Volume 87, Number 6,
June 2026

ISSN 0005-1179
CODEN: AURCAT

A JOURNAL OF RUSSIAN ACADEMY OF SCIENCES

# AUTOMATION AND REMOTE CONTROL

Automation and Remote Control

Vol. 87, No. 6, June 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. 6, 2026**

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 483–493.


**LINEAR SYSTEMS**

# Anisotropy–Based Estimation Design for Linear Discrete Time–Invariant Systems with Measurement Dropouts

**A. V. Yurchenkov**[*,a] **and R. R. Ryakhimov**[*,b]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
*e-mail:* [a]*alexander.yurchenkov@yandex.ru,* [b]*ryakhimov.rinat@gmail.com*

Received April 29, 2025
Revised October 30, 2025
Accepted November 5, 2025

**Abstract**—This paper considers a time-invariant system with a set of imperfect measurements, the useful component of which may be unavailable at certain times. Virtual objects duplicating the state of the original object are introduced, and an output estimate is constructed for each of them. Using vectorization, an extended model of the system is derived in the form of a linear discrete time-invariant system with multiplicative noises. The external disturbance is chosen from the class of sequences of random vectors with a bounded level of mean anisotropy. For the system describing the estimation error dynamics, conditions for the boundedness of the anisotropic norm are obtained, under which the chosen type of estimator exists. An invertible linearizing change of variables is indicated, which allows reducing the problem of finding the estimator matrices to checking the solvability of a special system of linear matrix inequalities with a convex constraint.

*Keywords*: anisotropy-based theory, systems with multiplicative noise, sensor network with dropouts, convex optimization

**DOI:** 10.7868/S1608303226060016

## 1. INTRODUCTION

The problems of external disturbances rejection considered in control theory were formulated back in the first half of the last century [1] and gradually transitioned from deterministic [2] and bounded disturbances [3] to stochastic ones [4, 5], interest in which persists to this day [6]. It is worth noting that the most frequently described classes of disturbances are Gaussian disturbances [7] or square-integrable (summable in the case of discrete time) [8]. The disturbance rejection problems associated with these classes, with an $\mathcal{H}_2$- or $\mathcal{H}_\infty$-optimal criterion, have both advantages and disadvantages. On the one hand, with known stochastic characteristics of the disturbance, $\mathcal{H}_2$-optimal estimators and controllers have proven themselves well; however, solutions to such problems do not possess the robustness properties inherent in $\mathcal{H}_\infty$-optimal control theory. On the other hand, $\mathcal{H}_\infty$-optimal controllers introduce excessive conservatism into the dynamics of a system with weakly colored disturbances at the input.

Anisotropy-based control theory proposes to view the problem of suppressing the influence of an external disturbance as a stochastic modification of $\mathcal{H}_\infty$-optimal theory, since the external disturbance is chosen from the class of random, rather than square-summable, sequences [9–11]. This expands the scope of anisotropy-based theory compared to $\mathcal{H}_\infty$-theory, since for deterministic and norm-bounded disturbances, the mean anisotropy level will tend to infinity, and the limiting value of the anisotropic norm will tend to the $\mathcal{H}_\infty$-norm. For random external disturbances, anisotropy-based theory will give results between $\mathcal{H}_2$- and $\mathcal{H}_\infty$-theories in terms of the cost function, depending on the value of the information criterion. The results obtained over the thirty years of existence of anisotropy-based theory cover the problems of robust stability analysis of time-invariant and time-varying systems, control synthesis, and filtering [11–13]. Later, works appeared devoted to systems whose state cannot be described by deterministic equations [14–16]. In particular, systems with multiplicative noises are a convenient form of describing the dynamics for systems representing mechanical, biological, ecological, financial, and hybrid objects [17]. Within the framework of this work, the problem of constructing an output estimate for a system with dropouts will be considered. An imperfect measurement is understood as such a measurable output of the system that, at an arbitrary time, may not contain useful information about the object of observation, but may be completely random. The case of random measurement is called dropout. It turns out that a system with one or several (possibly duplicating each other) imperfect measurable outputs can be described by difference equations with multiplicative noises, for which some results for time-varying systems have already been obtained in anisotropy-based theory.

Anisotropy-based theory allows combining the features of both $\mathcal{H}_2$- and $H_\infty$-optimal control theories, and in limiting cases even provide identical results, being in this respect a generalizing theory. Therefore, when constructing an output estimate or generating control, it makes sense to choose the anisotropy-based approach in the case where the stochastic characteristics of the external disturbance are not specified. If the expectation and covariance matrix of the disturbance are known, then the solution obtained within the framework of anisotropy-based theory is generally less conservative, since it is not designed for the so-called "worst case" of disturbance.

The paper is organized as follows: Section 2 contains preliminary information—the necessary minimum of concepts from anisotropy-based theory; Section 3 presents the problem statement; Section 4 derives the main result; Section 5 presents simulation results; Section 6 provides a brief conclusion.

## 2. PRELIMINARY INFORMATION

We list the main definitions from anisotropy-based theory that directly relate to the problem under consideration. More detailed material can be found in [18, 19].

### *2.1. Anisotropy of a Random Vector*

**Definition 1** [19]. The anisotropy of a random vector $w$ with values in $\mathbb{R}^m$ is the value equal to

$$\mathbf{A}(w) = \min_{\lambda>0} \mathbf{D}(f||p_{m,\lambda}) = \frac{m}{2}\ln\left(\frac{2\pi e}{m}\mathbf{E}[|w|^2]\right) - h(w),$$

where $f$ is the probability density function of the vector $w$ with respect to the Lebesgue measure in $\mathbb{R}^m$, $h(w) = -\mathbf{E}[\ln f(w)]$ is the differential entropy, $\lambda$ is a positive parameter defining the density

$$p_{m,\lambda}(x) = (2\pi\lambda)^{-m/2}\exp\left(-\frac{|x|^2}{2\lambda}\right), \quad x \in \mathbb{R}^m,$$

of the isotropic Gaussian distribution in $\mathbb{R}^m$ with zero mean and scalar covariance matrix $\lambda I_m$, where $I_m$ is the identity matrix of order $m$, and $\mathbf{E}[\cdot]$ is the expectation operator.

This definition is related to relative entropy and characterizes the minimum measure of difference between a random vector and the set of centered Gaussian vectors with a scalar covariance matrix.

### *2.2. Mean Anisotropy of a Sequence of Random Vectors*

**Definition 2** [9]. For a stationary ergodic sequence $W = \{w_k\}$ of Gaussian vectors, the concept of mean anisotropy is introduced, equal to the value

$$\overline{\mathbf{A}}(W) = \lim_{N\to\infty} \frac{\mathbf{A}(W_{0:N-1})}{N},$$

where $W_{0:N-1} = [w_0^{\mathrm{T}}, \ldots, w_{N-1}^{\mathrm{T}}]^{\mathrm{T}}$ is the extended vector.

The mean anisotropy level is used to describe the properties of an external disturbance within the framework of uncertainty in stochastic parameters when analyzing the robust stability of linear discrete time-invariant systems, since the extended vector of any distribution different from the reference one a centered Gaussian with a scalar covariance matrix will have an increasing anisotropy as the time horizon increases.

### *2.3. Anisotropic Norm of a Time-Invariant System*

Consider a linear system $F$ with input $W \in \mathcal{L}_2^m$ and output $Z \in \mathcal{L}_2^p$, where $\mathcal{L}_2^*$ denotes the Lebesgue space of vectors with finite norm. The sequence $W$ is generated by a shaping filter $G$ from standard Gaussian white noise $V$:

$$w_j = \sum_{k=0}^{\infty} g_k v_{j-k}, \quad j \in \mathbb{Z},$$

where $g_k \in \mathbb{R}^{m\times m}$. Let $G(z)$ denote the transfer function of the former filter:

$$G(z) = \sum_{k=0}^{\infty} g_k z^k, \quad |z| < 1, \quad z \in \mathbb{C}.$$

Let the set of former filters generating sequences of random vectors with bounded mean anisotropy be denoted as follows [20]:

$$\mathcal{G}_a = \{G \in \mathcal{H}_2^{m\times m} : W = GV,\ \overline{\mathbf{A}}(W) \leqslant a\},$$

where $V = \{v_k\}_{k\in\mathbb{Z}}$ is standard Gaussian white noise with zero mean and unit covariance matrix.

**Definition 3** [10]. The anisotropic norm $|\!|\!| F |\!|\!|_a$ of a linear system $F$ is the quantity

$$|\!|\!| F |\!|\!|_a = \sup_{G\in\mathcal{G}_a} \frac{\|FG\|_2}{\|G\|_2},$$

where $\|G\|_2 = \left(\sum_{k=0}^{\infty} \mathrm{tr}(g_k g_k^{\mathrm{T}})\right)^{1/2}$ is the $\mathcal{H}_2$-norm of the transfer function $G(z)$.

## 3. PROBLEM STATEMENT

Consider a linear discrete time-invariant system with a set of imperfect sensors

$$\begin{aligned} x_{k+1} &= Ax_k + B_w w_k, \\ z_k &= C_z x_k + D_z w_k, \\ y_{j,k} &= \lambda_{j,k} C_j x_k + D_j w_k, \quad j = \overline{1,n}, \end{aligned} \tag{1}$$

where

$$P(\lambda_{j,k} = 1) = p_j, \quad j = \overline{1,n},$$

characterize the probability that the current measurement $y_{j,k}$ contains information about the state vector. In the case $\lambda_{j,k} = 1$, the measured output contains useful data; otherwise, $\lambda_{j,k} = 0$, it contains only the disturbance vector $D_j w_k$. The matrices $A, B_w, C_z, D_z, C_j, D_j,\ j = \overline{1,n}$, of system (1) are known and have dimensions consistent with the vectors $x_k,\ w_k,\ z_k,\ y_{j,k}$. The external disturbance is a stationary Gaussian sequence of random vectors $W = \{w_k\}_{k\geqslant 0}$, for which a constraint on the mean anisotropy level of the entire sequence is known: $\overline{\mathbf{A}}(W) \leqslant a,\ a \geqslant 0$. It is assumed that all elements of the sequence $W$ and the random variables $\lambda_{j,k}$ are mutually independent.

*Problem 1.* For the chosen estimator model

$$\begin{aligned} \hat{x}_{j,k+1} &= \sum_{i=1}^{n} \mathfrak{a}_{ji} \left(A\hat{x}_{i,k} + H_{ji}(y_{i,k} - \hat{y}_{i,k})\right), \\ \hat{z}_{j,k} &= \sum_{i=1}^{n} \mathfrak{a}_{ji} C_z \hat{x}_{i,k}, \\ \hat{y}_{j,k} &= p_j C_j \hat{x}_{j,k}, \quad j = \overline{1,n}, \end{aligned} \tag{2}$$

where $\mathfrak{a} \in \mathbb{R}^{n\times n}$ is the adjacency matrix characterizing the information exchange between sensors, it is required to find such values of the parameters $H_{ji}$ for which the anisotropic norm of the system $F$ from the external disturbance to the estimation error vector is bounded by a minimum value $\gamma$:

$$|\!|\!| F |\!|\!|_a \leqslant \gamma.$$

It should be noted here that the adjacency matrix $\mathfrak{a}$ is not necessarily symmetric. This matrix describes a directed graph whose vertices are the sensors (detectors/sensors or any available measurements), and the directed edges correspond to data exchange between the vertices connected by these edges. As a result, the estimate can be constructed not based on each individual sensor, but by collecting a consensus variant of the estimate, where the contribution of each sensor is accounted for using a correction coefficient of the matrix $\mathfrak{a}$. Thus, the coefficients of the adjacency matrix have a simple physical meaning, which lies in the level of "trust" in the information coming from other sensors to the one forming the estimate. A natural requirement is the element-wise non-negativity and the equality to one of the sum of elements in each row:

$$\sum_{i=1}^{n} \mathfrak{a}_{ji} = 1, \quad \forall j = \overline{1,n}.$$

## 4. MAIN RESULT

This section provides a solution to Problem 1. The idea is to transform the object model into the form of a system with multiplicative noises and then check a sufficient condition for the boundedness of the anisotropic norm of such a system in terms of solving a convex optimization problem.

### *4.1. Construction of Models in Terms of Estimation Errors*

At the first step, we define virtual models as follows: for each existing estimated output $y_{j,k}$, we introduce a state $x_{j,k} \in \mathbb{R}^{n_x}$, duplicating the state $x_k$ of the object (1):

$$\begin{aligned} x_{j,k+1} &= Ax_{j,k} + B_w w_k, \\ z_{j,k} &= C_z x_{j,k} + D_z w_k, \\ y_{j,k} &= \lambda_{j,k} C_j x_{j,k} + D_j w_k, \end{aligned} \tag{3}$$

where $j = \overline{1,n}$. In models of the form (1), each object has only one measured output. For the states $x_{j,k}$ and outputs $z_{j,k}$, we introduce error vectors

$$\tilde{x}_{j,k} = x_{j,k} - \hat{x}_{j,k},$$
$$\tilde{z}_{j,k} = z_{j,k} - \hat{z}_{j,k},$$

which will have the following dynamics equations:

$$\begin{aligned}\tilde{x}_{j,k+1} = Ax_{j,k} - \sum_{i=1}^{n} \mathfrak{a}_{ji} A\hat{x}_{i,k} + \left(B_w - \sum_{i=1}^{n} \mathfrak{a}_{ji} H_{ji} D_i\right) w_k \\ - \sum_{i=1}^{n} \mathfrak{a}_{ji} H_{ji}(\lambda_{i,k} - p_i) C_i x_{i,k} + \sum_{i=1}^{n} \mathfrak{a}_{ji}(A - p_i H_{ji} C_i)\tilde{x}_{i,k}, \\ \tilde{z}_{j,k} = C_z\left(x_{j,k} - \sum_{i=1}^{n} \mathfrak{a}_{ji} x_{i,k}\right) + \sum_{i=1}^{n} \mathfrak{a}_{ji} C_z \tilde{x}_{i,k} + D_z w_k.\end{aligned}$$

### *4.2. Transformation of the Model to a System with Multiplicative Noises*

Let us compose extended vectors of state, state errors, and estimates from all vectors $x_{j,k}$, $\tilde{x}_{j,k}$, $\tilde{z}_{j,k}$, $j = \overline{1,n}$:

$$\overline{x}_k = (x_{1,k}^{\mathrm{T}}, \ldots, x_{n,k}^{\mathrm{T}})^{\mathrm{T}},$$
$$\tilde{x}_k = (\tilde{x}_{1,k}^{\mathrm{T}}, \ldots, \tilde{x}_{n,k}^{\mathrm{T}})^{\mathrm{T}},$$
$$\tilde{z}_k = (\tilde{z}_{1,k}^{\mathrm{T}}, \ldots, \tilde{z}_{n,k}^{\mathrm{T}})^{\mathrm{T}}.$$

Define the combined state from the vectors $\overline{x}_k$ and $\tilde{x}_k$:

$$\zeta_k = (\overline{x}_k^{\mathrm{T}}, \tilde{x}_k^{\mathrm{T}})^{\mathrm{T}}.$$

Now it is possible to write the dynamics equation in terms of errors depending on the external disturbance:

$$\begin{aligned}\zeta_{k+1} &= \left(\mathcal{A}_0 + \sum_{i=1}^{n} \xi_{i,k}\mathcal{A}_i\right)\zeta_k + \mathcal{B}w_k, \\ \tilde{z}_k &= \mathcal{M}\zeta_k + \mathcal{N}w_k,\end{aligned} \tag{4}$$

where the following notations are adopted:

$$\xi_{i,k} = \lambda_{i,k} - p_i, \quad i = \overline{1,n},$$

and the matrices included in expressions (4) have the form:

$$\mathcal{A}_0 = \begin{bmatrix} \overline{A} & 0_{nn_x \times nn_x} \\ \overline{A} - \overline{A}^{\mathfrak{a}} & \overline{A}^{\mathfrak{a}} - \overline{H}^{\mathfrak{a}}\overline{C}^{p} \end{bmatrix}, \quad \mathcal{A}_i = \begin{bmatrix} 0_{nn_x \times nn_x} & 0_{nn_x \times nn_x} \\ \overline{H}^{\mathfrak{a}}\overline{C}_i & 0_{nn_x \times nn_x} \end{bmatrix},$$
$$\mathcal{B} = \begin{bmatrix} \overline{B} \\ \overline{B} - \overline{H}^{\mathfrak{a}}\overline{D} \end{bmatrix},$$
$$\mathcal{M} = \begin{bmatrix} \overline{M} - \overline{M}^{\mathfrak{a}} & \overline{M}^{\mathfrak{a}} \end{bmatrix}, \quad \mathcal{N} = \overline{N},$$

with notations

$$\begin{aligned}
\overline{A} &= I_n \otimes A, & \overline{A}^{\mathfrak{a}} &= \mathfrak{a} \otimes A,\\
\overline{B} &= \mathrm{col}(n) \otimes B, & \overline{H}^{\mathfrak{a}} &= \underset{j,i=\overline{1,n}}{\mathrm{block}}(\mathfrak{a}_{ji} H_{ji}),\\
\overline{C}^p &= \underset{i=\overline{1,n}}{\mathrm{diag}}(p_i C_i), & \overline{C}_j &= \underset{i=\overline{1,n}}{\mathrm{diag}}(\delta_{ij} C_i),\\
\overline{D} &= \underset{j=\overline{1,n}}{\mathrm{col}}(D_j), & \overline{M} &= I_n \otimes C_z,\\
\overline{M}^{\mathfrak{a}} &= \mathfrak{a} \otimes C_z, & \overline{N} &= \mathrm{col}(n) \otimes D_z,
\end{aligned}$$

where $\mathrm{col}(n)$ denotes the column vector of length $n$ consisting of ones; $\underset{j=\overline{1,n}}{\mathrm{col}}(X_j)$ is the matrix obtained by arranging the matrices $X_j$, $j = \overline{1,n}$, into a column; the matrix $\underset{j,i=\overline{1,n}}{\mathrm{block}}(X_{ji})$ is a block matrix consisting of the corresponding blocks; $\underset{i=\overline{1,n}}{\mathrm{diag}}(X_i)$ is a diagonal matrix with the matrices $X_i$, $i = \overline{1,n}$, as blocks on the main diagonal, and all off-diagonal blocks are zero; $\otimes$ denotes the Kronecker product.

### *4.3. Condition for Boundedness of the Anisotropic Norm*

The following theorem provides sufficient conditions for the boundedness of the anisotropic norm for system (4).

**Theorem 1** [15]. *For a system $F$ of the form* (4) *with additional condition*

$$\lim_{k\to\infty} \rho\left(\left(E\left[\mathcal{A}^k\right]\right)^{\frac{1}{k}}\right) < 1,$$

*where $\mathcal{A} = \mathcal{A}_0 + \sum_{i=1}^{n} \xi_{i,k}\mathcal{A}_i$ given a known constraint $a$ on the mean anisotropy level of the external disturbance sequence $\{w_k\}$, the anisotropic norm will be bounded by $\gamma$ if there exists a solution with respect to the matrices $R$, $S$ and the scalar parameter $q \in \left[0, \|F\|_\infty^{-2}\right)$, $\sigma_i = p_i(1-p_i)$ to the system of inequalities*

$$R \succ \sum_{i=0}^{n} \sigma_i^2 \mathcal{A}_i^{\mathrm{T}} R \mathcal{A}_i + q\mathcal{M}^{\mathrm{T}}\mathcal{M} + L^{\mathrm{T}} S^{-1} L, \tag{5}$$

$$\begin{aligned}
S &= (I_{m_w} - q\mathcal{N}^{\mathrm{T}}\mathcal{N} - \mathcal{B}^{\mathrm{T}} R \mathcal{B})^{-1}, \\
L &= S(\mathcal{B}^{\mathrm{T}} R \mathcal{A} + q\mathcal{N}^{\top}\mathcal{M})
\end{aligned} \tag{6}$$

*subject to the convex constraint*

$$-\frac{1}{2}\ln\det\left((1-q\gamma^2)S\right) \geqslant a, \tag{7}$$

*there the expression $X \succ 0$ should be understood in the sense of positive definiteness of the matrix $X$.*

Theorem 1 contains sufficient conditions for the boundedness of the anisotropic norm of a system with multiplicative noises (4) in the form of inequalities. However, inequality (5) is nonlinear due to the dependence of the matrices $\mathcal{B}, \mathcal{A}_i, i = \overline{1,n}$, on the estimator matrices $H_{ji}, j, i = \overline{1,n}$, and their multiplication by the matrix $R$ of the solution of this inequality. To get away from this problem and remain within the constraints in the form of linear matrix inequalities, we formulate the solution to the problem as conditions for the solvability of a system of inequalities of a special form.

**Theorem 2.** *Suppose that for system* (1) *a constraint a is known on the mean anisotropy level of the disturbance sequence* $W = \{w_k\}_{k\geqslant 0}$. *An estimator of the form* (2) *will guarantee an upper bound by* $\gamma$ *for the anisotropic norm of the estimation error system if the following system of inequalities*

$$\begin{bmatrix} R-\mathcal{M}^{\mathrm{T}}\mathcal{M} & * & * & * & * & \cdots & * \\ \mathcal{N}^{\mathrm{T}}\mathcal{M} & \eta I_{m_w}-\mathcal{N}^{\mathrm{T}}\mathcal{N} & * & * & * & \cdots & * \\ R\mathcal{A}_{00}+X\mathcal{A}_{01} & -R\mathcal{B}_{00}-X\mathcal{B}_{01} & R & * & * & \cdots & * \\ \sigma_1 X\mathcal{A}_{11} & 0 & 0 & R & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_n X\mathcal{A}_{1n} & 0 & 0 & 0 & 0 & \cdots & R \end{bmatrix} \succ 0, \tag{8}$$

$$\begin{bmatrix} \eta I_{m_w}-\Psi-\mathcal{N}^{\mathrm{T}}\mathcal{N} & * \\ R\mathcal{B}_{00}+X\mathcal{B}_{01} & R \end{bmatrix} \succ 0, \tag{9}$$

$$\ln\det\Psi \geqslant 2a + m_w\ln(\eta-\gamma^2), \tag{10}$$

*has a solution with respect to the matrices* $R = R^{\mathrm{T}} \succ 0$, $X \in \mathbb{R}^{4nn_x\times 2np_y}$, *and the scalar parameter* $\eta > 0$, *where* $\sigma_i = p_i(1-p_i)$,

$$\mathcal{A}_{00} = \begin{bmatrix} \overline{A} & 0 \\ \overline{A}-\overline{A}^{\mathfrak{a}} & \overline{A}^{\mathfrak{a}} \end{bmatrix}, \quad \mathcal{A}_{01} = \begin{bmatrix} 0 & 0 \\ 0 & -\overline{C}^{p} \end{bmatrix},$$

$$\mathcal{A}_{j1} = \begin{bmatrix} 0 & 0 \\ -\overline{C}_j & 0 \end{bmatrix}, \quad j = \overline{1,n},$$

$$\mathcal{B}_{00} = \begin{bmatrix} \overline{B} \\ \overline{B} \end{bmatrix}, \ \mathcal{B}_{01} = \begin{bmatrix} 0 \\ -\overline{N} \end{bmatrix}.$$

The proof of Theorem 2 is given in Appendix.

It should be mentioned here that among the entire set of solution matrices $R$ of system (8)–(10), we will choose only block-diagonal ones, since this will be necessary for the numerical solution of the problem using standard tools.

### *4.4. Calculation of the Estimator Matrix Parameters*

In the event of a successful solution of the matrix inequalities (8)–(10), the set of estimator matrices $H_{ji}, j, i = \overline{1,n}$, can be calculated according to the inverse substitution

$$\overline{H}^{\mathfrak{a}} = [0_{2nn_x\times 2nn_x} \ \ I_{2nn_x}]R^{-1}X \begin{bmatrix} 0_{2np_y\times 2np_y} \\ I_{2np_y} \end{bmatrix}. \tag{11}$$

The problem of finding the minimum value of the anisotropic norm bound $\gamma$ can be formulated as a convex optimization problem:

$$\gamma^2 \underset{(8)-(10),\mathcal{R},\Psi,\eta,\gamma^2}{\longrightarrow} \min. \tag{12}$$

Direct calculation of the matrices $H_{ji}, j, i = \overline{1,n}$, is not necessary, since the spatial implementation of the estimator (2) depends on the blocks of matrix (11). The solution of the convex optimization problem (12) will correspond to the minimum constraint on the anisotropic norm of the estimation error system.

## 5. SIMULATION

As an example, consider a model consisting of a quadrotor aerial vehicle fixed on a gyroscopic stand. The object can rotate freely around three axes. It is necessary to estimate its roll angle using the gyroscope and camera available on the object, located frontally in front of the object. We assume that the object is closed by stabilizing control based on an LQG controller. The discrete model of such a system, according to the dynamics equations (1), in the state space has the following form:

$$A = \begin{bmatrix} 0.9988 & 0.0000 & 0.0000 & 0.0095 & 0.0000 & 0.0000 \\ 0.0000 & 0.9999 & 0.0000 & 0.0000 & 0.0099 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0100 \\ -0.2411 & 0.0000 & 0.0000 & 0.8926 & 0.0000 & 0.0000 \\ 0.0000 & -0.0141 & 0.0000 & 0.0000 & 0.9763 & 0.0000 \\ 0.0000 & 0.0000 & -0.0009 & 0.0000 & 0.0000 & 0.9957 \end{bmatrix}, \quad B_w = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \\ 1.0028 & 0.0000 & 0.0000 \\ 0.0000 & 0.0027 & 0.0000 \\ 0.0000 & 0.0000 & 0.0008 \end{bmatrix},$$

$$C_z = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad D_z = 0, \quad C_1 = \begin{bmatrix} I_3 & 0_{3\times 3} \end{bmatrix}, \quad C_2 = C_z, \quad D_1 = D_2 = 0.05 \cdot I_3,$$

and parameters $p_1 = 0.9$, $p_2 = 0.95$. The first measured output corresponds to data from the onboard gyroscope, the second from the frontal camera. It is assumed that the angle values are recovered from the angular velocity measurements by integration; at the initial time, the estimates are considered zero.

Figure 1 shows the simulation results for the mean anisotropy level equal to 30. The solid line corresponds to the estimated output. Estimates 1 and 2 correspond to the case where the readings from the other measured output were used with a weighting coefficient of 0.2; in this case, the

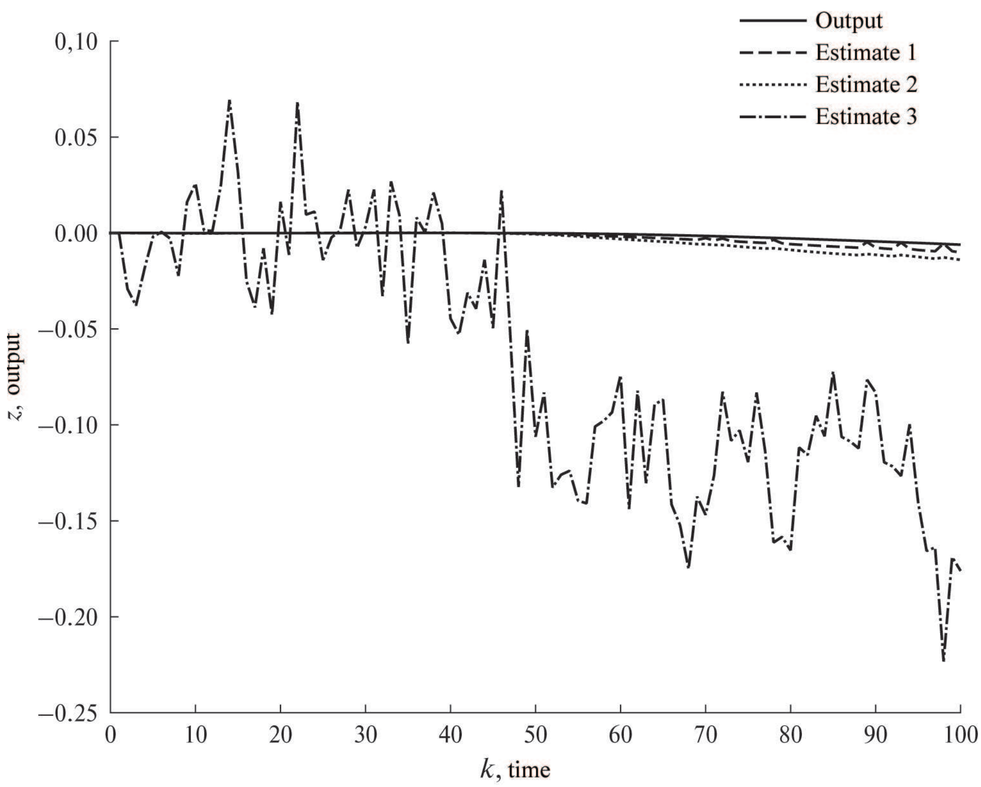


**Fig. 1.** Estimated output.

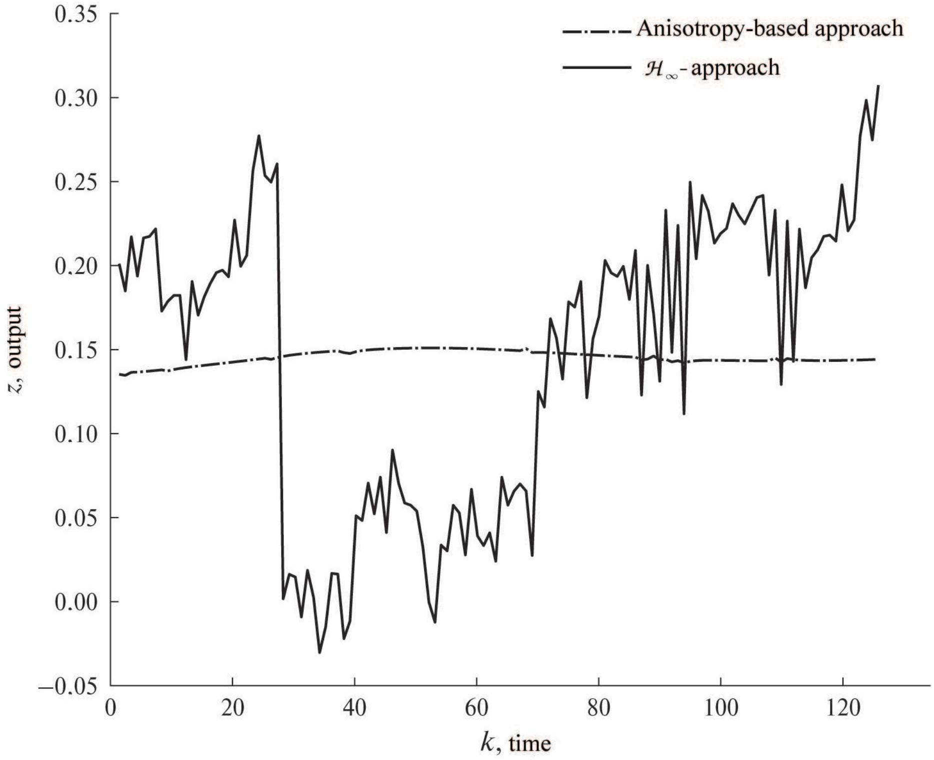


**Fig. 2.** Comparison of estimations.

adjacency matrix has the following form:

$$\mathfrak{a} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}.$$

Estimate 3 corresponds to the case where only the first measured output was used to construct the estimate.

For comparison, we present the values of the Euclidean norms of the estimated output $\|z\| = 0.0221$ and the three obtained estimates: $\|\hat{z}_1\| = 0.0385$, $\|\hat{z}_2\| = 0.0598$, $\|\hat{z}_3\| = 1.0222$.

Figure 2 presents the simulation results under the influence of an external disturbance with a fixed mean anisotropy level equal to 30. In this case, we compare only the consensus estimate obtained based on two measurements. The dash-dotted line denotes the estimate of the roll angle based on the estimator using the anisotropy-based approach, the solid line denotes the $\mathcal{H}_\infty$-optimal estimate. Although the values of the Euclidean norms of the anisotropy-based estimate $\|\hat{z}_{AB}\| = 1.6308$ and the $\mathcal{H}_\infty$ estimate $\|\hat{z}_{\mathcal{H}_\infty}\| = 1.9165$ differ by 15%, it is noticeable from the graphs that in the considered example with imperfect measurements, the nature of the obtained estimates differs significantly, which indicates the advantage of using the anisotropy-based estimator in this case.

## 6. CONCLUSION

The paper solves the problem of constructing an estimate for a system whose measured outputs may be unavailable at an arbitrary time. To this end, it is proposed to duplicate such outputs and build a consensus estimate for the output of the object of observation. Since the solution to the problem can be obtained using standard application software packages for computations, it seems possible to use the synthesized estimator, for example, for constructing output feedback control for dynamic objects.

FUNDING

This work was supported in part by the Russian Foundation for Basic Research, projects no. 24-21-20055.

*APPENDIX*

## A.1. PROOF OF THEOREM 2

Apply the Schur complement lemma [21] to inequality (5), which leads to the appearance of an inequality of the following form:

$$\begin{bmatrix} R-\mathcal{M}^{\mathrm{T}}\mathcal{M} & * & * & * & * & \cdots & * \\ \mathcal{N}^{\mathrm{T}}\mathcal{M} & \eta I_{m_w}-\mathcal{N}^{\mathrm{T}}\mathcal{N} & * & * & * & \cdots & * \\ R\mathcal{A}_0 & -R\mathcal{B} & R & * & * & \cdots & * \\ \sigma_1 R\mathcal{A}_1 & 0 & 0 & R & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_n R\mathcal{A}_n & 0 & 0 & 0 & 0 & \cdots & R \end{bmatrix} \succ 0,$$

where the substitution $R = q^{-1}R$ is taken into account. As can be seen, the last inequality is still nonlinear. To reduce it to a linear form, we introduce the following change of variables:

$$X = RU = R\begin{bmatrix} Y & 0 \\ 0 & \overline{H}^{\mathfrak{a}} \end{bmatrix},$$

where $Y$ is some matrix (it is worth noting in advance that this matrix does not affect the solution of the problem, namely, the values of the estimator matrices $H_{ji}, j, i = \overline{1,n}$). Taking into account the last change of matrix variables, we obtain inequality (8).

To derive inequality (9), we define an additional positive definite matrix $\Psi$ related to the matrix $S$ as follows:

$$0 \prec \frac{1}{\eta}\Psi \prec S^{-1},$$

where $\eta^{-1} = q$. Taking into account the positive definiteness of the matrix $\Psi$, the change of variables $X = RU$, and applying the Schur complement lemma, we obtain inequality (9). The convex constraint (10) follows from inequality (7) after the corresponding changes of variables. The theorem is proved.

## REFERENCES

1. Bulgakov, B.V., On the Accumulation of Disturbances in Linear Vibrational Systems with Constant Parameters, *Dokl. Akad. Nauk SSSR*, 1946, vol. 5, pp. 339–342.
2. Vidyasagar, M., Optimal Rejection of Persistent Bounded Disturbances, *IEEE Trans. Autom. Control*, 1986, vol. 31, pp. 527–535.
3. Yakubovich, E.D., Solving One Problem of Optimal Control for a Discrete Linear System, *Autom. Remote Control*, 1975, vol. 36, no. 9, pp. 73–79.
4. Doyle, J.C., Glover, K., Khargonekar, P.P., and Francis, B.A., State-space Solutions to Standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control Problems, *IEEE Trans. Autom. Control*, 1989, vol. 34, no. 8, pp. 831–847.
5. Saberi, A., Sanutti, P., and Chen, B.M., *$\mathcal{H}_2$-optimal control*, Prentice Hall, Englewood Cliffs, N.J., 1995.
6. Lee, H. and Park, J., Quantitative Controllability Metric for Disturbance Rejection in Linear Unstable Systems, *Mathematics*, 2025, vol. 13, no. 6, pp. 1–24.

7. Kalman, R.E., A New Approach to Linear Filtering and Prediction Problems, *J. Basic Engineering*, 1960, vol. 82, no. 1, pp. 35–45.

8. Zames, G., Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses, *IEEE Trans. Autom. Control*, 1981, vol. 26, no. 2, pp. 301–320.

9. Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V., Anisotropy of Signals and the Entropy of Linear Stationary Systems, *Doklady Math.*, 1995, vol. 51, pp. 388–390.

10. Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V., On Computing the Anisotropic Norm of Linear Discrete-time-invariant Systems, *Proc. 13 IFAC World Congr.*, 1996, pp. 179–184.

11. Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V., State-space Solution to Anisotropy-based Stochastic $\mathcal{H}_\infty$-Optimization Problem, *Proc. 13 IFAC World Congr.*, 1996, pp. 427–432.

12. Timin, V.N. and Kurdyukov, A.P., Suboptimal Anisotropy-based Filtering on a Finite Horizon, *Autom. Remote Control*, 2016, vol. 77, no. 1, pp. 5–29.

13. Tchaikovsky, M.M., Timin, V.N., Kustov, A.Yu., and Kurdyukov, A.P., Numerical Procedures for Anisotropic Analysis of Stationary Systems and Synthesis of Suboptimal Controllers and Filters, *Autom. Remote Control*, 2018, vol. 79, no. 1, pp. 162–181.

14. Kustov, A.Yu., State-Space Formulas for Anisotropic Norm of Linear Discrete Time Varying Stochastic Systems, *Proc. 15th Int. Conf. on Electrical Engineering, Computing Science and Automatic Control*, 2018, pp. 1–6.

15. Yurchenkov, A.V., An Anisotropy-based Boundedness Criterion for Time-invariant Systems with Multiplicative Noises, *Control Sciences*, 2022, no. 5, pp. 13–20.

16. Yurchenkov, A.V., Kustov, A.Yu., and Timin, V.N., The sensor network estimation with dropouts: Anisotropy-based approach, *Automatica*, 2023, vol. 151, art. no. 110924.

17. Shen, B., Wang, Z., and Hung, Y.S., Distributed $\mathcal{H}_\infty$-consensus Filtering in Sensor Networks with Multiple Missing Measurements: The Finite-horizon Case, *Automatica*, 2010, vol. 46, pp. 1682–1688.

18. Diamond, P., Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V., Anisotropy-based performance analysis of linear discrete time invariant control systems, *Int. J. Control*, 2001, vol. 74, no. 1, pp. 28–42.

19. Vladimirov, I.G., Diamond, P., and Kloeden, P., Anisotropy-based Robust Performance Analysis of Finite Horizon Linear Discrete Time Varying Systems, *Autom. Remote Control*, 2006, vol. 67, no. 8, pp. 92–111.

20. Semyonov, A.V., Vladimirov, I.G., and Kurdjukov, A.P., Stochastic approach to $\mathcal{H}_\infty$-optimization, *Proc. 33rd Conf. on Decision and Control*, 1994, vol. 3, pp. 2249–2250.

21. Boyd, S. and Vandenberghe, L., *Convex Optimization*, Cambridge: Cambridge University Press, 2004.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 494–509.


NONLINEAR SYSTEMS

# Energy Quasi-Optimal Angular Acceleration of a Spacecraft Based on the Poinsot Concept

**Ya. G. Sapunkov**[*,a] **and A. V. Molodenkov**[*,b]

[*] *Institute of Precision Mechanics and Control, Russian Academy of Sciences, Saratov, Russia*
*e-mail:* [a]*iptmuran@san.ru,* [b]*molalexei@yandex.ru*

Received December 27, 2024
Revised October 17, 2025
Accepted October 27, 2025

**Abstract**—This paper is devoted to the kinematic problem of the optimal, in terms of energy consumption, program angular acceleration of a spacecraft under arbitrary boundary conditions imposed on its angular position and angular velocity. A quasi-optimal analytical solution of the problem is obtained within the classical Poinsot concept of the angular motion of a solid body as a generalized coning motion and Pontryagin's maximum principle. This solution is brought to an algorithm. Supporting analytical and numerical examples are provided to show either the proximity of the quasi-optimal and optimal solutions or their complete coincidence, depending on the boundary conditions.

*Keywords*: spacecraft, solid body, angular acceleration, Poinsot concept, optimal and quasi-optimal control, analytical solution, algorithm

**DOI:** 10.7868/S1608303226060026

## 1. INTRODUCTION

In many spacecraft guidance and control problems, it is required to know the optimal, in some sense, program angular acceleration of a spacecraft under arbitrary (given) boundary conditions [1–5]. In the literature, as a rule, such acceleration is either numerically found or analytically constructed (e.g., [5]) based on polynomials (splines) by representing the spacecraft attitude quaternion via polynomials and expressing the angular velocity vector through this quaternion. However, no guarantees (rigorously proved theorems or considerations from theoretical mechanics) are given that these analytical solutions will approximate the optimal angular motion trajectory of the spacecraft well enough on the entire set of its angular motions under any boundary conditions.

In this paper, we study the kinematic problem of the optimal, in terms of energy consumption, program angular acceleration of a spacecraft under arbitrary boundary conditions for its angular position and angular velocity. It is difficult to obtain the optimal analytical solution of this problem under arbitrary boundary conditions: in the general case, the exact solution of the attitude problem of a solid body by its known angular velocity (the Darboux problem) remains unknown [6, 7]. Following the classical Poinsot concept of the angular motion of a solid body as a generalized coning motion, we reformulate the optimal angular acceleration problem of a spacecraft in this class of motions. The spacecraft trajectory is given by explicit expressions containing arbitrary quaternionic and scalar constants and two arbitrary scalar functions (the parameters of the generalized coning motion). These parameters and their derivatives are employed as variables to pose and solve an optimization problem in which the second derivatives of the parameters act as controls. Note that the generality of the original problem is almost not violated: the known exact solution of the classical optimal angular acceleration problem in the planar rotation case [4] or in the new special case of regular precession (considered below) and the similar solutions of the problem within the above concept completely coincide; in other cases, in numerical calculations of the solution of the original problem and the analytical solution proposed, the relative error between the values of an optimization criterion (the determinative characteristic of the problem) constitutes at most 1%, including spacecraft rotations to large angles. Therefore, the analytical solution proposed in this paper can be treated as quasi-optimal with respect to that of the classical optimal angular acceleration problem of a spacecraft. Explicit expressions for the attitude quaternion and the angular velocity vector of a spacecraft are derived. An explicit formula for the angular acceleration of a spacecraft is obtained by differentiating the expression for the angular velocity vector. Finally, an analytical algorithm for solving the problem is presented, which can be used aboard the spacecraft.

The analytical solution method proposed below has previously been successfully applied to dynamic optimal rotation problems for spacecraft of various configurations with different optimization criteria [8, 9], as well as as to the design of quasi-optimal (in terms of time) angular acceleration of a spacecraft under arbitrary boundary conditions [10].

## 2. CLASSICAL PROBLEM FORMULATION

The angular motion of a spacecraft as a solid body around its center of gravity is described by a system of differential equations in vector-matrix form:

$$2\begin{bmatrix}\dot{\lambda}_0\\ \dot{\lambda}_1\\ \dot{\lambda}_2\\ \dot{\lambda}_3\end{bmatrix} = \begin{bmatrix}-\lambda_1 & -\lambda_2 & -\lambda_3\\ \lambda_0 & -\lambda_3 & \lambda_2\\ \lambda_3 & \lambda_0 & -\lambda_1\\ -\lambda_2 & \lambda_1 & -\lambda_0\end{bmatrix}\begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\end{bmatrix}, \tag{2.1}$$

$$\begin{bmatrix}\dot{\omega}_1 & \dot{\omega}_2 & \dot{\omega}_3\end{bmatrix}^{\mathrm{T}} = \begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \varepsilon_3\end{bmatrix}^{\mathrm{T}}, \tag{2.2}$$

which is equivalent to the short representation [4]

$$2\dot{\mathbf{\Lambda}} = \mathbf{\Lambda} \circ \boldsymbol{\omega}, \tag{2.3}$$

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon}, \tag{2.4}$$

where $\mathbf{\Lambda}(t) = [\lambda_0(t),\, \lambda_1(t),\, \lambda_2(t),\, \lambda_3(t)]^{\mathrm{T}}$ or $\mathbf{\Lambda}(t) = \lambda_0(t) + \lambda_1(t)i_1 + \lambda_2(t)i_2 + \lambda_3(t)i_3$ is the normalized quaternion that describes the angular position of the spacecraft ($||\mathbf{\Lambda}|| = \lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$); $\boldsymbol{\omega} = [\omega_1(t),\, \omega_2(t),\, \omega_3(t)]^{\mathrm{T}}$ or $\boldsymbol{\omega}(t) = \omega_1(t)\mathbf{i}_1 + \omega_2(t)\mathbf{i}_2 + \omega_3(t)\mathbf{i}_3$ is the angular velocity vector of the spacecraft, which can be treated as a quaternion with the zero scalar part $\boldsymbol{\omega}(t) = [0,\, \omega_1(t),\, \omega_2(t),\, \omega_3(t)]^{\mathrm{T}}$; $\boldsymbol{\varepsilon} = [\varepsilon_1(t),\, \varepsilon_2(t),\, \varepsilon_3(t)]^{\mathrm{T}}$ is the angular acceleration vector of the spacecraft; the imaginary units $i_1$, $i_2$, and $i_3$ of the Hamiltonian correspond to the orthonormal basis vectors $\mathbf{i}_1$, $\mathbf{i}_2$, and $\mathbf{i}_3$, respectively, of the three-dimensional vector space; $\circ$ indicates the quaternion product; finally, $^{\mathrm{T}}$ means the vector transpose. The phase coordinates $\mathbf{\Lambda}$ and $\boldsymbol{\omega}$ are assumed to be continuous functions, and the angular acceleration $\boldsymbol{\varepsilon}$ (control) is assumed to be a piecewise continuous [11].

The boundary conditions imposed on the attitude and angular velocity of the spacecraft are arbitrary:

$$\mathbf{\Lambda}(0) = \mathbf{\Lambda}_0, \quad \mathbf{\Lambda}(T) = \mathbf{\Lambda}_T, \tag{2.5}$$

$$\boldsymbol{\omega}(0) = \boldsymbol{\omega}_0, \quad \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T. \tag{2.6}$$

It is required to find the optimal control $\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)$ that minimizes the energy consumption functional

$$J = \int\limits_0^T ||\boldsymbol{\varepsilon}||\, dt, \tag{2.7}$$

where the time $T$ is arbitrary and fixed.

The problem formulation (2.1)–(2.7) is kinematic since it neglects the dynamic configuration of the spacecraft. Note that the optimal angular acceleration of the spacecraft can be found as the derivative of its optimal angular velocity in the full dynamic optimal attitude problem of a spacecraft, where the Euler dynamic equations are used instead of the vector equation (2.4) and the control function is the vector of the external control moment applied to the spacecraft [8, 9].

## 3. DIMENSIONLESS VARIABLES

We reformulate the problem by replacing the original dimensional variables with dimensionless ones:

$$t^{\text{dimless}} = t^{\text{dim}}/T, \quad \boldsymbol{\omega}^{\text{dimless}} = \boldsymbol{\omega}^{\text{dim}}T, \quad \boldsymbol{\varepsilon}^{\text{dimless}} = \boldsymbol{\varepsilon}^{\text{dim}}T^2, \quad J^{\text{dimless}} = J^{\text{dim}}T^3.$$

In this case, all the expressions will not change, except for the optimization functional

$$J = \int_0^1 ||\boldsymbol{\varepsilon}||\, dt. \tag{3.1}$$

Below we will solve problem (2.3)–(2.6), (3.1) in the dimensionless variables for $T = 1$, with the removed superscripts in the problem formulation.

## 4. REDUCTION TO A BOUNDARY VALUE OPTIMIZATION PROBLEM

Following Pontryagin's maximum principle [4, 11], we introduce a quaternion $\boldsymbol{\Psi}(t)$, conjugate to the spacecraft attitude quaternion $\boldsymbol{\Lambda}(t)$, and a vector $\boldsymbol{\varphi}(t)$, conjugate to the spacecraft angular velocity vector $\boldsymbol{\omega}(t)$. The Hamilton–Pontryagin function is given by

$$H = -\left(\boldsymbol{\varepsilon},\, \boldsymbol{\varepsilon}\right) + \left(\boldsymbol{\Psi},\, \boldsymbol{\Lambda} \circ \boldsymbol{\omega}\right)/2 + \left(\boldsymbol{\varphi},\, \boldsymbol{\varepsilon}\right), \tag{4.1}$$

where (. , .) means the inner product of vectors.

The conjugate system of equations has the form

$$\begin{cases} 2\dot{\boldsymbol{\Psi}} = \boldsymbol{\Psi} \circ \boldsymbol{\omega} \\ \dot{\boldsymbol{\varphi}} = -\text{vect}\left(\widetilde{\boldsymbol{\Lambda}} \circ \boldsymbol{\Psi}\right)/2, \end{cases} \tag{4.2}$$

where vect(.) and ˜ are the vector part and conjugation of the quaternion, respectively. The linear differential systems of equations for the variables $\boldsymbol{\Psi}$ and $\boldsymbol{\Lambda}$ coincide, hence their solutions differ by a quaternion constant $\mathbf{C}$ :

$$\boldsymbol{\Psi} = \mathbf{C} \circ \boldsymbol{\Lambda}. \tag{4.3}$$

Therefore, in view of the notation adopted in [4]:

$$\mathbf{p} = \text{vect}\, \left(\widetilde{\boldsymbol{\Lambda}} \circ \boldsymbol{\Psi}\right) = \widetilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda}, \tag{4.4}$$

where $\mathbf{c}_v$ is the vector part of the quaternion $\mathbf{C}$, the expressions (4.2) become

$$\begin{cases} \mathbf{p} = \widetilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda} \\ \dot{\boldsymbol{\varphi}} = -\mathbf{p}/2. \end{cases} \tag{4.5}$$

Thus, due to the self-conjugation of the linear differential system of equations (2.1) (or (2.3)), the dimension of the boundary value problem of the maximum principle is reduced by four [4]. The maximum condition (3.1) yields the following continuous structure of optimal control:

$$\boldsymbol{\varepsilon}^{\text{opt}} = \boldsymbol{\varphi}/2. \tag{4.6}$$

From (2.3), (2.4), (4.4), and (4.5) it follows that $\dot{\boldsymbol{\omega}} = \boldsymbol{\varphi}/2$, $\ddot{\boldsymbol{\omega}} = -\mathbf{p}/4$, and $\dddot{\boldsymbol{\omega}} = -\dot{\mathbf{p}}/4 = \ddot{\boldsymbol{\omega}} \times \boldsymbol{\omega}$, where the symbol $\times$ means the vector product. Written as a single equation, they become

$$\dddot{\boldsymbol{\omega}} = \ddot{\boldsymbol{\omega}} \times \boldsymbol{\omega}. \tag{4.7}$$

As a result, the optimal angular acceleration of the spacecraft is found by solving the boundary value problem (2.3), (4.7), (2.5), (2.6).

Using the expression (4.4), we write the Hamilton–Pontryagin function (4.1) as

$$H = -\left(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}\right) + \left(\mathbf{p}, \boldsymbol{\omega}\right)/2 + \left(\boldsymbol{\varphi}, \boldsymbol{\varepsilon}\right). \tag{4.8}$$

Based on formulas (2.3), (2.4), and (4.5), we also indicate the vector first integral [12] relating the phase and conjugate variables of the problem: $\mathbf{p}/4 + \boldsymbol{\varphi} \times \boldsymbol{\omega} = \mathbf{const}\ \forall\ t \in [0, T]$.

Despite that this paper considers the problem without control constraints (the constraint on the magnitude of the angular acceleration is indirectly realized here by minimizing the integral control-quadratic functional), the expressions (4.4)–(4.6) can be used to estimate the magnitude of the optimal angular acceleration vector. For a continuously differentiable vector function $\boldsymbol{\varepsilon}^{\mathrm{opt}}$ of the optimal control, we have

$$| \boldsymbol{\varepsilon}^{\mathrm{opt}}(t) | = \left| \int\limits_0^t \dot{\boldsymbol{\varepsilon}}^{\mathrm{opt}}(\tau)\, d\tau + \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) \right| \le \int\limits_0^t | \dot{\boldsymbol{\varepsilon}}^{\mathrm{opt}}(\tau) |\ d\tau + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) |$$

$$= \int\limits_0^t | \mathbf{p}(\tau) |\ d\tau/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) | = | \mathbf{c}_v |\ t/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) | .$$

In other words, at any time instant $t \in [0, 1]$, the magnitude of the vector function $\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)$ must lie inside a sphere of variable radius

$$||\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)|| \le \left(| \mathbf{c}_v |\ t/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) |\right)^2\ \forall\ t \in [0, 1]. \tag{4.9}$$

Similarly, on the other side, we obtain

$$||\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)|| \le \left(| \mathbf{c}_v |\ (1-t)/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(1) |\right)^2\ \forall\ t \in [0, 1]. \tag{4.10}$$

Consequently, the magnitude of the optimal angular acceleration vector must belong to the intersection of the sets defined by inequalities (4.9), (4.10), i.e.,

$$\begin{gathered} ||\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)|| \le \gamma^2(t)\ \forall\ t \in [0, 1], \\ \gamma^2(t) = \min_t \left\{ \Big(| \mathbf{c}_v |\ t/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(0) |\Big)^2, \Big(| \mathbf{c}_v |\ (1-t)/4 + | \boldsymbol{\varepsilon}^{\mathrm{opt}}(1) |\Big)^2 \right\}, \end{gathered} \tag{4.11}$$

where the vector constant $\mathbf{c}_v$ is determined from the given constants of the problem statement.

Note that in the special case of the boundary conditions on the angular velocity (2.6) corresponding to a planar Euler rotation, when the exact solution of the problem is known, $| \mathbf{c}_v |$ in (4.11) is explicitly expressed through the given boundary conditions on the spacecraft angular position (2.5):

$$| \mathbf{c}_v | = | \operatorname{vect}(\boldsymbol{\Lambda}_T \circ \widetilde{\boldsymbol{\Lambda}}_0)/\sin(\arccos(\operatorname{scal}(\boldsymbol{\Lambda}_T \circ \widetilde{\boldsymbol{\Lambda}}_0))) |, \tag{4.12}$$

where scal(.) means the scalar part of the quaternion, and the magnitude of the vector function $\boldsymbol{\varepsilon}^{\mathrm{opt}}(t)$ can reach the boundary of the region (4.11), (4.12).

## 5. PARTIAL SOLUTION OF THE PROBLEM IN ELEMENTARY FUNCTIONS

In the class of regular coning motions, we obtain a new analytical partial solution of the optimal angular acceleration problem of a spacecraft, expressed in elementary functions. In this class of motions, the optimal angular velocity of the spacecraft has the form

$$\boldsymbol{\omega}(t) = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1 \beta \sin \gamma t + \mathbf{i}_2 \beta \cos \gamma t + \mathbf{i}_3 \gamma) \circ \mathbf{Q}, \tag{5.1}$$

where $\mathbf{Q}$ (quaternion), $\beta$, and $\gamma$ are arbitrary constants; in addition,

$$||\mathbf{K}|| = \mathrm{q}_0^2 + \mathrm{q}_1^2 + \mathrm{q}_2^2 + \mathrm{q}_3^2 = 1. \tag{5.2}$$

The quaternion $\mathbf{Q}$ defines the rotation of the vector $\mathbf{i}_1\beta \sin\gamma t + \mathbf{i}_2\beta\cos\gamma t + \mathbf{i}_3\gamma$ about some constant axis passing through a fixed point of the spacecraft. Let us show that the angular velocity (5.1) satisfies equation (4.7). Differentiating the expression (5.1) three times with respect to $t$ yields $\dot{\boldsymbol{\omega}}(t) = \beta\gamma\widetilde{\mathbf{Q}} \circ (\mathbf{i}_1\cos\gamma t - \mathbf{i}_2 \sin\gamma t) \circ \mathbf{Q}$, $\ddot{\boldsymbol{\omega}}(t) = -\beta\gamma^2\widetilde{\mathbf{Q}} \circ (\mathbf{i}_1 \sin\gamma t + \mathbf{i}_2\cos\gamma t) \circ \mathbf{Q}$, and $\dddot{\boldsymbol{\omega}}(t) = \beta\gamma^3\widetilde{\mathbf{Q}} \circ (-\mathbf{i}_1\cos\gamma t + \mathbf{i}_2 \sin\gamma t) \circ \mathbf{Q}$. By substituting (5.1) and the above derivatives into (4.7), we verify the desired equality with $\dddot{\boldsymbol{\omega}} = \ddot{\boldsymbol{\omega}} \times \boldsymbol{\omega} = (\ddot{\boldsymbol{\omega}} \circ \boldsymbol{\omega} - \boldsymbol{\omega} \circ \ddot{\boldsymbol{\omega}})/2$.

Based on (2.3) and (5.1), the spacecraft trajectory can be explicitly written in the form of a regular precession:

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{Q}} \circ \mathbf{exp}\{\mathbf{i}_2\beta t/2\} \circ \mathbf{exp}\{\mathbf{i}_3\gamma t/2\} \circ \mathbf{Q}, \tag{5.3}$$

where $\mathbf{exp}\{.\}$ denotes the quaternion exponential [4].

The expressions (5.1)–(5.3) contain five arbitrary constants $\beta$, $\gamma$, and $q_i$, $i = 0, 1, 2$. (The constant $q_3$ is related to condition (5.2).) Let us satisfy the boundary conditions (2.5), (2.6). Due to the insufficient number of arbitrary constants in (5.1), we impose requirements on $|\boldsymbol{\omega}_0|$ and $\boldsymbol{\omega}_T$ in the course of solving the problem, while the unit vector $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0/|\boldsymbol{\omega}_0|$ is arbitrary and given. At the point $t = 0$, by formula (5.1),

$$\boldsymbol{\omega}_0 = |\boldsymbol{\omega}_0|\,\boldsymbol{\omega}_0^e = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_2\beta + \mathbf{i}_3\gamma) \circ \mathbf{Q}, \tag{5.4}$$

$$\|\boldsymbol{\omega}_0\| = \left\|\widetilde{\mathbf{Q}} \circ (\mathbf{i}_2\beta + \mathbf{i}_3\gamma) \circ \mathbf{Q}\right\| = \left\|\widetilde{\mathbf{Q}}\right\| \|\mathbf{i}_2\beta + \mathbf{i}_3\gamma\| \; \|\mathbf{Q}\| = \beta^2 + \gamma^2; \tag{5.5}$$

at the right end of the spacecraft trajectory for $t = T = 1$, from (5.3) we have

$$\boldsymbol{\Lambda}_T = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{Q}} \circ \mathbf{exp}\{\mathbf{i}_2\beta/2\} \circ \mathbf{exp}\{\mathbf{i}_3\gamma/2\} \circ \mathbf{Q}, \tag{5.6}$$

with

$$\mathrm{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) = \mathrm{scal}(\mathbf{exp}\{\mathbf{i}_2\beta/2\} \circ \mathbf{exp}\{\mathbf{i}_3\gamma/2\}), \tag{5.7}$$

where scal(.) means the scalar part of the quaternion. Based on (5.2) and (5.4)–(5.6), we find $|\boldsymbol{\omega}_0|$, $\beta$, $\gamma$, and $\mathbf{Q}$.

Let us represent (5.4) and (5.6) in the form

$$(\mathbf{i}_2\beta + \mathbf{i}_3\gamma) \circ \mathbf{Q} - \mathbf{Q} \circ \boldsymbol{\gamma}_0 = 0,$$
$$\mathbf{exp}\{\mathbf{i}_2\beta/2\} \circ \mathbf{exp}\{\mathbf{i}_3\gamma/2\} \circ \mathbf{Q} - \mathbf{Q} \circ \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T = 0.$$

Equivalently, involving the $\boldsymbol{m}$- and $\boldsymbol{n}$-matrices isomorphic to quaternions [13], we have

$$\left(\begin{bmatrix} 0 & 0 & -\beta & -\gamma \\ 0 & 0 & -\gamma & \beta \\ \beta & \gamma & 0 & 0 \\ \gamma & -\beta & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -\omega_{01} & -\omega_{02} & -\omega_{03} \\ \omega_{01} & 0 & \omega_{03} & -\omega_{02} \\ \omega_{02} & -\omega_{03} & 0 & \omega_{01} \\ \omega_{03} & \omega_{02} & -\omega_{01} & 0 \end{bmatrix}\right) \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.8}$$

$$\left(\begin{bmatrix} \mu_0 & -\mu_1 & -\mu_2 & -\mu_3 \\ \mu_1 & \mu_0 & -\mu_3 & \mu_2 \\ \mu_2 & \mu_3 & \mu_0 & -\mu_1 \\ \mu_3 & -\mu_2 & \mu_1 & \mu_0 \end{bmatrix} - \begin{bmatrix} \nu_0 & -\nu_1 & -\nu_2 & -\nu_3 \\ \nu_1 & \nu_0 & \nu_3 & -\nu_2 \\ \nu_2 & -\nu_3 & \nu_0 & \nu_1 \\ \nu_3 & \nu_2 & -\nu_1 & \nu_0 \end{bmatrix}\right) \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.9}$$

where the coefficient matrix of the linear algebraic system (5.9) is determined by the components of the quaternions $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ :

$$\begin{cases} \boldsymbol{\mu} = \mathbf{exp}\{\mathbf{i}_2\beta/2\} \circ \mathbf{exp}\{\mathbf{i}_3\gamma/2\} \\ \mu_0 = \cos(\beta/2)\cos(\gamma/2), \quad \mu_1 = \sin(\beta/2)\sin(\gamma/2) \\ \mu_2 = \sin(\beta/2)\cos(\gamma/2), \quad \mu_3 = \cos(\beta/2)\sin(\gamma/2), \end{cases} \tag{5.10}$$

$$\boldsymbol{\nu} = \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T. \tag{5.11}$$

Their norms are $\|\boldsymbol{\mu}\| = 1$ and $\|\boldsymbol{\nu}\| = 1$, and the coefficient matrices of systems (5.8), (5.9) have rank 2. By choosing two linearly independent equations in (5.8) and in (5.9), we obtain a homogeneous system of the form

$$\begin{bmatrix} 0 & \omega_{01} & \omega_{02} - \beta & \omega_{03} - \gamma \\ -\omega_{01} & 0 & -(\omega_{03} + \gamma) & \omega_{02} + \beta \\ \mu_1 - \nu_1 & 0 & -(\mu_3 + \nu_3) & \mu_2 + \nu_2 \\ 0 & \nu_1 - \mu_1 & \nu_2 - \mu_2 & \nu_3 - \mu_3 \end{bmatrix} \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{5.12}$$

A nontrivial solution of the system of equations (5.12) is due to the zero determinant of its coefficient matrix. Using this fact and formulas (5.4), (5.5), (5.10), and (5.11), we get

$$|\boldsymbol{\omega}_0| = (\mu_2\beta + \mu_3\gamma)/(\nu_1\omega_{01}^e + \nu_2\omega_{02}^e + \nu_3\omega_{03}^e). \tag{5.13}$$

Based on (5.5), (5.7), (5.10), and (5.13), we compile a system of two nonlinear equations to determine the constants $\beta$ and $\gamma$ :

$$\begin{cases} (\beta^2 + \gamma^2)\,(\nu_1\omega_{01}^e + \nu_2\omega_{02}^e + \nu_3\omega_{03}^e)^2 - (\mu_2\beta + \mu_3\gamma)^2 = 0 \\ \mathrm{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) - \cos(\beta/2)\cos(\gamma/2) = 0. \end{cases} \tag{5.14}$$

The value of $|\boldsymbol{\omega}_0|$ is found from (5.3) using the solution of system (5.4).

The components of the quaternion $\mathbf{Q}$ are given by

$$\begin{array}{ll} q_3 = \pm\left[1 + (B_0/G)^2 + (B_1/G)^2 + (B_2/G)^2\right]^{-1/2}, & q_0 = B_0 q_3/G, \\ q_1 = B_1 q_3/G, & q_2 = B_2 q_3/G, \end{array} \tag{5.15}$$

where

$$\begin{cases} B_0 = -(\mu_2 + \nu_2)(\omega_{03} + \gamma) + (\mu_3 + \nu_3)(\omega_{02} + \beta) \\ B_1 = (\mu_1 - \nu_1)\omega_{01} + (\mu_3 + \nu_3)(\gamma - \omega_{03}) + (\mu_2 + \nu_2)(\beta - \omega_{02}) \\ B_2 = (\mu_1 - \nu_1)(\omega_{02} + \beta) + (\mu_2 + \nu_2)\omega_{01} \\ G = (\mu_1 - \nu_1)(\omega_{03} + \gamma) + (\mu_3 + \nu_3)\omega_{01}. \end{cases} \tag{5.16}$$

The boundary condition on the spacecraft angular velocity for $t = T = 1$ shall be

$$\boldsymbol{\omega}(T) = \boldsymbol{\omega}_T = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1\beta\sin\gamma + \mathbf{i}_2\beta\cos\gamma + \mathbf{i}_3\gamma) \circ \mathbf{Q}. \tag{5.17}$$

Finally, if the boundary conditions on the spacecraft angular velocity $\boldsymbol{\omega}$ satisfy the requirements (5.13) and (5.17), the spacecraft angular motion trajectory will lie in the class of regular coning motions and will be determined by the explicit expressions (5.1) and (5.2). (For all $t \in [0, T]$, $\boldsymbol{\omega}(t)$ will belong to a coning surface determined in space by arbitrary given boundary conditions on the position, $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Lambda}_T$, and the arbitrary given direction of its initial value vector $\boldsymbol{\omega}_0^e$.)

The optimal angular acceleration from (2.4), (5.1) is

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \beta\gamma\widetilde{\mathbf{Q}} \circ (\mathbf{i}_1\cos\gamma t - \mathbf{i}_2\sin\gamma t) \circ \mathbf{Q}, \tag{5.18}$$

$$||\boldsymbol{\varepsilon}|| = \beta^2\gamma^2 = \mathrm{const}. \tag{5.19}$$

The dimensionless optimization functional (3.1) takes the optimal value

$$J = \int_0^1 ||\boldsymbol{\varepsilon}||\, dt = \beta^2\gamma^2. \tag{5.20}$$

The conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ are determined using formulas (5.18) and (4.5). The problem with the above restrictions has been completely solved.

Now we present an algorithm for constructing the minimum-energy angular acceleration of a spacecraft in the class of regular coning motions:

1) Use the given constant conditions $T = 1$, $\boldsymbol{\Lambda}_0$, $\boldsymbol{\Lambda}_T$ (2.5), and $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0/\,|\boldsymbol{\omega}_0|$ and formulas (2.6), (5.10), (5.11), (5.15), and (5.16) to determine the unknowns $\beta$, $\gamma$, and $|\boldsymbol{\omega}_0|$.

2) Use $\boldsymbol{\Lambda}_0$, $\boldsymbol{\Lambda}_T$, $\beta$, $\gamma$, $T = 1$, and $|\boldsymbol{\omega}_0|$ and formulas (5.15) and (5.16) to find the quaternion $\mathbf{Q}$.

3) Use the expressions $\boldsymbol{\omega}_0^{\text{calc}} = |\boldsymbol{\omega}_0|\,\boldsymbol{\omega}_0^e$ and $\boldsymbol{\omega}_T^{\text{calc}} = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1\beta\sin\gamma + \mathbf{i}_2\beta\cos\gamma + \mathbf{i}_3\Omega) \circ \mathbf{Q}$ to find the conditions $\boldsymbol{\omega}_0^{\text{calc}}$ and $\boldsymbol{\omega}_T^{\text{calc}}$ at the ends of the trajectory.

4) Compare the values of $\boldsymbol{\omega}_0^{\text{calc}}$ and $\boldsymbol{\omega}_T^{\text{calc}}$ with the given ones (2.6).

5) If equality holds in item 4), then the optimal solution of the problem is in the class of regular coning motions; use formulas (5.1), (5.3), and (5.18) and item 1) to calculate the angular velocity, trajectory, and angular acceleration vector of the spacecraft; determine the optimal value of the functional based on (5.10).

6) Calculate the conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ by formulas (4.4), (4.5), (5.18).

This optimal solution of the problem in the class of regular coning motions and the previously known exact solution [4] in the class of planar Euler rotations (under the condition $\boldsymbol{\omega}_0\ \boldsymbol{\omega}_T\ ||\text{vect}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$) for the special cases of boundary conditions on the spacecraft angular velocity will be used as analytical confirmations when constructing a quasi-optimal solution of the spacecraft angular acceleration problem in the class of generalized coning motions. Also, we will provide some suggestive considerations based on the numerical solution of the optimal angular acceleration problem of a spacecraft with arbitrary boundary conditions.

## 6. JUSTIFICATION OF THE APPROACH BASED ON NUMERICAL SOLUTIONS OF THE PROBLEM

Let us numerically solve the boundary value problem of the maximum principle (Section 4) for the original kinematic optimal acceleration problem (2.3)–(2.6), (3.1) of a spacecraft:

$$\begin{cases} 2\dot{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \circ \boldsymbol{\omega} \\ \dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon} \\ \dot{\boldsymbol{\varphi}} = -\mathbf{p}/2 \\ \ \mathbf{p} = \widetilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda}, \ \ \mathbf{c}_v = \mathbf{const}, \end{cases} \tag{6.1}$$

$$\boldsymbol{\Lambda}(0) = \boldsymbol{\Lambda}_0, \ \ \boldsymbol{\omega}(0) = \boldsymbol{\omega}_0, \tag{6.2}$$

$$\boldsymbol{\Lambda}(T) = \boldsymbol{\Lambda}_T, \ \ \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T, \tag{6.3}$$

$$\boldsymbol{\varepsilon}^{\text{opt}} = \boldsymbol{\varphi}/2. \tag{6.4}$$

It is required to find $\boldsymbol{\varepsilon}^{\text{opt}}$, $T^{\text{opt}}$, $\boldsymbol{\Lambda}^{\text{opt}}$, $\boldsymbol{\omega}^{\text{opt}}$, and $\mathbf{c}_v$. The terminal conditions (6.3) are conveniently written in the phase space $\boldsymbol{\Lambda} \times \boldsymbol{\omega}$ of dimension 7 as

$$\text{vect}\ (\boldsymbol{\Lambda}(T) \circ \widetilde{\boldsymbol{\Lambda}}_T) = 0, \ \ \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T. \tag{6.5}$$

(For details, see [8, 9].)

A numerical solution method for such problems was described in [8, 9]. Also, for comparison, we provide kinematic characteristics based on the calculation results in the full dynamic optimal rotation problem of a spacecraft (in the sense of minimizing energy consumption) as a solid body of various dynamic configurations [9] in dimensionless variables ($T = 1$) under the same boundary conditions:

$$\boldsymbol{\Lambda}_0 = (0.7951,\ 0.2981,\ -0.3975,\ 0.3478),\ \ \boldsymbol{\omega}_0 = (0.2739,\ -0.2388,\ -0.3), \tag{6.6}$$

$$\boldsymbol{\Lambda}_T = (0.8443,\ 0.3984,\ -0.3260,\ 0.1485),\ \ \boldsymbol{\omega}_T = (0.0,\ 0.0,\ -0.59). \tag{6.7}$$

*Spacecraft 1.* An arbitrary spacecraft, $I_1 = 0.9869$, $I_2 = 1.1843$, and $I_3 = 0.7895$.

*Spacecraft 2.* An arbitrary spacecraft, $I_1 = 0.9506$, $I_2 = 1.3308$, and $I_3 = 0.5704$.

*Spacecraft 3.* The International Space Station (ISS) in its early version [14], $I_1 = 4\,853\,000$ kg $\times$ m$^2$, $I_2 = 23\,601\,000$ kg $\times$ m$^2$, and $I_3 = 26\,278\,000$ kg $\times$ m$^2$, or the dimensionless variables $I_1 = 0.2358$, $I_2 = 1.1466$, and $I_3 = 1.2766$.

*Spacecraft 4.* Space Shuttle, which has characteristics almost like a dynamically symmetric solid body: $I_1 = 3\,400\,648$ kg $\times$ m$^2$, $I_2 = 21\,041\,672$ kg $\times$ m$^2$ or $I_1 = 0.1967$, $I_2 = 1.2168$, and $I_3 \approx I_2$.

*Spacecraft 5.* An arbitrary spacecraft, $I_1 = 0.9116$, $I_2 = 1.3674$, and $I_3 = 0.5470$.

Table 1 shows the values of the attitude quaternion and angular velocity vector of the spacecraft obtained by solving the optimal angular acceleration problem (6.1)–(6.5) and the dynamic optimal rotation problem [9] for spacecraft 1–4 at the intermediate point $t = 0.5$ of the time interval. For comparison, the last line of Table 1 contains the data obtained by solving the quasi-optimal angular acceleration problem of a spacecraft considered in Sections 6 and 7 of the paper.

**Table 1**. Values of the attitude quaternion and angular velocity vector

| Spacecraft | $\lambda_0$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\omega_1$ | $\omega_2$ | $\omega_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.8095 | 0.3628 | −0.3766 | 0.2670 | −0.0499 | −0.0115 | −0.4941 |
| 2 | 0.8093 | 0.3631 | −0.3765 | 0.2674 | −0.0496 | −0.0116 | −0.4949 |
| 3 | 0.8077 | 0.3654 | −0.3773 | 0.2678 | −0.0506 | −0.0159 | −0.4949 |
| 4 | 0.8086 | 0.3634 | −0.3780 | 0.2668 | −0.0519 | −0.0137 | −0.4948 |
| Boundary value problem (6.1)–(6.5) | 0.8096 | 0.3625 | −0.3768 | 0.2668 | 0.0502 | −0.0114 | −0.4937 |
| Modified problem | 0.8099 | 0.3627 | −0.3756 | 0.2673 | −0.0488 | −0.0098 | −0.4938 |

For the spacecraft under consideration, Table 2 presents the components of the angular acceleration vectors $\boldsymbol{\varepsilon} = [\varepsilon_1(t),\ \varepsilon_2(t),\ \varepsilon_3(t)]^{\mathrm{T}}$ at the points $t = 0$, $t = 0.5$, and $t = T = 1$ of the optimal control process.

**Table 2**. Values of the angular acceleration

| Spacecraft | $\varepsilon_1(0)$ | $\varepsilon_2(0)$ | $\varepsilon_3(0)$ | $\varepsilon_1(0.5)$ | $\varepsilon_2(0.5)$ | $\varepsilon_3(0.5)$ | $\varepsilon_1(1)$ | $\varepsilon_2(1)$ | $\varepsilon_3(1)$ |
|---|---|---|---|---|---|---|---|---|---|
| 2 | −0.9649 | 0.7262 | −0.4736 | −0.3051 | 0.2053 | −0.2902 | 0.5357 | −0.1154 | −0.0965 |
| 3 | −0.9399 | 0.7235 | −0.4449 | −0.3207 | 0.2039 | −0.2942 | 0.5681 | −0.1076 | −0.0940 |
| 4 | −0.9262 | 0.6536 | −0.4422 | −0.3182 | 0.2586 | −0.2978 | 0.5591 | −0.2446 | −0.0886 |
| 5 | −0.7914 | 0.6537 | −0.4345 | −0.3968 | 0.2549 | −0.3002 | 0.7341 | −0.2381 | −0.0875 |
| Boundary value problem (6.1)–(6.5) | −0.9854 | 0.7259 | −0.4892 | −0.2917 | 0.2087 | −0.2878 | 0.5077 | −0.1272 | −0.0985 |
| Modified problem | −0.9647 | 0.7634 | −0.4932 | −0.3103 | 0.1687 | −0.2847 | 0.5350 | −0.0220 | 0.1024 |

We also describe the kinematic characteristics of the optimal motion when the initial and terminal states of the spacecraft are determined by the relations (6.6) and

$$\boldsymbol{\Lambda}_T = (0.79368,\ 0.49375,\ -0.26823,\ 0.23309),\ \ \boldsymbol{\omega}_T = (0.2,\ 0.3,\ -0.2), \tag{6.8}$$

respectively.

Table 3, with reference to problem (6.1)–(6.5) and the solution of the dynamic optimal rotation problem [9] for spacecraft 3–5, shows the components of the attitude quaternion and the angular velocity vector at $t = 0.5$; Table 4, the components of the angular acceleration vectors at $t = 0$, $t = 0.5$, and $t = T = 1$.

**Table 3.** Orientation quaternion and angular velocity vector

| Spacecraft | $\lambda_0$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\omega_1$ | $\omega_2$ | $\omega_3$ |
|---|---|---|---|---|---|---|---|
| 3 | 0.79080 | 0.41371 | −0.35618 | 0.27679 | 0.31768 | −0.01578 | −0.52304 |
| 4 | 0.79103 | 0.41317 | −0.35664 | 0.27635 | 0.31932 | −0.01491 | −0.52314 |
| 5 | 0.78999 | 0.41442 | −0.35860 | 0.27492 | 0.32414 | −0.01453 | −0.52402 |
| Boundary value problem (6.1)–(6.5) | 0.78995 | 0.41434 | −0.35808 | 0.28583 | 0.32168 | −0.01522 | −0.52451 |

**Table 4.** Angular acceleration

| Spacecraft | $\varepsilon_1(0)$ | $\varepsilon_2(0)$ | $\varepsilon_3(0)$ | $\varepsilon_1(0.5)$ | $\varepsilon_2(0.5)$ | $\varepsilon_3(0.5)$ | $\varepsilon_1(1)$ | $\varepsilon_2(1)$ | $\varepsilon_3(1)$ |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.36496 | 0.53821 | −1.00244 | −0.07017 | 0.44837 | 0.10585 | −0.54722 | 0.88317 | 1.18497 |
| 4 | 0.28551 | 0.51819 | −1.00680 | −0.04259 | 0.45769 | 0.10771 | −0.55803 | 0.86381 | 1.18517 |
| 5 | 0.19271 | 0.34271 | −1.05149 | −0.05270 | 0.54857 | 0.12304 | −0.41047 | 0.70409 | 1.16712 |
| Boundary value problem (6.1)–(6.5) | 0.26547 | 0.38677 | −1.00250 | −0.06733 | 0.52385 | 0.10458 | −0.43888 | 0.74958 | 1.18404 |

The same calculations were carried out for other boundary conditions. According to Tables 1–4 and other calculations, in the dynamic optimal rotation problem, the kinematic characteristics of the optimal motion depend significantly on the initial and terminal states, less significantly on its configuration, and are close enough to the results of the kinematic optimal angular acceleration problem of the spacecraft. Therefore, the kinematic optimal angular acceleration problem (2.3)–(2.7) has a general nature for spacecraft of arbitrary configurations. In this case, expressions for the attitude quaternion and angular velocity vector in the kinematic problem can be derived analytically in explicit form based on the solution of the so-called modified optimal reorientation problem of the spacecraft in the class of generalized coning motions. And the control angular acceleration of the spacecraft can be determined by differentiating its angular velocity vector. We consider this in more detail.

## 7. THE MODIFIED OPTIMAL ANGULAR ACCELERATION PROBLEM OF A SPACECRAFT

The general solution of the fundamental attitude problem of a solid body by its known angular velocity (2.1), (2.3), called the Darboux problem, is unknown. Therefore, we obtain a solution in the class of generalized coning motions. For this purpose, let the angular velocity vector $\boldsymbol{\omega}(t)$ be defined as

$$\boldsymbol{\omega}(t) = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1\dot{\mathrm{B}}(t)\sin\Gamma(t) + \mathbf{i}_2\dot{\mathrm{B}}(t)\cos\Gamma(t) + \mathbf{i}_3\dot{\Gamma}(t)) \circ \mathbf{Q}, \tag{7.1}$$

where the functions $\mathrm{B}(t)$ and $\Gamma(t)$ (the parameters of the generalized coning motion) are arbitrary. In this case, equation (2.3) has an exact solution [8, 9] satisfying the initial condition (2.5):

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{Q}} \circ \mathbf{exp}\{-\mathbf{i}_3\Gamma(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(\mathrm{B}(t) - \mathrm{B}(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Gamma(t)/2\} \circ \mathbf{Q}. \tag{7.2}$$

Clearly, formulas (7.1) and (7.2) include expressions for the angular velocity and spacecraft trajectory in the known exact solutions of the optimal angular acceleration problem when the angular velocity vector keeps a constant direction on the entire time interval of spacecraft motion (a planar

Euler rotation of a spacecraft [4]) or makes a regular precession (see Section 5 of this paper). Moreover, formulas (7.1) and (7.2) with arbitrary functions $\mathrm{B}(t)$ and $\Gamma(t)$ are a direct generalization of (5.1) and (5.3) with the linear functions $\beta t$ and $\gamma t$, respectively. The Darboux problem can generally be reduced, using bijective changes of variables [8, 9], to an equation of the form (2.3), where the angular velocity takes the form $\boldsymbol{\omega}^*(t) = -\boldsymbol{\omega}(t)$ and belongs to the class of similar motions. (The quaternionic differential equation (2.3) with the angular velocity $\boldsymbol{\omega}^*(t)$ still has no explicit solution.) Thus, the form of the angular velocity vector (7.1) corresponds to the classical Poinsot concept (kinematic interpretation) of the angular motion of a solid body as a generalized coning motion [7].

Let the second derivatives of the functions B and $\Gamma$ be treated as controls. With the designations

$$\dot{\mathrm{B}} = \mathrm{B}_1, \quad \dot{\Gamma} = \Gamma_1, \tag{7.3}$$

we compile the controlled system:

$$\dot{\mathrm{B}} = \mathrm{B}_1, \quad \dot{\Gamma} = \Gamma_1, \quad \dot{\mathrm{B}}_1 = u_1, \quad \dot{\Gamma}_1 = u_2, \tag{7.4}$$

where B, $\mathrm{B}_1$, $\Gamma$, and $\Gamma_1$ are the phase coordinates of this problem, and $u_1$, $u_2$ are the controls. The quaternion $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \mathbf{Q}_2 \circ \mathbf{Q}_1, \quad \mathbf{Q}_1 = \mathbf{exp}\{\mathbf{i}_1\alpha_1/2\}, \quad \mathbf{Q}_2 = \mathbf{exp}\{\mathbf{i}_2\alpha_2/2\}, \tag{7.5}$$

where $\alpha_1$ and $\alpha_2$ are arbitrary constants. Note that the quaternions $\mathbf{Q}_1$ and $\mathbf{Q}_2$ rotate the angular velocity vector $\boldsymbol{\omega}$ (7.1) about the axes $\mathbf{i}_1$ and $\mathbf{i}_2$, respectively. Due to the additive constant included in the function $\Gamma(t)$, the rotation about the axis $\mathbf{i}_3$ is already incorporated in the expression (7.1). The conjugation of the quaternion $\widetilde{\mathbf{Q}}$ has the form

$$\widetilde{\mathbf{Q}} = \widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2, \quad \widetilde{\mathbf{Q}}_1 = \mathbf{exp}\{-\mathbf{i}_1\alpha_1/2\}, \quad \widetilde{\mathbf{Q}}_2 = \mathbf{exp}\{-\mathbf{i}_2\alpha_2/2\}. \tag{7.6}$$

In view of (7.5) and (7.6), the functions $\boldsymbol{\omega}$ (7.1) and $\boldsymbol{\Lambda}$ (7.2) will satisfy the boundary conditions (2.5) and (2.6) if

$$\widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2 \circ (\mathbf{i}_1\mathrm{B}_1(0)\sin\Gamma(0) + \mathbf{i}_2\mathrm{B}_1(0)\cos\Gamma(0) + \mathbf{i}_3\Gamma_1(0)) \circ \mathbf{Q}_2 \circ \mathbf{Q}_1 = \boldsymbol{\omega}_0, \tag{7.7}$$

$$\widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2 \circ (\mathbf{i}_1\mathrm{B}_1(T)\sin\Gamma(T) + \mathbf{i}_2\mathrm{B}_1(T)\cos\Gamma(T) + \mathbf{i}_3\Gamma_1(T)) \circ \mathbf{Q}_2 \circ \mathbf{Q}_1 = \boldsymbol{\omega}_T, \tag{7.8}$$

$$\boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2 \circ \mathbf{exp}\{-\mathbf{i}_3\Gamma(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(\mathrm{B}(T) - \mathrm{B}(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Gamma(T)/2\} \circ \mathbf{Q}_2 \circ \mathbf{Q}_1 = \boldsymbol{\Lambda}_T. \tag{7.9}$$

For the controlled system (7.4), the optimization problem is posed as follows: find optimal controls $u_1(t)$ and $u_2(t)$ transferring system (7.4) from the state

$$\mathrm{B} = \mathrm{B}(0), \quad \mathrm{B}_1 = \mathrm{B}_1(0), \quad \Gamma = \Gamma(0), \quad \Gamma_1 = \Gamma_1(0) \tag{7.10}$$

to the state

$$\mathrm{B} = \mathrm{B}(T), \quad \Gamma_1 = \Gamma_1(T), \quad \Gamma = \Gamma(T), \quad \Gamma_1 = \Gamma_1(T) \tag{7.11}$$

that satisfy the relations (7.7)–(7.9), where the parameters $\alpha_1$ and $\alpha_2$ are to be determined, and implement the optimality criterion

$$J = \int_0^T (u_1^2 + u_2^2)\, dt. \tag{7.12}$$

We write the expressions (7.7)–(7.9) as

$$(\mathbf{i}_1\mathrm{B}_1(0)\sin\Gamma(0) + \mathbf{i}_2\mathrm{B}_1(0)\cos\Gamma(0) + \mathbf{i}_3\Gamma_1(0)) = \mathbf{Q}_2 \circ \mathbf{Q}_1 \circ \boldsymbol{\omega}_0 \circ \widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2, \tag{7.13}$$

$$(\mathbf{i}_1\mathrm{B}_1(T)\sin\Gamma(T) + \mathbf{i}_2\mathrm{B}_1(T)\cos\Gamma(T) + \mathbf{i}_3\Gamma_1(T)) = \mathbf{Q}_2 \circ \mathbf{Q}_1 \circ \boldsymbol{\omega}_T \circ \widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2, \tag{7.14}$$

$$\begin{gathered} \mathbf{exp}\{-\mathbf{i}_3\Gamma(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(\mathrm{B}(T) - \mathrm{B}(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Gamma(T)/2\} \\ = \mathbf{Q}_2 \circ \mathbf{Q}_1 \circ \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T \circ \widetilde{\mathbf{Q}}_1 \circ \widetilde{\mathbf{Q}}_2. \end{gathered} \tag{7.15}$$

The control angular acceleration of a spacecraft is determined from equation (2.4) by differentiation:

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon}. \tag{7.16}$$

This problem will be called the modified optimal angular acceleration problem of a spacecraft; its exact solution is acceptable as an approximate or quasi-optimal solution of the classical optimal problem (2.3)–(2.6), (3.1).

## 8. SOLUTION OF THE MODIFIED OPTIMAL ANGULAR ACCELERATION PROBLEM OF SPACECRAFT

The Hamilton–Pontryagin function of problem (7.4)–(7.12) is given by

$$H = -(u_1^2 + u_2^2) + \psi_1 \mathrm{B}_1 + \psi_2 \Gamma_1 + \psi_3 u_1 + \psi_4 u_2, \tag{8.1}$$

where the conjugate variables satisfy the differential system

$$\dot{\psi}_1 = 0, \;\; \dot{\psi}_2 = 0, \;\; \dot{\psi}_3 = -\psi_1, \;\; \dot{\psi}_4 = -\psi_2. \tag{8.2}$$

The general solution of system (8.2) has the form

$$\psi_1 = c_1, \;\; \psi_2 = c_2, \;\; \psi_3 = -c_1 t + c_3, \;\; \psi_4 = -c_2 t + c_4, \tag{8.3}$$

where $c_1, \ldots, c_4$ are arbitrary constants.

The maximum condition for the Hamilton–Pontryagin function (8.2) yields the following expressions for the optimal controls:

$$u_1 = \psi_3/2 = (-c_1 t + c_3)/2, \;\; u_2 = \psi_4/2 = (-c_2 t + c_4)/2. \tag{8.4}$$

By substituting formulas (8.4) into equations (7.4), we find their general solution containing eight unknown constants $c_1, \ldots, c_8$ :

$$\begin{aligned} &\mathrm{B} = -c_1 t^3/12 + c_3 t^2/4 + c_5 t + c_6, && \Gamma = -c_2 t^3/12 + c_4 t^2/4 + c_7 t + c_8, \\ &\mathrm{B}_1 = -c_1 t^2/4 + c_3 t/2 + c_5, && \Gamma_1 = -c_2 t^2/4 + c_4 t/2 + c_7. \end{aligned} \tag{8.5}$$

Based on (8.4), let $c_6 = 0$ in the expression for the function B (8.5). The nine unknown constants of the problem, $c_1, \ldots, c_5,\ c_7,\ c_8,\ \alpha_1$, and $\alpha_2$ are determined from the nine equations of system (7.13)–(7.15). (Due to the requirement $||\boldsymbol{\Lambda}|| = 1$, three scalar equations are independent in the quaternionic representation (7.15).) Note that these equations can be solved by the modified Newton method, where the initial approximations will be the values of the constants $c_1, \ldots, c_5,\ c_7,\ c_8$, $\alpha_1$, and $\alpha_2$ from the exact solution of the problem in the special case (see Section 5 of the paper).

By using (7.1), (7.2), and (8.5), we obtain explicit dependencies determining the laws of change of the angular velocity and trajectory of the spacecraft in the optimal angular acceleration problem in the class of generalized coning motions. In view of (7.1), (8.4), and (8.5), formula (7.16) gives an analytical expression for the angular acceleration vector $\boldsymbol{\varepsilon}$ :

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1(u_1 \sin\Gamma + \mathrm{B}_1\Gamma_1 \cos\Gamma) + \mathbf{i}_2(u_1 \cos\Gamma - \mathrm{B}_1\Gamma_1 \sin\Gamma) + \mathbf{i}_3 u_2) \circ \mathbf{Q}. \tag{8.6}$$

Due to (7.3), (7.5), and (7.6), the components of the vectors $\boldsymbol{\omega}$ (7.1) and $\boldsymbol{\varepsilon}$ (8.6) can be explicitly written as

$$\begin{aligned} &\omega_1 = \mathrm{B}_1 \sin\Gamma \cos\alpha_2 - \Gamma_1 \sin\alpha_2, \\ &\omega_2 = \mathrm{B}_1(\sin\Gamma \sin\alpha_1 \sin\alpha_2 - \cos\Gamma \cos\alpha_1) + \Gamma_1 \sin\alpha_1 \cos\alpha_2, \\ &\omega_3 = \mathrm{B}_1(\sin\Gamma \cos\alpha_1 \sin\alpha_2 - \cos\Gamma \sin\alpha_1) + \Gamma_1 \cos\alpha_1 \cos\alpha_2, \\ &\varepsilon_1 = u_1(\sin\Gamma + \mathrm{B}_1\Gamma_1 \cos\Gamma) \cos\alpha_2 - u_2 \sin\alpha_2, \\ &\varepsilon_2 = u_1(\sin\Gamma \sin\alpha_1 \sin\alpha_2 + \cos\Gamma \cos\alpha_1) \\ &\qquad + \mathrm{B}_1\Gamma_1(\cos\Gamma \sin\alpha_1 \sin\alpha_2 - \sin g \cos\alpha_1) + +u_2 \sin\alpha_1 \cos\alpha_2, \\ &\varepsilon_3 = u_1(\sin\Gamma \cos\alpha_1 \sin\alpha_2 - \cos\Gamma \sin\alpha_1) \\ &\qquad + f_1\Gamma_1(\cos\Gamma \cos\alpha_1 \sin\alpha_2 + \sin\Gamma \sin\alpha_1) + u_2 \cos\alpha_1 \cos\alpha_2. \end{aligned}$$

The norm of the control angular acceleration vector of a spacecraft, appearing in the functional (2.7) (or (3.1)) of the original problem, is expressed in terms of the elements of the modified problem as follows:

$$||\boldsymbol{\varepsilon}|| = u_1^2 + \mathrm{B}_1^2\Gamma_1^2 + u_2^2. \tag{8.7}$$

Thus, the modified optimal angular acceleration problem of a spacecraft has been completely solved.

If the boundary conditions on the angular velocity of a spacecraft are specified as described in Section 5, the solutions of the classical and modified optimal angular acceleration problems will coincide. The same is true when $\boldsymbol{\omega}_0\ \boldsymbol{\omega}_T\ ||\mathrm{vect}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$ (a planar Euler rotation of a spacecraft [4]); in this case, $\mathrm{B}_1^2\Gamma_1^2 = 0$ in (8.7), and the criterion (7.12) corresponds to the optimization criterion (2.7) (or (3.1)) of the classical problem. For arbitrary boundary conditions, assuming that $\int_0^1 \mathrm{B}_1^2\Gamma_1^2\,dt$ is small compared to $\int_0^1 ||\boldsymbol{\varepsilon}||\,dt$ in (8.7), the last term can be omitted. Then the modified optimal control problem of a spacecraft in the variables B, $\Gamma$, $\mathrm{B}_1$, $\Gamma_1$, $u_1$, and $u_2$ with the functional (7.12) and the expressions (7.1), (7.2), (7.13)–(7.15), and (8.4)–(8.6) corresponds to the classical optimal angular acceleration problem in the class of generalized coning motions. Based on the considerations of Sections 5 and 6 of this paper, the modified problem can be treated as a quasi-optimal angular acceleration problem of a spacecraft under arbitrary boundary conditions.

Similar to the classical problem (see formulas (4.11) and (4.12)), at any time instant, the magnitude $||\boldsymbol{\varepsilon}||$ in the modified problem lies inside a variable bounded domain determined by the constants of the problem conditions. Due to (8.4), (8.7), and the above assumptions, we obtain

$$||\boldsymbol{\varepsilon}(t)|| \le \left((c_1^2 + c_2^2)t^2 + 2 \mid c_1c_3 + c_2c_4 \mid t + c_3^2c_4^2\right)/4 \;\forall\; t \in [0,1]. \tag{8.8}$$

Here is a quasi-optimal algorithm for finding the minimum-energy angular acceleration of a spacecraft.

1. In view of (8.3)–(8.5), use the boundary conditions $\boldsymbol{\Lambda}_0$, $\boldsymbol{\Lambda}_T$ (2.5), $\boldsymbol{\omega}_0$, $\boldsymbol{\omega}_T$ (2.6), $T = 1$, formulas (7.5), (7.6), and the nine scalar equations from (7.13)–(7.15) to find the nine arbitrary constants $c_1, \ldots, c_5$, $c_7$, $c_8$, $\alpha_1$, and $\alpha_2$; then determine the functions $\mathrm{B}(t)$, $\mathrm{B}_1(t)$, $\Gamma(t)$, and $\Gamma_1(t)$.
2. Calculate the quaternion $\mathbf{Q}$ by formula (7.5).
3. Find the law of change of the angular velocity of the spacecraft by (7.1):

$$\boldsymbol{\omega}(t) = \widetilde{\mathbf{Q}} \circ \left(\mathbf{i}_1\dot{\mathrm{B}}(t)\sin\Gamma(t) + \mathbf{i}_2\dot{\mathrm{B}}(t)\cos\Gamma(t) + \mathbf{i}_3\dot{\Gamma}(t)\right) \circ \mathbf{Q}.$$

4. Find the law of change of the angular motion trajectory of the spacecraft by (7.2):

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{Q}} \circ \mathbf{exp}\{-\mathbf{i}_3\Gamma(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2(\mathrm{B}(t) - \mathrm{B}(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3\Gamma(t)/2\} \circ \mathbf{Q};$$

5. Construct the angular acceleration vector of the spacecraft by formula (8.6):

$$\boldsymbol{\varepsilon} = \widetilde{\mathbf{Q}} \circ (\mathbf{i}_1(u_1\sin\Gamma + \mathrm{B}_1\Gamma_1\cos\Gamma) + \mathbf{i}_2(u_1\cos\Gamma - \mathrm{B}_1\Gamma_1\sin\Gamma) + \mathbf{i}_3u_2) \circ \mathbf{Q}.$$

## 9. NUMERICAL EXAMPLES

In this section, we compare the numerical solution results of the original (classical) optimal angular acceleration problem of a spacecraft and the quasi-optimal solution of this problem using the analytical algorithm from Section 8. The values of $\alpha_1$, $\alpha_2$, $c_1, \ldots, c_5$, $c_7$, $c_8$ in the quasi-optimal solution of the spacecraft rotation problem with the boundary conditions (6.6), (6.7) are given by

$$\begin{aligned} &\alpha_1 = -0.0421, \quad \alpha_2 = -0.2226, \quad c_1 = 3.2902, \quad c_2 = -1.4885, \quad c_3 = 2.2113, \\ &c_4 = -1.45, \qquad c_5 = -0.4156, \quad c_6 = 0, \qquad c_7 = -0.2221, \quad c_8 = -0.9216. \end{aligned}$$

As it turned out, the solutions of the two problems are close enough. In this example, the values of the functional (3.1) of the original and modified problems are 0.4782 and 0.4797, respectively;

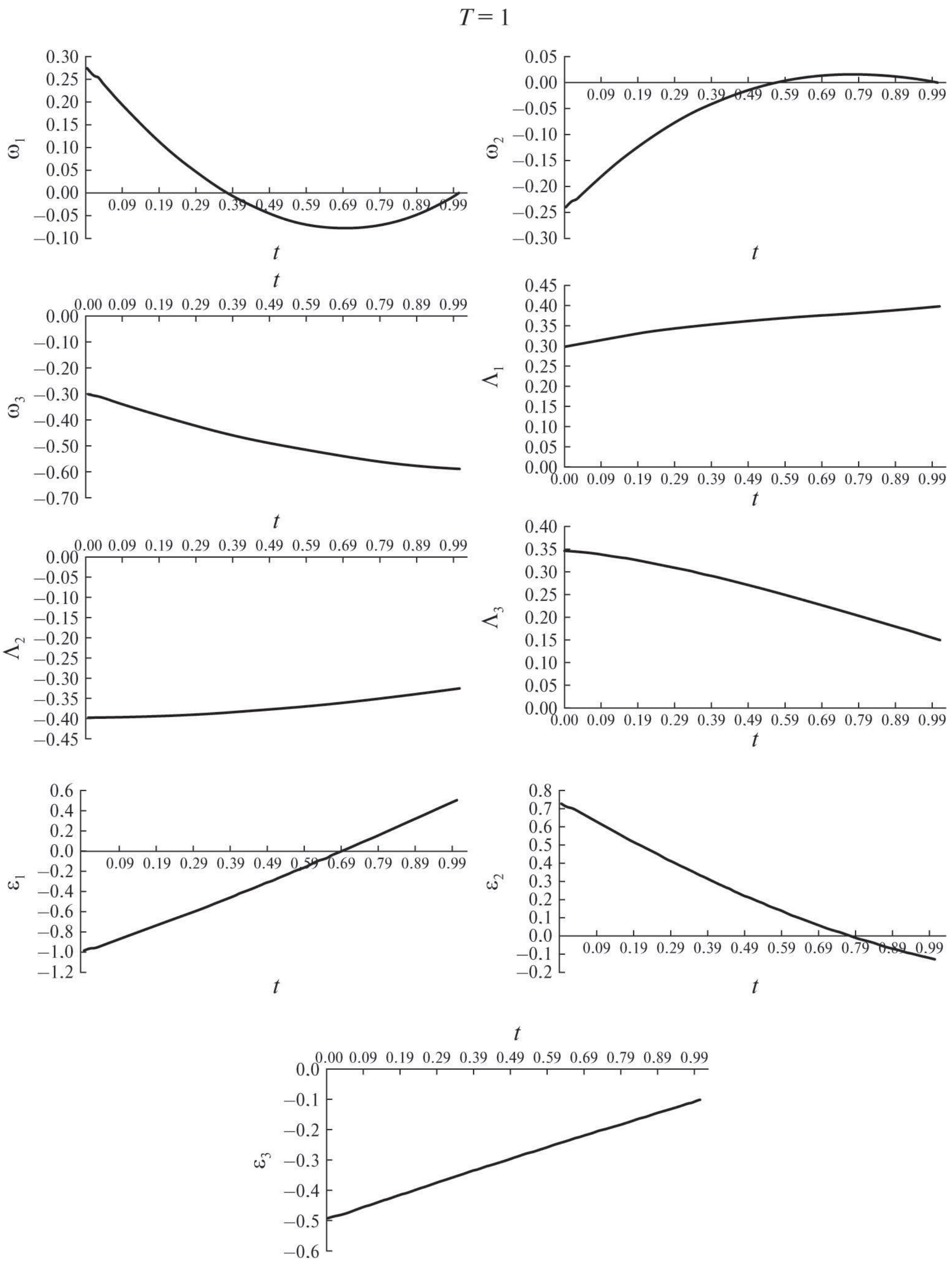


**Fig. 1.** The solution of the classical problem under arbitrary boundary conditions.

thus, the relative error between the values of (3.1) in the classical and modified problems is 0.3%. Figures 1 and 2 show the graphs of the projections of the spacecraft angular velocity vector $\omega_i(t)$, the components of the spacecraft attitude quaternion $\Lambda_i(t)$, and the projections of the spacecraft angular acceleration vector $\varepsilon_i(t)$, $i = 1, 2, 3$, depending on the time variable $t$; the graphs are quite similar for the two problems.

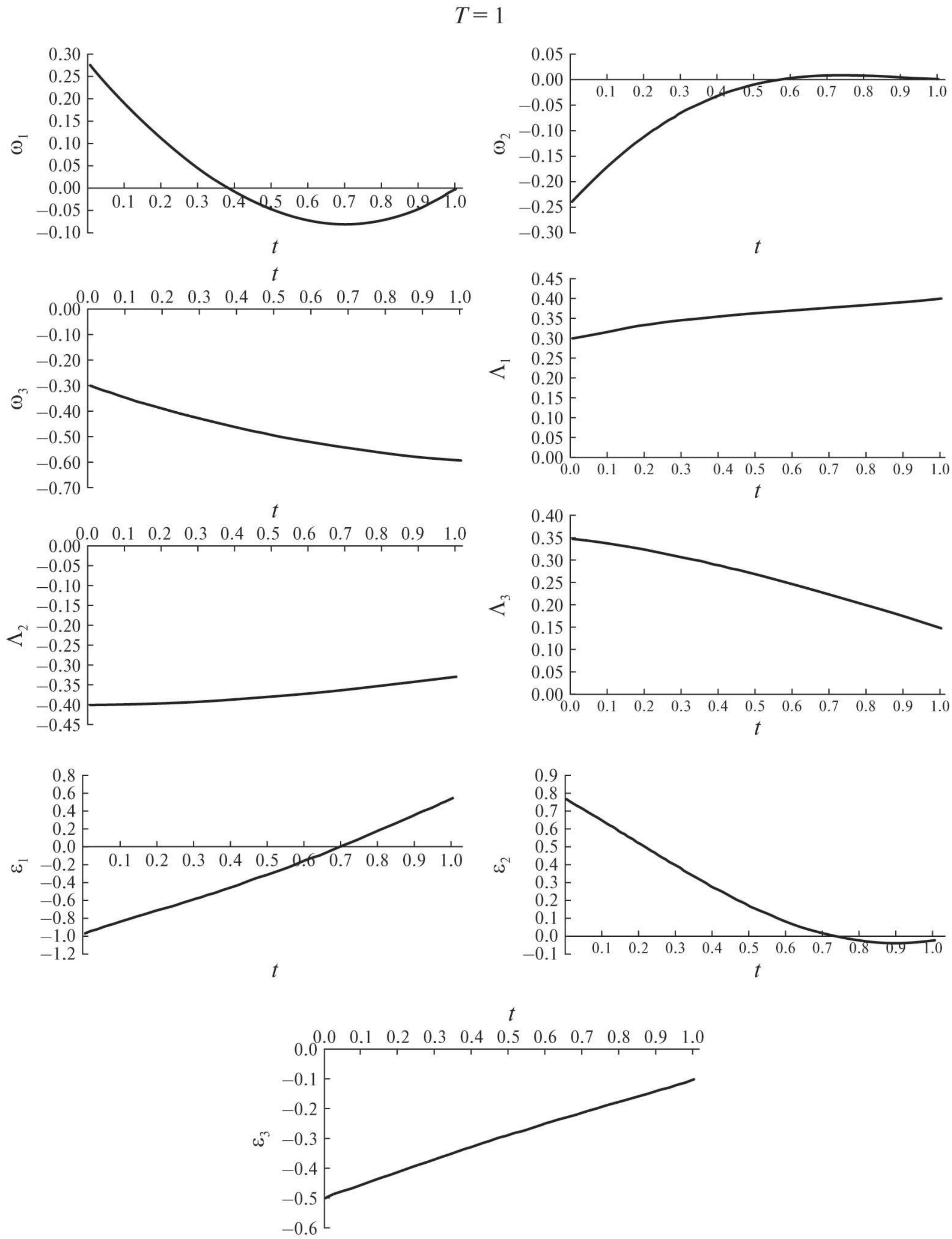


**Fig. 2.** The solution of the modified problem under arbitrary boundary conditions.

The energy-optimal angular acceleration problem of a spacecraft was also solved numerically in the following cases: the initial state is determined by (6.6), and the terminal position is set by rotating the spacecraft from the initial position to a certain angle about the Euler axis with the unit vector

$$(0.04500,\ -0.07519,\ -0.99615)\,. \tag{9.1}$$

Table 5 provides the components of the quaternions of the spacecraft terminal position for rotations to various Euler angles $\varphi$ (deg) about the vector (9.1).

**Table 5.** Formation of terminal orientation quaternions

| $\varphi^\circ$ | $\lambda_0(T)$ | $\lambda_1(T)$ | $\lambda_2(T)$ | $\lambda_3(T)$ |
|---|---|---|---|---|
| 30.0 | 0.84643 | 0.40650 | −0.31853 | 0.12982 |
| 60.0 | 0.84013 | 0.48716 | −0.21783 | −0.09703 |
| 90.0 | 0.77657 | 0.53461 | −0.10229 | −0.31727 |
| 120.0 | 0.66009 | 0.54564 | 0.02023 | −0.51589 |
| 150.0 | 0.49863 | 0.51948 | 0.14136 | −0.67935 |
| 180.0 | 0.30318 | 0.45792 | 0.25286 | −0.79652 |

Next, Table 6 shows the values of the functional (3.1) in the classical and modified optimal angular acceleration problems when transferring the spacecraft from the initial (6.6) to terminal state $\mathbf{\Lambda}_T$ according to Table 5, with $\boldsymbol{\omega}_T = (0.0,\ 0.0,\ -0.59)$, also indicating the absolute and relative differences between them, $\Delta J = J^{\text{modif}} - J^{\text{classic}}$ and $\left(\Delta J / J^{\text{classic}}\right) \times 100\%$.

**Table 6.** Values of the functionals and their differences

| Functionals and differences of their values | $\varphi = 30.0^\circ$ | $\varphi = 60.0^\circ$ | $\varphi = 90.0^\circ$ | $\varphi = 150.0^\circ$ | $\varphi = 180.0^\circ$ |
|---|---|---|---|---|---|
| $J^{\text{classic}}$ | 0.52385 | 4.63277 | 15.31437 | 56.39081 | 86.78094 |
| $J^{\text{modif}}$ | 0.52510 | 4.63724 | 15.32882 | 56.52086 | 87.51533 |
| $\Delta J$ | 0.00125 | 0.00447 | 0.01445 | 0.13005 | 0.73439 |
| % | 0.24 | 0.10 | 0.09 | 0.23 | 0.85 |

Finally, Table 7 provides the same data for the cases where the terminal angular velocity of the spacecraft is given by the vector $\boldsymbol{\omega}_T = (0,0,\ 0,0,\ 0,0)$, i.e., when transferring the spacecraft to a state of rest.

**Table 7.** Values of the functionals and their differences

| Functionals and differences of their values | $\varphi = 90.0^\circ$ | $\varphi = 120.0^\circ$ | $\varphi = 150.0^\circ$ | $\varphi = 180.0^\circ$ |
|---|---|---|---|---|
| $J^{\text{classic}}$ | 24.25074 | 45.17597 | 72.66431 | 106.71186 |
| $J^{\text{modif}}$ | 24.28745 | 45.19513 | 72.77169 | 107.40843 |
| $\Delta J$ | 0.03672 | 0.01916 | 0.10738 | 0.69657 |
| % | 0.15 | 0.04 | 0.15 | 0.65 |

The calculations for various boundary conditions show the closeness of the solutions of the classical and modified angular acceleration problems of a spacecraft. Therefore, the solution of the modified problem can be treated as a quasi-optimal solution of the classical optimal angular acceleration problem of a spacecraft.

## 10. CONCLUSIONS

The analytical quasi-optimal algorithm for solving the kinematic optimal, in terms of energy consumption, program angular acceleration problem of a spacecraft under arbitrary boundary conditions is theoretically justified and yields the optimal angular acceleration of a spacecraft with acceptable accuracy. This algorithm does not require a numerical solution of the boundary value problem of the maximum principle or any complex numerical solution; it provides ready-made analytical laws of quasi-optimal program control and program trajectory change that can be installed aboard a spacecraft. In this regard, the above results can be applied to nanoclass spacecraft with limited computing power. Recall that the Kotelnikov–Study transference principle [13] allows extending quaternionic formulas describing the control of angular motion to biquaternionic ones describing the control of the general spatial motion of a solid body. Based on the above results, this principle can be used to obtain an analytical quasi-optimal, in terms of energy consumption, program control algorithm for the spatial motion (maneuvering) of a spacecraft.

## REFERENCES

1. Scrivener, S.L. and Thompson, R.C., Survey of Time-Optimal Attitude Maneuvers, *J. Guidance, Control Dynam.*, 1994, vol. 17, no. 2, pp. 225–233.
2. Junkins, J.L. and Turner, J.D., *Optimal Spacecraft Rotational Maneuvers*, New York: Elsevier, 1986.
3. Crassidis, J.L. and Markley, F.L., *Fundamentals of Spacecraft Attitude Determination and Control*, New York: Springer, 2014.
4. Branets, V.N. and Shmyglevskii, I.P., *Primenenie kvaternionov v zadachakh orientatsii tverdogo tela* (Application of Quaternions to Solid Body Attitude Problems), Moscow: Nauka, 1973.
5. Somov, Ye.I., Butyrin, S.A., and Somova, T.E., Autonomous Guidance and Control of Mini-Satellites in a Low-Orbit Constellation during Areal Scanning Survey, *Proceedings of the 30th Saint Petersburg International Conference on Integrated Navigation Systems*, May 29–31, 2023 Saint Petersburg, Russia.
6. Ivanova, E.A., On One Approach to Solving the Darboux Problem, *Izv. Ross. Akad. Nauk. Mekh. Tverd. Tel.*, 2000, no. 1, pp. 45–52.
7. Molodenkov, A.V. and Perelyaev, S.E., Solution of Approximate Equation for Modified Rodrigues Vector and Attitude Algorithm Design, *J. Guidance, Control Dynam.*, 2021, vol. 44, no. 6, pp. 1224–1227.
8. Molodenkov, A.V. and Sapunkov, Y.G., Analytical Quasi-Optimal Solution of the Problem of the Time-Optimal Rotation of a Spacecraft, *J. Comput. Syst. Sci. Int.*, 2021, vol. 60, no. 4, pp. 639–653.
9. Sapunkov, Ya.G. and Molodenkov, A.V., Quasioptimal Spacecraft Attitude Control Constructed According to the Poinsot Concept, *Aerospace*, 2023, vol. 10, no. 5, pp. 402–417.
10. Sapunkov, Ya.G. and Molodenkov, A.V., Quasi-Optimal Angular Acceleration of a Spacecraft Based on the Poinsot Concept, *Autom. Remote Control*, 2025, vol. 86, no. 3, pp. 217–234.
11. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., and Mishchenko, E.F., *The Mathematical Theory of Optimal Processes*, Oxford: Pergamon Press, 1964.
12. Molodenkov, A.V. and Sapunkov, Ya.G., Solution of the Problem of the Optimal Re-Orientation of a Rotationing Spherically Symmetric Spacecraft with Pulse Control, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2007, vol. 10, pp. 78–84.
13. Chelnokov, Yu.N., *Kvaternionnye i bikvaternionnye modeli i metody mekhaniki tverdogo tela i ikh prilozheniya* (Quaternionic and Biquaternionic Models and Methods of Solid Mechanics and Their Applications), Moscow: Fizmatlit, 2006.
14. Banit, Yu.R., Belyaev, M.Yu., Dobrinskaya, T.A., et al., Estimating the Inertia Tensor of the International Space Station on the Base of the Telemetry Information, *Preprints of the Keldysh Institute of Applied Mathematics*, 2002, no. 57.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 510–518.


CONTROL IN TECHNICAL SYSTEMS

# Assessment of the Feasibility of Intercepting an Intensively Maneuvering High-Speed Aircraft by an Inertial Interceptor with a Time-Varying Control Law

**V. S. Verba**[*,a], **V. I. Merkulov**[*,b], **and D. V. Zakomoldin**[**,c]

[*] *JSC Concern Vega, Moscow, Russia*
[**] *Military Academy of Aerospace Defense named after Marshal of the Soviet Union G.K. Zhukov, Tver, Russia*
*e-mail:* [a]*vvs.msk@gmail.com,* [b]*ilya-zagrebelnyi@mail.ru,* [c]*denjuga68@yandex.ru*

Received July 30, 2025
Revised October 10, 2025
Accepted October 15, 2025

**Abstract**—This paper presents a time-varying control law designed for an interceptor's guidance to an intensively maneuvering aircraft (aerial target). The control law takes into account the mismatch between the dynamic properties of the interceptor and the target. Simulation results are provided, indicating the potential feasibility of intercepting the target and improving guidance accuracy.

*Keywords*: interceptor control law, maneuvering aerial target, mismatch of dynamic properties, miss distance

**DOI:** 10.7868/S1608303226060036

## 1. INTRODUCTION

An effective solution of the problem of intercepting aerospace attack means is necessitated by the requirement to protect important objects of the defending side from potential destruction.

In view of the continuous improvement trend of aerospace attack means toward increasing their speed and maneuverability characteristics [1–7], the inertia of interceptors can be a significant factor affecting their potentially achievable guidance performance to an intensively maneuvering high-speed aircraft (HSA) [8].

Two possible approaches to considering the interceptor's inertia were discussed in [8]. The first approach is to generate a control signal based on the predicted rather than current values of the HSA state coordinates. The second approach involves an appropriate transformation of input actions that compensates for the interceptor's inertia via control signal corrections covering high derivatives of the state coordinates. However, these approaches complicate the information support system due to the need to estimate high derivatives of the line-of-sight (LoS) angular rate [8].

Therefore, it is topical to design alternative control signals for an interceptor's guidance to an HSA with the following requirements [9]:

—The control signal should take into account the mismatches between the dynamic properties of the interceptor and HSA.

—The control signal should be time-varying to redistribute control priorities from eliminating angle guidance errors at the initial guidance stage to eliminating the LoS angular rate when approaching the HSA.

This paper aims to design a time-varying control signal for an inertial interceptor with the potential feasibility of intercepting an HSA.

## 2. PROBLEM STATEMENT. CONTROL SIGNAL DESIGN

The design procedure is based on the local optimization approach of the statistical theory of optimal control (STOC) within the linear-quadratic-Gaussian problem framework. For a given system described by [9]

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \xi(t), \tag{1}$$

which includes a subsystem (aerial target) of the form

$$\dot{\mathbf{x}}_{\text{tar}}(t) = \mathbf{F}_{\text{tar}}(t)\mathbf{x}_{\text{tar}}(t) + \xi_{\text{tar}}(t), \tag{2}$$

generating input actions for another subsystem (interceptor)

$$\dot{\mathbf{x}}_{\text{int}}(t) = \mathbf{F}_{\text{int}}\mathbf{x}_{\text{int}}(t) + \mathbf{B}_{\text{int}}\mathbf{u}(t) + \xi_{\text{int}}(t), \tag{3}$$

under available measurements

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \xi_{\text{z}}(t), \tag{4}$$

this approach yields the optimal control law

$$\mathbf{u}(t) = \mathbf{K}^{-1}\mathbf{B}_{\text{int}}^T\left[\mathbf{Q}\Delta\hat{\mathbf{x}}(t) - \mathbf{G}\hat{\mathbf{s}}_{\text{int}}(t)\right], \tag{5}$$

in terms of the minimum value of the performance functional

$$\begin{aligned} I = M\Bigg\{ &(\mathbf{x}_{\text{tar}}(t) - \mathbf{x}_{\text{int}}(t))^T\mathbf{Q}(\mathbf{x}_{\text{tar}}(t) - \mathbf{x}_{\text{int}}(t)) + 2(\mathbf{x}_{\text{tar}}(t) - \mathbf{x}_{\text{int}}(t))^T\mathbf{G}\mathbf{s}_{\text{int}}(t) \\ &+ \mathbf{s}_{\text{int}}^T(t)\mathbf{Q}\mathbf{s}_{\text{int}}(t) + \int_0^t \mathbf{u}^T(t)\mathbf{K}\mathbf{u}(t)\,dt \Bigg\}, \end{aligned} \tag{6}$$

where

$$\begin{gathered} \mathbf{x} = \begin{bmatrix}\mathbf{x}_{\text{tar}}^T & \mathbf{x}_{\text{int}}^T\end{bmatrix}^T, \quad \mathbf{F} = \begin{bmatrix}\mathbf{F}_{\text{tar}} & \mathbf{O}_1 \\ \mathbf{O}_2 & \mathbf{F}_{\text{int}},\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}\mathbf{O}_3\mathbf{B}_{\text{int}}\end{bmatrix}, \quad \mathbf{s}_{\text{int}} = \left(\mathbf{F}_{\text{tar}} - \mathbf{F}_{\text{int}}\right)\mathbf{x}_{\text{tar}}, \\ \xi = \begin{bmatrix}\xi_{\text{tar}}^T & \xi_{\text{int}}^T\end{bmatrix}^T, \quad \Delta\hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{tar}} - \hat{\mathbf{x}}_{\text{int}}, \end{gathered} \tag{7}$$

The other designations are as follows: $\mathbf{x}_{\text{int}}$ and $\mathbf{x}_{\text{tar}}$ are the $n$-dimensional state vectors of the interceptor and target, respectively; $\mathbf{F}_{\text{int}}$ and $\mathbf{F}_{\text{tar}}$ are the internal connection matrices of the subsystems (2) and (3), characterizing their dynamics; $\mathbf{u}$ is the control vector; $\mathbf{B}_{\text{int}}$ is the control efficiency matrix; $\mathbf{z}$ is the measurement vector; $\xi$ is the noise vector; $\xi_{\text{tar}}$, $\xi_{\text{int}}$, and $\xi_z$ are the noise vectors of the states (2), (3) and measurements (4), respectively; $\mathbf{H}$ is the connection matrix between $\mathbf{z}$ and $\mathbf{x}$; $\mathbf{Q}$ and $\mathbf{K}$ are nonnegative definite penalty matrices for control accuracy and control costs, respectively; $\mathbf{G}$ is the mutual connection matrix between $\Delta\mathbf{x}$ and $\mathbf{s}_{\text{int}}$; $\mathbf{s}_{\text{int}}$ is the mismatch vector between the dynamic properties of the subsystems; $t$ is the current time; $\mathbf{O}_1$–$\mathbf{O}_3$ are zero matrices of compatible dimensions; finally, $M\{\bullet\}$ denotes the expectation operator.

To shorten the mathematical derivations, the dependence on time in (1)–(7) will be omitted.

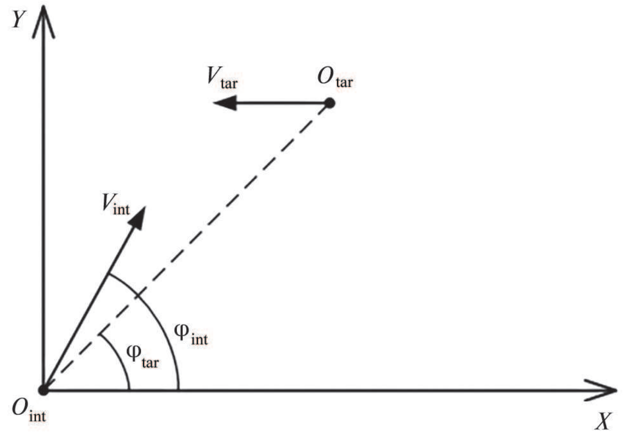


**Fig. 1.** The mutual arrangement of the target and interceptor.

The geometry of the target's and interceptor's angles is shown in Fig. 1. In this figure, $O_{\text{int}}$ and $O_{\text{tar}}$ are the positions of the interceptor and target, respectively; $\mathbf{V}_{\text{tar}}$ and $\mathbf{V}_{\text{int}}$ are the velocity vectors of the target and interceptor, respectively; $\varphi_{\text{int}}$ is the interceptor's line-of-sight (LoS) angle (the angle between its velocity vector and the horizontal axis); finally, $\varphi_{\text{tar}}$ is the HSA's flight direction angle (the angle between the interceptor–target LoS and the horizontal axis).

The control design problem is formulated as follows.

Let an HSA (target) be described by a system of equations of the form:

$$\begin{aligned} &\dot{\varphi}_{\text{tar}} = \omega_{\text{tar}}, && \varphi_{\text{tar}}(0) = \varphi_{\text{tar}0}, \\ &\dot{\omega}_{\text{tar}} = -\frac{2\dot{R}}{R}\omega_{\text{tar}} + \frac{1}{R}(j_{\text{tar}} - j_{\text{int}}) + \xi_{\text{tar}}, && \omega_{\text{tar}}(0) = \omega_{\text{tar}0}. \end{aligned} \tag{8}$$

For an inertial interceptor corresponding to the model

$$\begin{aligned} &\dot{\varphi}_{\text{int}} = \omega_{\text{int}}, && \varphi_{\text{int}}(0) = \varphi_{\text{int}0}, \\ &\dot{\omega}_{\text{int}} = -\frac{1}{T_{\text{int}}}\omega_{\text{int}} + \frac{b}{T_{\text{int}}} j_{\text{int}} + \xi_{\text{int}}, && \omega_{\text{int}}(0) = \omega_{\text{int}0}, \end{aligned} \tag{9}$$

with available measurements of the form

$$\varphi_z = \varphi_{\text{tar}} + \xi_z, \tag{10}$$

it is required to find a control signal $j_{\text{int}}$ minimizing the performance functional

$$\begin{aligned} I = M\Bigg\{ &\begin{bmatrix} \varphi_{\text{tar}} - \varphi_{\text{int}} \\ \omega_{\text{tar}} - \omega_{\text{int}} \end{bmatrix}^T \begin{bmatrix} \dfrac{q_{11}}{R} & q_{12} \\ q_{21} & \dfrac{q_{22}}{R} \end{bmatrix} \begin{bmatrix} \varphi_{\text{tar}} - \varphi_{\text{int}} \\ \omega_{\text{tar}} - \omega_{\text{int}} \end{bmatrix} \\ &+ 2\begin{bmatrix} \varphi_{\text{tar}} - \varphi_{\text{int}} \\ \omega_{\text{tar}} - \omega_{\text{int}} \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} 0 \\ \left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\dot{R}}{R}\right)\omega_{\text{tar}} \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\dot{R}}{R}\right)\omega_{\text{tar}} \end{bmatrix}^T \begin{bmatrix} \dfrac{q_{11}}{R} & q_{12} \\ q_{21} & \dfrac{q_{22}}{R} \end{bmatrix} \begin{bmatrix} 0 \\ \left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\dot{R}}{R}\right)\omega_{\text{tar}} \end{bmatrix} + \int_0^t j_{\text{int}}^2 k_{\text{j}}\, dt \Bigg\}. \end{aligned} \tag{11}$$

The expressions (8)–(11) have the following notation: $\varphi_{\text{tar}}$ and $\varphi_{\text{int}}$ are the target's LoS angle and the interceptor's flight direction angle, respectively, in the frame associated with the interceptor's center of mass; $\omega_{\text{tar}}$ and $\omega_{\text{int}}$ are the LoS angular rates of the target and interceptor, respectively; $\varphi_{\text{tar}0}$, $\varphi_{\text{int}0}$ and $\omega_{\text{tar}0}$, $\omega_{\text{int}0}$ are the initial values of the corresponding angles and their rates, respectively; $b$ and $T_{\text{int}}$ are the interceptor's gain and time constant, respectively; $j_{\text{tar}}$ and $j_{\text{int}}$ are the lateral accelerations of the target and interceptor, respectively; $R$ and $\dot{R}$ are the interceptor's range to the target and its rate of change, respectively; $\xi_{\text{tar}}$ and $\xi_{\text{int}}$ are the disturbances of the target and interceptor trajectories (centered Gaussian noises); $\frac{q_{11}}{R}$, $q_{12}$, $q_{21}$, and $\frac{q_{22}}{R}$ are nonnegative elements of the matrix $\mathbf{Q}$; finally, $g_{11}$, $g_{12}$, $g_{21}$, and $g_{22}$ are nonnegative elements of the matrix $\mathbf{G}$. Note that in practice, the most difficult task is to find the elements of the matrices $\mathbf{Q}$, $\mathbf{G}$, and $\mathbf{K}$; here, the solution is often empirical. The most well-known method involves the principle of equal strength: the products of the squared maximum permissible errors and the corresponding penalty coefficients are supposed to be the same for all coordinates. By setting the maximum permissible errors (variances) and a penalty coefficient for a single coordinate, one can preliminarily estimate the penalty coefficients for the other coordinates. Then the resulting coefficients are refined based on simulation.

The time-varying target model (8) is highly universal due to, first, consideration of the influence of the range and approach speed on the angular coordinates and, second, the realizability of maneuvers of almost any complexity by manipulating the laws of change of $j_{\text{tar}}$.

While being fairly simple, the kinematic models (8) and (9) describe the mutual arrangement (geometry) of the interceptor and HSA and the geometry's rate of change with sufficient accuracy for the problem under consideration [8].

The performance functional (11) possesses the following features. Besides the requirements for control accuracy and control costs (the first and fourth terms of the functional), it covers the requirement to reduce the mutual influence of control errors and the mismatch between the dynamic properties of the interceptor and HSA (the second term), as well as the requirement to compensate for the above mismatch (the third term). In addition, the first and third terms of the performance functional (11) contain the time-varying coefficients of the control accuracy penalty matrix, $\frac{q_{11}}{R}$ and $\frac{q_{22}}{R}$, which yield a time-varying control signal even when using the simplified time-invariant interceptor model (9); for details, see [9].

Models (8) and (9) are linear, the noises are Gaussian, and the performance functional (11) is quadratic. Therefore, based on the separation theorem and the statistical equivalence principle, control signals will be designed using the deterministic models (8) and (9) with $\xi_{\text{tar}} = 0$, $\xi_{\text{int}} = 0$, and the coordinates of (8) and (9) replaced by their optimal estimates.

By matching (8) with (2), (9) with (3), and (11) with (6) and (7), we obtain

$$\begin{gathered}
\mathbf{x}_T = \begin{bmatrix} \hat{\varphi}_{\text{tar}} \\ \hat{\omega}_{\text{tar}} \end{bmatrix}, \quad \mathbf{x}_{\text{int}} = \begin{bmatrix} \hat{\varphi}_{\text{int}} \\ \hat{\omega}_{\text{int}} \end{bmatrix}, \quad \mathbf{B}_{\text{int}} = \begin{bmatrix} 0 \\ \dfrac{b}{T_{\text{int}}} \end{bmatrix}, \quad \mathbf{u} = \hat{j}_{\text{int}}, \\
\mathbf{Q} = \begin{bmatrix} \dfrac{q_{11}}{R} & q_{12} \\ q_{21} & \dfrac{q_{22}}{R} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}, \quad \mathbf{F}_{\text{int}} = \begin{bmatrix} 0 & 1 \\ 0 & -\dfrac{1}{T_{\text{int}}} \end{bmatrix}, \\
\mathbf{F}_{\text{tar}} = \begin{bmatrix} 0 & 1 \\ 0 & -\dfrac{2\hat{\dot{R}}}{\hat{R}} \end{bmatrix}, \quad \mathbf{s}_{\text{int}} = \begin{bmatrix} 0 \\ \left( \dfrac{1}{T_{\text{int}}} - \dfrac{2\hat{\dot{R}}}{\hat{R}} \right) \hat{\omega}_{\text{tar}} \end{bmatrix}, \quad \mathbf{K} = k_j.
\end{gathered} \tag{12}$$

Based on direct analysis of $\mathbf{s}_{\text{int}}$, we draw the following conclusions: at long ranges (as $\hat{\omega}_{\text{tar}} \to 0$), the mismatch of the dynamic properties of the HSA and interceptor has almost no effect; at short ranges, characterized by an increasing value of $\hat{\omega}_{\text{tar}}$, the significance of $\mathbf{s}_{\text{int}}$ grows gradually.

Substituting (12) into (5) gives

$$\jmath_{\text{int}} = \frac{1}{k_{\text{j}}}\begin{bmatrix} 0 \\ \dfrac{b}{T_{\text{int}}} \end{bmatrix}^T \left[\begin{bmatrix} \dfrac{q_{11}}{\hat{R}} & q_{12} \\ q_{21} & \dfrac{q_{22}}{\hat{R}} \end{bmatrix}\begin{bmatrix} \hat{\varphi}_{\text{tar}} - \hat{\varphi}_{\text{int}} \\ \hat{\omega}_{\text{tar}} - \hat{\omega}_{\text{int}} \end{bmatrix} - \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}\begin{bmatrix} 0 \\ \left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\hat{\dot{R}}}{\hat{R}}\right)\hat{\omega}_{\text{tar}} \end{bmatrix}\right]$$

$$= \begin{bmatrix} 0 & \dfrac{b}{k_{\text{j}}T_{\text{int}}} \end{bmatrix}\begin{bmatrix} \dfrac{q_{11}}{\hat{R}}\Delta\hat{\varphi} + q_{12}\Delta\hat{\omega} - g_{12}\left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\hat{\dot{R}}}{\hat{R}}\right)\hat{\omega}_{\text{tar}} \\ q_{21}\Delta\hat{\varphi} + \dfrac{q_{22}}{\hat{R}}\Delta\hat{\omega} - g_{22}\left(\dfrac{1}{T_{\text{int}}} - \dfrac{2\hat{\dot{R}}}{\hat{R}}\right)\hat{\omega}_{\text{tar}} \end{bmatrix},$$

then

$$\jmath_{\text{int}} = \frac{bq_{21}}{k_{\text{j}}\,T_{\text{int}}}\Delta\hat{\varphi} + \frac{bq_{22}}{k_{\text{j}}\,T_{\text{int}}\hat{R}}\Delta\hat{\omega} - \frac{bg_{22}}{k_{\text{j}}\,T_{\text{int}}}\left(\frac{1}{T_{\text{int}}} - \frac{2\hat{\dot{R}}}{\hat{R}}\right)\hat{\omega}_{\text{tar}},$$
$$\Delta\hat{\varphi} = \hat{\varphi}_{\text{tar}} - \hat{\varphi}_{\text{int}}, \quad \Delta\hat{\omega} = \hat{\omega}_{\text{tar}} - \hat{\omega}_{\text{int}}. \tag{13}$$

Direct analysis of (13) allows us to conclude the following:

1. The control signal (13) is all-altitude. This property is ensured by using the interceptor's lateral acceleration, instead of the rudder deflection angle, as the control signal; recall that the rudder's efficiency depends on air density, which varies with altitude.

2. The control signal consists of three components, with the first term $\frac{bq_{21}}{k_{\text{j}}\,T_{\text{int}}}\Delta\hat{\varphi}$ representing a modification of the direct method, and the sum of the first and second terms $\frac{bq_{21}}{k_{\text{j}}\,T_{\text{int}}}\Delta\hat{\varphi} + \frac{bq_{22}}{k_{\text{j}}\,T_{\text{int}}\hat{R}}\Delta\hat{\omega}$ representing the method of sequential predictions. At long ranges as $\hat{\omega}_{\text{tar}} \to 0$ and $\frac{bq_{22}}{k_{\text{j}}\,T_{\text{int}}\hat{R}} \to 0$, the requirement to eliminate the angle errors $\Delta\varphi$ becomes most significant. At short ranges, along with the growing influence of the second term due to an increase of $\frac{bq_{22}}{k_{\text{j}}\,T_{\text{int}}\hat{R}}$ for smaller $R$, the significance of the third term also raises with increasing $\hat{\omega}_{\text{tar}}$. Therefore, the control signal is time-varying. The factor $\frac{1}{T_{\text{int}}} - \frac{2\dot{R}}{R}$ in the third term of (13), which characterizes the mismatch between the dynamic properties of the interceptor and HSA, actually account for this mismatch in the control signal. It increases for short ranges, predetermining the growing influence of the target's maneuver. Since the mismatch between the dynamic properties of the interceptor and HSA is included in the control signal, it can be compensated for by generating an appropriate control action.

3. To form the required control signal (13), it is necessary to estimate the range, approach speed, onboard bearing of the HSA, and the LoS angular rate, which imposes no restrictions on its practical realization.

4. The weight of errors in the control signal (13) is determined not by the absolute values of the penalty coefficients, but by their ratios $\frac{q_{21}}{k_{\text{j}}}$, $\frac{q_{22}}{k_{\text{j}}}$, and $\frac{g_{22}}{k_{\text{j}}}$.

## 3. EFFICIENCY ANALYSIS OF THE CONTROL SIGNAL

Computer simulations are intended to verify the potential feasibility of intercepting a complexly maneuvering HSA, with a change in the sign of the angular rate derivative, using an inertial interceptor with the control law (13).

The simulations were carried out for HSA interception on head-on crossing courses, provided that the target is faster than the interceptor, with a limitation on the magnitude of the control

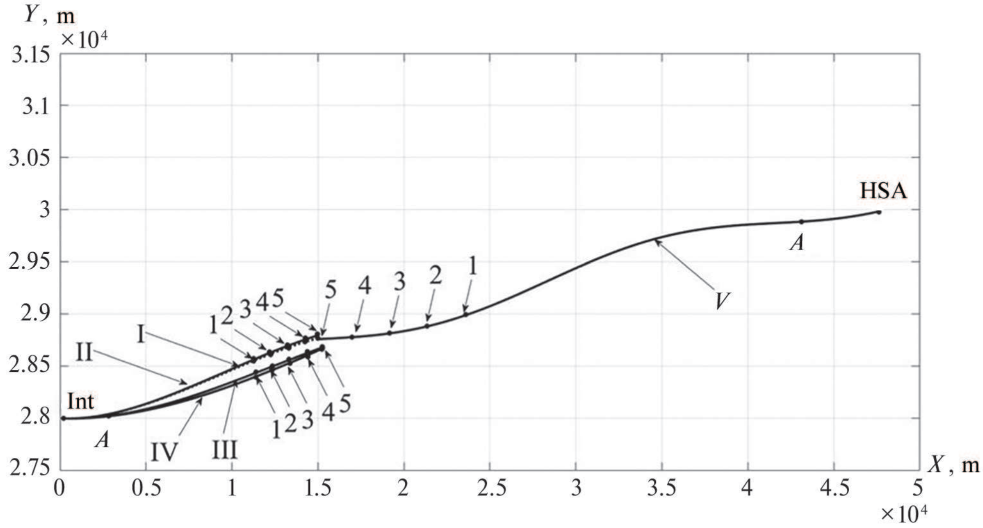


**Fig. 2.** Trajectories of HSA and interceptor during the interception process.

signal and without estimation errors, i.e., $\hat{R} = R$, $\hat{\dot{R}} = \dot{R}$, $\hat{\varphi}_{\text{tar}} = \varphi_{\text{tar}}$, $\hat{\varphi}_{\text{int}} = \varphi_{\text{int}}$, $\hat{\omega}_{\text{tar}} = \omega_{\text{tar}}$, and $\hat{\omega}_{\text{int}} = \omega_{\text{int}}$. This assumption is valid since the desired control signal ensures a potential interception (in the absence of noise and disturbances). When implementing this approach, the dynamic errors of the guidance system can be preliminarily estimated, in the form of current and final miss distances, from the simulation results. The estimation of fluctuation errors, when the noises are nonzero, requires additional studies.

To assess the possibility of simplifying the control law (13), we tested three schemes of the guidance law:

scheme I, defined by the relation (13);

scheme II, defined by the relation

$$\hat{j}_{\text{int}} = \frac{bq_{21}}{k_{\text{j}}\,T_{\text{int}}}\Delta\hat{\varphi} + \frac{bq_{22}}{k_{\text{j}}\,T_{\text{int}}\hat{R}}\Delta\hat{\omega}; \tag{14}$$

scheme III, defined by the relation

$$\hat{j}_{\text{int}} = \frac{bq_{21}}{k_{\text{j}}\,T_{\text{int}}}\Delta\hat{\varphi} \tag{15}$$

(under different values of the interceptor's time constant $T_{\text{int}}$).

The efficiency was assessed by the current and final miss distances, as well as the interceptor's lateral accelerations realized during the interception process.

The simulation results are shown in Figs. 2–5, namely:

—the motion of the HSA and interceptor along the entire interception trajectory in the vertical plane (Fig. 2);

—the trajectories of the HSA and interceptor at the final guidance stage (Fig. 3);

—the current $h$ and final $h_{\text{fin}}$ miss distances realized in the guidance schemes $j_1$, $j_2$, and $j_3$ (Fig. 4);

—the realized lateral accelerations (Fig. 5).

In Figs. 2 and 3 for the vertical plane $XOY$, numerals I, II, and III denote the interceptor's trajectories when realizing the guidance schemes $j_1$, $j_2$, and $j_3$, respectively, for a time constant of

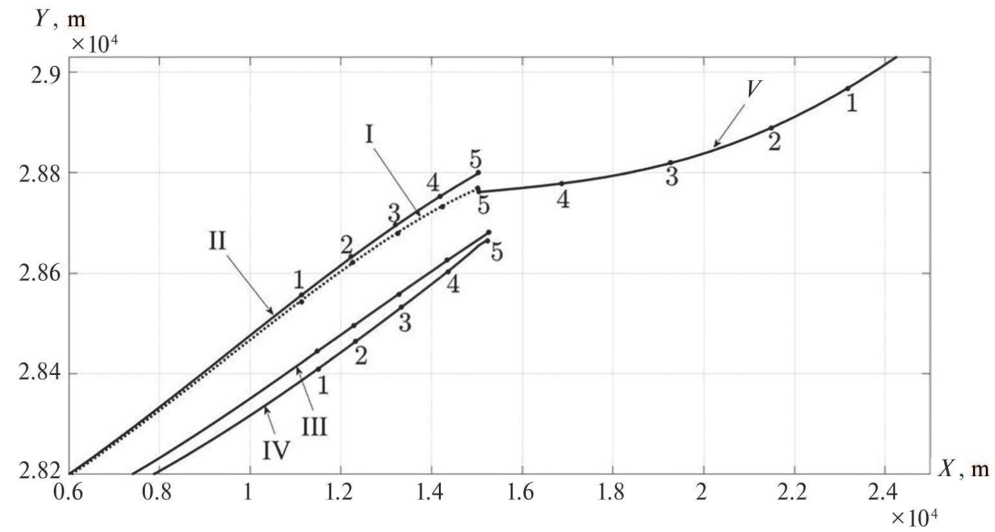


**Fig. 3.** Trajectories of HSA and interceptor at the final guidance stage.

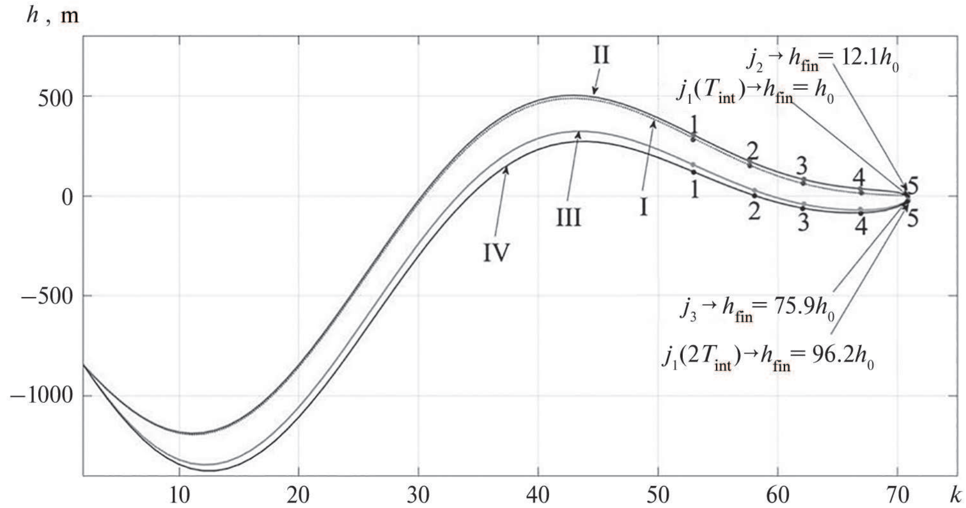


**Fig. 4.** The current and final miss distances of the interceptor with different guidance laws.

$T_{\text{int}} = 2$ and $b = 2$; the numeral IV corresponds to the interceptor's trajectory when realizing the guidance scheme $j_1$ for a time constant of $2\,T_{\text{int}}$; $V$ indicates the trajectory of the maneuvering HSA with a change in the sign of the derivatives of the state coordinates; points HSA and Int are the initial positions of the HSA and interceptor, respectively; finally, points $A$, 1, 2, 3, 4, and 5 are the current positions of the HSA and interceptor at the same time instants.

In Fig. 4, curves I–IV are the current miss distances $h$ when realizing the guidance schemes $j_1$, $j_2$, and $j_3$; $h_{\text{fin}}$ is the final miss distance; finally, $h_0$ is the final miss distance when realizing the guidance scheme $j_1$ for a time constant $T_{\text{int}}$.

In Fig. 5, curves I–IV are the current lateral accelerations arising during the interception process.

Direct analysis of the simulation results leads to the following conclusions:

1. The time-varying control law $j_1$ (13), which takes into account the mismatch between the dynamic properties of the target and interceptor, improves the guidance accuracy of the inertial interceptor. At the same time, an attempt to simplify the control law by using the lateral acceler-

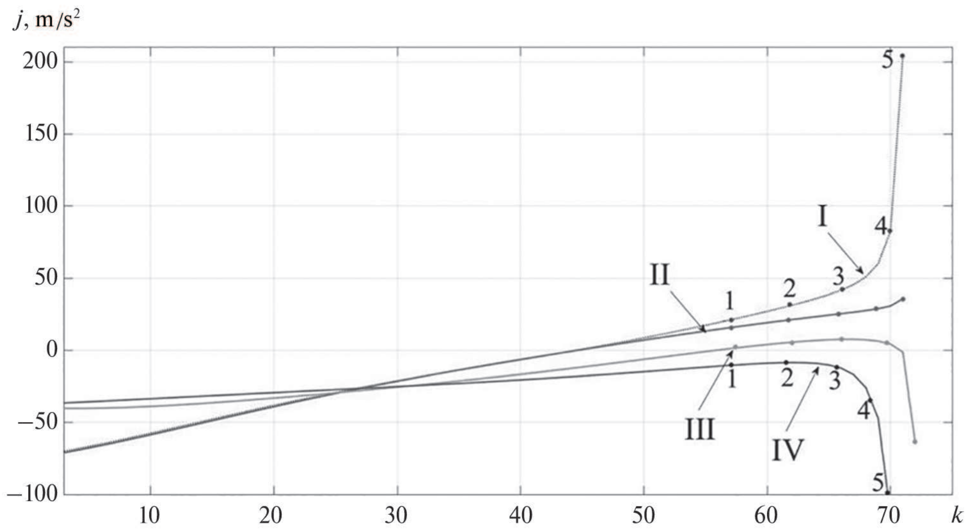


**Fig. 5.** The lateral accelerations of the interceptor with different guidance laws.

ations (14) and (15) decreases the guidance accuracy, in terms of the final miss distance $h_{\text{fin}}$ of the interceptor (see curves I–IV in Fig. 4).

2. The time-varying nature of the guidance law is significant for guidance efficiency. This is confirmed by curves I–III in Fig. 2. According to the analysis, the time-varying guidance laws along trajectories I and II with the above initial conditions, after the interceptor covers the first 3 km (point $A$ in Fig. 2, about 40 km to the target), are separated from curve III, which is the result of interception by the time-invariant guidance law.

3. The additional consideration of the mismatch between the dynamic properties of the interceptor and target in the law $j_1$ has an effect at the final interception stage. When the interceptor passes the mark of about 7 km, trajectories I and II in Fig. 3 are separated. This interceptor position corresponds to a range of about 25 km to the target. Therefore, it is crucial to account for the mismatch between the dynamic properties of the interceptor and target in the guidance law.

4. The current miss distances are convergent for all interceptor control laws $j_1$–$j_3$, as evidenced by the joint analysis of curves I–IV in Fig. 4.

5. Doubling the interceptor's inertia, characterized by the time constant $T_{\text{int}}$, reduces the interception efficiency, as evidenced by the joint analysis of curves I and IV in all figures.

6. There are no restrictions on the practical realization of the control law (13). Indeed, on the one hand, the existing range, speed, and angular coordinate sensors can be used for the information support of (13); on the other, guidance is performed within realizable lateral accelerations (curves I–IV in Fig. 5).

## 4. CONCLUSIONS

In this paper, based on the local optimization approach of the statistical theory of optimal control, a time-varying control law for an inertial interceptor's guidance to an intensively maneuvering aerial target has been designed. This law takes into account the mismatch between the dynamic properties of the interceptor and the target.

Due to the time-varying nature of the control law and consideration of the mismatch between the dynamic properties of the interceptor and target, the final miss distance of the interceptor has been reduced. At the same time, in the above simulation conditions, the time-varying control law has an effect starting from a range of about 40 km to the target; consideration of the mismatch between the dynamic properties of the interceptor and target, starting from a range of about 25 km to the target.

The time-varying property of the guidance law has been ensured by introducing time-varying penalty coefficients for control accuracy (depending on the application conditions) into the performance functional. The mismatch between the dynamic properties of the interceptor and HSA has been taken into account by using different mathematical models of the objects and by introducing the matrix of measured disturbances (characterizing the above mismatch) into the performance functional.

It has been established that the time-varying interceptor control law, which considers the mismatch between the dynamic properties of the interceptor and the target, potentially ensures an interception of the target.

## REFERENCES

1. Verba, V.S., Merkulov, V.I., and Zakomoldin, D.V., Problems of Interception of High Speed Aircraft Maneuvering According to Complex Laws. Part 1. Features of the Use of High Speed Aircraft, Complicating Their Interception, *Achievements of Modern Radioelectronics*, 2024, vol. 78, no. 3, pp. 5–12.
2. Verba, V.S., Merkulov, V.I., Zakomoldin, D.V., and Likhachev, V.P., Problems of Interception of High Speed Aircraft Maneuvering According to Complex Laws. Part 2. Analysis of Flight Paths of Foreign Aircraft, *Achievements of Modern Radioelectronics*, 2024, vol. 78, no. 4, pp. 5–14.
3. Yibo, D., Yue, X., Chen, G., et al., Review of Control and Guidance Technology on Hypersonic Vehicle, *Chinese Journal of Aeronautics*, 2022, vol. 35, no. 7, pp. 1–18.
4. Grant, M. and Antony, T., Rapid Indirect Trajectory Optimization of a Hypothetical Long Range Weapon System, *Proceedings of the AIAA Atmospheric Flight Mechanics Conference*, 2016, pp. 1–19.
5. Besser, H., Shaffer, A., Zimper, D., et al., Hypersonic Vehicles: Game Changers for Future Warfare, *Transforming Joint Air Power: J. JAPCC*, 2017, no. 24, pp. 11–27.
6. Shen, H., Liu, Y., Chen, B., et al., Control-Relevant Modeling and Performance Limitation Analysis for Flexible Air-Breathing Hyper-Sonic Vehicles, *Aerospace Science Technology*, 2018, no. 76, pp. 340–349.
7. Karako, T. and Dahlgren, M., Complex Air Defense Countering the Hypersonic Missile Threat, *A Report of the CSIS Missile Defense Project*, 2022.
8. Verba, V.S., Merkulov, V.I., and Ilchuk, A.R., Approaches to Optimizing Guidance Methods to High-Speed Intensively Maneuvering Targets. Part III. Guidance Optimization Considering the Dynamic Properties of Interceptors, *Autom. Remote Control*, 2025, vol. 86, no. 2, pp. 153–164.
9. Merkulov, V.I. and Verba, V.S., *Sintez i analiz aviatsionnykh radioelektronnykh sistem upravleniya. Kniga 1* (Design and Analysis of Aircraft Radioelectronic Control Systems. Book 1), Moscow: Radiotekhnika, 2023.

*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. 6, pp. 519–528.




# Optimal Insurance and Investment Strategies in Individual Risk Models with Background Risks

**A. Y. Golubin**[*,**,a] **and V. N. Gridin**[**,b]

[*]*National Research University Higher School of Economics, Moscow, Russia*
[**]*Design Information Technologies Center Russian Academy of Sciences, Moscow Region, Odintsovo, Russia*
*e-mail:* [a]*agolubin@hse.r*, [b]*info@ditc.ras.ru*

Received August 10, 2025
Revised September 25, 2025
Accepted November 5, 2025

**Abstract**—In the presented paper, a one-stage problem of simultaneous finding an optimal insurance and investment portfolio in the presence of background risks is studied. The goal functional is a functional of Markovitz type which depends on the first two moments of the final capital. It is shown that the optimal insurance necessarily belongs to a class of stop loss insurances, equations determining parameters of the stop loss insurance and investment portfolio are derived. A numerical example illustrating the theoretical results is given.

*Keywords*: optimal insurance, optimal investment, Markovitz functional, background risk

**DOI:** 10.7868/S1608303226060046

## 1. INTRODUCTION

One of the first results of study of models with background risks [1] showed that without the assumption of background risks a discrepancy arises between theoretical and empirical data, when solving various problems of financial mathematics. Later, a number of papers appeared on the choice of an investment portfolio that maximizes the expected utility of the final capital in the presence of background risks. As an example, we can cite the article [2] which showed that "unfair" external risk (with a negative mean) increases risk aversion in the investor's utility function. A characteristic feature of such problems is the lack of a dependence between background risks and asset returns.

However, in [3], it is proved that there is a significant correlation between asset returns and certain background risks. Paper [4] investigated the impact of background risk on portfolio choice and the efficient frontier within the mean-variance criteria. In [5], the author reviewed existing results within the mean-variance and expected utility approaches in the presence of additive external risks.

Insurance as a tool for reducing losses from background risks has been used in continuous-time dynamic models. Decisions on portfolio formation from risky assets and consumption management under background risk were studied in [6]. Using the dynamic programming equations, the authors found that optimal insurance hedges investment risks and balances consumption growth and volatility. In [7], the problem of choosing insurance that reduces the loss from an external piece-wise constant background risk process was investigated; a characterization of the function dual to the Bellman function as a solution to the Hamilton–Bellman–Jacobi equation was found. In a similar setting, the problem of a firm's choice of insurance strategy against a jump-like process of productivity failures and against natural disasters was studied in [8] using the dynamic programming method, but without allowing for investment.

In this paper, we consider the one-stage problem of selecting optimal insurance coverage for a group of policyholders (clients)—the so-called individual risk model—and the simultaneous choice of an investment portfolio in a securities market with a risk-free asset. The maximized functional is a Markowitz-type functional, i.e., it depends only on the first two moments of the final insurer's capital, which consists of the capital received from the insurance transaction and the capital earned on investments. It is assumed that there is a background risk which affects the amount of clients' losses, and a background risk which affects the overall return on investment. The results obtained differ from earlier findings in several respects. First, instead of insuring only background risks, an optimal risk-sharing function for the group of clients is derived. Second, the problem of choosing both insurance and an investment portfolio is investigated. Third, the normal approximation of aggregate risk, which is widely used in financial mathematics, is not employed here, and the results are therefore "exact" in this sense.

Section 2 describes insurance and investment markets, first without background risks and, then, with background risks. An optimization problem is formulated with the variance as a risk measure. Section 3 is devoted to finding its solution, where an optimal insurance and optimal portfolio being found as solutions to separate problems due to the specifics of the objective functional. Section 4 solves a similar optimization problem, but using the standard deviation as a risk measure. Necessary optimality conditions are obtained. Sections 5 and 6 contain a numerical example and concluding remarks, respectively.

## 2. PROBLEM FORMULATION

### *2.1. Description of Investment and Insurance Markets*

First, consider an investment market of $n+1$ assets (see, for example, [10]), where the return vector $\bar{R} = (R_0, \ldots, R_n)$ and there is no background risk. Let $R_0 = m_0$ almost sure (a.s.), i.e., it is a risk-free asset. Note that the presence of such an asset allows one to avoid spending the entire initial capital $w > 0$ on risk investments and, for example, when $m_0 = 1$, one can avoid investing the amount $wa_0$ and put it "in one's pocket" (without taking inflation into account). Let us denote by $\bar{a} = (a_0, \ldots, a_n) \in R^{n+1}$ the investment portfolio, where $a_i$ is the share in percentage of the initial capital $x_0 > 0$ invested into the $i$th asset. A typical budget constraint is $\sum_{i=0}^{n} a_i = 1$. This means self-financing of the investor-insurer and his ability to "short sell", i.e. borrowing some assets at current prices in order to invest money in others with subsequent reimbursement at the prices of the next quotation. The investor's capital $X_{\bar{a}} = w\sum_{i=0}^{n} a_i R_i = w(\sum_{i=1}^{n} a_i(R_i - m_0) + m_0)$, where $w > 0$ is the initial capital. After substituting $a_0 = 1 - \sum_{i=1}^{n} a_i$, we have:
the expected value and variance

$$\mu(a) = w\left(\sum_{i=1}^{n} a_i \triangle m_i + m_0\right), \text{ where } \triangle m_i = m_i - m_0,\ m_i = E\,R_i,\ a \in R^n,$$

$$\sigma^2(a) = w^2 a C a^T,$$

where $a^T$ denotes the transposed row vector $a$. Let's introduce natural assumptions:

$$0 < m_0 < \min_{i=1,\ldots,n} m_i,$$

the risk-return covariance matrix $C$ of dimension $n \times n$ is positive definite.

Let us now describe the insurance market. A group of $l$ clients is homogeneous, i.e., their losses $X_1, \ldots, X_l$ are independent and identically distributed random variables (i.v.) with a distribution function $F_1(x)$. The insurer chooses a Borel-measurable risk sharing function $I(x) : 0 \le I(x) \le x$, $x \in [0, \infty)$, then the share of the $i$th client's loss paid by the insurer is equal to $I(X_i)$. The insurer's initial capital is $w > 0$; the premium payed by the insured (see, for example, [11]) is $d = (1+\alpha)EI(X_1)$, where $\alpha > 0$ is a given loading coefficient.

### *2.2. Model with Background Risks*

Consider the insurance and investment model, taking into account the background risks. The insurer invests his initial capital $w$[1] so that his final capital consists of insurance income $S_1 = l(1+\alpha)E\,I(X_1+Z)-\sum_1^l I(X_j+Z)$ and income from investments $S_2 = w[\sum_1^n a_i(R_i-m_0)+m_0]+wV$. Here: $Z$ is the background insurance risk such that $X_j + Z \geq 0$ a.s. and $\sigma^2(I(X_1 + Z)) < \infty$, $V$ is the total investment background risk, $\sigma^2(V) < \infty$.

*Remark 1.* Here, the losses of clients are $X_j^Z \stackrel{def}{=} X_j+Z$, $j = 1,\ldots,l$ are assumed to be identically distributed, but are not independent, since $Z$ enters the expression for each loss. Moreover, $X_j$ and $Z$ are assumed to be dependent. Such a model is atypical for problems of "classical" actuarial mathematics [11], in which independence of losses is assumed. Background risks $W$ and $V$, affecting, respectively, the losses of policyholders and the total return on investing risky assets, are assumed to be independent. This condition seems justified, since here the insurance and investment markets are of different nature and are not related.

The Markowitz functional under the study has the form

$$J[I,a] \equiv E\,(S_1+S_2) - \beta\sigma^2(S_1+S_2) \to \max_{I,a},\ 0 \leq I(x) \leq x,\ a \in R^n. \tag{1}$$

Here, $\beta > 0$ has the meaning of the investor-insurer's caution coefficient: the larger $\beta$, the more cautious the insurer. Substituting the expressions for $S_1$ and $S_2$, we obtain

$$\begin{gathered} J[I,a] = l\alpha E\,I(X_1^Z) + w(a\triangle m^T + m_0 + E\,V) \\ -\beta\left(l\sigma^2[I(X_1^Z)] + l(l-1)cov(I(X_1^Z), I(X_2^Z)) + w^2aCa^T + 2aQ^T + \sigma^2(V)\right), \end{gathered} \tag{2}$$

where $Q^T = cov(R,V)$ is a column vector with components $cov(R_i,V)$, $i = 1,\ldots,n$.

## 3. MAIN RESULTS

Let us return to problem (1) and consider an auxiliary problem

$$J[I,a] \to \max_I,\ 0 \leq I(x) \leq x, \tag{3}$$

for a fixed $a \in R^n$, the expression for $J[I,a]$ is given in (2). The following statement gives the necessary conditions for optimality in (3), establishing: the form of the optimal $I^*(x) = x \wedge k^*$ (stop loss insurance with the so-called insurer retention level $k^*$), where $x \wedge y \stackrel{def}{=} \min\{x,y\}$, and the equation for finding the parameter $k^*$.

**Statement 1.** *Problem* (3) *has solution* , $I^*(x) = x \wedge k^*$, *where* $0 < k^* < \infty$ *is the only root of the optimality equation*

$$\xi(k) = 0, \quad \textit{where} \quad \xi(k) \stackrel{def}{=} \alpha - 2l\beta\left(k - \int\limits_0^k (1 - F_1^Z(x))dx\right). \tag{4}$$

**Proof.**

**Lemma 1.** *There is a solution to problem* (3).

The proof of the lemma is given in Appendix.

[1] Here it is possible to pay the total premium $ld$ in parts up to the end of the insurance period, so the capital available for investment at the beginning of the period is equal to $w$.

Denote $I_\rho(x) = \rho I^*(x) + (1-\rho)I(x)$, where $\rho \in [0,1]$, the function $I^*$ is a solution of (3) and $I$ is an arbitrary risk sharing function. Find the derivative $\frac{d}{d\rho} J[I_\rho, a]\Big|_{\rho=1}$, taking into account that

$$\frac{d}{d\rho} E\, I_\rho(X_1^Z)\Big|_{\rho=1} = \int\limits_0^\infty (I^*(x) - I(x)) dF_1^Z(x),$$

$$\frac{d}{d\rho} \sigma^2(I_\rho(X_1^Z))\Big|_{\rho=1} = 2\int\limits_0^\infty (I^*(x) - E\, I^*(X_1^Z))(I^*(x) - I(x)) dF_1^Z(x)$$

and

$$\frac{d}{d\rho} cov\,(I_\rho(X_1^Z), I_\rho(X_2^Z))\Big|_{\rho=1} = 2\int\limits_0^\infty\int\limits_0^\infty I^*(y)(I^*(x) - I(x)) dG(x,y) - 2\int\limits_0^\infty E\, I^*(X_1^Z)(I^*(x) - I(x)) dF_1^Z(x),$$

where the distribution functions are: $F_1^Z(x) = P(X_1^Z \le x)$ and $G(x,y) = P(X_1^Z \le x, X_2^Z \le y)$, the latter is obviously symmetric in its arguments. Since the set of risk sharing functions is convex, the necessary optimality condition for $I^*$ in (3) is the inequality

$$\frac{d}{d\rho} J[I_\rho, a]\Big|_{\rho=1} \ge 0$$

for any risk sharing function $I$. Taking into account expression (2), this condition can be reformulated as: $I^*$ maximizes the integral

$$\max_I \int\limits_0^\infty\int\limits_0^\infty I(x)\phi(x,y) dG(x,y), \tag{5}$$

where $\phi(x,y) = \alpha - 2\beta[I^*(x) - E\, I^*(X_1^Z) + (l-1)(I^*(y) - E\, I^*(X_1^Z)]$. According to the Neyman–Pearson lemma (see, for example, [12]) the function $I^*$ maximizes in (5) if and only if

$$I^*(x) = \begin{cases} x, & \phi(x,y) > 0, \\ 0, & \phi(x,y) < 0 \end{cases} \tag{6}$$

up to a set of zero measure $G$. Let us consider (6) not for all $(x,y) \in R_+^2$, but only on the subset $\{(x,y) : x = y,\ x \ge 0, y \ge 0\}$, and the integral in (5) is taken over the distribution function $G(x,x)$, so the integral is $\int_0^\infty I(x)\phi(x,x) dG_1(x)$, where the distribution function $G_1(x) = P(X_1^Z \le x, X_2^Z \le x)$. Then the necessary condition for fulfilling (6) is

$$I^*(x) = \begin{cases} x, & \psi(x) > 0, \\ 0, & \psi(x) < 0 \end{cases} \tag{7}$$

up to a set of zero measure $G_1$.

$$\text{Here the function } \psi(x) \overset{def}{=} \phi(x,x) = \alpha - 2l\beta[I^*(x) - E\, I^*(X_1^Z)]. \tag{8}$$

The value $\psi(0) = \alpha + 2l\beta E\, I^*(X_1^Z) > 0$. As $x$ increases from 0, the function $I^*(x) = x$ according to (7), therefore $\psi(x)$ decreases (see (8)) up to the point $x = k^*$ where the abscissa axis is touched. As $x$ increases from $k^*$, the function $\psi(x)$ can neither decrease nor increase. Indeed, when $\psi(x)$ decreases from zero, we obtain a contradiction with (7), since $I^*(x) = 0$ for such $x$. Suppose that $\psi(x)$ increases from zero, then (see (7)) $I^*(x) = x$ and $\psi(x)$ must decrease. Thus, $I^*(x) = x \wedge k^*$, $k^* \in (0,\infty)$. Obtain an equation for finding the level $k^*$ by substituting the function $I_k(x) = x \wedge k$ into expression (8) and $E\, X_1^Z \wedge k = \int_0^k (1 - F_1^Z(x)) dx$ instead of $E\, I^*(X_1^Z)$. As a result, we get the optimality equation (4) in Statement 1. Note that the found optimal stop loss insurance is distinct from the deductible obtained in [13] for a continuous-time background risk insurance model with a random horizon.

*Remark 2.* Let $0 < \mathrm{IS} \leq \infty$ denote the maximum possible value of $X_1^Z$ or, more precisely, $\mathrm{IS} = \sup \mathit{supp}\, F_1^Z$—the upper bound of the distribution support $F_1^Z$. The case when the retention level $k^* \geq IS$ means the full compensation for the client's loss.

To conclude the analysis of (1), let us study the problem

$$\max_a J[I,a],\ a \in R^n \tag{9}$$

for a fixed function $I(x)$. The objective function has the form

$$E\,S_1 + w(a\triangle m^T + m_0 + E\,V) - \beta[\sigma^2(S_1) + w^2(aCa^T + 2aQ^T + \sigma^2(V))].$$

Calculating its gradient and equating it to zero, we get $J_a'[I,a] = w\triangle m - 2w^2\beta(aC+Q) = 0$. From here

$$a^* = \left(\frac{\triangle m}{2w\beta} - Q\right)C^{-1}. \tag{10}$$

Due to the positive definiteness of the matrix $C$, the function $J[I,a]$ is strictly concave in $a$. Therefore, the found vector $a^*$ is the unique solution to (9). Thus, Statement 2 below is proved.

**Statement 2.** *The only solution to problem* (9) *is portfolio $a^*$ defined in* (10).

Note that the income from insurance $S_1$ depends only on $I$, and the income from investment $S_2$ depends only on $a$. Thus, we obtain from Statements 1 and 2: the solution to the problem of maximizing the Markowitz functional with background risks is: stop loss insurance $I^*(x) = x \wedge k^*$, where $k^*$ is defined in (4), and a portfolio $\bar{a}^* = (a_0^*, a^*) \in R^{n+1}$, in which $a_0^* = 1 - \sum_{i=1}^n a_i^*$.

## 4. OPTIMAL INSURANCE AND INVESTMENT IN A MODEL WITH STANDARD DEVIATION OF THE CAPITAL

The standard deviation $\sigma(S_1 + S_2)$, rather than the variance $\sigma^2(S_1 + S_2)$, is used as the risk measure in the Markowitz functional (see, for example, [10, 11] for its use in financial models). The expressions for $S_1 = S_2^I$ and $S_2 = S_2^a$ are given in Section 2.2. The studied problem is

$$J[I,a] \equiv E\,(S_1 + S_2) - \beta\sigma(S_1 + S_2) \to \max_{I,a},\ 0 \leq I(x) \leq x,\ a \in R^n. \tag{11}$$

The advantage of this approach is that both terms in (11) have the same dimension, for example, roubles. The detailed expression for the objective functional is (see (2))

$$\begin{gathered} J[I,a] = l\alpha E\,I(X_1^Z) + w(a\triangle m^T + m_0 + E\,V) \\ -\beta\sqrt{l\sigma^2[I(X_1^Z)] + 2l(l-1)cov(I(X_1^Z), I(X_2^Z)) + w^2(aCa^T + 2aQ^T + \sigma^2(V))}. \end{gathered} \tag{12}$$

Problems with a similar formulation, but without background risk, were studied in [14].
Let's study an auxiliary problem

$$J[I,a] \to \max_I,\ 0 \leq I(x) \leq x. \tag{13}$$

**Statement 3.** *Problem* (13) *has a solution $I_a^*(x) = x \wedge k_a^*$, where $0 < k_a^*$ is the minimal root of the equation*

$$\xi_a(k) = 0,\ \textit{where}\ \xi_a(k) = \alpha - l\beta[kE\,I^k(X_1^Z)]/(\sigma^2(S_1^{I^k} + \sigma^2(S_2^a))^{1/2}, \tag{14}$$

*with the sharing rule $I^k(x) = x \wedge k$.*

**Proof.** The existence proof for (13) repeats the arguments in the proof of Lemma 1.
As before, we denote by $I_\rho(x) = \rho I^*(x) + (1-\rho)I(x)$, where $\rho \in [0,1]$ and $I^*(x) = I_a^*(x)$ maximizes

$J[I,a]$. The necessary condition for the optimality of $I^*$ in (13)

$$\frac{d}{d\rho}J[I_\rho,a]\Big|_{\rho=1}\geq 0 \text{ for any risk sharing function } I.$$

This inequality can be rewritten taking into account the found expressions for the derivatives of $E\,I_\rho(X_1^Z)$, $\sigma^2(I_\rho(X_1^Z))$ and $cov\,(I_\rho(X_1^Z),I_\rho(X_2^Z))$ at the point $\rho=1$ (see the proof of Statement 1): $I^*(x)=I_a^*(x)$ is the solution to the problem of maximizing the integral

$$\max_I\int_0^\infty\int_0^\infty I(x)\phi_a(x,y)dG(x,y),$$

where $\phi_a(x,y)=\alpha-\beta[I^*(x)-E\,I^*(X_1^Z)+(l-1)(I^*(y)-E\,I^*(X_1^Z)]/(\sigma^2(S_1^{I^*}+\sigma^2(S_2^a))^{1/2}$. According to the Neyman–Pearson lemma, the function $I^*(x)$ gives maximum if and only if

$$I^*(x)=\begin{cases} x, & \phi_a(x,y)>0,\\ 0, & \phi_a(x,y)<0\end{cases} \tag{15}$$

up to a set of zero measure $G$. As in the proof of Statement 1, we consider (15) not for all $(x,y)\in R_+^2$, but only on the subset $\{(x,y):\,x=y,\,x\geq 0,y\geq 0\}$ Then, the necessary condition for fulfilling (15) is

$$I^*(x)=\begin{cases} x, & \psi_a(x)>0,\\ 0, & \psi_a(x)<0\end{cases} \tag{16}$$

up to a set of zero measure $G_1$. Here the function

$$\psi_a(x)\overset{def}{=}\phi_a(x,x)=\alpha-l\beta[I^*(x)-E\,I^*(X_1^Z)]/(\sigma^2(S_1^{I^*}+\sigma^2(S_2^a))^{1/2}. \tag{17}$$

Note that the denominator in (17) is always positive, since $\sigma^2(S_2^a)>0$. The value $\psi_a(0)=\alpha+2l\beta E\,I^*(X_1^Z)>0$. As $x$ increases from 0, the function $I^*(x)=x$ according to (16), therefore $\psi_a(x)$ decreases (see (17)) up to the point $x=k_a^*$ where the abscissa axis is touched. As $x$ increases from $k_a^*$, the function $\psi_a(x)$ neither decreases nor increases. Indeed, when $\psi_a(x)$ decreases from zero, we obtain a contradiction with (16), since $I^*(x)=0$ for such $x$. Suppose that $\psi_a(x)$ increases from zero, then (see (16)) $I^*(x)=x$ and $\psi_a(x)$ must decrease, not increase. Thus, $I^*(x)=x\wedge k_a^*$, where $k_a^*>0$.

For finding the retention level $k_a^*$, substitute the function $I^k(x)=x\wedge k$ instead of $I^*(x)$ into the expressions for $E\,I^*(X_1^Z)$ and $\sigma^2(S_1^{I^*}$ in (17). As a result, we obtain the optimality equation (14) in Statement 3. □

Note that the case $k_a^*=\infty$, i.e. (see (14)), the function $\xi(k)>0$ on $[0,\infty)$ cannot be excluded. This means that the insurer covers the entire client's loss.

Let us now study the optimal investment problem for a fixed risk sharing function $I$,

$$\max_a J[I,a],\; a\in R^n. \tag{18}$$

**Statement 4.** *Let $a_I^*$ be a solution to problem* (18). *Then this vector is the root of the equation*

$$\triangle m\left(\sigma^2(S_1^I)+w^2[aCa^T+2aQ^T+\sigma^2(V)]\right)^{1/2}-w\beta(aC+Q)=0. \tag{19}$$

**Proof.** The objective function is

$$E\,S_1^I+w(a\triangle m^T+m_0+E\,V)-\beta\sqrt{\sigma^2(S_1^I)+w^2(aCa^T+2aQ^T+\sigma^2(V))}.$$

Find its gradient in $a$ and equate it to zero,

$$J_a'[I,a]=w\triangle m-w^2\beta(aC+Q)/\left(\sigma^2(S_1^I)+w^2[aCa^T+2aQ^T+\sigma^2(V)]\right)^{1/2}=0.$$

This equation is obviously equivalent to (19). If $a_I^*$ is a solution to (18), then, taking into account the absence of restrictions on $a$, we obtain that $a_I^*$ satisfies (19). □

Taking into account the results obtained in Statements 3 and 4, we have the following optimality conditions in the original problem $J[I, a] \to \max$, where the maximum is taken over $I$, $a$.

**Statement 5.** *Let* $(I^{**}, a^{**})$ *be a solution to problem* (1). *Then* $I^{**}(x) = x \wedge k^{**}$, *where* $(k^{**}, a^{**})$ *is a solution to the system of n+1 nonlinear equations:* (19), *where instead of* $I(x)$ *the function* $I^k(x) = x \wedge k$ *is substituted, and* (14).

## 5. EXAMPLE

Let us give a numerical example of solving problem (1) with variance as a risk measure, which does not pretend to be of practical significance. Let the insureds number be $l = 10$, the number of risky assets $n = 2$, the vector of expected returns $E\,\bar{R} = (m_0, m_1, m_2) = (1, 1.1, 1.15)$, the vector of covariances of returns on risky assets and background risk $Q = (-0.3, -0.25)$ and the covariance matrix

$$C = \begin{pmatrix} 0.3 & 0.2 \\ 0.2 & 0.4 \end{pmatrix}.$$

The initial capital $w = 1$, the distribution function of client's loss with the background risk is equal to $F_1^Z(x) = 1 - p + p(1 - \exp(-2x))$, $x \geq 0$, where the probability of an insurance event

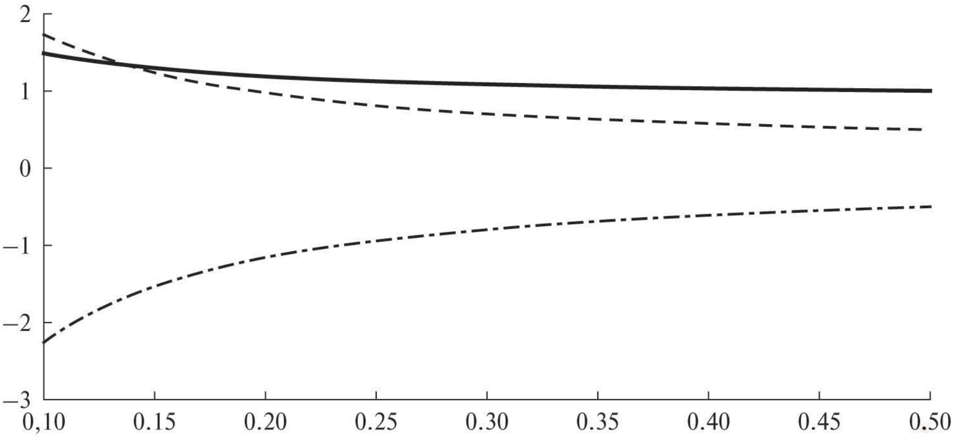


**Fig. 1.** Graphs of the dependence of the optimal portfolio components on $\beta$: $a_1^*$—solid line, $a_2^*$—dotted line, investment in the risk-free asset $a_0^* = 1 - a_1^* - a_2^*$—dash-dotted line.

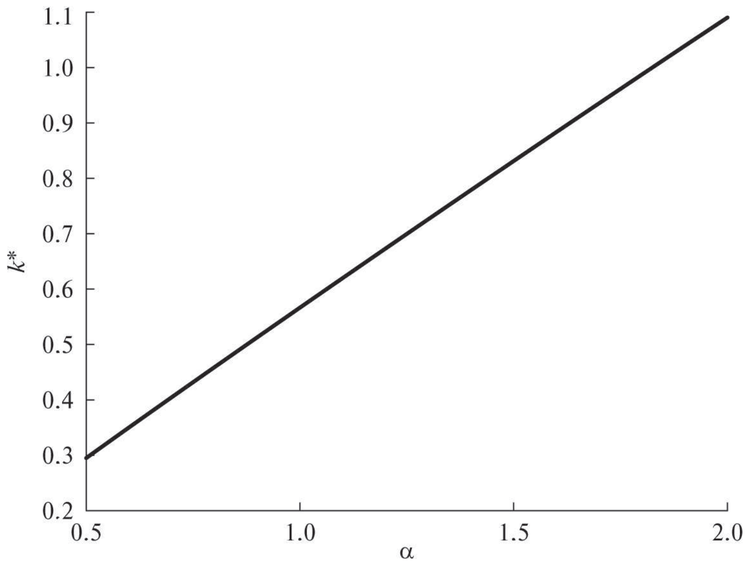


**Fig. 2.** Graph of the dependence of the retention level $k^*$ on $\alpha$.

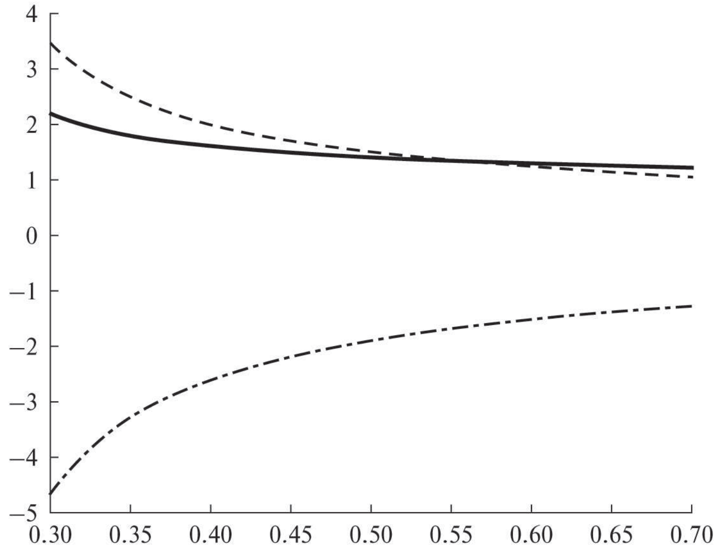


**Fig. 3.** Optimal portfolio components vs $\beta$: $a_1^*$—solid line, $a_2^*$—dotted line, $a_0^* = 1 - a_1^* - a_2^*$—dash-dotted line.

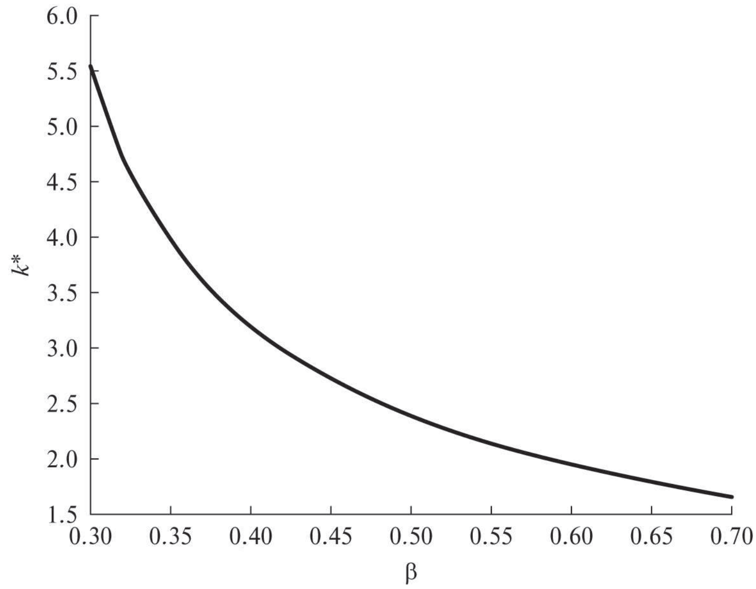


**Fig. 4.** Optimal level $k^*$ vs $\beta$.

$p = P(X_1^Z > 0) = 0.2$. The caution coefficient $\beta$ and the load coefficient $\alpha$ are variable parameters. We find the parameter $k^*$ in the risk sharing function $I^*(x) = x \wedge k^*$ and the portfolio of investments into risky assets $a^* = (a_1^*, a_2^*)$ by applying Statements 1 and 2. The results of the numerical calculations are shown in Figs. 1 and 2.

From Fig. 1 it is evident that with the growth of the caution coefficient $\beta$ investments in risky assets $a_1^*(\beta)$ and $a_2^*(\beta)$ decrease, and the function $a_0^*(\beta)$ increases, remaining negative. The latter means that the investor borrows money from a risk-free asset, but this amount decreases with the increase in $\beta$.

As quite expected, Fig. 2 shows that the level $k^*$ in stop loss insurance increases with an increase in the loading coefficient $\alpha$—as the insurance price grows, the insurer increases the share of the client's risk coverage. The function $k^*(\alpha)$ is close to linear one, which is explained by a smallness of the nonlinear part in equation (4) which determines $k^*$.

It is depicted on Fig. 3 that the investor borrows money from the risk-less asset and this sum decreases as well as investment into risky assets when $\beta$ grows. Compared with Fig. 1, here the borrowed and invested sums are greater, which is explained by a relatively slow growth of standard deviation vs variance of the final capital.

According to Fig. 4, if the caution coefficient $\beta$ increases then, naturally, the retention level $k^*$ decreases, i.e. the insurer lessens his/her share of risk $X_1^Z \wedge k^*$.

## 6. CONCLUSION

This article studies a one-stage problem for an insurer-investor to select optimal insurance, i.e. risk sharing functions, for a group of clients and an investment portfolio in the presence of additive background risks. Two Markowitz-type functionals are considered, using variance and standard deviation as risk measures, respectively. In the first case, the explicit expressions for the insurance risk sharing function and the portfolio were found; in the second case, the necessary optimality conditions were found in the form of the system of nonlinear equations. Numerical examples are provided to illustrate the theoretical results. The following research directions on this topic seem promising: using not only the insurer's initial capital but also the insurance premiums received from clients as the initial investment capital; employing other objective functionals, such as the expected utility of final capital; and studying a multi-stage optimization problem with additive and/or multiplicative background risks.

## FUNDING

This work was supported by State Program FFNR-2024-0003.

*APPENDIX*

### A.1. PROOF OF LEMMA 1

Let $\{I_m\}$ be the maximizing sequence of risk sharing in problem (3), i.e.

$$J[I_m, a] \to J^* = \sup_I J[I, a].$$

Applying Helly's theorem to the sequence of random vectors $\{(I_m(X_1^Z, I_m(X_2^Z)\}$ (see the expression in (2) for $J[I, a]$), we obtain that there exists a subsequence $\{(I_s(X_1^Z), I_s(X_2^Z)\}$ that weakly converges to some limit $(\xi, \eta)$. To prove that $\xi$ and $\eta$ are proper random variables, i.e. $P(\xi < \infty) = 1$ and $P(\eta < \infty) = 1$, it suffices to note that, by virtue of the definition $I_m(X_i^Z) \leq X_i^Z$ a.s., $i = 1, 2$. Since $X_1^Z$ and $X_2^Z$ are identically distributed and measurable with respect to the Borel sigma-algebra, the limit can be represented as $(\xi, \eta) = (I^*(X_1^Z), I^*(X_2^Z))$ for some measurable function $I^*(x)$, where $0 \leq I^*(x) \leq x$.

The equality $J^* = J[I^*, a]$ now follows from the weak convergence of $\{(I_s(X_1^Z, I_s(X_2^Z)\}$ to $\{(I^*(X_1^Z, I^*(X_2^Z)\}$, the continuity of the objective functional in (2) with respect to $E\, I(X_1^Z)$, $E\, I^2(X_1^Z)$, and $cov(I(X_1^Z),\, I(X_2^Z))$.

## REFERENCES

1. Bodie, Z., Merton, R.C., and Samuelson, W.F., Labour supply flexibility and portfolio choice in a life cycle model, *Journal of Economic Dynamics and Control*, 1992, no. 16, pp. 427–449.
2. Gollier, C. and Pratt J.W., Risk vulnerability and the tempering effect of background risk, *Econometrica: Journal of the Econometric Society*, 1996, vol. 64, no. 5, pp. 1109–1123.

3. Heaton, J. and Lucas, D., Portfolio choice in the presence of background risk, *The Economic Journal*, 2000, vol. 110, no. 460, pp. 1–26.

4. Jiang, C.H., Ma, Y.K., and An, Y.B., An analysis of portfolio selection with background risk, *Journal of Banking and Finance*, 2010, vol. 34, pp. 3055–3060.

5. Guo, X., Wagener, A., Wong, W.K., and Zhu, L.X., The two-moment decision model with additive risks, *Risk Management*, 2018, vol. 20, no. 1, pp. 77–94.

6. Lin, W. and Lu, J., Risky asset allocation and consumption rule in the presence of background risk and insurance markets, *Insurance: Mathematics and Economics*, 2012, vol. 50, no. 1, pp. 150–158.

7. Mnif, M., Optimal risk control under marked point processes shocks: A dynamic programming duality approach, *International Journal of Theoretical and Applied Finance*, 2013, vol. 16, no. 7, pp. 350–360.

8. Hong, H., Wang, N., and Yang, J., Mitigating disaster risks in the age of climate changem, *Technical report, National Bureau of Economic Research*, 2020.

9. Markowitz H., Portfolio Selection, *Journal of Finance*, 1952, vol. 7, pp. 77–91.

10. Alexander, G.J. and Baptista, A.M., Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis, *Journal of Economic Dynamics and Control*, 2002, vol. 26, pp. 1159–1193.

11. Bowers, N.L., Gerber, H.U., Hickman, J.C., et al, *Actuarial Mathematics*, Itaca, Illinois: The Society of Actuaries, 1986.

12. Lehman, E.L., *Testing statistical hypotheses*, New York: Wiley and Sons, 1959.

13. Moore, K.S. and Young, V.R., Optimal insurance in a continuous-time model, *Insurance: Mathematics and Economics*, 2006, vol. 39, no. 1, pp. 47–68.

14. Golubin, A.Y. and Gridin, V.N., Optimal insurance strategy in the individual risk model under a stochastic constraint on the value of the final capital, *Automation and Remote control*, 2019, vol. 80, no. 4, pp. 708–717.

*This paper was recommended for publication by F.T.Aleskerov, a member of the Editorial Board*

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 529–541.




# Forecasting Epidemic Outbreaks of Acute Respiratory Viral Infections with Early Warning Signals and Machine Learning

**M. Yu. Lazutov**[*,a] **and V. N. Leonenko**[*,**,b]

[*]*ITMO University, St. Petersburg, Russia*
[**]*Smorodintsev Research Institute of Influenza, St. Petersburg, Russia*
*e-mail:* [a]*marklazutov@gmail.com,* [b]*vnleonenko@yandex.ru*

Received April 20, 2025
Revised October 120, 2025
Accepted November 6, 2025

**Abstract**—Forecasting critical transitions in epidemic incidence is important for epidemic surveillance. Earlier detection can support a more timely public-health response. This paper studies anomaly detection methods based on artificial intelligence and early warning signals (EWSs) in incidence time series. The goal is to predict transitions from seasonal infections to epidemic outbreaks. Two approaches are examined. The first is a classification approach that estimates whether a critical transition is near. The second is a regression approach that forecasts future infection dynamics. Several machine learning models are applied to two types of data. The models include ensemble methods (Easy Ensemble, RUSBoost, and Balanced Bagging) and deep learning architectures (Early Warning Signal Network (EWSNet), LSTM, and GRU). The first dataset contains influenza incidence with epidemic periods labeled by expert criteria. The second dataset contains unlabeled COVID-19 incidence. The results show that Easy Ensemble and EWSNet provide the best balance between precision and recall. Recurrent neural networks model the dynamics of mean values effectively, whereas variance forecasting remains more difficult. The results show that classical EWS methods and machine learning can be combined to improve epidemic forecasting and support public-health decision making.

*Keywords*: epidemic forecasting, early warning signals, machine learning, time-series analysis, critical transitions

**DOI:** 10.7868/S1608303226060056

## 1. INTRODUCTION

Early detection and forecasting of critical phenomena are important in many scientific fields [1], including epidemiology [2]. This task can also be viewed as a signal-recognition problem [3]. A relevant example is the emergence of a new respiratory viral strain. At early stages, such strains may be difficult to distinguish from seasonal acute respiratory viral infections in aggregated surveillance data. For this reason, specialized statistical indicators are used together with more general methods. These indicators are known as early warning signals (EWSs). They help estimate whether a system is approaching a transition point, or critical point. Many EWS variants have been proposed. They include spectral EWSs based on the spectrum of a time series and EWSs based on critical slowing down [4]. The latter group includes indicators based on increases in variance, autocorrelation, skewness, and related statistics. For epidemic processes, many studies have tested EWSs on both synthetic and real data [5].

This study compares methods for early detection of incidence outbreaks caused by pandemic respiratory viruses. The analysis uses incidence data for acute respiratory viral infections, including seasonal influenza and COVID-19. We compare specialized EWSs based on critical slowing down with more general statistical approaches. We also evaluate whether these methods can support epidemic-process management and control. Two forecasting settings are considered. In the first setting, an epidemic threshold is available. Then the task is binary classification: the critical transition point is either "near" or not near. The proximity interval is selected according to epidemic-surveillance requirements. In this study, it is one month before the critical point. In the second setting, epidemic-threshold labels are unavailable. This is the case for COVID-19, because the disease has not circulated long enough to estimate a multi-year baseline. Then outbreak detection is formulated as forecasting selected EWSs. The absolute indicator values are predicted, and the time to the critical point is estimated from the forecast incidence. This setting is a special case of predictive statistical methods based on multi-year incidence data (see, for example, [6]). Such methods, together with SIR models, are among the most widely used forecasting models in epidemiology.

## 2. BACKGROUND AND RELATED APPROACHES

Machine learning is now widely used in epidemic forecasting. This development follows broader progress in machine learning and artificial intelligence. Traditional models, such as SIR and SEIR models, are useful but have limitations. They may be too rigid for nonlinear factors and for the early stages of pandemics caused by novel viruses [2]. Classical machine learning algorithms can also be effective. For example, Random Forest and support vector machines have been used for early epidemic outbreak detection [7]. Recurrent neural networks have also shown good results in disease forecasting; see, for example, [8]. At the same time, many studies continue to use the theory of complex systems and EWSs. In these studies, EWSs may be statistical indicators or explanatory models, such as a risk index based on social-distancing measures [28]. Table 1 lists studies that use different EWS-based approaches to epidemic forecasting. Some studies use a hybrid approach that combines machine learning with EWSs. However, comparative studies of such combinations are still limited. To address this gap, this paper considers several models and two problem formulations: classification and regression.

**Table 1.** Related studies

| | 2025 Zhao et al. [9] | 2024 Bizzotto et al. [10] | 2025 Gao et al. [11] | 2024 Drake et al. [12] | 2021 O'Brien [13] |
|---|---|---|---|---|---|
| Context | 31 diseases, 134 regions | COVID-19, Italy | COVID-19, Singapore, Canada, Hong Kong | SARS-CoV-2, UK | COVID-19, 24 countries |
| Method | Critical-slowing-down resilience indicators + time-to-event analysis | Nowcasting reproduction number method ($R_e$) | Feature-based time-series classification (ML) | Phylogenetic logistic growth rate (LGR) analysis | Sequential EWS assessment using GAMs |
| Best-performing EWSs | Variance, ACF | Nowcasted $R_e$ | Variance, ACF, skewness | Maximum LGR of dominant clusters | Concatenation of variance, ACF, and skewness |
| Forecast horizon | 18–21 days | 6–23 days | 14 days | 6–7 days | 10–14 days |

## 3. DATA

Two input datasets were used in this study. The first dataset was obtained from the database of the Research Institute of Influenza. It contains weekly incidence of acute respiratory viral infections in St. Petersburg from January 1990 through October 2020. The dataset includes the absolute number of cases. It also includes the estimated city population for each year of the study period. This information makes it possible to convert incidence to relative measures. The dataset also indicates whether the seasonal epidemic threshold was exceeded. The threshold was calculated according to the methodology of the Research Institute of Influenza. This indicator separates periods with an officially declared epidemic situation from non-epidemic periods. The general form of the time series is shown in Fig. 1. Since the data contain epidemic and non-epidemic labels, they are suitable for an EWS-based classification problem.

The second time series was obtained from open data on COVID-19 incidence in St. Petersburg from March 2020 to March 2024 [14]. The general form of the time series is shown in Fig. 2. These incidence data were collected daily. For COVID-19, it is more difficult to define an "epidemic outbreak" than for seasonal and pandemic influenza. The reason is that coronavirus infection does not have clearly expressed seasonality. In this study, outbreaks are defined as sustained increases in incidence followed by declines. These increases form epidemic waves in the data and are associated with the emergence of new coronavirus strains.

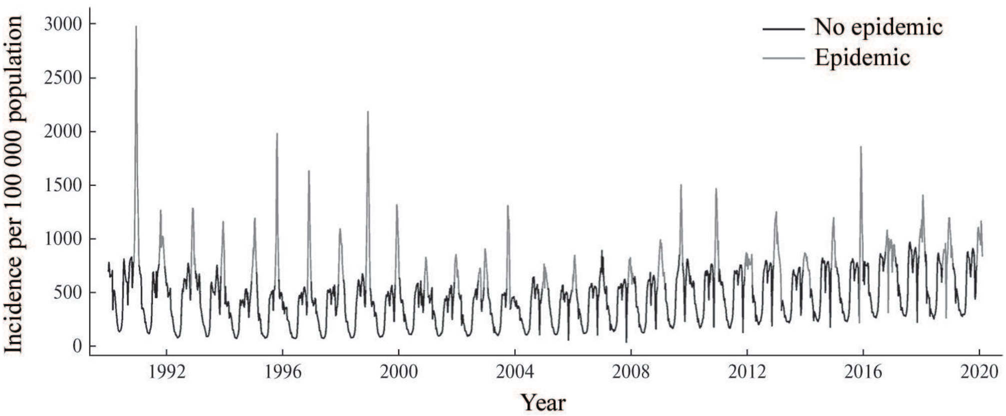


**Fig. 1.** Time series with labeled epidemics.

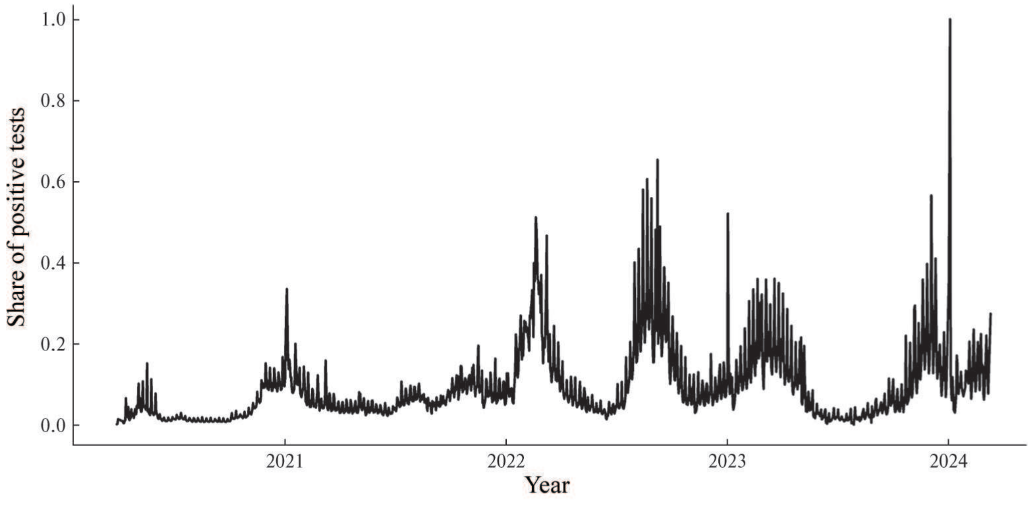


**Fig. 2.** Time series without labeled epidemics.

## 4. THEORETICAL BACKGROUND

We selected EWSs that have shown good performance for epidemic time series. According to Table 1 and [15], most studies use a forecasting horizon of about one to two weeks. Variance and the mean value are among the most useful indicators of proximity to a critical point. These indicators often increase when epidemic onset approaches. Figure 3 shows EWS calculations for acute respiratory viral infection incidence. Two cases are shown: one near influenza epidemic onset and one far from it. In this example, the most informative indicator is variance with a window length of 30. In the first case, this indicator increases monotonically. This is a sign of critical slowing down and may indicate an approaching phase transition [4]. In the second case, the indicator decreases monotonically. This means that the effect is absent, which is the expected behavior of an EWS far from a transition. Variance with window lengths of 15 and 20 does not show the same pattern. Thus, variance is sensitive to the window length. The window parameter should therefore be selected empirically before the indicator is used in epidemic-surveillance practice. In the classification task, we compare variance and skewness as EWSs. In the regression task, we forecast the variance and the mean incidence value and use them to predict epidemic increases. We selected longer decision horizons than those used in the reviewed studies: 4 and 5 weeks for classification and 14 and 28 days for regression.

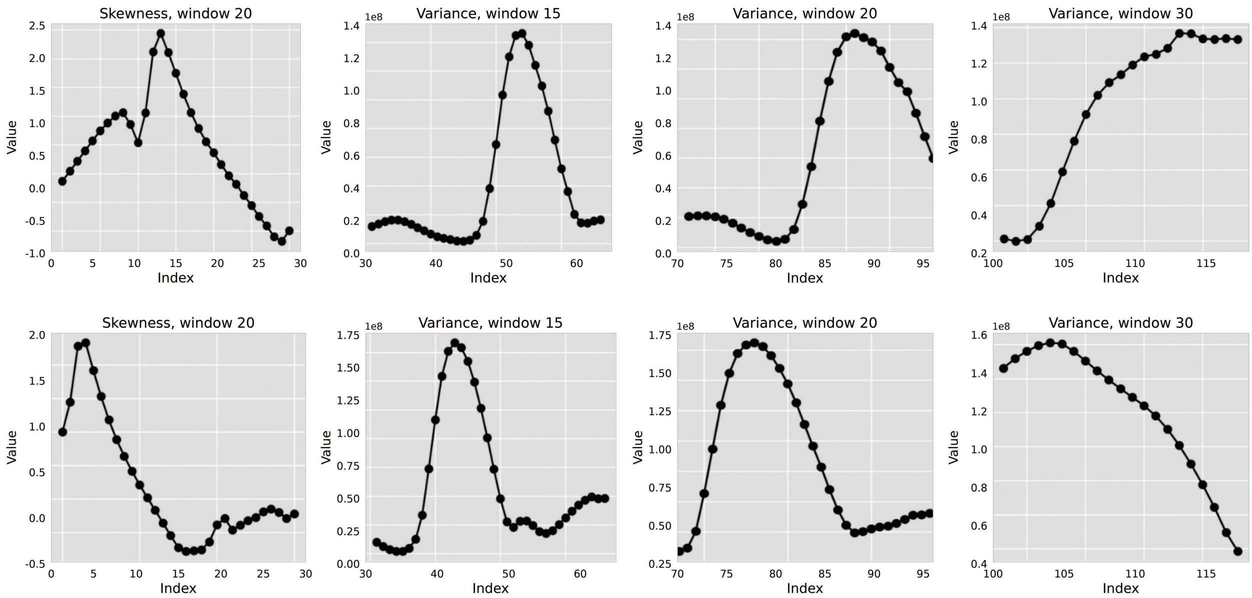


**Fig. 3.** EWSs near (upper row) and far from (lower row) the critical transition.

The next section describes the methods used to solve the classification and regression problems. The EWSs were calculated using the Python package `ewstools` [16].

### *4.1. Classification Approach*

The data were preprocessed before model training and testing. First, the time series was divided into segments without epidemics. Second, a time interval $t$ before epidemic onset was specified. This interval defines the warning lead time. After that, the training sample was constructed.

Data preprocessing was performed as follows. The time series with the number of incidence cases is represented as a vector $X = \{x_k\}$, $k = \overline{1, |X|}$ and is supplemented with a time series of binary labels $Y$:

$$Y = \begin{cases} 1, & \text{epidemic}, \\ 0, & \text{no epidemic}. \end{cases} \tag{1}$$

All non-epidemic segments are then selected:

$$X' = \{X_i \mid \forall j \in [0,\ |X_i| - 1]\ Y_j = 0\}, \quad i = \overline{0, M-1},$$

where $X_i$ is a subset of $X$ and $M$ is the number of non-epidemic segments.

The obtained segments are then transformed using the sliding-window rule:

$$X'' = \begin{pmatrix} X_0^{(0)} & X_0^{(1)} & \dots & X_0^{(w-1)} \\ X_0^{(1)} & X_0^{(2)} & & X_0^{(w)} \\ \vdots & & \ddots & \vdots \\ X_{M-1}^{N-w} & X_{M-1}^{N-w+1} & \dots & X_{M-1}^{N-1} \end{pmatrix}, \tag{2}$$

where $N = |X_{M-1}|$ and $w$ is the window length. In this study, $w = 30$ was used.

The choice of the threshold value $t$ is important. This value determines when the model should report an approaching epidemic. According to epidemic-surveillance specialists, a warning should be received no later than one month before epidemic onset. This corresponds to $t = 4$ or $t = 5$. The incidence time series for seasonal acute respiratory viral infections is recorded weekly. Therefore, $t = 4$ and $t = 5$ correspond to 28 and 35 days. The following binary array describes whether a given time-series point is close to epidemic onset:

$$Y'_l = \begin{cases} 1, & \text{if } E_l - I_l \le t, \\ 0, & \text{otherwise}. \end{cases} \tag{3}$$

Here $l = \overline{0, K}$, and $K$ is the number of rows in matrix $X''$. The value $I_l$ is the index of the last element of the $l$th time series in $X''$. The value $E_l$ is the first index of the nearest epidemic incidence value, where $E_l \in [0, |X|]$.

For each row in the obtained matrix, EWSs were calculated with specified values of the window $w_{ews}$. In this work, variance was calculated for $w_{ews} = \{15,\ 20, 30\}$. Skewness was calculated for $w_{ews} = 20$. The calculated EWS values were then combined into the final matrix:

$$X''' = \begin{pmatrix} S_0^{(0)} & S_0^{(1)} & \dots & V_0'^{,(0)} & V_0'^{,(1)} & \dots & V_0''^{(0)} & \dots & V_0'''^{(29)} \\ \vdots & & \ddots & & & & \vdots & & \\ S_{M-1}^{(0)} & S_{M-1}^{(1)} & \dots & V_{M-1}'^{,(0)} & V_{M-1}'^{,(1)} & \dots & V_{M-1}''^{(0)} & \dots & V_{M-1}'''^{(29)} \end{pmatrix}. \tag{4}$$

Here $S_i^{(j)}$ is the skewness value with index $j$ for time series $i$. The value $V_i'^{,(j)}$ is the variance value with $w_{ews} = 15$ and index $j$ for time series $i$. The values $V''^{(j)}_i$ and $V'''^{(j)}_i$ are variance values with $w_{ews} = 20$ and $w_{ews} = 30$, respectively.

Thus, the final dataset is obtained as

$$D_c = \left\{X_i''',\ Y_i'\right\}, \quad i = \overline{1, K}. \tag{5}$$

The resulting dataset was divided into training, validation, and test samples. The split was 70, 10, and 20%, respectively. The sample elements were not mixed because the data are ordered in time. The dataset is imbalanced: class 0 samples outnumber class 1 samples for all values of $t$. In this task, false negatives are especially important. A false negative means that the model incorrectly reports no epidemic threat. Therefore, recall plays the primary role:

$$Recall = \frac{TP}{TP + FN}. \tag{6}$$

Here $TP$ is the proportion of true positives, and $FN$ is the proportion of false negatives. Accuracy should also remain adequate, for example above 0.5. This condition helps ensure that the models do not classify randomly:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}. \tag{7}$$

Here $FP$ is the proportion of false positives, and $TN$ is the proportion of true negatives. The balance between $FN$ and $FP$ is also important. Therefore, the $F1$ score is used:

$$F1 = 2\frac{Precision \cdot Recall}{Precision + Recall}. \tag{8}$$

The area under the precision–recall curve (PR-AUC) is also useful for model evaluation. It does not depend on the number of true negatives. It also summarizes model performance over a range of threshold values. Therefore, it can be more informative than the $F1$ score.

The following models were used for classification: EWSNet, Easy Ensemble [19], RUSBoost [20], Balanced Bagging classifier [21], Decision Tree, and Random Forest [22]. Deep neural networks with convolutional and recurrent layers can be effective when enough training data are available. Such networks can classify upcoming critical transition points [17, 18]. For this reason, we tested the specialized EWSNet method [18] together with more general approaches. EWSNet combines two independently operating networks: a long short-term memory (LSTM) network and a fully convolutional neural network (FCNN). The authors trained EWSNet on a large synthetic dataset with several dynamical systems. The model shows good accuracy on empirical test data. However, it can still produce frequent false alarms. In the cited study, its test accuracy was about 75%. Its main advantage over classical machine learning models is invariance to the length of the input time series. The comparison results are presented in the next section.

### *4.2. Regression Approach*

For unlabeled COVID-19 data, the task was to forecast an epidemic increase in incidence. The time series of the share of positive COVID-19 tests was represented as an array $X = \{x_k\}$, $k = \overline{1, |X|}$. The time series was split into training, validation, and test sets in proportions of 70, 10, and 20%.

Let $T = \lceil 0.7 \cdot |X| \rceil$ and $V = \lceil 0.1 \cdot |X| \rceil$. Then

$$X_{train} = \{x_k\}, \quad k = \overline{1, T}, \tag{9}$$

$$X_{val} = \{x_k\}, \quad k = \overline{T+1, T+V}, \tag{10}$$

$$X_{test} = X \setminus (X_{train} \cup X_{val}). \tag{11}$$

A set of EWSs can then be calculated. These EWSs form additional time series obtained with a sliding window $w_{ews} = 14$. Each point $x_k$ in $X$ is matched to a point $e_X^k$ in the EWS time series $E_X$. This EWS point is calculated on the interval $[x_{k-w_{ews}},\ x_{k-1}]$. The modeling task is to calculate $\{e_X^{T+V+1}, e_X^{T+V+2}, \dots\}$. In this task, we use only the time series of mean values $E_{X_{mean}}$ and variances $E_{X_{var}}$.

For model training, the resulting time series must be transformed into matrices using a sliding window $m$. This parameter can be determined by evaluating metrics on the validation set. The metrics are described below.

Six metrics are used to evaluate modeling quality: $MSE$, $RMSE$, $MAE$, $NRMSE$, $NMAE$, and $R^2$:

$$MSE = \frac{1}{|X_{test}|} \sum_{k=1}^{|X_{test}|} \left( e_X^{T+V+k} - x_{T+V+k} \right)^2, \tag{12}$$

$$RMSE = \sqrt{MSE}, \tag{13}$$

$$MAE = \frac{1}{|X_{test}|} \sum_{k=1}^{|X_{test}|} |e_X^{T+V+k} - x_{T+V+k}|, \tag{14}$$

$$NRMSE = \frac{RMSE}{\max(X_{test}) \ - \min(X_{test})}, \tag{15}$$

$$NMAE = \frac{MAE}{\max(X_{test}) \ - \min(X_{test})}, \tag{16}$$

$$R^2 = 1 - \frac{\sum_k \left( x_{T+V+k} - e_X^{T+V+k} \right)^2}{\sum_k \left( x_{T+V+k} - \mathrm{mean}(X_{test}) \right)^2}. \tag{17}$$

The $MSE$ was selected as the loss function in the models. The $MAE$ was also calculated because it is easier to interpret from the subject-area perspective. It directly shows the magnitude of forecast errors. The $RMSE$ is stricter because it is more sensitive to outliers than $MAE$. This study analyzes time series with different scales, namely the mean and variance. Therefore, normalized metrics, such as $NRMSE$ and $NMAE$, were used. These metrics are invariant to scale changes. The coefficient of determination $R^2$ was used to estimate the proportion of variance in the reference values explained by the forecast.

Many studies use reservoir computing for short-term forecasting [23–25]. This forecast is short-term because the models are trained only on segments without critical transitions. If transition segments are included, the dynamics may be modeled incorrectly. For example, in [30], recurrent neural networks were used to forecast critical transitions in ecological systems. The networks were trained on raw time series without EWSs, and not all models produced correct estimates. In our case, information is available about some critical transitions that have already occurred. This information can be used for longer-term forecasting. Therefore, we used an LSTM network and a GRU network. GRU is a simplified version of LSTM with one fewer component in its cell. LSTM networks are considered robust and accurate for forecasting chaotic dynamical systems [31]. In the experiments, both LSTM and GRU networks had 128 cells. The selected sliding-window width $m$ was 20. This corresponds to 20 days, or 20 input neurons. In all cases, the models were trained for 100 epochs with the ADAM optimizer.

## 5. EXPERIMENTAL RESULTS

### *5.1. Classification Results*

Tables 2 and 3 present the test-set metrics for threshold values $t = 4$ and $t = 5$, respectively. The metrics were obtained with the optimal hyperparameters. Confusion matrices were calculated in the same way (Fig. 4).

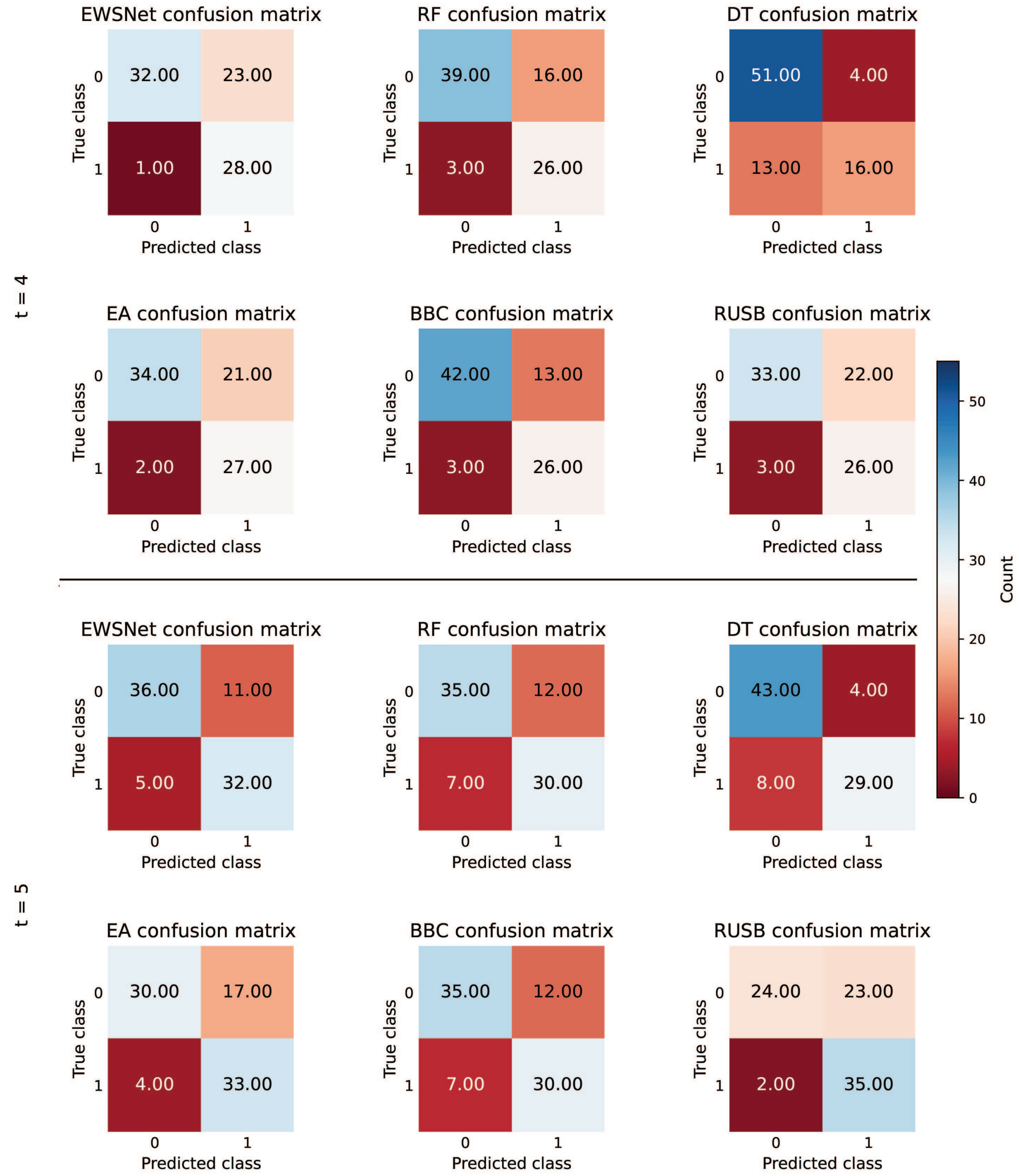


**Fig. 4.** Confusion matrices at $t = 4$ (two upper rows) and $t = 5$ (two lower rows).

The experiments support the following conclusions:

- For Decision Trees, accuracy varies from moderate to high values, approximately 0.75–0.85. However, recall and the $F1$ score vary strongly depending on the threshold value. This method shows unsatisfactory recall at the lower time-threshold value ($t = 4$). Its lower PR-AUC compared with other ensemble methods indicates weaker performance in ranking positive samples.

**Table 2.** Metric values for the four-week horizon ($t = 4$)

| Classifier | Accuracy | Recall | F1 score | PR-AUC |
|---|---|---|---|---|
| Decision Tree | 0.79 | 0.55 | 0.65 | 0.75 |
| Easy Ensemble | 0.72 | 0.93 | 0.70 | **0.87** |
| RUSBoost | 0.70 | 0.89 | 0.67 | 0.78 |
| Balanced Bagging | **0.80** | 0.89 | **0.76** | 0.84 |
| EWSNet | 0.71 | **0.96** | 0.70 | 0.86 |
| Random Forest | 0.77 | 0.89 | 0.73 | 0.85 |

**Table 3.** Metric values for the five-week horizon ($t = 5$)

| Classifier | Accuracy | Recall | F1 score | PR-AUC |
|---|---|---|---|---|
| Decision Tree | **0.85** | 0.78 | **0.82** | 0.89 |
| Easy Ensemble | 0.78 | 0.89 | 0.78 | **0.92** |
| RUSBoost | 0.75 | **0.90** | 0.73 | 0.90 |
| Balanced Bagging | 0.77 | 0.81 | 0.75 | 0.86 |
| EWSNet | 0.80 | 0.86 | 0.80 | 0.88 |
| Random Forest | 0.77 | 0.81 | 0.75 | 0.90 |

- Easy Ensemble shows consistently good results for different threshold values in terms of recall, PR-AUC, and the $F1$ score. Adaptive boosting combined with sampling effectively handles class imbalance. High recall indicates clear identification of minority-class samples. However, this sometimes reduces precision, so the $F1$ score is not always the highest.
- RUSBoost provides balanced results with good recall and PR-AUC values. The cost of the method's simplicity and robustness under class imbalance is a slightly lower accuracy at lower threshold values ($t = 4$) compared with Easy Ensemble.
- Balanced Bagging can be considered a reliable classification tool with moderate recall, PR-AUC, and $F1$ score values. The use of oversampling methods such as SMOTE at $t = 4$ improves the detection of minority-class samples.
- The specialized EWSNet method demonstrates high recall, but is slightly inferior in accuracy and the $F1$ score. It works reliably at $t = 4$, reaching the maximum recall value in the experiment, 0.965, but its lower precision limits its competitiveness in terms of the $F1$ score.
- Random Forest shows stable results. It is structurally simpler and easier to use than several competing methods. It also requires a limited number of features and trees. However, its PR-AUC usually lags behind ensemble methods such as Easy Ensemble and RUSBoost.

### *5.2. Regression Results*

Tables 4 and 5 present the test-set metrics for COVID-19 incidence, which is the second time series. The results are shown for two forecasting horizons: 14 and 28 days. The tables correspond to the GRU and LSTM networks, respectively. The forecasts are shown in Fig. 5.

**Table 4.** Results for GRU

| | 14 days | | 28 days | |
|---|---|---|---|---|
| | Mean | Variance | Mean | Variance |
| $NMAE$ | 0.038 | 0.058 | 0.074 | 0.077 |
| $NRMSE$ | 0.057 | 0.125 | 0.113 | 0.139 |
| $R^2$ | 0.960 | 0.697 | 0.838 | 0.494 |

**Table 5.** Results for LSTM

| | 14 days | | 28 days | |
|---|---|---|---|---|
| | Mean | Variance | Mean | Variance |
| $NMAE$ | 0.032 | 0.063 | 0.054 | 0.077 |
| $NRMSE$ | 0.055 | 0.131 | 0.094 | 0.123 |
| $R^2$ | 0.961 | 0.652 | 0.899 | 0.578 |

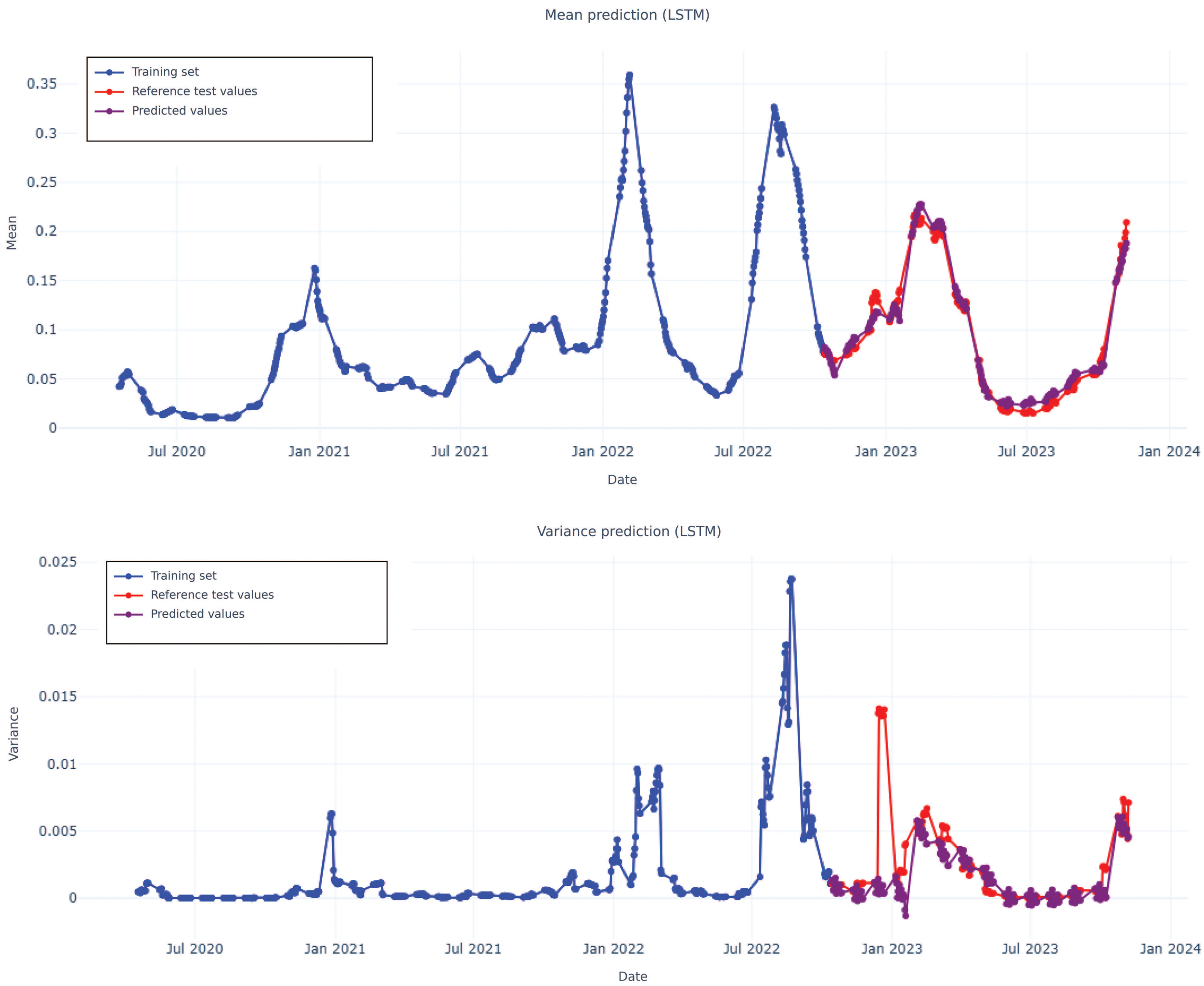


**Fig. 5.** Mean and variance prediction results for the COVID-19 incidence time series.

Tables 4 and 5 show that LSTM performs better than GRU. However, the difference between the models is small, so both approaches can be used. The mean value is forecast more accurately than the variance. As expected, the 28-day horizon gives worse results than the 14-day horizon. Nevertheless, the results remain satisfactory for mean-value prediction.

## 6. CONCLUSION

This paper considered methods for forecasting critical transitions in epidemic time series. The methods combine early warning signals with machine learning. Two main approaches were identified. The first is binary classification, which determines whether a critical transition point is near. The second is predictive modeling of the mean and variance of incidence indicators. In epidemic surveillance systems, predictive modeling is most useful at the early stages of circulation of novel pandemic strains. At this stage, the seasonality of the infection may not yet be known. Individual epidemic waves may also be difficult to distinguish. In this case, the next increase in incidence can be inferred from interval forecasts of the mean and variance of new infection cases. Later, a new disease may become seasonal, with clear periods of epidemic increase and decline. Then it becomes possible to estimate the threshold incidence that separates epidemic waves. Retrospective data can then be used to label time-series segments. In this case, we propose switching to simpler classification-based outbreak detection algorithms.

The experimental results show that Easy Ensemble and EWSNet are the most effective models for the classification task. Their metrics change only slightly across the considered threshold values. They also show high recall and satisfactory precision. For incidence forecasting, recurrent neural networks, including LSTM and GRU, model the mean value effectively. However, they have difficulty predicting variance. Variance is less smooth and more sensitive to fluctuations, so it is a harder target for these models. The values of $NRMSE$ and $R^2$ indicate satisfactory prediction of mean-value dynamics. LSTM slightly outperforms GRU.

This study shows that EWS calculation can be combined with machine learning to forecast critical transitions in epidemic systems. The results may be useful not only in public health but also in other fields where critical phenomena must be understood and forecast. Future work will expand the test datasets and examine improved model architectures. This should increase the robustness and applicability of the proposed methods in other scenarios. One promising direction is the use of multivariate machine learning methods, including methods based on multimodal data. Such models may account for temperature and humidity [29], disease-related user search queries [26], and retrospective incidence data. They may also include data on telemedicine consultations for cough and emergency-department visits for bronchiolitis [27]. In Russian epidemic surveillance, these approaches are limited by restricted data availability and low data quality. However, the models can already be implemented and tested on synthetic data. This would prepare them for future use as data collection improves and new demographic and epidemic datasets become available.

## FUNDING

This work was supported by the Russian Science Foundation under Agreement no. 24-25-20104 dated April 12, 2024. It was also supported by the St. Petersburg State Autonomous Institution "Fund for the Support of Scientific, Scientific-Technical, and Innovative Activities" through a regional grant for co-financing the project under Agreement no. 24-25-20104 dated May 22, 2024, agreement no. 169 dated June 3, 2024.

## REFERENCES

1. Ramadevi, B. and Bingi, K., Chaotic time series forecasting approaches using machine learning techniques: A review, *Symmetry*, 2022, vol. 14, no. 5, article no. 955.

2. Leonenko, V.N., When a pandemic comes to town: Combating propagation of novel viruses in Russian cities using statistical and mathematical modeling, in *Proc. 2024 IEEE 3rd Int. Conf. on Problems of Informatics, Electronics and Radio Engineering (PIERE)*, IEEE, 2024, pp. 990–995.

3. Bulychev, Yu.G., Signal Recognition without State Space Expansion Based on Observations Containing a Singular Interference: The Case of Nonlinear Parameters of Basis Functions, *Autom. Remote Control*, 2024, vol. 85, no. 2, pp. 147–161.

4. Bury, T., Bauch, C.T., and Anand, M., Detecting and distinguishing tipping points using spectral early warning signals, *J. R. Soc. Interface*, 2020, vol. 17, no. 170, article no. 20200482.

5. Southall, E., Brett, T.S., Tildesley, M.J., et al., Early warning signals of infectious disease transitions: A review, *J. R. Soc. Interface*, 2021, vol. 18, no. 182, article no. 20210555.

6. Zhang, X., Yang, L., Chen, T., et al., Predicting influenza-like illness trends based on sentinel surveillance data in China from 2011 to 2019: A modelling and comparative study, *Infect. Dis. Model.*, 2024, vol. 9, no. 3, pp. 816–827.

7. Alfred, R. and Obit, J.H., The roles of machine learning methods in limiting the spread of deadly diseases: A systematic review, *Heliyon*, 2021, vol. 7, no. 6, article no. e07371.

8. Adugna, A., Abebaw, D., Abebaw, A., and Jemal, M., Deep learning architectures for influenza dynamics and treatment optimization: A comprehensive review, *Front. Artif. Intell.*, 2025, vol. 8, article no. 1521886.

9. Zhao, Q., Moniz, N., Korotasz, A., et al., Early warning signals of emerging infectious diseases, *bioRxiv*, 2025, article no. 2025.03.03.641350.

10. Bizzotto, A., Guzzetta, G., Marziano, V., et al., Increasing situational awareness through nowcasting of the reproduction number, *Front. Public Health*, 2024, vol. 12, article no. 1430920.

11. Gao, S., Chakraborty, A.K., Greiner, R., et al., Early detection of disease outbreaks and non-outbreaks using incidence data: A framework using feature-based time series classification and machine learning, *PLOS Comput. Biol.*, 2025, vol. 21, no. 2, article no. e1012782.

12. Drake, K.O., Boyd, O., Franceschi, V.B., et al., Phylogenomic early warning signals for SARS-CoV-2 epidemic waves, *EBioMedicine*, 2024, vol. 100, article no. 104939.

13. O'Brien, D.A. and Clements, C.F., Early warning signals predict emergence of COVID-19 waves, *medRxiv*, 2021, article no. 2021.06.24.21259444.

14. Kouprianov, A., COVID-19.SPb.monitoring. Monitoring COVID-19 epidemic in St. Petersburg, Russia: Data and scripts, 2021. https://github.com/alexei-kouprianov/COVID-19.SPb.monitoring.

15. Brett, T., Ajelli, M., Liu, Q.H., et al., Detecting critical slowing down in high-dimensional epidemiological systems, *PLOS Comput. Biol.*, 2020, vol. 16, no. 3, article no. e1007679.

16. Bury, T.M., ewstools: A Python package for early warning signals of bifurcations in time series data, *J. Open Source Softw.*, 2023, vol. 8, no. 82, article no. 5038.

17. Bury, T., Sujith, R.I., Pavithran, I., et al., Deep learning for early warning signals of tipping points, *PNAS*, 2021, vol. 118, no. 39, article no. e2106140118.

18. Deb, S., Sidheekh, S., Clements, C.F., Krishnan, N.C., and Dutta, P.S., Machine learning methods trained on simple models can predict critical transitions in complex natural systems, *R. Soc. Open Sci.*, 2022, vol. 9, no. 2, article no. 211475.

19. Liu, X.Y., Wu, J., and Zhou, Z.H., Exploratory undersampling for class-imbalance learning, *IEEE Trans. Syst. Man Cybern. B Cybern.*, 2009, vol. 39, no. 2, pp. 539–550.

20. Seiffert, C., Khoshgoftaar, T.M., Van Hulse, J., and Napolitano, A., RUSBoost: A hybrid approach to alleviating class imbalance, *IEEE Trans. Syst. Man Cybern. A Syst. Hum.*, 2010, vol. 40, no. 1, pp. 185–197.

21. Maclin, R. and Opitz, D., An empirical evaluation of bagging and boosting, in *Proc. Fourteenth Natl. Conf. on Artificial Intelligence*, 1997, pp. 546–551.

22. Bishop, C., *Pattern Recognition and Machine Learning*, Berlin: Springer, 2006.

23. Deb, S., Sidheekh, S., Clements, C.F., Krishnan, N.C., and Dutta, P.S., Machine learning methods trained on simple models can predict critical transitions in complex natural systems, *R. Soc. Open Sci.*, 2022, vol. 9, no. 2, article no. 211475.

24. Lim, S., Giorgini, L., Moon, W., and Wettlaufer, J.S., Predicting critical transitions in multiscale dynamical systems using reservoir computing, *Chaos*, 2020, vol. 30, no. 12, article no. 123126.

25. Kong, L.-W., Fan, H.W., Grebogi, C., and Lai, Y.C., Machine learning prediction of critical transition and system collapse, *Phys. Rev. Res.*, 2021, vol. 3, no. 1, article no. 013090.

26. Su, K., Xu, L., Li, G., et al., Forecasting influenza activity using self-adaptive AI model and multi-source data in Chongqing, China, *EBioMedicine*, 2019, vol. 47, pp. 284–292.

27. Morbey, R.A., Todkill, D., Moura, P., et al., Using machine learning to forecast peak health care service demand in real-time during the 2022–23 winter season: A pilot in England, UK, *PLOS One*, 2025, vol. 20, no. 1, article no. e0292829.

28. Cho, G., Park, J.R., Choi, Y., et al., Detection of COVID-19 epidemic outbreak using machine learning, *Front. Public Health*, 2023, vol. 11, article no. 1252357.

29. Amendolara, A.B., Sant, D., Rotstein, H.G., and Fortune, E., LSTM-based recurrent neural network provides effective short term flu forecasting, *BMC Public Health*, 2023, vol. 23, no. 1, article no. 1788.

30. Lapeyrolerie, M. and Boettiger, C., Limits to ecological forecasting: Estimating uncertainty for critical transitions with deep learning, *Methods Ecol. Evol.*, 2022, vol. 14, no. 3, pp. 785–798.

31. Sangiorgio, M. and Dercole, F., Robustness of LSTM neural networks for multi-step forecasting of chaotic time series, *Chaos Solitons Fractals*, 2020, vol. 139, article no. 110045.

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

ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 542–568.




# Control in Linear Best-Response Games on Networks: A Review of Models and Methods

**I. V. Petrov**[*,**]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
[**] *National Research University Higher School of Economics, Moscow, Russia*
*e-mail: ivpetrov@ipu.ru*

Received February 11, 2025
Revised November 20, 2025
Accepted November 25, 2025

**Abstract**—Numerous studies are devoted to systems of connected agents and network control. Historically, the greatest interest of control theory has been in averaging systems, particularly the consensus problem. However, network interaction can be characterized by more specific functions reflecting the dependence on the actions of network neighbors, which is particularly evident in models of strategic interaction on a network, the subject of game theory. This paper surveys game-theoretic models of network interaction from the class of linear best-response games. The models are formally described, and formulations of control problems in this class of games are considered. Special attention is paid to the connection with consensus models and the well-known related control problems that can be stated in terms of a strategic interaction of agents. Despite the general similarity with the models of linear systems from control theory, this class of games allows capturing qualitative aspects of strategic interaction of connected agents and highlighting the role of structural characteristics.

*Keywords*: peer effects, games on networks, linear systems, network control

**DOI:** 10.7868/S1608303226060066

## 1. INTRODUCTION

The scientific interest in control of network interaction between different agents began to emerge in the mid-20th century, following the advent of the concept of networks. Networks were intended to describe systems of interacting agents; unlike a graph as a mathematical object strictly formalizing the nature of a link, networks do not specify the interaction mechanisms. In sociology [1, 2], economics [3, 4], and law [5, 6], the network nature of interaction implies the presence of formal and informal links between agents, which are not always possible to describe mathematically; but such links make the processes of decision-making, propagation/dissemination (e.g., of information, activity, diseases), etc., interdependent.

A peculiarity of many social systems is individual incentives of participants, which was reflected in the emergence, nearly during the same period, of game theory as a tool for describing conflicts. Since the end of the 20th century, the two approaches—network and game-theoretic—have begun to be used jointly. Games in which agents are connected by a network structure (or the network structure results from their interaction) use the results of graph theory and network analysis to determine the influence of network structure on strategic behavior.

Simultaneously, formal analysis and control methods of network systems were developed. In automatic control theory and its applications, the so-called averaging systems and consensus models [7–11] gained enormous popularity among researchers, particularly the classical DeGroot model [12] for the dynamics of opinion formation in a network structure. Mathematical methods for describing systems of strategically interacting agents have specific features, and the interaction result (outcome) often differs from consensus. Decisions made by agents under the influence of their interaction with other agents generate inefficiency, so it is topical to design and analyze various control methods.

Since the early 2000s, several control approaches have been proposed based on network characteristics of agents in socio-economic systems. Approaches to control the interaction of agents have become particularly popular in economics: models of competition and cooperation among firms, interbank interaction, and models of public goods consumption. Several game-theoretic models closely related to econometric models for identifying the influence of an agent's environment on their behavior have emerged; in such models, researchers describe the relationship between an agent's choice and the choices of neighbors. And the greatest interest has been attracted by games where the player's best response linearly depends on the actions of neighbors; formal statements of control problems in such models have arisen as well. This review is devoted to the models of the above class and corresponding control methods.

### *1.1. Terminology*

In mathematical models of the relationship between agents' actions, the starting point is the concept of endogenous social effects [13]. Depending on the context, sociologists and social psychologists use various terms, e.g., peer influences, social norms, neighborhood effects, social interactions, herd behavior, conformism, etc. In other words, interacting with each other, agents (people, organizations, etc.) begin to learn, imitate, or copy the behavior of their environment.

In microeconomics and econometrics research, the influence of such interaction on individual behavior is usually called peer effects [16]. They are defined as the correlation of actions among members of the same reference group. To identify peer effects, the Linear-in-Means (LIM) model [13] was proposed; it studies the influence of social interactions on the behavior and outcomes of individuals in a group. The model is used to analyze how the behavior or characteristics of a group (in the original formulation, average values) affect the behavior of an individual member.

Historically, groups were understood as parts of a single population (e.g., residents of one community, classmates, colleagues, or companies competing in the same market), which was reflected in econometric models of peer effects by an explanatory variable characterizing the behavior of one of the disjoint groups. With the conceptualization of networks, it became possible to consider individual links between people (network effects)[1] to refine existing models and resolve a number of technical identification problems [16]. Identification issues of such models were discussed in [13, 16, 22, 23], as well as in the reviews [24, 25] devoted to econometric studies and [26, 27] devoted to the identification of peer effects in field and laboratory experiments.

Linear best-response games can be treated as generative models of peer/network effects: whereas econometric models focus on identifying these effects and causal relationships from real data, network games provide a theoretical basis for explaining the formation of peer effects by modeling the strategic behavior of agents. In game-theoretic models, the main attention is paid to the role of the interaction structure: how the structure affects equilibrium in the game; what is the result of local interaction effects of nodes, in terms of aggregated characteristics; which nodes and links are key. The interest in linear models is mostly motivated by the crucial role of the network in the analysis of mutual influence of nodes, equilibrium, and control problems.

[1] In their work on the theory of industrial organization [17], the authors proposed applying the term "network effects" to markets with increasing returns to scale (e.g., see [18]), and the term "network externalities" to markets where agents' actions generate inefficient equilibria. While some of the models below do capture firm interactions in markets [19, 20], the terms are not always appropriate according to [17]. With this aspect in mind, the term "network effects" will be used here both as a synonym for peer effects and in a broader sense, referring to situations in which agent's actions depend on the actions of those connected to this agent in the network, as in [21].

### *1.2. Network Interventions*

In recent decades, many results have been obtained concerning the structure of real networks, their formation mechanisms, and the way networks affect the behavior of their participants. Examples of using network structure to elaborate economic policy have appeared [28]; and tools for estimating treatment effects in the cases of interconnected individuals have been proposed [29].

In the research of strategic interaction networks, authors examine how information about the network should be considered for solving control problems in such models.

Network effects are examples of externalities, i.e., situations where a human's well-being depends on the actions of other people without mutually agreed compensation [30]. In this vein, linear best-response games describe the influence of linear externalities in the actions of players connected to the environment. From the perspective of game-theoretic interaction models, this means that equilibrium is inefficient in the presence of externalities; for details, see subsection 2.1.

In economics, externalities have always been regarded as a source of inefficiency, and the classical way to use or counteract externalities is direct intervention in the economic process (economic interventionism). Similarly, the concept of network interventions has emerged, i.e., the process of using social interaction network data to change the outcomes of network participants.[2] The term "network interventions" gained particular popularity in the literature after the same-name paper [31], where it was formulated rather in the context of sociological studies. There exist various network intervention strategies, and assessing the applicability and effectiveness of these methods is of interest. As will be shown below, equilibrium in the game-theoretic models under consideration is always inefficient, which motivates researchers to develop control methods.

In statistical causal inference methods, the estimation of the average treatment effect traditionally assumes independent treatment effects across individuals [32]. Under network interaction, agents can influence each other, violating this assumption. In practical situations, the interaction outcomes of actors depend on the behavior of their network neighbors, and the presence of links influences the average treatment effect. This is usually described by the term "network interference": under network interaction effects, the average effect for treated individuals and the overall effect diverge, and ignoring these factors underestimates the final treatment effect [29].

### *1.3. Applications*

There are many quantitative assessment examples for the influence of group behavior on individual outcomes. Throughout this paper, the variable $x_i$ denotes the decision made by agent $i$. In most cases, $x_i$ is understood as the agent's effort or action in a process of interest to researchers. Standard examples in the literature include the following:

- educational processes, probably the most popular application area [14, 15, 33, 34]. The variable $x_i$ can be the student's average grade or the results of sociological surveys. While econometric models analyze how student's performance depends on the environment, the game-theoretic setting formalizes the strategic nature of efforts: for example, students choose the time spent on education or participation in additional activities, taking into account the choices of their classmates and friends;
- social norms and deviant behavior. There are many situations in which unethical behavior spreads in social networks. For example, some corporate cultures encourage weak ethical norms [35], and the presence of unethical peers in an academic context increases students' temptation to cheat [36, 37]. In this case, game-theoretic interaction models allow investigating the role of social norms in the choice of individual efforts and their position in the network structure. Here, the variable $x_i$ is an indicator function characterizing the performance of a certain action;

[2] Network interventions can be treated as an analog of the term "network control" used for technical systems.

- crime networks [38]. This is historically the first example[3] underlying the linear quadratic game model [39] (see subsection 2.2) and the key player problem (see subsection 3.2); here, $x_i$ is the crime effort profile. Econometric models assess how links between agents manifest deviant behavior [40]. Network games show how the effects of strategic complementarity (or substitutability, see subsection 2.1) in the efforts of interacting agents propagate through the network;
- competition and cooperation of firms [19]. Here, $x_i$ is the output (normalized sales volume). The authors investigate the impact of collaboration in R&D work. In general, the issue of interest is the role of interfirm relations in economic development: the effect of exogenous expansion of business networks on firm productivity and the underlying mechanisms [41].

The main criticism of econometric models for identifying network effects is associated with the problem of endogeneity [13, 23], namely, the correlation between the characteristics and actions of agents, as well as between the environment and the model error. Various methods have been proposed to combat endogeneity, such as two-stage least squares with instrumental variables, nonlinear regression models, etc. (The interested reader can find the details in the surveys [24, 25].)

The purpose of this paper is to review control models and methods proposed for the case of the linear dependence of players' best-response functions on the actions of neighbors within the game-theoretic interaction framework on a network. This paper is organized as follows. In the Introduction, we briefly describe the history of network models of strategic interaction and control problems, highlighting the significant role of identification models for network effects (peer effects). Section 2 is devoted to linear best-response games, including their classification and analysis tools, as well as examples of such games and control problems solved in comparative statics. In Section 3, the Principal's problem—change the equilibrium outcome of the game—is presented; a classification of corresponding control methods, examples of control problems, and a description of solutions to several control problems for strategically interacting agents are provided. In the Conclusions, we indicate some examples from topical directions for further research.

## 2. LINEAR BEST-RESPONSE GAMES

### *2.1. Player's Payoff and Linear Best Response*

Consider a noncooperative one-stage game on a network with complete information. A set of agents $N = \{1, \ldots, n\}$ strategically interact on the network defined by a symmetric adjacency matrix $G = \{g_{ij}\} \in \{0,1\}^{n\times n}$, $g_{ii} = 0$ : they try to benefit as much as possible, taking into account the actions of other players. Each participant $i$ chooses some nonnegative action $x_i$ and receives a payoff $v_i$ from the decision. The payoff (utility) function

$$v_i(x) = f(x_i, z(x_{-i}, G)) \tag{1}$$

depends on the player's action $x_i$, i.e., the benefits and costs from the chosen action, and network effects (the actions of other players $x_{-i} = \{x_j\}_{j\neq i}$); the latter dependence is represented as a function of the player's environment $z(x_{-i}, G)$ considering the network structure (see the details below). In addition to these two components, the model may include the so-called pure externalities, i.e., the dependence of the agent's payoff on other exogenous parameters (not associated with players' actions).

Agents maximize their individual payoffs; for this purpose, each agent $i$ needs to best respond to the actions of other agents $x_{-i}$ :

$$BR_i(x_{-i}, G) = \operatorname*{Argmax}_{x_i\geqslant 0} v_i(x_i, z(x_{-i}, G)). \tag{2}$$

[3] https://lens.monash.edu/2019/08/06/1375976/network-science-identifying-key-players-in-collective-dynamics

The function $BR_i(x_{-i})$ is called the best-response/reply function of players: being independent of the actions of player $i$, this function specifies the latter's optimal response to the actions of other participants. An outcome $x^* = \{x_i^*\}_{i \in N}$ is called a Nash equilibrium in the class of pure strategies of the game if, for all $i$,

$$x_i = BR_i(x_{-i}, G). \tag{3}$$

Each agent is subjected to direct and indirect network effects. A direct effect on agent $i$ is exerted by the network neighbors (the vertices adjacent to vertex $i$). Indirect effects arise due to the network nature of interaction: an agent is indirectly influenced by the neighbors of neighbors (second-level neighbors), those of neighbors (third-level neighbors), etc. Similar reasoning applies to the influence of network links on agents' actions: a direct effect of an edge of graph $G$ is received by the vertices connected by this edge, and an indirect effect is received by other vertices connected to the original pair of vertices.

A key aspect of the game is the particular dependence of player $i$ on the actions of network neighbors: in linear best-response games, the set of all best responses of player $i$ is a singleton, and the best-response function linearly depends on the player's environment $z_i(x_{-i}, G)$; and equilibrium in such a game satisfies a system of linear algebraic equations, which are the best responses of players under $x \geqslant 0$. Currently, the following classification of player environment functions is common:

- local aggregate functions:

$$z(x_{-i}, G) = \beta \sum_{j=1}^{n} g_{ij} x_j;$$

  each player's payoff depends on the sum of the actions of network neighbors;

- local average functions:

$$z(x_{-i}, G) = \beta \sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j,$$

  where $d_i$ is the degree of node $i$; each player's payoff depends on the average action of network neighbors.

The parameter $\beta \in \mathbb{R}$ has a multiplicative effect and characterizes the strength of network interaction (social multiplier [42]). In the absence of the network component ($g_{ij} = 1$ for all $i, j$), both classes reduce to additive aggregation games [43, 44], where each agent is affected by the same set of other agents' strategies; in games on networks, this set depends on the agent and the network.

The distinction between the above classes is how agent $i$ accounts for (or observes) the activity of network neighbors. Unlike local aggregation, where a change in the action of agent $j$ directly affects agent $i$, in the case of averaging, agent $i$ focuses on the average characteristic of the environment regardless of the (high or low) contribution of agent $j$. In other words, local aggregation models are intended to describe the so-called spillover effects (the propagation of contributions of some participants to others), while local averaging models reflect behavioral aspects related to the role of the environment in the agent's choice. As shown in [45, 46], local aggregation and averaging models[4] are different both in terms of payoff function properties and in comparative statics and solutions of control problems. In particular, the role of the parameter $\beta$ also differs: in local aggregation games, increasing the strength of network interaction monotonically increases the players' marginal payoff, which is not the case in the local averaging model (see subsection 2.2).

[4] In this connection, resource network models [47] deserve attention, where both mechanisms—aggregation and averaging—are, in essence, employed simultaneously, albeit not within a game-theoretic framework.

Another classification basis, proposed in [48] (including for non-network models), is the sign of the environment function. Depending on this sign, game-theoretic models receive different interpretations:

- models with strategic complementarity [49] and, accordingly, positive network effect, where the contribution of neighbor $j$ of player $i$ increases the marginal payoff of player $i$ (examples can be found in [39, 45]);
- models with strategic substitutability and, accordingly, negative network effect, where the effort of neighbor $j$ of player $i$ reduces the marginal payoff of player $i$ (examples can be found in [50, 51]).

Formally, the matter concerns the sign of the second cross partial derivative of the player's payoff with respect to the individual action and the action of the neighbors: if $g_{ij} = 1$, then $\frac{\partial^2 v_i}{\partial x_i x_j} > 0$ for games with strategic complementarity and $\frac{\partial^2 v_i}{\partial x_i x_j} < 0$ for games with strategic substitutability. In other words, positive (negative) local network externalities manifest in the actions of agents: for games with strategic complementarity (substitutability), an increase in the efforts of other players ($x_{-i}$) leads to higher actions of player $i$ bringing relatively greater (smaller, respectively) benefit compared to lower efforts of this player. Examples of models and players' best responses are presented in Table 1.

**Table 1.** Models of linear best-response games

| | Strategic complementarity | Strategic substitutability |
|---|---|---|
| Local aggregation | Linear quadratic game $BR_i = b_i + \beta \sum\limits_j g_{ij} x_j$ | Local public goods game $BR_i = b_i - \beta \sum\limits_j g_{ij} x_j$ |
| Local averaging | Social norms in networks $BR_i = (1-\beta) b_i + \beta \sum\limits_{j=1}^{n} \frac{g_{ij}}{d_i} x_j$ | Anticonformism $BR_i = (1-\beta) b_i - \beta \sum\limits_{j=1}^{n} \frac{g_{ij}}{d_i} x_j$ |

Within any economic model, an ideal state of the system (maximum social welfare) is commonly identified. In the case of control problems, this refers to the so-called first-best outcome/contract (social optimum), by analogy with Principal–agent problems in contract theory [52]. In works devoted to control problems in linear best-response games, social welfare is almost always understood as the total payoff of all agents:

$$W = \sum_{i=1}^{n} v_i$$

(see subsection 3.1 for details); and the social optimum (i.e., the solution of the problem $\max_x W$ and the corresponding equilibrium outcomes $x^O$ and payoffs $W^O$) is studied in comparative statics, assuming that agents seek to maximize not their individual payoff functions but the welfare of all agents in aggregate. As will be demonstrated below (on examples of particular games), the Nash equilibrium is never the social optimum: the network effect either underestimates or overestimates the actions of agents. Therefore, it is topical to develop control methods in such games.

### *2.2. Mathematical Models of Network Effects: Coordination, Competition, Cooperation, and Awareness*

**Coordination game**

The classical model of social influence and opinion dynamics in a network [53, 54] is a special case of the consensus problem, which has important applications in engineering control problems [8]. This model can also be treated as a microeconomic model of strategic decision-making by agents [55].

That is, let $n$ agents seek to maximize their individual payoff characterized by the function

$$v_i = -\sum_{j=1}^{n} \frac{g_{ij}}{d_i}(x_i - x_j)^2. \tag{4}$$

In other words, it is disadvantageous for an agent to make a choice (e.g., technology or language) differing from those of neighbors. Due to the concavity of the function $v_i(x)$, the first-order optimality conditions are sufficient to calculate the system of best responses: the optimal response of player $i$ to the actions of network neighbors,

$$BR(x_{-i}) = \sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j, \tag{5}$$

coincides with the opinion dynamics model. For instance, DeGroot's model can be interpreted as a game where participants exhibit behavior similar to that of network neighbors by best responding to their actions.

In such a game, any action profile where all agents take the same action constitutes a Nash equilibrium. Moreover, agents possess no individual characteristics distinguishing them from each other. Therefore, a model reflecting the interaction of agents with individual skills under the average behavior of their environment has become popular.

**Social norms in networks**

From a game-theoretic interaction standpoint, another modification of DeGroot's model is of interest, originally proposed in [56]. (The Friedkin–Johnsen model describes the best-response dynamics of players; the dynamics were analyzed, e.g., in [57].) In this modification, each node $i$ additionally maintains a constant internal opinion. There are several popular versions of the model in a game-theoretic setting, particularly the one with incomplete information where agents try to determine some state of the world [58]. In this case, the interpretation of the model is close to the well-known beauty contest game [59].

Below, we consider the local averaging model, also known as the social-norm model [46]. Each player receives a payoff[5] of the form

$$v_i = b_i x_i - \frac{1}{2}\left(\frac{\beta}{1-\beta}\right)\left(x_i - \sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j\right)^2 - \frac{1}{2}x_i^2, \tag{6}$$

where $b_i > 0$ reflects the individual productivity of agent $i$, and $\beta > 0$ is the taste for conformism. Agents undertake costly actions $(-x_i^2/2)$, receive the marginal payoff $b_i x_i$ independent of the neighbors' efforts, and compare their individual contribution $x_i$ with the average contribution of their network neighbors, seeking to minimize the difference between their individual action and the average action of their group. In view of the significant role of the term $\sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j = \overline{x}_i$, the authors called it the *social norm* of agent $i$ in other publications in the context of sociological studies.

[5] Other payoff functions can be considered, e.g.,

$$v_i = b_i x_i - \frac{\theta}{2}\sum_j \frac{g_{ij}}{d_i}(x_i - x_j)^2 - \frac{x_i^2}{2}$$

or

$$v_i = b_i x_i - \theta x_i \sum_j \frac{g_{ij}}{d_i} x_j - \frac{1+\theta}{2} x_i^2,$$

where $\theta = \frac{\beta}{1-\beta}$. According to an observation of the authors [46], these payoff functions share the same best-response function; however, even if the equilibrium efforts $x^*$ are identical, welfare analysis and its comparative statics may differ since the equilibrium utilities, and hence welfare, are not the same.

The first-order optimality conditions $\partial v_i/\partial x_i = 0$ yield the players' best responses

$$x_i = (1-\beta)b_i + \beta\overline{x}_i, \tag{7}$$

which can be equivalently written in the matrix form

$$x = (1-\beta)b + \beta\hat{G}x, \tag{8}$$

where $\hat{G} = \{\hat{g}_{ij} = g_{ij}/d_i\}^{n\times n}$. For $\beta < 1$, there exists a unique Nash equilibrium in pure strategies:

$$x^* = (1-\beta)(I-\beta\hat{G})^{-1}b. \tag{9}$$

First, note the effect of individual productivity. Although the influence of $b_i$ on $v_i^*$ is always positive, the mutual influence of agents' productivity can be multidirectional:

$$\frac{\partial v_i^*}{\partial b_j} \gtreqless 0 \iff x_i^* \gtreqless \overline{x}_i^*. \tag{10}$$

In other words, the equilibrium utility of agent $i$ increases (decreases) in response to a small change in the productivity $b_j$ of agent $j$ iff the equilibrium response of player $i$ is greater (smaller, respectively) than the equilibrium social norm of this player. Of main interest is the dependence of the game outcome on the parameter $\beta$. As shown by the authors, the dependence is generally nonmonotonic, and they considered two extreme cases as follows:

- pure individualism ($\beta = 0$), when each player's equilibrium response equals the individual productivity:

$$\sum_{j=1}^{n} x_j^* = \sum_{j=1}^{n} b_j;$$

- total conformism ($\beta \to 1$), when all agents choose the same effort equal to the weighted average of individual productivity:

$$\lim_{\beta\to 1}\sum_{j=1}^{n} x_j^* = n\sum_{j=1}^{n} \overline{d}_j b_j,$$

where $\overline{d}$ is the vector of normalized vertex degrees, $\overline{d_i} = d_i/\sum_{j=1}^{n} d_j$. When will the total contribution $\sum_j x_j^*$ be higher, under pure individualism ($\beta = 0$) or total conformism ($\beta \to 1$)? The answer depends on the correlation between the productivity $b$ and the vertex degree distribution in the graph $G$ :

$$\lim_{\beta\to 1}\sum_{j=1}^{n} x_j^* \gtreqless \sum_{j=1}^{n} x_j^*(\beta = 0) \iff Corr(\overline{d}, b) \gtreqless 0. \tag{11}$$

If $b$ and $\overline{d}$ are positively (negatively) correlated, i.e., the agents with higher productivity occupy more (less) central positions in the network, total conformism will increase (decrease, respectively) the aggregate efforts.

The social optimum in the model, i.e., the solution of the problem

$$\max_x \quad b^T x - \frac{1}{2}x^T H(\beta)x, \tag{12}$$

where $H(\beta) := I + \frac{\beta}{1-\beta}\left(I-\widehat{G}\right)^T\left(I-\widehat{G}\right)$, is achieved by solving the system of all best responses

$$x_i = (1-\beta)b_i + \beta\overline{x}_i + \beta\sum_{j=1}^{n}\frac{g_{ij}}{d_i}\left(x_j - \overline{x}_j\right). \tag{13}$$

The distinction between the optimum and the Nash equilibrium in the game is the last term in this system of the best responses of players (the influence of the agent's individual action on the social norm of neighbors). In the Nash equilibrium, when deciding on their individual efforts, agents neglect this factor, creating a (positive or negative) externality: if $i$ and $j$ are neighbors, then

$$\frac{\partial v_i}{\partial x_j} \gtreqless 0 \iff x_i \gtreqless \overline{x}_i. \tag{14}$$

In other words, by exerting the effort $x_j$, agent $j$ imposes a positive (negative) externality on the neighbor $i$ iff the effort of $i$ is above (below, respectively) the social norm of agent $i$. The optimum is achieved when the players' equilibrium responses match their individual productivity:

$$x^O = b. \tag{15}$$

In this case, a class of regular networks is distinguished where the aggregate efforts are always optimal: $x^* = x^O = b$. The reason is that in regular networks, the positive and negative externalities imposed by agents on their neighbors balance each other, so the aggregate effect becomes optimal.

Individual productivity is also crucial in the problem of adding and removing links between agents: the Principal takes into account only the productivity of agents, while information about the network structure does not matter:

(1) key-link adding: in any network, adding a link between two agents with high (low) productivity increases (decreases, respectively) their efforts and, moreover, increases (decreases, respectively) the efforts of all other agents in the network;
(2) key-link removing: regardless of the network structure, the Principal should remove the link between the two most productive agents.

The solution of the incentive problem will be discussed below when analyzing control in the linear quadratic game (subsection 3.2).

**Linear quadratic game**

The linear quadratic game on a network, proposed in [39], is a most well-known model of network games. The payoff function of player $i$,

$$v_i = x_i \left( b_i + \beta \sum_{j=1}^{n} g_{ij} x_j \right) - \frac{x_i^2}{2} \tag{16}$$

includes benefit from the individual action, $x_i b_i$, and the actions of neighbors, $\beta \sum_j g_{ij} x_j x_i$, as well as the quadratic cost of the decision, $x_i^2/2$. The parameter $b_i > 0$ is the standalone marginal return of player $i$. The parameter $\beta$ reflects the dependence on neighbors' actions: for $\beta > 0$, the actions of players are strategic complements; for $\beta < 0$, strategic substitutes.

In the linear quadratic game, the first-order optimality conditions

$$\frac{\partial v_i}{\partial x_i} = b_i + \beta \sum_{j=1}^{n} g_{ij} x_j - x_i = 0 \tag{17}$$

lead to the following best-response functions:

$$BR_i(x_{-i}) = b_i + \beta \sum_{j=1}^{n} g_{ij} x_j. \tag{18}$$

In matrix form, the system of the best responses can be written as

$$(I - \beta G)x = b,$$

where $I$ denotes an identity matrix of appropriate dimensions. If the matrix $(I - \beta G)$ is invertible (for details, see subsection 2.3), then there exists a unique Nash equilibrium: the equilibrium responses of players (in matrix form) are given by

$$x^* = (I - \beta G)^{-1} b. \tag{19}$$

The idea of the authors [39] was to create a bridge between the concept of Nash equilibrium and measures of centrality for network nodes. Consider a network with an adjacency matrix $G$ and a scalar value $\beta \geq 0$. Then the vector of centralities [60] of the parameter $\beta$,

$$\sum_{k=0}^{+\infty} \beta^k G^k 1 = (I - \beta G)^{-1} 1$$

reflects the total number of paths in $G$ starting at vertex $i$. The parameter $\beta$ is a decay factor reducing the relative weight of the longest paths. Thus, the Nash equilibrium in the linear quadratic game on a graph exactly coincides with the vector of Katz–Bonacich centralities.

A more general model was considered in [61]: an arbitrary weighted network, the set of agent's admissible actions $x_i \in [0, L]$, and the players' best response

$$BR_i(x_{-i}) = \min\left(b_i + \beta \sum_{j=1}^{n} g_{ij} x_j, L\right). \tag{20}$$

For $L = \infty$, the original model is obtained.

The comparative statics of the model were investigated in [62]: the social optimum (the solution of $\max_x \sum v_i$) leads to the following system of players' best responses:

$$b_i - x_i + 2\beta \sum_{j=1}^{n} g_{ij} x_j = 0. \tag{21}$$

It can be equivalently written as

$$x_i^O = BR_i(x_{-i}) + \beta \sum_{j=1}^{n} g_{ij} x_j, \tag{22}$$

i.e., the equilibrium efforts are too small since each agent ignores the positive influence (positive externality arising from complementarity) of the individual action on the choices of neighbors. As a result, the equilibrium in the game turns out to be inefficient. Due to this effect, agents' actions are always underestimated, and the Principal in the incentive problem will choose unidirectional interventions: all agents will either receive additional subsidies or, conversely, be subjected to fines/taxes. (For a detailed discussion of this issue, see subsection 3.2.)

Obviously, the network structure monotonically affects the social optimum as well. In the case of symmetric matrices, the following result was established in [63]: for two matrices $\Sigma$, $\Sigma'$ such that $\Sigma = -I + \beta G$ and $\Sigma' > \Sigma$, if $\sigma'_{ij} > \sigma_{ij}$ for all $i, j$, then

$$x^*(\Sigma') > x^*(\Sigma). \tag{23}$$

In the network formation model proposed in [64], agents choose other agents to create a link by maximizing their individual payoff according to formula (16). Due to the indicated properties of the function, agents with the highest centrality measure will receive the greatest benefit, and others will seek to create links with them.

**Local public goods game**

Games on networks with strategic substitutability describe situations where the action of one agent reduces the incentives for neighbors to increase their efforts. Such games model competitive scenarios, e.g., competition for limited resources, markets, or benefits, where players are connected through an interaction network, and their strategies are substitutable. This effect is opposite to strategic complementarity, where one agent's actions stimulate neighbors to increase efforts.

In the local public goods game on a network, agents decide on their individual contributions to the production of some non-excludable good (i.e., the agents cannot be excluded from its consumption), which benefits both themselves and their network neighbors. In many studies, this model was interpreted as a model of sectoral market theory [65], particularly an oligopoly [66]; and some versions were further developed in the field of business games and behavioral experiments [67]. This model is closely related to oligopoly theory because both consider strategic interaction among players with the strategic substitutability of actions.

Let firm $i$ produce some quantity of goods $x_i$, and let its profit be $v_i(x) = x_i p(\sum_j x_j) - cx_i$, where $p(\sum_j x_j)$ denotes the inverse demand function, and $c$ is the costs. In the case of linear inverse demand, the profit of firm $i$ can be written as

$$v_i = x_i\left(b - \left(x_i + \beta \sum_{j=1}^{n} g_{ij}x_j\right)\right) - cx_i, \tag{24}$$

where $g_{ij} = g_{ji} = 1$ means that the products of firm $i$ and firm $j$ are substitutable, $\beta$ is the degree of substitutability, and $b$ is the market size. More general network models were studied in [19, 68]. Similar models were considered in [69, 70], where they were supplemented by the parameter $r_i$ (type) of agent $i$ (reflecting the efficiency or qualification of the latter's activity). The payoff function of players has the form

$$v_i = x_i\left(b - \sum_{j=1}^{n} x_j\right) - \frac{x_i^2}{2r_i}, \tag{25}$$

and the best response on an arbitrary graph $G$ can be represented as

$$\left(\frac{1}{r}I + \beta G\right)x = b. \tag{26}$$

The model formulated in [19] simultaneously accounts for the effects of strategic complementarity and substitutability: firms compete in markets while cooperating in R&D projects (the so-called Cournot oligopoly game with the spillover effect of R&D collaborations). The player's best response in this model takes the form

$$x_i = b_i - \rho\sum_{j=1}^{n} b_{ij}x_j + \beta\sum_{j=1}^{n} g_{ij}x_j, \tag{27}$$

where $B = \{b_{ij}\}$ is the competition matrix between firms (a link between firms arises only if they compete in the same market), and $G = \{g_{ij}\}$ is the cooperation matrix in R&D. Under rather nontrivial conditions, there exists a unique equilibrium given by

$$x^* = (I + \rho B - \beta G)^{-1}b. \tag{28}$$

Up to this point, we have considered the case of common knowledge, i.e., a situation where each agent has complete information about the individual characteristics of other agents, the structure of links, etc. Reflexive games are one approach to analyzing information asymmetry. A reflexive game is a model of agents' decision-making based on a hierarchy of their beliefs about opponents' behavior [71]. If the awareness structure in such a game has finite complexity (a point-based or pointwise information structure [50] representing a finite hierarchy of beliefs), then a reflexive game graph $\overline{G}$ can be constructed, clearly demonstrating the relationship between the actions of agents participating in equilibrium. In the case of the linear best response of players[6], i.e.,

$$BR_i(x_{-i}) = b_i - \beta \sum_{j=1}^{n} \overline{g}_{ij} x_j, \tag{30}$$

the informational equilibrium can be found using the formula

$$x^* = (I + \beta\overline{G})^{-1} b, \tag{31}$$

where the vector $b$ reflects the awareness of agents, and $\overline{G}$ is the reflexive game graph [72]. The vertices of the graph $\overline{G}$ correspond to real and phantom agents participating in the reflexive game, and the arcs of $\overline{G}$ reflect the mutual awareness of agents.

In this model, solutions of many control problems (particularly in comparative statics) are close to those of problems in the linear quadratic game on a network. For the case of a unique equilibrium in the model, the authors [73] considered the incentive-targeting problem; see subsection 3.2.

### *2.3. Equilibrium Analysis of Linear Best-Response Games*

The existence and uniqueness of equilibrium in the above examples are usually established in two ways: via existence theorems for a solution of a system of linear algebraic equations (as in [39]) or via proofs of the existence of a potential function for a game.

The concept of potential games was pioneered in [74]. The main idea is to show the existence of a function describing the outcome of a strategic interaction of players using a scalar function.[7] A function $\varphi$ is a potential function (best-reply potential) of a game if, for any $x_i, x_i'$, and $x_{-i}$ and all $i \in N$,

$$\varphi(x_i, x_{-i}) - \varphi(x_i', x_{-i}) = v_i(x_i, x_{-i}) - v_i(x_i', x_{-i}). \tag{32}$$

Linear best-response games possess a potential function (generally, nonunique): for example [77], it can be represented in the matrix form

$$\varphi(x) = x^T 1 - \frac{1}{2} x^T (I \pm \beta G) x. \tag{33}$$

The maximum of a potential function corresponds to a Nash equilibrium. A sufficient condition for a unique solution is the concavity of a potential function. The matrix of second derivatives takes the form $\nabla^2 \varphi(x) = -(I \pm \beta G)$, and $\varphi(x)$ is strictly concave iff the matrix $I \pm \beta G$ is positive definite: for any $y \neq 0$,

$$y^T (I \pm \beta G) y > 0. \tag{34}$$

[6] The original paper considered the model

$$v_i = x_i \left( b - \sum_{j=1}^{n} \hat{g}_{ij} x_j \right) - \frac{x_i^2}{2} \tag{29}$$

and $g_{ii} = 1$.

[7] As demonstrated in [75], for some classes of games, the potential function is a Lyapunov function of the game dynamics; hence, the Lyapunov function method can be applied for equilibrium analysis [76].

For local aggregation games, depending on the sign of $\beta G$, a necessary and sufficient condition for the existence of a unique equilibrium is either $\beta\lambda_{\max}(G) < 1$ in the case of strategic complementarity $(-)$, or $\beta|\lambda_{\min}(G)| < 1$ in the case of strategic substitutability $(+)$, where $\lambda_{\max}(G)$ and $\lambda_{\min}(G)$ are the largest and smallest eigenvalues of the matrix $G$, respectively. In local averaging games, these conditions are not required: $G$ is a row-stochastic matrix ($\lambda_{\max}(G) = 1$), and the existence condition becomes $\beta < 1$.

A significant distinction of games with strategic substitutability is the existence of a corner equilibrium [78]: the potential function ceases to be concave, and by the game's property of nonnegative strategies $x^*$, there arises a game situation (strategy profile) where some players—passive agents—are inactive ($x_i = 0$) whereas the others—active ones—choose the maximum admissible action $x_j = b_j$. In this strategy profile, the equilibrium existence condition is related to the presence of a maximal independent set in the graph $G$, which, like the number of equilibria, may be nonunique: active players must not be connected among themselves.

In the case of an asymmetric matrix $G$, a sufficient condition for equilibrium existence is the positive definiteness of the Hermitian component of the matrix $G = \frac{G+G^T}{2}$ and, consequently, $1 + \delta\lambda_{\max(\min)}(\frac{G+G^T}{2}) > 0$. However, in the case of strategic substitutability, the analysis of best-response dynamics is more complicated (see the discussion below).

As shown in [63, 100], an equilibrium in a linear best-response game is a solution of the linear complementarity problem (LCP). According to [76], LCP is a subclass of variational inequality problems[8]; the analysis of games in this vein (including the case of nonlinear best response) was presented in [76, 79, 80] and the references therein. Using the formulation of Nash equilibrium as a solution of the variational inequality problem, the authors [76] investigated the relationship between the parameterization of the game and the matrix of partial derivatives of the best-response operator (the Jacobian of the game), as well as the influence of the Jacobian's properties on the existence and uniqueness of equilibrium and the convergence of best-response dynamics in discrete and continuous time.

### *2.4. Best-Response Dynamics*

Above, the interaction between agents has been formulated as a one-stage noncooperative game, but this interaction can be represented alternatively. In [81], system (18) was formulated as a linear affine dynamic system in continuous time:

$$\dot{x} = Ax + b, \tag{35}$$

where $A = \beta G - I$. For each agent $i$,

$$\dot{x_i} = b_i + \beta \sum_{j=1}^{n} g_{ij} x_j - x_i. \tag{36}$$

The solution of the system, or its steady state, is

$$\begin{aligned} \lim_{t\to\infty} x(t) &= -A^{-1}b \\ &= -(\beta G - I)^{-1}b \\ &= (I - \beta G)^{-1}b. \end{aligned} \tag{37}$$

[8] The formulation of Nash equilibrium as a solution of a variational inequality generalizes the game's potential property [74].

The difference approximation of equation (35) is also of interest[9]: it can be used to establish a relationship between the discretization step $\delta$ and the network effect $\beta$. The particular value of the factor at the interaction matrix is related both to the direct influence of the environment on the agent's choice and to the discretization step of the model. In addition, the best responses of players may be bounded due to the specifics of various real-life situations: in most cases (effort level in collective behavior models, output in production models of goods or services, etc.), the best response is nonnegative; in the popular analysis model of educational processes, agents choose effort levels often measured in hours, and their best responses are bounded (below by zero and above by some threshold) and are also supposed to be integer. All these aspects restrict the choice of appropriate tools for analyzing dynamics in such games [82]. The relationship between the game-theoretic interpretation and the theory of dynamic systems, particularly for the above model, was discussed in the book [83].

In game-theoretic terminology, the procedure for finding equilibrium by players is called tatonnement, e.g., Nash/Cournot tatonnement (borrowed from French). In games with strategic substitutability, an equilibrium that is stable with respect to discrete Nash tatonnement implies stability of equilibrium with respect to continuous best-response dynamics, but the converse is not true [78]. The convergence of continuous best-response dynamics to Nash equilibrium in such games was shown in [84]; the case of farsighted agents was considered in [85].

In a game on a directed graph $G$, best-response dynamics can cycle. Two classes of directed networks for which best-response dynamics converge to a unique equilibrium were introduced in [86]. One class is determined by the possibility of changing the game parameters in a special way so that the dynamics of the new game coincide with those of the original one; the other class of networks is closely related to the Kolmogorov criterion for reversibility of a Markov chain.[10]

### *2.5. Relation to Other Fields*

The main problems and applications in control of networks were considered in [7, 8]. In this field, control issues in networks (in particular, collective, cooperative, and multi-agent control) are studied. Let $x_i(t) \in \mathbb{R}^n$ be the state vector of network nodes at a time instant $t$, $i = 1, \ldots, n$; this vector consists of elements $g_{ij}$ reflecting the link between vertices $i$ and $j$ in a graph $G$. As noted by the authors, a wide class of multi-agent networks is described by continuous-time models of the form

$$\dot{x_i} = F_i(t, x_i, u_i) + \sum_{j=1}^{n} g_{ij}(t) f_{ij}(x_i, x_j), \tag{38}$$

where functions $F_i(\cdot)$ and $f_{ij}(\cdot)$ characterize the agent's local dynamics and interaction with other agents, respectively. In most widespread networks, the interactions of agents depend only on their disagreements (the divergence of their states); and linear networks with the function

$$f_{ij}(x_i, x_j) = x_i - x_j \tag{39}$$

are the best studied by now. The remainder of this review is devoted precisely to such models. At the same time, a formulation similar to (38) is valid not only for consensus models but also for game-theoretic interaction models, while the function $f_{ij}(x_i, x_j)$ has specific features (see subsections 2.1 and 2.4 of the paper).

Input–output (interindustry) models [89, 90] are closely related to the above ones as well; similar issues were studied in works devoted to dynamic input–output models [91–93] and optimal planning

[9] Formula (19) has at least three interpretations: it is the equilibrium of a game-theoretic model, the equilibrium of a linear affine dynamic system, and also a centrality measure of network nodes.

[10] The relationship between directed graphs and the Kolmogorov criterion was also addressed in [87, 88].

models [94]. Let $x_i$ be the output of (economic) sector $i$, and let $G$ be the direct cost matrix. Its elements reflect the costs (input) of sector $i$ to provide one unit of goods or services (output) of sector $j$. Then the total cost matrix is

$$\Lambda = (I - G)^{-1},$$

whose elements reflect the output of sector $i$ required to provide one unit of the final output of sector $j$. In the literature, a more general case of the matrix $\Lambda$, depending on a parameter $\beta$ (the Leontief inverse), is also considered [89]: $\Lambda = (I - \beta G)^{-1}$.

Various ways of describing game-theoretic interaction on a network were addressed in [95, 96]; control problems in the case of strategic network formation, in [97, 98]; and in models of bounded rationality, in [50, 71]. For details, the reader is referred to subsection 2.2. This review does not cover cooperative games [99], although several cooperative solutions of the games described above were proposed in [100–102].

Note separately the works devoted to control problems in social network models [12, 103] and in models of mixed social influence (the mechanisms of assimilative/dissimilative influence and bounded trust) [104]. As demonstrated above, such models can also be treated as models of strategic interaction.

## 3. CONTROL PROBLEMS

### *3.1. Principal's Objective Functions and Budget Constraints*

A control authority (also called Principal or central planner in the literature) receives information about the participants' payoffs and can influence their actions, generally speaking, at different time instants and in various ways. Consequently, the outcomes of agent interaction change, becoming more desirable for the Principal. The subject of control can be various elements of the system, i.e., agents' actions, their incentives, the interaction structure, and other characteristics.

In the case of linear best-response games, the control problem is posed as a general optimization problem of the form

$$\max_u W(x, u) \tag{40}$$

subject to the constraints

$$x \geqslant 0, \quad x = BR(x, u), \quad K(u, u_0) \leqslant C, \tag{41}$$

where $u$ is a game parameter influenced during the control process, with $u_0$ specifying its initial value; $K(u, u_0)$ is a cost function characterizing constraints on parameter variations relative to the initial value; and $C$ is a constant. As shown above, in the analysis of comparative statics and social optimum, the vector $x$ (the players' best responses) can be treated as a control action. Thus, a two-stage game is considered: at the first stage, the Principal applies control, and at the second stage, the agents play out an equilibrium.

Although the same best-response function may correspond to various payoff functions, the solution of the control problem may differ depending on the function maximized by the Principal:

- increasing/decreasing the aggregate outcomes of players:

$$W = \sum_{i=1}^{n} x_i \longrightarrow \max;$$

- increasing the total payoff of all participants, i.e., the Benthamite (utilitarian) social optimum:

$$W = \sum_{i=1}^{n} v_i \longrightarrow \max;$$

- increasing the payoffs of the least well-off participants, i.e., the Rawlsian (maximin) social optimum:

$$W = \min_i v_i \longrightarrow \max.$$

The challenge of choosing the Principal's objective functions in such problems has received little attention so far; for example, see [105]. And the most popular choice is the social welfare function (the sum of agents' payoffs).

There exist several classifications of control problems applicable to the models under consideration [31, 95, 106, 107]. In [31], the following network intervention strategies were presented, each with different tactical alternatives:

- individual (the identification of key network participants based on some network characteristics, see Table 2);
- segmentation (network division into groups of nodes for further intervention);
- alteration (the change of links between participants or the entire interaction network, see Table 2);
- induction (network excitation for activation of new interactions (cascades) between participants, which can be informational/behavioral, etc.; for example, see [108]);

**Table 2.** Control problems in linear best-response games

| Problem type | Description | Publications |
|---|---|---|
| Individual | Key network participants are identified based on some network characteristics | Galeotti and Goyal, 2009 [109]; Candogan, Bimpikis, and Ozdaglar, 2012 [110]; Bloch and Querou, 2013 [111]; Fainmesser and Galeotti, 2016 [112]; Demange, 2017 [113]; Galeotti, Golub, and Goyal, 2020 [73]; Kor and Zhou, 2023 [114]; Belhaj, Deroian, and Safi, 2023 [116]; Jeong and Shin, 2024 [115]; Dasaratha, Golub, and Shah, 2024 [117] |
| Alternation | Links between participants or the entire interaction network are changed | Borgatti, 2006 [118]; Ballester, Calvo-Armengol, and Zenou, 2006 [39]; Corbo, Calvo-Armengol, and Parkes, 2006 [119]; Corbo, Calvo-Armengol, and Parkes, 2007 [63]; Golub and Lever, 2010 [120]; Konig, Tessone, and Zenou, 2014 [64]; Belhaj, Bervoets, and Deroian, 2016 [121]; Matveenko and Korolev, 2016 [122]; Hiller, 2017 [123]; Harkins, 2020 [124]; Li, 2023 [125]; Sun, Zhao, and Zhou, 2023 [126] |

Several authors distinguished institutional interventions, emphasizing the specifics of changing the rules of interaction [95, 106]. In [95], staff control (control of the composition of participants) was singled out into a class of problems; in [31], this class was attributed to alteration problems (control of the interaction structure, structural control).

In works devoted to control in linear best-response games, $u$ can be the staff of players $i \in N$, the incentives $b$, or the interaction structure $G$. Solutions of some problems are described below; and the table lists the related publications according to the above classification.

**Control of agents' incentives**

In such control problems, the Principal changes the individual characteristics of agents' activities. Incentive problems in which the Principal seeks to maximize social welfare in models with strategic complementarity were considered in [73, 113] (including the case of the coordination game [115]). Monopolist's pricing problems under local consumption externalities were studied in [110, 111], particularly the relationship between consumer centrality in the network and the prices and quantities offered by the monopolist.

**Structural control**

Structural control adjusts the level of activity by changing the network structure. Creating or removing links with such control affects the centrality of agents, thereby changing the equilibrium. Researchers investigate optimal networks from the Principal's standpoint. According to [64], the class of nested split graphs (NSGs) stands out among simple networks maximizing a Principal's concave objective function. In any network that is not an NSG, links can be created between agents to improve welfare in the linear quadratic game on this network [121].

**Control under uncertainty**

Many authors resort to analyzing the stability of control methods with respect to exogenous disturbances or probabilistic uncertainty; for example, see [73, 78, 127, 128]. Importantly, data on the network structure are often difficult to obtain [129, 130]; this aspect has motivated the appearance of models for both probabilistic and game-theoretic uncertainty regarding the network structure.

For instance, in [112], the Principal chooses the optimal price level for agents, having only information about the degree distribution of network nodes. In another case considered, the network structure is observable, but the location or identity of agents is private information [131]. In other words, the Principal cannot distinguish networks that are identical up to node permutation. An essential feature of this game design is that the vector $x^*$ of equilibrium responses becomes a contract offered by the Principal to the agents. The authors identified classes of network structures for which the Principal's optimal contracts are incentive-compatible for agents and their coalitions.

**Network identification**

One of the most important problems is network identification. There are several methods for solving this problem based on data from a series of game outcomes, involving a statistical approach (graph regularization) [132], optimization [133], and machine learning [134]. A separate line of research is the solution of bilevel optimization problems[11] [136], where the observed equilibrium responses of players are supposed to arise from optimizing the interaction structure (by the agents) in terms of the social welfare function.

Another network identification approach is based on information about the best-response dynamics of players and the solution of a control problem [137]. The Principal does not know the network structure in the game but observes the players' best responses and manipulates the actions of some players. As demonstrated by the authors, the network structure identifiability criterion is equivalent to the rank controllability criterion of the system reduced to a special form; and they applied the algorithm [138], originally developed for identifying stable linear systems.

### *3.2. Control in the Linear Quadratic Game on a Network*

Perhaps the linear quadratic game on a network is the first and most thoroughly studied game from the perspective of control problems; for this game, there are many statements of control problems, see [39, 118] and the references above. Several problems will be considered below to demonstrate the features and specifics of control problems in linear best-response games. The

[11] Bilevel optimization problems stem from Stackelberg games, where the outer (or inner) problem optimizes the Leader's (or Follower's) actions [135].

payoff of players in this model has the form

$$v_i = x_i \left( b_i + \beta \sum_{j=1}^{n} g_{ij} x_j \right) - \frac{x_i^2}{2}, \tag{42}$$

and the equilibrium in the game is

$$x^* = (I - \beta G)^{-1} b.$$

Historically, the key player problem was the first. A key player is a player whose removal has the greatest impact on the aggregate outcome. Let $G$ be an original (symmetric) adjacency matrix and $G^{-i}$ be the new matrix obtained from $G$ by zeroing out the row and column corresponding to the $i$th vertex. Then the Principal's problem takes the form $\max_{i \in N} \left[ \sum_j x_j^*(G) - \sum_j x_j^*(G^{-i}) \right]$ or equivalently,

$$\min_{i \in N} \sum_{j=1}^{n} x_j^*(G^{-i}). \tag{43}$$

Let $M = \{m_{ij}\} = (I - \beta G)^{-1}$, and $k = \{k_i\} = (I - \beta G)^{-1} 1$. The authors introduced a special intercentrality measure to capture both the individual centrality of players and their contribution to the centrality of others:

$$c_i = \frac{k_i^2}{m_{ii}}.$$

As shown, the vertex with the largest value of $c_i$ solves problem (43).

The incentive-targeting problem was considered in [73]: the Principal seeks to maximize the social welfare function in equilibrium:

$$W = \sum_{i=1}^{n} v_i = \frac{1}{2} x^T x \longrightarrow \max_b. \tag{44}$$

Thus, the Principal changes the standalone marginal return $\widehat{b} \to b$ of players, which is interpreted as a variation in players' incentives (e.g., monetary subsidies to firms). The Principal's budget constraint has the form

$$K(b, \widehat{b}) = \sum_{i=1}^{n} (b_i - \widehat{b}_i)^2 \leqslant C \sim n. \tag{45}$$

The solution of (44) is the efficient network heuristic

$$\widehat{b}_{nh} = b + \sqrt{C} \mu_{1(n)}, \tag{46}$$

where $\mu_{1(n)}$ denotes the eigenvector of the matrix $G$ corresponding to the largest (smallest) eigenvalue in the case of positive (negative, respectively) values of the coefficient $\beta$.

As indicated above, the choice of the Principal's objective function significantly affects the outcome. If, in this problem, we pass from maximization of the total payoff (44) to

$$W = \sum_{i=1}^{n} x_i \longrightarrow \max_b$$

(maximization of the aggregate outcome of participants), the solution will be the budget allocation proportional to the individual contribution of players:

$$\widehat{b}_{nh} = b + \sqrt{C} x^*.$$

The distinction between the functions is that the sum of payoffs accounts for both the actions and their differences: $\sum_i u_i = \frac{1}{n}\sum_i x_i^2 = \hat{x}^2 + \sigma^2$, where $\hat{x}$ and $\sigma^2$ are the mean and variance, respectively.

Due to several properties of the cost function (45), the authors extended the results to some other functions [139]. One notable case is the Principal's linear costs:

$$K(b, \widehat{b}) = \sum_{i=1}^{n} |b_i - \widehat{b}_i| \leqslant C \sim n.$$

Here, the solution is to transfer the entire budget to a single player $i^*$, $\widehat{b}_{i^*} = b_{i^*} + C$, chosen based on the players' contributions to the outcome of the Principal's intervention.

It is of interest to compare the solutions of the incentive problem for local aggregation and local averaging models. To achieve the social optimum without budget constraints in the linear quadratic game

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} \left[ (b_i + s_i)x_i + \beta \sum_{j=1}^{n} g_{ij} x_i x_j - \frac{x_i^2}{2} \right] \longrightarrow \max_s, \tag{47}$$

the Principal needs to choose $s_i^O = \beta \sum_j g_{ij} x_j$. Due to the properties of the game, this value is always positive, and the Principal will subsidize more central players in the network.

As shown above, in the social norms model, agents create both positive and negative externalities for their neighbors. As a result, in an incentive problem [46] of the form

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} \left[ (b_i + s_i)x_i - \frac{\theta}{2} \left( x_i - \sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j \right)^2 - \frac{x_i^2}{2} \right] \longrightarrow \max_s, \tag{48}$$

the solution is the vector $s^O$ composed of

$$s_i^O = \beta \sum_{j=1}^{n} \frac{g_{ij}}{d_i} \left( x_j - \sum_{k=1}^{n} \frac{g_{jk}}{d_j} x_k \right): \tag{49}$$

the Principal subsidizes (taxes) the agents whose neighbors exert efforts above (below, respectively) their social norms. (This contrasts with the linear quadratic game, where the Principal is forced to incentivize agents because of their positive externalities.) In other words, the Principal needs to subsidize agents exerting efforts below the average of their neighbors and to tax those exerting efforts above their neighbors.

We emphasize that till recently, changing the initial vector $b$ of individual characteristics has been supposed to be a problem with a nonzero budget; in real applications, however, the Principal redistributes resources among agents without additional costs. The balanced budget condition (Pigouvian taxes) was addressed in [20].

The problem of finding the interaction structure maximizing the Principal's objective function was considered in [63, 119]. As $\beta$ grows, the solutions of the two Principal's problems—maximizing the aggregate outcomes and the sum of payoffs—coincide, and among all networks, the maximum is attained for those with the largest spectral radius. The case of controlling a combination of multiple activities (when $x_i$ becomes a vector) was addressed in [114]. The cases of simultaneously using different control strategies were studied, e.g., in [140, 141].

## 4. CONCLUSIONS

This paper has presented a detailed review of control problems in the strategic interaction of agents on a network when the agent's dependence on the actions of others is described by the linear best response function. The Nash equilibrium in such games is not the social optimum due to externalities, which motivates the development of control mechanisms (e.g., identifying key agents, incentivizing agents, and creating or removing links) to increase the efficiency of collective interaction.

Some directions for further research are being actively developed at present and deserve special attention.

### *4.1. Nonlinear Best Response*

Control problems in nonlinear best-response games were considered in [142–144]. A general model of network effects was proposed in [34]. In this model, the nonlinear best response includes the local averaging model as a special case; the nonlinearity is implemented via a CES function with an elasticity parameter $\beta$ :

$$\overline{x} = \left(\sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j^{\beta}\right)^{\frac{1}{\beta}}; \tag{50}$$

and the player's best response takes the form

$$BR_i = (1-\lambda_2)b_i + (\lambda_1+\lambda_2)\left(\sum_{j=1}^{n} \frac{g_{ij}}{d_i} x_j^{\beta}\right)^{\frac{1}{\beta}}, \tag{51}$$

where $\lambda_1$ and $\lambda_2$ correspond to the intensity of the spillover effect and conformism, respectively. For $\beta = 1$, the LIM model arises; if $\beta$ is very large, the model $\overline{x}_{-i} = \max_j\{x_j\}$, where the agent is oriented toward the network neighbor with the greatest effort; and when $\beta$ takes large negative values, the dominating model $\overline{x}_{-i} = \min_j\{x_j\}$, where the agent is oriented toward the neighbor with the smallest effort.

### *4.2. Games on Multiplex Networks*

In many cases, the relationship among agents was considered within a single network. However, agents often maintain multiple types of relations, such as cooperation, mutual assistance, borrowing, etc. [145]. One way to describe such interaction is multiplex networks, i.e., multilayer networks without links between vertices of different layers that reflect the coexistence of various relations among agents [146]. Such games were analyzed in [147, 148].

### *4.3. Large Networks and Graphons*

Some networks encountered in applied research are so large that analysis becomes difficult or even impossible [149]. A tool for analyzing large networks and the asymptotic behavior of sequences of graphs with a growing number of vertices was proposed in [150]. A graph function (graphon) is an object that generalizes discrete graphs to the case of large networks in the continuous-limit representation. Centrality measures for such objects were introduced in [151]. In [127], the above games were extended to the case of a limit object: players were represented as a population on the interval $[0, 1]$, and the probability of a link between them was described by a graphon $W(i, j)$ defined on the unit square. Dynamics in such games were investigated in [152], and control problems were considered in [141, 153].

## REFERENCES


1. Asch, S.E., *Social Psychology*, Oxford University Press, 1987.
2. Merton, R.K., *Social Theory and Social Structure*, Free Press, 1968.

3. Katz, M.L. and Shapiro, C., Network Externalities, Competition, and Compatibility, *The American Economic Review*, 1985, vol. 75, no. 3, pp. 424–440.

4. Farrell, J. and Saloner, G., Standardization, Compatibility, and Innovation, *RAND J. Econom.*, 1985, vol. 16, no. 1, pp. 70–83.

5. Klausner, M., Corporations, Corporate Law, and Networks of Contracts, *Va. L. Rev.* 1995, vol. 81, pp. 7–57.

6. Lemley, M.A. and McGowan, D., Legal Implications of Network Economic Effects, *Calif. L. Rev.*, 1998, vol. 86, pp. 4–79.

7. Amelina, N.O., Anan'evskii, M.S., Andrievskii, B.R., et al., *Problemy setevogo upravleniya* (Challenges in Network Control), Izhevsk: Regulyarnaya i Khaoticheskaya Dinamika, 2015.

8. Proskurnikov, A.V. and Fradkov, A.L., Problems and Methods of Network Control, *Autom. Remote Control*, 2016, vol. 77, no. 10, pp. 1711–1740.

9. Kozyakin, V.S., Kuznetsov, N.A., and Chebotarev, P.Yu., Consensus in Asynchronous Multiagent Systems. I. Asynchronous Consensus Models, *Autom. Remote Control*, 2019, vol. 80, no. 4, pp. 593–623.

10. Kozyakin, V.S., Kuznetsov, N.A., and Chebotarev, P.Yu., Consensus in Asynchronous Multiagent Systems. II. Method of Joint Spectral Radius, *Autom. Remote Control*, 2019, vol. 80, no. 5, pp. 791–812.

11. Kozyakin, V.S., Kuznetsov, N.A., and Chebotarev, P.Yu., Consensus in Asynchronous Multiagent Systems. III. Constructive Stability and Stabilizability, *Autom. Remote Control*, 2019, vol. 80, no. 6, pp. 989–1015.

12. Gubanov, D.A., Novikov, D.A., and Chkhartishvili, A.G., *Social Networks: Models of Information Influence, Control and Confrontation*, Studies in Systems, Decision and Control, vol. 189, Cham: Springer, 2019.

13. Manski, C.F., Identification of Endogenous Social Effects: The Reflection Problem, *Rev. Econom. Stud.*, 1993, vol. 60, no. 3, pp. 531–542.

14. Poldin, O. and Yudkevich, M., Peer-Effects in Higher Education: A Review of Theoretical and Empirical Approaches, *Voprosy obrazovaniya/Educational Studies Moscow*, 2013, no. 4, pp. 106–123.

15. Dokuka, S., Valeeva, D., and Yudkevich, M., How Academic Achievement Spreads: The Role of Distinct Social Networks in Academic Performance Diffusion, *PLOS One*, 2020, vol. 15, no. 7, art. no. e0236737.

16. Bramoulle, Y., Djebbari, H., and Fortin, B., Identification of Peer Effects Through Social Networks, *Journal of Econometrics*, 2009, vol. 150, no. 1, pp. 41–55.

17. Liebowitz, S.J. and Margolis, S.E., Network Externality: An Uncommon Tragedy, *J. Econom. Perspect.*, 1994, vol. 8, no. 2, pp. 133–150.

18. Polterovich, V.M., Internet, Civic Culture and Evolution of Coordination Mechanisms, *Herald of the Central Economics and Mathematics Institute*, 2018, vol. 1, no. 1.

19. Konig, M.D., Liu, X., and Zenou, Y., R&D Networks: Theory, Empirics, and Policy Implications, *Review of Economics and Statistics*, 2019, vol. 101, no. 3, pp. 476–491.

20. Galeotti, A., Golub, B., Goyal, S., Talamás, E., and Tamuz, O., Taxes and Market Power: A Principal Components Approach, *arXiv:2112.08153*, 2021.

21. Zenou, Y., Peer vs. Network Effects: Microfoundations, Identification, and Beyond, CEPR Discussion paper no. 20814, 2025.

22. Moffitt, R.A., Policy Interventions, Low-Level Equilibria, and Social Interactions, in *Social Dynamics*, Durlauf, S.N. and Young, H.P., Eds., MIT Press, 2001, pp. 45–82.

23. Angrist, J.D., The Perils of Peer Effects, *Labour Economics*, 2014, vol. 30, pp. 98–108.

24. Bramoulle, Y., Djebbari, H., and Fortin, B., Peer Effects in Networks: A Survey, *Annual Review of Economics*, 2020, vol. 12, pp. 603–629.

25. Advani, A. and Malde, B., Methods to Identify Linear Network Models: A Review, *Swiss Journal of Economics and Statistics*, 2018, vol. 154, pp. 1–16.

26. Villeval, M.C., Performance Feedback and Peer Effects, in *Handbook of Labor, Human Resources and Population Economics*, Zimmermann, K., Ed., Cham: Springer, 2020, pp. 1–38.

27. Krupka, E.L. and Weber, R.A., Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?, *Journal of the European Economic Association*, 2013, vol. 11, no. 3, pp. 495–524.

28. Elliott, M.L., Goyal, S., and Teytelboym, A., Networks and Economic Policy, *Oxford Review of Economic Policy*, 2019, vol. 35, no. 4, pp. 565–585.

29. Bhadra, S., and Schweinberger, M., Causal Inference under Network Interference, *arXiv:2508.06808*, 2025.

30. Jackson, M.O., Rogers, B.W., and Zenou, Y., Networks: An Economic Perspective, *arXiv:1608.07901*, 2016.

31. Valente, T.W., Network Interventions, *Science*, 2012, vol. 337, no. 6090, pp. 49–53.

32. Sokolov, B., Causal Estimands for Policy Evaluation and Beyond, *SocArXiv*, 2025.

33. Sacerdote, B., Peer Effects with Random Assignment: Results for Dartmouth Roommates, *Quart. J. Econom.*, 2001, vol. 116, no. 2, pp. 681–704.

34. Boucher, V., Rendall, M., Ushchev, P., and Zenou, Y., Toward a General Theory of Peer Effects, *Econometrica*, 2024, vol. 92, no. 2, pp. 543–565.

35. Dimmock, S.G., Gerken, W.C., and Graham, N.P., Is Fraud Contagious? Coworker Influence on Misconduct by Financial Advisors, *The Journal of Finance*, 2018, vol. 73, no. 3, pp. 1417–1450.

36. Carrell, S.E., Malmstrom, F.V., and West, J.E., Peer Effects in Academic Cheating, *J. Human Res.*, 2008, vol. 43, no. 1, pp. 173–207.

37. Krekhovets, E.V. and Poldin, O.V., Students' Social Media: Formation Factors and Influence on Studies, *Voprosy Obrazovaniya/Educational Studies Moscow*, 2014, no. 4, pp. 127–144.

38. Calvo-Armengol, A. and Zenou, Y., Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior, *International Economic Review*, 2004, vol. 45, no. 3, pp. 939–958.

39. Ballester, C., Calvo-Armengol, A., and Zenou, Y., Who's Who in Networks. Wanted: The Key Player, *Econometrica*, 2006, vol. 74, no. 5, pp. 1403–1417.

40. Dimant, E., Contagion of Pro- and Anti-Social Behavior among Peers and the Role of Social Proximity, *Journal of Economic Psychology*, 2019, vol. 73, pp. 66–88.

41. Cai, J. and Szeidl, A., Interfirm Relationships and Business Performance, *Quart. J. Econom.*, 2018, vol. 133, no. 3, pp. 1229–1282.

42. Glaeser, E.L., Sacerdote, B.I., and Scheinkman, J.A., The Social Multiplier, *J. Eur. Econom. Associat.*, 2003, vol. 1, no. 2/3, pp. 345–353.

43. Kukushkin, N., Nash Equilibrium in Games with Additive Aggregation, *Economics and Mathematical Methods*, 2000, vol. 36, no. 4.

44. Jensen, M.K., Aggregative Games, in *Handbook of Game Theory and Industrial Organization*, vol. I, Edward Elgar Publishing, 2018, pp. 66–92.

45. Patacchini, E. and Zenou, Y., Juvenile Delinquency and Conformism, *The Journal of Law, Economics, & Organization*, 2012, vol. 28, no. 1, pp. 1–31.

46. Ushchev, P. and Zenou, Y., Social Norms in Networks, *Journal of Economic Theory*, 2020, vol. 185, art. no. 104969.

47. Zhilyakova, L.Yu. and Kuznetsov, O.P., *Teoriya resursnykh setei* (Theory of Resource Networks), Moscow: RIOR, 2017.

48. Bulow, J.I., Geanakoplos, J.D., and Klemperer, P.D., Multimarket Oligopoly: Strategic Substitutes and Complements, *Journal of Political Economy*, 1985, vol. 93, no. 3, pp. 488–511.

49. Vives, X. and Vravosinos, O., Strategic Complementarity in Games, *Journal of Mathematical Economics*, 2024, vol. 113, no. 2, art. no. 103005.

50. Novikov, D.A. and Chkhartishvili, A.G., Information Equilibrium: Punctual Structures of Information Distribution, *Autom. Remote Control*, 2003, vol. 64, no. 10, pp. 1609–1619.

51. Bramoulle, Y. and Kranton, R., Public Goods in Networks, *Journal of Economic Theory*, 2007, vol. 135, no. 1, pp. 478–494.

52. Yudkevich, M., Podkolzina, E., and Ryabinina, A., *Osnovy teorii kontraktov: Modeli i zadachi* (Foundations of Contract Theory: Models and Problems), Moscow: Higher School of Economics, 2002.
53. French Jr., J.R., A Formal Theory of Social Power, *Psychological Review*, 1956, vol. 63, no. 3, pp. 181–194.
54. DeGroot, M.H., Reaching a Consensus, *Journal of the American Statistical Association*, 1974, vol. 69, no. 345, pp. 118–121.
55. Golub, B. and Jackson, M.O., How Homophily Affects the Speed of Learning and Best-Response Dynamics, *The Quarterly Journal of Economics*, 2012, vol. 127, no. 3, pp. 1287–1338.
56. Friedkin, N.E. and Johnsen, E.C., Social Influence and Opinions, *Journal of Mathematical Sociology*, 1990, vol. 15, no. 3/4, pp. 193–206.
57. Ghaderi, J. and Srikant, R., Opinion Dynamics in Social Networks with Stubborn Agents: Equilibrium and Convergence Rate, *Automatica*, 2014, vol. 50, no. 12, pp. 3209–3215.
58. Golub, B. and Morris, S., Expectations, Networks, and Conventions, *arXiv:2009.13802*, 2020.
59. Morris, S. and Shin, H.S., Social Value of Public Information, *Amer. Econom. Rev.*, 2002, vol. 92, no. 5, pp. 1521–1534.
60. Bonacich, P., Power and Centrality: A Family of Measures, *Amer. J. Soc.*, 1987, vol. 92, no. 5, pp. 1170–1182.
61. Belhaj, M., Bramoulle, Y., and Deroian, F., Network Games under Strategic Complementarities, *Games and Economic Behavior*, 2014, vol. 88, pp. 310–319.
62. Helsley, R.W. and Zenou, Y., Social Networks and Interactions in Cities, *Journal of Economic Theory*, 2014, vol. 150, pp. 426–466.
63. Corbo, J., Calvo-Armengol, A., and Parkes, D.C., The Importance of Network Topology in Local Contribution Games, in *Lecture Notes in Computer Science*, Deng, X. and Graham, F.C., Eds., Berlin–Heidelberg: Springer, 2007, vol. 4858, pp. 388–395.
64. Konig, M.D., Tessone, C.J., and Zenou, Y., Nestedness in Networks: A Theoretical Model and Some Applications, *Theoretical Economics*, 2014, vol. 9, no. 3, pp. 695–752.
65. Singh, N. and Vives, X., Price and Quantity Competition in a Differentiated Duopoly, *Rand J. Econom.*, 1984, vol. 15, no. 4, pp. 546–554.
66. Geraskin, M.I., A Survey of the Latest Advances in Oligopoly Games, *Autom. Remote Control*, 2023, vol. 84, no. 6, pp. 637–653.
67. Korgin, N., Introduction to Theory of Control in Organizations for Kids via Interactive Games, *IFAC-PapersOnLine*, 2015, vol. 48, no. 29, pp. 289–294.
68. Bimpikis, K., Ehsani, S., and Ilkılıc, R., Cournot Competition in Networked Markets, *Management Science*, 2019, vol. 65, no. 6, pp. 2467–2481.
69. Goubko, M.V. and Novikov, D.A., *Teoriya igr v upravlenii organizatsionnymi sistemami* (Game Theory in Control of Organizational Systems), Moscow: SINTEG, 2002.
70. Fedyanin, D. and Chkhartishvili, A., On a Problem of Information Control in Social Networks, *Large-Scale Systems Control*, 2010, no. 31, pp. 265–275.
71. Novikov, D.A. and Chkhartishvili, A.G., *Reflexion and Control: Mathematical Models*, Boca Raton: CRC Press, 2014.
72. Petrov, I.V. and Chkhartishvili, A.G., The Incentive-Targeting Problem in a Reflexive Game with a Point-Type Awareness Structure, *Control Sciences*, 2024, no. 5, pp. 35–40.
73. Galeotti, A., Golub, B., and Goyal, S., Targeting Interventions in Networks, *Econometrica*, 2020, vol. 88, no. 6, pp. 2445–2471.
74. Monderer, D. and Shapley, L.S., Potential Games, *Games and Economic Behavior*, 1996, vol. 14, no. 1, pp. 124–143.
75. Sandholm, W.H., *Population Games and Evolutionary Dynamics*, MIT Press, 2010.
76. Parise, F. and Ozdaglar, A., A Variational Inequality Framework for Network Games: Existence, Uniqueness, Convergence and Sensitivity Analysis, *Games and Economic Behavior*, 2019, vol. 114, pp. 47–82.

77. Jackson, M.O. and Zenou, Y., Games on Networks, in *Handbook of Game Theory with Economic Applications*, vol. 4, Elsevier, 2015, pp. 95–163.

78. Bramoulle, Y., Kranton, R., and D'amours, M., Strategic Interaction and Networks, *Amer. Econom. Rev.*, 2014, vol. 104, no. 3, pp. 898–930.

79. Melo, E., A Variational Approach to Network Games, Working paper no. 2018.05, 2018.

80. Zenou, Y. and Zhou, J., Sign-Equivalent Transformations and Equilibrium Systems: Theory and Applications, *Unpublished manuscript, Monash University*, 2024.

81. Sharkey, K.J., A Control Analysis Perspective on Katz Centrality, *Scientific Reports*, 2017, vol. 7, no. 1, art. no. 17247.

82. Bramoulle, Y. and Kranton, R., Games Played on Networks, in *Oxford Handbook on the Economics of Networks*, Bramoulle, Y., Galeotti, A., and Rogers, B., Eds., 2016, Chapter 5.

83. Opoitsev, V.I., *Ravnovesie i ustoichivost' v modelyakh kollektivnogo povedeniya* (Equilibrium and Stability in Collective Behavior Models), Moscow: Nauka, 1977.

84. Bervoets, S. and Faure, M., Stability in Games with Continua of Equilibria, *Journal of Economic Theory*, 2019, vol. 179, pp. 131–162.

85. Bayer, P., Herings, P.J.-J., and Peeters, R., Farsighted Manipulation and Exploitation in Networks, *Journal of Economic Theory*, 2021, vol. 196, art. no. 105311.

86. Bayer, P., Kozics, G., and Sz'oke, N.G., Best-Response Dynamics in Directed Network Games, *Journal of Economic Theory*, 2023, vol. 213, art. no. 105720.

87. Al'pin, Y.A., Harary's Theorem on Signed Graphs and Reversibility of Markov Chains, *J. Math. Sci.*, 2014, vol. 199, pp. 375–380.

88. Al'pin, Yu.A. and Bashkin, I.V., Nonnegative Chainable Matrices and Kolmogorov's condition, *Zap. Nauchn. Sem. POMI*, 2021, vol. 504, pp. 5–20.

89. Acemoglu, D., Carvalho, V.M., Ozdaglar, A., and Tahbaz-Salehi, A., The Network Origins of Aggregate Fluctuations, *Econometrica*, 2012, vol. 80, no. 5, pp. 1977–2016.

90. Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A., *Networks, Shocks, and Systemic Risk: Technical Report*, National Bureau of Economic Research, 2015.

91. Dyukalov, A.N. and Ilyutovich, A.E., Asymptotic Properties of Optimal Trajectories in Economic Dynamics, *Autom. Remote Control*, 1973, vol. 34, no. 3, pp. 423–434.

92. Dyukalov, A.N., Ivanov, Yu.N., and Tokarev, V.V., Control Theory and Economic Systems. I. Problem of Description, *Autom. Remote Control*, 1974, vol. 35, no. 5, pp. 797–810.

93. Dubovskii, S.V. and Uzdemir, A.P., Optimality Criteria and Variational Approaches in Dynamic Models of an Economy, *Autom. Remote Control*, 1974, vol. 35, no. 6, pp. 951–958.

94. Dubovskii, S.V., Dyukalov, A.N., Ivanov, Yu.N., et al., On the Design of Optimal Economic Plan, *Autom. Remote Control*, 1972, vol. 33, no. 8, pp. 1336–1349.

95. Novikov, D.A., Games and Networks, *Autom. Remote Control*, 2014, vol. 75, no. 6, pp. 1145–1154.

96. Goyal, S., *Connections: An Introduction to the Economics of Networks*, Princeton University Press, 2012.

97. Gubko, M.V., Control of Organizational Systems with Network Interaction of Agents. I. A Review of Network Game Theory, *Autom. Remote Control*, 2004, vol. 65, no. 8, pp. 1276–1291.

98. Gubko, M.V., Control of Organizational Systems with Network Interaction of Agents. II. Stimulation Problems, *Autom. Remote Control*, 2004, vol. 65, no. 9, pp. 1470–1485.

99. Mazalov, V.V. and Chirkova, J.V., *Networking Games: Network Forming Games and Games on Networks*, Elsevier Science, 2019.

100. Naghizadeh, P. and Liu, M., Provision of Public Goods on Networks: On Existence, Uniqueness, and Centralities, *IEEE Transactions on Network Science and Engineering*, 2017, vol. 5, no. 3, pp. 225–236.

101. Elliott, M. and Golub, B., A Network Centrality Approach to Coalitional Stability, 2012. http://www.mit.edu/bgolub/papers/centrality.pdf

102. Yan, C., Cooperative Solutions for Network Games with Quadratic Utilities, *Contributions to Game Theory and Management*, 2023, vol. 16, pp. 282–294.

103. Dorofeeva, Yu.A., Game-Theoretic Models of Opinion Dynamics, *Extended Abstract of Cand. Sci. (Phys.–Math.) Dissertation*, Petrozavodsk State University, Petrozavodsk, 2021.

104. Kozitsin, I.V., Optimal Control in Opinion Dynamics Models: Diversity of Influence Mechanisms and Complex Influence Hierarchies, *Chaos, Solitons & Fractals*, 2024, vol. 181, art no. 114728.

105. Jackson, M.O., Inequality's Economic and Social Roots: The Role of Social Networks and Homophily, *SSRN Scholarly paper no. 3795626*, 2021.

106. Siciliano, M.D. and Whetsell, T.A., Strategies of Network Intervention: A Pragmatic Approach to Policy Implementation and Public Problem Resolution Through Network Science, *arXiv:2109.08197*, 2021.

107. Robins, G., Lusher, D., Broccatelli, C., et al., Multilevel Network Interventions: Goals, Actions, and Outcomes, *Social Networks*, 2023, vol. 72, pp. 108–120.

108. Kempe, D., Kleinberg, J., and Tardos, E., Maximizing the Spread of Influence through a Social Network, *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, pp. 137–146.

109. Galeotti, A. and Goyal, S., Influencing the Influencers: A Theory of Strategic Diffusion, *The RAND Journal of Economics*, 2009, vol. 40, no. 3, pp. 509–532.

110. Candogan, O., Bimpikis, K., and Ozdaglar, A., Optimal Pricing in Networks with Externalities, *Operations Research*, 2012, vol. 60, no. 4, pp. 883–905.

111. Bloch, F. and Quérou, N., Pricing in Social Networks, *Games and Economic Behavior*, 2013, vol. 80, pp. 243–261.

112. Fainmesser, I.P. and Galeotti, A., Pricing Network Effects, *The Review of Economic Studies*, 2016, vol. 83, no. 1, pp. 165–198.

113. Demange, G., Optimal Targeting Strategies in a Network under Complementarities, *Games and Economic Behavior*, 2017, vol. 105, pp. 84–103.

114. Kor, R. and Zhou, J., Multi-Activity Influence and Intervention, *Games and Economic Behavior*, 2023, vol. 137, pp. 91–115.

115. Jeong, D. and Shin, E., Optimal Influence Design in Networks, *Journal of Economic Theory*, 2024, vol. 220, art. no. 105877.

116. Belhaj, M., Deroıan, F., and Safi, S., Targeting in Networks under Costly Agreements, *Games and Economic Behavior*, 2023, vol. 140, pp. 154–172.

117. Dasaratha, K., Golub, B., and Shah, A., Incentive Design with Spillovers, *SSRN Scholarly paper no. 4853054*, 2024.

118. Borgatti, S.P., Identifying Sets of Key Players in a Social Network, *Computational & Mathematical Organization Theory*, 2006, vol. 12, pp. 21–34.

119. Corbo, J., Calvo-Armengol, A., and Parkes, D., A Study of Nash Equilibrium in Contribution Games for Peer-to-Peer Networks, *ACM SIGOPS Operating Systems Review*, 2006, vol. 40, no. 3, pp. 61–66.

120. Golub, B. and Lever, C., The Leverage of Weak Ties: How Linking Groups Affects Inequality, Working paper, 2010. *http://www.stanford.edu/?bgolub/papers/intergroup.pdf*

121. Belhaj, M., Bervoets, S., and Deroıan, F., Efficient Networks in Games with Local Complementarities, *Theoretical Economics*, 2016, vol. 11, no. 1, pp. 357–380.

122. Matveenko, V.D. and Korolev, A.V., Equilibrium in a Network Game with Production and Knowledge Externalities, *Autom. Remote Control*, 2018, vol. 79, no. 7, pp. 1342–1360.

123. Hiller, T., Peer Effects in Endogenous Networks, *Games and Economic Behavior*, 2017, vol. 105, pp. 349–367.

124. Harkins, A., Network Comparative Statics, *SSRN Scholarly paper no. 3847211*, 2020.

125. Li, X., Designing Weighted and Directed Networks under Complementarities, *Games and Economic Behavior*, 2023, vol. 140, pp. 556–574.

126. Sun, Y., Zhao, W., and Zhou, J., Structural Interventions in Networks, *Int. Econom. Rev.*, 2023, vol. 64, no. 4, pp. 1533–1563.

127. Parise, F. and Ozdaglar, A., Graphon Games: A Statistical Framework for Network Games and Interventions, *Econometrica*, 2023, vol. 91, no. 1, pp. 191–225.

128. Galeotti, A., Golub, B., Goyal, S., et al., Robust Market Interventions, *arXiv:2411.03026*, 2024.

129. Breza, E., Chandrasekhar, A.G., McCormick, T.H., et al., Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data, *American Economic Review*, 2020, vol. 110, no. 8, pp. 2454–2484.

130. Viviano, D., Policy Targeting under Network Interference, *Review of Economic Studies*, 2025, vol. 92, no. 2, pp. 1257–1292.

131. Bloch, F. and Shabayek, S., Targeting in Social Networks with Anonymized Information, *Games and Economic Behavior*, 2023, vol. 141, pp. 380–402.

132. Lake, B. and Tenenbaum, J., Discovering Structure by Learning Sparse Graphs, *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, 2010, vol. 32, pp. 778–784.

133. Leng, Y., Dong, X., Wu, J., and Pentland, A., Learning Quadratic Games on Networks, *Proceedings of the International Conference on Machine Learning (PMLR)*, 2020, pp. 5820–5830.

134. Rossi, E., Monti, F., Leng, Y., et al., Learning to Infer Structures of Network Games, *Proceedings of the International Conference on Machine Learning (PMLR)*, 2022, pp. 18809–18827.

135. Hong, M., Wai, H.-T., Wang, Z., and Yang, Z., A Two-Timescale Stochastic Algorithm Framework for Bilevel Optimization: Complexity Analysis and Application to Actor-Critic, *SIAM Journal on Optimization*, 2023, vol. 33, no. 1, pp. 147–180.

136. Zhang, C., Liu, S., Wai, H.-T., et al., Network Games Induced Prior for Graph Topology Learning, *arXiv:2410.24095*, 2024.

137. Ding, K., Chen, Y., Wang, L., Ren, X., and Shi, G., Network Learning in Quadratic Games from Best-Response Dynamics, *IEEE/ACM Transactions on Networking*, 2024, vol. 32, no. 5, pp. 3669–3684

138. Jongeneel, W., Sutter, T., and Kuhn, D., Efficient Learning of a Linear Dynamical System with Stability Guarantees, *IEEE Transactions on Automatic Control*, 2023, vol. 68, no. 5, pp. 2790–2804.

139. Galeotti, A., Golub, B., and Goyal, S., Supplement to "Targeting Interventions in Networks", *Econometrica*, 2020, vol. 88, no. 6, pp. 2445–2471.

140. Kor, R. and Zhou, J., Welfare and Distributional Effects of Joint Intervention in Networks, *arXiv:2206.03863*, 2022.

141. Petrov, I., Structural Interventions in Linear Best-Response Games on Random Graphs, *IFAC-PapersOnLine*, 2023, vol. 56, no. 2, pp. 2830–2833.

142. Belhaj, M. and Deroian, F., Competing Activities in Social Networks, *The BE Journal of Economic Analysis & Policy*, 2014, vol. 14, no. 4, pp. 1431–1466.

143. Allouch, N., On the Private Provision of Public Goods on Networks, *Journal of Economic Theory*, 2015, vol. 157, pp. 527–552.

144. Cai, J., Zhang, C., and Wai, H.-T., Optimal Pricing for Linear-Quadratic Games with Nonlinear Interaction Between Agents, *IEEE Control Systems Letters*, 2024, vol. 8, pp. 1631–1636.

145. Chandrasekhar, A.G., Chaudhary, V., Golub, B., and Jackson, M.O., Multiplexing in Networks and Diffusion, *arXiv:2412.11957*, 2024.

146. Cheng, C., Huang, W., and Xing, Y., A Theory of Multiplexity: Sustaining Cooperation with Multiple Relations, *SSRN Scholarly paper no. 3811181*, 2021.

147. Ebrahimi, R. and Naghizadeh, P., United We Fall: On the Nash Equilibria of Multiplex and Multilayer Network Games, *arXiv:2402.06108*, 2024.

148. Zenou, Y. and Zhou, J., Games on Multiplex Networks, *American Economic Review*, 2026, vol. 116, no. 4, pp. 1415–1458.

149. D'Souza, R.M., di Bernardo, M., and Liu, Y.-Y., Controlling Complex Networks with Complex Nodes, *Nature Reviews Physics*, 2023, vol. 5, no. 4, pp. 250–262.

150. Lovász, L., *Large Networks and Graph Limits*, American Mathematical Society, 2012.

151. Avella-Medina, M., Parise, F., Schaub, M.T., and Segarra, S., Centrality Measures for Graphons: Accounting for Uncertainty in Networks, *IEEE Transactions on Network Science and Engineering*, 2020, vol. 7, no. 1, pp. 520–537.

152. Al Taha, F., Rokade, K., and Parise, F., Gradient Dynamics in Linear Quadratic Network Games with Time-Varying Connectivity and Population Fluctuation, *Proceedings of the 62nd IEEE Conference on Decision and Control (CDC)*, IEEE, 2023, pp. 1991–1996.

153. Parise, F. and Ozdaglar, A., Analysis and Interventions in Large Network Games, *Annual Review of Control, Robotics, and Autonomous Systems*, 2021, vol. 4, no. 1, pp. 455–486.

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