Volume 87, Number 4,
April 2026

ISSN 0005-1179
CODEN: AURCAT

AUTOMATION AND REMOTE CONTROL

Automation and Remote Control

Vol. 87, No. 4, April 2026

Editor-in-Chief
Andrey A. Galyaev

http://ait.mtas.ru

Automation and Remote Control

ISSN 0005-1179

SCOPE

Automation and Remote Control is one of the first journals on control theory. The scope of the journal is control theory problems and applications. The journal publishes reviews, original articles, and short communications (deterministic, stochastic, adaptive, and robust formulations) and its applications (computer control, components and instruments, process control, social and economy control, etc.).

Automation and Remote Control is abstracted and/or indexed in *ACM Digital Library, BFI List, CLOCKSS, CNKI, CNPIEC Current Contents/Engineering, Computing and Technology, DBLP, Dimensions, EBSCO Academic Search, EBSCO Advanced Placement Source, EBSCO Applied Science & Technology Source, EBSCO Computer Science Index, EBSCO Computers & Applied Sciences Complete, EBSCO Discovery Service, EBSCO Engineering Source, EBSCO STM Source, EI Compendex, Google Scholar, INSPEC, Japanese Science and Technology Agency (JST), Journal Citation Reports/Science Edition, Mathematical Reviews, Naver, OCLC WorldCat Discovery Service, Portico, ProQuest Advanced Technologies & Aerospace Database, ProQuest-ExLibris Primo, ProQuest-ExLibris Summon, SCImago, SCOPUS, Science Citation Index, Science Citation Index Expanded (Sci-Search), TD Net Discovery Service, UGC-CARE List (India), WTI Frankfurt eG, zbMATH.*

Journal website: http://ait.mtas.ru

0005-1179/26. *Automation and Remote Control* (ISSN: 0005-1179 print version, ISSN: 1608-3032 electronic version) is published monthly by Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow 117997, Russia.
Volume 87 (12 issues) is published in 2026.
Publisher: Trapeznikov Institute of Control Sciences, Russian Academy of Sciences.
65 Profsoyuznaya street, Moscow 117997, Russia; e-mail: redacsia@ipu.rssi.ru; http://ait.mtas.ru, http://ait-arc.ru

Contents

Automation and Remote Control
Vol. 87, No. 4, 2026

Special Issue

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 293–295.

Introductory Remarks to the Special Issue Devoted to MLSD 2024

DOI: 10.7868/S1608303226040014

This special issue of *Automation and Remote Control* presents selected papers of the 17th International Conference on Management of Large-Scale System Development (MLSD 2024). The conference was held on September 24–26, 2024, at the Trapeznikov Institute of Control Sciences, the Russian Academy of Sciences (ICS RAS), Moscow, Russia. The issue's papers are remarkable for their orientation toward solving topical control/management problems of large-scale systems. They demonstrate a progression from fundamental theoretical constructs (equations of mathematical physics and stability theory) through the development of new algorithms and methods (optimization, robust control, and machine learning) to the solution of particular applied problems in manufacturing, the energy sector, and the social sphere. In all the papers, the object of study is a system that can be characterized as large-scale:

- technical and technical-natural systems: power systems (analysis of oscillatory stability, I.B. Yadykin), oil and gas field development processes (two-phase filtration models, M.M. Volnykh and A.G. Kushner; well placement, A.I. Ermolaev et al.), chemical and oil refining enterprises (emergency response, A.F. Rezchikov et al.), and mechanical engineering enterprises (scheduling theory, E.N. Khobotov);
- organizational and socio-economic systems: research teams (labor productivity, A.S. Bogomolov et al.), the energy sector (forecasting indicators, V.A. Ivanyuk), and project programs (consideration of "soft dependencies," I.V. Burkova and A.V. Shchepkin).

The paper "A Mathematical Model for Increasing the Productivity of a Scientific Team" (A.S. Bogomolov, O.I. Dranko, V.A. Kushnikov, et al.) is at the focus of Russia's current scientific and technical policy. Here, the main novelty lies in formalizing the process of collective scientific research as an optimization problem. The authors treat the process of preparing publications not as an indivisible whole, but as a set of discrete tasks (writing the introduction, reviewing the previous results, performing computations, etc.) that can be distributed among team members with different efficiencies. Bogomolov et al. directly link their research work to "Labor Productivity," the National Project of the Russian Federation targeted at a significant growth of this indicator by 2030. The problem of increasing the productivity of intellectual/creative work is relevant for academic institutes, universities, corporate research centers, and any other organizations where the result depends on the efficiency of team-based research.

I.B. Yadykin's paper "Energy Stability Metrics of Linear Continuous Systems" is at the forefront of fundamental control theory, solving pressing problems of modern applications (the energy sector and complex networks). This research contains several major and interrelated original results: the concept of basic systems and invariant energy metrics, spectral decompositions for multiple roots, new frequency-domain formulas for solving Lyapunov equations, an efficient algorithm for computing the inverse Gramian, and a hybrid stability criterion. The author convincingly demonstrates the connection of his work with urgent applied problems, such as the analysis of oscillatory stability of electric power systems (weakly damped inter-area oscillations and the resonant interaction of modes), controllability analysis of complex networks (social, transport, and biological), and the creation of digital twins. Therefore, the results are in demand in high-tech industries.

M.M. Volnykh and A.G. Kushner present the paper "Cauchy Problem for the Buckley–Leverette Equation of Two-Phase Filtration with Variable Porosity." It is distinguished by strongly pronounced applied (petroleum engineering) and theoretical (mathematical physics) components. The paper contains several original scientific results: a new formulation of the classical problem (variable porosity), the application of geometric apparatus for its analysis, and the derivation of explicit formulas for characteristics and a multi-valued solution in an important special case. The authors significantly contribute both to the theory of partial differential equations and to the mathematical modeling of oil production processes. The obtained results (parabolic characteristics instead of straight lines) qualitatively differ from the classical case, demonstrating the importance of variable porosity.

In the paper "Robust Approach to Dynamic Compensation of Disturbances," S.A. Kochetkov and V.A. Utkin address a key problem in control theory—ensuring robustness to disturbances—and confirm this on a modern and complex object (a UAV with a load). This research is characterized by a high level of novelty and provides several original scientific results. Note the major ones: 1) a synthesis of dynamic compensation and sliding mode-based methods to ensure robustness in nonlinear systems, and 2) a new conceptual formulation of the problem where part of the plant's model acts as a controlled generator of uncertainties. It makes a considerable contribution to the development of the theory of robust and invariant control.

The paper "Hybrid Controllers in Control Problems of Real Objects" (A.F. Pashchenko and E.S. Duvanov) responds to the industry's demand for adaptive and accurate control systems for complex plants, confirming this with experiments. This work possesses scientific novelty at the engineering and methodological level. The main result lies in the original design and comparative analysis of two hybrid approaches (NNPC and FQR), their detailed implementation in MATLAB/Simulink, and successful validation on real laboratory plants with different dynamics.

A.F. Rezchikov, O.I. Dranko, et al. devote their paper, entitled "Models for Managing the Elimination of Consequences of Critical Situations at Oil Refining and Chemical Enterprises," to the problem of industrial safety and protection of the population and territories from man-induced critical situations. This theme matches priorities established by federal legislation (in particular, Federal Law no. 68-FZ of December 21, 1994 (as amended on August 8, 2024)) and GOST (state) standards. A strong point of the conducted research is the combination of a comprehensive system-dynamic approach to the elimination of consequences of critical situations with an original model correction algorithm. The authors propose an iterative algorithm for maintaining the model error at an acceptable level ($\leqslant 10\%$).

E.N. Khobotov's paper "Methods of Construction and Features of Work Schedules for Mechanical Engineering enterprises" solves a key and long-standing problem of scheduling theory, i.e., the transition from optimization at the shopfloor level to real planning at the level of a large mechanical engineering enterprise. Its undeniable scientific novelty lies in an original, multi-level, and adaptive sequential aggregation/decomposition methodology, which can be practically used to solve problems of previously unattainable dimensions.

The value of the paper "Using Machine Learning Methods for Analyzing and Forecasting of Small Samples of Macroeconomic Indicators in the Energy Sector of the Russian Federation" (V.A. Ivanyuk) is an original applied analysis. The author compares three fundamentally different classes of models (the classical linear regression, the exponential smoothing method, and the neural network approach) on carefully processed data with application of Bayesian ensembling (averaging), thorough consideration of constraints, industry-level detailing, and identification of critical sectors.

The paper "Consideration of Soft Dependencies under Stochastic Uncertainty in Multi-Project Program Implementation" (I.V. Burkova and A.V. Shchepkin) has high practical significance. In modern project management, especially in complex programs (construction, IT development, and R&D), managers constantly face a choice: is it worth implementing additional (non-mandatory) measures that can accelerate or reduce the cost of subsequent stages? Such dependencies are termed "soft" by the authors. The novelty of this research consists in the in-depth analysis of the influence of the probabilistic nature of soft dependencies.

A.I. Ermolaev, A.V. Akhmetzyanov, and A.R. Latipov present the paper "Controlled Random Search and Likelihood Ratio in Boolean Programming Problems," where they develop and test an approximate algorithm for solving linear Boolean programming problems of high dimension. This research is topical since, for many applied problems, exact methods (e.g., branch-and-bound) require prohibitive computational resources, and the value of an exact solution is reduced due to initial data errors. The authors propose an iterative controlled random search algorithm that tests, at each step, one of two statistical hypotheses using the sequential probability ratio test (SPRT). The efficiency of this algorithm is demonstrated on a practically important problem of oil and gas field development, i.e., the optimal placement of a given number of wells. Computational experiments were carried out for different grid sizes (up to 400 blocks) and different numbers of wells.

Chairman of the Program Committee of the MLSD Conference
S.N. Vassilyev, Academician of the Russian Academy of Sciences
ICS RAS
snv@ipu.ru

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 296–313.

SPECIAL ISSUE

A Mathematical Model for Increasing the Productivity of a Scientific Team

A. S. Bogomolov[*,**,a], **O. I. Dranko**[***,b], **V. A. Kushnikov**[*,c], **A. F. Rezchikov**[***,d], **and I. A. Stepanovskaya**[***,e]

[*] *Saratov Scientific Center, Russian Academy of Sciences, Saratov, Russia*
[**] *Saratov University named after N.G. Chernyshevsky, Saratov, Russia*
[***] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a]*bogomolov@ya.ru*, [b]*olegdranko@gmail.com*, [c]*kushnikoff@iptmuran.ru*, [d]*rw4cy@mail.ru*, [e]*irstepan@ipu.ru*

Received July 8, 2025
Revised October 16, 2025
Accepted November 20, 2025

Abstract—The study is based on the idea of introducing the necessary practical, productive component and optimization into the creative process of producing scientific publications. The productivity of a research team is considered as a relative assessment of its publication activity. The problem of distributing tasks among members of the research team in order to maximize labor productivity is formulated both substantively and formally. An appropriate distribution of tasks can be determined by solving an optimization problem. In this problem, the lower limits are formed based on the publication plan for the current state assignment, the upper limits are based on the maximum job load of team members, and the objective function is the ratio of the team's publication activity assessment to the total number of workload hours. The proposed approach is illustrated by calculations based on accumulated statistics. Calculations based on current criteria for assessing publication activity revealed that the most productive research, according to the assessment scale of the Ministry of Education and Science, will be primarily in the direction of publishing Level I articles from the "White List of Scientific Journals," while to ensure a high performance indicator for scientific activity according to the current scale of the V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences (ICS RAS) has found that preparing presentations for Scopus conferences and publications in Tier IV journals is the most preferable approach. However, recommendations are highly dependent on the allocation of employee time to various tasks. Calculations based on actual data show a significant increase in labor productivity in one of the ICS RAS laboratories as a result of applying the proposed approach.

Keywords: modeling, management, optimization, labor productivity, publication, scientific work

DOI: 10.7868/S1608303226040024

1. INTRODUCTION

Scientific research is a complex, multilayered process whose creative components are difficult to formalize and plan. Nevertheless, government and private grantors impose specific requirements on the results of this process to ensure they are obtained and applied in a timely manner. Therefore, researchers and their teams face the challenge of systematizing and planning the publication process to increase productivity. The issue of increasing labor productivity is currently relevant for many sectors of human activity. For example, the Russian Federation's National Project "Labor Productivity" [1] plans to increase this indicator by at least 40% by 2030 at enterprises in the basic and raw materials sectors of the economy and social institutions. Developing relevant models and methods for solving this problem is essential for maintaining and enhancing the country's competitiveness in the modern world. Labor productivity can be considered as an indicator of national development [2].

Applied and fundamental science should be considered among the priority areas for achieving this goal. However, in creative industries, which include scientific research, the task of increasing productivity has its own specific challenges. When evaluating a team's scientific work, it's generally assumed that the primary criterion is the theoretical and practical significance of the research results. However, understanding and recognition of this significance by the community may not come immediately. Therefore, in the information age, it is especially important to intensify the implementation of scientific results. To achieve this, the authors of this article believe that the process of scientific creativity must be accompanied by systematic and targeted efforts to promote it in leading specialized journals and scientometric systems. Successful promotion of scientific results not only facilitates their recognition but also provides the necessary incentive, including financial incentives, to continue the team's work. Systems for assessing scientific productivity proposed by the state, qualified clients, and expert councils are designed to address these issues. In Russia, the main such systems are the "List of peer-reviewed scientific publications in which the main scientific results of dissertations for the degree of candidate of science and for the degree of doctor of science must be published" [3] and the "White List of Scientific Journals" of the Russian Center for Scientific Information [4] (hereinafter BS RCSI). Internationally, these are Scopus [5], Web of Science [6], and other systems. The active representation of researchers and publications in these spaces is associated with well-known and accepted systems and methodologies for evaluating scientific work, such as the Hirsch index [7], the FWCI index of the ratio of the total number of citations to the average number of citations in a subject area [8], as well as other indicators, for example, based on bibliometric indicators [9] and the assessment of opinions on social networks [10]. However, these estimates and the algorithms for calculating them do not directly answer the questions of the collective organization of scientific work.

For these reasons, improving scientific productivity and publication activity in such systems requires systematic study and support with specialized mathematical models and algorithms. These models and algorithms will be aimed at managing the productivity of individual researchers, laboratories, institutes, and the country's scientific sectors as a whole. This article is devoted to the construction of such models and algorithms.

2. STATE OF THE ART

2.1. Methodological Foundations

The theory of inventive problem solving [11] can provide a number of useful recommendations for the rational preparation of collective publications:

- analysis and resolution of contradictions (immersion in the topic versus limited time, a uniform article style versus the personal writing style of different authors, thorough review versus prompt execution);
- application of the basic principles and tools of the theory of inventive problem solving (resources, automation, etc.);
- assignment of roles within the team (lead author, researcher, theorist, practitioner, editor).

However, this theory does not recommend using mathematical methods and calculations to distribute the workload in the form of tasks that best suit specific performers.

Similar recommendations can be derived from the concept of scientific work organization [12]. When creating publications, there should be a systematic approach to distributing responsibilities among process participants, with each specialist performing strictly defined functions:

- the project manager coordinates the work and performs final editing;
- the researcher collects and analyzes data;
- the methodologist develops the research approach;
- the writer forms the text;
- the editor is responsible for stylistic editing and proofreading.

It is important to consider the participants' qualifications: the supervisor provides expert review, co-authors develop the methodology, graduate students perform the initial data processing, and junior research fellows perform support tasks.

We believe that mathematical models and analytical calculations to support and justify decision-making when distributing publication assignments will be a development and refinement of these approaches. This will take into account the principles of Lean Manufacturing methodology [13], the Japanese methodology of continuous improvement [14], and the 5S system (5S—sorting, systematization, cleaning, standardization, improvement (self-discipline)) [15], which largely overlap with these concepts.

2.2. Mathematical Models and Methods

Finding an appropriate distribution of tasks when preparing publications can be based on classical mathematical programming methods [16] and their modifications [17], used in organizing and planning production, as well as for improving the predicted performance of organizational systems [18–20]. Both methods of individual incentives [21] and problems related to distributing work among performers are considered, where variations of classical methods for solving the assignment problem [22] should be mentioned. These methods are aimed at distributing work for its fastest completion, which partially corresponds to the goals of the problem under consideration.

2.3. Modern Research and Practices on Labor Distribution

The problem under consideration can be interpreted as a time management problem; however, modern time management techniques and methods primarily provide strategic, qualitative recommendations.

Improving human productivity is currently being explored in various industries, including using artificial intelligence: in logistics [23], in tourism supply chains [24], and elsewhere. The issue of balancing humans and robots [25] is also being widely explored in the field of intelligent manufacturing. [26] proposes using a system dynamics model to analyze the interrelations of factors influencing labor productivity in a team. Although this article does not provide recommendations for the appropriate distribution of work, we note the potential of system dynamics for analyzing and modeling the feasibility of such appropriate distribution plans.

Modern publications demonstrate that talent management can play a significant role in improving labor productivity [27]. However, it is used primarily in team formation rather than in assigning tasks to employees.

The proposed recommendations include increasing researcher mobility [28] and changing residence and organization after completing their studies [29]. The conclusions in such studies are based on a fairly extensive statistical analysis, but they do not provide specific recommendations for the work of research teams.

In other works, some strategic recommendations are made primarily for individual researchers. Regularity, consistency, and moderation are recommended: working on texts no more than 40–45 hours per week, avoiding procrastination and perfectionism, and taking "small steps" when writing articles [30].

The Institute of Control Sciences of the Russian Academy of Sciences is developing an information system for the analysis of scientific activity (ISAND) [31] in the field of control theory [32, 33]. As of January 15, 2025, 186 000 publications by 185 000 authors have been processed. The ISAND ontology is described, focused on representing and collecting knowledge in the field of control theory: both scientific knowledge (the control theory ontology) and knowledge related to the scientific activities of agents in this field (organizations, journals, conferences, and individual researchers).

The authors of various publications separately note the benefits of collaborative work when creating publications. It is believed that a scientist can increase their productivity by working on joint projects with other researchers. This suggests that the overall assessment of a research team's publication activity will be higher when distributing work on individual articles among researchers than when they only produce their own individual publications. This assumption is consistent with experience: high-ranking publications typically have a fairly large number of co-authors. However, collaborative work on publications, when the number of required publications is sufficiently large, in some cases requires some regulation in terms of the appropriate distribution of tasks and the transfer of information between the authors of individual papers. The issues of organizing research teams, whose members possess individual characteristics, varying abilities, and technical capabilities for performing a particular task, have apparently not been explored in open publications. This study is devoted to finding analytical models and methods for addressing such issues. The problem under consideration has the following formal and substantive formulation.

3. STATEMENT OF THE PROBLEM

Let there be a team of n employees who, over a period of time T, produce m types of publications. Each publication of the type $(i = 1, \ldots, m)$ is valued at $(p_i \geqslant 0)$ points. Further, unless otherwise specified, the index i will be used for publications and it will be assumed that $i = 1, \ldots, m$.

A set of publication part types $S = \{s_1, \ldots, s_k\}$ is given (for example: review, model, calculation, etc.) so that a publication of each type i can be represented as a certain set of parts $S_i \subseteq S$.

We call a part of a publication of a certain type a task, for example: introduction, conclusion, mathematical model, etc. Consider the universe U—the set of all tasks, i.e., parts of all types $s_1, \ldots, s_k$ of all publication types. We assume that different employees can complete tasks from U in any order. A publication is considered complete when all tasks in it are completed.

For each employee $j = 1, \ldots, n$, the time $t_{j,i}$ for successful completion of any task $u_i \in U$ is known. This value is determined empirically, taking into account the employee's position, experience, inclinations, and technical capabilities. It may reflect both the employee's creative qualities (e.g., "creativity" as the ability to quickly come up with a productive research idea), organizational skills (successfully coordinating publication parts from different co-authors), and executive abilities, as well as organization (e.g., quickly completing and submitting a new publication to a journal). Further, unless otherwise specified, the index j will be used to denote employees, and it will be assumed that $j = 1, \ldots, n$.

The task is to assign a set of tasks $U_j(T) \subseteq U$ to each employee j in such a way that the overall assessment of the completed publications of the members of the scientific team for the period T in relation to the regular working time for this period is maximum:

$$\frac{\sum_{i=1}^{m} k_i p_i}{\sum_{j=1}^{n} T_j} \to \max,$$

where $k_i \geqslant 0$ is the number of publications of type i produced by the team over period T, T_j is the employee's workload over period T.

The following constraints must be met:

- $k_i \geqslant a_i$, where $a_i \geqslant 0$ are specified numbers reflecting the publication plan from the application for the topic of the state contract or grant;
- the total workload of each employee j over period T does not exceed their workload T_j;
- each task is assigned to no more than one employee. This is ensured by a sufficiently detailed breakdown of publications into tasks.
- We assume that an integer number of publications are prepared over period T. We do not consider cases where work on a publication takes several years. In real-world situations, a publication may be prepared but not accepted. In subsequent periods, it may be reworked and submitted to other publications, and the results of the efforts and resources expended in period T are not effectively assessed.

The substantive task is to distribute the resources of research staff in the most productive way.

4. MODELS FOR SOLVING THE PROBLEM

4.1. Publication Preparation Model

The proposed general framework for publication preparation is shown in Fig. 1. The process is iterative, as reflected by the feedback loops in the diagram. In practice, these loops can be more complex during publication creation.

In this paper, when calculating the decomposition by performer, the following main publication components (tasks) were identified. Each of these components, as evidenced by experience with article writing by a team of authors, can be assigned to a researcher depending on their available time, qualifications, and aptitudes.

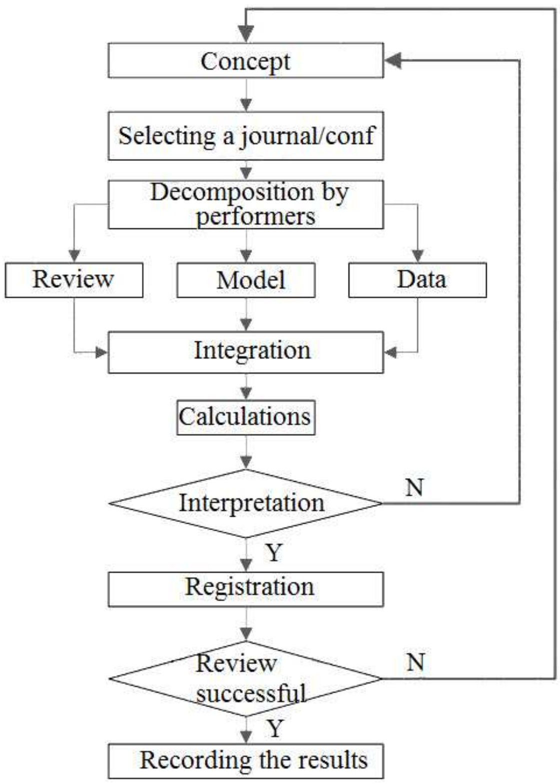

Fig. 1. High-level graph of the publication creation stages.

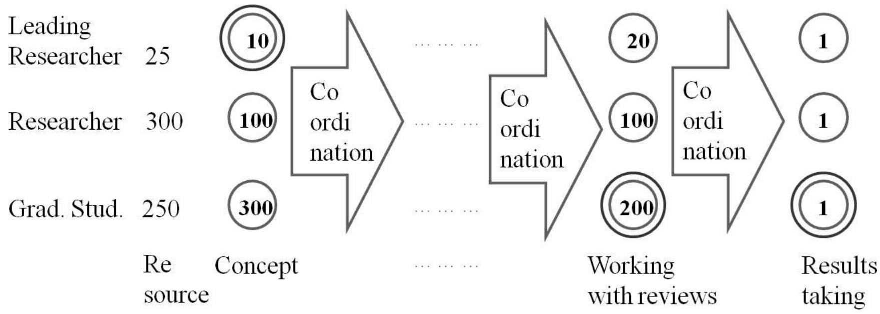

Fig. 2. A fragment of the task assignment process for a single publication.

1. Text
   1.1. Title
   1.2. Abstract
   1.3. Introduction
   1.4. Overview
   1.5. Model
   1.6. Algorithm
   1.7. Data
   1.8. Calculations
   1.9. Discussion of Calculations
   1.10. Conclusion
   1.11. References
   1.12. Section Coordination
   1.13. Design
2. Expertise. Report
3. Submission to a Journal/Conference
4. Working with Reviews
5. Presentation, Speech
6. Recording Results (Publication awards, Best Paper Competitions)

Figure 2 shows a typical fragment of the task distribution process for a single publication (a Level III article of the White List). The current resource is defined as the employee's remaining time in hours at the time of planning; the circles indicate the number of hours required to complete the task; the vertices selected within the optimal or quasi-optimal solution are double-circled.

Next, we describe the proposed mathematical model for the efficient distribution of publication preparation tasks.

4.2. Mathematical Model

In general, the solution involves finding the task distribution as a 0-1 matrix X from n rows and $m' = \sum_{i=1}^{m} p_i e_i$ columns, where $p_i \geqslant a_i$ is a realistic upper estimate of the number of publications of type i over the period T (for example, a small team will likely publish no more than one or two Level I White List articles per year); e_i is the number of tasks in the publication breakdown of type i.

Each row of the X matrix determines whether an employee has completed or not completed each task in each of the team publications under consideration. Matrix element $[X]_{j,i}$ equals 1, if an employee is assigned a task corresponding to the column i of matrix X, and equals 0 otherwise.

We consider publication A to be completed (scheduled for completion) if the sums $\sum_{j=1}^{n}[X]_{j,i}$ for all columns of the matrix X corresponding to this publication are equal to 1 (provided that each task is assigned to no more than one performer).

Let's denote by $I(A, X)$ the indicator of whether this condition is met: $I(A, X) = 1$ if each task into which publication A is divided has an assignee, and $I(A, X) = 0$ otherwise. Some properties of the function $I(A, X)$ are given by the following statement.

Statement 1. *The function I(A, X) is nonlinear with respect to the variables X, but can be represented using the multiplication operation.*

Proof of Statement 1. The nonlinearity of $I(A, X)$ follows from the fact that the matrix X_0, in which publication A is complete, i.e., $I(A, X_0) = 1$, can easily be represented as the sum of matrices $X_0 = X_1 + X_2$, in each of which publication A is not complete, and therefore, $I(A, X_1) = I(A, X_2) = 0$.

Linearity of the function $I(A, X)$ in the variables X would require the equality $I(A, X_0) = I(A, X_1) + I(A, X_2)$, which is not satisfied in this case.

Representability using the multiplication operation follows from the fact that $I(A, X)$ can be represented as

$$I(A, X) = \prod_{i\in[X]_A} \sum_{j=1}^{n}[X]_{j,i},$$

where $[X]_A$ is the set of column numbers of X associated with A, as required.

Further, taking into account the introduced notations, the number of publications of type i performed in period T will be $k_i = \sum_{j=1}^{p_i} I(A_{i,j}, X)$, where $A_{i,j}$ is the publication of type i under number j. Then the assessment of the publication activity of the team is equal to

$$Pub = \sum_{i=1}^{m} p_i \sum_{j=1}^{p_i} I(A_{i,j}, X).$$

Given the dimensionality of the variable matrix X introduced above, we will now consider a Boolean programming problem with nm' variables. Most of the variables will be equal to 0, reflecting the non-participation of employees in some publications. Furthermore, the problem dimensionality in practice is significantly reduced by making assumptions about the numbers p_i, as well as by assigning similar tasks to the same employees.

Let the time standards required by employees to complete tasks U be written in a matrix *Norm* of dimension $n \times \sum_{i=1}^{m} e_i$, where n is the number of employees, m is the number of publication types, and e_i is the number of tasks into which a publication of type i is broken down. An element of the matrix *Norm* is equal to the number of hours required by an employee in a row to complete a task in the corresponding column.

Considering that publications of each type i can be considered in numbers from 1 to p_i, we construct a corresponding extended normalization matrix $Norm'$ of dimension $n \times m'$, where $m' = \sum_{i=1}^{m} p_i e_i$. The columns of $Norm'$ correspond to the tasks into which publications $1, \ldots, p_1$ of type 1, $1, \ldots, p_2$ of type $2, \ldots, 1, \ldots, p_m$ of type m are divided. Then, the time spent by employee j on completing all of their tasks is equal to the scalar product of row j of the matrix X and row j of the matrix *Norm*. The condition for limiting an employee's time based on workload then takes the form

$$\sum_{i=1}^{m'} [M]_{j,i}[X]_{j,i} \leqslant T_j.$$

Thus, the mathematical formulation of the problem, taking into account the notation introduced above, is as follows: Find a Boolean matrix X such that

$$\frac{\sum_{i=1}^{m} p_i \sum_{j=1}^{p_i} I(A_{i,j}, X)}{\sum_{j=1}^{n} T_j} \to \max,$$

under restrictions

$$\sum_{j=1}^{p_i} I(A_{i,j}, X) \geqslant a_i, \quad i = 1, \ldots, m,$$
$$\sum_{i=1}^{m'} [M]_{j,i} [X]_{j,i} \leqslant T_j, \quad j = 1, \ldots, n,$$
$$\sum_{j=1}^{n} [X]_{j,i} \leqslant 1, \quad i = 1, \ldots, m'.$$

Note that in this problem, the first constraint and the objective function are nonlinear with respect to the problem variables, since they contain the publication indicator $I(A,\, X)$, the nonlinearity of which is shown above. Linearity can be achieved if the indicators of all publications under consideration are set to specific values, 0 or 1. However, this would also imply a certain selection of articles of different types, for example, one Level I article of the White List and two articles each of Levels II and III.

This means that the planned assessment of the team's publication activity will be known, and the team's productivity will have a specific value.

5. CRITERIA AND DATA

5.1. Criteria for Assessing Scientific Activity

Criteria for assessing scientific activity are typically based on the evaluation of a team's publications, and other results (rewards, grants, etc.) are usually derived from this evaluation. In the Russian Federation, the main criteria for assessing publications are currently provided by the guidelines of the Ministry of Education and Science [3, 4].

In addition to public motivation, it is advisable to consider possible personal criteria that influence the bonus component of remuneration. For example, the scientific performance indicator (hereinafter, SPI) of an employee of the Institute of Control Sciences of the Russian Academy of Sciences [34] is defined as

$$\mathrm{SPI}_j = \sum_{i=1}^{k(j)} \frac{p_i}{n_i},$$

where $k(j)$ is the number of publications of the employee, n_i is the number of co-authors of publication i, p_i is the publication weight in points from Table 1.

Qualification requirements for Institute employees can be based on the correspondence from such tables (e.g., at least a certain number of points per year for a senior research fellow), or in conjunction with them (e.g., at least 10 publications in 5 years for a chief research fellow, the required level of technical readiness of scientific results, etc.). Employees are interested in increasing the productivity of the entire team, but issues of subtle interaction and mutual reinforcement of employee and team motivation as a whole require further study [35, 36].

Table 1. Weights in the ICS RAS Publication Assessment System

No.	Publication Types	Score
1	Publication in a Level I publication of the White List	400
2	Publication in a Level II publication of the White List	300
3	Publication in a Level III publication of the White List	200
4	Publication in a Level IV publication of the White List	150
5	Articles in journals, collections of articles, or conference proceedings, book chapters	72
6	Articles in journals included in the HAC List	24
7	Articles in journals, collections of articles, book chapters from the list of journals	6
8	Abstracts of papers at a scientific conference	2
9	Plenary paper at a scientific conference / at a conference whose proceedings are indexed by WoS or Scopus	12/96
10	Paper at a conference (at least three pages long)	6
11	Brochure (does not have an ISBN or is less than 6 p.s.), per p.s.	2
12	Monograph (ISBN and volume over 6 p.s.) / WoS, Scopus Monograph, per p.s.	12/24
13	Russian Federation Patent	18
14	Certificate of Registration of Software, Database, or Integrated Circuit	5

5.2. Initial Data

Working time, a scarce resource within the "cost-benefit" logic, was considered the primary resource enabling employees to complete publication preparation tasks [37, 38]. The data source was management accounting tables of actual time expenditure associated with publication preparation by a Doctor of Engineering Sciences, a leading research fellow at one of the ICS RAS laboratories. Time tracking is performed in Outlook and uploaded to Excel for analysis. The database is then processed using big data (Business Intelligence, BI) methods. Processing time is up to 15 minutes per month. Analysis time, with an established planning and analysis process, is 0.5–1 hour per month. Summary results for uploading time to the ICS RAS are presented in Table 2.

Table 2. Workload by activity, hours

Year	Publications	Research	Org.	Road	Report	New	Cons.	Preparation	Total
2020	507	467	198	48	84	97	26	19	1444
2021	602	636	164	107	69	14	60	30	1681
2022	475	715	92	135	22	6	28	11	1483
2023	468	680	73	184	19	28	11	14	1475
Overall total	2051	2499	526	474	194	144	125	73	6084
Average	512	626	131	119	48	36	31	18	1521

The following breakdown by activity was used:

- "Publications"—time associated with a specific publication. "Research" item—time not tied to a specific publication: general discussions, collecting and processing organizational data, organizing conferences, reviews, etc. This data/calculations may or may not be published later, but this is not currently factored into time tracking.
  "Org." item—organizational issues.
  "New" item—developing new projects.
  "Cons." item—consultations. "Travel" item—includes one-way time.

6. ANALYSIS AND CALCULATIONS

6.1. Analysis of Publication Preparation Activities

38 of the author's publications from 2020–2023 were analyzed, including "ineffective" publications (rejections received, partial work completed, including work for future use). The total estimated preparation time for these ineffective publications was 499 hours over 4 years.

The distribution of publications by year is shown in Table 3.

The following is observed:

- increase in the number of publications;
- increase in SPI;
- stabilization in the number of hours;
- increase in SPI points per hour.

Note that the average number of SPI points per hour remains relatively low.

Table 3. Distribution of publications by year

Year	Number of publications	SPI of author	Stab. Number of h	Average Score/h
2020	9	183	440	0.42
2021	10	168	530	0.32
2022	15	278	490	0.57
2023	16	327	461	0.71
Total	50	956	1921	0.50

The distribution of publications by type is shown in Table 4.

Table 4. Distribution of publications by type for 2020–2023

General, journal, or conference	Number of publications	Author	Hours	Average, points/hour
Monograph	3	115	104	1.11
Conference	24	630	876	0.72
Journal	11	210	442	0.48
(empty)	12	0	499	0.00
Total	50	956	1921	0.50

Based on the analysis of the presented data, it was concluded that conferences (0.72 points/hour) are slightly more productive than journal articles (0.48 points/hour).

6.2. Calculations

Each task requires a specific amount of time, depending on the volume, complexity, and level of the publication. Table 5 shows a sample table for estimating time expenditure by publication type for a researcher with a Doctor of Engineering qualification, a presenting research fellow. "UBSi" in this and the following tables denotes White List level i.

It should be noted that at the time of preparation of this factual data, the concept of the White List had not yet been introduced into the publication activity assessment guidelines. Therefore, the White List is effectively accounted for as levels 2 and 4, along with the publication's submission to Scopus and the Higher Attestation Commission (HAC) of Russia.

The presented calculation, based on actual data, extrapolated the time standards for White List Level I articles. The current scenario assumes that the time and complexity of writing a Level I article increases significantly due to the additional complexity of formulating the problem and discussing the results, preparing the review, and meeting reviewer requirements. Other research teams may have different average time standards for preparing publications in various categories, to which the proposed method can be adapted. Further in the article we will designate Level I of the White List as LW1, Level II as LW2, Level III as LW3 and Level IV as LW4.

Table 5. Estimated time for publication preparation. Author 1

Section	LW1	LW2	LW3	LW4	Scopus IEEE	HAC
Volume, in thousands of characters	1170	565	314	191	135	126
Total resources, h	965	450	253	160	96	117
1. Text	20	15	10	10	8	10
1.1. Concept	10	5	3	2	1	1
1.2. Title	10	5	3	2	1	1
1.3. Abstract	10	8	6	5	3	3
1.4. Introduction	200	50	20	15	10	12
1.5. Overview	100	40	20	20	15	20
1.6. Model, Algorithm	100	60	30	20	15	20
1.7. Data	400	200	120	60	30	30
1.8. Calculations	80	40	20	10	5	5
1.9. Discussion of calculations	5	4	3	2	1	1
1.10. Conclusion	10	8	6	4	2	4
1.11. Literature						
1.12. Section Coordination	20	15	12	10	5	10
1.13. Design	40	30	20	10	3	3
2. Submission to Journal/Conference, Expertise Report	160	80	40	20	5	5
3. Working with Reviews	0	0	0	0	30	0
4. Presentation, Speech (in person)	5	5	1	1	1	1
5. Recording results (SPI, best paper competitions)	5	5	1	1	1	1

The optimization problem based on the data in Table 5 was solved using nonlinear programming. Analysis of the calculation results (Table 6) shows that optimization should select publication types with the highest ratio of score to time spent.

Table 6. Hourly efficiency for one author (leading researcher)

Section	LW1	LW2	LW3	LW4	Scopus	HAC
Hourly efficiency, Ministry of Education and Science criterion	0.017	0.018	0.016	0.013	0.007	0.001
Hourly efficiency, SPI criterion	0.342	0.531	0.637	0.785	0.533	0.190

For a small team with low transaction costs, each operation is assigned to the performer with the highest productivity (efficiency). If the performer is overloaded, the task is assigned to the next most efficient employee. At its core, this approach is based on the cost-effectiveness methodology.

Let's consider an example of activity optimization in two cases: for Author 1, a leading researcher, and for Author 2, a researcher.

Example 1 (single author). Below are the results of the optimization calculations for Author 1 (leading researcher) using the Ministry of Education and Science criterion (Table 7) and the SPI criterion (Table 8).

Table 7. Optimization results for one Author 1 (leading researcher), Ministry of Education and Science criteria

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
Criterion scenario Ministry of Education and Science	15.6	4.8	10.0	0.8	0.0	0.0	0.0
Additionally: SPI	428.1	96.2	300.0	31.9	0.0	0.0	0.0
Number of publications	1.4	0.2	1.0	0.2	0.0	0.0	0.0
Total number of hours	896	281.24	565.00	50.12	0.00	0.00	0.00

Table 8. Optimization results for one Author 1 (leading researcher), criterion SPI

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
SPI scenario: Ministry of Education and Science criterion	13.2	0.0	0.0	8.2	5.0	0.0	0.0
Criterion SPI	627.6	0.0	0.0	327.6	300.0	0.0	0.0
Number of publications	3.64	0.00	0.00	1.64	2.00	0.00	0.00
Total number of hours	896	0.00	0.00	514.36	382.00	0.00	0.00

The difference in achieving the maximum values of the Ministry of Education and Science criterion is approximately 18%, while for the SPI criterion it is 47%. Similar optimization calculations for Author 2 (researcher) for the Ministry of Education and Science criteria (Table 9) and SPI (Table 10) show a slightly smaller difference in achieving the maximum values of the criteria—approximately 20%.

Table 9. Optimization results for one Author 2 (research fellow), Criterion Ministry of Education and Science

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
Criterion Scenario Ministry of Education and Science	7.8	0.0	1.6	6.2	0.0	0.0	0.0
Additionally: SPI	296.3	0.0	48.9	247.4	0.0	0.0	0.0
Number of publications	1.4	0.0	0.2	1.2	0.0	0.0	0.0
Total number of hours	896	0.00	177.57	718.79	0.00	0.00	0.00

Table 10. Optimization results for one Author 2 (research fellow), Criterion SPI

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
Scenario SPI: Criterion Ministry of Education and Science	6.4	0.0	0.0	1.4	5.0	0.0	0.0
Criterion SPI	356.6	0.0	0.0	56.6	300.0	0.0	0.0
Number of publications	2.28	0.00	0.00	0.28	2.00	0.00	0.00
Total number of hours	896	0.00	0.00	164.36	732.00	0.00	0.00

An analysis of the obtained data allows us to conclude that the main priority for achieving the maximum results according to the Ministry of Education and Science methodology is publications at Level I of the White List, and when achieving the maximum number of publications at Level IV, as well as conferences with publications in journals included in Scopus, the main priority is achieved.

According to the ICS RAS publication assessment methodology, publications at Levels III and IV of the White List are identified as the main priority (in the previous version of the SPI, the main priority was articles from journals included in Scopus).

Example 2 (two authors).

In this example, we use the following assumptions:

- publications are jointly prepared by two authors;
- the authors are related by topic and have sufficiently effective communication, so coordination is practically unnecessary;
- assignment distribution (matrix X) is established uniformly for all types of publications; the results of joint work take into account the number of authors (fractality).

An analysis of the calculation results shows that, according to both the Ministry of Education and Science criterion (Table 11) and the SPI criterion (Table 12), it is advisable to focus efforts on preparing articles at levels II and III of the White List.

Table 11. Optimization results for two authors, Criterion Ministry of Education and Science

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
Criterion Scenario Ministry of Education and Science	17.5	0.0	20.0	15.0	0.0	0.0	0.0
Additionally: SPI	1 198	0	600	598	0	0	0
Number of publications	4.99	0.00	2.00	2.99	0.00	0.00	0.00
Total number of hours	1 793	0	980	812	0	0	0

Table 12. Optimization results for two authors, Criterion SPI

Chapter	Total	LW1	LW2	LW3	LW4	Scopus	HAC
Criterion Scenario SPI	15.4	0.0	0.9	20.0	10.0	0.0	0.0
Additionally: Criterion Ministry of Education and Science	1 426	0	26	800	600	0	0
Number of publications	8.09	0.00	0.09	4.00	4.00	0.00	0.00
Total number of hours	1 793	0	43	1 087	663	0	0

The summary results are shown in Table 13. For collaborative work, combining efforts is beneficial for the team according to the SPI criterion. According to the Ministry of Education and Science criterion, Author 1 loses significantly (in the personal version, one LW2 article and 0.24 LW1 articles; in the joint version, taking fractality into account, one LW2 article and 1.5 LW3 articles).

Table 13. Summary of optimization results for the team of authors

Criterion	Author 1 (leading research fellow)	Author 2 (research fellow)	Author 1 and 2, Ministry of Education and Science	Author 1 and 2, SPI
Criterion Ministry of Education and Science	15.5	7.8	8.75=17.5/2	7.7=15.4/2
Criterion SPI	428	296	599=1198/2	713=1426/2

The estimated distribution of responsibilities between the Authors is shown in Table 14.

Table 14. Distribution of time between authors

Chapter	Author 1	Author 2
1. Text		
1.1. Concept	1.00	0.00
1.2. Title	0.00	1.00
1.3. Summary	1.00	0.00
1.4. Introduction	1.00	0.00
1.5. Review	1.00	0.00
1.6. Model, algorithm	1.00	0.00
1.7. Data	0.00	1.00
1.8. Calculations	0.00	1.00
1.9. Discussion of calculations	0.88	0.12
1.10. Conclusion	1.00	0.00
1.11. Bibliography	0.00	1.00
1.12. Coordination of sections	1.00	0.00
1.13. Design	0.00	1.00
2. Submission to a journal/conference. Expertise report	0.54	0.46
3. Working with reviews	1.00	0.00
4. Presentation, speech	0.00	1.00
5. Recording of results (SPI, best works competitions)	0.00	1.00

7. TESTING AND DISCUSSION

7.1. Testing the results

To test the described approach, we present the results of the ICS RAS “Large-Scale Systems” laboratory. Figure 3 shows the connectivity graph for publications from the ISAND system. Half of the authors (five people) are connected to each other, while the other half (five people) are weakly

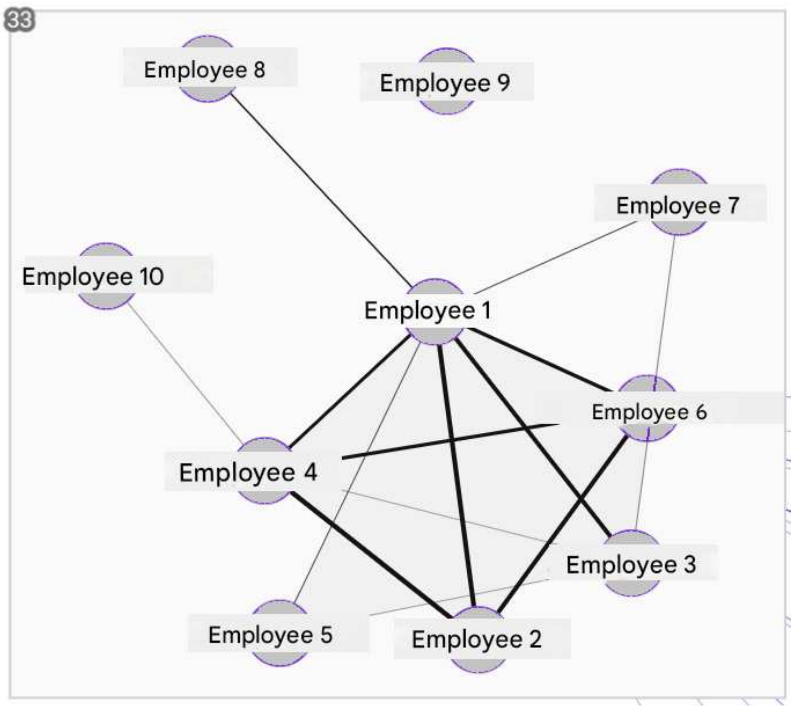

Fig. 3. Co-authorship connectivity network of the laboratory under study.

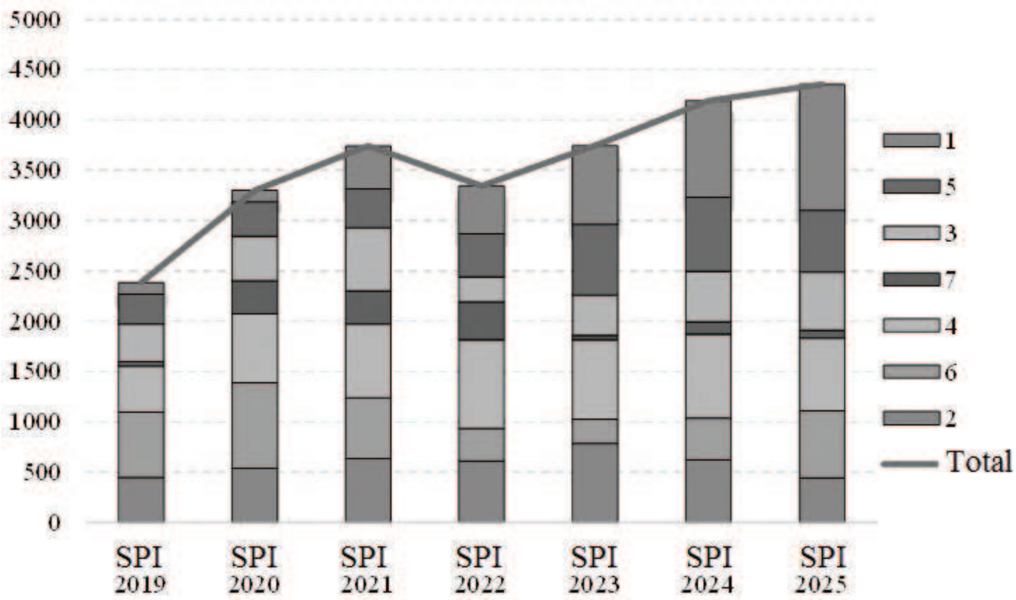

Fig. 4. Dynamics of authors' results according to the SPI methodology.

connected to each other. This group of authors can be considered moderately connected [32]. Figure 4 shows the SPI dynamics of the laboratory's authors (permanent staff). The results of some authors who recently joined the laboratory are not taken into account; their results were added in recent years. The calculations do not take into account the time spent by employees with short tenures. According to the SPI methodology for a given year, the results for the previous three years are taken into account. For example, the 2025 SPI is the results for 2022–2024. Figure 4 shows an increase in the laboratory's overall productivity. This increase was achieved by some authors who actively utilized the publication process with division of labor by section:

- Author 1: from 117 to 1251 points;
- Author 3: from 375 to 579 points;
- Author 5: from 297 to 614 points;
- Author 4: from 456 to 720 points.

7.2. Calculation Conclusions

- Teamwork is most effective, in which authors switch to tasks that best suit them (where they demonstrate the highest productivity based on their qualifications and inclinations);
- The next task in this type of work is assigned to the most effective and least busy author;
- The peak in publications corresponds to the highest efficiency in using the available working time of co-authors;
- For the full implementation of the proposed results, it is critical to track the time of various team members on various tasks;
- The higher the degree of team cohesion, the greater the opportunity to implement the obtained results and the greater the increase in scientific productivity resulting from this implementation.

7.3. Limitations of the Model and Approach

The proposed approach has a number of limitations, which are due to the following reasons:

- It requires time and effort to accumulate reliable statistics and determine time standards for authors to complete individual tasks;

- At the concept stage, which, according to Table 5, takes relatively little time, the researcher cannot always accurately assume how and to what extent this concept can be implemented. In some cases, implementation may prove ineffective;
- There may be significant transaction costs (information exchange, text approval, iterations during model formulation and calculations). In sufficiently large or relatively unconnected teams, it seems realistic to designate "cliques" of related authors, where each "clique" would work on its own set of publications using the proposed methodology.

7.4. Recommendations for Authors and Teams

In the case discussed above, the following options are proposed for increasing scientific productivity.

1. Reduction of ineffective time in "Research work" and "Organization"—shared work that does not directly lead to publications. With a possible 20 percent in time savings, this translates to 125 hours saved per year.
2. These freed-up hours are directed toward publication preparation. Change assessment: 512 hours previously + 125 additionally = 637 hours. Conduct a detailed analysis of individual tasks, with potential time reductions. Estimating potential time savings of 20 percent in time increases the target publication efficiency to $0.71 \times 1.2 = 0.85$ points/hour. Use software packages for calculations. Use generative artificial intelligence to prepare texts. This requires great care and has significant limitations. At its current stage, GAI can generate general descriptions (without taking into account the depth and specificity of the publication) and paraphrase the text, but it makes significant errors in calculations and generating references. Plagiarism checking systems limit the use of generative artificial intelligence. Write the publication text directly in the journal template (saving time on formatting).
3. When planning employee time, consider the feasibility of each employee having some time to develop more individual projects, the consideration of which is not always immediately communicated.

8. CONCLUSION

A substantive and analytical formulation of the problem of increasing the productivity of a team of researchers is presented. It is shown that the problem can be reduced to a nonlinear optimization problem. An approach to the solution is presented that utilizes the specifics of the subject area. Using an example of calculations with real data and an analysis of publication preparation times in one of the ICS RAS laboratories, a multiple increase in research productivity as a result of the proposed methodology is demonstrated. In the ICS RAS laboratory of "large-scale systems," the results of a subgroup of employees actively participating in collaborative projects with division of labor significantly increased. Overall, it can be concluded that the proposed methodology is effective, but not fully or with maximum efficiency. This is due to limited communication among authors within the team, as well as difficulties in defining time standards for completing various tasks by different authors. We believe that this problem formulation and solution methodology are suitable for application not only in ICS RAS laboratories but also in the scientific activities of other research centers (universities, corporate institutes, engineering centers, and other organizations with creative work). This will require appropriate adaptation of the model, restrictions, and data collection and structuring procedures. The ongoing implementation and use of such systematic procedures for monitoring and improving the productivity of intellectual labor is necessary to maintain and develop the country's achievements in the international competitive arena.

REFERENCES

1. National project "Labor Productivity" website. https://xn–80aapampemcchfmo7a3c9ehj.xn–p1ai/projects/proizvoditelnost-truda/. Accessed March 24, 2026.

2. Dranko, O.I., Rezchikov, A.F., Stepanovskaya, I.A., et al., Scenario modeling of economic growth based on indicative planning, *Control Sciences*, 2024, no. 5, pp. 25–41.

3. Higher Attestation Commission of Russia website. https://vak.minobrnauki.gov.ru. Accessed March 24, 2026.

4. List of journals of the Federal State Budgetary Institution "Russian Center for Scientific Information." https://journalrank.rcsi.science/ru/record-sources/. Accessed March 24, 2026.

5. Scopus. https://www.scopus.com/. Accessed March 24, 2026.

6. Web of Science. https://www.webofscience.com. Accessed March 24, 2026.

7. Hirsch, J. E., An index to quantify an individual's scientific research output, *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, no. 46, pp. 16569–16572.

8. Field-Weighted Citation Index (FWCI) Reference Page. https://helpcenter.pure.elsevier.com/en_US/data-sources-and-integrations/field-weighted-citation-impact-fwci-metr. Accessed March 24, 2026.

9. Henneken, E.A. and Kurtz, M.J., Usage bibliometrics as a tool to measure research activity, *Springer Handbook of Science and Technology Indicators*, Springer, Cham, 2019. https://doi.org/10.1007/978-3-030-02511-3_32

10. Hoffmann, C.P., Lutz, C., and Meckel, M., Impact factor 2.0: Applying social network analysis to scientific impact assessment, *Proc. 47th Hawaii Int. Conf. Syst. Sci.*, Waikoloa, HI, USA, 2014, pp. 1576–1585. https://doi.org/10.1109/HICSS.2014.202

11. Altshuller, G.S., *Find an Idea: Introduction to TIPS – Theory of Inventive Problem Solving*, 3rd ed., Moscow, Alpina Publisher, 2010.

12. Myasishchev, V.N., Principles of organizing the scientific study of labor. Ergology and ergotechnics, *Organ. Psychol. Labor Psychol.*, 2023, vol. 8, no. 3, pp. 190–211. https://doi.org/10.38098/ipran.opwp_2023_28_3_009

13. Popov, A., Modern approaches to the implementation and use of lean production model, *BIO Web Conf.*, 2023, vol. 76, pp. 02002.

14. Imai, M., *Kaizen: The Key to Japan's Competitive Success*, New York, McGraw-Hill Education, 1986, 260 p.

15. Randhawa, J. and Ahuja, I., 5S implementation methodologies: Literature review and directions, *Int. J. Product. Qual. Manag.*, 2017, vol. 20, no. 1, pp. 49. https://doi.org/10.1504/IJPQM.2017.080692

16. Karmanov, V.G., *Mathematical Programming*, Moscow, Fizmatlit, 2004.

17. Sinha, S.M., *Mathematical Programming: Theory and Methods*, Elsevier Science, 2006. https://doi.org/10.1016/B978-813120376-7/50002-6.

18. Tsvirkun, A.D., Rezchikov, A.F., Dranko, O.I., et al., Optimization and simulation approach to determining critical combinations of company parameters, *Autom. Remote Control*, 2024, vol. 85, no. 10, pp. 867–874. https://doi.org/10.1134/S0005117924700322

19. Bogomolov, A.S., Dvoryashina, M.M., Dranko, O.I., et al., Stress testing of nonfinancial organizations: An analytical approach to solving the reverse problem, *Control Sciences*, 2021, no. 6, pp. 13–24.

20. Rezchikov, A.F., Dranko, O.I., Bogomolov, A.S., et al., Definition of critical combinations of parameters in complex systems, *Proc. 17th Int. Conf. Manag. Large-Scale Syst. Dev. (MLSD'2024)*, Moscow, ICS RAS, 2024, pp. 420–425.

21. Kochieva, T.B. and Novikov, D.A., *Basic Incentive Systems*, Moscow, ICS RAS, 2000, 108 p.

22. Taha, H. A., *Introduction to Operations Research*, Moscow, Williams, 2007, 901 p.

23. Sun, Z., Tian, Z., Xie, X., et al., A meta-cognitive based logistics human resource modeling and optimal scheduling, *Eng. Appl. Artif. Intell.*, 2024, vol. 130, pp. 107760.

24. Torkabadi, A., Mamoudan, M.M., Erdebilli, B., et al., A multi-objective game theory model for sustainable profitability in the tourism supply chain: Integrating human resource management and artificial neural networks, *Syst. Soft Comput.*, 2025, vol. 7, pp. 200217.

25. Wagh, S.D., Implementation of manufacturing execution system (MES) to increase productivity and competitiveness in the manufacturing sector, *Int. J. Res. Appl. Sci. Eng. Technol.*, 2024, no. 12(4), pp. 3703–3714.

26. Alefari, M., Almanei, M., and Salonitis, K., A system dynamics model of employees' performance, *Sustainability*, 2020, vol. 12, pp. 6511.

27. Epebinu, O., Oguntuase, D., and Oruma, A., Talent management and organizational performance: A study in South-West Nigeria, *Bus. Perspect. Rev.*, 2023, vol. 5, no. 2, pp. 28–36. https://doi.org/10.38157/bpr.v5i2.600

28. Nefedova, A.I., Volkova, G.L., Dyachenko, E.L., et al., International mobility and publication activity of young scientists: What statistics, bibliometrics, and the employees themselves say, *J. New Econ. Assoc.*, 2021, no. 4(52), pp. 98–121.

29. Slepykhm, V., Academic inbreeding and productivity of STEM early career researchers in different environments, *Res. Policy*, 2025, vol. 54, no. 6, pp. 105240.

30. Murray, R., *Writing for Academic Journals*, New York, Open University Press-McGraw-Hill, 2013.

31. Information System for the Analysis of Scientific Activity website. https://isand.ipu.ru/. Accessed March 24, 2026.

32. Gubanov, D.A., Novikov, D.A., Chkhartishvili, A.G., et al., Information system for the analysis of scientific activity (ISAND) in the field of management theory, *Probl. Upravlen.*, 2024, no. 3, pp. 42–65.

33. Gubanov, D.A., Kuznetsov, O. P., Sukhoverov, V.S., and Chkhartishvili, A.G., On constructing profiles in the subject space of control theory, *Proc. 9th Int. Conf. "Knowledge-Ontologies-Theories" (ZONT-2023)*, Novosibirsk, Sobolev Inst. Math. SB RAS, 2023, pp. 89–94.

34. Order No. 53 of February 25, 2025, "Regulations on the Institute's Rating Incentive Bonuses from January 1, 2026," ICS RAS.

35. Dranko, O.I., Rezchikov, A.F., Stepanovskaya, I.A., et al., Assessment of goal synergies in managing cyber-physical production systems of a mechanical engineering enterprise, *Sens. Syst.*, 2024, no. 2(274), pp. 23–29. https://doi.org/10.25728/datsys.2024.2.4

36. Tsvirkun, A.D., Rezchikov, A.F., Kushnikov, V.A., et al., Models and methods for checking the attainability of goals and feasibility of plans in large-scale systems using the example of goals and plans for elimination of the consequences of flood, *Autom. Remote Control*, 2023, vol. 84, no. 12, pp. 1283–1293. https://doi.org/10.1134/S000511792312007X

37. Koch, R., *The 80/20 Revolution: Why the Era of Creative Individuals Is Coming and How You Can Achieve Wealth and Prosperity*, Moscow, Poppuri, 2004, 336 p.

38. Dranko, O.I. and Irikov, V.A., The cost-efficiency method as a tool for selecting priority enterprise projects, *Manag. Account.*, 2011, no. 4, pp. 15–20.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 314–335.

SPECIAL ISSUE

Energy Stability Metrics of Linear Continuous Systems

I. B. Yadykin

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
e-mail: jadykin1@mail.ru

Received July 8, 2025
Revised October 14, 2025
Accepted November 20, 2025

Abstract—The concept of basic systems as systems with a transfer function in the form of an inverse characteristic polynomial of the system is introduced. For these basic systems, we obtain spectral decompositions of energy-based stability metrics from their state-space representations in canonical forms, including Jordan, modal, controllability, and observability. We present explicit formulas for the spectral decomposition of the squared H_2-norm, specifically for the case where the transfer function possesses repeated poles. Furthermore, we develop efficient recursive algorithms to compute the spectral decompositions of the inverse matrices of gramians of continuous dynamical systems based on the repeated spectra of the original dynamics matrices. Spectral decompositions of energy stability metrics are obtained in the form of spectral decomposition of the square H_2-norm of the transfer function of the base system for the case of its simple and multiple roots. A new approach to analyzing the stability of linear systems based on an enhanced Routh–Hurwitz stability criterion is proposed, which includes forming Routh tables and Xiao matrices, computing the spectrum of the dynamics matrix, computing spectral decompositions of energy-based metrics, and jointly applying the Routh–Hurwitz stability criterion and the criterion of boundedness of energy-based stability metrics of basic systems. Finally, we conclude with practical recommendations for applying these theoretical and algorithmic results in systems analysis and design.

Keywords: spectral decompositions, canonical forms, basis systems, gramians, Lyapunov and Sylvester differential equations, Cauchy matrices, Xiao matrices, energy metrics

DOI: 10.7868/S1608303226040034

1. INTRODUCTION

It is well known that gramians are solutions to Sylvester and Lyapunov equations, which are the subject of a vast number of scientific publications [1–4]. These algebraic and differential matrix equations also play a fundamental role in control theory. In recent years, interest has grown in developing methods for computing various energy-related metrics to analyze the stability and degree of controllability, reachability, and observability of such systems. The initial spectral decompositions of gramians for linear continuous and discrete systems with simple spectra were derived in [1]. This work employed the spectral decomposition of the integral representation of solutions to Lyapunov or Sylvester equations to achieve this result. Furthermore, it pioneered the analysis of the temporal characteristics of gramians in controllability canonical forms, examining the role of eigenvectors and the kernels of Cauchy matrices in synthesizing optimal low-order dynamical models. Subsequently, a series of studies [5–12] proposed energy-based metrics for assessing the stability and reachability of both stable and unstable linear systems. In control theory practice, the gramian method as an approach to solving Lyapunov and Sylvester differential equations is applied to a range of practical problems, including:

- Optimal placement of sensors and actuators [7, 10–13];
- Simplification of linear and bilinear control system models using energy metrics [1–3];

- Analysis and synthesis of modal control systems [13–15, 17];
- Analysis of the accuracy of control systems under random disturbances [2, 4];
- Condition monitoring and technical diagnostics based on the identification of energy balance anomalies [13, 16, 18].

The gramian method provides both qualitative and quantitative evaluation of the energy localized in weakly damped oscillatory modes. Furthermore, the resonant interaction between closely spaced vibrational modes within a dynamical system can be significantly enhanced. By transforming the system's state equations into controllability and observability canonical forms, it becomes possible to extract invariant energy metrics that independent of coordinate transformations. The structural aspects of controllability and observability in relation to gramian computation and analysis are explored in studies [19, 20]. An important problem of optimal sensor and actuator placement based on various energy functionals, including invariant ellipsoids, is considered in [4, 7].

2. PROBLEM STATEMENT

Consider a stationary MIMO LTI dynamic system of the form

$$\dot{x}(t) = Ax(t) + Bu(t), x(0) = 0, y(t) = Cx(t), \tag{1}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m, y(t) \in \mathbb{R}^m$. In case of SISO LTI systems $u(t) \in \mathbb{R}^1, y(t) \in \mathbb{R}^1$. The Lyapunov and Sylvester equations play an important role in control theory and the design of control systems based on integral performance indicators, as well as in the synthesis of optimal systems using state observers. Two types of Lyapunov equations are used to solve these problems

$$AP + PA^T = -I_n. \tag{2}$$

$$AP + PA^T = -\sum_{j,\eta} e_j e_\eta^T. \tag{3}$$

Equations of the first type arise in MIMO LTI systems defined by equations of state of the general form (1). Equations of the second type arise in SISO LTI systems defined by equations of state in modal canonical form and in canonical forms of controllability and observability. In modal canonical form, we have the following expressions for the matrix A_d and vectors b, c

$$A_d = diag\begin{bmatrix} s_1 & \dots & s_n \end{bmatrix}, b = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T, c = \begin{bmatrix} r_1 & r_2 & \dots & r_n \end{bmatrix}^T, \tag{4}$$

where r_i is residue of the transfer function at the pole "i." For the base system we have

$$c = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T. \tag{5}$$

In the canonical forms of controllability in this case the equations of state have the form

$$\dot{x}_c(t) = A_c^F x_c(t) + b^F u(t), x_c(0) = 0, \tag{6}$$
$$y_c(t) = c^F x_c(t),$$

$$A_c^F = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}, b^F = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix}^T, \tag{7}$$

$$c^F = \begin{bmatrix} \xi_0 & \xi_1 & \dots & \xi_{n-2} & \xi_{n-1} \end{bmatrix},$$

where ξ_i is coefficient of the numerator polynomial of the transfer function

$$W(s) = \frac{\xi_{n-1}s^{n-1} + \cdots + \xi_1 s + \xi_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}.$$

For the base system we have

$$c^F = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \end{bmatrix}. \tag{8}$$

The classical criterion for the stability of continuous linear systems is based on solving equations of the form (2), (3) with a modified right-hand side of the form $-bb^T$. These matrices depend on the similarity transformation matrices, so the controllability gramians and the expressions for the squares of the H_2-norm of the transfer function obtained using the gramians are not invariant under various transformations of the state space. This drawback can be avoided by using right-hand sides of the form (2) or (3). We call these equations of the second type the basic Lyapunov equations, and the corresponding state equations, in which the vector c^F is chosen in the form $c^F = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \end{bmatrix}$ or $c = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$, the basic equations of state. These equations of state correspond to a transfer function (TF)

$$W_{base}\left(s\right) = \frac{1}{s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0} = \frac{1}{N(s)}, \tag{9}$$

where $N\left(s\right)$ is characteristic polynomial of the system.

Definition 1 [5, 6]. Let us call a Xiao matrix a square matrix that has a zero-plaid structure of the form

$$Y = \begin{bmatrix} y_1 & 0 & -y_2 & 0 & y_3 & \ldots \\ 0 & y_2 & 0 & -y_3 & 0 & \ldots \\ -y_2 & 0 & y_3 & 0 & \ldots & \ldots \\ 0 & -y_3 & 0 & \ldots & \ldots & \ldots \\ y_3 & 0 & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 & y_n \end{bmatrix} y_i \in \mathbb{C}, i = \overline{1,n}. \tag{10}$$

The elements of this matrix are calculated using the formulas

$$y_{j\eta} = \begin{cases} 0, \quad \text{if} \;\; j+\eta = 2k+1, \quad k = 1,\ldots,n, \\ y_n = \frac{1}{2Y_{n,1}}, \\ y_{n-l} = \frac{-\sum_{i=1}^{m-1}(-1)^i Y_{n-l,i+1} y_{n-l+i}}{Y_{n-l,1}}, \text{if } j+\eta = 2k, \;\; k = 1,\ldots,n, l = 1,\ldots,n-1, \end{cases}$$

It is known that basic equations play an important role in studying the stability of linear systems. Let us consider a differential equation of state of the form

$$\begin{gathered} y^{(n)}\left(t\right) + a_{n-1}y^{(n-1)}\left(t\right) + \ldots + a_0 y\left(t\right) = \delta\left(t\right), \\ y^{(n-1)}\left(0\right) = 1, \quad y^{(n-2)}\left(0\right) = 0, \quad \ldots, \quad y\left(0\right) = 0, \end{gathered} \tag{11}$$

where $\delta\left(t\right) - \delta$ is Dirac function. This equation corresponds to the transfer function of the base system. It is proved that the controllability gramian for the system (11) has the form of a Xiao matrix (10), in which the elements y_i are calculated using the measured quantities $y^{(i)}\left(t\right)$ [6]

$$y_i = \int_0^\infty \left[y^{(i)}(t)\right]^2 dt. \tag{12}$$

These estimates have the physical meaning of measured energies of the states of the system, and they can be calculated in various ways, including using state observers [19].

Definition 2. The energy metric of the stability of the system (1) is the square of the H_2-norm of its transfer function [10–12].

We will show that basic systems can be completely controllable if additional conditions are met. The controllability matrix for the modal base system (4), (5) is a Vandermonde matrix, which is a full rank matrix only if all eigenvalues of the dynamics matrix are distinct. The controllability matrix for the base system (7), (8) is not singular. Therefore, the basic system (7), (8) is completely controllable. The gramian controllability matrix for the base system (7), (8) in Xiao form has the form [5]

$$P^{cF} = \sum_{k=1}^{n} \sum_{\eta=0}^{n-1} \sum_{j=0}^{n-1} \frac{s_k^j \left(-s_k\right)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n} \left(s_k - s_\lambda\right) \prod_{\lambda=1,}^{n} \left(-s_k - s_\lambda\right)} 1_{j+1\eta+1}, \tag{13}$$

$$P^{cF} = \sum_{k=1}^{n} \sum_{\eta=0}^{n-1} \sum_{j=0}^{n-1} \frac{s_k^j \left(-s_k\right)^\eta}{\dot{N}\left(s_k\right) N\left(-s_k\right)} 1_{j+1\eta+1}. \tag{14}$$

It follows that the matrices of the gramians P^{cF} are positive definite.

Remark 1. It follows from (14) that the Xiao matrix depends on the quadratic forms $s_k^j(-s_k)^\eta$ formed by the eigenvalues of the original base and antistable base systems, on the residues of the transfer function of the base system, and on the values of the characteristic polynomial of the antistable system in its eigenvalues. Quadratic forms define the property of the zero blanket of Xiao matrices. The denominator of the expression (14), which depends on the characteristic polynomial $N(s)$, determines the value of the square of the H_2-norm of the transfer function, which is the energy metric of the system (1). The important role of the base system in the formation of energy sustainability metrics is obvious. Complete controllability of basic systems is one of the conditions for the positive definiteness of their controllability gramians. This simplifies the calculation of controllability gramians for the base system (7), (8), simplifies the transformation of gramians into Hadamard form and the calculation of the squared H_2-norm of the transfer function (9). Moreover, this guarantees the normality of the gramian matrix, which implies that the diagonal elements and traces of the corresponding gramians are positive. Their elements depend only on the eigenvalues of the dynamics matrix and its characteristic polynomial, which do not depend on similarity transformations and, therefore, are invariants under these transformations. As for the basic system (4), (5), the controllability gramian in it takes the simple form of Cauchy matrices [1, 21]

$$P_c = \frac{-1}{s_j + s_\eta^*}, \forall j, \eta = \overline{1, n}, \forall s_j, s_\eta \in \mathbb{C}^-. \tag{15}$$

The symmetric Cauchy matrix is defined as

$$C = \frac{-1}{x_j + x_{\eta,}}, \forall j, \eta = \overline{1, n}, \forall x_j, x_\eta \in \mathbb{R}^-. \tag{16}$$

Theorem 1 [21]. *A symmetric Cauchy matrix of the form* (16) *is positive definite if and only if the numbers x_j, x_η are positive and mutually distinct*

$$0 < x_1 < x_2 < \ldots < x_n, \quad 0 < x_n < x_{n-1} < \ldots < x_1.$$

Corollary 1. *Let us consider the Cauchy matrix of the form* (16)*, of the stable base system* (4), (5)*, or for the system* (7), (8)*. Let us assume that the systems are stable and the eigenvalues of their dynamics matrices are real and mutually distinct. Then a symmetric Cauchy matrix of the form* (16) *is positive definite if and only if the conditions are satisfied*

$$0 < -s_1 < -s_2 < \ldots < -s_n, \quad 0 < -s_n < -s_{n-1} < \cdots < -s_1.$$

In this case, the role of positive numbers $x_j, x_{\eta,}$ is played by the real parts of the eigenvalues of the dynamics matrix $-s_j, -s_\eta$, which may turn out to be complex-valued. This requires representing gramians as stable Hermitian matrices

$$[P_c]_{Herm} = [P_o]_{Herm}, \quad p_{j\eta Herm} = \frac{1}{2}\left(\frac{-1}{s_j + s_\eta^*} + \frac{-1}{s_j^* + s_{\eta,}}\right) = Re\left(\frac{-1}{s_j + s_\eta^*}\right).$$

In [1, Lemma 9.4] it is proved that the Cauchy matrix (15) of a stable diagonalized base system is positive definite. Note that for basic systems their controllability and observability gramians coincide

$$P_c = P_o.$$

In this case, we obtain an important property of the basic system

$$\sigma_k = (-2Re(s_k))^{-1}, k = \overline{1,n},$$

where σ_k are Hankel singular values of gramian matrices. In this case, we obtain a simple formula for the square of the H_2-norm of the transfer function of the base system

$$||W(s)||_2^2 = trP_C = \sum_{k=1}^{n} \sigma_k = \sum_{k=1}^{n} \frac{1}{-2Re(s_k)}.$$

An important role in the study of spectral decompositions of solutions of equations of the form (2), (3), including the gramians of the basic systems, is played by multiple eigenvalues of the dynamics matrices, which is associated with the consideration of equations of state in Jordan form.

The objectives of this study are as follows. The first goal of the paper is to obtain spectral decompositions of the gramian matrices of the underlying continuous dynamical systems from the multiple spectra of their dynamics matrices. Another objective of the paper is to develop recurrent algorithms for calculating the controllability gramian and its inverse gramian for MIMO LTI continuous systems. The third objective is to analyze the energy metrics of stability of linear stationary systems using the algebraic Routh–Hurwitz stability criterion.

3. MAIN RESULTS

Let us consider the problem of calculating spectral decompositions of MIMO LTI and SISO LTI systems using solutions of differential and algebraic Lyapunov equations in the frequency domain. As is well known, if a linear system is stable, its transfer function is strictly proper, but all poles are distinct, then the following formulas are valid for calculating the square of the H_2-norm $W(s)$[1]

$$\|W(s)\|_2^2 = \sum_{k=1}^{n} r_k W(-s_k), r_k = Res\left[N^{-1}(s), s_k\right]. \tag{17}$$

In the case of the base system (4), (5) energy stability metric is separable across the poles of the transfer function, with each spectral component being a positive number. The energy metrics of stability J_1, J_2 have the form [19]

$$J_1 = \sum_{k=1}^{n} \frac{r_k}{\dot{N}(s_k) N(-s_k)}, \tag{18}$$

$$J_2 = \sum_{k=1}^{n} \sum_{\rho=1}^{n} \frac{-1}{s_k + s_\rho} \frac{1}{\dot{N}(s_k)\dot{N}(s_\rho)}. \tag{19}$$

(18) can be represented in an equivalent form

$$J_1 = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{r_k}{\prod_{\lambda=1,\lambda\neq k}^{n} (s_k - s_\lambda) \prod_{\lambda=1,}^{n} (-s_k - s_\lambda)}.$$

Since the matrix of a stable diagonalized base system is positive definite, (19) can be represented as the trace of the Cauchy matrix

$$J_2 = tr \sum_{k=1}^{n} \sum_{\rho=1}^{n} Re\left(\frac{-1}{s_k + s_\rho^*}\right) e_k e_\rho^T \, .$$

The spectral decomposition of the transfer function TF by multiple poles of the basic SISO LTI system in general has the form [24]

$$W(s) = \sum_{k=1}^{n} \sum_{\nu=1}^{m_k} \frac{L_{k\nu}}{(s - s_k)^{m_k - \nu + 1}}, \tag{20}$$

where s_k is the pole of the transfer function, $L_{k\nu}$ is the residue of the transfer function of order "ν" at the pole s_k

$$L_{k\nu} = \frac{1}{(\nu - 1)!} \left[\frac{d^{\nu-1}}{ds^{\nu-1}} \left(\frac{(s - s_k)^\nu}{N(s)}\right)\right]_{s=s_k} . \tag{21}$$

For basic systems, calculating the residue of the transfer function TF over multiple poles is greatly simplified. Let the strictly eigentransfer function TF of the general form of a SISO LTI system have the form

$$W(s) = \frac{1}{N(s)} = \frac{1}{\prod_{i=1}^{n} (s - s_i)^{m_i}}.$$

Let us introduce the notation

$$N_{k1}(s) = (s - 1)^{m_1} \times \ldots \times (s - s_{k-1})^{m_{k-1}} \times (s - s_{k+1})^{m_{k+1}} \times \ldots (s - s_n)^{m_n} ,$$
$$N_{k1}(s_k) = (s_k - s_1)^{m_1} \times \ldots \times (s_k - s_{k-1})^{m_{k-1}} \times (s_k - s_{k+1})^{m_{k+1}} \times \ldots (s_k - s_n)^{m_n} .$$

Then the residue $L_{k\nu}$ of the transfer function $W(s)$ can be calculated using the following recurrent algorithm.

The first step of the algorithm $\nu = 1$

$$L_{k1}(s_k,\ s_\mu) = N_k^{-1}(s_k,\ s_\mu) \quad \forall k, \mu = \overline{1, n}.$$

The second step of the algorithm $\nu = 2$

$$L_{k2}(s_k,\ s_\mu) = -N_k^{-1}(s_k,\ s_\mu) \left(\frac{m_1}{s_k - s_1} + \ldots + \frac{m_{k+1}}{s_k - s_{k+1}} + \ldots + \frac{m_n}{(s_k - s_n)}\right), \quad \forall k, \mu = \overline{1, n},$$

or

$$L_{k2}(s_k,\ s_\mu) = -L_{k1}(s_k,\ s_\mu) \sum_{\mu=1,\mu\neq k}^{m} \frac{m_\mu}{(s_k - s_\mu)}, \quad \forall k, \mu = \overline{1, n},$$

$$L_{k3} = \frac{1}{2!}\left[L_{k2} - \frac{1}{N_{k1}(s)} \sum_{\mu=1,\mu\neq k}^{n} \frac{m_\mu}{(s - s_\mu)^2}\right]\Bigg|_{s=s_k}, \quad \forall k = \overline{1, n},\ \forall \nu = \overline{1, m_k}.$$

The limiting relations follow from the residue formulas

$$\lim_{s_k \to s_i} |L_{k\nu}| = \infty.$$

Theorem 2. *Spectral decompositions of energy metrics of stability of basic systems by multiple poles of the TF.*

Consider a basic general-purpose SISO LTI system with multiple poles and a transfer function of the form (18)*. We assume that the system is fully controllable, strictly eigenvalue-free, and stable.*

Then the following statements are true:

1) *The spectral decomposition of the energy metric of stability of the base system by multiple poles of the TF has the form*

$$J_1 = J_2 = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k} \frac{(-1)^{m_k-\nu}}{(m_k-\nu)!}\left\{ L_{k\nu}\left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left(\frac{1}{N(-s)}\right)\right]_{s=s_k}\right\}_{Herm}, \tag{22}$$

$$L_{k\nu} = \frac{1}{(\nu-1)!}\left[\frac{d^{\nu-1}}{ds^{\nu-1}}\left(\frac{(s-s_k)^{\nu}}{N(s)}\right)\right]_{s=s_k}. \tag{23}$$

2) *The spectral decomposition of the energy metric of the base system by combinational multiple poles of the TF has the form*

$$J_2 = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\sum_{\rho=1}^{n}\sum_{\mu=1}^{m_\rho} \frac{(m_k-\nu+m_\rho-\mu+1)!}{(m_k-\nu)!}\left[L_{k\nu}L_{\rho\mu}\frac{1}{(-s_k-s_\rho^*)^{m_k-\nu+m_\rho-\mu+1}}\right]_{Herm}, \tag{24}$$

where

$$L_{\rho\mu} = \frac{1}{(\mu-1)!}\left[\frac{d^{\mu-1}}{ds^{\mu-1}}\left(\frac{(s-s_\rho)^{\mu}}{N(s)}\right)\right]_{s=s_\rho}. \tag{25}$$

The energy metrics of the base system (4), (5) *or the system* (7), (8) J_1, J_2 *are positive numbers. They are invariants under various similarity transformations.*

Proof. Let us consider the following expansions of the transfer functions of the original system (1) for the case of its multiple poles and its antistable subsystem

$$W(-s) = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k} \frac{L_{k\nu}}{(-s-s_k)^{m_k-\nu+1}}, \tag{26}$$

where $L_{k\nu}$ is residue of order "ν" at the pole s_k of the TF

$$L_{k\nu} = \frac{1}{(\nu-1)!}\left[\frac{d^{\nu-1}}{ds^{\nu-1}}\left(\frac{(-s-s_k)^{\nu}}{N(-s)}\right)\right]_{s=s_k}.$$

Let us take advantage of the fact that in this case the transfer function, like its inverse Laplace transform, are scalar functions

$$J = \left\|N^{-1}(s)\right\|_2^2 = \int_0^{\infty} L^{-1}\left[N^{-1}(s)\right] L^{-1}\left[N^{-1}(-s^*)\right] dt. \tag{27}$$

We calculate the integral (27) as the limit at $T \to \infty$ of another integral

$$J = \lim_{T\to\infty}\int_0^{T} L^{-1}\left[N^{-1}(s)\right] L^{-1}\left[N^{-1}(-s)\right] dt.$$

Let us introduce the antiderivative function

$$H(t) = L^{-1}\left[N^{-1}(s)\right] L^{-1}\left[N^{-1}(-s)\right],$$

where

$$\left[N^{-1}(s)\right] = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k} L_{k\nu}\frac{1}{(s-s_k)^{m_k-\nu+1}},$$

$$L_{k\nu} = \frac{1}{(\nu-1)!}\left[\frac{d^{\nu-1}}{ds^{\nu-1}}\left(\frac{(s-s_k)^{\nu}}{N(s)}\right)\right]_{s=s_k},$$

$$L^{-1}\left[N^{-1}(s)\right] = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\frac{L_{k\nu}}{(m_k-\nu)!}t^{m_k-\nu}e^{s_k t},$$

$$\left[N^{-1}(-s)\right] = \sum_{\rho=1}^{n}\sum_{\mu=1}^{\rho} L_{\rho\mu}\frac{1}{(-s-s_k)^{m_\rho-\mu+1}},$$

$$L_{\rho\mu} = \frac{1}{(\mu-1)!}\left[\frac{d^{\mu-1}}{ds^{\mu-1}}\left(\frac{(-s-s_\rho)^{\mu}}{N(-s)}\right)\right]_{s=s_\rho},$$

$$L^{-1}\left[N^{-1}(-s)\right] = \sum_{\rho=1}^{n}\sum_{\mu=1}^{m_\rho}\frac{L_{\rho\mu}}{(m_\rho-\mu)!}t^{m_\rho-\mu}e^{-s_\rho^* t}.$$

The first statement of the theorem is related to the decomposition of the square of the H_2-norm of $N^{-1}(s)$, when in the primitive function only the function is subject to spectral decomposition $L^{-1}\left[N^{-1}(s)\right]$. In this case, the theorem on the Laplace transform of the product of complex functions of time, the image of which is a fractional rational fraction, takes the form

$$L\left[L^{-1}\left[N^{-1}(s)\right]\times N^{-1}(-s)\right] = L\left[H(t)\right]$$
$$= H(s) = \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\frac{(-1)^{m_k-\nu}L_{k\nu}}{(m_k-\nu)!}\left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left(\frac{1}{N(-s)}\right)\right]_{s=s-s_k}.$$

Thus

$$J_1 = \left\|N^{-1}(s)\right\|_2^2 = -H(0) = J_2$$
$$= \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\frac{(-1)^{m_k-\nu}}{(m_k-\nu)!}\left\{L_{k\nu}\left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left(\frac{1}{N(-s)}\right)\right]_{s=s_k}\right\}_{Herm}.$$

This proves the first statement of the theorem. If we add to this scheme the spectral decomposition of the second function $L^{-1}\left[N^{-1}(s)\right]$ and we use the theorem on the Laplace transform of the product of complex functions of time, the image of which is a fractional rational fraction, then we obtain

$$L\left[L^{-1}\left[N^{-1}(s)\right]\times N^{-1}(-s)\right] = L\left[H(t)\right]$$
$$= \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\frac{(-1)^{m_k-\nu}}{(m_k-\nu)!}L_{k\nu}\left\{\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left[\sum_{\rho=1}^{n}\sum_{\mu=1}^{m_\rho}L_{\rho\mu}\frac{1}{(-s-s_\rho^*)^{m_\rho-\mu+1}}\right]_{s=s_k}\right\}$$
$$= \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\sum_{\rho=1}^{n}\sum_{\mu=1}^{m_\rho}\frac{(m_k-\nu+m_\rho-\mu+1)!}{(m_k-\nu)!}\times\left[L_{k\nu}L_{\rho\mu}\frac{1}{(-s_k-s_\rho^*)^{m_k-\nu+m_\rho-\mu+1}}\right]_{Herm}.$$

Since $H(t)$ is a primitive function, then from (20), taking into account the stability of the system, it follows

$$J_2 = \left\| N^{-1}(s) \right\|_2^2 = H(t)]_0^{\infty} = -H(0)$$
$$= \sum_{k=1}^{n} \sum_{\nu=1}^{m_k} \sum_{\rho=1}^{n} \sum_{\mu=1}^{m_\rho} \frac{(m_k-\nu+m_\rho-\mu+1)!}{(m_k-\nu)!} \left[L_{k\nu} L_{\rho\mu} \frac{1}{(-s_k - s_\rho^*)^{m_k-\nu+m_\rho-\mu+1}} \right]_{Herm}.$$

This proves the second assertion of the theorem. Note that the dynamics matrices of the underlying system were not used in the proof of the theorem. It follows that its statements are valid both for the base system (4), (5) and for the system (7), (8). Moreover, the energy metrics J_1, J_2 are equal to each other, they are positive numbers, since the H_2-norm of the transfer function is a positive number. They are invariants under various similarity transformations, since they are functions of the eigenvalues of the system dynamics matrix.

Let us further consider the differential and algebraic Lyapunov equations of the form

$$\frac{dP(t)}{dt} = A_d P(t) + P(t) A_d^T + b_d b_d^T, \quad P(0) = 0_{nxn}, \quad t \in [0, T], \tag{28}$$

$$A_d P(t) + P(t) A_d^T = -b_d b_d^T. \tag{29}$$

Corollary 2. *Let the conditions of Theorem 2 be satisfied, but all poles of the TF be distinct. Then the following statement is true:*

$$\begin{gathered} J_1(s_1, \ldots, s_n) = \sum_{k=1}^{n} \frac{1}{\dot{N}(s_k)} \frac{1}{(-s_k)^n + \ldots + a_1(-s_k) + a_0} > 0, \\ J_2(s_1, \ldots, s_n) = \sum_{k=1}^{n} \sum_{\rho=1}^{n} J_{2,k\rho}(s_1, \ldots, s_n) > 0, \\ J_{2,k\rho}(s_1, \ldots, s_n) = 2Re \frac{r_k r_\rho}{(-s_k - s_\rho)} = 2\alpha_{k\rho} Re \frac{(r_k r_\rho)}{\alpha_{k\rho}^2 + \beta_{k\rho}^2}, \\ \alpha_{k\rho} = -Res_k - Res_\rho, \beta_{k\rho} = -\Im s_k - \Im s_\rho. \end{gathered} \tag{30}$$

Proof. Let us consider the basic system (6), (7) and formulas (11), (12). Since for the basic system the input and output vectors have the form (7), (8), we obtain

$$J_2(s_1, \ldots, s_n) = c^F P_c^F (c^F)^T = y_{n-1} = \int_0^{\infty} k^2(\tau)\, d\tau,$$

where $k(\tau)$ is the impulse response of the system. This is where the formulas for the statements follow.

Theorem 3. *Suppose that the SISO LTI system* (1) *is fully controllable, there exists a non-degenerate coordinate transformation $x_d = Tx$ such that*

$$A_d = TAT^{-1}.$$

Let us assume that all eigenvalues of the matrix A are different, do not belong to the imaginary axis, and the conditions are satisfied

$$s_i + s_j \neq 0, \qquad \forall i, j = \overline{1, n}.$$

Then the following statements are true:

1) *The solution of the equation* (28) *in the complex domain exists, is unique*

$$L\left[P\left(t\right)\right]=\frac{1}{s}\left\{\sum_{i,j=1}^{n}\frac{-1}{\left(s_i+s_j^*\right)}\left[Res\left(Is-A_d\right)^{-1},s_i\right]b_{di}b_{dj}\left[Res\left(Is-A_d^*\right)^{-1},s_j^*\right]\right\}_{Herm}, \qquad (31)$$

where b_{di} *is the "i" element of the vector* b_d, b_{dj} *is the "j" element of the vector* b_d.

2) *The solution of the equation* (28) *in the real domain is unique and has the form*

$$P\left(0,T\right)=\int_0^T\left\{\sum_{i,j=1}^{n}\frac{-1}{\left(s_i+s_j^*\right)}b_d b_d^T e^{\left(s_i+s_j^*\right)\tau}e_j e_\eta^T d\tau\right\}_{Herm}. \qquad (32)$$

3) *The solution of the algebraic equation* (29) *in the real domain exists, is unique, and has the form*

$$P\left(0,\infty\right)=\left\{\sum_{i,j=1}^{n}\frac{-b_{di}b_{dj}}{\left(s_i+s_j^*\right)}e_j e_\eta^T\right\}_{Herm}. \qquad (33)$$

4) *The matrix* $P^{-1}\left(0,\infty\right)$ *under the conditions of the theorem exists and is unique*

$$P^{-1}\left(0,\infty\right)=\frac{-P_0}{p_0}, \qquad (34)$$

where P_0 *is the free term of the Faddeev matrix polynomial,* p_0 *is the free term of the characteristic polynomial of the matrix* $P(0,T)$, *determined using the Faddeev–Leverrier recurrent algorithm* [22, 23].

Proof. Applying the Laplace transform to matrix functions of time has its own peculiarities. Let's consider diagonal matrices of size $n\times n$

$$F_1\left(t\right)=diag\left[\ldots\ \ldots\ f_{1,\nu}(t)\ \ldots\ \ldots\right],F_2\left(t\right)=diag\left[\ldots\ \ldots\ f_{2,\eta}(t)\ \ldots\ \ldots\right],$$

where the functions $f_{1,\nu}\left(t\right), f_{2,\eta}(t)$ are Laplace transformable, and their images $f_{1,\nu}\left(s\right), f_{2,\eta}\left(s\right)$ are rational fractions that have only n poles and no zeros

$$f_{1,i}\left(s\right)=\frac{1}{\left(s-s_{1,i}\right)},\quad f_{2,j}\left(s\right)=\frac{1}{\left(s-s_{2,j}\right)}.$$

Let us show that for each element "ij" of the Laplace image of the product of matrices $F_1\left(t\right)b_{di}b_{dj}F_2\left(t\right)$ the following formulas are valid

$$L\left[e_iF_1\left(t\right)b_{di}b_{dj}F_2\left(t\right)e_j^T\right]=\sum_{i,j=1}^{n}1_{ii}b_{di}b_{dj}\frac{1}{\left(s-s_{1,i}-s_{2,j}\right)}1_{jj},\quad 1_{ii}=e_ie_i^T, \qquad (35)$$

$$f_{1,i}\left(t\right)=e^{s_it},\quad f_{2,j}\left(t\right)=e^{s_j^*t}.$$

The identities for the residues of resolvents are also valid

$$\left[Res\left(Is-A_d\right)^{-1},s_i\right]=1_{ii},\quad\left[Res\left(Is-A_d^*\right)^{-1},s_j^*\right]=1_{jj}.$$

Using the theorem on the Laplace transform of a product of real functions of time whose image is a fractional rational algebraic function, we obtain (35). As is known, the solution to the Lyapunov differential equation with zero initial conditions is the Lyapunov integral of the form

$$P\left(t\right)=\int_0^T e^{At}b_d b_d^T e^{A^*t}dt.$$

Applying (35), we obtain

$$P(0,T) = \int_0^T \left\{ \sum_{i,j=1}^{n} \frac{-1}{\left(s_i + s_j^*\right)} b_{di} b_{dj} e^{\left(s_i + s_j^*\right)\tau} e_j e_\eta^T d\tau \right\}_{Herm}. \tag{36}$$

As shown in [13]

$$P(t) = \sum_{i,j=1}^{n} P_{ij}(t), P_{ij}(t) = \left\{ \frac{e^{(s_i+s_j^*)t}}{s_i + s_j^*} 1_{ii} b_{di} b_{dj} 1_{jj} \right\}_{Herm}.$$

The formula in statement 3) is proved by passing to the limit in (36). The existence and uniqueness of a solution to equation (29) in the time domain has been proven in many works, for example [3]. Let us prove the validity of this statement in the complex domain. Since the Lyapunov integral is a solution of the equation (29), it is sufficient to prove this statement for its Laplace transform. Note that the matrix $e^{At} b_d b_d^T e^{A^* t}$ is a smooth, non-zero function of time. For any "ij" element, the inequality

$$\lim_{t\to\infty, \varepsilon\leftrightarrow 0} \int_0^T e_j \left| e^{At} b_d b_d^T e^{A^* t} \right| e_\eta^T dt < \infty, \quad \forall t \in [0,T), \quad \forall i,j = \overline{1,n}.$$

It follows that for all time functions $e_j \left| e^{At} b_d b_d^T e^{A^* t} \right| e_\eta^T$ satisfying the conditions of the theorem, there exists an abscissa of absolute convergence σ, which proves the existence and uniqueness of the direct Laplace transform for the solution (32). On the other hand, if there exists a direct Laplace transform of the solution, then there also exists an inverse transform

$$\frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} e_\nu \left| e^{At} b_d b_d^T e^{A^* t} \right| e_\mu^T e^{ts} ds, \quad c > \sigma, \quad \forall \nu, \mu = \overline{1,n}.$$

Moreover, the inverse transformation is unique in the Lebesgue sense. The validity of assertion 4) follows from the complete controllability of the system [4]. This implies the existence and uniqueness of the inverse matrix $P^{-1}(0,T)$ for any T. We write the expansion of the resolvent of the matrix $P(0,T)$ in a Faddeev–Leverrier series

$$[Is - P(0,T)]^{-1} = \frac{P_{n-1}s^{n-1} + \ldots + P_1 s + P_0}{p_n s^n + p_{n-1}s^{n-1} + \ldots + p_1 s + p_0}. \tag{37}$$

Putting in (37) $s = 0$, we obtain the formula

$$P^{-1}(0,\infty) = \frac{-P_0}{p_0}, \tag{38}$$

where P_0 is the free term of the matrix Faddeev polynomial, p_0 is the free term of the characteristic polynomial of the matrix $P(0,T)$. They are determined using the Faddeev–Leverrier recurrent algorithm [22, 23]

$$\begin{gathered} P_{n-1}s^{n-1} = Ip_n s^{n-1}, p_n = 1, \\ P_{n-2}s^{n-2} = \left(Ip_{n-1} + p_{n-2}P(0,T)\right) s^{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_0 = \left(Ip_1 + p_2 P(0,T) + \ldots + p_n P(0,T)^{n-1}\right). \end{gathered} \tag{39}$$

The coefficients p_i are determined by the formulas

$$p_{n-i} = \frac{-1}{i} tr\left(P(0,T) P_{n-i}\right).$$

Remark 2. In [13] it was proved that the results of the theorem can be extended to the case of canonical Jordan forms of the equations of state (1) for the case when the characteristic equation of the dynamics matrix of the SISO LTI system (1) can be represented in the form

$$N(s) = \prod_{k=1}^{n} (s - s_k)^{m_k}, \sum_{k=1}^{n} m_k = q.$$

The solution of the equation (3) in the real domain for multiple eigenvalues s_k of the matrix A_d with multiplicity m_k has the form

$$P(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{j\eta}(t),$$

$$P_{j\eta}(t) = \sum_{j=0}^{n-1}\sum_{\eta=0}^{n-1} h_{j\eta}(t)\, b_d b_d^T e_j e_\eta^T,$$

$$h_{j\eta}(t) = L^{-1}\left\{ s^{-1} \sum_{k=1}^{n}\sum_{\nu=1}^{m_k} L_{k,\nu j} \frac{(-1)^{m_k-\nu}}{(m_k-\nu)!} \left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}} \left(\frac{s^\eta}{N(-s)} \right) \right]_{s=s-s_k} \right\},$$

$$L_{k,\nu j} = \frac{1}{(\nu-1)!} \left[\frac{d^{\nu-1}}{ds^{\nu-1}} \left(\frac{(s-s_k)^{m_k} s^j}{N(s)} \right) \right]_{s=k}.$$

Let us consider a simple basic SISO LTI system with dimension $n = 3$, represented by the equations of state in diagonal canonical form of the form

$$\dot{x}_d = A_d x_d + b_d u,$$

$$A_d = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, b_d = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

$$N(s) = a_3 s^3 + a_2 s^2 + a_1 s + a_0, a_3 = 1, a_2 = 6, a_1 = 11, a_0 = 6,$$

$$s_1 = -1, s_2 = -2, s_3 = -3.$$

Let's calculate the solution to the algebraic Lyapunov equation. Since in this example we are solving not a differential equation, but an algebraic Lyapunov equation for a stable system, (38) becomes the

$$P(\infty) = [p_{ij}(\infty)], \quad p_{ij}(\infty) = \frac{-b_{di} b_{dj}}{s_i + s_j}. \tag{40}$$

Substituting the parameters of the example into the formula, we obtain

$$P(\infty) = \begin{bmatrix} 0.5 & 0.666 & 0.75 \\ 0.666 & 1 & 1.2 \\ 0.75 & 1.2 & 1.5 \end{bmatrix}.$$

Putting $s = 0$ in (37), we obtain

$$P^{-1} = \frac{-P_0}{p_0}. \tag{41}$$

The first step of the recurrent algorithm. We set $P_2 = I_2, p_3 = 1$,

$$k = 0.$$

Step two. Calculate P_1, p_2

$$k = 1,$$

$$P_1 = p_2 I_3 + PP_2 = \begin{bmatrix} -2.5 & 0.666 & 0.75 \\ 0.666 & -2 & 1.2 \\ 0.75 & 1.2 & -1.5 \end{bmatrix},$$

$$p_2 = \frac{-1}{1} tr\,(PP_2) = -3.$$

Step three. Calculate P_0, p_1

$$k = 2,$$

$$P_0 = p_2 I_3 + PP_2 = \begin{bmatrix} 0.06 & -0.099 & 0.0492 \\ -0.099 & 0.1875 & -0.1005 \\ 0.0492 & -0.1005 & 0.5844 \end{bmatrix},$$

$$p_1 = \frac{-1}{2} tr\,(PP_1) = -0.30394.$$

Step four: Calculate p_0

$$k = 3,$$

$$p_0 = \frac{-1}{3} tr\,(PP_2) = -0.000966, p_0^{-1} = -1035.19.$$

We substitute the expressions obtained in the last and penultimate steps into (37)

$$P^{-1} = \begin{bmatrix} 62.114 & -102.4838 & 50.9313 \\ -102.4838 & -194.0981 & -104.36 \\ 50.9313 & -104.36 & 58.426 \end{bmatrix}.$$

It is not difficult to verify that the resulting matrix is the inverse matrix of the controllability gramian.

Theorem 4. *Let's consider the MIMO LTI system* (1) *in the canonical form of controllability and write the Lyapunov differential equation*

$$\frac{dP(t)}{dt} = A_c^F P(t) + P(t) \left(A_c^F\right)^T + B^F \left(B^F\right)^T, \quad P(0) = 0_{nxn}, \quad t \in [0, T]. \tag{42}$$

Let us assume that the system (1) *is completely controllable, there exists a non-degenerate coordinate transformation* $x_c = Tx$, *such that*

$$A_c^F = TAT^{-1}.$$

Let us assume that all eigenvalues of the matrix A_c^F *are different, do not belong to the imaginary axis, and the conditions* $s_i + s_j \neq 0, \forall i, j = \overline{1, n}$ *are satisfied.*

Then the following statements are true:

1) *The solution of the equation* (42) *in the complex domain exists, is unique, and has the form*

$$[P(s)] = \frac{1}{s} \left\{ \sum_{j=1}^{n} \left[Res \left(Is - A_c^F \right)^{-1}, s_j \right] B^F \left(B^F \right)^T \left\{ \left[\left[-Is - (A_c^F)^T \right]^{-1} \right] \right\} \right\}_{Herm}, \tag{43}$$

where

$$A_{cj}^F = a_{j+1} I + a_{j+2} A_c^F + \ldots + a_n \left(A_c^F\right)^{n-j-1}, \quad j = \overline{0, n-1},$$

$$A_{c\eta}^F = a_{\eta+1} I + a_{\eta+2} A_c^F + \ldots + a_n \left(A_c^F\right)^{n-\eta-1}, \ \eta = \overline{0, n-1}.$$

2) *The solution of the equation* (42) *in the real domain is unique and has the form*

$$P_{j\eta}(t) = \frac{\sum_{k=1}^{n} s_k^j (-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)}(e^{s_k t} - 1)A_{cj}^F B^F \left(B^F\right)^T \left(A_{c\eta}^F\right)^T \tag{44}$$
$$= \Omega_{c,j\eta}^F(t) \circ \Psi_{c,j\eta}^F,$$
$$\Omega_{c,j\eta}^F(t) = \frac{\sum_{k=1}^{n} s_k^j (-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)}(e^{s_k t} - 1)e_j e_\eta^T,$$
$$\Psi_{c,j\eta}^F = A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T.$$

3) *The infinite controllability gramian has the form*

$$P(0,\infty) = \left\{\sum_{j=0}^{n-1}\sum_{\eta=0}^{n-1}\frac{\sum_{k=1}^{n} s_k^j (-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)}e_j e_\eta^T\right\}_{Herm} \tag{45}$$
$$\times A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T.$$

4) *The matrix $P^{-1}(0,\infty)$ under the conditions of the theorem exists and is unique*

$$P^{-1}(0,\infty) = \frac{-P_0^F}{p_0}, \tag{46}$$

in which P_0^F is the free term of the matrix Faddeev polynomial for the equations of state in the canonical form of controllability, p_0 is the free term of the characteristic polynomial of the gramian, determined using the Faddeev–Leverrier recurrent algorithm [22, 23].

5) *The results of the theorem can be extended to the case of canonical Jordan forms of the equations of state* (42) *for the case when the characteristic equation of the dynamics matrix has "k" different roots of multiplicity m_k. The solution of the equation* (42) *in the real domain for multiple eigenvalues s_k of the matrix A_c^F with multiplicity m_k has the form*

$$P(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{j\eta}(t),$$
$$P_{j\eta}(t) = h_{j\eta}(t)\, A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T,$$
$$h_{j\eta}(t) = L^{-1}\left\{s^{-1}\sum_{k=1}^{n}\sum_{j=1}^{m_k} L_{k,j}\frac{(-1)^{m_k-\nu}}{(m_k-\nu)!}\left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left(\frac{s^\eta}{N(-s)}\right)\right]_{s=s-s_k}\right\}, \tag{47}$$
$$L_{k,j} = \frac{1}{(j-1)!}\left[\frac{d^{j-1}}{ds^{j-1}}\left(\frac{(s-s_k)^{m_k}s^j}{N(s)}\right)\right]_{s=k}. \tag{48}$$

Proof. The general solution method is based on the spectral decomposition of the solution of the Lyapunov integral. The proof of the existence and uniqueness of solutions repeats the analogous proof of Theorem 3. The validity of Assertions 2–4 and (39) are proved in exactly the same way. The formulas themselves have a different form, determined by the dynamics matrix in controllability form. The general solution for the case of different roots of the characteristic equation in the complex domain has the form [13]

$$L[P(t)] = \frac{1}{s}\left\{\sum_{j=1}^{n}\left[Res\left(Is - A_c^F\right)^{-1}, s_j\right] B^F \left(B^F\right)^T \left\{\left[\left[-Is - (A_c^F)^T\right]^{-1}\right]\right\}\right\}_{Herm}.$$

Let us introduce the matrix antiderivative function $H_1(t)$ and the scalar matrix function $H(t)$ of the form

$$H_1(t) = H(t)\sum_{j=0}^{n-1}\sum_{\eta=0}^{n-1} A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T,$$

$$H(t) = \sum_{k=1}^{n}\sum_{j=0}^{n-1}\sum_{\eta=0}^{n-1} \frac{s_k^j(-s_k)^\eta}{\dot{N}(s_k)N(-s_k)} 1_{j+1\eta+1},$$

$$A_{cj}^F = a_{j+1}I + a_{j+2}A_c^F + \ldots + a_n\left(A_c^F\right)^{n-j-1}, j = \overline{0, n-1},$$

$$A_{c\eta}^F = a_{\eta+1}I + a_{\eta+2}A_c^F + \ldots + a_n\left(A_c^F\right)^{n-\eta-1}, \eta = \overline{0, n-1}.$$

According to (14), the matrix $H(t)$ is a Xiao matrix. It follows that the matrix $P_{j\eta}(t)$ can be represented as a Hadamard product [13]

$$P_{j\eta}(t) = \frac{\sum_{k=1}^{n} s_k^j(-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)}(e^{s_k t} - 1)A_{cj}^F B^F \left(B^F\right)^T \left(A_{c\eta}^F\right)^T$$
$$= \Omega_{c,j\eta}^F(t) \circ \Psi_{c,j\eta}^F,$$

$$\Omega_{c,j\eta}^F(t) = \frac{\sum_{k=1}^{n} s_k^j(-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)}(e^{s_k t} - 1)e_j e_\eta^T, \Psi_{c,j\eta}^F = A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T.$$

From these formulas it follows that the infinite controllability gramian for the case of different roots of the characteristic equation in the real domain has the form

$$P(0,\infty) = \left\{\sum_{j=0}^{n-1}\sum_{\eta=0}^{n-1} \frac{\sum_{k=1}^{n} s_k^j(-s_k)^\eta}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k - s_\lambda)\prod_{\lambda=1}^{n}(-s_\rho - s_\lambda)} e_j e_\eta^T\right\}_{Herm}$$
$$\times A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T.$$

This proves the validity of statements 1–3. The validity of statement 4 is proved in the same way as Theorem 3. To construct an infinite gramian of controllability, the recurrent Faddeev–Leverrier algorithm is used

$$P^{-1}(0,\infty) = \frac{-P_0}{p_0},$$

in which P_0 is the free term of the Faddeev matrix polynomial, p_0 is the free term of the characteristic polynomial of the controllability gramian matrix

$$P(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{j\eta}(t),$$

$$P_{j\eta}(t) = h_{j\eta}(t) A_{cj}^F B^F \left(B^F\right)^T (A_{c\eta}^F)^T,$$

$$h_{j\eta}(t) = L^{-1}\left\{s^{-1}\sum_{k=1}^{n}\sum_{j=1}^{m_k} L_{k,j}\frac{(-1)^{m_k-\nu}}{(m_k-\nu)!}\left[\frac{d^{m_k-\nu}}{ds^{m_k-\nu}}\left(\frac{s^\eta}{N(-s)}\right)\right]_{s=s-s_k}\right\},$$

$$L_{k,j} = \frac{1}{(j-1)!}\left[\frac{d^{j-1}}{ds^{j-1}}\left(\frac{(s-s_k)^{m_k}s^j}{N(s)}\right)\right]_{s=k}.$$

Corollary 3. *Let us generalize the results of Theorem 4 to continuous linear SISO LTI systems. Assume that all the conditions of Theorem 4 are satisfied with MIMO LTI systems replaced by SISO LTI systems. Then the following spectral decomposition of the gramian of controllability of a SISO LTI system in the real domain holds*

$$P\left(t\right)=\sum_{k=1}^{n}\sum_{j=1}^{n}\sum_{\eta=1}^{n}\frac{s_k^{j-1}(-s_k)^{\eta-1}}{\prod_{\lambda=1,\lambda\neq k}^{n}\left(s_k-s_\lambda\right)\prod_{\lambda=1}^{n}\left(-s_\rho-s_\lambda\right)}(e^{s_k t}-1)e_j e_\eta^T.$$

Proof. For a system in canonical controllability form, the decomposition of the resolvents of the dynamics matrices A^F, A^{F*} has a simple form, therefore the formulas are valid

$$(Is-A^F)^{-1}b^F=\begin{bmatrix}0\\ \vdots\\ 0\\ N(s)^{-1}\end{bmatrix}+\begin{bmatrix}0\\ \vdots\\ N(s)^{-1}\\ 0\end{bmatrix}+\ldots+\begin{bmatrix}N(s)^{-1}\\ \vdots\\ 0\\ 0\end{bmatrix},$$

$$b^{FT}(-Is-A^{F*})^{-1}=\begin{bmatrix}0\\ \vdots\\ 0\\ N(-s)^{-1}\end{bmatrix}^T+\begin{bmatrix}0\\ \vdots\\ N(-s)^{-1}\\ 0\end{bmatrix}^T+\ldots+\begin{bmatrix}N\left(-s\right)^{-1}\\ \vdots\\ 0\\ 0\end{bmatrix}^T.$$

Substituting them into (43), using the theorem on the Laplace transform of the product of complex functions of time, the image of which is a fractional rational fraction, and the theorem on displacement in the complex domain, we prove Corollary 3 [13].

Corollary 4 [13, Theorem 1]. *Assume that the conditions of Theorem 4 are satisfied for a SISO LTI system. Furthermore, assume that the system is asymptotically stable. Then, the spectral decomposition of the inverse gramian of the LTI system's controllability in the complex domain is a solution to the systems of equations*

$$\vec{X}=\left(I_n\otimes P_c^F\right)^{-1}\vec{I_n}, \tag{49}$$

$$P_c^F=\sum_{k=1}^{n}\sum_{\eta=0}^{n-1}\sum_{j=0}^{n-1}\frac{s_k^j\left(-s_k\right)^\eta}{\dot{N}\left(s_k\right)N\left(-s_k\right)}\mathbf{1}_{j+1\eta+1}. \tag{50}$$

Proof. From the conditions of the theorem it follows that the solution of the system (49)–(50) exists and is unique. In [13] the following algorithm for calculating the matrices of inverse gramians in Xiao form was proposed. Let X be the matrix of the inverse gramian of controllability, x_i be the "ith" column of the matrix X. Then the inverse gramian is defined by the formula

$$P_c^F x_i=e_i, \qquad i=\overline{1,n},$$

where e_i-x_i is the "ith" column of the matrix I_n. These equations can be rewritten as solutions to Sylvester's algebraic equation

$$\left(I_n\otimes P_c^F\right)\vec{X}=\vec{I_n}, \quad X=\left(P_c^F\right)^{-1}.$$

Let's rewrite (14) as

$$P_c^F=\sum_{k=1}^{n}P_{ck}^F, \quad P_{ck}^F=\sum_{\eta=0}^{n-1}\sum_{j=0}^{n-1}\frac{s_k^j\left(-s_k\right)^\eta}{\dot{N}\left(s_k\right)N\left(-s_k\right)}\mathbf{1}_{j+1\eta+1},$$

where P_{ck}^F is the controllability subgramian matrix corresponding to the pole s_k

$$\vec{X} = \left(I_n \otimes P_c^F\right)^{-1} \vec{I_n}.$$

Finally, we obtain the desired spectral decomposition of the inverse gramian of controllability over the spectrum of the dynamics matrix in canonical controllability form in the form (49), (50). The inverse gramian has the structure of the zero plaid of the Xiao matrix ([5, 6]). The columns of the matrix P_c^F will have the same structure. Insignificant elements of the zero plaid are determined by the formula

$$x_{ij} = 0, \quad \forall i, j = i + j \neq 2k, \quad k = \overline{1, n}.$$

On the other hand, the final inverse gramian is defined as a solution of the Sylvester differential equation

$$X = -\int_0^t e^{P_c^F t} dt, \quad \text{or} \quad \frac{dX}{dt} = e^{P_c^F t},$$

which can be solved in the complex domain

$$\mathcal{L}\left(\frac{dX}{dt}\right) = \left(Is - P_c^F\right)^{-1}.$$

The solution of this equation using spectral decompositions of the gramian resolvent can be obtained based on the application of the Faddeev–Leverrier algorithm.

Remark 3. An analysis of the formulas of Theorems 3 and 4 shows that they are much more complicated for calculating gramians than the formulas of the energy metrics of Theorem 2 J_1, J_2. However, the latter are invariant under any similarity transformations, including those that transform the MIMO LTI system (1) into a new system whose state equations correspond to one of three canonical forms: Jordan, modal, and controllability.

4. ENERGY METRICS OF STABILITY AND THE ROUTH-HURWITZ STABILITY CRITERION

A necessary and sufficient condition for the asymptotic stability of the basic system of the second type has the form of boundedness of the squares of the norms of the TF of the system [1]

$$\left\|N^{-1}(s)\right\|_2^2 < \infty, \quad \left\|N^{-1}(s)\right\|_\infty^2 < \infty, \tag{51}$$

or in accordance with Theorem 2 of the boundedness of the squares of the norms of the PF of the base system for the case of multiple roots

$$J_1 < \infty, \quad J_2 < \infty, \quad \max_{s_i} J_1 < \infty, \quad \max_{s_i} J_2 < \infty. \tag{52}$$

We define the energy metric of stability of the basic system in the form $J_1(s_1, \ldots, s_n)$, $J_2(s_1, \ldots, s_n)$. Spectral decompositions of the energy metrics $J_1(s_1, \ldots, s_n)$, $J_2(s_1, \ldots, s_n)$ by simple roots of the characteristic equation have the form (3.2), (3.3). The mathematical correctness of these metrics follows from the fact that the metrics are equal to the squares of the H_2 or H_∞ norms of the transfer function of the base system, which are invariants under similarity transformations

$$\left\|N^{-1}(s)\right\|_{inf}^2 = \max_{s_k} \left\|N^{-1}(s)\right\|_2^2, \forall k = 1, \ldots, n.$$

Spectral decompositions of the energy metric $J_2(s_1,\ldots,s_n)$ over multiple roots of the characteristic equation have the form

$$J_2(s_1,\ldots,s_n) =$$
$$= \sum_{k=1}^{n}\sum_{\nu=1}^{m_k}\sum_{\rho=1}^{n}\sum_{\mu=1}^{m_\rho} \frac{(m_k-\nu+m_\rho-\mu+1)!}{(m_k-\nu)!} \times \left[L_{k\nu}L_{\rho\mu}\frac{1}{(-s_k-s_\rho^*)^{m_k-\nu+m_\rho-\mu+1}}\right]_{Herm}.$$

Note that the metrics of the base systems are independent of the Faddeev matrices in the decomposition of the controllability gramians. These formulas imply that as the system approaches the stability boundary, caused by the roots of the characteristic equation approaching the imaginary axis, the energy stability metrics tend to infinity. Let us define an acceptable value for the energy stability metric as a sufficiently large positive number N_{perm}

$$J_1(s_1,\ldots,s_n) = N_{1perm}, J_2(s_1,\ldots,s_n) = N_{2perm}. \tag{53}$$

We will consider any basic system to be conditionally unstable if all roots of its characteristic equation are in the left half-plane, but the energy stability metric exceeds the established acceptable value.

Weakly stable and resonant roots of the characteristic equation that determine the system's proximity to the stability boundary. Analysis of the formulas in Theorem 2 shows that two groups of roots are weakly stable. The first group consists of roots close to the imaginary axis. The second group consists of complex root pairs that are close to each other in norm (a resonant root group). As the weakly stable eigenvalues of the first group approach each other, the corresponding terms become positive definite, and their values become arbitrarily large positive numbers. Similar limiting relations hold for a resonant root group. In the latter case, the degree of oscillatory instability is much higher, since the absolute values of the residues $L_{kv}, L_{\rho\mu}$ are inversely proportional to the difference between the roots s_k, s_ρ. The energy assessment of the weakly stable components of the expansions of the energy metrics of stability allows us to determine and evaluate the energy reserve of stability in the form

$$R_{1stab} = 20lg\frac{N_{perm}}{J_1(s_1,\ldots,s_n)}, \tag{54}$$

$$R_{2stab} = 20lg\frac{N_{perm}}{J_2(s_1,\ldots,s_n)}. \tag{55}$$

The introduced stability metrics (52), (53) do not contradict the classical stability criterion based on the proximity of the nearest root of the characteristic equation to the imaginary axis, but supplement it by taking into account the energy accumulated in the system's states. The same applies to the algebraic stability criterion of Routh-Hurwitz linear stationary systems [5, 6]. It can be expected that the combined application of the Routh–Hurwitz criterion and spectral decompositions of gramians in the form of energy stability metrics (54), (55) for stability analysis will create a synergistic effect, enhancing the advantages of each approach and mitigating their disadvantages. The hybrid criterion includes the formation of Routh tables and Xiao matrices, the calculation of the spectrum of the dynamics matrix, the calculation of the spectral decompositions of the energy metrics $J_1(s_1,\ldots,s_n), J_2(s_1,\ldots,s_n)$, the formation of the stability criterion (54), (55), and the analysis of stability for various scenarios of the system's operation. It is obvious that the obtained criteria are valid not only for the base system, but also for a continuous stationary linear system with many inputs and many outputs, the characteristic polynomial of which coincides with the characteristic polynomial of the base system. In [9], a method for calculating the gramian of the base system in canonical controllability form is proposed, based on the use of the Routh table and without the need to calculate the spectrum of the dynamics matrix. The diagonal elements of the Xiao matrices are related to the elements of the Routh table by the relations

$$y_{j\eta}=\begin{cases} 0, \text{ if } j+\eta=2k+1, \quad k=1,\dots,n; \\ y_n=\frac{1}{2Y_{n,1}}, \\ y_{n-l}=\frac{-\sum_{i=1}^{m-1}(-1)^i Y_{n-l,i+1}y_{n-l+i}}{Y_{n-l,1}}, \text{if } j+\eta=2k, \quad k=1,\dots,n, l=\overline{1,n-1}, \end{cases} \tag{56}$$

where $R_{j\eta}$ are "$j\eta$" elements of the Routh table, m is the number of elements in the j row of the Routh table. Let us assume that, along with the Routh table, all eigenvalues of the dynamics matrix s_k have been pre-calculated and they are different. According to (13), (14) the controllability gramian of the base system has the form of Xiao matrices, the diagonal elements of which have the form

$$y_i=\sum_{k=1}^{n}\frac{1}{\prod_{\lambda=1,\lambda\neq k}^{n}(s_k-s_\lambda)\prod_{\lambda=1,}^{n}(-s_k-s_\lambda)}(-1), \ i=\overline{1,n}. \tag{57}$$

Expressions (57) show that the energy stability metric depends on the distribution density of the real parts of the eigenvalues s_k, s_λ on the real axis: the closer s_λ to s_k, the greater the value of the energy metric [14, 16]. This is the group effect of nonlinear mode interaction, which intensifies as the absolute values of the real parts of the roots of the characteristic equation and their combinations tend to zero. The advantage of the (56) algorithm is that it eliminates the need to calculate the eigenvalues of the dynamics matrix. This is also its disadvantage, since the coefficients of the characteristic equation in most cases have no physical meaning, as they are not associated with specific physical devices. The advantage of the (57) algorithm is that the diagonal elements of the gramian matrices express the energy of the measured physical variable y_i of the underlying system. At the same time, the values of the metric are inversely proportional to the products of the absolute values of the differences of all eigenvalues that are close to each other in norm. In this sense, the metric (57) estimates the energy of the combined interaction in a group of weakly stable oscillatory modes. The Routh–Hurwitz criterion does not have this property. Furthermore, the variables y_i depend on the eigenvalues s_k, each of which is associated with a specific physical device and has its label. These eigenvalues can be used to calculate the squares of the H_2- and H_∞-norms of transfer functions in problems of optimal synthesis of robust control systems based on energy efficiency criteria. For closed control systems, formulas (51) are used, in which $N^{-1}(s)$ is the characteristic polynomial of the closed system. Energy metrics also allow us to evaluate the the degree of achievability, the volume of attraction ellipsoids, the average minimum energy, and the centrality indices of dynamic networks [7, 10–12 , 15]. Joint use of algorithms (56), (57) in the considered hybrid Routh–Hurwitz criterion provides a more in-depth analysis of the aperiodic and oscillatory stability of the system, taking into account the multiplicity of the system's eigenvalues.

5. CONCLUSION

The method of spectral and singular gramian decompositions is an extension of frequency methods for the analysis and synthesis of control systems. The main tools for constructing decompositions are Laplace transformation theorems for spectral and singular decompositions of solutions of Lyapunov and Sylvester differential equations, decomposition of the resolvent of the dynamics matrix and the gramian matrix into a series of Faddeev–Leverier series, and recursive algorithms for calculating Faddeev matrices. These decompositions for canonical forms of controllability and observability have an explicit physical interpretation in terms of energy measures of quadratic forms formed by the state variables of the basic system. The paper shows that the direct and inverse gramians of controllability can be computed using a single general effective recursive procedure. The article proposes a new method for calculating Laplace images of finite gramians for simple and multiple eigenvalues based on numerically stable recursive algorithms for decomposing the image of the matrix resolvent into a Faddeev–Leverrier series [22, 23]. The application of transformations of system equations into canonical forms provides the following advantages when calculating spectral decompositions of direct and inverse gramians:

1) Gramian matrices are calculated as solutions of Lyapunov differential equations in the frequency and time domains, which allows simultaneous calculation of solutions to Lyapunov algebraic equations, both finite and infinite gramians, both direct and inverse gramians for continuous dynamic systems specified by state equations in canonical forms: Jordan, modal, controllability, observability.

2) Gramians in canonical forms of controllability and observability in the form of Xiao matrices are invariant under similarity transformations. To obtain them, it is sufficient to calculate only n diagonal elements of the gramian matrices instead of n^2 elements.

3) The energy metrics of stability of the basic systems J_1 and J_2 allow us to calculate the degree of stability of the system not only for different roots, but also for multiple roots of the characteristic equation. The presence of multiple roots significantly increases the threat of loss of stability.

4) Spectral decompositions of the gramians of controllability of a stable basic system in Cauchy form depend on the real parts of the simple and multiple eigenvalues of the dynamics matrix, on their distribution density on the real axis [25].

5) Recursive algorithms for spectral decomposition of gramians contain only matrix addition and multiplication operations and do not contain inversion and exponentiation operations. They contain only addition, multiplication, and exponentiation operations for scalar functions, and do not contain division operations for functions equal to zero. The subtractions of the transfer functions of the base systems, which are used in the calculation of energy criteria, can also be calculated using recursive algorithms. All this is a consequence of the decomposition of the resolvent matrices of the dynamics of the initial stable and dual unstable systems into the Faddeev–Leverrier series and the application of classical frequency methods of analysis and synthesis of dynamic systems, which makes this approach a new effective tool for calculating energy criteria of stability.

A certain disadvantage of the new methods and algorithms developed in this work is the problem of their computational implementation for high-dimensional systems. However, it is precisely the use of recursive versions of the developed algorithms for calculating gramians and energy metrics that makes their application effective for high-dimensional dynamic systems [26]. The article presents research results that may be useful in the design of digital twins, methods for optimizing energy costs in control systems, analysis of energy metrics of stability and reachability, condition monitoring based on the analysis of energy balance anomalies, the development of energy metrics for systems defined by state equations for dynamic networks in the form of graphs, and in the vibration analysis of technical objects [13,28].

FUNDING

This work was supported by the Russian Foundation for Basic Research, project no. 25-29-20158.

REFERENCES

1. Antoulas, A.C., *Approximation of Large-Scale Dynamical Systems*, SIAM: Philadephia, 2005.

2. Benner, P. and Damm, T., Lyapunov equations, Energy Functionals and Model Order Reduction of Bilinear and Stochastic Systems, *SIAM J. Control Optim.*, 2011, vol. 49, pp. 686–711.

3. Behr, M., Benner, P., and Damm, T., Solution formulas for differential Sylvester and Lyapunov equations, *Calcolo*, 2019, vol. 56, art. no. 51. https://doi.org/10.1007/s10092-019-0348-x

4. Polyak, B.T., Khlebnikov, M.V., and Rapoport, L.B., *Matematicheskaya teoriya avtomaticheskogo upravleniya* (Mathematical theory of automatic control), Moscow: LENAND, 2019. 504 pp.

5. Hauksdottir, A., Sigurdsson, S., Adalgeirsson, S., Porgilsson, H., and Herjolfsson, G., Closed form solutions of the Sylvester and the Lyapunov equations — closed form gramians, *2008 American Control Conference*, Seattle, WA, USA, 2008, pp. 2585–2590. https://doi.org/10.1109/ACC.2008.4586881

6. Hauksdottir, A. and Sigurdsso, S., The continuous closed form controllability gramian and its inverse, *2009 American Control Conference Hyatt Regency Riverfront*, St. Louis, MO, USA, 2009, pp. 5345–5350. https://doi.org/10.1109/ACC.2009.5160123

7. Liu, Y., Slotine, J., and Barabasi, A., Controllability of complex networks, *Nature*, 2011, vol. 473, pp. 167–173. https://doi.org/10.1038/nature10011

8. Xiao, C.S., Feng, Z.M., and Shan, X.M., On the Solution of the Continuous-Time Lyapunov Matrix Equation in Two Canonical Forms, *IEE Proc.*, 1992, vol. 139, no. 3, pp. 286–290. https://doi.org/10.1049/ip-d.1992.0038

9. Sreeram, V. and Agathoklis, P., Solution of Lyapunov equation with system matrix in companion form, *IEE Proc. D. Control. Theory Appl.*, 1991, vol. 138, no. 6, pp. 529–534. https://doi.org/10.1049/ip-d.1991.0074

10. Pasqualetti, F., Zampieri, S., and Bullo, F., Controllability metrics, limitations and algorithms for complex networks, *IEEE Transactions on Control of Network Systems*, 2014, vol. 1, no. 1, pp. 40–52. https://doi.org/10.1109/ACC.2014.6858621

11. Lindmark, G. and Altafini, C., Minimum energy control for complex networks, *Scientific Reports*, 2018, vol. 8, art. no. 3188, pp. 1–14. https://doi.org/10.1038/s41598-018-21398-7

12. Biryukov, D.S., Slita, O.V., and Ushakov, A.V., Estimation of control costs in the problem of ensuring the desired structure of modes and their robustness, *Izv. vuzov. Priborostroenie*, 2009, vol. 52. no. 11. pp. 38–43.

13. Yadykin, I.B., Spectral decompositions of inverse matrices of gramians and energy metrics of continuous dynamic systems, *Autom. Remote Control*, 2024, no. 10, pp. 80–107.

14. Yadykin, I.B. and Galyaev, A.A., On the methods for calculation of gramians and their use in analysis of linear dynamic systems, *Autom. Remote Control*, 2013, vol. 74, pp. 207–224.

15. Zhao, S. and Pasqualetti, F., Networks with diagonal controllability gramian: Analysis, graphical conditions, and design algorithms, *Automatica*, 2019, vol. 102, pp. 10–18.

16. Iskakov, A., Kutyakov, E., and Kataev, D., Locating the source of forced oscillations on the basis of Lyapunov modal analysis, *IFAC-PapersOnLine. Rabat: Elsevier*, 2024, vol. 58, no. 13, pp. 685–690.

17. Krasnova, S.A. and Utkin, V.A., *Kaskadnyi sintez nablyudatelei sostoyaniya dinamicheskikh sistem* (Cascade synthesis of state observers for dynamic systems), Moscow: Nauka, 2006, 272 pp.

18. Bakhtadze, N. and Yadykin, I., Discrete Predictive Models for Stability Analysis of Power Supply Systems, *Mathematics*, 2020, vol. 8, no. 11, pp. 1943. https://doi.org/10.3390/math8111943

19. Yadykin, I.B. and Galyaev, I.A., Spectral decompositions of gramians and energy metrics of continuous unstable control systems, *Autom. Remote Control*, 2023, no. 10, pp. 132–149.

20. Yadykin, I.B. and Galyaev, I.A., Structural spectral methods for solving continuous Lyapunov equations, *Autom. Remote Control*, 2023, no. 12, pp. 18–37.

21. Fiedler, M., Notes on Hilbert and Cauchy matrices, *Linear algebra and its Applications*, 2010, vol. 432, pp. 351–356.

22. Faddeev, D.K. and Faddeeva, V.N., *Vychislitel'nye metody lineinoi algebry* (Computational methods of linear algebra), Uchebnik-M: Lan', 2009, 726 pp.

23. Hanson, B. and Peeters, R., A Faddeev Sequence Method for solving Lyapunov and Sylvester Equations, *Linear Algebra and its Applications*, 1996, vol. 241–243, pp. 401–430.

24. Gardner, M.F. and Barns, J.L., *Transients in linear systems studied by the Laplace transformation*, vol. 1. Lumped-constant systems. New York, London. Wiley, Chapman and Hall, 1942.

25. Furtat, I.B., Density Systems. Analysis and Control, *Autom. Remote Control*, 2023, vol. 84, no. 11, pp. 1175–1190.

26. Bekhiti, B., Fragulis, G., Maraslidis, G., et al., A Novel Recursive Algorithm for Inverting Matrix Polynomials via a Generalized Leverrier-Faddeev Scheme: Application to FEM Modeling of Wing Vibrations in a 4th Generation Fighter Aircraft, *Mathematics*, 2025, vol. 13, art. no. 2101. https://doi.org/10.3390/math13132101

27. Mironovskii, L.A. and Solov'eva, T.N., Analysis and synthesis of modally balanced systems, *Autom. Remote Control*, 2013, vol. 74, no. 4, pp. 588–603.

28. Yarar, E., Erturk, A.T., Acikgez, C., and Karabay, S., Experimental and numerical vibration analysis of surface mechanical attrition treatment, *Journal of Vibration and Control*, 2022, vol. 29, no. 23–24, pp. 5737–5749.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 336–342.

SPECIAL ISSUE

Cauchy Problem for the Buckley–Leverett Equation of Two-Phase Filtration with Variable Porosity

M. M. Volnykh[*,a] **and A. G. Kushner**[*,**,b]

[*] *Lomonosov Moscow State University, Moscow, Russia*
[**] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a] *volnykh.mm17@physics.msu.ru,* [b] *kushner@physics.msu.ru*

Received July 8, 2025
Revised November 10, 2025
Accepted November 20, 2025

Abstract—The problem of constructing a multivalued solution for the Buckley–Leverett equation describing the process of two-phase filtration in a medium with a variable porosity coefficient is considered. Such an equation is used to calculate the evolution of the phase separation surface "oil–water" during the development of oil fields. It is shown how to construct the caustic of multivalued solutions for given initial conditions.

Keywords: filtration, porous media, multi-valued solutions, caustics, jets, contact structure, contact vector fields

DOI: 10.7868/S1608303226040044

1. INTRODUCTION

During oil field development, reservoir energy is maintained by injecting a displacing agent (water or solutions with active additives) through special wells. The movement of the oil displaced by this agent, and the agent itself, depends on the pressure difference between the injection and production zones. In homogeneous reservoirs, fluid movement between alternating rows of production and injection wells is often considered one-dimensional. This assumption is justified, for example, when injection and extraction are carried out through parallel wells [1].

One-dimensional fluid flow is particularly evident when using horizontal wells with parallel boreholes. The displacing agent is injected through injection wells and directed into horizontally oriented fractures formed by hydraulic fracturing. Under pressure, fluid flows from these fractures toward production wells, from where oil is extracted to the surface through vertical channels.

This production approach is particularly effective for fields with hard-to-recover reserves, where successful oil displacement requires physical and chemical methods. These methods include the use of active component solutions (e.g., polymers, surfactants, carbon dioxide, and others), thermal methods (such as hot water or steam injection), as well as wave stimulation (hydrodynamic, electromagnetic, ultrasonic), or combinations thereof.

2. MODEL DESCRIPTION

The Buckley–Leverett equation [2] describes one-dimensional two-phase filtration in a porous medium. It is based on Darcy's law, according to which the filtration rate is proportional to the difference in the partial pressures of the phases. In the classical Buckley–Leverett model, this difference is assumed to be constant. The permeability of the pore space, which is assumed to be non-deformable, is also assumed to be constant.

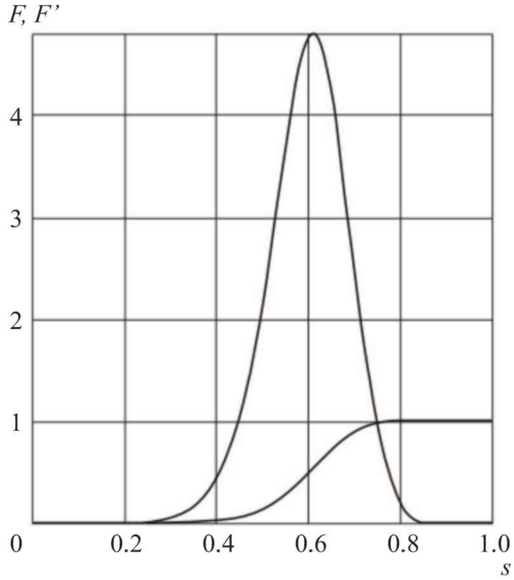

Fig. 1. Graphs of functions F and its derivative.

Here we consider the Buckley–Leverett equation in the following form:

$$m\frac{\partial s}{\partial t} + U\frac{\partial}{\partial x}(F(s)) = 0. \tag{1}$$

Here t is time, x is the spatial coordinate (the Ox axis is directed in the direction of fluid flow), $s = s(t, x)$ is water saturation (the volume fraction of pores occupied by water), m is the porosity coefficient, and $F(s)$ is the Buckley–Leverett function, indicating the volume fraction of water in the total flow U, which is assumed to be constant. An approximate graph of the Buckley-Leverett function and its derivative is shown in Fig. 1. The prime denotes the derivative of function F with respect to variable s.

It is usually assumed that the porous medium is homogeneous, i.e., in equation (1), the porosity coefficient m is constant. Here, we abandon this assumption and assume that function m depends on the spatial variable, i.e., $m = m(x)$. We assume that the function $m(x)$ is differentiable.

Let's consider the Cauchy problem for equation (1):

$$s(0, x) = S(x), \tag{2}$$

where $S(x)$ is a given function, which we assume to be differentiable.

As is well known, the classical solution to this problem collapses upon reaching a certain critical point in time, regardless of the initial water saturation distribution. The solution develops a discontinuity, and the phase boundary moves like a shock wave.

Therefore, instead of classical solutions, we will consider multivalued solutions [3].

This approach circumvents the difficulties associated with the collapse of solutions to equations and makes it possible to describe the evolution of shock waves.

3. CHARACTERISTICS OF THE BUCKLEY–LEVERETT EQUATION

We use a differential-geometric approach to equation (1). This method, as applied to the problem under consideration, is described in [4]. With this approach, the equation is considered as a hypersurface in the space of 1-jets of functions of two variables.

The space of 1-jets $J^1 = J^1(\mathbb{R}^2)$ of functions of two independent variables t and x is a five-dimensional space with canonical coordinates $t, x, u_{00}, u_{10}, u_{01}$. Let us indicate the coordinates of the 1-jet $\theta = j^1_a(s)$ of the function $s = s(t, x)$ at the point $a = (a_1, a_2) \in \mathbb{R}^2$:

$$t(\theta) = a_1,\ \ x(\theta) = a_2,\ \ u_{00}(\theta) = s(a),\ \ u_{10}(\theta) = \frac{\partial s}{\partial t}(a),\ \ u_{01}(\theta) = \frac{\partial s}{\partial x}(a).$$

The space J^1 is equipped with the four-dimensional contact distribution

$$\mathcal{C} : J^1 \ni \theta \mapsto \mathcal{C} = \ker \varkappa_\theta \subset T_\theta J^1.$$

It is generated by the differential 1-form (the Cartan form)

$$\varkappa = du_{00} - u_{10} dt - u_{01} dx.$$

Equation (1) corresponds to the function $f = m(x)u_{10} + UF'(u_{00})u_{01}$ and a hypersurface in the space of 1-jets is $\mathcal{E} = \{f = 0\}$. For brevity, we introduce the following notation:

$$H(u_{00}) = UF'(u_{00}). \tag{3}$$

Then

$$f = m(x)u_{10} + H(u_{00})u_{01}. \tag{4}$$

The multivalued solution of equation (1) is a two-dimensional surface $L \subset \mathcal{E}$, which is an integral manifold of the contact distribution $\mathcal{C}$, i.e., the restriction of the Cartan form to it is zero: $\varkappa|_L = 0$.

Let us introduce the vector field

$$X = -m(x)\frac{\partial}{\partial t} - H(u_{00})\frac{\partial}{\partial x} \tag{5}$$

on the space of 0-jets J^0. The value of the Cartan form on this field coincides with the function f, and the extension of the field X to the space of 1-jets is a contact vector field with generating function f:

$$X_f = -m(x)\frac{\partial}{\partial t} - H(u_{00})\frac{\partial}{\partial x} + H'(u_{00})u_{01}u_{10}\frac{\partial}{\partial u_{10}} + \left(m'(x)u_{10} + H'(u_{00})u_{01}^2\right)\frac{\partial}{\partial u_{01}}.$$

This field is called the *characteristic* vector field for equation (1), and its trajectories are called the *characteristics.*

This vector field is tangent to the hypersurface $\mathcal{E}$ and, moreover, is tangent to every multivalued solution of equation (1) (see [5]). It can be used to construct a multivalued solution to the generalized Cauchy problem. Let us consider this in more detail.

4. CAUCHY PROBLEM

Initial condition (2) generates a Cauchy curve in the space of 1-jets:

$$\mathcal{K} = \{t = 0, u_{00} = S(x), u_{01} = S'(x), m(x)u_{10} + H(u_{00})S'(x) = 0\} \subset \mathcal{E}.$$

Since $\varkappa|_{\mathcal{K}} = 0$, this curve is integral for the contact distribution.

Since the coefficient $m(x) \neq 0$, at each point $a \in \mathcal{K}$, the tangent vector of the field X_f does not lie in the tangent space to this curve: $X_f|_a \notin T_a\mathcal{K}$. By shifting the Cauchy curve along the trajectories of the vector field X_f, we obtain a two-dimensional surface

$$L = \Phi^{(1)}_\tau(\mathcal{K}).$$

Here τ is the shift parameter along the trajectories, and $\Phi_\tau^{(1)}$ is the shift transformation (the flow of the vector field X_f) from $\tau = 0$ to τ. This flow is an extension of the flow Φ_τ of the vector field (5) into the space of 1-jets.

Since the flow $\Phi_\tau^{(1)}$ preserves the contact distribution $\mathcal{C}$, the surface L is a multivalued solution.

Thus, the multivalued solution L is a disjunctive union of the characteristics passing through the curve $\mathcal{K}$. This multivalued solution will be called a *multivalued solution of the Cauchy problem.*

The projections of the characteristics onto the plane of independent variables t, x are integral curves of the following ordinary differential equation:

$$\frac{dx}{dt} = \frac{H(u_{00})}{m(x)}. \tag{6}$$

We also call these curves characteristic curves.

Since $X_f(u_{00}) = 0$, i.e., each characteristic of equation (1) lies entirely in the fiber $u_{00} = \text{const}$. Of course, each characteristic has its own constant. Therefore, the function $H(u_{00})$ is constant on the characteristic.

The integral curves of equation (6) have the form

$$t = \frac{1}{H(u_{00})} \int m(x)dx + C. \tag{7}$$

In this formula, the parameter u_{00} is a constant whose value is determined by the initial point of the characteristic.

Unlike the case of constant porosity, when the projections of the characteristics onto the plane of independent variables are straight lines, here a family of curves is obtained.

5. THE CASE OF LINEAR POROSITY

Let's consider the case of linear porosity:

$$m(x) = \alpha x + \beta.$$

In this case, the condition $m(x) > 0$ must be satisfied. As function (3), we choose the function

$$H(u) = a \exp\left(-\frac{(u-b)^2}{2c^2}\right) + d,$$

which well approximates the derivative F' of the Buckley–Leverett function. Here $\alpha, \beta, a, b, c, d$ are some real numbers, and β, a, c are not equal to zero. As the initial function (2), we choose the function

$$S(x) = \frac{1}{1+x^2},$$

which well approximates the water saturation function near the injection well located at $x = 0$.

For brevity, we denote the coordinate function u_{00} by u, omitting the subscripts. Then, vector field (5) takes the form

$$X = -(\alpha x + \beta)\frac{\partial}{\partial t} - \left(a \exp\Big(-\frac{(u-b)^2}{2c^2}\Big) + d\right)\frac{\partial}{\partial x}. \tag{8}$$

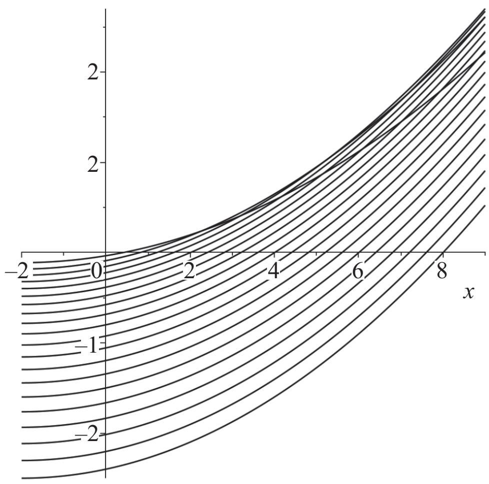

Fig. 2. Characteristics (9) for parameter values (10).

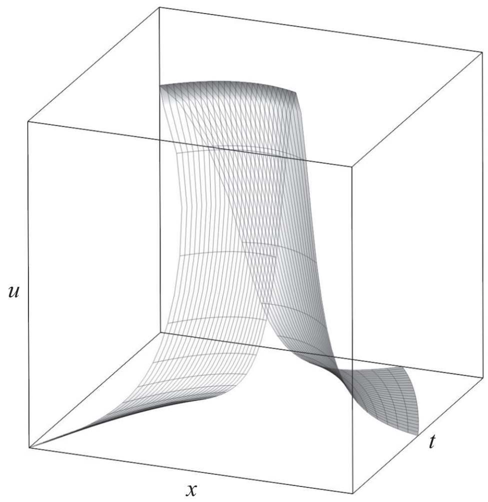

Fig. 3. The caustic Σ on the multivalued solution.

Its flow is found explicitly:

$$\Phi_\tau : \begin{cases} t \mapsto \alpha \left(a \exp\Big(-\dfrac{(u-b)^2}{2c^2} \Big) + d \right) \dfrac{\tau^2}{2} - (\alpha x + \beta)\tau + t, \\ x \mapsto -\left(a \exp\Big(-\dfrac{(u-b)^2}{2c^2} \Big) + d \right) \tau + x, \\ u \mapsto u. \end{cases}$$

Passing through the point (t_0, x_0) characteristic (7) has the form of a parabola (see Fig. 2).

$$\alpha(x^2 - x_0^2) + 2\beta(x - x_0) - 2(t - t_0) \left(a \exp\left(-\frac{(bx_0^2 + b - 1)^2}{2(x_0^2+1)^2 c^2} \right) + d \right) = 0. \tag{9}$$

Starting at some point in time $t = T$, the parabolas (9) intersect, and for $t > T$, the solution to equation (1) becomes multivalued (see Fig. 2).

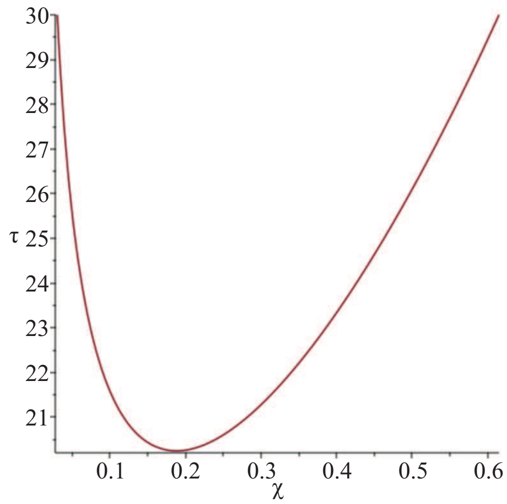

Fig. 4. The graph of multivalued solution.

At $t = T$, a so-called "gradient catastrophe" occurs: the partial derivative with respect to variable x of the classical single-valued solution becomes infinite. A similar phenomenon is observed for the case of constant porosity, however, the characteristics there are not parabolas, but straight lines.

This paper does not address the issue of overcoming this difficulty. The standard method is to construct discontinuous solutions from multivalued solutions using conservation laws and the Hugoniot–Rankine conditions [6].

The envelope of the family of characteristics (9) is the projection of the caustic $\Sigma \subset L$ onto the plane of independent variables. At the caustic points, this projection

$$\pi_{1,0} : L \to \mathbb{R}^2(t, x)$$

has a singularity: the tangent planes to the multivalued solution project onto curves.

Figures 2 and 3 show the characteristics and the caustic Σ for the following parameter values:

$$a = 0.8,\ b = 2,\ c = 0.5,\ d = 0.2,\ \alpha = -0.01,\ \beta = 0.02. \tag{10}$$

The projection of the multivalued solution onto the space $\mathbb{R}^3$ with coordinates t, x, u (see Fig. 4) in parametric form is

$$L : \begin{cases} t = \dfrac{\tau}{2}\left(\tau\alpha a \exp\left(-\dfrac{(b\chi^2 + b - 1)^2}{2(\chi^2+1)^2c^2)}\right) + (d\tau - 2\chi)\alpha - 2\beta\right), \\ x = -a\tau\exp\left(-\dfrac{(b\chi^2 + b - 1)^2}{2(\chi^2+1)^2c^2}\right) - d\tau + \chi, \\ u = \dfrac{1}{1+\chi^2}. \end{cases}$$

6. CONCLUSION

It should be noted that other two-phase flow models exist that take capillary forces into account and lead to second-order differential equations. One such example is the Rapoport–Leas equation [7]. However, their study raises higher-order difficulties, overcoming which requires other methods, such as those of finite-dimensional dynamics [8, 9].

Other examples of the application of contact geometry methods to problems in continuum mechanics and thermodynamics are given in [10]. In particular, for certain classes of Buckley–Leverett functions, exact solutions were found for the Barenblatt equations, which describe two-phase flow in porous media in the presence of surface-active reagents dissolved in water and changing the viscosity of the fluids [11].

In [12] the Buckley–Leverett model was constructed and investigated, in which the porosity coefficient in equation (1) is a random variable.

FUNDING

This work was particularly supported by the Russian Science Foundation, projects no. 25-21-00152.

REFERENCES

1. Entov, V.M. and Zazovsky, A.F., *Gidrodinamika protsessov povysheniya nefteotdachi* (Hydrodynamics of enhanced oil recovery processes), Moscow: Nedra, 1989 (in Russian).
2. Buckley, S.E. and Leverett, M.C., Mechanism of fluid displacement in sands, *Trans. AIME*, 1942, no. 146, pp. 107–116.
3. Krasil'shchik, I.S., Lychagin, V.V., and Vinogradov, A.M., *Geometry of Jet Spaces and Nonlinear Partial Differential Equations*, Philadelphia: Gordon and Breach Science Publishers, 1986.
4. Akhmetzyanov, A.V., Kushner, A.G., and Lychagin, V.V., Geometric theory of special modes in the distributed-parameter control systems, *Autom. Remote Control*, 2013, vol. 74, pp. 1786–1801. https://doi.org/10.1134/S0005117913110027
5. Bocharov, A.V. et al., *Symmetries and Conservation Laws for Differential Equations of Mathematical Physics*, Providence, R.I., American Mathematical Society, 1999.
6. Akhmetzyanov, A.V., Kushner, A.G., and Lychagin, V.V., *Matematicheskie modeli upravleniya razrabotkoi neftyanykh mestorozhdenii* (Mathematical Models of Oil Field Development Management), Moscow: IPU RAS, 2017.
7. Rapoport, L.A. and Leas, W.J., Properties of linear waterflood, *Trans. AIME*, 1953, vol. 198, pp. 139–148.
8. Akhmetzyanov, A.V., Kushner, A.G., and Lychagin, V.V., Attractors in models of porous media flow, *Dokl. Math.*, 2017, vol. 95, pp. 72–75.
9. Kushner, E.N., Classification of generalized Rapoport–Leas equations, *Proc. 11th Int. Conf. Manag. Large-Scale Syst. Dev. (MLSD)*, Moscow, Russia, 2018, pp. 1–3. https://doi.org/10.1109/MLSD.2018.8551940
10. Kushner, A., Mukhina, S., and Gavrilov, V., Contact structures in continuous media theory, *Proc. 17th Int. Conf. Manag. Large-Scale Syst. Dev. (MLSD)*, Moscow, Russia, 2024, pp. 1–4. https://doi.org/10.1109/MLSD61779.2024.10739520
11. Mukhina, S., Symplectic geometry of the oil displacement Barenblatt equations, *J. Geom. Phys.*, 2024, vol. 204, pp. 105277. https://doi.org/10.1016/j.geomphys.2024.105277
12. Isaeva, A.V. and Serdobolskaya, M.L., Solving the Buckley–Leverett equation with a random coefficient of porosity, *Numer. Methods Program.*, 2012, vol. 13, no. 4, pp. 517–524.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 343–358.

SPECIAL ISSUE

Robust Approach to Dynamic Compensation of Disturbances

S. A. Kochetkov[*,a] **and V. A. Utkin**[*,b]

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

Received July 8, 2025
Revised November 1, 2025
Accepted November 20, 2025

Abstract—In this paper, the classical problem statement of external disturbances compensation in linear systems is extended to the case of nonlinear systems in a robust formulation using the dynamic compensation method. The robust properties of closed-loop systems are provided through using of deep feedback loops and discontinuous control inputs. For the output variable regulation, it is proposed to consider part of the controlled plant as a generator of model uncertainties. In this case, unlike a disturbance generator, it is possible to correct the dynamics of the model generator and convergence rate. As an example of the application of the proposed approach, the problem of altitude control of an unmanned aerial vehicle with a spring-loaded suspension is solved.

Keywords: dynamic compensation, sliding modes, external disturbances, decomposition

DOI: 10.7868/S1608303226040054

1. INTRODUCTION

In the control theory, the method of dynamic disturbance compensation is widely used in the following problem formulation. If it is possible to describe the behavior of external disturbances using a known mathematical model, then invariance can be ensured using the dynamic compensation method [1–3], or using the theory of asymptotic observers [4–7]. By using the first approach, asymptotic observers calculate estimates of unknown disturbances, which are then used for feedback generation. By utilizing the second one, the dynamic feedback is synthesized with the help of dynamic compensators in accordance with the disturbance model, thereby compensating for the influence of disturbances on output variables through control channels. When using both approaches, the synthesis of combined control law is based on an extended model of the original plant by introducing a dynamic subsystem corresponding to the disturbance subspace.

Some of the disadvantages of the dynamic compensation method are, firstly, the excessive expansion of the dynamic order of the control plant model due to the inclusion of a reference generator and, secondly, the sensitivity of the closed-loop system to parametric uncertainties in both the control plant model itself and the disturbance generator. The latter circumstance makes the use of the dynamic compensation method ineffective in practice.

To solve the problem of dynamic disturbance compensation for high-dimensional mathematical models of control plants, decomposition methods based on back-stepping control, decentralized control methods, and the block approach have proven effective. These methods split the solution to the original problem into independent stages, each of which synthesizes local feedback [8] in a reduced-dimensional subsystem. On the other hand, methods for systems with discontinuous sliding mode controls [9] and deep feedback [9, 10], which have already become classic in the synthesis of robust and invariant systems, make it possible to ensure the robustness of closed-loop systems to parametric uncertainties.

Section 2 presents a well-known result of the dynamic disturbance compensation method applied to linear systems using the motion separation method, which allows the synthesis problem to be broken down into independent, lower-dimensional subproblems for synthesizing the control plant itself and independently selecting the disturbance compensator. Within the framework of the approaches described above, Section 3 extends the problem of dynamic disturbance compensation to nonlinear systems. Section 4 considers the problem of dynamic compensation in the case of a parametrically uncertain disturbance generator.

Section 5 is devoted the problem of horizontal attitude control for an unmanned aerial vehicle (UAV) carrying a payload suspended by a tether. A unique feature of the problem formulation in this example is that the disturbance generator is a model of the behavior of a load suspended on a weightless string, which, generally speaking, is an integral part of the control object model. Unlike a classical disturbance generator, the model of a load suspended on a weightless string considered in this article is a non-autonomous, yet controllable, subsystem. Feedback allows for influencing the dynamics of the load's behavior and compensating for external disturbances. The authors believe the proposed approach has significant potential in control theory.

2. PRELIMINARY DISCUSSION. FORMULATION OF THE PROBLEM

For the preliminary discussion let us consider the linear system

$$\dot{x} = Ax + Bu + Q\eta, \quad y = Dx, \tag{1}$$

combined with the model of the external perturbations η

$$\dot{\omega} = W\omega, \;\; \eta = H\omega, \;\; \omega \in R^r, \;\; \dim W = (r \times r), \;\; \dim H = (l \times r), \tag{2}$$

where $x \in R^n$ is the system state space vector, $u \in R^m$ is the vector of control inputs, $y \in R^p$ is the output vector, $\eta \in R^l$ is the disturbances vector, $A,\ B,\ Q,\ D,\ W,\ H$ are constant matrices of the appropriate dimensions.

Under condition, that the disturbances vector is not measured, the problem of its attenuation with respect to the output vector is stated in the paper.

Also, it is assumed, that for the plant (1)–(2) the matching condition is satisfied, under which the channels of action of the disturbances and control vectors coincide:

$$\mathrm{Im}Q \subset \mathrm{Im}B \Rightarrow \exists \Lambda : Q = B\Lambda, \;\; \dim \Lambda = (m \times l). \tag{3}$$

In this case, the equations (1) can be rewritten as

$$\dot{x} = Ax + B(u + \Lambda\eta), \;\; y = Dx. \tag{4}$$

It should be noted, that for the object (4), the invariance conditions for the whole state space vector are satisfied (full disturbances suppression). In [11], the decomposition approach is proposed, which provides the solution of the stated problem of the invariance in the plant (2), (4). According to this methodology, the synthesis task for the system with $n + r$ dimension to be broken down into independently solvable subproblems of $n - m$, m, r dimensions.

Using a non-singular change of variables, the system (2)–(4) is reduced to the regular controllability form [12]:

$$\begin{aligned} &\dot{x}_1 = A_{11}x_1 + A_{12}x_2, \\ &\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + B_2u + Q_2\omega, \\ &\dot{\omega} = W\omega, \end{aligned} \tag{5}$$

where $x_1 \in R^{n-m}$, $x_2 \in R^m$, $\mathrm{rank}B_2 = m$.

The dynamic compensator is formed like the following

$$\dot{z} = Wz - v, \ \ v, z \in R^r. \tag{6}$$

The relation for the control input is

$$u = -B_2^{-1} Q_2 z + \overline{u}, \tag{7}$$

After combining (5)–(7) one can get the system for the variables $x_1, \ x_2, \ \varepsilon = \omega - z$:

$$\begin{aligned} \dot{x}_1 &= A_{11} x_1 + A_{12} x_2, \\ \dot{x}_2 &= A_{21} x_1 + A_{22} x_2 + B_2 \overline{u} + Q_2 \varepsilon, \\ \dot{\varepsilon} &= W\varepsilon + v. \end{aligned} \tag{8}$$

After the coordinate transformation $\overline{x}_2 = x_2 + F_1 x_1$ and $\overline{\varepsilon} = \varepsilon + L x_2$ the last plant (8) is written in the form

$$\begin{aligned} \dot{x}_1 &= (A_{11} - A_{12} F) x_1 + \overline{x}_2, \\ \dot{\overline{x}}_2 &= \overline{A}_{21} x_1 + \overline{A}_{22} \overline{x}_2 + B_2 \overline{u} + Q_2 \overline{\varepsilon}, \\ \dot{\overline{\varepsilon}} &= (W + LQ_2)\overline{\varepsilon} + G_{11} x_1 + G_{12} \overline{x}_2 + LB_2 \overline{u} + v. \end{aligned}$$

The control inputs of the dynamic compensator and the system $v, \ \overline{u}$ are chosen according to the formulas

$$\begin{aligned} v &= -G_{11} x_1 - G_{12} \overline{x}_2 - LB_2 \overline{u}, \\ \overline{u} &= -B_2^{-1} (\overline{A}_{21} x_1 + (\overline{A}_{22} - Q_2 L) \overline{x}_2 - F_2 \overline{x}_2), \end{aligned} \tag{9}$$

The equations of the closed-loop system are

$$\begin{aligned} \dot{x}_1 &= (A_{11} x_1 - A_{12} F_1) x_1 + \overline{x}_2, \\ \dot{\overline{x}}_2 &= -F_2 \overline{x}_2 + Q_2 \overline{\varepsilon}, \\ \dot{\overline{\varepsilon}} &= (W + LQ_2)\overline{\varepsilon}. \end{aligned} \tag{10}$$

Assuming, that the original system is controllable and observable, the desired spectrum can be assigned by choosing the matrices $L, \ F_1, \ F_2$ in the system (10). The solution of this problem can be broken down into independently solvable subproblems of smaller dimensions $n - m$, m, r, respectively [11].

3. METHOD OF DYNAMIC COMPENSATION IN NONLINEAR SYSTEMS

In nonlinear case the equations of the control object and disturbances are

$$\begin{gathered} \dot{x} = f(x) + B(x)u + Q(x)\eta, \ \ \dot{\omega} = W\omega, \ \ \eta = H\omega, \\ x \in R^n, \ \ \ u \in R^m, \ \ \ \eta \in R^r. \end{gathered} \tag{11}$$

For the system (11) additionally requirement is introduced, that is

$$f(0) = 0. \tag{12}$$

The matching condition (3) is transformed to the representation

$$\mathrm{Im} Q(x) \subset \mathrm{Im} B(x), \forall x \Rightarrow \exists \Lambda(x) : Q(x) = B(x)\Lambda(x),$$

According to this matching condition, system (11) can be represented in regular controllability form [12] with the help of some nonsingular nonlinear coordinate transformation

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2), \\ \dot{x}_2 &= f_2(x_1, x_2) + B_2(x_1, x_2)[u + \Lambda_2(x_1, x_2)\omega], \end{aligned} \tag{13}$$

where $x_1 \in R^{n-m}, \ x_2 \in R^m, \ u \in R^m, \ \det B_2 \neq 0.$

Following the block approach [8], a step-by-step procedure is proposed below to solve the system stabilization problem (13).

Step 1. Introducing new variable $\overline{x}_2 = x_2 - \varphi(x_1)$ one can write from (13)

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \overline{x}_2 + \varphi(x_1)), \\ \dot{\overline{x}}_2 &= f_2(x_1, \overline{x}_2 + \varphi(x_1)) + B_2(x_1, \overline{x}_2 + \varphi(x_1))[u + \Lambda_2(x_1, \overline{x}_2 + \varphi(x_1))\omega] - \frac{\partial \varphi(x_1)}{\partial x_1}\dot{x}_1, \end{aligned} \tag{14}$$

Further, the notations for the functions $f_2(x_1, \overline{x}_2 + \varphi(x_1)),\ B_2(x_1, \overline{x}_2 + \varphi(x_1)),\ \Lambda_2(x_1, \overline{x}_2 + \varphi(x_1))$ are simplified to $f_2(\cdot),\ B_2(\cdot),\ \Lambda_2(\cdot)$.

In the first subsystem, $\varphi(x_1)$ is chosen to ensure its stabilization $x_1 \to 0$ under the assumption $\overline{x}_2 = 0$.

Step 2. The stability of the second subsystem (13) can be provided by choosing the control input u in the form

$$u = B_2^{-1}(\cdot)\left[F_2\overline{x}_2 - f_2(\cdot) + \frac{\partial \varphi(x_1)}{\partial x_1}\dot{x}_1\right] + \Lambda_2(\cdot)\left[-z_2 + L_2\overline{x}_2\right], \tag{15}$$

where

$$\dot{z}_2 = Wz_2 + v_2,\ v_2, z_2 \in R^r. \tag{16}$$

The second subsystem (14), closed by feedback (15), taking into account (16), with the variables $\overline{x}_2$, $\varepsilon_2 = \omega - z_2 + L_2\overline{x}_2$ takes the representation

$$\begin{aligned} \dot{\overline{x}}_2 &= F_2\overline{x}_2 + B_2(\cdot)\Lambda_2(\cdot)\varepsilon_2, \\ \dot{\varepsilon}_2 &= W\varepsilon_2 + L_2\left(F_2\overline{x}_2 + B_2(\cdot)\Lambda_2(\cdot)\varepsilon_2\right) + v_2. \end{aligned} \tag{17}$$

Step 3. By using the feedback $v_2 = -L_2F_2\overline{x}_2$, one can rewrite (17)

$$\begin{aligned} \dot{\overline{x}}_2 &= F_2\overline{x}_2 + B_2(\cdot)\Lambda_2(\cdot)\varepsilon_2, \\ \dot{\varepsilon}_2 &= [W + L_2B_2(\cdot)\Lambda_2(\cdot)]\varepsilon_2. \end{aligned} \tag{18}$$

The stabilization problem of the system (18) is solved by appropriate choice of the matrix $F_2,\ L_2$. The separation principle may be used. The elements of the matrix F_2 is chosen to provide negative real parts of the matrix F_2 eigen values. Then with $L_2 = L_{20}B_2^{-1}(0,0)$ $(L_{20} \in R^{r\times m})$ the stability of the second subsystem (18) may be provided

$$\dot{\varepsilon}_2 = [W + L_{20}\Lambda_2(0,0)]\varepsilon_2,$$

where the matrix $W + L_{20}\Lambda_2(0,0)$ is Hurwitz, if the pair $(\Lambda_2(0,0), W)$ is observable.

Thus, as in the linear case, the stabilization problem of the nonlinear system (11) can be solved based on the choice of local feedbacks in subsystems of lower dimension.

By applying the theory of the systems with separable motions based on deep feedback and/or sliding modes, it is also possible to decompose the solution of the object (11) stabilization problem, while the stability of the subsystem (16) is provided by using the methods of the modal control theory.

Let us consider a solution of the stabilization problem of the plant (13) by using sliding mode theory [9]. In the second subsystem (14), let us introduce the control input like the following

$$u = v - \Lambda_2(\cdot)\overline{z}_2 + B_2^{-1}(\cdot)\frac{\partial \varphi(x_1)}{\partial x_1}\dot{x}_1,\ v = -M(x)\operatorname{sign}\overline{x}_2,$$

where the variable $\overline{z}_2$ is described by the dynamic system

$$\dot{\overline{z}}_2 = W\overline{z}_2 + Lv, \quad \overline{z}_2 \in R^r. \tag{19}$$

For coordinates $\overline{x}_2$, $\overline{\varepsilon}_2 = \omega - \overline{z}_2$ from (14), taking into account (19), the following relations can be obtained

$$\begin{aligned} \dot{\overline{x}}_2 &= f_2(\cdot) + B_2(\cdot)[v + \Lambda_2(\cdot)\overline{\varepsilon}_2], \\ \dot{\overline{\varepsilon}}_2 &= W\overline{\varepsilon}_2 + Lv. \end{aligned} \tag{20}$$

Let us choose the matrix

$$M(x) = B_2^{-1}(x)\mathrm{diag}\{M_i\}, \quad M_i = \mathrm{const} > 0 \;\; (i = \overline{1,m}),$$

where diagM_i is diagonal matrix.

With the sufficiently large M_i, the sliding mode arises along the manifold

$$\overline{x}_2 = x_2 - \varphi(x_1) = 0.$$

By choosing a vector function $\varphi(x_1)$, that stabilizes the first equation (14), the problem of the stability of the first subsystem (14) is solved.

In the sliding mode, after convergence of the variables $x_1,\ x_2$ to the origin ($f_2(x_1, x_2) = 0$), one can use the equivalent control method [9] and write

$$v_{eq} = -B_2^{-1}(0,0)\Lambda_2(0,0)\overline{\varepsilon}_2.$$

Given that the pair of matrices $(\Lambda_2(0,0), W)$ is observable, by assigning the coefficients of the matrix $L = L_0B_2^{-1}(0,0)$, a spectrum with negative real parts of the eigenvalues is provided for the matrix of the internal dynamic for the variable $\overline{\varepsilon}_2$ from (20):

$$\dot{\overline{\varepsilon}}_2 = [W - L_0\Lambda(0,0)]\overline{\varepsilon}_2.$$

4. DISTURBANCES GENERATOR WITH UNKNOWN PARAMETERS

This section discusses one of the variants of the plant (1) with a control vector of the unit dimension. It is assumed, that the behavior of disturbances is described by a linear non-stationary dynamic system with a certain interval parameters:

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + A_{12}x_2, \\ \dot{x}_2 &= A_{21}x_1 + A_{22}x_2 + B_2(u + \eta), \end{aligned} \tag{21}$$

where $x_1 \in R^{n-1},\ \ x_2 \in R,\ \ u,\ \eta \in R^1$ and the disturbances are generated by a dynamic model of the form

$$\dot{\omega} = W\omega, \quad \eta = q^{\mathrm{T}}\omega, \;\; \omega \in R^r, \;\; \dim W = (r \times r). \tag{22}$$

The parameters of the matrix W are assumed to be unknown, but belonging to some intervals $|w_{ij}| \le W_{ij} = \mathrm{const}$.

In the previous sections, it was indicated, that the controllability of the pair (A_{11}, A_{12}) implies the controllability of the original system. Using a non-singular change of variables $\overline{x}_1 = x_2 - f_1^{\mathrm{T}}x_1$, the equations of the system (22) can be rewritten like the following

$$\begin{aligned} \dot{x}_1 &= (A_{11} + a_{12}f_1^{\mathrm{T}})x_1 + a_{12}\overline{x}_1, \\ \dot{\overline{x}}_1 &= \overline{a}_{21}^{\mathrm{T}}x_1 + \overline{a}_{22}\overline{x}_1 + b_2u + q^{\mathrm{T}}\omega, \end{aligned} \tag{23}$$

By choosing the vector f_1^{T}, the eigenvalues of the matrix $A_{11}+a_{12}f_1^{\mathrm{T}}$ are assigned, and the problem of stabilizing the (23) object is reduced to ensuring the stability of the second subsystem.

Since the parameters of the disturbances generator (22) are unknown, a dynamic compensator in the form of (6) cannot be selected to solve stated problem. Without loss of generality, under the assumption of observability of the pair (q^{T}, W), the disturbances generator model (22) can be transformed to the canonical representation

$$\dot{\omega}_{\mathrm{i}} = \omega_{i+1},\ \ i=\overline{1, r-1},\ \ \dot{\omega}_r = \sum_{\mathrm{i}=1}^{r} g_i\omega_i,\ \ |g_i| \le G_i = \mathrm{const},\ \ \omega_1 = q^{\mathrm{T}}\omega. \tag{24}$$

The second subsystem (23) with the control input $b_2u = -\overline{a}_{21}^{\mathrm{T}}x_1 - \overline{a}_{22}\overline{x}_1 + \overline{u}$ is written in the form

$$\dot{\overline{x}}_1 = \overline{u} + \omega_1. \tag{25}$$

Therefore, the original problem has been transformed into the problem of the dynamic compensation of the disturbances in the plant (25) with the perturbations described by the dynamic model (24).

This problem can be solved by choosing $\overline{u} = -z_1$, where z_1 is the output of the dynamic system

$$\dot{z}_i = z_{i+1},\ \ i=\overline{1, r-1},\ \ \dot{z}_r = \sum_{\mathrm{i}=1}^{r} d_iz_i + v,\ \ d_i = \mathrm{const}, \tag{26}$$

where v is the new control input.

Introducing the notation $e_i = \omega_i - z_i$, $i=\overline{1, r}$, one can rewrite the system (25) as

$$\dot{\overline{x}}_1 = e_1. \tag{27}$$

Differentiating the variable (27) r times, the expanded system can be obtained in the canonical form [13]

$$\dot{\overline{x}}_i = \overline{x}_{i+1},\ \ i=\overline{1, r-1},\ \ \dot{\overline{x}}_r = \sum_{i=1}^{r}[g_ie_i + (g_i - d_i)z_i] + v, \tag{28}$$

where $\overline{x}_{i+1} = e_i,\ i=\overline{1, r-1}$.

By choosing a sufficiently large amplitude of discontinuous control inputs [9], taking into account the constraints $|g_i| \le k_i$, the system (28) can be stabilized by the following control action:

$$v = -M\mathrm{sign}(s), \tag{29}$$

due to which a sliding mode is organized on the manifold

$$s = \overline{x}_r + \sum_{i=1}^{r-1} C_i\overline{x}_i = 0,$$

with motion along it described by the equivalent control method

$$\dot{\overline{x}}_i = \overline{x}_{i-1},\ \ i=\overline{1, r-1},\ \ \dot{\overline{x}}_r = \sum_{i=1}^{r} C_i\overline{x}_i.$$

If C_i are chosen as the coefficients of the Hurwitz polynomial, then sliding mode motion ensures asymptotic convergence of the system.

The proposed scheme for constructing invariant systems, using a canonical representation and sliding mode theory methods, provides the robustness properties of whole system, which includes plant dynamics and the disturbances model. Note, that the perturbation generator (24) allows the assumption of variable model parameters, that are bounded in absolute value: $|g_i(t)| \le G_i = \text{const}$.

To obtain the estimates of the system state space vector (28), necessary for the control input (29) synthesis, a state space observer must be constructed.

For order r, further a procedure for synthesizing an observer-differentiator of the variable (27) is synthesized.

1. Introduce a filter with stable eigen motions

$$\begin{aligned} &\dot{\xi}_i(t) = \xi_{i+1}(t) + \overline{x}_1,\ i = \overline{1, r-1}, \\ &\dot{\xi}_r(t) = c^{\mathrm{T}}\xi(t) + \overline{x}_1, \end{aligned} \tag{30}$$

where $\xi^{\mathrm{T}}(t) = (\xi_1(t), \ldots, \xi_r(t)) \in R^r$, $c \in R^r$ is the coefficient vector of Hurwitz polynomial.

2. Represent the system (30) in the canonical form

$$\begin{aligned} &\dot{\overline{\xi}}_i(t) = \overline{\xi}_{i+1}(t),\ i = \overline{1, r-1}, \\ &\dot{\overline{\xi}}_r(t) = c^{\mathrm{T}}\xi(t) + \overline{\xi}_{r+1}(\mathrm{t}), \end{aligned} \tag{31}$$

where

$$\begin{aligned} &\overline{\xi}_1(t) = \xi_1(t), \\ &\overline{\xi}_2(t) = \xi_2(t) + \overline{x}_1, \\ &\overline{\xi}_3(t) = \xi_3(t) + \overline{x}_1 + \overline{x}_2, \\ &\ldots \\ &\overline{\xi}_r(t) = \xi_r(t) + \sum_{i=1}^{r-1} \overline{x}_i, \\ &\overline{\xi}_{r+1}(t) = \sum_{i=1}^{r} \overline{x}_i. \end{aligned} \tag{32}$$

3. Let us construct an observer of the state vector $\overline{\xi}^{\mathrm{T}}(t) = (\overline{\xi}_1(t), \ldots, \overline{\xi}_3(t))$ and the signal $\overline{\xi}_{r+1}(t)$ as applied to the system (31):

$$\begin{aligned} &\dot{\hat{\xi}}_i(t) = \hat{\xi}_{i+1}(t) + \nu_i,\ i = \overline{1, r-1}, \\ &\dot{\hat{\xi}}_r(t) = c^{\mathrm{T}}\hat{\xi}(t) + \nu_r, \quad \hat{\xi}^{\mathrm{T}}(t) = (\hat{\xi}_1(t), \ldots, \hat{\xi}_r(t)). \end{aligned} \tag{33}$$

4. The equations (31) and (33) in terms of residuals $\overline{\varepsilon}_{\mathrm{i}}(t) = \overline{\xi}_i(t) - \hat{\xi}_i(t)$ are:

$$\begin{aligned} &\dot{\overline{\varepsilon}}_{\mathrm{i}}(t) = \overline{\varepsilon}_{i+1}(t) - \nu_i,\ i = \overline{1, r-1}, \\ &\dot{\overline{\varepsilon}}_r(t) = c^{\mathrm{T}}\overline{\varepsilon}(t) + \overline{\xi}_{r+1}(t) - \nu_r, \end{aligned} \tag{34}$$

where $\overline{\varepsilon}^{\mathrm{T}}(t) = (\overline{\varepsilon}_1(t), \ldots, \overline{\varepsilon}_r(t))$.

In papers [14–16], it was shown that the choice of the corrective input actions of the observer (34) in the class of linear functions with the saturation according to the hierarchy

$$\nu_1 = \mathrm{sat}_{L_1}(l_1\overline{\varepsilon}_1(t)),\ \ \nu_i = \mathrm{sat}_{L_i}(l_i v_{i-1}),\ \ i = \overline{2, r},\ L_1,\ \ L_i, l_i = \mathrm{const}$$

solves the stabilization problem of the system (34) with a given accuracy.

Definition 1. For $M = \text{const} > 0$ and an arbitrary time function $s(t)$, by definition

$$\mathrm{sat}_M(s) = \min(|s|, M)\mathrm{sign}(s(t)).$$

Indeed, under the assumptions

$$|\overline{\varepsilon}_i(t)| \le E_i,\ |\dot{\overline{\varepsilon}}_i(t)| \le \overline{E}_i;\quad \mathrm{E_i}, \overline{E}_i = \text{const},$$
$$\left|\overline{\xi}_{r+1}(t)\right| \le E_{r+1} = \text{const}$$

We schematically describe the stabilization procedure for the system (34) .

1. In the first subsystem

$$\dot{\overline{\varepsilon}}_1(\mathrm{t}) = \overline{\varepsilon}_2(t) - \mathrm{sat}_{L_1}(l_1\overline{\varepsilon}_1(t)),$$

with the amplitude

$$L_1 > |\overline{\varepsilon}_2(t)|$$

the variable $\overline{\varepsilon}_1(t)$ converges to the linear zone after a finite time, and the equality holds:

$$\dot{\overline{\varepsilon}}_1(\mathrm{t}) = \overline{\varepsilon}_2(t) - l_1\overline{\varepsilon}_1(t).$$

In this case, the relations $|\overline{\varepsilon}_1(t)| \le E_2/l_1,\ |\overline{\varepsilon}_2(t) - l_1\overline{\varepsilon}_1(t)| \le \overline{E}_2/l_1$ are satisfied, and by increasing the coefficient l_1 for the variable $\overline{\varepsilon}_1(t) \approx 0$, a given degree of stabilization can be ensured. As a result, the estimation with a given accuracy of the variable $\overline{\varepsilon}_2$ is $\overline{\varepsilon}_2(t) \approx l_1\overline{\varepsilon}_1(t)$.

Let us describe the first point with the help of logical diagram:

$$\begin{aligned}
&\dot{\overline{\varepsilon}}_1(t) = \overline{\varepsilon}_2(t) - \mathrm{sat}_{L_1}(l_1\overline{\varepsilon}_1(t)),\ L_1 > |\overline{\varepsilon}_2(t)| \Rightarrow \dot{\overline{\varepsilon}}_1(\mathrm{t}) = \overline{\varepsilon}_2(t) - l_1\overline{\varepsilon}_1(t)\\
&\Rightarrow |\overline{\varepsilon}_1(t)| \le \frac{E_2}{l_1},\ |\,\overline{\varepsilon}_2(t) - l_1\overline{\varepsilon}_1(t)| \le \frac{\overline{E}_2}{l_1} \Rightarrow l_1 \to \infty :\\
&\overline{\varepsilon}_i(t) \approx 0,\ \overline{\varepsilon}_{i+1}(t) \approx l_i\overline{\varepsilon}_i(t).
\end{aligned}$$

2. In the following stages $i = \overline{2, r-1}$, corrective actions are synthesized according to the following logical scheme:

$$\begin{aligned}
&\dot{\overline{\varepsilon}}_i(t) = \overline{\varepsilon}_{i+1}(t) - \mathrm{sat}_{L_i}(l_i\overline{\varepsilon}_i(t)),\ L_\mathrm{i} > |\overline{\varepsilon}_{i+1}(t)| \Rightarrow \dot{\overline{\varepsilon}}_\mathrm{i}(\mathrm{t}) = \overline{\varepsilon}_{i+1}(t) - l_i\overline{\varepsilon}_i(t) \Rightarrow\\
&\Rightarrow |\overline{\varepsilon}_i(t)| \le \frac{E_{i+1}}{l_i},\ |\,\overline{\varepsilon}_{i+1}(t) - l_i\overline{\varepsilon}_i(t)| \le \frac{\overline{E}_{i+1}}{l_i} \Rightarrow l_i \to \infty :\\
&\overline{\varepsilon}_i(t) \approx 0,\ \overline{\varepsilon}_{i+1}(t) \approx l_i\overline{\varepsilon}_i(t).
\end{aligned}$$

3. After consideration the last subsystem (34)

$$\dot{\overline{\varepsilon}}_r(t) = c^{\mathrm{T}}\overline{\varepsilon}(t) + \overline{\xi}_{r+1}(t) - \mathrm{sat}_{L_r}(l_r\overline{\varepsilon}_r(t)),$$

and taking into account $\overline{\varepsilon}_i(t) \approx 0$ by choosing $L_r > \left|\overline{\xi}_{r+1}(t)\right|$ the relation is obtained

$$\overline{\xi}_{r+1}(t) \approx l_r\overline{\varepsilon}_r(t).$$

The variables $\overline{x}_i,\ i = \overline{1, r}$, corresponding to the derivatives of the signal $\overline{x}_1$, can be calculated using the estimates $\overline{\xi}(t)$ and the signal $\overline{\xi}_{r+1}(t)$ by sequentially solving the equations (32) from top to bottom.

5. EXAMPLE OF THE USE OF THE DEVELOPED ALGORITHMS FOR UAV CONTROL

This section examines the applied problem of controlling the horizontal movement of an unmanned aerial vehicle (UAV) transporting a load suspended on a weightless cable (see Fig. 1). The

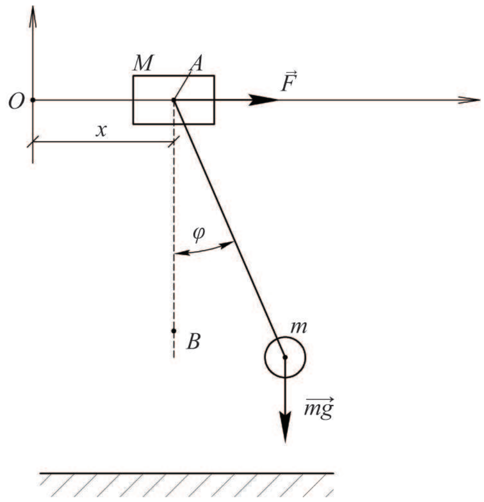

Fig. 1. The qualitative scheme of UAV with the transported cargo.

dynamic equations describing the behavior of the specified object can be represented as

$$\begin{aligned}
&\dot{x} = v,\\
&\dot{v} = -\frac{J_{\text{eq}}\alpha_v}{\Delta}v - \frac{m^2l^2g}{\Delta}\varphi - \frac{ml\alpha_\omega}{\Delta}\omega + \frac{J_{\text{eq}}}{\Delta}F,\\
&\dot{\varphi} = \omega,\\
&\dot{\omega} = -\frac{mglM_{\text{eq}}}{\Delta}\varphi - \frac{M_{\text{eq}}\alpha_\omega}{\Delta}\omega - \frac{ml\alpha_v}{\Delta}v + \frac{ml}{\Delta}F,
\end{aligned} \tag{35}$$

where x is the horizontal coordinate of the center of mass, v is the horizontal velocity of the UAV center of mass, φ is the angular deviation of the cable with the load, ω is the angular velocity of the cable with the load, $J_{\text{eq}} = J + ml^2$, J is the inertia of the suspended load during its rotation about its center of gravity, $M_{\text{eq}} = m + M$, m, M are the masses of the suspended body and the UAV, respectively, l is the distance from the point of suspension of the load to its center of gravity, $\Delta = M_{\text{eq}}J_{\text{eq}} - m^2l^2$, AB is the line of zero deflection of the cable with the load, OA is the line along which the UAV moves with the load (horizontal line), α_M is a coefficient of viscous friction of the medium for the aircraft, α_m is a coefficient of viscous friction of the medium for the suspended load, α_r is a coefficient taking into account the effect of losses on friction at the point of suspension of the cable, $\alpha_v = \alpha_m + \alpha_M$, $\alpha_\omega = \alpha_m l + \alpha_r$, F is the traction force in in the horizontal direction, the cable is assumed to be weightless compared to the load.

The problem of the horizontal position stabilization of UAV at a given point is stated in this section:

$$\lim_{t\to\infty}|x(t) - x_{\text{d}}| = \lim_{t\to\infty}|\overline{x}(t)| = 0,\ \ \overline{x}(t) = x(t) - x_{\text{d}},\ \ x_{\text{d}} = \text{const} > 0. \tag{36}$$

It is assumed, that only UAV position and velocity can be measured.

Using a change of variables

$$\begin{pmatrix} y \\ v_y \end{pmatrix} = \begin{pmatrix} -\dfrac{1}{J_{\text{eq}}} & 0 & \dfrac{1}{ml} & 0 \\ 0 & -\dfrac{1}{J_{\text{eq}}} & 0 & \dfrac{1}{ml} \end{pmatrix} \begin{pmatrix} x \\ v \\ \varphi \\ \omega \end{pmatrix},$$

one can rewrite the equations (35) in the form

$$
\begin{aligned}
&\dot{x} = v,\\
&\dot{v} = -\frac{gl^3m^3}{J_{\text{eq}}\Delta}x - \frac{\alpha_v J_{\text{eq}}^2 + \alpha_\omega l^2 m^2}{J_{\text{eq}}\Delta}v - \frac{gl^3m^3}{\Delta}y - \frac{\alpha_\omega l^2 m^2}{\Delta}v_y + \frac{J_{\text{eq}}}{\Delta}F,\\
&\dot{y} = v_y,\\
&\dot{v}_y = -\frac{mgl}{J_{\text{eq}}}y - \frac{\alpha_\omega}{J_{\text{eq}}}v_y - \frac{mgl}{J_{\text{eq}}^2}x - \frac{\alpha_\omega}{J_{\text{eq}}^2}v.
\end{aligned}
$$

Taking into account (36), the last system is transformed to the representation:

$$
\begin{aligned}
&\dot{\overline{x}} = v,\\
&\dot{v} = -\frac{gl^3m^3}{J_{\text{eq}}\Delta}\overline{x} - \frac{\alpha_v J_{\text{eq}}^2 + \alpha_\omega l^2 m^2}{J_{\text{eq}}\Delta}v - \frac{gl^3m^3}{\Delta}y - \frac{\alpha_\omega l^2 m^2}{\Delta}v_y + \frac{J_{\text{eq}}}{\Delta}F-\\
&\quad - \frac{gl^3m^3}{J_{\text{eq}}\Delta}x_{\text{d}},\\
&\dot{y} = v_y,\\
&\dot{v}_y = -\frac{mgl}{J_{\text{eq}}}y - \frac{\alpha_\omega}{J_{\text{eq}}}v_y - \frac{mgl}{J_{\text{eq}}^2}\overline{x} - \frac{\alpha_\omega}{J_{\text{eq}}^2}v - \frac{mgl}{J_{\text{eq}}^2}x_{\text{d}}.
\end{aligned}
\tag{37}
$$

According to the proposed approach, unmeasured state variables y, v_y may be treated as disturbances. Considering, that the second and fourth equations of the system contain constant terms $-\frac{gl^3m^3}{J_{\text{eq}}\Delta}x_{\text{d}}$ and $-\frac{mgl}{J_{\text{eq}}^2}x_{\text{d}}$, the complete disturbances model according to (37) can be represented as

$$
\dot{\omega} = W\omega + B_e e,
\tag{38}
$$

where $\omega = \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix}^{\text{T}} = \begin{pmatrix} x_{\text{d}} & y & v_y \end{pmatrix}^{\text{T}}$, $W = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -\frac{mgl}{J_{\text{eq}}^2} & -\frac{mgl}{J_{\text{eq}}} & -\frac{\alpha_\omega}{J_{\text{eq}}} \end{pmatrix}$,

$e = \begin{pmatrix} \overline{x} & v \end{pmatrix}^{\text{T}}$, $B_e = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -\frac{mgl}{J_{\text{eq}}^2} & -\frac{\alpha_\omega}{J_{\text{eq}}^2} \end{pmatrix}$.

Using (37)–(38), the transformed system (37) is

$$
\begin{aligned}
&\dot{\overline{x}} = v,\\
&\dot{v} = -\frac{gl^3m^3}{J_{\text{eq}}\Delta}\overline{x} - \frac{\alpha_v J_{\text{eq}}^2 + \alpha_\omega l^2 m^2}{J_{\text{eq}}\Delta}v + Q\omega + \frac{J_{\text{eq}}}{\Delta}F,\\
&\dot{\omega} = W\omega + B_e e,
\end{aligned}
\tag{39}
$$

where $Q = \begin{pmatrix} -\frac{gl^3m^3}{J_{\text{eq}}\Delta} & -\frac{gl^3m^3}{\Delta} & -\frac{\alpha_\omega l^2 m^2}{\Delta} \end{pmatrix}$.

The difference between the systems (39) and (5) representations is that the disturbances subsystem is non-autonomous (the $B_\omega e$ term is present). However, as will be shown below, this does not change the proposed methodology to feedback synthesis.

We will demonstrate by direct calculations, that the (W, Q) pair is unobservable. To do this, the rank of the observability matrix is calculated:

$$\operatorname{rank}\begin{pmatrix} Q \\ QW \\ QW^2 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} -\dfrac{gm^3l^3}{J_{\mathrm{eq}}\Delta} & -\dfrac{gm^3l^3}{\Delta} & -\dfrac{\alpha_\omega m^2l^2}{\Delta} \\ \dfrac{\alpha_\omega gl^3m^3}{\Delta J_{\mathrm{eq}}^2} & \dfrac{\alpha_\omega gl^3m^3}{\Delta J_{\mathrm{eq}}} & \dfrac{m^2l^2(\alpha_\omega^2 - gmlJ_{\mathrm{eq}})}{J_{\mathrm{eq}}\Delta} \\ \dfrac{gm^3l^3(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{J_{\mathrm{eq}}^3\Delta} & \dfrac{gm^3l^3(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{J_{\mathrm{eq}}^2\Delta} & \dfrac{\alpha_\omega m^2l^2(-\alpha_\omega^2 + 2gmlJ_{\mathrm{eq}})}{J_{\mathrm{eq}}^2\Delta} \end{pmatrix}$$

$$= \operatorname{rank}\begin{pmatrix} -\dfrac{gm^3l^3}{J_{\mathrm{eq}}\Delta} & -\dfrac{gm^3l^3}{\Delta} & -\dfrac{\alpha_\omega m^2l^2}{\Delta} \\ 0 & 0 & -\dfrac{gm^3l^3}{\Delta} \\ 0 & 0 & \dfrac{gm^3l^3}{J_{\mathrm{eq}}\Delta} \end{pmatrix} = \begin{pmatrix} -\dfrac{gm^3l^3}{J_{\mathrm{eq}}\Delta} & -\dfrac{gm^3l^3}{\Delta} & -\dfrac{\alpha_\omega m^2l^2}{\Delta} \\ 0 & 0 & -\dfrac{gm^3l^3}{\Delta} \\ 0 & 0 & 0 \end{pmatrix} = 2.$$

The system (39), using the substitution of functions,

$$\xi = H\omega,\ H = \begin{pmatrix} Q \\ QW \\ H_3 \end{pmatrix},\ H_3 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix},\ \xi = \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 \end{pmatrix}^{\mathrm{T}}$$

is transformed into the following

$$\begin{aligned} \dot{\overline{x}} &= v, \\ \dot{v} &= -\frac{gl^3m^3}{J_{\mathrm{eq}}\Delta}\overline{x} - \frac{\alpha_v J_{\mathrm{eq}}^2 + \alpha_\omega l^2m^2}{J_{\mathrm{eq}}\Delta}v + \xi_1 + \frac{J_{\mathrm{eq}}}{\Delta}F, \\ \dot{\xi} &= \widetilde{W}\xi + \widetilde{B}_e e, \end{aligned} \tag{40}$$

where $\widetilde{W} = HWH^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -\dfrac{gml}{J_{\mathrm{eq}}} & -\dfrac{\alpha_\omega}{J_{\mathrm{eq}}} & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

$$\widetilde{B}_e = HB_e = \begin{pmatrix} \dfrac{\alpha_\omega gm^3l^3}{\Delta J_{\mathrm{eq}}^2} & \dfrac{\alpha_\omega^2 m^2l^2}{\Delta J_{\mathrm{eq}}^2} \\ \dfrac{gm^3l^3(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{\Delta J_{\mathrm{eq}}^3} & \dfrac{\alpha_\omega m^2l^2(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{\Delta J_{\mathrm{eq}}^3} \\ 0 & 0 \end{pmatrix}.$$

In the structure (40) is highlighted unobservable subspace with respect to the variable ξ_4. When synthesizing feedback based on a dynamic compensator, the active variable suddenly does not contribute to the output error of the variable, so the active compensator constructs the observed components $\xi^* = (\ \xi_1 \ \ \xi_2\ )^{\mathrm{T}}$ of the the disturbances vector ξ.

The observability of the pair $(W^*,\ Q^*)$ is now revealed by the direct substitution

$$\operatorname{Rank}\begin{pmatrix} Q^* \\ Q^*W^* \end{pmatrix} = \operatorname{Rank}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2,$$

where

$$W^* = \begin{pmatrix} 0 & 1 \\ -\dfrac{gml}{J_{\mathrm{eq}}} & -\dfrac{\alpha_\omega}{J_{\mathrm{eq}}} \end{pmatrix}, \; Q^* = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

In the introduced notations, the disturbances model is

$$\dot{\xi}^* = W^*\xi^* + B_e^* e, \; B_e^* = \begin{pmatrix} \dfrac{\alpha_\omega g m^3 l^3}{\Delta J_{\mathrm{eq}}^2} & \dfrac{\alpha_\omega^2 m^2 l^2}{\Delta J_{\mathrm{eq}}^2} \\ \dfrac{g m^3 l^3(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{\Delta J_{\mathrm{eq}}^3} & \dfrac{\alpha_\omega m^2 l^2(-\alpha_\omega^2 + gmlJ_{\mathrm{eq}})}{\Delta J_{\mathrm{eq}}^3} \end{pmatrix}. \tag{41}$$

The horizontal control force is chosen like the following

$$F = -\frac{\Delta}{J_{\mathrm{eq}}} Q^* s + \frac{\Delta}{J_{\mathrm{eq}}} \overline{u} \tag{42}$$

with dynamic compensator according to (6), (41)

$$\dot{s} = W^* s + B_e^* e - \overline{v}. \tag{43}$$

The equations of the closed-loop system according to (40), (42)–(43):

$$\begin{aligned} &\dot{\overline{x}} = v, \\ &\dot{v}_= - \frac{g l^3 m^3}{J_{\mathrm{eq}}\Delta}\overline{x} - \frac{\alpha_v J_{\mathrm{eq}}^2 + \alpha_\omega l^2 m^2}{J_{\mathrm{eq}}\Delta} v + Q^*\varepsilon + \overline{u}, \\ &\dot{\varepsilon} = W\varepsilon + \overline{v}, \end{aligned} \tag{44}$$

where $\varepsilon = \xi^* - s$.

Selecting the control inputs $\overline{u}$, $\overline{v}$ like in the expressions (9)

$$\begin{aligned} &\overline{v} = -LKe + W^* L v, \\ &\overline{u} = \frac{g l^3 m^3}{J_{\mathrm{eq}}\Delta}\overline{x} + \left(\frac{\alpha_v J_{\mathrm{eq}}^2 + \alpha_\omega l^2 m^2}{J_{\mathrm{eq}}\Delta} + Q^* L\right) v + Ke, \end{aligned} \tag{45}$$

the equations of the closed-loop system according to (44)–(45) are derived

$$\begin{aligned} &\dot{\overline{x}} = v, \\ &\dot{v} = Ke + Q^*(\varepsilon + Lv), \\ &\dot{\varepsilon} = W^*\varepsilon - LKe + W^* L v, \end{aligned}$$

where $L = \begin{pmatrix} L_1 & L_2 \end{pmatrix}^{\mathrm{T}} \in R^2$, $K = \begin{pmatrix} K_1 & K_2 \end{pmatrix}^{\mathrm{T}} \in R^2$ are the matrices with constant coefficients.

By introducing the new variable $\overline{\varepsilon} = \varepsilon + Lv$, the last system is rewritten as

$$\begin{aligned} &\dot{\overline{x}} = v, \\ &\dot{v} = Ke + Q^*\overline{\varepsilon}, \\ &\dot{\overline{\varepsilon}} = (W^* + LQ^*)\overline{\varepsilon}. \end{aligned}$$

Since the pair (W^*, Q^*) is observable, the convergence of the variable $\overline{\varepsilon}$ can be assigned at a desired rate.

The numerical values of the plant parameters adopted in the simulation

Parameter	Value
l, [m]	5
J, [kg$\cdot$ m^2]	1.2
α_v, [kg/s]	10
α_ω, [m$^2\cdot$kg/s]	20
m, [kg]	3.5
M, [kg]	2
a_1	7.5
a_0	14
$x(t_0)$, [m]	10
$v(t_0)$, [m/s]	5
$\varphi(t_0)$, [rad]	–0.5
$\omega(t_0)$, [rad/s]	10
x_d, [m]	25

For the numerical experiment, the selected parameters are presented in the table.

The characteristic polynomial of the matrix $W^* + LQ^*$:

$$p^2 + \frac{-L_1 J_\mathrm{eq} + \alpha_\omega}{J_\mathrm{eq}} p + \frac{-L_2 J_\mathrm{eq} + gml - L_1 \alpha_\omega}{J_\mathrm{eq}} = 0. \tag{46}$$

To assign the coefficients of the last polynomial equal to the desired ones, the elements of the matrix L are chosen. The desired polynomial is

$$p^2 + a_1 p + a_0 = 0, \tag{47}$$

with the parameters from the table and the roots of the equation

$$p_1 = -3.5, \ p_2 = -4.$$

According to the reasoning, from (46)–(47), one can obtain

$$\frac{-L_1 J_\mathrm{eq} + \alpha_\omega}{J_\mathrm{eq}} = a_1 = 7.5, \quad \frac{-L_2 J_\mathrm{eq} + gml - L_1 \alpha_\omega}{J_\mathrm{eq}} = a_0 = 14.$$

Substituting the parameters from the table, the feedback coefficients of the dynamic compensator are calculated

$$L_1 = \frac{\alpha_\omega - a_1 J_\mathrm{eq}}{J_\mathrm{eq}} = -7.2745;$$

$$L_2 = \frac{-a_0 J_\mathrm{eq}^2 - (\alpha_\omega - a_1 J_\mathrm{eq})\alpha_\omega + J_\mathrm{eq} gml}{J_\mathrm{eq}^2} = -10.4243.$$

The following feedback matrix K is used in the control inputs (45)

$$K = \begin{pmatrix} -0.12 & -0.7 \end{pmatrix}.$$

The feedback with such parameters ensures stability and assigns time constants

$$\tau_1 = 0.4 \text{ s}^{-1}, \quad \tau_2 = 0.3 \text{ s}^{-1}$$

in the subsystem corresponding to the vector e.

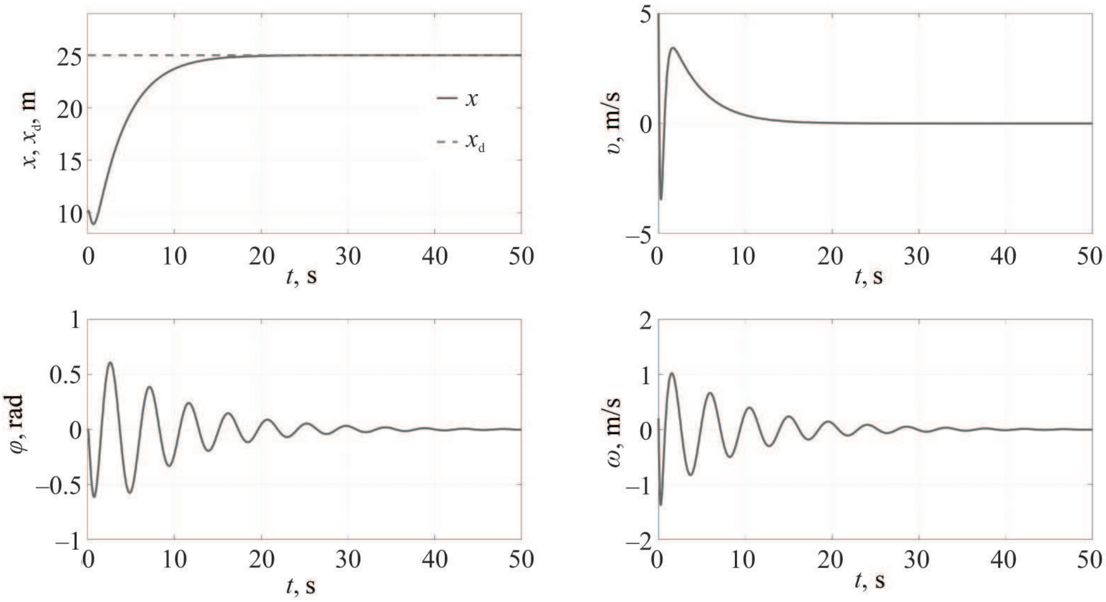

Fig. 2. Transient process for the state space of the plant (31).

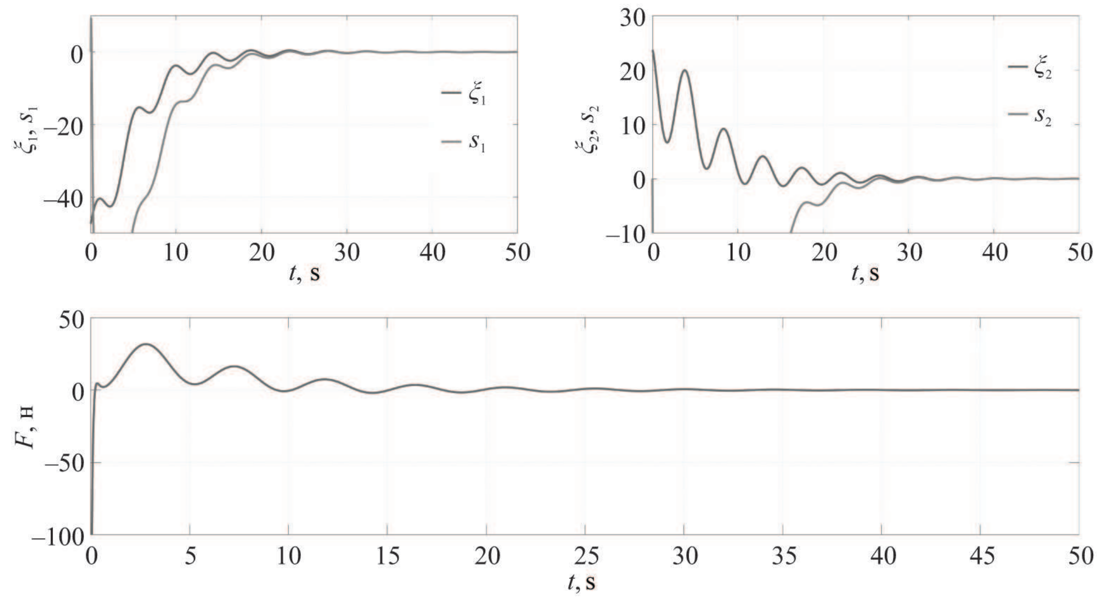

Fig. 3. Transient process for the components of the disturbances vector (36) and dynamic compensator (39).

The simulation results of the system (31) with the dynamic feedback (42)–(43), (45) and the initial conditions from the table are shown on Figs. 2 and 3.

As can be seen from Fig. 3, the current UAV altitude converges to the desired value, and the components of the dynamic compensator vector s converge to the corresponding components of the disturbances vector ξ^*. Note that, despite of the convergence of the output variable to zero, the control force F contains the damped oscillatory component (see Fig. 3). This is due to the fact, that the control input compensates the oscillations of the variables φ, ω, the damping rate of which is determined by the coefficient α_ω.

The control law (42)–(43), (45) was synthesized without correction of this damping, and for this reason, a long tail exists in the transient process (see Figs. 2 and 3). However, due to the controllability of the original system, (37) by correcting with the using the components $\overline{x}, v$ (see equations (37)), it would be possible to correct the damping of the components φ, ω, and, consequently, the length of the specified tail.

6. CONCLUSION

The decomposition methods are proposed for the problem of dynamic compensation of the external disturbances of a given class. The use of methods for synthesizing systems with separable motions within the framework of sliding mode theory and deep feedback loops has allowed us to extend existing dynamic compensation methods to a class of nonlinear systems and ensure robust properties under uncertainty in both, the control plant model and the disturbances generator. It should be noted, that the considered example of UAV control with an oscillating load, which can be considered both as part of the control plant model or as an external disturbance, demonstrates the potential for extending the problems of dynamic compensation of the external disturbances to the tasks, where the disturbances generator is considered as a part of the control plant. Together with the decomposition methods proposed in this paper, this approach also allows designers to decompose the feedback synthesis procedure into smaller subproblems.

FUNDING

The work was partially supported by Russian Science Foundation, project no. 25-11-00067, https://rscf.ru/project/25-11-00067/.

REFERENCES

1. Kulebakin, V.S., K teorii avtomaticheskikh vibratsionnykh regulyatorov elektricheskikh mashin, *Teoreticheskaya i eksperimental'naya electrotekhnika* (Theoretical and Experimental Electrotechnics), 1932, no. 4, pp. 3–21.

2. Wonhem, W.M., *Linear Multivariable Control: A Geometric Approach*, New York: Springer, 1985.

3. Francis, B.A. and Wonham, W.M., The Internal Model Principle of Control Theory, *Automatica*, 1979. vol. 12, no. 5, pp. 457–465.

4. Khalil, H.K., Cascade High-Gain Observers in Output Feedback Control, *Automatica*, 2017, vol. 80, pp. 110–118.

5. Krasnova, S.A., Utkin, V.A., and Mikheev, Yu.V., Cascade Design of State Observers, *Autom. Remote Control*, 2001, vol. 62, no. 2, pp. 207–226.

6. Pyrkin, A., Bobtsov, A., Ortega, R., and Isidori, A., An Adaptive Observer for Uncertain Linear Time-Varying Systems with Unknown Additive Perturbations, *Automatica*, 2023, vol. 147, no. 2, p. 110677.

7. Floquet, T., Edwards, C., and Spurgeon, S., On sliding Mode Observers for Systems with Unknown Inputs, *Journal of Adaptive Control and Signal Processing*, 2007, vol. 21, no. 8–9, pp. 638–656.

8. Lukianov, A.G., Utkin, V.A., Izosimov, D.B., et.al., Block Control Principle. I, *Autom. Remote Control*, 1990, vol. 51, no. 5, pp. 601–608.

9. Utkin, V., Guldner, J., and Shi, J., *Sliding Mode Control in Electro-Mechanical Systems*, CRC Press, 2009.

10. Filimonov, A.B. and Filimonov N.B., Robastanya correktsiya v sistemah upravleniya s bol'shimi coefficientami, *Mehatromica, avtomatizaciya, upravlenie* (Mechatronics, Automation, Control), 2014, vol. 12, pp. 3–10.

11. Utkin, V.A. and Utkin, V.I., Design of Invariant Systems by the Method of Motion Separation, *Autom. Remote Control*, 1983, vol. 44, no. 12, pp. 1559–1567.

12. Loukianov, A.G. and Utkin, V.I., Methods of Reducing Equations of Dynamic Systems to Regular Form, *Autom. Remote Control*, 1981, vol. 42, no. 4, pp. 413–420.

13. Floquet, T. and Barbot, J.P., A Canonical Form for the Design of Unknown Input Sliding Mode Observers, *Lecture Notes in Control and Information Sciences*, 2006, vol. 334, pp. 271–292.

14. Krasnova, S.A., Utkin, A.V., and Utkin, V.A., Synthesis of a Control System for a Waste Heat Boiler with Forced Circulation under Restrictions on Control Actions, *Mathematics*, 2022, vol. 10, no. 14, p. 2397.

15. *Krasnova, S.A., Utkin, A.V., and Utkin, V.A.*, Block Approach to Analysis and Design of the Invariant Nonlinear Tracking Systems, *Autom. and Rem. Control.*, 2017, vol. 78, no. 12, pp. 2120–2140.

16. Gulyukina, S.I. and Utkin, V.A., A Block Approach to CSTR Control Under Uncertainty, State-Space And Control Constraints, *Control Sciences*, 2021, no. 5, pp. 48–59.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 359–373.

SPECIAL ISSUE

Hybrid Controllers in Control Problems of Real Objects

A. F. Pashchenko[*,a] **and E. S. Duvanov**[**,b]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
[**] *Lipetsk State Technical University, Lipetsk, Russia*
e-mail: [a] *a.paschenko@ipu.ru,* [b] *evgenyduvanov@yandex.ru*

Received July 8, 2025
Revised October 15, 2025
Accepted November 20, 2025

Abstract—The paper continues the authors' research on control of complex technical facilities using hybrid methods. Two approaches are described and analyzed: neural network predictive and fuzzy hybrid controllers. Issues of adaptive tuning, control accuracy, and practical implementation of both methods are considered. Theoretical and practical results of the developed algorithms application in laboratory and industrial experiments are presented.

Keywords: hybrid control methods, neural network predictive controller, fuzzy quadratic regulator, biotechnical control systems, MATLAB

DOI: 10.7868/S1608303226040064

1. INTRODUCTION

Hybrid approaches combining classical algorithms with artificial intelligence methods are increasingly being adopted in modern control systems to enhance adaptability, accuracy, and reliability. Two of the most promising directions in hybrid control are neural network predictive controllers and fuzzy hybrid controllers. The first ones enable prediction of complex facility dynamics and optimization of control actions over a prediction horizon, while the latter effectively adapt to uncertainties and disturbances on the basis of expert rules. Both approaches demonstrate superiority over classical PI controllers in practice, yielding better control quality indices in the respective applied domains.

Artificial neural networks, capable of approximating complex nonlinear dependencies, are widely used for constructing predictive models of controlled devices. Within a conventional scheme, the model is pre-trained on historical data, and after that the neural network predicts future values of the output variable over a specified prediction horizon. Based on these predictions, an optimization algorithm generates the control action, minimizing a selected performance criterion. Thus, in the control system of a complex mechanical facility, a neural network predictive controller had demonstrated higher accuracy and faster response compared to a classical PI controller [1]. The application of a neural network model increases the sensitivity and responsiveness of the system to external disturbances, which is particularly relevant for problems with dynamic alterations. However, such methods require considerable computational resources and substantial training data, and this sophisticates their rapid practical deployment on industrial facilities.

Fuzzy hybrid controllers represent a combination of heuristic expert rules, formalized as membership functions and inference rules, with methods of optimal control. The hybrid architecture combines the robustness of fuzzy logic against uncertainties and nonlinearities with the precision and predictability of classical algorithms, providing adaptive real-time adjustment of control set points and improving the system's response to dynamic disturbances.

Fuzzy hybrid controllers find practical application due to their ability to consider complex interrelationships among various technological parameters. This approach reduces mode oscillations and ensures a smooth, stable transition to the desired set point compared with traditional PI controllers. Hybrid controllers in automatic systems demonstrate significantly higher accuracy in maintaining specified parameters, leading to improved product quality and reduced operational risks [2].

2. DESIGN AND ADJUSTMENT OF THE NEURAL NETWORK PREDICTIVE CONTROLLER

The neural network predictive controller ($NNPC$) is a model predictive control method adapted for nonlinear systems. A neural network trained on experimental data serves as the plant model; at each sampling instant it predicts the dynamics of the output signals. The control error optimization problem is then formulated and solved on the basis of these predictions and the prediction horizon. This approach combines the advantages of model-based predictive control with the flexibility of neural network models. The objective function takes the following form:

$$J = \sum_{i=N_1}^{N_2} \left(y_r(t+j) - y_m(t+j)\right)^2 + \rho \sum_{j=1}^{N_u} \left(u'(t+j-1) - u(t+j-2)\right)^2, \tag{1}$$

where N_1, N_2 are the minimum and maximum prediction horizons (typically $N_1 = 1$), N_u is the control horizon, y_r is the desired reference output, y_m is the network-predicted output, u' is the vector of proposed control inputs, and ρ is the control penalty weighting coefficient.

Figure 1 shows the structure of the control system based on the neural network predictive controller

The network is pre-trained offline in batch mode on experimental data from the actual plant or device, minimizing the prediction error—the difference between the true system output and the network prediction. A two-layer *MLP* (*Multilayer Perceptron*) is typically used for object identification. The network receives as inputs the most recent control and output values of the system (through the lag unit).

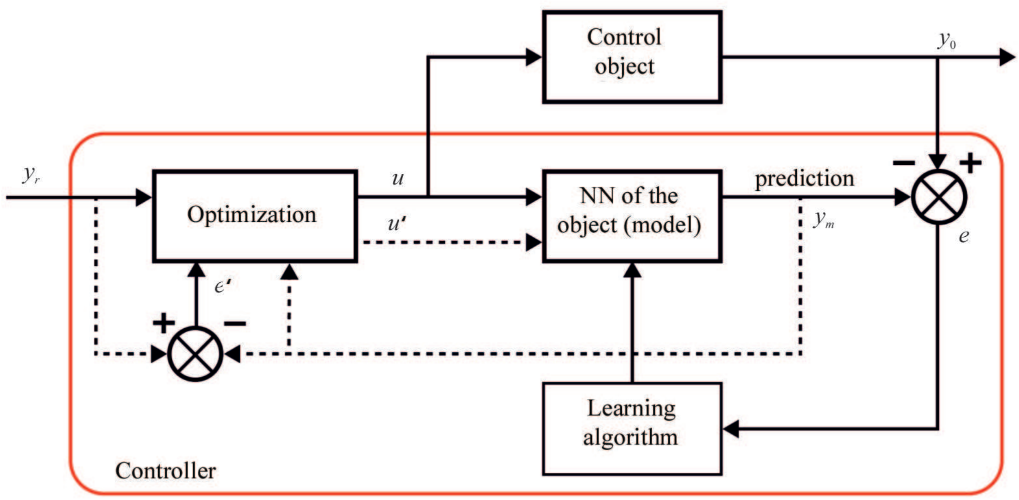

Fig. 1. Structure of the neural network predictive controller-based control system.

The network input vector comprises $u(t), u(t-1), \ldots$ and $y_p(t), y_p(t-1), \ldots$—the current and delayed values of the system's input and output. These inputs are followed by a single hidden layer with a nonlinear (sigmoidal) activation function. The output is a linear combination of the hidden layer, yielding the one-step-ahead prediction $y_m(t+1)$.

This approach rests on the universal approximation theorem [3]: a two-layer neural network can approximate any nonlinear function to sufficient accuracy. After training is complete, the network weights are stored and the model is integrated into the controller for sequential prediction of future system responses. At each time step the network processes input $u'(t)$ and the previous state, producing $y_m(t+1)$; this output is then fed back as an input when necessary to simulate multi-step-ahead behavior depending on the prediction horizon.

Suppose the network is trained so that at time instants $t, t-1, \ldots$, it produces a one-step-ahead prediction $y_m(t+1)$ as given inputs. Repeated iterative application of this network yields predictions $y_m(t+2), y_m(t+3), \ldots, y_m(t+N_2)$, provided that the future inputs $u'(t), u'(t+1)$ are known or assumed.

The first sum in (1) accumulates the squared tracking errors over the horizon $[t+N_1, t+N_2]$; the second penalizes the magnitude of control increments—the difference between the proposed new u' and the last applied u.

When solving the optimization problem, criterion J is minimized with respect to $u'(t), \ldots,$ $u'(t+N_u-1)$; however, in actual control only the first element of the resulting optimal vector is applied.

In the *NNPC* methodology typically $N_1 = 1$ is fixed (tracking error is counted from the very next step). The value $\rho > 0$ is chosen to balance controller responsiveness and control smoothness. The network model is represented as

$$y_m(t+1) = f(u(t), u(t-1), \ldots, y_p(t), y_p(t-1), \ldots), \tag{2}$$

where $y_p(t)$ is the current object's output. For example, the model is described by

$$y_m(t+1) = W^{(2)}\sigma(W^{(1)}[u(t), y_p(t)] + b^{(1)}) + b^{(2)}, \tag{3}$$

where $W^{(1)}$ and $W^{(2)}$ are the weight matrices of the first (input) and second (hidden) layers, $b^{(1)}$ and $b^{(2)}$ are the bias vectors, and σ is the sigmoidal activation function [4, 5].

The operation of this controller type is fundamentally based on solving an optimization problem at each sampling instant. The objective function J incorporates two components: the cumulative output tracking error over the prediction horizon and a penalty on the control increment $\Delta u(t)$. As the solution approaches u', the iterative optimization algorithm (Levenberg–Marquardt method) computes the gradient of J with respect to the control parameters and takes a step toward reducing the error [6].

In the practical implementation, at each step the controller performs a finite number of search iterations, after which it adopts and applies the currently found optimal u to the object. By virtue of the iterative scheme and the local approximation of the optimized function, the method converges to a local minimum of J, but does not guarantee a global optimum.

A neural network predictive controller interacting with the controlled device via the *OPC* (*Open Platform Communications*) protocol can be implemented. Figure 2 shows the respective connection diagram. Figure 3 shows the block diagram of the subsystem with the neural network predictive controller.

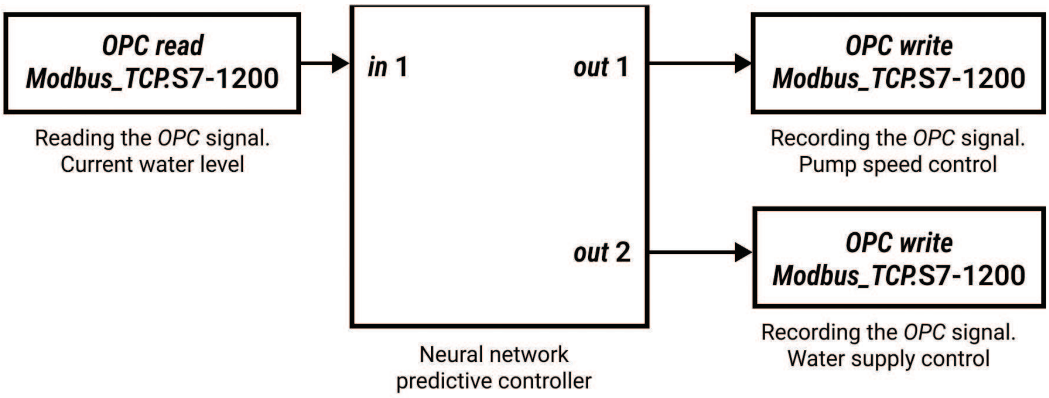

Fig. 2. Connection diagram to the controller via an *OPC* server.

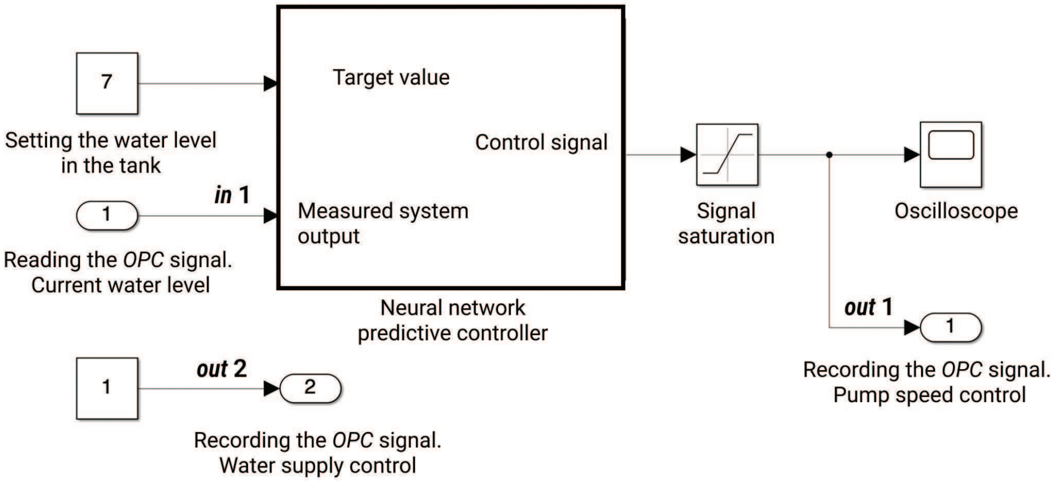

Fig. 3. Subsystem block with the neural network predictive controller.

3. DESIGN AND TUNING OF THE FUZZY QUADRATIC REGULATOR

In 1960, A.M. Letov and R.E. Kalman formulated and successfully solved the linear-quadratic control problem, which consists in finding the optimal control for linear systems using quadratic performance criteria. Letov developed the methodology for time-invariant linear systems, while Kalman extended this approach to non-stationary time-varying systems [7, 8].

The linear quadratic regulator (LQR) implements an optimal control method for linear dynamic systems, based on minimization of the quadratic loss function $e^2(t)$ by use of quadratic criteria. This method aims to minimize a quadratic functional, accounting for the system dynamics and for input and output control parameters, by finding the optimal state feedback gain matrix through the solution of the Riccati differential equation [9].

Consider a linear continuous-time object admitting a vector-matrix state equation representation:

$$\dot{x}(t) = A\,x(t) + B\,u(t), \quad x(t_0) = x_0, \tag{4}$$

where $x \in \mathbb{R}^n$ is the n-dimensional state vector, $A \in \mathbb{R}^{n\times n}$ is the $[n \times n]$ object parameter matrix, $B \in \mathbb{R}^{n\times m}$ is the $[n \times m]$ control matrix with $m \leqslant n$, and $u \in \mathbb{R}^m$ is the control vector.

The state equations of a linear time-invariant system can be written as [9]:

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t), \\ y(t) &= C\,x(t) + D\,u(t), \end{aligned} \tag{5}$$

where A is the $[n \times n]$ system state matrix, B is the $[n \times m]$ input matrix, C is the $[k \times n]$ output matrix, and D is the $[k \times m]$ feedthrough matrix.

The optimality criterion is as follows:

$$J(u(t)) = \min \int_0^\infty [x^T Q\,x + u^T R\,u]\,dt, \tag{6}$$

where matrices Q and R are the weighting matrices for the state variables and control inputs, respectively.

The synthesis objective is to determine the state feedback gain matrix K by minimizing the functional, which reduces to solving the Riccati equation. The Hamiltonian function takes the form

$$H(t, x, u, \lambda) = -\tfrac{1}{2}(x^T Q\,x + u^T R\,u) + \lambda^T (A\,x + B\,u). \tag{7}$$

The gradient conditions can be expressed as

$$\frac{\partial H}{\partial \lambda} = A\,x + B\,u = \dot{x}, \tag{8}$$

$$\frac{\partial H}{\partial x} = A^T \lambda - Q\,x = -\dot{\lambda}, \tag{9}$$

$$\frac{\partial H}{\partial u} = B^T \lambda - R\,u = 0. \tag{10}$$

Using the gradients of the linear form Gz and the quadratic form $z^T S z$, we can write:

$$\frac{\partial}{\partial z}(Gz) = G^T, \qquad \frac{\partial}{\partial z}(z^T S z) = 2\,S\,z,$$

where G is a row vector and S is a symmetric matrix.

From equation (10) it follows that $u = R^{-1} B^T \lambda$. Eliminating the control $u(t)$ from equation (8), we obtain:

$$\dot{x} = A\,x + B\,R^{-1} B^T \lambda, \tag{11}$$

$$\dot{\lambda} = Q\,x + A^T \lambda. \tag{12}$$

To find the optimal control, we solve the system by assuming the existence of a matrix $P(t)$ such that the identity $\lambda(t) = -P(t)\,x(t)$ holds; consequently:

$$u(t) = -R^{-1} B^T P(t)\,x(t). \tag{13}$$

The necessary and sufficient condition for the existence of $P(t)$ is then established, yielding the matrix algebraic Riccati equation for P:

$$PA + A^T P - PBR^{-1} B^T P + Q = 0. \tag{14}$$

Consequently, the linear-quadratic algorithm is a method aimed at minimizing the optimality criterion subject to an equality constraint that incorporates the model of an object. The optimal control law is then set by:

$$u(t) = -K\,x(t), \qquad K = R^{-1} B^T S. \tag{15}$$

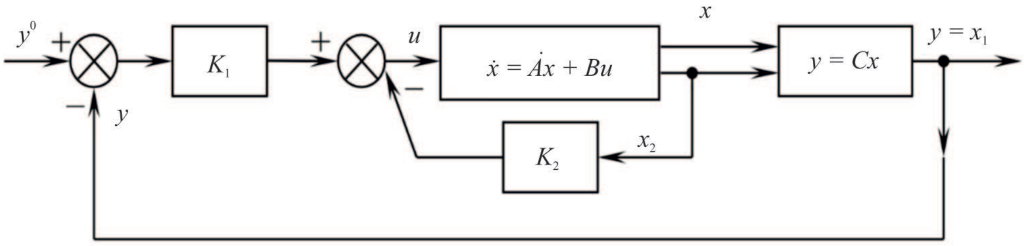

Fig. 4. Control system with the linear-quadratic control algorithm.

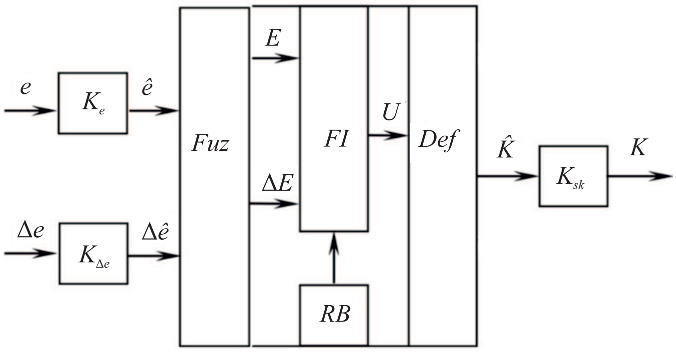

Fig. 5. Structure of the fuzzy quadratic regulator.

It should be noted that the performance criterion can be extended to a finite horizon or to tracking problems (with approximate targeted trajectories); however, in the basic formulation the stabilization-to-equilibrium problem is considered.

The principal advantages of classical LQR, whose schematic representation is shown in Fig. 4, are: guaranteed asymptotic stability of the closed-loop system under full state availability, uniqueness of the solution when the weighting matrices are positive definite, and a systematic (structured) approach to controller synthesis. This control algorithm provides "guaranteed" stability and robustness to small disturbances under an adequate system model. However, the classical LQR has limitations: it requires precise knowledge of the linear object's model and admits only a quadratic problem formulation. In the presence of strong nonlinearities or uncertainties this method may become ineffective; any change in the system structure or parameters requires recomputation of the controller.

To extend the capabilities of the LQR, fuzzy logic-based methods are introduced. The fuzzy quadratic regulator (*FQR*) combines the optimal control law with an adaptive fuzzy inference mechanism [10]. With such approaches the control structure is augmented by a fuzzy logic block that, based on the current system state or tracking error, provides "corrections" to the parameters of the classical controller. For example, a *Linear Fusion Function* may be introduced, combining the LQR output and the fuzzy correction term.

This architecture permits the theoretically grounded properties of optimal control (minimization of the quadratic functional) to be combined with the flexibility of fuzzy rules, providing improved stability, transient response quality, and fault tolerance in the presence of nonlinearities and external disturbances.

In the structure of the nonlinear part of the FQR there are two inputs: the control error $e(k)$ and its rate of change $\Delta e(k) = \Delta e(k)/\Delta t$, and one output K (Fig. 5). Using scaling factors K_e, $K_{\Delta e}$, the actual input values e, Δe are transformed into normalized values $\hat{e}$, $\Delta\hat{e}$ in the interval $[-1, 1]$.

At the fuzzification stage, the normalized input values $\hat{e}$, $\Delta\hat{e}$ are converted into fuzzy variables E, ΔE [1, 11].

The Mamdani fuzzy inference method yields the fuzzy output U from the fuzzy inputs E and ΔE. The defuzzification operation converts the fuzzy output U into a normalized crisp value $\hat{u} \in [L, -L]$ using the centroid method. This normalized value $\hat{u}$ is then multiplied by the scaling factor K_{sk} to obtain the actual value of K.

Adjusting the fuzzy component of the quadratic regulator is a nontrivial task, as it involves both coarse high-level coefficient adjustment and fine low-level tuning to improve control quality.

When constructing a fuzzy model with two inputs and one output, the centroid defuzzification method is typically applied at the defuzzification stage [1]:

$$\hat{y} = \frac{\int_Y y\, Y'(y)\, dy}{\int_Y Y'(y)\, dy}. \tag{16}$$

At the subsequent stage of establishing the input–output relationship of the fuzzy model, the rule base of the fuzzy inference system is applied. Each rule in this base consists of an antecedent and a consequent.

Writing the variables in continuous form: $\hat{e}(t)$, $\Delta\hat{e}(t)$, $u(t)$, $t \in [0, T]$. For a continuous fuzzy controller, the rules take the form:

$$R^\theta\colon \text{if } \hat{e}(t) \text{ is } E^\theta \text{ and } \Delta\hat{e}(t) \text{ is } \Delta E^\theta, \text{ then } u(t) \text{ is } U^\theta, \tag{17}$$

where E^θ, ΔE^θ, and U^θ are fuzzy sets characterizing the error $\hat{e}(t)$, its rate of change $\Delta\hat{e}(t)$, and the control $u(t)$ on $t \in [0, T]$, belonging to the respective linguistic term sets $E \in T_e$, $\Delta E \in T_{\Delta e}$, $U \in T_u$ with elements that are linguistic values.

Based on the analysis of the obtained transient and phase trajectories, the fuzzy rules for the controller are determined. The construction of the rule base for fuzzy quadratic control involves identifying the necessary linguistic variables for the inputs and output, as well as defining the linguistic terms describing the values of these variables.

It should be noted, however, that if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_Z, the current control value is maintained and K equals K_Z:

$$R: \text{if } \hat{e} \text{ is } E\_Z \text{ and } \Delta\hat{e} \text{ is } DE\_Z, \text{ then } K \text{ is } K\_Z. \tag{18}$$

If the error $\hat{e}$ is not specified, the rules should increase the transient response speed at the largest control error.

In the course of the investigation it was established that the fuzzy logic block output requires actions of the same nature but of different magnitudes for each state variable. Consequently, the fuzzy logic block yields a single coefficient y, while multiple scaling factors are required. The number of these factors depends on the model order. Furthermore, since the obtained model is normalized, the scaling factors must be determined.

Table 1 presents the linguistic variables for the fuzzy quadratic regulator. The abbreviations commonly used to denote linguistic values in fuzzy logic are as follows: *NVB* (*Negative Very Big*), *NB* (*Negative Big*), *NM* (*Negative Medium*), *NS* (*Negative Small*), *Z* (*Zero*), *PS* (*Positive Small*), *PM* (*Positive Medium*), *PB* (*Positive Big*), *PVB* (*Positive Very Big*) [1].

Table 1. Fuzzy Rule Selection for Gain K

$\hat{e}\,/\,\Delta\hat{e}$	*NB*	*NS*	*Z*	*PS*	*PB*
NB	*NVB*	*NB*	*NM*	*NS*	*Z*
NS	*NB*	*NM*	*NS*	*Z*	*PS*
Z	*NM*	*NS*	*Z*	*PS*	*PM*
PS	*NS*	*Z*	*PS*	*PM*	*PB*
PB	*Z*	*PS*	*PM*	*PB*	*PVB*

The fuzzy rules derived from Table 1 are listed below. The selection of these rules is governed by the nature of the control action, the error, and its rate of change:

R^1: if $\hat{e}$ is E_NB and $\Delta\hat{e}$ is DE_NB, then K is K_NVB,
R^2: if $\hat{e}$ is E_NB and $\Delta\hat{e}$ is DE_NS, then K is K_NB,
R^3: if $\hat{e}$ is E_NB and $\Delta\hat{e}$ is DE_Z, then K is K_NM,
R^4: if $\hat{e}$ is E_NB and $\Delta\hat{e}$ is DE_PS, then K is K_NS,
R^5: if $\hat{e}$ is E_NB and $\Delta\hat{e}$ is DE_PB, then K is K_Z,
R^6: if $\hat{e}$ is E_NS and $\Delta\hat{e}$ is DE_NB, then K is K_NB,
R^7: if $\hat{e}$ is E_NS and $\Delta\hat{e}$ is DE_NS, then K is K_NM,
R^8: if $\hat{e}$ is E_NS and $\Delta\hat{e}$ is DE_Z, then K is K_NS,
R^9: if $\hat{e}$ is E_NS and $\Delta\hat{e}$ is DE_PS, then K is K_Z,
R^{10}: if $\hat{e}$ is E_NS and $\Delta\hat{e}$ is DE_PB, then K is K_PS,
R^{11}: if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_NB, then K is K_NM,
R^{12}: if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_NS, then K is K_NS,
R^{13}: if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_Z, then K is K_Z,
R^{14}: if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_PS, then K is K_PS,
R^{15}: if $\hat{e}$ is E_Z and $\Delta\hat{e}$ is DE_PB, then K is K_PM,
R^{16}: if $\hat{e}$ is E_PS and $\Delta\hat{e}$ is DE_NB, then K is K_NS,
R^{17}: if $\hat{e}$ is E_PS and $\Delta\hat{e}$ is DE_NS, then K is K_Z,
R^{18}: if $\hat{e}$ is E_PS and $\Delta\hat{e}$ is DE_Z, then K is K_PS,
R^{19}: if $\hat{e}$ is E_PS and $\Delta\hat{e}$ is DE_PS, then K is K_PM,
R^{20}: if $\hat{e}$ is E_PS and $\Delta\hat{e}$ is DE_PB, then K is K_PB,
R^{21}: if $\hat{e}$ is E_PB and $\Delta\hat{e}$ is DE_NB, then K is K_Z,
R^{22}: if $\hat{e}$ is E_PB and $\Delta\hat{e}$ is DE_NS, then K is K_PS,
R^{23}: if $\hat{e}$ is E_PB and $\Delta\hat{e}$ is DE_Z, then K is K_PM,
R^{24}: if $\hat{e}$ is E_PB and $\Delta\hat{e}$ is DE_PS, then K is K_PB,
R^{25}: if $\hat{e}$ is E_PB and $\Delta\hat{e}$ is DE_PB, then K is K_PVB.

Figure 6 shows the block diagram of the closed-loop control system with the FQR in the MATLAB–Simulink environment.

The set point, accounting for the scaling factor, can be expressed as

$$y'(t) = y^0 K_{y0}, \qquad \text{control error } e(t) = y'(t) - y(t),$$

continuous-to-discrete signal conversion for the fuzzy logic block:

$$e_{ZOH}(t) = \mathrm{ZOH}(e(t),\, T),$$

differentiation of the discrete signal and rate of change of the error:

$$de(nT) = e_{ZOH}(nT) - e_{ZOH}(nT - T).$$

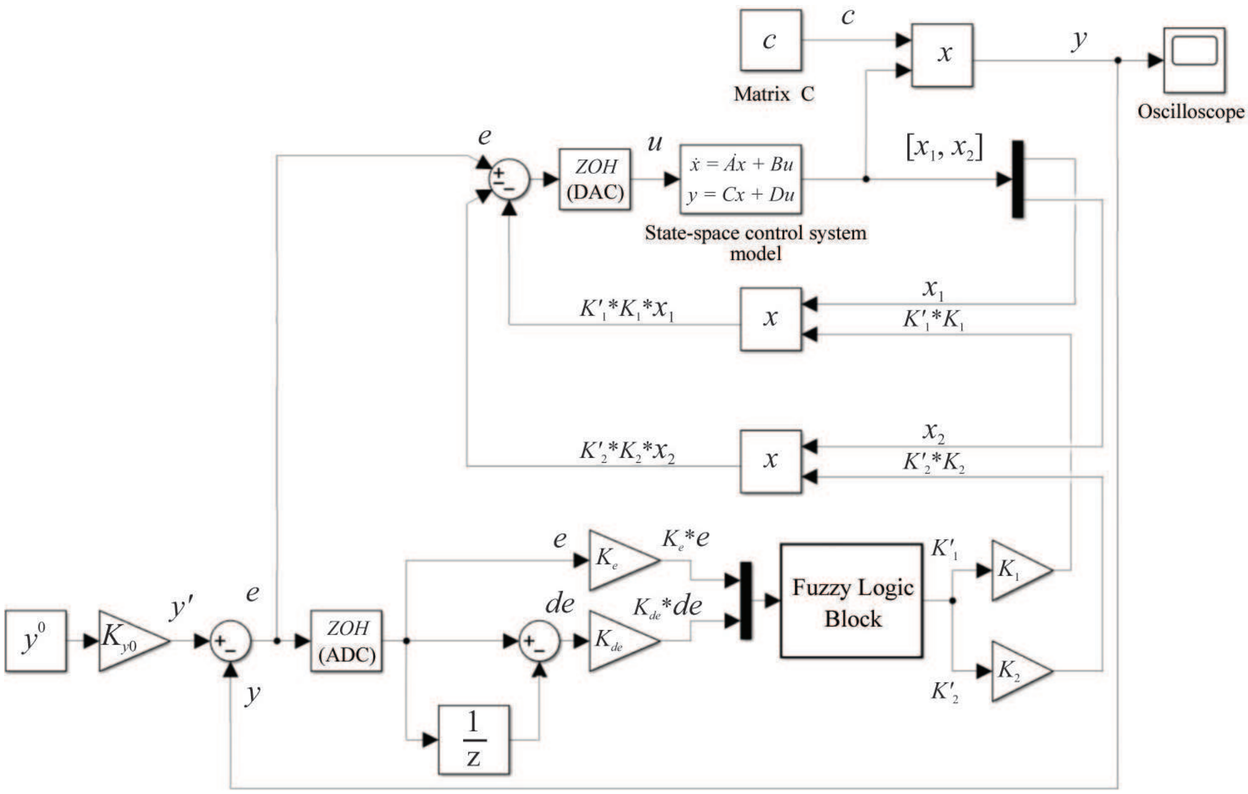

Fig. 6. Block diagram of the closed-loop control system with the fuzzy quadratic regulator.

The fuzzy logic block is described as follows:

$$\begin{bmatrix} K_1' \\ K_2' \end{bmatrix} = \tilde{f}_{\text{FLB}}(K_e\, e_{ZOH},\, K_{\Delta e}\, de). \tag{19}$$

State-space description of the controlled object's model:

$$\dot{X}(t) = A\,X(t) + B\,y(t), \qquad X(t) = C\int_t \dot{X}(t)\,dt.$$

Then

$$X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix},$$

consequently,

$$U(t) = e(t) - X_1(t)\,K_1'(t)\,K_1 - X_2(t)\,K_2'(t)\,K_2.$$

Inverse conversion from discrete to continuous signal:

$$u(t) = \text{ZOH}(U(t)), \qquad y(t) = C\,X(t)$$

is performed by the *Zero-Order Hold* block.

In the fuzzy logic block of the quadratic regulator, the gains K_1' and K_2' are adjusted in accordance with the linguistic rules. This gain adaptation procedure is carried out to optimize controller performance, thereby enhancing the efficiency and accuracy of system control.

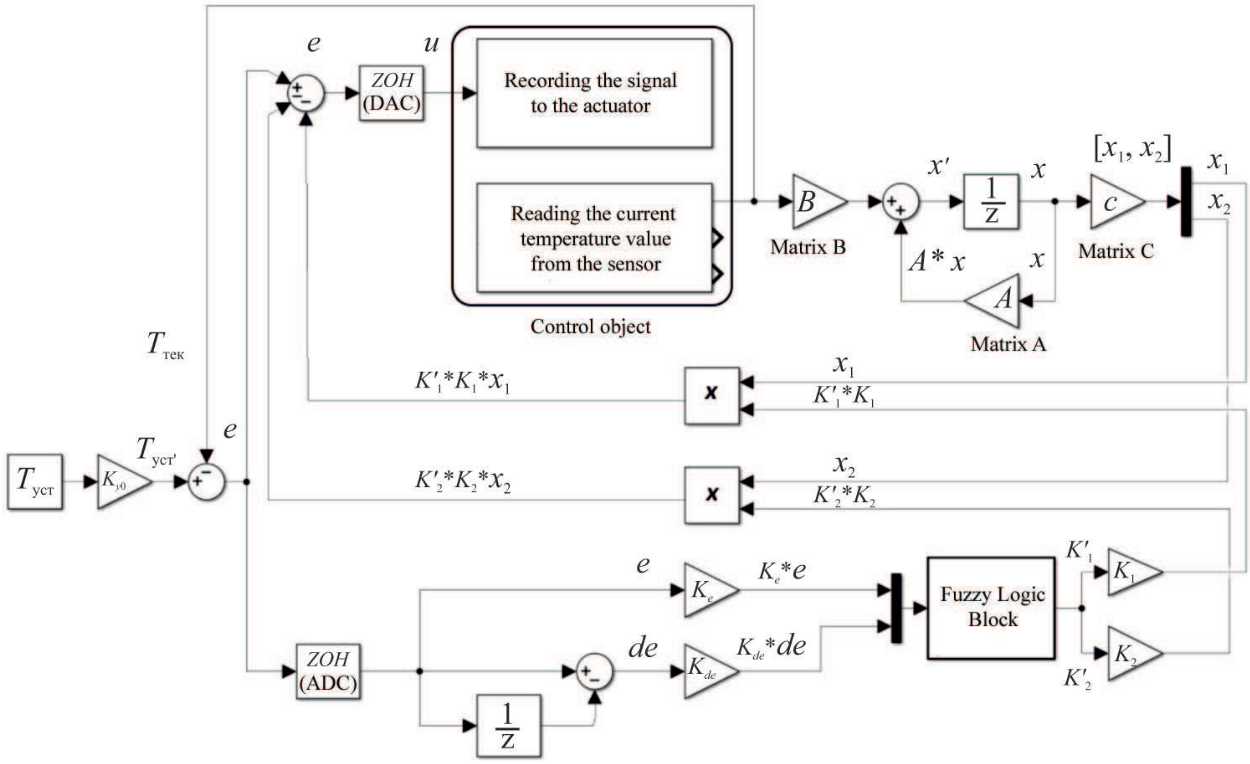

Fig. 7. Block diagram of the closed-loop control system with the fuzzy quadratic regulator and *OPC* server blocks.

Unlike the simulation scheme, in the OPC server scheme when connecting to the controlled device it is necessary to add a state observer block that acquires the object's state vector in real time (Fig. 7).

4. IMPLEMENTATION OF HYBRID CONTROLLERS AT REAL FACILITIES

The paper presents an extended theoretical justification for the application of the neural network predictive controller to control problems involving objects with complex dynamics. It is shown that the use of neural network models enables effective approximation of system behavior over the prediction horizon and, in combination with optimization methods, ensures high control accuracy. The mathematical formalization of the optimality criterion, the implementation of the prediction horizon strategy, and the controller architecture make it possible to consider various constraints, delays, and nonlinear properties of the controlled process.

At present, this type of controller has been successfully applied in research tasks connected with liquid level regulation in a hydraulic plant [12, 13]. MATLAB–Simulink software, a Simatic S7-1200 programmable logic controller, and an OPC server were used in the control implementation.

The MATLAB environment with the Deep Learning Toolbox library enables efficient implementation of complex neural network-based control systems. The *NN Predictive Controller* block (Fig. 3) serves as the built-in NNPC—a controller consisting of the reference input (*reference*), the measured system output input (*feedback*), and the control signal output (*control*).

This block is configured through dialog windows divided into two tabs: *Controller Design* (predictor parameters) and *Plant Identification* (network model parameters). The key block parameters are: prediction and control horizons (N_1, $N_2 = 1$), the weighting coefficient ρ, the optimization

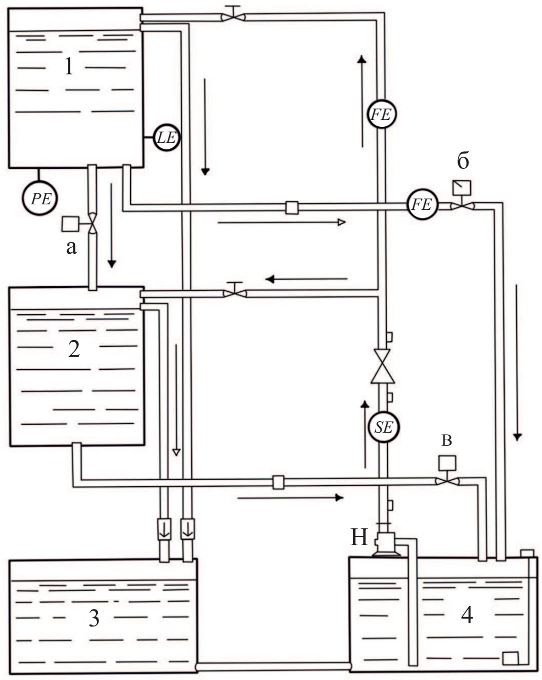

Fig. 8. Research bench "Hydraulic Plant."

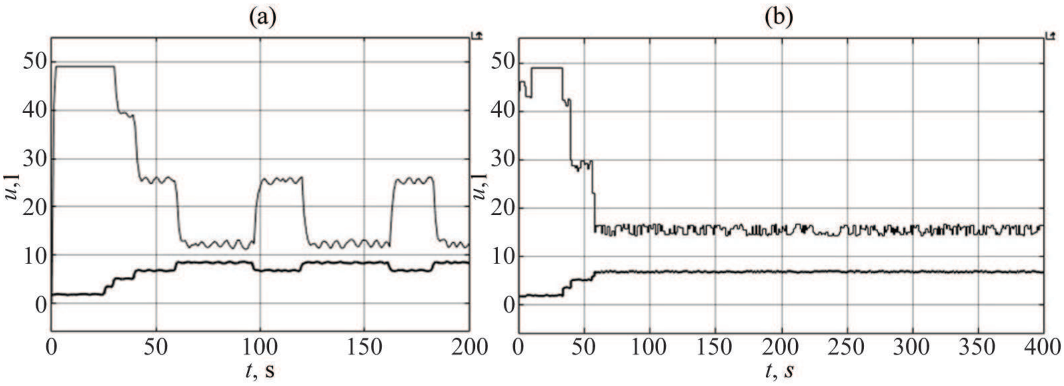

Fig. 9. Transient response using: (a)—classical PI controller, (b)—neural network predictive controller.

parameter α (which determines the required relative reduction of J at each search step), the optimization algorithm, and the number of iterations.

The *Plant Identification* tab specifies the parameters of the trainable object model network: the number of hidden layer neurons, the number of input and output variable delays, the training function, and the training data reading settings.

In the *Simulink Plant Model*, a pre-prepared model with the object's transfer function is selected, from which the neural network obtains a preliminary representation of the object's behavior.

Based on a series of studies, the major objective of which was to maintain the liquid level in tank 2 (Fig. 8) at 7 liters, it is concluded that the level control system based on a classical PI controller is unstable.

This is due to the fact that the parameters of the classical PI controller are unable to compensate for minimal accumulation of excess liquid in the tank, leading to spurious responses to changes (Fig. 9).

The results of the experimental study of the two controller types are presented in Table 2.

Table 2. Experimental Results

Controller type	Rise time, s	Transition time, s	Overshoot, 1	Stability behavior
Neural network predictive controller	54	57	0.04	Stable
PI controller	42	66	0.23	Unstable

It should be noted, however, that neural network controllers may be inferior to fuzzy controllers due to their complexity in training and adjustment, high computational cost, and lower predictability of behavior under disturbances and transient conditions.

Fuzzy controllers, by contrast, provide stability and fast response due to simple logical rules, minimal computation time, and transparency of the control algorithm formulation, making them preferable for tasks with precise stability requirements—in particular, in biotechnical control systems where specific process regulations must be strictly observed.

Table 3 presents the types and parameters of the membership functions of the fuzzy component of the quadratic regulator in the *Fuzzy Logic Toolbox* block of MATLAB–Simulink for the temperature regulation task in an incubation cabinet [9, 10, 14]. The fuzzy control rule base was synthesized using the phase plane method.

Table 3. Membership Functions of the Fuzzy Component of the Quadratic Regulator

Fuzzy variable	Membership function	MF type	Parameters
E	E_NB	*trapmf*	[−1.45 −1.05 −1 −0.493]
	E_NS	*trimf*	[−1 −0.5 0]
	E_Z	*trimf*	[−0.5 0 0.5]
	E_PS	*trimf*	[0 0.5 1]
	E_PB	*trapmf*	[0.506 0.999 1.16 1.34]
dE	DE_NB	*trapmf*	[−1.45 −1.05 −1 −0.493]
	DE_NS	*trimf*	[−1 −0.5 0]
	DE_Z	*trimf*	[−0.5 0 0.5]
	DE_PS	*trimf*	[0 0.5 1]
	DE_PB	*trapmf*	[0.506 0.999 1.16 1.34]
K	K_NVB	*trapmf*	[−1.23 −1.02 −0.987 −0.758]
	K_NB	*trimf*	[−1 −0.75 −0.5]
	K_NM	*trimf*	[−0.75 −0.5 −0.25]
	K_NS	*trimf*	[−0.5 −0.25 0]
	K_Z	*trimf*	[−0.25 0 0.25]
	K_PS	*trimf*	[0.00408 0.254 0.504]
	K_PM	*trimf*	[0.25 0.5 0.75]
	K_PB	*trimf*	[0.5 0.75 1]
	K_PVB	*trapmf*	[0.762 0.999 1.03 1.2]

The initial values of gain coefficients K_1 and K_2 (Figs. 6 and 7) are computed using the function $[K, S, e] = \mathrm{lqr}(W\_ss, q, r)$, which provides the baseline controller adjustment with the specified weighting matrices W_ss, q, and r, establishing the foundation for subsequent fuzzy correction according to the defined linguistic rules.

The conversion of the controlled object transfer function to the state-space representation is carried out using the built-in MATLAB function `tf2ss/W_ss`.

Figure 6 illustrates the formal verification procedure for system properties prior to LQR design using the MATLAB functions `ctrb` and `obsv`. First, the controllability matrix

$$C = [B \quad AB \quad A^2B \; \cdots \; A^{n-1}B]$$

is formed; its rank is determined using `ctrb`. If it equals the dimension n of the state vector, the system is declared fully controllable. Similarly, using `obsv` the observability matrix

$$Q = [C^T \; (CA)^T \; (CA^2)^T \; \cdots \; (CA^{N-1})^T]^T$$

is constructed; when it has full rank, all states can be reconstructed, ensuring correct observer operation. In the single-input case, matrix R reduces to a scalar parameter r, which in the optimizing functional

$$J = \min \int\limits_0^\infty (x^T Q\, x + u^T R\, u)\, dt$$

is assigned empirically and plays a key role in balancing control energy consumption and system responsiveness: increasing r reduces the amplitude of the control signal and improves robustness, but slows the dynamic response, whereas decreasing r makes the controller more aggressive, reducing settling time at the expense of increased risk of overshoot and actuator wear.

After the initial computation of the gain coefficients $K = \mathrm{lqr}(A, B, Q, R)$, controllability and observability are verified; then, through successive simulations and experimental tests, the parameter r is adjusted until the required transient response quality indices (settling time, overshoot) are achieved while full controllability and observability are maintained. This iterative optimization serves as an important link connecting the classical LQR approach with the subsequent fuzzy adaptation, enabling high accuracy and reliability in real operating conditions [9, 10].

Thus, the described approach combines rigid optimization via the classical LQR algorithm at the initialization stage with flexible adaptation via fuzzy logic at the real-time stage, ensuring high control accuracy and robustness over a wide range of operating conditions.

The active experiment on temperature regulation in the incubation cabinet was conducted in two stages: finding the optimal control algorithm and controlling a full incubation cycle. The experiments were carried out to study the effect of external factors on the control processes by fully opening the incubation cabinet door for a period of 1200 s, simulating a short-time presence of personnel inside the cabinet. The air temperature in the incubation cabinet for one hour prior to the experiment was 38 °C. The forced-circulation fan was switched on. The results of an experiment are presented in Table 4.

Table 4. Experimental Results for the Incubation Cabinet Temperature Control

Control type	Overshoot, °C	Transition time, s
	Built-in PI controller	
	1.9	3055
	Optimized PI controller	
	1.5	3000
	Fuzzy quadratic regulator	
	1.0	2730

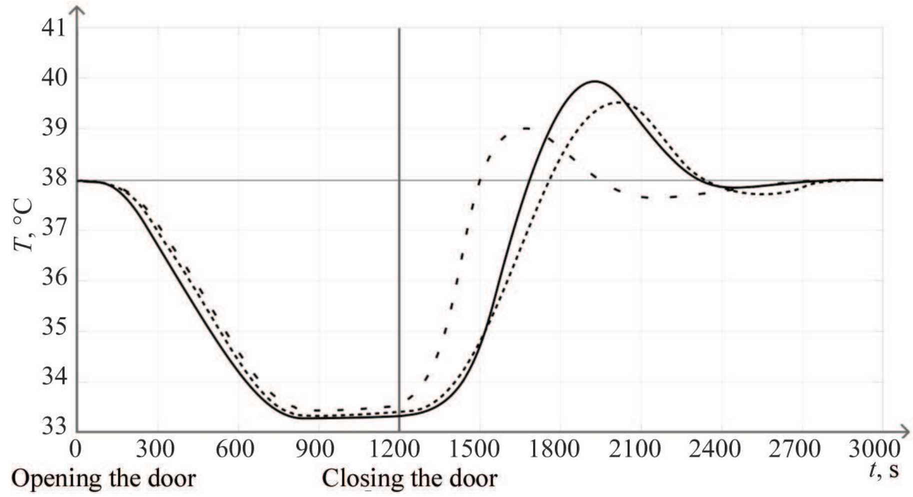

Fig. 10. Effect of external disturbances on the controlled plant in the system with the built-in (solid line), optimized (dense dashed line), and fuzzy quadratic (dashed line) controllers.

Figure 10 shows the response of the controlled object to an external disturbance.

5. CONCLUSION

The choice between a neural network predictive controller and a fuzzy controller should be governed by accuracy requirements, the nature of the object's dynamics, availability of training data and computational resources, and the technological specifics of the application.

Table 5 presents the key decision factors for the selection of a hybrid controller.

Table 5. Key Factors in the Selection of a Hybrid Controller

Characteristic	Neural network predictive controller	Fuzzy controller
Operating principle	Uses an ANN-based plant model; predicts future system behavior and solves the optimization problem via cost functional J.	Control is determined by a set of IF–THEN expert rules; input signals are fuzzified, a fuzzy control action is inferred, and then defuzzified to yield a crisp output value.
Plant model	Identified from data: an ANN is trained to approximate the plant dynamics from experimental data. Capable of representing complex nonlinear dependencies.	No explicit mathematical model; dynamics are defined by rules (described in natural language through linguistic terms).
Interpretability	Low: black-box model; difficult to explain why it produces a particular control action.	High: rules are human-readable and the control logic can be easily modified.
Computational complexity	High: solving an optimization problem in real time is required at each control step. Training can be resource-intensive.	Low: rule evaluation and defuzzification operations are fast; suitable for high-speed controllers.

REFERENCES

1. Pashchenko, F.F., Kudinov, Yu.I., Pashchenko, A.F., Kudinov, I.Yu., Kelina, A.Yu., and Duvanov, E.S., *Nechetkie modeli i systemy upravleniya* (Fuzzy Models and Control Systems), Moscow: Lenand, 2026.

2. Ignatyev, V.V. and Finaev, V.I., The use of hybrid regulator in design of control systems, *World Appl. Sci. J.*, 2013, vol. 23, no. 10, pp. 1291–1297.

3. Cybenko, G.V., Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.*, 1989, vol. 2, no. 4, pp. 303–314.

4. Akesson, B.M. and Toivonen, H.T., A neural network model predictive controller, *J. Process Control*, 2006, vol. 16, no. 9, pp. 937–946.

5. Draeger, A., Engell, S., and Ranke, H., Model predictive control using neural networks, *IEEE Control Syst. Mag.*, 1995, vol. 15, no. 5, pp. 61–66.

6. Vasickaninova, A., Bakosova, M., Oravec, J., et al., Neural network predictive controller design, *Chem. Eng. Trans.*, 2017, vol. 61, pp. 121–126.

7. Letov, A.M., Analytical controller design, *Avtom. Telemekh.*, 1960, vol. 21, no. 5, pp. 436–441.

8. Kalman, R., Falb, P., and Arbib, M., *Topics in Mathematical System Theory*, New York: McGraw-Hill, 1969.

9. Kudinov, Y.I., Pashchenko, F.F., Kelina, A.Y., et al., Analysis of control system models with conventional LQR and fuzzy LQR controller, *Procedia Comput. Sci.*, 2019, vol. 150, pp. 737–742.

10. Pashchenko, F.F., Kudinov, Y.I., Pashchenko, A.F., et al., Fuzzy quadratic control of thermal object, *Proc. 2019 1st Int. Conf. on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA)*, IEEE, 2019, pp. 288–293.

11. Vassilyev, S.N., Kudinov, Yu.I., Pashchenko, F.F., et al., Intelligent control systems and fuzzy controllers II, *Autom. Remote Control*, 2020, vol. 80, pp. 345–357.

12. Duvanov, E.S., Fedyanin, T.V., and Pashchenko, A.F., The feasibility of using a predictive neural network controller, *Proc. 17th Int. Conf. on Management of Large-Scale System Development (MLSD)*, IEEE, 2024, pp. 1–5.

13. Duvanov, E.S., Batishchev, R.V., Pashchenko, A.F., et al., Applying of predictive neural network controllers, *Proc. 4th Int. Conf. on Technology Enhanced Learning in Higher Education (TELE)*, IEEE, 2024, pp. 422–427.

14. Duvanov, E.S., *Analis i sintez gibridnych regulyatorov dlya prediktivnogo upravleniya teplovym protsessom na osnove nechetkoi logiki* (Analysis and Synthesis of Hybrid Controllers for Predictive Control of a Thermal Process Based on Fuzzy Logic), Extended Abstract of Cand. Sci. (Eng.) Dissertation, Lipetsk: LSTU, 2024.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 4, pp. 374–386.

SPECIAL ISSUE

Methods of Construction and Features of Work Schedules for Mechanical Engineering Enterprises

E.N. Khobotov[*,**]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
[**] *Russian Technological University MIREA, Moscow, Russia*
e-mail: e_khobotov@mail.ru

Received July 8, 2025
Revised September 8, 2025
Accepted November 20, 2025

Abstract—This paper considers the challenges and methods of constructing work schedules at mechanical engineering enterprises that manufacture large quantities of products and components for their assembly. The concepts and principles of information aggregation create the development basis of these methods. The considered methods make it possible to create coordinated work schedules for all production divisions of mechanical engineering enterprises. Such schedules make it possible to define the start and end times of processing for each components batch on all enterprise equipment.

Keywords: work schedules, work schedule construction methods, modeling methods, processing equipment, components, product assembly, manufacturing technology, production divisions, aggregation principles

DOI: 10.7868/S1608303226040074

1. INTRODUCTION

Tasks of scheduling theory typically arise when it is possible to choose one or another sequence for performing works [1–3]. The results of their solutions are highly effective and noticeable in industry. The creation and using work schedules created for production systems and production divisions allows for significant reductions in the time required to complete incoming orders without the need for additional investment.

A fairly large number of effective methods exist for constructing such work schedules [1–3]. Unfortunately, these methods turned out to be unsuitable for building work schedules in enterprises with a large number of equipment and the production of significant volumes of products. The unsuitability of these methods is due to the extremely large sizes of the schedules at these enterprises. However, it is precisely for such enterprises that it is possible to obtain the greatest economic benefit from the use of work schedules.

Enterprise efficiency significantly increases when using schedules created for manufacturing of incoming orders on all equipment within the enterprise. Such schedules, referred in [5] to as coordinated, ensure the manufacturing technology for each batch of components on all equipment at the enterprise. Furthermore, such schedules must avoid unforced equipment downtime, and, where possible, minimize forced equipment downtime.

Paper [4] proposes an approach using aggregation principles to create coordinated schedules in enterprises. Papers [5, 6] applied this approach to form work schedules in engineering enterprises with slipway and conveyor assembly of manufactured products. However, in cases where it is necessary to create schedules for processing a large number of components at enterprises with a significant amount of equipment, major problems arise. The methods proposed in [4] and developed in [5, 6] for constructing work schedules in such cases may prove insufficiently effective, and in some cases, completely unsuitable.

This paper proposes methods for constructing work schedules at such mechanical engineering enterprises. The proposed methods for constructing work schedules utilize the ideas and principles of sequential aggregation.

2. PROBLEM STATEMENT

Let us consider the formulation of the problem of manufacturing components in the production divisions of enterprises for the assembly of manufactured products from them.

Let us say that an enterprise has M production divisions. These divisions manufacture components for assembly L types of manufactured products. The enterprise manufactures products in batches, the sizes of which are N_l $(l = 1, \ldots, L)$.

The quantity and types of components, assemblies, and parts are known for each manufactured product. Furthermore, the production time and sequence of each batch of components for any type of manufactured product are known, as are the changeover times for the equipment required for their production. The enterprise manufactures products assembled from components, and units manufactured within the enterprise's production divisions, as well as from components purchased from third-party manufacturers.

The task requires the creation for enterprise such work schedule, in which any component it is necessary to manufacture in accordance with its technological route. Furthermore, it is desirable to complete the incoming order according to this schedule in the shortest possible time.

3. PRINCIPLES OF CONSTRUCTING SCHEDULES FOR PROCESSING COMPONENTS AT ENTERPRISES

Let us consider the principles and specifics of developing schedules for the production of components at mechanical engineering enterprises.

At discrete-production enterprises, it is advisable to create work schedules for all equipment within the enterprise. Only such work schedules will ensure the execution of component manufacturing technology at the enterprise and can reduce the lead-time for incoming orders.

As already noted, it is impossible to create work schedules at enterprises using existing methods, since the work schedule sizes, even at small enterprises, are too large for these methods.

In [4], the author proposed an approach to developing methods suitable for scheduling at mechanical engineering enterprises. The idea of such approach, which utilizes information aggregation principles, is to form special groups of components. Along with these component groups, we form equipment groups for processing of such components.

We note that the formation of equipment groups does not imply changes in the administrative structure of enterprises and necessary only for scheduling work. However, it is convenient to select enterprise production divisions as equipment groups. Each component group includes only those components that, in accordance with their manufacturing technology, necessary to deliver to the same production divisions in the same order.

However, within the enterprise production divisions, components in each group are processed using their own technological routes of processing. We can consider of components groups formed in this way as "generalized parts," and the equipment groups using for their processing as "generalized machines."

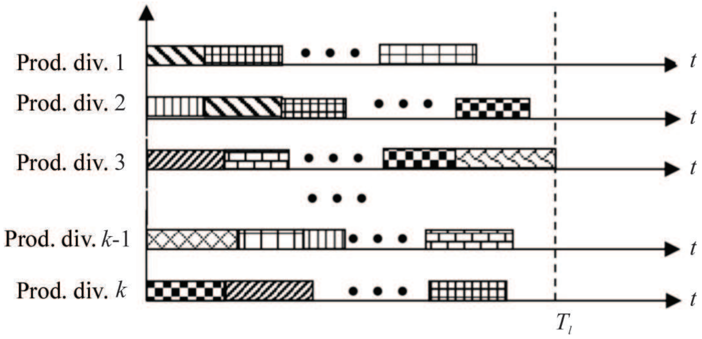

Fig. 1. Gantt chart of the "framework" schedule for processing the formed component groups of the Jth product.

Indeed, groups of components are delivered to the production divisions of enterprises for processing in accordance with their manufacturing technology, just as parts undergoing machining are delivered to the machining equipment.

For each manufactured part, the machining sequence is known on the equipment of the production division its machined. For each group of components, the processing sequence in the enterprise divisions is also known, according to the principles for forming such groups. For each manufactured part, the processing times are known or normative of its processing times are set for all equipment used. The processing time for a group of components in any division that processes we can to determine by creating a processing schedule of such group in that division. The processing time for a group of components in a division is the completion time of processing the last component of that group on the division's equipment.

It should be noted that in the future we will use the schedules for processing groups of components in divisions to generate schedules for the production of components at enterprises.

In [4], the author called the production schedule for groups of components in the enterprise's divisions a "framework" schedule.

The dimensionality of such a schedule is significantly smaller than that of the original schedule. This is because in the "framework" schedule, instead of several dozen pieces of equipment in each division, only one unit is considered—a "generalized machine" or enterprise division. Instead of several dozen pieces of components, only one unit is considered—a "generalized part" or group of components.

Therefore, such a schedule it is possible to construct using traditional scheduling methods if, after aggregation, the dimensionality of the problem of constructing a "framework" schedule is not too large for these methods.

Parts processing schedules in production systems and areas it is possible to represent using Gantt charts [1–3]. "Wireframe" schedules also it is possible to represent using these charts. Figure 1 shows a Gantt chart representing an example of a "wireframe" schedule for processing component groups for the lth product in an enterprise's production divisions.

The Gantt chart for such a schedule, as shown in Fig. 1, has virtually the same appearance as the Gantt chart for a parts processing schedule in a manufacturing division [1–3]. However, in a "wireframe" Gantt chart, instead of machines, the ordinate axis represents the enterprise's production divisions (production systems and areas, shops). Furthermore, in such a chart, the abscissa axis represents the processing times of groups of parts in the corresponding divisions, instead of the time of processing parts.

In "framework" schedules, just as with batch-based component production, it is possible to reduce forced downtime in production divisions by transferring some of the processed batches from

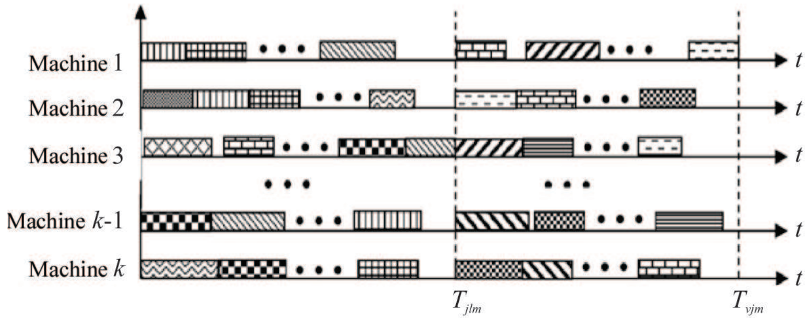

Fig. 2. Gantt chart of the processing of the ith and vth adjacent groups of components of the Jth product in the mth production division.

the processed group to the division next the technological route for processing this components group if if the division is already available.

The resulting "framework" schedule allows for the rapid creation of a preliminary, agreed-upon work schedule for the enterprise. The creation of this schedule, which will be referred to hereafter as an "expanded framework" or "unpacked framework" schedule, is conveniently explained by constructing a Gantt chart for it. A Gantt chart designed for an "expanded framework" or "unpacked framework" schedule has a separate x-axis for each equipment unit in the enterprise. Each axis represents the processing times of components manufactured on the equipment in the division corresponding to this x-axis.

It's worth remembering here that processing schedules for all component groups have already been created for the divisions where these groups are to be processed according to their manufacturing technology, since they determined their processing times for the "framework" schedule.

Therefore, an "expanded framework" or "unpacked framework" schedule is created by placing the previously created processing schedules for component groups in that division in the Gantt chart at the positions of each divisions and on the axes corresponding to the same equipment. The created processing schedules for component groups are placed at the positions of each division in the same order as was obtained when creating the "framework" schedule. The processing schedule for the next component group in each division can be placed on the "expanded framework" schedule diagram only after the processing of the last component from the previous group is completed.

Therefore, the completion times for the "expanded framework" and "framework" schedules will coincide. An "expanded framework" schedule generated using this scheme will be the agreed-upon schedule at the enterprise level. This is true because in this schedule, processing of each component begins on any equipment used by the enterprise only when its processing is completed on the preceding equipment along the manufacturing route for that component.

However, such a work schedule contains many "unforced" equipment downtimes during the processing of adjacent groups of components, making this schedule not entirely successful. This can be seen by plotting a Gantt chart containing a fragment of such a schedule (Fig. 2), which shows the processing schedule for the ith and vth adjacent groups of components of the lth product in the mth division.

Equipment downtime in such schedules arises due to the rules for constructing "framework" schedules. According to these rules, processing of the next group of components in a division can begin only after the last component from the previous group has been processed, and the equipment freed up from processing components from the previous group will remain idle. Therefore, the work duration obtained using such a schedule will be increased due to these downtimes, and using it as a schedule for the entire enterprise, as already noted, is impractical.

Such downtime it is possible to reduce by performing a "schedule gluing" operation. After this operation, equipment freed up from processing components of the current group in each production division is immediately used to process components of the next group, if possible.

As noted, the dimensionality of the "framework" schedule will be significantly smaller than that of the original schedule. Therefore, in many cases, such a schedule, unlike the original schedule, t is possible to construct using existing methods in a reasonable time.

Furthermore, as noted earlier, an "extended framework" schedule it is possible to construct based on a "framework" schedule, which is a coordinated schedule at the enterprise level. After "merging" the "extended framework" schedule, it is possible to convert into a coordinated schedule for the enterprise, defining a sequence that does not disrupt the processing technology of components across all equipment within the enterprise.

Here, we assumed that all the necessary work schedules it is possible to construct, since we were examining the use of aggregation concepts to construct work schedules at mechanical engineering enterprises.

However, in some cases, especially when schedules need to be constructed at enterprises with a large number of production division and equipment within them, as well as those processing a large number of components, significant problems may arise. One of these problems is the excessively large dimensionality of the resulting work schedules, which can lead to an unacceptably long time for scheduling.

4. BUILDING LARGE-SCALE PROCESSING SCHEDULES

Let us consider the principles and specifics of creating large-scale work schedules that arise at enterprises with a large number of production divisions and equipment in them, as well as at enterprises that process a large number of components.

The dimensionality of work schedules at such enterprises, as already noted, are extremely large. Therefore, existing methods turns out to be unsuitable for constructing such work schedules. It is necessary to develop special methods using the ideas and principles of the "framework" schedule construction method described above. These methods provide the ability to perform additional iterations when the dimensionality of the "framework" schedule remains too large. Each of these iterations involves a new aggregation step.

At each aggregation stage, the proposed method selects production divisions from the previous stage, combines them, and forms groups of components for processing in these divisions. It is advisable to combine between two and seven to ten production divisions. The production division formed as a result of combining the production divisions of the previous stage and the groups of components formed for processing in them receive the number of the current stage. We will discuss the formation of component groups in more detail in the next section.

After the formation of divisions, at the new stage of aggregation it is necessary to form groups of components, whose processing must occur in these divisions. Each group of components at the new stage includes groups of components from the previous stage, which, in accordance with the production technology, it is necessary to produce in the same divisions at the new stage.

Therefore, the number of component groups in a new stage cannot exceed the number of component groups in the previous stage. If some component groups in a new stage include even a few component groups from the previous stage, the number of component groups in the new stage will be significantly smaller than the number of component groups in the previous stage. Furthermore, due to the consolidation of production divisions that occurs at each aggregation stage, their number in the new stage will be smaller than in the previous stage.

Therefore, the dimension of the "framework" schedule for processing groups of components in production divisions after the aggregation stage can only be smaller than the dimension of the "framework" schedule before this stage, and in many cases, significantly smaller. During the implementation of the aggregation stages, it is advisable to select as combined production divisions the production divisions of the next-level enterprise that include several production divisions from the previous stage to be combined. For example, for production divisions such as production areas, the next-level production divisions are the production shops containing several production divisions to be combined.

For more clearly and conveniently description of the approach to constructing "framework" schedules of large dimensionality, it is convenient to introduce the following notation. We will call the initially formed production divisions or equipment groups Stage 1 production divisions or equipment groups. The groups of components processed in these production divisions or equipment groups we call Stage 1 component groups.

The production divisions formed in the second aggregation stage will be referred to as Stage 2 production divisions. The groups of components processed in these production divisions will be referred to as Stage 2 component groups. Each Stage 2 component group is formed according to the same rule as a Stage 1 component group. According to this rule, all components in a Stage 2 component group are delivered in the same order for processing to the same Stage 2 production divisions or equipment groups.

By analogy, the production divisions formed at the nth stage of aggregation will be called production divisions of the nth stage, and each group of components formed at the nth stage will be called a component group of the nth stage. Each group of components of the nth stage is formed according to the same rules as the group of components of the 1st, 2nd, and 3rd stages. The production divisions at each aggregation stage, starting with the second, are formed from the production divisions of the previous stage. The component groups at each stage of aggregation, starting with the second, are also formed from the component groups of the previous stage.

The principles of creating and using such methods it is conveniently examined in the following example, where the large dimensionality of the work schedule being constructed at an enterprise leads to insurmountable difficulties.

Furthermore, although schedules for processing component groups in the enterprise's production divisions can be built using existing methods in an acceptable time, the dimensionality of the "framework" schedule turns out to be too large, and building it using existing methods require an unacceptably long time.

After merging several production divisions and forming groups of components for processing in these divisions, it is advisable to evaluate the possibility of creating a "framework" schedule at the enterprise with using of existed methods. This assessment it is possible to make based on the number of production divisions and component groups, which will be known after their formation.

It should be noted here that a preliminary assessment of the possibility of constructing a "framework" schedule is extremely important and should be used when making decisions about scheduling. Existing scheduling methods allow for the rapid and reliable creation of good work schedules [1–3] when the number of parts and equipment units is not significant. As the number of parts and equipment units increases, the time required to create a work schedule will increase. However, starting from certain values of these quantities, the increase in schedule creation time will be almost exponential. Creating a schedule for large values of these quantities may require an unacceptably long time.

For each scheduling method, the values at which schedule generation time begins to increase sharply will vary. Understanding and estimating the time required to obtain a completed schedule during the scheduling process is very difficult. Therefore, using existing methods without preliminary scheduling feasibility assessments can lead to significant time expenditures, and very often, completing the required schedule within an acceptable timeframe is impossible.

Let us continue our consideration of an example in which schedules can be generated and processing times for all component groups determined in the divisions of the enterprise where they are processed. However, the extremely large dimensionality of the "framework" schedule precludes its construction using existing methods.

In this case, it is advisable to perform the next aggregation stage, which will be the second one. At this stage, the production divisions of the first stage are combined and the production divisions of the second stage are formed. Then, the component groups of the first stage are used to form the component groups of the second stage. The process of forming component groups will be described in more detail in the next section.

After forming the component groups, the feasibility of constructing a "framework" schedule is checked, since the dimensionality of the "framework" schedule after each aggregation stage can only decrease. If the check does not confirm the feasibility of constructing a "framework" schedule, the next aggregation stage is performed. Otherwise, a "framework" schedule is constructed, and after its construction, an "expanded framework" schedule is generated. Based on this schedule, after the "gluing" operation, a schedule is generated that can be used for planning and managing the enterprise's work.

For determination of the processing time of the formed group of components stage 1 in the division of the 1st stage it is possible to build schedule of processing such components group in this division. However, at the stages of aggregation, starting from the second stage, it is proposed to first build "framework" schedules in order to schedule the work of groups of components in the formed divisions. The following circumstances determine the feasibility of constructing a "framework" schedule.

First, existing methods may be unsuitable for constructing schedules for processing components from formed groups in production divisions of a new stage, consisting of many production divisions from the previous stage. With "framework" schedules, it is possible to construct schedules of very large dimensions.

Second, schedules have already been constructed for all groups of components from the previous stage, and their processing times are known in the production divisions of the previous stage where they are processed.

It is hardly practical to repeatedly construct schedules for the same components in the same production divisions. Constructing "framework" schedules does not require re-constructing existing schedules. In cases where schedule fragments have already been built for some of the components processed in a production division, these schedule fragments and their execution times are used to build "framework" schedules and are not built anew.

The feasibility of constructing a "framework" schedule is determined by the following circumstances. First, existing methods may be unsuitable for constructing schedules for processing components from formed groups in production divisions of a new stage, consisting of many production divisions from the previous stage. Using "framework" schedules, it is possible to construct schedules of very large dimensions. Second, schedules have already been constructed for all groups of components from the previous stage, and their processing times are known in the production divisions of the previous stage where they are processed.

It is hardly practical to repeatedly construct schedules for the same components in the same production divisions. Constructing "framework" schedules does not require re-constructing existing schedules. In cases where schedule fragments have already been built for some of the components processed in a production division, these schedule fragments and their execution times are used to build "framework" schedules and are not built anew.

Indeed, at each stage of aggregation, beginning with the second, new divisions and component groups are formed for that stage.

Such divisions, as already noted, are formed by combining divisions from the previous stage, the composition of which is known for each formed division.

Each component group for a new stage also consists of component groups from the previous stage, which are processed in the same order in the same divisions of the current stage. The composition of these groups from the previous stage is also known for each formed component group for that stage. For each component group from the previous stage, the processing times and sequence in the divisions of the previous stage are known

Therefore, to determine the processing time for a formed component group for a new stage in the formed division of the new stage in which this group is processed, a “framework” schedule can be constructed.

This “framework” schedule is a schedule for processing groups of components from the previous stage, which are included in the formed group of components of the new stage, in the production divisions of the previous stage, combined into this division of the new stage.

After constructing the “framework” schedule, an “expanded framework” schedule it is possible to generate based on it using the scheme described above. From this “framework” schedule, a reasonably good schedule of work in the corresponding division is obtained using the “gluing” operation. The time of completion of processing of the last component on the division’s equipment is selected as the component processing time in this division.

Furthermore, it is possible existing methods to used for construction of “framework” schedules in established divisions, since it is desirable to combine a very limited number of divisions from the previous stage when forming divisions.

As already noted, after each aggregation stage, the number of component groups cannot increase, which, as the number of divisions decreases, leads to a reduction in the dimensionality of the “framework” schedule. Therefore, after completing several aggregation stages, the number of component groups and the number of established divisions can be reduced so much that it will be possible to construct a “framework” schedule for virtually any mechanical engineering enterprise using existing methods.

Based on the constructed “framework” schedule, as already noted, without any principals difficulties it is possible to obtain a fairly good works schedule at any enterprise.

This schedule is suitable for managing and planning the enterprise’s operations, and will also allow for determining the processing sequence and start and end times for all components on any equipment at the enterprise.

If, after completing the aggregation step, the resulting estimates indicate that creating a “framework” schedule may be difficult or unacceptably time-consuming, the next aggregation step is performed. This step is performed similarly to the previous aggregation step described above.

In this situation, the obtained estimates indicate that creating work schedules in these divisions using existing methods may be difficult. This situation typically arises from the poor selection of certain divisions with excessive equipment. Furthermore, such a situation may arise due to an excessive number of components for which processing schedules must be created, as well as due to the simultaneous presence of these factors.

When this situation arises, it is advisable to divide the equipment of each division for which creating a work schedule requires excessive time. The number of subgroups into which such divisions t is necessary to divide hat equipment of the same type is not allocated to different subgroups of the divided division. Otherwise, problems will arise when forming subgroups of components sent to these subgroups of the divided division for processing.

In such situations, it is necessary to divide the equipment of each division, for which scheduling work requires excessive time, into subgroups. It is advisable to choose the number and composition of such subgroups for this division so that equipment of the same type falls into only one subgroup, and not into different ones. Otherwise, problems will arise when forming subgroups of components that are sent to these subgroups for processing. Let us remind you that the formation of equipment groups, as already noted, does not entail changes in the administrative structure of enterprises and is only necessary for building work schedules

In addition, it is advisable to select these subgroups in such a way that, using existing methods, it is possible to create schedules for processing the corresponding subgroups of components within an acceptable time frame.

For each subgroup of equipment from a divided division, it is necessary to form subgroups of components that this subgroup of equipment should handle. These subgroups of components are necessary to form according to the rule described earlier from components directed to the divided division for processing. According to this rule, all components in each subgroup must be sent for processing in the same order to these subgroups of equipment and to other divisions of the enterprise. Furthermore, these subgroups must include all components sent for processing to the divided division.

After this, it is necessary to create a schedule for processing these subgroups of components. As before, it is advisable to verify the feasibility of creating such schedules in the formed subgroups using the boundary values of the methods employed, as described above. In cases where the boundary values are not exceeded, it is advisable to create such schedules using the existing methods and determine the processing completion times from them.

If all these schedules can be created, then the possibility of creating a "framework" schedule to process the subgroups of components in the subgroups of the divided division is verified.

If there are no exceedances of the boundary values, it is advisable to construct such a "framework" schedule. If the "framework" schedule is successfully constructed, an "expanded framework" schedule is created from it. After the "gluing" operation, a coordinated schedule of activities in the split division can be obtained from the "expanded" schedule, and from received schedule the completion times of these activities are determined. If any subgroups of the split division exceed the boundary values when processing the corresponding subgroups of components, these subgroups are divided into smaller subgroups, and the procedure described above is repeated for them.

If, when checking the feasibility of constructing a "framework" schedule for processing subgroups of components in the split division using existing methods, it turns out that difficulties may arise, then It is advisable the next aggregation step to performe according to the scheme described above.

5. PRINCIPLES AND FEATURES OF FORMING COMPONENT GROUPS

Let us briefly review the basic principles and features of component groups formation.

It is convenient to examine at first these principles using the example of forming first-level component groups. The formation of these component groups begins with the preliminary distribution of components by levels and their numbering. The level number assigned to each distributed component corresponds to the number of divisions it manufactured. The first level is assigned to components processed in a single division. The last k_1th level is assigned to components processed in the largest number of divisions within the enterprise. For forming component groups, they are numbered sequentially by component level, starting with the last k_1th level and ending with the components of the lowest level.

It is convenient to examine at first these principles using the example of forming first-level component groups. The formation of these component groups begins with the preliminary distri-

bution of components by levels and their numbering. The level number assigned to each distributed component corresponds to the number of divisions it manufactured. The first level is assigned to components processed in a single division. The last k_1th level is assigned to components processed in the largest number of divisions within the enterprise. For forming component groups, they are numbered sequentially by component level, starting with the last k_1th level and ending with the components of the lowest level.

Component groups are typically formed by selecting a base component for the first group. The first base component for the first group being formed is typically the component of the last k_1th level, which is assigned the first number in the numbering process. Then, the next component of the last k_1th level is selected as the component to be inspected. After this, a comparison is made between the sequence of the enterprise divisions where the component being inspected and the base component of the group being formed are processed.

If the sequences of delivery for processing to the enterprise divisions of the base and inspected components match, then the component being inspected can be included in the group defined by this base component.

If these sequences differ, the sequences of delivery for processing to divisions of the base component and the next component to be inspected by number are compared. In this case, the previously inspected component will be the base component for the next group of components. Then, the first base component is compared with the next component to be inspected by number. If the sequence of delivery for processing to divisions of these components is the same, then the component being inspected is included in the group defined by this base component.

Otherwise, the sequence of delivery for processing to divisions of this inspected component it is necessary to compare with the sequence of delivery for processing for the base component.

If these sequences match, the inspected component It is advisable to include in the group of components defined by this base component, and to select the next component for inspecting.

The sequence of delivery for processing to divisions of each inspected component it is necessary to compare one by one with the sequence of delivery for processing to divisions of previously received base components.

If the delivery sequences for processing to divisions of the inspected component match the sequence of delivery of any of the previously received base components, the inspected component may be include in the group defined by this base component. Otherwise, the inspected component will be the base component for the next group.

This process continues until completion of the distribution into the formed groups of components of the k_1th level.

After completion of the distribution of level components into the formed groups, the distribution of components into the created and new groups of levels $(k_1 - 1), (k_1 - 2), \ldots,$ 2nd, and 1st components may be carry out almost in a similar manner. However, the sequence of delivery of components of lower levels for processing to the enterprise's divisions cannot coincide with the analogous sequences of basic components, since the latter are of a higher level. Therefore, it is advisable to consider the following options for including the inspected component in the formed group of components.

One option for including the inspected component in the formed group of components is that the sequence of its processing production divisions is a fragment in the similar sequence of the basic component.

A fragment in the sequence of divisions of a basic component here is understood to be the part of this sequence from the division with which the fragment begins to the division that ends it, without skipping divisions in the sequence. Another possibility for including the inspected component in the component group is that the sequence of delivery for processing to the enterprise divisions of the tested component is a subsequence of the similar sequence of the basic component.

One option for including the inspected component in the formed group of components is that the sequence of its processing production divisions is a fragment in the similar sequence of the basic component.

A fragment in the sequence of divisions of a basic component here is understood to be the part of this sequence from the division with which the fragment begins to the division that ends it, without skipping divisions in the sequence.

Another possibility for including the inspected component in the component group is that the sequence of delivery for processing to the enterprise divisions of the inspected component is a subsequence of the similar sequence of the basic component.

In this case, between divisions may be the divisions in which it is necessary to process the inspected and the main components, and divisions in which it is necessary to process the main components, but the inspected components does not.

When using the first principle of forming component groups, a larger number of such groups may be create. This can complicate the construction of a "framework" schedule for component processing at enterprises. In some cases, such difficulties can be quite significant.

Using the second principle of forming component groups, it is possible to obtain a smaller number of component groups, which simplifies the creation of a "framework" schedule for component processing at enterprises.

However, in this case, both when creating work schedules and during the manufacturing process, it is necessary to monitor components in groups that do not need to process in all divisions along the processing route for the basic components of these groups.

It is advisable to send such components to a warehouse if they do not require processing in the next division along the processing route for the group, and if they need to deliver in a timely manner from the warehouse to the divisions where they should continue processing as part of the group.

A very useful approach is to first apply the first principle of forming component groups. Then, if a very large number of component groups is obtained, combine some of them to reduce their number to an acceptable level. It is advisable to combine groups using the basic components of these groups.

A very useful approach is to first apply the first principle of forming component groups. Then, if it turned out that a number of component groups is very large, you should combine some of them to reduce their number to an acceptable level. It is advisable to combine groups using the basic components of these groups.

Furthermore, when forming component groups, it is advisable to take into account that many components, especially those processed in a small number of divisions, it is possible to include in different groups.

If such components to include in groups where it is necessary to process they in the divisions that equipment are most loaded with processing components already included in these groups, the processing time of the groups after their inclusion may increase significantly.

Let us consider the specifics of forming component groups in the second and subsequent stages of aggregation.

Component groups for a new aggregation stage are formed after the divisions in which they will be processed have been formed, and are formed from component groups from the previous stage.

However, at a new aggregation stage, due to the merger of several divisions, the level of the basic components of the previous stage and the component groups they define may change significantly.

Such a change may occur, for example, if a basic component of one group was processed in divisions of a certain stage, which, after aggregation, were included in different divisions of the next stage. Whereas a basic component of another group was processed in a larger number of divisions of the same stage, which were included in groups in a significantly smaller number of divisions or even in a single division of the next stage.

Therefore, these components, to form component groups for the new stage, it is necessary to distributed among new levels, taking into account the composition and number of divisions at that stage, and numbered.

In this case, as before, each level corresponds to the number of divisions at that stage in which the basic components distributed to that level are manufactured. The first level includes component groups processed in a single division at that stage. The last k_rth level, where r indicates the aggregation stage, includes components processed in the largest number of enterprise divisions formed at that aggregation stage.

The numbering of basic components begins with components at the last k_rth level, sequentially proceeds through lower component levels, and ends with components of the lowest level. Then, in accordance with the scheme described above, component groups for the new stage are formed.

6. SOME FEATURES AND PROPERTIES OF "FRAMEWORK" SCHEDULES AT MECHANICAL ENGINEERING ENTERPRISES

Let us consider the purpose and main properties of "framework" schedules for processing groups of components in mechanical engineering enterprise divisions.

Existing methods, as already noted, turned out unsuitable for creating work schedules at enterprises, especially those with a large amount of processing equipment and processed components.

Creating "framework" schedules allows us to solve the problems of creating work schedules at such enterprises.

The fact is that, according to the rules for drawing up "framework" schedules, the components of each formed group it is necessary to deliver for processing to divisions in accordance with the technological route of the group processing.

However, in each division along the processing route, the components of a group are processed according to its own technological processing route. Therefore, before starting processing in any division along the technological route of the group, all components of the group it is necessary preliminarily to process according to their technological processing route.

Furthermore, the dimensionality of "framework" schedules is significantly smaller than that of the original schedules, since in "framework" schedules, instead of several dozen units of equipment in each division, a single unit is considered—the division itself. Instead of dozens of units of components, only one unit is considered—the group of components processed in that division. Therefore, in many cases, such schedules, unlike original schedules, it is possible to construct using existing methods in a reasonable time.

"Framework" schedules have a very important property, allowing them to use quickly and easily generate a preliminary version of a work schedule for enterprises or large divisions.

Previously, the formation of such a schedule was described, which was called an "extended structure" or "unpacked structure" schedule. After the schedule "gluing" operation, this schedule can become a very good coordinated schedule for enterprises or large divisions. A coordinated schedule here, as already noted, is understood to mean a schedule that does not disrupt technology of component manufacturing.

The time required to complete work on such a "glued" schedule is significantly reduced compared to the time required to complete work on schedules prior to the "gluing" operation.

"Glued" schedules prove to be quite suitable for planning work at the enterprise and in its divisions, for obtaining substantiated information on the lead times of incoming orders, for organizing the work of the enterprise's support services, etc.

It should also be noted that using "framework" schedules, it is possible to quickly obtain quite good upper estimates for order fulfillment times.

For construction "framework" schedules, as already noted, it is necessary to determine the processing times of component groups in the enterprise divisions processed these groups. These times it is possible to determine from the schedules for processing component groups in the corresponding enterprise divisions. Such schedules it is possible to construct independently of each other and in any order. Therefore, it is possible to parallelize their construction. Such parallelization will significantly reduce the time required to determine the processing times of component groups in enterprise divisions, which will result in a reduction in the time required to construct the "framework" schedule.

Furthermore, in the section summarizing the main properties of "framework" schedules, it is necessary to highlight the properties already mentioned in this work.

"Framework" schedules, like regular part processing schedules in production systems and areas, it is possible to represent using Gantt charts.

In the "framework" schedules, some of the already processed batches of components from the processing group can be transferred to an exempt division for processing, if this division is the next in the technological processing route of this group.

7. CONCLUSION

This paper proposes a method that utilizes sequential aggregation principles to create very large-scale work schedules for mechanical engineering enterprises. The proposed method allows for relatively simple parallelization of computations, significantly reducing schedules creation time.

Creating work schedules for enterprises with large amounts of equipment and processing a significant number of components will allow:

—to utilize the capabilities of scheduling theory methods to improve the enterprises work efficiency and reduce lead times for incoming orders;

—to create coordinated work plans and schedules that define the start and end times of all operations on all equipment of each enterprises.

REFERENCES

1. Pinedo, M.L., *Scheduling. Theory, Algorithms, and Systems*, Sixth Edition, Leipzig: Springer, 2022.
2. Bruker, P., *Scheduling Algorithms*, Leipzig: Springer, 2007.
3. Lazarev, A.A., *Theory of schedules. Methods and algorithms*, Moscow: ICS RAS, 2019.
4. Khobotov, E.N., On some models and methods of the solution of scheduling problems in discrete enterprises, *Autom. Remote Control*, 2007, vol. 68, no. 12, pp. 2172–2186.
5. Khobotov, E.N., Models of Equipment Selection for Modernization of Enterprises with Conveyor Assembly of Products, *Autom. Remote Control*, 2024, no. 10, pp. 906–918.
6. Khobotov, E.N., Problems and methods of work scheduling for enterprises with slipway assembly of manufactured products, *Avtomat. i Telemekh.*, 2023, no. 12, pp. 80–95.

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