УДК 51-74 *05.11.13*

IS ANY TRENDLESS NOISE TRACK CAN SERVE AS A NEW SOURCE OF INFORMATION?

R.R. Nigmatullin, R.K. Sagdiev

Radioelectronics and Informative-Measurement Technics Department
Kazan National Research Technical University (KNRTU-KAI)
named after A.N. Tupolev
10, Karl Marx str., Kazan, 420111, Tatarstan, Russian Federation

Abstract. In this paper the authors want to prove that a trendless sequence (TLS) can be used as an additional source of information. This additional information can be extracted from random noise with the help of 3D-DGIs (discrete geometrical invariants) method that allows to reduce 3N random data points to 13 parameters composed from the combination of integer moments and their intercorrelations up to the fourth order inclusive. Actually, they form a «universal» 13-feature space for comparison of one random sequence with another one. Comparison of these parameters associated with different noise tracks allows to use this set of parameters for calibration and other purposes associated with «standard»/reference equipment. As an example, we treated the measured and nonfiltered TLS obtained from ELVIS II workbench (National Instrument Corporation). New method helps to find the differences between 4 types of the recorded TLS(s) samplings (forming the corresponding rectangle matrices) belonging to the chosen ADCs and choose the «best» one among them.

Keywords: Discrete Geometrical Invariants; 13-dimensional feature space; Trendless noise tracks; ADC(s) comparison.

Introduction

The idea that the chosen set of random fluctuations (after removing a possible trend and defined as a «noise») contains a «hidden» source of information is not a new one. Part of the researchers tries to extract an information based on some reasonable suppositions. In particular, the basics of the fluctuation noise spectroscopy (including also many useful references) is given in the book [1] and papers [2-3].

The approach based on the Mori-Zwanzig formalism was given in the papers [4-5]. Unfortunately, these approaches include some unjustified suppositions and contain some treatment errors. Therefore, they become *useless* in analysis of random sequences having different nature. These suppositions and their analysis are listed in paper [6]. The essential results were obtained from analysis of electrochemical noise [7-12], other type of ‘noises’ as understanding of the earthquakes phenomenon and their quantitative description, analysis of medical data is given in papers [3, 4, 13, 14]. Unfortunately, in spite of these promising attempts the general picture associated with analysis of arbitrary types of random sequences (especially TLS) is far from an ‘ideal’ one. Researches based on their own (and in many cases unjustified suppositions) try to process many types of random sequences and extract an «information» mixed with treatment/uncontrollable errors. Especially, they have problems with analysis of equipment «noise» (pure TLS) that in many cases is not the Gaussian, uniform and other type of «color» noises that are widely used in the mathematical statistics as an approximation of the real TLS(s).

Therefore, there is a vital task in creation of the reliable processing tool that: (a) should be free from the treatment errors and (b) rather «universal» for application to any TLS that is generated or contained in the given equipment.

In this paper, we propose for a wide group of researchers the desired and general tool based on the «modeless» reading of random sequences/fluctuations. The creation of this promising method is based on the ideas expressed by Yu. I. Babenko [15,16] associated with the generalization of the Pythagorean theorem. One of us (RRN) applied these ideas to random sequences and used the discrete geometrical invariants (DGI) in 2D space for differentiation of different types of olive oils [17] and equipment noise [18] reducing initially $2N$ random data points to 8 independent parameters representing a combination of integer moments and intercorrelations up to the fourth order inclusive. One can continue these ideas and consider the complete set of the DGI(s) in 3D space. Application of this method allows to transform the set of arbitrary chosen $3N$ data points, forming the complete fourth order form to 13 independent parameters forming the feature space of the corresponding dimension. Besides, these 13 parameters extracted from the considered random sequence represent themselves a specific «fingerprint» for observing the evolution of random sequence in time or against another external factor as concentration, electromagnet/acoustic field intensityetc., in 3D space. This reduction procedure reminds a procedure used in the statistical mechanics when with the help of the Gibbs partition function $3N$ trajectories of the microscopic particles are reduced to a finite set of thermodynamic parameters. Actually, the 3D-DGI(s) method realizes a similar procedure, i.e., it reduces $3N$ combination of an arbitrary random data points to 13 quantitative parameters forming the specific feature space. We want to stress here that this reduction procedure does not use *any* model and can be considered as the «modeless» method. Besides, it does not contain any treatment errors as well and keeps only the measurement/experimental errors.

The content of the paper is organized in five subsections and expressed clearly for reading of a potential and attentive reader.

1. Description of the 3D-DGI(s) method

In this section, we describe the mathematical details associated with the derivation of the *complete* DGI in 3D-space. We remind here that preliminary results based on the application of the *incomplete* DGI form of the fourth order in 3D space is outlined recently in [19]. Let us consider the power-law form of the fourth order:

$$\begin{gathered}
L_k^{(4)} = A_{40}^{(1,0)}\left(y_1 - r_{1k}\right)^4 + A_{40}^{(2,0)}\left(y_2 - r_{2k}\right)^4 + A_{40}^{(3,0)}\left(y_3 - r_{3k}\right)^4 - \\
-B_{22}^{(1,2)}\left(y_1 - r_{1k}\right)^2\cdot\left(y_2 - r_{2k}\right)^2 - B_{22}^{(1,3)}\left(y_1 - r_{1k}\right)^2\cdot\left(y_3 - r_{3k}\right)^2 - B_{22}^{(2,3)}\left(y_2 - r_{2k}\right)^2\cdot\left(y_3 - r_{3k}\right)^2 + \\
+C_{211}^{1(2,3)}\left(y_1 - r_{1k}\right)^2\cdot\left(y_2 - r_{2k}\right)\cdot\left(y_3 - r_{3k}\right) + C_{211}^{2(1,3)}\left(y_2 - r_{2k}\right)^2\cdot\left(y_1 - r_{1k}\right)\cdot\left(y_3 - r_{3k}\right) + \\
+C_{211}^{3(1,2)}\left(y_3 - r_{3k}\right)^2\cdot\left(y_1 - r_{1k}\right)\cdot\left(y_2 - r_{2k}\right) - \frac{1}{2}D_{31}^{(1,2)}\left(y_1 - r_{1k}\right)\cdot\left(y_2 - r_{2k}\right)\left[\left(y_1 - r_{1k}\right)^2 + \left(y_2 - r_{2k}\right)^2\right] - \\
-\frac{1}{2}D_{31}^{(1,3)}\left(y_1 - r_{1k}\right)\cdot\left(y_3 - r_{3k}\right)\left[\left(y_1 - r_{1k}\right)^2 + \left(y_3 - r_{3k}\right)^2\right] - \\
-\frac{1}{2}D_{31}^{(2,3)}\left(y_2 - r_{2k}\right)\cdot\left(y_3 - r_{3k}\right)\left[\left(y_2 - r_{2k}\right)^2 + \left(y_3 - r_{3k}\right)^2\right].
\end{gathered} \tag{1}$$

In expression (1) the upper indices define the combination of the variables y_α ($\alpha = 1, 2, 3$) fixing the location of an arbitrary point $M(y_1, y_2, y_3)$ in 3D-space, the low indices determine the values of the power-law exponents that correspond to the *complete* algebraic form of the fourth order. The choice of the sign's combination (±) before the constants in (1) will be explained below. Three random sequences are determined by the values $r_{\alpha k}$ ($\alpha = 1, 2, 3$; $k = 1, 2,\ldots, N$). Expression (1) represents itself the *complete* form of the fourth order that contains the combination of three variables associated with an arbitrary point $M(y_1, y_2, y_3)$ and three arbitrary sequences $r_{\alpha k}$. The desired DGI is obtained from the following requirement:

$$\frac{1}{N}\sum_{k=1}^{N} L_k^{(4)} = I_4, \tag{2}$$

In order to remove in expression (2) the cubic terms we introduce the variables:

$$Y_\alpha = y_\alpha - \langle r_\alpha \rangle,\ \langle r_\alpha \rangle = \frac{1}{N}\sum_{k=1}^{N} r_{\alpha k}\ , \tag{3}$$

and nullify the *linear* terms. This requirement helps us to separate the desired variables Ya from each other and keep only the terms of the second and fourth orders, correspondingly. In order to decrease the number of constants in (2) and derive the DGI *not* depending on some free constants one defines three key ratio constants $R^{(\alpha,\beta)}$, with combination $(\alpha, \beta) = (1, 2), (1, 3), (2, 3)$:

$$\begin{gathered} R^{(\alpha,\beta)} = \frac{B^{(\alpha,\beta)}}{A} = \frac{C^{\gamma(\alpha,\beta)}}{A} = \frac{D^{(\alpha,\beta)}}{A}; \\ A_{40}^{(\alpha)} = A_{40}^{(\beta)} = A_{40}^{(\gamma)} \equiv A,\ \alpha,\beta,\gamma = 1,2,3. \end{gathered} \tag{4}$$

It is convenient also to introduce the following notations for the integer moments and their intercorrelations and present them as:

$$\begin{gathered} Q_{\alpha^n\beta^m\gamma^l} = \frac{1}{N}\sum_{k=1}^{N}\left((\Delta r_{3k})^m (\Delta r_{2k})^n (\Delta r_{1k})^l\right) \equiv \left\langle (\Delta r_\alpha)^m (\Delta r_\beta)^n (\Delta r_\gamma)^l \right\rangle, \\ \alpha \ge \beta \ge \gamma,\ (\alpha,\ \beta,\ \gamma) = 1,\ 2,\ 3. \end{gathered} \tag{5}$$

In the result of the introduced notations (4) and (5), the system of linear equations for the finding of unknown ratios $R^{(\alpha,\beta)}$ from the nullification requirement of the entering linear terms accepts the form:

$$\begin{gathered} \left[2Q_{221} - Q_{332} + \frac{3}{2}Q_{211} + \frac{1}{2}Q_{222}\right]\cdot R^{(1,2)} + \\ + \left[2Q_{331} - Q_{322} + \frac{3}{2}Q_{311} + \frac{1}{2}Q_{333}\right]\cdot R^{(1,3)} - 2Q_{321}\cdot R^{(2,3)} = 4Q_{111}; \end{gathered}$$

$$\left[2Q_{211} - Q_{331} + \frac{3}{2}Q_{221} + \frac{1}{2}Q_{111}\right] \cdot R^{(1,2)} - 2Q_{321} \cdot R^{(1,3)} +$$
$$+\left[2Q_{332} - Q_{311} + \frac{3}{2}Q_{322} + \frac{1}{2}Q_{333}\right] \cdot R^{(2,3)} = 4Q_{222};$$

(6)

$$-2Q_{321} \cdot R^{(1,2)} + \left[2Q_{311} - Q_{221} + \frac{3}{2}Q_{331} + \frac{1}{2}Q_{111}\right] \cdot R^{(1,3)} +$$
$$+\left[2Q_{322} - Q_{211} + \frac{3}{2}Q_{332} + \frac{1}{2}Q_{222}\right] \cdot R^{(2,3)} = 4Q_{333}.$$

The linear system of equations helps to reduce 3 moments (Q_{333}, Q_{222}, Q_{111}) and 7 intercorrelations of the third order (Q_{332}, Q_{322}, Q_{221}, Q_{211}, Q_{331}, Q_{311}, Q_{321}) to calculation of three unknown ratios $R^{(\alpha,\beta)}$ only. We should notice also that the combination of the algebraic signs in (1) is chosen in that way for the keeping of the partial solution $R = 1$ of system (6a) in the case when all three random sequences $r_{\alpha k}$ are identical to each other, i.e. $r_{1k} = r_{2k} = r_{3k}$. It is natural to define it as the case of spherical symmetry. If only two sequences coincide with other (for example, $r_{1k} = r_{2k} \neq r_{3k}$) then we deal with the case of the cylindrical symmetry.In this case, the linear system (6a) is reduced to the couple of linear equations relatively the variables $R^{(1,2)} \neq R^{(1,3)} = R^{(2,3)}$. The number of triple correlations equals four in this case (Q_{111}, Q_{113}, Q_{133}, Q_{333}).

Equation (6) facilitate considerably the further calculations. After averaging procedure applied to expression (2) the structure of the fourth order form can be rewritten as:

$$K_4\left(Y_1, Y_2, Y_3\right) + K_2\left(Y_1, Y_2, Y_3\right) = I_4. \tag{7}$$

As before [17-19], we chose the value of the invariant I_4 as the double value of the free constant (FC) figuring in the left-hand side of (7), i.e. FC(left side) = $2FC$(right side) = I_4. After some algebraic manipulations the fourth and the second order forms entering to the left-hand side can be presented as:

$$K_4\left(Y_1, Y_2, Y_3\right) = Y_1^4 + Y_2^4 + Y_3^4 + R^{(1,2)}Y_1Y_2\left[Y_3^2 - \frac{1}{2}\left(Y_1 + Y_2\right)^2\right] +$$
$$+R^{(1,3)}Y_1Y_3\left[Y_2^2 - \frac{1}{2}\left(Y_1 + Y_3\right)^2\right] + R^{(2,3)}Y_2Y_3\left[Y_1^2 - \frac{1}{2}\left(Y_2 + Y_3\right)^2\right]; \tag{8a}$$

$$K_2\left(Y_1, Y_2, Y_3\right) = A_{11}Y_1^2 + A_{22}Y_2^2 + A_{33}Y_3^2 +$$
$$+A_{12}Y_1Y_2 + A_{13}Y_1Y_3 + A_{23}Y_2Y_3. \tag{8b}$$

The constants $A_{\alpha\beta}$ figuring in expression (8b) are defined as:

$$
\begin{aligned}
&A_{11} = 6Q_{11} - \left(Q_{22} + \frac{3}{2}Q_{21}\right)R^{(1,2)} - \left(Q_{33} + \frac{3}{2}Q_{31}\right)R^{(1,3)} + Q_{32}R^{(2,3)};\\
&A_{22} = 6Q_{22} - \left(Q_{11} + \frac{3}{2}Q_{21}\right)R^{(1,2)} + Q_{31}R^{(1,3)} - \left(Q_{33} + \frac{3}{2}Q_{32}\right)R^{(2,3)};\\
&A_{33} = 6Q_{33} + Q_{21}R^{(1,2)} - \left(Q_{11} + \frac{3}{2}Q_{31}\right)R^{(1,3)} - \left(Q_{22} + \frac{3}{2}Q_{32}\right)R^{(2,3)};\\
&A_{12} = -\left(4Q_{21} + \frac{3}{2}Q_{11} + \frac{3}{2}Q_{22} - Q_{33}\right)R^{(1,2)} + 2Q_{32}R^{(1,3)} + 2Q_{31}R^{(2,3)};\\
&A_{13} = 2Q_{32}R^{(1,2)} - \left(4Q_{31} + \frac{3}{2}Q_{11} + \frac{3}{2}Q_{33} - Q_{22}\right)R^{(1,3)} + 2Q_{12}R^{(2,3)};\\
&A_{23} = 2Q_{31}R^{(1,2)} + 2Q_{21}R^{(1,3)} - \left(4Q_{32} + \frac{3}{2}Q_{22} + \frac{3}{2}Q_{33} - Q_{11}\right)R^{(2,3)}.
\end{aligned}
\tag{9}
$$

The constant I_4 (defined by 3 moments and 12 intercorrelations of the fourth order) figuring in the right-hand side of (7) is defined as:

$$
\begin{aligned}
I_4 = Q_{1111} + Q_{2222} + Q_{3333} &- \left(Q_{2211} - Q_{3321} + \frac{1}{2}Q_{2111} + \frac{1}{2}Q_{2221}\right)R^{(1,2)} -\\
&- \left(Q_{3311} - Q_{3221} + \frac{1}{2}Q_{3111} + \frac{1}{2}Q_{3331}\right)R^{(1,3)} -\\
&- \left(Q_{3322} - Q_{3211} + \frac{1}{2}Q_{3222} + \frac{1}{2}Q_{3332}\right)R^{(2,3)}.
\end{aligned}
\tag{10}
$$

It is interesting to notice that in the case of the spherical symmetry ($r_{1k} = r_{2k} = r_{3k}$) all correlations coincide with each other and the value of I_4 equals zero. The form of the fourth order (7) admits the separation of the variables in the spherical system of coordinates. If one accepts the conventional notations:

$$
\begin{aligned}
&y_1 = \langle y_1 \rangle + R\sin\theta\cos\varphi;\\
&y_2 = \langle y_2 \rangle + R\sin\theta\sin\varphi;\\
&y_3 = \langle y_3 \rangle + R\cos\theta;\\
&0 \le \theta \le \pi, 0 \le \varphi \le 2\pi,
\end{aligned}
\tag{11}
$$

then substitution of these variables into (7) leads to the following biquadratic equation relatively the unknown radius $R(\theta, \varphi)$:

$$\left[R(\theta,\varphi)\right]^4+\left(\frac{P_2(\theta,\varphi)}{P_4(\theta,\varphi)}\right)\left[R(\theta,\varphi)\right]^2-\frac{I_4}{P_4(\theta,\varphi)}=0. \tag{12a}$$

The desired solution ($R(\theta, \varphi) > 0$) is written as:

$$R(\theta,\varphi)=\left[\frac{\sqrt{P_2^2(\theta,\varphi)+4I_4\cdot P_4(\theta,\varphi)}-P_2(\theta,\varphi)}{2P_4(\theta,\varphi)}\right]^{\frac{1}{2}}. \tag{12b}$$

The polynomials $P_{2,4}(\theta, \varphi)$ entering in (12) are defined by the following expressions:

$$\begin{aligned}P_4(\theta,\varphi)&=\sin^4\theta\cdot\cos^4\varphi+\sin^4\theta\cdot\sin^4\varphi+\cos^4\theta+\\&+R^{(1,2)}\sin^2\theta\sin\varphi\cos\varphi\left[\cos^2\theta-\frac{\sin^2\theta}{2}(\sin\varphi+\cos\varphi)^2\right]+\\&+R^{(1,3)}\sin\theta\cos\theta\cos\varphi\left[\sin^2\theta\sin^2\varphi-\frac{1}{2}(\sin\theta\cos\varphi+\cos\theta)^2\right]+\\&+R^{(2,3)}\sin\theta\cos\theta\sin\varphi\left[\sin^2\theta\cos^2\varphi-\frac{1}{2}(\sin\theta\sin\varphi+\cos\theta)^2\right];\end{aligned} \tag{13a}$$

$$\begin{aligned}P_2(\theta,\varphi)&=A_{11}\sin^2(\theta)\cos^2(\varphi)+A_{22}\sin^2(\theta)\sin^2(\varphi)+A_{33}\cos^2(\theta)+\\&+A_{12}\sin^2(\theta)\sin(\varphi)\cos(\varphi)+A_{13}\sin(\theta)\cos(\theta)\cos(\varphi)+A_{23}\sin(\theta)\cos(\theta)\sin(\varphi).\end{aligned} \tag{13b}$$

The last expressions (11) – (13) determine the final form of the DGI in 3D-space. It includes three surfaces determined by expressions (11). The further analysis shows that expression (12b) equals zero (because $I_4 = 0$) in the case of the coincidence of three compared random sequences ($r_{1k} = r_{2k} = r_{3k}$). The radius $R(\theta, \varphi)$ can contain the complex expression when the integrand in (12b) becomes *negative*. It accepts the negative values when the constant I_4 in the most cases defined by expression (10) becomes negative. In this case, it is convenient in many cases to rewrite expressions (11) in the following final form:

$$\begin{gathered}y_1=\langle y_1\rangle+|R(\theta,\varphi)|\sin\theta\cos\varphi;\\y_2=\langle y_2\rangle+|R(\theta,\varphi)|\sin\theta\sin\varphi;\\y_3=\langle y_3\rangle+|R(\theta,\varphi)|\cos\theta;\\|R(\theta,\varphi)|=\sqrt{\left[\operatorname{Re}(R(\theta,\varphi))\right]^2+\left[\operatorname{Im}(R(\theta,\varphi))\right]^2};\\0\le\theta<\pi,\ 0\le\varphi<2\pi.\end{gathered} \tag{14}$$

One can notice also that expressions representing the desired surfaces can be replaced by three plane functions, also. Really, if we subordinate the angular variables to condition:

$$y_1(j) = \langle r_1 \rangle + R(\varphi_j, \theta_j)\cos(\varphi_j)\sin(\theta_j);$$
$$y_2(j) = \langle r_2 \rangle + R(\varphi_j, \theta_j)\sin(\varphi_j)\sin(\theta_j);$$
$$y_3(j) = \langle r_3 \rangle + R(\varphi_j, \theta_j)\cos(\theta_j); \tag{15}$$
$$\varphi_j = 2\pi\left(\frac{j}{N}\right), \ \theta = \pi\left(\frac{j}{N}\right), R_j \geq 0, j = 0.1..., N,$$

then these three functions can serve for identification and quantitative description of the given random/deterministic sequences. The compact numbers of the parameters that determine their behavior are remained the *same* and equaled 13. It is interesting to notice also that if three compared random functions completely coincide with each other ($I_4 = 0$) then these functions are reduced to the gravity point $<r_1> = <r_2> = <r_3>$.

The curves defined by equation (15) facilitate considerably the numerical analysis of initial data, while the surfaces (14) give actually a demonstration of the compared sequences in 3D-space. We want to stress here again that the final expressions (14) and (15) do *not* use *any* model assumptions and are determined completely by the measured data together with their measurement errors. Finishing this section one can say that this method can be applied for reduction of initial data. This reduction procedure can be divided on the following stages:

1. Initially, any available data can be written in the form of rectangle matrix $[N \times M]$, where number N ($j = 1, 2,..., N$ – number of rows) determines the given data points and M ($m = 1,2,\ldots, M$ – columns) determines the number of the repeated measurements forming in total the statistically significant sampling. As the result of application of 3D-DGI method we obtain the reduced matrix $[M \times S]$, where each column of the reduced matrix ($Pr_{m,s}$: $<y_\alpha>$ (3), $R^{(\alpha,\beta)}$(3), $A_{\alpha\beta}$(6), I_4(1); $\alpha, \beta = 1, 2, 3$) determines the complete combination of the moments and their intercorrelations $(3 + 3 + 6 + 1 = 13)$ up to the fourth order inclusive. In the result of application of the 3D-DGI method we obtain $s = 1, 2,\ldots, S$ ($S = 13$) distributions that demonstrate the variations of each statistical parameter $Pr_s(m)$ with respect to the number of repeated measurements ($m = 1,2\ldots, M$).

2. The further reductionis possible if one takes into account that each random function $y_s(m) \equiv Pr_s(m)$ (belonging to the column m) is located inside the rectangle $M \times (\mathrm{Range}[y_s(m)])$, where $Range(f) = \max(f) - \min(f)$. For comparison one random function $y_{1,s}(m)$ with another $y_{2,s}(m)$ corresponding to the chosen parameter s($s = 1, 2,\ldots,S$) one can use the following simple formula:

$$Q_{1,2}(s) = \frac{Range(y_{1,s}) + Range(y_{2,s})}{\max(y_{1,s}, y_{2,s}) - \min(y_{1,s}, y_{2,s})}. \tag{16}$$

This expression in spite of its simplicity is really effective for comparison of the statistical closeness of a pair random functions belonging to the given/another sampling participating in comparison operation, when the real behavior of these random functions $y_{1,2,s}(m)$ are not known. Really, if the function $Q_{1,2}(s)$ is located in the interval [1,2] then the pair random functions are statistically *close* to each other. In the case when $Q_{1,2}(s)\in[0,1)$ one can conclude that the pair random functions compared are statistically *different*. Besides this important parameter (16), one can take into account the symmetry of the random function $y_1(m)$. Any random function located in the rectangle $M\times Range[y(m)]$ crosses the line $\langle Pr(m)\rangle$ coinciding with its mean value. Therefore, for evaluation of the symmetry of a random function one can introduce the value:

$$Sm(y) = \frac{\max(y) - mean(y)}{mean(y) - \min(y)}. \tag{17}$$

If the value $Sm(y)$ is located near the unit value ($Sm(y) \approx 1$) then the line $\langle y\rangle$ divides the rectangle $M\times Range[y(m)]$ on two almost equal parts. In other cases, the value $Sm(y)$ can be (>,<) 1 determines the measure of asymmetry. After application of expression (16) for comparing of similar columns (belonging to the same parameter Pr_s) one can receive finally the vector of the length S = 13 that contains information about the statistical closeness of two matrices compared. It is interesting to notice that simple expression (16) can be used also for comparison each successive measurement with another one in the given rectangle matrix $[N\times M]$. If one compares the vectors y_m forming the columns of the initial matrix with each other then in the result of application (16) one can obtain the symmetrical matrix $U(m_1, m_2)$ ($m_{1,2} = 1,2,\ldots, M$) with elements located in the interval $0 \le U(m_1, m_2) \le 2$. Only elements located in the interval $1 \le U(m_1, m_2) < 2$ will correspond to a «good» experiment; while the elements from the interval $0 \le U(m_1, m_2) < 1$ should be considered as possible «outliers» and correspond to «bad»/unsuccessful measurements.

3. How to find the parameters belonging only to one matrix in order to compare them with similar parameters of another tested matrix in cases when the reference / «pattern» matrix is absent? Initially, it is necessary to scale each column to the *same* interval:

$$Prn_s = \frac{DPr_s}{Range(Pr_s)} \equiv \frac{Pr_s - \langle Pr_s\rangle}{Range(Pr_s)};$$
$$-\frac{1}{2} \le Prn_s \le \frac{1}{2};\ \langle Pr_s\rangle = \frac{1}{M}\sum_{m=1}^{M} \mathrm{Pr}_{m,s}. \tag{18}$$

This normalization procedure makes all parameters Pr_s statistically close to each other with mean value equaled zero mean(Prn_s) = 0 and with the $Range(Prn_s) = 1$. If one integrates expressions (18) for each parameter one can receive the statistically different curves JP_s = Integral (Prn_s) for each initial parameter ($s = 1,2,\ldots,S$). The distributions of the ranges of these integral curves $P_1 = Range(JP_s)$ together with distribution of asymmetries $P_2 = Range(Sm(JP_s))$ calculated with the help of expression (17) give finally the matrix containing $[S\times 2]$ containing (S = 13) rows and two columns only. For these two functions $P_{1,2}$ one can add two other functions:

$$P_{3,4} = \left(\frac{1}{S}\right)\sum_{s=1}^{S} Y_s,\quad Y_s = Prn_s, JP_s. \tag{19}$$

These distributions for the chosen data are shown below in Figs 4-5. If we calculate the ranges of these 4 columns one obtains finally 4 values *only* that can characterize the initial matrix $[N\times M]$. If we have a set of matrices $[N\times M]_q$ ($q = 1,2,\ldots,Q$) then this simple and general procedure allows to select the «best» one having minimal values of these 4 key parameters $P_{1,2,3,4}$. It will characterize the stability of the initial sequence and their minimal values will serve as a criterion for selection of the «best» TLS among other TLS(s). We want to notice here that this final stage of treatment of «big» matrices is differed from the procedure used in paper [6]. Earlier, one of us (RRN) had the set of rectangle matrices that can be characterized as «normal/reference» ones and compared them with «strange/tested» matrices associated with defects. Data that will be analyzed below do not contain this information. Therefore, we propose here more general and common procedure described above for selection the «best» data expressed in the form of rectangle matrices.

2. Description of the experimental circuitry

ELVIS II workbench by National Instruments was used as ADC. ELVIS II has 1.25 MS/s (Mega samples per second) maximum sampling rate and 16-bit resolution [21]. As other ADC(s), ELVIS II contains some circuits for initial signal processing. The simplified block diagram of the channel for digitizing of an analog signal is shown in Fig. 1. This workbench represents itself an analog input circuitry. In Fig. 1 the abbreviation «MUX» defines a multiplexer.

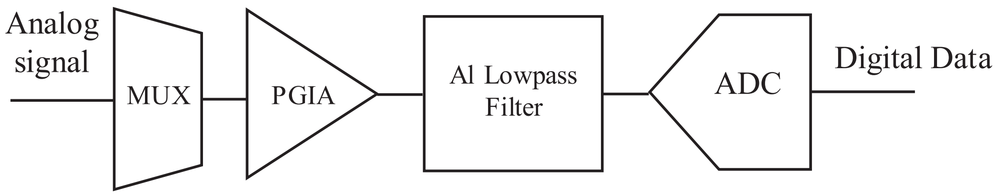


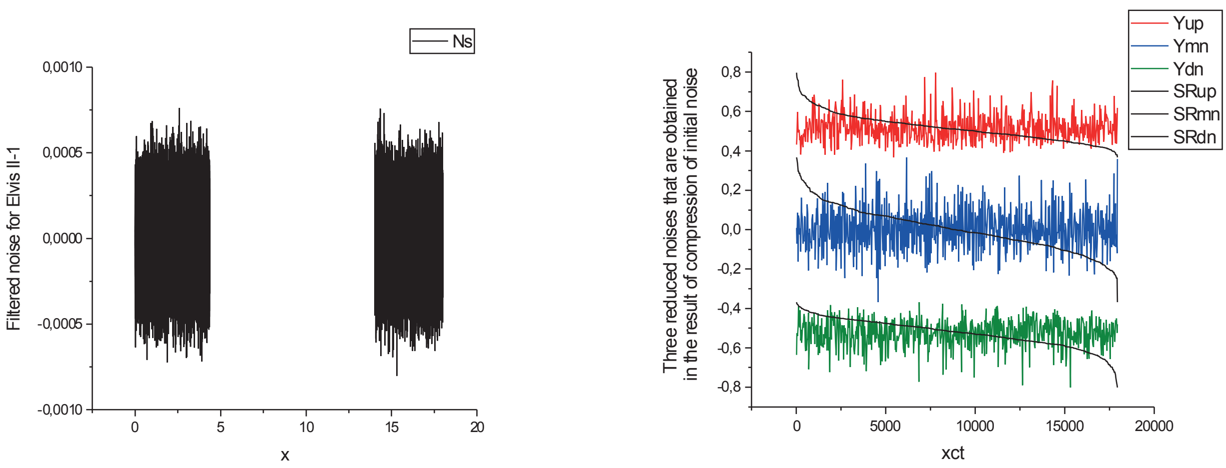
Fig. 1. The workbench ELVIS II analog input circuitry:
MUX – multiplexer; PGIA – programmable gain instrumentation amplifier; AILF – analog Input Lowpass Filter; ADC – analog to digital converter

It routs one Analog Input (AI) channel at a time to the ADC, because ELVIS II uses one ADC for several AI(s). PGIA serves as an abbreviation for the Programmable Gain Instrumentation Amplifier. The PGIA amplifier attenuates a signal for obtaining the maximum resolution of the ADC. Analog Input Lowpass Filter is a mandatory block of any ADC, it is necessary to suppress aliasing phenomenon. Any signal digitization contains anti-aliasing filtering. The passband of the used filter starts from «zero» Hz to the upper cutoff frequency that equals to the sampling frequency divided by two. The signal that is passed only through the anti-aliasing filter is defined in this paper as «unfiltered» signal. One can change the lower and upper cut-off frequencies to the desired values. After realization of this procedure, a digitized signal should be passed through a digital filter. If a signal is passing through an additional digital filtering, one can define it as the filtered signal. In this paper,we consider the filtered signals, non-filtered signals will be considered in another paper.We connected 50 Ohm resistor between analog input and the «ground». It simulates the output impedance of the circuit that can be connected to the given ADC. The input signal range has been programmed and finally the PGIA has a gain equals unit value. Such connection provides the conditions that the ADC itself becomes the *main* source of a noise in the presented circuitry. Four different types of ELVIS II were used in the experiment. In this paper, they are defined as «ELVIS II-L» (L = 1,2,3,4). Fifteen ($m = 1, 2, 3\ldots, M$; $M = 15$) successive measurements were realized using the selected type of the embedded ADC belonging to the chosen ELVISII-L. Each of these measurements was carried out with the fixed 50 Ohm resistor at the input (as it was mentioned earlier) and $1.8{\cdot}10^5$ data points were recorded with a sampling frequency 10 kHz.

3. Data treatment algorithm

In the result of the procedure described in the previous section one can receive 4 matrices. Each rectangle matrix includes N=1.8·10^5 data points and M = 15 successive measurements corresponding to the chosen «ELVIS II -L» (L =1, 2, 3, 4) workbench. The basic problem is that one can choose the «best» one among them, best on their TLS data. In our case, the pattern workbench is absent. In order to solve this problem one can propose the following algorithm. It is divided on the following steps.

S-1. In order to decrease the computational time associated with treatment of long sequences ($N > 10^5$) one can suggest the procedure of reduction to three incident points that was described and successfully used earlier in papers [22, 23]. This procedure helps to divide initial TLS on three independent parts Yup_j, Ymn_j, Ydn_j, where number of data points is compressed $j = 1,2,\ldots, [N_c = N/b]$, and the compression parameter is equaled $b = 300$, the rectangle brackets […] determine the integer part extraction operation. These three obtained random sequences [Yup_j, Ymn_j, Ydn_j] describe the distribution of the «up», «mean» and «down» amplitudes, accordingly. This procedure helps to obtain three similar TLS(s) from the initial sequence as it is illustrated in Fig. 2, where three reduced TLS(s) for the first sequence (ELVIS II-1) are shown.



a *b*

Fig. 2. Results of procedure, which helps to obtain three similar reduced TLS(s) from the initial sequence: *a* – on the left-hand side we show the initial trendless noise recorded for «Elvis II-1» (as an example, we took the first measurement from available M = 15); *b* – on the right-hand side we demonstrate the result of reduction to three incident points (The compression coefficient b = 300. The sequences of the ranged amplitudes (SRAs) are shown by bold black lines. Noise data recorded for other ELVIS II -2,3,4 is similar and, therefore, other data are not shown)

S-2. In the result of application of the first step one can prepare 4×[N_c = 1000×M = 15] matrices corresponding to each «ELVIS II-L» (L = 1, 2, 3, 4). Each device finally can be presented in the form of one reduced rectangle matrix (based on the calculation of three SRAs shown in Fig. 2) having the size (15×13) with the help of 3D-DGI method, described in section 2. Each calculated set combining 13 parameters corresponds as minimum to three surfaces that are defined by equation (14). The most *significant* surface is associated with the behavior of the radius modulus

$\left|R(\theta,\varphi)\right| = \sqrt{\left[\mathrm{Re}\left(R(\theta,\varphi)\right)\right]^2 + \left[\mathrm{Im}\left(R(\theta,\varphi)\right)\right]^2}$. It is depicted in Figs. 3 for four monotone SRAs curves shown in by solid black lines in Fig. 2.

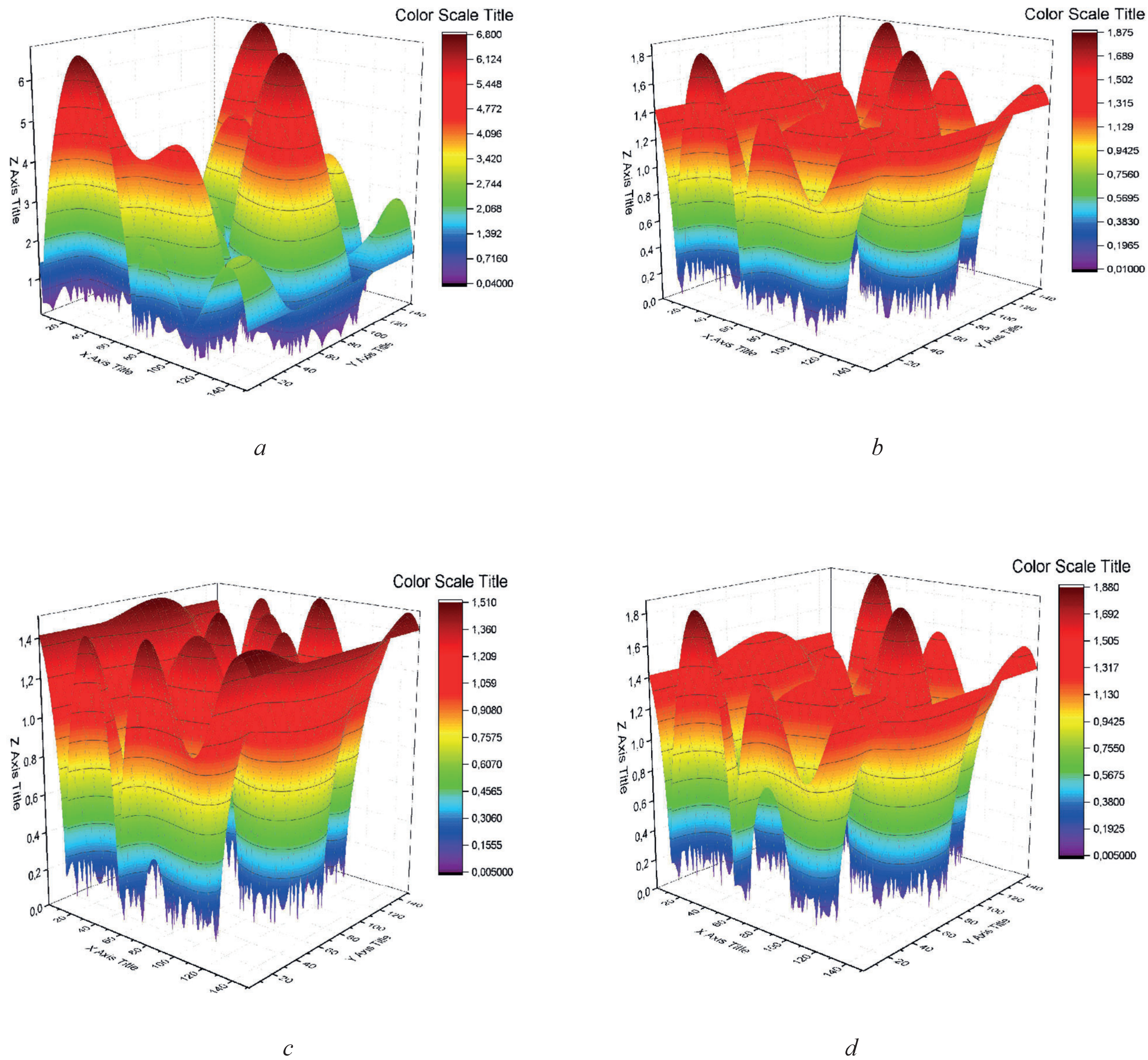


Fig. 3. Behavior of the radius modulus for four monotone SRAs curves shown in by solid black lines in Fig. 2:
a – on the up left-hand side we show the 3D surface recorded for «ELVIS II-L = 1»;
b – on the up right-hand side we show 3D-surface for «ELVIS II -L = 2»
(as one can notice from qualitative comparison their pictorial surfaces are differed from each other);
c – on the down left-hand side we show the 3D surface recorded for «ELVIS II-L=3»;
d – on the down right-hand side we show 3D-surface for «ELVIS II-L=4»
(as one can notice from qualitative comparison of these surfaces,
they are slightly differed from each other). We show surfaces only for the first measurements.
Other surfaces for other successive measurements m = 2, 3,…,15 are not shown

In order to save place for the key figures we show only the basic stages and demonstrate the differences between parameters obtained with the help of expressions (18)-(19). These curves play a key role in the final comparison of the given workbenches ELVIS II-L = 1, 2, 3, 4, when the pattern workbench is absent.

The final stage allows to select the «best» ELVIS II-L based on the range values of the parameters $P_{1,2,3,4}$. Figure 4, *a* demonstrates the behavior of the ranges of integrals $P_1 = Range(JP_s)$ for all ELVIS II -L = 1, 2, 3, 4. Figure 4, *b* shows the ranges of the asymmetrical coefficients $P_2 = Range(Sm(JP_s))$. Definitely, as a criterion we select the «best one» based on their ranges values.

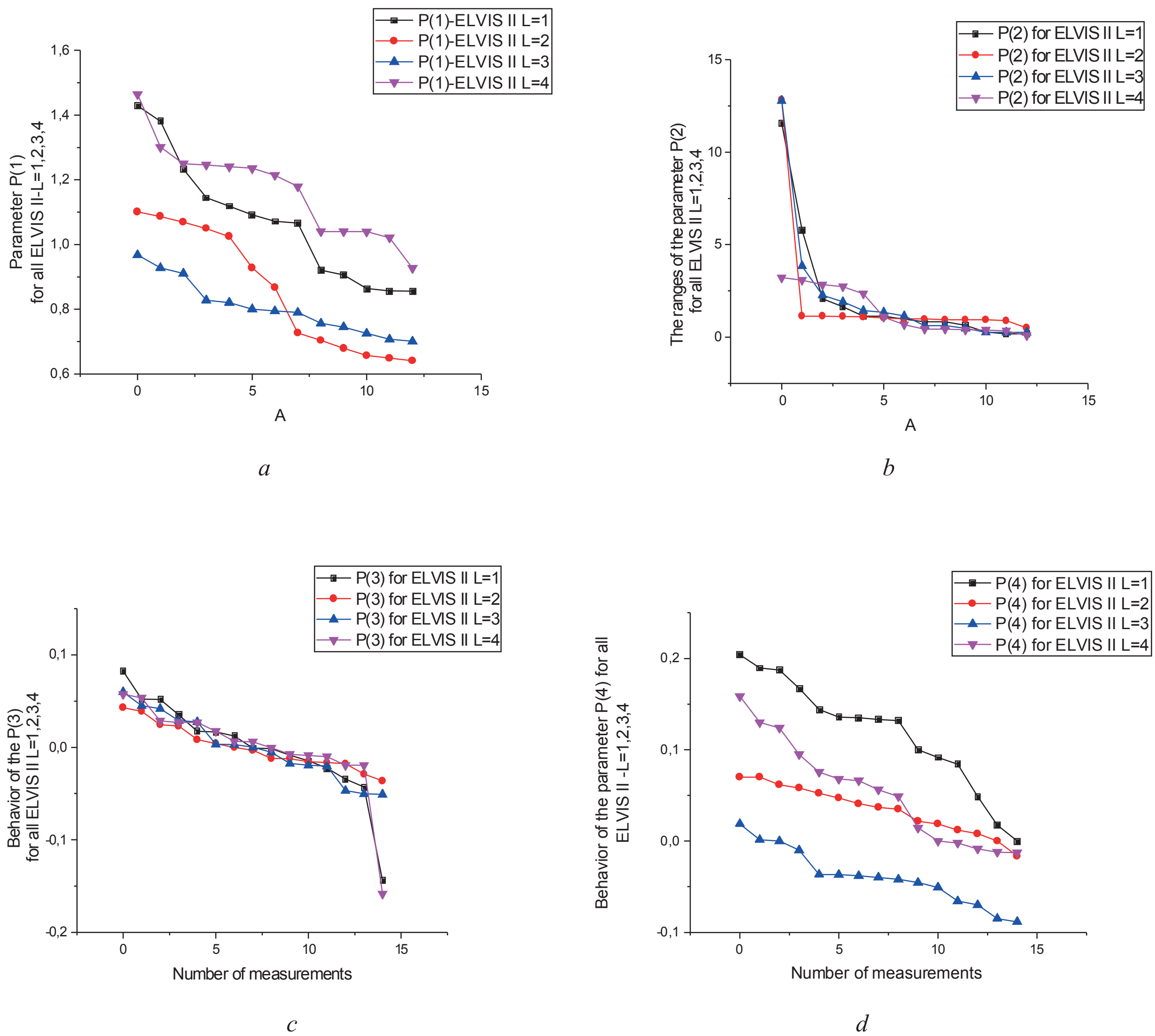


Fig. 4. Selection of the «best» ELVIS II-L based on the range values of the parameters $P_{1,2,3,4}$:
a – on the up left-hand side we show the behavior of the parameter P1 expressed in the form of the sequences of the ranged amplitudes for all ELVIS II-L = 1, 2, 3, 4;
b – on the up right-hand side we place a similar figure for the parameter P2 = Range(Sm(JPs));
c – the behavior of the parameter P3 defined by expression (19) is shown on the down left
(This is averaged value taken over all measurements equaled M = 15);
d – on the right-hand side we show the behavior of the averaged value of the parameter P4. These four parameters allow to select the most suitable workbench from the available ones based on their minimal values criterion

The ranges of the parameters $P_{1,2,3,4}$ allows to select the «best» Elvis II workbench. These four key parameters are listed in Table 1.

Table 1

In this table all ranges for all key parameters $P_{1,2,3,4}$ are collected

The workbenches (in vertical) and their ranges (in horizontal)	Ranges of P_1	Ranges of P_2	Ranges of P_3	Ranges of P_4
ELVIS II L=1	0.57387	0.46009	0.26714	0.53755
ELVIS II L=2	**11.38654**	**12.33249**	**12.53470**	**3.13543**
ELVIS II L=3	**0.20667**	**0.07926**	**0.10074**	**0.20596**
ELVIS II L=4	0.20404	0.08664	0.10725	0.17108

Analysis of these data show that the first place will belong to ELVIS-II-L = 3, the workbench ELVIS-II-L = 2 can be selected as the worst one. These parameters are highlighted by the bold line. The second place in this workbench «competition» can be awarded to the ELVIS-II-L = 4.

Results and discussion

In this paper, we obtained the complete form for the 3D-DGI surface. This surface contains 13 parameters that determine the desired feature space. These parameters, in turn, include the combination of the moments and intercorrelations up to the fourth order inclusive. From mathematical point of view, the 3D-DGI method is derived in the result of diagonalization of the fourth-order form. As is it known that the modern statistics is based on diagonalization of the quadratic forms only and the moments of the third and fourth order were not taken into account. This method is applicable to consideration of three arbitrary random sequences having N data points. The selection of an *optimal* triple combination (Fig. 2(*b*)) associated with the chosen random sequence represents a problem and merits a further research. In this paper, it is reasonable to use the procedure of reduction to three incident points. However, other options are also possible. Besides the description of 3D-DGI method, the general reduction procedure is proposed that was proved to be helpful in comparison of «big» rectangle matrices with each other with the help of expression (16). If all data can be dismembered onto two parts – «friend-or-foe» the simplest expression (16) becomes efficient, when this division onto reference/tested data) is possible. If the «pattern»/reference standard is absent then one can use the procedure based on new expressions (17) – (19). The attentive reader may notice that this approach can be applied successfully to analysis of different images as well. Really, let us suppose that some rectangle matrix corresponds to an initial «image». With the help of the 3D-DGI method described above one can reduce the initial image to an «effective» surface containing for its construction 13 parameters only. This effective surface can serve as an effective «fingerprint» differentiating the chosen «image» from another one. Besides, this effective image helps an operator/robot in acceptance of the right decision. This simple idea for its justification and optimization merits the further research. The TLS data obtained from four ELVIS II workbenches showed the effectiveness of the proposed method and allowed to select the «best» one (see Table 1).

One of us (RRN) does believe that potential researchers will receive a new original and rather general tool for treatment and comparison of different TLS data, especially associated with rectangle matrices having «large» sizes. This method enables to suggest more accurate scheme for classification of different «color» noises including also Gaussian, heat and flicker-noises based on the unified platform, containing integer moments and intercorrelations, up to the fourth order, inclusive. One can stress here that other platforms of such kind cannot be found and realized, because the *analytical* separation of the polynomial forms exceeding the fourth order becomes *impossible*. Therefore, the 3D-DGI method representing itself a «specific platform» for comparison of random functions is *general* and *unique*. This method opens new horizons in analysis of different random fluctuations/ functions and because of its importance merits a further research.

REFERENCES

1. *Timashev, S.F.* Flikker-noise spectroscopy. Information in Chaotic signals // Moscow: PhysMatLit publishing house, 2007. – 248 p. http://www.fml.ru

2. *Timashev, S.F. and Polyakov Yu. S.* Review of flicker noise spectroscopy in electrochemistry/ / Fluctuations and Noise Letters. – 2007. – Vol. 7. – № 2. – Pp. R15-R47.

3. *Timashev, S.F., Demin, S.A., Panischev, O.Yu., Polyakov, Yu.S., Kaplan, A.Ya., Nefedev, Yu. A.* Flicker-noise spectroscopy as a tool for the personalized medicine of the future//Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki. – 2012. – Vol. 154. – Book 4. – Pp. 161 – 177.

4. *Yulmetyev, R.M.* Stochastic dynamics of time correlation in complex systems withdiscrete time// Phys. Rev. – 2000. – E 62 6178.

5. *Yulmetyev, R. et al.* Quantication of heart rate variability by discrete nonstationarity non-Markov stochastic processes // Phys. Rev. – 2002. E 65 046107.

6. *Nigmatullin, R.R., Vorobev, A.S.* The «universal» Set of Quantitative Parameters for Reading of the Trendless Sequences// Fluctuation and Noise Letters. – 2019. – Vol. 18. – No. 4. 1950023.

7. *Tyagai,V.A.* Faradaic noise of complex electrochemical reactions // Electrochimica Acta. 1971. – Vol. 16. – P. 1647.

8. *Barker, G.C.* Faradaic reaction noise // Journal of Electroanalytical Chemistry and Interfacial Electrochemistry. – 1977. – Vol. 82. – P. 145.

9. *Bozdech, S., Krisher, K., Crespo-Yapur, D.A., Savonova, E., Bonnefont, F.* $1/f^2$ noise in bistable electrocatalytic reactions on mesoscale electrodes // Faraday discussions. – 2016. – Vol. 193. – P. 187.

10. *Nigmatullin, R.R., Martemianov, S., Evdokimov, Yu.K., Denisov, E., Thomas, F.* Adiatantov. New approach for PEMFC diagnostics based on quantitative description of quasi-periodic oscillations// International Journal of Hydrogen Energy. – 2016. – Vol. 41. – P. 12582.

11. Astafev E. Electrochemical noise measurements methodologies of chemical power sources// Instrumentation Science & Technology. – 2019. – V. 47. – P. 233.

12. Astafev E. and Uksche A. Peculiarities of hardware for electrochemical noise measurement in chemical power sources // IEEE Transactions on Instrumentation and Measurement. – 2019. – Vol. 68. – P. 4412.

13. *Gutenberg, B.* Richter. Seismicity of the Earth and associated phenomena – Princeton, NJ: Princeton University Press. – 1949. – 310 p.

14. *Nigmatullin, R.R., Vorobev, A.S., Nepeina, K.S., Alexandrov, P.N.* Fractal description of the complex beatings: How to describe quantitatively seismic waves? // Chaos, Solitons and Fractals 2019. – Vol. 120. – P. 171. https://doi.org/10.1016/j.chaos.2019.01.017. 2019.

15. *Babenko, Y.I.* Power-law Invariants of Discrete Sets, 'Professional' Publ. House, S-Petersburg. 2014. ISBN 978-5-91259-095-5. In Russian.

16. *Babenko, Y. I.* Power Relations in a Circumference and a Sphere, Norell Press, Division of Norell Press Inc., NJ 08326, USA, 1997.

17. *Nigmatullin, R.R., Vorobev, A.S., Budnikov, H.C., Sidelnikov A.V., Badikova, A.D., Maksyutova, E.I*. The usage of unremovable artefacts for the quantitative «reading» of nanonoises in voltammetry//New Journal of Chemistry. – Vol. 43. 6168. DOI: 10.1039/c9nj00159j, 2019.

18. *Nigmatullin, R.R., Vorobev, A.S., Nasybullin, A.R., Lino, P., Grande, M.* Advanced and sensitive method by discrete geometrical invariants for detection of differences between complex fluids // Communications in Nonlinear Science and Numerical Simulation. 2019. – Vol. 73. – P. 265. https://doi.org/10.1016/j.cnsns.2019.02.012.

19. *Nigmatullin, R.R.* Discrete Geometrical Invariants in 3D-space: how three random sequences can be compared in terms of «universal» statistical parameters? // Frontiers in Physics.2020. doi: 10.3389/fphy.2020.00076.

20. *Nigmatullin, R.R., Vorobev, A.S.* The «universal» Set of Quantitative Parameters for Reading of the Trendless Sequences // Fluctuation and Noise Letters.World Scientific Publishing Company. – 2019. – Vol. 18. – No. 4. 1950023 (19 p.), DOI: 10.1142/S0219477519500238.

21. NI Educational Laboratory Virtual Instrumentation Suite II Series (NI ELVISTM II Series) User Manual. https://www.ni.com/pdf/manuals/374629c.pdf

22. *Nigmatullin Raoul R., Domenico Striccoli, Gennaro Boggia, Cristiano Ceglie.* A novel approach for characterizing multimedia 3D video streams by means of quasiperiodic processes// Signal, Image and Video Processing. – 2016. – P. 1. DOI:10.1007/s11760-016-0866-9.

23. Nigmatullin Raoul R., Giniatullin Rashid A., Skorinkin Andrei I. Membrane current series monitoring: essential reduction of data points to finite number of stable parameters // Computational Neuroscience. – 2014. – Vol. 8. Article 120, 1.DOI: 10.3389/fncom.2014.00120.

МОЖЕТ ЛИ ПРОИЗВОЛЬНЫЙ БЕСТРЕНДОВЫЙ ШУМ СЛУЖИТЬ НОВЫМ ИСТОЧНИКОМ ИНФОРМАЦИИ?

Р.Р. Нигматуллин, Р.К. Сагдиев

Казанский национальный исследовательский технический университет
им. А.Н. Туполева-КАИ
Российская Федерация, 420111, г. Казань, ул. Карла Маркса, 10

Аннотация. Авторы хотят доказать, что случайная последовательность без тренда (БТСП) может быть использована в качестве дополнительного источника информации. Эта дополнительная информация может быть извлечена из случайного бестрендового шума с помощью метода 3D-DGIs (3-мерного метода дискретных геометрических инвариантов), который позволяет уменьшить 3N случайных точек, представленных в форме 3-х БТП до 13 параметров, составленных из комбинации целых моментов и их взаимных корреляций до четвертого порядка включительно. На самом деле эти параметры образуют «универсальное» 13-мерное функциональное пространство признаков для сравнения одной случайной последовательности с другой. Сравнение параметров, связанных с различными шумовыми дорожками, позволяет использовать этот набор параметров для калибровки и других целей, связанных со «стандартным» /эталонным оборудованием. В качестве примера рассматривали измеренные БТСП, полученные с рабочего стола ELVIS II (National Instrument Corporation). Новый метод помогает найти различия между четырьмя типами предварительно отфильтрованных выборок БТСП (образующих соответствующие прямоугольные матрицы), принадлежащих выбранным АЦП, и выбрать «лучший» из них.

Ключевые слова: Дискретные геометрические инварианты; 13-мерное пространство признаков; бестрендовые случайные последовательности (БТП); сравнение АЦП на основе БТП.

Статья отправлена в редакцию 07.05.2021