac{S_{j\ell(n)} - S_{(j-1)\ell(n)}}{\sqrt{\ell(n)}} \right|^{2+\delta} = C_2(n) \rightarrow 0 \text{ as } j \rightarrow +\infty.$$

In the sequel, it may be handy to use the notation

$$(3.9) \quad Y_{j,\ell} = \frac{S_{j\ell(n)} - S_{(j-1)\ell(n)}}{\sqrt{\ell(n)}}, 1 \leq j \leq m = m(n).$$

In the next section, we will state two *CLT*'s based on these hypotheses and next. A third one will be a refinement of Theorem 50. Next, the results are particularized into more specific versions.

#### 4. Results and Commentaries

In this section, we present general *CLT*'s for associated *rv*'s and next give different forms in specific types of independent and dependent data and finally make a comparison with available results.

**4.1. General *CLT*'s.** We have following results.

**THEOREM 51.** *Let  $X_1, X_2, \dots, X_n$  be an associated sequence of mean-zero random variables defined on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . If the sequence is stationary, then*

$$\frac{S_n}{\sqrt{n}} = \frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}} \rightsquigarrow \mathcal{N}(0, \sigma^2) \text{ as } n \rightarrow +\infty,$$

*In the general setting, if (L), (H0), (Ha), (Hb) and (Hc) hold, then*

$$\frac{S_n}{s_n} = \frac{X_1 + X_2 + \dots + X_n}{s_n} \rightsquigarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow +\infty.$$

**Remark.** The first statement of the Theorem is the classical *CLT* of **Newman and Wright (1981)**. The second statement corresponds to our effort to get the maximal result from the Newman's approach for non-stationary data.

Next, we state a Lynderberg-Feller-Levy type theorem given some assumptions.

**THEOREM 52.** *Let  $X_1, X_2, \dots, X_n$  be an associated sequence of mean-zero random variables defined on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Denote for each  $j \in \{1, \dots, m\}$ ,  $\tau_j^2 = \text{Var}(S_{j\ell} - S_{(j-1)\ell}) = \mathbb{E}(S_{j\ell} - S_{(j-1)\ell})^2$  and*

$$\nu_{m(n)}^2 = \tau_1 + \dots + \tau_{m(n)}, n \geq 1.$$

*Assume that the assumptions (L) and (Ha) hold and either (Hab) or (Hb) is true. Then we have the following equivalence result :*

$$\max_{1 \leq k \leq m(n)} \mathbb{E}(S_{j\ell} - S_{(j-1)\ell})^2 / s_n^2 \rightarrow 0 \text{ as } n \rightarrow +\infty$$

and

$$S_n / s_n \rightsquigarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow +\infty,$$

if and only if for any  $\varepsilon > 0$ ,

$$(4.1) \quad \frac{1}{s_n^2} \mathbb{E} \left( (S_{j\ell} - S_{(j-1)\ell})^2 1_{(|S_{j\ell} - S_{(j-1)\ell}| \geq \varepsilon \nu_{m(n)})} \right) \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

Moreover the sequence  $(\nu_{m(n)})_{n \geq 1}$  may be replaced by the sequence of  $(s_n)_{n \geq 1}$  in Condition (4.1) to give

$$(4.2) \quad \frac{1}{s_n^2} \mathbb{E} \left( (S_{j\ell} - S_{(j-1)\ell})^2 1_{|S_{j\ell} - S_{(j-1)\ell}| \geq \varepsilon s_n} \right) \rightarrow 0 \text{ as } n \rightarrow 0.$$

**THEOREM 53.** *Let  $X_1, X_2, \dots, X_n$  be an associated sequence of mean-zero random variables defined on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . If*

$$\left| \Psi_{\frac{s_{m\ell}}{s_n}}(t) - \prod_{j=1}^m \Psi_{Y_{j,n}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \right| \rightarrow 0 \text{ as } n \rightarrow +\infty,$$

then we have the following equivalence result :

$$\max_{1 \leq k \leq m(n)} \mathbb{E} (S_{j\ell} - S_{(j-1)\ell})^2 / s_n^2 \rightarrow 0 \text{ as } n \rightarrow +\infty$$

and

$$S_n/s_n \rightsquigarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow +\infty,$$

if and only if for any  $\varepsilon > 0$ , Formula 4.1 holds.

REMARK 15. Let us make the following remarks.

**(1)** The method of proving the above theorem will consist of decomposing the sums of variables into sums of blocks of variables and treating these as if they were independent. Naturally, we will need some control on the approximation between the sums of the dependent blocks and their independent counterparts. This control will be achieved using characteristic functions and is based on the inequality in Lemma 31 of Newman. Our approach is to go the farthest possible only using the moments conditions.

**(2)** Theorem 52 is not yet a Lyndeberg-Feller-Levy Theorem (LFLT). Using Lemma 31, its only say we have a LFLT provided assumptions that make the CLT problem into a CLT one concerning independent variables. A full LFLT can not be achieved as long as the proofs are based on the approximation of Lemma 31.

Before we proceed to the proofs in Subsection 4.3, we are now going to derive some consequences and particular cases of the theorem.

#### 4.2. Commentaries and Consequences.

4.2.1. The most general approach leading to a Feller-Levy CLT type. We begin a general comments of the approach.

Almost all the available CLT results used the Newman's method based on Lemma 31. The approach we used is intended to get the sharpest results we can get in that frame. In earlier versions of our results, we were not aware of the results of Oliveira (2012). However, with the knowledge of these results, our works still present a number of significant advantages we want to highlight here. Actually, Oliveira (2012) attains the best we can do in the Newman approach : the best and unique way to find out assumptions under which

$$S_n/s_n \rightsquigarrow N(0, 1),$$

holds, is to reduce it to

$$(4.3) \quad \prod_{j=1}^m \Psi_{Y_{j,\ell}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \rightarrow \exp(-t^2/2) \text{ as } n \rightarrow +\infty.$$

after are defined the random variables  $Y_{j,\ell}$  of characteristic functions  $\Psi_{Y_{j,\ell}}$  given in Formula (3.9) based on the decomposition  $(L)$  (We recall that our notation of  $Y_{j,\ell}$  are not the same as that of Oliveira (2012)).

This is the justification of Assumption (3.4) above, which corresponds to the equivalent one we used, which is

$$(4.4) \quad \left| \Psi_{\frac{s_{m\ell}}{s_n}}(t) - \prod_{j=1}^m \Psi_{Y_{j,\ell}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \right| \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

From there, the authors did not, as far as we know, capitalize this fact in order to have a Feller-Levy final *CLT* version, as we did in Theorem 53. In our view, this version is the starting point for new *CLT*'s out of the Newman approach.

4.2.2. *General condition.* Based on the best the Newman approach can give, it remains to have the most general conditions to ensure (4.3) and (4.4). If we wish to directly express (4.4) into the  $X_i$ 's, Oliveira (2012) proved in page 109 that their assumption (3.8) implies Formula (4.2) in Theorem 52 above. In general, authors usually provide *CTL*'s based on conditions ensuring (4.3) and (4.4).

In that specific case, we proceeded into two directions :

(1) Expressing general conditions based on moments. We will see in the next subsection, how the available *CLT*'s may be derived from Theorem 52.

(2) Keeping the notation of decomposition  $(L)$  in the assumptions. This will allow, in particular cases, to base methods on specific values of  $m(n)$  and  $\ell(n)$

4.2.3. *Comparisons.* Let us highlight some comparison results.

**(1) With Theorem 50. A possible gap in Theorems 50 of Oliveira (2012).**

By setting

$$Y_{j,\ell(n)}^* = Y_{j,\ell(n)} \text{ for } j = 1, \dots, m(n) \text{ and } Y_{m(n)+1,\ell}^* = \sum_{i=m(n)\ell(n)+1}^{r(n)} X_i / \sqrt{\ell(n)},$$

we have

$$S_n = \ell(n) \sum_{j=1}^{m(n)+1} Y_{j,\ell(n)}^*.$$

It comes that

$$\begin{aligned} s_n^2 = \text{Var}(S_n) &= \ell(n) \sum_{j=1}^{m(n)} \text{var}(Y_{j,\ell(n)}) + \ell^2 \text{Var}(Y_{m(n)+1,\ell(n)}^*) \\ &+ 2\ell(n) \sum_{1 \leq h < k \leq m(n)+1} \text{cov}(Y_{h,\ell(n)}^*, Y_{k,\ell(n)}^*). \end{aligned}$$

By definition of the Cox coefficient (as named such that by [Bulinski and Shashkin \(2007\)](#)), and since the indices of the  $X_i$ 's in  $\text{cov}(Y_{h,\ell}^*, Y_{k,\ell}^*)$  are distanced by  $1, \dots, \ell$  points in absolute values, and the  $X_i$ 's are therein normalized by  $\sqrt{\ell}$ , we have

$$\begin{aligned} \sum_{1 \leq h < k \leq m(n)+1} \text{cov}(Y_{h,\ell(n)}^*, Y_{k,\ell(n)}^*) &\leq \frac{1}{\ell} \sum_{h=1}^{\ell(n)} \sup_{i \geq 1} \sum_{k: |k-i| \geq h} \text{cov}(X_i, X_k) \\ &\leq \frac{1}{\ell} \sum_{i=1}^{\ell(n)} u(i), \end{aligned}$$

and then

$$s_n^2 = \text{Var}(S_n) \leq \ell \sum_{j=1}^{m(n)} \text{var}(Y_{j,\ell(n)}) + \ell(n)^2 \text{Var}(Y_{m(n)+1,\ell(n)}^*) + 2 \sum_{i=1}^{\ell(n)} u(i),$$

which gives

$$\begin{aligned} \left| 1 - \frac{\ell(n)}{s_n^2} \sum_{j=1}^m \text{var}(Y_{j,\ell(n)}) - \frac{\ell(n)^2}{s_n^2} \text{Var}(Y_{m(n)+1,\ell(n)}^*) \right| &\leq \frac{2\ell(n)}{s_n^2} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} u(i) \right\} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty, \end{aligned}$$

by Césaro's Lemma if

$$(4.5) \quad \limsup_{n \rightarrow +\infty} \frac{\ell(n)}{s_n^2} < +\infty.$$

Here, it seems to us that the authors of [Oliveira \(2012\)](#) might have not taken into account the term  $\ell Y_{m+1,\ell}^*$  in the line -7 of their page 108. At line -6 of the same page, their formula  $S_n = Y_{1,\ell_n} + \dots + Y_{m_n,\ell_n}$  also misses to include the remaining  $X_i$ 's corresponding to  $i \in [m_n \ell_n + 1, m_n \ell_n + r_n]$ , where  $m_n$  is the integer part of  $n/\ell_n$  and  $r_n = n - m_n \ell_n$ . And although it is possible to get rid of the term  $\ell Y_{m+1,\ell}^*$  in line -1 of their page 108 as we explained in the lines following the remark (R2) in the proof of Theorem 51 below, we still think it would be handled in the proof of Formula (4.4), as we did at the stage of Formula (4.8) of the same proof below.

Based on this remark, the hypotheses (3.6) and (3.7) are true and our (Hab) holds. Then Formula (4.5) holds and Formula (4.4) is true. The Feller-Levy theorem handles the remaining part. Further

$$\frac{\ell(n)^2}{s_n^2} \text{Var} (Y_{m(n)+1,\ell(n)}^*) \leq \frac{1}{s_n^2} \sum_{i=m(n)\ell(n)+1}^{r(n)} \text{var}(X_i) + \frac{2\ell(n)}{s_n^2} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} u(i) \right\}.$$

Then, if Assumption (3.6) and (4.5) hold, then (Hab) is implied by a general condition of the form

$$(4.6) \quad \frac{1}{s_n^2} \sum_{i=t_n}^{u_n} \text{var}(X_i) \rightarrow 0 \text{ as } n \rightarrow \infty,$$

for  $0 \leq t_n \leq u_n \leq n, u_n - t_n \leq \ell(n), (u_n - t_n)/n \rightarrow 0$  as  $n \rightarrow \infty$ .

## (2) With Cox and Grimmet Theorem 48.

It is immediate that the first part of Assumption (3.1) in that theorem, that is

$$\text{Var}(X_j) \geq c_1 > 0,$$

implies, by association, that

$$s_n^2 \geq \sum_{i=1}^n \text{var}(X_i) \geq nc_1$$

and (4.5) holds since

$$\limsup_{n \rightarrow +\infty} \frac{\ell(n)}{s_n^2} = \limsup_{n \rightarrow +\infty} \frac{n}{s_n^2} \times \frac{\ell(n)}{n} \leq c_1 \limsup_{n \rightarrow +\infty} \frac{\ell(n)}{n} \leq c_1.$$

Next, the second part, that is

$$\mathbb{E} |X_j|^3 \leq c_2 < +\infty, j \geq 1,$$

implies, by the formula  $|x|^p \leq 1 + |x|^q$  for  $1 \leq p \leq q$  (see [Loève \(1977\)](#), page 157), that for  $c_3 = 1 + c_2$ ,

$$E |X_j|^3 \leq c_3, j \geq 1,$$

and then Formula (4.6) above holds since

$$\frac{1}{s_n^2} \sum_{i=t_n}^{u_n} \text{var}(X_i) \leq c_3 \frac{(u_n - t_n)}{s_n^2} \leq \frac{c_3}{c_1} \frac{(u_n - t_n)}{n} \rightarrow 0.$$

Next, by re-making the considerations given in Subsubsection 4.2.2, Formula (4.2) of Theorem 52 holds if

$$\frac{1}{s_n^2} \sum_{j=1}^{m(n)} \int_{\{|X_j| \geq \varepsilon s_n\}} X_j^2 d\mathbb{P} \rightarrow 0 \text{ as } n \rightarrow +\infty,$$

for any  $\varepsilon > 0$ . But we have under Condition (3.1) of Theorem 48,

$$\begin{aligned} \frac{1}{s_n^2} \sum_{j=1}^{m(n)} \int_{\{|X_j| \geq \varepsilon s_n\}} X_j^2 d\mathbb{P} &= \frac{1}{s_n^2} \sum_{j=1}^{m(n)} \int_{\{|X_j| \geq \varepsilon s_n\}} \frac{|X_j|^3}{|X_j|} d\mathbb{P} \\ &\leq \frac{1}{\varepsilon s_n^3} \sum_{j=1}^{m(n)} \int_{\{|X_j| \geq \varepsilon s_n\}} |X_j|^3 \\ &\leq \frac{m(n)c_2}{\varepsilon s_n^3} \leq \frac{c_2}{c_1^{3/2}} \frac{m(n)}{n^{3/2}} \rightarrow 0. \end{aligned}$$

Hence, Condition (4.2) is true. Finally Condition (3.2) ensures (4.4) and the Cox Theorem 48 is obtained.

4.2.4. *Conclusion.* We conclude in two points.

**(A)** By combining our results with especially those of Oliveira (2012), we have proved that the Newman approach already gave the best results in a Lyndeberg-Feller-Levy type of *CLT*. It is still possible to find different more or less sharp expressions of Conditions (4.3) and (4.4), stated in Subsection 4.2. But no very different results cannot be expected there. Yet, the *CLT* problem is widely open since the current results use the Newman approach. Is it possible to get rid of this approach and to use another one more general to establish more general **CLT**'s? This seems to be the direction to be taken.

**(B)** In Fall *et al.* (2018), an associated sequence is studied as a particular case. Using a direct method based on the characteristic function method, it has been shown to satisfy the *CLT* property. Yet, this sequence did not satisfy the Cox-Grimmet condition  $\inf_{n \geq 1} \mathbb{E}X_i^2 \geq c_1 > 0$ . This kind of work may constitute a lead to more general *CLT*.

**4.3. Proof of Theorem 51.** As almost all the proofs of *CLT*'s for associated or weakly associated rv's, our proof is based on the three steps of the original method of Newman and Wright (1981). For compact notation sake, we simply set  $\ell(n) = \ell$  and  $m(n) = m$ . Let us define  $\Psi_{\frac{S_n}{s_n}}(t) = \mathbb{E}(e^{itS_n/s_n})$ ,  $t \in \mathbb{R}$ .

First, we have for  $t \in \mathbb{R}$ ,

$$\begin{aligned}
 \left| \Psi_{\frac{S_n}{s_n}}(t) - \Psi_{\frac{S_{m\ell}}{s_n}}(t) \right| &= \left| \mathbb{E}(e^{itS_n/s_n}) - \mathbb{E}(e^{itS_{m\ell}/s_n}) \right| \\
 &= \left| \mathbb{E} \left[ e^{itS_{m\ell}/\sqrt{m\ell}} (e^{it[(S_n/s_n)-(S_{m\ell}/s_n)]} - 1) \right] \right| \\
 (4.7) \qquad \qquad \qquad &\leq \mathbb{E} \left| e^{it \left( \frac{S_n}{s_n} - \frac{S_{m\ell}}{s_n} \right)} - 1 \right|.
 \end{aligned}$$

But for any  $x \in \mathbb{R}$ ,

$$|e^{ix} - 1| = |(\cos x - 1) + i \sin x| = \left| 2 \sin \frac{x}{2} \right| \leq |x|.$$

Thus the second member of (4.7) is, by the Cauchy-Schwarz's inequality, bounded by

$$|t| \mathbb{E} \left| \frac{S_n}{s_n} - \frac{S_{m\ell}}{s_n} \right| \leq |t| \text{Var} \left( \frac{S_n}{s_n} - \frac{S_{m\ell}}{s_n} \right)^{\frac{1}{2}}$$

and

$$\delta_{m,\ell} = \text{Var} \left( \frac{S_n}{s_n} - \frac{S_{m\ell}}{s_n} \right) = \frac{1}{s_n^2} \text{Var} (S_n - S_{m\ell}),$$

which tends to zero as  $n \rightarrow +\infty$  by (Hb) since

$$(4.8) \quad \delta_{m,\ell} = \frac{1}{s_n^2} \text{Var} \left( \sum_{i=1}^r X_{m\ell+i} \right)$$

$$(4.9) \quad \leq \frac{\ell}{s_n^2} \text{Var} \left( \frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} X_{m\ell+i} \right)$$

$$(4.10) \quad \leq C_1(n) \rightarrow 0.$$

This proves that

$$(4.11) \quad \left| \Psi_{\frac{S_n}{s_n}}(t) - \Psi_{\frac{S_{m\ell}}{s_n}}(t) \right| \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

**(R1)** Remark also for the purpose of Theorem 52 that the same conclusion holds when (Hab) is true and we do not need (Hb) in addition.

Next, remind that  $Y_{j,\ell} = (S_{j\ell} - S_{\ell(j-1)})/\sqrt{\ell}$ , for  $1 \leq j \leq m$ . Observe that

$$\frac{S_{m\ell}}{s_n} = \frac{\sqrt{\ell}}{s_n} \sum_{j=1}^m Y_{j,\ell}.$$

According to the Newman's inequality (see Lemma 31), we have

$$\left| \Psi_{\frac{S_{m\ell}}{s_n}}(t) - \prod_{j=1}^m \Psi_{Y_{j,\ell}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \right| \leq \frac{\ell t^2}{2s_n^2} \sum_{1 \leq j \neq k \leq m} \text{Cov}(Y_{j,\ell}, Y_{k,\ell}).$$

But,

$$\begin{aligned}
 \frac{\ell t^2}{2s_n^2} \sum_{1 \leq j \neq k \leq m} \text{Cov}(Y_{j,\ell}, Y_{k,\ell}) &= \frac{\ell t^2}{2s_n^2} \text{Var} \left( \sum_{j=1}^m Y_{j,\ell} \right) - \frac{\ell t^2}{2s_n^2} \sum_{j=1}^m \text{Var}(Y_{j,\ell}) \\
 &= \frac{t^2}{2} \left[ \text{Var} \left( \frac{\sqrt{\ell}}{s_n} \sum_{j=1}^m Y_{j,\ell} \right) - \frac{\ell}{s_n^2} \sum_{j=1}^m \text{Var}(Y_{j,\ell}) \right] \\
 &= \frac{t^2}{2} \left[ \text{Var} \left( \frac{1}{s_n} S_{m\ell} \right) - \frac{\ell}{s_n^2} \sum_{j=1}^m \text{Var} \left( \frac{S_{j\ell} - S_{\ell(j-1)}}{\sqrt{\ell}} \right) \right] \\
 &\leq \frac{t^2}{2} \left[ 1 - \frac{\ell}{s_n^2} \sum_{j=1}^m \text{Var} \left( \frac{S_{j\ell} - S_{\ell(j-1)}}{\sqrt{\ell}} \right) \right] \\
 &\quad - \frac{t^2}{2s_n^2} \text{Var} \left( \sum_{j=m\ell+1}^n X_j \right),
 \end{aligned}$$

which tends to zero as  $n \rightarrow +\infty$  by (Ha) and (Hb), that is

$$(4.12) \quad \left| \Psi_{\frac{S_{m\ell}}{s_n}}(t) - \prod_{j=1}^m \Psi_{Y_{j,n}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \right| \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

The proof will be completed by establishing that

$$(4.13) \quad \prod_{j=1}^m \Psi_{Y_{j,n}} \left( \frac{\sqrt{\ell}}{s_n} t \right) \rightarrow \exp(-t^2/2) \text{ as } n \rightarrow +\infty.$$

**(R2)** Here, we make a second remark which is relevant to the Proof of Theorem 52 and next to generalizations of the results. The above computations led to

$$\begin{aligned}
 0 &\leq \frac{\ell t^2}{2s_n^2} \sum_{1 \leq j \neq k \leq m} \text{Cov}(Y_{j,\ell}, Y_{k,\ell}) = \frac{t^2}{2} \left[ 1 - \frac{\ell}{s_n^2} \sum_{j=1}^m \text{Var} \left( \frac{S_{j\ell} - S_{\ell(j-1)}}{\sqrt{\ell}} \right) \right] \\
 &\quad - \frac{t^2}{2s_n^2} \text{Var} \left( \sum_{j=m\ell+1}^n X_j \right) \\
 &\leq \frac{t^2}{2} \left[ 1 - \frac{\ell}{s_n^2} \sum_{j=1}^m \text{Var} \left( \frac{S_{j\ell} - S_{\ell(j-1)}}{\sqrt{\ell}} \right) \right].
 \end{aligned}$$

Then only  $(Ha)$  is needed to ensure (4.13).

Now, we resume to the normal course of our demonstration. From this step, the conclusion on the weak law of  $S_n/s_n$ , comes uniquely from Formula (4.13) which expresses the weak convergence of sums of the form

$$(4.14) \quad T_{m(n)}^* = \frac{1}{s_n} \sum_{j=1}^{m(n)} V_j,$$

where the  $V_j$ 's are independent random variables such that for each  $j \in \{1, \dots, m\}$ ,  $V_j^*$  has the same law as  $S_{j\ell} - S_{(j-1)\ell}$ . Remind that, for each  $j \in \{1, \dots, m\}$ ,  $\tau_j^2 = Var(S_{j\ell} - S_{(j-1)\ell}) = \mathbb{E}(S_{j\ell} - S_{(j-1)\ell})^2$  and

$$\nu_{m(n)}^2 = \tau_1 + \dots + \tau_{m(n)}, n \geq 1.$$

By Assumption  $(Ha)$ , we have  $\nu_{m(n)}/s_n \rightarrow 1$  as  $n \rightarrow +\infty$  and by Slutsky theorem (see for example Proposition 15 in Lo (2018), page 60), the weak convergence, if it holds, would be the same as that of

$$T_{m(n)} = \frac{1}{\nu_{m(n)}} \sum_{j=1}^{m(n)} V_j.$$

Condition  $(Hb)$  is the Lyapounov's one for this problem (see Loève (1977), page 287, Point B), where  $\nu_{m(n)}$  is replaced by  $s_n$ . This completes the proof.

**4.4. Proof of Theorem 52.** Based of the remarks labelled by **(R1)** and **(R2)** in the body of the proof of Theorem 51, we conclude that if  $(L)$ ,  $(Ha)$  and  $(Hab)$  hold, the conclusion on the weak law of  $S_n/s_n$ , comes uniquely from Formula (4.13). At this step, the condition on the  $(2 + \delta)^{th}$  moments, that  $\mathbb{E}|X_j|^{2+\delta} < +\infty, j \geq 1$ , is not required. And, Formula (4.13) expresses the weak convergence of the sums defined in (4.14).

From there, the problem becomes the classical Lyndeberg-Levy-Feller Theorem. And we have the following conclusion :

**(a)**  $\max_{1 \leq k \leq m(n)} \{\tau_j/\nu_{m(n)}\} \rightarrow 0$  as  $n \rightarrow +\infty$  and

$$\frac{1}{\nu_{m(n)}} \sum_{j=1}^{m(n)} V_j \rightsquigarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow +\infty,$$

if and only if

(b) for any  $\varepsilon > 0$ ,

$$g_n(\varepsilon) = \frac{1}{v_{m(n)}^2} \int_{(|x| \geq \varepsilon v_{m(n)})} x^2 dF_{V_j} \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

These two conditions are exactly those given in the statement of the theorem, where the replacement of  $(|x| \geq \varepsilon v_{m(n)})$  by  $(|x| \geq \varepsilon s_n)$  in the expression of  $g$  is possible because of  $\nu_{m(n)}/s_n \rightarrow 1$  as  $n \rightarrow +\infty$ .

This finishes the proof on this theorem.

**4.5. Proof of Theorem 53.** The proof of Theorem 53 is based on that of Theorem 51 from Formula (4.12).

**Acknowledgment** The second author acknowledges support from the World Bank Excellence Center (CEA-MITIC) that is continuously funding his research activities starting 2014. The first author thanks the **Programme de formation des formateurs** of USSTB who financed his stays in the LERSTAD of UGB while preparing his Ph.D dissertation. Both authors acknowledge support from the **Réseau EDP - Modélisation et Contrôle**, of Western African Universities, that financed travel and accommodation of the second author while visiting USTTB in preparation of a series of works with his PhD students there.

## Bibliography

- Bulinski A. and Shashkin A.(2007). Limit theorems for associated random fields and related systems. World Scientific Publishing, Singapore.
- Burton, R.M., Dabrowski, A.R. and Dehling, H. (1986). An invariance principle for weakly associated random variables, *Stoch. Proc. Appl.*, 23, 301-306.
- Cox, J.T. and Grimmett, G. (1984) Central limit theorems for associated random variables and the percolation model, *Ann. Probab.*, 12, 514-528.
- Dabrowski, A.R. and Dehling, H. (1988). A Berry-Esseen theorem and a functional law of the iterated logarithm for weakly associated random variables, *Stochastic Process. Appl.*, 30, 247-289.
- Daley, D. J.(1968). Stochastically monotone Markov Chains, *Z. Wahrsch. theor. verw Gebiete*, 10, 305-317.
- Esary, J., Proschan, F. and Walkup, D.(1967). Association of random variables with application. *Ann. Math Statist.*, 38,
- Fortuin, C., Kastelyn, P. and Ginibre, J.(1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22, 89-103.
- Gut, A. (2005). *Probability : A Graduate Course*. Springer Science+Business Media, Inc. ISBN 0-387-22833-0.
- Jiazhu, P.(2002). Tail dependence of random variables from ARCH and heavy-tailed bilinear models. *Sciences in China*, 45 (6), Ser. A, 749-760.
- Le Cam, L.(1986). The Central Limit Theorem Around 1935. *Statistical Science.*, Vol. 1 (1), 78-91.
- Lo G.S., Fall A.M. and Harouna S.(2018). A Central limit Theorem of dependent sums of standard exponential functionals motivated by extreme value theory. *To come*.
- Lehmann, E. L.(1966). Some Concepts of dependence. *Ann. Math. Statist.*, 37, 1137-1153.
- Lo, G.S.(2018). Weak Convergence (IA). Sequences of random vectors. SPAS Books Series.(2016). Doi : 10.16929/sbs/2016.0001. Arxiv : 1610.05415
- Loève, M.(1977). *Probability Theory I*. Springer-Verlag. New-York.

- Louhichi, S.(2000). Weak convergence for empirical processes of associated sequences. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques* 36 (5), pp. 547–567.
- Newman C.M. (1980) Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.* 74, 119-128.
- Newman, C.M and Wright, A.L.(1981). An invariance principle for certain dependent sequences. *Ann. probab.*, **9** (4), 671-675.
- Newman, C.M and Wright, A.L.(1982). Associated random variables and martingale inequalities. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 59, 361-371.
- Oliveira, P.E.(2012). *Asymptotics for Associated Random Variables*. DOI 10.1007/978-3-642-25532-8, © Springer-Verlag Berlin Heidelberg.
- Pitt, L.(1982). Positively Correlated normal variables are associated. *Ann. Probab.*, 10, 496-499.
- Rao, P. B. L. S.(2012). *Associated sequences, Demimartingales and Non-parametric Inference. Probability and its applications* . Springer Basel Doredrecht, Heidelberg, London, New York.
- Sangaré, H. and Lo, G. S.(2016) A Review on asymptotic normality of sums of associated random variables. *Afrika Statistika*, 11 (1), pp.855-867. Doi : 10.16929/as/2016.855.79. Arxiv 1405.4316.
- Yu, H.(1993). A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences. *Probab. Theory Related Fields*. 95, 357-370