The Annals of Applied Probability
1998, Vol. 8, No. 3, 775–792

COINTEGRATED PROCESSES WITH INFINITE VARIANCE INNOVATIONS

BY VYGANTAS PAULAUSKAS AND SVETLOZAR T. RACHEV

Vilnius University and University of California, Santa Barbara

It is widely accepted that the Gaussian assumption is too restrictive to model either financial or some important macroeconomic variables, because their distributions exhibit asymmetry and heavy tails. In this paper we develop the asymptotic theory for econometric cointegration processes under the assumption of infinite variance innovations with different distributional tail behavior. We extend some of the results of Park and Phillips which were derived under the assumption of finite variance errors.

1. Introduction and summary of results. Stable non-Gaussian models have attracted the attention of economists, econometricians, probabilists, statisticians and time series analysts. The use of heavy-tailed distributions and of stable non-Gaussian models in econometrics and financial mathematics can be traced to the early papers of Mandelbrot (1963a, b) and Fama (1965) from the 1960's to the more recent papers of Akgiray and Booth (1988), Akgiray, Booth and Seifert (1988), Koedijk and Kool (1992), Rachev, Kim and Mittnik (1996). An extensive list of references and a review of various econometric and time series models under the heavy-tailed non-Gaussian hypothesis may be found in several recent papers [see Mittnik and Rachev (1993), Rachev, Kim and Mittnik (1996)].

The concept of cointegration was introduced in the seminal work of Granger (1981) and further developed by Engle and Granger (1987), Engle and Yoo (1987), Stock and Watson (1988) and Johansen (1988,1991). The classical method of removing the stochastic trend in nonstationary econometric time series is to use differencing procedures. However, it is now recognized that many econometric variables are "cointegrated," meaning that the linear combination of the integrated variables is stationary. The basic idea of a cointegrated process is that each of the processes considered is stochastically trended (i.e., nonstationary) but some linear combination of the processes reduces the order of integratedness and produces a stationary error process. More precisely, if all components of a multivariate econometric time series have a unit root, there may exist linear combinations of the components that are without a unit root. Typically, these linear combinations of the components are viewed as long-term relations between the underlying economic variables. An economical interpretation is that, even though most macroeconomic variables, like GNP and money stock, are themselves trended, the difference between them

Received April 1997; revised October 1997.
AMS 1991 *subject classifications*. 60F17, 60H05, 62M10.
Key words and phrases. Cointegrated processes, stable distributions, Lévy processes, ordinary least-squares estimators.

is constant in the long run. We may regard these series as defining a long-run equilibrium relationship, as the difference between them is stationary. Moreover, the representation theorem [see Engle and Granger (1987)] states that if a set of variables are cointegrated, then there exists a valid error correction representation of the data. Granger (1981) analyzed the integrated processes of economic time series and Nelson and Plosser (1982) reported strong evidence of nonstationarity in U.S. macroeconomic time series. Rachev, Kim and Mittnik (1996) studied the stochastic trends in economic univariate series under the Paretian assumption.

One of the convienent representations for a cointegrated process was introduced by Park and Phillips (1988). Let

$$Y_k = AX_k + u_k, \qquad k = 1, 2, \dots, n, \tag{1.1}$$

where Y_k is p-dimensional random vector, $A = \{a_{ij}\}$ is $p \times q$ matrix of coefficients and the q-dimensional vectors X_k, $k \geq 0$, are generated by a random-walk process

$$X_k = X_{k-1} + v_k. \tag{1.2}$$

We shall use the following notation for the coordinates of the introduced vectors

$$Y_k = (y_{k1}, \dots, y_{kp})', \qquad X_k = (x_{k1}, \dots, x_{kq})',$$
$$u_k = (u_{k1}, \dots, u_{kp})', \qquad v_k = (v_{k1}, \dots, v_{kq})'.$$

All vectors are vector columns and the sign $'$ stands for transposition. Since the results that we are interested in do not depend on the initial value X_0 of the process (1.2), we may assume that X_0 is an arbitrary random variable (or, in particular, a constant).

Let $r = p + q$ and $w_k = (w_{k1}, \dots, w_{kr})'$, where $w_{ki} = u_{ki}$ if $i = 1, \dots, p$ and $w_{k,\,p+j} = v_{kj}$ for $j = 1, \dots, q$.

The assumptions on the innovation process w_k, $k \geq 1$ in Park and Phillips (1988), roughly speaking, were the following [the exact formulation can be found in Phillips and Durlauf (1986), Theorem 2.1]. Vectors w_k were allowed to be weakly dependent, satisfying some mixing conditions such that the multivariate invariance principle for appropriately normed sums $\sum_k^{[nt]} w_k$ holds with multivariate Wiener process as a limit. This means that all coordinates of w_t has finite variances (and sometimes existence of moments of higher order is assumed, which is rather usual, considering weakly dependent summands). In our paper we relax this condition and consider the so-called case of "heavy-tailed" innovations. Although we assume a rather restrictive condition that w_k, $k = 1, 2, \dots$ are i.i.d. random vectors, we do not assume independence of coordinates of the innovation vector and allow different tail behavior for different coordinates. Thus we are dealing with a limit theorem with an operator-stable law as a limit and allow diagonal matrix normalization. The last step toward generality in this direction would be to consider innovations w_k in the domain of attraction of a general operator-stable distribution. Precisely stated, our assumptions on the sequence $\{w_k,\ k \geq 1\}$ generating the cointegrated model are as follows. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r)$ be an multiindex, satisfying $1 < \alpha_i \leq 2$, $i = 1, \ldots, r$, $S_n = D_n \sum_{i=1}^n w_i$, where $D_n = \operatorname{diag}(n^{-1/\alpha_1}, \ldots, n^{-1/\alpha_r})$ and $\operatorname{diag}(d_1, \ldots, d_k)$ stands for a diagonal $k \times k$ matrix with entries d_i in the diagonal. Since we assume that all $\alpha_i > 1$, then without loss of generality we may assume that $Ew_k = 0$, $k \geq 1$. Here it is worth mentioning that this assumption is made only because in practice it is difficult to interpret innovations having infinite mean values.

ASSUMPTION A. Here w_k, $k \geq 1$ is a sequence of i.i.d. random vectors belonging to the domain of normal attraction (DNA) of an $\overline{\alpha}$-stable r-dimensional vector $\xi(1)$; that is,

$$S_n \Rightarrow \xi(1) \quad \text{in } R^r, \tag{1.3}$$

where $\Rightarrow$ stands for the weak convergence of distributions (or random elements).

Let $D_r \equiv D([0, 1], R^r)$ be the space of R^r-valued cadlag functions defined on $[0, 1]$ and equipped with the Skorokhod topology. Let

$$Z_n(t) = S_{[nt]}, \qquad 0 \leq t \leq 1$$

and let $\xi(t)$, $0 \leq t \leq 1$, be an r-dimensional Lévy stable process, determined by the $\overline{\alpha}$-stable random vector $\xi(1)$, appearing in Assumption A. In the Appendix it is shown that (1.3) is equivalent to the following relation:

$$Z_n \Rightarrow \xi \quad \text{in } D_r. \tag{1.4}$$

In Assumption A, instead of DNA we can use the domain of attraction (DA); only the cumbersome calculations arising when dealing with slowly varying functions have forced us to state our results in the framework of DNA.

More information about relations (1.3) and (1.4), some facts about Lévy stable processes, and relations between spaces D_r and $(D_1)^r$ will be given in the Appendix [see, also, Samorodnitsky and Taqqu (1993), Gikhman and Skorokhod (1969), Jacod and Shiryaev (1987)]. Here we only mention that we do not exclude the case where some exponents, α_j with $j = i_m$, $1 \leq i_1 < \cdots < i_k \leq r$ are equal to 2. Then the vector

$$\left(\xi_{i_1}(t), \ldots, \xi_{i_k}(t)\right)$$

will be the k-dimensional Brownian motion and will be independent of the vector of the remaining coordinates $(\xi_i(t),\ i \neq i_j,\ j = 1, \ldots, k)$.

Since the vector w_k consists of two parts of lengths p and q, then when dealing with $r \times r$ matrices ($r = p + q$) it will be convenient to denote corresponding blocks of the matrix in the following way: if $B = \{b_{ij}\}_{i,\, j=1,\ldots,r}$ then

$$B = \begin{bmatrix} [B]_{11} & [B]_{12} \\ [B]_{21} & [B]_{22} \end{bmatrix},$$

for example, $[B]_{11} = \{b_{ij}\}_{i,\, j=1,\ldots,p}$, $[B]_{21} = \{b_{ij}\}_{i=p+1,\ldots,r}^{j=1,\ldots,p}$.

We consider the ordinary least-squares (OLS) regression estimator of the matrix A, which will be denoted by $\widehat{A}_n$. Where it will not cause misunderstanding, we suppress the subscript n and will simply write $\widehat{A}$; the same remark will apply to other variables as well. Denote

$$\begin{aligned}\mathbb{X}' &= \mathbb{X}'_n = (X_1, X_2, \dots, X_n),\\ \mathbb{Y}' &= \mathbb{Y}'_n = (Y_1, Y_2, \dots, Y_n),\\ \mathbb{U}' &= \mathbb{U}'_n = (u_1, u_2, \dots, u_n),\\ \mathbb{V}' &= \mathbb{V}'_n = (v_1, v_2, \dots, v_n).\end{aligned}$$

Here $\mathbb{X}'$ and $\mathbb{V}'$ are $q \times n$ matrices and $\mathbb{Y}'$ and $\mathbb{U}'$ are $p \times n$ matrices. Then [see, e.g., Park and Phillips (1988) or Phillips and Durlauf (1986)],

$$\widehat{A} = \mathbb{Y}'\mathbb{X}\big(\mathbb{X}'\mathbb{X}\big)^{-1} \tag{1.5}$$

and

$$\widehat{A} - A = \mathbb{U}'\mathbb{X}\big(\mathbb{X}'\mathbb{X}\big)^{-1}. \tag{1.6}$$

In the case where the innovations u_k and v_k have finite variances (all $\alpha_i = 2$) the normalization of the quantity (1.6) is simple $(n(\widehat{A}_n - A))$; in contrast, here for each entry of the matrix $\widehat{A}_n - A$ a different normalization is needed. To this end, we introduce the diagonal matrices

$$\begin{aligned}\mathbb{T}_1 &= \mathbb{T}_{1,n} = \operatorname{diag}\big(n^{-1/\alpha_1}, \dots, n^{-1/\alpha_p}\big),\\ \mathbb{T}_2 &= \mathbb{T}_{2,n} = \operatorname{diag}\big(n^{-1/\alpha_{p+1}}, \dots, n^{-1/\alpha_{p+q}}\big).\end{aligned} \tag{1.7}$$

It is not difficult to see that the proper normalizations for $\mathbb{U}'\mathbb{X}$ and $\mathbb{X}'\mathbb{X}$ are $\mathbb{T}_1\mathbb{U}'\mathbb{X}\mathbb{T}_2$ and $n^{-1}(\mathbb{T}_2\mathbb{X}'\mathbb{X}\mathbb{T}_2)$, respectively. Therefore the proper normalization for the quantity (1.6) is

$$n\big(\mathbb{T}_1(\widehat{A} - A)\mathbb{T}_2^{-1}\big). \tag{1.8}$$

It remains to introduce some notation describing the limit distribution of (1.8). We recall that $\xi(\cdot)$ is a limiting process in (1.4), $\xi_i(\cdot)$, $i = 1, 2, \dots, r$ being its coordinates. The notation

$$\int_0^t \xi_i^-(s)\, d\xi_j(s)$$

will stand for the Itô stochastic integral. Here $x^-(s)$ denotes the left limit of the function $x \in D[0,1]$ at point $0 < s \le 1$. For simplicity of writing, we shall suppress this superscript and write $\int_0^t \xi_i(s)\, d\xi_j(s)$ or simply $\int_0^t \xi_i\, d\xi_j$ when there is no ambiguity in such notation. Then $\int_0^t \xi(s)\, d(\xi(s))'$, or simply, $\int_0^t \xi\, d\xi'$ stands for the matrix with elements $\int_0^t \xi_i\, d\xi_j$, $i, j = 1, \dots, r$. Next we denote

$$\begin{aligned}\big[\xi_i, \xi_j\big]_t &= \xi_i(t)\xi_j(t) - \int_0^t \xi_j\, d\xi_i - \int_0^t \xi_i\, d\xi_j,\\ \big[\xi, \xi\big]_t &= \big\{\big[\xi_i, \xi_j\big]_t\big\}_{i,j=1,\dots,r}.\end{aligned}$$

Several monographs and textbooks are devoted to the theory of stochastic integration, for example, Protter (1990), Elliott (1982) and Kopp (1984).

Our main result follows.

THEOREM. *Suppose that in the cointegrated processes model* (1.1) *and* (1.2), *the sequence of innovations* $\{w_k, k \geq 1\}$ *satisfies Assumption* A, *and therefore the invariance principle* (1.4) *holds. Then for the estimated matrix* $\widehat{A}_n$, *given by* (1.5), *the following limit relation* (*weak convergence in* $R^{p\times q}$) *holds*:

$$
\begin{aligned}
&n(\mathbb{T}_1(\widehat{A}_n - A)\mathbb{T}_2^{-1}) \\
&\quad\Rightarrow \left[\xi(1)(\xi(1))' - \int_0^1 \xi(s)\, d(\xi(s))'\right]_{12} \left\{\left[\int_0^1 \xi(s)(\xi(s))'\, ds\right]_{22}\right\}^{-1},
\end{aligned}
\tag{1.9}
$$

as $n \to \infty$, *where the diagonal matrices* $\mathbb{T}_i$, $i = 1, 2$ *are defined in* (1.7).

REMARK 1. It is easy to see that the limit distribution in (1.9) can be written in a different, but equivalent form, namely as

$$
\left[\left(\int_0^1 \xi(s)\, d(\xi(s))'\right)' + [\xi, \xi]_1\right]_{12} \left\{\left[\int_0^1 \xi(s)(\xi(s))'\, ds\right]_{22}\right\}^{-1}.
$$

REMARK 2. The univariate model (1.1) with Y_{k-1} instead of X_k and with $p = 1$ and $A = 1$ was considered earlier by Chan and Tran (1989) [see also Phillips (1990)]. Caner (1995) studied the model (1.1) with $A = I$, $X_k = Y_{k-1}$, $p \geq 1$, assuming a more complicated structure of the innovations u_t. One of the goals of the paper is to provide a mathematically rigorous proof of the main asymptotic result in cointegration theory when the innovations are i.i.d. and are in the domain of attraction of $(\alpha_1, \dots, \alpha_r)$-stable law. The novelty of our approach is the use of general results for the convergence of stochastic integrals for semimartingales, proved by Jakubowski, Memin and Pages (1989) and Kurtz and Protter (1991a,b), (1996a,b). This approach is new even in the Gaussian case, and it promises to weaken the i.i.d. assumption and allow the investigation of weakly dependent and nonidentically distributed innovations. We believe that limit theorems based on verification of the UT condition (using results from the above-cited papers and references therein) may lead to simpler proofs of some existing limiting results in econometrics as well as provide approaches to new results, such as asymptotic analysis of Johansen's cointegration model with stable innovations [see Johansen (1991)].

REMARK 3. As was pointed out by an Associate Editor, there is another possible approach to our limiting problem; namely, instead of using limit theorems for partial sums, one can apply convergence results for point processes generated by the innovations. Although the functionals arrising in this approach are also discontinuous, there are various ways to deal with these types of obstacles; see Davis and Resnick (1985) and Resnick (1987).

2. Proofs. We can write

$$n\big(\mathbb{T}_1(\widehat{A}-A)\mathbb{T}_2^{-1}\big) = (\mathbb{T}_1\mathbb{U}'\mathbb{X}\mathbb{T}_2)\big(n^{-1}(\mathbb{T}_2\mathbb{X}'\mathbb{X}\mathbb{T}_2)\big)^{-1};$$

therefore it is clear that the main step in the proof of the theorem is the following lemma.

LEMMA 1. *Under Assumption A, we have as $n \to \infty$,*

$$\begin{aligned}&\big(\mathbb{T}_1\mathbb{U}'\mathbb{X}\mathbb{T}_2,\, n^{-1}(\mathbb{T}_2\mathbb{X}'\mathbb{X}\mathbb{T}_2)\big)\\ &\quad\Rightarrow \left[\xi(1)(\xi(1))' - \int_0^1 \xi(s)\,d(\xi(s))'\right]_{12}, \left\{\left[\int_0^1 \xi(s)(\xi(s))'\,ds\right]_{22}\right\}^{-1},\end{aligned} \tag{2.1}$$

in R^{qr}.

By the continuous mapping theorem and (2.1), taking into account that

$$P\left\{\det\left[\int_0^1 \xi(s)(\xi(s))'\,ds\right]_{22} = 0\right\} = 0,$$

we obtain (1.9). Before proving (2.1) we make the following comments. Without mentioning it, we shall use repeatedly the following well-known result [see, for example, Billingsley (1968), Theorem 4.4]: if $X_n \Rightarrow X_0$, $Y_n \to_p a$, then $(X_n, Y_n) \Rightarrow (X_0, a)$ and, in particular, if addition is defined and is a continuous operation, then $X_n + Y_n \Rightarrow X_0 + a$. Here random elements X_n and Y_n are with values in a separable metric space, and we shall use this result in the case of multidimensional spaces and $a = 0$.

The next comment concerns the idea of the proof of (2.1). This relation is of the form

$$W_n \Rightarrow W_0,$$

where W_n and W_0 are qr-dimensional random vectors. Since W_n is some function of the process Z_n and we have (1.4), one may get the impression that the continuous mapping theorem gives us (2.1) without difficulties. Unfortunately, this is not the case, for the following reason. The part of coordinates of the right-hand side vector in (2.1) are in the form of stochastic integrals

$$\int_0^1 \xi_i(t)\,d\xi_j(t)$$

and generally we cannot prove the limit relation using the continuous mapping theorem, unless the integrals under consideration can be understood as path-by-path Stieltjes integrals. In order for the last situation to hold, at least one of the processes ξ_i or ξ_j must be of finite variation with probability 1. It is known (unfortunately, we did not find a relevant reference for this fact; on the other hand it is not difficult to prove it) that Lévy stable processes with index $\alpha > 1$ are of unbounded variation on every bounded interval a.s., therefore the limit relation (2.1) cannot be obtained using the continuous mapping theorem only. There were earlier papers dealing with such a problem [see, e.g., Chan and Wei (1988), Jeganathan (1991)]. Here we propose the following approach to overcome the above dificulty, and we explain the idea (which is a simple one and sends a message: use the powerful results from stochastic analysis) in a simple two-dimensional situation. Suppose that we have a sequence (X_n, Y_n) of two-dimensional processes and $(X_n, Y_n) \Rightarrow (X_0, Y_0)$ in D_2. Let f: $D_2 \to R^2$, $f = (f_1, f_2)$, $Z_n = f(X_n, Y_n)$, $Z_0 = f(X_0, Y_0)$. We want to prove that $Z_n \Rightarrow Z_0$, but we know that the relation

$$f_1(X_n, Y_n) \Rightarrow f_1(X_0, Y_0)$$

cannot be obtained by the continuous mapping theorem [one reason for this can be that f_1 is not continuous on the support of (X_0, Y_0)]. Let us denote $U_n = f_1(X_n, Y_n)$, $U_0 = f_1(X_0, Y_0)$. Suppose that we can prove the relation

$$(X_n, Y_n, U_n) \Rightarrow (X_0, Y_0, U_0) \quad \text{in } D_2 \times R. \tag{2.2}$$

When the functional f_1 is a stochastic Itô integral (and this case is of interest to us) there is a vast literature devoted to limit theorems of such a type [see, for example, Kurtz and Protter (1991a), (1996a,b), Stricker (1985), Jakubowski, Memin and Pages (1989)]. Now we take

$$g: D_2 \times R \to R^2, \ g(x, y, u) = (u, f_2(x, y)), \qquad (x, y) \in D_2, \ u \in R. \tag{2.3}$$

If f_2 is a continuous functional on D_2, then g is the continuous mapping on $D_2 \times R$. Now (2.2), (2.3) and the continuous mapping theorem give us

$$g(X_n, Y_n, U_n) \Rightarrow g(X_0, Y_0, U_0)$$

and since $g(X_n, Y_n, U_n) = Z_n$ and $g(X_0, Y_0, U_0) = Z_0$ we have the wanted relation.

Now we can start with a proof of (2.1).

PROOF OF LEMMA 1. First we consider separately the matrices on the left-hand side of (2.1). If $A = \{a_{ij}\}$ then $(A)_{ij}$ will stand for a_{ij}. We have, for $i = 1, \dots, p$, $j = 1, \dots, q$,

$$\begin{aligned}
(\mathbb{T}_1 \mathbb{U}' \mathbb{X} \mathbb{T}_2)_{ij} &= n^{-(1/\alpha_i + 1/\alpha_{p+j})} \sum_{k=1}^{n} u_{ki} x_{kj} \\
&= n^{-(1/\alpha_i + 1/\alpha_{p+j})} \sum_{k=1}^{n} u_{ki} \left(x_{0j} + \sum_{m=1}^{k} v_{mj} \right) \\
&= \left(n^{-1/\alpha_i} \sum_{k=1}^{n} u_{ki} \right) \left(n^{-1/\alpha_{p+j}} \sum_{k=1}^{n} v_{kj} \right) \\
&\quad - \sum_{k=1}^{n} n^{-1/\alpha_{p+j}} v_{kj} \left(n^{-1/\alpha_i} \sum_{m=1}^{k-1} u_{mi} \right) + n^{-1/\alpha_{p+j}} x_{0j} n^{-1/\alpha_i} \sum_{k=1}^{n} u_{ki},
\end{aligned}$$

where $\sum_{i=1}^{0} a_i = 0$. Since

$$n^{-1/\alpha_i} \sum_{k=1}^{t} u_{ki} = Z_{ni}\left(\frac{t}{n}\right), \qquad n^{-1/\alpha_{p+j}} \sum_{k=1}^{t} v_{kj} = Z_{n,\, p+j}\left(\frac{t}{n}\right),$$

then we can write

$$(\mathbb{T}_1 \mathbb{U}' \mathbb{X} \mathbb{T}_2)_{ij} = Z_{ni}(1) Z_{n,\, p+j}(1) - \int_0^1 Z_{ni}(t) dZ_{n,\, p+j}(t) + o_p(1). \tag{2.4}$$

Now, for $i,\, j = 1, \dots, q$,

$$n^{-1}(\mathbb{T}_2 \mathbb{X}' \mathbb{X} \mathbb{T}_2)_{ij} = n^{-(1+1/\alpha_{p+i}+1/\alpha_{p+j})} \sum_{k=1}^{n} x_{ki} x_{kj}.$$

Since $x_{ki} = x_{0i} + \sum_{j=1}^{k} v_{ji}$, we obtain, after simple calculations,

$$n^{-1}(\mathbb{T}_2 \mathbb{X}' \mathbb{X} \mathbb{T}_2)_{ij} = \int_0^1 Z_{n,\, p+i}(s) Z_{n,\, p+j}(s)\, ds + o_p(1). \tag{2.5}$$

From (2.4) and (2.5) we see that the only terms for which we cannot apply the continuous mapping theorem are the stochastic integrals

$$\int_0^1 Z_{n,\, i}(s)\, dZ_{n,\, p+j}(s), \qquad i = 1, \dots, p, \;\; j = 1, \dots, q.$$

Therefore, according to our plan [see (2.2)], we need the following result.

PROPOSITION 2. *Under Assumption A we have the relation*

$$\begin{aligned} &\left(Z_n, \int_0^t Z_{ni}(u)\, dZ_{n,\, p+j}(u),\; i = 1, \dots, p,\; j = 1, \dots, q\right) \\ &\qquad \Rightarrow \left(\xi, \int_0^t \xi_i(u)\, d\xi_{p+j}(u),\; i = 1, \dots, p,\; j = 1, \dots, q\right) \quad \text{in } D_s, \end{aligned} \tag{2.6}$$

as $n \to \infty$, *where* $s = p + q + pq$.

PROOF. There are a large number of papers devoted to convergence of stochastic integrals, only a small part of which were mentioned above. The problem can be formulated as follows. We have two sequences of (real or vector-valued) stochastic processes X_n and Y_n and we know that $(X_n, Y_n) \Rightarrow (X_0, Y_0)$ in an appropriate space. Then we look at what conditions ensure that

$$\left(X_n, Y_n, \int X_n\, dY_n\right) \Rightarrow \left(X_0, Y_0, \int X_0\, dY_0\right).$$

It turns out that in a very general situation, when processes X_n and Y_n are semimartingales, the last relation holds if the so-called UT (uniform tightness) condition for the sequence Y_n is satisfied. This condition was introduced in Stricker (1985), and a general result was proved in Jakubowski, Memin and Pages (1989). In Kurtz and Protter (1991a) another condition was given and it was proved [see Kurtz and Protter (1991a, b), Memin and Slominski (1991)] that both conditions are equivalent and, furthermore, they are necessary in some sense.

We shall formulate the general result from Kurtz and Protter (1991a), which we shall use in our setting, although we do not define all concepts (such as adapted processes, local martingales, stopping times) which appear in the formulation. At the final stage of preparation of this paper, J. Memin sent us a letter with a sketch of the verification of the UT condition in a one-dimensional case using the form of the condition given in Jakubowski, Memin and Pages (1989).

Let X_n be a sequence of random processes with sample paths in $D([0,\infty), \mathbb{M}^{k,m})$, where $\mathbb{M}^{k,m}$ stands for real-valued $k \times m$ matrices, and let Y_n be another sequence of random processes from $D([0,\infty), R^m)$. For a process $A(t)$ with sample paths of finite variation on bounded time intervals, denote

$$\mathrm{Var}_t(A) = \sup \sum_i \|A(t_{i+1}) - A(t_i)\|,$$

where the supremum is taken over finite partitions of $[0,t]$. Here and in what follows we use the following notation for the norms in R^m: $\|x\| = (\sum_{i=1}^n x_i^2)^{1/2}$, $\|x\|_\infty = \max_{1\le i\le m} |x_i|$. Without explicitly mentioning it, we shall use the fact that all norms on a finite-dimensional space are equivalent.

THEOREM A [Kurtz and Protter (1991a), Theorem 2.7]. *For each n, let (X_n, Y_n) be an $\mathscr{F}_t^n$-adapted process with sample paths in $D([0,\infty), \mathbb{M}^{k,m} \times R^m)$ and let Y_n be an $\mathscr{F}_t^n$-semimartingale. Suppose that $Y_n = M_n + A_n + \tilde{Y}_n$ where M_n is a local $\mathscr{F}_t^n$-martingale, A_n is an $\mathscr{F}_t^n$-adapted, finite variation process and $\tilde{Y}_n$ is constant except for finitely many discontinuities in any finite time interval. Let $N_n(t)$ denote the number of discontinuities of $\tilde{Y}_n$ in the interval $[0,t]$. Suppose $\{N_n(t)\}$ is stochastically bounded for each $t > 0$ and (UT) for each $r > 0$, there exist stopping times τ_n^r such that $P\{\tau_n^r \le r\} \le 1/r$ and*

$$\sup_n E\{[M_n]_{t\wedge\tau_n^r} + \mathrm{Var}_{t\wedge\tau_n^r}(A_n)\} < \infty. \tag{2.7}$$

If $(X_n, Y_n, \tilde{Y}_n) \Rightarrow (X_0, Y_0, \tilde{Y}_0)$ in the Skorokhod topology on $D([0,\infty), \mathbb{M}^{k,m} \times R^m \times R^m)$, then Y_0 is a semimartingale with respect to a filtration to which X_0 and Y_0 are adapted and

$$\left(X_n, Y_n, \int_0^t X_n\, dY_n\right) \Rightarrow \left(X_0, Y_0, \int_0^t X_0\, dY_0\right)$$

in the Skorokhod topology on $D([0,\infty), \mathbb{M}^{k,m} \times R^m \times R^k)$. If $(X_n, Y_n, \tilde{Y}_n) \to (X_0, Y_0, \tilde{Y}_0)$ in probability, then the triple converges in probability.

The setting which we are interested in is a little bit different from the result formulated above; in our case X_n will be not a matrix-valued but rather a vector-valued random process and we deal with the interval $[0,1]$ instead of $[0,\infty)$. But it is not difficult to see that minor changes allow us to consider our setting and essentially the only thing we need to do is to verify a UT-type condition for the sequence of processes $\{Z_{n,\,p+j},\ \ j = 1, \dots, q\}$ ($\{Z_{n,\,i},\ i = 1, \dots, p\}$ in our case will be as $\{X_n\}$). In order not to write $p + j$ every time, we "rename" the second part of S_n, Z_n and ξ in the following way:

$$\overline{Z}_n = (\overline{Z}_{n1}, \dots, \overline{Z}_{nq}), \qquad \overline{S}_n = (\overline{S}_{n1}, \dots, \overline{S}_{nq}), \qquad \overline{\xi} = (\overline{\xi}_1, \dots, \overline{\xi}_q),$$

where $\overline{Z}_{nj} = Z_{n,\,p+j}$, $\overline{S}_{nj} = S_{n,\,p+j}$, $\overline{\xi}_j = \xi_{p+j}$. If $a_{nj} = n^{1/\alpha_{p+j}}$, $j = 1, \dots, q$, then

$$\overline{S}_{nj} = a_{nj}^{-1} \sum_{k=1}^{n} v_{kj},$$

$$\overline{Z}_{nj}(t) = \overline{S}_{[nt],\,j}, \qquad 0 \le t \le 1,$$

and, under Assumption A,

$$\overline{Z}_n \Rightarrow \overline{\xi} \quad \text{in } D_q.$$

We recall that we have assumed $Ev_{1j} = 0$ and that Assumption A implies

$$\Lambda_k = \sup_{t>0} t^{\alpha_{p+k}} P\big\{|v_{1k}| > t\big\} < \infty, \qquad k = 1, \dots, q.$$

As a first step in proving the UT condition, we separate large jumps of the process $\overline{Z}_n$ from small ones. Let $b > 0$ be some fixed number and

$$v_i^{(n)} = \big(v_{i,\,1}^{(n)}, \dots, v_{i,\,q}^{(n)}\big), \qquad v_{i,\,j}^{(n)} = a_{nj}^{-1} v_{i,\,j},$$

$$B(b) = \big\{x \in R^q \colon \|x\|_\infty \le b\big\},$$

$$a_n(b) = E v_1^{(n)} \mathbb{1}\big\{v_1^{(n)} \in B(b)\big\},$$

$$\tilde{v}_i^{(n)}(b) = v_i^{(n)} \mathbb{1}\big\{v_i^{(n)} \in B(b)\big\} - a_n(b),$$

$$\tilde{\tilde{v}}_i^{(n)}(b) = v_i^{(n)} \mathbb{1}\big\{v_i^{(n)} \in (B(b))^c\big\}.$$

Now we can write

$$\overline{Z}_n(t) = Y_n^{(1)}(t) + A_n(t) + Y_n^{(2)}(t),$$

where

$$Y_n^{(1)}(t) = \sum_{i=1}^{[nt]} \tilde{v}_i^{(n)}(b), \qquad Y_n^{(2)}(t) = \sum_{i=1}^{[nt]} \tilde{\tilde{v}}_i^{(n)}(b), \qquad A_n(t) = [nt] a_n(b).$$

PROPOSITION 3. *For any fixed $0 < b < \infty$, the function $A_n(t)$ is of finite variation, the process $Y_n^{(2)}$ is of finite variation a.s. and there exists a constant C, depending on b and on parameters of distribution of v_1 such that*

$$\sup_n \operatorname{Var}_1(A_n) \le \sup_n E \operatorname{Var}_1(Y_n^{(2)}) \le C.$$

If $N_n(t)$ denotes the number of discontinuities of $Y_n^{(2)}$ in the interval $[0, t]$, then $\{N_n(t), n \ge 1\}$ is stochastically bounded for each t.

PROOF. From the definition of variation it is easy to see that, due to the relation $a_n(b) = -E\tilde{\tilde{v}}_1^{(n)}(b)$, we have to prove that

$$\sup_n nE\left\|\tilde{\tilde{v}}_1^{(n)}(b)\right\| \le C. \tag{2.8}$$

Let F_n stand for the distribution of $\|v_1^{(n)}\|_\infty$. Then

$$E\left\|\tilde{\tilde{v}}_1^{(n)}(b)\right\| \le CE\left\|\tilde{\tilde{v}}_1^{(n)}(b)\right\|_\infty \le C\int_b^\infty x\,dF_n(x).$$

However,

$$bP\left\{\|v_1^{(n)}\|_\infty > b\right\} \le \sum_{i=1}^q bP\left\{|v_{1i}| > ba_{ni}\right\} \le \frac{1}{n}\sum_{i=1}^q \Lambda_i.$$

In a similar way we can estimate the term

$$\int_b^\infty P\left\{\|v_1^{(n)}\|_\infty > x\right\}dx \le \sum_{i=1}^q \int_b^\infty P\left\{|v_{1i}| > xa_{ni}\right\}dx \le \frac{1}{n}\sum_{i=1}^q C_i\Lambda_i,$$

where $C_i = \int_b^\infty x^{-\alpha_{p+i}}\,dx$. From these estimates (2.8) follows. Now since $N_n(t) = \sum_{i=1}^{[nt]} \mathbb{1}\{v_i^{(n)} \in (B(b))^c\}$ and

$$P\left\{N_n(t) > c\right\} \le c^{-1}EN_n(t) = c^{-1}[nt]P\left\{v_1^{(n)} \in \left(B(b)\right)^c\right\},$$

the same inequalities prove the stochastic boundedness of $\{N_n(t)\}$ for all t and the proposition is proved.

Since Proposition 3 shows that the part of our process Z_n which is of finite variation can be controlled uniformly with respect to n without using stopping times, we construct stopping times based on $Y_n^{(1)}$ only. Noting that the process $Y_n^{(1)}$ is a martingale with respect to the natural filtration,

$$\mathscr{F}_t^n = \sigma\left(v_1, \dots, v_{[nt]}\right), \qquad 0 \le t \le 1$$

and that it is a jump process with jumps at the points $t = k/n$, $k = 1, \dots, n$, we need to consider the triangular array

$$\left\{\tilde{S}_n^k = \sum_{i=1}^k \tilde{v}_i^{(n)}(b),\ k = 0, 1, \dots, n\right\}, \qquad n \ge 1,$$

with its natural filtration $\mathscr{F}_k^n = \sigma(v_1, \dots, v_k)$.We also need to construct stopping times τ_n^d such that for each d,

$$\sup_n E\left\|\tilde{S}_n^{k\wedge\tau_n^d}\right\|^2 \le C, \tag{2.9}$$

and for each k there exists $d = d(k)$ such that

$$P\{\tau_n^d > k\} \ge 1 - \frac{1}{k}. \tag{2.10}$$

For any $d > 0$ let us define

$$\tau_n^d = \min\left\{k\colon \|S_n^j\| \le d,\ \ j = 1, \dots, k-1,\ \ \|\tilde{S}_n^k\| > d\right\}$$

with the agreement that min{empty set} $= +\infty$. It is a stopping time, since it is the hitting time of an open set $\left\{x \in R^q\colon \|x\| > d\right\}$ for the random process $Y_n^{(1)}$; see Protter (1990). Now, if $\tau_n^d > k$, then $\|\tilde{S}_n^k\| \le d$ and if $\tau_n^d \le k$, then

$$\|\tilde{S}_n^{\tau_n^d}\| = \|\tilde{S}_n^{\tau_n^d - 1} + \tilde{v}_{\tau_n^d}(b)\| \le d + \|\tilde{v}_{\tau_n^d}(b)\| \le d + c\|\tilde{v}_{\tau_n^d}(b)\|_\infty \le d + c \cdot b.$$

Therefore

$$E\left\|\tilde{S}_n^{k \wedge \tau_n^d}\right\|^2 \le (d + cb)^2$$

and we have (2.9). It is interesting to note that if we were able somehow to control the "size" of the jumps at a time of hitting the set $\{x\colon \|x\| > d\}$, there would be no need to separate large jumps. However, if we multiply $\tilde{\tau}_n^d$ (constructed in the same way, only with $\overline{S}_n$ instead of $\tilde{S}_n$) by the indicator function of the event $\{\|v_{\tilde{\tau}_n^d}^{(n)}\| \le b\}$, the multiplied random variable is no longer a stopping time. In order to get (2.10) we note that, for $0 < \alpha \le 1$,

$$P\left\{\tau_n^d > n\alpha\right\} = P\left\{\max_{j \le n\alpha} \|\tilde{S}_n^j\| \le d\right\}.$$

As $n \to \infty$, this probability approaches the probability

$$P\left\{\sup_{0 \le t \le \alpha} \|\xi^{(1)}(t)\| \le d\right\},$$

and this probability can be made arbitrarily close to 1 by taking sufficiently large d. Here $\xi^{(1)}$ is a stable Lévy process obtained as a limit for the sequence $Y_n^{(1)}$. That the latter sequence is convergent, as well as the convergence $(Z_n, Y_n^{(2)}) \Rightarrow (\xi, \xi^{(2)})$, which is required in Theorem A (we recall that here Z_n stands for both X_n and Y_n, $\overline{Z}_n$ playing the role of Y_n), can be proved in a way similar to Gikhman and Skorokhod (1969), Theorems 9.6.1 and 9.6.2, where truncation of summands in a triangular array is used, or by the point processes technique, as in Davis and Resnick (1985). Proposition 2 is proved.

Having (2.6) and taking the appropriate mapping $h\colon D_s \to R^{qr}$, we get (2.1). Therefore, Lemma 1, as well as the theorem is proved.

APPENDIX

We collect here some (known) facts about relations (1.3) and (1.4), Lévy processes and Skorokhod spaces.

Although we shall deal with relations (1.3) and (1.4), we introduce new notation, independent of the notations of the previous section. Thus let $X_i = (X_{i1}, \dots, X_{id})$, $i \ge 1$ be i.i.d. random vectors in R^d. In order not to deal with centering, we assume that $EX_{1j} = 0$ for those j for which the expectations exist.

Denote

$$\text{(A.1)}\qquad \begin{aligned} S_n &= (S_{n1},\dots,S_{nd}), & S_{nj} &= a_{nj}^{-1}\sum_{i=1}^{n} X_{ij},\\ Z_n(t) &= \big(Z_{n1}(t),\dots,Z_{nd}(t)\big), & Z_{nj} &= S_{[nt],\,j}, \qquad 0\le t\le 1. \end{aligned}$$

Although it is possible to state results with general norming vectors $a_n = (a_{n1},\dots,a_{nd})$, for simplicity of writing and clarity of understanding we take the special case where

$$\text{(A.2)}\qquad a_{nj} = n^{1/\alpha_j}, \qquad 1\le j\le d,$$

and $\overline{\alpha} = (\alpha_1,\alpha_2,\dots,\alpha_d)$ is a multiindex with $0<\alpha_i\le 2$, $i = 1,2,\dots,d$. This setting can be reformulated as follows: we consider sums of i.i.d. d-dimensional vectors in the DNA of operator-stable random vectors, restricting normalization only to diagonal matrices and thus we assume that marginal distributions of summands belong to the DNA of some stable univariate law, including the Gaussian one.

When dealing with stochastic process Z_n, two spaces are appropriate. One is $D([0,1], R^d)$, which has already been introduced. Another one is $(D[0,1])^d = D[0,1]\times\cdots\times D[0,1]$—the usual product of Skorokhod spaces $D[0,1]$ with product topology. For the first one we shall use the introduced notation D_d and for the second one we use the notation D^d. The relations between these spaces are as follows [see, e.g., Jacod and Shiryaev (1987)]. As "abstract" sets, they coincide; σ-fields of Borel sets on both spaces coincide, too, but as topological spaces they are different—the topology of D_d is strictly finer than product topology on D^d. Therefore, from the weak convergence of measures in space D_d there follows the weak convergence of these measures in D^d, but it is possible to give examples, showing that the converse statement is not true.

Let $\{Y(t),\ t\ge 0\}$ be the Lévy process with values in R^d, that is, a stochastically continuous process with independent and strictly stationary increments. Then it is well known [see, e.g., Protter (1990) or Gikhman and Skorokhod (1969)] that there exist a vector $a\in R^d$, a symmetric nonnegative defined matrix Γ and a measure ν on R^d satisfying

$$\nu\{0\} = 0, \qquad \int_{R^d} \|x\|^2\big(1+\|x\|^2\big)^{-1}\nu(dx) < \infty,$$

such that for any $z\in R^d$,

$$\text{(A.3)}\qquad \begin{aligned} E\exp\{i(z,Y(t))\} = \exp\Big\{t\Big[& i(z,a) - \tfrac{1}{2}(\Gamma z,z)\\ & + \int_{\|x\|\le 1}\big(\exp(i(x,z)) - 1 - (z,x)\big)\nu(dx)\\ & + \int_{\|x\|>1}\big(\exp(i(x,z)) - 1\big)\nu(dx)\Big]\Big\}. \end{aligned}$$

The measure ν is called the Lévy measure for the process Y. Matrix Γ corresponds to the Gaussian part of the Lévy process Y.

Lévy processes have cadlag sample paths, all are semimartingales, and they have good integrability properties. For example, we use in the main text the fact that one-dimensional α-stable Lévy processes on a finite interval are in L_p for any $p > 0$ [see Samorodnitsky and Taqqu (1993), page 510, for this fact].

We are interested in the following statements

$$S_n \Rightarrow Y(1) \quad \text{in } R^d, \tag{A.4}$$

$$Z_n(\cdot) \Rightarrow Y(\cdot) \quad \text{in } D^d, \tag{A.5}$$

$$Z_n(\cdot) \Rightarrow Y(\cdot) \quad \text{in } D_d. \tag{A.6}$$

Here for the multiindex $\overline{\alpha}$ we shall assume that $0 < \alpha_i \le \alpha_{i+1} \le 2$ for all i (although in the main text such an assumption generally is not acceptable, since we cannot rearrange the coordinates of innovation vectors).

Let us consider two cases. First:

$$0 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_d < 2. \tag{a}$$

PROPOSITION 4. *Let S_n, Z_n be defined by* (A.1) *and* (A.2) *and let Y be a Lévy process with Lévy spectral measure ν and $\Gamma \equiv 0$ in* (A.3). *Then all statements* (A.4)–(A.6) *are equivalent and each of them is equivalent to the following statement*:

$$\lim nP\{X_1^{(n)} \in A\} = \nu(A) \tag{A.7}$$

for all $A \in \mathscr{B}(R^d \setminus \{0\})$ such that $\nu(\partial A) = 0$, $\nu(A) < \infty$, here ∂A denotes the boundary of a set A and

$$X_1^{(n)} = \left(n^{-1/\alpha_1} X_{11}, \dots, n^{-1/\alpha_d} X_{1d}\right). \tag{A.8}$$

REMARK 1. The Lévy measure ν of the process Y can be described as follows. Let τ: $R^d \to R^d$, $\tau(x) = (\operatorname{sgn} x_1 |x_1|^{1/\alpha_1}, \dots, \operatorname{sgn} x_d |x|^{1/\alpha_d})$ and $\tilde{\nu} = \nu \circ \tau$. Then

$$\tilde{\nu}\left\{x\colon \|x\| \ge r, \frac{x}{\|x\|} \in B\right\} = r^{-1} H(B), \tag{A.9}$$

where H is some finite measure on the unit sphere of R^d, $S^d = \{x\colon \|x\| = 1\}$ and $B \in \mathscr{B}(S^d)$. It is easy to see that in the case $\alpha_1 = \alpha_2 = \cdots = \alpha_d = \alpha$ we get the usual result about the DNA and in this case

$$\nu\left\{x\colon \|x\| > r,\ \frac{x}{\|x\|} \in B\right\} = r^{-\alpha} H(B).$$

REMARK 2. Components of the limiting process Y [and of $Y(1)$, in particular] will be independent if the measure H in (A.9) is discrete and concentrated at points $\pm e_k$, $k = 1, 2, \dots, d$, where e_k is the kth element of the standard orthonormal basis in R^d. This happens if the coordinates of X_1 are independent, but this is not a necessary condition for independence of the coordinates of Y.

Now let us consider the second case:

(b) $$0 < \alpha_1 \le \cdots \le \alpha_k < \alpha_{k+1} = \cdots = \alpha_d = 2, \qquad 1 \le k < d.$$

In this case it is convienent to divide all vectors under consideration into two parts; therefore we shall write

$$S_n = (S_n^{(1)}, S_n^{(2)}), \qquad Z_n = (Z_n^{(1)}, Z_n^{(2)}),$$
$$Y_n = (Y^{(1)}, Y^{(2)}), \qquad Y(1) = \big(Y^{(1)}(1), Y^{(2)}(1)\big),$$
$$X_1^{(n)} = \big(X_1^{(n,1)}, X_1^{(n,2)}\big),$$

where, for example,

$$Y^{(1)}(t) = \big(Y_1(t), \dots, Y_k(t)\big),$$
$$Y^{(2)}(t) = \big(Y_{k+1}(t), \dots, Y_d(t)\big),$$
$$X_1^{(n,1)} = \big(n^{-1/\alpha_1} X_{11}, \dots, n^{-1/\alpha_k} X_{1k}\big).$$

Here $Y^{(1)}$ is a Lévy process in R^k and $Y^{(2)}$ is Brownian motion in R^{d-k}; namely, for $z \in R^{d-k}$,

$$E \exp(i(z, Y^{(2)}(t))) = \exp\{-\tfrac{1}{2} t(\Gamma z, z)\},$$

where Γ is a symmetric nonnegative definite $(d-k) \times (d-k)$ matrix. The processes $Y^{(1)}$ and $Y^{(2)}$ are independent; see Sharpe (1969). Let us consider one more relation:

(A.10) $$S_n^{(1)} \Rightarrow Y^{(1)}(1) \quad \text{in } R^k, \qquad S_n^{(2)} \Rightarrow N(0, \Gamma) \quad \text{in } R^{d-k},$$

where as usual $N(0,\Gamma)$ stands for normal distribution with mean zero and covariance matrix Γ.

PROPOSITION 5. *Let Γ be some symmetric nonnegative definite $(d-k) \times (d-k)$ matrix and let ν be a Lévy measure on R^k, defined in the same way as in Proposition* 1 [$\nu \circ \tau = \tilde{\nu}, \tilde{\nu}\{x \in R^k\colon \|x\| > r,\ x/\|x\| \in B\} = r^{-1} H(B)$, $B \in \mathscr{B}(S^k)$]. *Then the statements* (A.4)–(A.6) *and* (A.10) *are equivalent and each of them is equivalent to the following condition: for every $A \in \mathscr{B}(R^k \setminus \{0\})$ such that $\nu(\partial A) = 0$, $\nu(A) < \infty$, we have*

$$\lim n P\{X_1^{n,1} \in A\} = \nu(A)$$

and for $\forall\ z \in R^{d-k}$ and for $\forall\ \varepsilon > 0$,

$$\lim n \Big\{ E\big|(z, X_1^{(n,2)})\big|^2 \mathbb{1}\{\|X_1^{(n,2)}\| \le \varepsilon\} - \big(E(z, X_1^{(n,2)}) \mathbb{1}\{\|X_1^{(n,2)}\| \le \varepsilon\}\big)^2 \Big\} = (\Gamma z, z).$$

PROOF OF PROPOSITIONS 4 AND 5. The case $d = 2$ [but without the statement (A.6)] was considered in Resnick and Greenwood (1979). [Here it is necessary to mention that in order to apply the UT condition (see Theorem A in Section 2) (A.5) is not sufficient and we need (A.6).] Certainly, Propositions 4 and 5 should be credited to Skorokhod (1957); in fact, in Skorokhod (1957), where an essentially one-dimensional case was considered, it is written: "...we wish to point out that all results of Sections 1–3 can be carried over to the case of a finite-dimensional Banach space." Thus the following several lines can be considered as the explanation of how this can be done.

Since the proof of equivalency of all statements except (A.6) in the general case $d > 2$ causes only notational difficulties, we need only to show that one of the statements, say (A.4), implies (A.6). Convergence of finite-dimensional distributions of Z_n to the corresponding finite-dimensional distributions of Y follows in a standard way; thus it remains to prove the tightness of the sequence of distributions of Z_n. In contrast to the proof of implication (A.4)$\to$(A.5), when tightness of $\{Z_n,\ n \geq 1\}$ in D^d is simply implied by the tightness of coordinates $\{Z_{ni},\ n \geq 1\}$ $i = 1, \dots, d$, generally the tightness of coordinates does not imply the tightness in D_d. The straightforward approach in proving the tightness of Z_n in D_d would be as follows. Since the modulus of continuity in the space D_d is defined in the same way as in D except that the absolute value sign is changed by the norm, one can repeat all steps (with necessary changes) in obtaining the bound for this modulus of continuity for Z_n. We can propose the following approach, which seems to be a little bit simpler (at least, notationally). We can use the following fact [see Problem 22 on page 153 of Ethier and Kurtz (1986)]: $\{(X_n^1, X_n^2, \dots, X_n^d),\ n \geq 1\}$ is relatively compact in $D([0,\infty], R^d)$ if and only if $\{X_n^k,\ n \geq 1\}$ and $\{X_n^k + X_n^l,\ n \geq 1\}$ are relatively compact in $D[0,\infty)$ for all $k, l = 1, \dots, d$. Since we have the tightness of coordinates of the process Z_n, we need to establish the tightness of processes $\{Z_{nk} + Z_{nl},\ n \geq 1\}$ for all possible combinations of $k, l = 1, \dots, d$. (Here it is appropriate to note that addition is not a continuous operation in D spaces; therefore we cannot apply the continuous mapping theorem.) Let us fix k and l, $1 \leq k \neq l \leq d$, and let us denote

$$\zeta_{ni} = a_{nk}^{-1} X_{ik} + a_{nl}^{-1} X_{il}, \qquad V_n = \sum_{i=1}^{n} \zeta_{ni},$$

$$U_n(t) = V_{[nt]}, \qquad 0 \leq t \leq 1.$$

Thus we have a triangular array $\{\zeta_{ni},\ i = 1, \dots, n\}$, $n \geq 1$, and for each n, random variables ζ_{ni}, $i = 1, \dots, n$ are independent and identically distributed. From (A.7), taking a special set $A_x = \{z \in R^d\colon |z_k + z_l| > x\}$ we get a measure ν_{kl} on line, defined by the relation

$$\nu_{kl}(y\colon |y| > x) = \int_{A_x} \nu(dz).$$

For this measure we have

$$\lim nP\{|\zeta_{n1}| > x\} = \nu_{kl}\{y\colon\ |y| > x\}. \tag{A.11}$$

Having (A.11), we can easily check that the conditions of Theorem 9.6.1 from Gikhman and Skorokhod (1969) are satisfied; thus we get the tightness of the sequence $\{U_n,\ n \geq 1\}$, which is exactly the sequence $\{Z_{nk} + Z_{nl},\ n \geq 1\}$. Applying the above-mentioned fact, we get the tightness of Z_n, $n \geq 1$ in $D([0, 1], R^d)$. Then the relation (A.6) follows and we have completed the proof of Propositions 4 and 5. $\square$

Acknowledgments. The paper was written while the first-named author was visiting the Department of Statistics and Applied Probability at UCSB; the hospitality of the faculty is appreciated. Also we thank H. Ahn and V. Mackevičius for very useful conversations on stochastic integration theory.

REFERENCES

AKGIRAY, V. and BOOTH, G. G. (1988). The stable law model of stock returns. *J. Bus. Econ. Statist.* **6** 51–57.

AKGIRAY, V., BOOTH, G. G. and SEIFERT, B. (1988). Distribution properties of Latin American black market exchange rates. *Journal of International Money and Finance* **7** 37–48.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

CANER, M. (1995). Tests for cointegration with infinite variance errors. Preprint.

CHAN, N. H. and TRAN, L. T. (1989). On the first order autoregressive process with infinite variance. *Econometric Theory* **5** 354–362.

CHAN, N. H. and WEI, C. Z. (1988). Limiting distribution of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16** 367–401.

DAVIS, R. and RESNICK, S. (1985). Limit theorems for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13** 179–195.

ELLIOTT, R. J. (1982). *Stochastic Calculus and Applications*. Springer, New York.

ENGLE, R. F. and GRANGER, C. W. J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica* **55** 251–276.

ENGLE, R. F. and YOO, S. B. (1987). Forecasting and testing in cointegrated systems. *J. Econometrics* **35** 143–159.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes. Characterization and Convergence*. Wiley, New York.

FAMA, E. (1965). The behavior of stock market prices. *J. Business* **38** 34–105.

GIKHMAN, I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes*. W. B. Saunders, Philadelphia.

GRANGER, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *J. Econometrics* **16** 121–130.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.

JAKUBOWSKI, A., MEMIN, J. and PAGES, G. (1989). Convergence en loi des suites d'integrales stochastiques sur l'espace $\mathbb{D}^1$ de Skorokhod. *Probab. Theory Related Fields* **81** 111–137.

JEGANATHAN, P. (1991). On the asymptotic behavior of least-squares estimators in AR time series with roots near the unit circle. *Econometric Theory* **7** 269–306.

JOHANSEN, S. (1988). Statistical analysis of cointegration vectors. *J. Econom. Dynam. Control* **12** 231–254.

JOHANSEN, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59** 1551–1580.

KOEDIJK, K. and KOOL, C. M. J. (1992). Tail estimates of East European exchange rates. *J. Bus. Econom. Statist.* **10** 83–96.

KOPP, P. E. (1984). *Martingales and Stochastic Integrals*. Cambridge Univ. Press.

KURTZ, T. G. and PROTTER, P. (1991a). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** 1035–1070.

KURTZ, T. G. and PROTTER, P. (1991b). Characterization of the weak convergence of stochastic integrals. *Stochastic Analysis. London Math. Soc. Lecture Note Ser.* **167** 255–259. Cambridge Univ. Press.

KURTZ, T. G. and PROTTER, P. (1996a). Weak convergence of stochastic integrals and differential equations. *Probabilistic Models for Nonlinear Partial Differential Equations*. *Lecture Notes in Math.* **1627** 1–41. Springer, Berlin.

KURTZ, T. G. and PROTTER, P. (1996b). Weak convergence of stochastic integrals and differential equations II: infinite dimensional case. *Probabilistic Models for Nonlinear Partial Differential Equations*. *Lecture Notes in Math.* **1627** 197–285. Springer, Berlin.

MANDELBROT, B. B. (1963a). New methods in statistical economics. *J. Political Economy* **71** 421–440.

MANDELBROT, B. B. (1963b). The variation of certain speculative prices. *J. Business* **26** 394–419.

MEMIN, J. and SLOMINSKI, L. (1991). Condition UT et stabilité en loi des solutions d'equations differentielles stochastiques. *Séminaire de Probabilités XXV. Lecture Notes in Math.* **1485** 162–177. Springer, Berlin.

MITTNIK, S. and RACHEV, S. T. (1993). Modeling asset returns with alternative stable distributions. *Econometric Rev.* **12** 261–330.

NELSON, C. R. and PLOSSER, C. I. (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *J. Monetary Economics* **10** 129.

PARK, J. Y. and PHILLIPS, P. C. B. (1988). Statistical inference in regressions with integrated processes I. *Econometric Theory* **4** 468–497.

PHILLIPS, P. C. B. (1990). Time series regression with a unit root and infinite-variance errors. *Econometric Theory* **6** 44–62.

PHILLIPS, P. C. B. and DURLAUF, S. N. (1986). Multiple time series regression with integrated processes. *Rev. Econom. Stud.* **53** 473–495.

PROTTER, P. (1990). *Stochastic Integration and Differential Equations*. Springer, Berlin.

RACHEV, S. T., KIM, J-R. and MITTNIK, S. (1997). Econometric modeling in the presence of heavy-tailed innovations: a survey of some recent advances. *Comm. Statist. Stochastic Models* **13** 841–866.

RESNICK, S. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

RESNICK, S. and GREENWOOD, P. (1979). Bivariate stable characterization and domains of attraction. *J. Multivariate Anal.* **9** 206–221.

SAMORODNITSKY, G. and TAQQU, M. S. (1993). *Stable Non-Gaussian Random Processes: Models with Infinite Variance*. Chapman and Hall, London.

SHARPE, M. (1969). Operator-stable probability distributions on vector groups. *Trans. Amer. Math. Soc.* **136** 51–65.

SKOROKHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theory Probab. Appl.* **2** 138–171.

STOCK, J. H. and WATSON, W. W. (1988). Testing for common trends. *J. Amer. Statist. Assoc.* **83** 1097–1107.

STRICKER, C. (1985). Lois de semimartingales et criteres de compacite. *Séminaire de Probabilités XIX. Lecture Notes in Math.* **1123** 209–217. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
VILNIUS UNIVERSITY
NAUGARDUKO 24
VILNIUS 2006
LITHUANIA
AND
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106-3110
E-MAIL: vpaul@ieva.maf.vu.lt
vpaul@pstat.ucsb.edu

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106-3110
E-MAIL: rachev@pstat.ucsb.edu