ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 1, pp. 89–96.

INTELLECTUAL CONTROL SYSTEMS, DATA ANALYSIS

On Some Properties of Spline Grids in Classification Problems

V. G. Babikov[*,a] **and A. A. Galyaev**[*,b]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a] *babikov@ipu.ru,* [b] *galaev@ipu.ru*

Received August 3, 2025
Revised September 24, 2025
Accepted October 6, 2025

Abstract—This work supplements the authors' series of publications on the topic of detection and classification of weak signals. Continuing this theme, the current study presents results of an applied nature. Here the authors focus on adaptive metric grids applied to information diagrams and the development of new classification features.

Keywords: information entropy, discrete Fourier transform, statistical disequilibrium, statistical complexity, spectral complexity, metric grid

DOI: 10.7868/S1608303226010078

1. INTRODUCTION

Classical methods for solving detection problems use additional information about signal properties. For instance, solving the optimal filtering problem requires knowledge of signal properties: periodicity, frequency band, etc. [1]. Solving the problem of distinguishing between two hypotheses relies on the Neyman–Pearson lemma, determines the fact of exceeding an optimal threshold for a given false alarm probability, and requires estimating the statistical properties of the sample distributions of noise and the signal-noise mixture [2]. Reference [3] presents a variation of the Neyman–Pearson lemma, whose peculiarity is that it does not allow errors when accepting one hypothesis instead of another. Solving the change-point problem requires tuning the algorithm to changes in the unknown statistical characteristics of the distributions of noise and the signal-noise mixture, as does the anomaly search problem [4]. All listed methods demonstrate qualitative and reliable performance when the signal exceeds the noise, but at low signal-to-noise ratios or in heavy noise environments, they often give incorrect answers.

Regarding the use of information characteristics for solving problems in statistical physics, classifying medical tests (e.g., for studying breast histopathology image textures), etc., this approach already has a long history. As early as 1999, [5] introduced a complexity measure based on the average information gain for quantitatively describing the complexity of two-dimensional patterns. [6] raises issues related to developing a general measure of complexity — or regularity, or structure — for two-dimensional systems. Such tasks arise, for example, when studying surfaces in geology.

Article [7] is devoted to studying the possibility of texture classification via two-dimensional ordinal patterns, particularly via the statistical complexity diagram. In [8], the authors confirm that the statistical complexity diagram is a popular tool for distinguishing stochastic signals (noise) from deterministic chaos (in English-language literature, the term for this tool is the complexity-entropy (CE) plane). However, the authors discovered that both high-dimensional deterministic time series and stochastic surrogate data (noises) can be located in the same region of the diagram, and their representations show very similar behavior for different signal-noise mixtures or even signal types. Therefore, classifying this data based on their position in the diagram plane can be difficult or even misleading.

The question of the complexity of distinguishing noise and noise-like deterministic signals is also covered in detail in [9]. That work is devoted to studying statistical and spectral complexity diagrams. The diagrams are constructed as follows: information entropy is plotted on the abscissa axis, statistical (spectral) complexity on the ordinate axis. In the text, statistical complexity diagrams are denoted as (S, C_{TV}) and spectral complexity diagrams as (S, C_S). The work provides relevant definitions, makes significant remarks about diagram sensitivity, and formulates and proves lemmas on estimating upper and lower bounds on statistical and spectral complexity diagrams for various signal-noise mixtures. Some important patterns of behavior of noise-like and weak signals, and possibilities for their detection under white and blue noise conditions, are identified.

In the listed and numerous other works devoted to two-dimensional information analysis methods using information diagrams, the possibility of solving detection and classification problems is demonstrated. While the authors' article [9] presents an alternative interpretation of these methods, which is that statistical complexity diagrams (like some other two-dimensional information diagrams) are in one-to-one correspondence with the number of discrete spectral components of the signal and the signal-to-noise ratio. Quite unexpectedly, it turns out that by calculating the values of information entropy and statistical complexity for the obtained signal-noise mixture, one can simultaneously determine the number of effective spectral discretes and the energy shares of the signal and noise. Nevertheless, this is the case — the article consistently shows, via the construction of analytical grids, that statistical and spectral complexity diagrams can be reduced to a two-dimensional diagram of discreteness and signal-to-noise ratio. Specifically, the mentioned work presents analytical estimates of metric grids for statistical and spectral complexity diagrams. The tool of metric grids (under certain conditions) allows determining a unique correspondence between the tuple of information characteristics (S, C_{TV}) and the tuple of parameters (d – number of discretes in the spectrum, SNR – signal-to-noise ratio). However, when constructing metric grids, it is first necessary to determine what kind of noise is present in the signal-noise mixture (white noise, blue noise, or some other noise). Natural noise, of course, differs from deterministic white, blue, and other noises. In this regard, the construction of so-called adaptive grids is required. This is discussed in the “Adaptive Grids” section of this work.

2. DIAGRAMS OF SPECTRAL AND STATISTICAL COMPLEXITY

In previously published works on information characteristics, the authors focused mainly on new sensitive information criteria, which is important for solving the detection problem [10–12], as well as on constructing metric grids, which play a role in solving the classification problem [3, 9].

The new spectral measure $C_S(p)$ (1), proposed in [10], is especially effective for detecting weak signals with a small number of discretes ($d < 50$ for $N \geqslant 2048$) in a mixture with white noise. [9] provides relevant definitions (essentially, d is the number of harmonic oscillations of equal amplitude in the so-called d-signal). Below, expressions (1)–(4) present a number of notations required for solving the detection/classification problem of weak signals:

$$C_S(p) = -\frac{4}{\log_2 N}\left(\sum_{k=1}^{N} p_k \log_2 p_k\right)\left(\sum_{k=1}^{N} p_k + \frac{K_N}{\ln N} + 1\right), \tag{1}$$

$$C_{TV}(p) = -\frac{4}{\log_2 N}\left(\sum_{i=1}^{N} p_i \log_2 p_i\right)\left(\sum_{i=1}^{N} p_i - \frac{1}{N}\right)^2. \tag{2}$$

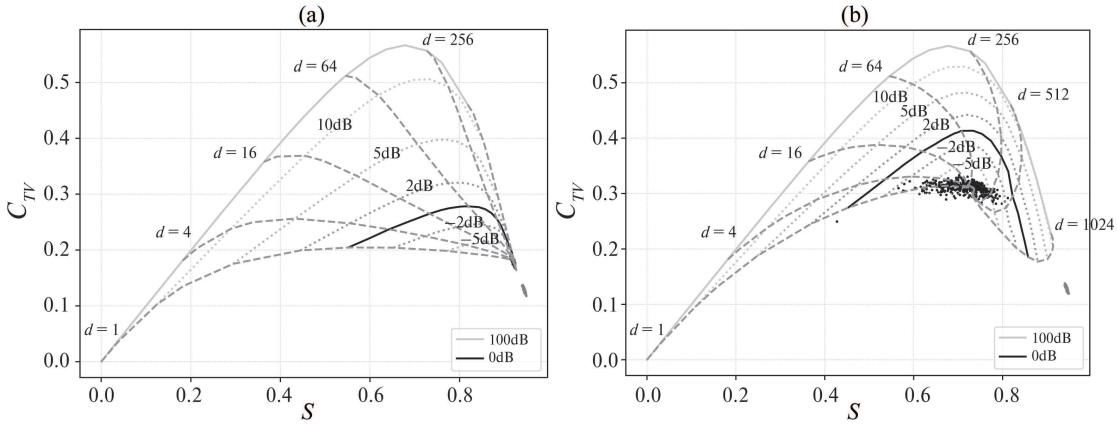

Fig. 1. Comparison of metric grids for blue and pink noises for $N = 2048$ (for white noise: $S = 0.945$, $C_{TV} = 0.125$); (a) – grid for blue noise, (b) – grid for pink noise.

In turn, statistical complexity is the product of information entropy $S(p)$ and statistical disequilibrium D_{TV} ($C_{TV}(p) = S(p) \cdot D_{TV}(p)$):

$$D_{TV}(p) = \frac{N}{4}\left(\sum_{i=1}^{N} p_i - \frac{1}{N}\right)^2, \tag{3}$$

$$S(p) = -\frac{1}{\log_2 N}\left(\sum_{i=1}^{N} p_i \log_2 p_i\right), \tag{4}$$

where N is the length of the spectral data series, p_k is the power spectrum, K_N is a normalization factor close to one and depending on N.

Previously published works have shown that statistical and spectral complexity diagrams are an effective tool for solving detection and classification problems. Moreover, information diagrams based on spectral complexity $C_S(p)$ are less applicable to solving the classification problem than information diagrams based on statistical complexity $C_{TV}(p)$ (2). Thus, an information metric that has an advantage in solving the detection problem is inferior to another metric when solving the classification problem. Therefore, the authors further develop the information diagram method based on statistical complexity, which is more sensitive for determining the number of discretes in the signal spectrum.

Work [9] is devoted to constructing analytical grids for statistical and spectral complexity diagrams (white noise). Note that analytical grids have a great advantage — the possibility of constructing inverse continuous functions and, as a consequence, obtaining a quick result for high-precision estimation of discretes d and the signal-to-noise ratio from the obtained estimates of entropy and statistical complexity. In cases where the noise differs significantly from white, and the signals are weak, it is necessary to build adaptive grids.

Figure 1 shows examples of metric grids on the statistical complexity diagram under blue and pink noise conditions. Note that under pink noise compared to blue or white noise, it becomes possible to "discern" noise-like signals, i.e., classify signals with a large number of discretes, particularly short pulses. Thus, in Fig. 1a (blue noise), signals with the number of discretes $d > 256$ are practically indistinguishable, while in Fig. 1b (pink noise), it becomes possible to distinguish signals with the number of discretes $d > 1024$, i.e., when the number of spectral components of the signal is larger than the size of the entire spectrum (i.e., when the signal is noise-like).

It should also be said that alternative diagrams constructed based on entropy and statistical disequilibrium D_{TV} (3) allow in most cases to more effectively separate discretes and solve the classification problem than diagrams based on statistical complexity C_{TV}. This is because when

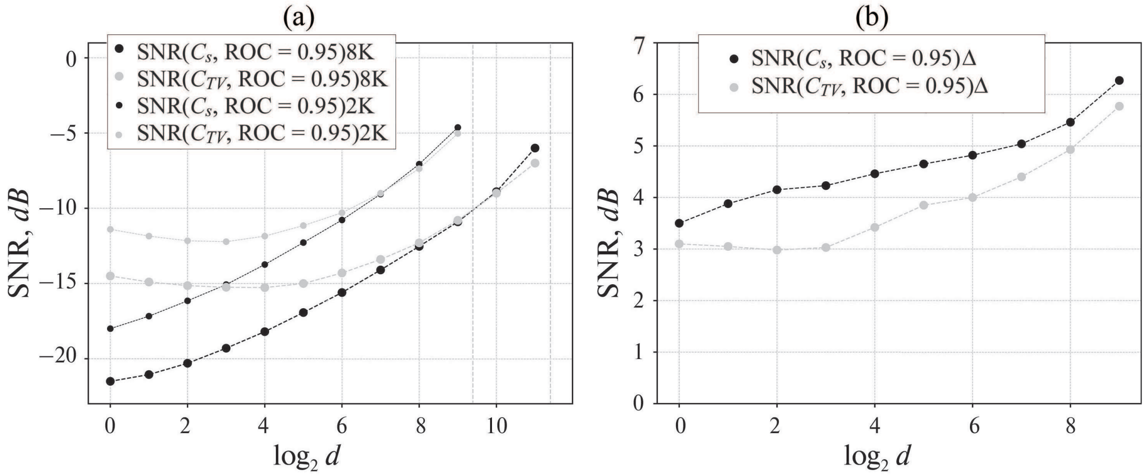

Fig. 2. Comparison of ultimate signal detection capabilities depending on the number of discretes d and spectrum size N; (a) – limiting SNR levels depending on discreteness (AUC = 0.95), (b) – effect of increasing dimensionality (8K vs 2K), change in limiting SNR.

a signal appears in white noise, entropy S (4) and disequilibrium D_{TV} change in opposite directions, and when multiplied ($C_{TV} = S \cdot D_{TV}$), they partially suppress each other. The authors' works consider statistical complexity diagrams because this method is known and often used in practice. Subsequent research will also be devoted to diagrams based on entropy and statistical disequilibrium.

For solving detection and classification problems in practice, the construction of so-called adaptive grids is required. The point is that when solving applied problems, the researcher faces the problem of determining the type of noise, which can differ significantly from white, blue, etc. It is important to have a tool that allows quickly adapting the metric grid to real conditions. This is discussed in more detail later in the article.

Separately, it is worth noting that the spectrum size N determines the capabilities for estimating the number of discretes in the spectrum. Previously, in [3], an estimate was given for determining the maximum number of discretes in white noise: $d < N/4$. Figure 2 presents a comparison of the ultimate capabilities for signal detection via statistical and spectral complexities for $N = 2048$ and for $N = 8192$ depending on the number of signal discretes in white noise. Figure 2b shows a comparison of white metric grids for $N = 2048$ and $N = 8192$. The graphs show that the capabilities for classification and detection of signals using metric grids increase with N [9]. The gray curve demonstrates a decrease in SNR for signals with a small number of discretes by approximately 3 dB when increasing the spectrum size from $N = 2048$ to $N = 8192$ for the statistical complexity diagram and by approximately 4 dB for the spectral complexity diagram. This is an important observation because it becomes clear that with increasing N, the sensitivity of the spectral complexity diagram grows faster than the sensitivity of the statistical complexity diagram. The result of the dependence of the limiting SNR on the number of discretes in the spectrum was obtained via numerical experiment (ROC = 0.95).

3. ADAPTIVE GRIDS

In practice, the construction of so-called adaptive grids is required, which can be obtained as follows: from accumulated observations along the bearings of a phased array where a minimum of energy is observed (conditionally, no signal), the average entropy S and statistical complexity C_{TV} are estimated. Then, noise is generated such that in each window, the mathematical expectation of entropy equals S, and the mathematical expectation of statistical complexity equals C_{TV}. This noise is mixed with simple signals of different numbers of discretes (a combination of sinusoidal

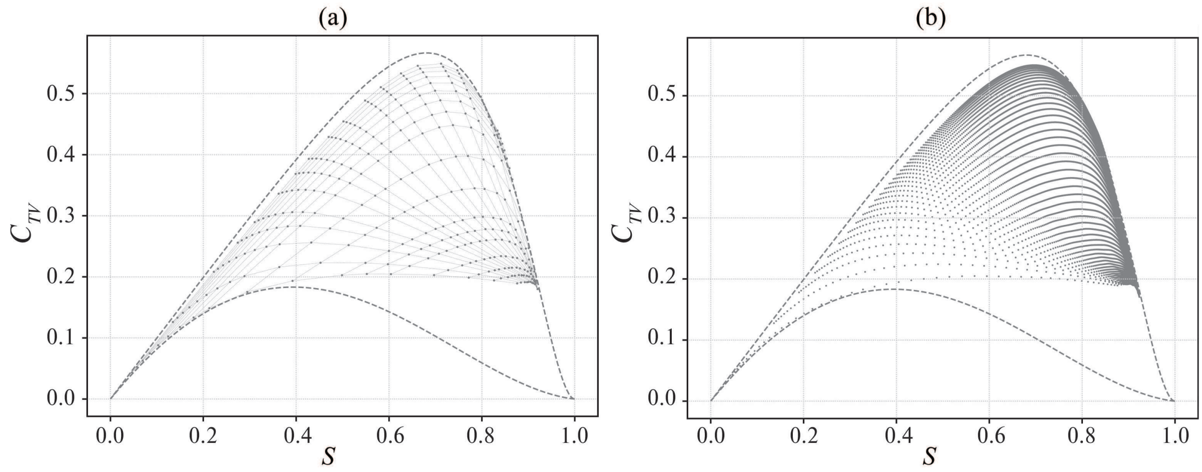

Fig. 3. (a) Estimation of blue grid nodes and (b) calculation of spline grid for the statistical complexity diagram ($N = 2048$).

signals of equal amplitude) to obtain d-signals with given characteristics. Then, by averaging over 200 cases, the entropy and complexity of each node of the metric grid are estimated (each node is characterized by the number of discretes d and the signal-to-noise ratio level SNR). Then, a spline grid is constructed based on the nodes of the coarse grid (obtaining a detailed grid). This method of constructing metric grids is mentioned in [9] for comparing analytical grids with numerical experiments.

Figure 3a presents a metric grid for the statistical complexity diagram ($N = 2048$) for detecting and classifying signals in blue noise, constructed according to the algorithm described above for adaptive grids. The grid nodes are defined by the following discreteness values and signal-to-noise ratios: $d \in \{1, 2, 4, 6, 8, 12, 16, 20, 28, 36, 52, 68, 96, 128, 192, 256, 384, 512\}$, $\text{SNR} \in \{-10, -7, -5, -3, -1, 0, 1, 3, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$. Each node is built by averaging over 200 cases (calculating the grid took about 20 minutes). The Appendix provides an example of constructing an analytical metric grid for blue noise.

Figure 3b presents a metric spline grid, which is built on the basis of the coarse grid 3,a (the figure shows a grid of size 500×50) using two-dimensional interpolation RectBivariateSpline (Python also has a simpler analogue, interp2d). Two-dimensional spline interpolation when constructing a metric grid significantly increases its level of detail (under limited calculation time), which, in turn, allows estimating discreteness and the signal-to-noise ratio with high accuracy [9].

4. FEATURES GENERATED BY DISCRETENESS AND SIGNAL-TO-NOISE RATIO

The classification problem is associated with various information criteria. As shown in [9], through information characteristics such as entropy and statistical complexity, one can obtain estimates of the discreteness level d and the signal-to-noise ratio SNR, which, in turn, generate a whole family of strong classification criteria. Let's focus on some of them.

1) **Correlation level of d and SNR.** A high level of correlation between discreteness and the signal-to-noise ratio within a certain time interval may indicate a change in the distance to the object.
2) **Difference in discreteness** (absolute and relative) between the mean discreteness value d of the signal-noise mixture, measured over k windows (duration of each window τ), and the D estimate of the discreteness of the signal-noise mixture, measured over the entire time window $k\tau$: $D - d$, $D/d - 1$. The level of this difference allows determining how stable the discrete frequencies are. Such a criterion can be useful, for example, in determining heart rate variability.

3) **Discreteness levels and signal-to-noise ratio** for the signal-noise mixture after subtracting the spectral envelope. More on methods of envelope subtraction and line spectrum extraction is discussed in [13].
4) **Energy-weighted frequency** of the d main spectral discretes (having the highest energy). This indicator can be useful for estimating the frequency of the peak of cavitation noise.
5) **Spectrum localization level** and other criteria.

Let us explain the concept of "Spectrum localization level" in more detail. As already mentioned, the discreteness diagram [9] has unique properties that allow extracting additional information for solving the classification problem. In particular, according to the discreteness diagram, the researcher gets the opportunity to formalize the search for the number of "effective" discretes d in the spectrum. This, in turn, allows building a distribution for this random variable. Note also that discreteness d and the signal-to-noise ratio, being unique characteristics of the signal, are also the most important classification features, but this is far from all. One of the most important criteria generated by discreteness is the spectrum localization level. The essence of this criterion is to determine the indices of the d discretes that were identified when forming the discreteness criterion d through the application of the Discreteness Diagram tool. Let Z be the total distance of the d discretes with the maximum energy from the mean value over the sample of these d discretes (the distance is defined as the modulus of the difference between the index of the discrete in the unordered spectrum and the mean value of all indices of these d discretes). So, let's define the minimum and maximum possible total distance of d discretes from their mean:

$$Z_{\min}(N,d) = \begin{cases} \dfrac{d^2-1}{4}, & d = 2k+1, \\ \dfrac{d^2}{4}, & d = 2k. \end{cases}$$

$$Z_{\max}(N,d) = \begin{cases} \dfrac{d(N-d+3)}{4} + \dfrac{1}{2}\left(N - d - \dfrac{N}{d} - 2\right), & d = 2k+1,\ k > 1, \\ \dfrac{dN}{4}, & d = 2k. \end{cases}$$

Let's define the frequency localization criterion based on the above formulas (the criterion value interval should be within $(0,1)$; for this, we use the estimates $Z_{\min}, Z_{\max}$ for even values of discreteness d). The spectrum localization level is the following expression:

$$K_z = \frac{Z - Z_{\min}}{Z_{\max} - Z_{\min}} = \frac{4Z - d^2}{2(Nd - d^2)}.$$

This criterion allows answering the question of how much the spectrum of the observed signal is localized in a separate spectral region. And the average frequency over the d discretes allows determining the localization region. This criterion allows classifying many types of signals. As a practical application, consider this example: for white noise (as well as for broadband interference or an explosion), the value of this criterion is close to 0.5; for marine vessels, depending on deadweight, this indicator varies within the interval 0.1–0.4; and for the calls of some marine animals, this criterion is close to zero.

5. CONCLUSION

This work is the third in the authors' series of works devoted to metric grids for the complexity diagram. The applied nature of the work is primarily related to the concept of an adaptive grid and the identification of a number of informative classification features. Collectively, all these articles form the method of metric grids and constructing the discreteness diagram, the essence of which is to estimate the number of discretes in the spectrum and the signal-to-noise ratio from a single time window under low signal-to-noise ratio conditions.

FUNDING

The work was partially supported by the Russian Science Foundation grant (project no. 23-19-00134).

APPENDIX

Previously, in [10], an estimate of entropy depending on the size of the spectral series N for white noise was obtained:

$$S(N) \approx 1 + \frac{\gamma - 1}{\ln N} \approx 1 - \frac{0.422}{\ln N}, \tag{A.1}$$

where γ is the Euler–Mascheroni constant.

An approximate estimate of entropy for blue noise, which also gives a good approximation, can be obtained using the formula

$$S(N) \approx 1 - \frac{0.616}{\ln N}. \tag{A.2}$$

It has been experimentally confirmed that this formula for blue noise estimates the true entropy value with an accuracy up to the fourth decimal place ($128 < N < 16\,384$). Figure 4 shows the distribution density of normalized ordered power spectra of white and blue noises for $N = 2048$. Formula (A.2) allows obtaining an approximate metric grid for blue noise:

$$S(q, d, N) = S_0(N)(1 - q) - \frac{q}{\ln N}, \tag{A.3}$$

$$D_{TV}(q, d, N) = \frac{1}{4}\left(\alpha(d, N) + 1 + q + \beta(1 - q^2) + (1 - q)\left(-1 + 2\exp(1 - q^{-1})\right)\right)^2, \tag{A.4}$$

$$C_{TV}(q, d, N) = S(q, d, N) \cdot D_{TV}(q, d, N), \tag{A.5}$$

where $N \geqslant 1000$ is the size of the spectral series, and $S_0(N)$ is the value of information entropy for blue noise $S_0(N) \approx 1 - \frac{0.616}{\ln N}$ (for $N = 2048$ $S_0 \approx 0.9192$), $\alpha(d, N) = -\frac{2d}{N}$, $\beta \approx 0.12$ is an empirically obtained coefficient, q is the energy share of the d-signal in the signal-noise mixture.

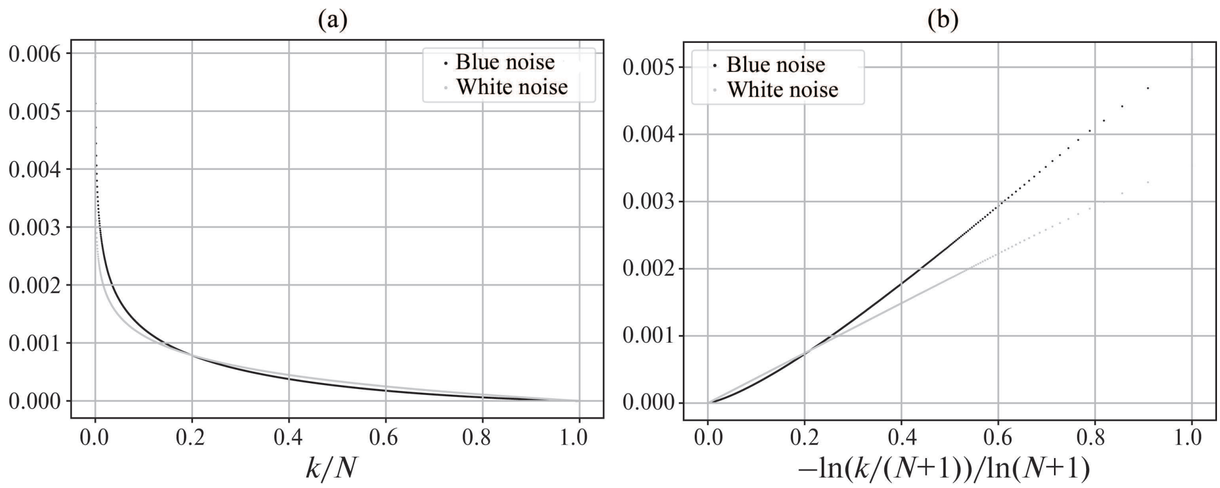

Fig. 4. Explanation for formula (A.1); (a) – ordered spectrum of one realization of a frame of length $N = 2048$ for white and blue noise, (b) – the same realizations for white and blue noises, horizontal axis in logarithmic scale.

REFERENCES

1. Ronald, L.A. and Duncan, W.M., *Signal analysis: time, frequency, scale, and structure*, N.J.: Wiley-IEEE Press, 2004.

2. Shiryaev, A.N., *Probabilistic-statistical methods in decision theory. Modern mathematics*, M.: MTSNMO: NMU, 2020.

3. Babikov, V.G. and Galyaev, A.A., Information diagrams and their capabilities for classifying weak signals, *Probl. Inform. Transmiss.*, 2024, vol. 60, no. 2, pp. 127–140.

4. Kishan, G.M., Chilukuri, K.M., and Huang HuaMing, *Anomaly Detection Principles and Algorithms*, Cham. Springer, 2017.

5. Andrienko, Y.A., Brilliantov, N.V., and Kurths, J., Complexity of two-dimensional patterns, *Eur. Phys. J. B.*, 2000, vol. 15, no. 3, pp. 539–546.

6. Feldman, D.P. and Crutchfield, J.P., Structural information in two-dimensional patterns: Entropy convergence and excess entropy, *Physical Review E.*, 2003, vol. 67, no. 5, pp. 1–12.

7. Bandt, C. and Wittfeld, K., Two new parameters for the ordinal analysis of images, *Chaos*, 2023, vol. 33, no. 7, pp. 1–20.

8. Kottlarz, I. and Parlitz, U., Ordinal pattern-based complexity analysis of high-dimensional chaotic time series, *Chaos*, 2023, vol. 33, no. 7, pp. 1–11.

9. Babikov, V.G. and Galyaev, A.A., Complexity-entropy plane analytical representation, *Probl. Inform. Transmiss.*, 2025, vol. 61, no. 1, pp. 27–40.

10. Galyaev, A.A., Babikov, V.G., Lysenko, P.V., and Berlin, L.M., New spectral measure of complexity and its capabilities for detecting signals in noise, *Dokl. RAN. Matematika, informatika, protsessy upravleniya*, 2024, vol. 518, pp. 80–88.

11. Galyaev, A.A., Babikov, V.G., Lysenko, P.V., and Berlin, L.M., Addition to the article "a new spectral measure of complexity and its capabilities for detecting signals in noise," *Dokl. Math.*, 2024, vol. 518, pp. 89–92.

12. Galyaev, A.A., Berlin, L.M., Lysenko, P.V., and Babikov, V.G., Order Statistics of the Normalized Spectral Distribution for Detecting Weak Signals in White Noise, *Autom. Remote Control*, 2024, no. 12, pp. 49–69.

13. Babikov, O.V. and Babikov, V.G., Estimating the fundamental frequency of a propeller shaft-blade harmonic series using the hilbert transform and autocorrelation, *Mathematical Problems of Control*, 2025, no. 3, pp. 12–23.

This paper was recommended for publication by A.I. Kibzun, a member of the Editorial Board