TAR(p)/ARCH(1) Process with an Arbitrary Threshold: Guaranteed Parameter Estimation and Change-Point Detection

Sergei Vorobeychikov, Yulia Burkatovskaya and Ekaterina Sergeeva

***Abstract*—A sequential method of unknown autoregressive parameters estimation of TAR(p)/ARCH(1) model with an arbitrary threshold is presented. This procedure is based on the construction of the special stopping rule and weights for weighted least square estimation method, allowing guarantee the prescribe accuracy of the estimation. Also a sequential procedure of change point detection is proposed. Upper bounds for its basic characteristics, such as the probability of false alarm and the delay probability, are obtained. The ergodicity region of TAR(2)/ARCH(1) model is studied and asymptotic properties of the proposed method for ergodic TAR(p)/ARCH(1) process are investigated.**

***Index Terms*—AR/ARCH, guaranteed parameter estimation, change point detectionAR/ARCH, guaranteed parameter estimation, change point detectionT**

I. Introduction

Threshold autoregressive (TAR) models proposed by Tong [1] definitely are one of the most popular classes of nonlinear time series models for conditional mean, because they do not only provide a better fit than linear models, but also reveal a strictly nonlinear behavior (e.g. limit cycles, jump resonance, harmonic distortion) which linear models cannot duplicate [2]. Though sometimes such models have to be completed by a specification of the conditional variance. ARCH/GARCH type models first introduced by Engle [3] are often considered for the conditional variance. One of the most popular applications of the models is analysis and modeling of stock market. In particular, they are used to describe the volatility. A lot of authors note that the classical ARCH/GARCH models do not explain some peculiarities of the volatility behavior, such as asymmetry and response for news. Consequently, rather complicated models based on ARCH/GARCH are proposed and used. Sidorov and others in [4] describe volatility by GARCHJumps models with two separate components (normal and unusual news), which cause two types of innovation (smooth and jump-like innovations). The first component is the GARCH process, and the second one reflects the result of unexpected events and describe jumps in volatility. Guo and Cao in [5] propose a new smooth transition GARCH model, which allows for an asymmetric response of volatility to the size and sign of shocks, and an asymmetric transition dynamics for positive and negative shocks. The authors apply the model to the empirical financial data: the NASDAQ index and the individual stock IBM daily returns. TAR/ARCH models also allows us to describe some non-linear effects, such as clustering and different behavior subject to the sign of a stock return.

Estimators of the unknown parameters based on the idea of the usage of a special stopping rule in order to guarantee precisely their quality in a special sense were first proposed by Wald in [6] and are also very popular. So, Lee and Sriram in [7] proposed a sequential procedure for estimation of TAR(1) parameters, which allows one to construct least squares asymptotically risk efficient estimator. Sriram in [8] used sequential sampling methods to construct confidence intervals for unknown parameters with the fixed size and prescribed coverage probability. Konev and Galtchouk in [9] proposed the sequential least square estimator with the stopping rule determined by the trace of the observed Fisher information matrix, which is asymptotically normally distributed in the stability region. In [10] we developed a sequential procedure for the estimation of unknown parameters of the TAR(1)/ARCH(1) process, which can guarantee the precise accuracy of estimators.

The problem of the change point detection for autoregressive processes with conditional heteroskedasticity is well known and extremely interesting. With different assumption and for different types of models such problem was recently considered for example by Vorobeychikov and others in [10], Fryzlewicz and Subba Rao in [11], Na, Lee and Lee in [12]. Properties of commonly used algorithms are studied asymptotically or by simulations, as theoretical investigation is extremely complicated or hardly ever possible. The usage of the special stopping rule for the least squares estimators with the guaranteed accuracy of unknown parameters allows us to investigate both asymptotic and non-asymptotic properties of algorithms, such as false alarm and delay probabilities.

This paper proposes the guaranteed weighted least square estimators of unknown autoregressive parameters of the TAR(p)/ARCH(1) process with an arbitrary threshold. Asymptotic properties for the estimators are considered and the upper bounds for the standard deviation (asymptotic and non-asymptotic) are constructed. The authors present the procedure of change point detection with guaranteed characteristics for this process.

Manuscript received 12 July, 2016. This paper is supported by The National Research Tomsk State University Academic D.I. Mendeleev Fund Program (NU 8.1.55.2015 L) in 2014–2015 and by RFBR Grant 16-01-00121.

S. Vorobeychikov is with the Department of Applied Mathematics and Cybernetics and International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, 36 Lenin Prospekt, 634050 Tomsk, Russia, e-mail: sev@mail.tsu.ru.

Yu. Burkatovskaya is with the Institute of Cybernetics, Tomsk Polytechnic University, 30 Lenin Prospekt, 634050 Tomsk, Russia; Department of Applied Mathematics and Cybernetics and International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, Tomsk State University, 36 Lenin Prospekt, 634050 Tomsk, Russia, e-mail: tracey@tpu.ru.

E. Sergeeva is with the Institute of Cybernetics, Tomsk Polytechnic University, 30 Lenin Prospekt, 634050 Tomsk, Russia; e-mail: sergeeva_e_e@mail.ru.

II. Problem Statement

We consider the TAR(p)/ARCH(1) autoregressive process specified by the equation

$$\begin{gathered} x_k = X_k \Lambda^1 \mathbf{1}_{\{x_{k-1} \geq a\}} + X_k^T \Lambda^2 \mathbf{1}_{\{x_{k-1} < a\}} \\ + \sqrt{\omega + \alpha^2 x_{k-1}^2} \xi_k; \\ X_k = [x_{k-1}, ..., x_{k-p}]; \\ \Lambda^j = [\lambda_1^j, ..., \lambda_p^j]^T, \quad j = 1, 2, \end{gathered} \tag{1}$$

where $\{\xi_k\}_{k \geq 0}$ is a sequence of independent identically distributed random variables with zero mean and unit variance, $\omega > 0$, $0 < \alpha^2 < 1$, a is an arbitrary known constant. So, the process under consideration is the p-order autoregressive process with ARCH noise and the autoregressive parameters depending on the previous value of the process. The value of the parameter vector $\Lambda = [\Lambda^1, \Lambda^2]$ changes from $\mu = [\mu_1^1, ..., \mu_p^1, \mu_1^2, ..., \mu_p^2]$ to $\beta = [\beta_1^1, ..., \beta_p^1, \beta_1^2, ..., \beta_p^2]$ at the change point θ. The parameters values before and after θ are supposed to be unknown. The difference between μ and β satisfies the condition

$$(\mu^j - \beta^j)^T (\mu^j - \beta^j) \geq \Delta, \quad j = 1, 2, \tag{2}$$

where Δ is the known value defining the minimum difference between the parameters before and after the change point. The problem is to detect the change point θ from the observations x_k.

In [19] and [20], we proposed to detect the instant of the parameters change in the autoregressive process by making use of guaranteed sequential estimators. The sequence of estimators is constructed and the estimators obtained on different time intervals were compared. In paper [21], we applied this approach to more complicated TAR(p)/ARCH(1) model with the unbounded noise variance. In this paper, we extend our approach to TAR(p)/ARCH(1) model with an arbitrary threshold. The ergodicity region of the process is investigated. Besides, more precise asymptotic results are obtained.

III. Ergodicity of the Process

For investigation of asymptotic properties of estimators of unknown parameters of the given models it is important to obtain necessary and sufficient conditions for ergodicity or even strongly for geometric ergodicity of such models . There are distinguished three main approaches to establish geometrical ergodicity in nonlinear conditionally heterockedastic autoregression [2]. The first approach is based on the assumption that the linear regression part becomes main part for the stability research due to the usage of infinite number of values of the process considered; the assumption that the radius of the companion matrix of this linear regression part is less than one [13], [14] for the AR/ARCH model. The second one uses the concept of the Lyapunov exponent for the AR/ARCH [15] and the TAR/ARCH [16] and allows to obtain geometric ergodicity within more general assumptions in much larger parameter space than in [13], [14] but the assumptions appear much more difficult to validate. The last one is the approach, which was proposed first by Liebscher [17] and then extended for AR/ARCH model [3], based on the concept of the joint spectral radius of a set of matrices and also allows to obtain the geometric ergodicity in lager regions of parameter space than [13], [14].

In this paper, we obtained sufficient conditions for ergodicity of process (1) based on one of the theorems given by Mein and Tweedie [22]. We can reduce proving geometric ergodicity of a Markov chain $\{X_n\}$ by verifying the following condition: there is a positive test function $g(X)$ such as and a compact K such as

$$\begin{gathered} a) \; E\left[g\left(X_{k+1}\right)|X_k = X\right] < g(X) - c, \\ c > 0, \quad X \notin K; \\ b) \; E\left[g\left(X_{k+1}\right)|X_k = X\right] < R < \infty, \\ X \in K; \end{gathered} \tag{3}$$

The main problem of this approach is the choice of the function $g(X)$. Let us choose it as a linear function of the X. For the TAR(2)/ARCH(1) process we can write it in the form

$$g(X_{k+1}) = C_1^+ x_k^+ + C_1^- x_k^- + C_2^+ x_{k-1}^+ + C_2^- x_{k-1}^-,$$

where C_1^+, C_1^-, C_2^+, C_2^- are some positive constants,

$$x^+ = \max\{x, 0\}, \quad x^- = \max\{-x, 0\}.$$

We choose the compact $K = [-M, M] \times [-M, M]$, where $M > |a|$.

At first, we consider condition (3a). For $x_{k-1} > M$ using (1) we obtain

$$\begin{gathered} g(X_{k+1}) = C_1^+ \left(\lambda_1^1 x_{k-1} + \lambda_2^1 x_{k-2} + \sqrt{\omega + \alpha^2 x_{k-1}^2} \xi_k\right)^+ \\ + C_1^- \left(\lambda_1^1 x_{k-1} + \lambda_2^1 x_{k-2} + \sqrt{\omega + \alpha^2 x_{k-1}^2} \xi_k\right)^- \\ + C_2^+ x_{k-1}; \end{gathered}$$

Consequently (here and then, $X = (x_{k-1}, x_{k-2})$),

$$\begin{gathered} E\left[g(X_{k+1})|X_k = X\right] \\ = C_1^+ \left(\lambda_1^1 x_{k-1} + \lambda_2^1 x_{k-2}\right) \int_{D_k}^{+\infty} f(x)dx \\ + C_1^+ \sqrt{\omega + \alpha^2 x_{k-1}^2} \int_{D_k}^{+\infty} x f(x)dx \\ - C_1^- \left(\lambda_1^1 x_{k-1} + \lambda_2^1 x_{k-2}\right) \int_{-\infty}^{D_k} f(x)dx \\ - C_1^1 \sqrt{\omega + \alpha^2 x_{k-1}^2} \int_{-\infty}^{D_k} x f(x)dx + C_2^+ x_{k-1}; \\ D_k = \frac{-\lambda_1^1 x_{k-1} - \lambda_2^1 x_{k-2}}{\sqrt{\omega + \alpha^2 x_{k-1}^2}}. \end{gathered}$$

Here $f(x)$ is the density of distribution of the noise ξ_k. By introducing the following notations

$$\int_{D_k}^{+\infty} x f(x)dx = F_k^1, \quad \int_{D_k}^{+\infty} x f(x)dx = F_k^2,$$

and taking into account that $E\xi_k = 0$, we can rewrite the last equation in the form

$$\begin{gathered}E\left[g(X_{k+1})|X_k=X\right]=C_1^+\left(\lambda_1^1x_{k-1}+\lambda_2^1x_{k-2}\right)F_k^1\\+C_1^+\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2\\-C_1^-\left(\lambda_1^1x_{k-1}+\lambda_2^1x_{k-2}\right)(1-F_k^1)\\-C_1^-\sqrt{\omega+\alpha^2x_{k-1}^2}(-F_k^2)+C_2^+x_{k-1}\\=(C_1^++C_1^-)\left(\lambda_1^1x_{k-1}+\lambda_2^1x_{k-2}\right)F_k^1\\-C_1^-\left(\lambda_1^1x_{k-1}+\lambda_2^1x_{k-2}\right)+C_2^+x_{k-1}\\+(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2.\end{gathered}$$

Note that if the function $f(x)$ is symmetric than

$$0\le F_k^2\le\int\limits_0^\infty xf(x)dx=D. \tag{4}$$

The function $g(X)$ has the following form

$$g(X)=\begin{cases}C_1^+x_{k-1}+C_2^+x_{k-2}, & \text{if } x_{k-2}>0;\\ C_1^+x_{k-1}-C_2^-x_{k-2}, & \text{if } x_{k-2}\le 0.\end{cases}$$

Hence, for $x_{k-2}>0$

$$\begin{gathered}E\left[g(X_{k+1})|X_k=X\right]-g(X)\\=(C_1^++C_1^-)\lambda_1^1x_{k-1}F_k^1-C_1^-\lambda_1^1x_{k-1}+C_2^+x_{k-1}\\-C_1^+x_{k-1}\ +(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2\\+(C_1^++C_1^-)\lambda_2^1x_{k-2}F_k^1-C_1^-\lambda_1^2x_{k-2}-C_2^+x_{k-2}.\end{gathered}$$

As x_{k-1} and x_{k-2} can take any value, to fulfill condition (3a) the last expression should satisfy the following inequalities

$$\begin{gathered}1)\ (C_1^++C_1^-)\lambda_1^1x_{k-1}F_k^1-C_1^-\lambda_1^1x_{k-1}+C_2^+x_{k-1}\\-C_1^+x_{k-1}+(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2<0;\\2)\ (C_1^++C_1^-)\lambda_2^1x_{k-2}F_k^1-C_1^-\lambda_1^2x_{k-2}-C_2^+x_{k-2}\le 0.\end{gathered}$$

The further reasonings depend on the signs of the parameters λ_1^1, λ_2^1. Consider all possible cases.

- $\lambda_1^1>0$. In expression 1), the first summand is positive, and $F_k^1<1$; besides, the last summand is also positive and $F_k^2<D$, where D is defined in (4). So, to fulfill 1) we need

$$\begin{gathered}(C_1^++C_1^-)\lambda_1^1x_{k-1}-C_1^-\lambda_1^1x_{k-1}+C_2^+x_{k-1}\\-C_1^+x_{k-1}+(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}D<0;\end{gathered}$$

hence, we obtain the following condition of ergodicity

$$\alpha<\frac{C_1^+(1-\lambda_1^1)-C_2^+}{(C_1^++C_1^-)D}.$$

- $\lambda_1^1\le 0$. In expression 1), the first summand is non-positive; so, to fulfill 1) we need

$$\begin{gathered}-C_1^-\lambda_1^1x_{k-1}+C_2^+x_{k-1}-C_1^+x_{k-1}\\+(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}D<0;\end{gathered}$$

hence, we obtain the following condition of ergodicity

$$\alpha<\frac{C_1^-\lambda_1^1+C_1^+-C_2^+}{(C_1^++C_1^-)D}.$$

To fulfill it, we need the following condition

$$C_1^-\lambda_1^1+C_1^+-C_2^+>0.$$

Introducing the following notations

$$\frac{C_1^-}{C_1^+}=t,\quad \frac{C_2^+}{C_1^+}=s,$$

we can rewrite the ergodicity conditions for λ_1^1 in the form

$$\begin{gathered}\alpha<\frac{(1-\lambda_1^1)-s}{(1+t)D},\quad \lambda_1^1>0;\\ \alpha<\frac{t\lambda_1^1+1-s}{(1+t)D}.\quad \lambda_1^1<0;\end{gathered} \tag{5}$$

Consider now condition 2). Taking into account that $x_{k-2}>0$, we obtain

$$(C_1^++C_1^-)\lambda_2^1F_k^1-C_1^-\lambda_2^1-C_2^+\le 0.$$

- $\lambda_2^1>0$. The first summand is positive; hence, the condition takes the form

$$(C_1^++C_1^-)\lambda_2^1-C_1^-\lambda_2^1-C_2^+\le 0,$$

consequently,

$$\lambda_2^1\le\frac{C_2^+}{C_1^+}=s.$$

- $\lambda_2^1\le 0$. The first summand is non-positive; hence, the condition takes the form

$$-C_1^-\lambda_2^1-C_2^+\le 0,$$

consequently,

$$\lambda_2^1\ge-\frac{C_2^+}{C_1^-}=-\frac{s}{t}.$$

Consider now the case $x_{k-2}<0$. Following the same line of reasoning, we obtain the ergodicity conditions

$$\begin{gathered}1)\ (C_1^++C_1^-)\lambda_1^1x_{k-1}F_k^1-C_1^-\lambda_1^1x_{k-1}+C_2^+x_{k-1}\\-C_1^+x_{k-1}+(C_1^++C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2<0;\\2)\ (C_1^++C_1^-)\lambda_2^1x_{k-2}F_k^1-C_1^-\lambda_2^1x_{k-2}+C_2^-x_{k-2}\le 0.\end{gathered}$$

Condition 1) is the same, and condition 2) can be rewritten in the form

$$-(C_1^++C_1^-)\lambda_2^1F_k^1+C_1^-\lambda_2^1-C_2^-\le 0$$

- $\lambda_2^1>0$. The first summand is non-positive; hence, the condition takes the form

$$C_1^-\lambda_2^1-C_2^-\ge 0,$$

consequently,

$$\lambda_2^1\le\frac{C_2^-}{C_1^-}.$$

- $\lambda_2^1\le 0$. The first summand is positive; hence, the condition takes the form

$$-(C_1^++C_1^-)\lambda_2^1+C_1^-\lambda_2^1-C_2^-\le 0$$

consequently,

$$\lambda_2^1\ge-\frac{C_2^-}{C_1^+}.$$

Introducing the following notation

$$\frac{C_2^-}{C_1^+}=q,$$

we can rewrite the ergodicity conditions for λ_2^1 in the form

$$\max\left\{-\frac{s}{t},-q\right\}\le\lambda_2^1\le\min\left\{s,\frac{q}{t}\right\}. \tag{6}$$

For $x_{k-1} < -M$ using (1) we obtain

$$g(X_{k+1}) = C_1^+ \left(\lambda_1^2 x_{k-1} + \lambda_2^2 x_{k-2} + \sqrt{\omega + \alpha^2 x_{k-1}^2}\xi_k\right) \\ -C_2^- x_{k-1};$$

Consequently,

$$\begin{gathered} E\left[g(X_{k+1})|X_k = X\right] \\ = (C_1^+ + C_1^-)\left(\lambda_1^2 x_{k-1} + \lambda_2^2 x_{k-2}\right) F_k^1 \\ -C_1^-\left(\lambda_1^2 x_{k-1} + \lambda_2^2 x_{k-2}\right) - C_2^- x_{k-1} \\ +(C_1^+ + C_1^-)\sqrt{\omega + \alpha^2 x_{k-1}^2} F_k^2. \end{gathered}$$

The function $g(X)$ has the following form

$$g(X) = \begin{cases} -C_1^- x_{k-1} + C_2^+ x_{k-2}, & \text{if } x_{k-2} > 0; \\ -C_1^- x_{k-1} - C_2^- x_{k-2}, & \text{if } x_{k-2} \le 0. \end{cases}$$

Hence, for $x_{k-2} > 0$ the last expression should satisfy the following conditions

$$\begin{aligned} &1)\ (C_1^+ + C_1^-)\lambda_1^2 x_{k-1} F_k^1 - C_1^- \lambda_1^2 x_{k-1} - C_2^- x_{k-1} \\ &\quad +C_1^+ x_{k-1} + (C_1^+ + C_1^-)\sqrt{\omega + \alpha^2 x_{k-1}^2} F_k^2 < 0; \\ &2)\ (C_1^+ + C_1^-)\lambda_2^2 x_{k-2} F_k^1 - C_1^- \lambda_2^2 x_{k-2} - C_2^+ x_{k-2} \le 0. \end{aligned}$$

For $x_{k-2} < 0$ the conditions are

$$\begin{aligned} &1)\ (C_1^+ + C_1^-)\lambda_1^2 x_{k-1} F_k^1 - C_1^- \lambda_1^2 x_{k-1} - C_2^- x_{k-1} \\ &\quad +C_1^- x_{k-1} + (C_1^+ + C_1^-)\sqrt{\omega + \alpha^2 x_{k-1}^2} F_k^2 < 0; \\ &2)\ (C_1^+ + C_1^-)\lambda_2^2 x_{k-2} F_k^1 - C_1^- \lambda_2^2 x_{k-2} + C_2^- x_{k-2} \le 0. \end{aligned}$$

Conditions 2) are practically the same as in the previous case $x_{k-1} > M$, we only need to use λ_2^2 instead of λ_2^1; so, we obtain

$$\max\left\{-\frac{s}{t}, -q\right\} \le \lambda_2^2 \le \min\left\{s, \frac{q}{t}\right\}. \tag{7}$$

Consider now condition 1) for different signs of λ_1^2.

• $\lambda_1^2 \ge 0$. The first summand in 1) is non-positive; so, the condition has the form

$$\begin{gathered} -C_1^- \lambda_1^2 x_{k-1} - C_2^- x_{k-1} + C_1^- x_{k-1} \\ +(C_1^+ + C_1^-)\sqrt{\omega + \alpha^2 x_{k-1}^2} F_k^2 < 0 \end{gathered}$$

and implies the inequality

$$\alpha < \frac{-C_1^- \lambda_1^2 + C_1^- - C_2^-}{(C_1^+ + C_1^-)D}.$$

• $\lambda_1^2 < 0$. The first summand in 1) is positive; so, the condition has the form

$$\begin{gathered} (C_1^+ + C_1^-)\lambda_1^2 x_{k-1} - C_1^- \lambda_1^2 x_{k-1} - C_2^- x_{k-1} + C_1^- x_{k-1} \\ +(C_1^+ + C_1^-)\sqrt{\omega + \alpha^2 x_{k-1}^2} F_k^2 < 0 \end{gathered}$$

and implies the inequality

$$\alpha < \frac{C_1^+ \lambda_1^2 + C_1^- - C_2^-}{(C_1^+ + C_1^-)D}.$$

So, we obtain the ergodicity conditions for λ_1^1 in the form

$$\begin{aligned} \alpha &< \frac{t(1-\lambda_1^2) - q}{(1+t)D}, \quad \lambda_1^2 \ge 0; \\ \alpha &< \frac{\lambda_1^2 + t - q}{(1+t)D}, \quad \lambda_1^2 < 0. \end{aligned} \tag{8}$$

Finally, by making use (5) and (6) we combine all ergodicity conditions for the parameter α

$$\begin{aligned} \alpha &< \frac{(1-\lambda_1^1) - s}{(1+t)D}, \quad \lambda_1^1 > 0; \\ \alpha &< \frac{t\lambda_1^1 + 1 - s}{(1+t)D}, \quad \lambda_1^1 \le 0; \\ \alpha &< \frac{t(1-\lambda_1^2) - q}{(1+t)D}, \quad \lambda_1^2 \ge 0; \\ \alpha &< \frac{\lambda_1^2 + t - q}{(1+t)D}, \quad \lambda_1^2 < 0. \end{aligned} \tag{9}$$

Note that the upper bound for α should be positive. Taking into account that $t > 0$ we obtain additional conditions

$$\begin{aligned} (1-\lambda_1^1) - s > 0, &\quad \lambda_1^1 > 0; \\ t\lambda_1^1 + 1 - s > 0, &\quad \lambda_1^1 \le 0; \\ t(1-\lambda_1^2) - q > 0, &\quad \lambda_1^2 \ge 0; \\ \lambda_1^2 + t - q > 0, &\quad \lambda_1^2 < 0; \end{aligned} \tag{10}$$

By making use (8) and (7) we combine all ergodicity conditions for the parameter λ_2^l

$$\max\left\{-\frac{s}{t}, -q\right\} \le \lambda_2^l \le \min\left\{s, \frac{q}{t}\right\}. \tag{11}$$

To obtain the widest ergodicity region for α, we should minimize s and q taking into account conditions (11). There can be several cases subject to the parameters λ_2^l.

Case 1: $0 \le \lambda_2^2 < \lambda_2^1$. In this case, condition (11) has the form

$$\lambda_2^1 \le \min\left\{s, \frac{q}{t}\right\},$$

and minimum values of the constants are $s = \lambda_2^1$, $q = t\lambda_2^1$. Note that the case $0 \le \lambda_2^1 < \lambda_2^2$ is in fact the same; we only need to replace further λ_2^1 by λ_2^2. Using this, we can construct the ergodicity region for α subject to the signs of the parameters λ_1^l.

$\lambda_1^1 > 0$, $\lambda_1^2 \ge 0$. The ergodicity region takes the form

$$\alpha < \min\left\{\frac{1-\lambda_1^1-\lambda_2^1}{(1+t)D}, \frac{t(1-\lambda_1^2-\lambda_2^1)}{(1+t)D}\right\} \tag{12}$$

with the additional conditions

$$\begin{gathered} 1-\lambda_1^1-\lambda_2^1 > 0; \\ 1-\lambda_1^2-\lambda_2^1 > 0, \end{gathered}$$

which can be generalized in the form

$$\lambda_1^j + \lambda_2^l < 1, \quad j,l \in \{1,2\}. \tag{13}$$

We obtain the widest region if the parameter t satisfies the equation

$$1-\lambda_1^1-\lambda_2^1 = t(1-\lambda_1^2-\lambda_2^1).$$

In this case both upper bound constants for α are equal; hence, choosing minimum between them, we obtain the maximum value of the upper bound. By expressing t from this and using it in (12), we obtain

$$\alpha < \frac{(1-\lambda_1^1-\lambda_2^1)(1-\lambda_1^2-\lambda_2^1)}{(2-\lambda_1^1-\lambda_1^2-2\lambda_2^1)D}. \tag{14}$$

Due to condition (13), the right hand side of (14) is positive.

$\lambda_1^1 > 0$, $\lambda_1^2 < 0$. The ergodicity region takes the form

$$\alpha < \min\left\{\frac{1-\lambda_1^1-\lambda_2^1}{(1+t)D}, \frac{\lambda_1^2 + t - t\lambda_2^1}{(1+t)D}\right\} \tag{15}$$

with the additional conditions

$$1-\lambda_1^1-\lambda_2^1>0$$

which has already obtained in (13). The parameter t providing the biggest region of ergodicity satisfies the equation

$$1-\lambda_1^1-\lambda_2^1=\lambda_1^2+t-t\lambda_2^1.$$

By expressing t from this and using it in (15), we obtain

$$\alpha<\frac{(1-\lambda_1^1-\lambda_2^1)(1-\lambda_2^1)}{(2-\lambda_1^1-\lambda_1^2-2\lambda_2^1)D}. \qquad (16)$$

Due to condition (13), the right hand side of (16) is positive.

Similarly, we can consider two remaining cases.

$\lambda_1^1\le 0$, $\lambda_1^2\ge 0$. The ergodicity region is

$$\alpha<\frac{(1-\lambda_1^2-\lambda_2^1)(1-\lambda_2^1)}{(2-\lambda_1^1-\lambda_1^2-2\lambda_2^1)D} \qquad (17)$$

if conditions (13) fulfill.

$\lambda_1^1\le 0$, $\lambda_1^2<0$. The ergodicity region is

$$\alpha<\frac{(1-\lambda_2^1)^2-\lambda_1^1\lambda_1^2}{(2-\lambda_1^1-\lambda_1^2-2\lambda_2^1)D} \qquad (18)$$

with the additional conditions

$$(1-\lambda_2^l)^2-\lambda_1^1\lambda_1^2>0. \qquad (19)$$

So, for $\lambda_2^1\ge 0$ and $\lambda_2^2\ge 0$ combining (13)–(19) we obtain the following ergodicity region

$$\begin{gathered}\lambda_2^l\ge 0,\quad l\in\{1,2\};\\ \lambda_1^j+\lambda_2^l<1,\quad j,l\in\{1,2\};\\ \lambda_1^1\lambda_1^2<(1-\lambda_2^l)^2,\quad l\in\{1,2\};\\ \alpha<\frac{\min\{A_1,A_2,A_3,A_4\}}{(2-\lambda_1^1-\lambda_1^2-2\max\{\lambda_2^1,\lambda_2^2\})D}=\alpha^*,\end{gathered} \qquad (20)$$

where

$$\begin{gathered}A_1=(1-\lambda_1^1-\max\{\lambda_2^1,\lambda_2^2\})(1-\lambda_1^2-\max\{\lambda_2^1,\lambda_2^2\});\\ A_2=(1-\lambda_1^1-\max\{\lambda_2^1,\lambda_2^2\})(1-\max\{\lambda_2^1,\lambda_2^2\});\\ A_3=(1-\lambda_1^2-\max\{\lambda_2^1,\lambda_2^2\})(1-\max\{\lambda_2^1,\lambda_2^2\});\\ A_4=(1-\max\{\lambda_2^1,\lambda_2^2\})^2-\lambda_1^1\lambda_1^2.\end{gathered}$$

Case 2: $\lambda_2^1<\lambda_2^2\le 0$. In this case, condition (11) has the form

$$\min\left\{-q,-\frac{s}{t}\right\}\le\lambda_2^1$$

and minimum values of the constants are $s=-t\lambda_2^1$, $q=-\lambda_2^1$. Note that the case $\lambda_2^2<\lambda_2^1\le 0$ is in fact the same; we only need to replace further λ_2^1 by λ_2^2. Using this, we can construct the ergodicity region for α as described above for the case 1. The region has the following form

$$\begin{gathered}\lambda_2^l\le 0,\quad l\in\{1,2\};\\ (1-\lambda_1^1)(1-\lambda_1^2)-\left(\lambda_2^l\right)^2>0,\quad l\in\{1,2\};\\ \lambda_1^j+\lambda_2^l(\lambda_1^{3-j}+\lambda_2^l)<1,\quad j,l\in\{1,2\};\\ \lambda_1^1\lambda_1^2+\lambda_2^l(\lambda_1^1+\lambda_1^2+\lambda_2^l)<1,\quad l\in\{1,2\};\\ \alpha<\frac{\min\{B_1,B_2,B_3,B_4\}}{(2-\lambda_1^1-\lambda_1^2-2\min\{\lambda_2^1,\lambda_2^2\})D}=\alpha^*,\end{gathered} \qquad (21)$$

where

$$\begin{gathered}B_1=(1-\lambda_1^1)(1-\lambda_1^2)-\left(\min\{\lambda_2^1,\lambda_2^2\}\right)^2;\\ B_2=1-\lambda_1^1-\min\{\lambda_2^1,\lambda_2^2\}(\lambda_1^2+\min\{\lambda_2^1,\lambda_2^2\});\\ B_3=1-\lambda_1^2-\min\{\lambda_2^1,\lambda_2^2\}(\lambda_1^1+\min\{\lambda_2^1,\lambda_2^2\});\\ B_4=1-\lambda_1^1\lambda_1^2-\min\{\lambda_2^1,\lambda_2^2\}(\lambda_1^1+\lambda_1^2+\min\{\lambda_2^1,\lambda_2^2\}).\end{gathered}$$

If we choose

$$M>\frac{\omega}{(\alpha^*)^2-\alpha^2},$$

and

$$g(X_{k+1})=x_k^+ +tx_k^- +sx_{k-1}^+ +qx_{k-1}^-,$$

where s, q and t are constructed as above, subject to signs of λ_1^j, λ_2^l, $j,l\in\{1,2\}$, condition (3a) will be fulfilled.

Case 3: $\lambda_2^1<0<\lambda_2^2$. In this case, evaluation s and q is more difficult then in the previous one, because we should take into account not only signs of λ_2^j but their absolute values. As a result, we have three possible sets of values for s and q, and we should consider these sets subject to signs of λ_1^l, i.e., for four cases. These calculations are rather cumbersome; so, we do not consider this case in the paper.

To prove condition (3b) we have to bound from above $E\left[g(X_{k+1})|X_k=X\right]$ when $X\in K$; we have

$$\begin{gathered}E\left[g(X_{k+1})|X_k=X\right]\\ =\left(C_1^+ + C_1^-\right)\left(\lambda_1^j x_{k-1}+\lambda_2^j x_{k-2}\right)F_k^1\\ -C_1^-\left(\lambda_1^j x_{k-1}+\lambda_2^j x_{k-2}\right)+C_2^+x_{k-1}\\ +(C_1^+ + C_1^-)\sqrt{\omega+\alpha^2x_{k-1}^2}F_k^2\\ \le\left(C_1^+ + C_1^-\right)\left(|\lambda_1^j|+|\lambda_2^j|\right)M\\ +C_2^+M+(C_1^+ + C_1^-)\sqrt{\omega+\alpha^2M^2}D<\infty.\end{gathered}$$

Note that when $\lambda_2^1=\lambda_2^2=0$, the ergodicity regions (20) and (21) matches the one obtained in [10] for TAR(1)/ARCH(1). It differs from regions described in [18] and [14], because it includes negative values of the parameters $\lambda_1^1<-1$ or $\lambda_1^2<-1$; in [18] and [14] absolute values of all parameters are less then one. Cline and Pu in [16] obtained the exact ergodicity region for more general model TAR(p)/ARCH(p), but it should be calculated numerically for $p>1$; in our paper, we proposed explicit expressions.

Fig. 1 demonstrates an example of an ergodic TAR(2)/ARCH(1) process behavior, and Fig. 2 presents an example of non-ergodic one. Both trajectories have cluster effect and outliers, but the latter one is more chaotic and maximum values of the process are greater then for the former one.

IV. Guaranteed parameter estimator

Since the parameters both before and after the change point are unknown, it is logical to apply the estimators of the unknown parameters in the change point detection procedure. In this section we construct guaranteed sequential parameter estimators for the parameter vectors $\Lambda^j, j=1,2$. Such estimators were proposed in [20] for an autoregressive process. The main advantage of the estimators is their preassigned mean square accuracy depending on the parameter of the estimation procedure.

It should be noted that if parameters ω and α are unknown then the process (1) has unknown and unbounded from above noise variance. To obtain a process with bounded noise variance we denote $\max\{1,|x_{k-1}|\}$ as m_k and rewrite the

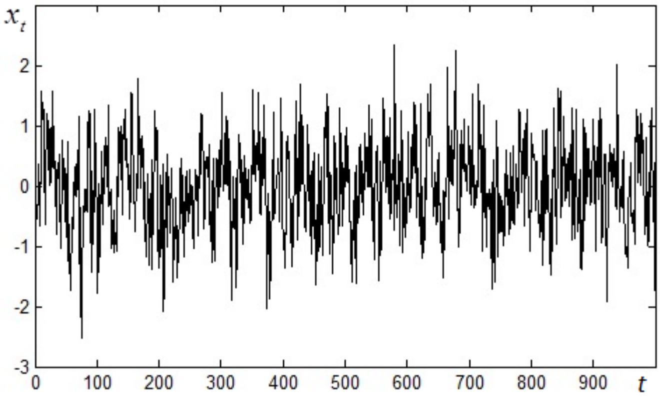

Fig. 1. Ergodic TAR(2)/ARCH(1) process

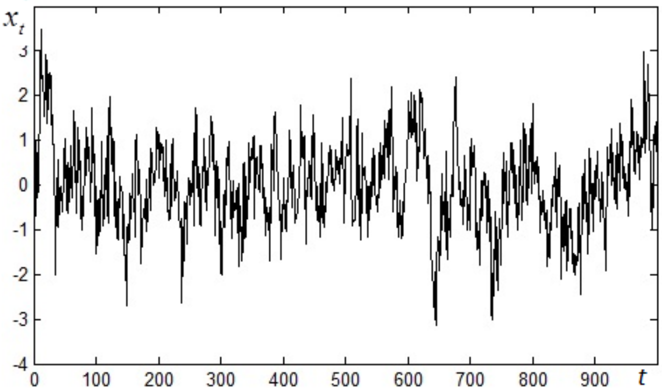

Fig. 2. Non-ergodic TAR(2)/ARCH(1) process

process in the form

$$\begin{gathered} y_k = Y_k^1 \Lambda^1 + Y_k^2 \Lambda^2 + \gamma_k \xi_k; \\ Y_k^1 = \frac{1}{m_k} X_k \mathbf{1}_{\{x_{k-1} \geq a\}}. \quad Y_k^2 = \frac{1}{m_k} X_k \mathbf{1}_{\{x_{k-1} < a\}}; \\ y_k = \frac{x_k}{m_k}, \quad \gamma_k = \frac{\sqrt{w + \alpha^2 x_{k-1}^2}}{m_k}. \end{gathered} \tag{22}$$

The noise variance of the process $\{y_k\}$ is bounded from above by the unknown value $(\omega + \alpha^2)$. To eliminate the influence of the unknown constant in [10] we proposed to use the special factor Γ_N constructed by first N observations in the following form

$$\begin{gathered} \Gamma_N = C_N \sum_{k=1}^{N} \left(\frac{x_k}{\min\{1, |x_k|\}} \right)^2; \\ C_N = E \left(\sum_{k=1}^{N} \xi_k^2 \right)^{-1}, \end{gathered} \tag{23}$$

where N observations are taken at the interval where all the values $|x_k|$ are sufficiently large. It was proved in [20] that for the process AR(p) the compensating factor satisfies the following condition

$$E \frac{1}{\Gamma_N} \leq \frac{1}{\omega + \alpha^2}. \tag{24}$$

This proof can be generalized for our case with minimum changes so we omit it.

If the random variables $\{\xi_k\}$ have standard normal distribution, then the sum $\sum\limits_{k=1}^{N} \xi_k^2$ has χ^2 distribution with N degrees of freedom. In this case

$$C_N = \frac{1}{2^{N/2} \Gamma(N/2)} \int_0^{+\infty} x^{N/2-3} e^{-x/2} dx = \frac{1}{(N-2)(N-4)}.$$

This constant is defined for $N \geq 5$.

Let us consider now a weighted least squares estimator for process (22). The process can be rewritten in the form

$$\begin{gathered} y_k = Y_k \Lambda + \gamma_k \xi_k; \\ Y_k = [Y_k^1, Y_k^2]. \end{gathered} \tag{25}$$

So the weighted least squares estimator has the following form

$$\begin{gathered} \hat{\Lambda} = C^{-1}(m) \sum_{k=N+1}^{m} v_k (Y_k)^T y_k; \\ C(m) = \sum_{k=N+1}^{m} v_k (Y_k)^T Y_k, \end{gathered} \tag{26}$$

where $0 < v_k \leq 1$. According to definition (22) $(Y_k^j)^T Y_k^i = O_p$ for $i \neq j$ (here O_p stands for a zero matrix of the order p). Hence, taking into account (25), one obtains that the matrix $C(m)$ has a block structure

$$\begin{gathered} C(m) = \begin{bmatrix} C(1,m) & O_p \\ O_p & C(2,m) \end{bmatrix} \\ C^{-1}(m) = \begin{bmatrix} C^{-1}(1,m) & O_p \\ O_p & C^{-1}(2,m) \end{bmatrix} \\ C(j,M) = \sum_{k=N+1}^{m} v_k (Y_k^j)^T Y_k^j, \quad j = 1,2. \end{gathered}$$

Using this result and (22), (25) one can obtain

$$\begin{aligned} \hat{\Lambda}^1 &= [C^{-1}(1,m) \; O_p] \sum_{k=N+1}^{m} v_k (Y_k)^T y_k \\ &= [C^{-1}(1,m) \; O_p] \sum_{k=N+1}^{m} v_k (Y_k)^T (Y_k \Lambda + \gamma_k \xi_k) \\ &= [C^{-1}(1,m) \; O_p] \begin{bmatrix} C(1,m)\Lambda_1 \\ C(2,m)\Lambda_2 \end{bmatrix} \\ &+ [C^{-1}(1,m) \; O_p] \sum_{k=N+1}^{m} v_k (Y_k)^T \gamma_k \xi_k. \end{aligned}$$

Hence,

$$\begin{gathered} \hat{\Lambda}^1 = \Lambda_1 + C^{-1}(1,m) \eta(1,m); \\ \eta(1,m) = \sum_{k=N+1}^{m} v_k (Y_k^1)^T \gamma_k \xi_k. \end{gathered} \tag{27}$$

The same result can be obtained for $\hat{\Lambda}^2$. It allows us to construct estimators for Λ^1 and Λ^2 separately, i.e.

$$\hat{\Lambda}^j = C^{-1}(j;m) \sum_{k=N+1}^{m} v_k (Y_k^j)^T y_k, \quad j = 1,2. \tag{28}$$

The obtained estimator can be modified in order to bound the standard deviation of the estimator from above. To do so, we change the sample size m for a special random stopping time τ^j. Also, we use special weights $v_{j,k}$ for every estimator.

Let $H > 0$ be a parameter of the estimation procedure. Further we prove that it defines the accuracy of the proposed parameter estimators. Then the estimators are constructed by using the sequential weighted least squares method and consequently can be written in the following form:

$$\hat{\Lambda}^j = \hat{\Lambda}^j(H) = C^{-1}(j;\tau^j)\sum_{k=N+1}^{\tau^j} v_{j,k}(Y_k^j)^T y_k; \\ C(j,M) = \sum_{k=N+1}^{m} v_{j,k}(Y_k^j)^T Y_k^j, \quad j=1,2. \tag{29}$$

Let $\nu_{\min}(j,M)$ be the minimum eigenvalue of the matrix $C(j,M)$. Then the stopping instants $\tau^j = \tau^j(H)$ are defined by the following conditions

$$\tau^j = \inf\left(M > N:\ \nu_{\min}(j,M) \geq H\right). \tag{30}$$

Now we consider the choice of the weights $v_{j,k}$. Let the matrix $C(j,M)$ be degenerate for $M = N+1,...,N+\sigma^j$ and $C(j,\sigma^j+1)$ be non-degenerated. The weights on the interval $[N+1, N+\sigma^j]$ are taken in the following form:

$$v_{j,k} = \begin{cases} \dfrac{1}{\sqrt{\Gamma_N Y_k^j (Y_k^j)^T}}, & \text{if } Y_N^j,\dots,Y_k^j \text{ are linearly independent;} \\ 0, & \text{otherwise.} \end{cases} \tag{31}$$

The weights on the intervals $[N+\sigma^j+1, \tau^j-1]$ are found from the following condition:

$$\frac{\nu_{\min}(j,k)}{\Gamma_N} = \sum_{l=N+\sigma}^{k} v^2(j,l) Y_l^j (Y_l^j)^T. \tag{32}$$

At the instants τ^j, the weights are determined by the condition:

$$\frac{\nu_{\min}(j,\tau^j)}{\Gamma_N} \geq \sum_{l=N+\sigma+1}^{\tau^j} v^2(j,l) Y_l^j (Y_l^j)^T; \\ \nu_{\min}(j,\tau^j) = H. \tag{33}$$

Theorem 1. Let the parameter vector Λ^j in (1) be constant. Then the stopping time τ^j (30) is finite with probability one and the mean square accuracy of estimator (29) is bounded from above

$$E||\hat{\Lambda}^j(H) - \Lambda^j||^2 \leq \frac{H+p-1}{H^2}. \tag{34}$$

Proof. According to the definition of the instant τ^j (30) it is finite with probability one if

$$\sum_{l=N+\sigma}^{k} v^2(j,l) Y_l^j (Y_l^j)^T \to \infty \text{ as } k \to \infty. \tag{35}$$

The series converges if and only if $\forall \varepsilon > 0$ as $M \to \infty$ (see [23])

$$P\left\{\sum_{l\geq M} v^2(j,l) Y_l^j (Y_l^j)^T \geq \varepsilon\right\} \to 0. \tag{36}$$

The factor $Y_l^j(Y_l^j)^T$ does not tend to zero because the absolute value of the first component is equal to $|x_{l-1}|\mathbf{1}_{[x_{l-1}\geq a]}/\max\{1,|x_{l-1}|\}$ for $j=1$, and it is equal to $|x_{l-1}|\mathbf{1}_{[x_{l-1}<a]}/\max\{1,|x_{l-1}|\}$ for $j=2$. According to equation (1), $|x_{l-1}|$ exceeds unity with a non-zero probability and can be both greater and smaller then a; hence, the absolute value of the first component is equal to 1 with non-zero probability. So condition (36) can hold true only because of the choice of the weights $v(j,l)$.

Suppose that the matrix $C(j, M-1)$ is not diagonal. According to the definition of the minimum eigenvalue of a matrix

$$\nu_{\min}(j,M) = \min_{x:||x||=1}(x, C(j,M)x),$$

where (x,y) is the scalar product of the vectors x and y. Then by using (29), we obtain

$$\nu_{\min}(N+1,N) \\ = \min_{x:||x||=1}(x, ((C(j,M-1) + v_{j,M}(Y_M^j)^T Y_M^j)x) \\ = \min_{x:||x||=1}((x, C(j,M-1)x) + v_{j,M}(Y_M^j x)^2).$$

Let z_M be the argument of the minimum in the above equation. According to (32), we obtain

$$(z_M, C(j,M-1)z_M) + v_{j,M}(Y_M^j z_M)^2 \\ = \nu_{\min}(j,M-1) + v_{j,M}^2 Y_M^j (Y_M^j)^T.$$

So we have derived the quadratic equation for $v_{j,M}$ with roots in the form

$$v_{1,2} = \frac{1}{2Y_M^j(Y_M^j)^T}\left[(Y_M^j z_M)^2 \pm \sqrt{D}\right]; \\ D = (Y_M^j z_M)^4 \\ +4Y_M^j(Y_M^j)^T[(z_M, C(j,M-1)z_M) - \nu_{\min}(j,M-1)].$$

It is obvious that

$$(z_M, C(j,M-1)z_M) - \nu_{\min}(j,M-1) \geq 0.$$

Thus the following two cases are possible.

Case 1. The equation has two zero roots: $v_1 = v_2 = 0$. This is possible if and only if z_M is the eigenvector of the matrix $C(j,M-1)$ corresponding to $\nu_{\min}(j,M-1)$ and $Y_M^j z_M = 0$. However, the first component of Y_M^j depends on the random variable ξ_M, which is independent on the $\{Y_k^j\}_{k<M}$. Hence the vector Y_M^j is orthogonal to the given eigenvector of the matrix $C(j,M-1)$ with zero probability.

Case 2. The equation has one non-positive and one positive root. Taking the major root as $v_{j,M}$, one obtains

$$v_{j,M}^2 Y_M^j (Y_M^j)^T \geq \frac{(Y_M^j z_M)^4}{Y_M^j(Y_M^j)^T} \\ +(z_M, C(j,M-1)z_M) - \nu_{\min}(j,M-1). \tag{37}$$

The first term in (37) is equal to $Y_M^j(Y_M^j)^T\cos^4(\alpha_M)/2$, where α_M is the angle between Y_M^j and z_M. Since $Y_M^j(Y_M^j)^T$ does not converge to zero, the first term converges to zero if and only if $\cos(\alpha_M) \to 0$ when $M \to \infty$. On the other hand, if the second term in (37) converges to zero then z_M converges to the eigenvector of the matrix $C(j,M-1)$ corresponding to $\nu_{\min}(j,M-1)$. If $v_{j,M} \to 0$, then the matrix $C(j,M)$ changes slightly with the increasing M. Hence, the eigenvectors of the matrix change slightly too, and z_M converges to a certain vector z^*. Therefore, the right side of (37) converges to zero if the cosine of the angle between Y_M^j and z^* converges to zero. However, the first component of Y_M^j depends on x_{M-1} which can take any value, this cosine can be sufficiently large with non-zero probability.

Note that condition (36) can hold true if all eigenvalues of the matrix $C(j, M-1)$ for certain M are equal. It is possible if and only if the matrix $C(j, M-1)$ is diagonal. The matrix $C(j, N+k) = v_{j,N+k}(Y_{N+k}^j)^T Y_{N+k}^j$, where k is the least number such as Y_{N+k}^j is non-zero, is not diagonal. It can easily be proved that if the matrix $C(j, M-1)$ is not diagonal, then the next matrix $C(j, M)$ is diagonal with zero probability.

Hence, condition (36) does not hold true for TAR(p)/ARCH(1) process (1), and this implies (35).

According to (27), one can obtain

$$\hat{\Lambda}^j = \Lambda^j + C^{-1}(j, \tau^j)\eta(j, \tau^j);$$
$$\eta(j, \tau^j) = \sum_{k=N+1}^{\tau^j} v_{j,k}(Y_k^j)^T \gamma_k \xi_k.$$

By using the norm properties and (32), one obtains

$$\begin{aligned} ||\hat{\Lambda}^j(H) - \Lambda^j||^2 &\leq (\nu_{\min}(j, \tau^j))^{-2}||\eta(j, \tau^j)||^2 \\ &\leq \frac{||\eta(j, \tau^j)||^2}{H^2}. \end{aligned} \quad (38)$$

Let $F_k = \sigma\{\xi_1, ..., \xi_k\}$ be a sigma-algebra generated by the random variables $\{\xi_1, ..., \xi_k\}$ and $\tau^j(M) = \min\{\tau^j, M\}$ is a truncated stopping instant. According to (30) the instant $\tau^j(M)$ satisfy the condition $\{\tau^j(M) = k\} \in F_{k-1}$. Using the properties of conditional expectations one obtains

$$\begin{aligned} &E||\eta(j, \tau^j(M))||^2 \\ &= E\sum_{k=N+1}^{M} E[v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2 \xi_k^2 \mathbf{1}_{\tau^j \leq k}|F_{k-1}] \\ &+2E\sum_{k=N+2}^{M}\sum_{l=N+1}^{k-1} E[v_{j,k}v_{j,l} Y_k^j (Y_l^j)^T \gamma_k \gamma_l \xi_k \xi_l \mathbf{1}_{\tau^j \leq k}|F_{k-1}] \\ &= E\sum_{k=N+1}^{M} v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2 \mathbf{1}_{\tau^j \leq k} E[\xi_k^2|F_{k-1}] \\ &+2E\sum_{k=N+2}^{M}\sum_{l=N+1}^{k-1} v_{j,k}v_{j,l} Y_k^j (Y_l^j)^T \gamma_k \gamma_l \xi_l \mathbf{1}_{\tau^j \leq k} E[\xi_k|F_{k-1}]. \end{aligned}$$

Since ξ_k does not depend on F_{k-1}, the second summand is equal to zero and one obtains

$$E||\eta(j, \tau^j(M))||^2 = E\sum_{k=N+1}^{\tau^j(M)} v_{j,k}^2 Y_k^j (Y_k^j)^T \gamma_k^2.$$

Due to the choice of the weights $v_{j,k}$ (31–32) one obtains

$$\begin{aligned} E\sum_{k=N+1}^{\tau^j} v_{j,k}^2 Y_k^j (Y_k^j)^T &= E\sum_{k=N+1}^{N+\sigma^j} v_{j,k}^2 Y_k^j (Y_k^j)^T \\ +E\sum_{k=N+\sigma^j+1}^{\tau^j} v_{j,k}^2 Y_k^j (Y_k^j)^T &\leq \frac{p-1}{\Gamma_N} + \frac{H}{\Gamma_N}. \end{aligned}$$

According to (22) one can see that $\gamma_k^2 \leq \omega + \alpha^2$. Note that $\tau^j(M) \to \tau^j$ as $M \to \infty$, so

$$E||\eta(j, \tau^j)||^2 \leq (\omega + \alpha^2)(H + p - 1)E\frac{1}{\Gamma_N}.$$

Due to property (24) of the factor Γ_N and inequality (38) one obtains

$$||\hat{\Lambda}^j(H) - \Lambda^j||^2 \leq \frac{H + p - 1}{H^2}.$$

The theorem has been proved.

Further we establish asymptotic properties of the constructed procedures. We need the following auxiliary result.

Lemma 1. *Let $\xi_1, \ldots, \xi_n$ be independent identically distributed standard Gaussian variables. Then for any $\lambda_1, \ldots, \lambda_n$, $\lambda_i \geq 0$, $\lambda_1 + \ldots + \lambda_n = 1$ and for sufficiently large C*

$$\mathcal{P}\left\{\lambda_1\xi_1^2 + \ldots + \lambda_n\xi_n^2 > C\right\} \leq \mathcal{P}\left\{\xi_1^2 > C\right\}. \quad (39)$$

Proof. First we give the proof for the case of $n = 2$. We need to minimize the function

$$J(\lambda_1) = \mathcal{P}\left\{\lambda_1\xi_1^2 + (1-\lambda_1)\xi_2^2 < C\right\} \to \min_{0\leq\lambda_1\leq 1/2}. \quad (40)$$

Here we take into account that the variables ξ_1 and ξ_2 are independent identically distributed, thus $J(\lambda_1) = J(1-\lambda_1)$. We rewrite the last expression in the form

$$\begin{aligned} J(\lambda_1) &= \int_0^{C/\lambda_1} \mathcal{P}\left\{\lambda_1 y + (1-\lambda_1)\xi_2^2 < C\right\} f(y)dy \\ &= \int_0^{C/\lambda_1} \mathcal{P}\left\{\xi_2^2 < \frac{C - \lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\ &= \int_0^{C/\lambda_1} F\left(\frac{C - \lambda_1 y}{1-\lambda_1}\right) f(y)dy, \end{aligned}$$

where $f(\cdot)$ is the density of distribution, $F(\cdot)$ is the distribution function of the variable ξ_i^2. Differentiating $J(\lambda_1)$, one obtains

$$\begin{aligned} J'(\lambda_1) &= \int_0^{C/\lambda_1} f\left(\frac{C - \lambda_1 y}{1-\lambda_1}\right) \frac{C - y}{(1-\lambda_1)^2} f(y)dy \\ &= \int_0^{C} f\left(\frac{C - \lambda_1 y}{1-\lambda_1}\right) \frac{C - y}{(1-\lambda_1)^2} f(y)dy \\ &- \int_C^{C/\lambda_1} f\left(\frac{C - \lambda_1 y}{1-\lambda_1}\right) \frac{y - C}{(1-\lambda_1)^2} f(y)dy. \end{aligned}$$

Both integrals in the last expression are positive. The second one tends to zero when $C \to \infty$. Thus

$$\lim_{C\to\infty} J'(\lambda_1) \geq 0 \quad \forall\lambda_1 \in [0, 1/2].$$

For $\lambda_1 = 0$ one can obtain

$$J'(0) = \int_0^{\infty} f(C)(C - y)f(y)dy = f(C)(C - 1)$$

If $C > 1$ then $J'(0) > 0$, and $J(\lambda_1)$ increases with λ_1. For $\lambda_1 = 1/2$ we can obtain

$$\begin{aligned} J'(1/2) &= 4\int_0^{2C} f(2C - y)(C - y)f(y)dy \\ &= 2\int_0^{2C} f(2C - y)(2C - y)f(y)dy \\ &-2\int_0^{2C} f(2C - y)yf(y)dy = 0. \end{aligned}$$

Here we use the change of the variable $y = 2C - z$ in the second integral. For an arbitrary $\lambda_1 \in (0, 1/2]$ one obtains

$$J'(\lambda_1) = \int_0^C f\left(\frac{C-\lambda_1 y}{1-\lambda_1}\right)\frac{C-y}{(1-\lambda_1)^2} f(y)dy - \int_C^{C/\lambda_1} f\left(\frac{C-\lambda_1 y}{1-\lambda_1}\right)\frac{y-C}{(1-\lambda_1)^2} f(y)dy.$$

Both integrals in the last expression are positive. The second one tends to zero when $C \to \infty$. Thus

$$\lim_{C\to\infty} J'(\lambda_1) \geq 0 \quad \forall \lambda_1 \in [0, 1/2].$$

Hence using (40) one obtains

$$\mathcal{P}\left\{\lambda_1\xi_1^2 + (1-\lambda_1)\xi_2^2 < C\right\} \geq \mathcal{P}\left\{\xi_1^2 < C\right\}. \quad (41)$$

Let $\lambda_1 \leq \ldots \leq \lambda_n$, $\lambda_1 + \ldots + \lambda_n = 1$, hence $\lambda_1 \in [0, 1/n]$. We need to prove that for sufficiently large C

$$\mathcal{P}\left\{\lambda_1\xi_1^2 + \ldots + \lambda_n\xi_n^2 > C\right\} \geq \mathcal{P}\left\{\xi_1^2 < C\right\}. \quad (42)$$

Inequality (41) gives us this result for $n = 2$. Suppose that (42) holds for 2, 3, ..., $n-1$ summands. Then

$$\mathcal{P}\left\{\lambda_1\xi_1^2 + \ldots + \lambda_n\xi_n^2 > C\right\} = \int_0^{C/\lambda_1} \mathcal{P}\left\{\frac{\lambda_2\xi_2^2}{1-\lambda_1} + \ldots + \frac{\lambda_n\xi_n^2}{1-\lambda_1} > \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy.$$

Taking into account that $\lambda_2 + \ldots + \lambda_n = 1 - \lambda_1$, and using (42) for the case $(C - \lambda_1 y)(1 - \lambda_1) \geq C$ (i.e. $y < C$) for sufficiently large C, one obtains

$$\mathcal{P}\left\{\frac{\lambda_2\xi_2^2}{1-\lambda_1} + \ldots + \frac{\lambda_n\xi_n^2}{1-\lambda_1} > \frac{C-\lambda_1 y}{1-\lambda_1}\right\} \geq \mathcal{P}\left\{\xi_2^2 > \frac{C-\lambda_1 y}{1-\lambda_1}\right\}.$$

Thus

$$\begin{aligned}
&\mathcal{P}\left\{\lambda_1\xi_1^2 + \ldots + \lambda_n\xi_n^2 > C\right\} \\
&\geq \int_0^C \mathcal{P}\left\{\xi_2^2 < \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\
&+ \int_C^{C/\lambda_1} \mathcal{P}\left\{\frac{\lambda_2\xi_2^2}{1-\lambda_1} + \ldots + \frac{\lambda_n\xi_n^2}{1-\lambda_1} > \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\
&= \int_0^{C/\lambda_1} \mathcal{P}\left\{\xi_2^2 < \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\
&- \int_C^{C/\lambda_1} \mathcal{P}\left\{\xi_2^2 < \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\
&+ \int_C^{C/\lambda_1} \mathcal{P}\left\{\frac{\lambda_2\xi_2^2}{1-\lambda_1} + \ldots + \frac{\lambda_n\xi_n^2}{1-\lambda_1} > \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy
\end{aligned}$$

The last two integrals tend to zero as $C \to \infty$. So one obtains

$$\begin{aligned}
&\lim_{C\to\infty} \mathcal{P}\left\{\lambda_1\xi_1^2 + \ldots + \lambda_n\xi_n^2 > C\right\} \\
&\geq \lim_{C\to\infty} \int_0^{C/\lambda_1} \mathcal{P}\left\{\xi_2^2 < \frac{C-\lambda_1 y}{1-\lambda_1}\right\} f(y)dy \\
&= \lim_{C\to\infty} \mathcal{P}\left\{\lambda_1\xi_1^2 + (1-\lambda_1)\xi_2^2 > C\right\}
\end{aligned}$$

This and (41) imply (39). The Lemma has been proved.

Theorem 2. If process (1) is ergodic, and the sample volume N to construct the compensating factor Γ_N satisfies the following conditions

$$N \to \infty, \quad N/H \to 0 \text{ as } H \to \infty,$$

then for sufficiently large H

$$\mathcal{P}\left\{\left|\left|\hat{\Lambda}^j - \Lambda^j\right|\right|^2 > x\right\} \leq 2\left(1 - \Phi\left(\sqrt{\frac{xH^2}{H+p-1}}\right)\right) \quad (43)$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Proof. We consider estimator (29). According to (38),

$$\left|\left|\hat{\Lambda}^j - \Lambda^j\right|\right|^2 \leq \frac{||\eta(j,\tau^j)||^2}{H^2},$$

i.e.

$$\left|\left|\hat{\Lambda}^j - \Lambda^j\right|\right|^2 \leq \frac{1}{H^2}\left|\left|\sum_{k=N+1}^{\tau^j} \upsilon_{j,k}(Y_k^j)^T\gamma_k\xi_k\right|\right|^2. \quad (44)$$

Denote $Z = [z_1, \ldots, z_p]$, $||Z|| = 1$, and consider a linear combination of the components of the vector from the last equation

$$X(\tau^j) = \frac{1}{\sqrt{P(H)}}\sum_{k=N+1}^{\tau^j} \upsilon_{j,k}Z(Y_k^j)^T\gamma_k\xi_k,$$

where $P(H) = H + p - 1$.

Further we find the limit distribution of $X(\tau^j)$ along the lines of the proof of the martingale central limit theorem (see [23]). Let us calculate the characteristic function of X_τ. Denote

$$\begin{aligned}
\varepsilon_k = \varepsilon_k(H) &= \frac{1}{\sqrt{P(H)}}\upsilon_{j,k}Z(Y_k^j)^T\gamma_k\xi_k\chi_{[\tau^j \geq k]}, \\
X(n) &= \sum_{k=N+1}^{n} \varepsilon_k.
\end{aligned} \quad (45)$$

It is evident that under the assumptions of Theorem 1 as $n \to \infty$

$$|X(\tau^j) - X(n)| \to^{\mathcal{P}} 0.$$

Thus, in order to find the characteristic function of $X(\tau^j)$, one needs to find the limit of the characteristic function of $X(n)$. Denote

$$\mathcal{E}^n(\eta) = \prod_{k=N+1}^{n} E\left(e^{i\eta\varepsilon_k}\middle|\mathcal{F}_{k-1}\right),$$

Lemma ([23]). *If (for given η) $|\mathcal{E}^n(\eta)| \geq c(\eta) > 0$, $n > 1$, then convergence in probability $\mathcal{E}^n(\eta) \to E\left(e^{i\eta X}\right)$ is sufficient for convergence $E\left(e^{i\eta X(n)}\right) \to E\left(e^{i\eta X}\right)$.*

Check the lemma conditions for the process TAR(p)/ARCH(1)

$$\begin{aligned}
|\mathcal{E}^n(\eta)| &= \prod_{k=N+1}^{n}\left|E\left[e^{i\eta\varepsilon_k}\middle|\mathcal{F}_{k-1}\right]\right| \\
&= \prod_{k=N+1}^{n}\left|1 + E\left[e^{i\eta\varepsilon_k} - 1 - i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right|.
\end{aligned}$$

By using the inequality $\left|e^{i\eta x}-1-i\eta x\right| \leq (\eta x)^2/2$, we obtain

$$\begin{aligned}|\mathcal{E}^n(\eta)| &\geq \prod_{k=N+1}^{n} \left(1-E\left[\left|e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\right|\middle|\mathcal{F}_{k-1}\right]\right)\\ &\geq \prod_{k=N+1}^{n}\left(1-\frac{1}{2}E\left[(\eta\varepsilon_k)^2\middle|\mathcal{F}_{k-1}\right]\right)\\ &= \prod_{k=N+1}^{n}\left(1-\frac{(\eta v_{j,k}Z(Y_k^j)^T\gamma_k)^2\chi_{[\tau^j\geq k]}}{2P(H)}E\left[\xi_k^2\middle|\mathcal{F}_{k-1}\right]\right)\\ &= \exp\left\{\sum_{k=N+1}^{n}\ln\left(1-\frac{(\eta v_{j,k}Z(Y_k^j)^T\gamma_k)^2\chi_{[\tau^j\geq k]}}{2P(H)}\right)\right\}.\end{aligned}$$

Model (22) implies $(v_{j,k}Z(Y_k^j)^T\gamma_k)^2 \leq ||Z||^2p(\omega+\alpha^2)$; hence, as $H\to\infty$

$$\frac{(v_{j,k}Z(Y_k^j)^T\gamma_k)^2}{P(H)} \to 0.$$

By using the inequality $\ln(1-x)\geq -2x$, where $x\in(0,1/2]$, for any $H\geq H_0(\eta)$, and taking into account that $\gamma_k^2\leq\omega+\alpha^2$ one obtains

$$\begin{aligned}|\mathcal{E}^n(\eta)| &\geq \exp\left\{-\sum_{k=N+1}^{\min(n,\tau^j)}\frac{(\eta v_{j,k}Z(Y_k^j)^T\gamma_k)^2}{P(H)}\right\}\\ &\geq \exp\left\{-\frac{\eta^2(\omega+\alpha^2)}{P(H)}\sum_{k=N+1}^{\tau}\left(v_{j,k}Z(Y_k^j)^T\right)^2\right\}.\end{aligned}$$

Taking into account (31) and (33), we obtain

$$\begin{aligned}|\mathcal{E}^n(\eta)| &\geq \exp\left\{-\frac{\eta^2(\omega+\alpha^2)}{P(H)}\frac{||Z||^2P(H)}{\Gamma_N}\right\}\\ &= \exp\left\{-\frac{\eta^2(\omega+\alpha^2)}{\Gamma_N}\right\}.\end{aligned}$$

If the process (1) is ergodic, then the random variable $1/\Gamma_N$ tends to its expectation as $N\to\infty$. Consequently, in the conditions of the theorem and taking into account (24), for sufficiently large H we obtain $(\omega+\alpha^2)/\Gamma_N\leq 1$, and

$$|\mathcal{E}^n(\eta)| \geq \exp\left\{-\eta^2\right\} > 0.$$

The lemma conditions hold true.

Further we investigate an asymptotic behavior of $\mathcal{E}^n(\eta)$. Write this function in the form

$$\begin{aligned}&\mathcal{E}^n(\eta)\\ &= \exp\left\{\sum_{k=N+1}^{n}E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right\}\\ &\times\exp\left\{-\sum_{k=N+1}^{n}E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right\}\\ &\times\prod_{k=N+1}^{n}\left(1+E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right).\end{aligned} \tag{46}$$

Then we show that the product of the last two factors tends to 1. Denote $\alpha_k = E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]$. Using the inequality $|e^x-1|\leq e^{|x|}|x|$, we have

$$\begin{aligned}&\left|\prod_{k=N+1}^{n}(1+\alpha_k)e^{-\alpha_k}-1\right|\\ &= \left|\exp\left\{\ln\prod_{k=N+1}^{\tau}(1+\alpha_k)e^{-\alpha_k}\right\}-1\right|\\ &\leq \exp\left\{\left|\ln\prod_{k=N+1}^{n}(1+\alpha_k)e^{-\alpha_k}\right|\right\}\\ &\left|\ln\prod_{k=N+1}^{n}(1+\alpha_k)e^{-\alpha_k}\right|.\end{aligned}$$

Taking into account the inequalities $|\ln(1+x)-x|\leq 2|x|^2$ for $|x|<1/2$ and $\left|e^{i\eta x}-1-i\eta x\right|\leq(\eta x)^2/2$, as $H>H_0(\eta)$, we have

$$\begin{aligned}&\left|\ln\prod_{k=N+1}^{n}(1+\alpha_k)e^{-\alpha_k}\right| \leq \sum_{k=N+1}^{n}\left|\ln(1+\alpha_k)-\alpha_k\right|\\ &\leq 2\sum_{k=N+1}^{n}|\alpha_k|^2 = 2\sum_{k=N+1}^{n}\left(E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right)^2\\ &\leq \frac{\eta^4}{P^2(H)}\sum_{k=N+1}^{\tau^j}\left(v_{j,k}Z(Y_k^j)^T\gamma_k\right)^4.\end{aligned}$$

Taking into account that

$$(v_{j,k}Z(Y_k^j)^T)^2\gamma_k^4 \leq ||Z||^2p(\omega+\alpha^2)^2,$$

with the usage of (31–33) one obtains

$$\begin{aligned}&\left|\ln\prod_{k=N+1}^{n}(1+\alpha_k)e^{-\alpha_k}\right|\\ &\leq \frac{\eta^4||Z||^2p(\omega+\alpha^2)^2}{P^2(H)}\sum_{k=N+1}^{\tau^j}(v_{j,k}Z(Y_k^j)^T)^2\\ &\leq \frac{\eta^4||Z||^4p(\omega+\alpha^2)^2}{P(H)\Gamma_N}\to 0.\end{aligned}$$

Thus the product of the last two multipliers in (46) tends to 1 in probability as $n\to\infty$, $H\to\infty$.

Consider the first multiplier

$$\begin{aligned}&\exp\left\{\sum_{k=N+1}^{n}E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k\middle|\mathcal{F}_{k-1}\right]\right\}\\ &= \exp\left\{-\frac{1}{2}\sum_{k=N+1}^{n}E\left[(\eta\varepsilon_k)^2\middle|\mathcal{F}_{k-1}\right]\right\}\\ &\times\exp\left\{\sum_{k=N+1}^{n}E\left[e^{i\eta\varepsilon_k}-1-i\eta\varepsilon_k+\frac{(\eta\varepsilon_k)^2}{2}\middle|\mathcal{F}_{k-1}\right]\right\}.\end{aligned}$$

Let us prove that the second multiplier in this equation tends to 1. By using the inequality

$$\left|e^{i\eta x}-1-i\eta x+(\eta x)^2/2\right| \leq |\eta x|^3/6$$

and (33), one can rewrite it as

$$\begin{aligned}&\left|\sum_{k=N+1}^{n}E\left[e^{i\eta\varepsilon_t}-1-i\eta\varepsilon_k+\frac{(\eta\varepsilon_k)^2}{2}\middle|\mathcal{F}_{k-1}\right]\right|\\ &\leq \frac{1}{6P^{3/2}(H)}\\ &\times\sum_{k=N+1}^{n}E\left[\left|\eta v_{j,k}Z(Y_k^j)^T\gamma_k\xi_k\right|^3\chi_{[\tau^j\geq k]}\middle|\mathcal{F}_{k-1}\right]\\ &= \frac{B^3|\eta|^3E|\xi_k|^3}{6P^{3/2}(H)}\sum_{k=N+1}^{\tau}\left|v_{j,k}Z(Y_k^j)^T\gamma_k\right|^3\chi_{[\tau^j\geq k]}.\end{aligned}$$

By using (31)–(33), and $\left|v_{j,k}Z(Y_k^j)^T\gamma_k\right| \leq \sqrt{p(\omega+\alpha^2)}$, one obtains that the last expression tends to 0. So the second multiplier in the previous expression tends to 1 as $H\to\infty$. Consider the first multiplier

$$\exp\left\{-\frac{1}{2}\sum_{k=N+1}^{n} E\left[(\eta\varepsilon_k)^2\middle|\mathcal{F}_k\right]\right\}$$
$$=\exp\left\{-\frac{\eta^2}{2P(H)}\sum_{k=N+1}^{\min(n,\tau^j)}(v_{j,k}Z(Y_k^j)^T\gamma_k)^2\right\}$$
$$=\exp\left\{-\frac{\eta^2}{2}\langle X_n\rangle\right\}.$$

Note that according to (22) and (31)–(33), $\langle X_n\rangle$ is a bounded submartingale. Thus, the limit $\langle X_\infty\rangle=\lim_{n\to\infty}\langle X_n\rangle$ exists almost surely, and $\langle X_\infty\rangle \leq (\omega+\alpha^2)/\Gamma_N$. On the other hand, $\langle X_n\rangle\to\langle X_\tau\rangle$ as $n\to\infty$. So the distribution X_τ is asymptotically normal. Thus, the random vector

$$Y=\frac{1}{\sqrt{P(H)}}\sum_{k=N+1}^{\tau^j} v_{j,k}(Y_k^j)^T\gamma_k\xi_k \tag{47}$$

is asymptotically normal with the parameters $(0,\Sigma)$, where the covariance matrix

$$\Sigma=E\frac{1}{P(H)}\sum_{k=N+1}^{\tau^j} v_{j,k}^2(Y_k^j)^T Y_k^j\gamma_k^2 \tag{48}$$

possess the following property

$$\begin{aligned}\mathrm{tr}\Sigma&=E\frac{1}{P(H)}\sum_{k=N+1}^{\tau^j} v_{j,k}^2 Y_k^j(Y_k^j)^T\gamma_k^2\\ &\leq E\frac{\omega+\alpha^2}{\Gamma_N}\leq 1,\end{aligned} \tag{49}$$

which can be proved by using (31)–(33).

Now turn to estimation the probability (43). By using (44) and (47), one obtains

$$\mathcal{P}\left\{\left|\left|\hat\Lambda^j-\Lambda^j\right|\right|^2>x\right\}\leq\mathcal{P}\left\{\frac{P(H)}{H^2}\,||Y||^2>x\right\}.$$

Using the Fubini's theorem to change the order of integration and denoting $xH^2/P(H)$ as C_H, one obtains

$$\mathcal{P}\left\{||Y||^2>\frac{xH^2}{P(H)}\right\}$$
$$=\int\limits_{YY^T>C_H}\int\limits_{-\infty}^{\infty}\frac{\exp\left\{-i\lambda^TY\right\}}{2\pi}E\exp\left\{-\frac{1}{2}\lambda^T\Sigma\lambda\right\}d\lambda dY$$
$$=E\int\limits_{YY^T>C_H}\int\limits_{-\infty}^{\infty}\frac{\exp\left\{-i\lambda^TY\right\}}{2\pi}\exp\left\{-\frac{1}{2}\lambda^T\Sigma\lambda\right\}d\lambda dY$$
$$=E\frac{1}{\sqrt{(2\pi)^m|\Sigma|}}\int\limits_{YY^T>xH}\exp\left\{-\frac{1}{2}Y\Sigma^{-1}Y^T\right\}dY.$$

The matrix Σ is symmetric and positive definite, hence an orthogonal transformation T, resulting in a matrix Σ to diagonal form Σ', exists. Thus, $T\Sigma T^T=\Sigma'$, $TT^T=T^TT=I$, where I is the identity matrix. Using the change of variables $S=Y\Sigma^{-1/2}T^T$, one obtains

$$\mathcal{P}\left\{||Y||^2>C_H\right\}$$
$$=E\frac{1}{\sqrt{(2\pi)^m}}\int\limits_{\sum_{j=1}^m\nu_js_j^2>C_H}\exp\left\{-\frac{1}{2}SS^T\right\}dS=$$
$$E\mathcal{P}\left\{\sum_{j=1}^m\nu_js_j^2>C_H\right\},$$

where $\nu_1,\ldots,\nu_m$ are the eigenvalues of the matrix Σ', and $s_1,\ldots,s_m$ are the independent components of the Gaussian vector S. Using inequality (49), one obtains

$$\sum_{j=1}^m\nu_j=\mathrm{tr}\Sigma'=\mathrm{tr}\Sigma\leq 1.$$

This and Lemma 1 imply (43). The Theorem has been proved.

V. Change Point Detection Procedure

Let us consider now the change point detection problem for process (1). At the first stage, we define intervals $[\tau_{n-1}^j+1,\tau_n^j]$, $n\geq 1$. The estimators $\hat\Lambda_n^j$ of the parameters of process (1) are constructed on each interval. Then the estimators on intervals $[\tau_{n-l-1}^j+1,\tau_{n-l}^j]$ and $[\tau_{n-1}^j+1,\tau_n^j]$, where $l>1$ is an integer, are compared. If the interval $[\tau_{n-l-1}^j+1,\tau_n^j]$ does not include the change point θ, then the vector Λ^j on this interval is constant. It can be equal to the initial value μ^j or the final value β^j. Thus for certain n, if $\tau_{n-l}^j<\theta<\tau_{n-1}^j+1$, the difference between values of the parameters on intervals $[\tau_{n-l-1}^j+1,\tau_{n-l}^j]$ and $[\tau_{n-1}^j+1,\tau_n^j]$ is no less then Δ. This is the key property for the change point detection.

We construct a set of sequential estimation plans

$$(\tau_n^j,\hat\Lambda_n^j)=(\tau_n^j(H),\hat\Lambda_n^j(H)),\quad n\geq 1,\quad j=1,2,$$

where $\{\tau_n^j\},\ n\geq 0$ is the increasing sequence of the stopping instances ($\tau_0=N$), and $\hat\Lambda_n^j$ is the guaranteed parameter estimator on the interval $[\tau_{n-1}^j+1,\tau_n^j]$. Then we choose an integer $l>1$ and define the statistics I_n^j

$$I_n^j=||\hat\Lambda_n^j-\hat\Lambda_{n-l}^j||^2. \tag{50}$$

This statistic is the squared deviation of the estimators with numbers n and $n-l$. Statistics properties are given in the following theorem.

Theorem 3. The expectation of the statistics I_n^j (50) satisfies the following inequality:

$$\begin{aligned}&E\left[I_n^j\middle|\tau_n^j<\theta\right]\leq\frac{4(H+p-1)}{H^2};\\ &E\left[I_n^j\middle|\tau_{n-l}^j<\theta\leq\tau_{n-1}^j\right]\\ &\geq\Delta-4\sqrt{\Delta\frac{H+p-1}{H^2}}.\end{aligned} \tag{51}$$

Proof. Denote the deviation of the estimator $\hat\Lambda_n^j$ from the true value of the parameter Λ^j as ζ_n^j. Let the parameter value remain unchanged until the instant τ_n^j, i.e., $\theta>\tau_n^j$. In this case, $\hat\Lambda_n^j=\mu^j+\zeta_n^j$, $\hat\lambda_{n-l}^j=\mu^j+\zeta_{n-l}^j$ and statistic (50) can be written in the form

$$I_n^j=\left|\left|\zeta_n^j-\zeta_{n-l}^j\right|\right|^2.$$

According to Theorem 1,

$$E||\zeta_n^j||^2 \leq \frac{H+p-1}{H^2} \quad (52)$$

To estimate the expectation of the statistic, we use property (52) and the inequality $||a-b||^2 \leq 2||a||^2 + 2||b||^2$:

$$EI_n^j \leq E\left(2||\zeta_n^j||^2 + 2||\zeta_{n-l}^j||^2\right) \leq 4\frac{H+p-1}{H^2}. \quad (53)$$

Let the change of the parameter take place on the interval $[\tau_{n-l}^j, \tau_{n-1}^j]$ i.e. $\tau_{n-l}^j < \theta \leq \tau_{n-1}^j$. In this case, $\hat{\Lambda}_n^j = \beta^j + \zeta_n^j$, $\hat{\Lambda}_{n-l}^j = \mu^j + \zeta_{n-l}^j$, and statistic (50) is

$$I_n^j = ||\beta^j - \mu^j + \zeta_n^j - \zeta_{n-l}^j||^2.$$

To estimate the expectation of the statistics, we take advantage of the inequality $||a-b|| \geq ||a|| - ||b||$ and condition (52)

$$\begin{aligned} EI_n^j &\geq E\left(||\beta^j - \mu^j|| - ||\zeta_n^j - \zeta_{n-l}^j||\right)^2 \\ &\geq ||\beta^j - \mu^j||^2 - 2||\beta^j - \mu^j||\, E||\zeta_n^j - \zeta_{n-l}^j|| \\ &\geq \Delta - 4\sqrt{\Delta\frac{H+p-1}{H^2}}. \end{aligned}$$

The theorem has been proved.

Hence, the change of the expectation of the statistic I_n^j allows us to construct the following change point detection algorithm. The I_n^j values are compared with a certain threshold δ, where

$$\frac{4(H+p-1)}{H^2} < \delta < \Delta - 4\sqrt{\Delta\frac{H+p-1}{H^2}}. \quad (54)$$

When the value of the statistic exceeds δ then the change point is considered to be detected. If at least one parameter of the vector $\Lambda = [\Lambda_0, \Lambda_1]$ changes, then the change point θ can be detected.

The probabilities of false alarm and delay in the change point detection in any observation cycle are important characteristics of any change point detection procedure. Due to the application of the guaranteed parameter estimators in the statistics, we can bound these probabilities from above.

Theorem 4. The probability of false alarm $P_{0,n}$ and the probability of delay $P_{1,n}$ in n-th observation cycle $[\tau_{n-1}^j + 1, \tau_n^j]$ are bounded from above

$$\begin{aligned} P_{0,n} &\leq \frac{2(H+p-1)}{\delta H^2}; \\ P_{1,n} &\leq \frac{2(H+p-1)}{(\sqrt{\Delta}-\sqrt{\delta})^2 H^2}. \end{aligned} \quad (55)$$

Proof. First, we consider the false alarm probability, i.e. the probability that the statistic J_i exceeds the threshold before the change point. Using the norm properties and the Chebyshev inequality, we obtain

$$\begin{aligned} P_{0,n} &= \mathcal{P}\left\{ I_n^j > \delta \middle|\, \tau_n^j < \theta \right\} \\ &= \mathcal{P}\left\{||\zeta_n^j - \zeta_{n-l}^j||^2 > \delta\right\} \leq \frac{2E\left(||\zeta_n^j||^2 + ||\zeta_{n-l}^j||^2\right)}{\delta}. \end{aligned}$$

This and (52) imply the first inequality from (55).

Then we consider the delay probability, i.e., the probability that the statistic I_n^j does not exceed the threshold after the change point

$$\begin{aligned} P_{1,n} &= \mathcal{P}\left\{ I_n^j < \delta \middle|\, \tau_{n-l}^j < \theta < \tau_{n-1}^j \right\} \\ &= \mathcal{P}\left\{||\beta^j - \mu^j + \zeta_n^j - \zeta_{n-l}^j||^2 < \delta\right\} \\ &= \mathcal{P}\left\{||\zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta}\right\}. \end{aligned}$$

Taking into account that $||\beta^j - \mu^j||^2 > \Delta$ and using the norm properties and the Chebyshev inequality, one obtains

$$\begin{aligned} P_{1,n} &\leq \mathcal{P}\left\{||\beta^j - \mu^j|| - ||\zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta}\right\} \\ &\leq \mathcal{P}\left\{\sqrt{\Delta} - ||\zeta_n^j - \zeta_{n-l}^j|| < \sqrt{\delta}\right\} \\ &= \mathcal{P}\left\{||\zeta_n^j - \zeta_{n-l}^j|| > \sqrt{\Delta} - \sqrt{\delta}\right\} \\ &\leq \frac{2E\left(||\zeta_n^j||^2 + ||\zeta_{n-l}^j||^2\right)}{(\sqrt{\Delta}-\sqrt{\delta})^2}. \end{aligned}$$

This and (52) imply the second inequality from (55).

The theorem has been proved.

Then we consider asymptotic properties of the proposed change point detection procedure for $H \to \infty$ if process (1) is ergodic, i.e. the asymptotic inequalities for the probabilities of false alarm and delay.

Theorem 5. If process (1) is ergodic, and the compensating factor Γ_N satisfies the following conditions $N \to \infty$, $N/H \to 0$ as $H \to \infty$, then for sufficiently large H the probabilities of false alarm and delay in n-th observation cycle $[\tau_{n-1}^j + 1, \tau_n^j]$ are bounded from above

$$\begin{aligned} P_{0,n} &\leq 2\left(1 - \Phi\left(\sqrt{\frac{\delta H^2}{2(H+p-1)}}\right)\right); \\ P_{1,n} &\leq 2\left(1 - \Phi\left(\sqrt{\frac{\left(\sqrt{\Delta}-\sqrt{\delta}\right)^2 H^2}{2(H+p-1)}}\right)\right) \end{aligned} \quad (56)$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Proof. First, we consider the false alarm probability. Along the lines of the proof of Theorem 4, we obtain

$$P_{0,n} = \mathcal{P}\left\{||\zeta_n^j - \zeta_{n-l}^j||^2 > \delta\right\}.$$

Note that ζ_n^j is the difference between the estimator $\hat{\Lambda}_n^j$ and the true value of the parameter Λ^j; hence, the vector

$$Z_n^j = \frac{1}{\sqrt{2P(H)}}(\zeta_n^j - \zeta_{n-l}^j)$$

has the same properties that the vector Y (47) and, according to (43),

$$\begin{aligned} &\mathcal{P}\left\{\left|\left|\zeta_n^j - \zeta_{n-l}^j\right|\right|^2 > x\right\} \\ &\leq 2\left(1 - \Phi\left(\sqrt{\frac{xH^2}{2(H+p-1)}}\right)\right) \end{aligned} \quad (57)$$

This implies the first inequality from (56).

Then we consider the delay probability. Along the lines of the proof of Theorem 4, we obtain

$$P_{1,n} \leq \mathcal{P}\left\{||\zeta_n^j - \zeta_{n-l}^j|| > \sqrt{\Delta} - \sqrt{\delta}\right\}.$$

This and (57) imply the second inequality from (56).

The theorem has been proved.

These estimators can be used instead of (55) for sufficiently large H.

TABLE I
PARAMETER ESTIMATION FOR THE TAR(2)/ARCH(1) PROCESS

H	$\hat{\mu}_1^1$	$\hat{\mu}_2^1$	$\hat{\mu}_2^1$	$\hat{\mu}_2^2$	T	d	D
50	0.501	0.303	0.303	0.492	724	0.009	0.02
100	0.506	0.304	0.293	0.497	1444	0.0046	0.01
200	0.498	0.295	0.304	0.497	2686	0.0028	0.005
250	0.503	0.301	0.300	0.500	13811	0.0005	0.001

VI. SIMULATION RESULTS AND THEIR DISCUSSION

This section presents the simulation results for the described algorithms. For every experiment 100 replications were conducted.

First, we considered the parameter estimation problem for the TAR(2)/ARCH(1) process (1) with the parameters

$$\Lambda^1 = [\ 0.5, \quad 0.1\ ], \quad \Lambda^2 = [\ 0.1, \quad 0.3\ ];$$
$$\omega = 0.6, \quad \alpha^2 = 0.4,$$

and with the standard Gaussian noise. The compensating factors Γ_n was constructed with additional conditions $x_{k-1}^2 > 0.1$ and $x_k^2/\max\{1, x_{k-1}^2\} < 5$. We add these conditions in order to bound from above the compensating factor and to avoid increasing the estimation interval. The noise variance of the process in the special form (22) is bounded from above by the value $0.6 + 0.4 = 1$. The number n was chosen as the integral part of $\sqrt{H}$.

Table I presents the simulation results. Here H is the parameter of the procedure, $\hat{\mu}_1^1$ and $\hat{\mu}_2^1$ are the mean estimators of the corresponding parameters $\mu_1^1 = 0.5$ and $\mu_2^1 = 0.3$ $\hat{\mu}_2^1$ and $\hat{\mu}_2^2$ are the mean estimators of the corresponding parameters $\mu_2^1 = 0.3$ and $\mu_2^2 = 0.5$ calculated by 100 replications, T is the mean number of observations used to calculate the estimator, $d = ||\hat{\Lambda} - \Lambda||^2$ averaged over 100 realizations, D is the theoretical upper bound for the mean square accuracy of the estimator given by inequality (34).

One can see that the mean number of the observation increases linearly by H. This property is important for sequential estimators (see [24]) because it characterizes the optimality of the procedure in the case of independent observation.

The sample mean square error of the estimation is about four times less then the theoretical one. It is connected with the complicated structure of the TAR/ARCH process. It has the unbounded noise variation, so we divide the equation by the number not less then unity. As a result $\nu_{\min}(j, m)$ in (30) grows slowly and the estimation interval increases. Besides, the compensating factor exceeds the real upper bound of the noise variance about two times. It implies decreasing of the mean square error and increasing of the mean estimation interval in the same proportion.

Fig. 3 – Fig. 6 demonstrate examples of the sequences of TAR(2)/ARCH(1) parameters estimates for $H = 100$. Here solid lines indicate true values of the parameters, and dotted lines shows the behavior of estimates. Every time unity corresponds to 10000 observations.

Further we conducted simulations of the proposed change point detection procedure. The simulations were conducted for the TAR(2)/ARCH(1) process. Before the instant θ it was specified by the equation (1) with the parameters

$$\Lambda^1 = [\ 0.5, \quad 0.3\ ], \quad \Lambda^2 = [\ 0.3, \quad 0.5\ ];$$
$$\omega = 0.4, \quad \alpha = 0.1,$$

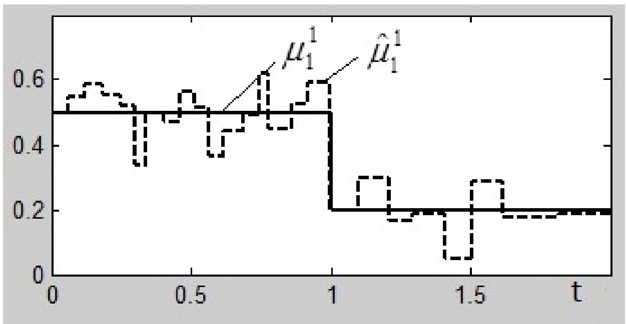

Fig. 3. Estimator for μ_1^1.

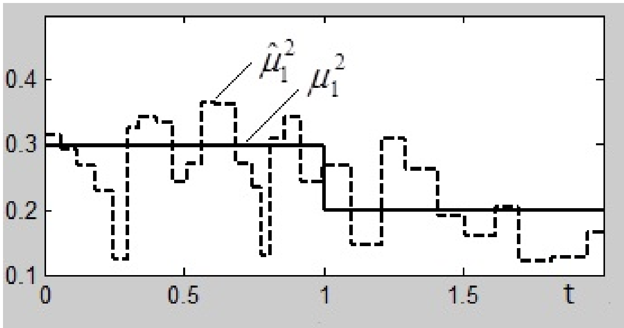

Fig. 4. Estimator for μ_1^2.

After the instant $\theta = 10000$ he parameters are

$$\Lambda^1 = [\ 0.2, \quad 0.2\ ], \quad \Lambda^2 = [\ 0.8, \quad 0.1\ ];$$
$$\omega = 0.4, \quad \alpha^2 = 0.6,$$

In this process in form (22) the noise variance is bounded from above by unity both before and after the change point. The change point $\theta = 10000$ and $\delta = 0.025$. Note that we choose the change point as a rather big number in order to have possibility to estimate the mean number of observation between false alarm using a sufficient sample size.

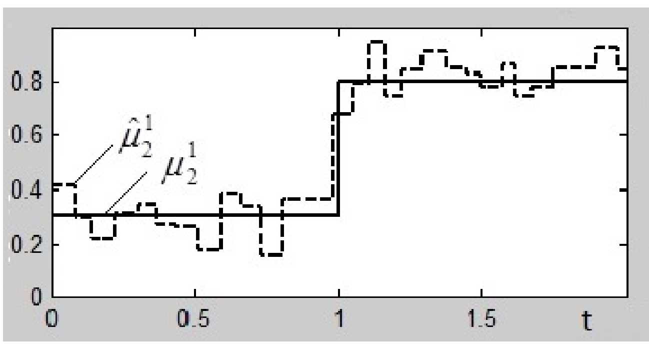

Fig. 5. Estimator for μ_2^1.

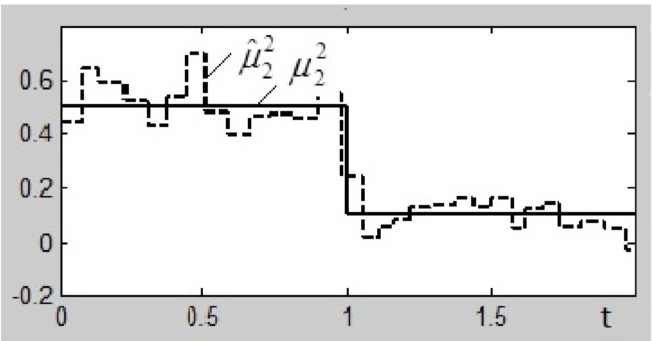

Fig. 6. Estimator for μ_2^2.

TABLE II
CHANGE-POINT DETECTION FOR THE TAR(2)/ARCH(1) PROCESS

H	T_0	T_1	$\hat{p}_0$	P_0a	P_0	$\hat{p}_1$	P_1a	P_1
100	10901	1237	0.63	0.46	0.81	0.0	0.46	0.81
200	196750	2242	0.05	0.22	0.4	0.0	0.22	0.4
400	1000000	5151	0.0	0.05	0.2	0.0	0.05	0.2

Table II presents the results of the simulations. Here H and δ are the parameters of the procedure, T_1 is the mean delay in the change point detection, T_0 is the mean interval between false alarms (if the cell is empty then there were no false alarms), $\hat{p}_0$ and $\hat{p}_1$ are the sample probabilities of the false alarms and of the delay, respectively, P_0a, P_1a and P_0, P_1 are the asymptotic and non-asymptotic upper bounds for the probabilities expressed by formulas (55) and (56).

One can see that when the parameter H increases then the mean delay increases too. It is explained by the fact that increase of H leads to rise of the number of observation which are necessary to construct more accurate parameter estimator. On the other hand, the difference between the estimator and the exact parameter becomes less and hence, the error probabilities also decrease. When parameter $H = 100$ one can see, that the asymptotic upper bound for probability of false alarm does not work well, this arises from complexity of the model and the estimation procedure needs bigger value of H to construct more accurate estimations of the parameters.

When the parameter H increases then the mean interval between false alarms increases too, what leads to the decrease of probabilities of delay. Actually, in this experiment, the sample probability of delay is equal to 0. It can be explained in connection with value of T_1. Hence, we have zero sample probability of delay, but also we always have mean delay, which grows linearly with H.

VII. CONCLUSION

The results in this paper were derived with strong mathematical evidence and are theoretical. Besides, the efficiency of the algorithms is checked via simulation. It can be very interesting to test them on the real data.

REFERENCES

[1] Tong, H. "On the Threshold Model", in Chen. C.H.(Ed.), Pattern Recognition and Signal Processing. Sijthoff and Noordhoff, Amsterdam.1978.

[2] Meitz, M., Saikkonen, P. "A note on the geometric ergodicity of nonlinear AR-ARCH model," Statistics and Probability Letters, vol.80,pp. 631-638, 2010.

[3] Engle, R.F. "Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kindom inflation," Econometrica, vol.50,pp.987-1008,1982.

[4] Sidorov, S.P., Revutskiy, A., Faizliev, A., Korobov, E., Balash, V. "Stock Volatility Modelling with Augmented GARCH Model with Jumps," IAENG International Journal of Applied Mathematics, vol.44, no.4, pp.212–220, 2014

[5] Guo, Z.-F., Cao, L. "An Asymmetric Smooth Transition GARCH Model," IAENG International Journal of Applied Mathematics, vol.41, no.4, pp.349–351, 2011.

[6] Wald, A. "Sequential analysis," New York, John Wiley & Sons, 1947.

[7] Lee, S. T.N.Sriram, T.N. "Sequential point estimation of parameters in a threshold AR(1) model," Stochastic Processes and their Applications, vol.84, pp.343–355, 1999.

[8] Sriram, T.N. "Fixed size confidence regions for parameters of threshold AR(1) models," Journal of Statistical Planning and Inference, vol.97, pp.293–304, 2001.

[9] Konev, V.V., Galtchouk, L.I. "On asymptotic normality of sequential LS-estimate for unstable autoregressive process AR(2)," Journal of Multivariate Analysis, vol.101, pp.2616–2636, 2010.

[10] Burkatovskaya, Y.B., Sergeeva, E.E., Vorobeychikov, S.E. "On Guaranteed Sequential Change Point Detection for TAR(1)/ARCH(1) process", Communications in Computer and Information Science, vol.487, pp.59–68, 2014.

[11] P. Fryzlewicz, S.Subba Rao, "Multipue-change-point detection for auto-regressive conditional heteroscedastic processes", Journal of the Royal Statistical Society.,Series B,Statistical Methodolog.,pp.1-22,2014.

[12] Na, O., Lee, J., Lee, S. "Change Point Detection in Copula ARMA-GARCH Models," Journal of Time Series Analysis, vol.33, no.4, pp.554–569 ,2012.

[13] Chen, G. Chen, M. "On a class of nonlinear AR(p) models with nonlinear ARCH errors," Australian and New Zealand Journal of Statistics, vol.43, pp.445–454, 2001.

[14] Lu, Z., Jiang, Z. "L_1 geometric ergodicity of multivariate nonlinear AR model with an ARCH term," Statistic & Probability Letters, vol. 51, pp.121–130, 2001.

[15] Cline, D.B.H. "Evaluating the Lyapounov exponent and existence of moments for threshold AR-ARCH models," Journal of Time Series Analysis, vol.28, pp.241–260, 2007.

[16] Cline, D.B.H., Pu, H.H. "Stability and the Lyapounov exponent of threshold AR-ARCH models," Annals of Applied Probability, vol.14, pp.1920–1949, 2004.

[17] Liebscher, E. "Towards a unified approach for proving geometrical ergodicity and mixing properties of nonlinear autoregressive processes," Journal of Time Series Analysis,vol.26, pp.669-689, 2005.

[18] Lu, Z. "On the geometric ergodicity of a non-linear autoregressive model with an autoregressive conditional heteroscedastic term," Statistica Sinica, vol. 8, pp.1205–1217, 1998.

[19] Sergeeva, E.E., S. Vorobejchikov, S.E.. "An Efficient Algorithm for Detecting a Change Point of Autoregressive Parameters of AR(p)/ARCH(q) Process," In:Proceedings of the 11th International Conference of Pattern Recognition and Information Processing, 156–159 (2011).

[20] Vorobeychikov, S.E., Burkatovskaya, Y.B. "Change point detection of autoregressive process with unknown parameters," (2011) IFAC Proceedings Volumes (IFAC-PapersOnline), 18 (PART 1), pp. 13215-13220.

[21] Burkatovskaya, Y.B., Vorobeychikov, S.E., Sergeeva, E.E.: "TAR(p)/ARCH(1) Process: Guaranteed Parameter Estimation and Change-Point Detection," Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2016, IMECS 2016, 16-18 March, 2016, Hong Kong, pp 936–941.

[22] Mein, S., Tweedie, R.: "Markov Chains and Stochastic Stability," Springer Verlag, 1993.

[23] Shiryaev, A.N.: "Probability," Moscow: Nauka, Chief Editorial Staff of Physical-Mathematical Literature, (1980).

[24] Lorden, G.: "Procedures for reacting to a change in distribution," Annals. Math. Statist., No. 42, P. 1897–1971 (1971).