On a Non-Classical Invariance Principle

Youri DAVYDOV[1] and Vladimir ROTAR [2]

[1] Laboratoire Paul Painlevé - UMR 8524
Université de Lille I - Bat. M2
59655 Villeneuve d'Ascq, France
Email: youri.davydov@univ-lille1.fr

[2] Department of Mathematics and Statistics
of the San Diego State University, USA and
the Central Economics and Mathematics Institute
of the Russian Academy of Sciences, RF
Email: vrotar@math.ucsd.edu

Abstract. We consider the invariance principle without the classical condition of asymptotic negligibility of individual terms. More precisely, let r.v.'s $\{\xi_{nj}\}$ and $\{\eta_{nj}\}$ be such that

$$E\{\xi_{nj}\} = E\{\eta_{nj}\} = 0,\ E\{\xi_{nj}^2\} = E\{\eta_{nj}^2\} = \sigma_{nj}^2,\ \sum_j \sigma_{nj}^2 = 1,$$

and the r.v.'s $\{\eta_{nj}\}$ are normal. We set

$$S_{kn} = \sum_{j=1}^{k} \xi_{nj},\ \ Y_{kn} = \sum_{j=1}^{k} \eta_{nj},\ \ t_{kn} = \sum_{j=1}^{k} \sigma_{nj}^2.$$

Let $X_n(t)$ and $Y_n(t)$ be continuous piecewise linear (or polygonal) random functions with vertices at (t_{kn}, S_{kn}) and (t_{kn}, Z_{kn}), respectively, and let P_n and Q_n be the respective distributions of the processes $X_n(t)$ and $Y_n(t)$ in $\mathbb{C}[0,1]$.

The goal of the present paper is to establish necessary and sufficient conditions for convergence of $P_n - Q_n$ to zero measure not involving the condition of the asymptotic negligibility of the r.v.'s $\{\xi_{nj}\}$ and $\{\eta_{nj}\}$.

AMS 1991 Subject Classification:
Primary 60F17, Secondary 60G15.

Keywords: Invariance principle, non-classical invariance principle, non-classical limit theorem, asymptotic confluence of distributions.

1 Introduction and results

1.1 Background and Motivation

The term "non-classical" concerns various limit theorems not involving the condition of asymptotic negligibility of the individual random variables (r.v.'s). To our knowledge, the convergence of the distributions of sums of r.v.'s to the normal distribution in the general situation, that is, without the condition mentioned, was first considered by P. Lévy [7] and M. Loéve [9, Chapter VIII, Section 28]. A developed theory with necessary and sufficient conditions was built by V.M. Zolotarev and his followers, V.M. Kruglov and Yu.Yu. Machis; see, e.g., [18], [6], [10], the monograph [19], the review part in [13], and references therein.

A somewhat different approach - see also comments below - that uses different types of conditions, was suggested in [12] and [13]. In this paper, we proceed mainly from the framework of [12] and [13].

In the case of normal convergence and finite variances, the simplest result from [12] and [13] may be stated as follows.

Let $\{\xi_{jn}\}$ be an array of independent r.v.'s such that $E\{\xi_{jn}\} = 0$, $E\{\xi_{jn}^2\} = \sigma_{jn}^2 < \infty$, and for each n,

$$\sum_j \sigma_{jn}^2 = 1. \tag{1.1.1}$$

Without loss of generality, we assume all $\sigma_{jn} \neq 0$.

Let $F_{jn}(x)$ be the distribution function (d.f.) of ξ_{jn}, and $\Phi_{jn}(x)$ be the normal d.f. with the same zero expectation and the same variance; that is, $\Phi_{jn}(x) = \Phi(x/\sigma_{jn})$, where $\Phi(x)$ is the standard normal d.f. Set $S_n = \sum_j \xi_{jn}$.

Proposition 1 *([12]) For*

$$P(S_n \leq x) \to \Phi(x), \text{ for all } x, \text{ as } n \to \infty, \tag{1.1.2}$$

it is necessary and sufficient that

$$\sum_j \int_{|x|>\varepsilon} |x| \cdot |F_{jn}(x) - \Phi_{jn}(x)| dx \to 0, \textit{ as } n \to \infty, \textit{ for any } \varepsilon > 0. \tag{1.1.3}$$

(This particular result is presented also in [17] and [14].) It is easy to show (see, for example, [14, p.310]) that in the classical case where $\max_j \sigma_{jn} \to 0$, the Lindeberg condition implies (1.1.3), so Lindeberg's theorem follows from Proposition 1. On the other hand, condition (1.1.3) takes into account possible proximity of the distributions of the r.v.'s to normal ones. In particular, if $F_{jn} \equiv \Phi_{jn}$ and hence $P(S_n \leq x) \equiv \Phi(x)$, then (1.1.3) becomes trivial.

It is worthwhile to note also that Proposition 1 is equivalent to Zolotarev's non-classical theorem from [18] proved much earlier. In the framework of [18], the summands were directly divided into two groups: those with "small" variances, and the rest. For the r.v.'s from the former group, Lindeberg's condition was imposed, while the summands from the latter group were required to be close to the corresponding normal r.v.'s in Lévy's metric. Such a division into two groups reflects the essence of the matter: "small" summands should be in the framework of the classical CLT, while "large" summands should be themselves close to normals. On the other hand, condition (1.1.3) allows to treat the summands in a unified way. Another difference between the theorem from [18] and Proposition 1 is that the latter uses an integral metric.

In the sufficiency case, the result of Proposition 1 was generalized to the case of semi-martingales in Liptser and Shiryaev's paper [8]; see also Jacod and Shiryaev's book [4, VII, 5b; VIII, 4c].

To generalize the result above to the case of convergence to distributions different from normal, one may proceed as follows. Consider another array of independent r.v.'s $\{\eta_{jn}\}$. We assume that for each n, the numbers of terms for ξ's and η's in the arrays $\{\xi_{jn}\}$ and $\{\eta_{jn}\}$ are the same and, just for simplicity, are finite. Let $E\{\eta_{jn}\} = 0$, $E\{\eta_{jn}^2\} = \sigma_{jn}^2$, and let G_{jn} denote the distribution of η_{jn}. The problem is to establish conditions under which

$$\prod_j F_{jn} - \prod_j G_{jn} \Rightarrow 0 \text{ as } n \to \infty, \tag{1.1.4}$$

where product of distributions is understood in the sense of convolution, and convergence $\Rightarrow$ itself is weak convergence (with respect of all continuous bounded functions). At least formally, this is a more general setup, since (1.1.4) does not presuppose the existence of limits for $\prod_j F_{jn}$ and $\prod_j G_{jn}$ separately. On the other hand, in the particular case when $G_{jn} \equiv \Phi_{jn}$, (1.1.4) clearly coincides with (1.1.2) in view of (1.1.1).

In the general situation (1.1.4), instead of (1.1.3), we consider the condition

$$\sum_j \int\limits_{||x|>\varepsilon|} |x| \cdot |F_{jn}(x) - G_{jn}(x)|dx \to 0, \text{ as } n \to \infty, \text{ for any } \varepsilon > 0. \tag{1.1.5}$$

In [5], it was shown that when G_{jn} are Poisson, (1.1.5) remains to be a necessary and sufficient condition for the fulfillment of (1.1.4), however attempts to obtain a similar result in the general case failed. The situation became clear when in [15] and [16] it was proved that in general, relation (1.1.5) is necessary for a more stronger type of convergence. Namely, (1.1.5) proves to be true if and only if

$$\prod_{j\in B_n} F_{jn} - \prod_{j\in B_n} G_{jn} \Rightarrow 0 \text{ as } n \to \infty, \tag{1.1.6}$$

for any sequence $\{B_n\}$ of subsets of the indices j. See [15] and [16] for detail; note also that in [16] the case of infinite variances is considered as well.

The fact that in the normal case, (1.1.4) and (1.1.6) occur to be equivalent is connected with the fact that normal distributions are only possible components of the decomposition of the normal law. The same concerns the Poisson case, however in general, relations (1.1.4) and (1.1.6) are certainly not equivalent.

Next, note that (1.1.6) deals with all possible partial sums, so if we manage to establish the validity of this relation, it is natural to continue and consider a more sophisticated problem, namely, the asymptotic proximity of the distributions of the partial-sum-processes based on the r.v.'s $\{\xi_{jn}\}$ and $\{\eta_{jn}\}$.

The main goal of this note is to point out the fact that condition (1.1.3) is *necessary and sufficient* for the validity of invariance principle in the case of Gaussian limiting processes in the general, that is, non-classical setup. To our knowledge, this fact has not been aired yet, though as we will see, in view of already known results, the proof turns out to be not very difficult.

Note also that, as a matter of fact, we consider a slightly more general problem of proximity of the distributions of the polygonal process generated by the above r.v.'s ξ_{jn} and the polygonal process generated by the corresponding normal r.v.'s. In the classical case, when $\max_j \sigma_{jn} \to 0$, such a result clearly corresponds to the classical invariance principle of Donsker-Prokhorov ([2], [11]), however without the condition mentioned we deal with a somewhat more complicated situation.

We hope to consider a more general case of non-normal limiting distributions in the next publication.

1.2 Results

As was mentioned, we assume for simplicity that for each n, the numbers of terms in each array, $\{\xi_{jn}\}$ or $\{\eta_{jn}\}$, are finite. Suppose all η's are normal, so $G_{jn}(x) = \Phi_{jn}(x) = \Phi(x/\sigma_{jn})$. We again assume (1.1.1) to hold, and set

$$\begin{aligned} S_n &= \sum_j \xi_{jn}, \quad Y_n = \sum_j \eta_{jn}, \\ S_{kn} &= \sum_{j=1}^{k} \xi_{jn}, \quad Y_{kn} = \sum_{j=1}^{k} \eta_{jn}, \\ t_{kn} &= \sum_{j=1}^{k} \sigma_{jn}^2. \end{aligned} \tag{1.2.1}$$

Let $X_n(t)$ and $Y_n(t)$ be continuous piecewise linear (or polygonal) random functions with vertices at (t_{kn}, S_{kn}) and (t_{kn}, Y_{kn}), respectively. Let $\mathcal{P}_n$ and $\mathcal{Q}_n$ be the respective distributions of the processes $X_n(t)$ and $Y_n(t)$ in $\mathbb{C} = \mathbb{C}[0,1]$.

Theorem 2 *Condition (1.1.3) is necessary and sufficient for*

$$\mathcal{P}_n - \mathcal{Q}_n \Rightarrow 0 \tag{1.2.2}$$

(more precisely, to zero measure) weakly.

Below, we show that the sequences $\{\mathcal{P}_n\}$ and $\{\mathcal{Q}_n\}$ are relatively compact, and hence in our case the above convergence is equivalent to that in the Lévy-Prokhorov's metric π, that is, $\pi(\mathcal{P}_n, \mathcal{Q}_n) \to 0$. In general, when compactness does not take place, and so to speak, “parts of the distributions move to infinity”, asymptotic proximity of distributions even in the one-dimensional case may be defined in different ways, so the very notion of proximity requires further analysis. We consider this question separately in [1].

We supplement Theorem 2 by the following simple proposition. Let for each n, the function $\sigma_n^2(t) = E\{X_n^2(t)\}$. Clearly, $\sigma_n(t)$ is continuous on $[0,1]$,

$$\sigma_n^2(t_{kn}) = \sum_{j=1}^{k} \sigma_{jn}^2,$$

and in each segment $[t_{(k-1)n}, t_{kn}]$, the function $\sigma_n^2(t)$ is a quadratic function.

Proposition 3 *The process $Y_n(t)$ converges in distribution to a Gaussian process $Y(t)$ on $[0,1]$ such that $E\{Y(t)\} = 0$ and $E\{Y^2(t)\} = \sigma^2(t)$ if and only if for each $t \in [0,1]$,*

$$\sigma_n(t) \to \sigma(t).$$

If $\max_j \sigma_{jn} \to 0$, then $\sigma^2(t) = t$, and $Y(t)$ is the standard Wiener process. In general, the segment $[0,1]$ may be divided into two sets, A and B, with the following properties.

The set A is a union of a finite or countable number of segments, and on each such a segment the process $Y(t)$ is linear.

The set $B = [0,1] \diagdown A$, and if a segment $[a,b] \subset B$, then the process $Y(a+s) - Y(a)$ is the standard Wiener process for $s \in [0, b-a]$.

2 Proofs

The main issue is to prove the relative compactness of the measure sequences $\{\mathcal{P}_n\}$ and $\{\mathcal{Q}_n\}$ (with respect to weak convergence of distributions in $\mathbb{C}$). For brevity, we omit sometimes the adjective "relative".

2.1 Compactness in the normal case

For the proof below, we need to consider a modification of the process $Y_n(t)$. For each $n = 1, 2, \ldots$, consider a partition of $[0,1)$ into some intervals $[s_{(j-1)n}, s_{jn})$ where $j = 1, \ldots, m_n \leq \infty$, and $0 = s_{0n} < s_{1n} < \ldots$. The number of intervals may be infinite, points s_{jn} may differ from the points t_{jn} above.

Let $W_n(t)$ be a continuous piecewise linear process such that $W_n(0) = 0$, on each interval $[s_{(j-1)n}, s_{jn})$ the trajectory of the process is linear, and each increment $W_n(s_{jn}) - W_n(s_{(j-1)n})$ is either equal to zero, or to a normal r.v. ζ_{jn} with zero mean and a variance of $s_{jn} - s_{(j-1)n}$. We prove the relative compactness of the family of the distributions of $W_n(t)$.

In accordance with a well known criterion (see, e.g., [3]), it suffices to prove that

(A) $\sup_n P\{|W_n(0)| > A\} \to 0$ as $A \to \infty$;

(B) there exist constants $a, b, c > 0$ such that for any n and $t, s \in [0,1]$,

$$E\{|W_n(t) - W_n(s)|^a\} \leq c|t-s|^{1+b}.$$

In our case, (A) is obvious. We verify (B) with $a=4$, $b=1$.

Set $v_{jn}^2 = E\{(W_n(s_{jn}) - W_n(s_{(j-1)n}))^2\}$. By the definition of W_n, either $v_{jn}^2 = 0$, or $v_{jn}^2 = s_{jn} - s_{(j-1)n}$.

If both points $t, s \in [s_{(k-1)n}, s_{kn}]$ for some $k \geq 1$, and $v_{jn}^2 \not= 0$, then

$$E\{|W_n(t) - W_n(s)|^4\} = E\left\{\left(\frac{|t-s|}{v_{kn}^2}\zeta_{kn}\right)^4\right\} \leq \frac{|t-s|^4}{v_{kn}^8} 3v_{kn}^4 \leq 3|t-s|^2 \quad (2.1.1)$$

since in this case $|t-s| \leq v_{kn}^2$. On the other hand, if $v_{kn} = 0$, then $W_n(t) - W_n(s) = 0$, and (2.1.1) is clearly true.

If $t = s_{kn}$ and $s = s_{mn}$ for some k and $m > k$, then the r.v. $W_n(t) - W_n(s)$ is normal with a variance that does not exceed $s_{mn} - s_{kn}$. Then

$$E\{|W_n(t) - W_n(s)|^4\} \leq 3\left(s_{mn} - s_{kn}\right)^2 = 3|t-s|^2.$$

In general, if $t \in [s_{(k-1)n}, s_{kn}]$ and $s \in [s_{(m-1)n}, s_{mn}]$ for some k and $m > k$, then in view of the above bounds,

$$\begin{aligned} E\{|W_n(t) - W_n(s)|^4\} &\leq E\left\{\left(|W_n(t) - W_n(s_{kn})| + |W_n(s_{kn}) - W_n(s_{(m-1)n})|\right.\right. \\ &\qquad \left.\left. + |W_n(s_{(m-1)n}) - W_n(s)|\right)^4\right\} \\ &\leq 27\left\{E\left\{|W_n(t) - W_n(s_{kn})|^4\right\} + E\left\{|W_n(s_{kn}) - W_n(s_{(m-1)n})|^4\right\}\right. \\ &\qquad \left. + E\left\{|W_n(st_{(m-1)n}) - W_n(s)|^4\right\}\right\} \leq 243|t-s|^2. \blacksquare \end{aligned}$$

2.2 Compactness of $\{\mathcal{P}_n\}$

First, note that in [8, Lemma 2], relative compactness in the non-classical situation was established in the general case of local *martingales* with respect to weak convergence in $\mathbb{D}$. However, it is not exactly what we need since we consider convergence in $\mathbb{C}$.

Certainly, once we consider continuous processes, and if limiting processes are also continuous (which is true in our case), compactness in $\mathbb{D}$ implies convergence in $\mathbb{C}$. However, when considering piecewise linear processes like $X_n(t)$ we loose the martingale property even when the r.v.'s ξ_{jn} are independent. On the other hand, if we switch to piecewise constant processes, we have to consider convergence in $\mathbb{D}$, which is not enough for us.

We believe that this is a technical obstacle and it may be somehow fixed, but in any case, in our opinion, a self contained (and relatively short) proof for the situation of independent summands would have an intrinsic value. So, we provide this proof.

Thus, we establish relative compactness of $\{\mathcal{P}_n\}$ in $\mathbb{C}$ under condition (1.1.3).

Set $\mathbf{k}_{jn} = [t_{(j-1)n}, t_{jn}]$, where the points t_{jn} are defined as in (1.2.1). For $\delta > 0$, we define the process $X_n(t;\delta)$ as a result of replacement of the r.v.'s ξ_{jn} by the r.v.'s $\tilde{\xi}_{jn} = \xi_{jn}\mathbf{1}\{\sigma^2_{jn} > \delta\}$ in the definition of $X_n(t)$. (As usual, $\mathbf{1}\{A\}$ is the indicator of a condition A.)

First, we show that for a fixed $\delta > 0$, the family of the distributions of $X_n(t;\delta)$ is compact. Indeed, denote by $\widetilde{\mathbf{k}}^{\delta}_{mn} = [r^{\delta}_{(m-1)n}, r^{\delta}_{mn}]$ the segments $\mathbf{k}_{jn}$ where the process $X_n(t;\delta)$ is not constant. We assume that $\widetilde{\mathbf{k}}^{\delta}_{mn}$ is on the left of $\widetilde{\mathbf{k}}^{\delta}_{(m+1)n}$. Since $\delta > 0$, the number of the segments $\widetilde{\mathbf{k}}_{mn}$ is finite. Denote this number by $q(n,\delta)$. Clearly, $q(n,\delta) \leq q = [1/\delta]$ where $[a]$ stands for the integer part of a. It is convenient to think that always $m = 1,...,q$, setting $\widetilde{\mathbf{k}}^{\delta}_{mn} = [1,1]$ for $m > q(n,\delta)$.

Clearly, there exists a subsequence $\widetilde{\mathbf{n}} = \{\widetilde{n}_i\}$ and segments $\widetilde{\mathbf{k}}^{\delta}_{m} = [r^{\delta}_{(m-1)}, r^{\delta}_{m}]$, $m = 1,...,q$, such that

$$\widetilde{\mathbf{k}}^{\delta}_{m\widetilde{n}_i} \to \widetilde{\mathbf{k}}^{\delta}_{m} \text{ as } \; i \to \infty,$$

(that is, the corresponding endpoints of the segments converge).

On the other hand, for each $\widetilde{\mathbf{k}}^{\delta}_{mn}$, the distribution of the increment $X_n(r^{\delta}_{mn};\delta) - X_n(r^{\delta}_{(m-1)n};\delta)$ is equal to a distribution F_{jn} for some j. Then from the main condition (1.1.3) it follows that the distribution of $X_{\widetilde{n}_i}(r^{\delta}_{m\widetilde{n}_i};\delta) - X_n(r^{\delta}_{(m-1)\widetilde{n}_i};\delta)$ weakly converges to the normal distribution with zero mean and the variance equal to the length of $\widetilde{\mathbf{k}}^{\delta}_{m}$. (We skip a formal proof of this fact. Because (1.1.3) is true for any $\varepsilon > 0$, we have convergence in the corresponding integral metric on any segments $[\varepsilon,\infty)$ and $(-\infty,-\varepsilon]$. This implies weak convergence. Since the limiting distribution is continuous, we have as a matter of fact uniform convergence, but we do not need it.)

Since the distribution of the process $X^{\delta}_n(\cdot)$ is uniquely specified by the finite dimensional distribution of the increments on the segments $\widetilde{\mathbf{k}}^{\delta}_{m}$, we finally conclude that the distribution of $X_{\widetilde{n}_i}(\cdot;\delta)$ weakly converges to the distribution of a continuous piecewise linear Gaussian process $W(t;\delta)$ having points of growth only in the segments $\widetilde{\mathbf{k}}^{\delta}_{m}$ and such that the increments $W(r^{\delta}_{m};\delta) - W(r^{\delta}_{(m-1)};\delta)$ are normal with zero mean and variance $r^{\delta}_{m} - r^{\delta}_{(m-1)}$.

Now, we proceed to a direct proof of compactness. Consider a sequence of positive numbers $\delta_n \to 0$. As was shown, there exists a subsequence $\mathbf{n}^{(1)} = \{n^{(1)}_i\}$

such that

$$X_{n_i^{(1)}}(\,\cdot\,;\delta_1) \stackrel{d}{\Rightarrow} W(\,\cdot\,;\delta_1) \text{ as } i\to\infty,$$

where $\stackrel{d}{\Rightarrow}$ stands for weak convergence of the corresponding distributions, and $W^{\delta_1}(\cdot)$ is a Gaussian process of the type $W^{\delta}(\cdot)$ described above.

Similarly, we can choose a subsequence $\mathbf{n}^{(2)}$ of the sequence $\mathbf{n}^{(1)}$ such that

$$X_{n_i^{(2)}}(\,\cdot\,;\delta_2) \stackrel{d}{\Rightarrow} W(\,\cdot\,;\delta_2) \text{ as } i\to\infty,$$

where $W^{\delta_2}(\cdot)$ is a Gaussian process with the same properties as above. Continuing to reason in the same fashion, we come to a nested sequence of subsequences $\mathbf{n}^{(1)} \supseteq \mathbf{n}^{(2)} \supseteq \dots$ such that for all $k = 1,2,\dots,$

$$X_{n_i^{(k)}}(\,\cdot\,;\delta_k) \stackrel{d}{\Rightarrow} W(\,\cdot\,;\delta_k) \text{ as } i\to\infty.$$

Next, consider the sequence of the Gaussian processes $\{W(\,\cdot\,;\delta_1),\ W(\,\cdot\,;\delta_2),\dots\}$. By the result of Section 2.1, there exists a subsequence m_j such that

$$W(\,\cdot\,;\delta_{m_j}) \stackrel{d}{\Rightarrow} W(\cdot),$$

where $W(\cdot)$ is a Gaussian process.

Now, we censor the sequence $\mathbf{n}^{(1)} \supseteq \mathbf{n}^{(2)} \supseteq \dots$, choosing only $\mathbf{n}^{(m_1)} \supseteq \mathbf{n}^{(m_2)} \supseteq \dots$. By construction, we can choose a sequence $n_1, n_2, \dots$ such that

$$n_1 \in \mathbf{n}^{(m_1)},\ \ n_2 \in \mathbf{n}^{(m_2)}, \dots, n_i \in \mathbf{n}^{(m_i)},\ \ \dots$$

and

$$X_{n_i}(\,\cdot\,;\delta_{m_i}) \stackrel{d}{\Rightarrow} W(\cdot) \text{ as } i\to\infty.$$

At the last step of the proof, we set $Z_n(t;\delta) = X_n(t) - X_n(t;\delta)$, and consider the sequence of the processes $U_i(t) = Z_{n_i}(t;\delta_{m_i})$. Each process $U_i(t)$ is a continuous process that is linear on each segment $\mathbf{k}_{jn_i}$ and such that the variance of the increment of the process on each $\mathbf{k}_{jn_i}$ does not exceed δ_{m_i}. Since $\delta_{m_i} \to 0$ as $i\to\infty$, all increments are asymptotically negligible. Formally, the processes $\{U_i(t)\}$ are not exactly of the type appearing in the classical invariance principle since for a finite number of segments $\mathbf{k}$ (with appropriate indices), the increments equals zero rather than having a variance equal the length of $\mathbf{k}$. Nevertheless, the proof of compactness may run exactly as, e.g., in the classical proof from Prokhorov's paper [11, Section 3.1].

Thus, the sequence of the distributions of $U_i(\,\cdot\,)$ is compact, and so does the sequence of the distributions of $X_{n_i}(\,\cdot\,;\delta_{m_i})$. It remains to observe that the processes $U_i(\,\cdot\,)$ and $X_{n_i}(\,\cdot\,;\delta_{m_i})$ are independent.

2.3 Proof of Theorem 2

2.3.1 Necessity

Let

$$\mathcal{P}_n - \mathcal{Q}_n \Rightarrow 0 \tag{2.3.1}$$

weakly in $\mathbb{C}[0,1]$. As was shown in Section 2.1, the sequence $\{\mathcal{Q}_n\}$ is compact. Then $\{\mathcal{P}_n\}$ is compact either.

Now, since $\sum_j \sigma_{jn}^2 \equiv 1$, the marginal distribution function for Y_n, i.e., $P(Y_n \leq x) \equiv \Phi(x)$. Hence, in view of (2.3.1),

$$P(S_n \leq x) \to \Phi(x).$$

By virtue of Proposition 1, this implies the validity of (1.1.3).

2.3.2 Sufficiency

Assume that condition (1.1.3) holds. Then, as was proved above, both sequences, $\{\mathcal{P}_n\}$ and $\{\mathcal{Q}_n\}$, are compact. Hence, it suffices to establish the convergence of the differences of all finite-dimensional marginal distributions.

Let $t_1 < t_2 < ... < t_k$ be points in $[0,1]$. Set $\mathbf{X}_n(t_1,...,t_k) = (X_n(t_1),...,X_n(t_k))$ and $\mathbf{Y}_n(t_1,...,t_k) = (Y_n(t_1),...,Y_n(t_k))$ and denote by $P_n(t_1,...,t_k)$ and $Q_n(t_1,...,t_k)$ the distributions of the random vectors $\mathbf{X}_n(t_1,...,t_k)$ and $\mathbf{Y}_n(t_1,...,t_k)$, respectively. Both sequences, $\{P_n(t_1,...,t_k)$ and $Q_n(t_1,...,t_k)$, are compact.

We should prove that

$$P_n(t_1,...,t_k) - Q_n(t_1,...,t_k) \Rightarrow 0. \tag{2.3.2}$$

Let the half interval $\mathbf{r}(j,n) = [t_{(j-1)n}, t_{jn})$, and the relations $t_i \in \mathbf{r}(m_{in},n)$, $i = 1.,,,.k$, define the integers m_{in}. Then for $i = 1,...,k$,

$$\begin{aligned} X_n(t_i) &= S_{(m_{in}-1)n} + \frac{t_i - t_{m_i n}}{\sigma^2_{m_{in}}} \xi_{m_{in}n}, & (2.3.3) \\ Y_n(t_i) &= Z_{(m_{in}-1)n} + \frac{t_i - t_{m_i n}}{\sigma^2_{m_{in}}} \eta_{m_{in}n}. & (2.3.4) \end{aligned}$$

For each n, consider the random vectors

$$\Bigg(\sum_{j=1}^{m_{1n}-1}\xi_{jn},\ \xi_{m_1n},\ \sum_{j=m_{1n}+1}^{m_{2n}-1}\xi_{jn},\ \xi_{m_2n}\mathbf{1}(m_2>m_1),\\ \ldots,\sum_{j=m_{(k-1)n}+1}^{m_{kn}-1}\xi_{jn},\ \xi_{m_kn}\mathbf{1}(m_k>m_{k-1})\Bigg), \tag{2.3.5}$$

and

$$\Bigg(\sum_{j=1}^{m_{1n}-1}\eta_{jn},\ \eta_{m_1n},\ \sum_{j=m_{1n}+1}^{m_{2n}-1}\eta_{jn},\ \eta_{m_2n}\mathbf{1}(m_2>m_1),\ldots\\ \ldots,\sum_{j=m_{(k-1)n}+1}^{m_{kn}-1}\eta_{jn},\ \eta_{m_kn}\mathbf{1}(m_k>m_{k-1})\Bigg) \tag{2.3.6}$$

where, by convention, $\sum_a^b=0$ for $a>b$.

Vectors (2.3.5) and (2.3.6) are those with independent coordinates and are of the fixed dimension $2k$. Denote the jth coordinates of these vectors by Ψ_{jn}, and Υ_{jn}, respectively, and set $\boldsymbol{\Psi}_n=(\Psi_{1n},\ldots,\Psi_{2k,n})$, $\boldsymbol{\Upsilon}_n=(\Upsilon_{1n},\ldots,\Upsilon_{2k,n})$. Let the symbol $\mathcal{P}_X$ denote the distribution of a r.v. or a random vector X.

First, note that the families of the distributions $\{\mathcal{P}_{\boldsymbol{\Psi}_n}\}$ and $\{\mathcal{P}_{\boldsymbol{\Upsilon}_n}\}$ are compact. Second, by results of [15]-[16] mentioned in the Introduction, condition (1.1.3) implies that

$$\prod_{j\in B_n}F_{jn}-\prod_{j\in B_n}\Phi_{jn}\Rightarrow 0$$

weakly for any sequence $\{B_n\}$ of sets of indices. In particular, this means that

$$\mathcal{P}_{\Psi_{jn}}-\mathcal{P}_{\Upsilon_{jn}}\Rightarrow 0$$

weakly for each j. Since the coordinates of the vectors $\boldsymbol{\Psi}_n$ and $\boldsymbol{\Upsilon}_n$ are independent, this implies that

$$\mathcal{P}_{\boldsymbol{\Psi}_n}-\mathcal{P}_{\boldsymbol{\Upsilon}_n}\Rightarrow 0.$$

On the other hand, in view of (2.3.3) and (2.3.4), each r.v. $X_n(t_i)$ is a linear combination of the r.v.'s Ψ_{jn}, and each r.v. $Y_n(t_i)$ is the linear combination of the r.v.'s Υ_{jn} with *the same coefficients* as for $X_n(t_i)$. Together with the compactness of $\mathcal{P}_{\boldsymbol{\Psi}_n}$ and $\mathcal{P}_{\boldsymbol{\Upsilon}_n}$, this leads to (2.3.2). ■

Since the sequence of the distributions $\{Q_n\}$ is compact, the proof of Proposition 3 is straightforward, and we skip it.

References

[1] Davydov, Yu.A and Rotar, V.I., *On asymptotic proximity of distributions*, to appear.

[2] Donsker, M.D., *An invariance principle for certain probability limit theorems*, Mem.Amer.Math.Soc., 6, 1951.

[3] Gikhman, I.I. and Skorokhod, A.V., *The Theory of Stochastic Processes*, Springer-Verlag, 1970.

[4] Jacod, J. and Shiryaev, A.N., *Limit Theorems for Stochastic Processes*, Springer, 1987, 2003.

[5] Kiryanova, L.V. and Rotar, V.I., *On non-classical conditions or convergence of convolutions to the Poisson distribution*, in "Stochastic processes and applications", MIEM, 1987.

[6] Kruglov, V.M., *Limit theorems for sums of independent random variables with values in Hilbert's space*, Theory of Probab. and its Applic., XVII, 2, 1972.

[7] Lévy, P., *Thèorie de l'addition des variables alèatoires,* Paris: Gauthier-Villars, 1954.

[8] Liptser, R.Sh. and Shiryaev, A.N., *On the invariance principle for semi-martingales: the "non-classical case"*, Theory of Probab. and its Applic., XXVIII, 1, 1983.

[9] Loève, M., *Probability Theory*, 3rd edition, Princeton, N.J., Van Nostrand , 1963.

[10] Machis, Yu.Yu., *Limit theorems in the non-classical setup*, Theory of Probab. and its Applic., XVI, 1, 1971.

[11] Prokhorov, Yu.,V., *Convergence of random processes and limit theorems in probability theory*, Theory Probabl. Appl., I, 2, 1956.

[12] Rotar, V.I., *On a generalization of the Lindeberg-Feller theorem*; Math.Notes, 1975, 1.

[13] Rotar, V.I., *On summation of independent variables in the nonclassical situation*; Russian Mathematical Surveys, 37, 6, 1982.

[14] Rotar, V.I., *Probability Theory*, World Scientific, 1998.

[15] Rotar, V.I. and Sholomitsky, A.G., *Necessary and sufficient conditions for proximity of convolution*s; Proceedings of the Sixth USSR-Japan Symposium on Probability Theory and Mathematical Statistics, World Scientific Publishing, 1992.

[16] Rotar, V.I. and Sholomitsky, A.G., *On proximity of convolutions*; Theory Probabl. Appl., XXXVII, 2, 1992,.

[17] Shiryaev, A.N., *Probability*, 2nd edition, Springer, 1996.

[18] Zolotarev, V.M., *A generalization of the Lindeberg-Feller theorem*, The Theory of Probab. and its Applic., XII, 4, 1967.

[19] Zolotarev, V.M., *Modern Theory of Summation of Random Variables*, V.S.P. Intl Science, 1997.