DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Volume 32 Number 1 (2024)

Founded in 1993

Founder: PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA

DOI: 10.22363/2658-4670-2024-32-1

Edition registered by the Federal Service for Supervision of Communications, Information Technology and Mass Media

Registration Certificate: ПИ № ФС 77-76317, 19.07.2019

ISSN 2658-7149 (Online); 2658-4670 (Print)
4 issues per year.
Language: English.

Publisher

Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University).

Indexed by

– Scopus (https://www.scopus.com),
– Ulrich's Periodicals Directory (http://www.ulrichsweb.com),
– Directory of Open Access Journals (DOAJ) (https://doaj.org),
– Russian Index of Science Citation (https://elibrary.ru),
– CyberLeninka (https://cyberleninka.ru).

Aim and Scope

Discrete and Continuous Models and Applied Computational Science arose in 2019 as a continuation of RUDN Journal of Mathematics, Information Sciences and Physics. RUDN Journal of Mathematics, Information Sciences and Physics arose in 2006 as a merger and continuation of the series "Physics", "Mathematics", "Applied Mathematics and Computer Science", "Applied Mathematics and Computer Mathematics".

Discussed issues affecting modern problems of physics, mathematics, queuing theory, the Teletraffic theory, computer science, software and databases development.

It's an international journal regarding both the editorial board and contributing authors as well as research and topics of publications. Its authors are leading researchers possessing PhD and PhDr degrees, and PhD and MA students from Russia and abroad. Articles are indexed in the Russian and foreign databases. Each paper is reviewed by at least two reviewers, the composition of which includes PhDs, are well known in their circles. Author's part of the magazine includes both young scientists, graduate students and talented students, who publish their works, and famous giants of world science.

The Journal is published in accordance with the policies of COPE (Committee on Publication Ethics). The editors are open to thematic issue initiatives with guest editors. Further information regarding notes for contributors, subscription, and back volumes is available at http://journals.rudn.ru/miph

E-mail: miphj@rudn.ru, dcm@sci.pfu.edu.ru

© RUDN University, 2024

Editorial board

Editor-in-Chief

Yury P. Rybakov, Doctor of Sciences in Physics and Mathematics, Professor, Honored Scientist of Russia, Professor of the Institute of Physical Research & Technologies, RUDN University, Moscow, Russia

Vice Editors-in-Chief

Leonid A. Sevastianov, Doctor of Sciences in Physics and Mathematics, Professor, Professor of the Department of Computational Mathematics and Artificial Intelligence, RUDN University, Moscow, Russia

Dmitry S. Kulyabov, Doctor of Sciences in Physics and Mathematics, Docent, Professor of the Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia

Members of the editorial board

Konstantin E. Samouylov, Doctor of Sciences in Technical Sciences, Professor, Head of Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia
Yulia V. Gaidamaka, Doctor of Sciences in Physics and Mathematics, Professor, Professor of the Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia
Gleb Beliakov, PhD, Professor of Mathematics at Deakin University, Melbourne, Australia
Michal Hnatič, DrSc, Professor of Pavol Jozef Safarik University in Košice, Košice, Slovakia
Datta Gupta Subhashish, PhD in Physics and Mathematics, Professor of Hyderabad University, Hyderabad, India
Olli Erkki Martikainen, PhD in Engineering, member of the Research Institute of the Finnish Economy, Helsinki, Finland
Mikhail V. Medvedev, Doctor of Sciences in Physics and Mathematics, Professor of the Kansas University, Lawrence, USA
Raphael Orlando Ramírez Inostroza, PhD, Professor of Rovira i Virgili University (Universitat Rovira i Virgili), Tarragona, Spain
Bijan Saha, Doctor of Sciences in Physics and Mathematics, Leading Researcher in Laboratory of Information Technologies of the Joint Institute for Nuclear Research, Dubna, Russia
Ochbadrah Chuluunbaatar, Doctor of Sciences in Physics and Mathematics, Leading Researcher in the Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Mongolia

English Text Editors: *Nikolay E. Nikolaev, Ivan S. Zaryadov, Konstantin P. Lovetskiy*

Address of editorial board:
3 Ordzhonikidze St., 115419 Moscow, Russia
Tel. +7 (495) 955-07-16, e-mail: publishing@rudn.ru
Editorial office:
Tel. +7 (495) 952-02-50, miphj@rudn.ru, dcm@sci.pfu.edu.ru
site: http://journals.rudn.ru/miph

Paper size 70×108/16. Offset paper. Offset printing. Typeface "Adobe Source".
Conventional printed sheet 14.17. Printing run 500 copies. Open price. The order 2.
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA
6 Miklukho-Maklaya St., 117198 Moscow, Russia
Printed at RUDN Publishing House:
3 Ordzhonikidze St., 115419 Moscow, Russia,
Ph. +7 (495) 952-04-41; e-mail: publishing@rudn.ru

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 4
http://journals.rudn.ru/miph

Contents

Ermolayeva A. M. Statistical methods for estimating quartiles of scientific conferences 5

Zaryadov I. S., Viana H. C.C., Korolkova A. V., Milovanova T. A. Chronology of the development of Active Queue Management algorithms of RED family. Part 2: from 2006 up to 2015 18

Ayriyan A., Grigorian H., Papoyan V. Sampling of integrand for integration using shallow neural network 38

Stepa C. A., Fedorov A. V., Gevorkyan M. N., Korolkova A. V., Kulyabov D. S. Solving the eikonal equation by the FSM method in Julia language 48

Vasilyeva I. I., Demidova A. V., Druzhinina O. V., Masina O. N. Computer research of deterministic and stochastic models "two competitors—two migration areas" taking into account the variability of parameters 61

Lovetskiy K. P., Sergeev S. V., Kulyabov D. S., Sevastianov L. A. Application of the Chebyshev collocation method to solve boundary value problems of heat conduction 74

Zhivtsova A. A., Beschastnyy V. A. A new link activation policy for latency reduction in 5G integrated access and backhaul systems 86

Nikonov E. G., Nazmitdinov R. G., Glukhovtsev P. I. Computer studies of a dependence of equilibrium state structure on a number of particles for a two-dimensional system of charged particles confined in a disk potential 99

Zemlyanaya E., Bogolubskaya A., Bashashin M., Alexeeva N. Numerical study of the φ^4 standing waves in a ball of finite radius 106

Khankhasaev V. N., Bairov S. A. The numerical solution of the nonlinear hyperbolic-parabolic heat equation 112

Karnilovich S. P., Lovetskiy K. P., Sevastianov L. A., Strashnova S. B., Shaar Y. N. On cyclotron damping of longitudinal wave 122

© RUDN University, 2024

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 5–17
http://journals.rudn.ru/miph

Research article

UDC 519.23
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-5-17 EDN: BAUTRQ

Statistical methods for estimating quartiles of scientific conferences

Anna M. Ermolayeva

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: January 27, 2024; revised: February 16, 2024; accepted: March 1, 2024)

Abstract. The article presents the results of the evaluation of quartiles of scientific conferences presented by leading rating agencies. The estimates are based on the use of three methods of multivariate statistical analysis: linear regression, discriminant analysis and neural networks. A training sample was used for evaluation, including the following factors: age and frequency of the conference, number of participants and number of reports, publication activity of the conference organizers, citation of reports. As a result of the study, the linear regression model confirmed the correctness of the quartiles exposed for 77% of conferences, while the methods of neural networks and discriminant analysis gave similar results, confirming the correctness of the quartiles exposed for 81 and 85% of conferences, respectively.

Key words and phrases: evaluation of quartiles of scientific conferences, discriminant analysis, neural networks, linear regression

1. Introduction

As it is known [1], quartile (quarter) is a category of scientific publications, which is determined by bibliometric indicators reflecting, first of all, the level of citation, that is, the relevance of the publication by the scientific community. And if the procedure for assigning quartiles to scientific journals has long been developed and successfully applied in practice [2–5]. In addition, many metrics have been introduced to assess the impact of journals, such as impact factor, 5-year impact factor, immediacy index, and impact factor without self cites, median impact factor, aggregate impact factor and others [6]. At the same time, this issue remains the subject of research for scientific conferences [7–11]. Some rating agencies have already begun to rank scientific conferences without disclosing the details of this procedure. For example, there is a CORE conference ranking [12], a CCF conference ranking [13], and a Microsoft Academic conference ranking (has been deleted) [14]. The disadvantages of the first two ratings are that they are expert, regional and do not fully disclose the procedure for ranking conferences. They also rank only computer science conferences.

Researchers use various methods to compile new conference rankings, such as correlation analysis [7, 15], statistical analysis [15, 16], calculation of indicators similar to journal ones [9], graph and tree analysis [8, 17], regression analysis [11, 16]. Many of these studies involved the use of several of the listed methods. There were also works devoted to the search for methods for predicting the rating of a conference or predicting the impact of works presented at a particular conference [18]. Machine learning was used for these purposes [19, 20]. Therefore, this study is devoted to comparing two popular methods for predicting conference rankings, and I also included in the study such a statistical method as discriminant analysis, which is essentially a mathematical prerequisite for machine learning.

We managed to find data on some conferences via the Internet, including their quartiles and a number of other indicators, which will be discussed below. As a result, we received a training sample from 23 conferences, on the basis of which we will try to assess the adequacy of the quartiles exposed using three methods of multidimensional statistical analysis: linear regression, discriminant analysis and neural networks.

© Ermolayeva A. M., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

2. Training sample

Let's introduce the notation:

- Y is a random variable (r.v.), taking the values 1, 2, 3 or 4 is the quartile of a scientific conference;
- X_1 is a non–negative r.v., which takes values from a set of real numbers is the average citation of conference materials (the number of citations per report over the 10 years from 2011 to 2020);
- X_2 is an integer positive r.v. is the number of conference participants;
- X_3 is an integer positive r.v. is the number of reports at the conference;
- X_4 is an integer positive r.v. is the number of participants who submitted more than one report;
- X_5 is an integer value that takes two values: 0 or 1 is an indicator of the publication activity of the conference organizers (1 — if the organizers submitted a report to the conference and 0 — otherwise);
- X_6 is a non–negative r.v., which takes values from a set of real numbers is an indicator of the publication activity of the conference organizers, equal to the average citation of scientific publications per conference organizer.

The table 1 shows a training sample of the values of r.v. Y, X_1–X_6, compiled from the materials of the websites [21–23].

3. Linear regression model

Based on the data presented in the table 1, we will build a linear regression model reflecting the dependence of Y on the factors listed above. We will carry out the construction using the SPSS statistical package.

At the beginning, we will estimate the degree of linear dependence of Y on factors from X_1 to X_6 by constructing a Pearson pair correlation matrix. The study showed that a significant relationship is observed between Y and factors X_1, X_2, X_3 (table 2). The X_4–X_6 factors have little effect on the Y values, so we will not take them into account in the future. At the same time, a strong relationship is observed between factors X_2 and X_3. To avoid the negative impact of multicollinearity, we excluded factor X_3 from consideration and construct a two-factor regression model $Y(X_1, X_2)$.

As a result, the equation is obtained (see the table 3):

$$Y = -0.049 \cdot X_1 + 0.012 \cdot X_2 + 2.237. \tag{1}$$

Note that table 3 shows not only the absolute values of the coefficients of the model, but also the results of checking their significance using the T-criterion. According to the data from the last column of the table, all coefficients are significant with a significance level not exceeding 10^{-3}. In the second column of the table, estimates of the standard deviation σ_j of the coefficients of the model are calculated, as well as their values after standardization. According to these data, a change in the j-th coefficient of the model by one σ_j entails a change in Y by approximately $0.62\sigma_j$ downward for the coefficient at X_1 and by $0.517\sigma_j$ upward for the coefficient at X_2.

According to the data presented in the table 4, the constructed model reflects by 85.1% the real dependence of the quartile on the citation of materials and the number of conference participants. At the same time, 72.4% of the Y variation in our model is due to the variability of factors X_1 and X_2. The model itself is significant at a significance level not exceeding 10^{-3} (see the table 5).

There i s no autocorrelation of residues in the constructed model, because the Durbin–Watson statistics, equal to 1.722 (see the table 4), falls into the interval $(d_u; 4 - d_u)$, where $d_u = 1.33$ (according to the table of critical values for the significance level $\alpha = 0.05$).

The absence of auto-correlation of the residuals in combination with the condition of independence of the observational results actually means that the conditions of the Gauss–Markov theorem are fulfilled, on the basis of which it is true.

Statement 1. Model (1) is a model with minimal variance among all linear models of a fixed level of significance α.

Let's determine the estimate of the variance of the errors of the model (1). To do this, first solve the question of the normality of the residuals. We will check the normality using the Frosini criterion [24].

Table 1

Training sample

Number	Y	X_1	X_2	X_3	X_4	X_5	X_6
1	1	55.20	55	146	2	1	36.17
2	1	37.85	58	126	8	0	32.50
3	1	25.62	79	153	3	1	9.33
4	2	18.93	74	139	9	0	16.21
5	2	16.38	48	132	0	1	7.17
6	2	14.39	95	143	6	1	12.21
7	2	7.06	31	48	5	0	8.83
8	2	7.03	30	87	0	1	15.83
9	2	6.87	59	105	8	0	17.32
10	2	6.46	33	94	9	1	16.58
11	2	6.04	30	78	0	0	4.83
12	2	5.95	31	66	1	0	23.21
13	3	5.34	26	33	0	1	8.50
14	2	3.69	25	48	0	1	10.51
15	3	3.42	17	34	0	1	10.67
16	3	3.39	13	26	2	0	18.67
17	4	3.20	95	255	9	1	22.37
18	2	3.07	52	100	5	1	25.87
19	4	2.48	110	301	5	0	14.83
20	4	2.42	157	345	0	1	20.17
21	4	2.04	99	282	8	0	16.71
22	4	1.89	135	380	7	1	25.60
23	4	1.76	169	382	2	1	10.50

Table 2

Pearson Pair Correlation Matrix

	Y	X_1	X_2	X_3
Y	1	−0.678	0.588	0.666
X_1	−0.678	1	−0.114	−0.141
X_2	0.588	−0.114	1	0.959
X_3	0.666	−0.141	0.959	1

Table 3

Coefficients

Model	Non-standard	Standard error	Standard	t	Signific.
Constant	2.237	0.246		9.082	0.000
X_1	−0.049	0.009	−0.62	−5.244	0.000
X_2	0.012	0.003	0.517	4.377	0.000

Table 4

Summary for the model

Model	R	R^2	Adjust. R^2	Standard estim. error	Durbin–Watson
1	0.851	0.724	0.697	0.572	1.722

Table 5

Analysis of variance

Model	Sum of squar	Degr. of freed.	Stand. deviat.	F	Signific.
Regression	17.196	2	8.598	26.283	0.000
Residual	6.543	20	0.327		
Total	23.739	22			

To do this, you need to calculate the statistics:

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \Phi(z_i) - \frac{i - 0.5}{n} \right|, \tag{2}$$

where $z_i = \dfrac{x_i - \overline{x}}{s}$; $\overline{x} = \dfrac{1}{n} \sum\limits_{i=1}^{n} x_i$; $s^2 = \dfrac{1}{n} \sum\limits_{i=1}^{n} (x_i - \overline{x})^2$; $\Phi(z_i)$ is the distribution function $N(0, 1)$.

The results of the calculations are presented in the table 6.

Fixing the significance level $\alpha = 0.01$ and considering that $C_{\text{cr}}(0.01) = 0.341$ [24], we obtain:

$$B_n = 0.306 < C_{\text{cr}}(0.01) = 0.341. \tag{3}$$

Therefore, the residuals are distributed normally.

And finally, considering that the variance estimate σ^2 is determined by the formula:

$$\sigma^2 = \frac{1}{n - (p + 1)} (Y - Y^*)^T (Y - Y^*) \tag{4}$$

and is equal in our case to 0.327 (see the last column of the table 6), we come to the following result.

Statement 2. The model of linear regression of quartiles of scientific conferences, based on the data presented in the table 1, has the form:

$$Y = -0.049 \cdot X_1 + 0.012 \cdot X_2 + 2.237 + \epsilon, \tag{5}$$

where ϵ is the r.v. having a normal distribution with parameters $m = 0$ and $\sigma = 0.57$.

Table 6

Calculation of B_n statistics

Num.	Y	X_1	X_2	Y^*	$Y - Y^*$	$F(Z_i)$	$B(i)$	$(Y - Y^*)^{**}2$
1	1	25.620	79.000	1.930	−0.930	0.040	0.018	0.864
2	2	3.070	52.000	2.711	−0.711	0.090	0.025	0.505
3	2	14.390	95.000	2.672	−0.672	0.102	0.007	0.451
4	2	6.870	59.000	2.608	−0.608	0.125	0.027	0.370
5	2	3.69	25	2.356	−0.356	0.250	0.054	0.127
6	2	5.95	31	2.317	−0.317	0.274	0.035	0.101
7	2	6.46	33	2.316	−0.316	0.274	0.009	0.100
8	2	6.04	30	2.301	−0.301	0.284	0.042	0.091
9	2	7.06	31	2.263	−0.263	0.309	0.061	0.069
10	2	7.03	30	2.253	−0.253	0.316	0.097	0.064
11	2	18.93	74	2.197	−0.197	0.359	0.097	0.039
12	4	1.76	169	4.179	−0.179	0.370	0.130	0.032
13	1	37.85	58	1.078	−0.078	0.440	0.103	0.006
14	2	16.38	48	2.010	−0.010	0.494	0.093	0.000
15	4	2.42	157	4.002	−0.002	0.500	0.130	0.000
16	4	1.89	135	3.764	0.236	0.670	0.004	0.056
17	4	2.48	110	3.435	0.565	0.857	0.140	0.319
18	4	2.04	99	3.325	0.675	0.898	0.137	0.456
19	3	5.34	26	2.287	0.713	0.910	0.106	0.508
20	3	3.42	17	2.273	0.727	0.915	0.067	0.528
21	3	3.39	13	2.227	0.773	0.928	0.037	0.598
22	4	3.20	95	3.220	0.780	0.929	0.006	0.608
23	1	55.20	55	0.192	0.808	0.936	0.043	0.653
						$Sum =$	1,467	6,543
						$B(n) =$	0,306	0,327
						$C_{\text{cr}}(0,01)=$	0,341	

Further, based on the data for the three “new” conferences, we obtained the predicted values of their quartiles using the model 1 (see the table 7).

As we can see from the results presented in the table 7, conferences numbered 24 and 25 should be assigned the 1st and 2nd quartiles, respectively. With conference number 26, the picture is not so clear, because the predicted value of Y lies approximately in the middle between numbers 3 and 4, which suggests that this conference should be assigned the 4th quartile with a probability of 0.55 or the 3rd quartile with a probability of 0.45.

Table 7

The results of the calculation of the predicted quartile values

Num.	Quartile (forecast. significant.) Y	Citation X_1	Quantity participants X_2
24	1.0323	38.30	56
25	2.23101	5.51	22
26	3.55425	2.75	121

4. Discriminant analysis

Discriminant analysis is a classification method, the purpose of which is to divide the objects of observation into classes according to the values of the effective feature, depending on a number of controlled factors [24]. In our case, the effective feature is the quartile, and the controlled factors are the citation of the conference materials and the number of its participants. Our further goal is to classify new conferences using discriminant analysis, the data for which are presented in the table 7, based on the training sample presented in the table 1. To solve this problem, we still use the SPSS statistical package.

First of all, we pay attention to the data shown in the table 8. This table shows the results of checking the significance of differences in the average values of discriminant functions in data groups corresponding to factors X_1 and X_2 using the Wilkes' Lambda criterion. In our case, the significance levels for each factor do not exceed 0.05, which proves the existence of discriminating features of these factors and confirms the possibility of their use for discriminant analysis.

Table 8

The criterion of equality of group averages

Function	Wilkes' Lambda	F	Degr. of freedom 1	Degr. of freedom 2	Sgn.
X_1	0.190	26.972	3	19	0.000
X_2	0.232	20.913	3	19	0.000

According to the data presented in the table 9, the first discriminant function takes into account 67.5% of the variance of the effective feature, and the correlation between the training sample data and the data obtained by the model is 0.918, which is a fairly high indicator. For the second discriminant function, these indicators are 32.5% and 0.849, respectively. The significance of discriminant functions was assessed using the Wilkes' Lambda criterion. According to the results presented in the table 10, the significance of both discriminant functions does not exceed 0.05.

Table 9

Eigenvalues

Function	Proper. value	% of var. explained	Cumulative %	Canonical correl.
X_1	5.378^a	67.5	67.5	0.918
X_2	2.586^a	32.5	100.0	0.849

According to the table 10, we obtain the following expressions for discriminant functions:

$$D_1(X_1, X_2) = 0.141 \cdot X_1 - 0.028 \cdot X_2 + 0.352, \tag{6}$$

$$D_2(X_1, X_2) = 0.083 \cdot X_1 + 0.034 \cdot X_2 - 3.100. \tag{7}$$

Table 10

Non-normalized coefficients of canonical discriminant functions

Function	1	2
X_1	0.141	0.083
X_2	−0.028	0.034
(Constant)	0.352	−3.100

Statement 3. Discriminant functions (6) and (7) are significant at the significance level $\alpha = 0.01$.

Proof. We will evaluate the significance of discriminant functions using the Wilkes' Lambda criterion [25], according to which it is necessary to calculate statistics:

$$\chi_k^2(m_k) = -(n - ((p+g)/2 - 1)\ln \Lambda_k, \quad k = 1, 2, \dots, \tag{8}$$

where $\Lambda_1 = \frac{1}{1+\lambda_1} \cdot \frac{1}{1+\lambda_2}$, $\Lambda_2 = \frac{1}{1+\lambda_2}$, $p = 2$ is the number of discriminant features; $g = 4$ is number of groups $m_1 = p + g$; $m_2 = p$ is number of degrees of freedom.

The calculation results are presented in the table 11.

Table 11

Lambda – Wilkes Statistics

Function	Λ_k	χ_k^2	m_k
1	0.044	59.470	6
2	0.279	24.265	2

It is known [25] that statistics $\chi_k^2(m_k)$ have a χ^2 distribution with m_k degrees of freedom. Fixing $\alpha = 0.01$ and considering that $(1 - \alpha)$ are quantiles χ^2 are distributions with degrees of freedom $m_1 = 6$ and $m_2 = 2$ are 16.8 and 9.21, respectively, we arrive at the following result:

1) since $59.470 > 16.8$, the hypothesis of the significance of the discriminant function (6) is accepted;

2) since $24.265 > 9.21$, the hypothesis of the significance of the discriminant function (7) is accepted.

Thus, the statement is proved. □

The results of the analysis are presented in the table 12. As a result, quartiles 1, 2 and 4 were assigned to the "new" conferences, respectively. At the same time, quartiles were predicted for conferences numbered 24 and 26 with probabilities of 1 and 0.996. For conference number 25, the picture was not so unambiguous. It was predicted the second quartile with a probability of 0.673, or the 3*rd* quartile with a probability of 0.327.

In addition, the quartiles of conferences from the training sample were recalculated. As a result, conferences with numbers 3, 13, 15 and 16 received new quartile values. The quartiles of the remaining conferences, amounting to 82.6%, were found to be correct.

5. Neural network

To solve the classification problem, a neural network called a multilayer perceptron is best suited [26]. Typically, a network consists of one input layer, one or more hidden layers, and one output layer. Each layer consists of several neurons. The neuron processes its inputs and generates one output value, which is transmitted to the neurons in the subsequent layer. Each neuron in the input layer represents the values of one predictor from the vector $x = (x_1, x_2)$. In our case, x_1 and x_2 are the citation and the number of participants in the scientific conference.

Table 12

Classification results

Num.	Actual group	1st most likely predicted group	Group probability	2nd most likely predicted group	Group probability
1	1	1	1.000	2	0.000
2	1	1	1.000	2	0.000
3	1	2**	0.524	1	0.473
4	2	2	0.983	3	0.011
5	2	2	0.951	3	0.048
6	2	2	0.828	4	0.166
7	2	2	0.785	3	0.215
8	2	2	0.777	3	0.223
9	2	2	0.927	3	0.069
10	2	2	0.791	3	0.209
11	2	2	0.760	3	0.240
12	2	2	0.767	3	0.233
13	3	2**	0.710	3	0.290
14	2	2	0.667	3	0.333
15	3	2**	0.572	3	0.428
16	3	2**	0.524	3	0.476
17	4	4	0.786	2	0.210
18	2	2	0.868	3	0.128
19	4	4	0.981	2	0.019
20	4	4	1.000	2	0.000
21	4	4	0.905	2	0.094
22	4	4	1.000	2	0.000
23	4	4	1.000	2	0.000
24	not grouped	1	1.000	2	0.000
25	not grouped	2	0.673	3	0.327
26	not grouped	4	0.996	2	0.004

To build the network, we use the “neural networks” section of the SPSS package, in which we specify the quartile of the conference as the dependent variable, and the citation and number of participants as the covariant, and set the data division into three subsets: training, control and verification in a ratio of 20 : 3 : 3. We set the network architecture manually, fixing the presence of one hidden layer with four neurons. We select the sigmoid as the activation function for the hidden and output layers. Then we select the interactive type of training using the gradient descent method and set the time and the rule for stopping the learning process. The network parameters are shown in the figure 1, and its configuration is shown in the figure 3.

In the report presented in the figure 2, we pay attention to the lines “sum of squares error” and “relative error” in the section “test sample”. The error values turned out to be 0.016 and 0.045. These values are quite small, which indicates that the neural network is well trained.

Input layer	Covariates	1	X1
		2	X2
		Number of neurons[a]	2
		Scaling method for covariates	Standardized
Hidden layers		Number of hidden layers	1
		The number of neurons in hidden layer1[a]	4
		Activation function	Sigmoid
Output layer	Dependent variables	1	Y
		Number of neurons	1
		Method of changing the scale for quantitative dependent variables	Normalized
		Activation function	Sigmoid
		Error function	Sum of the squares

Figure 1. Network Parameters

Training sample	Sum of squares error	0,115
	Relative error	0,100
	Stop rule used	Number of consecutive steps without reducing the error: 1
	Training time	0:00:00.002
Verification sample	Sum of squares error	0,016
	Relative error	0,045

Figure 2. Summary for the model

The predicted quartile values for both “new” conferences and conferences from the training sample are contained in the fourth column of the table 13. Note that for “new” conferences, the quartiles obtained using a neural network coincide with the quartiles obtained using discriminant analysis.

Table 13

Quartile values

Num.	The actual quartile value	The value of the quartile according to the regression method	The quartile value obtained by discriminant analysis	The quartile value predicted by the neural network
1	1	1	1	1
2	1	1	1	2*
3	1	2*	2**	2*
4	2	2	2	2
5	2	2	2	2
6	2	3*	2	2
7	2	2	2	2
8	2	2	2	2
9	2	2	2	2
10	2	2	2	2*
11	2	2	2	2*
12	2	2	2	2*
13	3	2*	2**	3
14	2	2	2	2
15	3	3	2**	3
16	3	3	2**	3
17	4	4	4	4
18	2	3*	2	2
19	4	3*	4	4
20	4	4	4	4
21	4	3*	4	4
22	4	4	4	4
23	4	4	4	4
24	1	1	1	1
25	2	2	2	2
26	4	4	4	4
Num.	discrep.	6	4	5
of %	matches	76.92	84.61	80.77

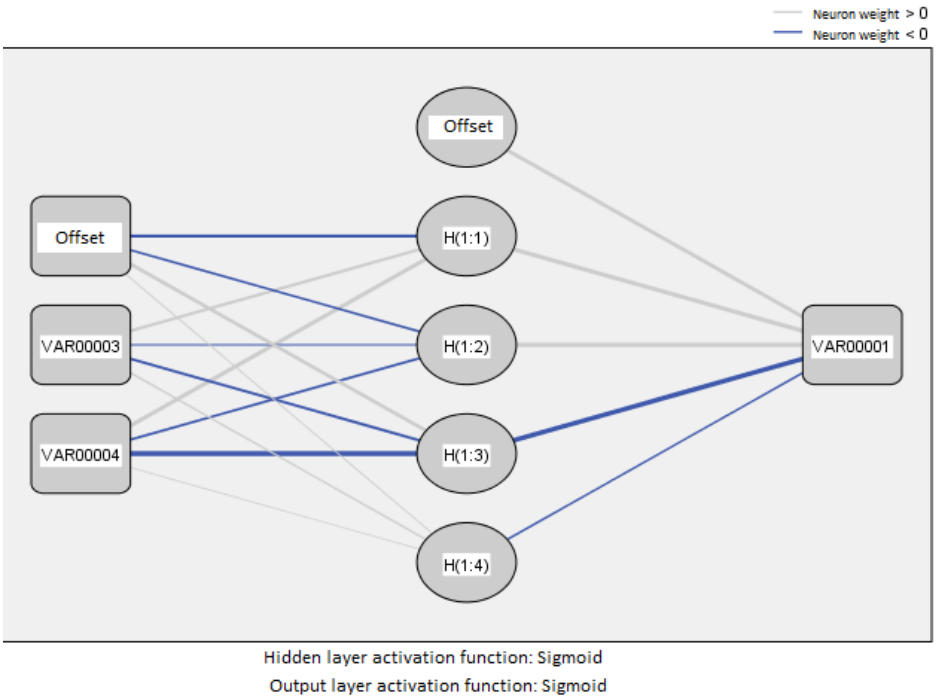


Figure 3. Neural Network configuration

6. Conclusion

As a result of the conducted research, we calculated quartiles of scientific conferences using three different methods. The results of the calculations are shown in the table 13. The quartiles marked with asterisks do not match those that were put up by rating agencies and which we called actual.

The last row of the table 13 shows the percentage of matches of the actual quantiles and quartiles calculated using the appropriate method. As we can see, the best indicator is for the discriminant analysis (4 discrepancies). In second place, with a difference of one conference, is the neural network. In third place is the linear regression method, which revealed 6 discrepancies.

Funding: The publication has been prepared with the support of the RUDN University Strategic Academic Leadership Program.

References

1. Prakash, B. *Quartiles of the journals and the secret of publishing* https://www.manuscriptedit.com/scholar-hangout/quartiles-of-the-journals-and-the-secret-of-publishing/.
2. Garfield, E. Citation indexes for science: A new dimension in documentation through association of ideas. *Science* **122,** 108–111 (1955).
3. Bergstrom, C. T., West, J. D. & Wiseman, M. A. The eigenfactor metrics. *Journal of neuroscience* **28,** 11433–11434 (2008).
4. Moed, H. F. Measuring contextual citation impact of scientific journals. *Journal of informetrics* **4,** 265–277. doi:10.1016/j.joi.2010.01.002 (2010).
5. González-Pereira, B., Guerrero-Bote, V. P. & Moya-Anegón, F. A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of informetrics* **4,** 379–391. doi:10.1016/j.joi.2010.03.002 (2010).
6. Kim, K. & Chung, Y. Overview of journal metrics. *Science Editing* **5,** 16–20 (2018).

7. Freyne, J., Coyle, L., Smyth, B. & Cunningham, P. Relative status of journal and conference publications in computer science. *Communications of the ACM* **53,** 124–132. doi:10.1145/1839676.1839701 (2010).
8. Jahja, I., Effendy, S. & Yap, R. H. Experiments on rating conferences with CORE and DBLP. *D-Lib Magazine* **20.** doi:10.1045/november14-jahja (2014).
9. Meho, L. I. Using Scopus's CiteScore for assessing the quality of computer science conferences. *Journal of Informetrics* **13,** 419–433. doi:10.1016/j.joi.2019.02.006 (2019).
10. Effendy, S. & Yap, R. H. C. *Investigations on rating computer sciences conferences: an experiment with the Microsoft Academic Graph Dataset* Apr. 2016. doi:10.1145/2872518.2890525.
11. Lee, D. H. Predictive power of conference-related factors on citation rates of conference papers. *Scientometrics* **118,** 281–304. doi:10.1007/s11192-018-2943-z (2019).
12. *Core conference ranking* http://portal.core.edu.au/conf-ranks/.
13. *CCF conference ranking* https://www.ccf.org.cn/en/.
14. *Microsoft Academic's field ratings for conferences* https://www.microsoft.com/en-us/research/project/academic/articles/microsoft-academic-analytics/.
15. Vrettas, G. & Sanderson, M. Conferences versus journals in computer science. *Journal of the Association for Information Science and Technology* **66,** 2674–2684 (2015).
16. Li, X., Rong, W., Shi, H., Tang, J. & Xiong, Z. The impact of conference ranking systems in computer science: A comparative regression analysis. *Scientometrics* **116,** 879–907. doi:10.1007/s11192-018-2763-1 (2018).
17. Küngas, P., Karus, S., Vakulenko, S., Dumas, M., Parra, C. & Casati, F. Reverse-engineering conference rankings: what does it take to make a reputable conference? *Scientometrics* **96,** 651–665. doi:10.1007/s11192-012-0938-8 (2013).
18. Steck, H. *Evaluation of recommendations: rating-prediction and ranking* in *Proceedings of the 7th ACM conference on Recommender systems* (2013), 213–220. doi:10.1145/2507157.2507160.
19. Chowdhury, G. R., Al Abid, F. B., Rahman, M. A., Masum, A. K. M. & Hassan, M. M. *Prediction of upcoming conferences ranking in Bangladesh based on analytic network process and machine learning* in *2018 International Conference on Innovations in Science, Engineering and Technology (ICISET)* (2018), 463–467. doi:10.1109/ICISET.2018.8745590.
20. Udupi, P. K., Dattana, V., Netravathi, P. & Pandey, J. *Predicting global ranking of universities across the world using machine learning regression technique* in *SHS Web of Conferences* **156** (2023), 04001.
21. *Scopus* https://www.scopus.com.
22. *DBLP* https://dblp.org/.
23. *Google Scholar* https://scholar.google.com/.
24. Kobzar, A. I. *Applied mathematical statistics* (Fizmatlit, 2006).
25. Orlova, I. V., Kontsevaya, N. V., Turundaevsky, V. B., Urodovskikh, V. N. & Filonova, E. S. Multidimensional statistical analysis in economic problems: computer modeling in SPSS (textbook). *International Journal of Applied and Fundamental Research,* 248–250 (2014).
26. Gafarov, F. M., Galimyanov, A. F., *et al. Artificial neural networks and applications* (Kazan Publishing House University, Kazan, 2018).

To cite: Ermolayeva A. M., Statistical methods for estimating quartiles of scientific conferences, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)5–17.DOI: 10.22363/2658-4670-2024-32-1-5-17.

Information about the authors

Ermolayeva, Anna M.—Assistant of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: ermolaeva-am@rudn.ru, ORCID: https://orcid.org/0000-0001-6107-6461)

УДК 519.23
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-5-17 EDN: BAUTRQ

Статистические методы оценки квартилей научных конференций

А. М. Ермолаева

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. В статье представлены результаты оценки квартилей научных конференций, выставленных ведущими рейтинговыми агентствами. Оценки получены на основе применения трёх методов многомерного статистического анализа: линейной регрессии, дискриминантного анализа и нейронных сетей. Для оценки использовалась обучающая выборка, включающая следующие факторы: возраст и периодичность конференции, количество участников и количество докладов, публикационная активность организаторов конференции, цитируемость докладов. В результате проведённого исследования линейная регрессионная модель подтвердила верность выставленных квартилей для 77% конференций, в то время как методы нейронных сетей и дискриминантного анализа дали близкие результаты, подтвердив верность выставленных квартилей для 81 и 85% конференций соответственно.

Ключевые слова: оценка квартилей научных конференций, дискриминантный анализ, нейронные сети, линейная регрессия

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 18–37
http://journals.rudn.ru/miph

Review

UDC 519.872
DOI: 10.22363/2658-4670-2024-32-1-18-37
EDN: CAMXUN

Chronology of the development of Active Queue Management algorithms of RED family. Part 2: from 2006 up to 2015

Ivan S. Zaryadov[1,2], Hilquias C.C. Viana[1], Anna V. Korolkova[1], Tatiana A. Milovanova[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[2] *Federal Research Center "Computer Science and Control" of the RAS, 44 Vavilova St, bldg 2, Moscow, 119333, Russian Federation*

(received: February 12, 2024; revised: March 10, 2024; accepted: March 12, 2024)

Abstract. This work is the second part of a large bibliographic review of active queue management algorithms of the Random Early Detection (RED) family, presented in the scientific press from 1993 to 2023. This part provides data on algorithms published from 2006 to 2015.

Key words and phrases: active queue management, AQM, random early detection, RED, congestion control

1. Introduction

This work is the second part of the brief bibliographic review of algorithms of the Random Early Detection (RED) family, compiled according to the dates of publication of scientific works (articles and conference proceedings) in which the algorithms in question were presented to the public. The first part was presented in [1].

The authors do not claim that the prepared review includes all existing algorithms, but is the most complete of those published previously, since it includes bibliographic data on 240 algorithms.

The characteristics of the RED algorithm are the following:

- The algorithm is extremely simple.
- The algorithm is designed to use as few computing resources as possible.
- The main computational complexity comes from calculating the reset (drop) function.
- Due to the use of a moving average in the algorithm, RED handles burst traffic well.
- In TCP/IP networks, the RED algorithm helps eliminate the global synchronization problem. It occurs when multiple sources operating over the same congested network segment experience packet loss. As a consequence, these sources simultaneously reduce the speed and then (also simultaneously) gradually increase it, which leads to new congestion, packet loss and repetition of the entire procedure. The network state periodically changes from idle to overloaded.
- RED allows to avoid global synchronization by selectively destroying packets from specific sources.

The mathematical model of RED-type algorithms is quite complex, so it is extremely problematic to substantiate the influence of different elements of the algorithm on the quality of its work.

The key disadvantage of RED, which many experts report, is the lack of a strict algorithm for setting RED parameters [2]. The RED parameters set:

- thresholds;
- reset (drop) function form;
- buffer size.

Changing some parameters immediately affects others.

Modifications of the RED algorithm consisted, as a rule, either in changing the number and/or value of thresholds (multiple thresholds to define different functions on different segments may be used for this purpose), or in changing the type of drop function (a single linear function was replaced

© Zaryadov I. S., Viana H. C.C., Korolkova A. V., Milovanova T. A., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

by several linear or nonlinear ones, or combinations of linear and nonlinear functions, in order to obtain the behavior as desired under the default configuration), or in replacing the average queue size $\hat{Q}$ by the current (instant) queue size q, either in the simultaneous use of the average $\hat{Q}$ and current q queue lengths, or in the dynamic change of one or several parameters (threshold values $Q_{\min}$ and $Q_{\max}$, maximum drop probability $p_{\max}$) depending on control parameters (queue size, incoming rate, rate of queue size change), or in the use of methods of fuzzy logic, Q-learning, neural networks to determine the optimal algorithm parameter values.

The changes also affected whether the new algorithm was being developed to manage a single incoming traffic flow or multiple incoming flows with different priorities.

Since quite a lot of similar functions can be specified (with minimal effort), and there is no adequate mathematical model yet, research to improve RED will continue indefinitely.

The review is structured as follows. Section 2 provides a brief description of the RED algorithm. Each subsequent section is dedicated to one year, and it presents algorithms of the RED family, scientific publications (articles in scientific journals, conference proceedings, technical reports, etc.) on which were presented this year. In Section 13 the authors discussed the results and the future research directions are highlighted.

2. The classic RED algorithm

In 1993 the classical Random Early detection (RED) algorithm was introduced in [3].

The classic RED algorithm is a queueing discipline with two thresholds ($Q_{\min}$ and $Q_{\max}$) and a low-pass filter to calculate the average queue size $\hat{Q}$ [3]:

$$\hat{Q}_{k+1} = (1 - w_q)\hat{Q}_k + w_q\hat{Q}_k, \quad k = 0, 1, 2, \dots, \tag{1}$$

where w_q, $0 < w_q < 1$ is a weight coefficient of the exponentially weighted moving-average and determines the time constant of the low-pass filter. As said in [3] RED monitors the average queue size and drops (or marks when used in conjunction with ECN) packets based on statistical probabilities $p(\hat{Q})$:

$$p(\hat{Q}) = \begin{cases} 0, & 0 \leqslant \hat{Q} < Q_{\min}, \\ \dfrac{\hat{Q} - Q_{min}}{Q_{\max} - Q_{\min}} p_{\max}, & Q_{\min} \leqslant \hat{Q} < Q_{\max}, \\ 1, & \hat{Q} \geqslant Q_{\max}, \end{cases} \tag{2}$$

where $p_{\max}$ is the fixed maximum value of drop (marking) probability if the threshold $Q_{\max}$ is overcome.

The principles of RED work:

- RED tracks the average queue size and dropped packets based on statistical probability;
- RED can also use ECN mark tracking;
- if the buffer is almost empty, then all packets are passed through as normal;
- as the queue begins to grow, the probability of packets being dropped also begins to increase;
- when the buffer is completely full, the probability becomes one and all incoming packets are discarded;
- we can say that when a router's buffer fullness exceeds some threshold value, the probability of dropping an incoming packet depends on the extent to which this threshold value is exceeded.

Analysis and criticism of proposed AQM algorithm are presented in the works [2, 4–9].

Suggestions for tuning and optimizing the key parameters of the algorithm are proposed in the following works [10–21]. The implementation of RED in the Next Generation Passive Optical Network (NG-PON) was presented in [22].

3. 2006

The State Dependent Random Early Detection (SDRED) algorithm [23] to improve delay and jitter performances by adjusting such RED parameters the maximum threshold $Q_{\max}$ and queue weight w_q to four different levels according to queue status.

In [24] the new Active Queue Management algorithm for 3G Radio Network Controllers, called Time-to-live based RED (TTLRED), was introduced and compared with three other AQM mechanisms: RED [3], ARED [25] and Packet Discard Prevention Counter (PDPC) [26]. Two version of TTLRED were given. The first one is TTLRED for incoming packets, where the averaged queue size in the Gentle RED algorithm [27] was replaced with a packet lifetime and the drop counter was also modified (only every N-th packet can be dropped even if the $Q_{\max}$ is exceeded. The second version is TTLRED for buffered packets, where packets were given a random dropping (or marking) time while entering the PDCP buffer.

In Nonlinear RED (NLRED) algorithm [28] in order to make the packet dropping function more flexible (packet dropping becomes gentler than RED at light traffic load but more aggressive at heavy load) it was proposed to replace the linear packet dropping function in RED [3] by a judicially designed nonlinear quadratic function and the rest of the original RED remained unchanged. The other works on NLRED are [29], where the optimal 3-rd order polynomial packet dropping function for NLRED in the presence of self-similar traffic was presented, and [30], where the comparative analysis of NLRED with other queueing mechanisms was made.

In order to improve the bandwidth fairness of classic RED [3] in [31] the new AQM algorithm called Bandwidth-Fair RED (BF-RED) was proposed. This algorithm, according to the authors, guarantees the bandwidth fairness by controlling the throughput of high bandwidth flows to protect the rest of the traffics when the router tends to be congested. The main idea of BF-RED algorithm is to evaluate the bandwidth usage by monitoring the queue length and to increase the drop probability for the high bandwidth flows when the bandwidth of the router is not enough.

In [32] the new AQM algorithm, Hybrid Random Early Detection (HRED), combining the more effective elements of such algorithms as BLUE [33, 34] and the proportional controller [35] with RED core by extending RED with a second probability parameter to provide stability under steady load, was proposed. HRED maps instantaneous queue length to a drop probability, automatically adjusting the slope and intercept of the mapping function to account for changes in traffic load and to keep queue length within the desired operating range.

The modified version of Adaptive RED [25], named Refined Adaptive RED (Re-ARED), was introduced in [36]. This algorithm rearranges queue size near the specified target queue size.

The another version of Adaptive RED [25] was Stabilized ARED (SARED) algorithm, proposed in [37]. In this algorithm the Exponential Weighted Moving Average (EWMA) of the instantaneous queue size from ARED was modified by using two different queue weights in order to hold queue stable and closer to target area.

Pre-estimation RED (PERED) algorithm was proposed in [38] as a new adaptive RED algorithm, which adjusts the RED parameters based on the previous estimation to gain better performance. DiffServ PERED (version of RERED for Differentiated Service (DiffServ) network) adjust the parameters not only according to the estimation made by PERED, but also based on the priorities of service subscribers.

In [39] the Queue Variation Adaptive RED (QVARED) algorithm, based on the variation of a queue per hour in order to be responded to bursty traffic more actively, was presented. This algorithm enhanced link efficiency and dynamically dropped packets due to have high throughput in case of bursty traffic.

The novel server-based congestion control algorithm, called SF-RED [40], was designed to provide inter-server fairness service in a simple and scalable manner by using multiple virtual RED queues, which are differently parameterized and are maintained in a single physical FIFO (First In First Out) queue.

The robust two-timescale simultaneous perturbation stochastic approximation algorithm with deterministic perturbation sequences for optimization of RED parameters was proposed in [41] and named Optimized RED (O-RED).

4. 2007

Another modification of the ARED algorithm [25], depending on changing the maximum dropping probability $p_{\max}$ by calculating the dropping probability at every arrival on the basis of a linear equation that reflects the slope of the drop probability — mean queue length and traffic load characteristics, was presented in [42]. The variation of the maximum probability in the proposed algorithm depends on the incoming arrival rate and the current average mean queue length. The

slope of the curve of the drop probability function is varied according to the variation of incoming arrival rate.

Adaptive Virtual Queue RED (AVQRED), described in [43, 44], was designed as modification of VQ-RED (VQRED) [45] for satellite networks. The AVQRED algorithm constructs a virtual queue and feeds the virtual queue sizes to the RED algorithm instead of feeding the weighted average queue sizes to it. AVQRED reshapes the incoming traffic according to the desired link utilization, so the RED parameters w_q and $p_{\max}$ are no longer in the algorithm because their functionalities are replaced by the desired link utilization.

In order to provide QoS in a multi-rate WLAN, where the throughput of every wireless node should be independent of other nodes'transmission cost (defined as the channel usage time), the Temporal Fair RED (TFRED) algorithm was introduced in [46] to address the congestion, efficiency, and fairness problems in WLAN (wireless local area network) by setting different drop probabilities for each flow going through the access point.

Fuzzy control RED (FCRED or FconRED) algorithm [47] was developed to overcome the drawbacks of the original RED by using a fuzzy controller to adjust the maximum drop probability $p_{\max}$ to stabilize the average queue length around the target queue length.

The new policy which favors packets with higher distance from source was introduced in [48] and named as Distance-Dependent RED (DDRED). The idea of DDRED is to eliminate packets having consumed fewer physical resources and thus coming from sources nearer to the congested router. By eliminating such packets, retransmission will be faster.

The new algorithm, Stochastic RED (StoRED), based on stochastic fair queuing and classic RED, was proposed in [49]. StoRED tries to enforce inter-flow fairness by distinguishing the active flows into a limited number of groups stochastically based on the bandwidth share obtained by flows. The algorithm tunes the packet dropping probability in such a way that the packets of the flow with higher transmission rate will more likely be dropped than flows with lower rate. For unresponsive flows (UDP flows) the modification of StoRED, named StoRED+, was given.

The modification of Adaptive RED (ARED) [25, 50], based on the multiplicative-increase multiplicative-decrease (MIMD) approach instead of additive-increase multiplicative-decrease (AIMD) policy and therefore named MIMD ARED, was described in [51].

In [52] the new RED algorithm named Preferential Dynamic Threshold RED (PDT-RED) and its optimized version OPDT-RED were proposed. These algorithms dynamically adjusted values of $Q_{\min}$ and $Q_{\max}$ thresholds by taking into account packets priority (all incoming packets are divided into several types with assigned priorities, packets with higher priority are dropped (marked) with a lower probability) and unused buffer space.

The idea of using neural networks for early congestion prediction (a prediction tool to determine the future values of the queue size (based on current and previous values of the queue length) and the necessity of packets to be dropped (marked) if the predicted queue size goes beyond the targeted value) is at the heart of the Neural Networks RED (NN-RED) algorithm [53].

The Dual RED algorithm for active congestion control specific for the handover traffic was proposed in [54]. The proposed mechanism is called Dual RED, because when the handover is finished the original RED (oRED) is used unchanged and an addon RED (aRED) is only used temporarily during a handover to handle the handed-over traffic. The structure of Dual RED in some way is similar to RIO [55].

In [56] the novel method using stochastic learning automata and thus named Learning automata RED (LARED) was introduced. In LARED the RED thresholds $Q_{\min}$ and $Q_{\max}$ in order to guarantee the target value of the delay are dynamically adjusted.

In [57] the authors put forward a fuzzy self-tuning PD-RED algorithm that uses a fuzzy controller to regulate the Proportional Derivative (PD) controller's parameters in real-time, which allows to gain rapid response and keep the queue length within the target values.

5. 2008

The linear model of Gentle RED (GRED) [27], named GRED Linear, was presented in [58]. The discrete-time queuing model of GRED algorithm was constructed and analyzed the performance of a single queue node. It was proposed to decreases linearly the packet arrival probability value in order to control congestion. The comparison of GRED Linear with other versions of RED algorithm was conducted in [59].

The improved RED algorithm named Enhanced RED (ERED,EnRED) based on congestion detection and congestion avoidance was introduced in [60]. In order to adapt RED parameters to the dynamic changes of network in the practical application it was proposed to combine the current queue length Q and the average queue length $\hat{Q}$ to a single parameter $\hat{Q}$ by using the weight w_q.

In [61] it was proposed to use a fuzzy logic (FL) controller technique based on the traditional RED algorithm to discover the congested router buffer as soon as the congestion occurs in the network. The new algorithm was named as Fuzzy-logic Controller-based RED (FConRED). This algorithm uses the average queue length and the packet loss rate as input linguistic variables and produces the packet dropping probability as a single output linguistic variable.

The new version of Modified RED (MRED) [62], named Progressive RED (PRED), was introduced in [63]. PRED not only adopts the instantaneous queue size to adjust the threshold $Q_{\max}$ dynamically to make adaptive response toward the instantaneous network condition, but also regulates the packet dropping probability non-linearly and progressively to adapt to various network conditions by comparing the instantaneous queue size with the progressive maximum queue threshold parameters.

The Kohonen-RED (KRED) algorithm, based on Kohonen neural network model, was described in [64]. It was proposed to compute the value of $p_{\max}$ parameter by using a Kohonen neural network.

In [65] two multiple RED routers algorithms of congestion control for two layers' network were suggested. The first algorithm — fair AQM algorithm, which extends the classic RED algorithm to edge routers and core routers, implements fair link capacities of edge routers. The second algorithm (more precisely, two versions of the algorithm) — unequal AQM algorithm tries to control the link capacities for the accessed edge routers and implements unequal link capacities of edge routers, but requires the core routers to have multi-queues buffers. Two versions of unequal AQM algorithm for the core router are considered: if the number of queues equals that of accessing edge routers and if the number of queues is far smaller than that of accessing edge routers.

In [66] the Autonomous RED (AURED) algorithm allowing a complete autonomous $p_{\max}$ adjustment process by tuning the packet drop probability according to the performance variation between two consecutive sampling periods was proposed.

6. 2009

Two new version of classic RED — Barrier optimized RED (B-RED) and Penalty optimized RED (P-RED) were introduced in [67] as obtained by barrier (Barrier optimized RED) and the penalty (Penalty optimized RED) function approaches solutions to probabilistic constrained optimization problem by assuming a nonlinear relationship between the RED average queue length and its parameters as the variables of the optimization problem. The proof of convergence of proposed B-RED and P-RED algorithms to RED was given in [68].

In [69] the new active queue management algorithm, Effective RED (ERED), was presented. The aim of ERED was to reduce packet loss rates in a simple and scalable manner by changing $Q_{\min}$ and $Q_{\max}$ thresholds and by controlling instantaneous queue size Q together with average queue size $\hat{Q}$. Also it was proposed to control average queue size $\hat{Q}$ when connections immediately reduce their sending rate in the case of no congestion. The performance evaluation and comparison of ERED with some other Active Queue Management algorithms was presented in [70].

The idea of using hazard rate function for estimation of RED packet drop probabilities was proposed in [71] and the new algorithm was named Hazard rate estimated RED (HERED). The usage of hazard rate function for a packet dropping function allows to make packet dropping gentler at light traffic load but more aggressive at heavy load.

The Priority Self-adaptive RED (PSRED) algorithm designed for Ad hoc networks with flows with different priorities was introduced in [72, 73]. The aim of the algorithm is to establish a balance between the queue length and the packet loss probability by self-adapting the values of $p_{\max}$.

The mechanism for RED parameters (thresholds $Q_{\min}$ and $Q_{\max}$, maximum drop probability $p_{\max}$ and queue weight w_q) simple, scalable and systematic tuning in response to changing network conditions (traffic load, link capacity and round-trip time), named as Auto-Parameterization RED (APRED), was proposed in [74]. This algorithm based on nonlinear dynamic model of TCP behavior which was developed by using fluid-flow and stochastic differential equation analysis.

The packet dropping algorithm, based on a self-tuning proportional and Integral feedback controller, which considers not only the average queue length at the current time point, but also

the past queue lengths during a round-trip time to smooth the impact caused by short-lived traffic dynamics, was introduced in [75] and named Self-tuning Proportional and Integral RED (SPI-RED).

The modified version of RED algorithm [3], called Multi RED (MRED), which was proposed in [76] as an effective way to reduce packet loss rates in the internet by computing the packet drop probability based on described heuristic method and enabling bandwidth allocation over a single link to different types of traffics.

The enhanced version of classic RED, named Light-weight Flow information RED (LwFRED), was designed in [77] for improving the fairness of RED aided by light-weight flow information. The basic idea of LwFRED was to classify flows into groups (like in [49] but in a different way) and adjust the dropping probability of each group according to its behavior.

The robust AQM algorithm for wireless Ad hoc networks was proposed in [78] and named Adhoc Hazard RED (AHRED). In order to prevent congestion which behaves rapidly according to the density of traffic load the described algorithm uses a packet drop probability function based on Weibull model of a hazard rate function (the conditional probability that a system will fail instantaneously some time alter time t given that it has survived up to time t).

7. 2010

In [79] the new version of RED algorithm mechanism for congestion avoidance in wired networks based on a learning-automata-like (LAL) philosophy and therefore named learning-automata-like RED (LALRED) was introduced. The main aim of LALRED was to optimize the value of the average queue size used as a control parameter for congestion avoidance and to consequently reduce the total loss of packets at the queue. It was achieved by stationing a LAL algorithm at the gateways and by discretizing the drop probabilities. The application of LALRED for optical networks was presented in [80].

The Fuzzy Q-learning RED (FQL-RED) algorithm based on Adaptive RED (ARED) [25] and Fuzzy RED [81, 82] was proposed in [83]. This algorithm uses a Q-learning method enhanced with a fuzzy inference system in order to ensure RED with self-adaptation and improved performance.

In [84] the novel autonomous Proportional and Differential RED algorithm (NPD-RED) was proposed as an extension of RED for the TCP/RED dynamic time-delayed model in wired network and wired–wireless network routers. NPD-RED is based on a self-tuning feedback proportional and differential controller, which not only considers the instantaneous queue length at the current time point, but also takes into consideration the ratio of the current differential error signal to the buffer size. Also the new probability packet dropping function, based on changes in the instantaneous queue and the differential in queue length, is introduced.

The novel version of RIO-C (RED with In/Out and Couple queue) [85], named RIO Based On Priority and Fair (PFRIO), was presented in [86]. The main change was made in the calculation of average queue length in RIO-C algorithm to decrease the influence of the number of low drop-precedence packet on high drop precedence packet by introducing the new parameter, which makes the influence change dynamically in accordance to the number of packets.

In [87] the a discrete-time dynamical model of TCP and UDP congestion control at the hosts coupled with RED active queue management at the routers, named TCP-UDP-RED, for controlling bifurcations and chaos in the internet congestion control system for Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) network was proposed.

The Robust RED (RRED) algorithm to improve the TCP throughput against Low-rate Denial-of-Service (LDoS) attacks was presented in [88]. The basic idea behind the RRED was to detect and filter out attack packets (by using the proposed new detection algorithm) before a normal RED algorithm was applied to incoming flows.

The new active queue management algorithm for Assured Forwarding (AF) services guaranteeing minimum bandwidth provided based on RIO algorithm [55] and Weighted Fair Queuing (WFQ) was introduced in [89]. In the proposed scheme, the control functions for RIO and WFQ are enhanced under the assumption that they are used for input (RIO) and output (WFQ) queue management.

In [90] the novel algorithm called RIO based on Fractional Exponent Coupling (RIO-FEC), based on RIO-C ((RED with In/Out and Coupled queue)) and RIO-D (RED with In/Out and Decoupled queues) algorithms [85] was proposed. The novelty of RIO-FEC lies with the fact that a controllable scheme for determining the coupling level among virtual queues was introduced by expressing the coupling level as a polynomial function with a fractional exponent as the power of the polynomial.

8. 2011

The new modification of the ARED algorithm [25] was introduced in the [91] and named Self-tuning RED (St-RED). In the proposed algorithm the appropriate value of the parameter $p_{\max}$ was dynamically obtained based on history information of both $p_{\max}$ and the average queue size in a some period of time. Also the parameter w_q was properly chosen according to a linear stability condition of the average queue length.

The modified version of RIO-C [85] algorithm with only change in calculation of average queue length was proposed in [92] and named Weighted-queue-length RIO-C (WRIO). WRIO realizes relative discrimination of different drop precedence, and improves link utilization, especially for the networks with TCP traffics.

The improved RED algorithm with sinusoidal packet-marking probability and dynamic weight (SW-RED) was introduced in [93]. The proposed algorithm dynamically adjusts the average queue weight and packet drop probability, making it smaller when approaching to $Q_{\min}$, and greater when close to $Q_{\max}$.

Another version of the ARED algorithm [25], New Adaptive RED (NARED), was presented in [94]. In NARED the packet drop probability function is nonlinear smoothed by using the membership function of the ascend demi-cauchy of fuzzy distribution in order to make the speed of growth of packet loss rate relatively slow near $Q_{\min}$ threshold and relatively faster near the $Q_{\max}$ threshold. Also for adapting to the changes of network environment the parameter $p_{\max}$ is dynamically adjusted by the length of average queue.

In [95] the new approach to modify ARED algorithm [25] in order to maintain the instantaneous queue length in the buffer and estimate the packet dropping probability was developed. The proposed algorithm named as Hazard rate based Heuristic ARED (HHA) because the original ARED packet drop function was extended by using the random probability Weibull distribution with a non-linear hazard rate (instantaneous failure rate) function.

In [96] the hop-to-hop controlled hierarchical multicast congestion control mechanism combining RED [3] (the ability to keep low delay while pursuit larger throughput, but also the strong sensitivity to parameters) and hop-to-hop (HTH) algorithm (the ability to rapidly respond to congestion and effectively improve throughput as well as utilization of link, but the need to use too much router resources in case of many streams) was introduced and analyzed.

The Weighted RED (WTRED) algorithm for congestion handling in TCP networks and increasing network performance by dynamically adjusting RED's maximum threshold, minimum threshold and weight parameters (based on the actual buffer size) was proposed in [97].

The novel adaptive version of Gentle RED (GRED) [27] was proposed in [98] and named Adaptive GRED (AGRED). Adaptive GRED detects congestion at router buffers in an preliminary stage, and enhances the parameters setting of the $Q_{\max}$ threshold and the probability $p_{\max}$. The performance analysis of AGRED and comparison with RED [3] and (GRED) [27] was conducted in [99], the comparison with GRED-Linear [58] was made in [59].

In order to deal with RED such problems as sensitivity to traffic load and parameters configuration and the variety of the equilibrium queue length because of congestion degree and parameter settings the improved RED (named IRED) algorithm was presented in [100]. The proposed algorithm uses auto-tuning proportional integral (PI) probability as an adaptation mechanism designed to adjust the maximum packet drop probability $p_{\max}$ for stable average queue length.

The modification of GRED [27] and NLRED [28] was proposed in [101] and named Modified RED (MRED). It is quite similar to GRED except that the linear packet dropping function was replaced by a nonlinear (quadratic) function as in NLRED algorithm.

9. 2012

The new per user AQM policy designed for blind network optimization, named the User Random Early Drop (URED) algorithm, which is a modification of the Flow RED (FRED) [102] and enforces fair resource allocation among users (tunnels), was introduced and evaluated in [103].

The Modified Random Early Detection (ModRED) algorithm was presented in [104]. It was proposed to restrict the TCP transmission window with the flow control window instead of the congestion control window, thus controlling the transmission window with a finer granularity.

For the purpose of countering slow start that causes rapid increase in load of the bottleneck router, the Harsh RED (HRED) algorithm was proposed in [105]. By taking advantage of well-defined TCP

slowstart behavior, the HRED queue average was made to cross HRED minimum threshold $Q_{\min}$ on a timely manner in order to start "a count-down" for actual HRED dropping by setting the HRED parameters w_q and $p_{\max}$ to a very large value compared with the recommended values for RED [3]. The evaluation and comparison of HRED with other AQM algorithms was conducted in [106].

The novel algorithm to achieve bandwidth fairness of RED with packet size consideration, named Bandwidth fair considering packet size RED (BF-PS-RED) as modification of BF-RED algorithm [31], was proposed in [107]. BF-PS-RED gathers packet size information (in order to the compensate unfairness resulted from the packet size difference) and uses it for adjusting the maximum drop probability $p_{\max}$ of a high throughput flow together with drop-weight.

The modification of Refined Adaptive RED (Re-ARED) [36] algorithm, named Fast Adapting RED (FARED), was presented in [108]. This algorithm efficiently varies the maximum drop probability $p_{\max}$ to improve the overall performance of the network/ FARED algorithm retains the target range as specified in Re-ARED algorithm [36] but modifies the upper bound and lower bound for parameters α and β respectively.

The modification of ARED [25] and Re-ARED [36] algorithms, named Cautious Adaptive Random Early Detection (CARED) algorithm that dynamically varies maximum drop probability $p_{\max}$ either conservatively or aggressively based on the level of traffic load to improve the overall performance of the network, was proposed in [109]. The detailed study of the proposed CARED algorithm was carried out in [110].

10. 2013

The new version of RED algorithm, named Velocity RED (VRED) and based on using the queue length growth velocity in order to measure the congestion level in router, was introduced in [111]. VRED triggers the drop probability according to proposed metric. The usage of the queue length growth velocity, according to the authors, leads to fast reaction and hence improves the network performance because the packet dropping becomes gentler at light traffic load but more aggressive at heavy traffic.

In [112] the novel RED-based active queue management algorithm, called Full Information Feedback RED (FIF-RED). This algorithm for packet drop probability function not only considers the average queue length $\hat{Q}$ but also takes into account the growth rate of the instantaneous queue length.

The new version of RED algorithm, based on ideas of Robust RED [88] and designed to defend against DoS attacks by using the flows trust values for identification and dropping of malicious packets, was introduced in [113] and named as RED with Flow Trust (RED-FT).

The new AQM algorithm, with minimal changes to the classic RED algorithm, for providing the effective solution to avoid congestion collapse of network services by introducing new threshold U_{th} (Upper Threshold) and for better use of buffer space, was presented in [114] and named Upper threshold RED (URED).

The new modification of Gentle RED (GRED) [27] algorithm, designed or early stage congestion detection at the router buffer and named Dynamic Gentle Random Early Detection (DGRED), was introduced in [115]. The proposed DGRED algorithm depends on the stability of the average queue length at a specific level between $Q_{\min}$ and dynamically changing $Q_{\max}$ thresholds values. Also the dynamical threshold $2Q_{\max}$ is introduced. The simulation based performance evaluation of DGRED and comparison with several Active Queue Management algorithms for computer network was conducted in [116].

In [117] the packet correlated RED (PCRED) algorithm was proposed in order to forward the video packets in an efficient way by properly utilizing the unused bits of the differential service (DS) field of the IP header and providing some heuristic information for packet early discarding.

The version of RED algorithm with two nonlinear quadratic drop probability functions instead of a single linear function was introduced in [118].

The new robust RED algorithm, based on ideas of ARED [25], NLRED [28] and Re-ARED [36] algorithms and named Adaptive Nonlinear RED (ANLRED), was presented in [119]. This algorithm minimizes the parameter sensitivity of RED by making minimal algorithmic modifications without introducing some new parameters for better performance. ANLRED varies $p_{\max}$ adaptively based on the change in average queue length $\hat{Q}$.

The new version of RED, named Modified RED (MDRED), was introduced in [120]. In the proposed algorithm the queue between minimum threshold $Q_{\min}$ and maximum threshold $Q_{\max}$ was virtually

divided into smaller subparts and calculation of packet drop probability was based on the average queue size $\hat{Q}$.

11. 2014

The extension of Gentle RED (GRED) algorithm [27], depending on a fuzzy logic system which reduces the large dependency on parameter settings, was proposed in [121] and named Fuzzy Logic Controller of Gentle Random Early Detection (GREDFL). The proposed algorithm uses the average queue length and the delay rate as input linguistic variables for a fuzzy logic system. The utilized fuzzy logic system produces a single output that represents a packet dropping probability, which in turn control and prevent congestion in early stage.

The new technique for network congestion avoidance and control, based on Balanced RED algorithm (BRED) [122, 123], was presented in [124] and named Neural Network Dynamic Balanced RED (NN-DBRED) because of using time delay line neural network as system's core to detect and separate adaptive and non-adaptive flows and also limit receive rate from them, to provide fairness between flows and avoid occurring congestion and buffer overflow.

The new enhanced algorithm, based on URED [114], was proposed in [125] to reduce parameter sensitivity and to improve the network performance in a congested networks by introducing a new threshold (the upper threshold) for better use of buffer space and to queue more packets which reduces packet drops. This algorithm was named Upper Threshold RED (UT-RED).

In [126] the Modified Gaussian Function based RED (MGF-RED) algorithm was introduced. The Modified Gaussian function was used for calculating the likelihood of packet drop as the average queue length increases the minimum threshold.

The new version of classic RED algorithm [3] with a modified weighted moving average ased on a difference equation (a recursive equation) was proposed in [127]. It was shown that Depending on a particular optimality criterion, the proper values of the modified weighted moving average function can be chosen.

The modification of RED with new congestion metric (to forecast when the queue will be overloaded and to use for dynamically setting of RED parameters) based on the rate of which the queue is occupied, was considered in [128] and named QRTRED.

Two versions (Modified1 and Modified2) of Modified RED for stabilized queue (MRED-QS) were presented in [129]. In both versions the probability drop function was changed from linear to non-linear: logarithmic function for Modified1 version and exponential function for Modified2 version.

The new RED algorithm, named Random Early Dynamic Detection (REDD), was proposed in [130] in order to identify and control congestion and to enhance RED's performance in regards to mean queue length and packets waiting time. In the REDD the maximum threshold $Q_{\max}$ position is adaptively changed depending of values of average queue length $\hat{Q}$. The comparison of REDD with Adaptive GRED (AGRED) [98] and GRED-Linear [58] was conducted in [59].

In order to provide better congestion control over the network while maintaining the advantage of RED the new algorithm called Enhanced Random Early Detection (ENRED) was proposed in [131]. The algorithm depends on enhancement of the average queue size on a way that limits queue size to minimize the delay and packet loss rate as compared to RED queue by introducing the new parameter — target queue (the difference between the current queue size and the average of the maximum threshold $Q_{\max}$ and minimum threshold $Q_{\min}$).

In [132] the new RED-based AQM algorithm with modified probability drop function was presented, it was named Curvilinear Random Early Detection (CLRED). Instead of RED single linear function the two-segment (a quadratic and a linear) dropping function was proposed.

As the modification of the RIO-C algorithm [85] the Improved nonlinear RIO-C (INRIO-C) algorithm was proposed in [133]. This algorithm takes nonlinear characteristics between the average length of the queue and packets loss probability into consideration, and proposes piecewise smooth discarding functions. The algorithm adopts coupled approach to calculate the virtual queue, and sets priorities for different discarding threshold.

12. 2015

The new version of RED with better fairness in dropping packets, promoting the real-time data transfers by considering the initial occupancy of queue and frame admission control based on this information, was considered in [134] and called fairRED.

The modified version of LALRED algorithm [79], based on the concept of a Learning Automata-Like (LAL) mechanism, was introduced in [135] and named Efficient LALRED (ELALRED). ELALRED algorithm was designed for congestion avoidance in wired networks by optimizing the value of the average queue size $\hat{Q}$.

The version of RED algorithm less biased against the User Datagram Protocol (UDP) packets was proposed in [136]. It was found out by authors that modification of the maximum threshold $Q_{\max}$ and the final drop probability gives a considerable reduction in the UDP packet drop for a relatively lesser reduce in the TCP throughput.

The modified version of the Fair RED (FRED) algorithm [137], named Enhancement Fair RED (EFRED), was proposed in [138]. In EFRED the FRED probability drop function was modified and replaced by the hazard rate function.

The new version of RED with nonlinear cubic probability drop function was presented in [139] and named as Adaptive Sigmoid RED (ASRED).

The modified version of Cautious Adaptive RED (CARED) algorithm [109], designed for Heterogeneous network and based on fuzzy logic ideas, was developed in [140] and called as Fuzzy Cautious Adaptive RED (Fuzzy–CARED). In this algorithm the value of maximum drop probability $p_{\max}$ is cautiously increased and decreased based on current traffic load after detecting congestion in the network.

13. Conclusions

The presented bibliographical chronological review of active control algorithms of the RED family is the most complete both in terms of the number of algorithms reviewed (more than two hundred) and in terms of the number of scientific publications analyzed and presented. This review will be useful to researchers in the field of the congestion control.

Active queue management algorithms of the RED family are not something new for the authors of this work, as evidenced by the publications presented below [141–149].

In the future, the authors plan not only to classify the considered algorithms based on the classification criteria presented in [141, 150, 151], but also to review and classify other active queue management algorithms.

References

1. Zaryadov, I. S., Viana, H. C., Korolkova, A. V. & Milovanova, T. A. Chronology of the development of Active Queue Management algorithms of RED family. Part 1: from 1993 up to 2005. *Discrete and Continuous Models and Applied Computational Science* **31,** 305–331. doi:10.22363/2658-4670-2023-31-4-305-331 (2023).
2. May, M., Bolot, J., Diot, C. & Lyles, B. *Reasons not to deploy RED* in *Seventh International Workshop on Quality of Service. IWQoS'99. (Cat. No.98EX354)* (1999), 260–262. doi:10.1109/IWQOS.1999.766502.
3. Floyd, S. & Jacobson, V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking* **1,** 397–413. doi:10.1109/90.251892 (1993).
4. Bonald, T., May, M. & Bolot, J.-C. *Analytic evaluation of RED performance* in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)* **3** (IEEE, 2000), 1415–1424. doi:10.1109/INFCOM.2000.832539.
5. De Cnodder, S., Elloumi, O. & Pauwels, K. *RED behavior with different packet sizes* in *Proceedings ISCC 2000. Fifth IEEE Symposium on Computers and Communications* (IEEE, 2000), 793–799. doi:10.1109/ISCC.2000.860741.
6. Christiansen, M., Jeffay, K., Ott, D. & Smith, F. D. Tuning RED for Web traffic. *IEEE/ACM Transactions on Networking* **9,** 249–264. doi:10.1109/90.929849 (2001).

7. Hollot, C., Misra, V., Towsley, D. & Gong, W.-B. *A control theoretic analysis of RED* in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)* **3** (IEEE, 2001), 1510–1519. doi:10.1109/INFCOM.2001.916647.
8. Brandauer, C., Iannaccone, G., Diot, C., Ziegler, T., Fdida, S. & May, M. *Comparison of tail drop and active queue management performance for bulk-data and Web-like Internet traffic* in *Proceedings. Sixth IEEE Symposium on Computers and Communications* (IEEE, 2001), 122–129. doi:10.1109/ISCC.2001.935364.
9. Joo, C. & Bahk, S. Scalability problems of RED. *Electronics Letters* **38,** 1297–1298. doi:10.1049/el:20020744 (2002).
10. Vaidya, R. & Bhatnagar, S. Robust optimization of Random Early Detection. *Telecommunication Systems* **33,** 291–316. doi:10.1007/s11235-006-9020-2 (Dec. 2006).
11. Tan, L., Zhang, W., Peng, G. & Chen, G. Stability of TCP/RED systems in AQM routers. *IEEE Transactions on Automatic Control* **51,** 1393–1398. doi:10.1109/TAC.2006.876802 (2006).
12. Zheng, B. & Atiquzzaman, M. A framework to determine bounds of maximum loss rate parameter of RED queue for next generation routers. *Journal of Network and Computer Applications* **31,** 429–445. doi:10.1016/j.jnca.2008.02.003 (2008).
13. Zheng, B. & Atiquzzaman, M. A framework to determine the optimal weight parameter of RED in next-generation Internet routers. *International Journal of Communication Systems* **21,** 987–1008. doi:10.1002/dac.932 (2008).
14. Min, G. & Jin, X. *Performance Modelling of Random Early Detection Based Congestion Control for Multi-Class Self-Similar Network Traffic* in *IEEE International Conference on Communications* (IEEE, 2008), 5564–5568. doi:10.1109/ICC.2008.1043.
15. Chen, X., Wong, S.-C. & Tse, C. K. Adding Randomness to Modeling Internet TCP-RED Systems with Interactive Gateways. *IEEE Transactions on Circuits and Systems II: Express Briefs* **57,** 300–304. doi:10.1109/TCSII.2010.2043388 (2010).
16. Woo, S. & Kim, K. Tight Upper Bound for Stability of TCP/RED Systems in AQM Routers. *IEEE Communications Letters* **14,** 682–684. doi:10.1109/LCOMM.2010.07.100375 (2010).
17. Shahram, J. & Seyed, R. Z. An Active Queue Management for High Bandwidth-Delay Product Networks. *International Journal of Computer Theory and Engineering* **5,** 763–767. doi:10.7763/IJCTE.2013.V5.792 (2013).
18. Hendrawan & Hernandia, P. *Random Early Detection utilizing genetics algorithm* in *8th International Conference on Telecommunication Systems Services and Applications (TSSA)* (IEEE, Kuta, Bali, Indonesia, 2014), 1–7. doi:10.1109/TSSA.2014.7065952.
19. Waheed, A., Habib Khan, N., Zareei, M., Ul Isla, S., Jan, L., Iqbal Umar, A. & Ehab, M. M. Traffic queuing management in the Internet of Things: an optimized RED algorithm based approach. *Computers, Materials & Continua* **66,** 359–372. doi:10.32604/cmc.2020.012196 (2021).
20. Goudru, N. G. *Tuning Pmax in RED Gateways for QoS Enhancement in Wireless Packet Switching Networks* in *Mathematical Modeling, Computational Intelligence Techniques and Renewable Energy* (eds Sahni, M., Merigó, J. M., Sahni, R. & Verma, R.) (Springer Singapore, Singapore, 2022), 321–334. doi:10.1007/978-981-16-5952-2_28.
21. Basheer, A., Hassan, H. J. & Muttasher, G. *Intelligent Parameter Tuning Using Deep Q-Network for RED Algorithm in Adaptive Queue Management Systems* in *Micro-Electronics and Telecommunication Engineering* (eds Sharma, D. K., Peng, S.-L., Sharma, R. & Zaitsev, D. A.) (Springer Nature Singapore, Singapore, 2022), 439–446. doi:10.1007/978-981-16-8721-1_42.
22. Xu, X., Liu, B., Zhang, L., Mao, Y., Wu, X., Ren, J., Han, S., Jiang, L. & Xin, X. *Self-adaptive bandwidth scheduling based on improved Random Early Detection for NG-PON* in *18th International Conference on Optical Communications and Networks (ICOCN)* (IEEE, Huangshan, China, 2019), 1–3. doi:10.1109/ICOCN.2019.8934251.
23. Ryoo, I.-t. & Yang, M. A State Dependent RED: An Enhanced Active Queue Management Scheme for Real-Time Internet Services. *IEICE Trans. Commun.* **89-B,** 614–617. doi:10.1093/ietcom/e89-b.2.614 (2006).
24. Lakkakorpi, J. & Cuny, R. *Comparison of different active queue management mechanisms for 3G radio network controllers* in *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006* **1** (IEEE, Las Vegas, NV, USA, 2006), 80–85. doi:10.1109/WCNC.2006.1683445.
25. Floyd, S., Gummadi, R. & Shenker, S. *Adaptive RED: An Algorithm for Increasing the Robustness of RED's Active Queue Management* tech. rep. (AT&T Center for Internet Research at ICSI, 2001).

26. Sagfors, M., Ludwig, R., Meyer, M. & Peisa, J. *Queue management for TCP traffic over 3G links* in *IEEE Wireless Communications and Networking, 2003. WCNC 2003* **3** (IEEE, New Orleans, LA, USA, 2003), 1663–1668. doi:10.1109/WCNC.2003.1200636.
27. Floyd, S. *Recommendation on using the "gentle variant of RED"* tech. rep. (The ICSI Networking ans Security Gropup, 2000).
28. Zhou, K., Yeung, K. L. & Li, V. O. Nonlinear RED: A simple yet efficient active queue management scheme. *Computer Networks* **50,** 3784–3794. doi:10.1016/j.comnet.2006.04.007 (2006).
29. Domańska, J., Augustyn, D. & Domański, A. The choice of optimal 3-rd order polynomial packet dropping function for NLRED in the presence of self-similar traffic. *Bulletin of the Polish Academy of Sciences: Technical Sciences* **60,** 779–786. doi:10.2478/v10175-012-0090-x (Dec. 2012).
30. Rastogi Shubhangi ans Zaheer, H. Comparative analysis of queuing mechanisms: Droptail, RED and NLRED. *Social Network Analysis and Mining* **6,** 70. doi:10.1007/s13278-016-0382-5 (2016).
31. Jie, Y., Jun, L. & Zhenming, L. *BF-RED: A Novel Algorithm for Improving Bandwidth Fairness of RED* in *IEEE International Conference on Networking, Sensing and Control* (IEEE, Ft. Lauderdale, FL, USA, 2006), 1001–1005. doi:10.1109/ICNSC.2006.1673287.
32. Joo, C., Bahk, S. & Lumetta, S. S. A hybrid active queue management for stability and fast adaptation. *Journal of Communications and Networks* **8,** 93–105. doi:10.1109/JCN.2006.6182909 (2006).
33. Feng, W., Kandlur D. D.and Saha, D. & Shin, K. G. *BLUE: A New Class of Active Queue Management Algorithms* tech. rep. (The University of Michigan, 1999).
34. Feng, W.-c., Shin, K. G., Kandlur, D. D. & Saha, D. The BLUE active queue management algorithms. *IEEE/ACM Transactions on Networking* **10,** 513–528. doi:10.1109/TNET.2002.801399 (2002).
35. Hollot, C. V., Misra, V., Towsley, D. & Gong, W.-B. *On designing improved controllers for AQM routers supporting TCP flows* in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society* **3** (IEEE, Anchorage, AK, USA, 2001), 1726–1734. doi:10.1109/INFCOM.2001.916670.
36. Kim, T.-h. & Lee, K.-h. *Refined Adaptive RED in TCP/IP Networks* in *SICE-ICASE International Joint Conference* (IEEE, Busan, South Korea, 2006), 3722–3725. doi:10.1109/SICE.2006.314633.
37. Javam, H. & Analoui, M. *SARED: Stabilized ARED* in *International Conference on Communication Technology* (IEEE, Guilin, China, 2006), 1–4. doi:10.1109/ICCT.2006.341669.
38. Cao, Z. & Xiao, Y. *A New DiffServ Supported AQM Algorithm* in *8th international Conference on Signal Processing* **3** (IEEE, Guilin, China, 2006). doi:10.1109/ICOSP.2006.345807.
39. Seol, J.-H., Lee, K. Y. & Hong, Y. S. *Performance Improvement of Adaptive AQM Using the Variation of Queue Length* in *TENCON 2006 - 2006 IEEE Region 10 Conference* (IEEE, Hong Kong, China, 2006), 1–4. doi:10.1109/TENCON.2006.343824.
40. Wu, H.-M., Wu, C.-C. & Lin, W. *SF-RED — a novel server-based AQM to provide inter-server fairness service* in *12th International Conference on Parallel and Distributed Systems - (ICPADS'06)* **1** (IEEE, Minneapolis, MN, USA, 2006), 6. doi:10.1109/ICPADS.2006.99.
41. Vaidya, R. & Bhatnagar, S. Robust optimization of Random Early Detection. *Telecommunication Systems* **33,** 291–316. doi:10.1007/s11235-006-9020-2 (Dec. 2006).
42. Al-Raddady, F. & Woodward, M. *A New Adaptive Congestion Control Mechanism for the Internet Based on RED* in *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)* **2** (IEEE, Niagara Falls, ON, Canada, 2007), 934–939. doi:10.1109/AINAW.2007.31.
43. Byun, D. J. & Baras, J. S. *Adaptive virtual queue random early detection in satellite networks* in *Wireless Telecommunications Symposium* (IEEE, Pomona, CA, USA, 2007), 1–9. doi:10.1109/WTS.2007.4563336.
44. Byun, D. J. & Baras, J. S. *A New Rate-based Active Queue Management: Adaptive Virtual Queue RED* in *Fifth Annual Conference on Communication Networks and Services Research (CNSR'07)* (IEEE, Fredericton, NB, Canada, 2007), 389–396. doi:10.1109/CNSR.2007.3.
45. Lin, X., Chang, X. & Muppala, J. K. *VQ-RED: An efficient virtual queue management approach to improve fairness in infrastructure WLAN* in *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)l* (IEEE, Sydney, NSW, Australia, 2005), 1–7. doi:10.1109/LCN.2005.136.

46. Huang, J., Wang, J. & Jia, W. *Downlink Temporal Fairness in 802.11 WLAN Adopting the Virtual Queue Management* in *IEEE Wireless Communications and Networking Conference* (IEEE, Hong Kong, China, 2007), 3035–3040. doi:10.1109/WCNC.2007.562.
47. Sun, J., Zukerman, M. & Palaniswami, M. *Stabilizing RED using a Fuzzy Controller* in *IEEE International Conference on Communications* (IEEE, Glasgow, UK, 2007), 266–271. doi:10.1109/ICC.2007.52.
48. Linck, S., Dedu, E. & Spies, F. *Distance-Dependent RED Policy (DDRED)* in *Sixth International Conference on Networking (ICN'07)* (IEEE, Sainte Luce, Martinique, France, 2007), 51–51. doi:10.1109/ICN.2007.37.
49. Chen, S., Zhou, Z. & Bensaou, B. *Stochastic RED and Its Applications* in *IEEE International Conference on Communications* (IEEE, Glasgow, UK, 2007), 6362–6367. doi:10.1109/ICC.2007.1053.
50. Feng, W.-C., Kandlur, D. D., Saha, D. & Shin, K. G. *Techniques for Eliminating Packet Loss in Congested TCP/IP Networks* tech. rep. (The University of Michigan, 1997).
51. Marquez, R., González, I., Carrero, N. & Sulbarán, Y. *Revisiting Adaptive RED: Beyond AIMD Algorithms* in *Network Control and Optimization* (eds Chahed, T. & Tuffin, B.) **4465** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007), 74–83. doi:10.1007/978-3-540-72709-5_8.
52. Sun, L. & Wang, L. *A Novel RED Scheme with Preferential Dynamic Threshold Deployment* in *International Conference on Computational Intelligence and Security Workshops (CISW 2007)* (IEEE, Harbin, China, 2007), 854–857. doi:10.1109/CISW.2007.4425629.
53. Hariri, B. & Sadati, N. NN-RED: an AQM mechanism based on neural networks. *Electronics Letters* **43,** 1053–1055. doi:10.1049/el:20071791 (19 2007).
54. Li, D., Theunis, J., Sleurs, K., Potemans, J., van Lil, E. & van de Capelle, A. *Improving RED Performance during Handovers in Wireless IP Networks* in *4th International Symposium on Wireless Communication Systems* (IEEE, Trondheim, Norway, 2007), 441–445. doi:10.1109/ISWCS.2007.4392378.
55. Clark, D. & Fang, W. Explicit allocation of best-effort packet delivery service. *IEEE/ACM Transactions on Networking* **6,** 362–373. doi:10.1109/90.720870 (1998).
56. Jahanshahi, M. & Meybodi, M. R. *An Adaptive Congestion Control Method for Guaranteeing Queuing Delay in RED-Based Queue Using Learning Automata* in *International Conference on Mechatronics and Automation* (IEEE, Harbin, China, 2007), 3360–3365. doi:10.1109/ICMA.2007.4304102.
57. Wei, T. & Hu, G. *Fuzzy Self-Tuning PD Control Algorithm for RED* in *Spring Congress on Engineering and Technology* (IEEE, 2012). doi:10.1109/scet.2012.6341893.
58. Abdel-jaber, H., Thabtah, F. & Woodward, M. *Traffic management for the gentle random early detection using discrete-time queueing* in *International Business Information Management Association (9th IBIMA) Conference. The Conference Proceedings* (IBIMA, Marrakech, Morocco, 2008), 289–298.
59. Abdel-jaber, H. Performance study of Active Queue Management methods: Adaptive GRED, REDD, and GRED-Linear analytical model. *Journal of King Saud University - Computer and Information Sciences* **27,** 416–429. doi:10.1016/j.jksuci.2015.01.003 (2015).
60. Que, D., Chen, Z. & Chen, B. *An improvement algorithm based on RED and its performance analysis* in *9th International Conference on Signal Processing* (IEEE, Beijing, China, 2008), 2005–2008. doi:10.1109/ICOSP.2008.4697538.
61. Abdel-jaber, H., Mahafzah, M., Thabtah, F. & Woodward, M. *Fuzzy logic controller of Random Early Detection based on average queue length and packet loss rate* in *International Symposium on Performance Evaluation of Computer and Telecommunication Systems* (IEEE, Edinburgh, UK, 2008), 428–432.
62. Feng, G., Agarwal, A., Jayaraman, A. & Siew, C. K. Modified RED gateways under bursty traffic. *IEEE Communications Letters* **8,** 323–325. doi:10.1109/LCOMM.2004.827427 (2004).
63. Su, G.-Y. & Ho, C. C. *Random Early Detection Improved by Progressive Adjustment Method* in *6th National Conference on Telecommunication Technologies and 2008 2nd Malaysia Conference on Photonics* (IEEE, Edinburgh, UK, 2008), 250–253. doi:10.1109/NCTT.2008.4814282.
64. Lochin, E. & Talavera, B. *Managing Network Congestion with a Kohonen-Based RED Queue* in *IEEE International Conference on Communications* (IEEE, Beijing, China, 2008), 5586–5590. doi:10.1109/ICC.2008.1047.
65. Lu, L., Xiao, Y., Woo, S. & Kim, K. *Nonlinear AQM for Multiple RED Routers* in *Third International Conference on Convergence and Hybrid Information Technology* **2** (IEEE, Busan, Korea (South), 2008), 122–127. doi:10.1109/ICCIT.2008.68.

66. Ho, H.-J. & Lin, W.-M. *AURED — Autonomous Random Early Detection for TCP Congestion Control* in *Third International Conference on Systems and Networks Communications* (IEEE, Sliema, Malta, 2008), 79–84. doi:10.1109/ICSNC.2008.22.
67. Patro, R. K. & Bhatnagar, S. A probabilistic constrained nonlinear optimization framework to optimize RED parameters. *Performance Evaluation* **66,** 81–104. doi:10.1016/j.peva.2008.09.003 (2009).
68. Bhatnagar, S. & Patro, R. K. A proof of convergence of the B-RED and P-RED algorithms for random early detection. *IEEE Communications Letters* **13,** 809–811. doi:10.1109/LCOMM.2009.091276 (2009).
69. Abbasov, B. & Korukoglu, S. Effective RED: An algorithm to improve RED's performance by reducing packet loss rate. *Journal of Network and Computer Applications* **32,** 703–709. doi:10.1016/j.jnca.2008.07.001 (2009).
70. Almomani, O., Saaidah, A., Balas, F. A. & Al-Qaisi, L. *Simulation Based Performance Evaluation of Several Active Queue Management Algorithms for Computer Network* in *10th International Conference on Information and Communication Systems (ICICS)* (IEEE, Irbid, Jordan, 2019), 7–12. doi:10.1109/IACS.2019.8809117.
71. Abbasov, B. & Korukoğlu, S. An active queue management algorithm for reducing packet loss rate. *Mathematical and Computational Applications* **14,** 65–72. doi:10.3390/mca14010065 (2009).
72. Kong, C., Guo, Z., Ping, L. & Peng, X. PSRED: a queue management algorithm with priority self-adaptive Random Early Detection for Ad Hoc network. *Journal of Jinggangshan University,* MAG: 2118414706 (2009).
73. Kong, C., Guo, Z., Ping, L. & Peng, X. *PSRED: a queue management algorithm with priority self-adaptive Random Early Detection for Ad Hoc network* in *Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009)* (Academy Publisher, Qingdao, China, 2009), 557–560.
74. Chen, W. & Yang, S.-H. The mechanism of adapting RED parameters to TCP traffic. *Computer Communications* **32,** 1525–1530. doi:10.1016/j.comcom.2009.05.008 (2009).
75. Xiong, N., Pan, Y., Jia, X., Park, J. H. & Li, Y. Design and analysis of a self-tuning feedback controller for the Internet. *Computer Networks* **53,** 1784–1797. doi:10.1016/j.comnet.2009.02.005 (2009).
76. Qadeer, M. A., Sharma, V., Agarwal, A. & Husain, S. S. *Differentiated services with multiple random early detection algorithm using ns2 simulator* in *2nd IEEE International Conference on Computer Science and Information Technology* (IEEE, Beijing, China, 2009), 144–8. doi:10.1109/ICCSIT.2009.5234732.
77. Hanlin, S., Yuehui, J., Yidong, C., Hongbo, W. & Shiduan, C. *Improving fairness of RED aided by lightweight flow information* in *2nd IEEE International Conference on Broadband Network & Multimedia Technology* (IEEE, Beijing, China, 2009), 335–339. doi:10.1109/ICBNMT.2009.5348505.
78. Abbasov, B. *AHRED: A robust AQM algorithm for wireless ad hoc networks* in *International Conference on Application of Information and Communication Technologies* (IEEE, Baku, Azerbaijan, 2009), 1–4. doi:10.1109/ICAICT.2009.5372588.
79. Misra, S., Oommen, B. J., Yanamandra, S. & Obaidat, M. S. Random Early Detection for Congestion Avoidance in Wired Networks: A Discretized Pursuit Learning-Automata-Like Solution. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **40,** 66–76. doi:10.1109/TSMCB.2009.2032363 (2010).
80. Kumakech, M. & Bulega, T. *Modeling A Threshold Flow-Oriented Traffic Routing Technique In Optical Networks* tech. rep. (Department of Networks, College of Computing and Information Sciences, Makerere University Kampala, Uganda, 2021).
81. Loukas, R., Kohler, S., Andreas, P. & Phuoc, T.-G. *Fuzzy RED: congestion control for TCP/IP Diff-Serv* in *10th Mediterranean Electrotechnical Conference. Information Technology and Electrotechnology for the Mediterranean Countries. Proceedings. MeleCon 2000 (Cat. No.00CH37099)* **1** (IEEE, Lemesos, Cyprus, 2000), 19–22. doi:10.1109/MELCON.2000.880358.
82. Rossides, L., Sekercioglu, A., Pitsillides, A., Vasilakos, A., Kohler, S. & Tran-Gia, P. *Fuzzy RED: Congestion Control for TCP/IP Diff-Serv* in *Advances in Computational Intelligence and Learning: Methods and Applications. International Series in Intelligent Technologies* (Springer Netherlands, Dordrecht, 2002). doi:10.1007/978-94-010-0324-7_24.

83. Masoumzadeh, S. S., Meshgi, K., Ghidari, S. S. & Taghizadeh, G. FQL-RED: an adaptive scalable schema for active queue management. *International Journal of Network Management* **21,** 147–167. doi:10.1002/nem.755 (2010).
84. Xiong, N., Vasilakos, A. V., Yang, L. T., Wang, C.-X., Kannan, R., Chang, C.-C. & Pan, Y. A novel self-tuning feedback controller for active queue management supporting TCP flows. *Information Sciences* **180,** 2249–2263. doi:10.1016/j.ins.2009.12.001 (2010).
85. Seddigh, N., Nandy, B., Pieda, P. S., Salim, J. H. & Chapman, A. *Experimental study of assured services in a diffserv IP QoS network* in *Internet Routing and Quality of Service* (eds Onvural, R. O., Civanlar, S., Doolan, P. J., Luciani, J. V., Civanlar, S., Doolan, P. J. & Luciani, J. V.) **3529** (SPIE, 1998), 217–230. doi:10.1117/12.333712.
86. Cui, Y.-b., Cui, Y., Hou, Z.-g., Kang, Y.-n. & Fu, Q. *An Active Queue Management algorithm based on DiffServ model* in *Second International Workshop on Education Technology and Computer Science* **1** (IEEE, Wuhan, China, 2010), 455–458. doi:10.1109/ETCS.2010.447.
87. Liu, F., Guan, Z.-H. & Wang, H. O. Controlling bifurcations and chaos in TCP–UDP–RED. *Nonlinear Analysis: Real World Applications* **11,** 1491–1501. doi:10.1016/j.nonrwa.2009.03.005 (2010).
88. Zhang, C., Yin, J., Cai, Z. & Chen, W. RRED: robust RED algorithm to counter low-rate denial-of-service attacks. *IEEE Communications Letters* **14,** 489–491. doi:10.1109/LCOMM.2010.05.091407 (2010).
89. Minagawa, T. & Ikegami, T. *Controlling user flows with RIO and WFQ* in *10th International Symposium on Communications and Information Technologies* (IEEE, Tokyo, Japan, 2010), 87–92. doi:10.1109/ISCIT.2010.5664907.
90. Lai, W.-P. & Liu, Z.-H. *Fractional Exponent Coupling of RIO* in *13th IEEE International Conference on Computational Science and Engineering* (IEEE, Hong Kong, China, 2010), 208–213. doi:10.1109/CSE.2010.35.
91. Jianyong, C., Cunying, H. & Zhen, J. Self-tuning Random Early Detection algorithm to improve performance of network transmission. *Mathematical Problems in Engineering* **2011,** Article ID 872347. doi:10.1155/2011/872347 (2011).
92. Xiaoping, Y., Hong, C. & Zhenyu, Z. *A Queue Management Algorithm for Differentiated Services* in *Fourth International Conference on Intelligent Computation Technology and Automation* **2** (IEEE, Shenzhen, China, 2011), 941–944. doi:10.1109/ICICTA.2011.521.
93. Zhang, S., Sa, J., Liu, J. & Wu, S. *An improved RED algorithm with sinusoidal packet-marking probability and dynamic weight* in *International Conference on Electric Information and Control Engineering* (IEEE, Wuhan, China, 2011), 1160–1163. doi:10.1109/ICEICE.2011.5777269.
94. Zhang, J., Xu, W. & Wang, L. An improved adaptive Active Queue Management algorithm based on nonlinear smoothing. *Procedia Engineering* **15.** CEIS 2011, 2369–2373. doi:10.1016/j.proeng.2011.08.444 (2011).
95. Fan, X., Wang, J., Guan, L., Gao, L. & Wang, X. *Heuristic Active Queue Management with Hazard rate function* in *Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* **4** (IEEE, Shanghai, China, 2011), 2281–2285. doi:10.1109/FSKD.2011.6019958.
96. Zhang, J.-C., Zhao, R.-X. & Chen, J.-J. *A Hop to Hop controlled hierarchical multicast congestion control mechanism* in *Electronics and Signal Processing* (ed Hu, W.) **97** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), 363–369. doi:10.1007/978-3-642-21697-8_46.
97. Hamadneh, N., Murray, D., Dixon, M. & Cole, P. *Weighted RED (WTRED) strategy for TCP congestion control* in *Informatics Engineering and Information Science* (eds Abd Manaf, A., Zeki, A., Zamani, M., Chuprat, S. & El-Qawasmeh, E.) **252** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), 421–434. doi:10.1007/978-3-642-25453-6_37.
98. Abdel-jaber, H., Ababneh, J., Thabtah, F., Daoud, A. M. & Baklizi, M. *Performance analysis of the proposed Adaptive Gentle Random Early Detection method under noncongestion and congestion situations* in *Digital Enterprise and Information Systems* (eds Ariwa, E. & El-Qawasmeh, E.) **194** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), 592–603. doi:10.1007/978-3-642-22603-8_52.
99. Baklizi, M., Abdel-Jaber, H., Ramadass, S., Abdullah, L. & Anbar, M. Performance assessment of AGRED, RED and GRED congestion control algorithms. *Information Technology Journal* **11,** 255–261. doi:10.3923/itj.2012.255.261 (Feb. 2012).
100. Wang, H., Ye, Z. & Wang, B. *Using auto-tuning proportional integral probability to improve random early detection* in *IEEE 13th International Conference on Communication Technology* (IEEE, Jinan, China, 2011), 1107–1111. doi:10.1109/ICCT.2011.6158054.

101. Zhang, Y., Ma, J., Wang, Y. & Xu, C. *MRED: an improved nonlinear RED algorithm* in *International Conference Proceedings on Computer and Automation Engineering (ICCAE 2011)* **44** (2011), 6–11.
102. Lin, D. & Morris, R. Dynamics of Random Early Detection. *SIGCOMM Comput. Commun. Rev.* **27,** 127–137. doi:10.1145/263109.263154 (1997).
103. Nossenson, R. & Maryuma, H. *Active Queue Management in blind access networks* in *The Third International Conference on Access Networks (ACCESS 2012)* (IARIA, Venice, Italy, 2012).
104. Ramachandra, G. A., Banu, R. & Ali Ahammed, G. F. *Analyzing marking Mod RED Active Queue Management scheme on TCP applications* in *International Conference on Information and Network Technology (ICINT 2012)* (IACSIT Press, Singapore, 2012), 251–257.
105. Järvinen, I., Ding, Y., Nyrhinen, A. & Kojo, M. *Harsh RED: Improving RED for Limited Aggregate Traffic* in *IEEE 26th International Conference on Advanced Information Networking and Applications* (IEEE, Fukuoka, Japan, 2012), 832–840. doi:10.1109/AINA.2012.103.
106. Järvinen, I. & Kojo, M. *Evaluating CoDel, PIE, and HRED AQM techniques with load transients* in *39th Annual IEEE Conference on Local Computer Networks* (IEEE, Edmonton, AB, Canada, 2014), 159–167. doi:10.1109/LCN.2014.6925768.
107. Yu, C.-h. & Lin, C.-l. *A novel algorithm to achieve bandwidth fairness of RED with packet size consideration* in *7th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (IEEE, Singapore, 2012), 659–662. doi:10.1109/ICIEA.2012.6360808.
108. Tahiliani, M. P., Shet, K. C. & Basavaraju, T. G. *FARED: Fast Adapting RED gateways for TCP/IP networks* in *Advanced Computing, Networking and Security* (eds Thilagam, P. S., Pais, A. R., Chandrasekaran, K. & Balakrishnan, N.) **7135** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), 435–443. doi:10.1007/978-3-642-29280-4_51.
109. Tahiliani, M. P., Shet, K. & Basavaraju, T. G. CARED: Cautious Adaptive RED gateways for TCP/IP networks. *Journal of Network and Computer Applications* **35.** Simulation and Testbeds, 857–864. doi:h10.1016/j.jnca.2011.12.003 (2012).
110. Tahiliani, M. P. & Shet, K. C. *Analysis of cautious adaptive RED (CARED)* in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (IEEE, Mysore, India, 2013), 1029–1034. doi:10.1109/ICACCI.2013.6637318.
111. Jamali, S., Alipasandi, B. & Alipasandi, N. VRED: an improvement over RED algorithm by using queue length growth velocity. *Journal of Advances in Computer Research* **4,** 31–38 (2013).
112. Jamali, S., Seyyed Hashemi, S. N. & Eftekhari Moghadam, A. M. On the use of a full information feedback to stabilize RED. *Journal of Network and Computer Applications* **36,** 858–869. doi:h10.1016/j.jnca.2012.11.002 (2013).
113. Jiang, X., Yang, J., Jin, G. & Wei, W. RED-FT: a scalable Random Early Detection scheme with flow trust against DoS attacks. *IEEE Communications Letters* **17,** 1032–1035. doi:10.1109/LCOMM.2013.022713.122652 (2013).
114. Patel, C. M. *URED: upper threshold RED an efficient congestion control algorithm* in *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (IEEE, Tiruchengode, India, 2013), 1–5. doi:10.1109/ICCCNT.2013.6726469.
115. Baklizi, M., Abdel-Jaber, H., Abualhaj, M., Abdullah, N., Ramadass, S. & Almomani, D. Dynamic stochastic early discovery: A new congestion control technique to improve networks performance. *International Journal of Innovative Computing, Information and Control (IJICIC)* **9,** 1113–1126 (2013).
116. Almomani, O., Saaidah, A., Balas, F. A. & Al-Qaisi, L. *Simulation based performance evaluation of several Active Queue Management algorithms for computer network* in *10th International Conference on Information and Communication Systems (ICICS)* (IEEE, Irbid, Jordan, 2019), 7–12. doi:10.1109/IACS.2019.8809117.
117. Peng, I.-H., Lin, M.-H., Chen, Y.-W., Yang, F.-M. & Su, A. Y. S. *Improvement of streaming video in Differential Service networks by using opportunity RED mechanism* in *Seventh International Conference on Complex, Intelligent, and Software Intensive Systems* (IEEE, Taichung, Taiwan, 2013), 644–648. doi:10.1109/CISIS.2013.116.
118. Feng, W., Ye, J. & Chend, Y. Improved packet discarded probability calculation method — PRED. chinese. *Computer Engineering and Applications* **49,** 92–96 (2013).
119. Manasa, S. ANLRED: a robust AQM mechanism for congestion avoidance. *International Journal of Computer Applications* **81,** 1–9. doi:10.5120/14196-2259 (2013).
120. Soni, H. & Mishra, P. Reducing packet loss in Active Queue Management. *International Journal of Computer Applications* **81,** 25–28. doi:10.5120/14208-2447 (Nov. 2013).

121. Baklizi, M., Abdel-Jaber, H., Adel, A., Abualhaj, M. & Ramadass, S. Fuzzy Logic Controller of Gentle Random Early Detection based on average queue length and delay rate. *International Journal of Fuzzy Systems* **16,** 9–19 (2014).
122. Anjum, F. & Tassiulas, L. *Fair bandwidth sharing among adaptive and non-adaptive flows in the Internet* in *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)* **3** (IEEE, New York, NY, USA, 1999), 1412–1420. doi:10.1109/INFCOM.1999.752161.
123. Anjum, F. & Tassiulas, L. *Balanced RED: an algorithm to achieve fairness in the Internet* tech. rep. (The Center for Satellite and Hybrid Communication Networks, 1999). doi:https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=439ed764c7cc8cd5af53201bda4e68b0facdb5c4.
124. Diva, M. A. & Teshnehleb, M. *Dynamic queue management using neural network based on balanced RED* in *Iranian Conference on Intelligent Systems (ICIS)* (IEEE, Bam, Iran, 2014), 1–4. doi:10.1109/IranianCIS.2014.6802597.
125. Sharma, R. & Dixit, G. Experimental study of RED performance by regulating Upper Threshold parameter. *International Journal of Computer Science and Information Technologies* **5,** 6202–6204 (2014).
126. Mahajan, M. & Singh, T. P. The Modified Gaussian function based RED (MGF-RED) algorithm for congestion avoidance in mobile Ad hoc networks. *International Journal of Computer Applications* **91,** 39–44. doi:10.5120/15889-5112 (2014).
127. Domańska, J., Domański, A., Augustyn, D. & Klamka, J. A RED modified weighted moving average for soft real-time application. *International Journal of Applied Mathematics and Computer Science* **24,** 697–707. doi:10.2478/amcs-2014-0051 (2014).
128. Jamali, S., Alipasandi, N. & Alipasandi, B. An improvement over Random Early Detection algorithm: A self-tuning approach. *Journal of Electrical and Computer Engineering Innovations (JECEI)* **2,** 57–61. doi:10.22061/jecei.2014.242 (2014).
129. Patel, S. *Performance analysis of RED for stabilized queue* in *Seventh International Conference on Contemporary Computing (IC3)* (IEEE, New York, NY, USA, 2014), 306–311. doi:10.1109/IC3.2014.6897191.
130. Abdeljaber, H., Thabtah, F., Woodward, M., Jaffar, A. & Al Bazaar, H. Random early dynamic detection approach for congestion control. *Baltic Journal of Modern Computing* **2,** 16–31 (2014).
131. Hamdy, A., El-Sayed, A., Elsaghir, Z. & Morsi, I. Enhanced Random Early Detection (ENRED). *International Journal of Computer Applications* **92,** 25–28. doi:10.5120/16039-5015 (2014).
132. Oluwatope, A. & Hassan, S. *Curvilinear RED: An improved RED algorithm for internet routers* in *Proceedings of the IASTED International Conference on Modelling and Simulation, AfricaMS 2014* (Acta Press, Gaborone, Botswana, 2014). doi:10.2316/P.2014.813-025.
133. Xu, J., Song, J., He, Q. & Xu, J. *A novel queue management algorithm INRIO-C based on Differentiated Services* in *IEEE 17th International Conference on Computational Science and Engineering* (IEEE, Chengdu, China, 2014), 541–546. doi:10.1109/CSE.2014.124.
134. John, J. K. & Siva Balan, R. V. *An enhanced router based queuing policy providing better fairness to real-time data transfers in Internet* in *International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]* (IEEE, Nagercoil, India, 2015), 1–6. doi:10.1109/ICCPCT.2015.7159172.
135. Mahajan, S. G. *Efficient LALRED for congestion avoidance using Automata-like solution* in *International Conference on Emerging Information Technology and Engineering Solutions* (IEEE, Mahashtra, India, 2015), 1–6. doi:10.1109/EITES.2015.11.
136. Mukund, Y. R., Rohit, C. & Chandavarkar, B. R. *Improving RED for reduced UDP packet-drop* in *International Conference on Communications and Signal Processing (ICCSP)* (IEEE, Melmaruvathur, India, 2015), 0026–0030. doi:10.1109/ICCSP.2015.7322887.
137. Kim, W.-J. & Byeong, G. FRED — fair random early detection algorithm for TCP over ATM networks. *Electronics Letters* **34,** 152–154. doi:10.1049/el:19980049 (1998).
138. Abdulkareem, M., Akil, K., Kalakech, A. & Kadry, S. EFRED: Enhancement of Fair Random Early Detection Algorithm. *International Journal of Communications, Network and System Sciences* **8,** 282–294. doi:10.4236/ijcns.2015.87028 (2015).
139. Zhang, D.-G., Ma, Z., Zhao, D.-X., Song, J.-J. & Liu, S. *Novel adaptive queue intelligent management algorithm* in *IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems* (IEEE, Dallas, TX, USA, 2015), 483–485. doi:10.1109/MASS.2015.98.

140. Ravindra, S. R. & Patil, A. *Fuzzy Cautious Adaptive Random Early Detection for Heterogeneous network* in *2015 Fifth International Conference on Advances in Computing and Communications (ICACC)* (IEEE, Kochi, India, 2015), 65–68. doi:10.1109/ICACC.2015.106.
141. Korolkova, A. V., Kulyabov, D. S. & Tchernoivanov, A. I. On the classification of RED algorithms. Russian. *RUDN Journal of Mathematics, Information Sciences and Physics* **3,** 34–46 (2009).
142. Korolkova, A. V. & Zaryadov, I. S. *The mathematical model of the traffic transfer process with a rate adjustable by RED* in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* (IEEE, Moscow, Russia, Oct. 2010), 1046–1050. doi:10.1109/ICUMT.2010.5676505.
143. Velieva, T. R., Korolkova, A. V. & Kulyabov, D. S. *Designing installations for verification of the model of active queue management discipline RED in the GNS3* in *The 6th International Congress on Ultra Modern Telecommunications and Control Systems. Saint-Petersburg, Russia. October 6-8, 2014* (IEEE Computer Society, 2015), 570–577. doi:10.1109/ICUMT.2014.7002164.
144. Korolkova, A. V., Kulyabov, D. S. & Sevastianov, L. A. Combinatorial and operator approaches to RED modeling. *Mathematical Modelling and Geometry* **3,** 1–18. doi:10.26456/mmg/2015-331 (2015).
145. Korolkova, A. V. & Zaryadov, I. S. *The mathematical model of the traffic transfer process with a rate adjustable by RED* in *International Congress on Ultra Modern Telecommunications and Control Systems* (IEEE, Moscow, Russia, 2010), 1046–1050. doi:10.1109/ICUMT.2010.5676505.
146. Zaryadov, I. S., Korolkova, A. V., Kulyabov, D. S., Milovanova, T. & Tsurlukov, V. *The survey on Markov-Modulated Arrival Processes and their application to the analysis of active queue management algorithms* in *Distributed Computer and Communication Networks. DCCN 2017. Communications in Computer and Information Science* (eds Vishnevskiy, V. M., Samouylov, K. E. & Kozyrev, D. V.) 417–430 (Springer International Publishing, Cham, 2017). doi:10.1007/978-3-319-66836-9_35.
147. Viana C. C., H., Zaryadov, I. S., Tsurlukov, V. V., Milovanova, T. A., Bogdanova, E. V., Korolkova, A. V. & Kulyabov, D. S. *The general renovation as the active queue management mechanism. Some aspects and results* in *Distributed Computer and Communication Networks. DCCN 2019* (eds Vishnevskiy, V., Samouylov, K. & Kozyrev, D.) 488–502 (Springer, Cham, 2019). doi:10.1007/978-3-030-36625-4_39.
148. Apreutesey, A. M. Y., Korolkova, A. V. & Kulyabov, D. S. *Modeling RED algorithm modifications in the OpenModelica* in *Proceedings of the Selected Papers of the 8th International Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems" (ITTMM-2019), Moscow, Russia, April 15–19, 2019* (eds Kulyabov, D. S., Samouylov, K. E. & Sevastianov, L. A.) **2407** (CEUR-WS, 2019), 5–14.
149. Viana Carvalho Cravid, H., Zaryadov, I. S. & Milovanova, T. A. Queueing systems with different types of renovation mechanism and thresholds as the mathematical models of active queue management mechanism. *Discrete and Continuous Models and Applied Computational Science* **28,** 305–318. doi:10.22363/2658-4670-2020-28-4-305-318 (2020).
150. Adams, R. Active queue management: a survey. *Communications Surveys & Tutorials, IEEE* **15,** 1425–1476. doi:10.1109/SURV.2012.082212.00018 (2013).
151. Abbas, G., Halim, Z. & Abbas, Z. H. Fairness-driven queue management: a survey and taxonomy. *IEEE Communications Surveys & Tutorials* **18,** 324–367. doi:10.1109/COMST.2015.2463121 (2016).

To cite: Zaryadov I. S., Viana H. C.C., Korolkova A. V., Milovanova T. A., Chronology of the development of Active Queue Management algorithms of RED family. Part 2: from 2006 up to 2015, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)18–37.DOI: 10.22363/2658-4670-2024-32-1-18-37.

Information about the authors

Zaryadov, Ivan S.—Candidate of Physical and Mathematical Sciences, Assistant Professor of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: zaryadov-is@rudn.ru, ORCID: https://orcid.org/0000-0002-7909-6396)

Viana, C.C. Hilquias—Ph.D. student of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: hilvianamat1@gmail.com, ORCID: https://orcid.org/0000-0002-1928-7641)

Korolkova, Anna V.—Candidate of Physical and Mathematical Sciences, Associate Professor of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: korolkova-av@rudn.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0001-7141-7610)

Milovanova, Tatianna A.—Candidate of Physical and Mathematical Sciences, Assistant Professor of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: milovanova-ta@rudn.ru, ORCID: https://orcid.org/0000-0002-9388-9499)

УДК 519.872
DOI: 10.22363/2658-4670-2024-32-1-18-37 EDN: CAMXUN

Хронология развития алгоритмов активного управления очередями семейства RED. Часть 2: 2006–2015

И. С. Зарядов[1,2], К. К. И. Виана[1], А. В. Королькова[1], Т. А. Милованова[1]

[1] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация
[2] *Федеральный исследовательский центр «Информатика и управление» РАН,*
ул. Вавилова, д. 44, корп. 2, Москва, 119333, Российская Федерация

Аннотация. Данная работа является второй частью большого библиографического обзора по алгоритмам семейства RED, представленных в научной печати с 1993 по 2023 год. В этой статье приведены данные по алгоритмам, опубликованным с 2006 по 2015 год.

Ключевые слова: активное управление очередями, AQM, RED, управление перегрузками

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 38–47
http://journals.rudn.ru/miph

Research article

UDC 519.65:519.217
PACS 07.05.Tp, 07.05.Mh, 02.70.–c
DOI: 10.22363/2658-4670-2024-32-1-38-47 EDN: GFROYO

Sampling of integrand for integration using shallow neural network

Alexander Ayriyan[1, 2, 3], Hovik Grigorian[1, 2, 3, 4], Vladimir Papoyan[1, 2, 3]

[1] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*
[2] *Alikhanyan National Science Laboratory (YerPhI), 2 Alikhanyan Brothers St, Yerevan, 0036, Republic of Armenia*
[3] *Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation*
[4] *Yerevan State University, 1 Alex Manoogian St, Yerevan, 0025, Republic of Armenia*

(received: November 13, 2023; revised: December 15, 2023; accepted: January 12, 2024)

Abstract. In this paper, we study the effect of using the Metropolis–Hastings algorithm for sampling the integrand on the accuracy of calculating the value of the integral with the use of shallow neural network. In addition, a hybrid method for sampling the integrand is proposed, in which part of the training sample is generated by applying the Metropolis–Hastings algorithm, and the other part includes points of a uniform grid. Numerical experiments show that when integrating in high-dimensional domains, sampling of integrands both by the Metropolis–Hastings algorithm and by a hybrid method is more efficient with respect to the use of a uniform grid.

Key words and phrases: Shallow Neural Network, Numerical Integration, Metropolis–Hastings Algorithm

1. Introduction

In the recent study [1], an algorithm for numerical integration was proposed based on the use of a neural network with one hidden layer. In this approach, the neural network approximates the integrand function within a bounded region that includes the integration domain. A training the neural network may certainly require a significant amount of time and computational resources. However, when the training is completed, the neural network architecture allows for the analytical integration of the approximated integrand. Furthermore, the integral of the neural network's function can be computed in any other subregion without the need for retraining. Thus, the neural network integration approach is efficient for tasks where it is necessary to repeatedly calculate the integral of the same function in different regions.

The neural network training is the main challenge of an integrand approximation. During supervised learning training data plays a significant role, and consequently an approach to their sampling. In the paper [1], a uniform grid-based discretization of the domain was used as the function sampling method. However, this approach is inefficient for integrands with significant variations in certain subregions.

In this article, the impact of using the Metropolis–Hastings algorithm for shape-based sampling of an integrand on the integration accuracy is considered. A hybrid approach for forming the training dataset is proposed, in which a portion of the training dataset (with a relative volume fraction denoted as ρ) is generated using the Metropolis–Hastings algorithm, while the other part includes nodes of a uniform grid.

© Ayriyan A., Grigorian H., Papoyan V., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

2. The neural network method of approximate integration

In many areas of science and engineering there is a need for approximate calculation of an integral for a given continuous real function $f : \mathbb{R}^n \to \mathbb{R}$ over a region S

$$I[f] = \int_S f(\mathbf{x}) d\mathbf{x}. \tag{1}$$

According to the universal approximation theorem [2] and Theorem 2 in [1], any function $f(\mathbf{x})$ as defined above can be approximated arbitrarily accurately using a shallow (single hidden layer) neural network $\widehat{f}(\mathbf{x})$ with a logistic sigmoid activation function (3). This network can be analytically integrated within a bounded convex region S.

The mathematical expression for such a neural network can be represented by the formula:

$$\widehat{f}(\mathbf{x}) = b^{(2)} + \mathbf{W}_2^T \sigma(\mathbf{b}^{(1)} + \mathbf{W}_1 \mathbf{x}) = b^{(2)} + \sum_{j=1}^{k} w_j^{(2)} \sigma\left(b_j^{(1)} + \sum_{i=1}^{n} w_{ij}^{(1)} x_i\right), \tag{2}$$

where:

- $\mathbf{W}_1$ and $\mathbf{W}_2$ are weight matrices for the first and second layers of the neural network, respectively;
- $\mathbf{b}^{(1)}$ and $b^{(2)}$ are the bias vectors for the first and second layers of the neural network, respectively;
- $\mathbf{x}$ is an n-dimensional vector of input values for the function f;
- σ is the logistic sigmoid function:

$$\sigma(z) = \frac{1}{1 + e^{-z}}. \tag{3}$$

The logistic sigmoid function and its antiderivatives are expressed in terms of polylogarithmic functions Li_m of different orders m: Li_0:

$$\mathrm{Li}_0(z) = -\frac{1}{1 - z^{-1}}. \tag{4}$$

In particular, the sigmoid function itself is expressed in terms of the zeroth-order polylogarithm

$$\sigma(z) = \frac{1}{1 + e^{-z}} = -\mathrm{Li}_0(-e^z). \tag{5}$$

Note that the representation of each subsequent antiderivative $\sigma(z)$ increases the order of the polylogarithm by one. This representation of the sigmoid function allows us to integrate $\widehat{f}(\mathbf{x})$ in accordance with Theorem 2 from [1].

The integration region S can be extended to $\tilde{S}$ such that $S \subseteq \tilde{S}$ and $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \tilde{S} \setminus S$. $\tilde{S}$ is an n-dimensional hyperrectangle in $\mathbb{R}^n$, specifically $\tilde{S} = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times \cdots \times [\alpha_n, \beta_n]$.

Thus, expression for the neural network integral $\hat{I}(f, \boldsymbol{\alpha}, \boldsymbol{\beta})$ is defined as:

$$\hat{I}(f, \boldsymbol{\alpha}, \boldsymbol{\beta}) = I[\widehat{f}] = b_2 \prod_{i=1}^{n} (\beta_i - \alpha_i) + \sum_{j=1}^{k} w_j^{(2)} \left[\prod_{i=1}^{n} (\beta_i - \alpha_i) + \frac{\Phi_j}{\prod_{i=1}^{n} w_{ij}^{(1)}} \right], \tag{6}$$

where Φ_j is defined as:

$$\Phi_j = \sum_{r=1}^{2^n} \xi_r \mathrm{Li}_n \left(-\exp\left[-b_j^{(1)} - \sum_{i=1}^{n} w_{ij}^{(1)} \ell_{i,r} \right] \right). \tag{7}$$

Here, ξ_r is the sign in front of the r-th term of the sigmoid integration, and $\ell_{i,r}$ represents the corresponding integration limit for the i-th dimension. These limits are defined by:

$$\xi_r = \prod_{d=1}^{n} (-1)^{\lfloor r/2^{n-d} \rfloor}, \tag{8}$$

$$\ell_{i,r} = \begin{cases} \alpha_i, & \text{if } \lfloor r/2^{n-i} \rfloor \text{ is even,} \\ \beta_i, & \text{otherwise.} \end{cases} \tag{9}$$

It is worth noting that an alternative to polylogarithmic functions can be the Fermi–Dirac integral:

$$F_n\mathbf{x} = \frac{1}{\Gamma(n+1)} \int_0^\infty \frac{t^n}{e^{t-x}+1} dt, \tag{10}$$

which is related to the polylogarithm as:

$$F_n\mathbf{x} = -\mathrm{Li}_{n+1}(-e^x). \tag{11}$$

Then the function (7) takes on a new form:

$$\Phi_j = \sum_{r=1}^{2^n} -\xi_r F_{n-1}\left(-b_j^{(1)} - \sum_{i=1}^{n} w_{ij}^{(1)} \ell_{i,r}\right). \tag{12}$$

This alternative representation can be useful when the weight coefficients $w_{ij}^{(1)}$ and biases $b_j^{(1)}$ acquire large values due to training, causing potential data type overflow issues. However, the effectiveness of this depends on the specific implementation of the Fermi–Dirac integral calculation.

3. Sampling of integrand

The main difficulty of the considered integration approach lies in the neural network training that approximates the integrand. In other words, the main problem is to calculate the weight matrices $\mathbf{W}_1$, $\mathbf{W}_2$, and biases $\mathbf{b}^{(1)}$, $b^{(2)}$ to achieve the minimum value of the objective function. A successful solution to a supervised learning problem depends on several factors, the most important one is the generation of a training dataset.

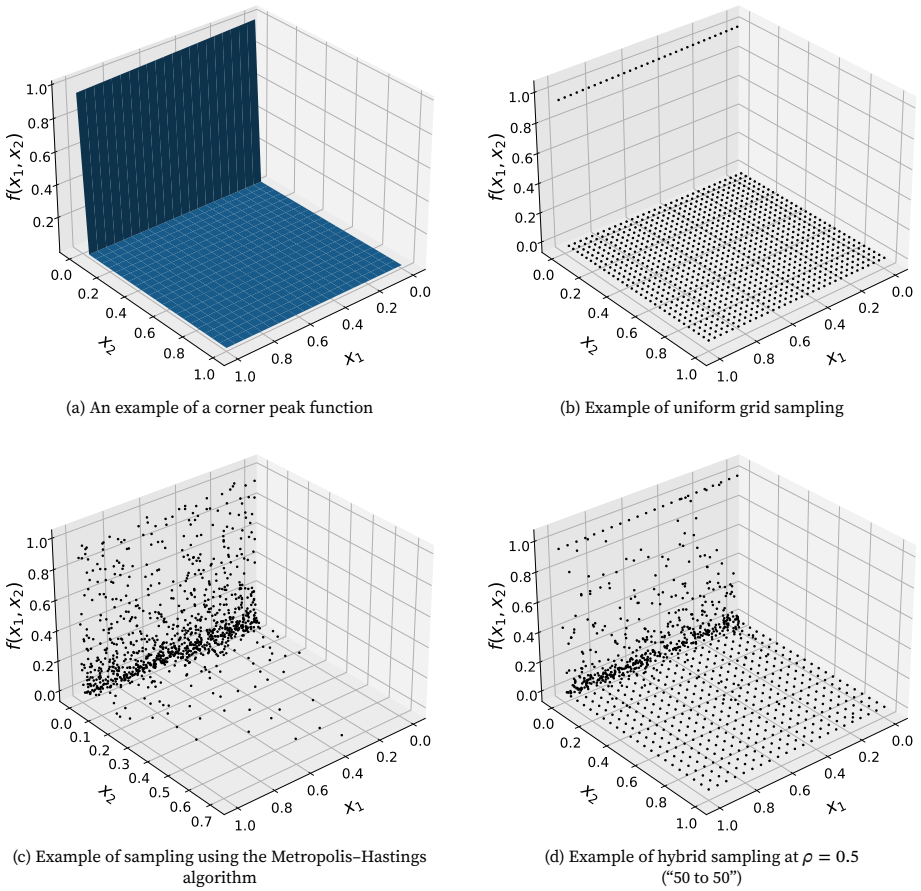
The training set, denoted as $D = \{(\mathbf{x}, f(\mathbf{x})) \mid \mathbf{x} \in S_N \subset \tilde{S}\}$ (where S_N is an N-element finite subset of $\tilde{S}$), includes the argument vector $\mathbf{x}$ and the corresponding function values $f(\mathbf{x})$. In other words, the training set results from sampling the integrand.

In [1], a uniform grid of nodes is used as the training dataset, which leads to insufficiently accurate approximation for families of integrands with sharp value changes in certain subregions, hence resulting in unsatisfactory integration results. In particular, there would be insufficient learning points in regions where the shape of the function is drastically sharp for instance for functions with explicitly pronounced peaks. This case is illustrated in the figure 1b.

In this research, a hybrid approach is proposed for forming the training dataset, where a part of the training data D_{MH} is generated using the Metropolis–Hastings algorithm [3] and [4], which allows sampling any probability distribution function. Another part of the dataset D_{UG} consists of nodes from a uniform grid. As a result both data sets united to the final training set $D = D_{MH} \cup D_{UG}$.

The Metropolis–Hastings algorithm is based on constructing a converging Markov chain, where each iteration involves generating a new random point $\mathbf{x}$ from an auxiliary distribution, followed by deciding whether to accept or reject this point, using information about the value of the integrand $f(\mathbf{x})$. Applying such an approach to functions with a narrow and high peak will increase the point density in areas where the function value increases, thereby improving both the integrand approximation and integration accuracy. Examples of point generation using the hybrid method and the Metropolis–Hastings algorithm are shown in figures 1d and 1c, respectively.

The main challenge in applying the Metropolis–Hastings algorithm lies in the absence of a universally efficient approach to determine the algorithm's parameters. Additionally, it is necessary to establish the number of points to generate using the Metropolis–Hastings algorithm relative to the total number of points for effective integrand approximation. This paper empirically investigates the impact of parameter determination and the proportion of points generated by the Metropolis–Hastings algorithm on the integration results.



(a) An example of a corner peak function

(b) Example of uniform grid sampling

(c) Example of sampling using the Metropolis–Hastings algorithm

(d) Example of hybrid sampling at $\rho = 0.5$ ("50 to 50")

Figure 1. Different ways to sample a function with a clearly defined peak with fixed parameters $c_1 = 0.0146162197$ and $c_2 = 299.985384$

4. Implementation and testing

4.1. Implementation

The implementation of the approach was carried out in the Python programming language within the ML/DL ecosystem in the computational component called `jhub2` [5] using the `Keras` [6] and `mpmath` [7] Python libraries.

The number of neurons in the hidden layer of the network was determined according to the expression:

$$k = \left\lfloor \left(\log_{10}(N)\right)^{-K_1} \frac{K_2 N}{(n+2)} \right\rfloor, \tag{13}$$

where $K_1 = 4.33$ and $K_2 = 16$ [1], N is the number of elements in the sample.

The mean squared error (MSE) was used as the objective function:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \left(f_i - \hat{f}_i\right)^2. \tag{14}$$

The neural network was trained using the Levenberg–Marquardt backpropagation algorithm [8], [9] for 5000 epochs. The Levenberg–Marquardt method is one of the fastest and has high convergence for small-sized neural networks [10]. For the neural network training 90% of the total number of points D were used, with the remaining 10% used for validation.

The inputs and outputs of the neural network were transformed into the range $[d_{\min}, d_{\max}]$ through Min-Max normalization [11]. After the training process, based on the obtained weight coefficients $\mathbf{W}_1$, $\mathbf{W}_2$, and biases $\mathbf{b}^{(1)}$, $b^{(2)}$ related to the normalized function f', the integral is calculated according to formula (6). The integration limits $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ must be scaled to $\boldsymbol{\alpha}'$ and $\boldsymbol{\beta}'$ according to the transformation of the neural network arguments. Then, according to formula (6), a scaled value of the integral $\hat{I}(f', \boldsymbol{\alpha}', \boldsymbol{\beta}')$ is obtained. To obtain the integral $\hat{I}(f, \boldsymbol{\alpha}, \boldsymbol{\beta})$, it is necessary to rescale $\hat{I}(f', \boldsymbol{\alpha}', \boldsymbol{\beta}')$ according to the expression:

$$\hat{I}(f, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{V(\tilde{S})(f_{\max} - f_{\min})}{V(S')(d_{\max} - d_{\min})} \hat{I}(f', \boldsymbol{\alpha}', \boldsymbol{\beta}') + \left(f_{\min} - \frac{f_{\max} - f_{\min}}{d_{\max} - d_{\min}} d_{\min} \right) V(\tilde{S}).$$

Here $f_{\max}$ and $f_{\min}$ are the maximum and minimum values of the function in the training set D, and $V(\tilde{S})$ and $V(S')$ are hypervolumes of the integration domain $\tilde{S}$ before and S' after the scale transformation, respectively, and since the latter are hyperrectangles, the volumes are:

$$V(\tilde{S}) = \prod_{i=1}^{n} (\beta_i - \alpha_i), \tag{15}$$

$$V(S') = \prod_{i=1}^{n} (\beta_i' - \alpha_i'). \tag{16}$$

In further calculations, $d_{\max} = 1$, and $d_{\min} = -1$.

The accuracy of integration is assessed by determining the number of correct digits (CD) in the approximate value of the integral obtained using the neural network:

$$CD(I, \hat{I}) = -\log_{10} \left| \frac{I - \hat{I}}{I} \right|. \tag{17}$$

4.2. Functions for testing

The testing of the hybrid sampling was performed on three classes of integrands, f_1, f_2, f_3, defined within the unit cube $[0, 1]^n$. All three classes of parameterized functions were taken from a set of functions for testing multidimensional integration algorithms compiled by Alan Genz [12]:

Oscillatory function:

$$f_1(x) = \cos\left(2\pi u_1 + \sum_{i=1}^{n} c_i x_i \right). \tag{18}$$

Corner Peak function:

$$f_2(x) = \left(1 + \sum_{i=1}^{n} c_i x_i \right)^{-(n+1)}. \tag{19}$$

Continuous function:

$$f_3(x) = \exp\left(-\sum_{i=1}^{n} c_i |x_i - u_i| \right). \tag{20}$$

Here, u_i are shift parameters, with values uniformly randomly distributed in the interval $[0, 1]$. The vector of parameters $\mathbf{c}$ can be used to control the complexity of integration. It is determined for each family of functions f_j separately:

$$\mathbf{c} = \left(\frac{h_j}{n^{e_j} \sum_{i=1}^{n} c_i'} \right) \mathbf{c}', \tag{21}$$

where $\mathbf{c}'$ is a vector of size n, with its components uniformly randomly distributed in the range $[0, 1]$. The values of h_j and e_j are fixed for each class of functions and are presented in the table 1.

Table 1

Values of integration complexity parameters

n	Parameters	f_1	f_2	f_3
$n = 2$	h_j	100	600	100
	e_j	1	1	1
$n = 6$	h_j	300	1000	200
	e_j	1.75	1.75	1.75

4.3. Testing of the hybrid approach

Each class of integrands (18)–(20) was investigated in two spatial dimensions, $n = 2$ and 6, with a corresponding number of points $N = 10^3$, 10^4, and 10^5. The results of the computations are illustrated in the figure 2.

Each point on the graph corresponds to the average value of 20 approximate integrals with varying parameters u and c. For $N = 10^5$, the number of computed integrals was reduced from 20 to 5 due to the lengthy neural network training.

For functions with a clear peak, the proposed method of defining the training set increases the accuracy of integration compared to training on the nodes of a uniform grid, for any number of points generated by the Metropolis–Hastings algorithm. On the other hand, a value of ρ near 1 can deteriorate the function approximation due to a lack of points near the small function values, thereby reducing integration accuracy. In the case of oscillatory and continuous functions, applying hybrid sampling does not significantly increase accuracy when $n = 2$. Furthermore, for large ρ values of 0.9 and 1.0, integration accuracy significantly decreases compared to training on the nodes of a uniform grid. However, for the case when $n = 6$, integration accuracy increases when ρ is set to 0.2 and 0.7 for the oscillatory function, and 0.1 and 0.2 for the continuous function, relative to $\rho = 0$.

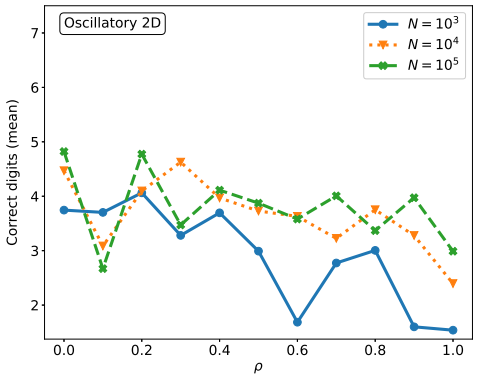
It is worth noting that for functions with a clear peak, with a small number of points ($N = 10^3$) and $n = 2$, increasing ρ the accuracy is rising by 2 digits. However, for the oscillatory function, the situation is reversed with increasing ρ, accuracy decreases from 4 to 2 digits. In the case of the continuous function, for most ρ values, accuracy remains nearly unchanged, but for large ρ values, the average CD value decreases significantly due to the low number or complete absence of points in the training set, which are nodes of the uniform grid. A similar trend of decreasing integration accuracy for large ρ values is also observed for the oscillatory function when $n = 6$. However, for the function with a clear peak, the behavior of the average CD as a function of ρ remains the same for small N values.

With an increase in the value of N, accuracy increases for almost all ρ values, mirroring the change in CD, but less abruptly. It is important to note that the use of the Metropolis–Hastings algorithm for generating the training set reduces computational costs. In particular, for a function with a clear peak, the accuracy at $\rho = 0.8$ and $N = 10^3$ is comparable to the result at $\rho = 0$ and $N = 10^5$. A similar effect exists for other function classes, but only for $n = 6$. The choice of the optimal value of ρ remains an open question.

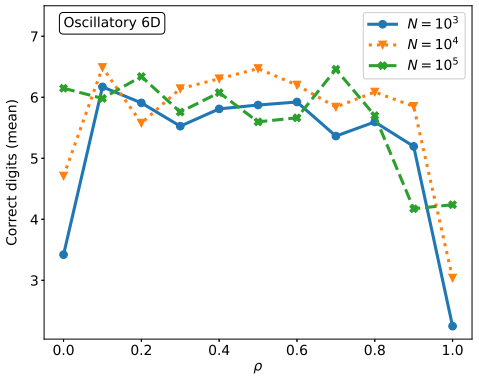
In general, the dependence of accuracy on the proportion of points may not be monotonic since it is determined by the nature of the function itself and the training set derived from this nature. Furthermore, neural network training also depends on weight initialization. Therefore, the non-monotonicity has a statistical nature and is dependent on the type of integrands. An assessment of statistical divergence requires a much larger volume of computations and could be discussed in a separate study.

5. Conclusion

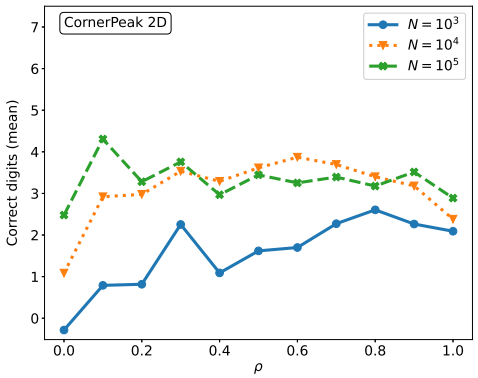
In the context of this study, it was found that the application of the Metropolis–Hastings algorithm improves the accuracy of the neural network integration compared to using a uniform grid of nodes.



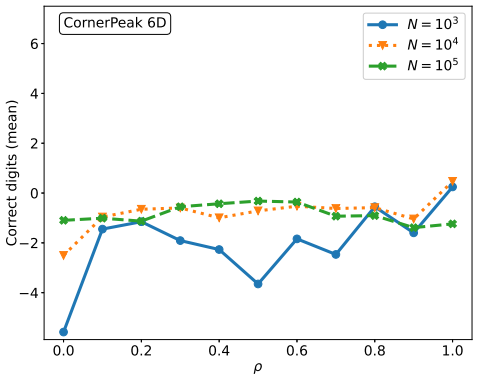
(a) Results for the oscillatory function in two-dimensional space



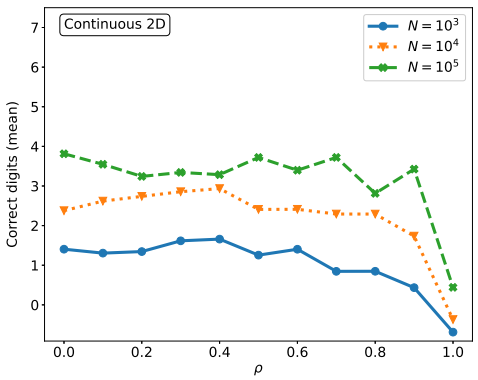
(b) Results for the oscillatory function in six-dimensional space



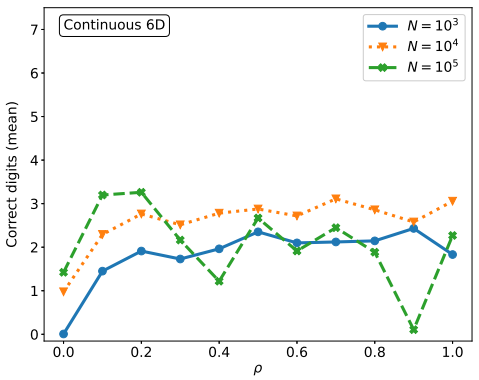
(c) Results for the corner peak function in two-dimensional space



(d) Results for the corner peak function in six-dimensional space



(e) Results for the continuous function in two-dimensional space



(f) Results for the continuous function in six-dimensional space

Figure 2. Results of the validation of hybrid sampling for functions (18)–(20). Each curve represents N. Each point on the graphs is the average value of 20 approximate integrals at a given ρ. For large values of $N = 10^5$, the number of trials was reduced to 5 due to the lengthy training time

Thus, a hybrid method for creating a training dataset has been proposed and tested. Part of the dataset is generated based on the function's values using the aforementioned algorithm, while another part includes nodes of a uniform grid within the function's domain and their corresponding values. To characterize the hybrid method, the concept of the relative proportion of the sample volume, denoted as ρ, obtained through pseudo-random generation, was introduced. The relationship between the accuracy of approximate integration and this parameter was investigated.

The testing was performed on three families of functions with two and six variables, proposed for testing integral computation methods. It was shown that the best results, on average, were obtained when ρ ranged from 0.1 to 0.3. It's also worth noting the improvement in results using the hybrid approach for higher dimensions, denoted as n, and a larger number of points, denoted as N.

Neural network integration appears promising in certain classes of problems, as the analytical formula for its integration as a function of integration domain parameters allows storing the integral form of the function and analytically computing an approximate value in any subdomain without the need for retraining the network. This work demonstrated that sampling using methods to generate points based on the values of integrands in combination with uniform grid nodes can improve the results of approximate integration by a neural network. Nevertheless, the choice of the optimal ratio of the first set of points to the second in the training dataset remains an open question.

Acknowledgments: The authors express their gratitude to the HybriLIT heterogeneous computing platform team for the opportunity to perform calculations in an ecosystem for machine learning, deep learning and data analysis problems. The authors thank Dr. Jan Buśa for valuable comments while reading the manuscript.

References

1. Lloyd, S., Irani, R. A. & Ahmadi, M. Using neural networks for fast numerical integration and optimization. *IEEE Access* **8,** 84519–84531. doi:10.1109/access.2020.2991966 (2020).
2. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control Signals and Systems* **2,** 303–314. doi:10.1007/bf02551274 (Dec. 1989).
3. Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57,** 97–109. doi:10.1093/biomet/57.1.97 (Apr. 1970).
4. Chib, S. & Greenberg, E. Understanding the Metropolis–Hastings algorithm. *The American Statistician* **49,** 327. doi:10.2307/2684568 (Nov. 1995).
5. *Ecosystem for tasks of machine learning, deep learning and data analysis* http://hlit.jinr.ru/en/access-to-resources_eng/ecosystem-for-ml_dl_bigdataanalysis-tasks_eng/. Accessed: 2023-10-10.
6. Chollet, F. *et al. Keras* https://keras.io.
7. Johansson, F. *et al. mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18)* http://mpmath.org.
8. Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **11,** 431–441. doi:10.1137/0111030 (June 1963).
9. Marco, F. D. *Tensorflow Levenberg–Marquardt* https://github.com/fabiodimarco/tf-levenberg-marquardt.
10. Kişi, Ö. & Uncuoğlu, E. Comparison of three back-propagation training algorithms for two case studies. *Indian Journal of Engineering and Materials Sciences* **12,** 434–442 (Oct. 2005).
11. Jiawei Han, M. K. & Pei, J. *Data mining: concepts and techniques* Third Edition. 703 pp. doi:10.1016/c2009-0-61819-5 (Elsevier Inc., 225 Wyman Street, Waltham, MA 02451, USA, 2012).
12. Genz, A. *A package for testing multiple integration subroutines* in *Numerical Integration* 337–340 (Springer Netherlands, 1987). doi:10.1007/978-94-009-3889-2_33.

To cite: Ayriyan A., Grigorian H., Papoyan V., Sampling of integrand for integration using shallow neural network, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)38–47.DOI: 10.22363/2658-4670-2024-32-1-38-47.

Information about the authors

Ayriyan, Alexander—PhD in Physics and Mathematics, Head of sector of the Division of Computational Physics of JINR, Assistant professor of Department of Distributed Information Computing Systems of Dubna State University; Senior Researcher of AANL (YerPhI) (e-mail: ayriyan@jinr.ru, phone: +7(496)216-35-98, ORCID: https://orcid.org/0000-0002-5464-4392)

Grigorian, Hovik—Candidate of Physical and Mathematical Sciences, Senior Researcher of JINR; Senior Researcher of AANL (YerPhI); Assistant professor of Dubna State University; assistant professor of Yerevan State University; (e-mail: hovik.grigorian@gmail.com, phone: +7(496)216-27-46, ORCID: https://orcid.org/0000-0002-0003-0512)

Papoyan, Vladimir—Junior researcher of JINR, Junior researcher of AANL (YerPhI), PhD student of Dubna State University (e-mail: vlpapoyan@jinr.ru, phone: +7(496)216-35-98, ORCID: https://orcid.org/0000-0003-0025-5444)

УДК 519.65:519.217
PACS 07.05.Tp, 07.05.Mh, 02.70.–c
DOI: 10.22363/2658-4670-2024-32-1-38-47 EDN: GFROYO

Способ формирования обучающей выборки для вычисления интеграла с использованием нейронной сети

А. С. Айриян[1, 2, 3], О. А. Григорян[1, 2, 3, 4], В. В. Папоян[1, 2, 3]

[1] *Объединённый институт ядерных исследований,*
ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация
[2] *Национальная научная лаборатория им. А. Алиханяна (ЕрФИ),*
ул. братьев Алиханян, д. 2, Ереван, 0036, Республика Армения
[3] *Государственный университет «Дубна»,*
ул. Университетская, д. 18, Дубна, 141980, Российская Федерация
[4] *Ереванский государственный университет,*
ул. Алекса Манукяна, д. 1, Ереван, 0025, Республика Армения

Аннотация. В настоящей работе исследуется применение алгоритма Метрополиса–Гастингса при формировании обучающей выборки для нейросетевой аппроксимации подынтегральной функции и его влияние на точность вычисления значения интеграла. Предложен гибридный способ формирования обучающего множества, в рамках которого часть выборки генерируется посредством применения алгоритма Метрополиса–Гастингса, а другая часть включает в себя узлы равномерной сетки. Численные эксперименты показывают, что при интегрировании в областях больших размерностей предложенный способ является более эффективным относительно узлов равномерной сетки.

Ключевые слова: нейронная сеть, приближенное интегрирование, алгоритм Метрополиса–Гастингса

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 48–60
http://journals.rudn.ru/miph

Research article

UDC 519.65:519.217
PACS 07.05.Tp, 02.70.–c
DOI: 10.22363/2658-4670-2024-32-1-48-60 EDN: HEYUGO

Solving the eikonal equation by the FSM method in Julia language

Christina A. Stepa[1], Arseny V. Fedorov[1], Migran N. Gevorkyan[1], Anna V. Korolkova[1], Dmitry S. Kulyabov[1,2]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

(received: December 15, 2023; revised: January 25, 2024; accepted: February 14, 2024)

Abstract. There are two main approaches to the numerical solution of the eikonal equation: reducing it to a system of ODES (method of characteristics) and constructing specialized methods for the numerical solution of this equation in the form of a partial differential equation. The latter approach includes the FSM (Fast sweeping method) method. It is reasonable to assume that a specialized method should have greater versatility. The purpose of this work is to evaluate the applicability of the FSM method for constructing beams and fronts. The implementation of the FSM method in the Eikonal library of the Julia programming language was used. The method was used for numerical simulation of spherical lenses by Maxwell, Luneburg and Eaton. These lenses were chosen because their optical properties have been well studied. A special case of flat lenses was chosen as the easiest to visualize and interpret the results. The results of the calculations are presented in the form of images of fronts and rays for each of the lenses. From the analysis of the obtained images, it is concluded that the FSM method is well suited for constructing electromagnetic wave fronts. An attempt to visualize ray trajectories based on the results of his work encounters a number of difficulties and in some cases gives an incorrect visual picture.

Key words and phrases: eikonal equation, geometric optics, wave optics, Julia language, Fast Sweeping Method

1. Introduction

In this article, we use the FSM (Fast Sweeping Method) method to solve the Eikonal equation using the example of three classical lenses: Luneburg, Maxwell and Eaton. These examples illustrate the limitations of the FSM method—it does a good job of calculating wave fronts, but is poorly applicable for calculating the trajectory of rays.

To model lenses, we use the Julia language and the Eikonal library, which implements the FSM method. The simulation results are visualized using the Mackie.jl library. Schematic illustrations are created using a separate vector graphics language, Asymptote.

1.1. Structure of the article

The paper consists of an introduction, a theoretical part, a description of the FSM (Fast sweeping method) [1–5], a description of implementation this method, a visualisation and discussion of the results.

In the theoretical part, the eikonal equation in Cartesian coordinate system is given, and the spherical lenses used for numerical experiments are schematically described.

In the next part, the numerical scheme of the FSM method with detailed formulas for the two-dimensional case is described, and its advantages and disadvantages compared to the feature method are briefly analyzed. Below is a brief description of the Eikonal library [6] for the Julia language [7] and a description of the program we have written that implements a numerical experiment.

In the final part, images of fronts and rays are presented, the results are analysed and conclusions are made about the advantages and disadvantages of the FSM method concerning the visualisation of the calculations.

© Stepa C. A., Fedorov A. V., Gevorkyan M. N., Korolkova A. V., Kulyabov D. S., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

1.2. Designations and agreements

For the purposes of this paper, we have followed standard notation, centred on the classic monograph [8]. All vector quantities are in bold, e.g., the position of a point X is denoted as $\mathbf{x} = (x^1, x^2, x^3)^T = (x, y, z)^T$. All vectors are considered columns and their components are numbered with upper indices. The eikonal function is denoted as $u(\mathbf{x})$, the refractive index is denoted as $n(\mathbf{x})$.

2. Eikonal equation

Previously, the authors in the paper [9] gave a detailed derivation of the eikonal equation for curvilinear coordinates, and in the paper [5] the characteristics method was considered. In this paper, however, we use the FSM method, which uses a rectangular grid, so the eikonal equation must be written in Cartesian coordinates.

We will consider the two-dimensional case of the eikonal equation in Cartesian coordinates with a boundary condition:

$$\begin{cases} \left(\dfrac{\partial u}{\partial x}\right)^2 + \left(\dfrac{\partial u}{\partial y}\right)^2 = n^2(x, y), \\ u(x, y) = 0, \quad (x, y) \in \Gamma \subset \mathbb{R}^2. \end{cases} \tag{1}$$

By specifying the points of the set Γ, the position of the radiation source and the boundary of the medium are specified. Written in this form, the eikonal equation is a nonlinear boundary value problem for the hyperbolic partial derivative equation.

Two approaches to the numerical solution of the eikonal equation can be distinguished:

- transformation to a system of ordinary differential equations (a system of Hamilton equations) by the method of characteristics [10, 11], and then applying one of numerous methods for the numerical solution of such equations;
- approach to the problem as a stationary boundary value problem: development of an efficient numerical algorithm for solving the system of nonlinear equations obtained by discretization (this type of methods includes, for example, the fast marching method).

In this paper, we focus on a method called FSM (*Fast Sweeping Method*). The authors do not know the standard translation of the name of this method, so the abbreviation FSM is used throughout the text.

The method was proposed in 2000 [1]. The basic idea of the method is to use Godunov's counter-flow difference scheme and Gauss–Zeidel iterative scheme with variable order of passing the mesh nodes. A detailed description of the numerical scheme is given in the Section 4.

FSM is simple to implement and requires a finite number of iterations. The complexity of the algorithm is $O(N)$ for N grid points. The number of iterations is independent of the number of grid nodes (of the grid size). The FSM method can be extended to the general case of the Hamilton–Jacobi equation.

3. Luneburg, Maxwell and Eaton lenses

3.1. General description of the lenses

Consider a lens, which is a sphere with centre at point X_0 and radius vector $\mathbf{X}_0$. The source of electromagnetic waves is placed at a point with radius vector $\mathbf{x}_0$.

It should be emphasised immediately that the method used for the numerical solution of the eikonal equation entails a different mathematical description of the electromagnetic wave.

- When using the method of characteristics, it is natural to interpret the radiation in the form of rays. Each ray in this case is a solution of the ODE system for given initial values of the generalised coordinates $\mathbf{x}$ and impulses $\mathbf{p}$. The initial values of the coordinates $\mathbf{x}_0$ specify the position of the source, i.e., the beginning of the ray, and the initial values of the impulses $\mathbf{p}_0$ specify the direction of the ray.
- The FSM method uses the wave interpretation of optics and assumes that the eikonal function $u(\mathbf{x})$ is initially defined at each point in space, or more precisely at each grid point (see section 4).

The location of the source is given by the boundary condition $u(\mathbf{x}_0) = 0$ of the system (1), where $\mathbf{x}_0$ is the radius vector of the points belonging to the source.

Figures 1 and 2 show lenses that receive electromagnetic radiation from a point source (figure 1) and a flat extended source (figure 2). The radiation in the figures is represented as rays, but in the FSM method there is no way to explicitly specify the direction and source of the rays, as it is assumed that the radiation is already present at every point in the region under consideration. This causes certain difficulties when it is necessary to visualise exactly the ray optical pattern.

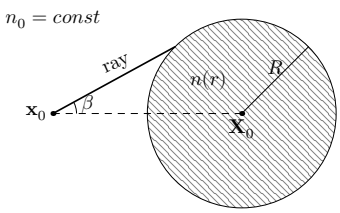


Figure 1. Lens with a point source

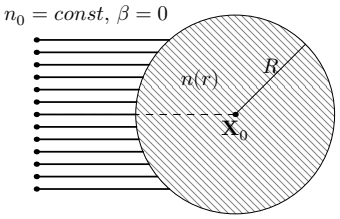


Figure 2. Lens with a flat light source

The refractive index outside the lens is constant and equal to n_0, while inside the lens it is a function of the distance from the centre of the lens to the current point $n(r)$, where $r = \|\mathbf{x} - \mathbf{X}_0\|$. To simplify the calculations, the centre of the lens should be placed at the origin. This especially simplifies the solution of the problem in cylindrical and spherical coordinates.

It may be convenient to define the location of a point source relative to the lens. Then its coordinates are determined by the lens radius R, the distance from the lens to the source d, and the angle θ, which is set off in a counterclockwise direction in the right-hand coordinate system, as shown in the figure 3. Then the radius of the source vector is given by the parametric equation of a circle with radius $R + d$

$$\mathbf{x}_0 = \mathbf{X}_0 + (d + R)(\cos\theta, \sin\theta)^T.$$

In particular, if the source lies on the lens as shown in the figure 4, its coordinates are given by the radius vector

$$\mathbf{x}_0 = \mathbf{X}_0 + R(\cos\theta, \sin\theta)^T.$$

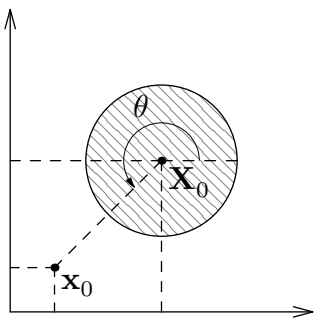


Figure 3. The location of the point source is given relative to the lens

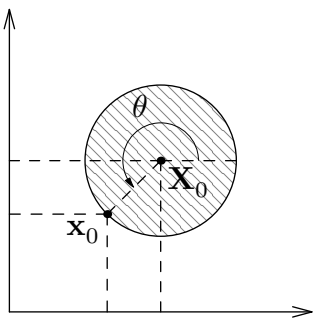


Figure 4. The point source is placed on the lens

The numerical scheme of the FSM method does not require to know the derivatives of the refractive index function $n(\mathbf{x})$, which can be considered as an advantage of this method over the characteristic method.

3.2. Luneburg lens

A Luneburg lens [8, 12] is a spherical lens of radius R centred at the point (X_0, Y_0, Z_0) with a refractive index of the following form

$$n(r) = \begin{cases} n_0\sqrt{2 - \left(\frac{r}{R}\right)^2}, & r \leqslant R, \\ n_0, & r > R, \end{cases} \tag{2}$$

where in Cartesian coordinates $r(x, y, z) = \sqrt{(x - X_0)^2 + (y - Y_0)^2 + (z - Z_0)^2}$ is the distance from the centre of the lens to an arbitrary point (x, y, z). It follows from the formula that the coefficient n varies continuously from $n_0\sqrt{2}$ to n_0 starting from the centre of the lens and ending at its boundary. For calculations it is more convenient to rewrite the expression for r in index form:

$$r(x^1, x^2, x^3) = \sqrt{(x^1 - X_0^1)^2 + (x^2 - X_0^2)^2 + (x^3 - X_0^3)^2} = \sqrt{\sum_{i=1}^{3}(x^i - X_0^i)^2}.$$

3.3. Maxwell lens

Maxwell's lens [8, 13] is also a spherical lens of radius R centred at the point $\mathbf{X}_0$ with a refractive index of the following kind:

$$n(r) = \begin{cases} \dfrac{n_0}{1 + \left(\frac{r}{R}\right)^2}, & r \leqslant R, \\ n_0, & r > R. \end{cases} \tag{3}$$

Figure 5 plots the change in refractive index for Maxwell and Luneburg lenses as a function of the radius of the point vector.

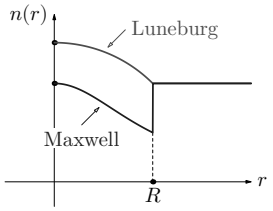


Figure 5. Refractive index of the Luneburg lens and the Maxwell lens

3.4. Eaton Lens

In the flat case, Eaton's lens [14] is a disc formed by two circles with radii R and $2R$, which is shown in the diagram of the lens in the figure 6

$$n(r) = \begin{cases} n_0\sqrt{\dfrac{2R}{r} - 1}, & r \in [R, 2R], \\ n_0, & r \notin [R, 2R]. \end{cases} \tag{4}$$

3.5. Source location

The optical properties of lenses are clearly shown by placing the radiation source at a specific point.

– For a Luneburg lens, the point source is usually placed on the surface of the lens so that the outgoing rays passing through the lens are parallel to each other, as shown in the figure 7.

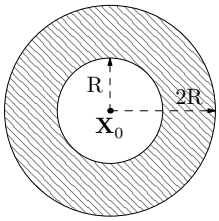


Figure 6. The scheme of the Eaton lens

– For Maxwell lenses, a point source is also placed on the surface of the lens. In this case, the outgoing rays are focused and converge at a point diametrical to the source, as shown in the figure 8.
– For the Eaton lens, the source is placed inside a sphere (circle) of small diameter R. In this case, all radiation does not extend beyond the lens and the rays are focused at a point symmetrical about the centre of the lens, as shown in the figure 9.

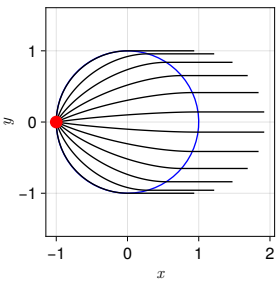


Figure 7. Luneburg lens

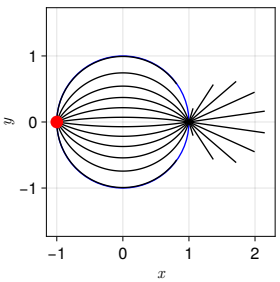


Figure 8. Maxwell lens

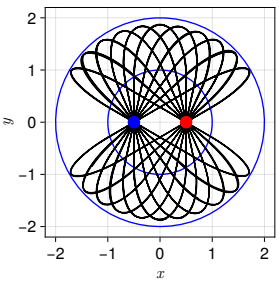


Figure 9. Eaton lens

The ray paths in figures 7, 8, and 9 are calculated using the characteristic [5] method.

If we consider the above source configuration, some regions of space will be completely free of rays. For an Eaton lens, all rays will be enclosed inside the large circle of the lens. For the Luneburg and Maxwell lenses, the areas where rays can pass through are shown by the hatching in the figures 10 and 11. No ray can penetrate the white colour regions.

The figures 10 and 11 are only schemes that conventionally show the areas where rays are present and absent. However, the ray paths are not shown inside the lens.

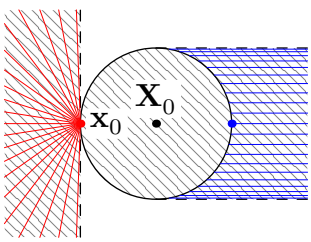


Figure 10. Luneburg lens

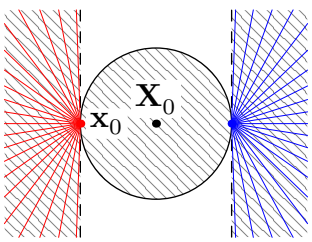


Figure 11. Maxwell lens

4. Fast sweeping method

4.1. FSM for two-dimensional eikonal equation

Let us proceed to the description of the numerical scheme. Let us divide the whole integration region into discrete nodes using the rectangular grid shown in the figure 12, where

- along the axis Ox we have I partition points $x_1 < x_2 < x_3 < \ldots < x_{I-1} < x_I$,
- along the Oy-axis, we have J partition points $y_1 < y_2 < y_3 < \ldots < y_{J-1} < y_J$.

The grid will consist of $I \times J$ nodes with coordinates (x_i, y_j), where $i = 1, \ldots, I$, and $j = 1, \ldots, J$.

Assume that the partitioning is chosen such that the grid spacing h is the same for both axes. The grid function u_{ij} approximates the function $u(x, y)$ at the grid nodes, i.e., $u_{ij} \approx u(x_i, y_j)$ and only at the point (x_0, y_0) does exact equality hold (figure 13).

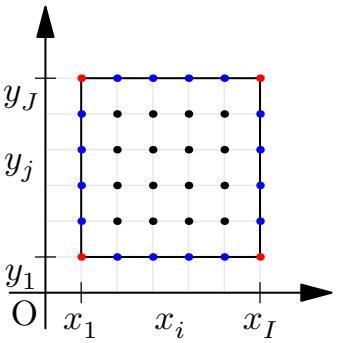


Figure 12. Different grid points of partitioning of the integration domain

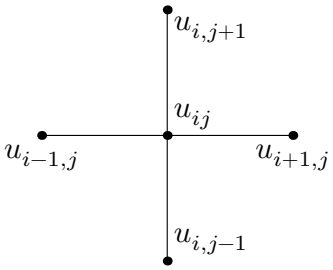


Figure 13. Numerical scheme template. u_{ij} are the grid functions

All grid points can be divided into three groups, which are highlighted in the figure 12 with different colours:

1. The internal grid points with indices $i = 2, \ldots, I-1$ and $j = 2, \ldots, J-1$ are shown in black in the diagram 12.
2. Points of the four grid boundaries with the indices
   - left boundary: $i = 1, \quad j = 2, \ldots, J-1$,
   - right boundary: $i = I, \quad j = 2, \ldots, J-1$,
   - lower boundary $i = 2, \ldots, I-1, \quad j = 1$,
   - upper boundary $i = 2, \ldots, I-1, \quad j = J$,

   are shown in blue in the diagram 12.
3. Corner points with the following fixed indices:
   - lower left corner point $i = 1, \quad j = 1$,
   - upper left corner point $i = 1, \quad j = J$,
   - lower right corner point $i = I, \quad j = 1$,
   - upper right corner point $i = I, \quad j = J$,

   are marked in red in the diagram 12.

Let's move on to presenting the algorithm. The Godunov' scheme (counter-current difference scheme) is used as the sampling scheme for the interior points of the domain. Let us introduce the following notations:

$$u_{xmin} = \min(u_{i-1,j}, u_{i+1,j}), \; u_{ymin} = \min(u_{i,j-1}, u_{i,j+1}), \; n_{ij} = n(x_i, y_j),$$

and also an indicator function:

$$(x)^+ = \begin{cases} x, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

To initialise the calculations, first of all, it is necessary to set the values of the grid function $u_{ij} = 0$ on the boundary Γ. These values will remain unchanged in the subsequent calculations. For all

other points of the grid function u_{ij}, it is necessary to assign sufficiently large positive values, which it obviously cannot reach. During the operation of the numerical scheme these values will be recalculated.

The computational process consists of four sweeps of the entire rectangular region. Each such sweep is two nested loops, where the indices are run in the following order:

- $i = 1, \dots, I$ and $j = 1, \dots, J$ — direct order,
- $i = I, \dots, 1$ and $j = 1, \dots, J$ — mixed order,
- $i = I, \dots, 1$ and $j = J, \dots, 1$ — reverse order,
- $i = 1, \dots, I$ and $j = J, \dots, 1$ — mixed order.

At each step, we should solve a nonlinear equation whose coefficients for each group of points will change slightly.

Group I Points $i = 2, \dots, I-1$ and $j = 2, \dots, J-1$

$$[(u_{ij} - u_{xmin})^+]^2 + [(u_{ij} - u_{ymin})^+]^2 = n_{ij}^2 h^2$$

Group II The following points belong to this group:

- left boundary: $i = 1, \quad j = 2, \dots, J-1$:

$$[(u_{1j} - u_{2j})^+]^2 + [(u_{1j} - u_{ymin})^+]^2 = n_{1j}^2 h^2,$$

- right boundary: $i = I, \quad j = 2, \dots, J-1$:

$$[(u_{Ij} - u_{I-1,j})^+]^2 + [(u_{Ij} - u_{ymin})^+]^2 = n_{Ij}^2 h^2,$$

- lower boundary $i = 2, \dots, I-1, \quad j = 1$:

$$[(u_{i1} - u_{xmin})^+]^2 + [(u_{i1} - u_{i2})^+]^2 = n_{i1}^2 h^2,$$

- upper boundary $i = 2, \dots, I-1, \quad j = J$:

$$[(u_{iJ} - u_{xmin})^+]^2 + [(u_{iJ} - u_{i,J-1})^+]^2 = n_{iJ}^2 h^2.$$

Group III The following points belong to this group:

- left lower corner point $i = 1, \quad j = 1$:

$$[(u_{11} - u_{21})^+]^2 + [(u_{11} - u_{12})^+]^2 = n_{11}^2 h^2,$$

- upper left corner point $i = 1, \quad j = J$:

$$[(u_{1J} - u_{2J})^+]^2 + [(u_{1J} - u_{1,J-1})^+]^2 = n_{1J}^2 h^2,$$

- lower right corner point $i = I, \quad j = 1$:

$$[(u_{I1} - u_{I-1,1})^+]^2 + [(u_{I1} - u_{I2})^+]^2 = n_{I1}^2 h^2,$$

- upper right corner point $i = I, \quad j = J$:

$$[(u_{IJ} - u_{I-1,J})^+]^2 + [(u_{IJ} - u_{I,J-1})^+]^2 = n_{IJ}^2 h^2.$$

Each of these equations differs only in the numerical coefficients and has the following form:

$$\left[(x-a)^+\right]^2 + \left[(x-b)^+\right]^2 = n_{ij}^2 h^2.$$

It can be reduced to a quadratic equation and the solution can be written as

$$x = \begin{cases} \min(a,b) + n_{ij}h, & |a-b| \geqslant n_{ij}h, \\ \dfrac{a+b+\sqrt{2n_{ij}^2h^2 - (a-b)^2}}{2}, & |a-b| < n_{ij}h. \end{cases}$$

4.2. FSM calculation in Julia language

For numerical modelling of lenses using the FSM method, we used the Eikonal [6] package for the Julia [15, 16] programming language. The Eikonal package is a small library that implements the Fast Sweeping and Fast Marching methods. The package is registered in the official Julia package repository and can be installed using standard methods.

Both methods are implemented for arbitrary dimensionality and their source code can fit in a single source file. For such a small package, the documentation is quite detailed and allows you to understand the functionality of the package relatively quickly. There is also a set of tests, which can also serve as illustrative examples.

Let's consider the use of the Eikonal library to compute fronts in the case of Maxwell and Luneburg planar lenses. In addition to this library, we will use our Lenses module described above, and also SVector package.

First of all, let's import all the necessary modules. We read the source code of the Lenses module using `include` and then add it to the common namespace using `using`. Import modules from the official repository using `using`.

```
include("../src/lense.jl")

# We use the FSM method implemented in the Eikonal library
using Eikonal
using StaticArrays: SVector
using .Lenses
```

We need a parametric equation of the circle that returns the coordinates as integers, because the Eikonal module uses an integer grid. We define it as a one-line function. Using the syntax of a dot with an operator or function (e.g. `.+`) allows you to apply the function and operator to all elements of an array or tuple at once.

```
circle_xy(center, R, ϕ) = round.(Int, center .+ (R*cos(ϕ), R*sin(ϕ)))
```

We define the required parameters in the form of constants. Each constant is provided with a documentation line, so no additional comments are needed. The lens type is selected using the `select_lense` function from the Lenses module, which allows you to choose which lens to calculate for right at the beginning of the program.

```
const centre = (500, 500)
"Lens radius"
const R = 300
"Refractive index"
const n₀ = 1.0
"Position of the source relative to the lens"
const source = circle_xy(center, R, π)
"Lens type"
const LENSE = select_lense(length(ARGS)>=1 ? ARGS[1] : "")(R, n₀,
 ↪  SVector(center..., 0.0))
"To draw rays or not"
const RAYS = true
```

Next, we set the grid size (I, J) for the approximating function u_{ij}, initialise the method using the function `FastSweeping` and initialise the array `v`, which in this library denotes the values of the refractive index n_{ij} in the grid nodes. To calculate the values of n_{ij} we use the function `n` from Lenses.

```
const tsize = (1000, 1000)

fsm = FastSweeping(Float64, tsize)

for x=1:tsize[1], y=1:tsize[2]
    fsm.v[x, y] = Lenses.n(SVector(x, y, 0.0), LENSE)
end
```

At the next step we start the calculations. The function `init!` allows you to set the boundary values, which in our case consist of a single point - the ray source. The result of calculations will be written in the attribute `fsm.t`, in our notations it is the approximating grid function u_{ij}.

```
println("Sweeping")
sweep!(fsm, verbose=false)
```

After obtaining the computational results, it remains to visualise the fronts and rays. We use the Makie library, while the Eikonal examples use the Plots library. We create images, axes on it, lens contour as a circle and display it on the coordinate plane. The source is visualised as a point. The fronts are visualised trivially, using the function `contour!`, which displays level lines for a function from two variables u_{ij}. The coordinate arrays x and y may not be passed to it, as they are the same as the indices i and j since we have defined an integer grid.

```
fig01 = Figure(size=(500, 500))
ax01 = Axis(fig01[1, 1])

# Contour of the lens as a circle
const lense_contour = Circle(Point2(center .|> Float64), R)

# Draw the contour of the lens
lines!(ax01, lense_contour, color=:blue)

# Source as a point
scatter!(ax01, source..., color=:red)

# Wavefront as contours of a function from two variables u(x, y)
contour!(ax01, fsm.t, levels=100)
```

Visualising rays is not such a trivial task anymore. The author of the Eikonal library has provided the `ray` function, which calculates ray points using the fastest gradient descent method. To use it, you need to specify the end point of the ray. The ray source will be selected as the ray origin.

```
  for θ = pi/6:0.025:pi/3
    pos = circle_xy(source, min(tsize...) - min(source...), θ)
    r1 = ray(fsm.t, pos)
    lines!(ax01, r1, color=:green)
  end
end
```

It should be noted that the `ray` function is unstable, because often its execution is accompanied by an exit beyond the `fsm.t` array boundaries, even if the ray end point is specified inside the rectangular area.

The result of the visualisation is shown in the figures 14 and 15. It should be noted here that in the case of the Maxwell lens, the application of the `ray` function gives rise to the following problem.

Rays emanating from a source on the surface of the lens must focus to a point (focus) located on the diametrically opposite side of the lens from the source and can only exit the lens after passing that point, as shown in the figure 16. So, when modelling electromagnetic radiation in the form of rays, only rays that have passed the lens and left the point of focus can exist outside the lens. It should be specified that we consider a beam of rays coming out of a point source at an angle in the interval $(-\pi/2, \pi/2)$.

Since the FSM method approximates the eikonal function $u(\mathbf{x})$ at each grid point, electromagnetic radiation will be present everywhere outside the lens in this modelling, as can be seen from the image of the fronts. Therefore, if you set the `ray` parameter `pos` to a point outside the lens surface when calling the `ray` function, the rays will be drawn incorrectly, as shown in the figure 18.

It is possible to obtain a conventionally correct image by specifying as the final position only those grid points that lie exclusively on the surface of the lens.

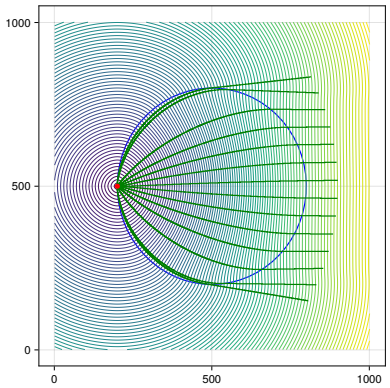


Figure 14. Fronts for the Luneburg lens

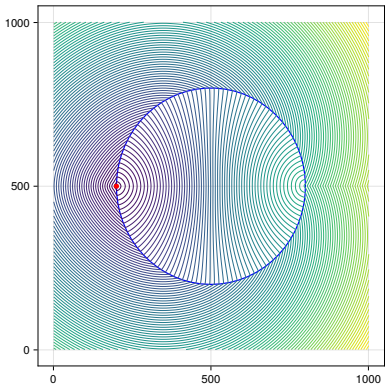


Figure 15. Fronts for the Maxwell lens

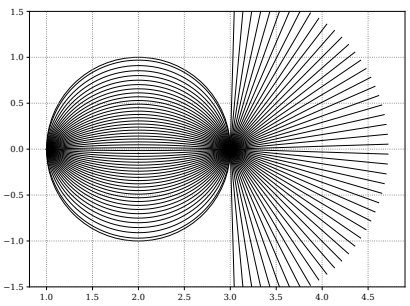


Figure 16. Correct image of the rays of Maxwell lens

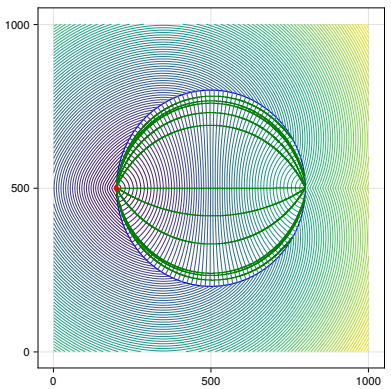


Figure 17. Conditionally correct image of rays of Maxwell lens

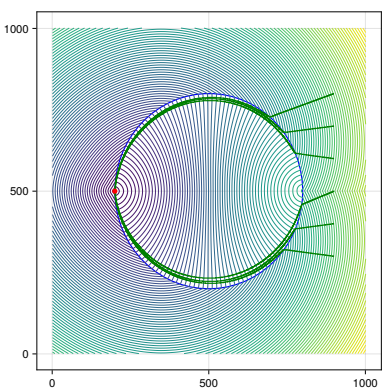


Figure 18. Error image of rays of Maxwell lens modeled by FSM method

5. Discussion

When solving the eikonal equation by the method of characteristics, it is reduced to a system of ODEs. Each solution of the system gives the trajectory of one ray. To obtain a beam of rays, it is necessary to repeatedly solve the obtained ODE system, changing each time the initial values. It can be said that this approach corresponds more to geometrical optics than to wave optics. It is not possible to calculate the points of wave fronts directly by the method of characteristics and requires additional manipulation with the coordinates of the points of the ray trajectory.

In contrast, the FSM method works directly with the Eikonal equation and does not require its transformation into any other form. It also does not require calculations of the derivatives of the refractive index function $n(\mathbf{x})$. The result of the method are approximated values of the function $u(\mathbf{x})$ for all grid points. Having these values, it is rather easy to visualize fronts by depicting level lines, but the visualization of rays has the problems described above.

Another feature of the FSM method is the fact that the electromagnetic field is initially assumed to be present at every point in the modeled region. As shown in figures 10 and 11 when interpreting electromagnetic radiation as rays, there are regions where the radiation does not penetrate.

6. Conclusion

The application of the FSM method for the numerical solution of the eikonal equation has been considered on the example of three well-studied lenses. The following points can be emphasized as advantages of the FSM method.

- It is not necessary to convert the eikonal equation to any other form.
- It is not necessary to find the derivatives of the refractive index function $n(\mathbf{x})$.
- Electromagnetic wave fronts can be constructed directly from the numerical solution.

Difficulties in ray construction and inconsistency with geometrical optics may be pointed out as a disadvantage, which was mentioned in the Discussion section 5.

Funding: This publication has been supported by the RUDN University Scientific Projects Grant System, project No 021934-0-000 (recipients Anna V. Korolkova; Migran N. Gevorkyan) and has been supported by the RUDN University Strategic Academic Leadership Program (recipient Dmitry S. Kulyabov; Arseny V. Fedorov; Christina A. Stepa).

References

1. Zhao, H. A fast sweeping method for Eikonal equations. *Mathematics of Computation* **74,** 603–627. doi:10.1090/s0025-5718-04-01678-3 (May 2004).
2. Gremaud, P. A. & Kuster, C. M. Computational study of fast methods for the eikonal equation. *SIAM Journal on Scientific Computing* **27,** 1803–1816. doi:10.1137/040605655 (Jan. 2006).
3. Jeong, W. & Whitaker, R. A fast eikonal equation solver for parallel systems. *SIAM conference* **84112,** 1–4 (2007).
4. Kulyabov, D. S., Gevorkyan, M. N. & Korolkova, A. V. *Software Implementation of the Eikonal Equation* in *Proceedings of the Selected Papers of the 8th International Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems" (ITTMM-2018), Moscow, Russia, April 16, 2018* (eds Kulyabov, D. S., Samouylov, K. E. & Sevastianov, L. A.) **2177** (Moscow, Apr. 2018), 25–32.
5. Kulyabov, D. S., Korolkova, A. V., Velieva, T. R. & Gevorkyan, M. N. *Numerical analysis of eikonal equation* in *Saratov Fall Meeting 2018: Laser Physics, Photonic Technologies, and Molecular Modeling* (ed Derbov, V. L.) **11066** (SPIE, Saratov, June 2019), 56. doi:10.1117/12.2525142. arXiv: 1906.09467.
6. Févotte, F. *Fast Sweeping and Fast Marching methods for the solution of eikonal equations* version 0.2.0. https://github.com/triscale-innov/Eikonal.jl (2023).
7. Lauwens, B. & Downey, A. *Think Julia* How to Think Like a Computer Scientist. 229 pp. (O'Reilly Media, Inc., 2019).
8. Born, M. & Wolf, E. *Principles of optics* 7th. 952 pp. (Cambridge University Press, 1999).
9. Fedorov, A. V., Stepa, C. A., Korolkova, A. V., Gevorkyan, M. N. & Kulyabov, D. S. Methodological derivation of the eikonal equation. *Discrete and Continuous Models and Applied Computational Science* **31,** 399–418. doi:10.22363/2658-4670-2023-31-4-399-418 (Dec. 2023).

10. Ivanov, D. I., Ivanov, I. E. & Kryukov, I. A. Hamilton–Jacobi equation-based algorithms for approximate solutions to certain problems in applied geometry. *Computational Mathematics and Mathematical Physics* **45,** 1297–1310 (8 2005).
11. Kabanikhin, S. I. & Krivorotko, O. I. Numerical solution eikonal equation. *Siberian Electronic Mathematical Reports* **10,** 28–34 (2013).
12. Fonseca, N. J. G., Tyc, T. & Quevedo–Teruel, O. A solution to the complement of the generalized Luneburg lens problem. *Communications Physics* **4.** doi:10.1038/s42005-021-00774-2 (2021).
13. Abbasi, M. A. B. & Fusco, V. F. Maxwell fisheye lens based retrodirective array. *Scientific Reports* **9.** doi:10.1038/s41598-019-52779-1 (Nov. 2019).
14. Zeng, Y. & Werner, D. H. Two-dimensional inside-out Eaton Lens. Design technique and TM-polarized wave properties. *Optical Express* **20,** 2335–2345. doi:10.1364/OE.20.002335 (Jan. 2012).
15. Gevorkyan, M. N., Kulyabov, D. S. & Sevastyanov, L. A. *Review of Julia programming language for scientific computing* in *The 6th International Conference "Distributed Computing and Grid-technologies in Science and Education"* (2014), 27.
16. Phillips, L. *Practical Julia. A hands-on introduction for scientific minds* 528 pp. (No Starch Press, Oct. 31, 2023).

To cite: Stepa C. A., Fedorov A. V., Gevorkyan M. N., Korolkova A. V., Kulyabov D. S., Solving the eikonal equation by the FSM method in Julia language, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)48–60.DOI: 10.22363/2658-4670-2024-32-1-48-60.

Information about the authors

Stepa, Christina A. (Russian Federation)—PhD student of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1042210111@pfur.ru, phone: +7 (495) 952-02-50, ORCID: https://orcid.org/0000-0002-4092-4326, ResearcherID: GLS-1445-2022)

Fedorov, Arseny V. (Russian Federation)—PhD student of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1042210107@rudn.ru, phone: +7 (495) 955-09-27, ORCID: https://orcid.org/0000-0002-3036-0117, ResearcherID: AGY-9849-2022, Scopus Author ID: 57219092618)

Gevorkyan, Migran N. (Russian Federation)—Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: gevorkyan-mn@rudn.ru, phone: +7 (495) 955-09-27, ORCID: https://orcid.org/0000-0002-4834-4895, ResearcherID: E-9214-2016, Scopus Author ID: 57190004380)

Korolkova, Anna V. (Russian Federation)—Docent, Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: korolkova-av@rudn.ru, phone: +7(495) 952-02-50, ORCID: https://orcid.org/0000-0001-7141-7610, ResearcherID: I-3191-2013, Scopus Author ID: 36968057600)

Kulyabov, Dmitry S. (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor of the Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru, phone: +7 (495) 952-02-50, ORCID: https://orcid.org/0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

УДК 519.65:519.217
PACS 07.05.Tp, 02.70.−c
DOI: 10.22363/2658-4670-2024-32-1-48-60 EDN: HEYUGO

Решение уравнения эйконала методом FSM на языке Julia

К. А. Штепа[1], А. В. Фёдоров[1], М. Н. Геворкян[1], А. В. Королькова[1], Д. С. Кулябов[1,2]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*
[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. Существует два основных подхода к численному решению уравнения эйконала: сведение его к системе ОДУ (метод характеристик) и конструирование специализированных методов для численного решения данного уравнения в виде дифференциального уравнения в частных производных. К последнему подходу относится метод FSM (Fast sweeping method). Резонно предположить, что специализированный метод должен обладать большей универсальностью. Цель данной работы — оценка применимости метода FSM для построения лучей и фронтов. Использовалась реализация метода FSM в библиотеке Eikonal языка программирования Julia. Метод применялся для численного моделирования сферических линз Максвелла, Люнеберга и Итона. Данные линзы были выбраны так как их оптические свойства хорошо изучены. Был выбран частный случай плоских линз, как наиболее простых для визуализации и интерпретации результатов. Результаты вычислений представлены в виде изображений фронтов и лучей для каждой из линз. Из анализа полученных изображений сделан вывод, что метод FSM хорошо подходит для построения фронтов электромагнитных волн. Попытка же по результатам его работы визуализировать траектории лучей наталкивается на ряд трудностей и в некоторых случаях дает неправильную визуальную картину.

Ключевые слова: уравнение эйконала, геометрическая оптика, волновая оптика, Julia language, FSM

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 61–73
http://journals.rudn.ru/miph

Research article

UDC 519.6:004.942
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-61-73

EDN: GZIFWL

Computer research of deterministic and stochastic models "two competitors—two migration areas" taking into account the variability of parameters

Irina I. Vasilyeva[1], Anastasia V. Demidova[2], Olga V. Druzhinina[3], Olga N. Masina[1]

[1] *Bunin Yelets State University, 28 Kommunarov St, Yelets, 399770, Russian Federation*
[2] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[3] *Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 44 Vavilov St, bldg 2, Moscow, 119333, Russian Federation*

(received: February 14, 2024; revised: March 1, 2024; accepted: March 14, 2024)

Abstract. The analysis of trajectory dynamics and the solution of optimization problems using computer methods are relevant areas of research in dynamic population-migration models. In this paper, four-dimensional dynamic models describing the processes of competition and migration in ecosystems are studied. Firstly, we consider a modification of the "two competitors—two migration areas" model, which takes into account uniform intraspecific and interspecific competition in two populations as well as non-uniform bidirectional migration in both populations. Secondly, we consider a modification of the "two competitors—two migration areas" model, in which intraspecific competition is uniform and interspecific competition and bidirectional migration are non-uniform. For these two types of models, the study is carried out taking into account the variability of parameters. The problems of searching for model parameters based on the implementation of two optimality criteria are solved. The first criterion of optimality is associated with the fulfillment of such a condition for the coexistence of populations, which in mathematical form is the integral maximization of the functions product characterizing the populations densities. The second criterion of optimality involves checking the assumption of the such a four-dimensional positive vector existence, which will be a state of equilibrium. The algorithms developed on the basis of the first and second optimality criteria using the differential evolution method result in optimal sets of parameters for the studied population-migration models. The obtained sets of parameters are used to find positive equilibrium states and analyze trajectory dynamics. Using the method of constructing self-consistent one-step models and an automated stochastization procedure, the transition to the stochastic case is performed. The structural description and the possibility of analyzing two types of population-migration stochastic models are provided by obtaining Fokker–Planck equations and Langevin equations with corresponding coefficients. Algorithms for generating trajectories of the Wiener process, multipoint distributions and modifications of the Runge–Kutta method are used. A series of computational experiments is carried out using a specialized software package whose capabilities allow for the construction and analysis of dynamic models of high dimension, taking into account the evaluation of the stochastics influence. The trajectory dynamics of two types of population-migration models are investigated, and a comparative analysis of the results is carried out both in the deterministic and stochastic cases. The results can be used in the modeling and optimization of dynamic models in natural science.

Key words and phrases: one-step processes, population dynamics models, stochastic differential equations, optimality criteria, differential evolution, stochastization, trajectory dynamics, computer modeling, software package

1. Introduction

The study of mathematical models of population systems began to develop actively in the twenties of the last century, thanks to the works of A. Lotka [1] and V. Volterra [2]. Currently, this direction includes the wide class study of models taking into account various interactions between populations

© Vasilyeva I. I., Demidova A. V., Druzhinina O. V., Masina O. N., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

(See, for example, [3–6]). Significant progress in the study is achieved due to the analysis of the dynamic stability models of ecological systems using the theory of differential equations, numerical methods and optimization methods [3, 4, 7–9].

A four-dimensional model of two competing species with migration between two ranges, taking into account the asymmetry coefficient, is considered in [10]. It is shown that the choice of the migration area is carried out depending on the value of this coefficient. The coefficient of asymmetry affects which of the habitats species migrate to first. Two-species Lotka–Volterra competition patch model is studied in [11]. It's shown that in the long time, either the competition exclusion holds that one species becomes extinct, or the two species reach a coexistence equilibrium, and the outcome of the competition is determined by the strength of the inter-specific competition and the dispersal rates.

The transition to the non-deterministic case based on the design stochastic self-consistent models (DSSM) method allows us to identify new qualitative properties of models and carry out a comparative analysis [12–16] and in the other works. For various types of population models, the DSSM method is used in [12, 17, 18]. In [18], a formalized description of the four-dimensional model "two competitors—two migration areas" and its modifications are proposed, taking into account the case when population growth coefficients are different (non-uniform reproduction of species). Using the implementation of the evolutionary algorithm, a set of parameters is obtained that ensure the coexistence of populations in the conditions of competition between two species in the general area, taking into account the migration of these species. Stochastization of the model "two competitors—two migration areas" (under conditions of non-uniform species reproduction) is carried out on the basis of the method of constructing self-consistent stochastic models. The dynamics of trajectories for deterministic and stochastic cases is studied, a comparative analysis is performed.

This paper is a continuation of [18] and contains the construction and analysis of such modifications of the model "two competitors—two migration areas", which allows us to study the influence of non-uniform migration flows and the influence of non-uniform interspecific competition on the trajectory dynamics both in the deterministic case and in the stochastic case.

In section 2 of this paper, we consider the construction of two modifications of the model "two competitors—two migration areas" with bidirectional non-uniform migration (to two refuges), taking into account the uniformity and non-uniformness of the interspecific competition coefficients. In section 3, the search for model parameters is carried out using an evolutionary algorithm taking into account different optimality criteria. In section 4, a study of the obtained deterministic four-dimensional models is carried out, two-dimensional and three-dimensional projections of phase portraits are constructed. In section 5, the transition to stochastic models "two competitors—two migration areas" is made on the basis of the constructing self-consistent stochastic models method, the dynamics of trajectories in the stochastic case is studied. The results of computer experiments are presented and the interpretation of these results is given taking into account the comparison of stochastic and deterministic models. The developed in Python [19] software package [20] is used to study the models. Section 7 discusses the results.

2. Description of the model modifications "two competitors—two migration areas" taking into account non-uniform migration

Ref. [18] describes a general four-dimensional deterministic model, which takes into account the influence of interspecific and intraspecific competition in two populations with bidirectional migration of both populations, and the non-uniform growth of population reproduction.

We describe further such a model "two competitors—two migration areas", for which the growth of population reproduction, interspecific and intraspecific competition are uniform, and migration is non-uniform. The specified model is given by a system of equations of the form

$$\begin{aligned}
\dot{x}_1 &= ax_1 - px_1^2 - rx_1x_3 + \beta x_2 - \gamma x_1,\\
\dot{x}_2 &= ax_2 - px_2^2 + \gamma x_1 - \beta x_2,\\
\dot{x}_3 &= ax_3 - px_3^2 - rx_1x_3 + \varepsilon x_4 - \delta x_3,\\
\dot{x}_4 &= ax_4 - px_4^2 + \delta x_3 - \varepsilon x_4,
\end{aligned} \tag{1}$$

where the incoming values are explained in the table 1.

Table 1

Variables and parameters of model (1)

Variable, parameter	Explanation of the variable, parameter
x_1	the competing population density of the first species in the general area
x_2	population density of the first species in the first refuge
x_3	the competing population density of the second species in the general area
x_4	population density of the second species in the second refuge
a	coefficient of natural growth
r	coefficient of interspecific competition
p	coefficient of intraspecific competition
β	coefficient of migration of the first species from the general area to the first refuge
γ	coefficient of migration of the first species from the first refuge to the general area
δ	coefficient of migration of the second species from the general area to the second refuge
ε	coefficient of migration of the second species from the second refuge to the general area

Let's move from model (1) to a model that takes into account the non-uniformness of the coefficient of interspecific competition r. We denote the estimated parameter of the competitive impact of the second type on the first by p_{13}. Accordingly, we denote the coefficient of the impact of the first type on the second by p_{31}. Thus we obtain a system of differential equations of the following form:

$$\begin{aligned} \dot{x}_1 &= ax_1 - px_1^2 - p_{13}x_1x_3 + \beta x_2 - \gamma x_1, \\ \dot{x}_2 &= ax_2 - px_2^2 + \gamma x_1 - \beta x_2, \\ \dot{x}_3 &= ax_3 - px_3^2 - p_{31}x_1x_3 + \varepsilon x_4 - \delta x_3, \\ \dot{x}_4 &= ax_4 - px_4^2 + \delta x_3 - \varepsilon x_4, \end{aligned} \tag{2}$$

where a is the coefficient of natural growth, β is the coefficient of the first species migration from the general area to the first refuge, γ is the coefficient of the first species migration from the first refuge to the general area, δ is the coefficient of the second species migration from the general area to the second refuge, ε is the coefficient of the second species migration from the second refuge to the general area, p_{ij} $(i \neq j)$ are coefficients of interspecific competition.

In the future, we will search for optimal sets of parameters that ensure the coexistence of species in the general area and the existence of species in refuges. In addition, we will carry out a comparative analysis of the trajectory dynamics of models (1) and (2), taking into account the considered initial conditions and optimal sets of parameters, as well as construct and study stochastic models corresponding to (1) and (2).

3. Search for model parameters using differential evolution

We consider optimization problems associated with finding parameter sets that guarantee the coexistence of all species in the general area and the existence of migratory species in refuges. The sets of parameters to be searched correspond to stationary modes of the system. We use such a numerical optimization method, which reduces to the implementation of the differential evolution algorithm [21–23].

In this paper we use two optimality criteria. The first optimality criterion is associated with maximizing the integral of the product of functions characterizing population densities. The specified maximization ensures the fulfillment of the selected condition for the coexistence of populations. It should be noted that an algorithm of differential evolution is used to solve problem of maximizing the integral, a program in Python is developed [18]. The second criterion of optimality involves checking the assumption of the existence of a positive equilibrium state of the four-dimensional model [24].

Based on the use of the first and second criteria, algorithms are developed using the differential evolution method. The algorithms are aimed at obtaining optimal sets of parameters of a four-dimensional system, which makes it possible to find approximate values of the components of the vector corresponding to a positive equilibrium state and to study the trajectory dynamics near equilibrium states. It is important to note that these two types of algorithms are based on intelligent search methods [25–27]. The implementation of these methods in the case under consideration makes it possible to find the numerical values of the components of the positive equilibrium state and identify such system parameters that provide the required properties of the equilibrium states.

In [18, 24], the implementation of algorithms for modifications of the system "two competitors—two migration areas" is considered without taking into account the non-uniform competition and migration. Here we consider more complex modifications with non-uniform bidirectional migration of two populations. Table 2 shows the characteristics of models (1) and (2) taking into account the first and second optimality criteria. The set of parameters obtained using the first optimality criterion is denoted by i-I, $i = 1, 2$. The set of parameters obtained using the second optimality criterion is denoted by i-II, $i = 1, 2$.

Table 2
Summary table of characteristics of models (1) and (2) under initial conditions $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$.

Equilibrium States	Parameters
Model (1) with set 1-I	
$x_1 = 58.04$, $x_2 = 90.75$, $x_3 = 56.42$, $x_4 = 91.04$	$a = 10.00$, $p = 0.10$, $r = 0.10$, $\beta = 5.67$, $\delta = 9.01$, $\gamma = 7.42$, $\varepsilon = 6.48$
Model (1) with set 1-II	
$x_1 = 58.27$, $x_2 = 90.79$, $x_3 = 56.09$, $x_4 = 91.17$	$a = 10.00$, $p = 0.10$, $r = 0.10$, $\beta = 7.32$, $\delta = 9.99$, $\gamma = 9.97$, $\varepsilon = 7.03$
Model (2) with set 2-I-a	
$x_1 = 57.78$, $x_2 = 90.46$, $x_3 = 56.89$, $x_4 = 90.70$	$a = 10.00$, $p = 0.10$, $p_{13} = 0.50$, $p_{31} = 0.70$, $\beta = 5.67$, $\delta = 9.01$, $\gamma = 7.42$, $\varepsilon = 6.48$
Model (2) with set 2-I-b	
$x_1 = 58.31$, $x_2 = 90.39$, $x_3 = 56.58$, $x_4 = 90.71$	$a = 10.00$, $p = 0.10$, $p_{13} = 0.70$, $p_{31} = 0.50$, $\beta = 5.67$, $\delta = 9.01$, $\gamma = 7.42$, $\varepsilon = 6.48$
Model (2) with set 2-II-a	
$x_1 = 71.34$, $x_2 = 98.80$, $x_3 = 6.06$, $x_4 = 43.60$	$a = 10.00$, $p = 0.10$, $p_{13} = 0.50$, $p_{31} = 0.70$, $\beta = 7.32$, $\delta = 9.99$, $\gamma = 9.97$, $\varepsilon = 7.03$
Model (2) with set 2-II-b	
$x_1 = 6.16$, $x_2 = 41.57$, $x_3 = 69.73$, $x_4 = 99.63$	$a = 10.00$, $p = 0.10$, $p_{13} = 0.70$, $p_{31} = 0.50$, $\beta = 7.32$, $\delta = 9.99$, $\gamma = 9.97$, $\varepsilon = 7.03$

Table 2 presents the sets of parameters we use in the process of computer experiments related to the analysis of the trajectory dynamics of models (1) and (2).

4. Results of computational experiments and comparative analysis of trajectory dynamics for deterministic models

This section presents the results of computational experiments for models (1) and (2), taking into account the selected initial conditions and the found parameters. Figure 1 shows the trajectories of the system (2) for the set of parameters 2-I-a in comparison with the corresponding trajectories of the system (1) with the resulting set of parameters 1-I.

Figure 2 shows the trajectories of system (2) for the set of parameters 2-I-b in comparison with the corresponding trajectories of system (1) with the resulting set of parameters 1-I. According to figures 1, 2, the densities of the corresponding populations for models (1) and (2) are kept at the same level for the selected time interval. Figure 3 shows the trajectories of system (1) with sets of parameters 1-I and 1-II.

Figure 4 shows the trajectories of system (2) for a set of parameters 2-I-a in comparison with the corresponding trajectories of system (2) with the resulting set of parameters 2-II-a.

Next, we will consider some projections of the phase portraits of model (2) on the plane and in space. The projection of the phase portrait on the plane (x_1, x_2) for the system (2) taking into account $x_3 = 56.89$, $x_4 = 90.70$ is shown in figure 5. The projection of the phase portrait in space (x_1, x_2, x_3) taking into account $x_4 = 90.70$ for the model (2) is shown in figure 6.

The presented analysis of models (1) and (2) is carried out with two sets of each model parameters and is aimed at studying two modes of species coexistence for a state of equilibrium determined by the method of differential evolution. In the future, we will study and analysis of the qualitative properties of the models, taking into account the introduction of stochastics.

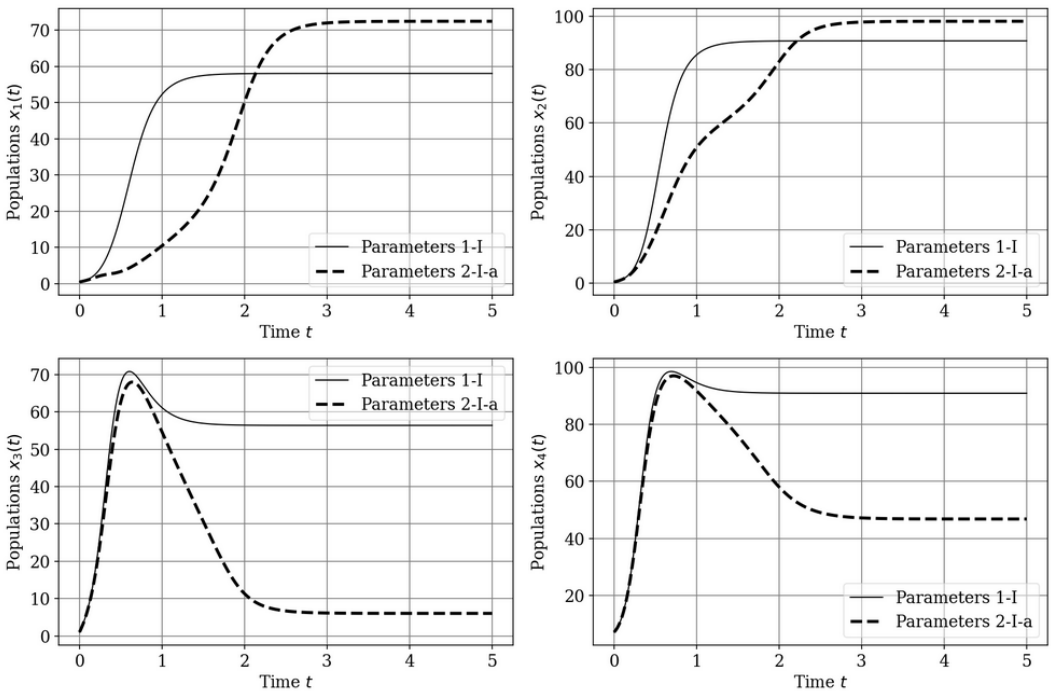


Figure 1. Trajectories of (1) and (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account the parameters sets 1-I and 2-I-a

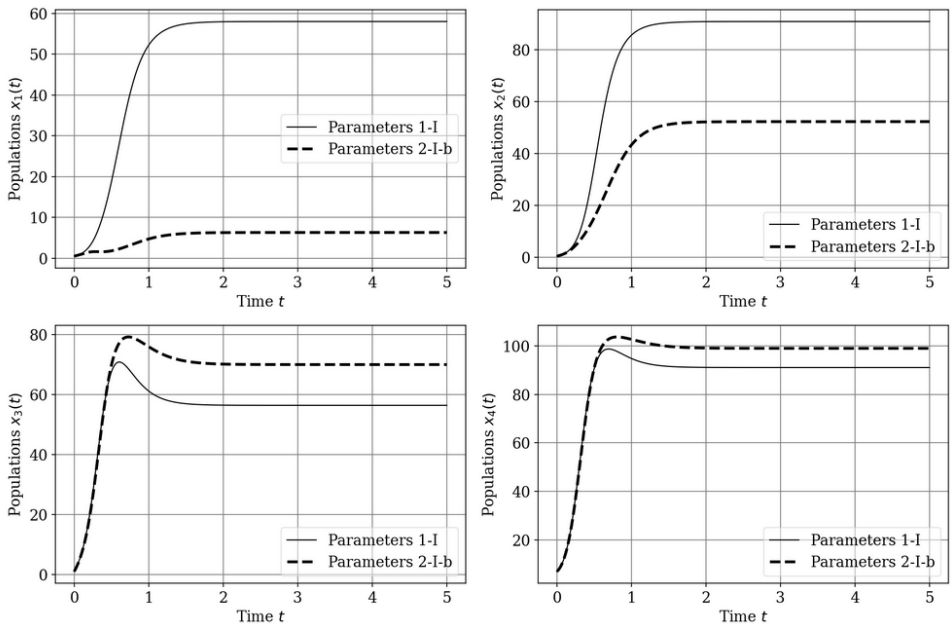


Figure 2. Trajectories of (1) and (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account the parameters sets 1-I and 2-I-b

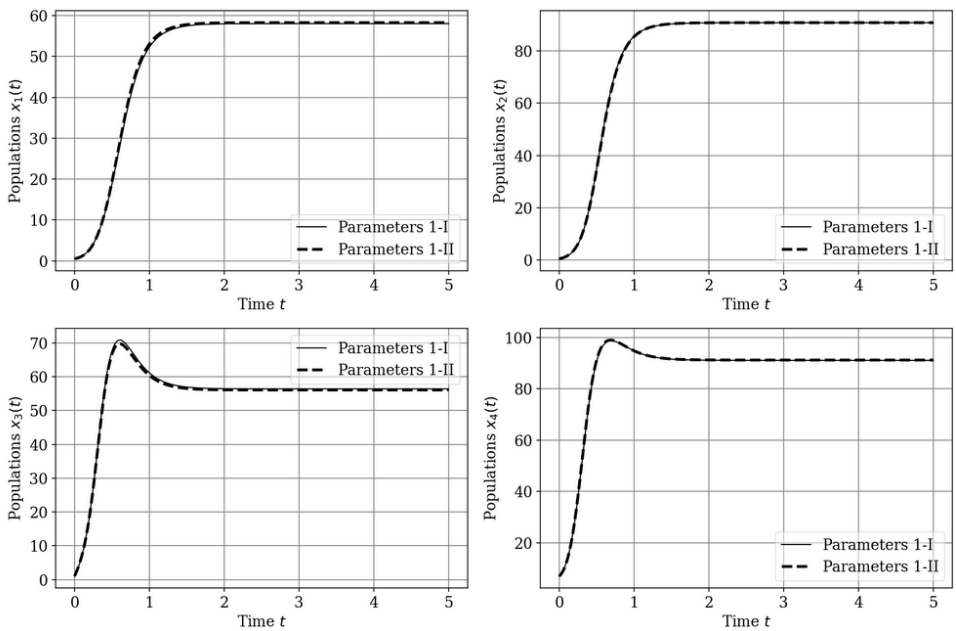


Figure 3. Trajectories of (1) and (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account the parameters sets 1-I and 1-II

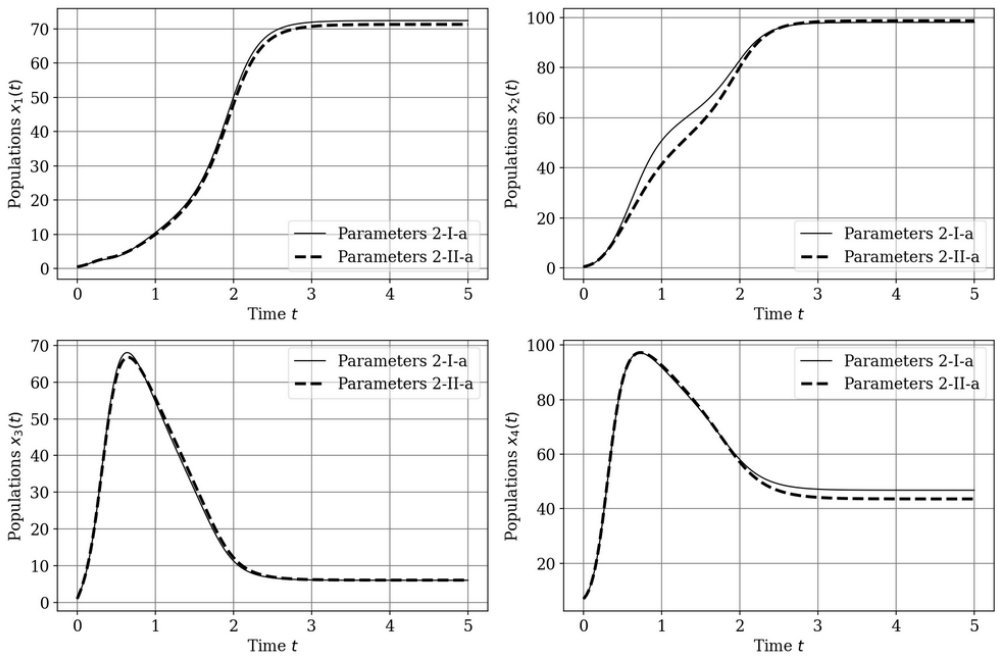


Figure 4. Trajectories of (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account the parameters sets 2-I-a and 2-II-a

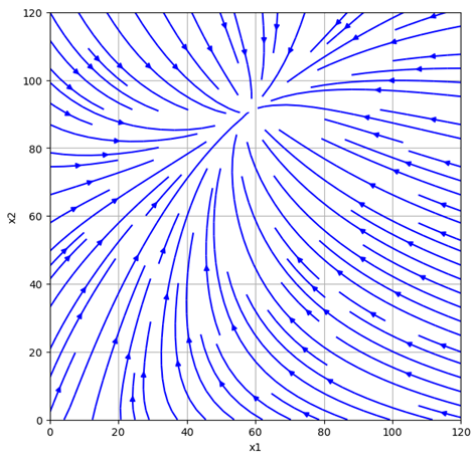


Figure 5. Projection of the phase portrait on the plane (x_1, x_2) for system (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$, $a = 10.00$, $p = 0.10$, $p_{13} = 0.10$, $p_{31} = 0.10$, $\beta = 2.56$, $\delta = 2.19$, $\gamma = 2.52$, $\varepsilon = 2.30$

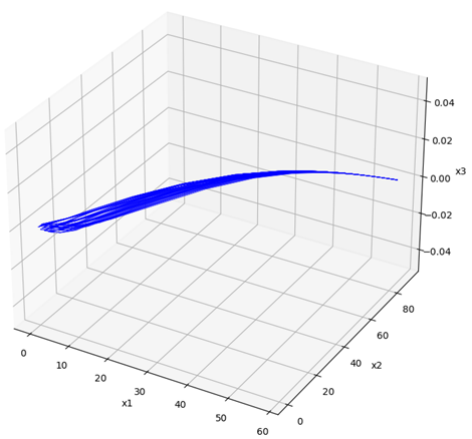


Figure 6. Projection of the phase portrait in the space (x_1, x_2, x_3) for system (2) at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$, $a = 10.00$, $p = 0.10$, $p_{13} = 0.10$, $p_{31} = 0.10$, $\beta = 2.56$, $\delta = 2.19$, $\gamma = 2.52$, $\varepsilon = 2.30$

5. Modifications of the stochastic model "two competitors—two migration areas"

To construct stochastic models corresponding to models (1) and (2), it is proposed to use the DSSM method [12–16]. Using the software implementation of this method allows:

(*i*) to construct a stochastic model of a dynamic system taking into account the description of interactions;
(*ii*) to construct an appropriate deterministic model;
(*iii*) to obtain numerical solutions of ODEs and SDUs and graphical representations of solutions.

To describe a stochastic system, according to the DSSM method, it is enough to write the Fokker–Planck equation. The coefficients of the Fokker–Planck equation for models (1) and (2) are obtained using a software package and are presented respectively in figures 7, 8.

```
Ввод [19]: f = de.drift_vector(XX, k_plus_1, model_1.matr_N(), model_1.matr_M())
           sp.Matrix(f)
```

Out[19]:

$$\begin{bmatrix} ax_1 - px_1^2 - rx_1x_3 - x_1\gamma + x_2\beta \\ ax_2 - px_2^2 + x_1\gamma - x_2\beta \\ ax_3 - px_3^2 - rx_1x_3 - x_3\delta + x_4\varepsilon \\ ax_4 - px_4^2 + x_3\delta - x_4\varepsilon \end{bmatrix}$$

```
Ввод [20]: g=de.diffusion_matrix(XX, k_plus_1, model_1.matr_N(), model_1.matr_M())
           g
```

Out[20]:

$$\begin{bmatrix} ax_1 + px_1^2 + rx_1x_3 + x_1\gamma + x_2\beta & -x_1\gamma - x_2\beta & 0 & 0 \\ -x_1\gamma - x_2\beta & ax_2 + px_2^2 + x_1\gamma + x_2\beta & 0 & 0 \\ 0 & 0 & ax_3 + px_3^2 + rx_1x_3 + x_3\delta + x_4\varepsilon & -x_3\delta - x_4\varepsilon \\ 0 & 0 & -x_3\delta - x_4\varepsilon & ax_4 + px_4^2 + x_3\delta + x_4\varepsilon \end{bmatrix}$$

Figure 7. Coefficients of the Fokker–Planck equation for the model (1)

```
Ввод [65]: f = de.drift_vector(XX, k_plus_2, model_2.matr_N(), model_2.matr_M())
           sp.Matrix(f)
```

Out[65]:

$$\begin{bmatrix} ax_1 - px_1^2 - p_{13}x_1x_3 - x_1\gamma + x_2\beta \\ ax_2 - px_2^2 + x_1\gamma - x_2\beta \\ ax_3 - px_3^2 - p_{31}x_1x_3 - x_3\delta + x_4\varepsilon \\ ax_4 - px_4^2 + x_3\delta - x_4\varepsilon \end{bmatrix}$$

```
Ввод [66]: g=de.diffusion_matrix(XX, k_plus_2, model_2.matr_N(), model_2.matr_M())
           g
```

Out[66]:

$$\begin{bmatrix} ax_1 + px_1^2 + p_{13}x_1x_3 + x_1\gamma + x_2\beta & -x_1\gamma - x_2\beta & 0 & 0 \\ -x_1\gamma - x_2\beta & ax_2 + px_2^2 + x_1\gamma + x_2\beta & 0 & 0 \\ 0 & 0 & ax_3 + px_3^2 + p_{31}x_1x_3 + x_3\delta + x_4\varepsilon & -x_3\delta - x_4\varepsilon \\ 0 & 0 & -x_3\delta - x_4\varepsilon & ax_4 + px_4^2 + x_3\delta + x_4\varepsilon \end{bmatrix}$$

Figure 8. Coefficients of the Fokker–Planck equation for the model (2)

Graphs of the numerical solution for the deterministic and stochastic case taking into account the sets of parameters 1-I and 2-I-a from the table 1 are shown in the (figures 9, 10). For the numerical solution of ODE systems, we use a software implementation of standard four-order Runge–Kutta methods. To solve the corresponding SDE, we use a specially developed library a detailed description of which is contained in [13].

For sets of parameters 1-II, 2-I-a, 2-I-b, 2-II-a, numerical experiment shows that in the stochastic case the trajectories fluctuate near deterministic trajectories, similar to the one shown in figures 9, 10 taking into account sets of parameters 1-I and 2-II-b respectively. In the next section, we present a comparative analysis of the results obtained for models (1) and (2) with different sets of parameters.

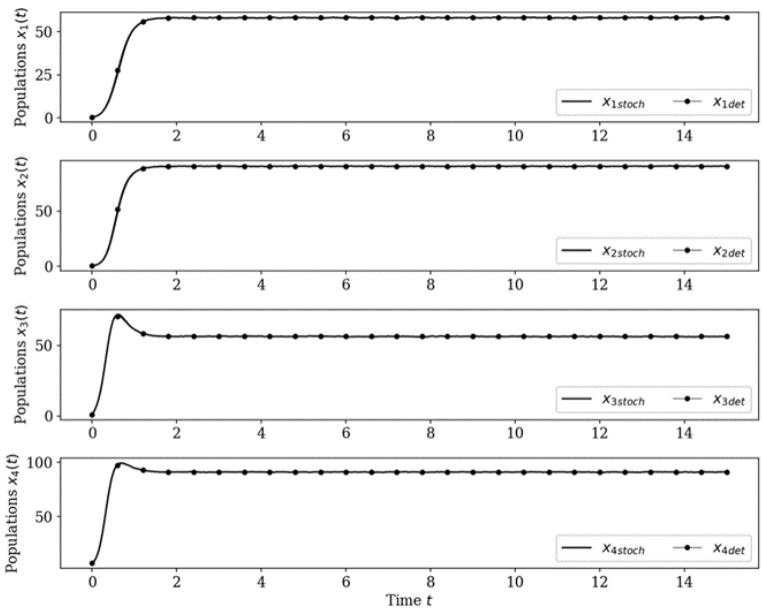


Figure 9. Trajectories of the system (1) and the corresponding stochastic system at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account parameters set 1-I: $a = 10.00$, $p = 0.10$, $r = 0.10$, $\beta = 5.67$, $\delta = 9.00$, $\gamma = 7.42$, $\varepsilon = 6.48$

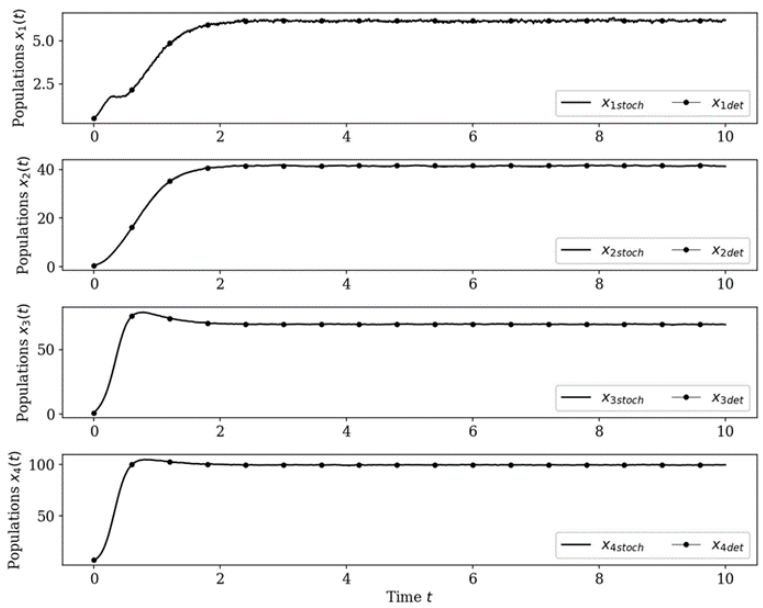


Figure 10. Trajectories of the system (2) and the corresponding stochastic system at $(x_1(0), x_2(0), x_3(0), x_4(0)) = (0.5, 0.5, 1, 7)$ taking into account parameters set 2-II-b: $a = 10.00$, $p = 0.10$, $p_{13} = 0.70$, $p_{31} = 0.50$, $\beta = 7.32$, $\delta = 9.99$, $\gamma = 9.97$, $\varepsilon = 7.03$

6. Discussion of the results

Let us first consider the results of a computational experiment for the case of parameters 1-I and parameters 2-I-a at $p_{13} < p_{31}$. A comparative analysis of the trajectories' behavior for models (1) and (2) according to figure 1 shows that:

1) there is a co-existence of all species corresponding to the stationary regime in the general area, as well as the existence of migratory species in refuges;
2) the non-uniformness of the interspecific competition coefficients affects the population density x_3 in the general area, while the population density x_3 of the model 2 decreases compared to the population density x_3 of the model 1.

Let us further consider the results for the case of parameters 1-I and 2-I-b at $p_{13} > p_{31}$. A comparative analysis of the solution trajectories of models (1) and (2) presented in figure 2 shows that:

1) there is a coexistence of all species in the general area and the existence of migratory species in refuges;
2) the population density x_1 of the model (2) decreases compared to the population density x_1 of the model (1).

Figure 3 shows the trajectories of the model (1) solutions taking into account parameters 1-I and 1-II and two optimization methods. Figure 4 shows the trajectories of solutions for model (2) taking into account the parameters 2-I-a and 2-II-a and two optimization methods. Comparison of trajectories allows us to conclude that the choice of the first or second optimality criterion does not significantly affect the population density both in the general area and in refuges. Computational experiments show the consistency of the two selected optimality criteria (the trajectories have a similar character).

Taking into account the transition to the stochastic case using the Fokker–Planck equations (figures 7, 8), computer experiments are carried out to identify trajectory dynamics. The results are presented in figures 9, 10. Computer experiments show that the introduction of stochastics has no effect on the behavior of systems described by the systems of equations (1) and (2). As in the deterministic case, solutions of stochastic differential equations reach the stationary mode. In Fig. 9, where the range from 0 to 10 corresponds to the change in population density x_1 in the general area, the fluctuating nature of the trajectory dynamics in the stochastic case is visually observed. In order to observe such a character for the other variables, it is possible to choose an enlarged scale of the drawing. For example, a fragment of the deterministic and stochastic trajectories of model (2) for the population density x_4 in the migration area, taking into account the enlarged scale, is shown in figure 11.

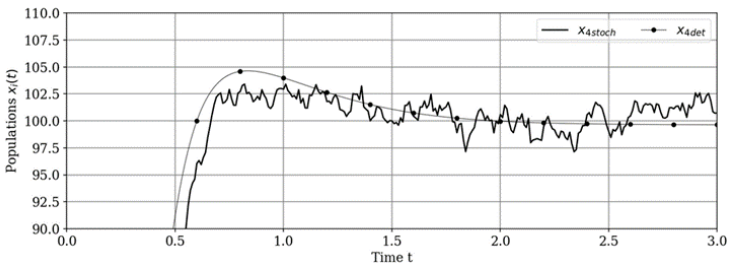


Figure 11. Fragment of deterministic and stochastic trajectories for model (2) with parameters set 2-II-b

Stochastic modeling revealed the similar nature of the trajectories of models (1) and (2) with the considered sets of parameters, and the initial set of parameters is obtained taking into account the conditions guaranteeing the coexistence of two populations in the general area and the positive abundance of each species in refuges. Thus, with the considered sets of parameters of models (1) and (2), it is sufficient to carry out a computational experiment only in the deterministic case to study the trajectory dynamics.

7. Conclusion

The paper presents a computer study of deterministic and stochastic models "two competitors—two migration areas". Computational experiments are based on the use of evolutionary algorithms and optimization methods for finding parameters taking into account their variability. The optimization problem is solved using the differential evolution method, which made it possible to find optimal parameters for population-migration models. Approximate stable equilibrium states corresponding to the obtained parameters are found for these models. The assessment of the changes influence in the coefficients of interspecific competition on the trajectory dynamics of four-dimensional models with non-uniform migration flows in the deterministic case is considered.

In this paper, the transition to stochastic models of population dynamics, taking into account competition and migration, based on the DSSM method, is carried out. The stochastization algorithm used made it possible to analyze the trajectory dynamics of stochastic models in comparison with deterministic analogues. The analysis demonstrates a negligible effect of introducing stochastics into deterministic models (1) and (2) taking into account the parameters and allows us to conclude that the computational experiment is sufficient only in the deterministic case.

As directions for further research, we can indicate the construction of population-migration models of the form "n competitors—n migration areas", where $n > 2$, as well as the study of multidimensional stochastic models taking into account the effects of additive random perturbations of the equations right-hand sides.

References

1. Lotka, A. J. *Elements of Physical Biology* (Williams and Wilkins Company, Baltimore, MD, USA, 1925).
2. Volterra, V. Fluctuations in the abundance of a species considered mathematically. *Nature,* 558–560. doi:10.1038/118558a0 (1926).
3. Bazykin, A. D. *Nonlinear dynamics of interacting populations* (Institute of Computer Research, Moscow–Izhevsk, 2003).
4. Pykh, Y. A. *Generalized Lotka–Volterra systems: theory and applications* (SPbGIPSR, St. Petersburg, 2017).
5. Li, A. *Mathematical modelling of ecological systems in patchy environments* in *Electronic Thesis and Dissertation Repository* (2021), 8059.
6. Hsu, S. B., Ruan, S. & Yang, T. H. Analysis of three species Lotka–Volterra food web models with omnivory. *Journal of Mathematical Analysis and Applications* **426,** 659–687. doi:10.1016/j.jmaa.2015.01.035 (2015).
7. Shestakov, A. A. *Generalized direct method for systems with distributed parameters* (URSS, Moscow, 2007).
8. Moskalenko, A. I. *Methods of nonlinear mappings in optimal control. Theory and applications to models of natural systems* (Nauka, Novosibirsk, 1983).
9. Zhou, Q. *Modelling Walleye population and its cannibalism effect* in *Electronic Thesis and Dissertation Repository* (2017), 4809.
10. Sadykov, A. & Farnsworth, K. D. Model of two competing populations in two habitats with migration: Application to optimal marine protected area size. *Theoretical Population Biology* **142,** 114–122. doi:10.1016/j.tpb.2021.10.002 (2021).
11. Chen, S., Shi, J., Shuai, Z. & Wu, Y. Global dynamics of a Lotka–Volterra competition patch model. *Nonlinearity* **35,** 817. doi:10.1088/1361-6544/ac3c2e (2021).
12. Demidova, A. V. Equations of population dynamics in the form of stochastic differential equations. *RUDN Journal of Mathematics, Information Sciences and Physics* **1,** 67–76 (2013).
13. Gevorkyan, M. N., Velieva, T. R., Korolkova, A. V., Kulyabov, D. S. & Sevastyanov, L. A. *Stochastic Runge–Kutta software package for stochastic differential equations* in *Dependability Engineering and Complex Systems* (eds Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T. & Kacprzyk, J.) (Springer International Publishing, Cham, 2016), 169–179. doi:10.1007/978-3-319-39639-2_15.
14. Korolkova, A. & Kulyabov, D. One-step stochastization methods for open systems. *EPJ Web of Conferences* **226,** 02014. doi:10.1051/epjconf/202022602014 (2020).
15. Gardiner, C. W. *Handbook of stochastic methods: for Physics, Chemistry and the Natural Sciences* (Springer, Heidelberg, 1985).

16. Van Kampen, N. G. *Stochastic processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992).
17. Demidova, A. V., Druzhinina, O. V., Masina, O. N. & Petrov, A. A. Synthesis and computer study of population dynamics controlled models using methods of numerical optimization, stochastization and machine learning. *Mathematics* **9,** 3303. doi:10.3390/math9243303 (2021).
18. Vasilyeva, I. I., Demidova, A. V., Druzhinina, O. V. & Masina, O. N. Construction, stochastization and computer study of dynamic population models "two competitors – two migration areas". *Discrete and Continuous Models and Applied Computational Science* **31,** 27–45. doi:10.22363/2658-4670-2023-31-1-27-45 (2023).
19. Sullivan, E. *Numerical methods – an inquiry-based approach with Python* 2021.
20. Demidova, A. V., Druzhinina, O. V., Masina, O. N. & Petrov, A. A. Development of algorithms and software for modeling controlled dynamic systems using symbolic computations and stochastic methods. *Programming and Computer Software* **49,** 108–121. doi:10.1134/S036176882302007X (2023).
21. Storn, R. & Price, K. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization* **23** (1995).
22. Das, S. & Suganthan, P. N. Differential evolution: a survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation* **15,** 4–31. doi:10.1109/TEVC.2010.2059031 (2011).
23. Eltaeib, T. & Mahmood, A. Differential evolution: a survey and analysis. *Applied Sciences* **8,** 1945. doi:10.3390/app8101945 (2018).
24. Petrov, A. A., Druzhinina, O. V., Masina, O. N. & Vasilyeva, I. I. The construction and analysis of four-dimensional models of population dynamics taking into account migration flows. *Uchenye zapiski UlGU. Series: Mathematics and Information Technology",* 43–55 (2022).
25. Vucetic, D. *Fuzzy differential evolution algorithm* in *Electronic Thesis and Dissertation Repository* (2012), 503.
26. Karpenko, A. P. *Modern search engine optimization algorithms. Algorithms inspired by nature* 2nd ed. (N.E. Bauman MSTU, Moscow, 2016).
27. Simon, D. *Evolutionary Optimization Algorithms. Biologically-Inspired and Population-Based Approaches to Computer Intelligence* (John Wiley & Sons, Inc., New York, 2013).

To cite: Vasilyeva I. I., Demidova A. V., Druzhinina O. V., Masina O. N., Computer research of deterministic and stochastic models "two competitors—two migration areas" taking into account the variability of parameters, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)61–73.DOI: 10.22363/2658-4670-2024-32-1-61-73.

Information about the authors

Vasilyeva, Irina I.—Assistant professor of Department of Mathematical Modeling, Computer Technologies and Information Security of Bunin Yelets State University (e-mail: irinavsl@yandex.ru, ORCID: https://orcid.org/0000-0002-4120-2595)

Demidova, Anastasia V.—Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia (RUDN University) (e-mail: demidova-av@rudn.ru, ORCID: https://orcid.org/0000-0003-1000-9650)

Druzhinina, Olga V.—Doctor of Physical and Mathematical Sciences, Chief Researche of Federal Research Center "Computer Science and Control" of Russian Academy of Sciences (e-mail: ovdruzh@mail.ru, ORCID: https://orcid.org/0000-0002-9242-9730)

Masina, Olga N.—Doctor of Physical and Mathematical Sciences, Professor of Department of Mathematical Modeling, Computer Technologies and Information Security of Bunin Yelets State University (e-mail: olga121@inbox.ru, ORCID: https://orcid.org/0000-0002-0934-7217)

УДК 519.6:004.942
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-61-73 EDN: GZIFWL

Компьютерное исследование детерминированных и стохастических моделей «два конкурента—два ареала миграции» с учетом вариативности параметров

И. И. Васильева[1], А. В. Демидова[2], О. В. Дружинина[3], О. Н. Масина[1]

[1] *Елецкий государственный университет им. И.А. Бунина, ул. Коммунаров, д. 28, Елец, 399770, Российская Федерация*
[2] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*
[3] *Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ул. Вавилова, д. 44, корп. 2, Москва, 119333, Российская Федерация*

Аннотация. Анализ траекторной динамики и решение задач оптимизации с применением компьютерных методов относится к актуальным направлениям исследования динамических популяционно-миграционных моделей. В настоящей работе изучаются четырехмерные динамические модели, описывающие процессы конкуренции и миграции в экосистемах. Во-первых, мы рассматриваем модификацию модели «два конкурента — два ареала миграции», в которой учитывается равномерная внутривидовая и межвидовая конкуренция в двух популяциях, а также неравномерная двунаправленная миграция обеих популяций. Во-вторых, мы рассматриваем модификацию модели «два конкурента — два ареала миграции», в которой внутривидовая конкуренция является равномерной, а межвидовая конкуренция и двунаправленная миграция являются неравномерными. Для указанных двух типов моделей исследование проводится с учетом вариативности параметров. Решены задачи поиска модельных параметров на основе реализации двух критериев оптимальности. Первый критерий оптимальности связан с выполнением такого условия сосуществования популяций, которое в математической форме представляет собой максимизацию интеграла от произведения функций, характеризующих плотности популяций. Второй критерий оптимальности включает в себя проверку предположения о существовании такого четырехмерного положительного вектора, который будет являться состоянием равновесия. Результатом работы алгоритмов, разработанных на основе первого и второго критериев оптимальности с применением метода дифференциальной эволюции, являются оптимальные наборы параметров изучаемых популяционно-миграционных моделей. Полученные наборы параметров используются для нахождения положительных состояний равновесия и для анализа траекторной динамики. С помощью метода построения самосогласованных одношаговых моделей и автоматизированной процедуры стохастизации выполнен переход к стохастическому случаю. Структурное описание и возможность анализа двух типов популяционно-миграционных стохастических моделей обеспечиваются получением уравнений Фоккера–Планка и уравнений в форме Ланжевена с соответствующими коэффициентами. Использованы алгоритмы генерирования траекторий винеровского процесса и многоточечных распределений и модификации метода Рунге–Кутты. Проведена серия вычислительных экспериментов с применением специализированного программного комплекса, возможности которого позволяют выполнять построение и анализ динамических моделей высокой размерности с учетом оценки влияния стохастики. Исследована траекторная динамика двух типов популяционно-миграционных моделей и выполнен сравнительный анализ результатов как в детерминированном, так и в стохастическом случае. Результаты могут найти применение в задачах моделирования и оптимизации динамических моделей естествознания.

Ключевые слова: одношаговые процессы, модели динамики популяций, стохастические дифференциальные уравнения, критерии оптимальности, дифференциальная эволюция, стохастизация, траекторная динамика, компьютерное моделирование, программный комплекс

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 74–85
http://journals.rudn.ru/miph

Research article

UDC 519.6:004.94
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-74-85
EDN: BUEBFE

Application of the Chebyshev collocation method to solve boundary value problems of heat conduction

Konstantin P. Lovetskiy[1], Stepan V. Sergeev[1], Dmitry S. Kulyabov[1,2], Leonid A. Sevastianov[1,2]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

(received: December 3, 2023; revised: January 12, 2024; accepted: January 25, 2024)

Abstract. For one-dimensional inhomogeneous (with respect to the spatial variable) linear parabolic equations, a combined approach is used, dividing the original problem into two subproblems. The first of them is an inhomogeneous one-dimensional Poisson problem with Dirichlet–Robin boundary conditions, the search for a solution of which is based on the Chebyshev collocation method. The method was developed based on previously published algorithms for solving ordinary differential equations, in which the solution is sought in the form of an expansion in Chebyshev polynomials of the 1st kind on Gauss–Lobatto grids, which allows the use of discrete orthogonality of polynomials. This approach turns out to be very economical and stable compared to traditional methods, which often lead to the solution of poorly defined systems of linear algebraic equations. In the described approach, the successful use of integration matrices allows complete elimination of the need to deal with ill-conditioned matrices.

The second, homogeneous problem of thermal conductivity is solved by the method of separation of variables. In this case, finding the expansion coefficients of the desired solution in the complete set of solutions to the corresponding Sturm–Liouville problem is reduced to calculating integrals of known functions. A simple technique for constructing Chebyshev interpolants of integrands allows to calculate the integrals by summing interpolation coefficients.

Key words and phrases: initial boundary problems, pseudo spectral collocation method, Chebyshev polynomials, Gauss–Lobatto sets, numerical stability, separation of variables

1. Introduction

Many important physics problems that involve two or more independent variables are solved using mathematical models that include partial differential equations, not limited to models based on ordinary differential equations. Models of this kind also include the description of heat propagation in solids using the heat equation with various boundary and initial conditions. An important method for solving partial differential equations known as the separation of variables method will be discussed below. Its essential feature is the reduction of the original partial differential equation to a system of simpler ordinary differential equations, which can be successfully solved based on given initial or boundary conditions.

The desired solution to a partial differential equation is expressed as an infinite series, which is the sum of solutions of individual ordinary differential equations. In many cases, it is convenient to represent the required solutions in the form of a series of sines, cosines, or polynomial functions, for example Chebyshev polynomials. This approach allows an effective use of the collocation method – a projection method for solving both integral and differential equations. Therefore, the first part of the present paper is devoted to a discussion of the Chebyshev interpolation method for solving the one-dimensional heat equation.

The Chebyshev collocation method has proven itself in solving a wide class of problems [1–5]. Particularly, in [6–9] its effectiveness was demonstrated in solving ODEs and problems of restoring functions from known first- and second-order derivatives. In Ref. [10] stable spectral methods for

© Lovetskiy K. P., Sergeev S. V., Kulyabov D. S., Sevastianov L. A., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

solving the Poisson equation with Dirichlet–Dirichlet, Dirichlet–Neumann and Neumann–Neumann boundary conditions were analyzed in detail. The present paper describes an algorithm for solving the 1-D Poisson equation with Dirichlet–Robin boundary conditions. The algorithm is based on the developed new method of spectral collocation and illustrates the effectiveness of the variable separation method in solving inhomogeneous heat conduction problems.

Spectral methods have proven themselves to be excellent in solving homogeneous boundary value problems for a wide class of partial differential equations using the method of separation of variables. In cases of inhomogeneous problems, methods for separating variables are not directly applicable. However, in this paper we show how the Chebyshev collocation method can be effectively applied in a two-stage solution scheme for a certain class of inhomogeneous boundary value problems for a 1-D linear parabolic equation.

2. Mathematical model of heat conduction

Thermal conductivity is the property of a material to conduct heat, which is assessed primarily from the point of view of Fourier's law of thermal conductivity. Heat conduction, also called diffusion, is the direct microscopic exchange of kinetic energy of particles (such as molecules) or quasiparticles (such as lattice waves) across a boundary between two systems. On a microscopic scale, thermal conduction occurs when hot, fast-moving, or vibrating atoms and molecules interact with neighboring atoms and molecules, transferring some of their energy (heat) to those neighboring particles. In other words, heat is transferred by conduction when neighboring atoms vibrate relative to each other or when electrons move from one atom to another.

When an object has a different temperature than another body or its surroundings, heat flows so that the body and surroundings reach the same temperature, at which point they are in thermal equilibrium. This spontaneous transfer of heat always occurs from a region of high temperature to another region of lower temperature, as postulated by the second law of thermodynamics. Thermodynamic and mechanical heat transfer are calculated using the heat transfer coefficient – the proportionality between heat flow and the thermodynamic driving force of heat flux. Heat flux is a quantitative vector representation of the movement of heat through a surface [11]. In an engineering context, the term "heat" is perceived as synonymous with thermal energy.

The heat conduction equation models diffusion processes [12], including thermal energy in solids, solutes in liquids, and biological populations. We will consider the heat conduction equation describing the temperature change in a one-dimensional rod of a finite length. Let us also consider several possible types of boundary conditions that can be used when modeling temperature changes.

A commonly used method for solving the heat conduction equation is the complete separation of variables method, which results in the solution of two ordinary differential equations generated by the separation of variables method. To solve one of the emerging subproblems, a uniform approach to solving the heat conduction equation for almost any of the frequently used (Dirichlet–Neumann–Robin) sets of boundary conditions is considered. A technique is proposed for constructing a general solution to the inhomogeneous Poisson equation – the heat conduction equation – regardless of the type of boundary conditions. Concretization of the solution – the determination of a pair of missing coefficients of expansion of the solution according to the selected polynomial basis occurs at the second stage, considering the specified (distinct types and combinations) boundary conditions.

3. Inhomogeneous boundary value problems

Let us consider the solution of an inhomogeneous initial-boundary value problem for a one-dimensional parabolic equation, including a time-independent inhomogeneous part of the equation and time-independent boundary conditions.

Two-sided Dirichlet–Dirichlet conditions

$$\begin{gathered} k\frac{\partial^2 u}{\partial x^2}(x,t) - \frac{\partial u}{\partial t}(x,t) = -F(x), \quad 0 < x < L, \quad t > 0, \\ u(0,t) = u_0, \quad u(L,t) = u_1, \quad t > 0, \\ u(x,0) = f(x), \quad 0 < x < L. \end{gathered} \tag{1}$$

The boundary conditions of problem (1) mean that the left and right edges of the rod have different constant temperatures due to ideal contact with heat baths having respectively temperatures u_0 and u_1.

Recall that a boundary value problem (BVP) is called inhomogeneous if either the partial differential equation or the boundary conditions are inhomogeneous. The well-known method of separating variables is not applicable to such inhomogeneous boundary value problems. However, in some cases it is possible to change the variables in such a way that the inhomogeneous boundary value problem transforms into two problems. One of which is a relatively simple inhomogeneous BVP for an ordinary differential equation (ODE), and the other is a homogeneous BVP for a partial differential equation (PDE).

Assume a function $F(x)$ to describe the intensity of heat generation inside the rod, u_0 and u_1 being constant. Replacing the unknown $u(x,t)$ with a new variable by means of substitution $u(x,t) = v(x,t) + \psi(x)$, we reduce the solution of problem (1) to a subsequent solution of two subproblems:

Problem A, an inhomogeneous problem with two-sided Dirichlet–Dirichlet conditions.

$$k\psi''(x) + F(x) = 0, \quad \psi(0) = u_0, \quad \psi(L) = u_1. \tag{2}$$

Problem B, a homogeneous boundary value problem for a partial differential equation.

$$\begin{cases} \dfrac{\partial^2 v}{\partial x^2}(x,t) = \dfrac{\partial v}{\partial t}(x,t), \\ v(0,t) = 0, \quad v(L,t) = 0, \\ v(x,0) = f(x) - \psi(x). \end{cases} \tag{3}$$

It is important to note that problem A is *a simple one-dimensional Poisson problem* — an ordinary differential equation of the second order with given inhomogeneous boundary value Dirichlet conditions. At the same time, this problem is a *problem of restoring a function from its known second-order derivative* $F(x)$ with two additional conditions, in this case, the Dirichlet boundary conditions.

The second auxiliary subproblem B is a homogeneous BVP that can be solved based on the traditional separation of variables method. The solution to the original problem (1) will be the "sum" of the solutions to problems A and B.

Let us illustrate the considered approach to the analytical solution of the inhomogeneous boundary value problem (1) using a pair of examples.

Example 1. Assume that $F(x) = r = const > 0$. It is required to solve the problem (1) under the following boundary and initial conditions for $L = 1$:

$$u(0,t) = 0, \quad u(1,t) = u_0, \quad t > 0,$$
$$u(x,0) = f(x), \quad 0 < x < 1.$$

In our example, both the partial differential equation and the boundary condition at point $x = 1$ are inhomogeneous. Let us perform the change of variables $u(x,t) = v(x,t) + \psi(x)$, then

$$\frac{\partial^2 u}{\partial x^2}(x,t) = \frac{\partial^2 v}{\partial x^2}(x,t) + \psi''(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x,t) = \frac{\partial v}{\partial t}(x,t).$$

We substitute these expressions into Eq. (1), so that the equation takes the form

$$k\frac{\partial^2 v}{\partial x^2}(x,t) + k\psi''(x) + r = \frac{\partial v}{\partial t}(x,t).$$

Let the function ψ satisfy the equation

$$k\psi''(x) + r = 0 \quad \text{or} \quad \psi''(x) = -\frac{r}{k}, \tag{4}$$

with the boundary conditions

$$\psi(0) = 0 \quad \text{and} \quad \psi(1) = u_0. \tag{5}$$

Then for $v(x,t)$ we obtain a problem of solving the homogeneous parabolic equation

$$k\frac{\partial^2 v}{\partial x^2}(x,t) - \frac{\partial v}{\partial t}(x,t) = 0, \tag{6}$$

with the boundary conditions

$$v(0,t) = 0 \quad \text{and} \quad v(1,t) = 0 \tag{7}$$

and the initial condition

$$v(x,0) = f(x) - \psi(x). \tag{8}$$

Equation (4) is integrated two times, as a result of which we arrive at its solution in the general form

$$\psi(x) = -\frac{r}{2k}x^2 + c_1 x + c_2. \tag{9}$$

Then to determine the constants c_1 and c_2, we use the boundary conditions

$$\psi(0) = 0 \quad \text{and} \quad \psi(1) = u_0,$$

distributed over the summands of the sought solution. Substituting these values into the general solution (9), we calculate the values of constants c_1 and c_2:

$$c_2 = 0, \quad c_1 = \frac{r}{2k} + u_0.$$

Therefore, the final partial solution of the Dirichlet–Dirichlet problem (4)–(5) for the simplest Poisson equation (4) has the form:

$$\psi(x) = -\frac{r}{2k}x^2 + \left(\frac{r}{2k} + u_0\right)x. \tag{10}$$

From the initial condition $u(x,0) = v(x,0) + \psi(x)$ it follows that $v(x,0) = u(x,0) - \psi(x) = f(x) - \psi(x)$. Therefore, to determine $v(x,t)$, we solve a new boundary value problem

$$\begin{gathered} k\frac{\partial^2 v}{\partial x^2}(x,t) = \frac{\partial v}{\partial t}(x,t), \quad 0 < x < 1, \quad t > 0, \\ v(0,t) = 0, \quad v(1,t) = 0, \quad t > 0, \\ v(x,0) = f(x) + \frac{r}{2k}x^2 - \left(\frac{r}{2k} + u_0\right)x, \quad 0 < x < 1, \end{gathered} \tag{11}$$

using the method of separation of variables. This method yields a solution to this problem in the form

$$v(x,t) = \sum_{n=1}^{\infty} A_n e^{-kn^2\pi^2 t} \sin n\pi x, \tag{12}$$

where the coefficients A_n are calculated by the formula

$$A_n = 2\int_0^1 \left[f(x) + \frac{r}{2k}x^2 - \left(\frac{r}{2k} + u_0\right)x\right] \sin n\pi x dx. \tag{13}$$

As a result, the solution to the original problem (1) is obtained by summing the solutions $\psi(x)$ and $v(x,t)$ of the homogeneous boundary value problem (4)–(5) and the boundary value problem (11)

$$u(x,t) = -\frac{r}{2k}x^2 + \left(\frac{r}{2k} + u_0\right)x + \sum_{n=1}^{\infty} A_n e^{-kn^2\pi^2 t} \sin n\pi x. \tag{14}$$

Note that $u(x,t) \to \psi(x)$ as $t \to \infty$ in expression (14). That is why the function $\psi(x)$ as a part of the solution of the heat conduction equation that does not change with time is called a steady solution. Whereas the function $v(x,t) \to 0$ as $t \to \infty$ is therefore called a transition solution.

The determination of the coefficients A_n according to formula (13) can be implemented using the integration technique based on the interpolation of the integrand with Chebyshev polynomials of the first kind on Gauss–Lobatto grids [13].

Example 2. In the next example, we consider an approach to solving an inhomogeneous boundary value problem when the Dirichlet condition is specified at the left end, i.e. the left end is in perfect contact with the heat bath (welded or screwed to a massive holder having a constant temperature u_0), and at the right end of the interval the Robin condition is imposed, i.e. this end exchanges heat with the environment at temperature u_m (hangs freely in the environment). Robin boundary conditions arise, e.g., when the ends are immersed in some liquid or gaseous medium. The initial temperature distribution $f(x)$ along the rod length $0 < x < L$ is also assumed to be given.

In accordance with the decomposition method, it is proposed to represent the solution to this inhomogeneous boundary value problem under given boundary and initial conditions

$$\begin{gathered} k\frac{\partial^2 u}{\partial x^2}(x,t) - \frac{\partial u}{\partial t}(x,t) = -F(x), \quad 0 < x < L, \quad t > 0, \\ u(0,t) = u_0, \quad u_0 = const, \\ \left.\frac{\partial u}{\partial x}\right|_{x=L} = -h(u(L,t) - u_m), \quad h > 0 \text{ and } u_m = const, \\ u(x,0) = f(x), \qquad 0 < x < L, \end{gathered} \tag{15}$$

as a combination of two terms

$$u(x,t) = v(x,t) + \psi(x), \tag{16}$$

each of them being a solution to a separate boundary value problem, respectively:

Problem A2. The function of the spatial variable $\psi(x)$ is a solution to an inhomogeneous ordinary differential equation with the Dirichlet boundary condition at the left end of the interval and the Robin boundary condition at the right end, namely:

$$\begin{gathered} k\frac{\partial^2 \psi(x)}{\partial x^2}(x) + F(x) = 0, \quad 0 < x < L, \quad t > 0, \\ \psi(0) = u_0, \quad u_0 = const, \\ \left.\frac{\partial \psi}{\partial x}\right|_{x=L} + h\psi(L) = hu_1, \quad h > 0 \text{ and } u_1 = const. \end{gathered} \tag{17}$$

Problem B2. The second term, the function of two variables $v(x,t)$, is a solution to a homogeneous boundary value problem with zero Dirichlet–Robin boundary conditions and given initial condition:

$$\begin{gathered} k\frac{\partial^2 v}{\partial x^2}(x,t) - \frac{\partial v}{\partial t}(x,t) = 0, \quad 0 < x < L, \quad t > 0, \\ v(0,t) = 0, \\ \left.\frac{\partial v}{\partial x}\right|_{x=L} + hv(L,t) = 0, \quad h > 0, \\ v(x,0) = f(x) - \psi(x), \quad 0 < x < L. \end{gathered} \tag{18}$$

Let us consider sequentially the methods for solving each of these problems.

4. Solving Problem A2 by the Chebyshev collocation method

By analogy with the method of approximate solution of the Poisson problem with various boundary conditions, such as the Dirichlet conditions on both ends of the interval $u(a) = \alpha$, $u(b) = \beta$, Neumann–Dirichlet condition $u'(a) = \alpha$, $u(b) = \beta$, or Dirichlet–Neumann condition $u(a) = \alpha$, $u'(b) = \beta$ thoroughly studied in Ref. [13], let us consider the solution of problem (17) based on the Chebyshev collocation method and, therefore, polynomial interpolation of the solution.

The polynomial interpolation based on using the basis of Chebyshev polynomials, leads to a necessity to formulate the problem in the interval $[-1, 1]$ instead of the initial interval $[a, b]$. In

the new interval, the polynomial interpolant of the function $y(x)$ is specified in the form of a series expansion $y(x) \sim \sum_{k=0}^{n} c_k T_k(x)$ in Chebyshev polynomials of the first kind $T_k(x)$ with the domain of definition on the segment $x \in [-1, 1]$.

The transition $x \Rightarrow t$ to new arguments, from $x \in [a, b]$ to $t \in [-1, 1]$, is implemented using a linear transformation

$$t = \frac{2x - (b + a)}{b - a}$$

and, if necessary, the inverse transformation

$$x = \frac{t(b - a) + (b + a)}{2}.$$

In this case, the values of the function, the integrals and derivatives are recalculated using the formulas

$$f(x) \Leftrightarrow f\left(\frac{b-a}{2}t + \frac{b+a}{2}\right), \quad x \in [a, b], \quad t \in [-1, 1],$$

$$\int_a^b f(x)dx = \int_{-1}^{1} f\left(\frac{b-a}{2}t + \frac{b+a}{2}\right)\frac{dx}{dt}dt, \quad x \in [a, b], \quad t \in [-1, 1],$$

$$\frac{dx}{dt} = \frac{b-a}{2},$$

$$f'_x(x) \Leftrightarrow \frac{df}{dt}(t) \Big/ \frac{dx}{dt}, \quad x \in [a, b], \quad t \in [-1, 1],$$

and [14] the upper estimate of the interpolation error has the form

$$\left| y(x) - \sum_{k=0}^{n} c_k T_k(x) \right| \le \frac{1}{2^n (n+1)!} \left| \frac{b-a}{2} \right|^{n+1} \max_{\xi \in [a,b]} \left| y^{n+1}(\xi) \right|.$$

Thus, Chebyshev interpolation provides an almost optimal approximation in the sense of the L_∞ norm and almost optimal to the L_2 norm. In addition, the use of Gauss–Lobatto nodes as interpolation nodes leads to optimal integration formulas.

As in Ref. [6], the method for approximate solution of problem (17) for a second-order ODE consists of sequential solution of several subproblems.

1. calculation of spectral coefficients of polynomial Chebyshev interpolation of the second derivative of the solution – the function of the right-hand side (17) on the Gaussian–Lobatto grid – interpolation of $F(x)$ in the basis of Chebyshev polynomials of the first kind;
2. calculation of those coefficients of the desired solution (except for the first two) that are determined from the differential conditions of the problem (allowing the solution to satisfy the differential conditions) – multiplication of the inverse (with respect to the matrix of spectral Chebyshev differentiation) matrix by the vector of interpolation coefficients;
3. additional determination of the first few coefficients of the solution based on boundary (or other independent additional) conditions.

Let us represent the approximate solution in the form of a finite series of orthogonal Chebyshev polynomials

$$p(x) = \sum_{k=0}^{n} c_k T_k(x), \quad x \in [-1, 1]. \tag{19}$$

Let us differentiate function (19) twice. The expression for the second derivative has the form

$$p''(x) = \sum_{k=0}^{n} c_k T''_k(x) = \sum_{k=0}^{n} b_k T_k(x), \quad x \in [-1, 1]. \tag{20}$$

Using recurrence relations satisfied by Chebyshev polynomials of the first kind and their derivatives [5], [15] and equating the coefficients for identical polynomials in (20), we arrive [5]

at the following dependence of the coefficients c_i, $i = 2, 3, \dots, n$, b_k, $k = 0, \dots, n$:

$$\mathbf{D}^{+}\mathbf{D}^{+}\mathbf{b} = \mathbf{c} \tag{21}$$

where $\mathbf{D}^{+}$ is a generalized inverse matrix with respect to the Chebyshev differentiation matrix in the spectral space [15], [16].

$$\mathbf{D}^{+}\mathbf{D}^{+}\mathbf{b} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \vdots & 0 & 0 & 0 & 0 \\ 1 & 0 & -\frac{1}{2} & 0 & 0 & \vdots & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & -\frac{1}{4} & 0 & \vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 & -\frac{1}{6} & \vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{8} & 0 & \vdots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \ddots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 & \frac{-1/2}{(n-3)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \vdots & \frac{1/2}{(n-2)} & 0 & \frac{-1/2}{(n-2)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 & \frac{1/2}{(n-1)} & 0 & \frac{-1/2}{(n-1)} \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 & 0 & \frac{1/2}{n} & 0 \end{bmatrix}^{2} \times \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ \vdots \\ b_{n-3} \\ b_{n-2} \\ b_{n-1} \\ b_n \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ \vdots \\ c_{n-3} \\ c_{n-2} \\ c_{n-1} \\ c_n \end{bmatrix} \tag{22}$$

Hence, the vector of coefficients $\{c_2, c_3, \dots, c_n\}$ is the result of double multiplication of a simple tridiagonal matrix $\mathbf{D}^{+}$ (inverse of the differentiation matrix) by vector $\{b_0, b_1, \dots, b_n\}$.

At the third stage of solving the problem, the first two coefficients of the expansion of the desired solution in Chebyshev polynomials are determined.

5. Dirichlet–Robin boundary conditions

For a one-dimensional problem considered on the interval $[-1, 1]$, the Dirichlet–Robin conditions look as follows

$$\begin{aligned} \alpha p(-1) &= g(-1), \\ \beta p(1) + \gamma p'(1) &= g(1). \end{aligned} \tag{23}$$

Here α, β, γ are given constants. The sign in front of the term with the derivative at the right boundary point is positive, since the outer normal to the domain of definition at the right boundary point is directed to $= \infty$, i.e., in the positive direction.

Let us take into account that the derivatives of Chebyshev polynomials of the first kind are simply expressed in terms of polynomials of the second kind:

$$\frac{dT_n}{dx} = nU_{n-1},$$

and, in addition, the relations

$$T_n(-1) = (-1)^n, \; T_n(1) = 1, \; U_n(-1) = (-1)^n(n+1), \; U_n(1) = (n+1),$$

are valid. In this case, the system of equations for calculating the unknown expansion coefficients of the solution has the form

$$\begin{aligned} \alpha\left(c_0 - c_1 + \sum_{k=2}^{n} c_k(-1)^k\right) &= g(-1), \\ \beta\left(c_0 + c_1 + \sum_{k=2}^{n} c_k\right) + \gamma\left(c_1 + \sum_{k=2}^{n} c_k k^2\right) &= g(1). \end{aligned} \tag{24}$$

Let us introduce the notation

$$\begin{gathered}\alpha(c_0 - c_1) = g(-1) - \alpha\sum_{k=2}^{n} c_k(-1)^k \equiv G(-1),\\ \beta(c_0 + c_1) + \gamma c_1 = g(1) - \beta\sum_{k=2}^{n} c_k - \gamma\sum_{k=2}^{n} c_k k^2 \equiv G(1),\end{gathered} \tag{25}$$

or

$$\begin{gathered}\alpha(c_0 - c_1) = G(-1),\\ \beta c_0 + (\beta + \gamma)c_1 = G(1).\end{gathered} \tag{26}$$

Remove the brackets

$$\begin{gathered}\alpha c_0 - \alpha c_1 = G(-1),\\ \beta c_0 + \beta c_1 + \gamma c_1 = G(1)\end{gathered} \tag{27}$$

and multiply the first equation by β and the second equation by α:

$$\begin{gathered}\beta\alpha c_0 - \beta\alpha c_1 = \beta G(-1),\\ \beta\alpha c_0 + \beta\alpha c_1 + \gamma\alpha c_1 = \alpha G(1).\end{gathered} \tag{28}$$

To calculate the coefficient c_1, subtract the first equation from the second one:

$$2\beta\alpha c_1 + \gamma\alpha c_1 = \alpha G(1) - \beta G(-1), \tag{29}$$

from which it follows that

$$c_1 = \frac{\alpha G(1) - \beta G(-1)}{\alpha(2\beta + \gamma)}. \tag{30}$$

To calculate the coefficient c_0, we substitute the calculated value of c_1 into the first equation (27):

$$\begin{gathered}c_0 = \frac{G(-1)}{\alpha} + c_1 = \frac{G(-1)}{\alpha} + \frac{\alpha G(1) - \beta G(-1)}{\alpha(2\beta + \gamma)},\\ c_0 = \frac{(\beta + \gamma)G(-1) + \alpha G(1)}{\alpha(2\beta + \gamma)}.\end{gathered} \tag{31}$$

6. Solving the problem B2. Method of separation of variables

Recall the formulation of problem (18). It is required to find a solution of the homogeneous boundary value problem with zero Dirichlet–Robin conditions and a given initial condition:

$$\begin{gathered}k\frac{\partial^2 v}{\partial x^2}(x,t) - \frac{\partial v}{\partial t}(x,t) = 0, \quad 0 < x < L, \quad t > 0,\\ v(0,t) = 0,\\ \left.\frac{\partial v}{\partial x}\right|_{x=L} + hv(L,t) = 0, \quad h > 0,\\ v(x,0) = f(x) - \psi(x), \quad 0 < x < L.\end{gathered}$$

where k is the heat conductivity coefficient depending on the material properties.

The method of separation of variables yields a solution to this problem (see, e.g., [17]) in the form

$$v(x,t) = \sum_{n=1}^{\infty} A_n e^{-k\lambda_n t}\sin\sqrt{\lambda_n}Lx, \tag{32}$$

where the coefficients A_n are calculated as

$$A_n = 2\int_0^L [f(x) + \psi(x)] \sin\sqrt{\lambda_n} Lx dx, \tag{33}$$

where $\lambda_n > 0$ is the set of positive solutions of the equation $\sqrt{\lambda} = -h\tan\left(\sqrt{\lambda}L\right)$, $n \in \mathbb{N}$.

Finally, the solution of the initial problem 2 (15) is obtained by adding the solution $\psi(x)$ of the boundary value problem (17) and the solution $v(x,t)$ of the homogeneous boundary value problem (18):

$$u(x,t) = \psi(x) + \sum_{n=1}^{\infty} A_n e^{-k\lambda_n t} \sin\sqrt{\lambda_n} Lx. \tag{34}$$

The determination of coefficients A_n by formula (33) can be implemented using the technique of integration based on the interpolation of the integrand by the Chebyshev polynomials of the first kind on Gauss–Lobatto grids. In this case, the calculation of the coefficients A_n reduces to elementary summation of the weighted even coefficients of the interpolant.

7. Conclusion

Among the numerical algorithms for solving initial and boundary value problems for linear ODEs of the first and second order, there are many methods that use the initial approximation (boundary conditions) as the initially active condition that determines all further solution of the problem. These are methods such as Euler, Adams–Bashforth, Runge–Kutta, etc. [18]. Other methods, based on approximation of the solution using global functions [1–5], are based on the construction of systems of equations that simultaneously include both initial (boundary) conditions and conditions that specify the behavior of the derivatives of the desired solution.

The solution to the main inhomogeneous initial-boundary value problem for a one-dimensional parabolic equation is presented in the form of a sequential solution of several subproblems. As a method for solving one of the subproblems – an inhomogeneous ordinary differential equation with Dirichlet–Robin boundary conditions – it is proposed to use the stable and efficient spectral method of Chebyshev collocation.

Polynomial interpolation of the desired solution by Chebyshev polynomials is carried out in several stages. At the first stage, a general solution is identified, i.e., a set of solutions that satisfies the differential equation, but does not necessarily satisfy the initial (boundary) conditions. Considering the initial (boundary) conditions is carried out at the last stage of solving the problem and reduces to solving a linear equation with two unknown coefficients.

The search for a general solution to an inhomogeneous ODE reduces to multiplying the transposed matrix of values of Chebyshev functions on the Gauss–Lobatto grid by the vector of the function values that specifies the right-hand side of the original differential equation to determine the interpolation coefficients for the expansion of the solution derivative. Next, multiplying the bi-diagonal integration matrix [5], [15] by the vector of these coefficients leads to obtaining all the coefficients of the desired solution, except for first ones. At the final stage, the use of the initial (boundary) condition makes it possible to determine the first two coefficients of the polynomial expansion of the solution.

The solution of the second homogeneous subproblem is carried out by the traditional method of separation of variables. In this case, to calculate the coefficients of expansion of the solution according to the basis of the Sturm–Liouville problem, an effective and stable spectral method of Chebyshev collocation is used. Thus, the authors expand the scope of applicability of the developed 2-stage Chebyshev collocation method.

Author Contributions: Conceptualization, software: Konstantin P. Lovetskiy; software: Stepan V. Sergeev; formal analysis: Dmitry S. Kulyabov; theoretical proof of the consistency of Chebyshev collocation method in the case of complex potentials: Leonid A. Sevastianov. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the RUDN University Scientific Projects Grant System, project No 021934-0-000 (Konstantin P. Lovetskiy). This research was supported by the RUDN University Strategic Academic Leadership Program (Dmitry S. Kulyabov). The work of Leonid A. Sevastianov was supported by the Russian Science Foundation (grant No. 20-11-20257).

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Boyd, J. P. *Chebyshev and Fourier spectral methods* second (Dover Books on Mathematics, 2013).
2. Mason, J. C. & Handscomb, D. C. *Chebyshev polynomials* doi:10 . 1201 / 9781420036114 (Chapman and Hall/CRC Press, New York, 2002).
3. Greengard, L. Spectral integration and two-point boundary value problems. *SIAM Journal on Numerical Analysis* **28,** 1071–1080. doi:10.1137/0728057 (1991).
4. Shen, J., Tang, T. & Wang, L. *Spectral methods* doi:10.1007/978-3-540-71041-7 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011).
5. Fornberg, B. *A practical guide to pseudospectral methods* (Cambridge University Press, New York, 1996).
6. Sevastianov, L. A., Lovetskiy, K. P. & Kulyabov, D. S. *Multistage collocation pseudo-spectral method for the solution of the first order linear ODE* in *VIII International Conference on Information Technology and Nanotechnology (ITNT)* (2022), 1–6. doi:10.1109/ITNT55410.2022.9848731.
7. Sevastianov, L. A., Lovetskiy, K. P. & Kulyabov, D. S. A new approach to the formation of systems of linear algebraic equations for solving ordinary differential equations by the collocation method. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics* **23.** in Russian, 36–47. doi:10.18500/1816-9791-2023-23-1-36-47 (2023).
8. Lovetskiy, K. P., Kulyabov, D. S. & Hissein, W. Multistage pseudo-spectral method (method of collocations) for the approximate solution of an ordinary differential equation of the first order. *Discrete and Continuous Models and Applied Computational Science* **30,** 127–138. doi:10.22363/2658-4670-2022-30-2-127-138 (2022).
9. Sevastianov, L. A., Lovetskiy, K. P. & Kulyabov, D. S. *Numerical integrating of highly oscillating functions: effective stable algorithms in case of linear phase* 2021. doi:10.48550/arXiv.2104.03653.
10. Lovetskiy, K. P., Kulyabov, D. S., Sevastianov, L. A. & Sergeev, S. V. Chebyshev collocation method for solving second order ODEs using integration matrices. *Discrete and Continuous Models and Applied Computational Science* **31,** 150–163. doi:10.22363/2658-4670-2023-31-2-150-163 (2023).
11. Lienhard, J. H. I. & Lienhard, J. H. V. *A heat transfer textbook fifth edition* 2020.
12. Tikhonov, A. N. & Samarskii, A. A. *Equations of Mathematical Physics* (Nauka, M., 2004).
13. Sevastianov, L. A., Lovetskiy, K. P. & Kulyabov, D. S. A new approach to the formation of systems of linear algebraic equations for solving ordinary differential equations by the collocation method. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics* **23,** 36–47. doi:10.18500/1816-9791-2023-23-1-36-47 (2023).
14. Stewart, G. W. *Afternotes on numerical analysis* (Society for Industrial and Applied Mathematics, USA, 1996).
15. Amiraslani, A., Corless, R. M. & Gunasingam, M. Differentiation matrices for univariate polynomials. *Numerical Algorithms* **83,** 1–31. doi:10.1007/s11075-019-00668-z (2020).
16. Rezaei, F., Hadizadeh, M., Corless, R. & Amiraslani, A. Structural analysis of matrix integration operators in polynomial bases. *Banach Journal of Mathematical Analysis* **16,** 5. doi:10.1007/s43037-021-00156-4 (2022).
17. Boyce, W. E. & DiPrima, R. C. *Elementary differential equations and boundary value problems* 9th Edition (Wiley, New York, 2009).
18. Planitz, M. *et al. Numerical recipes: the art of scientific computing* 3rd Edition (Cambridge University Press, New York, 2007).

To cite: Lovetskiy K. P., Sergeev S. V., Kulyabov D. S., Sevastianov L. A., Application of the Chebyshev collocation method to solve boundary value problems of heat conduction, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)74–85.DOI: 10.22363/2658-4670-2024-32-1-74-85.

Information about the authors

Lovetskiy, Konstantin P.—Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Computational Mathematics and Artificial Intelligence of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: lovetskiy-kp@rudn.ru, phone: +7(495)952-25-72, ORCID: https://orcid.org/0000-0002-3645-1060)

Kulyabov, Dmitry S.—Professor, Doctor of Sciences in Physics and Mathematics, Professor at the Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru,

phone: +7(495)952-02-50, ORCID: https://orcid.org/0000-0002-0877-7063)

Sevastianov, Leonid A.—Professor, Doctor of Sciences in Physics and Mathematics, Professor at the Department of Computational Mathematics and Artificial Intelligence of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), Leading Researcher of Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research (e-mail: sevastianov-la@rudn.ru, phone: +7(495)952-25-72, ORCID: https://orcid.org/0000-0002-1856-4643)

Sergeev, Stepan V.—PhD student of Department of Computational Mathematics and Artificial Intelligence of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1142220124@rudn.ru, ORCID: https://orcid.org/0009-0004-1159-4745)

УДК 519.6:004.94
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-74-85 EDN: BUEBFE

Применение метода коллокации Чебышева для решения граничных задач теплопроводности

К. П. Ловецкий[1], С. В. Сергеев[1], Д. С. Кулябов[1,2], Л. А. Севастьянов[1,2]

[1] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация
[2] *Объединённый институт ядерных исследований,*
ул. Жолио-Кюри, д. 6, Дубна, Московская область, 141980, Российская Федерация

Аннотация. Для одномерных неоднородных (по пространственной переменной) линейных параболических уравнений используется комбинированный подход, разбивающий исходную задачу на две подзадачи. Первая из них – неоднородная одномерная задача Пуассона с граничными условиями Дирихле–Робена, поиск решения которой основан на методе чебышевской коллокации. Метод разработан на основе ранее опубликованных алгоритмов решения обыкновенных дифференциальных уравнений, в которых решение ищется в виде разложения по полиномам Чебышева 1-го рода на сетках Гаусса–Лобатто, что позволяет использовать дискретную ортогональность полиномов. Такой подход оказывается весьма экономичным и стабильным по сравнению с традиционными методами, приводящими к решению часто плохо определенных систем линейных алгебраических уравнений. В описываемом подходе удачное применение матриц интегрирования позволяет вообще избавиться от необходимости работы с плохо обусловленными матрицами.

Вторая, однородная задача теплопроводности решается методом разделения переменных. При этом отыскание коэффициентов разложения искомого решения по полному набору решений соответствующей задачи Штурма–Лиувилля сводится к вычислению интегралов от известных функций. Простая методика построения чебышевских интерполянтов подынтегральных функций позволяет вычислять интегралы суммированием интерполяционных коэффициентов.

Ключевые слова: начально-краевые задачи, псевдоспектральный метод коллокации, полиномы Чебышева, множества Гаусса–Лобатто, численная устойчивость, разделение переменных

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 86–98
http://journals.rudn.ru/miph

Research article

UDC 004.2, 004.7
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2024-32-1-86-98

EDN: CCHVBS

A new link activation policy for latency reduction in 5G integrated access and backhaul systems

Anna A. Zhivtsova, Vitaly A. Beschastnyy

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: February 22, 2024; revised: March 12, 2024; accepted: March 25, 2024)

Abstract. The blockage of the propagation path is one of the major challenges preventing the deployment of fifth-generation New Radio systems in the millimeter-wave band. To address this issue, the Integrated Access and Backhaul technology has been proposed as a cost-effective solution for increasing the density of access networks. These systems are designed with the goal of avoiding blockages, leaving the question of providing quality-of-service guarantees aside. However, the use of multi-hop transmission negatively impacts the end-to-end packet latency. In this work, motivated by the need for latency reduction, we design a new link activation policy for self-backhauled Integrated Access and Backhaul systems operating in half-duplex mode. The proposed approach utilizes dynamic queue prioritization based on the number of packets that can be transmitted within a single time slot, enabling more efficient use of resources. Our numerical results show that the proposed priority-based algorithm performs better than existing link scheduling methods for typical system parameter values.

Key words and phrases: 5G, IAB, millimeter wave, half-duplex, link scheduling, network control

1. Introduction

The digitalization of many areas of human activity relies upon a communication system capable of providing a wide range of services. The 5th generation (5G) mobile networks enable the provision of different services including Enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and Massive Machine-Type Communications (mMTC).

The services provided by 5G networks require improvements in various performance indicators. For example, eMBB needs to offer high throughput (up to 10 Gbps) and support high mobility devices (up to 500 km/h). URLLC requires delay reduction down to one millisecond. Finally, for mMTC services, the number of connected devices must be increased to up to 10 million per square kilometer, while also improving their energy efficiency [1].

In order to provision the required performance indicators in 5G, significant changes have been made to the architecture and operations of the 5G core (5GC) and radio access networks (RAN). For example, flexibility and adaptability in synchronization procedures, as well as the allocation and splitting of bands into subcarriers, have been increased. Additionally, modulation, coding, and error correction have been improved [2].

In addition to enhancing the RAN functionality, an important technical innovation of 5G is its substantially expanded frequency range. This allows for higher throughput by allocating vast bandwidth at high frequencies (greater than 24 GHz), while maintaining wide coverage through the utilization of lower frequencies. It is worth noting though that communications in the new high-frequency spectrum suffer from high propagation losses and require significant capital expenditures for upgrading and expanding network hardware infrastructure. In particular, as the coverage area of a base station is reduced due to propagation issues, network densification is necessary, which involves increasing the number of access points (APs) per unit area.

One way to densify 5G networks is to utilize the Integrated Access and Backhaul (IAB) technology. It employs relay nodes that are not wired connected to the core network as additional APs. The

© Zhivtsova A. A., Beschastnyy V. A., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

interference issues in the resulting multi-hop wireless network call for the half-duplex transmission, meaning that no network node can receive and transmit data at the same time. In turn, a half-duplex system requires an efficient link activation policy, which determines over which links data can be transmitted at any given time.

In this paper, we aim to design a new link activation policy for 5G IAB networks that allows for packet delay reduction and throughput maximization and can be employed in both centralized and distributed manners. The rest of the paper is structured as follows. First, in Section 2, we discuss the IAB technology and briefly overview the related work. Then, we formalize the model of an IAB network in Section 3 and propose a new link activation policy in Section 4. Next, in Section 5, we obtain realistic simulation parameters and numerically evaluate performance of the proposed policy in comparison with well-known link activation algorithms. Conclusions are drawn in the last section.

2. Background and related work

To minimize capital expenditures in deploying dense 5G networks, the 3GPP (3rd Generation Partnership Project) standardization body has proposed the IAB [3]. IAB allows to use relay nodes that are not directly connected to the core network as relaying APs. As depicted in figure 1, there are two types of APs in an IAB network: an IAB donor directly connected to the core network by a wired link, and one or more IAB nodes which transmit traffic from or to the core network through the IAB donor. The wireless links in the IAB network are divided into two types: access links between an AP and a User Equipment (UE), and backhaul links between APs. Both types of links use a shared time-frequency resource, as the name of the technology implies.

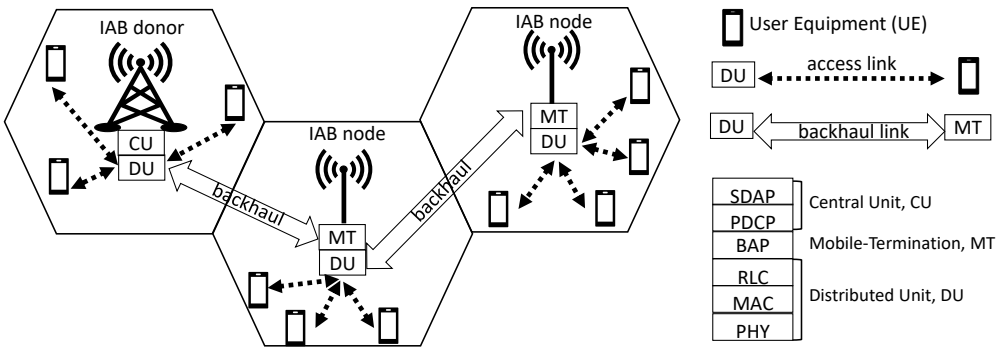


Figure 1. The main components of the IAB network

The IAB technology is based on the distributed architecture of 5G networks. This architecture separates the layers of the data transfer protocol stack between central and distributed units, as shown in figure 1. A Distributed Unit (DU) implements Radio Link Control (RLC), Medium Access Control (MAC), and Physical Layer (PHY). The DU is present at each AP and ensures the establishment, maintenance, and termination of radio connections. The Central Unit (CU) implements Service Data Adaptation Protocol (SDP) and Packet Data Convergence Protocol (PDCP). The CU is only present in the donor and provides connection with the core network. Each IAB node contains a Mobile Termination (MT). This component supports the Backhaul Adaptation Protocol (BAP), which forwards data streams that travel through multiple IAB nodes to and from the IAB donor.

In the first IAB standardization document [3] released in 2018, the IAB network was defined as a multi-hop wireless network with static APs and the ability of path selection. Also, the standard provides a list of possible options for implementation. For example, either in-band or out-of-band backhauling can be used. The use of time, frequency, or spatial multiplexing is permitted, as is end-to-end or hop-by-hop automatic repeat request (ARQ). The resource allocation is not fully determined by the standard, and has been explored in various research projects. For an extensive review, see [4].

As previously mentioned, the IAB standard allows the simultaneous operation of access and backhaul links within the same frequency band. This reduces the downtime for the radio resources,

but also increases interference [5]. Each transmitter interferes with all other active receivers in the network, except the one it is communicating with. The high-frequency 5G spectrum allows for directional transmission, reducing interference in many channels. Nevertheless, interference that occurs during simultaneous reception and transmission remains significant [6].

To eliminate interference caused by simultaneous reception and transmission in the IAB network, the standard [3] recommends using the half-duplex mode. This mode helps to reduce interference by limiting the number of channels on which transmission occurs at any given time. More precisely, half-duplex mode prevents any AP in the IAB network from receiving and transmitting data simultaneously. Although the half-duplex mode limits the network throughput and increases delays, it is an effective and simple way to reduce interference.

To efficiently implement half-duplex, it is essential to schedule transmission over links. This can be done by dividing time into slots and marking each link with 1 (ON) if it is allowed to transmit in the slot and 0 (OFF) otherwise [7–10]. Such link scheduling permits to ensure that the half-duplex constraints are met and to optimize selected performance metrics. For example, in [9] the link scheduling algorithm maximizes minimal user throughput, in [10] it optimizes the sum of user throughputs, and in [7, 8, 11] it targets some convex function of user throughput (such as the sum of logarithms).

In [7–11] the link scheduling is performed by solving an optimization problem with the objective function of throughput. On the other hand, constructing a queuing model of the studied network allows to evaluate and optimize the delay [12, 13], as well as to prove the stability of the network under some scheduling algorithms with any acceptable rates of incoming traffic [11, 14]. This approach was used to derive a number of link scheduling algorithms for general multi-hop wireless networks with interference, and in particular several throughput optimal greedy dynamic algorithms for efficient centralized control of multi-hop networks, which choose a transmission mode based on the current system state via argmin or argmax. *Backpressure* [15] is the most recognized throughput-oriented algorithm for network control and can be utilized for link scheduling, routing or flow control problems [11, 16–18]. While *backpressure* handles queue lengths, such algorithms as the *largest weighted delay first* [19, 20], *oldest cell first* [21] and *delay-based backpressure* [22] use packet delays to specify the system state. The latter is the delay-based version of *backpressure* and allows to reduce the maximum packet delay in the original *backpressure* algorithm. The *α-algorithm* [23] is a modification of *backpressure* aimed at reducing the total delay while remaining optimal in throughput. It uses a constant $\alpha \geqslant 1$ as a per-component power in the *backpressure* algorithm to point up the longest queues. The *αβ-algorithm* [24] algorithm aims to reduce the probability of buffer overflow and thus to provide shorter queue lengths and smaller delays. The activation of a link in this algorithm depends on the lengths of all queues that packets have passed before this link and will pass after.

The introduction of the IAB technology has revived interest in existing link scheduling methods for multi-hop wireless networks, however they should be analysed and modified by taking into account the specifics of IAB and the needs of 5G services. The present paper provides a step in this direction.

3. Model formalization

We consider a half-duplex IAB network where transmission takes place over either access or backhaul links at any given time. Furthermore, a link may be activated in either the uplink or downlink direction. We assume that the throughput of each link is constant. Additionally, we assume that all data packets traversing the network have the same size, and thus, in what follows, a packet is used as a unit of data.

We represent the considered IAB network as a directed graph consisting of four vertices as shown in figure 2. The vertices represent the IAB donor (the circle), the IAB node (the square), the UEs connected to the IAB donor (modeled as a single vertex and depicted by the left triangle), and the UEs connected to the IAB node (also modeled by a single vertex, the right triangle). The edges correspond to the communication links for direct wireless transmission. In what follows we use the terms vertices and nodes, as well as edges and links interchangeably.

The links are divided into uplink, which carry packets from UEs to the IAB donor, and downlink, carrying data from the IAB donor to UEs. Furthermore, a link can be either backhaul, responsible for data transmission between the IAB donor and node, or access, connecting UE nodes to their access points, see figure 2.

Each link of the IAB network graph can be viewed as a server accompanied by a queue of unlimited size where packets awaiting transmission are stored. The system can thus be represented by a queuing network depicted in figure 3. It consists of $I = 6$ service nodes (or queues) with queues 1, 3 and 5

corresponding to the downlink links, and 2, 4 and 6 – to the uplink. Queues 1 and 2 are coupled with backhaul links, and the rest – with the access links. Packets departing queue 1 enter queue 3, and packets departing queue 4 enter queue 2, which describes the two-hop transmission. Packets departing queues 2, 3, 5, 6 leave the system. The set of all queues is denoted by $\mathcal{I}$.

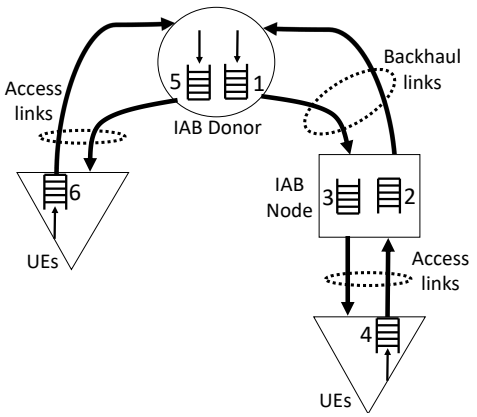


Figure 2. The considered IAB network as a directed graph

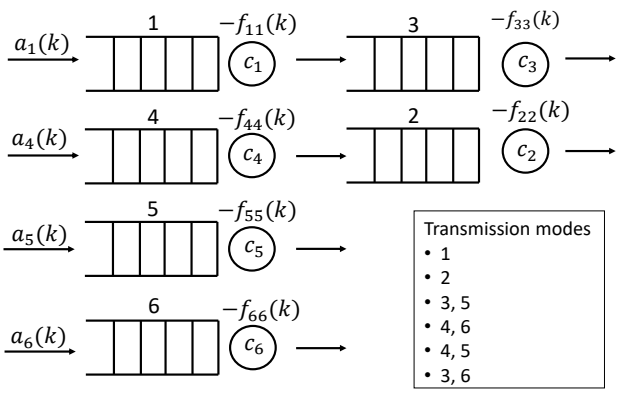


Figure 3. The queuing network corresponding to the modeled IAB network

The system is considered in discrete time indexed by $k = 0, 1, 2, \ldots$. We denote by $a_i(k)$ the number of packets exogenously arriving to the i-th queue in time slot $k \geqslant 0$. We have $a_2(k) = a_3(k) = 0$ for all $k \geqslant 0$, because packets entering queues 2 and 3 are first serviced in stations 4 and 1, respectively. For each of the remaining queues $i \in \mathcal{I}_0 = \{1, 4, 5, 6\}$ it is assumed that $a_i(k)$, $k \geqslant 0$, are independent and identically distributed (i.i.d.) random variables with finite first and second moments. We denote the row vector of arrivals in time slot k by $\mathbf{a}(k) = (a_i(k))_{i\in\mathcal{I}}$. The arrival rate to queue i is equal to the expectation of $a_i(k)$ and denoted by $\lambda_i = \mathbb{E}a_i(k)$.

We say that a packet is served when it is transmitted over a link, and that a queue is served (or active) when the packets it holds are serviced. The service duration is assumed exactly one time slot. Packets are served in batches. The maximum size of a batch that can be served in queue i in one time slot is fixed and denoted by $c_i \in \mathbb{N}$. The column vector $\mathbf{c} = (c_i)_{i\in\mathcal{I}}$ is called the link capacity vector. If the number of packets in an active queue i is fewer than c_i, then all packets in the queue are served in the time slot, otherwise packets are taken for service according to the discipline First Come First Served (FCFS), i.e., in the order of arrival.

The IAB specifics impose constraints on simultaneous activation of queues. By a transmission mode we understand a feasible combination of simultaneously active queues. Queues i and j such that $i \in \{1, 2\}$ and $j \in \{3, 4, 5, 6\}$ and the queues 1 and 2, 3 and 4, and 5 and 6, pairwise cannot be active in the same time slot due to the half-duplex constraints. Moreover, to maximize resource utilization, we do not consider transmission modes that activate fewer queues than allowed by the constraints. This results in the following transmission modes for the system: {1}, {2}, {3, 5}, {4, 6}, {4, 5}, {3, 6}. We denote the set of these transmission modes by Θ and assume they are indexed by l=1, ..., L, $L = |\Theta| = 6$, in the above order.

To specify the connectivity corresponding to the transmission modes listed above, we define, for each $\theta \in \Theta$, an $I \times I$ matrix $\mathbf{F}$ with elements

$$f_{i,j} = \begin{cases} 1, & \text{if } i \in \theta, \quad (i, j) \in \{(1, 3), (4, 2)\}, \\ -1, & \text{if } i \in \theta, \quad j = i, \quad i \in \mathcal{I}, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

We denote the set of such matrices by $\mathcal{F}$ and let them be ordered and indexed as in Θ. Since there is a one-to-one correspondence between the sets $\mathcal{F}$ and Θ, in what follows, we will specify a transmission mode by either $\theta \in \Theta$ or $\mathbf{F} \in \mathcal{F}$ interchangeably.

We assume that in each time slot only one transmission mode can be applied by a controller. Thus, in each time slot k, the system operates according to $\mathbf{F}(k) \in \mathcal{F}$.

Figure 4 shows the timing of events in a time slot $k \geqslant 0$, by which we understand the time $[t_k,\ t_{k+1})$, where $\Delta = t_{k+1} - t_k$ is a constant time slot duration. At the beginning of time slot k the system assumes a transmission mode $\mathbf{F}(k)$ for the time slot. Then, the queues activated by $\mathbf{F}(k)$ are served. Served packets from queues 2, 3, 5 and 6 depart the system, and served packets from queues 1 and 4 move, respectively, to queues 3 and 2. Then, before the end of time slot k, new packets arrive into the system and join queues 1, 4, 5, 6. Thus, no packet can join and depart a queue in one time slot.

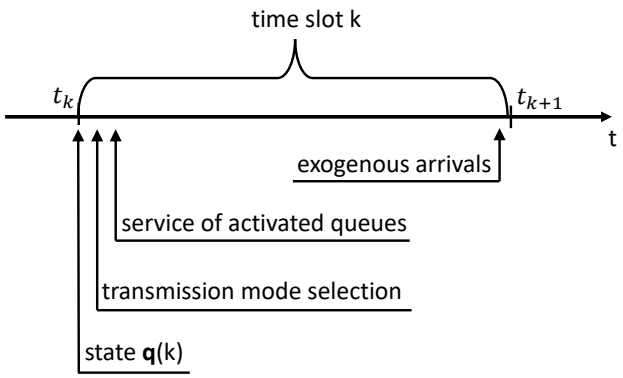


Figure 4. Timing of events in the considered model

Denote by $\mathbf{q}(k) = (q_i(k))_{i \in \mathcal{I}}$ a row vector whose entry $q_i(k)$ is the number of packets in queue i at the beginning of time slot k. Let $\mathbf{q}(0) = \mathbf{0}$ be a zero row vector of length I. Let a row vector $\mathbf{s}(k) = (s_i(k))_{i \in \mathcal{I}}$ with entries

$$s_i(k) = \min(c_i, q_i(k)), \quad i \in \mathcal{I}, \tag{2}$$

represent the number of packets that will be served in queue i in time slot k if the queue is active in this slot. Now, vector $\mathbf{q}(k+1)$ defining the system state in time slot $k+1$ relates to $\mathbf{q}(k)$ and the transmission mode $\mathbf{F}(k)$ as

$$\mathbf{q}(k+1) = \mathbf{q}(k) + \mathbf{s}(k)\mathbf{F}(k) + \mathbf{a}(k), \quad k \geqslant 0. \tag{3}$$

In what follows, we also assume that the transmission mode $\mathbf{F}(k) \in \mathcal{F}$ chosen in time slot k depends only on the system state at time k given by $\mathbf{q}(k)$. A function $\pi(\mathbf{q}(k)) = \mathbf{F}(k)$ will be referred to as the link scheduling (or control) policy.

The capacity region of the system is defined as the set of all combinations of arrival rates $(\lambda_1, \lambda_4, \lambda_5, \lambda_6)$ such that there exists a control policy that provides a finite time-average number of packets in the system operating with these rates as $k \to \infty$. Having a finite average number of packets in all queues is considered as a network stability criterion. A control policy providing network stability for all sets of arrival rates in the capacity region is called throughput optimal [25].

For the considered model, the network capacity region can be obtained as follows. Let p_l be the fraction of time when transmission mode θ_l, $l = 1, \dots, L$, is applied given some control policy π. Note that $\sum_{l=1}^{L} p_l = 1$ as only one transmission mode can be applied in each time slot. Now, the condition for the system to have a finite average number the packets can be written as

$$\begin{aligned} \lambda_6 &\leqslant c_6(p_4 + p_6), \quad \lambda_4 \leqslant c_4(p_4 + p_5), \\ \lambda_5 &\leqslant c_5(p_3 + p_5), \quad \lambda_1 \leqslant c_3(p_3 + p_6). \end{aligned} \tag{4}$$

By dividing each inequality by the capacity of the corresponding link and then summing up, we obtain the capacity region of the system in the form

$$\frac{\lambda_6}{c_6} + \frac{\lambda_4}{c_4} + \frac{\lambda_5}{c_5} + \frac{\lambda_1}{c_3} + \frac{2\lambda_1}{c_1} + \frac{2\lambda_4}{c_2} \leqslant 2. \tag{5}$$

As a key performance indicators we consider the average end-to-end delay $\overline{D}$ and the 99th percentile of the end-to-end delay probability distribution, denoted by P_{99}. The end-to-end delay is defined for each packet that has departed the system as its sojourn time in the system. We also consider such important aspects of every control policy as its throughput optimality and the control-induced overhead.

4. Centralized and distributed priority-based link scheduling

The idea behind the proposed priority-based link scheduling algorithm is as follows. To chose the transmission mode, we first prioritize the transmission modes according to whether the activated thereby queues hold more packets than can be served in one time slot. Then, among the transmission modes with the highest priority, we choose the one providing transmission of the greatest number of packets. This approach is similar to the *P-TREE* algorithm [26], which is a low-complexity scheduling algorithm designed for the multi-hop tree-shaped networks with only uplink traffic.

Recall, that for $\mathbf{F} \in \mathcal{F}$ a diagonal entry $f_{i,i}$ is -1 or 0 depending on whether queue i is activated or not in the transmission mode specified by $\mathbf{F}$. In the proposed *priority-based* algorithm, in each time slot k we obtain a set of priority transmission modes $\mathcal{F}^*(k)$ by the following procedure consisting of three steps:

1. Let the priority set $\mathcal{F}^*(k)$ include all transmission modes for which the maximum possible number of packets is served in all active queues, i.e., let

$$\mathcal{F}^*(k) := \{\mathbf{F} \in \mathcal{F} : s_i(k) f_{i,i} = c_i f_{i,i} \ \forall i \in \mathcal{I}\}. \tag{6}$$

2. If after Step 1 the set $\mathcal{F}^*(k)$ is empty, then let it include all transmission modes for which the maximum possible number of packets is served in at least one queue, i.e., let

$$\mathcal{F}^*(k) := \{\mathbf{F} \in \mathcal{F} : f_{i,i} = -1, \ s_i(k) = c_i \text{ for some } i\}. \tag{7}$$

3. If the set $\mathcal{F}^*(k)$ is still empty, then let $\mathcal{F}^*(k) := \mathcal{F}$.

Now, among the transmission modes of set $\mathcal{F}^*(k)$ we choose the one that results in serving the most packets in time slot k. Let $\mathrm{diag}(\mathbf{F}) = (f_{i,i})_{i \in \mathcal{I}}$ denote the column vector of diagonal elements of matrix $\mathbf{F}$. Since the number of packets served in time slot k under transmission mode $\mathbf{F}$ is $-\mathbf{s}(k)\mathrm{diag}(\mathbf{F})$,

the sought transmission mode is given by

$$\pi_{pb}(\mathbf{q}(k)) = \underset{\mathbf{F}\in\mathcal{F}^*(k)}{\operatorname{argmin}}\, \mathbf{s}(k)\operatorname{diag}(\mathbf{F}). \tag{8}$$

The choice of a transmission mode based on the current network state requires significant signaling. Next, in this section we propose an approach to designing a policy for distributed link scheduling whose performance is close to that of the centralized algorithm. The method is based on the use of *shadow queues* introduced in [27] and then implemented for delay reduction in multi-hop networks in [18]. We assign to each queue $i \in \mathcal{I}$ a shadow queue, which is a variable $\tilde{q}_i(k)$ such that the row vector $\tilde{\mathbf{q}}(k) = (\tilde{q}_i(k))_{i\in\mathcal{I}}$ evolves as

$$\tilde{\mathbf{q}}(k+1) = \tilde{\mathbf{q}}(k) + \min(\tilde{\mathbf{q}}(k), \mathbf{c})\mathbf{F}(k) + \tilde{\lambda}, \quad k \geq 0, \tag{9}$$

where min represents a per-component minimum. Here

$$\tilde{\lambda} = \big((1+\epsilon_1)\lambda_1, \ldots, (1+\epsilon_I)\lambda_I\big), \tag{10}$$

is a row vector in which ϵ_i, $i \in \mathcal{I}$, are positive constants such that $((1+\epsilon_i)\lambda_i)_{i\in\mathcal{I}_0}$ belongs to the system's capacity region.

The dynamics of the *shadow queues* (9) differ from that of the actual queues $\mathbf{q}(k)$ given by (3) in the use of the fixed $\tilde{\lambda}$ instead of the random disturbance $\mathbf{a}(k)$ representing the actual numbers of arrivals. Arrival rates λ_i and constants ϵ_i may not be integers, hence the components of $\tilde{\mathbf{q}}(k)$ may not be integers either, unlike the components of $\mathbf{q}(k)$.

As previously for the actual queues, we let $\tilde{\mathbf{q}}(0) = \mathbf{0}$. Then, to obtain $\tilde{\mathbf{q}}(k+1)$ by (9), its value in time slot k, $\tilde{\mathbf{q}}(k)$, is used in some given centralized control policy π_c to select a transmission mode, i.e., $\mathbf{F}(k) = \pi_c(\tilde{\mathbf{q}}(k))$. The chosen transmission mode is then substituted in (9). Thus, transmission mode selection does not depend on the actual network state and can be implemented in a distributed manner, where all nodes locally use the same policy π_c with the same fixed disturbance $\tilde{\lambda}$ and obtain the same controls, which they apply to the network. It was shown in [18] that such a control ensures a finite average number of packets in all actual queues as long as the non-zero elements of (10) are interior to the capacity region and the policy π_c is throughput optimal.

5. Numerical results

We now proceed illustrating the performance of the proposed approach. We assume that the capacities of the backhaul links, downlink access links, and uplink access links are all pairwise equal. That is, we let $c_1 = c_2$, $c_3 = c_5$, and $c_4 = c_6$. We also assume that IAB network is using the FR2 band with 200 MHz of bandwidth and a subcarrier spacing of 120 kHz, which corresponds to the NR numerology 3. Thus, the number of primary resource blocks, $N_{PRB}^{BW,\mu}$, is equal to 132, and the symbol duration T_s^{μ} is equal to 8.92×10^{-6}. Additionally, the uplink and downlink overheads, as defined by [28], are $OH_{UL} = 0.1$ for the uplink and $OH_{DL} = 0.1$ for the downlink.

We consider three different scenarios, each with a different set of parameter values. In the *maximum UL/DL* scenario, there are no hardware limitations for UEs in both the uplink or downlink directions. In the *limited UL* scenario, the capabilities of UEs are limited in the uplink direction only. Finally, in the *limited UL/DL*, UEs have limitations in both the uplink and downlink directions. Table 1 provides the scenario-specific values for the parameters used in our analysis.

According to 3GPP [28] the data rates of the access links can be estimated as

$$C_{X[\text{Mbps}]} = 10^{-6}\nu_{L,X}Q_{m,X}fR_X\frac{12N_{PRB}^{BW,\mu}}{T_s^{\mu}}(1-OH_X), \quad X \in \{UL, DL\}. \tag{11}$$

Let the time slot duration be 1 ms and let the packet size be 1500 bytes. Thus, to calculate, e.g., the capacity of the downlink access link, $c_3 = c_5$, we first compute C_{DL} in Mbps by (11) and then convert the value to packets per time slot as

$$c_{DL\,[\text{pkts/ms}]} = 10^{-3}(C_{DL\,[\text{Mbps}]} \times 10^6)/(8 \times 1500), \tag{12}$$

after which c_{DL} is rounded down to an integer and assigned to $c_3 = c_5$.

Table 1

Scenario-specific UE parameters

Parameter	**Notation**	**Maximum UL/DL**	**Limited UL**	**Limited UL/DL**
UL number of multiplexed layers	$\nu_{L,UL}$	4	2	1
DL number of multiplexed layers	$\nu_{L,DL}$	6	6	1
UL modulation order	$Q_{m,UL}$	6	4	4
DL modulation order	$Q_{m,DL}$	6	6	6
Scaling factor	f	1	1	0.75
UL error coding rate	R_{UL}	948/1024	490/1024	490/1024
DL error coding rate	R_{DL}	948/1024	948/1024	438/1024
UL rate, Mbps	C_{UL}	3547	611	229
DL rate, Mbps	C_{DL}	4848	4848	280

Since backhaul links are characterized by a higher transmission power and hence a high-order modulation scheme can be used, we take the backhaul link capacities one and a half times as large as the access downlink capacities, i.e., $c_B = 1.5c_{DL}$. Then c_B is also rounded down to an integer and assigned to $c_1 = c_2$. Thus, we obtain three vectors of link capacities $\mathbf{c}$: (606, 606, 404, 295, 404, 295) for *maximum UL/DL*, (606, 606, 404, 50, 404, 50) for *limited UL*, and (34, 34, 23, 19, 23, 19) for *limited UL/DL*.

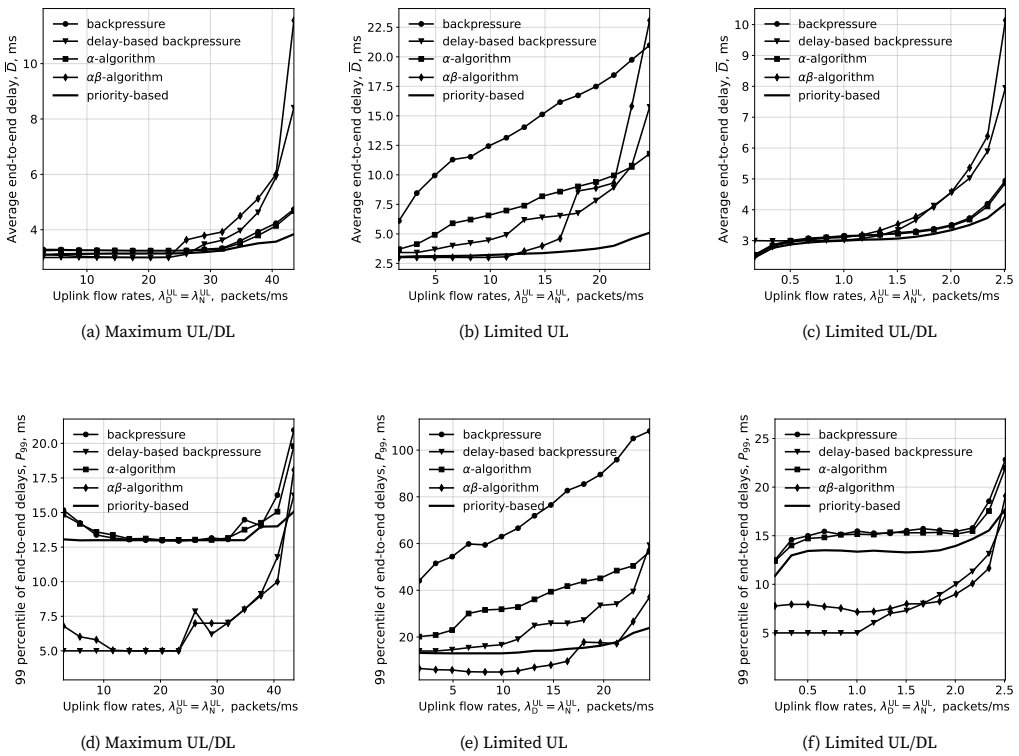
Finally, by following the recommendations for traffic modeling in the standard [3], we assume that the number of packets $a_i(k)$ arriving to queue $i \in \mathcal{I}_0$ in each time slot $k \geqslant 0$ are i.i.d. random variables distributed according to Poisson law with mean λ_i.

We start by comparing the centralized algorithms discussed in Section 2, namely *backpressure, delay-based backpressure, α-algorithm* and *αβ-algorithm*, with the centralized *priority-based* implementation in terms of the average delay $\overline{D}$ and the 99th delay percentile P_{99}. For a convenient presentation of results, we denote the uplink arrival rates from the donor- and node-associated UEs, respectively, as $\lambda_6 = \lambda_{\mathrm{D}}^{\mathrm{UL}}$ and $\lambda_4 = \lambda_{\mathrm{N}}^{\mathrm{UL}}$, and the downlink arrival rates to the donor- and node-associated UEs as $\lambda_5 = \lambda_{\mathrm{D}}^{\mathrm{DL}}$ and $\lambda_1 = \lambda_{\mathrm{N}}^{\mathrm{DL}}$. In all presented figures, at each point, 50 simulation runs, each having 1000 time slots, were generated and then averaged to obtain $\overline{D}$ and P_{99}.

The comparison of the centralized schemes is shown in figure 5, where the arrival rates at each AP are equal and the ratios of the downlink to uplink arrival rates are fixed to four, i.e., $\lambda_{\mathrm{N}}^{\mathrm{DL}} = \lambda_{\mathrm{D}}^{\mathrm{DL}} = 4\lambda_{\mathrm{N}}^{\mathrm{UL}} = 4\lambda_{\mathrm{D}}^{\mathrm{UL}}$. With such parameters, figure 5 shows $\overline{D}$ and P_{99} as functions of the uplink arrival rates for the three studied scenarios.

By analyzing the results in figure 5 we observe that the lowest average delay value is provided by the proposed *priority-based* algorithm. The closest result is demonstrated by the *backpressure* and *α-algorithm* in the *maximum UL/DL* and *limited UL/DL* scenarios. In terms of the 99th percentile P_{99}, *delay-based backpressure* and *αβ-algorithm* show the best performance in *maximum UL/DL* and *limited UL/DL*, whereas in *limited UL* the *αβ-algorithm* performs the best. Moreover, from figure 5 we can see that the *priority-based* policy provides network stability wherever the throughput optimal policies do, i.e., wherever it is possible.

We note that the qualitative behavior of all the algorithms in *maximum UL/DL* and *limited UL/DL* is similar. The rationale is that the elements of the link capacity vectors in these scenarios are closely proportional. Moreover, the ratios of the largest and smallest capacities therein are 2 and 1.8, while this ratio in *limited UL* is 12.1. The range of link capacities' values in *limited UL* is thus considerably wider, and the performance ranking of control policies it yields is different.



(a) Maximum UL/DL (b) Limited UL (c) Limited UL/DL

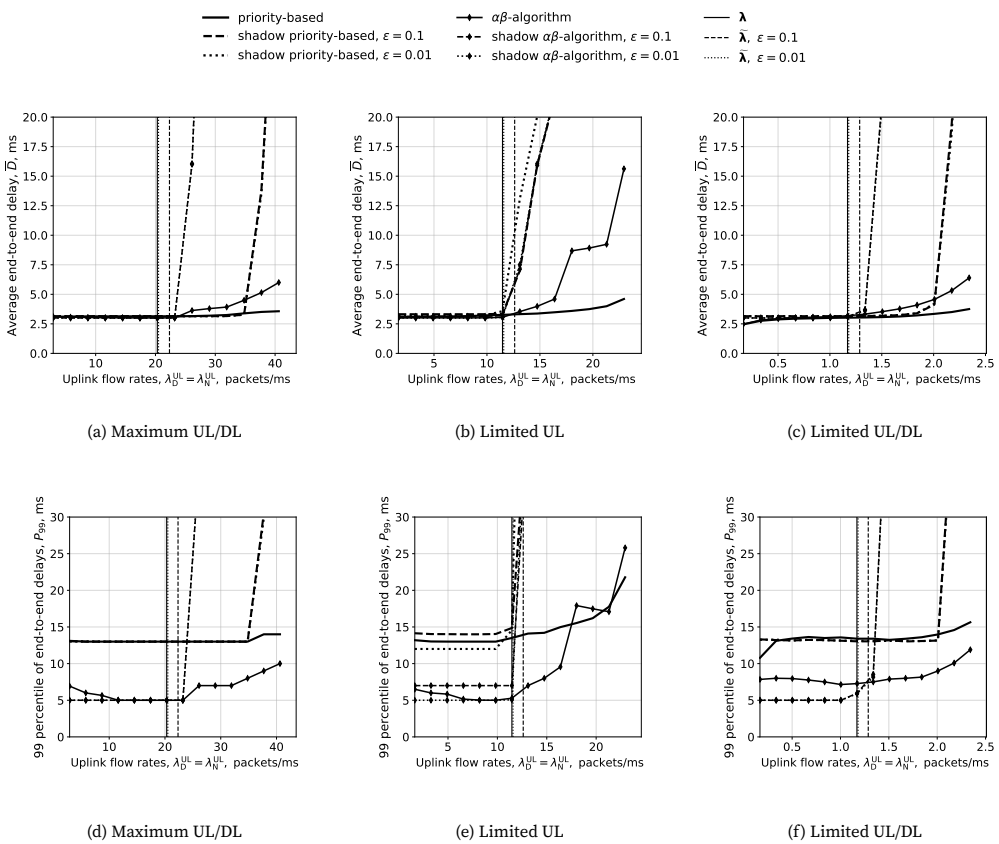
(d) Maximum UL/DL (e) Limited UL (f) Limited UL/DL

Figure 5. Performance evaluation of the centralized link activation policies in terms of the mean delay $\overline{D}$ (top) and the 99th delay percentile P_{99} (bottom) vs. the uplink arrival rates $\lambda_N^{UL} = \lambda_D^{UL}$ with $\lambda_N^{DL}/\lambda_N^{UL} = \lambda_D^{DL}/\lambda_D^{UL} = 4$.

Having identified the *αβ* and *priority-based* algorithms as performing best in terms of the 99th percentile and average end-to-end delay, respectively, we now evaluate their distributed implementations constructed using *shadow queues*. Recall, that a policy choosing the transmission mode based on the *shadow queue* lengths ensures network stability as long as $\tilde{\lambda}$ defined in (10) lies within the capacity region. This means that a larger ϵ can cause instability at high arrival rates but prevents it if the actual arrival rates increase slightly (no more than by $100\epsilon\%$) while $\tilde{\lambda}$ is fixed.

Similarly to figure 5, figure 6 shows the delay metrics $\overline{D}$ (top) and P_{99} (bottom) as functions of the arrival rates. Assuming that the system initially operates with some arrival rates λ, shown in figure 6 by the solid vertical lines, we fix two sets of shadow arrival rates: $\tilde{\lambda}_1$ defined by (10) using $\epsilon_i = 0.1$ for all i and indicated by the dashed vertical lines, and $\tilde{\lambda}_2$ defined using $\epsilon_i = 0.01$ for all i and shown by the dotted vertical lines. Then, we let the actual arrival rates vary along the horizontal and evaluate the system's performance under the centralized control (the results shown by solid lines) and using the *shadow queues* with the arrival rates $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ fixed previously (dashed and dotted lines, respectively). Thus, to the right from the dashed and dotted vertical lines, the actual arrival rates are greater than the corresponding shadow arrival rates $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$, and to the left they are smaller.

As it could be expected, the *shadow-queues*-controlled network is stable when the actual arrival rates are less than the shadow arrival rates. Additionally, we note that the delay performance is very close to that in a network with centralized control. Interestingly, in the *maximum UL/DL* and *limited UL/DL* scenarios the network is stable even if the actual arrival rates are slightly higher, than the shadow arrival rates. We note that the *priority-based* algorithm maintains the system stable over a wider range of real arrival rates than the *αβ-algorithm*.



(a) Maximum UL/DL (b) Limited UL (c) Limited UL/DL

(d) Maximum UL/DL (e) Limited UL (f) Limited UL/DL

Figure 6. Performance evaluation of the distributed link activation policies in terms of the mean delay $\overline{D}$ (top) and the 99th delay percentile P_{99} (bottom) vs. the uplink arrival rates $\lambda_N^{UL} = \lambda_D^{UL}$ with $\lambda_N^{DL}/\lambda_N^{UL} = \lambda_D^{DL}/\lambda_D^{UL} = 4$.

6. Conclusions

In this paper, we considered the IAB technology enabling cost-effective deployment of dense 5G networks operating in high frequency bands. Specifically, we concentrated on the half-duplex regime and focused on link activation as a critical task for this type of networks. By identifying throughput and delay as relevant performance criteria, we designed a *priority-based* link activation policy for 5G IAB networks, which allows for packet delay reduction and throughput maximization. The proposed policy can be implemented either by the network controller in a centralized way or in a distributed manner by each network node using the proposed *shadow queue* mechanism.

By using a model of an IAB network with a basic topology consisting of one IAB donor, one IAB node, and two groups of UEs we evaluated link activation policies for three scenarios, each with different hardware capabilities. Performance of the proposed policy was compared numerically to the well-known *backpressure* policy and its delay-oriented modifications. We have shown that the centralized *priority-based* policy provides the lowest average end-to-end delay in the considered simulation setup. It also outperforms some of the other studied policies in terms of the 99th delay percentile and achieves stability in the entire capacity region.

Finally, our results also demonstrate that a distributed implementation using *shadow queues* leads to approximately the same delays as the centralized implementation.

Funding: This paper has been supported by the Russian Science Foundation, project no. 23-79-10084, `https://rscf.ru/project/23-79-10084`.

References

1. Molchanov, D. A., Begishev, V. O., Samuilov, K. E. & Kucheryavy, E. A. *5G/6G networks: architecture, technologies, methods of analysis and calculation* 516 pp. (PFUR, 2022).
2. Holma, H., Toskala, A. & Nakamura, T. *5G Technology: 3GPP New Radio* (Wiley, 2020).
3. 3GPP. *Study on Integrated Access and Backhaul* Technical Report (TR) 38.874. Version 16.0.0 (3GPP, Dec. 2018).
4. Monteiro, V., Lima, F., Moreira, D., Sousa, D., Maciel, T., Behrooz, M. & Hannu, H. Paving the Way Toward Mobile IAB: Problems, Solutions and Challenges. *IEEE Open Journal of the Communications Society* **PP,** 1–1. doi:10.1109/OJCOMS.2022.3224576 (Jan. 2022).
5. Sadovaya, Y., Moltchanov, D., Mao, W., Orhan, O., Yeh, S.-p., Nikopour, H., Talwar, S. & Andreev, S. Integrated Access and Backhaul in Millimeter-Wave Cellular: Benefits and Challenges. *IEEE Communications Magazine* **60,** 81–86. doi:10.1109/MCOM.004.2101082 (2022).
6. Hong, S., Brand, J., Choi, J. I., Jain, M., Mehlman, J., Katti, S. & Levis, P. Applications of self-interference cancellation in 5G and beyond. *IEEE Communications Magazine* **52,** 114–121. doi:10.1109/MCOM.2014.6736751 (2014).
7. Ford, R., Gómez-Cuba, F., Mezzavilla, M. & Rangan, S. *Dynamic time-domain duplexing for self-backhauled millimeter wave cellular networks* in *2015 IEEE International Conference on Communication Workshop (ICCW)* (2015), 13–18. doi:10.1109/ICCW.2015.7247068.
8. Ahmed, I. & Mohamed, A. *On the joint scheduling and intra-cell interference coordination in multi-relay LTE uplink* in *2012 IEEE Globecom Workshops* (2012), 111–115. doi:10.1109/GLOCOMW.2012.6477554.
9. Wang, L., Ai, B., Niu, Y., Jiang, H., Mao, S., Zhong, Z. & Wang, N. Joint User Association and Transmission Scheduling in Integrated mmWave Access and Terahertz Backhaul Networks. *IEEE Transactions on Vehicular Technology,* 1–11. doi:10.1109/TVT.2023.3293788 (2023).
10. Qiao, J., Cai, L. X., Shen, X. & Mark, J. W. *STDMA-based scheduling algorithm for concurrent transmissions in directional millimeter wave networks* in *2012 IEEE International Conference on Communications (ICC)* (2012), 5221–5225. doi:10.1109/ICC.2012.6364219.
11. Gómez-Cuba, F. & Zorzi, M. Optimal Link Scheduling in Millimeter Wave Multi-Hop Networks With MU-MIMO Radios. *IEEE Transactions on Wireless Communications* **19,** 1839–1854. doi:10.1109/TWC.2019.2959295 (2020).
12. Yarkina, N., Moltchanov, D. & Koucheryavy, Y. Counter Waves Link Activation Policy for Latency Control in In-Band IAB Systems. *IEEE Communications Letters* **27,** 3108–3112. doi:10.1109/LCOMM.2023.3313233 (2023).
13. Gupta, M., Rao, A., Visotsky, E., Ghosh, A. & Andrews, J. G. Learning Link Schedules in Self-Backhauled Millimeter Wave Cellular Networks. *IEEE Transactions on Wireless Communications* **19,** 8024–8038. doi:10.1109/TWC.2020.3018955 (2020).
14. Gopalam, S., Hanly, S. V. & Whiting, P. Distributed and Local Scheduling Algorithms for mmWave Integrated Access and Backhaul. *IEEE/ACM Transactions on Networking* **30,** 1749–1764. doi:10.1109/TNET.2022.3154367 (2022).
15. Tassiulas, L. & Ephremides, A. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control* **37,** 1936–1948. doi:10.1109/9.182479 (1992).
16. Neely, M., Modiano, E. & Li, C.-P. *Fairness and optimal stochastic control for heterogeneous networks* in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.* **3** (2005), 1723–1734 vol. 3. doi:10.1109/INFCOM.2005.1498453.
17. Li, Q. & Negi, R. *Scheduling in Wireless Networks under Uncertainties: A Greedy Primal-Dual Approach* in *2011 IEEE International Conference on Communications (ICC)* (2011), 1–5. doi:10.1109/icc.2011.5963357.
18. Bui, L., Srikant, R. & Stolyar, A. *Novel Architectures and Algorithms for Delay Reduction in Back-Pressure Scheduling and Routing* in *IEEE INFOCOM 2009* (2009), 2936–2940. doi:10.1109/INFCOM.2009.5062262.
19. Stolyar, A. & Ramanan, K. Largest Weighted Delay First Scheduling: Large Deviations and Optimality. *The Annals of Applied Probability* **11.** doi:10.1214/aoap/998926986 (Feb. 2001).
20. Andrews, M., Kumaran, K., Ramanan, K., Stolyar, A., Vijayakumar, R. & Whiting, P. Scheduling in a queuing system with asynchronously varying service rates. *Probability in the Engineering and Informational Sciences* **18,** 191–217. doi:10.1017/S0269964804182041 (Apr. 2004).

21. McKeown, N., Mekkittikul, A., Anantharam, V. & Walrand, J. Achieving 100% throughput in an input-queued switch. *IEEE Transactions on Communications* **47,** 1260–1267. doi:10.1109/26.780463 (1999).
22. Ji, B., Joo, C. & Shroff, N. B. *Delay-based Back-Pressure scheduling in multi-hop wireless networks* in *2011 Proceedings IEEE INFOCOM* (2011), 2579–2587. doi:10.1109/INFCOM.2011.5935084.
23. Venkataramanan, V. J. & Lin, X. On Wireless Scheduling Algorithms for Minimizing the Queue-Overflow Probability. *IEEE/ACM Transactions on Networking* **18,** 788–801. doi:10.1109/TNET.2009.2037896 (2010).
24. Venkataramanan, V. J., Lin, X., Ying, L. & Shakkottai, S. *On Scheduling for Minimizing End-to-End Buffer Usage over Multihop Wireless Networks* in *2010 Proceedings IEEE INFOCOM* (2010), 1–9. doi:10.1109/INFCOM.2010.5462117.
25. Neely, M. *Stochastic Network Optimization with Application to Communication and Queueing Systems* doi:10.2200/S00271ED1V01Y201006CNT007 (Jan. 2010).
26. Venkataramanan, V. J. & Lin, X. *Low-complexity scheduling algorithm for sum-queue minimization in wireless convergecast* in *2011 Proceedings IEEE INFOCOM* (2011), 2336–2344. doi:10.1109/INFCOM.2011.5935052.
27. Bui, L., Srikant, R. & Stolyar, A. *Optimal resource allocation for multicast flows in multihop wireless networks* in *2007 46th IEEE Conference on Decision and Control* (2007), 1134–1139. doi:10.1109/CDC.2007.4434451.
28. 3GPP. *User Equipment (UE) radio access capabilities* Technical Specification (TS) 38.306. Version 17.2.0 (3rd Generation Partnership Project (3GPP), Sept. 2022).

To cite: Zhivtsova A. A., Beschastnyy V. A., A new link activation policy for latency reduction in 5G integrated access and backhaul systems, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)86–98.DOI: 10.22363/2658-4670-2024-32-1-86-98.

Information about the authors

Zhivtsova, Anna A.—bachelor's degree student of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia (RUDN University) (e-mail: aazhivtsova@sci.pfu.edu.ru, phone: +7(910)484-71-44, ORCID: https://orcid.org/0009-0007-8438-6850)

Beschastnyy, Vitaly A.—Candidate of Physical and Mathematical Sciences, assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia (RUDN University) (e-mail: vbeschastny@sci.pfu.edu.ru, phone: +7(905)776-38-58, ORCID: https://orcid.org/0000-0003-1373-4014, Scopus Author ID: 57192573001)

УДК 004.2, 004.7
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2024-32-1-86-98 EDN: CCHVBS

Стратегия активации каналов для снижения задержки пакетов в сетях интегрированного доступа и транзита 5G

А. А. Живцова, В. А. Бесчастный

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. Блокировка путей распространения радиоволн является одним из основных препятствий на пути развертывания сетей сотовой связи пятого поколения (Fifth Generation) Новое Радио (New Radio) в диапазоне миллиметровых волн (30–100 ГГц). Возможным решением данной проблемы является уплотнение сетей радиодоступа, однако оно связано высокими капитальными затратами операторов связи. Экономически эффективное уплотнение может быть достигнуто с помощью технологии интегрированного доступа и транзита (Integrated Access and Backhaul), использующей ретрансляционные узлы между абонентом и базовой станцией. Такие системы были разработаны главным образом для борьбы с блокировками без учета показателей качества обслуживания (Quality of Service). При этом использование ретрансляционных узлов отрицательно влияет на сквозную задержку пакета. В данной работе предлагается новая стратегия активации каналов направленная на сокращение задержек в системах интегрированного доступа и транзита, учитывающая органичения полудуплексной передачи. Предлагаемый подход основан на динамической приоритезации очередей на базе количества пакетов, которые могут быть переданы в одном временно́м слоте. Результаты имитационного моделирования с использованием реалистичных исходных данных показывают, что предлагаемый алгоритм обеспечивает наименьшую среднюю задержку по сравнению с известными подходами для различных значений нагрузки восходящей и нисходящей передачи.

Ключевые слова: 5G, интегрированный доступ и транзит, миллиметровые волны, полудуплекс, управление активацией каналов

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 99–105
http://journals.rudn.ru/miph

Research article

UDC 538.911, 538.915, 519.6
PACS 07.05.Tp, 02.70.Ns, 71.90.+q, 21.60.Jz, 24.60.Lz, 27.80.+w
DOI: 10.22363/2658-4670-2024-32-1-99-105 EDN: BXKOPH

Computer studies of a dependence of equilibrium state structure on a number of particles for a two-dimensional system of charged particles confined in a disk potential

Eduard G. Nikonov[1,2,3], Rashid G. Nazmitdinov[1,2], Pavel I. Glukhovtsev[2]

[1] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, 141980, Dubna, Russian Federation*
[2] *Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation*
[3] *HSE University, 34 Tallinskaya St, Moscow, 123458, Russian Federation*

(received: February 12, 2024; revised: March 10, 2024; accepted: March 24, 2024)

Abstract. The problem of finding equilibrium configurations of one-component charged particles, induced by external electrostatic fields in planar systems, is a subject of active studies in fundamental as well in experimental investigations. In this paper the results of numerical analysis of the equilibrium configurations of charged particles (electrons), confined in a circular region by an infinite external potential at its boundary are presented. Equilibrium configurations with minimal energy are searched by means of special calculation scheme. This computational scheme consists of the following steps. First, the configuration of the system with the energy as close as possible to the expected energy value in the ground equilibrium state is found using a model of stable configurations. Next, classical Newtonian molecular dynamics is used using viscous friction to bring the system into equilibrium with a minimum energy. With a sufficient number of runs, we obtain a stable configuration with an energy value as close as possible to the global minimum energy value for the ground stable state for a given number of particles. Our results demonstrate a significant efficiency of using the method of classical molecular dynamics (MD) when using the interpolation formulas in comparison with algorithms based on Monte Carlo methods and global optimization. This approach makes it possible to significantly increase the speed at which an equilibrium configuration is reached for an arbitrarily chosen number of particles compared to the Metropolis annealing simulation algorithm and other algorithms based on global optimization methods.

Key words and phrases: Thomson atom, Wigner crystal, molecular dynamics

1. Introduction

The question of how charged particles arrange themselves in a restricted planar geometry attracted continuous attention for many decades (for a review see [1]). Modern technology allows us to study such phenomena on the same scale, from Bose condensates with some thousand atoms to quantum dots with a few electrons, providing rich information about specific features of correlation effects in mesoscopic systems (see, for example, [2, 3]). However, finding the exact analytical equilibrium charge distribution (the one that makes the body an equipotential) is not a simple problem. The existence of the symmetry for considered system may simplify the task. Thomson was the first to suggest an instructive solution for interacting electrons, reducing the 3D harmonic oscillator confinement to a circular (2D) harmonic oscillator [4]. He developed an analytical approach, which enables us to trace a self-organization for a small number of electrons ($n < 10$) in a family of rings (shells) with a certain number of electrons in each shell.

Nowadays, many ideas and concepts introduced in condensed matter physics can be realized and analyzed with high accuracy as a function of particle number and boundary properties. In fact, nanotechnology gave rise to emergence of lateral quantum dots creation which properties where not so obvious in ninetieth. Assuming a circular symmetry of such quantum dots, the first attempt to understand the distribution of electrons in such confined systems were based on the Monte Carlo

© Nikonov E. G., Nazmitdinov R. G., Glukhovtsev P. I., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

(MC) calculations for charged particles (ions and electrons) confined by 2D parabolic and hard-wall potentials (see, e.g., [5–7]). The results of these calculations confirmed the predictions based on the Thomson model for $N \leq 52$. The next step in the attempt to find the analytical description of the distribution of charged particles in the disc geometry have been done in [8]. In this paper the basic principles of self-organization of one-component charged particles, confined in disk have been proposed. A system of equations was derived, which allows to determine equilibrium configurations for an arbitrary, but finite, number of charged particles that are distributed over several rings. This approach reduces significantly the computational effort in minimizing the energy of equilibrium configurations and demonstrates a remarkable agreement with the values provided by molecular dynamics calculations. This paper gave a new impetus to activity in finding the bridge between the distribution of finite number of confined charged particles to their continuous limit (e.g., [9–15]).

From the analysis, based on MC and MD calculations for a relatively small number of charged particles, it follows that the number of stable configurations grows very rapidly with the number of particles. There are many local minima that have energies very close to the global minimum. These metastable states with lower (or higher) symmetry are found with much higher probabilities than the true ground state [9, 10]. This picture is akin to a liquid-solid transition, when a rapid cooling gives rise to a glass-like disordered solid rather than a crystal with lower energy. Therefore, the description of this phenomenon requires the development of various not only analytical approaches but as well the effective numerical recipes with growing number of confined particles.

In this paper, we present a new approach for numerical-analytical analysis of the equilibrium configurations of charged particles (say, electrons) confined in a disk geometry. Using a model of stable configurations, which takes into account the interaction between shells of charged particles [8, 11], we obtain functional dependencies of the total number of particles of the system on the number of rings and those of the energy of the equilibrium configuration on the total number of particles. These dependencies make it possible to significantly simplify the search for the absolute minimum of the system for a given total number of one component charged particles.

2. System description

The physical formulation of the problem can be as follows. A system of similarly charged particles is given; they are located in a region with a cylindrical confining potential at the boundary. The configurations of the particles are determined by the Hamiltonian, in which the potential energy of the interparticle interaction dominates over the kinetic energy. It is necessary to find a stable configuration of N particles with the minimal energy inside the given region.

Thus, we consider two-dimensional system on a plane consisting of identically charged particles with mutual Coulomb interaction in the confining disk potential with the radius R. The Hamiltonian of such a system can be written as follows:

$$H = \sum_{i=1}^{N} V_{\text{ext}}(r_i) + \alpha \sum_{\substack{i,j=1 \\ i<j}}^{N} \frac{1}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|} + \sum_{i=1}^{N} T_i, \tag{1}$$

where $r_i = |\vec{\mathbf{r}}_i|$ is the distance to the center of the region bounded by the potential $V_{\text{ext}}(r)$, $\alpha = e^2/4\pi\varepsilon_0\varepsilon_r$ is a quantity characterizing the interaction of charges in the medium, and T_i is the particle kinetic energy. The confining potential $V_{\text{ext}}(r)$ is defined as follows:

$$V_{\text{ext}}(r) = \begin{cases} 0, & r < R; \\ \infty, & r \geqslant R. \end{cases} \tag{2}$$

To avoid a large number of metastable states (local minima), the system is considered at temperatures that are close to zero; at these temperatures, the potential energy dominates over the kinetic one. As a result, it is possible to rewrite the function of the total energy of the system (1) as follows:

$$H = \sum_{i=1}^{N} V_{\text{ext}}(r_i) + \alpha \sum_{\substack{i,j=1 \\ i<j}}^{N} \frac{1}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|}. \tag{3}$$

The problem is to find a stable configuration of the particle inside the given region with the minimally possible energy.

Using his model, Thomsom obtained equation (4) for analytical calculation of the coordinates of particles in the equilibrium state for one ring

$$E_n(r) = \frac{\alpha}{2r}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{1}{\sin\frac{\pi}{n}(|i-j|)} = \frac{\alpha n S_n}{4r}, \tag{4}$$

$$S_n = \sum_{i=1}^{n-1}\frac{1}{\sin\frac{\pi}{n}k}.$$

Here, $E_n(r)$ is the Coulomb energy of n particles with the charge e uniformly distributed along a circle with the radius r, $\alpha = e^2/4\pi\varepsilon_0\varepsilon_r$ is the quantity characterizing the interaction of charged particles in the medium. Without loss of generality, electrons with the charge e are considered as charged particles below. An original approach to the calculation of equilibrium configurations and the corresponding energy was formulated in [8, 11]; in this approach, the interaction between shells consisting of charged particles is also taken into account in addition to the energy of one ring (one shell). As a result of solving the (5)

$$\mathcal{F}_i = 0, \quad i = 2, \dots, p.$$

$$\mathcal{F}_i = r_i^2\sum_{j=i+1}^{p}\frac{n_j E\left[\left(r_j/r_i\right)^2\right]}{r_j^2 - r_i^2} - \frac{\pi}{8}S_{n_j} + r_i\sum_{j=1}^{i-1}n_j\left(\frac{r_j E\left[\left(r_i/r_j\right)^2\right]}{r_j^2 - r_i^2} - \frac{K\left[\left(r_i/r_j\right)^2\right]}{r_j}\right). \tag{5}$$

Here, $K = X_{-1}(E = X_1)$ are the complete elliptic integrals of the first (second) kind: $X_p(x) = \int_0^{\pi/2} dt(1 - x\sin^2 t)^{p/2}$; r_i is the value of the ith optimal radius for the given stable configuration of charged particles; and n_i is the number of particles in the ith shell.

As preliminary analysis showed, the solution of these equations makes it possible to significantly reduce the amount of computational work by finding the state that is closest to the equilibrium configuration. It is worth noting that this approach allows us to almost exactly determine both the equilibrium configuration and the total energy of the equilibrium state for $N < 52$.

In this paper, we develop a new approach and methods for calculating the coordinates of particles and the energy of the equilibrium configuration for an arbitrary finite number of particles using the obtained analytic dependencies of the distribution of particles and the energy of the equilibrium configuration on the total number of particles in the system for $N < 1000$.

3. Computational scheme

To further increase the efficiency and reduce the calculation time, the following modification of the traditional approach based on the molecular dynamics method [16] is proposed in this paper.

The computational scheme consists of the following steps.

1. We use interpolation formulas (6) for calculation initial particle distribution

$$N_i(N) = a_i N^{\frac{2}{3}} - b_i, \tag{6}$$

i	1	2	3	4	5	6
a_i	2.7948	1.3439	1.1323	1.0127	0.9482	0.8517
b_i	3.9444	7.2999	10.845	14.850	19.128	21.732

Formulas (6) were obtained by interpolating solutions [17, 18] to equations (3) of the equilibrium configuration model [8, 11].

2. Using the initial distribution of particles inside the circular region obtained in the previous step, we run ab initio calculations using Newtonian molecular dynamics. The dissipation of energy for cooling the system upon reaching the ground state with a minimum of potential and zero kinetic energy is modeled by adding to the equations of motion a viscous friction force

proportional to the speed (7), which realizes the corresponding outflow of energy from the system.

$$m\vec{\mathbf{r}}_i'' = -\nabla_i V\left(\vec{\mathbf{r}}_i\right) - b_f \vec{\mathbf{r}}_i', \quad \forall i \in \{1, 2, 3, \dots, N\}. \tag{7}$$

Here $V\left(\vec{\mathbf{r}}_i\right) = V_{\text{ext}}(r_i) + \alpha \sum_{\substack{i,j=1 \\ i<j}}^{N} \frac{1}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|}$ includes the external confining potential plus the Coulomb terms, and the friction coefficient b_f is the parameter controlling the quenching of velocities.

For calculations we used quenched velocity Verlet algorithm. Let at some moment of time t the coordinates $\vec{\mathbf{r}}_i(t)$ and velocities $\vec{\mathbf{v}}_i(t)$ of the particles be given, then at the moment of time $t + \Delta t$, where Δt is some fixed time step, we can obtain the coordinates $\vec{\mathbf{r}}_i(t + \Delta t)$ and velocities $\vec{\mathbf{v}}_i(t + \Delta t)$ using the following formulas (8), (9), (10).

$$\vec{\mathbf{f}}_i(t) = -\nabla_i V\left(\vec{\mathbf{r}}_i(t)\right). \tag{8}$$

$$\vec{\mathbf{r}}_i(t + \Delta t) = \vec{\mathbf{r}}_i(t) + \vec{\mathbf{v}}_i(t)\Delta t + \frac{1}{2}(\vec{\mathbf{f}}_i(t) - b_f \vec{\mathbf{v}}_i(t))\Delta t^2. \tag{9}$$

$$\vec{\mathbf{v}}_i(t + \Delta t) = \frac{\left(2 - b_f \Delta t\right)\vec{\mathbf{v}}_i(t) + \left(\vec{\mathbf{f}}_i(t) + \vec{\mathbf{f}}_i(t + \Delta t)\right)\Delta t}{2 + b_f \Delta t}. \tag{10}$$

This is the so-called quenched molecular dynamics method.

3. The final stage of calculations includes a fairly large sequence of runs of molecular dynamics calculations described at the previous step, which is due to the existence of a number of metastable states that exponentially increases with the number of particles near the point of global minimum of energy.

   In order to estimate the minimum number of runs N_{runs} to obtain the ground state energy value $E_{\min}$ as close as possible to the global minimum, the following procedure is used. First, a certain number of runs are performed depending on the number of particles in the system. Then the probability of finding the minimum $P(E_{\min} < E_{\text{val}})$ is calculated. Then the estimate for the minimum number of runs has the following form $N_{\text{runs}} = 1/P(E_{\min} < E_{\text{val}})$.

For a sufficiently large number of runs N_{runs}, in accordance with the Central Limit Theorem, the energy values of metastable states calculated in each run are distributed in accordance with the Gaussian distribution. Typically, the value $E_{\text{avr}} - 3\sigma$ is used as E_{val}. In this case, for a system consisting, for example, of 1000 particles, $E_{\text{avr}} - 3\sigma = 736980.1734$ in reduced units (E_{avr} is an average value for the general sample for all runs. And σ is the corresponding standard deviation.). At a given value of E_{val}, the minimum number of runs N_{runs} to obtain the global minimum energy must be greater than 741.

However, we use a different version of probability estimation and, accordingly, a minimum number of runs to practically guarantee obtaining the minimum energy value. This value was obtained using the asymptotic formula (11) for the minimum energy of the ground state, obtained by the authors [17, 18] based on an analysis of solutions to the equations of the equilibrium configuration model [8, 11].

$$E_{min}(N) = \frac{\pi}{4}N^2 - \frac{\pi}{2}N^{\frac{3}{2}} + \sqrt{\frac{\pi}{2}}N \tag{11}$$

In this case, the estimate for E_{val} for 1000 particles will be equal to 736978.54, which is approximately equal to $E_{\text{avr}} - 4\sigma$. The corresponding minimum number of runs N_{runs} to obtain the global minimum energy must be greater than 25237. This value for the number of runs N_{runs} looks more realistic for the case of the initial configuration of particles distributed randomly within the disk.

As an initial approximation, the particle distribution obtained by solving equations of the model of equilibrium configurations (5) for a certain number of rings is taken. After that, calculations are started using the quenched molecular dynamics method under the condition of a gradual decrease in the system temperature. When the zero temperature is reached, the calculations of the time evolution of the system are considered to be completed, after which the energy of the resulting equilibrium configuration of particles in the system is calculated.

For example, for the above-mentioned system of 1000 particles, we fixed the distribution of particles on the outer ring in the initial configuration $N_1 = 276$ in accordance with the formula (6). As a result, we needed only 500 runs to obtain the minimum ground state energy equal to 736979.7283. The best published value known to us is 736977.7079. This value can be obtained, according to our estimates, for a number of runs no less than 25237 [14].

4. Conclusions

The effectiveness of the approach proposed in this article for calculating the global minimum energy for the ground state of the above-mentioned systems of charged particles is based on the following modifications of the traditional molecular dynamics method.

First, the initial configuration for molecular dynamics calculations was calculated using analytic solutions of the equations of the equilibrium configuration model [8, 11]. Secondly, to cool the system, energy dissipation was carried out due to viscous friction forces. For this purpose, the Verlet velocity scheme with quenching was used. Finally, to estimate the number of molecular dynamics runs with different initial conditions, we used an asymptotic formula [17, 18] for the ground state energy, obtained from an analysis of solutions to the equations of the equilibrium configuration model [8, 11]. As a result, the computational efficiency in terms of computation time has increased by more than two orders of magnitude.

The algorithms and programs developed by us can be used to numerically study the stability of systems of charged particles in various fields of physics, chemistry, molecular biology, and nanotechnology, including, for example, the study of nano-objects, such as quantum dots.

Funding: The work was supported within the framework of the Joint Institute for Nuclear Research project no. 05-6-1118-2014/2023 with the use of resources of the HybriLIT heterogeneous cluster of the Laboratory of Information Technologies, Joint Institute for Nuclear Research. Work by E.G.N. partially supported by the Basic Research Program of the National Research University "Higher School of Economics".

References

1. Bowick, M. J. & Giomi, L. Two-dimensional matter: order, curvature and defects. *Advances in Physics* **58,** 449–563. doi:10.1080/00018730903043166 (2009).
2. Saarikoski, H., Reimann, S. M., Harju, A. & Manninen, M. Vortices in quantum droplets: Analogies between boson and fermion systems. *Rev. Mod. Phys.* **82,** 2785–2834. doi:10.1103/RevModPhys.82.2785 (3 Sept. 2010).
3. Birman, J., Nazmitdinov, R. & Yukalov, V. Effects of symmetry breaking in finite quantum systems. *Physics Reports* **526,** 1–91. doi:10.1016/j.physrep.2012.11.005 (2013).
4. Thomson, J. XXIV. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Phil. Mag.* **7,** 237–265. doi:10.1080/14786440409463107 (39 1904).
5. Lozovik, Y. E. & Mandelshtam, V. A. Classical and quantum melting of a Coulomb cluster in a trap. *Physics Letters A* **165,** 469–472. doi:10.1016/0375-9601(92)90349-Q (1992).
6. Bolton, F. & Rössler, U. Classical model of a Wigner crystal in a quantum dot. *Superlattices and Microstructures* **13,** 139–146. doi:10.1006/spmi.1993.1026 (1993).
7. Bedanov, V. M. & Peeters, F. M. Ordering and phase transitions of charged particles in a classical finite two-dimensional system. *Phys. Rev. B* **49,** 2667–2676. doi:10.1103/PhysRevB.49.2667 (4 Jan. 1994).
8. Cerkaski, M., Nazmitdinov, R. G. & Puente, A. Thomson rings in a disk. *Phys. Rev. E* **91,** 032312. doi:10.1103/PhysRevE.91.032312 (2015).
9. Amore, P. Comment on "Thomson rings in a disk". *Phys. Rev. E* **95,** 026601. doi:10.1103/PhysRevE.95.026601 (2 Feb. 2017).
10. Puente, A., Nazmitdinov, R. G., Cerkaski, M. & Pichugin, K. N. Reply to "Comment on 'Thomson rings in a disk' ". *Phys. Rev. E* **95,** 026602. doi:10.1103/PhysRevE.95.026602 (2 Feb. 2017).
11. Nazmitdinov, R. G., Puente, A., Cerkaski, M. & Pons, M. Self-organization of charged particles in circular geometry. *Phys. Rev. E* **95,** 042603. doi:10.1103/PhysRevE.95.042603 (2017).
12. Ash, B., Chakrabarti, J. & Ghosal, A. Static and dynamic properties of two-dimensional Coulomb clusters. *Phys. Rev. E* **96,** 042105. doi:10.1103/PhysRevE.96.042105 (4 Oct. 2017).

13. Jami, P., Ash, B. & Ghosal, A. Significance of the nature of disorder on the universal features of the spatio-temporal correlations of two-dimensional Coulomb-clusters. *Eur. Phys. J. B* **96,** 45. doi:10.1140/epjb/s10051-023-00510-5 (4 2023).
14. Amore, P. & Zarate, U. Thomson problem in the disk. *Phys. Rev. E* **108,** 055302. doi:10.1103/PhysRevE.108.055302 (5 Nov. 2023).
15. Batle, J., Vlasiuk, O. & Ciftja, O. Correspondence between Electrostatics and Contact Mechanics with Further Results in Equilibrium Charge Distributions. *Annalen der Physik,* 2300269. doi:10.1002/andp.202300269 (Nov. 2023).
16. Frenkel, D. & Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications* 664 pp. (Academic, New York, 2001).
17. Nikonov, E. G., Nazmitdinov, R. G. & Glukhovtsev, P. I. Molecular dynamics studies of equilibrium configurations of equally charged particles in planar systems with circular symmetry. *Computer research and modeling* **14,** 609–618. doi:10.20537/2076-7633-2022-14-3-609-618 (2022).
18. Nikonov, E. G., Nazmitdinov, R. G. & Glukhovtsev, P. I. On the Equilibrium Configurations of Charged Ions in Planar Systems with Circular Symmetry. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques* **17,** 235–239. doi:10.1134/S1027451023010354 (2023).

To cite: Nikonov E. G., Nazmitdinov R. G., Glukhovtsev P. I., Computer studies of a dependence of equilibrium state structure on a number of particles for a two-dimensional system of charged particles confined in a disk potential, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)99–105.DOI: 10.22363/2658-4670-2024-32-1-99-105.

Information about the authors

Nikonov, Eduard G.—Doctor of Physical and Mathematical Sciences, Head of Sector MLIT JINR (e-mail: e.nikonov@jinr.ru, phone: +7(496)2164722, ORCID: https://orcid.org/0000-0001-7162-0344, ResearcherID: C-4841-2016, Scopus Author ID: 6603099928)

Nazmitdinov, Rashid G.—Doctor of Physical and Mathematical Sciences, Leading Researcher BLTP JINR (e-mail: rashid@theor.jinr.ru, ORCID: https://orcid.org/000-0003-0489-7858, ResearcherID: G-4860-2016, Scopus Author ID: 7004352313)

Glukhovtsev Pavel I.—Master's degree student of Department of distributed information computing systems of Dubna State University (e-mail: pavelgl2018@gmail.com, ORCID: https://orcid.org/0009-0005-6424-4455)

УДК 538.911, 538.915, 519.6
PACS 07.05.Tp, 02.70.Ns, 71.90.+q, 21.60.Jz, 24.60.Lz, 27.80.+w
DOI: 10.22363/2658-4670-2024-32-1-99-105 EDN: BXKOPH

Компьютерные исследования зависимости от числа частиц структуры основного состояния двумерной системы заряженных частиц ограниченных круговым потенциалом

Э. Г. Никонов[1, 2, 3], Р. Г. Назмитдинов[1, 2], П. И. Глуховцев[2]

[1] *Объединенный институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*
[2] *Государственный университет «Дубна», ул. Университетская, д. 19, Дубна, 141980, Российская Федерация*
[3] *Национальный исследовательский университет «Высшая школа экономики», ул. Таллинская, д. 34, Москва, 123458, Российская Федерация*

Аннотация. Проблема нахождения равновесных конфигураций однокомпонентных заряженных частиц, индуцированных внешними электростатическими полями в планарных системах, является предметом активных исследований как в фундаментальных, так и в экспериментальных исследованиях. В данной работе представлены результаты численного анализа равновесных конфигураций заряженных частиц (электронов), удерживаемых в круговой области бесконечным внешним потенциалом на ее границе. Равновесные конфигурации с минимальной энергией ищутся с помощью специальной схемы расчета. Данная вычислительная схема состоит из следующих шагов. Сначала с помощью стационарной модели находится конфигурация системы с энергией, максимально близкой к ожидаемому значению энергии в основном состоянии равновесия. Далее используется классическая ньютоновская молекулярная динамика с использованием вязкого трения, чтобы привести систему в равновесие с минимальной энергией. При достаточном количестве прогонов мы получаем устойчивую конфигурацию со значением энергии, максимально близким к глобальному минимальному значению энергии для основного устойчивого состояния для заданного числа частиц. Наши результаты демонстрируют значительную эффективность использования метода классической молекулярной динамики (МД) при использовании интерполяционных формул по сравнению с алгоритмами, основанными на методах Монте-Карло и глобальной оптимизации. Такой подход позволяет существенно повысить скорость достижения равновесной конфигурации для произвольно выбранного числа частиц по сравнению с алгоритмом моделирования отжига Метрополиса и другими алгоритмами, основанными на методах глобальной оптимизации.

Ключевые слова: атом Томсона, кристалл Вигнера, молекулярная динамика

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 106–111
http://journals.rudn.ru/miph

Research article

UDC 519.63, 519.688
PACS 07.05.Tp, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-106-111
EDN: CBNJTD

Numerical study of the φ^4 standing waves in a ball of finite radius

Elena Zemlyanaya[1,2], Alla Bogolubskaya[1], Maxim Bashashin[1,2], Nora Alexeeva[3]

[1] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*
[2] *Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation*
[3] *University of Cape Town, 7701 Rondebosch, South Africa*

(received: February 7, 2024; revised: February 28, 2024; accepted: March 14, 2024)

Abstract. Study of spherically symmetric time-periodic standing waves of the ϕ^4 model in a ball of finite radius was carried out based on the numerical solution of a boundary value problem on a cylindrical surface for a wide range of values of the oscillation period. The standing waves in a ball of finite radius can be considered as an approximation of weakly radiating spherically symmetric oscillons in the ϕ^4 model. Stability analysis the waves obtained is based on the calculation of the corresponding Floquet multipliers. In the paper, mathematical formulation of the problem is given, the numerical approach is described, including the method of parallel implementation of the calculation of Floquet multipliers on the computing resources of the HybriLIT platform of the Multifunctional Information and Computing Complex of the Joint Institute for Nuclear Research (Dubna). The results of the study of the space-time structure and bifurcation of coexisting standing waves of various types are presented.

Key words and phrases: oscillons, numerical study, parallel computing

1. Introduction

We consider spherically symmetric standing waves in the ϕ^4 equation

$$\Phi_{tt} - \Delta\Phi - \Phi + \Phi^3 = 0, \quad \Delta = \frac{d^2}{dr^2} + \frac{2}{r}. \tag{1}$$

Equation (1) can be used as a model for a wide range of nonlinear wave processes within various physical contexts. The localized long-lived pulsating states (pulsons, oscillons) in the three-dimensional ϕ^4 theory are known since 1975 [1]. Subsequent computer simulations [2, 3] revealed the formation of long-lived pulsating structures of large amplitude and nearly unchanging width, see a fragment of such structure in figure 1. Renewed interest to oscillons is inspired by applications in cosmology and high energy physics (see, i.e. [4–7]).

Due the permanent loss of energy to the second-harmonic radiation, the oscillons are not exactly time-periodic. In contrast, the standing waves are periodic and can be determined as solutions of a boundary-value problem on the cylindrical surface. In a recent paper [8], the standing waves in a ball of large finite radius are considered as approximation of infinite-space weakly radiating spherically symmetric oscillons in the ϕ^4 model. In the present contribution, we provide details of the numerical approach and present results on the spatio-temporal structure and bifurcation of the standing waves. We outline the numerical scheme that was utilised for that purpose as well as its parallel computer implementation.

© Zemlyanaya E., Bogolubskaya A., Bashashin M., Alexeeva N., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

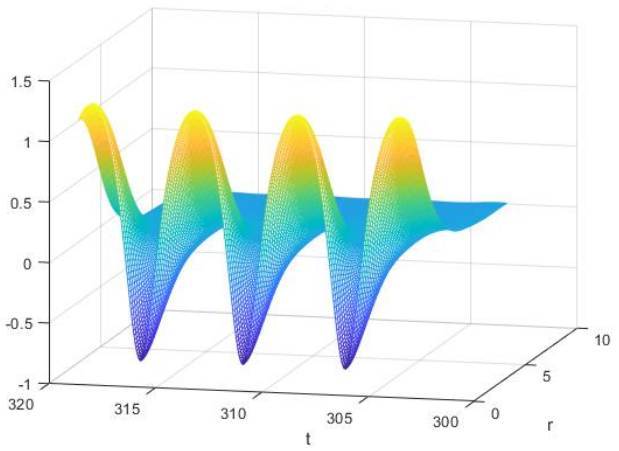


Figure 1. A fragment of pulsating radial component of spherically symmetric solution of Eq. (1)

2. Mathematical problem and numerical approach

Let $\phi = \Phi - \Phi_0$ where $\Phi(r,t)$ is a spherically symmetric solution of equation (1) approaching $\Phi_0 = -1$ (one of two vacuum solutions) as $r \to \infty$. Thus, equation (1) takes a form:

$$\phi_{tt} - \phi_{rr} - \frac{2}{r}\phi_r + 2\phi - 3\phi^2 + \phi^3 = 0. \tag{2}$$

We are searching for solutions of the equation (2) satisfying the boundary conditions

$$\phi_r(0,t) = \phi(R,t) = 0, \quad \phi(r,T) = \phi(r,0). \tag{3}$$

Letting $\tau = t/T$ and defining $\psi(r,\tau) = \phi(r,t)$ yields

$$\psi_\tau(r,\tau) = T\phi_t(r,t), \quad \psi_{\tau\tau}(r,\tau) = T^2\phi_{tt}(r,t).$$

Thus, a boundary-value problem in the two-dimensional domain $[0,1]\times[0,R]$ takes a form:

$$\psi_{tt} + T^2\cdot\left[-\psi_{rr} - \frac{2}{r}\psi_r + 2\psi - 3\psi^2 + \psi^3\right] = 0, \tag{4}$$

$$\psi_r(0,t) = \psi(R,t) = 0, \quad \psi(r,1) = \psi(r,0). \tag{5}$$

For each value of T the boundary-value problem (4),(5) was solved by means of the Newtonian iteration [9] with the 2nd order finite difference approximation for the derivatives. The t and r discrete steps were taken to be 0.01 and 0.1, respectively.

Any solution of equation (4) can be characterised by its energy

$$E = 4\pi\int_0^R \left(\frac{\phi_t^2}{2} + \frac{\phi_r^2}{2} + \phi^2 - \phi^3 + \frac{\phi^4}{4}\right) r^2 dr \tag{6}$$

and its corresponding frequency $\omega = \frac{2\pi}{T}$. If the solution with frequency ω does not change appreciably as R is increased — in particular, if the energy (6) does not change — this standing wave provides a fairly accurate approximation for the periodic solution in an infinite space.

Solutions of equation (4) were numerically continued in the parameter T and the energy-frequency diagram was constructed. Numerical continuation was organized as described in [10].

To classify the stability of the resulting standing waves against spherically symmetric perturbations we consider the linearised equation

$$y_{tt} - y_{rr} - \frac{2}{r}y_r - y + 3(\phi - 1)^2 y = 0 \tag{7}$$

with the boundary conditions $y_r(0,t) = y(R,t) = 0$ [8]. Expanding $y(r,t)$ in Fourier sine series produces a system of $2N$ ordinary differential equations for the coefficients:

$$\dot{u}_m = v_m, \qquad \dot{v}_m + F = 0. \tag{8}$$

Here

$$F = (2 + k_m^2)u_m - 3\sum_{n=1}^{N}(A_{m-n} - A_{m+n})u_n + \frac{3}{2}\sum_{n=1}^{N}(A_{m-n} - A_{m+n})u_n,$$

and A_n, B_n are T-periodic functions of t:

$$A_n(t) = \frac{2}{R}\int_0^R \phi(r,t)\cos(k_n r)dr, \quad B_n(t) = \frac{2}{R}\int_0^R \phi^2(r,t)\cos(k_n r)dr.$$

A set of $2N$ linearly independent solutions, evaluated numerically at $t = T$, forms a monodromy matrix of the system (8). The standing-wave solution $\phi(r,t)$ is deemed stable if all of its $2N$ eigenvalues μ of the monodromy matrix (Floquet multipliers) lie on the unit circle and unstable if there are multipliers outside the circle, see example of stability and instability case at the figure 2.

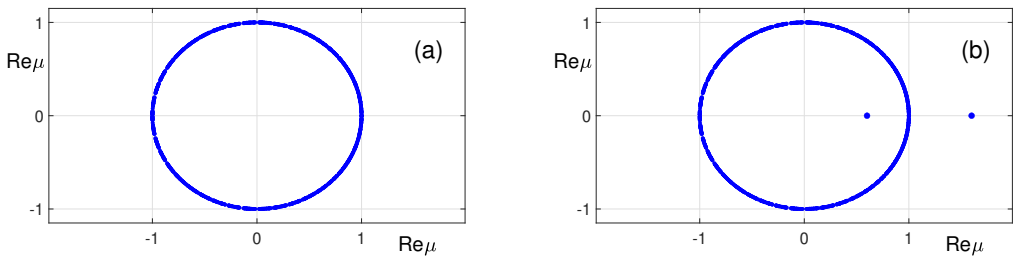


Figure 2. Floquet multipliers at the (Reμ,Imμ) plane. Stability case (a) T=4.7206; instability case (b): T=5.025. Here R=100

The above numerical approach was implemented using the MATLAB *ode45* function with the tolerance value 10^{-7}. A cubic spline interpolation was employed for the calculation of the $A_{m\pm n}(t)$ and $B_{m\pm n}(t)$ coefficients at a set of t points. With N=1000 Fourier harmonics, the calculation of the Floquet multipliers for each individual value of T takes about 48 hours and 24 hours on the *HybriLIT* cluster and the *Govorun* supercomputer, respectively.

To speed up the computations, a parallel algorithm was implemented based on the *parfor* operator. This produces an automated splitting of the solution of $2N$ Cauchy problems into available parallel threads, or "workers". The speedup of calculations in parallel mode can reach 20 times compared to the single- thread calculations.

3. Numerical results and conclusions

It was pointed out in [8] that the energy-frequency dependence, $E(\omega)$, obtained by a numerical continuation of solutions of equation (4) looks like a sequence of spikes triggered by the resonance of frequencies of two coexisting solutions. Positions of the spikes are R-sensitive. It can be seen in figure 3 where fragments of this diagram for R=100 and R=150 are shown.

The numerical study shows that the boundary value problem (4),(5) has a set of two coexisting spherically symmetric standing wave solutions. They are shown in figure 4(a,b). The first one is the Bessel-like wave without an explicitly localised in space core decaying in proportion to r^{-1} as $r \to R$) that branches off the zero solution, see figure 4a. The second type wave is characterised by an

exponentially localised in space and pulsating in time core with a small- amplitude slowly decaying in space second-harmonic tail as shown in figure 4b. The corresponding curves $\phi(r,0)$ for both waves are shown in figure 4c. Both solutions are shown at the point marked by the arrow in figure 4d: $\omega = 0.9802\,\omega_0$, E=45.585. Figure 4d demonstrates structure of a resonance spike and interconnection between the $E(\omega/\omega_0)$ branches of two types of waves. The Bessel-like branch extends to zero energy (red dashed curve in figure 4d). Both slopes of the resonance spike (blue solid curves in figure 4d) join the Bessel-like branch at points of period-doubling bifurcations.

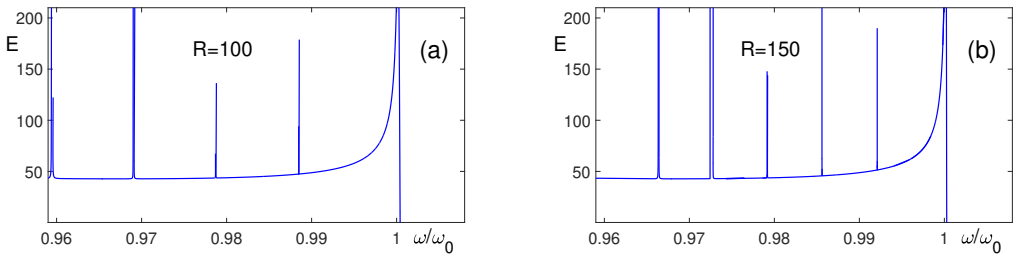


Figure 3. The energy-frequency diagram produced by numerical continuation of solutions of the boundary value problem (4), (5) for the cases R=100 (a) and R=150 (b)

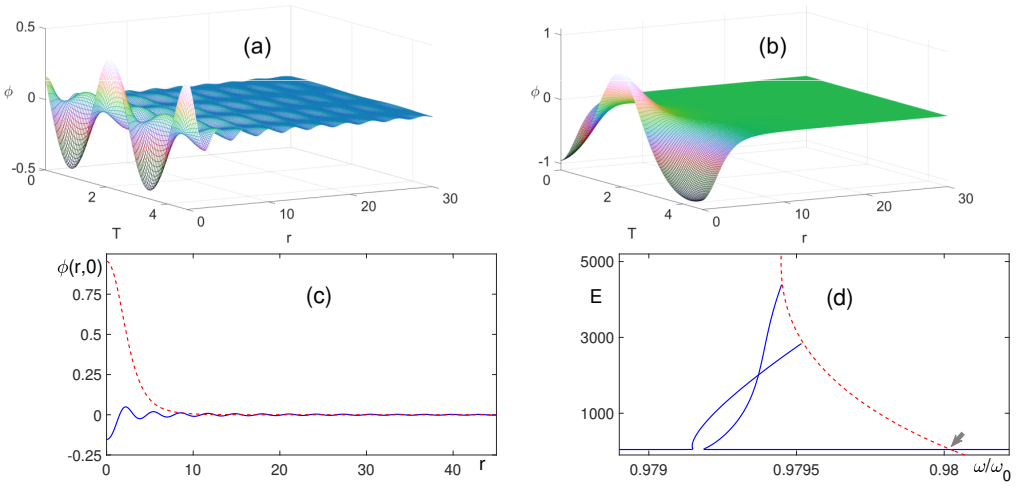


Figure 4. Coexisting standing waves in a ball of radius R=150 at the point $\omega = 0.9802\,\omega_0$: the Bessel-like wave (a) and the wave with exponentially localised pulsating core (b). (c): Curves $\phi(0,r)$ for the waves of the first (blue solid) and the second type (red dashed). (d): The $E(\omega/\omega_0)$ structure of the resonance spike region. The red dashed curve corresponds the Bessel-like wave. The arrow marks the point where coexisting waves (a) and (b) are shown

Stability analysis of the Bessel waves against the spherically symmetric perturbations shows that the Bessel waves are stable for low energy values starting from E=0 up to the bifurcation point at their intersection with either slope of the resonance spike. As for the second type of exponentially localised standing waves, the calculations in [8] show they have only short stability intervals between ω=0.9428ω_0 and ω=0.9435ω_0 inside the ball of radius R=40. In the case of a ball of larger radius, there are sections of stable frequencies of greater width. Note that the Bessel-like curve in figure 4d (red dashed) is stable for the values of E below the intersection with the left slope of the spike at the point $\omega = 0.97952\,\omega_0$, E=2879.3. All other solutions in figure 4d are found to be unstable.

It is important to emphasize that the curve $E(\omega)$ has a single minimum at $\omega_{\min} = 0.967\omega_0$ for all values of the radius R despite the fact that the number and the positions of the spikes are R-sensitive. Thus, we can consider the energy curve at least in the neighbourhoods away from the spikes to be

the approximation of the radius-independent envelope of the nearly-periodic oscillons in the infinite space.

Funding: This study was supported by the JINR-NRF Scientific Cooperation Program.

Acknowledgments: We thank the HybriLIT team for their help in organization of computations with the *HybriLIT* and *Govorun* resources.

References

1. Voronov, N. A., Kobzarev, I. Y. & Konyukhova, N. B. Possibility of the existence of X mesons of a new type. Russian. *JETP Letters* **22,** 590–594 (1975).
2. Bogolyubskii, I. & Makhankov, V. Lifetime of pulsating solitons in certain classical models. *JETP Letters* **24,** 12–14 (1976).
3. Bogolyubskii, I. & Makhankov, V. Dynamics of spherically symmetrical pulsons of large amplitude. *JETP Letters* **25,** 107–110 (1977).
4. Gleiser, M. Pseudostable bubbles. *Phys Rev D* **49,** 2978–2981. doi:10.1103/PhysRevD.49.2978 (6 1994).
5. Copeland, E., Gleiser, M. & Müller, H.-R. Oscillons: Resonant configurations during bubble collapse. *Phys Rev D* **52,** 1920–1933. doi:10.1103/PhysRevD.52.1920 (4 1995).
6. Honda, E. & Choptuik, M. Fine structure of oscillons in the spherically symmetric φ^4 Klein-Gordon model. *Phys Rev D* **65,** 084037. doi:10.1103/PhysRevD.65.084037 (8 2002).
7. Fodor, G., Forgácz, P., Grandclément, P. & Rácz, I. Oscillons and quasibreathers in the ϕ^4 Klein-Gordon model. *Phys Rev D* **74,** 124003. doi:10.1103/PhysRevD.74.124003 (12 2006).
8. Alexeeva, N., Barashenkov, I., Bogolubskaya, A. & Zemlyanaya, E. Understanding oscillons: Standing waves in a ball. *Phys Rev D* **107,** 076023. doi:10.1103/PhysRevD.107.076023 (7 2023).
9. Zhanlav, T. & Puzynin, I. Evolutionary Newton Procedure for Solving Nonlinear Equations. Russian. *Computational Mathematics and Mathematical Physics* **32,** 1–9. doi:10.1134/1.953099 (1992).
10. Puzynin, I., Boyadjiev, T., Vinitsky, S., Zemlyanaya, E., Puzynina, T. & Chuluunbaatar, O. Methods of computational physics for investigation of models of complex physical systems. *Physics of Particles and Nuclei* **38,** 70–116. doi:10.1134/S1063779607010030 (2007).

To cite: Zemlyanaya E., Bogolubskaya A., Bashashin M., Alexeeva N., Numerical study of the φ^4 standing waves in a ball of finite radius, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)106–111.DOI: 10.22363/2658-4670-2024-32-1-106-111.

Information about the authors

Zemlyanaya, Elena V.—Doctor of Physical and Mathematical Sciences, head of sector (e-mail: elena@jinr.ru, phone: +7(49621)2164728, ORCID: https://orcid.org/0000-0001-8149-9533, Scopus Author ID: 6701729810)

Bogolubskaya, Alla A.—Candidate of Physical and Mathematical Sciences, Senior Researcher (e-mail: abogol@jinr.ru, phone: +7(49621)2164015, ORCID: https://orcid.org/0000-0003-4356-8336)

Bashashin, Maxim V.—Junior Researcher (e-mail: bashashinmv@jinr.ru, phone: +7(49621)2163954, ORCID: https://orcid.org/0000-0002-2706-8668)

Alexeeva, Nora V.—Professor (e-mail: nora.alexeeva@uct.ac.za, phone: +27216503191, ORCID: https://orcid.org/0000-0001-9068-6023)

УДК 519.63, 519.688
PACS 07.05.Tp, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-106-111 EDN: CBNJTD

Численное исследование стоячих волн модели φ^4 в шаре конечного радиуса

Е. В. Земляная[1, 2], А. А. Боголюбская[1], М. В. Башашин[1, 2], Н. В. Алексеева[3]

[1] *Объединенный институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*
[2] *Государственный университет Дубна, ул. Университетская, д. 19, Дубна, 141980, Российская Федерация*
[3] *Кейптаунский университет, 7701, Рондебош, Южная Африка*

Аннотация. Проведено численное исследование сферически симметричных периодических по времени стоячих волн модели ϕ^4 в шаре конечного радиуса на основе вычисления решений сформулированной нелинейной краевой задачи на цилиндрической поверхности в широком диапазоне значений периода осцилляций и последующего анализа устойчивости полученных таким образом решений путем расчета соответствующих множителей Флоке. При этом стоячие волны в шаре конечного радиуса могут рассматриваться как аппроксимацияя слабоизлучающих сферически-симметричных осциллонов в модели ϕ^4. В работе описывается математическая постановка задачи и метод ее численного решения, обсуждается метод параллельной реализации расчета множителей Флоке на вычислительных ресурсах платформы HybriLIT Многофункционального информационно-вычислительного комплекса Объединенного института ядерных исследований (Дубна). Представлены результаты по исследованию пространственно-временной структуры и бифуркации сосуществующих стоячих волн различного типа.

Ключевые слова: осциллоны, компьютерное моделирование, параллельные вычисления

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 112–121
http://journals.rudn.ru/miph

Research article

UDC 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-112-121
EDN: BSGQHY

The numerical solution of the nonlinear hyperbolic-parabolic heat equation

Vladislav N. Khankhasaev[1,2], Safron A. Bairov[1]

[1] *East Siberia State University of Technology and Management, 40V Klyuchevskaya St, Ulan-Ude, 670013, Russian Federation*
[2] *Buryat State University, 24a Smolina St, Ulan-Ude, 670000, Russian Federation*

(received: January 15, 2024; revised: January 30, 2024; accepted: February 14, 2024)

Abstract. The article discusses a mathematical model and a finite-difference scheme for the heating process of an infinite plate. The disadvantages of using the classical parabolic heat equation for this case and the rationale for using the hyperbolic heat equation are given. The relationship between the hyperbolic thermal conductivity equation and the theory of equations with the retarded argument (delay equation) is shown. The considered mixed equation has 2 parts: parabolic and hyperbolic. Difference schemes use an integro-interpolation method to reduce errors. The problem with a nonlinear thermal conductivity coefficient was chosen as the initial boundary-value problem. The heat source in the parabolic part of the equation is equal to 0, and in the hyperbolic part of the equation sharp heating begins. The initial boundary-value problem with boundary conditions of the third kind in an infinite plate with nonlinear coefficients is formulated and numerically solved. An iterative method for solving the problem is described. A visual graph of the solution results is presented. A theoretical justification for the difference scheme is given. Also we consider the case of the nonlinear mixed equation of the fourth order.

Key words and phrases: hyperbolic-parabolic equation, delay equations, initial boundary-value problem, finite difference schemes, equations of the high order

1. Introduction

In the V.N. Khankhasaev's paper [1–3], which is bound up with the problem of mathematical modeling of the process of switching off the electric arc in the flue gas flow, various mathematical models bound up with the hyperbolic equation of thermal conductivity (obtained by generalization of the Fourier hypothesis [4]) were studied both analytically and numerically. In course of investigations bound up with the transfer processes in the case of high-intensity influence of the gas, the earlier hypotheses presuming the proportionality of the flow density to the vector of the potential gradient, which are based on the known physics laws, lead to an infinite rate of distribution of the perturbations, what contradicts to fundamental laws of nature.

The set known physics laws constructed on basis of the given theory includes the following laws:

$\overline{q}(x,y,z,t) = -\lambda \operatorname{grad} T(x,y,z,t)$ – the Fourier law; λ – thermal conductivity coefficient; $\overline{q}$ – heat flow density; T – temperature;

$\overline{q}(x,y,z,t) = -D \operatorname{grad} C(x,y,z,t)$ – the Fick law; D– diffusion coefficient; $\overline{q}$ – flow density of diffusion; C – concentration;

$\overline{q}(x,y,z,t) = -K \operatorname{grad} H(x,y,z,t)$ – the Darcy law; K – filtering coefficient; $\overline{q}$ – volume flow (or the filtering rate); H – pressure.

All these laws are written in the general form as follows:

$\overline{q}(x,y,z,t) = -A \operatorname{grad} U(x,y,z,t)$, i.e. the generalized law of transfer, where A is the transfer coefficient; $\overline{q}$ is the flow density; U is the potential.

© Khankhasaev V. N., Bairov S. A., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

The differential equation of transfer obtained from the given generalized law has the following form in the one-dimensional case:

$$\frac{\partial U}{\partial t} = a\frac{\partial^2 U}{\partial x^2}.$$

It is sufficient to differentiate the fundamental solution of this equation with respect to variable t and tend the time to zero to see that the rate of transfer of the potential at the initial time moment equals to infinity. The approximation of solid medium used in above laws and presuming the absence of its internal molecular structure implies that it is possible to undertake a limit transition in the integral laws of conservation for this medium, when the volume tends to zero. Such a limit transition allows one to obtain the equation of energy conservation in the differential form. Meanwhile, this procedure – from the viewpoint of contemporary physics – is incorrect because the environment is known to be composed of molecules. The environment has a discrete internal structure.

In order to avoid this paradox J.C. Maxwell [5], C. Cattaneo [6], P. Vernotte [7], who worked within the frames of the theory of thermal conductivity, based his reasoning on the molecular-kinetic conception, used the hypothesis of finiteness of duration of molecular collisions and proceeded from a new conception of the molecules' length of free path, obtained a new law of thermal conductivity. There appeared an additional addend $\tau\frac{\partial q}{\partial t}$ in the law, which took account of the discreteness of the environment's molecular structure and was responsible for the inertial character of heat. In this addend, is the relaxation time, i.e. the time of reaching some thermodynamic equilibrium between the heat flow and the temperature gradient. This generalized law of transfer may be written in the following form:

$$\tau\frac{\partial q}{\partial t} + q = -A\,\mathrm{grad}\,U. \tag{1}$$

In the process of solving the differential equation obtained from this law, observed was the first-kind discontinuity of the potential, which distributes from the source. Therefore, law (1) describes the appearance of waves in case of some high-intensity influence, which leads to some local system's non-equilibrium. These effects are most frequently observed when a body is impacted with short energy impulses, in shock waves or under high temperature gradients. The local equilibrium, which is obvious in cases of application of earlier physics laws, is valid for the time moments (intervals), which are in excess of the relaxation time. Therefore, classical transfer theories are valid, when the rate of processes is substantially smaller than the rate of distribution of perturbations in the medium under scrutiny [8, 9].

2. The relationship between the hyperbolic thermal conductivity equation and the theory of equations with the retarded argument (delay equations)

To the end of inference of the transfer differential equation in the one-dimensional case we have used the equation of thermal balance:

$$\frac{\partial q}{\partial x} = -g\frac{\partial U}{\partial t}. \tag{2}$$

Having substituted (1) into (2), we obtain:

$$A\frac{\partial^2 U}{\partial x^2} + \tau\frac{\partial^2 q}{\partial x \partial t} = g\frac{\partial U}{\partial t}. \tag{3}$$

Now, change the order of differentiation for the second addend in (3)

$$A\frac{\partial^2 U}{\partial x^2} + \tau\frac{\partial}{\partial t}\left(\frac{\partial q}{\partial x}\right) = g\frac{\partial U}{\partial t}.$$

On account of (2) we obtain

$$A\frac{\partial^2 U}{\partial x^2} - \tau g\frac{\partial}{\partial t}\left(\frac{\partial U}{\partial t}\right) = g\frac{\partial U}{\partial t}.$$

Finally, we obtain the following:

$$\tau\frac{\partial^2 U}{\partial t^2}+\frac{\partial U}{\partial t}=\frac{A}{g}\frac{\partial^2 U}{\partial x^2}. \tag{4}$$

Equation (4) belongs to the class of linear hyperbolic partial differential equations because it contains the second derivative with respect to time. If internal sources of perturbations are taken into account, then equation (2) assumes the following form:

$$-g\frac{\partial U}{\partial t}+F(U)=\frac{\partial q}{\partial x},$$

Hence equation (4) writes as follows:

$$\tau\frac{\partial^2 U}{\partial t^2}+\left(1-\frac{\tau}{g}\frac{dF}{dU}\right)\frac{\partial U}{\partial t}=\frac{A}{g}\frac{\partial^2 U}{\partial x^2}+\frac{F(U)}{g}. \tag{5}$$

Here $\frac{\partial F}{\partial U}$ may have any sign.

While turning back to equation (1), one can see that $-A\,\mathrm{grad}\,U$ represents an expansion of the flow into the Tailor series with respect to the powers of τ, where taken are only the first two members of the expansion. Hence if all the terms of the expansion are taken into account, then the series shall have the following form:

$$q+\tau\frac{\partial q}{\partial t}+\frac{\tau^2}{2}\frac{\partial^2 q}{\partial t^2}+\ldots=-A\,\mathrm{grad}\,U. \tag{6}$$

Having gathered the terms of series, we can rewrite expression (6) in the following form:

$$q(t+\tau)=-A\,\mathrm{grad}\,U. \tag{7}$$

Having replaced the variables $t+\tau=t^1$ in (7), and, next, again transferring to variable t, we obtain:

$$q(t)=-A\,\mathrm{grad}\,U(t-\tau). \tag{8}$$

The physics sense of expression (8) implies that the transfer process in the locally non-homogeneous media possesses inertial properties: the system reacts to the influence not at the same time moment but with a delay equal to the relaxation time τ, i.e. the flow density retards from the gradient of potential. From the technical viewpoint, expression (8), unlike that for (1), allows one to take account of all the terms of the expansion with respect to τ. While continuing the above inference technique with the use of (8), one can easily obtain the following equation with the retarded (with respect to time) argument[10–12]:

$$\frac{\partial U}{\partial t}=\frac{A}{g}\frac{\partial^2 U(x,t-\tau)}{\partial x^2}.$$

Therefore, the linear hyperbolic thermal conductivity equation (5) represents the second, more correct stage in the theory of mathematical modeling of heat transfer for the fast running processes with high-intensity perturbations.

The investigation of thermal conductivity processes using the generalized Fourier law is most relevant for rapidly occurring physical phenomena (for example, with nano- and fempto-second laser pulses) in the study of high-intensity processes of heating bodies (plasma, laser processing of materials, high-intensity heating of contact connections in electrical installations and etc.) [13, 14].

3. The nonlinear mixed equation of thermal conductivity

An unbounded plate is given in the form of an infinite strip, the size of which along the x axis is equal to the segment $[0, X]$, and the size along the y axis is equal to $(-\infty, \infty)$. We consider the properties of the plate along the y axis to be homogeneous and we will not mention y in the list of variables. The initial temperature distribution in the plate is given by some function $u(x, T_1) = u_0(x)$; at the plate boundaries the temperature of the medium is constant. Heat exchange with the environment occurs according to Newton's law (boundary conditions of the third kind). The thermophysical characteristics

$c_v, \lambda, \rho, \alpha, c$ are specified – specific heat capacity, thermal conductivity coefficient, specific density, heat transfer coefficient and heat output coefficient. It is required to find the temperature distribution over the thickness of the plate, i.e. by variable x, at any time $t \in [T_1, T_2]$. The differential equation and boundary conditions will then be written as:

$$k(x,t)u_{tt} + c_v(x,t)\cdot\rho(x,t)u_t = (\lambda(u,x,t)u_x)_x + c(x,t)u + f(x,t). \tag{9}$$

In the rectangular domain $G = [0,X]\times[T_1,T_2]$, $T_1 < 0$, $T_2 > 0$. Furthermore, $\forall(x,t)\in G$, $k(x,t) = 0$, $t \le 0$; $k(x,t) > 0$, $t > 0$; i.e. when $t \le 0$ — the equation(1) is parabolic, and when $t > 0$ — the equation(1) is hyperbolic. Let us formulate the problem with the following experimental data [15].

The initial boundary-value problem. Find the temperature field in an infinite plate homogeneous in variable y with $X = \pi$ and calculation time: $T_1 = -5$, $T_2 = 20$.

Initial condition and boundary conditions:

$$u(x,t)|_{t=T_1} = u_0(x) = 10\sin(x);$$

$$\left[\mp\lambda(u,x,t)\frac{\partial u(x,t)}{\partial x} + \alpha_{0,L}(x,t)\,u(x,t)\right]_{x=0,L} = \begin{cases} q_0(t), \\ q_L(t). \end{cases} \tag{10}$$

Coefficients: $k(x,t) = 0$ for $t \le 0$, $k(x,t) = 1$ for $t > 0$; $c(x,t) = 0$; $\lambda(u,x,t) = 0.5\cdot u^2 + 2$; $c_v(x,t)\cdot\rho(x,t) = a(x,t) = 672$; $q_0(t) = 5$; $q_L(t) = 10$; $\alpha_0(x,t) = 3.5$; $\alpha_L(x,t) = 3.5$.

Heat sources $f(x,t)$ change over time. In the parabolic part $f(x,t) = f1(x,t) = 0$ and in the hyperbolic part $f(x,t) = f2(x,t) = 100000\sin(x)\sin(t)$.

In the quasi-linear scheme, the coefficients λ are calculated from the temperatures $U_{i,j}$ of the previous time layer j, while in the essentially non-linear scheme, which is being implemented now:

$$\begin{aligned} k(i\cdot h, T_1 + j\cdot\tau)\frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\tau^2} + a(i\cdot h, T_1 + j\cdot\tau)\frac{u_{i,j+1} - u_{i,j}}{\tau} = \\ = \left(\lambda\left(u_{i+\frac{1}{2},j+1}, i\cdot h, T_1 + j\cdot\tau\right)_{i+\frac{1}{2}}\frac{u_{i+1,j+1} - u_{i,j+1}}{h} - \right. \\ \left. - \lambda\left(u_{i-\frac{1}{2},j+1}, i\cdot h, T_1 + j\cdot\tau\right)_{i-\frac{1}{2}}\frac{u_{i,j+1} - u_{i-1,j+1}}{h}\right)\frac{1}{h} + \\ + c(i\cdot h, T_1 + j\cdot\tau)u_{i,j+1} + \int_{t_j}^{t_{j+1}}\int_{x_{i-1/2}}^{x_{i+1/2}} f(x,t)dxdt. \end{aligned} \tag{11}$$

The coefficient λ is calculated as follows:

$$\lambda(u_{i\pm\frac{1}{2},j+1}, i\cdot h, T_1 + j\cdot\tau) = \frac{[\lambda(u_{i\pm1,j}, (i\pm1)\cdot h, T_1 + j\cdot\tau) + \lambda(u_{i,j}, i\cdot h, T_1 + j\cdot\tau)]}{2}.$$

It is clear that the resulting system of equations is nonlinear, so to solve this system we will use the simple iteration method. This method is as follows—at each time step we will determine the temperature field until it stops changing with changes:

$$\frac{\max\limits_i|u_{i,s+1} - u_{i,s}|}{\max\limits_i|u_{i,s+1}|} < \epsilon, \tag{12}$$

where s is the iteration number, ϵ is the calculation accuracy. When condition (12) is satisfied, then $u_{i,s+1} = u_{i,j+1}$. The following can be considered as an initial approximation: $u_{i,s=0} = u_{i,j}$.

It can be seen that the system (11) is already linear with respect to $u_{i,s+1}$, which makes it possible to use the sweep method and determine the unknown temperature field. But in this case the system is solved until the temperature field ceases to differ [16]. In such a scheme, the volume of calculations increases compared to a quasi-linear scheme, since at each time step it is necessary to solve the system of difference equations by the sweep method not once, but $s_{\max}$ times. However, the nonlinear

scheme gives a smaller error in the numerical solution of the original problem (9), (10) than the quasilinear one [17]. This is explained by the fact that the coefficients in the expressions for the grid analogues of heat flows are calculated at the same time as the temperatures. To reduce the error of a quasilinear scheme, the step size should be reduced, i.e., the number of time steps in the considered interval should be increased. Therefore, in many cases it turns out to be more profitable, even from the point of view of computer time costs, to use a nonlinear scheme and take larger time steps, performing several iterations at each [18].

The fields of temperatures for the scrutinized process have been obtained at various time moments (fig.1) in the environment of Mathcad-15 having a comfortable graphic interface. Similar to the works[19, 20] the following theorems is proved:

Theorem 1. Let function $c(x,t) < 0$ is sufficiently large with respect to the modulus,

$$2a - |k_t| \geqslant \delta > 0.$$

Hence for any function $f \in W_2^1(G)$ there exists a unique solution $u(x,t)$ of the initial boundary-value problem (9), (10) in space $W_2^2(G)$.

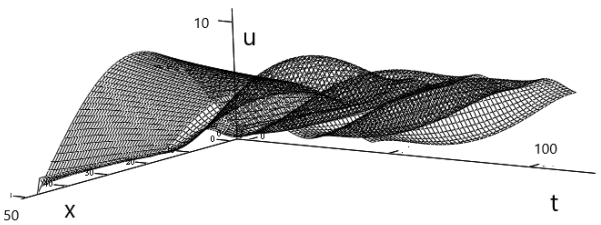


Figure 1. The result of the program solution

Theorem 2. Under the conditions of Theorem 1 the difference scheme (11) is stable, and interpolations $u_h^\tau(x,t)$ of solutions of this difference scheme converge weakly in $W_2^1(G)$ (when $h \to 0, \tau \to 0$) to the solution $u(x,t)$ of initial boundary-value problem (9), (10) from space $W_2^2(G)$.

4. The nonlinear mixed equation of the fourth order

In bounded domain D from R^n, consider the first boundary value problem for the fourth order nonlinear mixed partial differential equation:

$$Lu \equiv \sum_{i=0}^{n+1} L_i^* A_i(x, u, u_{x_1}, \dots, u_{x_n}, Ku) = h(x), \tag{13}$$

$$u|_\Gamma = f_1(x), \frac{\partial u}{\partial \nu}|_{\Gamma/\Gamma_0} = \sum_{i,j=1}^{n} \left(a_{i,j} \frac{\partial u}{\partial x_i} \nu_j \right)\Big|_{\Gamma/\Gamma_0} = f_2(x),$$

here: L^* is an operator formally Lagrange conjugate to the operator L; L_0 — identity operator; $L_i = \frac{\partial}{\partial x_i}$; $i = 1, \dots, n$; L_{n+1} — operator K of the form:

$$Ku \equiv \sum_{i,j=1}^{n} a_{i,j}(x) u_{x_i,x_j} + \sum_{i=1}^{n} b_i(x) u_{x_i} + c(x)u$$

with sufficiently smooth coefficients, satisfying the inequality:

$$\|Ku\|_{L_m(D)} \geqslant \alpha_1 \|u\|_{W_r^1(D)}, \quad \alpha_1 > 0, \quad m \geqslant 2, \quad r \geqslant 2, \tag{14}$$

for any functions $u(x)$ from C_K the class of twice continuously differentiable functions vanishing on the boundary Γ of the domain D, Γ_0 is the characteristic part of the boundary of Γ for the operator K,

$\vec{\nu} = (\nu_1, \dots, \nu_n)$ is the vector of the internal normal to Γ:

$$\Gamma_0 = \left\{ x \in \Gamma : \left(\sum_{i,j=1}^{n} a_{i,j} \nu_i \nu_j \right)(x) = 0 \right\}. \tag{15}$$

As the operator K, we can take the linear hyperbolic-parabolic heat conduction operator described above in (9).

Lemma 1. For any function $u(x)$ from C_K we derive inequality (14) with parameters $m = 2$ and $r = 2$ if the condition is met:

$$2a - |k_t| \geqslant \delta > 0. \tag{16}$$

Lemma 2. For any function $u(x)$ from C_K, if condition (16) is met, the following inequality with parameter $m = 2$ is deduced:

$$\|Ku\|_{L_m(D)} \geqslant \alpha_2 \left\| \frac{\partial u}{\partial N} \right\|_{L_2(\Gamma)}. \tag{17}$$

Let us define the Banach spaces H_+ and $H_\#$ with standards: $\|u\|_+ = \|Ku\|_{L_m(D)}$; $\|u\|_\oplus = \|Ku\|_{L_m(D)} + \|u\|_{W_e^1(D)}$, obtained by closing a set of functions from

$$C_L = \left(u \in C_K : \left. \frac{\partial u}{\partial N} \right|_{\Gamma/\Gamma_0} = 0 \right).$$

From (14) it follows that these are indeed the norms and spaces H_+ and $H_\#$ are obviously separable. With the help of the Clarkson's inequalities is proved the next lemma.

Lemma 3. Spaces H_+ and $H_\#$ reflective.

From the embedding theorems for Sobolev spaces it follows that functions from the spaces H_+ and $H_\#$ vanish on the entire boundary of Γ. Equality (15) means that on Γ_0 the derivative along the conormal is the tangent derivative to the boundary Γ and on functions from C_K vanishes on Γ_0, therefore inequality (17) actually means

$$\|u\|_+ = \|Ku\|_{L_m(D)} \geqslant \alpha_3 \left\| \frac{\partial u}{\partial N} \right\|_{L_2(\Gamma/\Gamma_0)}.$$

After introducing a continuous trace operator based on inequality (17) of Lemma 2 on functions from C_K and extending it by continuity to the spaces H_+ and $H_\#$, we find that for functions from H_+ and $H_\#$ the derivative with respect to the conormal vanishes in the space $L_2(\Gamma/\Gamma_0)$.

Suppose that the functions $f_1(x)$, $f_2(x)$ admit continuation $f(x)$ inside the region D from the space $W_m^2(D) \cap W_k^1(D)$, where $k = \max(r, e)$. Then a collection of functions of the form $u(x) = z(x) + f(x)$, where $z(x)$ from $H_+(H_\#)$, forms the space $H_+(f)[H_\#(f)]$.

Definition 1. Function $u(x)$ from $H_+(f)[H_\#(f)]$ let's call it a weak generalized solution to problem (13) if the identity holds:

$$\sum_{i=0}^{n+1} \int_D A_i(x, u, u_{x_1}, \dots, u_{x_n}), Ku) L_i \nu, \quad dD = \sum_{i=0}^{n+1} (A_i(x, L_i u), L_i \nu) = (h, \nu),$$
$$\forall \nu(x) \in C_L, \quad j = \overline{0, n+1}.$$

Definition 2. Function $u(x)$ from $H_+(f)[H_\#(f)]$ we call a strong generalized solution to problem (13) if there is a sequence of functions $z_i(x) \in C_L$ such that

$$\lim_{i \to \infty} \|z_i + f - u\|_{+[\oplus]} = \lim_{i \to \infty} \|L(z_i + f) - h\|_{-[\ominus]} = 0,$$

where $H_-(D)[H_\ominus(D)]$ are the negative spaces to $H_+(f)[H_\#(f)]$, constructed with respect to the Hilbert space $L_2(D)$.

Let us present a number of assumptions for various equations of the form (13), which essentially mean conditions on the behavior of nonlinear functions $A_i(x, \xi_j)$, $i, j = \overline{0, n+1}$, $\xi \in R^{n+1}$.

1. Conditions of limitation and continuity: $L : H_+(f) \to H_-(D)$.

The functions $A_i(x, \xi_j), i, j = \overline{0, n+1}$ satisfy the Caratheodory conditions, i.e. for almost all x from D are continuous in the set of variables ξ_j, for all values ξ_j are measurable in x and satisfy the inequalities:

$$A_i(x, \xi_j) \leqslant \alpha_4 \left(a(x) + \sum_{j=0}^{n+1} |\xi_j|^{p_{ij}} \right),$$

where $p_{i,j}$ are selected indicators.

2. Condition for coercivity of the operator Lu. For any function $u(x)$ from $H_+(f)[H_\#(f)]$ the following inequality holds:

$$\sum_{i=0}^{n+1} (A_i(x, L_j u), L_i u) \leqslant \alpha_5 \|u\|_+^m - \alpha_6, j = \overline{0, n+1},$$

$$\left[\sum_{i=0}^{n+1} (A_i(x, L_j u), L_i u) \leqslant \alpha_7 (\|u\|_+^m + \|u\|_{W_e^1(D)}^e) - \alpha_8 \right].$$

3. Condition for the definiteness of the variation of the operator Lu. For any functions $u(x)$, $v(x)$ from $H_+(f)[H_\#(f)]$ the following inequality holds:

$$(Lu - Lv, u - v) \geqslant \alpha_9 \|u - v\|_+^m,$$

$$[(Lu - Lv, u - v) \geqslant \alpha_{10} (\|u - v\|_+^m + \|u - v\|_{W_e^1(D)}^e)].$$

Similar to the work of Dubinsky Yu.A. the following theorem is proved:

Theorem 3: If assumptions 1) – 3) are met, then the first boundary value problem (13) for any function $h(x) \in H_-(D)[H_\ominus(D)]$ is set correctly, the weak solution coincides with the strong one, i.e. a mapping $Lu = h(x) \in H_-(D)[H_\ominus(D)]$ is a homeomorphism[21].

5. Conclusion

A program has been written to solve the mixed heat equation using the simple iteration method. A calculation was carried out with similar boundary conditions. The results coincide with the results of the first miscalculation.

There are other methods that have become widespread in practice for constructing an iterative process for solving systems of nonlinear difference equations. For example, Newton's method is based on the linearization of equations and is usually used in the case when the dependences of the coefficients on temperature are specified by analytical dependencies that can be differentiated. In further work, this method will be used to solve essentially nonlinear equations of mixed type.

Along with numerous methods for solving inverse coefficient problems for linear and nonlinear second order equations $K(u) = h$ you can also use the one proposed by Yu.A. Dubinsky. approach when this equation, which is generally unsolvable for an arbitrary right-hand side h, is associated with some 4th order equation $K^* K(u) = K^* h$, which is always solvable. Then the equation $K(u) = h$ is solvable up to the kernel of the operator K^*.

This construction can also be considered as a technique for describing the range of values of the mixed heat operator corresponding to an ill-posed problem with overdetermination. The presence of these additional boundary conditions, taking into account the release of some of them on the characteristic surfaces of the operator $K(u)$, is necessary for the numerical solution of the well-posed Dirichlet problem for the equation $K^* K(u) = K^* h$, if the operator $K^* K$ implements a homeomorphism.

Funding: The work was carried out with the financial support of the Russian Science Foundation grant No. 23-21-00269, https://rscf.ru/project/23-21-00269.

References

1. Buyantuyev, S. L. & Khankhasaev, V. N. About one generalization of Navier-Stokes equation in mathematical models an electric arc. *Elektrichestvo ISSN: 0013-5380 eISSN: 2411-1333* (1996).
2. Khankhasaev, V. N. & Buyantuyev, S. L. The numerical computation of one mathematical model of electric arc in the gas flow. *Proceedings of International Conference "Energy Conserving and Nature Protecting Technologies for the Baikal Lake Territory"* (2001).
3. Khankhasaev, V. N. & Darmakheev, E. V. On certain applications of the hyperbolic heat transfer equation and methods for its solution. *Mathematical Notes NEFU.* doi:10.25587/SVFU.2018.1.12772 (2018).
4. Shashkov, A. G., Bubnov, V. A. & Yanovskiy, S. Y. *Wave phenomena of thermal conductivity. System-structural approach* 298 pp. (URSS, Moscow, 2004).
5. Maxwell, J. C. On the Dynamic theory of gases. *Philosophical Transactions* **157** (1867).
6. Cattaneo, C. Sulla conduzione de calore. *Atti del seminario matematico e fisiko dell'Universita di Modena* **3** (1948).
7. Vernotte, P. Les paradoxes de la theorie continue de l'equation de la chauleur. *Compt. Rendus.* **246** (1958).
8. Bubnov, V. A. Molecular-kinetics basis for the heat-transfer equation. *Journal of engineering physics* **28** (1975).
9. Sobolev, S. L. Local non-equilibrium transport models. *Successes of Physical Sciences* **167** (1997).
10. Formalev, V. F., Selin, I. A. & Kuznetsova, E. L. Appearance and distribution of heat waves in the nonlinear anisotropic space. *Izvestiya of Russian Academy of Sciences. Ser. "Energy Problems"* (2010).
11. Pimenov, V. G. & Lozhnikov, A. B. Numerical method for modeling controlled heat conduction equation with delay. *Bulletin of Russian Universities. Mathematics.* **18** (2013).
12. Khankhasaev, V. N., Mizhidon, G. A. & Kozhanov, A. I. On the connection between some heat and mass transfer models and the theory of equations with the relarded argument. *Proceedings of the VII International Conference "Mathematics, its Applications and Education in Mathematics" ISBN 978-5-907331-29-7* (2020).
13. Petrova, L. S. Mathematical modeling of heating processes of piecewise homogeneous bodies taking into account the relaxation of heat flow. *Bulletin of Eurasian Science* **9** (2017).
14. Scharf, G. Approach to steady state in the heat equation and the hyperbolic heat transfer equation. *Internet journal "arXiv"* (2018).
15. Voskoboinikov, Y. E. *et al. Solving engineering problems in MathCAD package* 190 pp. (State architecture-builds University (Sibstrin), Novosibirsk, 2013).
16. Kuznetsov, G. V. & Sheremet, M. A. *Difference methods for solving heat conduction problems* (2007).
17. Khankhasaev, V. N. & Bairov, S. A. Solution of a mixed equation nonlinear with respect to the thermal conductivity coefficient by the integro-interpolation method with the third boundary condition. *Mathematics, its applications and mathematical education (MAME23): Proceedings of the VIII International Conference.* doi:10.53980/9785907599970 (2023).
18. Dulnev, G. D., Parfenov, V. G. & Sigalov, A. V. *Application of computers to solve heat transfer problems* 207 pp. (High School, Moscow, 1990).
19. Terehov, A. N. Boundary value problem for a mixed type equation. *Collection "Application of methods of functional analysis to problems in mathematics, physical and calc. math."* (1979).
20. Khankhasaev, V. N. & Mestnikova, N. N. An algorithm of numerical solving of heat transfer hyperbolic-parabolic equation one-dimensional with respect to the space variable in the environment of Mathcad 15. *Proceedings of the International conference "Nano-materials and Technologies–5"* (2014).
21. Khankhasaev, V. N. On the theory of nonlinear equations of mixed type of the fourth order. *Application of functional analysis methods to non-classical equations of mathematical physics* (1988).

To cite: Khankhasaev V. N., Bairov S. A., The numerical solution of the nonlinear hyperbolic-parabolic heat equation, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)112–121.DOI: 10.22363/2658-4670-2024-32-1-112-121.

Information about the authors

Khankhasaev, Vladislav N. —Candidate of Physical and Mathematical Sciences, assistant professor of Department of "Mathematics named after Ts.B. Shoynzhurov" of East Siberia State University of Technology and Management, assistant professor of Department of Fundamental Mathematics of Buryat State University (e-mail: hanhvladnick@mail.ru, phone: +7(914)9862967, ORCID: https://orcid.org/0000-001-6840-8270, ResearcherID: HKV-0681-2023, Scopus Author ID: 6504419222)

Bairov, Safron A.—Postgraduate student of Department of "Mathematics named after Ts.B. Shoynzhurov" of East Siberia State University of Technology and Management (e-mail: bairov.sofron@gmail.com, phone: +7(983)4563505, ORCID: https://orcid.org/0009-0002-6638-5073)

УДК 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2024-32-1-112-121 EDN: BSGQHY

Численное решение нелинейного гиперболо-параболического уравнения теплопроводности

В. Н. Ханхасаев[1,2], С. А. Баиров[1]

[1] *Восточно-Сибирский государственный университет технологий и управления, ул. Ключевская, д. 40В, стр. 1, г.Улан-Удэ, 670013, Российская Федерация*
[2] *Бурятский государственный университет имени Доржи Банзарова, ул. Смолина, д. 24А, Улан-Удэ, 670000, Российская Федерация*

Аннотация. В статье рассматривается математическая модель и конечно-разностная схема процесса нагрева бесконечной пластины. Приводятся недостатки использования классического параболического уравнения теплопроводности для данного случая и обоснования для использования смешанного уравнения. Показана связь гиперболического уравнения теплопроводности с теорией уравнений с запаздывающим аргументом (уравнением с запаздыванием). В смешанном уравнении присутствуют 2 части: параболическая и гиперболическая. В разностных схемах применяется интегро-интерполяционный метод для уменьшения погрешностей. В качестве краевой задачи выбрана задача с нелинейным коэффициентом теплопроводности. Источник тепла в параболической части уравнения равен 0, а в гиперболической части уравнения начинается резкий нагрев. Поставлена и численно решена начально-краевая задача с краевыми условиями третьего рода в бесконечной пластине с линейными и с нелинейными коэффициентами. Описан итерационный метод для решения задачи. Представлен наглядный график результатов решения. Дано теоретическое обоснование для разностной схемы. Также рассмотрен случай нелинейного смешанного уравнения четвертого порядка.

Ключевые слова: гиперболо-параболическое уравнение, уравнения с запаздыванием, начально-краевая задача, конечно-разностные схемы, уравнения высокого порядка

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (1) 122–127
http://journals.rudn.ru/miph

Research article

UDC 519.21
PACS 52.25.Fi
DOI: 10.22363/2658-4670-2024-32-1-122-127

EDN: BBLNGK

On cyclotron damping of longitudinal wave

Sergey P. Karnilovich[1], Konstantin P. Lovetskiy[1], Leonid A. Sevastianov[1,2], Svetlana B. Strashnova[1], Yahya N. Shaar[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*
[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

(received: February 28, 2024; revised: March 20, 2024; accepted: March 30, 2024)

Abstract. Average equations of motion of relativistic charged particles in the field of HF (high frequency) wave packets are obtained in the range of cyclotron resonance in the case of strong LF (low frequency) electric field. Strong electric field means that the characteristic velocity of the particle comparable with the electric drift velocity ($v \sim v_E$). It is shown that with taking into account the electric drift velocity, new mechanisms of damping of longitudinal waves become possible. The effect of a strong electrostatic field on the resonant interaction of relativistic particles with high-frequency waves, as well as the relativistic effect, on cyclotron resonance for a longitudinal wave, is analyzed. The analytical solution of the averaged system of equations in the quasi-relativistic approximation is analyzed, as well as a numerical experiment for the Langmuir wave under the condition of cyclotron resonance in the case of a strong electric field.

Key words and phrases: Electric drift velocity, damping, relativistic charged particles, strong electric field, longitudinal waves, high frequency wave packets, cyclotron resonance

1. Introduction

In the drift theory of the motion of charged particles in electromagnetic fields, two cases are distinguished: a "weak" electric field, when the velocity of electric drift is $V_E \sim \varepsilon V$, and a "strong" electric field, when $V_E \sim V$. Here V is the characteristic velocity of the particle, ε is a small parameter equal to the ratio of the gyroradius of the particle to the characteristic scale of inhomogeneity of a strong magnetic field. The case of a strong electrostatic field is fraught with certain difficulties even in the non-relativistic approximation [1]. In [1–4], a theory of the motion of charged particles in the field of wave packets in crossed electric and magnetic fields was constructed taking into account weakly relativistic effects.

In this paper, the interaction of cyclotron resonance for longitudinal waves propagating along a strong magnetic field is considered. It is assumed that the rate of electric drift is small compared to the speed of light in vacuum. Such a proposal is quite sufficient for solving many applied problems. The equations of motion of an advertising charged particle averaged over fast oscillations are obtained, taking into account the effects of quasi-stationary electric drift under the condition of cyclotron resonance.

2. Basic equations

The motion of a particle with charge q and rest mass m under the influence of HF (high frequency) electromagnetic field $\vec{E}_{\sim},\vec{B}_{\sim}$ and LF fields $\vec{E}_0,\vec{B}_0$ is described by the equations [1]

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \alpha\vec{F} - \frac{1}{c^2}\vec{v}(\vec{v}' \cdot \vec{F}) + \alpha[\vec{\Omega} \cdot \vec{v}], \tag{1}$$

© Karnilovich S. P., Lovetskiy K. P., Sevastianov L. A., Strashnova S. B., Shaar Y. N., 2024

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

where $\vec{F}_{0,s} = \frac{q}{m}\vec{E}_{0,s}$, $\vec{\Omega}_{0,s} = \frac{q\vec{B}_{0,s}}{mc}$, $\alpha = 1 - \frac{v^2}{2c^2}$, $\vec{F} = \vec{F}_0 + \sum \vec{F}_s$, $\vec{\Omega} = \vec{\Omega}_0 + \sum \vec{\Omega}_s$, c – the speed of light.

The velocity vector has the form

$$\vec{v}' = v_{||}\vec{e}_1 + \vec{v}_E + \vec{v}_\perp(\vec{e}_2 \cos\theta_0 + \vec{e}_3 \sin\theta_0), \tag{2}$$

where Θ_0 is the gyrophase, $\vec{e}_1(\vec{r},t) = \vec{B}_0/B_0$, $\vec{e}_2(\vec{r},t)$, $\vec{e}_3(\vec{r},t)$ are the unit vectors. $\vec{v}_E = c[\vec{E}_0 \cdot \vec{e}_1]/B_0$ is the electric drift velocity. The electromagnetic fields $\vec{E}_\sim$,$\vec{B}_\sim$, are considered in eikonal approximation as:

$$\vec{E}_s = \sum \vec{\mathcal{E}}_s e^{i\Theta_s} + c.c., \quad \vec{B}_s = \sum \vec{\mathcal{B}}_s e^{i\Theta_s} + c.c., \quad 1 \leqslant s \leqslant m. \tag{3}$$

Here $\vec{\mathcal{E}}_s$,$\vec{\mathcal{B}}_s$ are slowly varying complex amplitudes and Θ_s is the fast phase (eikonal) of the s^{th} wave packets ($s = 1, 2, 3, ..., M$).

Phases Θ_s are considered as the independent variables which are described by the equations:

$$\frac{d\Theta_s}{dt} = \omega_s + \frac{d\vec{r}}{dt} \cdot \vec{k}_s = v_s + \frac{v_\perp}{2}(\vec{e}_- e^{i\Theta_0} + c.c.)\vec{k}_s + \vec{k}_s\vec{v}_E. \tag{4}$$

The quantities

$$\omega_s(\vec{r},t) = -\frac{d\Theta_s}{dt}, \quad \vec{k}_s(\vec{r},t) = \nabla\Theta_s, \tag{5}$$

are the local frequency and the wave vector of the s^{th} wave packets, respectively.

Equations (1)–(4) together with Eq. (3) constitute a multi–periodic system, which can be simplified by smoothing over fast and nonresonant phases [1].

3. Average equations

In the range of cyclotron resonance, the corresponding combination of the phases $\Psi_{\text{res}} = \Theta_0 + \Theta_s$, shold be corresponded as an "semifast" variable and an equation for resonant phases Ψ_{res} should be added to the equations for slow dynamic variables of particles. Smoothed equations of motion for a single particle interacting with the arbitrary s^{th} wave packet at the condition of the cyclotron resonance $v + \omega \cong 0$ have the form:

$$\frac{d\vec{R}}{dt} = \vec{v}_E + \vec{e}_1 v_{||} \equiv \vec{u}_0, \tag{6}$$

$$\frac{dv_{||}}{dt} = \vec{v}_E\vec{e}_1' + \frac{v_\perp}{2}\text{div}\vec{e}_1 + \left(\Gamma - \frac{v_{||}^2}{c^2}\right)F_{0||} - \\ - \left\{\frac{i}{2}v_\perp\vec{\Omega} + \frac{v_\perp}{2c^2}(F_{s||} + \vec{e}_1[\vec{\Omega}_{s1}\vec{v}_E])\vec{v}_E + \frac{v_\perp v_{||}}{2c^2}\vec{F}_s\right\}\vec{e}_- e^{i\psi_{\text{res}}} + c.c., \tag{7}$$

$$\frac{dv_\perp}{dt} = \frac{v_\perp}{2}(\text{div}\vec{u}_0 - (\vec{e}_1(\vec{e}_1\nabla)\vec{u}_0)) + \frac{v_\perp v_{||}}{c^2}F_{0||} + \\ + \left\{\frac{1}{2}\Gamma\left(\vec{F}_s + \frac{i}{\omega_s}\left(v_{||}[\vec{k}_s\vec{F}_s] + i\left[\vec{v}_E[\vec{k}_s\vec{F}_s]\right]\right)\right) - \frac{1}{2c^2}\left(v_\perp^2\vec{F}_s + (\vec{u}_0F_s)\vec{v}_E\right)\right\}\vec{e}_- e^{i\psi_{\text{res}}} + c.c., \tag{8}$$

$$\frac{d\psi_{\text{res}}}{dt} = \omega_0 + v - \frac{1}{2}\vec{e}_1 \cdot rot\vec{u}_0 - \frac{i}{2}\vec{e}_- \cdot \vec{e}_+' - \Omega_0\frac{v_E^2}{2c^2} + \vec{k}_s \cdot \vec{v}_E + \\ + \left\{\frac{i}{v_\perp}\left(\frac{1}{2}\Gamma\left(\vec{F}_s + \frac{i}{\omega_s}\left(v_{||}[\vec{k}_s,\vec{F}_s] + i\left[\vec{v}_E[\vec{k}_s,\vec{F}_s]\right]\right)\right) - \frac{v_\perp}{2c^2}\left(\frac{[\vec{e}_1,\vec{k}_s]}{\omega_s} + i\vec{\omega}_0\right)\vec{F} \cdot \vec{v}_E\right\}\vec{e}_- e^{i\psi_{\text{res}}} + c.c., \tag{9}$$

where

$$\omega_0 = \Gamma\Omega_0, \quad \Gamma = 1 - \frac{v_{||}^2 + v_\perp^2 + v_E^2}{2c^2}, \quad F_{s,0||} = \vec{F}_{s,0} \cdot \vec{e}_1, \quad k_{s||} = \vec{k}_s\vec{e}_1,$$

$$v_s = -\omega_s + k_{s||}v_{||}, \quad \vec{e}_\pm = \vec{e}_2 \pm i\vec{e}_3, \quad (...)' = \left(\frac{\partial}{\partial t} + v_{||}\vec{e}_1 \cdot \nabla + \vec{v}_E \cdot \nabla\right)(...)$$

note, that Eqs. (6)–(9) take place only in the case of quasilongitudinal propagation of the wave with respect to $\vec{B}_0$.

4. Cyclotron resonance for a longitudinal wave

For simplicity let us consider Eqs. (6)–(9) in the case of the cyclotron resonance $\Psi_{\text{res}} = \Theta_0 + \Theta_s$ for a longitudinal wave $\vec{\mathcal{E}}_s||\vec{k}_s||\vec{B}_0$, $\psi_\Gamma = \theta_0 - \theta_s$, $\vec{B}_0 = const$, $\vec{E}_0 = const$, $\vec{k}_0 = const$, $\vec{\mathcal{E}}_s = \vec{\mathcal{E}}e^{i\alpha}$, $\vec{\mathcal{E}} = (0, 0, \mathcal{E})$.

Then

$$\frac{d\vec{R}}{dt} = \vec{u}_0, \quad \frac{dv_{||}}{dt} = \frac{e\mathcal{E}v_E v_\perp}{2me^2}\sin(\psi_r + \alpha), \quad \frac{dv_\perp}{dt} = \frac{e\mathcal{E}v_E v_{||}}{2me^2}\sin(\psi_r + \alpha),$$
$$\frac{d\psi_r}{dt} = \Omega_0\left(\Gamma - \frac{v_E^2}{2c^2}\right) + \omega - kv_{||} + \frac{v_{||}v_E}{2c^2 v_\perp} \cdot \frac{e\mathcal{E}}{m}\cos(\psi_\Gamma + \alpha). \tag{10}$$

System (10) shows that the cyclotron resonance is possible, when the relativistic effects and the $\vec{v}_E$ drift velocity are taken into account, such a resonance is impossible in the case of a weak electric field $v \gg v_E$. To explain physical mechanism of this resonance it is necessary to use a new system, which is moving with a drift velocity. One can get the energy integral from Eqs. (10):

$$\varkappa^{-2} = (\dot{\xi}^2\tau^2 + \sin^2\xi)^{-\frac{1}{2}}, \quad \xi = \frac{1}{2}(\Psi_r + \alpha), \tag{11}$$

where

$$\varkappa^{-2} = \frac{2U}{H+U}, \quad \frac{1}{\tau^2} = 2U, \quad U = \frac{ke\mathcal{E}v_{0\perp}v_E}{2mc^2} \cdot v_{0\perp}$$

is an initial velocity of the transversal velocity $v_\perp$, H is the Hamilton function of the system (10).

If $|\varkappa| > 1$, the particle is trapped by the wave and if $|\varkappa| < 1$, the particle is untrapped by the wave. The sign of $\varkappa$ is chosen to coincide with the sign of $\dot{\xi}$.

The equation (11) has the same form as the equation for resonant particles in the case of electrostatic wave [2].

Then by standard methods [1–6] one can calculate the coefficient of cyclotron damping of the longitudinal wave under consideration:

$$\gamma(t) = \gamma_L \sum_n \frac{64}{\pi}\int_0^1 d\mathcal{X}\left\{\frac{-2\pi n \cdot \sin\left(\frac{\pi nt}{\mathcal{X}k\tau}\right)}{\mathcal{X}^5k^2(1-q^{2n})(1+q^{-2n})} + \frac{(2n+1)\pi^2\sin\left(\frac{(2n+1)\pi t}{\mathcal{X}k\tau}\right)}{k^2(1+q^{2n+1})(1+q^{-2n-1})}\right\}, \tag{12}$$

where

$$q = \exp\left(\frac{\pi K}{K}\right), \quad K = K(\varkappa)(1-\varkappa^2)^{\frac{1}{2}}, \quad K(\varkappa) = F\left(\frac{\pi}{2}, \varkappa\right)$$

is the complete of elliptic integral of the first kind, γ_L is Landau damping coefficient.

5. Conclusion

Numerical solving the equation system with initial conditions and parameters match the ones in works [7–11], Langmuir wave was selected with frequency in the range from $\omega = 1.38 \times 10^9 s^{-1}$ to $2.39 \times 10^{11} s^{-1}$ and girophase $\Omega_0 = 4.60 \times 10^9 s^{-1}$ to $4.60 \times 10^{11} s^{-1}$ with a wavelength of $\lambda = 2.28$ cm to 2.5 cm. The electric drift velocity between the values $v_E = 3.0 \times 10^5 \frac{\text{cm}}{\text{s}}$ and $5.0 \times 10^8 \frac{\text{cm}}{\text{s}}$.

The results of the numerical solutions of the averaged equations for cyclotron resonance confirm the possibility of cyclotron resonance in longitudinal wave in the case of a relativistic particle. The

development of energy the graph in figure 1, shows the resonant leg with growth of energy ($|\varkappa| > 1$), which is unstable so the particle goes shortly out of resonance ($|\varkappa| < 1$) with partly retention of the gained energy.

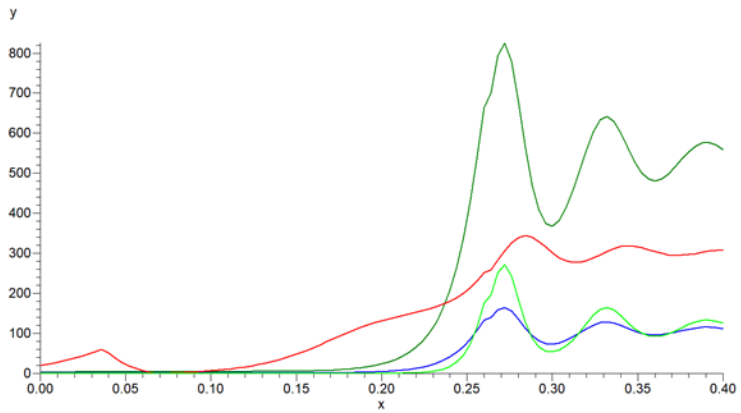


Figure 1. The graphs show the behavior of the particle's energy (bright green line), as well as the longitudinal and transverse velocities of the particle along the x axes (green and blue lines, respectively). The phase ψ_r is shown in graph by red line. The X-axis for velocity, energy, and resonant phase has dimension $t\omega_0$. The Y-axis for velocity has dimension v/c and for energy has dimension v^2/c^2

Funding: This research was funded by the RUDN University Scientific Projects Grant System, project No. 021934-0-000 (K.P. Lovetskiy, L.A. Sevastianov). This research was supported by the RUDN University Strategic Academic Leadership Program (S.P. Karnilovich, S.B. Strashnova, Yahya N. Shaar).

References

1. Karnilovich, S. P. & Milantiev, V. P. Effect of strong electric-field on resonance interaction between particles and HF waves. *Journal of Experimental and Theoretical Physics* **94,** 537–545 (1989).
2. Karnilovich, S. P. & Milantiev, V. P. *On cyclotron damping of longitudinal wave* in *XIX International Conference on Phenomena in Ionized Gases* **4** (1989), 810–811.
3. O'Neil, T. Collisionless Damping of Nonlinear Plasma Oscillations. *The Physics of Fluids* **8,** 2255–2262. doi:10.1063/1.1761193 (Dec. 1965).
4. Karnilovich, S. P. & Milantiev, V. P. *The effect of a strong quasi-stationary electric field on the resonant wave-particle interaction in a magnetized plasma* in *Mat. of the 6th All-Union Conference on the interaction of electromagnetic radiation with plasma, Dushanbe* (1991).
5. Karnilovich, S. P. & Milantiev, V. P. *Collection of scientific papers "Problems of quantum. and the stat. Physics"* in. Chap. Dynamics of resonant particles in RF fields in the MHD approximation (UDN, Moscow, 1989).
6. Karnilovich, S. P., Milantiev, V. P. & Bespalov, S. V. *Effects of the electric drift on the resonant wave-particle interaction* in *Proc. XXIV Int. Conf. On Phenomena in Ionized Gases — ICPIG-99, Warsaw* **2** (1999), 81–82.
7. Karnilovich, S. P., Milantiev, V. P. & Konovalseva, L. V. *The influence of the electric drift on resonant wave-particle interaction* in *Proc. 1992 ICPP Innsdruck. Austria* (1992).
8. Karpman, V. I., Istomin, J. N. & Shklyar, D. Nonlinear theory of a quasi-monochromatic whistler mode packet in inhomogeneous plasma. *Plasma Physics* **16,** 685–692. doi:10.1088/0032-1028/16/8/001 (1974).
9. Pocobelli, G. Damping of an electron plasma wave with detrapping of the electrons. *The Physics of Fluids* **24,** 2177–2182. doi:10.1063/1.863334 (1981).

10. Nishida, Y. & Sato, N. Observation of high-energy electrons accelerated by electrostatic waves propagating obliquely to a magnetic field. *Physical Review Letters* **59,** 653–656. doi:10.1103/PhysRevLett.59.653 (1987).
11. Sugawa, M. & Sugaya, R. Nonlinear Interaction between Electrostatic Electron Cyclotron Harmonic Waves and Electrons. *Journal of the Physical Society of Japan* **54,** 1339–1347. doi:10.1143/JPSJ.54.1339 (1985).

To cite: Karnilovich S. P., Lovetskiy K. P., Sevastianov L. A., Strashnova S. B., Shaar Y. N., On cyclotron damping of longitudinal wave, Discrete and Continuous Models and Applied Computational Science 32 (1)(2024)122–127.DOI: 10.22363/2658-4670-2024-32-1-122-127.

Information about the authors

Karnilovich, Sergey P.—Candidate of Sciences in Physics and Mathematics, Assistant professor of Institute of Physical Research and Technology of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: karnilovich-sp@rudn.ru, ORCID: https://orcid.org/0000-0002-7669-7030)

Lovetskiy, Konstantin P.—Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Computational Mathematics and Artificial Intelligence of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: lovetskiy-kp@rudn.ru, phone: +7(495)952-25-72, ORCID: https://orcid.org/0000-0002-3645-1060)

Sevastianov, Leonid A.—Professor, Doctor of Sciences in Physics and Mathematics, Professor at the Department of Computational Mathematics and Artificial Intelligence of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), Leading Researcher of Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research (e-mail: sevastianov-la@rudn.ru, phone: +7(495)952-25-72, ORCID: https://orcid.org/0000-0002-1856-4643)

Strashnova, Svetlana B.—Candidate of Chemical Sciences, Associate Professor at the Department of General and Inorganic Chemistry of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: strashnova-sb@rudn.ru, ORCID: https://orcid.org/0000-0002-2588-504X)

Yahya N. Shaar—PhD in Physics and Mathematics, Assistant professor of Institute of Physical Research and Technology of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: al-shaar-ya@rudn.ru, ORCID: https://orcid.org/0000-0003-4951-5657)

УДК 519.21
PACS 52.25.Fi
DOI: 10.22363/2658-4670-2024-32-1-122-127 EDN: BBLNGK

О циклотронном затухании продольной волны

С. П. Карнилович[1], К. П. Ловецкий[1], Л. А. Севастьянов[1,2], С. Б. Страшнова[1], Я. Н. Шаар[1]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*
[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. Выведены усредненные уравнения движения для релятивистских заряженных частиц в ВЧ поле (высокочастотных) волновых пакетов в диапазоне циклотронного резонанса в случае НЧ (низкочастотного) сильного электрического поля, где сильное электрическое поле означает, что характерная скорость частицы сравнима со скоростью электрического дрейфа ($v \sim v_E$). Показано, что при учете скорости электрического дрейфа становятся возможными новые механизмы затухания продольных волн. Проведен анализ влияния сильного электростатического поля на резонансное взаимодействие релятивистских частиц с высокочастотными волнами, а также влияние релятивизма на циклотронный резонанс для продольной волны. Получено аналитическое решение усредненной системы уравнений в квазирелятивистском приближении, а также проведен численный эксперимент для Ленгмюровской волны в случае циклотронного резонанса с учетом сильного электрического поля.

Ключевые слова: Скорость электрического дрейфа, затухание, релятивистские заряженные частицы, сильное электрическое поле, продольные волны, высокочастотные волновые пакеты, циклотронный резонанс