On $\alpha$-covariance, long, short and negative memories for sequences of random variables with infinite variance

Vygartas Paulauskas

July 17, 2013

Vilnius University

Abstract

We consider a measure of dependence for symmetric $\alpha$-stable random vectors, which was introduced by the author in 1976. We demonstrate that this measure of dependence can be extended for much more broad class of random vectors (up to regularly varying vectors in separable Banach spaces). This measure is applied for linear random processes and fields with heavy-tailed innovations, for some stable processes, and these applications show that this dependence measure, named as $\alpha$-covariance is a good substitute for the usual covariance.

Also we discuss a problem of defining long, short, and negative memories for stationary processes and fields with infinite variances.
1 Introduction

The importance of the notion of independence in probability theory is well-known, sometimes it is even stressed that this is the main feature which distinguishes the probability from the general measure theory. Therefore, the notion of dependence in probability theory as being in some sense opposite to independence, is also important, moreover, it is much more complicated for the following reason. Let \( X \) and \( Y \) be two random variables (or a random vector \((X, Y)\)) defined on some probability space. They are independent if their joint distribution is a product of marginal distributions. But if they are not independent, we would like to know what kind of dependence between \( X \) and \( Y \) we are facing with, how strong this dependence is. This means that in the case of dependence we want to measure this dependence, and we would like to have the property that for independent random variables this measure would be zero, while for the "strongest" dependence it would be one. Here "measure" is used not in the measure-set theoretical meaning, here it stands for some function (even not necessarily non-negative, as in the case of a correlation coefficient, which can be both positive and negative) on the set of all bivariate distributions. Again, what is the "strongest" dependence it is not quite clear, one possible candidate for such dependence can be the case where one random variable is a function of another, i.e., \( Y = f(X) \) where \( f \) is one-to-one function from a support of \( X \) to a support of \( Y \). In probability theory and mathematical statistics a lot of measures or concepts of dependence are introduced, among them classical Pearson correlation coefficient, Kendall’s \( \tau \), Spearman’s \( \rho \), more recent functional or physical dependence measure, and many others. We refer for recent survey papers [3] and [36] with big lists of references for measures of dependence.

In the case of random variables having finite second moments one of the most popular measures of dependence is correlation coefficient, defined by the following formula

\[
\text{Corr}(X, Y) = \frac{E(X - EX)(Y - EY)}{\sigma(X)\sigma(Y)},
\]

where \( \sigma^2(X) = E(X - EX)^2 \). For independent random variables correlation coefficient is zero, always \(-1 \leq \text{Corr}(X, Y) \leq 1 \) and if \( Y = aX + b \) for some real numbers \( a, b \), then \(|\text{Corr}(X, Y)| = 1 \). Unfortunately, equality \( \text{Corr}(X, Y) = 0 \) holds not only for independent \( X \) and \( Y \), it can be even for random vector \((X, Y)\) concentrated on some curve, the case, which we would like to attribute as the "strongest" dependence. Thus, it is possible to say that correlation coefficient measures linear dependence and random variables satisfying \( \text{Corr}(X, Y) = 0 \) are called uncorrelated. This notion is very important in the so-called \( L_2 \)-theory of random variables (uncorrelated means orthogonal in this theory). These properties and the simplicity of the notion explain why measures of dependence based on correlation and covariation are so popular and are used in many areas of probability and statistics, in particular, in time series analysis. One of the main ways to define memory properties (long, short, and negative memories)
for a covariance stationary mean zero process \( Y(t) \) is to use the decay and some other properties of the covariance function, see the precise Definition 8. In its turn, these properties are important when considering functional limit theorems for these processes. One class of stationary processes, for which these memory properties are studied most deeply are linear processes. Let \( \varepsilon_i, i \in \mathbb{Z} \), be independent, identically distributed (i.i.d.) random variables with finite second moment (without loss of generality we may assume \( E\varepsilon_1 = 0, E\varepsilon_1^2 = 1 \)) and let \( a_k, k \geq 0 \) be a sequence of real numbers satisfying \( \sum_{k=0}^{\infty} a_k^2 < \infty \) (this sequence sometimes is called a filter of a linear process, while random variables \( \varepsilon_i, i \in \mathbb{Z}, \) are called innovations). Then a linear process

\[
X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}, \quad t \in \mathbb{Z},
\]

is a stationary sequence, and the dependence is reflected in the covariance function

\[
\gamma(k) = E X_0 X_k, \quad k \in \mathbb{Z}.
\]

Since there is a simple expression of \( \gamma(k) \) via coefficients of the filter \( (\gamma(k) = \sum_{j=0}^{\infty} a_j a_{j+k}) \), properties of the filter define memory properties of the linear process under consideration. In Section 4 we discuss this question in detail and argue that memory properties can not be defined only by the decay of covariance function.

The situation is quite different if random variables \( X \) and \( Y \) have infinite variance, one can say that in this case there is no good substitute for correlation coefficient. Of course, in statistics there are above mentioned Kendall’s \( \tau \) or Spearman’s \( \rho \), which are based on rank statistics and, therefore, do not require any moments of random variables under consideration. But such statistics are not convenient for investigation dependence in more theoretical problems, for example, they are not convenient to measure dependence in linear processes with innovations without second moment. On the other hand, during last decades the role of the so-called heavy tailed distributions had increased both in theoretical and applied probability, therefore the problem of measuring the dependence between random variables having infinite variance remains an important (and difficult) problem.

The main aim of this paper is to revive interest to one measure of dependence which was introduced by the author more than 30 years ago in [26] for a specific, but rather important, class of symmetric \( \alpha \)-stable (\( S\alpha S \)) distributions. The importance of this class of distributions can be explained by the fact that trying to build a model involving random variables with infinite variance, as a first step, one takes stable random variables, or random variables in the domain of their attraction. Although at the beginning we shall deal mainly with bivariate random vectors, we shall define general \( d \)-dimensional \( S\alpha S \) random vectors. Let \( S_d = \{ x \in \mathbb{R}^d : ||x|| = 1 \} \) be the unit sphere in \( \mathbb{R}^d \), here \( ||x|| \) stands for the Euclidean norm in \( \mathbb{R}^d \). Random vector \( X = (X_1, \ldots X_d) \) is \( S\alpha S \) with parameter \( 0 < \alpha < 2 \) if there exists a unique symmetric finite measure \( \Gamma \) on \( S_d \) such that
the characteristic function (ch.f.) of $X$ is given by formula

$$E \exp \{i(t, X)\} = \exp \left\{ - \int_{S^d} |(t, s)|^\alpha \Gamma(ds) \right\}. \tag{1}$$

$\Gamma$ is called the spectral measure of the $S\alpha S$ random vector $X$. The Gaussian case $\alpha = 2$ is excluded from this definition, since in the Gaussian case there is no uniqueness of the spectral measure $\Gamma$: many different measures $\Gamma$ will give the same ch.f. About forty years ago (for historical details we refer to monograph [33]) two measures of dependence between coordinates of a bivariate $S\alpha S$ random vector $X = (X_1, X_2)$ with spectral measure $\Gamma$ were introduced. The first one, called covariation of $X_1$ on $X_2$ and denoted by $[X_1, X_2]_\alpha$, is defined for $\alpha > 1$ as follows:

$$[X_1, X_2]_\alpha = \int_{S^2} s_1 s_2^{(\alpha-1)} \Gamma(ds),$$

where $a^{(p)} = |a|^p \text{sign } a$. Although in the case $\alpha = 2$ this quantity is equal to the half of covariance between $X_1$ and $X_2$, which is symmetric, it is not symmetric in its arguments, i.e., in general (for $1 < \alpha < 2$)

$$[X_1, X_2]_\alpha \neq [X_2, X_1]_\alpha$$

All properties of covariation, including equivalent definition, are given in Chapter 2.7 of [33]. Main shortcoming of this measure of dependence, apart from the just mentioned non-symmetricity, is that it is not defined for $\alpha < 1$ (for $\alpha = 1$ it is possible to define the covariation, see Exercise 2.22 in [33]). Here it is appropriate to mention that recently in [8] symmetric covariation coefficient was introduced.

Another measure of dependence for $S\alpha S$ random vectors is the codifference, defined by formula

$$\tau(X_1, X_2) = \int_{S^2} (|s_1|^\alpha + |s_2|^\alpha - |s_1 - s_2|^\alpha) \Gamma(ds).$$

One can note that the codifference can be defined for a general bivariate random vector $Y = (Y_1, Y_2)$ by means of the following formula

$$\tau(Y_1, Y_2) = \ln f_Y(1, -1) - \ln f_Y(1, 0) - \ln f_Y(0, -1), \tag{2}$$

where

$$f_Y(t, s) = E \exp \{i(t Y_1 + s Y_2)\}.$$
ch.f. of a vector $\mathbf{Y}$ with independent coordinates is a product of marginal ch.f. of components.

Earlier than the codifference and about the same time as covariation were introduced, the author in [26] had proposed one more measure of dependence for $S\alpha S$ random vectors. Let $X = (X_1, X_2)$ be a $S\alpha S$ random vector with the spectral measure $\Gamma$, and let $\mathbf{Y} = (Y_1, Y_2)$ be a random vector on $S_2$ with the distribution $\tilde{\Gamma}(A) = (\Gamma(S_2))^{-1}\Gamma(A)$. Then the generalized association parameter (g.a.p.) of the random vector $X$ is defined as usual correlation coefficient for the random vector $\mathbf{Y}$:

$$
\tilde{\rho} = \tilde{\rho}(X_1, X_2) = \frac{EY_1Y_2}{\sqrt{EY_1^2EY_2^2}} = \frac{\int_{S_2} s_1s_2\tilde{\Gamma}(ds)}{\left(\int_{S_2} s_1^2\tilde{\Gamma}(ds)\int_{S_2} s_2^2\tilde{\Gamma}(ds)\right)^{\frac{1}{2}}},
$$

Also we shall use the following analog of the covariance between $X_1$ and $X_2$

$$
\rho = \rho(X_1, X_2) = \int_{S_2} s_1s_2\Gamma(ds),
$$

and we shall call it as $\alpha$-covariance of a $S\alpha S$ random vector $(X_1, X_2)$. Strictly speaking we should use the normalized measure $\tilde{\Gamma}$ instead of $\Gamma$ in the definition of $\alpha$-covariance (then it would be possible to say that $\rho(X_1, X_2)$ is simply covariance between random variables $Y_1$ and $Y_2$), but since $\Gamma$ is finite measure, $\rho(X_1, X_2) = \Gamma(S_2)EY_1Y_2$.

Here some remarks about the terminology is appropriate. The term ”generalized association parameter”, clearly, is very unsuccessful, it was introduced before the notions and terms ”covariation of $X_1$ on $X_2$” and ”codifference” were invented, at the time when the term ”association” was popular (see, for example papers [6], [7]). Now we suggest to call this parameter as $\alpha$-correlation coefficient for a $S\alpha S$ random vector, and from now, in what follows we shall write $\alpha$-correlation coefficient ($\alpha$-cc) instead of g.a.p. Thus we have three notions (or quantities)- covariation, codifference, and $\alpha$-covariance, which all become usual covariance in the case $\alpha = 2$.

The following proposition was proved in [26]

**Proposition 1** [26] The introduced $\alpha$-cc of a random vector $\mathbf{X} = (X_1, X_2)$ with ch.f. (1) (with $d = 2$) has the following properties:

(i) $|\tilde{\rho}| \leq 1$, and if the coordinates of $\mathbf{X}$ are independent then $\tilde{\rho} = 0$;

(ii) if $|\tilde{\rho}| = 1$, then the distribution of $\mathbf{X}$ is concentrated on a line, i.e., coordinates $X_1$ and $X_2$ are linearly dependent;

(iii) if $\alpha = 2$, $\tilde{\rho}$ coincides with a correlation coefficient of a Gaussian random vector with characteristic function (1);
(iv) $\hat{\rho}$ is independent of $\alpha$ and depends only on the spectral measure $\Gamma$.

Also in [26] it was shown that if a random vector $X$ is sub-Gaussian with ch.f.

$$\exp \left\{ -\left( \sigma_1^2 t_1^2 + 2r\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2 \right)^{\alpha/2} \right\}, \quad (5)$$

where $\sigma_1^2$, $\sigma_2^2$, and $r$ are variances and correlation coefficient, respectively, of underlying Gaussian vector, then the $\alpha$-cc $\hat{\rho} = r$.

This notion is easily generalized to $d$-dimensional $S\alpha S$ random vectors (see Proposition 3 in [26]) by defining the $\alpha$-correlation matrix $\hat{\Lambda}$ and $\alpha$-covariance matrix $\Lambda$ as usual correlation and covariance matrices, respectively, of a random vector on $S_d$ with a distribution $\hat{\Gamma}(A) = (\Gamma(S_d))^{-1} \Gamma(A)$.

Despite of the simplicity of definition and the fact that $\alpha$-cc of $S\alpha S$ random vectors satisfies main requirements for measures of dependence, it was almost not used. Only recently the interest to this measure of dependence was revived in [9] and [8], where the so-called symmetric covariation coefficient was introduced, it was compared with $\alpha$-cc, and estimations of $\alpha$-cc and this new symmetric covariation coefficient were proposed. The main goal of the present paper is to demonstrate that in the case of random vectors without variance these simple notions of $\alpha$-covariance and $\alpha$-cc are quite natural substitutes for covariance and correlation. We shall show that this measure of dependence can be extended to more wide class of heavy-tailed random vectors, is very suitable when considering stable processes, in particular, linear random processes and fields with heavy-tailed innovations.

Also we discuss the memory properties for stationary sequences without finite variance. We propose a unified approach to define memory properties both for processes and for fields, based on the rate of the growth of partial sums, formed by these processes or fields. The exponent $1/\alpha$ (which is mentioned as boundary between short and long memory in several papers) characterizes the growth of partial sums of sequences with no memory at all, that is, i.i.d. random variables, and this value is attributed to short memory and serves as a boundary between negative memory (when the exponent of the growth is smaller than $1/\alpha$) and long memory (exponent is bigger than $1/\alpha$).

By simple examples of linear processes and fields we show that this approach is more natural even in the case of finite variance. We want to stress that it is important to separate notions of memory and dependence, that is, to separate long and short-range dependence from memory properties, therefore it would be logical to call these properties negative, zero and positive memories, and even we suggest to introduce strongly negative memory (the case where the volatility of partial sums stays bounded and do not grow with number of summands in partial sums).

At the same time we agree with the attitude propagated by G. Samorodnitsky in his several papers (see, for example, [31] and [32]), that memory phenomenon is a complicated one and, most probably, there is no way to give one definition of memory which would be good for all cases. In different context the definition of memory can be different. For example, considering limit theorems for partial sums the notions of positive (long), zero (short), and negative
memories are very natural, while considering maximum operation instead of summation or problem of large deviations, classification of memory properties can be different - it seems that there is impossible to introduce negative memory considering partial maxima operation, see [31], where the same value $1/\alpha$ serves as boundary in the growth of partial maxima of stationary $S\alpha S$ random variables.

The rest of the paper is organized as follows. In the second section we consider $\alpha$-covariance for linear processes and fields with infinite variance. The third section is devoted to $\alpha$-covariance for random vectors defined as stochastic integrals and for stable processes. In the fourth section we consider memory properties of stationary random processes and fields. The last, fifth, section is, may be, the most important, since we show that the notion of $\alpha$-covariance, which was introduced for $S\alpha S$ random vectors can be extended and generalized for much more broad class of random vectors.

2 Linear processes and fields

2.1 Linear processes

We start with the case of linear processes. Let $\varepsilon_i, i \in \mathbb{Z}$ i.i.d. standard $S\alpha S$ random variables with ch.f. $\exp(-|t|^\alpha), 0 < \alpha \leq 2$ (in the case of Gaussian distribution variance will be not 1, but 2). We consider linear random process

$$X_k = \sum_{j=0}^{\infty} c_j \varepsilon_{k-j}, \quad k \in \mathbb{Z},$$

where $c_j, j \geq 0$ are real numbers satisfying condition

$$A := \sum_{j=0}^{\infty} |c_j|^\alpha < \infty.$$  

(7)

We get a stationary sequence of $S\alpha S$ random variables $X_k, \quad k \in \mathbb{Z}$ with ch.f. $\exp(-A|t|^\alpha)$, and the main question is how to measure the dependence between $X_0$ and $X_n$. Since bivariate random vector $(X_0, X_n)$ is jointly $S\alpha S$, we can apply as a measure of dependence $\alpha$-cc and $\alpha$-covariance. Let us denote $\tilde{\rho}_n = \tilde{\rho}(X_0, X_n)$ and $\rho_n = \rho(X_0, X_n)$. To formulate our result we need some more notations. Let

$$a_{n,j} = (c_j, c_{n+j}), \quad \|a_{n,j}\|^2 = (c_j^2 + c_{n+j}^2), \quad \tilde{a}_{n,j} = (c_j, c_{n+j})||a_{n,j}||^{-1}$$

$$A_{1,n} = \sum_{j=0}^{\infty} \frac{c_j^2}{||a_{n,j}||^{2-\alpha}}, \quad A_{2,n} = \sum_{j=0}^{\infty} \frac{c_{j+n}^2}{||a_{n,j}||^{2-\alpha}}, \quad A_n = \sum_{j=0}^{n} |c_j|^\alpha.$$ 

The convergence of the two above written series easily follows from (7), for example,

$$A_{1,n} = \sum_{j=0}^{\infty} \frac{c_j^2}{(c_j^2 + c_{n+j}^2)(2-\alpha)/2} \leq \sum_{j=0}^{\infty} c_j^2 |c_j|^\alpha - 2 = A.$$ 

7
Theorem 2. For a linear process $X_k$ from (6), satisfying (7), we have

$$
\hat{\rho}_n = \frac{\sum_{j=0}^{\infty} c_j c_{j+n} \|a_{n,j}\|^{\alpha - 2}}{\sqrt{A_{1,n}(A_{2,n} + A_{n-1})}},
$$

(8)

and

$$
\rho_n = \sum_{j=0}^{\infty} c_j c_{j+n} \|a_{n,j}\|^{\alpha - 2}.
$$

(9)

Proof of Theorem 2. We must find the spectral measure of the $S_{\alpha}S$ random vector $(X_0, X_n) = (\sum_{k=0}^{\infty} c_k \varepsilon_{-k}, \sum_{k=0}^{\infty} c_k \varepsilon_{n-k})$. Denoting $t = (t_1, t_2)$ and using the notations introduced before the formulation of Theorem 2, we can write

$$
E \exp \{i(t_1 X_0 + t_2 X_n)\} = E \exp \left\{ i t_1 \sum_{k=0}^{\infty} c_k \varepsilon_{-k} + i t_2 \sum_{k=0}^{\infty} c_k \varepsilon_{n-k} \right\}
$$

$$
= E \exp \left\{ \sum_{k=0}^{\infty} i(t_1 c_k + t_2 c_{k+n})\varepsilon_{-k} + \sum_{k=0}^{n-1} i t_2 c_k \varepsilon_{n-k} \right\}
$$

$$
= \exp \left\{ - \sum_{k=0}^{\infty} \left( |t_1| |\tilde{a}_{k,n}|^{\alpha} \|a_{n,k}\|^{\alpha} + \sum_{k=0}^{n-1} |c_k|^{\alpha} |t_2|^{\alpha} \right) \right\}
$$

From this expression we see that the bivariate $S_{\alpha}S$ random vector $(X_0, X_n)$ has the symmetric spectral measure $\Gamma_n$ concentrated at points $(0, \pm 1), \pm \tilde{a}_{k,n}, k \geq 0$, namely,

$$
\Gamma_n(0, \pm 1) = \frac{1}{2} A_{n-1}, \quad \Gamma_n(\pm \tilde{a}_{k,n}) = \frac{1}{2} \|a_{n,k}\|^{\alpha}.
$$

(10)

Due to (7) this measure is finite:

$$
\Gamma(S_2) = A_{n-1} + \sum_{k=0}^{\infty} \|a_{n,k}\|^{\alpha} \leq 2A.
$$

Having (10) we obtain (8) and (9) by simple calculations using definitions (3) and (4). The theorem is proved. \[\square\]

For two sequences $\{a_n\}$ and $\{b_n\}$ we shall write $a_n \sim b_n$, if $\lim a_n b_n^{-1} = 1$, and $a_n \simeq b_n$, if there exist two constants $0 < K_1 < K_2 < \infty$ such, that for sufficiently large $n$, $K_1 \leq a_n b_n^{-1} \leq K_2$. We can formulate two simple properties of the introduced measures of dependence of a linear process.

Proposition 3. For any $c \neq 0$ we have

$$
\tilde{\rho}(cX_0, cX_n) = \tilde{\rho}(X_0, X_n), \quad \rho(cX_0, cX_n) = |c|^\alpha \rho(X_0, X_n).
$$

(11)
If \( \tilde{c}_j \sim c_j \), then
\[
\rho(X_0, X_n) \simeq \rho(\tilde{X}_0, \tilde{X}_n),
\]
where \( \tilde{X}_n \) is defined by (6), only with coefficients \{\tilde{c}_j\} instead of \{c_j\}. If, additionally to (7), the following mild condition
\[
|c_{j+n}| \leq k|c_j|, \text{ for all } j, n, \text{ and for some } k > 0,
\]
is satisfied, then
\[
\tilde{\rho}_n \simeq \rho_n.
\]

Due to the last property, as in the case of finite variance, we shall deal mainly with \( \alpha \)-covariance, although there is a small difference between these two cases with finite and infinite variances: for a stationary sequence with finite variance, correlation and covariance for all lags differs by a constant (equal to the variance of the marginal distribution), while in the case of a stationary sequence (6) we have only (14).

Proof of Proposition 3. The equalities (11) are obvious, since it is easy to see that if \( \Gamma_n \) and \( \Gamma_{n,c} \) are the spectral measures of random vectors \((X_0, X_n)\) and \((cX_0, cX_n)\), respectively, then \( \Gamma_{n,c}(ds) = |c|^\alpha \Gamma_n(ds) \).

Let us denote \( \tilde{a}_{n,j} = (\tilde{c}_j, \tilde{c}_{n+j}) \). Having relation \( \tilde{c}_j = c_j(1 + \delta(j)) \) with \( \delta(j) \to 0, \text{ for } j \to \infty \) and \( |\delta(j)| \leq a \) for all \( j \) and for a sufficiently small \( 0 < a < 1 \), we can get
\[
||\tilde{a}_{n,j}||^2 = ||a_{n,j}||^2(1 + \delta_1(j, n)),
\]
where \( \delta_1(j, n) \to 0, \text{ for } j \to \infty, \) uniformly with respect to \( n \), and \( |\delta_1(j, n)| \leq a_1 \). Here \( a_1 \) can be expressed by \( a \) and will be small for small \( a \). Now we can write
\[
\rho(\tilde{X}_0, \tilde{X}_n) = \sum_{j=0}^{\infty} \tilde{c}_j \tilde{c}_{j+n} ||\tilde{a}_{n,j}||^{\alpha-2} = \sum_{j=0}^{\infty} c_j c_{j+n} ||a_{n,j}||^{\alpha-2}(1 + \delta_2(j, n)),
\]
where \( \delta_2 \) is obtained from the formal relation
\[
1 + \delta_2(j, n) = \frac{(1 + \delta(j))(1 + \delta(j + n))}{(1 + \delta_1(j, n))(2-\alpha)/2}.
\]
Again, it can be shown that \( \delta_2 \) has the same properties as \( \delta_1 \) and can be bounded \( |\delta_2(j, n)| \leq a_2 < 1 \) if \( a \) is chosen sufficiently small. The relation (12) follows from (15).

To prove (14) we need to show that the quantity \( A_{1,n}(A_{2,n} + A_{n-1}) \) is bounded from above and below by some constants. The bound from above
\[
A_{1,n}(A_{2,n} + A_{n-1}) \leq A^2
\]
is easy, while for the lower bound for \( A_{1,n} \) we use (13):
\[
A_{1,n} \geq \sum_{j=0}^{\infty} \frac{c_j^2}{(c_j^2(1 + k^2))(2-\alpha)/2} = (1 + k^2)^{(\alpha-2)/2} A.
\]
For any $\epsilon > 0$, we can find $N$ such that $\sum_{j=n}^{\infty} |c_j|^\alpha < \epsilon$ for all $n > N$, therefore

$$A_{2,n} + A_{n-1} \geq A_{n-1} = A - \sum_{j=n}^{\infty} |c_j|^\alpha \geq (1 - \epsilon)A.$$ 

From these estimates the relation (14) follows, and the proposition is proved. □

In special cases of the filter of a linear process (6) we have the following corollaries.

**Corollary 4** Let $c_j \sim 2^{-j}$ then $\rho_n \sim C(\alpha)2^{-n}$.

If $c_j \sim j^{-\beta}$, $\beta > 1/\alpha$, then: in the case $0 < \alpha \leq 1$

$$\rho_n \sim C(\alpha, \beta)n^{1-\beta\alpha},$$

in the case $1 < \alpha \leq 2$

$$\rho_n \sim \begin{cases} C(\alpha, \beta)n^{1-\beta\alpha}, & \text{if } 1/\alpha < \beta < 1/(\alpha - 1), \\ C(\alpha)n^{-\beta}(1 + I(\beta = 1/(\alpha - 1))\ln n), & \text{if } \beta \geq 1/(\alpha - 1), \end{cases}$$

where, as usual, $I(A)$ stands for the indicator function of an event $A$.

**Proof of Corollary 4.** Taking $c_j = 2^{-j}$ we simply have

$$\rho_n = \sum_{j=0}^{\infty} 2^{-j}2^{-j-n}(2^{-j}(1 + 2^{-2n})^{1/2})^{\alpha-2} = 2^{-n}\frac{(1 + 2^{-2n})^{(\alpha-2)/2}}{1 - 2^{-\alpha}}.$$ 

Now let us take $c_j = j^{-\beta}$, $j \geq 1$, $c_0 = 1$ and in the definition (9) of $\rho_n$ we separate two first terms

$$c_0c_n||a_{n,0}||^{\alpha-2} + c_1c_{n+1}||a_{n,1}||^{\alpha-2} \sim n^{-\beta},$$

then we can write

$$I_n := \sum_{j=2}^{\infty} c_jc_{j+n}||a_{n,j}||^{\alpha-2} \sim \int_1^{\infty} \frac{x^{-\beta}(x+n)^{-\beta}dx}{(x-2\beta + (x+n)-2\beta)^{(2-\alpha)/2}}.$$ 

After change of variables we get

$$I_n = n^{1-\beta\alpha}(I_{n,1} + I_{n,2}),$$

where

$$I_{n,1} = \int_1^{1/n} \frac{y^{-\beta}(1+y)^{-\beta}dy}{(y-2\beta + (1+y)-2\beta)^{(2-\alpha)/2}},$$

and $I_{n,2}$ is the integral of the same function over interval $(1, \infty)$. Therefore, $I_{n,2}$ is independent of $n$, and due to the condition $\beta\alpha > 1$, is a constant depending on $\alpha$ and $\beta$ only:

$$I_{n,2} = \int_1^{\infty} \frac{dy}{y^{\beta\alpha}g(y)} = C(\alpha, \beta).$$
Here \( g(y) \) is a function, bounded by positive constants from below and above for all \( 1 \leq y < \infty \). Now we estimate \( I_{n,1} \). Again, it is easy to see that

\[
I_{n,1} = \int_{1/n}^{1} \frac{dy}{y^{\beta(\alpha-1)}h(y)},
\]

where \( 1 \leq h(y) \leq 2^\beta (1 + 2^{-2\beta})^{(2-\alpha)/2} \) for \( 0 \leq y \leq 1 \). If \( \alpha \leq 1 \) then \( \alpha - 1 \leq 0 \), and we get that \( I_{n,1} \) is a constant, depending on \( \alpha \) and \( \beta \). In the case of \( 1 < \alpha \leq 2 \) and \( 1/\alpha < \beta < 1/(\alpha - 1) \), we have again that \( I_{n,1} \) is a constant, while if \( \beta = 1/(\alpha - 1) \) we get that \( I_{n,1} \) is of the order \( \ln n \). The corollary is proved.

Now we can compare our result for \( \alpha \)-covariance of a linear process (6) with other measures of dependence. There were several papers dealing with measures of dependence of linear processes with innovations with infinite variance, see, for example [16], [17], [18] and references there. In these papers the expressions of the covariation and the codifference for the process (6) were given:

\[
\tau(X_0, X_n) = \sum_{j=0}^{\infty} (|c_j|^\alpha + |c_{j+n}|^\alpha - |c_{j+n} - c_j|^\alpha),
\]

\[
[X_0, X_n]_\alpha = \sum_{j=0}^{\infty} c_{j+n} c_j^{(\alpha-1)}.
\]

The asymptotic of these quantities was investigated for \( FARIMA(p, q, d) \) process (see (2.2), (2.6) and (2.7) formulas in [16]), which is of the form (6) with specific coefficients \( c_j, j \geq 0 \). Namely, in [16] it was shown that these coefficients satisfy the following relation (see Lemma 3.2 and Corollary 3.1 there): if \( \alpha(d-1) < -1 \), then

\[
c_j = C(p, q, d) j^{|d-1|} (1 + O(j^{-1})).
\]  

(18)

Quantity \( d \) here and \( \beta \) used in Corollary 4 are related by equality \( \beta = 1 - d \). In [18] general case of (6) is investigated under conditions which are slightly different from those used in Corollary 4: the coefficients of the filter satisfy more general condition \( c_j = U(j) \) where \( U \) is regularly varying function with the index \( -\beta \), but there are some conditions of the type (18). Therefore in [18] the asymptotic relation \( \sim \) is obtained, while we have weaker relation \( \simeq \), but the order of decay of the quantity \( \tau_n := \tau(X_0, X_n) \) is the same as of the \( \alpha \)-covariance in Corollary 4. We can mention that in both papers [16] and [18] the case \( \beta = 1/(\alpha - 1) \) is excluded from formulation, only mentioning that there is ”phase transition” (see the remark before Theorem 4.1 in [16]). Also it is interesting to note that in the case of exponentially decreasing filter (\( c_i \sim 2^{-i} \)) the codifference is decreasing as \( 2^{-\alpha n} \) for \( 0 < \alpha < 1 \), while from Corollary 4 we have \( \rho_n \simeq C(\alpha) 2^{-n} \) for all \( 0 < \alpha \leq 2 \).
2.2 Linear fields

As it was mentioned in the introduction, \(\alpha\)-covariance, as the measure of dependence, can be easily applied for linear random fields on \(\mathbb{Z}^d\) with \(d \geq 2\) and \(\alpha\)-S random innovations. But since the notation and formulations become more complicated we restrict ourselves to the case \(d = 2\) and formulation of expression of \(\alpha\)-covariance via coefficients of a filter of a random field. Let \(\varepsilon_{i,j}, (i, j) \in \mathbb{Z}^2\) be i.i.d. standard \(\alpha\)-S random variables with c.h.f. \(\exp(-|t|^\alpha), 0 < \alpha \leq 2\). Now we consider linear random field

\[
X_{k,l} = \sum_{i,j=0}^{\infty} c_{i,j} \varepsilon_{k-i,l-j}, \quad (k, l) \in \mathbb{Z}^2,
\]

where \(c_{i,j}, \ i \geq 1, j \geq 0\), are real numbers satisfying condition

\[
A_1 := \sum_{i,j=0}^{\infty} |c_{i,j}|^\alpha < \infty.
\]

We are interested how strongly dependent are random variables \(X_{0,0}\) and \(X_{n,m}\). Let us denote \(\tilde{\rho}_{n,m} = \tilde{\rho}(X_{0,0}, X_{n,m}), \ \rho_{n,m} = \rho(X_{0,0}, X_{n,m})\). We shall consider two cases: \(n > 0, m > 0\) and \(n > 0, m < 0\), since due to the stationarity the remaining two cases can be transformed into the previous, for example, \((X_{0,0}, X_{n,m})\) has the same distribution as \((X_{-n,-m}, X_{0,0})\) and \(\rho_{n,m} = \rho_{|n|,|m|}\) for \(n < 0, m < 0\). Similarly, \(\tilde{\rho}_{n,m} = \rho_{-n,-m}\), for \(n > 0, m < 0\). Comparing with the case of linear processes now we need more complicated notations. First, let us consider the case \(n > 0, m > 0\). Let us denote

\[
a_{i,j}^{(n,m)} = (c_{i,j}, c_{i+n,j+m}), \quad \|a_{i,j}^{(n,m)}\|^2 = c_{i,j}^2 + c_{i+n,j+m}^2,
\]

\[
a_{i,j}^{(n,m)} = (c_{i,j}, c_{i+n,j+m}) \|a_{i,j}^{(n,m)}\|^{-1}
\]

Denote the following four regions of \(\mathbb{Z}_+^2\):

\(I_1 = \{(i, j) : 0 \leq i \leq n-1, \ 0 \leq j \leq m-1, \}\), \(I_2 = \{(i, j) : 0 \leq i \leq n-1, \ j \geq m, \}\), \(I_3 = \{(i, j) : i \geq n, \ 0 \leq j \leq m-1, \}\), \(I_4 = \{(i, j) : i \geq n, \ j \geq m, \}\).

Also denote \(\sum_k = \sum_{(i,j) \in I_k}, \ k = 1, 2, 3, 4\), and \(\sum_0 = \sum_{(i,j) \in \mathbb{Z}_+^2}\). Then we define

\[
A_{1,n,m} := \left(\sum_1 + \sum_2 + \sum_3\right) |c_{i,j}|^\alpha, \quad A_{2,n,m} := \sum_0 c_{i,j}^2 \|a_{i,j}^{(n,m)}\|^\alpha - 2
\]

\[
A_{3,n,m} := \sum_0 c_{i+n,j+m}^2 \|a_{i,j}^{(n,m)}\|^\alpha - 2
\]

In the case \(n > 0, m < 0\) we need the following notation:

\(J_1 = \{(i, j) : 0 \leq i < \infty, \ 0 \leq j \leq |m| - 1, \}\), \(J_2 = \{(i, j) : 0 \leq i < \infty, \ j \geq 0, \}\),
and $\sum^{(k)} = \sum_{(i,j) \in J_k} k = 1, 2$. Then we define

$$B_{n,m}^{(i)} = \sum |ci,j|^\alpha, \quad i = 1, 2,$$

$$B_{n,m}^{(3)} = \sum_0 c_{i,j+m}(c_{i,j+m}^2 + c_{i+n,j}^2)^{(\alpha-2)/2}, \quad B_{n,m}^{(4)} = \sum_0 c_{i,j+m}(c_{i,j+m}^2 + c_{i+n,j}^2)^{(\alpha-2)/2}.$$

Now we are able to formulate our result for linear random field (19).

**Theorem 5** For a linear field $X_{k,l}$ from (19), satisfying (20), for $n > 0, m > 0$, we have

$$\tilde{\rho}_{n,m} = \frac{\sum_0 c_{i,j+m}||a_{i,j}^{(n,m)}||^\alpha}{A_{2,n,m}(A_{1,n,m} + A_{3,n,m})},$$

and

$$\rho_{n,m} = \frac{\sum_0 c_{i,j+m}||a_{i,j}^{(n,m)}||^\alpha}{||a_{i,j}^{(n,m)}||^\alpha}.$$

If $n > 0, m < 0$, then

$$\tilde{\rho}_{n,m} = \frac{\sum_0 c_{i,j+m}||a_{i,j}^{(n,m)}||^\alpha}{\sqrt{(B_{n,m}^{(1)} + B_{n,m}^{(2)} + B_{n,m}^{(3)})}},$$

and

$$\rho_{n,m} = \frac{\sum_0 c_{i,j+m}||a_{i,j}^{(n,m)}||^\alpha}{||a_{i,j}^{(n,m)}||^\alpha}.$$

Clearly, for linear fields we can formulate the same properties of $\tilde{\rho}_{n,m}$ and $\rho_{n,m}$ as in Proposition 3 for linear processes. Also we can easily calculate these quantities for filters with regular behavior. As an example we provide one such result. Suppose that coefficients of the filter are hyperbolically decaying: $c_{i,j} \sim i^{-\beta_1} j^{-\beta_2}$ with $\beta_k > 1/\alpha, \ k = 1, 2$. For the convenience let us denote $\beta = (\beta_1, \beta_2)$.

**Corollary 6** Let $c_{i,j} \sim i^{-\beta_1} j^{-\beta_2}$ with $\beta_k > 1/\alpha, \ k = 1, 2$, and $m > 0, n > 0$. Then, if $0 < \alpha \leq 1$,

$$\rho_{n,m} \simeq C(\alpha, \beta)n^{1-\beta_1/\alpha}m^{1-\beta_2/\alpha}, \quad \beta_i > 1/\alpha;$$

if $1 < \alpha \leq 2$, then

$$\rho_{n,m} \simeq \begin{cases} C(\alpha, \beta)n^{1-\beta_1/\alpha}m^{1-\beta_2/\alpha}, & \text{if } 1/\alpha < \beta_i < (\alpha - 1)^{-1}, i = 1, 2, \\ C(\alpha, \beta)n^{1-\beta_1/\alpha}m^{-\beta_2}, & \text{if } \beta_i > 1/(\alpha - 1), i = 1, 2. \end{cases}$$

If $\beta_1 \geq 1/(\alpha - 1), 1/\alpha < \beta_2 < (\alpha - 1)^{-1}$, then

$$\rho_{n,m} \simeq C(\alpha, \beta)n^{-\beta_1}m^{1-\beta_2/\alpha}(1 + I(\beta_1 = (\alpha - 1)^{-1}) \ln n),$$

and if $\beta_2 \geq 1/(\alpha - 1), 1/\alpha < \beta_1 < (\alpha - 1)^{-1}$, then

$$\rho_{n,m} \simeq C(\alpha, \beta)n^{1-\beta_1/\alpha}m^{-\beta_2}(1 + I(\beta_2 = (\alpha - 1)^{-1}) \ln n).$$

This result is in accordance with the result of Corollary 4. We omit the calculations needed to prove these relations, since they are very similar to those used for linear processes.
3 $\alpha$-covariance for stochastic integrals and stable processes

It is well-known what important role in the theory of stable random vectors and processes play $\alpha$-stable stochastic integrals, that is, integrals of non-random functions with respect to an $\alpha$-stable random measures. The biggest part of the monograph [33] is devoted to these integrals, therefore we do not provide definitions of these notions (but we shall try to keep the same notation as in [33]), refereing a reader to this monograph. Let $(\mathbb{E}, \mathcal{F}, m)$ be a measurable space with a $\sigma$-finite measure $m$, and let $M$ be an $S\alpha S$ random measure, that is, we take the so-called skewness intensity function $\beta(x) \equiv 0$ in general Definition 3.3.1 in [33] of $\alpha$-stable random measure. This is done for the reason that $\alpha$-covariance we defined (till now, see Section 5 for extension of definition of $\alpha$-covariance) only for $S\alpha S$ random vectors. Taking $f \in L^\alpha(\mathbb{E}, \mathcal{F}, m)$, we get a $S\alpha S$ random variable

$$X = \int_{\mathbb{E}} f(x) M(dx),$$

while taking a collection $f_i \in L^\alpha(\mathbb{E}, \mathcal{F}, m)$, $i = 1, \ldots, k$, we get a $S\alpha S$ random vector

$$(X_1, \ldots, X_k), \quad X_i = \int_{\mathbb{E}} f_i(x) M(dx).$$

Taking a family of functions $\{f_t, t \in T\} \subset L^\alpha(\mathbb{E}, \mathcal{F}, m)$ we get a $S\alpha S$ random process

$$X(t) = \int_{\mathbb{E}} f_t(x) M(dx), \quad t \in T.$$  \hspace{1cm} (21)

Many well-known $S\alpha S$ random processes are obtained in this way. Namely, a moving average process is obtained with $\mathbb{E} = \mathbb{R}$, $m$=Lebesgue measure, and $f_t(x) = f(t - x)$:

$$X(t) = \int_{\mathbb{R}} f(t - x) M(dx), \quad t \in \mathbb{R}.$$ \hspace{1cm} (21)

An Ornstein-Uhlenbeck process

$$X(t) = \int_{-\infty}^{t} \exp\{-\lambda(t - x)\} M(dx), \quad t \in \mathbb{R},$$ \hspace{1cm} (22)

is obtained from (21) by taking $f(x) = \exp\{-\lambda x\} \mathbf{1}(x \geq 0)$. In a similar way, i.e., by choosing appropriate family of functions $f_t(x)$, we can get a symmetric linear fractional stable motion, a log-fractional stable motion (see Ch. 3.6 in [33]). It is worth to mention that linear processes and fields, considered above can be obtained in the same way from (21), taking $E = \mathbb{Z}$ (or $E = \mathbb{Z}^2$ in the case of fields), $m$ as a counting measure, and $f(k) = c_k \mathbf{1}(k \geq 0)$. The covariation and the codifference were defined for these more general objects, and they were extensively studied during last three decades. Most of these results are given in [33], see Ch. 4.7 therein, where the codifference function is calculated for many stationary $S\alpha S$ processes. Our goal is to show that $\alpha$-covariance is equally good
measure of dependence, more easily dealt with, and in some cases even better reflect dependence.

Let \((X_1, X_2)\) be a bivariate \(\alpha\)-\(S\) random vector, defined by means of stochastic integrals, i.e.,

\[
(X_1, X_2) \overset{D}{=} \left( \int_E f_1(x)M(dx), \int_E f_2(x)M(dx), \right)
\]

where \(\overset{D}{=}\) stands for equality in distribution. In Ch. 3.2 in [33] it is shown how to express spectral measure \(\Gamma\) of the random vector \((X_1, X_2)\) via control measure \(m\) and functions \(f_i\), also the expressions of the covariation (in the case \(1 < \alpha \leq 2\)) and the codifference are given:

\[
[X_1, X_2]_{\alpha} = \int_E f_1(x)f_2(x)^{\alpha-1}m(dx),
\]

\[
\tau(X_1, X_2) = \int_E (|f_1(x)|^\alpha + |f_2(x)|^\alpha - |f_1(x) - f_2(x)|^\alpha)m(dx).
\]

It is not difficult to write the expressions of \(\alpha\)-covariance and \(\alpha\)-correlation in these terms:

\[
\rho(X_1, X_2) = \int_E f_1(x)f_2(x)||\bar{f}(x)||^{\alpha-2}m(dx), \quad (23)
\]

\[
\bar{\rho}(X_1, X_2) = \frac{\int_E f_1(x)f_2(x)||\bar{f}(x)||^{\alpha-2}m(dx)}{\left(\int_E \frac{f_1^2(x)}{||f(x)||^{2-\alpha}}m(dx)\int_E \frac{f_2^2(x)}{||f(x)||^{2-\alpha}}m(dx)\right)^{1/2}},
\]

where \(||\bar{f}(x)|| = \left(f_1^2(x) + f_2^2(x)\right)^{1/2}. Formally in the above written formulae we should integrate over \(E_+ = \{x \in L^\alpha : ||\bar{f}(x)|| > 0\}\), but here we agree that integrand is equal to zero if \(||\bar{f}(x)|| = 0\). Comparing expressions of the codifference and \(\alpha\)-covariance, we see that the integrand in (23) is more simple to deal with. We shall demonstrate this by calculating \(\alpha\)-covariance function for Ornstein-Uhlenbeck process (22). Let us denote \(\rho(t) = \rho(X(0), X(t))\) and the normalized \(\alpha\)-covariance function \(\bar{\rho}(t) := \rho(t)/\rho(0)\). Since \(\rho(-t) = \rho(t)\), for \(t > 0\), it is sufficient to consider the case \(t > 0\).

**Proposition 7** Let \(X\) be the process defined in (22). For \(t > 0\) we have

\[
\rho(t) = \frac{1}{\alpha \lambda (1 + \exp(-2\lambda t))^{(2-\alpha)/2}} e^{-\lambda t}, \quad (24)
\]

and

\[
\bar{\rho}(t) = \frac{1}{(2^{-1}(1 + \exp(-2\lambda t)))^{(2-\alpha)/2}} e^{-\lambda t}. \quad (25)
\]

Thus, as \(t \to \infty\),

\[
\bar{\rho}(t) \sim 2^{(2-\alpha)/2} e^{-\lambda t}. \quad (26)
\]
Proof of Proposition 7. From (22) we see that we must apply (23) with
\[ f_1(x) = \exp(\lambda x) I(x \leq 0), \quad f_2(x) = \exp(\lambda x - \lambda t) I(x \leq t). \]

Then
\[ \rho(t) = \int_{-\infty}^{\infty} \frac{\exp(\lambda x) I(x \leq 0) \exp(\lambda x - \lambda t) I(x \leq t)}{(\exp(2\lambda x) I(x \leq 0) + \exp(2\lambda x - 2\lambda t) I(x \leq t))^{(2-\alpha)/2}} dx, \]

and simple integration gives us (24). Since \( \rho(0) = (\alpha\lambda)^{-1/2} \), the equality (25) is obtained from (24), and the relation (26) is obvious. \( \square \)

We can compare these results with corresponding results for the codifference, provided in Example 4.7.1 in [33]. If we denote by \( \tau(t) = \tau(X(0), X(t)) \) and normalized the codifference function by \( \bar{\tau}(t) = \tau(t)(||X(0)||^2)^{-1} \), where \( ||X(0)||^\alpha \) stands for the scale parameter of SoS random variable \( X(0) \), then
\[ \tau(t) = \frac{1}{\alpha\lambda} \left( 1 - (1 - \exp(-\lambda t))^\alpha + e^{-\alpha\lambda t} \right) \] (27)

and
\[ \bar{\tau}(t) \sim \begin{cases} 
\alpha \exp(-\lambda t), & \text{if } 1 < \alpha < 2, \\
2 \exp(-\lambda t), & \text{if } \alpha = 1, \\
\exp(-\alpha\lambda t), & \text{if } 0 < \alpha < 1. 
\end{cases} \] (28)

Comparing (24) with (27) we see that expression for \( \alpha \)-covariance is more simple and, the most important, gives exponential decay independent of \( \alpha \), while for the codifference in (28), in the range \( 0 < \alpha < 1 \), there is \( \alpha \) in the exponent. We see the same effect as in the case of linear processes with exponentially decaying filters, see the discussion at the end of subsection 2.1. The constant in the asymptotic of the normalized \( \alpha \)-covariance function in (26) is continuous in \( \alpha \) and varies in the small interval \((1/2, 1]\), while in (28) dependence of the constant on \( \alpha \) is discontinuous at \( \alpha = 1 \). Dependence of the \( \alpha \)-covariance function on the parameter \( \lambda \) is the same as of the codifference function, namely, if we denote by \( \tau(t, \lambda) \) and \( \rho(t, \lambda) \) the codifference and \( \alpha \)-covariance functions, respectively, of the Ornstein-Uhlenbeck process \( X(t) \) with parameter \( \lambda \), then we have
\[ \tau(t, \lambda_1) < \tau(t, \lambda_2), \quad \rho(t, \lambda_1) < \rho(t, \lambda_2), \] (29)

for \( \lambda_2 < \lambda_1 \) and for all \( 0 < \alpha \leq 2 \). To prove the second relation in (29) (the first one is proved in [33], see p. 210 therein) we consider (for a fixed \( t > 0 \)) the function
\[ h(\lambda) = \frac{\exp(-\lambda t)}{\lambda(1 + \exp(-2\lambda t))^{(2-\alpha)/2}}, \]

and it is easy to show that \( h'(\lambda) < 0 \) for all \( \lambda > 0, \ t > 0, \ 0, \alpha \leq 2 \). Here it is worth to mention that in the case \( \alpha = 2 \) the constants obtained from formulas (24) and (26) do not coincide with constants, given in [33], see p. 210:
if $\alpha = 2$, then $\rho(t) = \tau(t)/2$ (this equality can be seen also from the relation $|s_1|^2 + |s_2|^2 - |s_1 - s_2|^2 = 2s_1s_2$), and this difference comes from the fact that characteristic function of the standard $S\alpha S$ random variable is $\exp(-|t|^\alpha)$, while for Gaussian standard random variable this function is $\exp(-t^2/2)$.

We took the Ornstein-Uhlenbeck process as an example from large class of processes, whose finite dimensional distributions are $S\alpha S$. In the theory of stochastic processes there are important classes of $\alpha$-stable processes, such as sub-Gaussian, moving averages, harmonizable processes, fractional stable noises, etc., rather detailed study of such processes is presented in the fundamental monograph [33]. As the main tool in [33] to study dependence for these processes is used the codifference function. We believe (and this belief is based on the extensions and generalizations given in the last section) that in the case of infinite variance $\alpha$-covariance function is a better substitute for usual covariance function, although we admit that a lot of work must be done - during thirty years there were a lot of papers devoted to the codifference and covariation, while this paper is the first one after 1976 paper [26] (where this notion was only introduced) devoted to $\alpha$-covariance function.

4 Short, long and negative memories

4.1 Memory properties for random processes

The importance of notions of long-range and short-range dependence and notion of memory in the theory of stochastic processes and fields and, in particular, in time series analysis is well-known. The number of monographs and papers devoted to these notions are growing steadily, and this can be explained from one hand, by usage of these notions in many areas, ranging from econometrics and finance to hydrology and climate studies, on the other hand, by the complexity of these notions, complicated relations with other notions. We refer to important survey paper [32] and recent monograph [10] which gave impetus to look at these notions for sequences with infinite variance.

If we consider a stationary sequence with finite variance, there are several ways to define long-range dependence, in [11] there are provided even 8 definitions, but the three of them are main (other 5 are only modifications): via covariance function; via spectral density; and via the growth of partial sums (the so-called Allen variance). Let us note that in many papers the notions "long-range dependence" and "long memory" are used as synonyms. In the above cited monograph [10] in the subject list there is no notions "long-range dependence" or "short-range dependence" and stationary processes with finite second moment are divided into processes with short, long and negative memory; this is done by means of the covariance function. Namely, the following definition is given in [10], see Definition 3.1.2 there.

**Definition 8** A covariance stationary mean zero process $Y_t$, $t \in \mathbb{Z}$, with covariance function $\gamma_k = EY_0Y_k$ has: long memory if $\sum_{k \in \mathbb{Z}} |\gamma_k| = \infty$; short memory if
\[ \sum_{k \in \mathbb{Z}} |\gamma_n| < \infty \quad \text{and} \quad \sum_{j \in \mathbb{Z}} |\gamma_j| > 0; \quad \text{and negative memory (or antipersistence)} \]

if \[ \sum_{k \in \mathbb{Z}} |\gamma_n| < \infty \quad \text{and} \quad \sum_{j \in \mathbb{Z}} \gamma_j = 0. \]

One of the main applications of this definition is to linear processes with white-noise innovations (defined or by the relation (6), either by the analogous relation with summation over all \( \mathbb{Z} \)), and this is due to the rather simple expression of \( \gamma_j \) via coefficients of the filter and the relation \( \sum_{j \in \mathbb{Z}} \gamma_j = (\sum_{j \in \mathbb{Z}} c_j)^2 \). This relation allows to use the sum \( \sum_{j \in \mathbb{Z}} c_j \) for separation of short and negative memories. Therefore, it is tempting by means of the \( \alpha \)-covariance to define the same notions for linear processes (6), namely, we would say that the process (6) has: short memory if \[ \sum_{k \in \mathbb{Z}} |\rho_n| < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} c_j \neq 0; \]
long memory if \( \sum_{k \in \mathbb{Z}} |\rho_n| = \infty \); and negative memory if \[ \sum_{k \in \mathbb{Z}} |\rho_n| < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} c_j = 0. \]

Unfortunately, such classification of linear processes is unappropriate and useless for the following reason. We know that dependence and, particularly, memory properties play an important role in establishing limit properties of partial sum processes constructed from stationary sequences under consideration. These relations between memory properties and limit theorems for partial sum processes are deeply investigated, and we refer to the monograph [10] where these relations are given in details. Here we shall mention only that if \( S_n \) stands for the partial sum of a linear process with innovations with a finite variance and regularly varying filter \( c_k \sim k^{1-d} \), then \( \text{Var}S_n \sim n \) in the case of short memory, and \( \text{Var}S_n \sim n^{1+2d} \) in the case of long memory \((0 < d < 1/2)\) and in the case of negative memory \((-1/2 < d < 0 \text{ and } \sum_{j=0}^{\infty} c_j = 0)\). The similar situation is in the case of linear processes with infinite variance innovations. In [1] general limit theorems for the partial sum process formed by a linear process with innovations belonging to the domain of attraction of a stable law are proved, and these limit theorems can be taken as a basis for classification of linear processes with respect to memory properties. We shall provide here simplified version of the results from [1] avoiding more complicated formulations involving slowly varying functions. Let us consider a linear process

\[ X_k = \sum_{j=0}^{\infty} c_j \eta_{k-j}, \quad k \in \mathbb{Z}, \tag{30} \]

where \( \{\eta_i, i \in \mathbb{Z}\} \) are i.i.d. random variables belonging to the normal domain of attraction of a standard \( S\alpha S \) random variable with ch.f. \( \exp(-|t|^\alpha) \), \( 0 < \alpha < 2 \), and a filter \( \{c_j, j \geq 0\} \) satisfies the relation \( |c_j| \sim j^{-\beta} \). Let us consider the convergence of finite-dimensional distributions of the process

\[ Y_n(t) = A_n^{-1} \sum_{k=1}^{[nt]} X_k. \tag{31} \]

In [1] three cases are separated.

(i) If \( \sum_j |c_j| < \infty \) and \( \sum_j c_j \neq 0 \), then \( A_n \) grows as \( n^{1/\alpha} \), and the limit process is \( \alpha \)-stable Lévy motion.
(ii) If $\alpha > 1$ and $1/\alpha < \beta < 1$, then $A_n$ grows as $n^{1/\alpha+1-\beta}$ (more rapidly comparing with the case (i)) and the limit process is a linear fractional stable motion.

(iii) Let $0 < \alpha < 2$, $\max(1, 1/\alpha) < \beta < 1 + 1/\alpha$, and

$$\sum_{j=0}^{n} c_j \sim (\beta - 1)^{-1} n^{1-\beta},$$

(32)

than $A_n$ grows as $n^{1/\alpha+1-\beta}$ (now more slowly comparing with the case (i)) and again the limit process is a linear fractional stable motion.

It is necessary to note, that condition (32) is stronger then condition $\sum_{j=0}^{\infty} c_j = 0$. Also from these results and Corollary 4 it is clear that in the case $\alpha < 2$ memory properties can not be characterized by the convergence or divergence of series $\sum_{n} \rho_n$, as it was proposed above. Therefore, it seems more natural in the case $\alpha < 2$ memory properties to define according the growth of normalizing constants in limit theorems for partial sums, and this can be done not only for linear processes but for general stationary sequences.

Let $\{\xi_i, i \in \mathbb{Z}\}$ be a strictly stationary sequence which is jointly regularly varying with the index $0 < \alpha < 2$ (for the definition of jointly regularly varying sequence see, for example [2]). In order not to deal with centering we additionally assume that $E\xi_0 = 0$ if $\alpha > 1$ and that $\xi_0$ is symmetric if $\alpha = 1$. Let us denote

$$S_n(t) = \sum_{k=1}^{[nt]} \xi_k,$$

and we suppose that there exists a sequence of normalizing constants $A_n$ such that finite-dimensional distributions (f.d.d) of the process $A_n^{-1} S_n(t)$ converges weakly to corresponding f.d.d. of some stable processes (in particular, distribution of $A_n^{-1} S_n(1)$ converges to an $\alpha$-stable law).

**Definition 9** We say that the sequence $\{\xi_i, i \in \mathbb{Z}\}$ is: of short memory if $A_n = n^{1/\alpha} L(n)$ with some slowly varying function $L$ and the limit process is the Lévy stable motion; of long memory if $A_n = n^{1/\alpha+\delta} L(n)$ with some $0 < \delta < 1 - 1/\alpha$ and the limit process is a linear fractional stable motion; of negative memory if $A_n = n^{1/\alpha+\delta} L(n)$ with some $-1/\alpha < \delta < 0$ and the limit process is a fractional stable motion.

It is worth to note that in this definition the condition $0 < \delta < 1 - 1/\alpha$ means that long memory can be only in the case $\alpha > 1$. Heuristically it can be explained as follows: in the case $0 < \alpha < 1$ and independent $\{\xi_i, i \in \mathbb{Z}\}$ the normalizing sequence satisfies $A_n^\alpha \sim n$, and it is clear that for any stationary sequence the scale parameter of $S_n$ can not grow faster due to the moment inequality $E|\sum_{k=1}^{n} \xi_k|^\beta \leq n E|\xi_1|^\beta$ for any $\beta < \alpha < 1$.

In the case of linear processes the above given three cases from [1], formulated above, exactly gives us three cases of memory, defined in Definition 9, namely, 

(i) we have long memory if $\alpha > 1$, $1/\alpha < \beta < 1$ ($A_n \sim n^{1/\alpha+\delta}$ with $0 < \delta = 1 - \beta < 1 - 1/\alpha$),
A growth of \( \sum \) need only the natural condition sense of Definition 9. The most important fact is that in order to show this we which says that if \( b \) and apply the particular case of Toeplitz lemma (see, for example, [20], p 250) (such choice of \( c \) logical. We consider the growth of the variance

In [10] the asymptotic of this variance is obtained investigating behavior of covariances \( \gamma \), but for our purpose it is more convenient to write explicit expression of \( A^2 \) via coefficients \( c_k \), namely

\[
A^2_n = Var \sum_{k=1}^{n} X_k.
\]

In [10] the asymptotic of this variance is obtained investigating behavior of covariances \( \gamma \), but for our purpose it is more convenient to write explicit expression of \( A^2 \) via coefficients \( c_k \), namely

\[
A^2_n = \sum_{k=0}^{\infty} \left( \sum_{j=1}^{n} c_{j+k} \right)^2 + \sum_{k=1}^{n} \left( \sum_{j=0}^{n-k} c_j \right)^2.
\]

**Example 1.** Let us take at first the case \( 1/2 < \beta = 1-d < 1 \), \( 0 < d < 1/2 \).

If all \( c_k \) have the same sign we know (see [10]) that covariances \( \gamma \) decay as \( n^{-1+2d} \), \( \sum_{n \in \mathbb{Z}} |\gamma_n| = \infty \) and \( A^2 \) grows as \( n^{3-2\beta} = n^{1+2d} \) (long memory in the sense of both definitions). But if we take \( c_k = (-1)^k k^{-\beta} \), \( k \geq 1 \), \( c_0 = 0 \) (such choice of \( c_0 \) gives us \( \sum_{n=0}^{\infty} c_k > 0 \)), then it is not difficult to verify that for \( n = 2m \), \( m \geq 1 \) all \( \gamma_{2m} \) are positive, while all \( \gamma_{2m-1} \), \( m \geq 1 \) are negative but the decay remains the same: \( |\gamma_n| \) tends to zero as \( n^{-1+2d} \), therefore \( \sum_{n \in \mathbb{Z}} |\gamma_n| = \infty \) and we have long memory in the sense of Definition 8. But the growth of \( A^2 \) is only linear, i.e., as in the case of the short memory in the sense of Definition 9. The most important fact is that in order to show this we need only the natural condition \( \sum_k c_k^2 < \infty \), simple conditions of alternation \( c_k = -c_{k+1} \) and monotonicity \( |c_k| > |c_{k+1}| \), which allow to apply Leibnitz theorem for convergence of alternating series. Let us denote \( C = (\sum_{k=0}^{\infty} c_k)^2 \) and apply the particular case of Toeplitz lemma (see, for example, [20], p 250) which says that if \( b_n = \sum_{k=1}^{n} a_k \uparrow \infty \) and \( x_n \to x \), then

\[
\frac{1}{b_n} \sum_{k=1}^{n} a_k x_k \to x.
\]
Rewriting the second sum in (33) as $\sum_{n=0}^{n-1} \left( \sum_{j=0}^{k} c_j \right)^2$ and applying the above formulated statement with $a_k \equiv 1$ and $x_k = \left( \sum_{j=0}^{k} c_j \right)^2$, we have that

$$\frac{1}{n} \sum_{k=1}^{n} \left( \sum_{j=0}^{n-k} c_j \right)^2 \rightarrow C.$$ (34)

Applying the estimate $| \sum_{j=1}^{n} c_{j+k} | < |c_{1+k}|$ we easily get

$$\frac{1}{n} \sum_{k=1}^{n} \left( \sum_{j=1}^{n} c_{j+k} \right)^2 \rightarrow 0.$$ (35)

From (33), (34) and (35) we get $A_n^2 \sim Cn$, which means that the growth of normalizing sequence is the same as in the case of short memory in the sense of Definition 9. Moreover, the situation can be even worse. Taking the same sequence $c_k = (-1)^k k^{-\beta}$, $k \geq 1$, we can take $c_0 = - \sum_{k=1}^{\infty} c_k$. Now the sequence $A_n^2$ even does not grow to infinity. Namely, using the condition $\sum_{k=0}^{\infty} c_k = 0$ and using the property of alternation we can write

$$\left| \sum_{j=0}^{k} c_j \right| = \left| \sum_{j=k+1}^{\infty} c_j \right| < |c_{k+1}|.$$ (36)

Again, rewriting the second sum in (33) as earlier and applying (36), we see that this sum is bounded by the partial sum of convergent series $\sum_{k=0}^{\infty} c_k^2$. Also, as in (35), we see that the first sum in (33) is bounded by the same convergent series. It remains to show that in the case under consideration still we have the relation $\sum_{n \in \mathbb{Z}} |\gamma_n| = \infty$. To this aim using the condition $\sum_{k=0}^{\infty} c_k = 0$ we can write

$$\gamma_n = \sum_{k=0}^{\infty} c_k c_{k+n} = \sum_{k=1}^{\infty} c_k (c_{k+n} - c_n) = I_n^{(1)} + I_n^{(2)},$$

where

$$I_n^{(1)} = \sum_{k=1}^{\infty} c_{2k} (c_{2k+n} - c_n), \quad I_n^{(2)} = \sum_{k=0}^{\infty} c_{2k+1} (c_{2k+1+n} - c_n).$$

Let us consider the case $n = 2m$, $m \geq 1$. Separating the term with $k = 0$ in the second sum and substituting the particular values of $c_k$ we get

$$\gamma_{2m} = \left( \frac{1}{(2m+1)^{\beta}} + \frac{1}{(2m)^{\beta}} \right) - \frac{1}{(2m)^{\beta}} J_1 + J_2(m) + J_3(m),$$

where

$$J_1 = \sum_{k=1}^{\infty} \left( \frac{1}{(2k)^{\beta}} - \frac{1}{(2k+1)^{\beta}} \right), \quad J_2(m) = \sum_{k=1}^{\infty} \frac{1}{(2k(2k+2m))^{\beta}},$$

and so on.
and

\[ J_3(m) = \sum_{k=1}^{\infty} \frac{1}{(2k+1)(2k+2m+1)\beta}. \]

The series \( J_1 \) is alternating, therefore, converging and \( J_1 < 2^{-\beta} < 1 \), while series in the expressions \( J_i(m) \), \( i = 2, 3 \) are absolutely converging since \( 2\beta > 1 \). Integral criterion gives us that both these two series decay as \( (2m)^{1-\beta} \). Since for \( \beta < 1 \) we have \( 2\beta - 1 < \beta \), therefore \( \gamma_{2m} \) for all \( m \geq 1 \) are positive and decay as \( (2m)^{1-\beta} \), therefore \( \sum_{m=1}^{\infty} |\gamma_{2m}| = \infty \). Although we do not need, but it is possible to show that for \( n = 2m+1 \) all \( \gamma_{2m+1} \) are negative, have the same order of decay as \( \gamma_{2m} \) and there is monotonicity: \( |\gamma_{2m+1}| < |\gamma_{2m}| \).

**Example 2.** Now let us consider the case \( 1 < \beta = 1-d < 3/2 \), \( (-1/2 < d < 0) \). In this case \( \sum_{k=0}^{\infty} |c_k| < \infty \) and we have \( \sum_{k=0}^{\infty} |\gamma_k| < \infty \). A linear process with such filter, according Definition 8 can be of short memory if \( |\sum_{k=0}^{\infty} c_k| > 0 \), or of negative memory if \( \sum_{k=0}^{\infty} c_k = 0 \). Again, let us take \( |c_k| = k^{-\beta} \), \( k \geq 1 \) with \( \beta = 1-d > 1 \) and \( c_k = 0, k < 0 \). Let us choose \( c_0 = -\sum_{k=1}^{\infty} c_k \), thus we have the case of negative memory (Definition 8). Consider two extreme cases in this situation. First, let us take all \( c_k \), \( k \geq 1 \) of one sign, let’s say, positive. Using the relation \( \sum_{j=0}^{\infty} c_j = -\sum_{j=n+1}^{\infty} c_j \) it is not difficult (we omit the simple calculations) to show that \( A_n^2 \) grows as \( n^{3-2\beta} = n^{1+2d} \) for \( -1/2 < d < 0 \) \( (1 < \beta < 3/2) \) and we get that the limit process for (31) is fractional Brownian motion with the Hurst parameter \( H = 1/2 + d \). Thus, we have the case of negative memory in the sense of Definition 9, too. If \( \beta > 3/2 \) \( (d < -1/2) \), then the sequence \( A_n^2 \) is bounded and there will be no convergence of f.d.d.; if \( \beta = 3/2 \) then \( A_n^2 \) will grow logarithmic, but this growth does not allow to apply Lamperti theorem (see Theorem 3.4.1 in [10]).

Now let us consider another extreme case: we take all \( c_k \) alternating, that is, \( c_k = (-1)^k k^{-\beta} \), \( k \geq 1 \) (with the same fixed \( \beta \)) and \( c_0 = -\sum_{k=1}^{\infty} c_k \) (negative memory in the sense of Definition 8). In this case it is not difficult to show that \( A_n^2 \) stays bounded. Thus, we see that under conditions \( |c_k| = k^{-1+d} \), \( k \geq 1 \), (for a fixed \(-1/2 < d < 0\)) and \( c_0 = -\sum_{k=1}^{\infty} c_k \) we can get that \( A_n^2 \) grows as \( n^{1+2d} \) or stays bounded. It is an interesting question if it is possible for a fixed \(-1/2 < d < 0\) and any given \( 0 < \delta < 1 + 2d \) to choose the signs of coefficients \( c_k \) so that \( A_n^2 \) would grow as \( n^{\delta} \).

These two examples and considerations before them suggest two conclusions. First one is that notion of dependence should be separated from the memory properties, leaving for the expressions “long-range dependence” and ”short-range dependence” only the meaning that any measure of dependence is decaying slowly or quickly, respectively. The second one is that the memory properties in the case of finite variance should be defined in the same way as in Definition 9, namely, the case \( \alpha = 2 \) (only in this case we must cover two possibilities for a stationary sequence \( \{\xi_i, \ i \in \mathbb{Z}\} \): it can be jointly regularly varying with the index 2 or it can be with finite variance) should be included into Definition 9. Such definition allows to treat memory properties uniquely in both cases of finite and infinite variances. Also it is easy to give explanation for such classification. Long memory means that a stationary process "remem-
ber" the past values in such a way that the volatility of partial sums of this process are growing more rapidly comparing with the sequence of i.i.d. random variables, while negative memory means contrary - volatility of partial sums of this process are growing more slowly. And short memory means that the partial sums of this process behave in the same way as in the case of the sequence of i.i.d. random variables, which has no memory at all. From this explanation it seems that the terms "long memory" and "short memory" are not quite logical, if we would like to leave the term "negative memory", since words "long" and "short" has opposite meanings, while from arguments given above it follows that "long" and "negative" should be as opposite. Thus, more logical terms would be "positive memory", "zero memory", and "negative memory", these terms would be coherent with memory parameters \(0 < d < 1/2\), \(d = 0\), \(-1/2 < d < 0\) (in the case \(\alpha = 2\), see [10], p. 36; in the case \(\alpha < 2\) these intervals would be \(0 < d < 1 - 1/\alpha\), \(-1/\alpha < d < 0\)). Also these new terms fit well with the explanation of properties of increments of limit processes for (31): this process of partial sums always is with dependent increments, but in the case of zero memory it "forgets" this dependence and the limit process is with independent increments, while in the case of memories (both positive and negative) the limit process remains with dependent increments.

Considering linear processes with finite variance we saw that for some filters the normalization constants for partial sums can stay bounded. The similar situation can be in the case of linear processes with infinite variance (\(0 < \alpha < 2\)), since typical normalizing constants are of the form (again we do not take into account slowly varying functions)

\[
A_n^\alpha = \sum_{k=0}^{\infty} \left| \sum_{j=1}^{n} c_{j+k} \right|^\alpha + \sum_{k=1}^{\infty} \left| \sum_{j=0}^{n-k} c_j \right|^\alpha,
\]

and similar analysis as in the case \(\alpha = 2\) reveals the possibility for \(A_n^\alpha\) to stay bounded. Therefore it is reasonable to suggest to call such stationary sequences having strongly negative memory (memory is so strong that it prevents of growing the volatility of partial sums of the process). Although for general stationary sequences with strongly negative memory the problem of limits for partial sums has no sense (see, for example [14], Ch. 18 ), for linear processes even with strongly negative memory this problem is not trivial if we assume that infinitely many coefficients of a filter are non-zero.

Defining the memory properties by means of Definition 9, the next step will be to clarify what behavior of covariances (the case of finite variances) and \(\alpha\)-covariances (the case of infinite variances) give us the particular memory property. For general stationary sequences, without doubt, it is a difficult problem, but even for linear processes in the case of finite variance it is not easy one. If we leave the partition based on convergence or divergence of the series \(\sum_{k\in\mathbb{Z}} |\gamma_k|\) to separate long-range dependence and short-range dependence, then one would guess that short memory (or in new terminology zero memory) will be under the short-range dependence, namely \(\sum_{k\in\mathbb{Z}} |\gamma_k| < \infty\) and additional condition
\[ \sum_{k \in \mathbb{Z}} \gamma_k > 0. \] But the cases of positive and negative memories are more complicated, as Examples 1 and 2 show. For example, from the Example 1 follows that a linear process with long-range dependence can be of zero memory or even strongly negative memory. This means that the condition \( \sum_{k \in \mathbb{Z}} |\gamma_k| = \infty \) is insufficient for positive memory, stronger condition \( |\sum_{k \in \mathbb{Z}} \gamma_k| = \infty \) probably is needed. In the case of infinite variance the situation is much more complicated, even for linear processes the relation between memory properties defined in Definition 9 and the behavior of \( \alpha \)-covariances is not clear.

At the beginning of this subsection we mentioned that it is impossible to characterize memory properties by the convergence or divergence of the series \( \sum_{k \in \mathbb{Z}} |\rho_n| \), but one can try the series \( \sum_{k \in \mathbb{Z}} |\rho_n|^{f(\alpha)} \) with some function \( f(\alpha) \) for \( 0 < \alpha \leq 2 \). It turns out that such approach is successful with the function \( f(\alpha) = (\alpha - 1)^{-1} \) in the case \( 1 < \alpha \leq 2 \). Namely, from Corollary 4 we have that, in the case \( 1 < \alpha \leq 2 \) and \( 1/\alpha < \beta < 1 \),

\[ \rho_n \approx C(\alpha, \beta) n^{1-\beta} \]

therefore, \( \sum_{k \in \mathbb{Z}} |\rho_n|^{(\alpha-1)^{-1}} = \infty \) and we have long memory. If \( 1 < \alpha \leq 2 \), \( 1 < \beta \), and \( \sum_j c_j \neq 0 \), then

\[ \rho_n \approx C(\alpha, \beta) n^{-\max(\beta\alpha-1, \beta)}. \]

Now \( \sum_{k \in \mathbb{Z}} |\rho_n|^{(\alpha-1)^{-1}} < \infty \) and, according results from [1], formulated above, we have short memory. But these are very particular results, since in Corollary 4 we had investigated the behavior of \( \rho_n \) only in the case of regularly varying coefficients of a filter having constant sign. Further research involving the effect of alternation is needed, especially the case of negative memory in the case \( \alpha < 2 \) remains unclear. In the case \( \alpha = 2 \) we know that if we have additional condition that \( \sum_j c_j = 0 \), covariances decay more quickly comparing with the case \( \sum_j c_j \neq 0 \) under the same decay of \( |c_j| \). The same effect should be in the case \( \alpha < 2 \), but at present we have only conjecture that in the case \( 1 < \alpha < 2 \), \( \sum_j c_j = 0 \), and \( c_j = j^{-\beta}(1 + O(j^{-g(\alpha)}) \) with some function \( g \), we should get \( \rho_n \approx C(\alpha, \beta) n^{1-\beta} \).

### 4.2 Memory properties for random fields

The memory properties for random fields are less investigated comparing with the case of processes, even in the case of finite variance. In this case usually for a stationary random field \( X_k, k \in \mathbb{Z}^d \) long memory (which sometimes is used as synonym for long-range dependence) is defined as the property that covariance function \( \gamma_k := E X_0 X_k \) is not absolutely summable: \( \sum_{k \in \mathbb{Z}^d} |\gamma_k| = \infty \), while summability of this series means short memory. An alternative approach (albeit not equivalent) is to define memory properties via spectral density - roughly speaking, a random field has long memory if its spectral density is unbounded (and has singularity at zero). One of the most popular assumptions on the
behavior of covariance function is the following its growth at infinity 

\[ \gamma_k \sim \|k\|^{-\beta} L(\|k\|) b \left( \frac{k}{\|k\|} \right), \quad 0 < \beta < d, \]

where \( L \) is slowly varying at infinity function and \( b \) is non-negative continuous function defined on unit sphere of \( \mathbb{R}^d \). Exactly such condition was used in one of the pioneering works on long range dependence \([5]\), later on it was used with some modifications (changing the norm, dropping the assumption that \( b \) is non-negative, etc). Similar (in form) condition was used to describe the growth of spectral density at origin (see, for example, \([19]\)). Both such conditions (via covariance function and spectral density) gives us the so-called isotropic long memory, also there are papers dealing with non-isotropic long memory of stationary random fields. But both these two approaches (via covariance function and spectral density) has the following shortcomings. As we saw in previous subsections, it is almost impossible to introduce long and short memories by using substitutes of covariance such as \( \alpha \)-covariance or other similar measures of dependence in the case of infinite variance. It seems that negative memory for fields is not introduced even for fields with finite variance (at least the author have not seen any paper on this topic). Therefore it seems quite natural to suggest the same approach which was suggested for stationary processes and which is unified both for random processes with finite and infinite variance, namely, to use the growth of partial sums formed from the random field under consideration. But before giving the strict definitions we shall introduce some notations and shall give some explanations. For any set \( A \subset \mathbb{Z}^d \) let \( \#A \) stands for the cardinality of the set \( A \). If for processes we form partial sums by summing the values of a process over intervals of the increasing length, situation is more complicated when we pass to random fields, since now summation is possible over arbitrary sequence of finite increasing sets \( A_n \subset A_{n+1} \) only with requirement that \( \#A_n \to \infty \) or even over some system of set indexed by multi-indices. If a random field \( X_k, k \in \mathbb{Z}^d \) consists of i.i.d. random variables belonging to the domain of the normal attraction of \( \alpha \)-stable law, \( 0 < \alpha \leq 2 \), then the sequence \( b_n^{1/\alpha} \), where \( b_n = \#A_n \), presents the right normalization for \( \sum_{k \in A_n} X_k \) giving \( \alpha \)-stable law as a limit (again, for simplicity of writing we do not take into account slowly varying functions, since for classification of memory properties only the exponent in the normalizing sequence is important). Passing to general stationary random fields we shall require the joint regular \( \alpha \)-variation (or finite variance in the case \( \alpha = 2 \)) and limit \( \alpha \)-stable law (in order to avoid such trivial situation \( X_k = X \)). Then we would like to take the exponent \( 1/\alpha \) as characterization of short memory and a boundary value between long (or, as we wrote, more logical name it would be "positive") and negative memories, namely, if the normalizing sequence is \( b_n^{1/\alpha + \delta} \) with some \( \delta > 0 \) then we have long memory, while if \( \delta < 0 \) then there is negative memory of the field under consideration. But one can easily notice that such definition of memory properties would be incorrect in a sense that for a given random field memory properties may be dependent on the chosen sequence of sets \( A_n \). Looking more carefully at the Def-
inition 9, one can notice that the same situation is for processes: usually we take summation of the values of a process over intervals \( A_n = \{ k \in \mathbb{Z} : 1 \leq k \leq n \} \), but if we take instead of intervals the sets \( B_n = \{ k = 2m \in \mathbb{Z} : 1 \leq m \leq n \} \), a process with short memory (with respect to sets \( A_n \)) may became of long memory (with respect to sets \( B_n \)). Therefore, trying to define memory for random fields, we must choose some system of sets in \( \mathbb{Z}^d \), and, clearly, in \( \mathbb{Z}^d \) there are many possibilities for such choice, and the classification of memory of stationary fields, generally speaking, will depend on this choice. Although from the first glance it seems as unpleasant factor, on the other hand, such choice gives us more opportunities to investigate memory properties. It is clear that dependence for random fields is much more complicated comparing with dependence for processes (it can be different in different directions). Memory property is even more complicated, since, as we noted speaking about processes, the term "long-range dependencies" only means that a stationary process (or a field) "remember" old (or distant in the case of a field) values, while memory properties also characterize how a process or a field remember these values: due to the memory the volatility of partial sums of the sequence under consideration can be bigger (long memory, or positive memory in the new terminology) or smaller (negative memory), comparing with the sequence which has no memory at all (i.i.d. random variables). Short memory (zero memory) means that the volatility is the same as in the case of i.i.d. random variables. Thus, if we suspect that a random field has the so-called isotropic memory, we will take balls (in Euclidean norm) in \( \mathbb{Z}^d \), but, if we want to look if there is difference in memory properties along axes, we will take rectangles (or even we can rotate rectangles, if we suspect that axes with different memory properties do not coincide with coordinate axes). We shall demonstrate such possibility by simple example of a linear field, and for the simplicity of writing we consider the case \( d = 2 \).

**Example 3.** Let us take a linear field (19) with innovations having finite variances, and let

\[
Z_{n,m} = \sum_{t=1}^{n} \sum_{s=1}^{m} X_{t,s}.
\]

This means that we take the sets \( A_{n,m} = \{(t, s) \in \mathbb{Z}^2 : 1 \leq t \leq n, 1 \leq s \leq m \} \) with the cardinality \( \#A_{n,m} = nm \) and we assume that \( \min(n, m) \to \infty \). Let us take the filter of special structure: \( c_{i,j} = a_i b_j \), where \( a_i, b_j \) are real and \( \sum_{i=0}^{\infty} a_i^2 < \infty, \sum_{j=0}^{\infty} b_j^2 < \infty \). Although such structure of the filter does not mean that the random field is factorized into product of two processes, it turns out that the variance of \( Z_{n,m} \) can be factorized, and this means that memory properties along \( t \) and \( s \) axis can be different. One can easily verify that the following formula, analogous to (33) is true

\[
\text{Var} Z_{n,m} = (D_1 + D_2)(E_1 + E_2),
\]

where

\[
D_1 = \sum_{u=0}^{\infty} A^2_{n,u}, \quad D_2 = \sum_{u=1}^{n} A^2_{n-u}, \quad E_1 = \sum_{v=0}^{\infty} B^2_{m,v}, \quad E_2 = \sum_{v=1}^{m} B^2_{m-v},
\]

26
and

\[ A_k = \sum_{t=0}^{k} a_t, \quad A_{n,k} = \sum_{t=1}^{n} a_{t+k}, \quad B_k = \sum_{t=0}^{k} b_t, \quad B_{n,k} = \sum_{t=1}^{n} b_{t+k}. \]

Comparing (38) with (33) we see that each factor in (38) has exactly the same structure as the right-hand side of (33), only with \( a_i \) or \( b_i \) instead of \( c_i \).

This example allows us to use the analysis carried for linear processes and to get that for the random field with such particular filter we can have all sixteen possible combinations of memory properties (four for each axis, long (positive), short (zero), negative, and strongly negative), for example the random field can have long memory with respect to \( t \) (horizontal) axis and negative memory with respect to \( s \) (vertical) axis. To get such combination it is sufficient to take \( a_0 = 1, \quad a_i = i^{-1+d_1}, \quad 0 < d_1 < 1/2, \quad b_i = i^{-1+d_2}, \quad -1/2 < d_2 < 0, \quad i \geq 1, \quad b_0 = \sum_{i=1}^{\infty} b_i \), then the variance of \( Z_{n,m} \) will grow as \( n^{1+2d_1}m^{1+2d_2} \) (up to the constant, depending on \( d_1, d_2 \)). Also, taking alternating coefficients \( a_i \) or \( b_i \) (or even both) we can face the situation when the variance will not grow with respect to one or even both axes. One more consequence from this simple example is that in the case where the memory is non-isotropic, the cardinality of a set over which is taken partial summation is not appropriate characteristic: if in the above example we take \( d_1 = |d_2| \), then the growth of \( \text{Var} Z_{n,m} \) will be proportional to the cardinality of the rectangle of summation (i.e. \( nm \)) showing the short memory, while in reality we have long and negative memories with respect to corresponding coordinate axes.

Based on these considerations we can propose the following general definition of directional memory, analogous to Definition 9 for processes. For simplicity of writing we shall take again the case \( d = 2 \) (generalization to general case \( d > 2 \) is obvious). Let \( X = (X_{i,j}, i, j \in \mathbb{Z}) \) be a stationary random field with marginal distribution of \( X_{0,0} \) belonging to the domain of attraction of a stable law with index \( 0 < \alpha \leq 2 \), \( EX_{0,0} = 0 \) if \( \alpha > 1 \) and \( X_{0,0} \) is symmetric if \( \alpha = 1 \). Let \( Z_{n,m} \) be defined as in (37).

**Definition 10** We say that a stationary random field \( X \) defined above has directional \((\delta_1, \delta_2)\)-memory, if there exist slowly varying functions \( L_i, \; i = 1, 2 \) such that \( A_{n,m} Z_{n,m} \) converge weakly, as \( \min(m,n) \to \infty, \) to non-degenerate bivariate \( \alpha \)-stable law, \( 0 < \alpha \leq 2 \) and

\[ A_{n,m} = \frac{1}{n^{1/\alpha+\delta_1}m^{1/\alpha+\delta_2}L_1(n)L_2(m)}, \quad -\frac{1}{\alpha} < \delta_i < 1 - \frac{1}{\alpha}. \]

Positive value of corresponding \( \delta \) means positive (long) memory in corresponding direction, similarly, negative value of \( \delta \) shows negative memory, while zero value of \( \delta \) corresponds to zero (short) memory. The first step in application of this definition would be to prove the limit theorems for linear random fields, generalizing results in [1].
5 Extensions, generalizations and open problems

In the last section we provide several possible extensions or generalizations of \( \alpha \)-covariance function.

1) Examining more carefully the paper [26] it is possible to notice that the restriction of the definition of \( \alpha \)-cc to \( S_\alpha S \) random vectors is superfluous and without difficulty the notion of \( \alpha \)-covariance can be extended from \( S_\alpha S \) random vectors to more general \( \alpha \)-stable vectors. In the above cited paper the reason of this restriction was explained by an example of \( \alpha \)-stable non-symmetric random vector \((X_1, X_2)\) with independent coordinates, for which both \( \alpha \)-covariance and \( \alpha \)-correlation, defined by means of centered random variables \( Y_1 \) and \( Y_2 \) (see construction before formulas (3) and (4)) generally do not vanish. For a random vector \((X_1, X_2)\) with independent coordinates and finite second moments covariance is equal to zero only for centered coordinates, since \( E((X_1 - EX_1)(X_2 - EX_2)) = E(X_1 - EX_1)E(X_2 - EX_2) \). For \( \alpha \)-stable non-symmetric random vector \((X_1, X_2)\) with independent coordinates, contrary, \( E(Y_1 - EY_1)(Y_2 - EY_2) = -EY_1EY_2 \neq 0 \), if both expectations are non-zero, while \( EY_1Y_2 = 0 \), since \( \tilde{\Gamma} \) is concentrated on the axes. Thus, definitions of \( \alpha \)-covariance and \( \alpha \)-correlation by formulas (3) and (4) (without centering) can be extended to general \( \alpha \)-stable vectors. It is not difficult to verify that all properties of Proposition 1 remains valid.

2) Since the notion of covariance is defined not only for Gaussian random vectors, but for all vectors having finite second moments (that is, belonging to the domain of normal attraction of a Gaussian law) it is natural to try to extend \( \alpha \)-covariance for random vectors belonging to the domain of normal attraction of a \( S_\alpha S \) random vector. We propose to do this in the following way. Let \((\xi_1, \xi_2)\) be a random vector satisfying the following condition: there exists a finite symmetric measure \( \Gamma \) on \( S_2 \) such that for any Borel set on \( S_2 \)

\[
\lim_{t \to \infty} t^{\alpha} P \left( \frac{\| (\xi_1, \xi_2) \| > t}{(\| (\xi_1, \xi_2) \|)^{-1}} \in A \right) = \Gamma(A).
\]

This condition means that the random vector \((\xi_1, \xi_2)\) belongs to the domain of normal attraction of a \( S_\alpha S \) random vector \( X = (X_1, X_2) \) with the spectral measure \( \Gamma \). We suggest to define \( \alpha \)-covariance and \( \alpha \)-correlation of \((\xi_1, \xi_2)\) by means of the measure \( \Gamma \), as these quantities are defined for \( S_\alpha S \) random vector \( X = (X_1, X_2) \):

\[
\rho(\xi_1, \xi_2) = \rho(X_1, X_2) = \int_{S_2} s_1s_2 \Gamma(ds)
\]

and similarly for \( \tilde{\rho}(\xi_1, \xi_2) \) (taking into account the first generalization, given above, the assumption of the symmetry of \( \Gamma \) can be dropped). The reason for such definition is the following. If we consider sums of i.i.d two-dimensional random vectors with second moment (i.e., \( \alpha = 2 \)), appropriately normalized by scalars, then the covariance matrix of the limit Gaussian distribution is the same as that of summands. The similar situation is in the case \( \alpha < 2 \), when we consider sums of i.i.d. random vectors in the domain of \( \alpha \)-stable random vector:
measure $\Gamma$ which is the main characteristic of summands serves as the spectral measure of a limit stable distribution, and if we agree that $\Gamma$ is "responsible" for dependence properties between components of the limit law, it is natural that the same measure $\Gamma$ defines dependence for summands. Such extension of the notion of $\alpha$-covariance is very useful for linear processes (and fields, too), since considering linear processes (6) usually it is assumed that innovations are only in the normal domain of attraction of $\alpha$-stable random variable. Thus, let us consider a linear process

$$Z(k) = \sum_{j=0}^{\infty} c_j \eta_{k-j}, \quad k \in \mathbb{Z},$$

where $\{\eta_i, i \in \mathbb{Z}\}$, are i.i.d. random variables belonging to the normal domain of a standard $S\alpha S$ random variable with ch.f. $\exp(-|t|^\alpha), \quad 0 < \alpha < 2$, and a filter $\{c_j, \ j \geq 0\}$ is such that (40) is defined correctly. Then it is easy to see that finite dimensional distributions of the process $Z$ belong to the normal domain of corresponding distributions of the process $X$ from (6), therefore, taking into account (39), we get

$$\rho(Z(0), Z(n)) = \rho(X(0), X(n)),$$

and we can use the expressions given in Theorem 5. The same approach can be used for linear fields, too. Here it is appropriate to mention that the codifference for the random vector $(\xi_1, \xi_2)$ can be defined directly by the formula (2), but then $\tau(\xi_1, \xi_2)$ would not be the same as $\tau(X_1, X_2)$. It seems that to calculate $\tau(Z(0), Z(n))$ by means of (2) would be rather difficult.

3) Measures of dependence can be considered not only for finite-dimensional random vectors, but also for random elements with values in infinite-dimensional Banach (or even more general topological vector) spaces. Just after appearance of [26] the author spent a year at Gothenburg university studying infinitely divisible and stable laws in Banach and Hilbert spaces, the results of this work were presented in two preprints [24] and [25]. Part of these results were published later in [27], [28], but part remains unpublished till now. In [24] (see the end of the paper [27]) the analog of the correlation matrix $\Lambda_1$, defined for $k$-dimensional $S\alpha S$ random vector (see Proposition 3 in [26]) was introduced for $S\alpha S$ random vectors with values in a separable Banach space. This analog was named pseudo-correlation operator (it is an operator from $B^*$ (conjugate space of $B$) to $B$, as usual covariance operator), now, adopting terminology of this paper, it would be called $\alpha$-covariance operator. Let us note that with passing from finite-dimensional space to infinite-dimensional spaces one faces the principal difficulty: in the case $1 < \alpha$ not all finite measures on the unit sphere of a Banach space can be spectral measures of an $\alpha$-stable measure on $B$, and, as far as I know, the complete description of such spectral measure in Banach spaces still is not available. In [24] and [27] some properties of $\alpha$-covariance operators of $S\alpha S$ random vectors with values in separable Banach spaces, such as compactness and relations with Gaussian covariance operators, were described, but
also a lot of open problems were formulated, among them description of $\alpha$-stable measures in the space $C(0,1)$ in terms of $\alpha$-covariance operators. It is necessary to stress that during the last decades interest in stationary sequences of random elements in infinite-dimensional spaces has increased, mainly due to functional data analysis. Regularly varying time series in Banach spaces are intensively investigated, the list of references on this topic is growing very rapidly, see, for example, [2], [12], [13], [22] and references therein. We hope that the notion of $\alpha$-covariance operator, introduced in [24] and [27] will be useful in this context, also the approach, proposed to define memory properties in this paper could be applied for stationary regularly varying sequences in Banach spaces, only now one more dimension of complexity will appear - the geometry of Banach spaces. We intend to devote a separate paper to all these problems.

4) As it was mentioned in the introduction, the codifference can be defined for general bivariate random vectors, in particular, for infinitely divisible vectors, containing stable vectors as particular case. In papers [29] and [30] it was demonstrated that the codifference is very useful tool investigating mixing and ergodicity properties of infinitely divisible processes. It turns out that it is possible to introduce the notion of $\alpha$-covariance for a bivariate infinitely divisible random vector $X = (X_1, X_2)$ without finite variance and with the Lévy measure $Q$ (without Gaussian component) in the following way. If the vector $X$ has infinite second moment, the same can be said about the second moments for the measure $Q$, therefore we define the analog of $\alpha$-covariance for $X$ (may be it can be called $Q$-covariance, stressing that dependence between coordinates of an infinitely divisible vector is reflected by the Lévy measure $Q$) as usual covariance for radially re-scaled measure $Q$:

$$\kappa(X_1, X_2) = \int_{\mathbb{R}^2} x_1 x_2 \frac{Q(dx_1dx_2)}{\max(1, x_1^2 + x_2^2)}. \quad (41)$$

Since the Lévy measure $Q$ has similar properties as the spectral measure $\Gamma$ of $\alpha$-stable measures ($X_1$ and $X_2$ are independent if and only if (iff) measure $Q$ is concentrated on axes; coordinates are linearly dependent iff the measure $Q$ is concentrated on a line going through 0), such measure of dependence has the main properties of usual covariance. We note that in the case of a stable vector the quantity, defined in (41) will be equal to $\alpha$-covariance defined in (4) multiplied by the constant $\int_0^{\infty} (r^{\alpha-1} \max(1, r^2))^{-1} dr$. Clearly, we can extend this notion for stationary infinitely divisible sequences and processes $X = (X_t)$, $t \in \mathbb{Z}$, or $t \in \mathbb{R}$. If we denote $\kappa(t) = \kappa(X_0, X_t)$, then preliminary considerations based on Maruyama result from [21] show that vanishing of this function as $t \to \infty$ will be necessary condition for mixing of the process $X$, while (under mild condition on the Lévy measure $Q_0$ of $X_0$) vanishing of the codifference function is necessary and sufficient condition, see Corollary 1 in [29].

We provided four possible generalizations or extensions of the notion of $\alpha$-covariance and $\alpha$-covariance function, this list can be prolonged, but it seems that even the results formulated above allow to ascertain that these notions can be very useful in the case of infinite variance and could be good substitute for
usual covariance function. At the same time one must keep in mind that in case of infinite variance $\alpha$-covariance is not so universal, as covariance is in the $L_2$ (Hilbertian) theory. For example, in linear regression most probably covariation (introduced in [15] and applied for regression and filtration in [23] and [4]) is natural and probably can not be changed by $\alpha$-covariance. Also the interesting questions are about relations of $\alpha$-covariance with James orthogonality (the covariation is directly related, see [33]), association, mixing, distance covariance. All these questions are subjects for the future research.

References

[1] Astrauskas, A. (1983) Limit theorems for sums of linearly generated random variables. Lithuanian Math. J. 23, 127–134.
[2] Baarak, B. and Segers, J. (2009) Regularly varying multivariate time series. Stochastic Process. Appl. 119, 1055–1080.
[3] Bradley, R.C. (2005) Basic properties of strong mixing conditions. A survey and some open questions. Probab. Surveys 2, 107–144.
[4] Cambanis, S. and Miller, G. (1981) Linear problems in $p$th order and symmetric stable processes. SIAM J. Applied Math. 41, 43–69.
[5] Dobrushin, R.L. and Major, P. (1979) Non-central limit theorems for nonlinear functionals of Gaussian fields. Z. Wahr. verw. Geb. 50, 27–52.
[6] Esary, J.D., Proshan, F. and Walkup, D.W. (1967) Association of random variables with applications. Ann. Math. Statist. 38, 1466–1474.
[7] Esary, J.D. and Proshan, F. (1972) Relationships between some concepts of bivariate dependence. Ann. Math. Statist. 43, 651–655.
[8] Garel, B. and Kodia, B. (2013) Estimation and comparison of signed symmetric covariation coefficient and generalized association parameter for alpha-stable dependence modeling. To appear in Communications in Stat. Theory and Methods.
[9] Garel, B. and Kodia, B. (2013) Signed symmetric covariation coefficient for alpha-stable dependence modeling. C.R. Acad. Sci. Paris, Ser. I 347, 315–320.
[10] Giraitis, L., Koul, H.L. and Surgailis, D. (2012) Large Sample Inference for Long Memory Processes. Imperial College Press, London.
[11] Guégan, D. (2005) How can we define the concept of long memory? An econometric survey. Econometric reviews 24, 113–149.
[12] Hult, H. and Linskog, F. (2005) Extremal behavior of regularly varying stochastic processes. Stochastic Process. Appl. 115, 249–274.
[13] Hult, H. and Samorodnitsky, G. (2008) Tail probabilities for infinite series of regularly varying random vectors. Bernoulli 14, 838–864.
[14] Ibragimov, I.A. and Linnik, Yu.V. (1971) Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen.
[15] Kanter, M. (1972) Linear sample spaces and stable processes. J. Funct. Anal. 9, 441–456.
[16] Kokoszka, P.S. and Taqqu, M.S. (1994) Fractional ARIMA with stable innovations. Stoch. Processes Appl. 60, 19–47.
[17] Kokoszka, P.S. and Taqqu, M.S. (1994) Infinite variance stable ARMA processes. J. of Time Series Anal. 15, 203–220.
[18] Kokoszka, P.S. and Taqqu, M.S. (1996) Infinite variance stable moving averages with long memory. J. of Econometrics 73, 79–99.
[19] Lavansier, F. Long memory random fields. Lecture Notes in Stat. 187, 195–220.
[20] Loève, M. (1997) Probability Theory I, Springer, New York, 4 edition.
[21] Maruyama, G. (1970) Infinitely divisible processes. Theory Probab. Appl. 15, 1–22.
[22] Meinguet, T. and Segers, J. (2010) Regularly varying time series in Banach spaces, arXiv:1001.3262v1 [math.PR].
[23] Miller, G. (1978) Properties of certain symmetric stable distributions. J. Multivariate Anal. 8, 346–360.
[24] Paulauskas, V. (1976) Infinitely divisible and stable measures on separable Banach spaces. Preprint, Chalmers university of Tech., No. 1976-15, 1–41.
[25] Paulauskas, V. (1977) The rates of convergence to stable laws and the Law of the Iterated Logarithm in Hilbert space. Preprint, Chalmers university of Tech., No. 1977-5, 1–35.
[26] Paulauskas, V. (1976) Some remarks on multivariate stable distributions. J. Multivariate Anal. 6, 356–368.
[27] Paulauskas, V. (1978) On infinitely divisible and stable measures on separable Banach spaces I. Lithuanian Math. J. (Litovsk. Matem. Sb) 18, 517–527.
[28] V. Paulauskas and Paulauskas, V. and Račkauskas, A. (1980) On infinitely divisible and stable measures on separable Banach spaces II. Lithuanian Math. J. (Litovsk. Matem. Sb) 20, 305–316.
[29] J. Rosinski and T. Žak, Simple conditions for mixing of infinitely divisible processes, Stochastic Process. Appl., 61:277–288, 1996.
[30] Rosinski, J. and Žak, T. (1997) The equivalence of ergodicity and weak mixing for infinitely divisible processes. J. Theor. Probab. 10, 73–86.
[31] Samorodnitsky, G. (2004) Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes. Ann. Probab. 32, 1438–1468.
[32] Samorodnitsky, G. (2006) Long range dependence. Found. Trends in Stoch. Syst. 1, 163–257.
[33] Samorodnitsky, G. and Taqqu, M. (1994) Stable non-Gaussian Random Processes. Models with Infinite Variance. Chapman and Hall, New York.
[34] Székely, G.J. and Rizzo, M.L. (2009) Brownian distance covariance. Ann. Appl. Statist. 3, 1236–1265.
[35] Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007) Measuring and testing independence by correlation of distances. Ann. Statist. 35, 2769–2794.
[36] Wu, W.B. and Mielenzuz, J. (2010) A new look at measuring dependence. Lecture Notes in Statist. 200, 123–142.