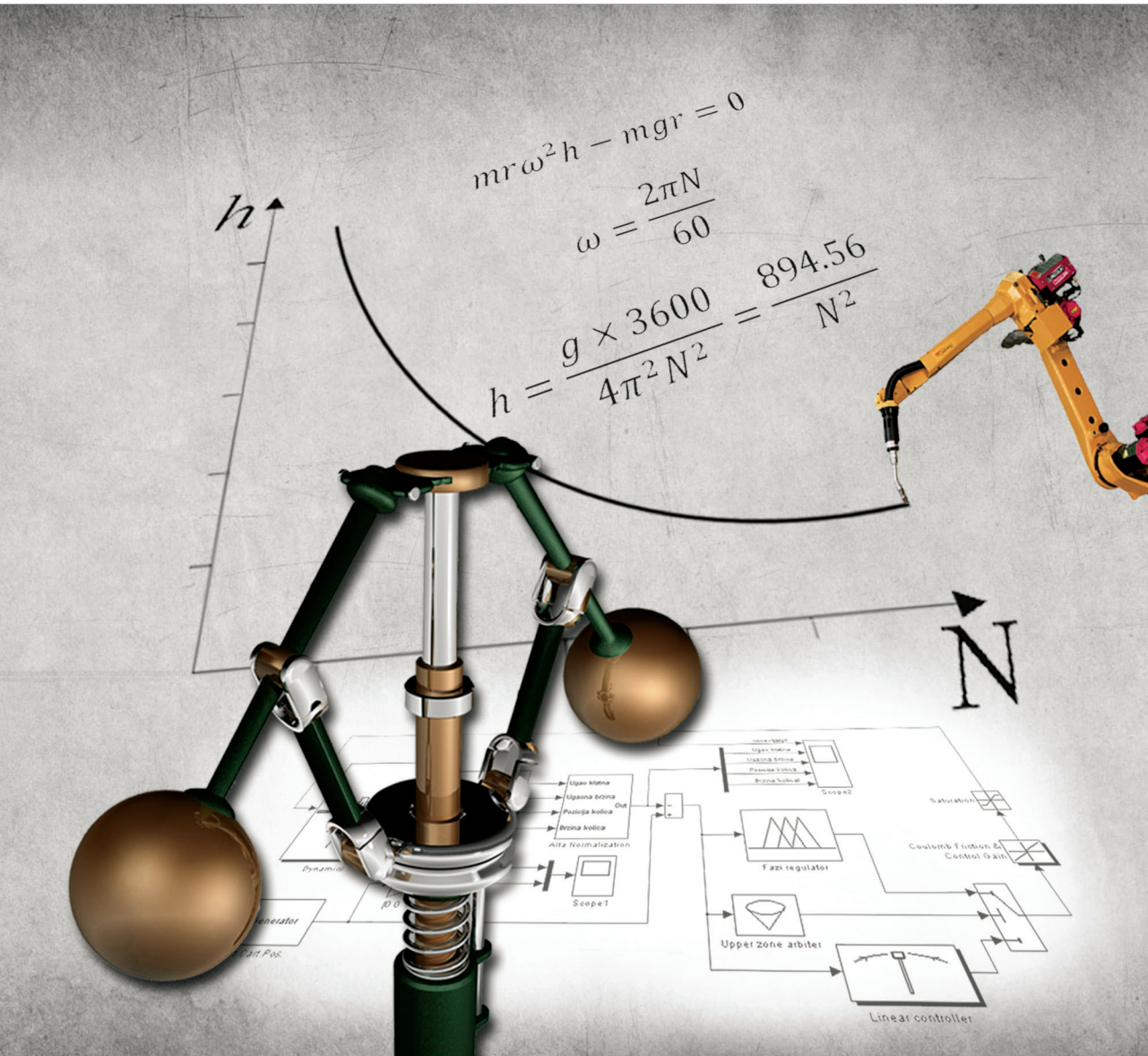
UNIVERSITY OF NIŠ

ISSN 1820 – 6417 (Print)
ISSN 1820 – 6425 (Online)
COBISS.SR-ID 158108940
UDC 62

FACTA UNIVERSITATIS

Series
AUTOMATIC CONTROL AND ROBOTICS
Vol. 21, Nº 1, 2022



Scientific Journal **FACTA UNIVERSITATIS**
UNIVERSITY OF NIŠ
Univerzitetski trg 2, 18000 Niš, Republic of Serbia
Phone: +381 18 257 095 Telefax: +381 18 257 950
e-mail: facta@ni.ac.rs http://casopisi.junis.ni.ac.rs/

Scientific Journal FACTA UNIVERSITATIS publishes original high scientific level works in the fields classified accordingly into the following periodical and independent series:

Architecture and Civil Engineering
Automatic Control and Robotics
Economics and Organization
Electronics and Energetics
Law and Politics
Linguistics and Literature
Mathematics and Informatics
Mechanical Engineering
Medicine and Biology
Philosophy, Sociology, Psychology and History
Physical Education and Sport
Physics, Chemistry and Technology
Teaching, Learning and Teacher Education
Visual Arts and Music
Working and Living Environmental Protection

SERIES: **AUTOMATIC CONTROL AND ROBOTICS**

Editor-in-Chief: **Marko Milojković**, e-mail: *fuacred@junis.ni.ac.rs*
University of Niš, Faculty of Electronic Engineering
Republic of Serbia, 18000 Niš, Aleksandra Medvedeva 14
Phone: +381 18 529 363, Fax: +381 18 588 399

Technical Assistance: **Staniša Lj. Perić**, University of Niš, Faculty of Electronic Engineering,
Department of Control Systems, Niš, Republic of Serbia, e-mail: *fuacrts@junis.ni.ac.rs*

Publication frequency – one volume, three issues per year.
Published by the University of Niš, Republic of Serbia

Financial support: Ministry of Education, Science and Technological Development of the Republic of Serbia
Printed by ATLANTIS DOO, Niš, Republic of Serbia
Circulation 60

Previous title: Scientific Journal FACTA UNIVERSITATIS, Series Mechanics, Automatic Control and Robotics. – Vol. 1, No 1 (1991) – Vol. 6, No 1(2007). – ISSN 0354 – 2009
Since 2007 devided in two Series:
Scientific Journal FACTA UNIVERSITATIS, Series Mechanics. – Vol. 7, No 1(2008) – . – ISSN
Scientific Journal FACTA UNIVERSITATIS, Series Automatic Control and Robotics. – Vol. 7, No 1(2008) – . – ISSN 1820 – 6417

ISSN 1820 – 6417 (Print)
ISSN 1820 – 6425 (Online)
COBISS.SR-ID 158108940
UDC 62

FACTA UNIVERSITATIS

SERIES **AUTOMATIC CONTROL AND ROBOTICS**
Vol. 21, Nº 1, 2022

UNIVERSITY OF NIŠ

INSTRUCTION FOR AUTHORS

As an author, you are kindly advised to use the paper template available for downloading on journal web site (section Download documents). This way you have nothing to change in terms of paper and text format. Simply applying the styles defined in the document will be sufficient.

Paper submitted for publication may be written in English, French or German (preferably in English) and submitted in the final camera-ready form. It is mandatory to submit your original work in Microsoft Word format (.doc not .docx) by using our online manuscript submission system. You have to make "New Submission" and upload your paper by using the online interface. All subsequent versions should be uploaded by using the same paper ID and your defined password. We are unable to process files sent by E-mail. We will do the final formatting and all necessary format conversions of your paper.

Articles are usually 10 to 25 type-written pages long. However, in special cases, shorter or longer articles may be accepted with appropriate reasoning. Author name, affiliation and complete address are to be placed underneath the title. Each paper should be preceded by a brief summary (50-150 words) in the same language. The text should be concise.

Letters, figures, and symbols should be clearly denoted so that no doubts about their meaning can arise. Symbols which may lead to confusion (e.g. letter I and figure 1, figure 0 and letter O) should be distinguished by marks which are explained in "Remarks for the typesetter".

Equations should be typewritten using MathType add-on (http://www.mathtype.com). For equations in your paper (Insert | Object | Create New | or MathType Equation), and, with the number, placed in parentheses at the right margin. Reference to equations should use the form "Eq. (2)" or simply (2). Each formula should occupy one line.

All figures should be numbered with consecutive Arabic numbers, have descriptive captions, and be mentioned in the text. Figures submitted must be of a standard high enough for direct reproduction. Line drawings should be prepared in electronic form. Figures should be planned in advance, so as to allow reduction to 12.75 cm in column width.

References should be quoted in the text by the corresponding number in square brackets and listed at the end of the manuscript in the order as they shown in the paper, in the same way as the following examples (for a journal, book, unpublished paper, proceeding, thesis, user manual, internet document):

[1] B. M. Danković, "A class of almost orthogonal filters," Journal of Circuits, Systems, and Computers, vol. 18, no. 5, pp. 923–931, 2009. [Online]. Available: http://dx.doi.org/10.1142/S0218126609005447

[2] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, 1975.

[3] M. T. Milojković, D. S. Antić, S. S. Nikolić, Z. D. Jovanović, S. Lj. Perić, "On a new class of quasi-orthogonal filters," International Journal of Electronics, [Online]. Available: http://dx.doi.org/10.1080/00207217.2012.743087, to be published.

[4] M.-B. Radac, R.-A. Achimescu, R.-E. Precup, S. Preitl, C.-A. Dragos, A.-I. Stinean, "Design and experiments for model-free PI control of DC drives," in *Proceedings of 8th IEEE International Symposium on Applied Computational Intelligence and Informatics*, Timisoara, Romania, pp. 103–108, 2013. [Online]. Available: http://dx.doi.org/10.1225/sc.2013.018

[5] D. Mitić, Digital variable structure systems based on input-output model. PhD thesis, University of Niš, Faculty of Electronic Engineering, 2006.

[6] Inteco, "The laboratory anti-lock braking system controlled from PC," User's manual, 2008. [Online]. Available: www.inteco.com.pl

[7] MATLAB, The Language of Technical Computing, 2013. [Online]. Available: http://www.mathworks.com/products/matlab [Accessed on December 2013].

Electronic submission. Papers for consideration should be submitted to the Series Editor in electronic form via the Journal's home page: http://casopisi.junis.ni.ac.rs/index.php/FUAutContRob/index.

FACTA UNIVERSITATIS

Series

Automatic Control and Robotics

Vol. 21, Nº 1, 2022

Contents

Regular Papers

FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 21, Nº 1, 2022, pp. 1 - 14
https://doi.org/10.22190/FUACR211220001M

Regular Paper

PERFORMANCE ANALYSIS OF MRC-SC MACRODIVERSITY RECEPTION OVER GENERALIZED FADING CHANNELS

UDC (621,395,38:519.724)

Nenad Milošević, Dejan Milić, Daniela Milović, Jelena Anastasov

University of Niš, Faculty of Electronic Engineering,
Department of Telecommunications, Republic of Serbia

Abstract. *This paper shows a detailed statistical characterization of a specific system configuration consisting of one multibranch maximal-ratio-combining (MRC) and one selection-combining (SC) micro-level base station, and SC back processing unit at macro level. Primarily, the scenario of the independent and identically distributed generalized-K fading channels is investigated. After that, the correlated branches at SC-based micro-level are assumed. The outage probability and the error probability performance for both cases are defined. According to the presented analytical analysis, numerical results are obtained. Also, the impact of the number of MRC and SC input branches, the impact of the fading/shadowing factor, the predefined outage threshold, the average signal-to-noise ratios and the correlation coefficient on the specified system performance is shown. Simulations validate the accuracy of the proposed analytical analysis.*

Key words: *Micro-diversity, macro-diversity, outage probability, error performance, fading, shadowing*

1. INTRODUCTION

In general, wireless channels are inevitably accompanied by multipath fading and shadowing phenomena effects, which consequently impair the reliability and the overall system performance [1]. Multipath fading arises when a signal propagates over different paths until reaching the destination point [2]. The shadowing phenomenon occurs when large obstacles such as massive hills, buildings, walls, or other objects veil the propagation path of the transmitted signal. In the engineering literature, there are many proposed statistical models that describe fading and shadowing, whether individually or simultaneously. The

Received December 20, 2021 / Accepted May 15, 2022
Corresponding author: Nenad Milošević
University of Niš, Faculty of Electronic Engineering, Department of Telecommunications, Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
E-mail: nenad.milosevic@elfak.ni.ac.rs

gamma-shadowed Nakagami-*m* fading i.e., generalized-*K* fading distribution has been proposed in [3], [4] as quite general composite model to define the simultaneous effects of fading and shadowing.

Over the years, problems caused by fading/shadowing have been addressed in a variety of ways. One of the well-known processing methods is the signal combining [5]. There are spatial, time, frequency, and polarization diversity schemes. Spatial diversity schemes, micro- and macro-diversity, are more commonly implemented to combat the impact of fading and shadowing, according to the fact that their application does not require the additional transmit power and bandwidth in comparison to other diversity techniques [6]. Furthermore, micro- and macro-diversity are employed to combat the effect of fading and shadowing simultaneously [7]-[11].

In the selection-combining (SC) scheme, a receiver selects the input branch with the strongest signal and bridges its input to the output. In the maximal-ratio-combining (MRC) technique, the output is formed as a linear combination of the input signals. A better performance is expected to be obtained by the MRC combining scheme in contrast to the SC scheme. In overall, MRC is widely recognized as the optimal combining scheme while the SC receiver is convenient because of its lower complexity and ease of implementation.

In micro-diversity schemes, a number of antennas can be utilized within a single base station with spacing between of an order of a wavelength or even shorter. Thus, micro-diversity antennas can experience different fading conditions and the received signals are usually mutually correlated. Macro-diversity involves the combination signals from two or more base stations that are separated by a certain distance, so the statistical independence of the received signals is easily maintained. In addition, in systems based on the macro-diversity, antennas are positioned at spatially allocated base stations, hence experiencing different shadowing conditions.

A detailed performance analysis in terms of the average channel capacity, the average symbol error probability and the outage probability over composite generalized-*K* fading channels has been obtained in [12]. Authors in [13] have evaluated the level crossing rate and average fade duration at the output of SC type macro-diversity system consisting of two multibranch MRC micro-diversity receivers in the presence of the correlative Nakagami-*m* fading. In [14], the moment generating function-based performance analysis of multibranch MRC diversity receiver over various modulation schemes in the generalized-*K* fading channel has been shown. The utilization of the threshold SC-based macro-diversity to overcome shadowing for outdoor wearable communications has been systematically investigated in [15]. The second order statistics of macro SC based diversity system for radio-frequency vehicle-to-infrastructure communications over interference limited fading environment have been shown in [16].

In this paper, we consider the standard MRC processing and so the general form of the receiver output and the initial steps in the performance evaluation are well-known. However, the macro-diversity layout creates a new channel structure which is far more complex than the micro-diversity channel. Hence, the MRC output has a completely new statistical distribution and a novel, more advanced analysis is required for a system performance evaluation. In particular, we consider a combining configuration with multibranch MRC and multibranch SC micro- and SC macro-level combining hierarchy over independent and identically distributed (i.i.d.) fading channels. Generalization to independent *M*-branch MRC and correlated two-branch SC micro level scenario is also presented in this paper. The outage and system error performance have been investigated

for these specific scenarios. In comparison to the works presented in [12]-[14], the system under consideration in our paper has a higher level of complexity, involving the MRC combining at micro-level. Also, a detailed analysis is performed assuming quite general shadowed fading channels, generalized-*K* fading channels, even when the input branches at macro-level receiver are correlated. For an uncorrelated scenario, the analysis is extended for the system configuration when micro-level receivers are equipped by multiple antennas.

In Section 2 the system model under consideration is explained and an equivalent non-hierarchical logical schematic of configuration is shown. A statistical analysis of *M*-branch MRC and *N*-branch i.i.d. micro level scenario, as well as the analysis of the correlated two-branch SC micro level scenario is given in Subsection 2.1. The expressions for the outage probability and average bit error rate (BER) analysis for both scenarios are derived in Subsection 2.2. In order to accomplish high generality of presented analytical analysis, the composite generalized-*K* fading channels were assumed. Numerical results and a discussion are indicated in Section 3 and final concluding remarks are listed in Section 4.

2. Problem Formulation

The system model under consideration is shown in Fig. 1. In order to increase the coverage area and to avoid the influence of interferences, base stations (BSs) are allocated in such manner to be closer to each other with the capability to individually cover shorter ranges. Consequently, in the uplink wireless communication in Fig 1, the two nearest BSs can receive copies of the same signal from a mobile station (MS). We assume that one of the BSs is equipped with an MRC microreceiver, while the other BS uses a simpler SC microreceiver. Therefore, there is no need to assume that all BSs are identical, which in turn enables an analysis even in more complex networks of heterogeneous nature. The received signals from both BSs are further processed in the back processing unit (BPU).

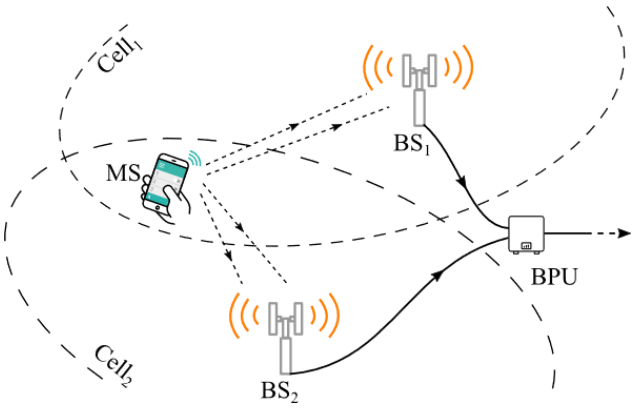


Fig. 1 System model

Simplification of this system model in terms of MRC and SC blocks is given in Fig 2. The input signals of the MRC microreceiver are x_1 and x_2, while the input signals of the SC micro-diversity antennas are x_3 and x_4, respectively. Also, the signal-to-noise ratio (SNR) parameters exist as signal quality measures and input factors for further analysis. The SNR is

a measure that compares the useful signal level to the noise level. The SNRs at the inputs of the MRC microreceivers are γ_1 and γ_2, while the SNR at its output is denoted as γ_5. SNRs at the input of the SC and equivalently – at the inputs of macro level combiner are γ_3 and γ_4, respectively. These three SNRs, γ_5, γ_3 and γ_4, are simultaneously inputs to an equivalent SC macroreceiver (see Fig. 3). Namely, the two 2:1 SCs can be replaced with one equivalent 3:1 SC without affecting the operation of the overall receiving system. Output of the SC macroreceiver located at BPU has a resulting signal whose SNR is denoted as γ_6.

The standard way of analysis concentrates on probability density functions (pdfs) at the output of micro level combiners [7], [8], [13] and continues in the same hierarchical manner to the macro level and its output. Depending on the fading statistics and type of combiners used, an analysis can be tricky and can lead to significant mathematical complexity. The proposed equivalent combiner approach enables a more elegant analysis without compromising the accuracy of the obtained results.

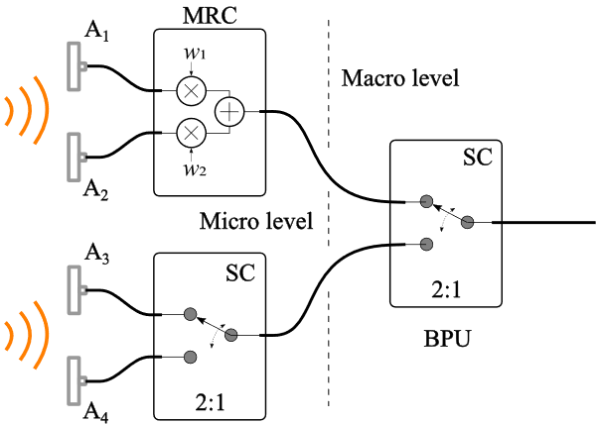


Fig. 2 Combining configuration with micro- and macro-level combining hierarchy

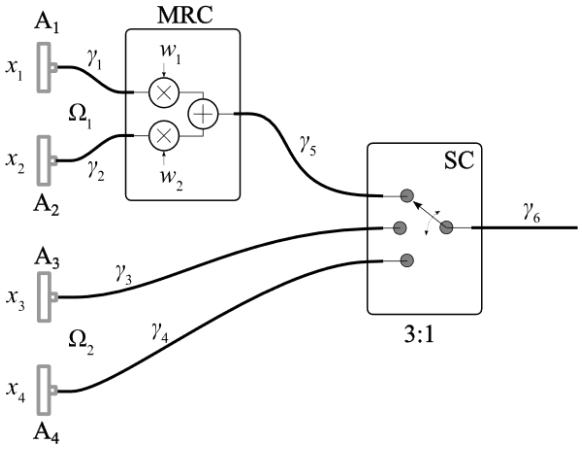


Fig. 3 Equivalent non-hierarchical logical schematic of configuration from Fig. (2). The two 2:1 SCs are replaced with one 3:1 SC.

2.1. Proposed approach in statistic analysis

Macro-level combiner switches the input with highest SNR value to its output, so the output SNR can be expressed as

$$\gamma_6 = \max(\gamma_3, \gamma_4, \gamma_5) \tag{1}$$

Using the conditional probability, the pdf of the output SNR value can be derived as [12]

$$p_{\gamma_6}(x) = p_{\gamma_3}(x)F_{\gamma_4}(x)F_{\gamma_5}(x) + p_{\gamma_4}(x)F_{\gamma_3}(x)F_{\gamma_5}(x) + p_{\gamma_5}(x)F_{\gamma_3}(x)F_{\gamma_4}(x) \tag{2}$$

where $F(\cdot)$ designates the cumulative distribution function (cdf). We assume that signals x_3 and x_4 are mutually independent and identically distributed (i.i.d.), so their pdfs and cdfs are $p_{\gamma_{3/4}}(x)$, and $F_{\gamma_{3/4}}(x)$, respectively. Therefore, we rewrite the previous equation as

$$p_{\gamma_6}(x) = p_{\gamma_5}(x)F^2_{\gamma_{3/4}}(x) + 2p_{\gamma_{3/4}}(x)F_{\gamma_{3/4}}(x)F_{\gamma_5}(x) \tag{3}$$

The interesting configuration of this system enables us to formulate the following lemma:

Lemma 1. When the mentioned cdfs of input signals are expressed in a closed form, and are continuously differentiable, then also the resulting pdf at the ouput of macro-level combiner is expressed in a closed form. The same stands for the cdf of output signal.

Proof. Pdfs are already expressed in a closed form, following from the previous derivation. We now proceed to prove this for cdfs also. By definition, the cdf is $F_{\gamma_6}(x) = \int_{-\infty}^{x} p_{\gamma_6}(t)\mathrm{d}t$, and for the specified case it transforms into

$$F_{\gamma_6}(x) = \int_{-\infty}^{x} p_{\gamma_5}(t)F^2_{\gamma_{3/4}}(t)\mathrm{d}t + 2\int_{-\infty}^{x} p_{\gamma_{3/4}}(t)F_{\gamma_{3/4}}(t)F_{\gamma_5}(t)\mathrm{d}t \tag{4}$$

We proceed with integration by parts on the first addend, taking $u(x) = F^2_{\gamma_{3/4}}(x)$, and $\mathrm{d}v(x) = p_{\gamma_5}(x)\mathrm{d}x$. It follows that $\mathrm{d}u(x) = 2F_{\gamma_{3/4}}(x)p_{\gamma_{3/4}}(x)$ and $v(x) = F_{\gamma_5}(x)$. Then the integral can be rewritten as

$$\int_0^x p_{\gamma_5}(t)F^2_{\gamma_{3/4}}(t)\mathrm{d}t = (F^2_{\gamma_{3/4}}(t)F_{\gamma_5}(t))\Big|_0^x - 2\int_0^x p_{\gamma_{3/4}}(t)F_{\gamma_{3/4}}(t)F_{\gamma_5}(t)\mathrm{d}t \tag{5}$$

We have used the property $p_{\gamma_{3/4}}(x) = 0 \text{ for } x \le 0$ to replace the lower limit of integration $-\infty$ with 0.

We notice that the second addend conveniently cancels out with the opposite sign addend form (4). Consequently, we can directly write the resulting cdf in symbolic form as

$$F_{\gamma_6}(x) = F^2_{\gamma_{3/4}}(x)F_{\gamma_5}(x)\,. \tag{6}$$

The result does not depend on the specific type of combiner used in the upper arm, and applies to arbitrary combining technique in this arm, as long as the cdf function of its output is continuously differentiable, as is the case in the majority of fading models.

Corollary 1.1. Single 2:1 SC combiner is obtained when $x_1 = x_2 = 0$. Then $p_5(x) = \delta(x)$ is simply a Dirac δ-function, and $F(x) = h(x)$ is a Heaviside step-function. Then the 3:1 combiner acts as a simple 2:1 combiner with inputs x_3 and x_4. Using (6), the cdf function

at its output is $F(x) = F^2_{\gamma_{3/4}}(x)$, whereas the pdf is simply its derivative: $p_6(x) = \mathrm{d}F(x)/\mathrm{d}x = 2p_{\gamma_{3/4}}(x)F_{\gamma_{3/4}}(x)$.

Corollary 1.2. The 3-inputs SC combiner results are obtained when $p_5(x) = p_{\gamma_{3/4}}(x)$, and $F(x) = F_{\gamma_{3/4}}(x)$. Then the 3:1combiner acts on three i.i.d. signals yielding output $F(x) = F^3_{\gamma_{3/4}}(x)$, and $p_6(x) = 3p_{\gamma_{3/4}}(x)F^2_{\gamma_{3/4}}(x)$.

Corollary 1.3. Using the previous two corollaries, it recursively follows that an *N*-input SC combiner yields $F(x) = F^N_{\gamma_{3/4}}(x)$, and $p_6(x) = Np_{\gamma_{3/4}}(x)F^{N-1}_{\gamma_{3/4}}(x)$.

Corollary 1.4. SC macro-diversity using an *M*-input MRC micro-diversity and an *N*-input SC micro-diversity: It follows that the output cdf is $F(x) = F^N_{\gamma_{3/4}}(x)F_{\gamma_5}(x)$, whereas pdf is $p_6(x) = p_{\gamma_5}(x)F^N_{\gamma_{3/4}}(x) + Np_{\gamma_{3/4}}(x)F^{N-1}_{\gamma_{3/4}}(x)F_{\gamma_5}(x)$. Here γ_5 refers to SNR at the output of the *M*-branch MRC micro combiner.

Remark 1.1. Results also apply to a case where signals x_3 and x_4 (see Fig. 3) are coming from outputs of two separate micro combiners of an arbitrary type, that are subjected to i.i.d. fading at an earlier combining stage.

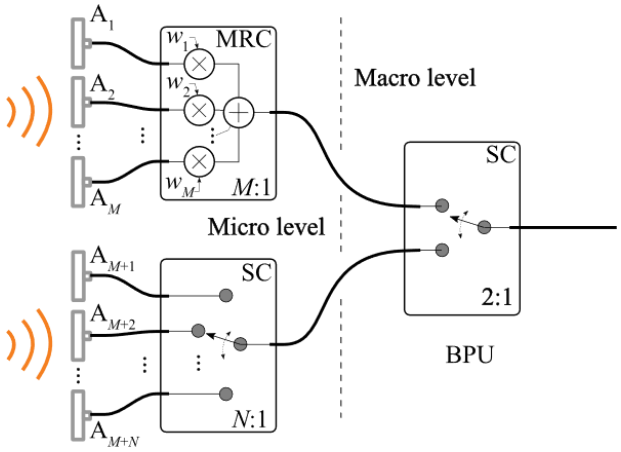


Fig. 4 More general combining configuration with *M*-branch MRC and *N*-branch SC at micro level, corresponding to corollary 1.4.

Lemma 2. Generalization to non-i.i.d. signals at individual BSs.

Proof. Let us assume that the signals at the two base stations are uncorrelated. On the other hand, signals at each antenna on any particular base station may be correlated, and thus non-independent, and also possibly non-i.i.d. In that case, referring to Fig. 3, we can describe signal statistics using joint pdfs and cdfs $p_{\gamma_3\gamma_4}(x,y)$, and $F_{\gamma_3\gamma_4}(x,y)$, respectively. The output SNR is then defined as

$$p_{\gamma_6}(x) = p_{\gamma_3}(x)P(\gamma_4,\gamma_5 < x) + p_{\gamma_4}(x)P(\gamma_3,\gamma_5 < x) + p_{\gamma_5}(x)P(\gamma_3,\gamma_4 < x), \tag{7}$$

where $p_{\gamma_3}(x)$ and $p_{\gamma_4}(x)$ are marginal pdfs of the joint pdf with respect of subscript random variable, and $P(x, y < z)$ represents probability that realizations of random variables *x* and *y* are both lower than value *z*. Using the joint pdf, we can rewrite the previous equation as

$$p_{\gamma_6}(x) = F_{\gamma_5}(x)\int_0^x p_{\gamma_3\gamma_4}(x,t)\mathrm{d}t + F_{\gamma_5}(x)\int_0^x p_{\gamma_3\gamma_4}(s,x)\mathrm{d}s + p_{\gamma_5}(x)\int_0^x\int_0^x p_{\gamma_3\gamma_4}(s,t)\mathrm{d}s\mathrm{d}t\,. \tag{8}$$

The cdf at the output of macro-diversity is

$$\begin{aligned} F_{\gamma_6}(x) = &\int_0^x\int_0^y F_{\gamma_5}(y)\,p_{\gamma_3\gamma_4}(y,t)\mathrm{d}t\mathrm{d}y + \int_0^x\int_0^y F_{\gamma_5}(y)\,p_{\gamma_3\gamma_4}(s,y)\mathrm{d}s\mathrm{d}y \\ &+\int_0^x\int_0^y\int_0^y p_{\gamma_5}(y)\,p_{\gamma_3\gamma_4}(s,t)\mathrm{d}s\mathrm{d}t\mathrm{d}y. \end{aligned} \tag{9}$$

We proceed with analyzing the last addend, via the use of integration by parts. We use $u = \int_0^y\int_0^y p_{\gamma_3\gamma_4}(s,t)\mathrm{d}s\mathrm{d}t$, and $\mathrm{d}v = p_{\gamma_5}(y)\mathrm{d}y$, yielding $\mathrm{d}u = \mathrm{d}y\left(\int_0^y p_{\gamma_3\gamma_4}(s,y)\mathrm{d}s + \int_0^y p_{\gamma_3\gamma_4}(y,t)\mathrm{d}t\right)$ and $v = F_{\gamma_5}(y)$. Then we get

$$\begin{aligned} \int_0^x p_{\gamma_5}(y)F_{\gamma_3\gamma_4}(y,y)\mathrm{d}y &= \left(F_{\gamma_5}(y)F_{\gamma_3\gamma_4}(y,y)\right)\Big|_0^x \\ &\quad -\int_0^x \mathrm{d}y\left[F_{\gamma_5}(y)\left(\int_0^y p_{\gamma_3\gamma_4}(s,y)\mathrm{d}s + \int_0^y p_{\gamma_3\gamma_4}(y,t)\mathrm{d}t\right)\right] \end{aligned} \tag{10}$$

After canceling the addends with opposite signs in (9) and (10), the final result is

$$F_{\gamma_6}(x) = F_{\gamma_5}(x)F_{\gamma_3\gamma_4}(x,x)\,. \tag{11}$$

2.2. Performance analysis at micro and macro-level over generalized-*K* fading channels

In the system under consideration, we assume that radio channels are corrupted by the composite generalized-*K* fading due to presence of simultaneous effects of the multipath fading and shadowing. Micro-diversity basic application purpose is to diminish influences of aforementioned channel phenomena, which justifies the assumption regarding the fading channel and even shows the generality of derived equations according to the generality of the generalized-*K* distribution. Further, mitigation of shadowing requires the use of macro-diversity.

Let the fading envelope over the *l*-th branch be described by the generalized-*K* random variable, r_l, following the probability density function (pdf) as [3]

$$p_{R_l}(x) = \frac{4x^{k_l+m_l-1}}{\Gamma(m_l)\Gamma(k_l)}\left(\frac{m_l}{k_l}\right)^{\frac{m_l+k_l}{2}} K_{k_l-m_l}\left(2x\sqrt{\frac{m_l}{\Omega_l}}\right), \tag{12}$$

with k_l and m_l determining the shadowing severity sharpness and fading depth, respectively, and $K_c(.)$ being a modified Bessel function of the second kind and the *c*th order [17, (8.407)]. $\Gamma(\cdot)$ stands for the Gamma function [17].

The pdf of the instantaneous SNR at the input of *l*th micro-diversity branch can be written in the form [12]

$$p_{\gamma_l}(\gamma)=\frac{2\gamma^{\frac{m+k-2}{2}}}{\Gamma(m)\Gamma(k)}\left(\frac{mk}{\bar{\gamma}}\right)^{\frac{m+k}{2}}K_{k-m}\left(2\sqrt{\frac{mk\gamma}{\bar{\gamma}}}\right) \tag{13}$$

When $k_l \to \infty$, the generalized-*K* distribution reduces to the Nakagami-*m* distribution, and for $k_l \to \infty$ and $m_l = 1$, it reduces to the Rayleigh distribution. For m_l = 1, it coincides with *K*, i.e. the Rayleigh–gamma distribution, and very well approximates the Rayleigh–lognormal distribution. When $k_l \to \infty$ and $m_l \to \infty$, the obtained distribution can describe a channel that is only under the influence of the white Gaussian noise. In addition, the cdf is defined as [12]

$$F_{\gamma}(\gamma)=\frac{1}{\Gamma(m)\Gamma(k)}G_{1,3}^{2,1}\left(\frac{mk\gamma}{\bar{\gamma}}\middle|\begin{matrix}1\\k,m,0\end{matrix}\right). \tag{14}$$

In order to study the outage probability and the average bit error rate (BER) of the overall system, the cdf of SNRs at the output of MRC and SC receivers will be obtained.

The pdf of SNR at the *M*-branch MRC output can be calculated as [14]

$$p_{\gamma_{\text{mrc}}}(\gamma)=2\frac{\gamma^{\frac{mM+k}{2}-1}}{\Gamma(mM)\Gamma(k)}\left(\frac{mk}{\bar{\gamma}}\right)^{\frac{mM+k}{2}}K_{k-mM}\left(2\sqrt{\frac{mk\gamma}{\bar{\gamma}}}\right). \tag{15}$$

Further, the cdf of γ_{mrc} can be obtained by transforming the Bessel *K* function into the Meijer's *G* function [19] and utilizing [20], in the following way

$$F_{\gamma_{\text{mrc}}}(s)=\frac{1}{\Gamma(mM)\Gamma(k)}\left(\frac{mk\gamma}{\bar{\gamma}}\right)^{\frac{mM+k}{2}}G_{1,3}^{2,1}\left(\frac{mk\gamma}{\bar{\gamma}}\middle|\begin{matrix}1-\frac{mM+k}{2}\\\frac{k-mM}{2},\frac{mM-k}{2},-\frac{k+mM}{2}\end{matrix}\right). \tag{16}$$

In the case of presence of correlation among branches, the expressions should be modified. According to the fact that the bivariate pdf of the input correlated SNRs over generalized-*K* fading channels, can be evaluated as [18]

$$p_{\gamma_1\gamma_2}(\gamma_1,\gamma_2)=\frac{4}{\Gamma(m)\Gamma(k)}\sum_{a,b=0}^{\infty}\frac{\rho_n^a}{a!\Gamma(m+a)}\frac{\rho_g^b}{b!\Gamma(k+b)}\frac{\prod_{\ell=1,2}\left(mk\frac{\gamma_\ell}{\sqrt{\bar{\gamma}_\ell}}\right)^{\xi}K_{\psi}\left(2\sqrt{mk\frac{\gamma_\ell}{\sigma_\ell}}\right)}{\gamma_1\gamma_2(1-\rho_n)^{k+a+b}(1-\rho_g)^{m+a+b}}, \tag{17}$$

where $\xi = k+m+a+b$, $\psi = k+b-m-a$, $\sigma_{\gamma_l}=(1-\rho_n)(1-\rho_g)\bar{\gamma}_l$, with ρ_n and ρ_g being correlation parameters. In this section, we analyse the case of an independent *M*-branch MRC and a correlated two-branch SC micro level scenario. Thus, the bivariate cdf of two correlated generalized-*K* SNRs is defined as [18]

$$F_{\gamma_1\gamma_2}(\gamma_1,\gamma_2)=\frac{1}{\Gamma(m)\Gamma(k)(1-\rho_g)^m(1-\rho_n)^k}$$
$$\sum_{a,b=0}^{+\infty}\frac{\rho_n^a}{a!\Gamma(m+a)}\frac{\rho_g^b}{b!\Gamma(k+b)}\frac{1}{((1-\rho_g)(1-\rho_n))^{a+b}} \tag{18}$$
$$\prod_{\ell=1,2}\left(\frac{mk\gamma_\ell}{\overline{\gamma}_\ell}\right)^{\xi/2}G_{1,3}^{2,1}\left(\frac{mk\gamma_\ell}{\sigma_{\gamma_\ell}}\middle|\begin{matrix}1-\frac{\xi}{2}\\ \frac{\psi}{2},-\frac{\psi}{2},-\frac{\xi}{2}\end{matrix}\right).$$

Recalling (6) or (11), the overall outage performance of the system under consideration, can be directly obtained as $P_{out}=F_{\gamma_6}(\gamma_{th})$ for independent i.d. or correlated i.d. branch scenario, respectively, where γ_{th} is the predefined outage threshold. According to the fact that cdf of SNRs at SC and MRC output as well as the cdf of SNR at macro level output are known, we can also evaluate the average BER of the overall system, utilizing

$$P_e=\frac{d^c}{2\Gamma(c)}\int_0^\infty e^{-d\gamma}\gamma^{c-1}F_\gamma(\gamma)\mathrm{d}\gamma\,, \tag{19}$$

where the parameters c and d are defined as (c, d) = (0.5, 1) for coherent binary phase-shift-keying (BPSK) and (c, d) = (0.5, 1) for differential BPSK (DBPSK). By substituting (6) or (11) in (18), the system error performance for two specified scenarios can be evaluated.

The previous integral is very complex and easy–to–follow solution, for both uncorrelated/ correlated scenarios, cannot be obtained. Thus, we determine the average BER approximated expression, for uncorrelated SC branches, in the range from medium-to-high SNRs, by representing the Meijer's G function in (16), into infinite series [21, eq.(07.34.06.0001.01)]. By substituting (16) and (14) into (6), and then (6) into (19), we solved the integral of three Meijer's G functions, with the help of [21, eq.(07.34.21.0081.01)], in the following form

$$P_e=\frac{d^c\overline{\gamma}}{2mk\Gamma(c)(\Gamma(m)\Gamma(k))^2\Gamma(mM)\Gamma(k)}$$
$$\left[\frac{\Gamma(mM-k)}{k}G_{3,1:1,3:0,1}^{1,2:2,1:1,0}\left(\begin{matrix}-k,-k-m,-k\\ -k-1\end{matrix}\middle|\begin{matrix}1\\ k,m,0\end{matrix}\middle|-\middle|\begin{matrix}0\\ \end{matrix}\middle|\frac{mk}{d\overline{\gamma}},\frac{\overline{\gamma}}{dmk}\right)+\right. \tag{20}$$
$$\left.\frac{\Gamma(k-mM)}{mM}G_{3,1:1,3:0,1}^{1,2:2,1:1,0}\left(\begin{matrix}-mM,-mM-m,-mM\\ -mM-1\end{matrix}\middle|\begin{matrix}1\\ k,m,0\end{matrix}\middle|-\middle|\begin{matrix}0\\ \end{matrix}\middle|\frac{mk}{d\overline{\gamma}},\frac{\overline{\gamma}}{dmk}\right)\right],$$

with $G_{p,q:p_1,q_1:p_2,q_2}^{m,n:m_1,n_1:m_2,n_2}(\,)$ being the bivariate Meijer's G function [21].

3. NUMERICAL RESULTS

Recalling previously obtained analytical results various numerical results are presented and discussed in this section. In particular, we have analysed the outage probability and the average BER of the system under consideration under different generalized-*K* fading and shadowing conditions. For the sake of simplicity, results for i.i.d channel conditions are shown, although can be used for evaluating generalized scenarios of correlated and not specifically i.i.d. environmental conditions. All numerical results are accompanied with independent Monte Carlo simulations.

In Fig. 5, the outage probability is plotted as a function of the normalized outage threshold varying number of input branches at macro-diversity SC receiver. The results are evaluated based on corollary 1.4. The increase in number of micro-diversity SC (*N*) and MRC (*M*) input branches evidently decreases the outage probability. It can be noticed that the lower values of the outage threshold indicate lower values of outage probability i.e. better system performance, regardless of the number of input branches.

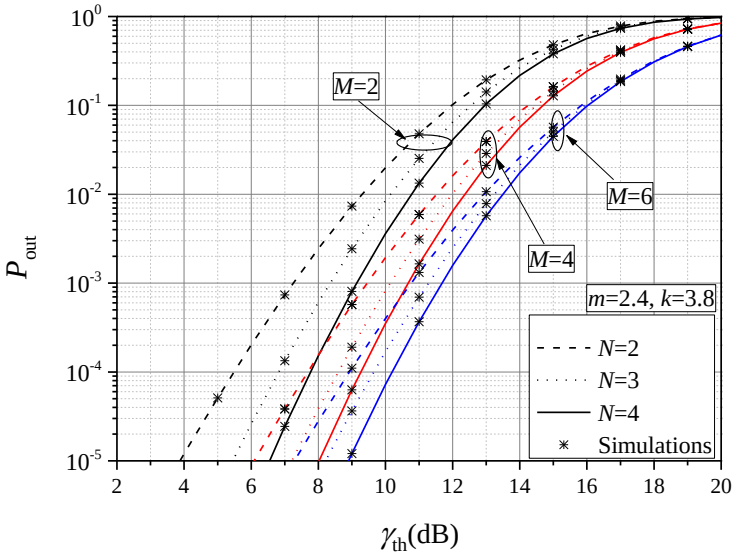


Fig. 5 Outage probability dependence on threshold for uncorrelated scenario

The outage probability dependence on the shadowing shaping factor, *k*, under various fading conditions is presented in Fig. 6. The outage threshold and number of branches at micro and macro levels remain constant. When fading and/or shadowing parameter increases the outage probability value improves. For higher average SNRs, the impact of fading shaping factor *m* is more pronounced. For instance, when $k = 2.5$ and the fading depth decreases i.e., *m* increases from $m = 0.5$ up to $m = 3.5$, the outage probability decreases one order of magnitude for $\bar{\gamma} = 10\text{dB}$ and even three orders of magnitude for $\bar{\gamma} = 15\text{dB}$.

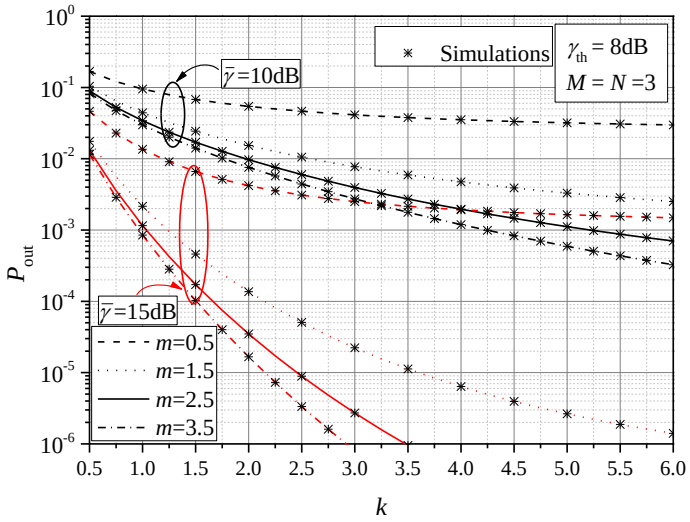


Fig. 6 Outage probability dependence on shadowing shaping factor

Fig. 7 depicts the impact of the average SNR and the fading depth factor (different values of parameter m) on the average BER for two modulation schemes. It can be noticed that the average BER for BPSK signal transmission improves, in comparison to BDPSK signaling. Also, one can notice that changing of the fading depth parameter shows a larger impact on the BDPSK error performance. For instance, by increasing the parameter m from m=0.5 to m=1.5, when $\bar{\gamma} = 16\text{dB}$, the average BER improves for an order of magnitude in the case of DBPSK, and even two orders of magnitude in the case of BDPSK transmission, for the same decrease of fading depth.

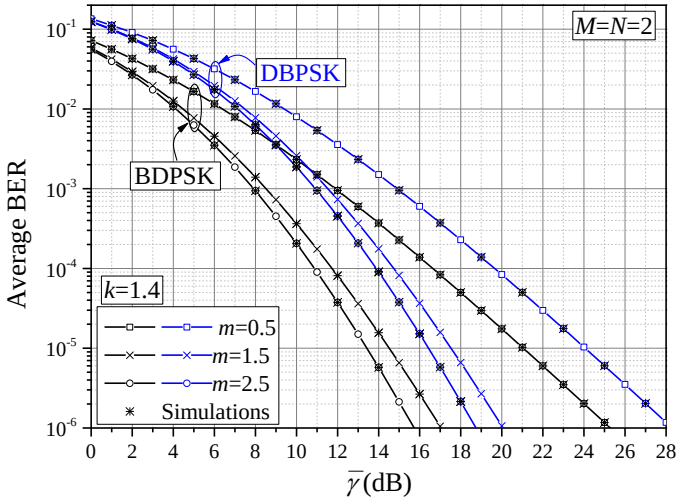


Fig. 7 Average BER versus the average SNR under different fading depth channel conditions

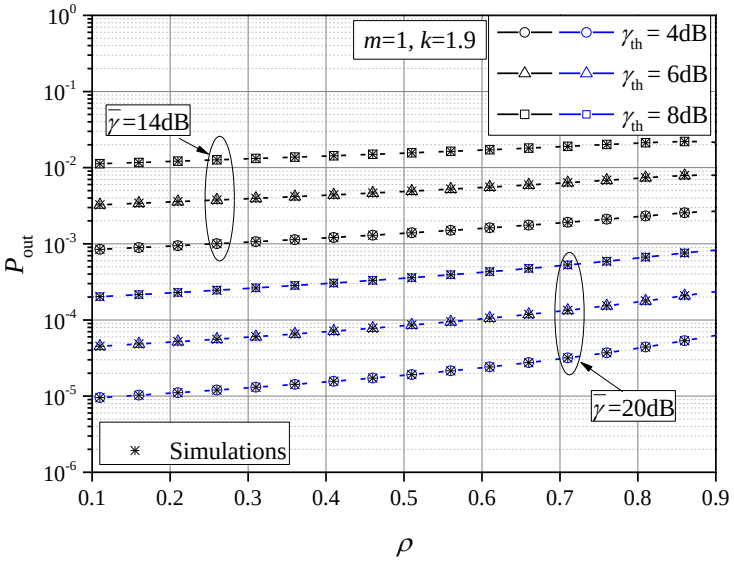


Fig. 8 Outage probability dependence on correlation factor ρ

In Fig. 8, the outage probability dependence on correlation, of the MRC micro level, is shown. The impact of the outage threshold and the average SNR on the outage probability is also depicted. For higher average SNR and/or smaller outage threshold values, better system performance is obtained. It is evident that the correlation coefficient increasing degrades the outage performance. Also, we can notice a more significant influence of correlation coefficient on the outage probability in high SNR regimes (higher slope of blue curves, for $\overline{\gamma} = 20\text{dB}$, is noticeable in comparison to black ones, for $\overline{\gamma} = 14\text{dB}$).

Simulation results show good agreement with analytical results, in all figures.

4. CONCLUSION

In this paper, the outage and error probability performance of a system consisting of MRC and SC receivers at micro- and SC unit at macro-level was investigated. The increase in number of MRC and SC micro-diversity input branches showed to improve the overall system performance. For higher values of fading and/or shadowing parameter, the outage as well as the error performance showed improvement. In the range of higher average SNR values, the impact of fading severity parameter was more pronounced. The average BER improved for the BPSK signal transmission in comparison to BDPSK. For higher average SNR and/or smaller outage threshold values, a better system performance was obtained. Results also showed that the increase in correlation among SC input branches degrades the outage probability values.

Acknowledgement: *This work has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.*

REFERENCES

[1] P. M. Shankar, *Fading and Shadowing in Wireless Systems*.Springer Science Business Media, New York, NY, USA, 2012.

[2] H. Suzuki, “A statistical model for urban radio propagation,” IEEE Transactions on Communications, vol. 25, no. 7, pp. 673–680, 1977. https://doi.org/10.1109/TCOM.1977.1093888

[3] I. Kostic, “Analytical approach to performance analysis for channel subject to shadowing and fading,” IEE Proceedings - Communication, vol. 152, no. 6, pp. 821-827, 2005. https://doi.org/10.1049/ip-com:20045126

[4] P. S. Bithas, N. C. Sagias, P. T. Mathiopoulos, G. K. Karagiannidis and A. A. Rontogiannis, "On the performance analysis of digital communications over generalized-K fading channels," IEEE Communications Letters, vol. 10, no. 5, pp. 353-355, 2006. https://doi.org/10.1109/LCOMM.2006.1633320.

[5] C. B. Dietrich, K. Dietze, J. R. Nealy, and W. L. Stutzman, “Spatial, polarization, and pattern diversity for wireless handheld terminals,” IEEE Transactions Antennas Propagations, vol. 49, no. 9, pp. 1271–1281, 2001. https://doi.org/10.1109/8.947018

[6] J. Boutros and E. Viterbo, “Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel,” IEEE Transactions on Information Theory, vol. 44, no. 4, pp. 1453–1467, 1998. https://doi.org/10.1109/18.681321

[7] A. A. Abu-Dayya and N. C. Beaulieu, "Performance of micro- and macro-diversity on shadowed Nakagami fading channels," in *Proceedings of the IEEE Global Telecommunications Conference GLOBECOM '91: Countdown to the New Millennium. Conference Record*, Phoenix, AZ, USA, pp. 1121-1124 vol.2, 1991, https://doi.org/10.1109/GLOCOM.1991.188549

[8] P. Shankar, “Analysis of microdiversity and dual channel macrodiversity in shadowed fading channels using a compound fading model,” AEU - International Journal of Electronics and Communications, vol. 62, no. 6, pp. 445449, 2008. http://dx.doi.org/10.1016/j.aeue.2007.06.008

[9] P. M. Shankar, "Macrodiversity and Microdiversity in Correlated Shadowed Fading Channels," IEEE Transactions on Vehicular Technology, vol. 58, no. 2, pp. 727-732, Feb. 2009. https://doi.org/10.1109/TVT.2008.926622

[10] P. M. Shankar, "Outage Probabilities of a MIMO Scheme in Shadowed Fading Channels with Micro- and Macrodiversity Reception," IEEE Transactions on Wireless Communications, vol. 7, no. 6, pp. 2015-2019, 2008. https://doi.org/10.1109/TWC.2008.070053

[11] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "Performance Analysis of Macrodiversity MIMO Systems with MMSE and ZF Receivers in Flat Rayleigh Fading," IEEE Transactions on Wireless Communications, vol. 12, no. 5, pp. 2240-2251, 2013. http://doi.org/10.1109/TWC.2013.032113.120798

[12] P. S. Bithas, N. C. Sagias, P. T. Mathiopoulos, G. K. Karagiannidis and A. A. Rontogiannis, "On the performance analysis of digital communications over generalized-K fading channels," IEEE Communications Letters, vol. 10, no. 5, pp. 353-355, 2006, https://doi.org/10.1109/LCOMM.2006.1633320.

[13] D. Stefanovic, S. Panic, P. Spalevic, “Second-order statistics of SC macrodiversity system operating over Gamma shadowed Nakagami-m fading channels,” AEU - International Journal of Electronics and Communications, vol. 65, no. 5, pp. 413–418, 2011. https://doi.org/10.1016/j.aeue.2010.05.001

[14] V. K. Dwivedi, G. Singh, “Moment Generating Function Based Performance Analysis of Maximal-Ratio Combining Diversity Receivers in the Generalized-K Fading Channels,” Wireless Personal Communications, vol. 77, no. 3, pp. 1959–1975, 2014. https://doi.org/10.1007/s11277-014-1618-1

[15] S. K. Yoo, S. L. Cotton, W. G. Scanlon, and G. A. Conway, “An experimental evaluation of switched combining based macro-diversity for wearable communications operating in an outdoor environment,” IEEE Transactions on Wireless Communications, vol. 16, no. 8, pp. 5338–5352, 2017. https://doi.org/10.1109/TWC.2017.2709298

[16] C. Stefanovic, S. Veljkovic, M. Stefanovic, S. Panic and S. Jovkovic, “Second Order Statistics of SIR based Macro Diversity System for V2I Communications over Composite Fading Channels,” in *Proceedings of the 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, Jalandhar, India, pp. 569-573, 2018. https://doi.org/10.1109/ICSCCC.2018.8703293

[17] I. S. Gradshteyn, I. M. Ryzhik, *Tables of integrals, series, and products, fifth edition*. Academic Press, New York, USA, 1994.

[18] P. S. Bithas, N. C. Sagias and P. T. Mathiopoulos, “The bivariate generalized-K (K_G) distribution and its application to diversity receivers,” *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2655-2662, 2009. https://doi.org/10.1109/TCOMM.2009.09.080039.

[19] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series: Vol. 3: More Special Functions*. CRC Press, New York, NY, USA, 1992.

[20] S. Adamchik and O. I. Marichev, “The algorithm for calculating integrals of hypergeometric type functions and its realization in REDUCE system,” in *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, New York, NY, USA, pp. 212–224, 1990. https://doi.org/10.1145/96877.96930

[21] The Wolfarm Functions Site, 2008. [Online] Available: http://functions.wolfram.com/

FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 21, Nº 1, 2022, pp. 15 - 24
https://doi.org/10.22190/FUACR211223002R

Regular Paper

LINEAR RECURRENCE RELATONS AND ORDINARY GENERATING FUNCTIONS APPLIED ON MODELING PROCESSES IN CONTROL THEORY

UDC (512.62:004)

Branislav Ranđelović[1], Saša Nikolić[2], Aleksandra Milovanović[2], Ivana Ilić[3]

[1]University of Niš, Faculty of Electronic Engineering, Department of Mathematics, Republic of Serbia
[2]University of Niš, Faculty of Electronic Engineering, Department of Control Systems, Republic of Serbia
[3]University of Niš, Faculty of Medicine, Department of Mathematics and Computer Science, Republic of Serbia

Abstract. *In this paper we apply multistep recurrence relations, as one of very simple and useful mathematical models. It is an efficient tool for solving many problems in mathematics, science, and technics. We also use generating functions, as a connection between real number sequences and real functions, and as a very smooth and efficient connection between the discrete mathematics and (continual) mathematical analysis. We present an application of multistep homogenous linear recurrence relations for modelling some processes in the control theory. Further on, we use the ordinary generating function aiming to find appropriate formulae for calculating members of an appropriate recurrence sequence. Finally, we show the application of this novel mathematical approach on one real example in the control theory.*

Key words: *Recurrence relation, generating function, control theory*

1. INTRODUCTION

There are a lot of problems and processes in mathematics, science, technology and other fields, where the recurrence relation is the most appropriate mathematical model for describing it (see [1]-[3]). It is sufficient that some problem or process can be described

Received December 23, 2021 / Accepted March 22, 2022
Corresponding author: Branislav Ranđelović
University of Niš, Faculty of Electronic Engineering, Department of Mathematics, Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
E-mail: branislav.randjelovic@elfak.ni.ac.rs

or valued with some sequence of real or complex numbers, for example values calculated, or sampled or measured in discrete moments in time, and we can use such a mathematical model. Those values must be mutually connected via some relation (see [4]-[6]), expressed with an explicit mathematical formula. This relation usually involves $k (k \in N)$ consecutive members of a real sequence, and it can be used for calculating the next value in a sequence, based on k previous values (see [6]-[8]).

A typical problem with recurrence relations is to determine an explicit formula for calculation of any member of sequence $a_n = a(n)$ (see [4]-[6]). The most common approach is using the characteristic equation of a given recurrence relation, which is precise, but not always easy to implement and not appropriate enough for the algorithmic approach and for programming.

In [10], we already introduced and showed another approach, using the ordinary generating function for a sequence of numbers (see [4], [9], [11]-[13]).

2. MATHEMATICAL BACKGROUND

The recurrence relation for some real or complex sequence of numbers $(a_n)_{n\in N}$ is a mathematical term, given with

$$F(a_{n+k},\ a_{n+k-1},\ a_{n+k-2},\ldots,\ a_n) = 0, \tag{1}$$

(see, for example, [1], [6]) which is the relation in an implicit form, or

$$a_{n+k} = f\left(a_{n+k-1}, a_{n+k-2},\ldots, a_n\right), \tag{2}$$

which is the relation in an explicit form. In both formulas, n are index of this sequence and k is order of this relation. If mappings F in (1) and f in (2) are linear, then we have a linear recurrence relation. Without loss of generality, as in [10], in the rest of this research paper, we will consider only linear recurrence relations (see [1], [5]- [6]).

The algorithm for calculating members of the real sequence is shown on Fig. 1:

```
read (a0, a1… ak-1);
for (i=k, i<=n, i++)
{ai = f(ai-1, ai-2, …, ai-k)=0;     % find next sequence member
 write (ai);
 for (j=0, j<i, j++)
    aj = aj+1;                      % relocate sequence members
}
```

Fig. 1 Algorithm in meta language (see [13])

The recurrence relation is a very useful model in both cases, for real and for complex sequences. In this paper, without loss of generality, we will consider only real sequences.

In order to show our method and main results, also without loss of generality, we will use recurrence relations with small values $k = 2$ (two-step recurrence relation) or $k = 3$ (three-step recurrence relation). For a successful use of the recurrence relation, it is necessary to have k starting values, i.e., to know values of $a_0, a_1, \ldots a_{k-1}$.

The most common approach for obtaining an explicit formula for calculating members of a recurrence relation is **by using the characteristic equation** of a given recurrence relation. For example, if a linear recurrence relation

$$\beta_{n+k}a_{n+k} + \beta_{n+k-1}a_{n+k-1} + \beta_{n+k-2}a_{n+k-2} + \ldots + \beta_n a_n = 0, \tag{3}$$

with starting conditions

$$a_0 = A_0,\ a_1 = A_1,\ \ldots\ a_{k-1} = A_{k-1},\ (A_0, A_1, \ldots A_{k-1} \in R), \tag{4}$$

is given, characteristic equation is

$$\beta_{n+k}\lambda^{n+k} + \beta_{n+k-1}\lambda^{n+k-1} + \beta_{n+k-2}\lambda^{n+k-2} + \ldots + \beta_n\lambda^n = 0,$$

i.e.,

$$\beta_{n+k}\lambda^{k} + \beta_{n+k-1}\lambda^{k-1} + \beta_{n+k-2}\lambda^{k-2} + \ldots + \beta_n = 0,$$

Solutions of this algebraic equation are real or complex numbers $\lambda_1, \lambda_2, \ldots \lambda_k$ and the formula for calculating any member of this sequence have a form

$$a_n = C_1\lambda_1^n + C_2\lambda_2^n + \ldots C_k\lambda_k^n. \tag{5}$$

Constants $C_1, C_2, \ldots C_k$ can be obtained from starting conditions (4).

Another approach (see [10]) to a problem of obtaining the formula for calculation of members of the recurrence sequence is **by using the ordinary generating function**, defined with (see [6], [9], [12])

$$F(t) = \sum_{n=0}^{\infty} a_n \cdot t^n\ .$$

For a linear recurrence relation (3), we can obtain ordinary generating function using following steps

$$\beta_{n+k}a_{n+k} + \beta_{n+k-1}a_{n+k-1} + \beta_{n+k-2}a_{n+k-2} + \ldots + \beta_n a_n = 0, \quad / \ t^{n+k}$$

$$\beta_{n+k}a_{n+k}t^{n+k} + \beta_{n+k-1}a_{n+k-1}t^{n+k} + \beta_{n+k-2}a_{n+k-2}t^{n+k} + \ldots + \beta_n a_n t^{n+k} = 0, \quad / \sum_{n=0}^{\infty}$$

$$\beta_{n+k}\sum_{n=0}^{\infty}a_{n+k}t^{n+k} + \beta_{n+k-1}\sum_{n=0}^{\infty}a_{n+k-1}t^{n+k} + \beta_{n+k-2}\sum_{n=0}^{\infty}a_{n+k-2}t^{n+k} + \ldots + \beta_n\sum_{n=0}^{\infty}a_n t^{n+k} = 0,$$

Then we have

$$\beta_{n+k}\sum_{n=0}^{\infty}a_{n+k}t^{n+k} + \beta_{n+k-1}t\sum_{n=0}^{\infty}a_{n+k-1}t^{n+k-1} + \beta_{n+k-2}t^2\sum_{n=0}^{\infty}a_{n+k-2}t^{n+k-2} + \ldots + \beta_n t^k\sum_{n=0}^{\infty}a_n t^n = 0,$$

$$\beta_{n+k}\sum_{n=k}^{\infty}a_n t^n + \beta_{n+k-1}t\sum_{n=k-1}^{\infty}a_n t^n + \beta_{n+k-2}t^2\sum_{n=k-2}^{\infty}a_n t^n + \ldots + \beta_n t^k\sum_{n=0}^{\infty}a_n t^n = 0,$$

$$\beta_{n+k}(F(t) - a_0 \ldots - a_{k-1}t^{k-1}) + \beta_{n+k-1}t(F(t) - a_0 \ldots - a_{k-1}t^{k-2}) + \ldots + \beta_n t^k F(t) = 0,$$

$$F(t)(\beta_{n+k}+\beta_{n+k-1}t+\ldots+\beta_n t^k)=\beta_{n+k}(a_0\ldots+a_{k-1}t^{k-1})+\ldots+\beta_{n+1}t^{k-1}a_0,$$

and the generating function is

$$F(t)=\frac{\beta_{n+k}(a_0\ldots+a_{k-1}t^{k-1})+\beta_{n+k-1}t(a_0\ldots+a_{k-2}t^{k-2})+\ldots+\beta_{n+1}t^{k-1}a_0}{\beta_{n+k}+\beta_{n+k-1}t+\ldots+\beta_n t^k},$$

where starting conditions are $a_0=A_0,\ a_1=A_1,\ \ldots\ a_{k-1}=A_{k-1},\ (A_0, A_1,\ \ldots A_{k-1}\in R)$. This function is rational and after the decomposition on simple fractions, using well-known summation of geometric series

$$\sum_{n=0}^{\infty}t^n=\frac{1}{1-t}, \tag{6}$$

we can obtain an explicit formula for calculating any member of this sequence, that will have the same form like in (5) (see [10]).

3. METHOD

There are various systems and processes in control theory (as well as in other areas and fields), that we usually divide into two main groups: continual systems and discrete systems. We should also notice a big difference between discrete systems (discrete by their nature) and discretized models of continual systems (see [14]). Anyway, the mathematical apparatus for analysis of discrete systems and analysis of discretized systems are the same. Mathematical models of continual systems are discretized when applying a discrete control [14]. The identification of the process and discretization can be done in two ways: by applying some method for identification of continual processes, and then to discrete it; or to discrete continual model during identification of process. Discrete processes are identified to mathematical models directly. Discrete processes in industry are, in fact, very rare. Their discrete character is a consequence of some discretization mechanism, for example by embedding some discrete measure instruments [14]. The result of such identification is the linear discrete mathematical model

$$y(n)=-\sum_{i=1}^{n}a_i\cdot y(k-1)+\sum_{i=1}^{n}b_i\cdot x(k-1-m)+\eta(k)$$

This model is with noise $\eta(k)$ and delays, where m is the number of delay cycles. Further on, we will deal with processes and systems in an ideal case, so models will be without noise and without delay, in the simplified form

$$y(n)=\sum_{i=1}^{n}a_i\cdot y(k-1).$$

Suppose that we have some process (P) and that this process is described with some real values, calculated in discrete moments of time. Let those values be states and/or outputs of this process. This process can be shown in a form of diagram (see Fig. 2).

State and output of this model in each moment is described with a number a_n, $n = k$, where $k = 0,\ 1,\ 2...$ in particular successive moments t_k. Each subsequent state is

determined by k previous states and mathematical model of this process are described with some relation

$$a_n = f\left(a_{n-1}, a_{n-2}, \ldots, a_{n-k}\right), \tag{7}$$

with starting conditions (4), that is of same type as relation (2).

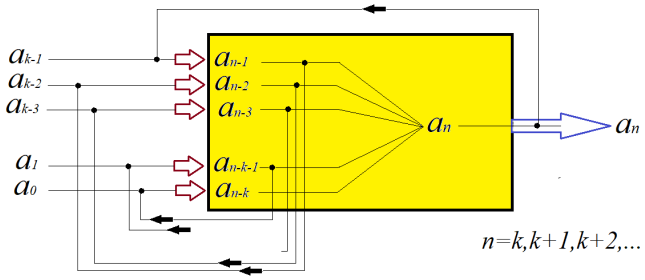


Fig. 2 Block diagram for system characterized with relation (7)

Without loss of generality, we will take $k = 2$, aiming to have the simpler form of recurrence relation, so called "two-step recurrence relation".

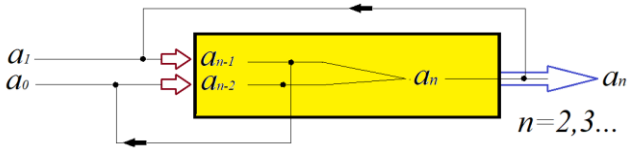


Fig. 3 Block diagram for system characterized with relation with $k=2$

Mathematical model of this process is

$$a_n = f(a_{n-1}, a_{n-2}), a_0 = A_0,\ a_1 = A_1 (A_0, A_1 \in R),$$

or with "shifted" indexes,

$$a_{n+2} = f(a_{n+1}, a_n), a_0 = A_0,\ a_1 = A_1 (A_0, A_1 \in R),$$

or in the implicit form

$$\beta_{n+2} a_{n+2} + \beta_{n+1} a_{n+1} + \beta_n a_n = 0, a_0 = A_0,\ a_1 = A_1 (A_0, A_1 \in R),$$

For a linear recurrence relation (7), we can obtain an ordinary generating function using following steps

$$\beta_{n+2} a_{n+2} + \beta_{n+1} a_{n+1} + \beta_n a_n = 0, \qquad / \, t^{n+2}$$

$$\beta_{n+2} a_{n+2} t^{n+2} + \beta_{n+1} a_{n+1} t^{n+2} + \beta_n a_n t^{n+2} = 0, \qquad / \sum_{n=0}^{\infty}$$

$$\beta_{n+2} \sum_{n=0}^{\infty} a_{n+2} t^{n+2} + \beta_{n+1} \sum_{n=0}^{\infty} a_{n+1} t^{n+2} + \beta_n \sum_{n=0}^{\infty} a_n t^{n+2} = 0,$$

Then we have

$$\beta_{n+2}\sum_{n=0}^{\infty}a_{n+2}t^{n+2}+\beta_{n+1}t\sum_{n=0}^{\infty}a_{n+1}t^{n+1}+\beta_n t^2\sum_{n=0}^{\infty}a_n t^n=0,$$

$$\beta_{n+2}\sum_{n=2}^{\infty}a_n t^n+\beta_{n+1}t\sum_{n=1}^{\infty}a_n t^n+\beta_n t^2\sum_{n=0}^{\infty}a_n t^n=0,$$

$$\beta_{n+2}(F(t)-a_0-a_1t)+\beta_{n+1}t(F(t)-a_0)+\beta_n t^2F(t)=0,$$

$$F(t)(\beta_{n+2}+\beta_{n+1}t+\beta_n t^2)=\beta_{n+2}(a_0+a_1t)+\beta_{n+1}ta_0,$$

and generating function is

$$F(t)=\frac{\beta_{n+2}(a_0+a_1t)+\beta_{n+1}ta_0}{\beta_{n+2}+\beta_{n+1}t+\beta_n t^2}, \tag{8}$$

where starting conditions are $a_0=A_0,\ a_1=A_1, (A_0,\ A_1\in R)$.

This function is rational and after the decomposition on simple fractions, using (6), we have $F(t)$ in an explicit form. We will assume, without loss of generality, that poles of function (8) are real (not complex) and unique, and that all coefficients in (7) are real. Other cases could be a topic for some further research.

4. APPLICATION

In order to validate the proposed method, G.U.N.T. Flow Control Trainer RT522 (Fig. 4) is a comprehensive structure equipped with modern industrial components. The pump delivers water from the tank through a piping system. The fluid flow is measured using an electromagnetic sensor, which allows further processing of the measured quantity by giving a standardized current signal at the output. The flow indicator is a rotometer. An industrial digital controller is used for control. The actuator, connected in a closed non-return loop, is an electromotive valve. The manual ball valve, allows defining the disturbances that are introduced into the system. The controlled parameter K_s and the size to be manipulated and written directly to the two-channel line recorder.

The system also contains management software (RT650.50) connected to a computer. The tank has a capacity of 30*l*, the centrifugal pump has a power of 250*V* with a maximum flow rate of 150 *l/min* and a speed of 2800 *rpm*. The maximum flow rate of the electromagnetic sensor is 6000 *l/h*. The control cabinet contains a power switch, a safety STOP button, a pump start button, a control terminal for monitoring output variables and manual control of the system. The control cabinet also has a printer that automatically prints the values of the output variables. A closer view of the entire system with the tank is shown in Fig. 5. The system is connected via a computer, which supports the *LabView* software package, and with auxiliary applications it is possible to set, control and monitor the operation of the system.

Fig. 4 G.U.N.T. Flow Control Trainer RT522

Fig. 5 A closer view of the entire system

Several experiments have been performed to identify the system in the form of a transfer function. The idle system is excited by setting the desired system response to a value of 1400 *l/h*. Experimental results are given in Fig. 6.

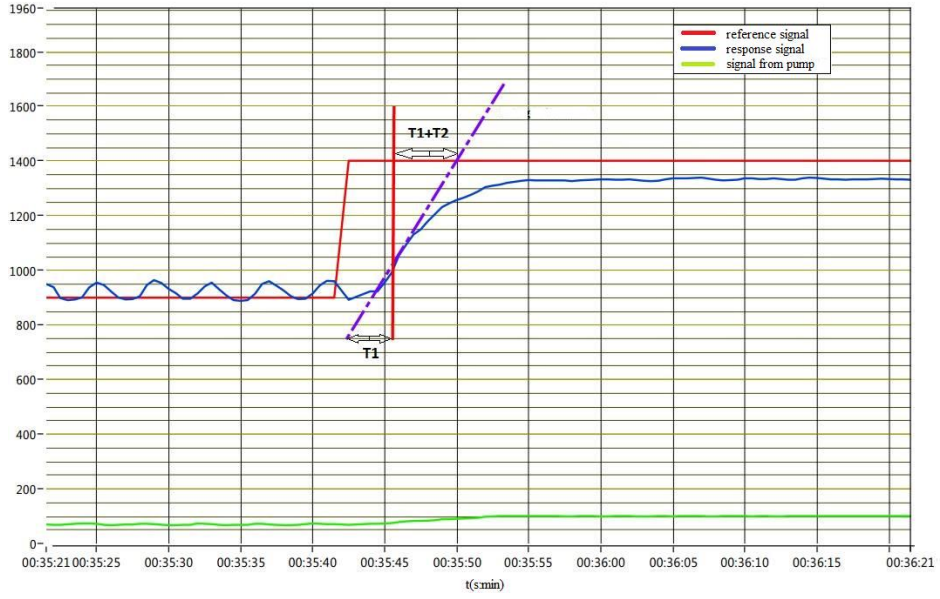


Fig. 6 Experimental results

Using the graph-analytical method [14] of the bounce response of the system for the transfer function, the following function was obtained:

$$W(s)=\frac{1}{5.20386s^2+4.59700s+1}$$
$$5.20386s^2Y(s)+4.59700sY(s)+Y(s)=X(s)$$

which leads us to a recurrence relation

$$5.20386a_{n+2}+4.59700a_{n+1}+a_n=1. \tag{9}$$

Now we will apply the algorithm introduced in [10]. Having in mind that tanks are empty at the beginning, we will assume that $a_0 = 0$, $a_1 = 0$, as start conditions. Although our problem is non-homogenous the linear recurrence relation (right side of equation (9) is not equal to 0), we can apply same method as with homogenous relations, to obtain an explicit formula for a_n.

$$5.20386a_{n+2}+4.59700a_{n+1}+a_n=1 \qquad / \; t^{n+2}$$
$$5.20386a_{n+2}t^{n+2}+4.59700a_{n+1}t^{n+2}+a_nt^{n+2}=t^{n+2} \qquad / \sum_{n=0}^{\infty}$$

and we have

$$5.20386\sum_{n=0}^{\infty} a_{n+2}t^{n+2}+4.59700\sum_{n=0}^{\infty} a_{n+1}t^{n+2}+\sum_{n=0}^{\infty} a_nt^{n+2}=\sum_{n=0}^{\infty} t^{n+2}$$
$$5.20386\sum_{n=0}^{\infty} a_{n+2}t^{n+2}+4.59700t\sum_{n=0}^{\infty} a_{n+1}t^{n+1}+t^2\sum_{n=0}^{\infty} a_nt^n=\sum_{n=0}^{\infty} t^{n+2}$$
$$5.20386\sum_{n=2}^{\infty} a_nt^n+4.59700t\sum_{n=1}^{\infty} a_nt^n+t^2\sum_{n=0}^{\infty} a_nt^n=t^2\sum_{n=0}^{\infty} t^n$$
$$5.20386F(t)+4.59700tF(t)+t^2F(t)=t^2\frac{1}{1-t}$$
$$F(t)(5.20386\ +\ 4.59700\ t\ +\ t^2)=\frac{t^2}{1-t},$$

so ordinary generating function is

$$F(t)=\frac{t^2}{(1-t)(5.20386+4.59700t+\ t^2)}=\frac{-1}{5.20386}\cdot\frac{t^2}{(1-t)(0.38759+t)(0.49579-t)}$$

Suppose that we can make decomposition

$$F(t)=\frac{A}{1-t}+\frac{B}{0.38759+t}+\frac{C}{0.495787-t},$$

then we will come to the system of linear equations

$$\begin{cases} -A+B-C=-0.19216, \\ 0.10820A-1.49579B+0.61241C=0, \\ 0.19216A+0.49579B+0.38759C=0, \end{cases}$$

with solutions $A=0.27466$, $B=0.02355$ and $C=0.10605$. So, we have

$$F(t)=\frac{0.27466}{1-t}+\frac{0.02355}{0.38759+t}+\frac{0.10605}{0.495787-t},$$

$$F(t)=\frac{0.27466}{1-t}+\frac{0.06076}{\left(1-\frac{-t}{0.38759}\right)}+\frac{0.21390}{\left(1-\frac{t}{0.49579}\right)}.$$

Generating function is

$$F(t)=0.27466\sum_{n=0}^{\infty}t^n+0.06076\sum_{n=0}^{\infty}\left(\frac{-t}{0.38759}\right)^n+0.21390\sum_{n=0}^{\infty}\left(\frac{t}{0.49579}\right)^n,$$

$$F(t)=\sum_{n=0}^{\infty}\left(0.27466+\frac{0.06076\cdot(-1)^n}{0.38759^n}+\frac{0.21390}{0.49579^n}\right)t^n.$$

Explicit formula for calculating numbers of real sequence a_n is

$$a_n=0.27466+\frac{0.06076\cdot(-1)^n}{0.38759^n}+\frac{0.21390}{0.49579^n}.$$

We can use the obtained formula for the calculation of any member of this real sequence, i.e., we can calculate value that characterizes our system.

5. CONCLUSION

Multistep recurrence relation is one of the useful mathematical models and also a very simple tool for many problems in mathematics, science and technology. So, there is a possibility to apply multistep linear recurrence relations for modelling problems in the control theory, what is the goal of this paper. An ordinary generating function of real sequence is used, in order to obtain formulae for calculating members of a sequence. Generating functions are just one of mathematical tools for the connection between real number sequences and real functions.

The main purpose of our method is to obtain function, expressed with an explicit formula (continual model) that represents a recurrence sequence of real numbers, which is a problem with a completely discrete nature. This method is some kind of a “D2C” (discrete-to-continual) smooth transformation.

This approach, which is here applied on some simple problems in the control theory, is just an introduction into a wide variety of possible applications for solving problems from

other fields (computer science, economy, biology, digital signal processing…), that have similar mathematical properties. It also opens wide new frontiers for further research in this field of applied mathematics.

REFERENCES

[1] R. Askey, Mourad Ismail, *Recurrence Relations*, Continued Fractions and Orthogonal Polynomials, AMS, 1984.
[2] S. Barnett, *Introduction to Mathematical Control Theory*, 1975.
[3] J. Zabczyk, *Mathematical Control Theory*, BirkHauser, 2010.
[4] J. Matousek, J. Nesetril, *Invitation to Discrete Mathematics*, Oxford University Press, Oxford, UK, 2008
[5] T. Koshy, *Discrete Mathematics with Applications*, Elseive Academic Press, Barlington USA, 2004.
[6] I.Z. Milovanović, et. al., *Diskretna matematika*, Univerzitet Niš, Pelikan Niš, 2000.
[7] W. Gautschi, *Computational Aspects of Three-Term Recurrence Relations*, SIAM Rev., vol. 9(1), pp. 24–82.
[8] A.J. Durán, W.Van Assche, *Orthogonal matrix polynomials and higher-order recurrence relations*, Linear Algebra and Its Applications, vol. 219, no. 1, pp. 261-280, 1995.
[9] G. Levitin, *The Universal Generating Function in Reliability Analysis and Optimization*, Springer-Verlag London Limited, 2005.
[10] B.M. Randjelovic, I.D. Ilic, S.S .Nikolic, V.V. Mitic, *Application od Homogenous Linear Recurrence Relations and Ordinary Generating Functions for Modeling Precesses in Control Theory, Proceedings of the XV International Conference on Systems, Automatic Control and Measurements, SAUM 2021,* Niš, Serbia, September 09.-10., 2021., pp. 62–65.
[11] P. Catarino and P. Vasco, *Modified k-Pell Sequence: Some Identities and Ordinary Generating Function*, Appl.Math.Sci., vol.7, pp. 6031–6037, 2013.
[12] S.K. Lando, *Lectures on Generating Functions*, AMS, 2002.
[13] D. Zeitlin, *Generating Functions for Products of Recursive Sequences*, Transactions of the American Math.Society vol. 116, pp. 300-315, 1965.
[14] S. Nikolic, B. Dankovic, D. Antic, Z. Jovanovic, M. Milojkovic, *Identifikacija procesa*, Elektronski fakultet, Nis, 2020.

FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 21, Nº 1, 2022, pp. 25 - 35
https://doi.org/10.22190/FUACR211223003C

Regular Paper

INVESTIGATING THE IMPACT OF TREE-BASED NETWORK TOPOLOGY ON THE SDN CONTROLLER PERFORMANCE

UDC ((004.3/.4:0.034.2)+004.78)

Danijel Čabarkapa[1], Dejan Rančić[2], Petar Pavlović[1], Miodrag Milićević[1]

[1]Academy of Professional Studies Šabac,
Department of Medical and Business-Technological Studies, Republic of Serbia
[2]University of Niš, Faculty of Electronic Engineering,
Department of Computer Science, Republic of Serbia

Abstract. *Software Defined Networking (SDN) is an important technology that enables a new approach to how we develop and manage networks. SDN divides the data plane and control plane and promotes logical centralization of network control so that the controller can schedule the data in the network effectively through the OpenFlow protocol. The performance and capabilities of the controller itself are important. The impact of network topology type on controller performance can be very significant. In order to have better communication in SDN, it is essential to have an analysis of the performance of specific network topologies. In this paper, we simulate ONOS and RYU controllers and compare their different network parameters under the proposed complex custom Tree-based topology. A network topology has been designed using a Mininet emulator, and the code for topology is executed in Python. From the throughput, packet transmission rate, and latency analysis, the ONOS controller displayed better results than RYU, showing that it can respond to requests more efficiently under complex SDN topologies and traffic loads. On the contrary, the RYU controller provides better results for the less complex SDN networks.*

Key words: *Software defined networking, OpenFlow, software switching, Mininet, ONOS controller, RYU controller*

Received December 23, 2021 / Accepted April 05, 2022
Corresponding author: Danijel Čabarkapa
Academy of Professional Studies Šabac, Department of Medical and Business-Technological Studies, Hajduk Veljkova 10, 15000 Šabac, Republic of Serbia
E-mail: d.cabarkapa@elfak.rs

1. INTRODUCTION

To facilitate new networking evolution, a concept of programmable networks has been proposed. The fundamental idea has evolved into what today is called Software Defined Network (SDN). Compared to traditional networks, SDN decouples the control logic from network layer devices and centralizes it for efficient traffic forwarding and flow management. SDN is having the ability to separate the data and control functions of core devices and consolidates all the control in a single node called the network controller [1]. This centralized entity provides programmable control of the whole network and enables real-time control of all the underlying devices. The SDN controller is comprised of logically centralized "network intelligence" software and has a global view of the network. The controller architecture can be centralized or distributed.

Centralized SDN controllers are mostly used in small-scale SDN networks, whereas distributed controllers can span across multiple domains. A single-threaded SDN controller is more suitable for less complex SDN deployments. In contrast, multi-threaded controllers are suitable for commercial purposes such as 4G/5G, SDN-WAN, ISP, and optical networks.

The controller in a SDN is the core and critical component responsible for making decisions on managing traffic in the underlying network. The core functions of the controller are mainly related to network topology and traffic flow. In SDN networks, Ethernet switches are replaced by OpenFlow switches. Each OpenFlow switch dynamically maintains a flow table, which consists of flow rules that determine the handling of network packets [2]. In an SDN architecture, the infrastructure devices (switches and routers) have been designed to act like modules that process the incoming data packets to forward these packets towards their destinations. The forwarding of these packets is based upon the logic-based set of rules programmed into the SDN controller(s).

Although the fundamental function of an SDN controller is flow management, several different metrics can be used for its performance analysis. In this paper, we presented a performance evaluation in both throughput and latency perspectives for the current well-known OpenFlow controllers: RYU and ONOS. The controller benchmarking tool was implemented for the incremental number of switches connected to the controller under the simulation environment. We have compared RYU and ONOS controllers using the Mininet emulator, along with a detailed analysis of their performance in a custom Tree-based network topology named as Fat-Tree. Topology influence on the SDN network performance is done by analysis and evaluation of different network parameters such as throughput, packet transmission rate, and latency.

From the TCP throughput simulation, the ONOS controller displayed better results than RYU, showing that it can respond to requests more promptly under complex Fat-Tree topology traffic loads. Simulation outcomes indicate that in round-trip propagation delay between end nodes ONOS exhibited better results than the RYU controller. The ONOS controller was found to outperform RYU in the proposed Fat-Tree topology environment regarding the throughput and time between packets sent to the end hosts.

The rest of this paper is organized as follows. Section II presents the SDN and controller architecture and gives an overview of OpenFlow, RYU, and ONOS features. Section III presents the simulation environment, research methodology, and metrics. Section IV shows the results of the SDN controller analysis. The paper is then concluded in Section V.

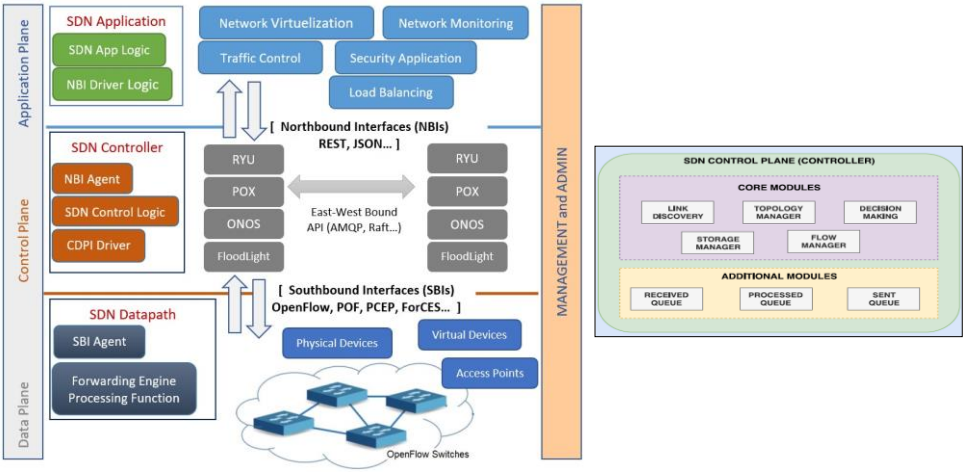


Fig. 1 Overview of a typical layered SDN architecture (left) and SDN controller architecture (right) [3]

2. Background and Related Work

A. SDN Network Architecture

The architecture of a SDN network can be divided into three planes: data plane, control plane, and application plane. The work in [3] discussed a three-layer SDN architecture model, as we can see in Fig. 1 (left). SDN separates the control and data plane of a traditional network device. The control plane is implemented through one or more logically centralized controllers. Control functionality is removed from network devices, that will become simple packet forwarding network nodes. The application plane interacts with the control layer to program the whole network and enforce different policies. The interaction among these layers is done through interfaces that work as communication protocols. SDN Application consists of one Application Logic module and one or more Northbound Interface (NBI) Drivers. The SDN is programmable through applications that interact with the underlying data plane devices. Higher-level logic can be implemented directly through these applications on top of controllers, which communicate through NBI Agents (REST, JSON, etc.) [4]. The SDN Datapath is a logical network device that comprises a Southbound Interface (SBI) Agent and a set of one or more traffic forwarding engines and traffic processing functions. The SBI defined as an interface between controller and Datapath, provides event notification, statistic reporting, capabilities advertisement, and high-level control of all forwarding operations. SBI interface is generally implemented using the OpenFlow protocol.

OpenFlow is the most widely accepted and deployed SBI standard for SDN and represents a protocol that is used for the communication between the controller and forwarding devices. OpenFlow modifies the SDN network in the sense that data plane nodes become simple devices that forward packets according to rules given by the controller. The main components of a SDN network are OpenFlow switches which can communicate with the controller via an OpenFlow protocol. An OpenFlow protocol can handle high-level routing, packet forwarding, and secure connection between the control plane and data plane. OpenFlow switch consists of one or more flow tables. Flow tables

determine data processing and forwarding with the help of flow entries. Each flow entry determines how data will be processed and forwarded in a network. The current version of OpenFlow is 1.5.1. as described in [5].

Fig. 1 (right) shows the architecture of the SDN controller. The core functions of the SDN controller are mainly related to topology and traffic flow, device management, and statistics tracking. All the controller functions are implemented via changeable modules, and the feature set of the controller may be adjusted to specific requirements of SDN networks. The topology itself is maintained by the topology manager. This provides the decision-making module to find optimal paths between nodes of the network. The controller tracks the topology by learning the existence of OpenFlow switches and other SDN devices and tracking the connectivity between them. Currently, there is a variety of open-source SDN controllers available for the community: POX, RYU, FloodLight, ONOS, ODL, OpenDayLight, etc. [6]. There is a lack of a standard for NBI, which has been implemented in several different forms. The ONOS and FloodLight controllers both use a Java and REST/RESTful API, while RYU and POX use Python API, etc. [7].

B. ONOS and RYU Controller

The Open Network Operating System (ONOS) is an open-source distributed SDN control platform, developed by the Open Networking Lab (ON Lab) [8]. The specific protocol feature of the system core makes ONOS available to be used for networks for various purposes such as company networks, campus networks, data center networks, etc. ONOS is specifically oriented to Internet Service Provider (ISP) networks, due to its distributed architecture and natively supports a distributed version of the controller, running on a cluster of servers. Each controller in the cluster is responsible for managing the OpenFlow switches under its domain. Each switch can connect to multiple ONOS controllers for reliability, but only one will be its master with full control over it in terms of read/write capabilities on the switch forwarding tables. The other controllers are denoted as slaves and one of them takes the control of a switch whenever the master controller fails. Generally, ONOS consists of NBI, SBI, and the Distributed Core. As a multi-threaded controller, ONOS is convenient for commercial purposes such as ISP and Data Center networks, SDN-WAN, and optical networks.

RYU controller is an open-source and component-based SDN framework implemented entirely in Python. RYU provides software components with well-defined APIs that make it simple to create control applications and SDN network management. RYU allows an event-driven programming paradigm in which the flow of the program is determined by events, and supports various protocols for managing network infrastructure, such as OpenFlow, Netconf (RFC 6241), OF-config, etc. [9]. The controller uses NBI APIs such as Restful, REST, REST/RPC user-defined API, etc. RYU provides a set of specific components such as OpenStack/Quantum virtualization, Firewall, OFREST, etc. for SDN applications.

C. Related works

Research in [10] gives a comprehensive investigation of open-source controllers RYU, POX, ONOS, and ODL. The authors were focused on parameters such as throughput and latency using Cbench tool. In [11], five controllers (RYU, POX, Trema, Floodlight, OpenDayLight) are compared, and the authors collect properties of each controller under specific evaluation: REST API support, modularity, virtualization, etc. In [12] authors

describe and evaluate two ONOS prototypes. The first version implemented logically centralized global network view, scale-out, and fault tolerance modules. The second version focused on improving ONOS performance in core network traffic engineering and scheduling. In [13] authors use a comparison of performance metrics such as delay, bandwidth, and received packets using network monitoring tools like IPERF and D-ITG to analyze the functionality of the POX controller. The results of this research were the recommendation of using a POX controller.

In [14], the authors have shown the performance comparison performed between the NOX, POX, Trema, and Floodlight in reactive and proactive modes. The results showed that the best performance is achieved when the controller is operating in the proactive mode because forwarding rules are installed on the switch. Authors in [15] presented a framework named HCprobe to compare seven controllers: RYU, Beacon, Maestro, MUL, FloodLight, NOX, and POX. To evaluate the efficiency of these controllers, the authors performed additional measurements like reliability, scalability, and security along with throughput and latency. The results show that FloodLight, Beacon, and MUL obtained minimum latency, while Beacon performed good results in the throughput test. Authors in [16] presented the crucial advantages and challenges of SDN security, flexibility, and performance against traditional TCP/IP networks.

3. Simulation Environment

The simulation hardware and software specifications are shown in Table 1. Performance analysis and network topology development were performed in an environment of a Windows 10 PC. The VirtualBox 6.1.22 hypervisor is used to instantiate two separate Virtual Machines (VM). Each VM is allocated 6 GB of RAM, and runs on Ubuntu 18.04, as the host Operating System (OS). Each VM separately contains Mininet with predefined ONOS and RYU controllers.

Table 1 Simulation hardware and software specifications

		PC	VM
Hardware	Processor	AMD Ryzen 5 3600, 3.6 GHz (6-core)	1 CPU, 6-core
	RAM	16 GB DDR4	6 GB DDR4
Software	OS	Windows 10 (64-bit), ver. 21H1	Ubuntu 18.04 (64-bit)
	VirtualBox	-	6.1.22
	Mininet	-	2.3.0d6
	RYU	-	4.3.2
	ONOS	-	2.5.0
	Python	-	3.8.5

The topology used in the simulation was based on Fat-Tree topology, as shown in Fig. 2. The Fat-Tree topology can be thought of as a reference data center topology. In this research paper, all OpenFlow switches are interconnected to each other, forming 3-level architecture: core, aggregation, and edge. This simulation deals with the results obtained

by Mininet emulating a custom Fat-Tree topology with 12 OpenFlow switches (*s1-s12*) and 32 hosts (*h1-h32*), where the 8 edge switches (*s5-s12*) have four emulated hosts each. The ONOS or RYU controller has the role of the SDN controller in the proposed topology. Host *h1* is denoted as Server, and *h2* is denoted as Client. After controller initialization, Mininet loads a Python script to instantiate the custom topology. Each OpenFlow switch is assigned a unique port for keeping track of network traffic.

Python 3.8 is used to write the network topology. Mininet Command Line Interface (CLI) is used to create the topology. Due to the topology complexity, we used a Mininet high-level API. We use *Topo* as the base class that provides the ability to create reusable and parametrized topology. We use methods *self.addSwitch()* and *self.addHost()* to import switches and hosts into topology and connect them. Further, method *self.addLink(node1, node2,**link_options)* is used for adding a bidirectional link that contains host and switch names and the number of options such as bandwidth or delay.

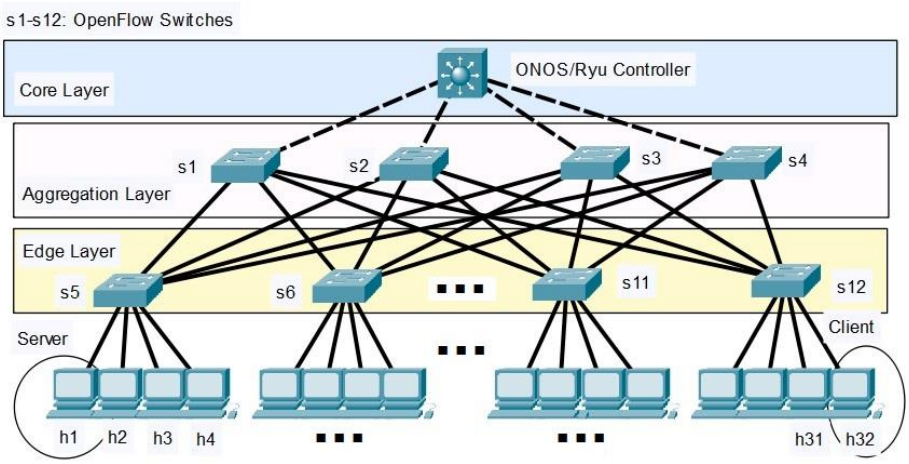


Fig. 2 The Fat-Tree SDN network topology

The Mininet CLI option *--mac* will automatically assign MAC addresses that match the host's names. All the nodes have been assigned a unique IP address from the 10.0.0.0/24 address range and a unique MAC address. To make switches connect to ONOS or RYU controller, we have used localhost 127.0.0.1 loopback IP address. The RYU controller uses port number 6633 to send messages to the OpenFlow switches. ONOS requires the 8181 port number to access the CLI and for GUI and REST API.

Spanning-Tree Protocol (STP) is a link management protocol that provides path redundancy while preventing undesirable loops in the network. Due to 64 links, 12 switches, and 32 hosts in our proposed topology “broadcast storms” are frequent. This undesirable network traffic circles endlessly in the network, due to the destination address in an unknown network. Both ONOS and RYU have a Python script for which the STP function is achieved using OpenFlow.

Latency Parameter: Latency is an important parameter to consider in the operation of a network, especially if it is used to transmit data from applications sensitive to delay or jitter. Round-Trip-Time (RTT) parameter identifies the time the packet spent on the up and downlinks, to and from the switches, and the time as a delay between end nodes. A performance comparison between RYU and ONOS controller in previously defined network topology is achieved by execution of ICMP query (Echo Request and Reply) connectivity test using the *ping* command. A *ping* test is performed between Client and Server hosts (*h1* and *h32* in topology). We make a latency analysis through the average RTT time for the first packet of the flow. We have chosen the average RTT of the first packet of a flow because the first packet going to the SDN network is first processed as a controller. Based on the first packet processing rule, the next packets are processed without connecting to the controller. Therefore, the first packet's RTT time is critical.

Throughput Parameter: Scalability in SDN can be achieved by improving the network throughput and creating a distributed network for data transmission. To evaluate throughput performance in the proposed SDN topology, we have considered TCP traffic between end hosts. A SDN throughput is generally defined as a rate for processing flow requests by the controller. For the performance analysis, we need to generate the TCP traffic between the Client and Server host and log the events using the IPERF networking tool. A typical IPERF output contains a timestamped report of the amount of data transferred through the network. With the IPERF tool, the TCP throughput and data loss are measured by sending and receiving TCP packets between pair of hosts (Client and Server host). Also, the time taken by TCP is calculated for data packets sent and received. For the OpenFlow packet capturing and analysis, we use Wireshark software [17-19].

4. Simulation Results

Considering SDN controller throughput, single-threaded RYU and multi-threaded ONOS show different results in the proposed Fat-Tree topology, where network traffic is significantly intensive. At first, we concluded that TCP throughput depends on the capabilities of the controller itself. The graph in Fig. 3 shows the results obtained by performing transmission between the farthest TCP Client *h1* and TCP Server *h32* in the proposed SDN topology. To measure ONOS and RYU controller throughput, the IPERF test has executed in 75 sec. on the Client, and data have been collected every 1 sec. on the Server host. TCP packets are sent from *h1* to *h32* with a default 1Mb/s sending rate. Let's analyze the observed results. It is calculated the average throughput stays at 18.6 Gbps for RYU, and 29.1 Gbps for the ONOS controller. According to the observation, the average throughput in RYU is 36.08% less than the ONOS controller. The graph also shows that the TCP throughput variations moderately fluctuate within the duration of the simulation. For the ONOS controller, the throughput variations are scientifically uniform to the RYU case. There are a few instances of excessive variations in the throughput for the RYU. Dropping instances were observed frequently between 41-51 sec. of simulation time, leading to degraded performance of the simulation run. This large drop now occurs again at 67 sec. of simulation, and the value of throughput was only 11.3 Gbps.

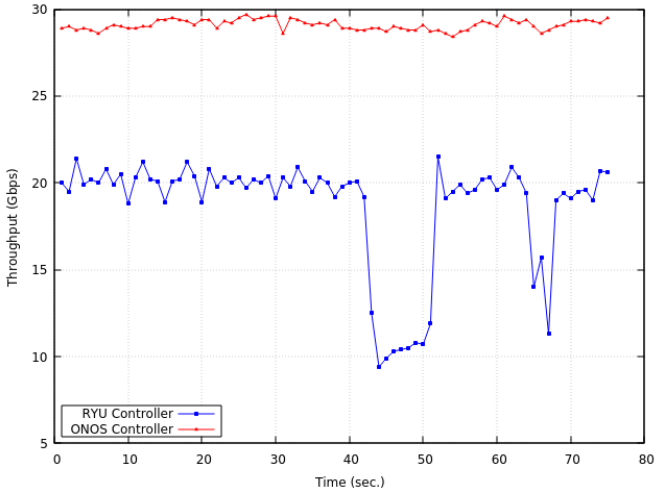


Fig. 3 TCP throughput between the farthest hosts (*h1* and *h32*) for the proposed topology with RYU or ONOS controller

This TCP throughput behavior can be caused by two phenomena. At first, this result shows that the RYU controller has a packet broadcast storm problem when controlling a complex proposed network with loops and a large number of OpenFlow switches and hosts. By default, the STP Python script is not built into the RYU controller. STP function for RYU is achieved using OpenFlow, and we used the *ryu.app.simple_switch_stp_13.py* script to enable STP. On the other hand, it appears that RYU requires more hardware resources (CPU utilization and memory) than ONOS. Here, the ONOS controller exhibits a better throughput performance. This is likely because of its inherent support for very large-scale networks. We decided to do the same simulation three more times to confirm whether this Ryu throughput behavior is accidental or not, and Mininet always generates similar results.

We use Wireshark to represent network traffic between the RYU/ONOS controller and the SDN nodes. Fig. 4 provides the results generated based on packets captured on the proposed Fat-Tree topology. As the controller and switch share the same VM guest the control channel is via the loopback interface, so monitoring the *loopback lo0* interface will give access to these packets. In this simulation, the messaging is using OpenFlow 1.3 so using the filter *openflow_v4* will show the communications between the hosts *h1* and *h32*. We have created a graph of the real-time captured network OpenFlow packets (*ofpt_packet_in*, *ofpt_packet_out*, *ofpt_stats_reply*). From the statistical analysis results, the continuous polling of data causes controller overhead. Fig. 4. shows the I/O graph for the proposed topology with ONOS (above graph), and RYU controller (bottom graph). Y-axis defines the number of transmitted packets, whereas X-axis denotes the time of simulation in seconds. As we can see from the graphs, the ONOS controller enables more network traffic and faster processing of packets. By making use of the I/O graph statistic evaluation, it was found that ONOS has a transmission rate that exceeds 1000 packets/sec, while RYU has a transmission rate that slightly exceeds only 100 packets/sec. This is because having a large number of OpenFlow switches causes conflict at the RYU controller data layer which demands high processing power and a worst packet transmission rate than ONOS. Furthermore, ONOS shows significantly better results in packet processing operations than Ryu, and the main reason is a multithread feature.

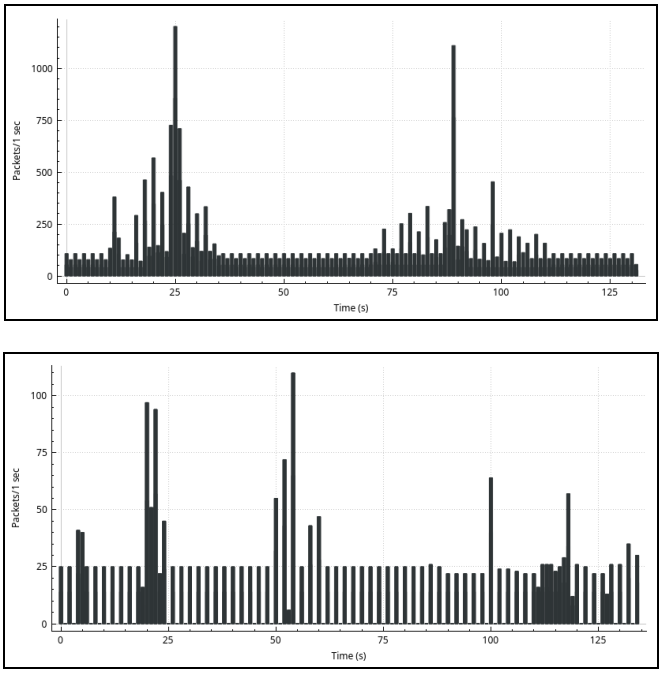


Fig. 4 I/O graph of OpenFlow packets transmitted per second for the proposed topology with ONOS (above) and RYU controller (below)

In the second part of the simulation, we observe the latency against ICMP packet size, from 100B to 1000B. In each performance test, we send 8 ping packets with the following sizes: 100, 500, and 1000 Bytes. We send the *ping* command from Server *h1* to Client *h32*, and the test is performed between the farthest nodes in the proposed topology. These *ping* commands are cycled 10 times for each test. Then we calculated the average values of measured RTTs of the first packet of the flow per test. These average values of the measured RTTs for both ONOS and RYU controllers are shown graphically in Fig. 5. It is visible that the propagation delay between nodes in the proposed topology is very different. From Fig. 5, the average RTT is the minimum for the ONOS controller and ICMP=100B packet size (RTT=7.89 ms). Therewith, the highest RTT value (RTT=15.3 ms) was measured for a RYU topology and ICMP=1000B.

From the obtained results, it is clear that a RYU controller is taking more time for transmission of the packet to its destination. The fact is that the RYU controller introduces larger latency in the network. In the RYU controller, the initial process of establishing network flow consumes time that introduces latency in the network. When the first packet sent by *h1* arrives at the OpenFlow switch, this switch does not know how to route it, encapsulate it, and forwards all the contents of the incoming packet to the controller, being responsible for managing the installation of the flow tables in each switch. From the above results, it can be concluded that longer ICMP packets require a longer processing time for the RYU controller, which affects the overall SDN network performance.

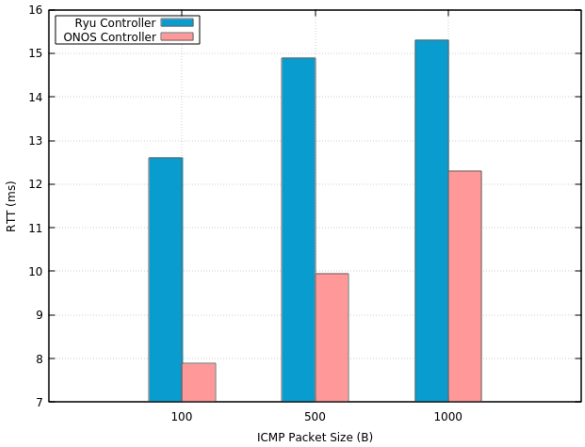


Fig. 5 Average RTT time for the first packet of the flow in proposed topology with RYU and ONOS controller

5. CONCLUSIONS AND FUTURE WORKS

RYU and ONOS are the two most powerful and widely used SDN controllers. There are many researchers currently working on the evaluation and comparison of these controllers. Both of them have their advantages and disadvantages. This paper can help researchers to choose between the RYU and ONOS in different use cases, especially for data centers and complex Tree-Based network environments with a large number of SDN controllers, switches, and links. The results of the simulation proved that ONOS performs better than RYU based on selected parameters.

In total, the ONOS controller was found to outperform RYU in the proposed Fat-Tree topology environment, especially regarding the throughput, time between packets sent to the end hosts, and response received at the OpenFlow switch. More sophisticated embedded ONOS algorithms, distributed architecture, and proactive installation of rules on the whole path that the packet will take, certainly lead to an overwhelming behavior of the ONOS controller in our proposed topology.

This research work opens opportunities for many other research directions. In the future, we plan to keep extending this study with the ONOS cluster multi-controller network environment and with some other SBI APIs. Furthermore, there were different sets of experiments that are left for future research. Some of the variations in experiments that can be conducted in the future to expand the scope of the investigations may include varying the size and/or numbers of the files being communicated. Moreover, it would also be interesting to investigate the results with different sizes of data packets and multi-controller topology in the experiments.

REFERENCES

[1] V. Nguyen, A. Brunstrom, K. Grinnemo, and J. Taheri, "SDN/NFV- Based Mobile Packet Core Network Architectures: A Survey," IEEE Communications Surveys Tutorials, vol. 19, no. 3, pp. 1567–1602, 2017. doi: 10.1109/COMST.2017.2690823

[2] N. Gude, T. Koponen, J. Pettit, "Nox: Towards an Operating System for Networks," ACM SIGCOMM Computer Communication Review, vol. 38, no. 3, p. 105, 2008.

[3] L. Zhu, Md M. Karim, K. Sharif, Fan Li, X. Du, M. Guizani, "SDN Controllers: Benchmarking & Performance Evaluation", ACM Computing Surveys, Vol. 53, Issue 6, Article No. 133, pp. 1–40, 2020, https://doi.org/10.1145/3421764

[4] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford: "OpenFlow: enabling innovation in campus networks", ACM SIGCOMM Comp. Communication Review, vol. 38, no. 2, pp. 69–74, 2008. doi: 10.1145/1355734.1355746

[5] ONF Foundation: "OpenFlow Switch Specification", Version 1.5.1 (Protocol version 0x06), 2015, https://opennetworking.org/wp-content/uploads/2014/10/openflow-switch-v1.5.1.pdf

[6] T. Zhang, F. Hu: "Controller architecture and performance in software-defined networks", in Network Innovation through OpenFlow and SDN, CRC Press, 1st edition, 2014, doi: https://doi.org/10.1201/b16521

[7] W. Zhou, Li Li, Min Luo, Wu Chou: "REST API Design Patterns for SDN Northbound API", 28th International Conference on Advanced Information Networking and Applications Workshops, 2014, doi:10.1109/WAINA.2014.153

[8] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O. Connor, P. Radoslavov, W. Snow: "ONOS: towards an open, distributed SDN OS", in Proceedings of the third workshop on Hot topics in software defined networking, pp. 1-6, 2014, https://doi.org/10.1145/2620728.2620744

[9] Md. T. Islam, N. Islam, Md. Al Refat: "Node to Node Performance Evaluation through RYU SDN Controller", Wireless Personal Comm., issue 1/2020, pp. 550-570, 2020, doi: 10.1007/s11277-020-07060-4

[10] P. Bispo, D. Corujo and R. L. Aguiar: "A Qualitative and Quantitative assessment of SDN Controllers", International Young Engineers Forum (YEF-ECE), pp. 6-11, 2017, doi: 10.1109/YEF-ECE.2017.7935632

[11] R. Khondoker, A. Zaalouk, R. Marx, K. Bayarou: "Feature-based comparison and selection of Software Defined Networking (SDN) controllers", WCCAIS, 2014, doi: 10.1109/WCCAIS.2014.6916572

[12] P. Berde, M. Gerola, J. Hart, Yuta Higuchi, M. Kobayashi, T. Koide, Bob Lantz, B. O'Connor, P. Radoslavov, "ONOS: Towards an Open, Distributed SDN OS", HotSDN '14: Proceedings of the third workshop on Hot topics in software defined networking, pp.1–6, 2014, https://doi.org/10.1145/2620728.2620744

[13] H. M. Noman, M. N. Jasim: "POX Controller and Open Flow Performance Evaluation in Software Defined Networks (SDN) Using Mininet Emulator," 3rd International Conference on Sustainable Engineering Techniques (ICSET 2020), vol. 881, 2020, doi:10.1088/1757-899X/881/1/012102

[14] M. P. Fernandez: "Comparing openflow controller paradigms scalability: reactive and proactive", IEEE 27th International Conference on Advanced Information Networking and Applications (AINA), 2013, doi: 10.1109/AINA.2013.113

[15] A. Shalimov, D. Zuikov, D. Zimarina, V. Pahskov, R. Smeliansky: "Advanced Study of SDN/OpenFlow Controllers," Proceedings of the Central Eastern European Software Engineering Conference CEE-SECR '13, no. 1, pp. 1-6, 2013, doi:10.1145/2556610.2556621

[16] S. H. Haji1, S. R. M. Zeebaree, R. H. Saeed, S. Y. Ameen et. all: "Comparison of Software Defined Networking with Traditional Networking," Asian Journal of Research in Computer Science, 9(2), 1-18. https://doi.org/10.9734/ajrcos/2021/v9i230216

[17] Wireshark User's Guide, Version 3.5.0, https://www.wireshark.org/docs/wsug_html_chunked

[18] R. Barrett, A. Facey: "Dynamic traffic diversion in SDN: test bed vs Mininet," in International Conference on Computing, Networking and Communications (ICNC): Network Algorithms and Performance Evaluation (2017). https://doi.org/10.1109/iccnc.2017.7876121

[19] S. H. Yeganeh, A. Tootoonchian and Y. Ganjali, "On scalability of software-defined networking," in IEEE Communications Magazine, vol. 51, no. 2, pp. 136-141, 2013, doi: 10.1109/MCOM.2013.6461198.

FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 21, Nº 1, 2022, pp. 37 - 46
https://doi.org/10.22190/FUACR220321004A

Regular Paper

ONE-BIT QUANTIZER PARAMETRIZATION FOR ARBITRARY LAPLACIAN SOURCES

UDC (621.391:517.988)

Danijela Aleksić, Zoran Perić

University of Niš, Faculty of Electronic Engineering,
Department of Telecommunications, Republic of Serbia

Abstract. *In this paper we suggest an exact formula for the total distortion of one-bit quantizer and for the arbitrary Laplacian probability density function (pdf). Suggested formula additionally extends normalized case of zero mean and unit variance, which is the most applied quantization case not only in traditional quantization rather in contemporary solutions that involve quantization. Additionally symmetrical quantizer's representation levels are calculated from minimal distortion criteria. Note that one-bit quantization is the most sensitive quantization from the standpoint of accuracy degradation and quantization error, thus increasing importance of the suggested parameterization of one-bit quantizer.*

Key words: *Laplacian source, one-bit quantization, symmetric quantizer design*

1. INTRODUCTION

Over the last few decades, numerous quantization methods have been suggested, which try to find a manner for minimizing the number of bits for real-valued presentations, striving for as higher as possible presentation accuracy [1]-[5]. The area of quantization application has constantly spread starting from traditional areas, i.e. information theory and digital signal processing, to come to the forefront in neural network (NN) field, primarily in resource-constrained environments [6]-[8]. While quantization in digital signal processing, as the signal compression method, fundamentally tries to minimize the difference between the quantized and the original signal, this minimization of inevitable quantization error is not recognized as the main target for quantized NN (QNN) models [6]-[8]. Instead, the main goal in indisputably attractive QNN area is to find the appropriate reduced-precision presentations, still generalizing well and attaining the high or satisfactory accuracy of the

Received March 21, 2022 / Accepted April 23, 2022
Corresponding author: Danijela Aleksić
University of Niš, Faculty of Electronic Engineering, Department of Telecommunications,
Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
E-mail: danijelaal@telekom.rs

applied NN model. Application of quantization, individually or jointly with some other method that significantly shrink an NN model size, one can see as an important driving force for the NN deployment on edge devices, especially on IoT devices [9]. Generally, to benefit from quantization deployment in hardware-dependent NN solutions, it is important to select a preferable quantizer model and to leverage the knowledge about quantizers and their parameterization [1].

Moreover, here it is very important to emphasize that all data or datasets in many real-world scenarios are not accessible due to privacy, proprietary or security reasons. Zero shot quantization (ZSQ) anticipates the quantization scheme that does not require an access to the original data [6]. Recently, much attention has been given to automatic speech recognition (ASR) models [10]. In paper [10], to calibrate and finetune the quantization model, synthetic data are emerged or generated. These synthetic data directly accommodate to the internal data statistics by the minimization of the Kullback-Leibler (KL) divergence. Here synthetic data, that capture data pattern are leveraged to overcome the drawback of original data. If the access to the original data is infeasible, an obvious question arises - how to quantize this data if ultra low-bit presentations are required in hardware-constrained environments?

In the scope of this paper for ultra low-bit quantizer solution of particular interest are quantizer design and its optimization to the input data. We want to tackle the quantizer efficiency challenge, primarily important in extremely low-bit quantization, due to only two representations available. To address one-bit quantization, that poses a real challenge, we do not use the ACIQ method for the analytical clipping range determination [11], since clipping effect nullify the overload distortion. Instead, we indeed calculate the total distortion for the unrestricted Laplacian probability density function (pdf) of arbitrary mean and standard deviation. The reason why we deal with the Laplacian distribution with heavy tails is because it well describes many real-world scenarios or phenomena [12]. In particular, our contributions are as follows:

- We propose an exact formula for the total distortion when the one-bit quantizer is used, while the input data are well described with an unrestricted arbitrary Laplacian pdf, especially having in mind that most of the data or datasets do not necessarily tend to the zero mean and unit variance.
- Our framework supports the additional optimization of the symmetric one-bit quantizer for the arbitrary Laplacian distribution.

The rest of paper is organized as follows: Section 2 describes the main motivation for accepting an arbitrary Laplacian pdf in our analysis, while Section 3 offers one-bit quantizer parameterization for the assumed Laplacian pdf. Section 4 provides the discussion on the performance achieved with the proposed quantizer. Section 5 summarizes and concludes on our research results.

2. WHY ARBITRARY LAPLACIAN PDF?

The Laplasian pdf, given by (1) is unimodal, log-concave pdf with a more pronounced peak and heavy tails [1]:

$$p(x|\mu,\sigma)=\frac{1}{\sqrt{2}\sigma}\exp\left\{-\frac{\sqrt{2}|x-\mu|}{\sigma}\right\}, \tag{1}$$

where μ is the mean and σ is the variance of input x.

Namely, the Laplacian distribution plays a prominent role in the probability theory, statistics and data modelling, since there is a widespread opinion that the Laplacian distribution fits many natural, economical and social phenomena [13]. Since in many real life situations, there is no prior information about data distribution, application of some well-known pdf becomes unavoidable. This explains our motivation to consider the arbitrary Laplacian distribution in the analysis presented in this paper.

Broadly speaking, many datasets exhibit some skewness and do not conform to the symmetry rules. Since we want to qualify our one-bit quantizer for special purposes where the low-complexity of hardware is one of the prerequisites, we will suggest the symmetric quantization model.

Recall that μ is a measure of the central tendency, determining where the values of x tend to be clustered, while σ points how x samples are spread out from μ to form the measure of dispersion [1]. As the pronounced peak is a specific feature of the Laplacian pdf, the largest number of samples x is concentrated around the mean value μ. Referring to our paper [14] we came to interesting conclusions for medium and high bit-rates when forming two granular regions for the restricted Laplacian pdf - Central Granular Region (CGR) and Peripheral Granular Region (PGR). We have shown that in general stands - the higher the bit rate, the higher the percentage of samples falls in the narrower CGR area. More precisely, for the amplitude dynamic defined by [1] and bit-rate 5bit/sample, 85.97% of the samples were concentrated in the CGR covering 31% of the granular region, while for the 8bit/sample, 92.37% of the samples falls in the CGR covering 25.89% of the granular region. As we want here to address the one-bit quantizer with only one pair of the quantization cells and symmetrically placed represents, we believe that we can expect that represents are close enough to the mean.

To the best of the authors' knowledge, an analysis of the influence of the low-bit quantizer's parameterization for the Laplacian pdf on Signal to Quantization Noise Ratio (SQNR) or NN model's accuracy has been addressed in numerous papers [15]-[18]. By applying two-bit and three-bit uniform quantization on the same NN model during the post-training phase, as in [15] and [16], the pure influence of the applied quantization to the NN model's accuracy is isolated. For the known Laplacian-like distribution of weights and MNIST dataset, in [15] and [16] we have proved that the quantizer design and relevant quantizer's representation levels had a stronger impact on the NN model's accuracy for the two-bit quantization case. Our anticipation is that the quantizer's representation levels determination is even more prominent in the one-bit quantization case.

Relying on a plethora of previous conclusions about uniform or nonuniform quantization [14], [19]-[21], further enhancements of one-bit quantizer parameterization are intuitively motivated by the better perceiving of the mean and variance for the arbitrary Laplacian pdf, particularly when the pdf of amplitudes being quantized was known in advance. Moreover, as a unique contribution of this paper we emphasize an analysis that outputs exact formulas for the simpler design and performance assessment of the one-bit quantizer.

3. One-bit Quantizer Parameterization

At the very beginning of this section, we recall briefly the basics of the quantization theory. An one-level quantizer Q_2 is defined by mapping $Q_2 : \mathbb{R} \rightarrow Y$ [1], where $\mathbb{R}$ is a set of real numbers, $Y = \{y_1, y_2\} \subset \mathbb{R}$ is the code book of size 2 containing representation

levels y_i, (i = 1, 2). With the one-bit quantizer Q_2, $\mathbb{R}$ is partitioned into 2 one-side unbounded in width quantization cells $\Re_i$ (i = 1, 2), where y_i specifies the i-th codeword and is the only representative for all real values x from $\Re_i$. Note that these representation levels are symmetrically placed around 0, since we address symmetric one-bit quantizer.

Let us calculate for arbitrary $\mu \geq 0$ and σ the total distortion, composed of D_- and D_+, where D_- and D_+ are distortions in negative and positive axis parts of the assumed pdf, or in quantization cells $\Re_1$ and $\Re_2$, respectively. Foremost, we will give an exact formula for arbitrary chosen represents y_i, (i = 1, 2), later adapted to our case with the assumed symmetry of representation levels.

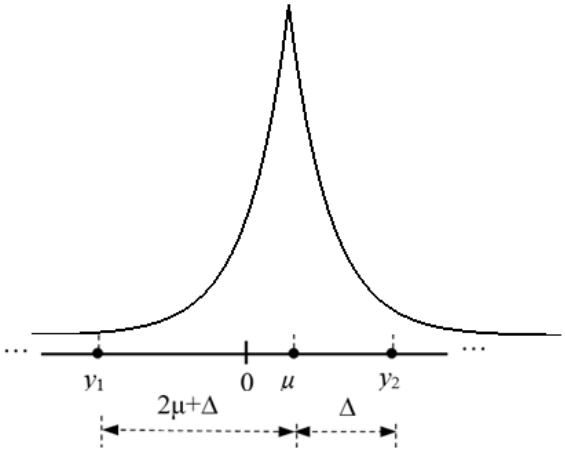


Fig. 1 Symmetrically placed representation levels for an arbitrary Laplacian pdf

$$D_- = \int_{-\infty}^{0} (x - y_1)^2 \, p(x|\mu,\sigma) \, dx, \; x < 0 \,, \tag{2}$$

$$D_+ = \int_{0}^{\mu} (x - y_2)^2 \, p(x|\mu,\sigma) \, dx + \int_{\mu}^{\infty} (x - y_2)^2 \, p(x|\mu,\sigma) \, dx, \; x \geq 0 \,, \tag{3}$$

Substituting (1) in (2) and (3) yields:

$$D_- = \frac{1}{\sqrt{2}\sigma} \int_{-\infty}^{0} (x - y_1)^2 \exp\left\{-\frac{\sqrt{2}(x-\mu)}{\sigma}\right\} dx, \; x < 0 \,, \tag{4}$$

$$D_+ = \frac{1}{\sqrt{2}\sigma}\left[\int_{0}^{\mu} (x - y_2)^2 \exp\left\{\frac{\sqrt{2}(x-\mu)}{\sigma}\right\} dx + \int_{\mu}^{\infty} (x - y_2)^2 \exp\left\{\frac{-\sqrt{2}(x-\mu)}{\sigma}\right\} dx\right], \tag{5}$$

By further reorganization of formulas (4) and (5) we obtain:

$$D_- = \frac{1}{2} \exp\left\{-\frac{\sqrt{2}\mu}{\sigma}\right\} (y_1^2 + \sqrt{2}\sigma y_1 + \sigma^2) \,. \tag{6}$$

$$D_{+}=\frac{1}{2}\left[(y_2{}^2+\sigma^2)\left(2-\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}\right)-y_2\left(4\mu+\sqrt{2}\sigma\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}\right)+2\mu^2\right]. \quad (7)$$

Finally, we find out the total distortion as:

$$D=D_{-}+D_{+}=\mu^2+\sigma^2-2\mu y_2+y_2{}^2-\frac{1}{2}\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}[(y_2{}^2-y_1{}^2)+\sqrt{2}\sigma(y_2-y_1)]. \quad (8)$$

Further we define SQNR for the one-bit quantization case and the arbitrary unrestricted Laplacian pdf as:

$$\mathrm{SQNR}=10\log_{10}\frac{\sigma^2}{D}. \quad (9)$$

$$\mathrm{SQNR}=10\log_{10}\frac{\sigma^2}{\mu^2+\sigma^2-2\mu y_2+y_2{}^2-\frac{1}{2}\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}[(y_2{}^2-y_1{}^2)+\sqrt{2}\sigma(y_2-y_1)]}. \quad (10)$$

In the special case of zero mean and unit variance (μ = 0 and σ^2=1), or in normalized case, denoting distortion by D^{n} we calculate:

$$D^{\mathrm{n}}=D_{-}+D_{+}\big|_{\mu=0,\sigma^2=1}=1+y_2{}^2-\frac{1}{2}[(y_2{}^2-y_1{}^2)+\sqrt{2}(y_2-y_1)]. \quad (11)$$

In our case with the symmetry of representation levels stands that y_2 = - y_1, distortion D^{s} for assumed symmetry condition becomes:

$$D^{\mathrm{s}}=D_{-}+D_{+}\big|_{y_2=-y_1}=\mu^2+\sigma^2-2\mu y_2+y_2{}^2-\sqrt{2}\sigma y_2\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}. \quad (12)$$

Finally, we calculate the total distortion for the normalized case with symmetrically placed representation levels as:

$$D^{\mathrm{ns}}=D_{-}+D_{+}\big|_{\mu=0,\sigma^2=1,y_2=-y_1}=1-\sqrt{2}y_2+y_2{}^2. \quad (13)$$

By minimizing the distortion, that is, by setting the first derivative of so obtained distortion D^{ns} with respect to y_2 equal to zero:

$$\frac{\partial D^{\mathrm{ns}}}{\partial y_2}=2y_2-\sqrt{2}=0. \quad (14)$$

we can find $y_2{}^{ns}=\frac{\sqrt{2}}{2}$, while for D^{s} we derive the minimal distortion condition as:

$$\frac{\partial D^{\mathrm{s}}}{\partial y_2}=-2\mu+2y_2-\sqrt{2}\sigma\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}=0. \quad (15)$$

Here the positive representation level y_2 is determined as:

$$y_2 = \mu + \frac{\sqrt{2}}{2}\sigma \exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\} = \mu + \Delta\,. \tag{16}$$

whereas Δ shows how far the representation level y_2 is from mean μ, while y_1 is distant from mean for $2\mu+\Delta$ (see Fig.1).

$$\Delta = \frac{\sqrt{2}}{2}\sigma \exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}. \tag{17}$$

By further reorganization of (16) we have:

$$y_2 = \mu\left(1+\frac{\sqrt{2}\sigma}{2\mu}\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}\right) = \mu\left(1+\frac{\Delta}{\mu}\right). \tag{18}$$

Substituting (16) in (10) yields:

$$\mathrm{SQNR}^{s} = 10\log_{10}\frac{1}{1-\frac{1}{2}\exp\left\{\frac{-2\sqrt{2}\mu}{\sigma}\right\}-\frac{\sqrt{2}\mu}{\sigma}\exp\left\{\frac{-\sqrt{2}\mu}{\sigma}\right\}}. \tag{19}$$

4. NUMERICAL RESULTS

In this section, we discuss achieved results as the outcomes of the parameterization for symmetrical one-bit quantizer. Let us remind that in previous section we have given an exact formula for SQNR calculation (19) when the symmetrical one-bit quantizer is used, while the input data are well described with an unrestricted arbitrary Laplacian pdf. Therefore, we find an interest to analyze how quantization of data having arbitrary Laplacian pdf affects data representation levels for the assumed symmetry of quantizer design.

Paying attention to the Eq. (19), one can note that SQNR^{s} for symmetrically placed representation levels depends on μ/σ (see Fig.2)

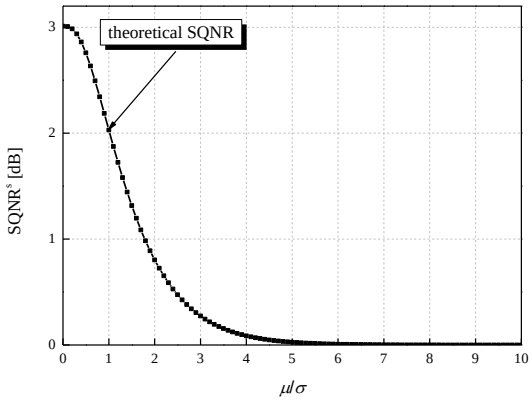


Fig. 2 SQNR^{s} dependence on μ/σ for symmetrical quantizer design

Table 1 SQNR^s for some specific μ and σ.

μ	σ	SQNR [db]
0	arbitrary	3.0103
1	1	2.0299
1	2	2.7590
1	5	2.9856
2	1	0.8036
2	2	2.0299
2	5	2.8616
5	1	0.0262
5	2	0.4743
5	5	2.0299

Table 1 shows SQNR^s for some specific μ and σ values. The highest SQNR^s is achieved for zero mean and arbitrary variance, while SQNR^s rapidly decreases with an increase of mean and for unit variance. If μ and σ are the same, SQNR^s amounts 2.0299.

For zero mean pdfs, representation level $y_2 = \sigma / \sqrt{2}$ depends only on σ. In case of non-zero mean pdfs, $y_2 = \mu$ (1+ Δ/μ) and according to Δ/μ dependence on μ/σ (see Fig. 3) we can conclude that μ predominantly influences y_2, as well as its symmetrical pair y_1.

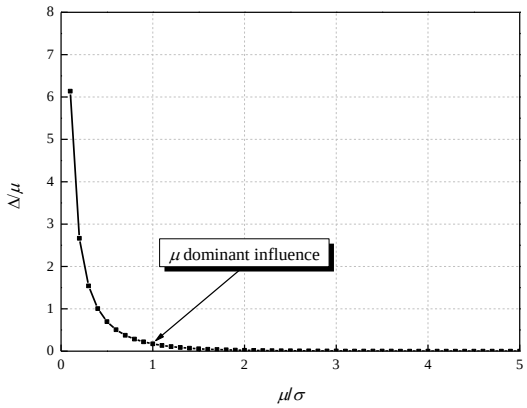


Fig. 3 Δ/μ dependence on μ/σ

For pre-trained QNN, customized for some specific task, further reused as the foundation for an another task, in case of aggressive one-bit quantization, QNN performance can be tremendously degraded. To alleviate this inevitable degradation, more precise mean and standard deviation assessment one can set as an indispensable prerequisite. In most of the performance analysis, zero mean and unit variance are accepted as the most common conditions of the ground-up quantizer design. Here we want to analyse an influence of two main parameters, μ and σ, that well describe Laplacian pdf, since their determination holds the key to improving the quantization's efficiency.

In doing so, we will analyse the Kullback-Leibler (KL) divergence for two arbitrary Laplacian pdfs [22], [23] (see Table 2 Case 1), while in general stands that:

$$\mathrm{KL}(\,p(x|\mu_1,\sigma_1)\|\,p(x|\mu_2,\sigma_2)) \neq \mathrm{KL}(\,p(x|\mu_2,\sigma_2)\|\,p(x|\mu_1,\sigma_1))\,. \tag{20}$$

Case 1 shows the measure or divergence of $p(x|\mu_1,\sigma_1)$ from $p(x|\mu_2,\sigma_2)$, while Eq. (20) underlines the inequality of reciprocal divergences of the same pair of functions $\{p(x|\mu_1,\sigma_1), p(x|\mu_2,\sigma_2)\}$. For two identical pdfs the KL divergence is zero, while for a large deviation of μ_1 from μ_2 and σ_1 from σ_2, KL divergence is large.

Kullback-Leibler divergence, as a similarity measure, is also used to specify extra bits required to describe a pair of distributions. If pdfs differ significantly, there is a need to provide additional bits that is not desirable in low-bit presentations.

If one of the pdfs implies a normalized case of zero mean and unit variance, we can find a KL divergence in cases where this normalization characterizes the first or the second pdf (Case 2 and Case 3, respectively). Note that the KL divergence given in Case 5, for a specific case of $\sigma_1 = \sqrt{2}$ does not depend on μ_1. In contrast, applying similar condition $\sigma_2 = \sqrt{2}$ in Case 4, one can notice that the KL divergence indeed depends on μ_2, that additionally confirms an inequality statement given by (20).

Table 2 KL divergence for some specific cases.

Case	μ_1	σ_1	μ_2	σ_2	KL divergence
1	μ_1	σ_1	μ_2	σ_2	$\dfrac{\sigma_1 \exp\left\{\dfrac{-\sqrt{2}\lvert\mu_1-\mu_2\rvert}{\sigma_1}+\lvert\mu_1-\mu_2\rvert\right\}}{\sigma_2}+\log\dfrac{\sigma_2}{\sigma_1}-1$
2	0	1	μ_2	σ_2	$\dfrac{\exp\{(1-\sqrt{2})\lvert\mu_2\rvert\}}{\sigma_2}+\log\sigma_2-1$
3	μ_1	σ_1	0	1	$\sigma_1 \exp\left\{\dfrac{-\sqrt{2}\lvert\mu_1\rvert}{\sigma_1}+\lvert\mu_1\rvert\right\}-\log\sigma_1-1$
4	0	1	μ_2	$\sqrt{2}$	$\dfrac{1}{\sqrt{2}}\exp\{(1-\sqrt{2})\lvert\mu_2\rvert\}+\log\sqrt{2}-1$
5	μ_1	$\sqrt{2}$	0	1	$\sqrt{2}-\log(\sqrt{2})-1$
6	μ_1	σ	μ_2	σ	$\exp\left\{\dfrac{(\sigma-\sqrt{2})\lvert\mu_1-\mu_2\rvert}{\sigma}\right\}-1$
7	μ	σ_1	μ	σ_2	$\dfrac{\sigma_1}{\sigma_2}+\log\dfrac{\sigma_2}{\sigma_1}-1$

If both pdfs have the same $\sigma_1 = \sigma_2 = \sigma$ (Case 6) the KL divergence depends on σ and difference of means, while for $\mu_1 = \mu_2 = \mu$, the KL divergence depends only on σ_1 / σ_2 (Case 7).

Since pre-trained model can be accommodated to the normalized case, by calculating the KL divergence similar as in paper [10], one can estimate μ, σ or μ/σ and specify symmetrical representation levels of one-bit quantization model according to the suggested formula (16).

As the total bit rate R in signal processing area according to formula $R = r_{fixed} + r_{add}/M$ [24], shows a strong dependence on r_{add} and M, where r_{add} – is the bit rate for additional information (here for μ and σ) and M – denotes the frame size, we should also pay attention to the analysis of R for the suggested one-bit design. ($r_{fixed} = 1$, see Table 3). Our NN model architecture specified in [15,16] consists of three FC (dense) layers. MNIST training and testing datasets are loaded and then flattened into one-dimensional vectors of 784 (28*28) elements. First two FC layers consist of 512 nodes, where the first layer accepts an input shape of (784,), so the overall number of bits for that level can be calculated as M = 784x512, and the total bit rate is R = 1+32/784x512 = 1,00008 ≈ 1. We can notice that neural networks are less demanding in terms of additional bits for side information transmitting, as we concluded - the larger the NN model, the closer the bit rate is to r_{fixed} or to $R = 1$, in one-bit quantizer case.

For the most common used presentations with $M = 32$ and $M^* = 16$ bits, required total bit-rates are shown in Table 3. As for low-bit presentations, it is desirable that additional information is displayed with as few bits as possible, so that $M \geq 120$ frame sizes should be preferred.

Table 3 Total bit rate R for one-bit quantizer in signal processing area.

r_{fixed}	r_{add}	M	R	r_{add}^*	M^*	R^*
1	32	20	2.6	16	20	1.8
1	32	40	1.8	16	40	1.4
1	32	120	1.267	16	120	1.133
1	32	240	1.133	16	240	1.067

5. CONCLUSION

In this paper, an importance of determining of the exact formula for one-bit quantizer's total distortion is highlighted in the case when the data are well described by an arbitrary Laplacian pdf. Aiming for the simplicity of quantization design, symmetrical quantizer is proposed, as well as the symmetry of representational levels. Note that precise determination of the mean and variance is a prerequisite to achieve the high quantization model's accuracy, since quantization accuracy can be tremendously degraded in the case when the mean and variance are not well adjusted. We have shown that the mean has a predominantly important influence on the representation level determination in the non-normalized case, since one of the representation levels is close to the mean. One-bit quantizers are especially convenient in highly proliferated resource-constrained devices thus increasing the need for the simple but accurate low-bit quantizer design.

Acknowledgement: *This research has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia and by the Science Fund of the Republic of Serbia, 6527104, AI-Com-in-AI.*

REFERENCES

[1] N. S. Jayant, P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[2] D. Hui, D. L. Neuhoff, "Asymptotic analysis of optimal fixed-rate uniform scalar quantization," *IEEE Transaction on Information Theory*, vol. 47, no. 3, pp. 957–977, 2001, doi: 10.1109/18.915652.

[3] S. Na, D. Neuhoff, "On the convexity of the MSE distortion of symmetric uniform scalar quantization," *IEEE Transaction on Information Theory*, vol. 64, no.4, pp. 2626−2638, 2018, doi: 10.1109/TIT.2017.2775615.

[4] S. Na, D. Neuhoff, "Monotonicity of step sizes of MSE-optimal symmetric uniform scalar quantizers," *IEEE Transaction on Information Theory*, vol. 65, no. 3, pp. 1782−1792, 2019, doi: 10.1109/TIT.2018.2867182.

[5] J. Lee, S. Na, "A rigorous revisit to the partial distortion theorem in the case of a Laplacian source," *IEEE Commun. Lett.*, vol. 21, pp. 2554–2557, 2017, doi: 10.1109/LCOMM.2017.2749218.

[6] A. Gholami, S. Kim, Z. Dong, Z. Yao, M. W. Mahoney, K. Keutzer, "A survey of quantization methods for efficient neural network inference," arXiv:2103.13630

[7] Y. Guo, "A survey on methods and theories of quantized neural networks," arXiv:1808.04752

[8] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, Y. Bengio, "Quantized neural networks: training neural networks with low precision weights and activations," *J. Mach. Learn. Res.*, vol. 18, pp. 6869–6898, 2017.

[9] D. Liu, H. Kong, X. Luo, W. Liu, R. Subramaniam, "Bringing AI to edge: from deep learning's perspective," arXiv:2011.14808.

[10] S. Kim, A. Gholami, Z. Yao, N. Lee, P. Wang, A. Nrusimha, B. Zhai, T. Gao, M. W. Mahoney, K. Keutzer, "Integer-only zero-shot quantization for efficient speech recognition," arXiv:2103.16827v2

[11] R. Banner, Y. Nahshan and D. Soudry, "Postraining 4-bit quantization of convolutional networks for rapid-deployment," *33rd Conference on Neural Information Processing Systems* (*NeurIPS 2019*), *Vancouver, Canada*, pp. 7948-7956, 2019.

[12] S. Gazor and W. Zhang, "Speech probability distribution," IEEE Signal Processing Letters, vol. 10, no. 7, pp. 204-207, 2003, doi: 10.1109/LSP.2003.813679.

[13] S. Kotz, T. Kozubowski and K. Podgorski, "The Laplace distribution and generalizations," *Birkhäuser*, Boston, 2001.

[14] J. Nikolić, D. Aleksić, Z. Perić and M. Dinčić, "Iterative algorithm for parameterization of two-region piecewise uniform quantizer for the Laplacian source," *Mathematics*, vol. 9, pp. 3091, 2021, doi: 10.3390/math9233091.

[15] Stefan Tomić, Jelena Nikolić, Zoran Perić, Danijela Aleksić, "Performance of Post-Training Two-Bits Uniform and Layer-Wise Uniform Quantization for MNIST Dataset from the Perspective of Support Region Choice,",*Mathematical Problems in Engineering*, Volume 2022, Article ID 1463094, 2022.J.

[16] Nikolić, Z. Perić, D. Aleksić, S. Tomić, A. Jovanović, "Whether the support region of three-bit uniform quantizer has a strong impact on post-training quantization for MNIST dataset?," *Entropy*, vol. 23, pp. 1699, 2021, doi: 10.3390/e23121699.

[17] W. Zhao, M. Teli, X. Gong, B. Zhang, D. Doermann, "A review of recent advances of binary neural networks for edge computing," *IEEE J. Miniat. Air Space Syst.*, vol. 2, pp. 25–35. 2021, doi: 10.1109/JMASS.2020.3034205.

[18] P. E. Novac, G. B. Hacene, A. Pegatoquet, B. Miramond, V. Gripon, "Quantization and deployment of deep neural networks on microcontrollers," *Sensors*, vol. 21, pp. 2984, 2021, doi: 10.3390/s21092984.

[19] D. Aleksić, Z. Perić, "Analysis and design of robust quasilogarithmic quantizer for the purpose of traffic optimisation," *Inf. Technol. Control*, vol. 47, pp. 615-622, 2018, doi: http://dx.doi.org/10.5755/j01.itc.47.4.20668.

[20] Z. Perić, J. Nikolić, D. Aleksić, A. Perić, "Symmetric quantile quantizer parameterization for the Laplacian source: qualification for contemporary quantization solutions, " *Math. Probl. Eng.*, vol. 2021, Article ID 6647135, 2021, doi: 10.1155/2021/6647135

[21] Z. Perić, D. Aleksić, "Quasilogarithmic quantizer for Laplacian source: support region ubiquitous optimization task," *Rev. Roum. Sci. Techn.*, vol. 64, pp. 403−408, 2019.

[22] G. Dehaune, "A deterministic and computable Bernstein-von Mises theorem," arXiv:1904.02505.

[23] S. Kullback, R. Leibler, "On information and sufficiency. Annals of mathematical statistics," vol. 22, pp. 79–86, 1951.

[24] S. Tomić, Z. Perić, J. Nikolić, "Modified BTC algorithm for audio signal coding," *Advances in Electrical and Computer Engineering*, vol. 16, 2016.

FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 21, Nº 1, 2022, pp. 47 - 58
https://doi.org/10.22190/FUACR220401005N

Regular Paper

NOVEL EXPONENTIAL TYPE APPROXIMATIONS OF THE *Q*-FUNCTION

UDC ((621.391:004)+519.21)

Jelena Nikolić, Zoran Perić

University of Niš, Faculty of Electronic Engineering,
Department of Telecommunications, Republic of Serbia

Abstract. *In this paper, we propose several solutions for approximating the Q-function using one exponential function or the sum of two exponential functions. As the novel Q-function approximations have simple analytical forms and are therefore very suitable for further derivation of expressions in closed forms, a large number of applications are feasible. The application of the novel exponential type approximations of the Q-function is especially important for overcoming issues arising in designing scalar companding quantizers for the Gaussian source, which are caused by the non-existence of a closed form expression for the Q-function. Since our approximations of the Q-function have simple analytical forms and are more accurate than the approximations of the Q-function previously used for the observed problem in the scalar companding quantization of the Gaussian source, their application, especially for this problem is of great importance.*

Key words: *Gaussian source, Q-function, exponential type approximations*

1. INTRODUCTION

Estimating the performance of digital communication systems, typically, signal to quantization noise ratio (SQNR) and the symbol error probability (SEP), plays a very important role in designing these systems. Performance evaluation of the average SEP of digital modulations in additive white Gaussian noise as well as fading channels and performance (distortion and SQNR) evaluation of scalar quantizers for a Gaussian source includes calculation of the improper integral whose solution cannot be obtained in a closed form [1]-[3]. Namely, in these calculations, the Gaussian *Q*-function and the erfc function, given by [4]:

Received April 01, 2022 / Accepted May 03, 2022
Corresponding author: Jelena Nikolić
University of Niš, Faculty of Electronic Engineering, Department of Telecommunications,
Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
E-mail: jelena.nikolic@elfak.ni.ac.rs

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\left\{-\frac{t^2}{2}\right\} dt = 1/2 \operatorname{erfc}(x/\sqrt{2}), \tag{1}$$

cannot be expressed in the form of elementary integrals. Motivated by this fact, numerous researches have been conducted with the aim of determining approximations of the *Q*-function having simple analytical forms and preferably high accuracies.

The infinite upper limit of the integral given in Eq. (1) poses the restriction on the application of numerical integration methods. The presence of the *Q*-function argument as the lower limit of the integral causes additional difficulties in analytical derivations in which this argument depends on some other random parameters so that statistical averaging is necessary [2]-[13]. For this reason, *Q*-function approximations that have simple analytical forms suitable for further derivations [5], [6], [13]-[17] are especially important. It is also preferable to determine the *Q*-function approximations of high approximation accuracies. However, from the requirement for simple analytical forms of the *Q*-function approximations, approximations of relatively low accuracies that is, of relatively high relative errors are usually outputted [13]. By taking into account these conflicting requirements, numerous approximations of the *Q*-function have been specified in literature, which, depending on the research goals and application, strive for a compromise between the simplicity of analytical form and accuracy (see for example, [2], [3], [5]-[26]).

In general, absolute and relative error (*Relative Eror*-RE) of approximating the *Q*-function with $F(x)$, is calculated from:

$$\Delta_Q(x) \triangleq |F(x) - Q(x)|, \tag{2}$$

$$\delta_Q(x) \triangleq \frac{|F(x) - Q(x)|}{Q(x)}. \tag{3}$$

It is important to point out that some of the approximations of the *Q*-function provide small RE for small values of the argument, but high RE for large argument values, while for some other approximations of the *Q* function, the opposite conclusion holds. Due to the mentioned shortcomings of the *Q*-function approximations, this problem is still very current.

This interesting research topic is especially important in the scalar quantization of a memoryless Gaussian source, since, as shown in [18], [21], [27]-[37], to design and evaluate the performance of these quantizers, closed form expressions cannot be derived for the presumed Gaussian source. Therefore, analytically simple approximate expressions for the *Q*-function that do not degrade greatly the accuracy of the *Q*-function are very necessary. The authors of papers in the field of design and performance evaluation of scalar quantizers for a memoryless Gaussian source commonly use approximations of the *Q*-function from [15]. To apply the approximations of the *Q*-function having the simplest analytical forms, they utilize the approximations of the *Q*-function from [15] determined by a relatively small number of terms of the asymptotic expansion of the erfc function. Although relatively simple closed-form expressions for distortion and SQNR of scalar quantizers designed for a Gaussian memoryless source have been derived by using approximations from [15], the impact of the *Q*-function approximation accuracy on the performance of the considered quantizer has not been analyzed. This inspired the research presented in [18], in which, in addition to proposing a novel analytical form of the *Q*-function approximation, a very important analysis of the influence of the *Q*-function approximation accuracy on the assessment of Gaussian source scalar quantization performance is presented.

In [21], a particularly significant problem that arises in the design of scalar companding quantizers for the Gaussian source is pointed out, which has motivated numerous studies in this field (for instance, [28], [29], [32], [34], [36]). Namely, it was emphasized that in order to perform the companding quantization for the assumed Gaussian source, it is first necessary to perform a numerical integration, and then solve integral equations to design the considered model of the companding quantizer. To simplify the process of designing this model of quantizer, various methods of linearization of the compressor function have been proposed in the literature [28], [29], [32], [34], which enabled a much simpler design and performance evaluation by means of linearization. The authors of this paper, to the best of their knowledge, are the only ones to approach the solution of the observed problem by applying the *Q*-function approximations in determining approximate compressor functions, in the manner described in [21]. Namely, in [21], we came up with the idea to assume a novel class of the *Q*-function approximations that enables the derivation of expressions in closed form for representation levels and decision thresholds and to analyze the effects achieved by applying the *Q*-function approximation from the proposed class to approximate the compressor function. Since only the first step in this direction has been made in the research presented in [21], and approximations of the *Q*-function of simple analytical forms, but of relatively low accuracies, have been proposed, further progress in this research can justifiably be expected. For this reason, the aim of this paper is to determine some novel approximations of the *Q*-function of simple analytical forms, such as ones from [21], which will be more accurate compared to those from [21]. Finally, it should be noted that in [21] (see Eq. (12) from [21]), the description of the approximate compressor function has been given for an arbitrarily selected approximation of the *Q*-function, $F(x)$. Therefore, this paper references to [21], avoiding the repetition of the description of the compressor function and the way of its approximation by means of the *Q*-function approximations. In other words, the focus of this paper is on determining some novel improved solutions of the *Q*-function approximations in terms of accuracy, which have comparable or equal complexities of analytical forms as the approximations from [21].

The rest of the paper is structured as follows: First, a short description of the related work in the field of exponential type *Q*-function approximations is provided in Section 2. Afterwards, in Section 3, novel exponential type approximations of the *Q*-function are specified, and their performance are presented and discussed in Section 4. Finally, in Section 5, the conclusions of our results are derived.

2. RELATED WORK

The group of exponential type approximations of the *Q*-function of relatively simple analytical forms encompasses approximations of the *Q*-function that can be represented by one exponential function or by the sum of two or more exponential functions [5], [6], [14], [16]. A special importance in this section should be given to the approximations from [5], which due to their simplicity are highly cited in literature. Namely, Chiani and Simon used the following form of the erfc function in [5] and [22]:

$$\operatorname{erfc}(x) = \frac{2}{\pi}\int_0^{\pi/2} \exp\left\{-\frac{x^2}{\sin^2\theta}\right\} d\theta,\ x \geq 0. \tag{4}$$

In particular, Craig came to this form of the erfc function in [9]. It is important to notice that unlike the standard definition of the erfc function, in which the argument of the function occurs as one of the integral limits, in the last expression, the argument of the erfc function occurs in the integrand. Simon utilized the fact that the integrand in Eq. (4) is monotonically increasing function of θ for $0 \le \theta \le \pi/2$ and he simply concluded that the upper bound of the erfc function can be determined from Eq. (4) by replacing the integrand with its maximum occurring at $\theta = \pi/2$:

$$\operatorname{erfc}(x) \le \frac{2}{\pi} \int_0^{\pi/2} \exp\{-x^2\}\, d\theta = \exp\{-x^2\},\ x \ge 0\,. \tag{5}$$

Chiani also considered significant this feature of the integrand from Eq. (4). Considering that it holds $0 \le \theta \le \pi/2$, he selected arbitrarily N+1 values for θ_0, θ_1,..., θ_N, arranged in the ascending order $0 = \theta_0 \le \theta_1 \le ... \le \theta_N = \pi/2$, and then showed that it holds:

$$\operatorname{erfc}(x) \le \frac{2}{\pi} \sum_{i=1}^{N} \int_{\theta_{i-1}}^{\theta_i} \exp\left\{-\frac{x^2}{\sin^2\theta_i}\right\} d\theta\,. \tag{6}$$

With further introducing

$$a_i = \frac{2(\theta_i - \theta_{i-1})}{\pi}\,, \tag{7}$$

$$b_i = \frac{1}{\sin^2\theta_i}\,, \tag{8}$$

Eq.(6) was transformed into:

$$\operatorname{erfc}(x) \le \sum_{i=1}^{N} a_i \exp\{-b_i x^2\}\,. \tag{9}$$

Although a more accurate approximation of the upper bound of the erfc function is obtained by increasing the value of N, the complexity of the analytical form of the obtained approximation certainly increases. For this reason, in [5], only the following values for N were assumed: $N = 1$, $N = 2$, $N = 3$ and $N = 6$, and the resulting approximations of the Q-function were determined from Eqs. (6)-(8) for equispaced points $\theta_i = i\pi/(2N)$, $i = 1,2,...,N$. Also, for $N = 2$, Chiani determined the following upper bound approximation of the erfc function in [5]:

$$\operatorname{erfc}(x) \simeq g(x,\theta_{\text{opt}}) = \frac{1}{6}\exp\{-x^2\} + \frac{1}{2}\exp\left\{-\frac{4}{3}x^2\right\}, \tag{10}$$

that is, the following upper bound approximation of the Q-function:

$$\overline{F}^{[5]}(x) = \frac{1}{12}\exp\left\{-\frac{x^2}{2}\right\} + \frac{1}{4}\exp\left\{-\frac{2x^2}{3}\right\},\ x > 0.5\,. \tag{11}$$

Namely, he applied the trapezoidal rule of numerical integration for an arbitrarily chosen point θ from $0 \le \theta \le \pi/2$

$$\operatorname{erfc}(x) \simeq g(x,\theta) = \left(\frac{1}{2} - \frac{\theta}{\pi}\right)\exp\{-x^2\} + \frac{1}{2}\exp\left\{-\frac{x^2}{\sin^2\theta}\right\}, \tag{12}$$

and performed a numerical optimization procedure to determine the particular value of the parameter θ

$$\theta_{\text{opt}} = \arg\min_{\theta} \frac{1}{R}\int_0^R \frac{\left|g(x,\theta) - \operatorname{erfc}(x)\right|}{\operatorname{erfc}(x)}\,dx, \tag{13}$$

so that the minimum of the integral of the absolute relative error for the interval [0, R] was achieved, where he assumed $20\log_{10}R = 13$ dB. As a result, $\theta_{\text{opt}} \approx \pi/3$ was determined. With further substitution of this value for θ in Eq. (12), Eq.(10) was than derived.

Chiani showed that the approximation of the erfc function or *Q*-function determined in this way is much more accurate than the Chernoff-type approximation [16]

$$F^{[16]}(x;a,b) = a\exp\{-b\,x^2\},\ a,b \in \mathbb{R}. \tag{14}$$

The function from Eq. (14), depending on the choice of real values of parameters a and b, as well as the interval of the argument values to which it is applied, can represent either the lower or upper bound approximation of the *Q*-function or the *Q*-function approximation itself. For comparative purposes, Chiani used the following upper bound approximation of the *Q*-function:

$$Q(x) \le \overline{F}^{[16]}(x) = \exp\left\{-\frac{x^2}{2}\right\},\ x \ge 0, \tag{15}$$

that is, the following upper bound approximation of the erfc function:

$$\operatorname{erfc}(x) \le 2\exp\{-x^2\},\ x \ge 0. \tag{16}$$

As highlighted in [5] and [8], the tightest possible Chernoff-type upper bound approximation of the *Q*-function can be determined if the right side of the inequality from Eq. (15) is multiplied by 0.5

$$Q(x) \le \overline{F}^{[6]}(x) = \frac{1}{2}\exp\left\{-\frac{x^2}{2}\right\},\ x \ge 0. \tag{17}$$

In [6], mathematical proof was given that the upper bound approximation of the *Q*-function, specified as in Eq. (17), is the most accurate bound approximation of the Chernoff-type for $x \ge 0$. Note that the same form of the upper bound approximation of the *Q*-function was reported many years later in [17], where a somewhat different approach to determining this approximation was applied.

Research from [14], published after one from [5], continued to address the problem of optimizing Chernoff-type approximations of the *Q*-function. In particular, by narrowing the interval of argument values and by performing the numerical optimization of parameters a and b of the Chernoff-type approximation, the following upper and lower bound approximations of the *Q*-function are specified in [14]:

$$Q(x) \le \overline{F}^{[14]}(x) = 0.28 \exp\left\{-1.275\frac{x^2}{2}\right\},\ x > 0.5\,, \tag{18}$$

$$Q(x) \ge \underline{F}^{[14]}(x) = 0.3 \exp\left\{-1.01\frac{x^2}{2}\right\},\ x > 0.535\,. \tag{19}$$

The authors of [14] pointed out that the bound approximations of the *Q*-function of higher accuracies can be determined if the interval of argument values in which these approximations are indeed bound approximations of the *Q*-function is narrowed. Note that the dashes placed above and below the $F(\cdot)$ function are introduced to distinguish the upper and lower bound approximations of the *Q*-function.

Inspired by the conclusion from [21] that further improvements in the accuracy of the applied approximations are very necessary, as well as by the fact that the authors of this paper, to the best of their knowledge, were the only ones who initiated solving the observed problem of designing compressor functions, in what follows we present the concept of how to achieve this improvement, with the restriction that relatively simple approximations of the *Q*-function from the class of exponential type approximations are applied.

3. NOVEL APPROXIMATIONS OF THE *Q* FUNCTION

Let us first recall that the bound approximations of the *Q*-function from [17] and [38] are given by:

$$\overline{F}^{[17]}(x) = \frac{1}{2}\exp\left\{-\frac{x^2}{2}\right\}, \tag{20}$$

$$\overline{F}^{[38]}(x) = \exp\left\{-\frac{x^2}{2}\right\}. \tag{21}$$

These bound approximations of the *Q*-function are characterized by very low accuracies, which degrade with increasing the value of the argument *x*. As highlighted in [6], it was proved that the upper bound approximation of the *Q*-function given by Eq. (20) is the most accurate Chernoff-type bound approximation of the *Q*-function for $x \ge 0$. Let us also recall that the *Q*-function approximation from [17], given by Eq. (20), was utilized in [21] as the initial solution to the compressor function approximation problem.

As described in detail in the previous section, Chiani applied the trapezoidal rule of numerical integration, optimized the choice of the parameter θ, obtained $\theta_{\text{opt}} \approx \pi/3$, and determined the approximation of the *Q*-function that represents the sum of two exponential functions [5]. If in the derivation conducted by Chiani we assume $\theta = \pi/2$, and reduce one exponential term accordingly, we can determine our first novel exponential type (Chernoff-type) approximation of the *Q*-function as follows:

$$F^{\text{novel 1}}(x) = \frac{1}{4}\exp\left\{-\frac{x^2}{2}\right\}. \tag{22}$$

In order to further determine novel approximations of the *Q*-function, it is important to highlight that in [14], it was pointed out that the bound approximations of the *Q*-

function of higher accuracies can be determined for narrower interval of argument values in which these bounds are approximations. If we apply these approximations to solve the problem specified in [21], we can expect a greater accuracy of the compressor function approximation because, as we will show in the next section, a much smaller relative error is introduced compared to that from [17]. However, as the Chernoff-type bound approximation of the *Q*-function are generally characterized by relatively low accuracies, it is interesting to consider the performance of the *Q*-function approximations that are not the *Q*-function approximation bounds and that are defined as the geometric and arithmetic mean of the *Q*-function approximation bounds from [14]. In accordance with the above mentioned, we have formulated the following two novel approximations of the *Q*-function of exponential type:

$$F^{\text{novel 2}}(x) = \sqrt{\underline{F}^{[14]}(x)\overline{F}^{[14]}(x)} = 0.2898 \exp\left\{-1.1375 \frac{x^2}{2}\right\}, \tag{23}$$

$$F^{\text{novel 3}}(x) = \frac{\underline{F}^{[14]}(x) + \overline{F}^{[14]}(x)}{2} = 0.14 \exp\left\{-1.275 \frac{x^2}{2}\right\} + 0.15 \exp\left\{-1.01 \frac{x^2}{2}\right\}. \tag{24}$$

Although for most of the considered values of argument *x*, by applying the approximations of the *Q*-function determined as the geometric and arithmetic mean of these two approximations, one can expect greater accuracies compared to the approximation of the *Q*-function given by Eq. (19), for some values of argument *x*, we will show that a further increase of the accuracy can be achieved by utilizing the benefits of applying the *Q*-function approximation given by Eq. (18). Therefore, in this paper we propose the following approximation of the *Q*-function:

$$F^{\text{novel 4}}(x; k) = \frac{k\overline{F}^{[14]}(x) + \underline{F}^{[14]}(x)}{k+1}, \; k \in \mathbb{N}, \tag{25}$$

which gives more weight in averaging to the *Q*-function approximation of the higher accuracy specified by Eq. (18) than the one specified by Eq. (19). Note that for $k = 1$, the last approximation of the *Q*-function is identically equal to the approximation specified by Eq. (24):

$$F^{\text{novel 4}}(x; k)_{|k=1} = F^{\text{novel 3}}(x). \tag{26}$$

Based on the results presented in the following section, we will conclude that introducing such a modification of the *Q*-function approximation is very meaningful, since changing the value of parameter *k* can increase the accuracy of the *Q*-function approximation given by Eq. (24). The choice of one value of parameter *k* will depend on some additional criteria that need to be introduced, which will be left for our future research.

4. NUMERICAL RESULTS

This section discusses the results (RE) determined by applying the proposed *Q*-function approximations and presents an adequate comparison with the *Q*-function approximations from literature, having the same or similar complexities of analytical forms as the proposed approximations. Let us first consider Figure 1, which compares REs determined for the case of applying the upper bound approximations of the *Q*-function from [17] and [38] and our

novel 1 *Q*-function approximation (specified by Eq. (22)). Note that we compare here the Chernoff-type approximations of the *Q*-function, hence, of the same complexities of the analytical forms. It can be noticed that the novel 1 approximation of the *Q*-function achieves significantly higher accuracy compared to the ones determined by applying the upper bound approximations of the *Q*-function from [17] and [38], for most of the values of argument *x* taken into consideration.

Figure 2 confirms the conclusion from [14]. Namely, in [14], it is anticipated that higher accuracies of the *Q*-function approximation bounds can be achieved if the interval of argument values in which these *Q*-function approximations are approximation bounds is narrowed. Figure 3 compares the dependences of the relative errors on the value of argument *x* for the case of applying the approximations of the *Q*-function from [14] and [17] with our novel 2 approximations of the *Q*-function. Based on these dependences, it can be noticed that our novel 2 approximation of the *Q*-function provides a significantly higher accuracy than the approximation of the *Q*-function from [17], while for one interval of the argument values, the

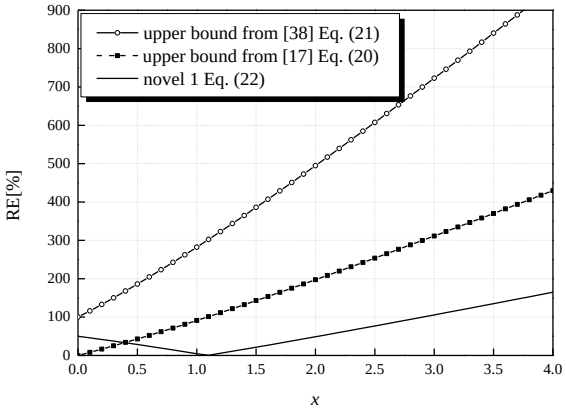


Fig. 1 Relative error: Application of the *Q*-function approximations from [17] and [38] and novel 1 *Q*-function approximation of the Chernoff-type.

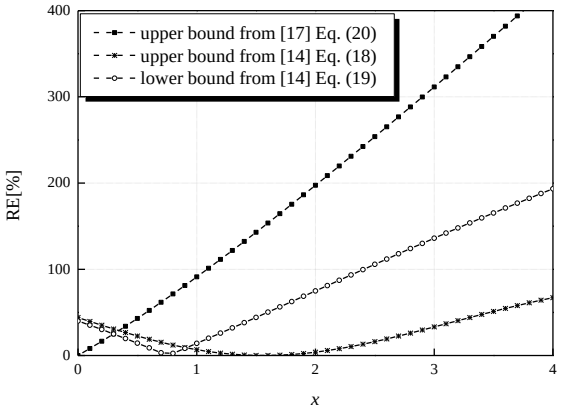


Fig. 2 Comparison of relative errors: Application of the *Q*-function approximations from [14] and [17].

advantage in terms of accuracy has the upper bound approximation from [14]. A similar conclusion can be derived based on the results shown in Figure 4, where the performance comparison of the novel 2 and novel 3 approximations of the *Q*-function and the approximation of the *Q*-function from [14] is presented.

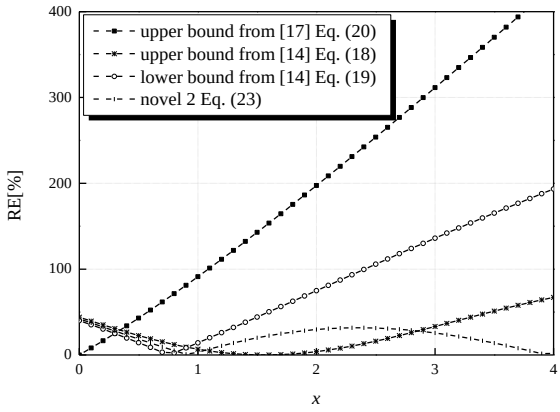


Fig. 3 Relative error: application of the approximations from [17] and [14] and novel 2 approximations determined as geometric mean of approximations from [14].

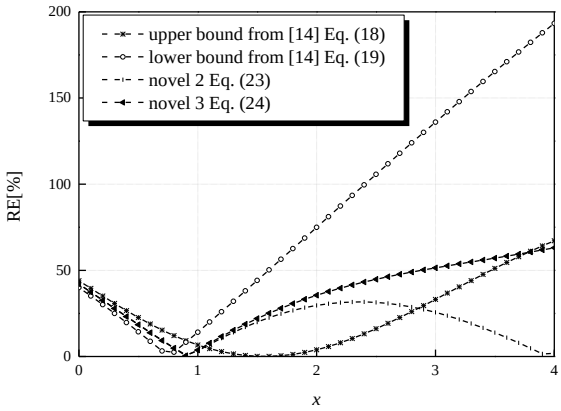


Fig. 4 Relative error: application of the approximations of the *Q*-function from [14], novel 2 and novel 3 approximations of the *Q*-function defined as the geometric mean and arithmetic mean of the approximations from [14].

Figure 5 compares the performance of novel 2, novel 3, and novel 4 approximations of the *Q*-function, with a different choice of parameter k being considered for novel 4 approximation, where it holds $k \in \mathbb{N}$. We can conclude that by increasing the value of parameter k, in the greater part of the interval of the argument values which were taken into consideration in our analysis, the relative error significantly improves, that is, RE decreases. The choice of the specific value of parameter k can be done by comparing the average RE,

and choosing the lowest value among the considered ones, which will certainly depend on the specific application, i.e. the interval of argument values taken into consideration, which are important for the specific application. Note that in this paper, the interval of values of argument x is specified as in [21] for the purposes of applying the scalar companding quantization of the Gaussian source. The results show that our modification of the Q-function approximation is very meaningful, since the change in the value of parameter k can increase the accuracy of the approximation given by Eq. (25).

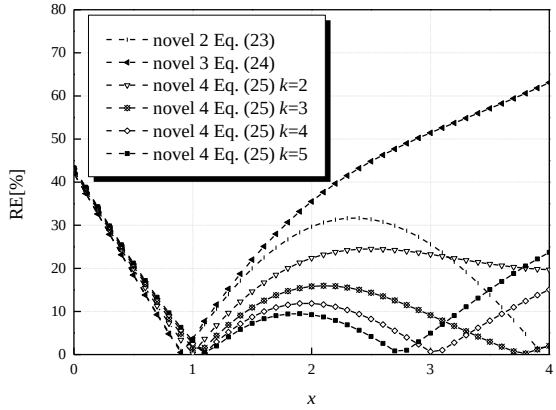


Fig. 5 Relative error: performance comparison of novel2, novel 3 and novel 4 approximations of Q-function.

In short, due to the stated properties of the proposed approximations of the Q-function of exponential type, it can be anticipated that the scope of applications of these approximations is indeed wide. The authors believe that the application of novel solutions for the approximation of the Gaussian Q-function is of great importance in neural networks as well, since the network coefficients are most often modeled by the Laplacian or Gaussian probability density function.

4. CONCLUSION

In this paper, we have highlighted one significant problem that arises in designing scalar companding quantization for the Gaussian source, which has motivated numerous studies in this field. We indicated the fact that in order to simplify the design process of this quantizer model, various methods of linearization of the compressor function have been proposed in literature, which enabled the design and performance evaluation of the linear solutions thus obtained. The authors of this paper are, to the best of their knowledge, the only ones to approach the solution of the observed problem by applying the Q-function approximations in determining the approximate compressor function. Due to early phase of this research, and due to the fact that it was noticed that it is possible to achieve further improvements, in this paper, we have defined some novel approximations of the Q-function that can be useful for overcoming the issue in designing the scalar companding quantizer for the Gaussian source. In addition to the fact that our proposed approximations are characterized by very simple analytical forms, these approximations are also characterized

by higher accuracies in relation to the approximations of the *Q*-function of comparable complexities of the analytical forms. Due to the stated properties of the proposed *Q*-function approximations of exponential type, we can anticipate that the scope of applications of these approximations is indeed wide. However, the versatility of the novel proposed *Q*-function approximations in numerous engineering problems and analyses involving the computation of the *Q*-function has been left for our future research.

Acknowledgement: *This research has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia and by the Science Fund of the Republic of Serbia, 6527104, AI-Com-in-AI.*

REFERENCES

[1] S. Na, "Asymptotic formulas for variance-mismatched fixed-rate scalar quantization of a Gaussian source," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2437–2441, May 2011.

[2] G. K. Karagiannidis and A. S. Lioumpas, "An improved approximation for the Gaussian *Q*-function," *IEEE Communications Letters*, vol. 11, no. 8, pp. 644–646, August 2007.

[3] S. Aggarwal, "Application of Gaussian *Q*-function approximations in fluctuating Beckmann fading model," *National Academy Science Letters*, April 2020.

[4] M. K. Simon, Probability distributions involving Gaussian random variables: A handbook for engineers and scientists, Kluwer Academic Publishers, 2006.

[5] M. Chiani, D. Dardari and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840–845, July 2003.

[6] S. Chang, P. C. Cosman and L. B. Milstein, "Chernoff-type bounds for the Gaussian error function," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 2939–2944, November 2011.

[7] P. A. ShirinAbadi and A. Abbasi, "On approximation of Gaussian *Q*-function and its applications," *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York City, NY, USA, pp. 0883-0887, October 2019.

[8] A. Mastin and P. Jailleth, "Log-quadratic bounds for the Gaussian *Q*-function," arXiv:1304.2488, 2013.

[9] J. W. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," *IEEE MILCOM Conference record*, Boston, MA, vol. 2, pp. 571–575, November 1991.

[10] D. Sadhwani and R. N. Yadav, "A comprehensible form of the product of two Gaussian *Q* functions and its usefulness in $\kappa-\mu$ shadowed fading distribution," *Journal of the Franklin Institute*, vol. 357, no. 8, pp. 5110–5123, May 2020.

[11] I. M. Tanash and T. Riihonen, "Global Minimax Approximations and Bounds for the Gaussian *Q*-Function by Sums of Exponentials," *IEEE Transactions on Communications*, vol. 68, no. 10, pp. 6514–6524, October 2020.

[12] F. E. Bouanani, Y. Mouchtak and G. K. Karagiannidis, "New Tight Bounds for the Gaussian *Q*-Function and Applications," *IEEE Access*, vol. 8, pp. 145037–145055, August 2020.

[13] S. Aggarwal, "A survey-cum-tutorial on approximations to Gaussian *Q*-function for symbol error probability analysis over Nakagami- *m* fading channels," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2195–2223, March 2019.

[14] N. Ermolova and S. G. Haggman, "Simplified bounds for the complementary error function; application to the performance evaluation of signal-processing systems," *12th European Signal Processing Conference*, Vienna, pp. 1087–1090, September 2004.

[15] C. Tellambura and A. Annamalai, "Efficient computation of erfc(x) for large arguments," IEEE Transactions on Communications, vol. 48, no. 4, pp. 529–532, April 2000.

[16] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, December 1952.

[17] A. Gasull and F. Utzet, "Approximating Mills ratio," *Journal of Mathematical Analysis and Applications*, vol. 420, no. 2, pp. 1832–1853, December 2014.

[18] J. Nikolić, Z. Perić and A. Jovanović, "Novel approximations for the *Q*-function with application in SQNR calculation," *Digital Signal Processing*, vol. 65, pp. 71–80, June 2017.

[19] Z. H. Perić, J. R. Nikolić and M. D. Petković, "Class of tight bounds on the *Q*-function with closed-form upper bound on relative error," *Mathematical Methods in the Applied Sciences*, John Wiley & Sons, vol. 42, no. 6, pp. 1786–1794, March 2019.
[20] J. Nikolić, Z. Perić and A. Marković, "Proposal of simple and accurate two-parametric approximation for the *Q*-function," *Mathematical Problems in Engineering*, vol. 2017, ID 8140487, 10 pages, December 2017.
[21] J. Nikolić and Z. Perić, "Optimal compressor function approximation utilizing *Q*-function approximations," *Facta Universitatis Series Automatic Control and Robotics*, vol. 15, no. 2, pp. 85-94, 2016.
[22] M. K. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Transactions on Communications*, vol. 46, pp. 200–210, February 1998.
[23] S. G. From, "Some new upper and lower bounds for the Mills ratio," *Journal of Mathematical Analysis and Applications*, vol. 486, no.1, 123872, 25 pages, June 2020.
[24] M. López-Benítez and D. Patel, "Sigmoid approximation to the Gaussian *Q*-function and its applications to spectrum sensing analysis," *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, pp. 1-5, April 2019.
[25] Z.-H. Yang and Y.-M. Chu, "On approximating the error function," *Journal of Inequalities and Applications*, 2016:311, 17 pages, November 2016.
[26] D. Sadhwani, R. N. Yadav and S. Aggarwal, "Tighter bounds on the Gaussian *Q* function and its application in Nakagami-*m* fading channel," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 574–577, October 2017.
[27] Z. Perić and J. Nikolić, "Asymptotic analysis of switched uniform polar quantization for memoryless Gaussian source," *IEEE Signal Processing Letters*, vol. 20, no. 1, pp. 75–78, January 2013.
[28] J. Nikolić, Z. Perić and L. Velimirović, "Simple solution for designing the piecewise linear scalar companding quantizer for Gaussian source," *Radioengineering*, vol. 22, no. 1, pp. 194-199, April 2013.
[29] J. Nikolić, Z. Perić, D. Aleksić and D. Antić, "Linearization of optimal compressor function and design of piecewise linear compandor for Gaussian source," *Advances in Electrical and Computer Engineering*, vol. 13, no. 4, pp. 73-78, 2013.
[30] J. Nikolić, Z. Perić and A. Jovanović, "Variance mismatch analysis of unrestricted polar quantization for Gaussian source," *IEEE Signal Processing Letters*, vol. 21, no. 5, pp. 540-544, May 2014.
[31] Z. Perić and J. Nikolić, "Design of asymptotically optimal unrestricted polar quantizer for Gaussian source," *IEEE Signal Processing Letters*, vol. 20, no. 10, pp. 980–983, October 2013.
[32] J. Nikolić, Z. Perić and A. Jovanović, "Two forward adaptive dual-mode companding scalar quantizers for Gaussian source," *Signal Processing*, vol. 120, no. 3, pp. 129-140, March 2016.
[33] D. Hui and D. Neuhoff, "Asymptotic analysis of optimal fixed-rate uniform scalar quantization", *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 957–977, March, 2001.
[34] Z. Perić, J. Lukić, J. Nikolić and D. Denić, "Application of mean-square approximation for piecewise linear optimal compander design for Gaussian source and Gaussian mixture model," *Information Technology and Control*, vol. 42, no. 3, pp. 277-285, 2013.
[35] D. Marco and D. Neuhoff, "Low-resolution scalar quantization for Gaussian sources and squared error", *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1689–1697, April, 2006.
[36] Z. Perić, N. Simić and J. Nikolić, "Design of single and dual-mode companding scalar quantizers based on piecewise linear approximation of the Gaussian pdf," *Journal of the Franklin Institute*, vol. 357, no. 9, pp. 5663–5679, June 2020.
[37] D. Marco and D. Neuhoff, "Low-resolution scalar quantization for Gaussian sources and absolute error", *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1177–1179, March, 2007.
[38] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, 1st ed. London, U.K.: Wiley, 1965.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

62

FACTA Universitatis. Series, Automatic Control and Robotics /
University of Niš ; editor-in-chief Marko Milojković. - Vol. 7, no. 1
(2008)- . - Niš : University of Niš, 2008- (Niš : Atlantis). - 24 cm

Dostupno i na: http://casopisi.junis.ni.ac.rs/index.php/FUAutContRob. -
Tri puta godišnje. - Delimično je nastavak: Facta Universitatis. Series: Automatic
Control and Robotics = ISSN 0354-2009. - Drugo izdanje na drugom medijumu:
Facta Universitatis. Series: Automatic Control and Robotics (Online) = ISSN 1820-6425
ISSN 1820-6417 = Facta Universitatis. Series: Automatic Control and Robotics
COBISS.SR-ID 158108940