CORRELATION DEFENSE
FOR QUANTUM RANDOMNESS

N.S. Perminov\textsuperscript{a,b}, O.I. Bannik\textsuperscript{a},
D.Yu. Tarankova\textsuperscript{a}, R.R. Nigmatullin\textsuperscript{a}

\textsuperscript{a}Kazan National Research Technical University named after A.N. Tupolev,
Kazan, 420111 Russia
\textsuperscript{b}Zavoisky Physical-Technical Institute, FRC Kazan Scientific Center,
Russian Academy of Sciences, Kazan, 420029 Russia

Abstract

New nonparametric methods were developed for verification and monitoring of quantum randomness based on the ranged correlation function (RCF) and a sequence of the ranged amplitudes (SRA). RCF analysis of different topology subsamples from the raw data of the prototype of a quantum random number generator on homodyne detection was carried out. It was shown that in the real system there are weak local regression relations, for which it is possible to introduce a robust criterion of significance. Precise SRA identification of the long samples statistics was carried out. The obtained results extend the traditional entropy methods of the useful randomness analysis and open the way for creation of new strict quality quantum standards and defense for physical random number generators.

Keywords: quantum information, ranged correlation functions, RCF defense, quantum randomness, physical random number generator

Introduction

Quantum randomness, as a phenomenon within modern quantum physics [1] and mathematical logic [2], has a very special status, allowing to attribute it to the class of effects in partially deterministic complex systems [3]. In addition, experimental physics has no ideal measuring technique capable of directly determining the “ideal” quantum randomness. Therefore, it is necessary to develop stable methods for nonparametric analysis of quantum time series [4, 5] in order to elaborate the fundamental criteria for the quantitative parameterization of quantum randomness.

In quantum communications [6] and the problems of generating true random numbers [7, 8], the question of the quantitative assessment of the quality of the initial raw randomness is one of the central ones. Common methods for analyzing the raw randomness in quantum systems often come down to identifying the empirical frequencies of long samples and calculating the autocorrelation function, but poorly predict methods for local analysis [9]. For this reason, we believe that new precise methods for the correlation analysis of time series [10] expand the possibilities of error correction in quantum information science at the post-processing stage and will make it possible to enhance the secrecy and defense of communications at the physical level.

In this paper, we developed new non-parametric methods for verifying and monitoring quantum chance based on the ranged correlation functions (RCF) and its analogues using a sequence of ranged amplitudes (SRA). A RCF analysis of various topology subsamples from the raw data of the prototype quantum generator of random numbers
on homodyne detection was performed. The data source in our work is the experimental prototype of a quantum random number generator based on the homodyne detection of vacuum fluctuations of laser radiation [8]. For the studied series, a method of precise SRA identification of the statistics of short samples was proposed, as well as weak local regression relations were found, for which stable criteria of significance were introduced. The results obtained contribute significantly to the traditional entropy analysis methods and promote the development of common quantum standards and defense for physical randomness.

1. Ranging

Any time series \( \{X_k\} \), \( k = 1, 2, \ldots, N \), consisting of real or complex numbers, can be ranged according to the chosen measure by decreasing (or increasing) and get a sequence of ranged amplitudes (SRA) of the form \( \{x_n\} \), \( n = 1, 2, \ldots, N \), where \( n \) is the index in SRA [10]. According to this definition, the sequence of SRA \( \{x_n\} \) is composed of exactly the same elements as the original sequence \( \{X_k\} \); therefore, SRA is a non-invasive (without loss of information) statistical quantitative characteristic of a data sample [10]. SRA is related to the distribution function by the following approximate relation (where \( N \) is the sample size) [4, 5, 11, 12]:

\[
F(x_n; N) = (N + 1 - n(x_n))/N.
\]

(1)

Note also that any (even non-smooth and infinite) statistical functions (statistical averages over the initial sample) of a given sample \( \{X_k\} \) and SRA \( \{x_n\} \) strictly coincide. Mathematically, this can be written as a condition \( G([X_k]) = \sum_k G(X_k) = \sum_n G(x_n) = G([x_n]) \) for any function \( G(x) \) (for example, \( G(x) \) can be some entropy measure). Hence, any entropic measure and sum-functions of the sample (including the SRA) on the information capacity are equivalent or superior to the original SRA. In this sense, SRA analysis significantly expands the possibilities of entropy analysis.

Many of the generalized correlation functions [10] of a pair of samples \( \{X_k\}, \{Y_k\} \) are in fact the prototype of three SRA: SRA \( \{X_k\} \), SRA \( \{Y_k\} \), and SRA of the direct product of samples \( \{w_n^2 = (X_kY_k)_n\} \). We will consider this triple as the base of the ranged correlation functions (RCF). In the future, for all samples by default, we will use the normalized scale – \( \{X_k \to (X_k - \min(X_k))/\text{max}(X_k) - \min(X_k))\} \). Such a normalization, the only mapping that does not destroy the structure of linear regression relations, allows to correctly define the generalized correlation functions [10] on the domain of complex variables, which is necessary to identify nonlinear regressions [3, 9, 10, 12]. Normalization makes it possible to select only one sample \( \{w_n\} \) among the three described SRA to verify redundant correlations based on RCF.

2. RCF analysis of subsamples

Quantitative analysis of internal correlations and randomness can be most effectively implemented on the basis of significantly different topology sub-samples of the initial sample, to which the quality criteria should be presented. For the original sample \( X_j \) of the size \( 2N \), \( j = 1, 2, \ldots, 2N \), we selected four subsets of the same size \( N \): \( X_{1,k} = X_k, X_{2,k} = X_{k+N}, X_{3,k} = X_{2k-1}, X_{4,k} = X_{2k}, k = 1, 2, \ldots, N \). Note that the properties of these four subsamples are homogeneous with respect to time, and two former (1, 2) and latter (3, 4) samples form disjoint covers of the initial sample with different topology. As a result, it is possible to consider the pairs independently: 1 & 2 and 3 & 4. The construction of the pair 1 & 2 is indicative of ultra-long correlations (typical, for example, for mathematical generators of pseudo-random numbers), while
the pair 3 & 4 enables one to “see” local linear regression connections, thereby expanding the possibilities of autocorrelation function for the initial sample of the length 2N.

The motivation for the in-depth study of this four subsamples was a statistically significant (by an order of magnitude) experimental observation of the difference in the Pearson correlation coefficients in the pair 1 & 2 (\( R_{12}^2 = 6.9 \cdot 10^{-4} \)) and pair 3 & 4 (\( R_{34}^2 = 8.4 \cdot 10^{-3} \)) for raw data of the size \( N = 10^6 \) obtained on the prototype of the quantum random number generator on homodyne detection. We discovered a local regression to the challenge of sustainable criteria of significance for nonrandomness in the source sample. The problem may be solved on the basis of the methods of SRA-identification [3] statistics, which, as we have shown earlier, can be effectively applied even to the short samples of quantum data [4, 5].

3. W statistics of the product of samples

Traditional methods for analysis of raw quantum randomness are often reduced to the identification of empirical frequencies or their histograms for long samples [7, 8], which causes certain identification errors associated with the invasiveness of these methods. For raw data obtained as a result of the homodyne detection of quantum randomness [8], it is considered correct to determine normally distributed empirical frequencies. However, a question immediately arises about the distribution of the cumulative frequencies [8], it is considered correct to determine normally distributed empirical frequencies. Therefore, we will further use the normality conjecture in the sense of the parameterization of the empirical frequency distribution by the Gaussian normal distribution. Our calculations show that the accuracy of the normalized SRA fitting due to the error function \( n(x) = (1 + \text{erf}((x - x_0)/dx))/2 \) is higher than the accuracy of the parameterization of the empirical frequency distribution by the Gaussian normal distribution. Therefore, we will further use the normality conjecture in the sense of the error function for other ranged data as well.

The main mathematical task of the traditional correlation statistical analysis of a pair of samples \( \{X_k\}, \{Y_k\} \) is to identify the symmetric relations given by the sum-functions \( G[[X_k; Y_k]] = G[[Y_k; X_k]] \) (in the spirit of entropy analysis), which can be written directly on the basis of the symmetric functions \( G = G_{\text{sym}}[W_k; R_k; R_{k2}] = X_k Y_k; R_k = X_k^2 + Y_k^2 \) or the angular symmetric functions (spherical coordinates) \( G_{\text{angle}}[\phi_k = \text{arcsin}(2W_k/R_k^2)/2; R_k^2] \). On the basis of symmetric functions, we usually further consider the variable-split sum-functions, which leads us in the analysis of \( G_{\text{sym}} \) to independent consideration of the series \( \{W_k\} \) and \( \{R_k\} \) and, hence, their SRA \( \{w_n\} \) and \( \{r_n\} \). In this paper, we restrict ourselves to the consideration of SRA \( \{w_n\} \) built on the normalized series \( \{X_k\} \) and \( \{Y_k\} \). It is important to emphasize that the variance which SRA statistics of the product of samples (SRA \( \{w_n\} \)) in its construction can be associated with the standard Pearson match criterion \( R^2 \) expanding its capabilities to a great extent.

Previously, we found the difference between \( R_{12}^2 = 6.9 \cdot 10^{-4} \) and \( R_{34}^2 = 8.4 \cdot 10^{-3} \) for two different subsamples from the raw data of the prototype of a quantum random number generator on homodyne detection. To see the difference between the subsamples...
1 & 2 and 3 & 4 at the level of W statistic and to demonstrate the sensitivity of RCF analysis, we built and parameterized at the expense of the error function $z = (1 + \text{erf}((w - w_0)/dw))/2$ sets $\{z_n = n/N; \omega_{12,n}\}$ and $\{z_n = n/N; \omega_{34,n}\}$. For the curves in Fig. 1, the fitting accuracy was approximately 0.9993; the obtained parameters of fitting $\{w_0 = 0.5144; dw = 0.1566\}$ for 1 & 2 and $\{w_0 = 0.5304; dw = 0.1524\}$ for 3 & 4 indicate a statistically significant difference between the two pairs of subsamples.

The criterion for the significance of correlations in this case is not one Pearson parameter $R^2$ and two parameters $\{w_0; dw\}$ of the model error function. The sensitivity of the RCF analysis technique can also be increased by improving the model fitting function [3, 10] through introducing an additional degree of freedom $\theta$ in the form of an additional non-extensiveness parameter to the error function of the $z = A + B \cdot \text{erf}((w - w_0)\theta/dw)$ type, which corresponds to the availability of effective memory in time series. However, taking into account additional parameters of fitting with the help of the standard methods of mathematical analysis is a nontrivial task in many practical situations, and the question of the implementation of algorithms for accounting non-extensiveness in statistical distributions requires a separate in-depth study [3, 10].

4. Angle analysis of randomness

An additional method for the correlation analysis of randomness on the basis of symmetric functions is to consider the distributions of the angles $\{\phi_k = \arcsin(2\omega_k^2/R_k^2)/2\}$ and the radii $\{R_k = (X_k^2 + Y_k^2)\}^{1/2}$ constructed from the centered data $\{X_k - X_k - (X_k)\}$. The SRA $\{r_n; z = n/N\}$ distribution is normal in the sense of the error function with a fit accuracy of about 0.9999, and the $\{\varphi_n; z = n/N\}$ angle distribution has the character of a uniform distribution ($\varphi_n \equiv n/N$) with almost the same degree of accuracy. Therefore, to identify potential regression links, we used a more subtle criterion based on the SRA analysis of discrete derivatives of the initial distributions $\{\phi_k^\prime = \phi_{k+1} - \phi_k\}$ and $\{R_k^\prime = R_{k+1} - R_k\}$ characterizing the heterogeneity of the angular distributions.

In Fig. 2, we constructed relative dependencies of the velocities of the radius-angle $\{r_{12,n}^\prime; \varphi_{12,n}^\prime\}$ and $\{r_{34,n}^\prime; \varphi_{34,n}^\prime\}$ for two subsamples 1 & 2 and 3 & 4 of different topology (size $N = 10^9$), for which the structure of the normal distribution was obtained (with the fitting accuracy of 0.9995) with differing adjustable parameters.

Since the distribution of the angles in both pairs 1 & 2 and 3 & 4 has the structure of a uniform random variable, it was possible to represent the series $\{\phi_k; R_k\}$ in a split form $\{\text{sign}(\phi_k^\prime); \text{abs}(\phi_k^\prime); R_k^\prime\}$, where randomness analysis can be carried out simulta-
Fig. 2. The distribution of the inhomogeneity of the angular variables for 1&2 subsamples \( \{ \theta_{12,n} \} (a) \) and 3&4 subsamples \( \{ \theta_{34,n} \} (b) \) (dotted line – experiment, solid line – theory).

Fig. 3. Basic NIST tests [13] for split quantum randomness, given by the inhomogeneities of the angular distributions \( \{ \text{sign}(\theta_{12,k}) \} \) and \( \{ \text{sign}(\theta_{34,k}) \} \) successively for all three components. At the same time, the series \( \{ \text{sign}(\theta_{12,k}) \} \) for both 1&2 and 3&4 contains an extremely small fraction of zeros \( (\sim 2.5 \cdot 10^{-4} \pm 3 \cdot 10^{-6}) \) and consists of numbers \( \{-1; 1\} \), which, after replacing \(-1 \rightarrow 0\), give us a bit sequence. For these sequences, the standard set of NIST cryptographic tests [13] can be applied in two different cases and a \( p \)-value criterion of significance is obtained \( (p\text{-value} > 0.01) \) indicates that the test has passed.

We used subsamples of the length \( N = 10^6 \) and found that a part of quantum randomness, when split off in such a way, satisfies the cryptographic criteria of true random bits, which can be immediately used in quantum cryptography applications [6, 8]. The results of passing the main tests of NIST [13] are shown in Fig. 3.

For the presented NIST tests, the standard notation was used: FREQUENCY – frequency test; BLOCK FREQUENCY – frequency test in Blocks; RUNS – check the “holes”; LONGEST RUNS – check the “holes” in the subsamples; RANK – check of matrix ranks; FFT – spectral test; UNIVERSAL – Maurer’s universal statistical test; CUMULATIVE SUMS – check of cumulative sums. Both sets of binary numbers \( \{ \text{sign}(\theta_{12,k}) \} \) for the subsamples 1&2 and \( \{ \text{sign}(\theta_{34,k}) \} \) for the subsamples 3&4, as shown by additional studies, successfully pass all randomness tests. Thus, we managed to present the initial randomness through its subsamples in such a three-component
form \( \{ \text{sign}(\phi'_k); \text{abs}(\phi'_k); R'_k \} \) that the first component is a true random variable, as verified by the NIST tests, and proved that it is possible for the second and third components (in parallel with the first component) to apply the criteria based on the SRA and RCF methods. These circumstances are important for the implementation of an effective procedure for extracting the final binary randomness from the raw data of a physical random number generator, which will have a regulated structure with the possibility of reliable statistical monitoring of internal security parameters and self-protection of the physical generator of random numbers.

**Conclusions**

The development of general criteria for the quality of randomness in view of the absence of regression equations that clearly distinguish non-randomness remains a task for further investigations. However, now we can present stable multi-parameter intermediate criteria based on the parameterization of the SRA curves, the product of the samples, and the SRA distributions of the angular variables of the subsamples of different topology, which extended the traditional methods of analyzing quantum randomness. A significant advantage of composite angular analysis with splitting of a species \( \{ \text{sign}(\phi'_k); \text{abs}(\phi'_k); R'_k \} \) is the possibility for accurately separating high-quality binary randomness \( \{ \text{sign}(\phi'_k) \} \) from the initial data set and verifying it using the standard set of NIST testing methods [13].

Due to non-invasiveness, the method of SRA and RCF can also be applied in the area of identification of various signal sources [3, 14, 15] and noise [10, 16, 17]. The demonstrated advantages of the ranged analysis open up new possibilities for a quantitative description of the quality of useful quantum randomness and the introduction of universal quantum standards and defense in the area of security of optical and quantum communications.

**Acknowledgments.** The research of noise in the area of photonics and quantum technologies was funded by the Government of the Russian Federation (project no. 14.Z50.31.0040; Feb. 17, 2017 (experimental part)). The work was also supported in part by the grant for young scientists of the Republic of Tatarstan (project no. 06-36-ts-G 2018 (theoretical part)).

**References**

1. Pusey M.F., Barrett J., Rudolph T. On the reality of the quantum state. *Nat. Phys.*, 2012, vol. 8, no. 6, pp. 475–478. doi: 10.1038/nphys2309.
2. Paterek T., Kofler J., Prevedel R., Klimek P., Aspelmeyer M., Zeilinger A., Brukner Č. Logical independence and quantum randomness. *New J. Phys.*, 2010, vol. 12, no. 1, art. 013019, pp. 1–9. doi: 10.1088/1367-2630/12/1/013019.
3. Nigmatullin R.R., Evdokimov Yu.K., Denisov E.S., Zhang W. New methods of complex systems inspection: Comparison of the ADC device in different operating modes. In: *Computational Problems in Science and Engineering*. Springer, 2015, pp. 187–204, doi: 10.1007/978-3-319-15765-8_9.
4. Smirnov M.A., Perminov N.S., Nigmatullin R.R., Talipov A.A., Moiseev S.A. Sequences of the ranged amplitudes as a universal method for fast noninvasive characterization of SPAD dark counts. *Appl. Opt.*, 2018, vol. 57, no. 1, pp. 57–61. doi: 10.1364/AO.57.000057.
5. Perminov N.S., Smirnov M.A., Nigmatullin R.R., Talipov A.A., Moiseev S.A. Comparison of the capabilities of histograms and a method of ranged amplitudes in noise analysis of single-photon detectors. *Comput. Opt.*, 2018, vol. 42, no. 2, pp. 338–342. doi: 10.18287/2412-6179-2018-42-2-338-342.
6. Gisin N., Ribordy G., Tittel W., Zbinden H. Quantum cryptography. *Rev. Mod. Phys.*, 2002, vol. 74, no. 1, pp. 145–195. doi: 10.1103/RevModPhys.74.145.

7. Balygin K.A., Zaitsev V.I., Klimov A.N., Kulik S.P., Molotkov S.N. Implementation of a quantum random number generator based on the optimal clustering of photocounts. *JEPT Lett.*, 2017, vol. 106, no. 7, pp. 470–476. doi: 10.1134/S0021364017190043.

8. Gabriel C., Wittmann C., Sysh D., Dong R., Mauerer W., Anderson U.L., Marquardt C., Leuchs G. A generator for unique quantum random numbers based on vacuum states. *Nat. Photonics*, 2010, vol. 4, pp. 711–715. doi: 10.1038/nphoton.2010.197.

9. Box G.F., Jenkins G.M., Reinsel G.C., Ljung G.M. *Time Series Analysis: Forecasting and Control*. Wiley & Sons, 2015. 712 p.

10. Nigmatullin R.R. The statistics of the fractional moments: Is there any chance to “read quantitatively” any randomness? *Comput. Probl. Sci. Eng. Signal Process.*, 2006, vol. 86, no. 10, pp. 2529–2547. doi: 10.1016/j.sigpro.2006.02.003.

11. Stephens M.A. Introduction to Kolmogorov (1933) on the empirical determination of a distribution. In: *Breakthroughs in Statistics*. New York, Springer, 1992, pp. 93–105. doi: 10.1007/978-1-4612-4380-9_9.

12. Kamps U. A concept of generalized order statistics. *J. Stat. Plann.*, 1995, vol. 48, no. 1, pp. 1–23. doi: 10.1016/0378-3758(94)00147-N.

13. Heckert A., Dray J., Vo S., Rukhin A., Soto J., Nechvatal J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D. A statistical test suite for random and pseudorandom number generators for cryptographic applications. *Special Publication (NIST SP)*, Rep. 800-22 Rev 1a.

14. Spitsyn V.G., Bolotova Yu.A., Phan N.H., Bui T.T.T. Using a haar wavelet transform, principal component analysis and neural networks for OCR in the presence of impulse noise. *Comput. Opt.*, 2016, vol. 40, no. 2, pp. 249–257. doi: 10.18287/2412-6179-2016-40-2-249-257.

15. Umnov A.V., Krylov A.S. Research of sparse representation method for ringing suppression. *Comput. Opt.*, 2016, vol. 40, no. 6, pp. 895–903. doi: 10.18287/2412-6179-2016-40-6-895-903.

16. Latypov R.Kh., Stolov E.L. Theory of ternary jitter-based true random number generators composed of identical gates. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2017, vol. 159, no. 2, pp. 246–262.

17. Nigmatullin R.R. Nanonoise: New non-invasive methods to read random sequences quantitatively. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2007, vol. 149, no. 3, pp. 71–83.

Received
November 11, 2018

**Perminov Nikolay Sergeevich**, Junior Researcher of the Kazan Quantum Center; Junior Researcher of the Quantum Optics and Informatics Laboratory

Kazan National Research Technical University named after A.N. Tupolev ul. K. Marx, 10, Kazan, 420111 Russia

Zavoisky Physical-Technical Institute, FRC Kazan Scientific Center, Russian Academy of Sciences

Sibirsky Tract, 10/7, Kazan 420029 Russia

E-mail: qm.kzn@ya.ru
Корреляционная защита квантовой случайности

Н.С. Перминов1,2, О.И. Банник1, Д.Ю. Таранкова1, Р.Р. Нигматуллин1

1Казанский национальный исследовательский технический университет имени А.Н. Туполева, г. Казань, 420111, Россия
2Казанский физико-технический институт имени Е.К. Завойского ФИЦ Казанский научный центр РАН, г. Казань, 420029, Россия

Аннотация

Предложены новые непараметрические методы верификации квантовой случайности на основе ранжированных корреляционных функций (РКФ) и последовательности ранжированных амплитуд (ПРА). Проведен РКФ-анализ различных по топологии подвыборок из сырьих данных прототипа квантового генератора случайных чисел на гомодинном детектировании. Показано, что в реальной системе существуют слабые локальные регрессионные связи, для которых можно ввести устойчивый критерий значимости, а также выполнена прецизионная ПРА-идентификация статистики длинных выборок. Полученные результаты расширяют традиционные энтропийные методы анализа полезной случайности и открывают путь для создания новых строгих квантовых стандартов качества и защиты для физических генераторов случайных чисел.

Ключевые слова: квантовая информатика, ранжированные корреляционные функции, РКФ-защита, квантовая случайность, физический генератор случайных чисел

Поступила в редакцию 14.11.2018

Перминов Николай Сергеевич, младший научный сотрудник Казанского квантового центра; младший научный сотрудник лаборатории квантовой оптики и информатики Казанский национальный исследовательский технический университет имени А.Н. Туполева ул. К. Маркса, д. 10, г. Казань, 420111, Россия Казанский физико-технический институт имени Е.К. Завойского ФИЦ Казанский научный центр РАН ул. Сибирский тракт, д. 10/7, г. Казань, 420029, Россия E-mail: qm.kzn@ya.ru
Ванник Олег Игоревич, научный сотрудник Казанского квантового центра
Казанский национальный исследовательский технический университет имени
А.Н. Туполева
ул. К. Маркса, д. 10, г. Казань, 420111, Россия
E-mail: olegbannik@gmail.com

Таранкова Диана Юрьевна, студентка кафедры радиоэлектроники и информационно-
измерительной техники
Казанский национальный исследовательский технический университет имени
А.Н. Туполева
ул. К. Маркса, д. 10, г. Казань, 420111, Россия
E-mail: tarankovadyu@ya.ru

Нигматуллин Равиль Рашидович, доктор физико-математических наук, профессор
кафедры радиоэлектроники и информационно-измерительной техники
Казанский национальный исследовательский технический университет имени
А.Н. Туполева
ул. К. Маркса, д. 10, г. Казань, 420111, Россия
E-mail: renigmat@gmail.com

*For citation: Perminov N.S., Bannik O.I., Tarankova D.Yu., Nigmatullin R.R. Correlation defense for quantum randomness. Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki, 2020, vol. 162, no. 1, pp. 98–106. doi: 10.26907/2541-7746.2020.1.98-106.*

*Для цитирования: Perminov N.S., Bannik O.I., Tarankova D.Yu., Nigmatullin R.R. Correlation defense for quantum randomness // Учен. зап. Казан. ун-та. Сер. Физ.-мат. науки. – 2020. – Т. 162, кн. 1. – С. 98–106. – doi: 10.26907/2541-7746.2020.1.98-106*