# PARAMETER IDENTIFICATION OF THE LINEAR DISCRETE-TIME STOCHASTIC SYSTEMS WITH UNKNOWN EXOGENOUS INPUTS

**Yulia Tsyganova**
Department of Mathematics,
Information and Aviation Technology
Ulyanovsk State University
Russian Federation
tsyganovayuv@ulsu.ru

**Andrey Tsyganov**
Department of Mathematics,
Physics and Technology Education
Ulyanovsk State University of Education
Russian Federation
tsyganov.av@ulspu.ru

Article history:
Received 10.11.2023, Accepted 26.11.2023

**Abstract**

The paper addresses a parameter identification problem for linear discrete-time stochastic systems with unknown exogenous inputs. Such systems are considered when solving practical problems related to the measurements processing in the case when it is impossible to do any assumptions about the evolution of unknown input signal or its statistical characteristics that can change over time. We consider a class of discrete-time linear stochastic systems with unknown exogenous inputs, where an additional source of a priori uncertainty of the system model is introduced, namely, the unknown parameter, on the elements of which the system model matrices can depend. This formulation of the parameter identification problem under the conditions of unknown inputs and the presence of random noises describes a high degree of uncertainty of a discrete-time linear stochastic system. We propose a novel solution to this problem based on the construction of a new instrumental identification criterion. Minimization of this criterion allows for evaluating the unknown system model parameters simultaneously with the estimating of the state vector and unknown exogenous inputs of the system. Numerical experiments confirm the validity and efficiency of the proposed parameter identification method.

**Key words**

discrete-time stochastic system, parameter identification, unknown exogenous input, instrumental identification criterion, sensitivity model, gradient-based optimization, metaheuristic optimization

## 1 Introduction

Methods for simultaneous input signal and state vector estimation for linear stochastic systems have attracted a lot of attention in recent decades due to their practical applications in modern research areas such as faults detection and isolation, estimating geophysical processes in case when it is impossible to do any assumptions about the evolution of unknown input signal or its statistical characteristics that can change over time.

For continuous-time systems necessary and sufficient conditions for the existence of an optimal system state estimate are known from [Kudva et al., 1980; Hou and Müller, 1992; Darouach et al., 1994]. Some algorithms for recovering unknown inputs were investigated in [Hou and Patton, 1998; Xiong and Saif, 2003]. In early papers [Darouach et al., 1994], construction of the observer and conditions for existence and stability were obtained for the class of deterministic systems. Also, several approaches have been developed for stochastic systems. The most common early approach is to treat the unknown input as a random process with a known characteristics (for example, the mean and covariance are known) or as a constant bias. This approach was introduced by Friedland [Friedland, 1969] and further discussed in [Ignagni, 1990].

For the class of discrete-time systems, the earliest approaches were based on the inclusion of an unknown input into the state vector of the stochastic system. It was assumed that the evolution model of the input signal is known. In this case, the well-known extended Kalman filter was used to solve the problem. To reduce the computational cost of the extended filter, Friedland [Friedland, 1969] proposed a two-stage Kalman filter in which the state estimate and the unknown input are separated. Although both methods are successfully used in many applications, they require knowledge of a evolution model describing unknown inputs.

Another approach to solving the problem was proposed by Kitanidis [Kitanidis, 1987], who developed an optimal recurrent filter for estimating the state vector, which is based on the assumption that a priori information about unknown inputs is not available. His result was extended in [Darouach and Zasadzinski, 1997] where the authors established stability and convergence conditions, and also developed a new method for filtering the system state vector. Further, Hsieh [Hsieh, 2000] established a connection between the two-stage Friedland filter and the Kitanidis filter, showing that the Kitanidis result can be obtained by making the two-stage filter independent of the basic input model. In addition, his method makes it possible to compute estimates of the unknown input. However, the optimality of this estimate has not been proven.

S. Gillijns and B. De Moor [Gillijns and De Moor, 2007a] made a great contribution to the development of the theory for discrete-time filtering in stochastic systems with unknown inputs. They extended the results obtained in [Kitanidis, 1987; Darouach and Zasadzinski, 1997] and proposed a recurrent filtering algorithm for the simultaneous estimation of the system state and unknown input vectors. Moreover, the obtained estimates have the minimum error variance. They also proved the optimality of computed estimates.

The next step was to solve the discrete-time filtering problem with unknown inputs included both in the state and sensor equations. Such a problem arises when the model contains systematic measurement errors and the uncertainty of the model itself, caused by unknown disturbances and/or due to the unmodeled dynamics of the process under investigation. A rigorous and simple method for estimating the state vector in the presence of unknown inputs was developed in [Hou and Müller, 1994; Hou and Patton, 1998]. The proposed approach was to first construct an equivalent system that is decoupled from the unknown inputs and then compute an unbiased estimate with a minimum error variance for that equivalent system.

Another approach is to parameterize the filter equations and then calculate the optimal parameter estimates by minimizing the trace of the covariance matrix of the estimation error. An optimal filter of this type was first developed by Kitanidis [Kitanidis, 1987]. His solution is limited to linear systems without direct transfer of unknown input to the system output. In addition, this solution does not allow for obtaining an estimate of the unknown input signal. Gillijns and De Moor [Gillijns and De Moor, 2007b] proposed a solution to the problem of joint estimation of the unknown input vector and the state vector for linear systems with direct feedthrough. Using linear unbiased minimum error variance estimator, they developed a three-stage recurrent filter in which the estimate of the system state vector and the unknown input signal are correlated. The input vector estimation is based on the least squares method developed by Gillijns and De Moor, while the state vector estimation problem is solved using the method developed by Kitanidis. In [Yong et al., 2013] authors presented a variation of an optimal filter that simultaneously estimates the states and unknown inputs providing the best linear unbiased estimate (BLUE) of the unknown input for linear discrete-time stochastic systems with direct feedthrough. They argued that in contrast to previous filters the information about the unknown input can be obtained from the current time step as well as the previous one, making it possible to estimate the unknown in different ways. A solution to the problem of simultaneous estimation of an unknown input signal and state vector for the linear system with a rank-deficient distribution matrix was proposed in [Hua et al., 2021]. The problem of unknown input estimation for systems which do not satisfy the matching, minimum phase, and detectability conditions was studied in [Zhirabok et al., 2023]. The suggested method is based on the reduced order model of the original system insensitive to the disturbance.

In recent years, research on estimating the state vector of a system with an unknown input has also focused on nonlinear systems. Based on the extended Kalman filter structure, some new filters have been developed that estimate the state of a nonlinear system [Xiao et al., 2018; Varshney et al., 2019].

In all above mentioned methods, it was assumed that all the system matrices are perfectly known for the optimal filter design. Thus, they do not solve the parameter identification problem. To the best of our knowledge, there is a small number of existing works on identification of stochastic systems with unknown exogenous inputs (see [Lan et al., 2013; Yu and Chakravorty, 2016; Kong et al., 2023], for example).

In this paper, we consider a class of discrete-time linear stochastic systems with unknown exogenous inputs, in which an additional source of a priori uncertainty of the system model is introduced, namely, the unknown system model parameter $\theta \in \mathbb{R}^p$, on the elements of which the system matrices can depend. This formulation of the parameter identification problem, under the condition of unknown inputs and the presence of random noises, describes a high degree of uncertainty of a discrete-time linear stochastic system. We propose a novel solution to the problem of parameter identification based on the construction of the new instrumental identification criterion. Minimization of this criterion allows for evaluating the unknown system model parameter simultaneously with the estimating of the state vector and unknown exogenous input.

It should be noticed that the problem considered in this paper is also related to the unknown input observer design problem which is well-known and has a long history in research and applications [Tranninger et al., 2023]. However, its significant differences from the parameter identification problem, which we solve in this paper, are as follows. Firstly, we consider a discrete-time linear time invariant (LTI) stochastic system in which the system matrices depend on an unknown parameter. In almost all papers devoted to unknown input observer design, the system matrices are assumed to be completely known. Thus, the problem of parameter identification is not solved there. Secondly, we assume that available measurements of the system state vector contain additive noises, whereas the observer design problem with unknown input considers unnoised measurements as a rule.

The results obtained in our paper can be applied to solving various practical problems related to the parameter identification of mathematical models represented by the discrete-time stochastic LTI systems in state space, based on the output-only measurements data. Examples of such applications include, but are not limited to, the tracking problem of a remotely piloted aircraft [He and Liu, 2023], the problem of the position and velocity tracking of multiple vehicles [Yong et al., 2016], the fault identification and state estimation problem when the system dynamics is afflicted by disturbances or faults [Yong et al., 2014], the heat conduction problem [Gillijns and De Moor, 2007c], the estimation, detection and control problems of the DC and step motor systems [Darouach et al., 2003; Yong et al., 2013; Furtat et al., 2023], parameter estimation problems of discrete-time models of the Van der Pol oscillator [Palanthandalam-Madapusi and Bernstein, 2007], the linearized model of a simplified longitudinal flight control system [Hmida et al., 2010], the spring-mass-damper system [Teymouri et al., 2020].

The remainder of the paper has the following structure. Section 2 provides basic definitions associated with the optimal filtering algorithm of simultaneous input and state estimation for linear discrete-time stochastic systems with unknown inputs. Also, the problem of parameter identification is stated and described. Section 3 contains the main results of the paper — the new instrumental identification criterion (IIC) and the system state sensitivity equations for evaluation of the IIC gradient.

Section 4 demonstrates how the newly proposed instrumental identification criterion and method for calculation the IIC gradient can be applied for solving the parameter identification problem of the considered stochastic system model with unknown periodic input signal. The comparison results of four different methods for minimizing the proposed identification criterion are presented. Section 5 concludes the paper.

## 2 Preliminaries and Problem Statement

Consider the discrete-time linear time invariant (LTI) stochastic system

$$\begin{cases} x_k = Fx_{k-1} + Bu_{k-1} + Gw_k, \\ z_k = Hx_k + v_k, \quad k = 1, 2, \dots, N \end{cases} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the system state vector; $u_k \in \mathbb{R}^r$ is an unknown exogenous input; $z_k \in \mathbb{R}^m$ is the measurements vector; matrices $F \in \mathbb{R}^{n\times n}$; $B \in \mathbb{R}^{n\times r}$; $G \in \mathbb{R}^{n\times q}$; $H \in \mathbb{R}^{m\times n}$; $N$ is the number of measurements; initial state $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$, additive model disturbance $w_k \in \mathbb{R}^q \sim \mathcal{N}(0, Q)$ and measurement noise $v_k \in \mathbb{R}^m \sim \mathcal{N}(0, R)$ are mutually independent. Covariance matrices $Q$ and $R$ of $w_k$ and $v_k$ are positive semidefinite.

In this paper, we consider system (1) in the case when exogenous input $u_k$ is completely unknown, i. e., there is no prior knowledge about the dynamics or statistic characteristics of $u_k$.

It is worth to note that in case input signal $u_k$ is known, is zero or is a zero-mean white random vector with known covariance matrix, the optimal discrete-time filtering problem for the system (1) reduces to the Kalman filtering problem [Grewal and Andrews, 2015]. If $u_k$ is the deterministic input and its dynamics is known, then suboptimal estimates of $u_k$ and $x_k$ can be obtained using a well-known extended Kalman filter [Anderson and Moore, 1979].

The problem of simultaneous input and state estimation for linear stochastic systems with completely unknown exogenous input was solved in [Kitanidis, 1987; Gillijns and De Moor, 2007a]. The derivation of corresponding optimal filters was based on unbiased minimum-variance estimation. In [Gillijns et al., 2007], this problem was investigated from the viewpoint of recursive least-squares estimation.

Throughout the paper, we assume that $\operatorname{rank} H = n$ and the following sufficient condition for the existence of an unbiased state estimator is satisfied [Kitanidis, 1987; Darouach and Zasadzinski, 1997]

$$\operatorname{rank} HB = \operatorname{rank} B = r. \tag{2}$$

Assumption (2) implies $n \geq r$ and $m \geq r$.

### 2.1 The Algorithm of Simultaneous Input and State Estimation for Linear Discrete-Time Stochastic Systems with Unknown Exogenous Inputs

The optimal filtering algorithm of simultaneous input and state estimation for linear discrete-time stochastic systems consists of three sequential steps repeated for each measurement $z_k$ ($k = 1, \dots, N$):

1) time update of the state vector estimate;
2) input signal estimation;
3) measurement update of the state vector estimate.

In [Gillijns and De Moor, 2007a], the authors suggested two variants of the filtering algorithm, in which steps 1 and 2 are the same, but step 3 is different. The first variant of the algorithm allows for computing an MVU estimate (MVU — minimum-variance unbiased) of the input vector $\hat{u}_{k-1}$ at step 2 and an unbiased estimate of the state vector $\hat{x}_k$ at step 3. In the second version of the algorithm at step 3, due to more complex calculations, an MVU estimate $\hat{x}_k$ of the state vector $x_k$ is obtained.

Let $I_n$ denote the identity matrix of size $n$. Then the filtering algorithm for the system (1) can be written as follows.

---

**Algorithm 1.** Gillijns and De Moor algorithm (GDM).
**Initialization.** $P_0 = \Pi_0$, $\hat{x}_0 = \bar{x}_0$.
**For** $k = 1, 2, \ldots, N$ **do**

**I. Time Update step.** Find a priori estimation error covariance matrix $P_{k|k-1}$ and a priori estimate of the state vector $\hat{x}_{k|k-1}$ as follows:

$$\hat{x}_{k|k-1} = F\hat{x}_{k-1}, \tag{3}$$
$$P_{k|k-1} = FP_{k-1}F^T + GQG^T. \tag{4}$$

**II. Input Estimation step.** Find an unknown input estimate $\hat{u}_{k-1}$ as follows:

$$\tilde{R}_k = HP_{k|k-1}H^T + R, \tag{5}$$
$$D_{k-1} = \left(B^T H^T \tilde{R}_k^{-1} HB\right)^{-1}, \tag{6}$$
$$M_k = D_{k-1}B^T H^T \tilde{R}_k^{-1} = (HB)^+, \tag{7}$$
$$\hat{u}_{k-1} = M_k(z_k - H\hat{x}_{k|k-1}). \tag{8}$$

**III. Measurement Update step.** Using the a priori estimates $P_{k|k-1}$, $\hat{x}_{k|k-1}$ and input estimate $\hat{u}_{k-1}$, find a posteriori values $P_k$ and $\hat{x}_k$ as follows:

$$K_k = P_{k|k-1}H^T\tilde{R}_k^{-1}, \tag{9}$$
$$\hat{x}_k^* = \hat{x}_{k|k-1} + B\hat{u}_{k-1}, \tag{10}$$
$$P_k^* = (I_n - K_kH)P_{k|k-1}, \tag{11}$$
$$\hat{x}_k = \hat{x}_k^* + K_k(z_k - H\hat{x}_k^*), \tag{12}$$
$$P_k = P_k^* + (I_n - K_kH)BD_{k-1}B^T \times (I_n - K_kH)^T. \tag{13}$$

**End**

---

### 2.2 Parameter Identification Problem

Now suppose that the matrices $F$, $G$, $Q$, $R$ defining system model (1) are known up to the value of parameter $\theta \in \mathbb{R}^p$, i. e.,

$$\begin{cases} x_k = F(\theta)x_{k-1} + Bu_{k-1} + G(\theta)w_k, \\ z_k = Hx_k + v_k, \quad k = 1, 2, \ldots, N \end{cases} \tag{14}$$

Let us state the parameter identification problem of system model (1) by available measurements $Z_1^N = \{z_1, \ldots, z_k, \ldots, z_N\}$.

Consider a process of discrete-time filtering providing by Algorithm 1. The estimation error $e_k(\theta) = x_k - \hat{x}_k(\theta)$ will depend on the value of parameter $\theta$, which is specified in the filtering algorithm equations. The minimum value of the error $e_k(\theta)$ can be obtained under the condition of a minimum by $\theta$ of the original identification criterion (OIC) in the form of quadratic functional

$$\mathcal{J}_e(\theta) = \mathbf{E}\left\{e_k^T(\theta)e_k(\theta)\right\}. \tag{15}$$

The problem is that functional (15) is not suitable for solving the parameter identification problem due to this functional is not instrumental, i.e., it is not practically feasible because the estimation errors, $e_k(\theta)$, are not available for direct observation. The most popular approach to solving this problem are MPE (minimum prediction error) methods [Astrom, 1980; Ljung, 1999] based on minimizing an identification criterion that depends on the observed measurement residuals. Such criteria include the well-known least squares and maximum likelihood criteria. Thus, the algorithm of numerical minimization of the original functional (15) by the parameter $\theta$ is replaced by the algorithm of numerical minimization of the selected instrumental criterion, which is practically feasible.

In this paper, for solving the parameter identification problem, we use an alternative approach that is the Active Principle of Adaptation (APA) [Semushin, 2011a; Semushin, 2011b; Semushin and Tsyganova, 2013; Semushin, 2014]. The main idea for it is to construct an auxiliary identification criterion $\mathcal{J}_\varepsilon(\theta)$ [Tsyganova, 2011; Semushin and Tsyganova, 2013; Semushin et al., 2018] which is instrumental because it depends on only directly observed values and can be minimized with the use of a known numerical optimization methods.

The APA approach to system adaption within the parameter uncertainty differs in the fact that it suggests an indirect state prediction error control in the form of $\mathcal{J}_\varepsilon(\theta)$. It has to satisfy two main requirements:

- it depends on the system observable values only;
- it attains its minimum coincidentally with the OIC (15).

Both the original and instrumental (auxiliary) identification criteria satisfy a relation

$$\mathcal{J}_\varepsilon(\theta) = \mathcal{J}_e(\theta) + \text{Const}. \tag{16}$$

Thus they have one and the same minimizer $\theta^\dagger$, i. e.,

$$\theta^\dagger = \underset{\theta}{\operatorname{argmin}}\, \mathcal{J}_e(\theta) = \underset{\theta}{\operatorname{argmin}}\, \mathcal{J}_\varepsilon(\theta).$$

## 3 Main Results

Now, we are ready to present the main result of the paper — the newly constructed instrumental identification criterion which allows for solving parameter identification problem for the class of discrete-time LTI stochastic systems with unknown exogenous inputs. Thus, we develop the existing APA approach for the systems with the high level of uncertainty. In all previous related works [Semushin, 2011a; Semushin, 2011b; Tsyganova, 2011; Semushin and Tsyganova, 2013; Semushin, 2014; Semushin et al., 2018] it was assumed that exogenous inputs are known or they are Gaussian white noises.

### 3.1 Constructing a New Instrumental Identification Criterion for Discrete-Time LTI Stochastic Systems with Unknown Exogenous Inputs

Consider as an OIC

$$\mathcal{J}_e(\theta) = \mathbf{E}\left\{e_k^T(\theta)e_k(\theta)\right\} = \operatorname{tr}\mathbf{E}\left\{e_k(\theta)e_k^T(\theta)\right\} \quad (17)$$

where $e_k(\theta) = x_k - \hat{x}_k^*(\theta)$ is an estimation error of the state vector $x_k$ that is evaluated within the GDM algorithm for a given $\theta$. If $\theta^\dagger = \operatorname{argmin}\mathcal{J}_e(\theta)$ then $e_k(\theta^\dagger)$ must be minimal.

Suppose that

$$\operatorname{rank} H = n \text{ and } \operatorname{rank} HB = \operatorname{rank} B = r.$$

Let us construct an observable process $\varepsilon_k(\theta)$ in the form

$$\varepsilon_k(\theta) = W^+ z_k - \hat{x}_k^*(\theta) \quad (18)$$

where $W^+ = (H^T H)^{-1}H^T$. Then

$$\begin{aligned}\varepsilon_k(\theta) &= W^+(Hx_k + v_k) - \hat{x}_k^*(\theta)\\ &= (x_k - \hat{x}_k^*(\theta)) + W^+v_k = e_k(\theta) + W^+v_k.\end{aligned} \quad (19)$$

Using (18) we construct IIC in the form

$$\mathcal{J}_\varepsilon(\theta) = \mathbf{E}\left\{\varepsilon_k^T(\theta)\varepsilon_k(\theta)\right\} = \operatorname{tr}\mathbf{E}\left\{\varepsilon_k(\theta)\varepsilon_k^T(\theta)\right\}. \quad (20)$$

**Theorem 1.** *Let matrices $H$ and $B$ in (14) not depend on $\theta$ and* $\operatorname{rank} H = n$, $\operatorname{rank} HB = \operatorname{rank} B = r$. *Then $\mathcal{J}_e(\theta)$ and $\mathcal{J}_\varepsilon(\theta)$ have one and the same minimizer and the following relation holds*

$$\mathcal{J}_\varepsilon(\theta) = \mathcal{J}_e(\theta) + \text{Const} \quad (21)$$

*where*

$$\text{Const} = \operatorname{tr}\left\{W^+R(W^+)^T\right\} - 2\operatorname{tr}\left\{W^+RM^TB^T\right\}$$

*does not depend on $\theta$.*

*Proof.* Taking into account (19) we may rewrite (20) as follows:

$$\begin{aligned}\mathcal{J}_\varepsilon(\theta) &= \operatorname{tr}\mathbf{E}\left\{\varepsilon_k(\theta)\varepsilon_k^T(\theta)\right\}\\ &= \operatorname{tr}\mathbf{E}\left\{(e_k(\theta) + W^+v_k)(e_k(\theta) + W^+v_k)^T\right\}\\ &= \operatorname{tr}\mathbf{E}\left\{e_k(\theta)e_k^T(\theta) + e_k(\theta)v_k^T(W^+)^T\right.\\ &\quad \left. + W^+v_ke_k^T(\theta) + W^+v_kv_k^T(W^+)^T\right\}\\ &= \operatorname{tr}\mathbf{E}\left\{e_k(\theta)e_k^T(\theta)\right\} + \operatorname{tr}\left\{\mathbf{E}\left\{e_k(\theta)v_k^T\right\}(W^+)^T\right\}\\ &\quad + \operatorname{tr}\left\{W^+\mathbf{E}\left\{v_ke_k^T(\theta)\right\}\right\}\\ &\quad + \operatorname{tr}\left\{W^+R(W^+)^T\right\}\\ &= \mathcal{J}_e(\theta) + \operatorname{tr}\left\{\mathbf{E}\left\{e_k(\theta)v_k^T\right\}(W^+)^T\right\}\\ &\quad + \operatorname{tr}\left\{W^+\mathbf{E}\left\{v_ke_k^T(\theta)\right\}\right\}\\ &\quad + \operatorname{tr}\left\{W^+R(W^+)^T\right\}.\end{aligned}$$

Let us find

$$\begin{aligned}\mathbf{E}\left\{e_k(\theta)v_k^T\right\} &= \mathbf{E}\left\{(x_k - \hat{x}_k^*(\theta))v_k^T\right\}\\ &= \mathbf{E}\left\{x_kv_k^T\right\} - \mathbf{E}\left\{\hat{x}_k^*(\theta)v_k^T\right\}.\end{aligned}$$

$\mathbf{E}\left\{x_kv_k^T\right\} = 0$ since $x_k$ and $v_k$ are independent and $\mathbf{E}\left\{v_k\right\} = 0$. Thus

$$\mathbf{E}\left\{e_k(\theta)v_k^T\right\} = -\mathbf{E}\left\{\hat{x}_k^*(\theta)v_k^T\right\}. \quad (22)$$

Let us find $\mathbf{E}\left\{\hat{x}_k^*(\theta)v_k^T\right\}$ using equations of Algorithm 1:

$$\begin{aligned}\mathbf{E}\left\{\hat{x}_k^*(\theta)v_k^T\right\} &= \mathbf{E}\left\{(\hat{x}_{k|k-1}(\theta) + B\hat{u}_{k-1}(\theta))v_k^T\right\}\\ &= \mathbf{E}\left\{\hat{x}_{k|k-1}(\theta)v_k^T\right\} + B\,\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\}.\end{aligned} \quad (23)$$

$\mathbf{E}\left\{\hat{x}_{k|k-1}(\theta)v_k^T\right\} = 0$ since $\hat{x}_{k|k-1}(\theta)$ and $v_k$ are independent and $\mathbf{E}\left\{v_k\right\} = 0$. Thus,

$$\mathbf{E}\left\{\hat{x}_k^*(\theta)v_k^T\right\} = B\,\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\}.$$

Now let us find $\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\}$:

$$\begin{aligned}\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\} &= \mathbf{E}\left\{M(z_k - H\hat{x}_{k|k-1}(\theta))v_k^T\right\}\\ &= M\,\mathbf{E}\left\{z_kv_k^T\right\} - MH\,\mathbf{E}\left\{\hat{x}_{k|k-1}(\theta)v_k^T\right\}.\end{aligned}$$

Since $\mathbf{E}\left\{\hat{x}_{k|k-1}(\theta)v_k^T\right\} = 0$ then

$$\begin{aligned}\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\} &= M\,\mathbf{E}\left\{z_kv_k^T\right\}\\ &= M\,\mathbf{E}\left\{(Hx_k + v_k)v_k^T\right\}\\ &= MH\,\mathbf{E}\left\{x_kv_k^T\right\} + M\,\mathbf{E}\left\{v_kv_k^T\right\}.\end{aligned}$$

Since $\mathbf{E}\left\{x_kv_k^T\right\} = 0$ then

$$\mathbf{E}\left\{\hat{u}_{k-1}(\theta)v_k^T\right\} = MR \text{ where } M = (HB)^+. \quad (24)$$

Using (22)–(24) we obtain

$$\begin{aligned}\mathcal{J}_\varepsilon(\theta) &= \mathcal{J}_e(\theta) - \operatorname{tr}\left\{BMR(W^+)^T\right\}\\ &\quad - \operatorname{tr}\left\{W^+RM^TB^T\right\} + \operatorname{tr}\left\{W^+R(W^+)^T\right\}\\ &= \mathcal{J}_e(\theta) + \operatorname{tr}\left\{W^+R(W^+)^T\right\}\\ &\quad - 2\operatorname{tr}\left\{W^+RM^TB^T\right\}.\end{aligned}$$

Denoting

$$\text{Const} = \operatorname{tr}\left\{W^+R(W^+)^T\right\} - 2\operatorname{tr}\left\{W^+RM^TB^T\right\}$$

we come to (21).

From (21) it follows that $\forall\theta \in D(\theta)$

$$\mathcal{J}_\varepsilon(\theta) = \mathcal{J}_e(\theta) + \text{Const}$$

where Const does not depend on $\theta$.

Therefore

$$\operatorname*{argmin}_\theta \mathcal{J}_e(\theta) = \operatorname*{argmin}_\theta \mathcal{J}_\varepsilon(\theta).$$

□

Now, replacing the expectation operator $\mathbf{E}\{\cdot\}$ in $\mathcal{J}_\varepsilon(\theta)$ by the uniform time averaging, we move to the workable IIC which can be applied in practice

$$\mathcal{J}_\varepsilon(\theta, N) = \frac{1}{N}\sum_{k=1}^{N} \varepsilon_k^T(\theta)\varepsilon_k(\theta). \tag{25}$$

The identification criterion (25) is calculated on the basis of the values obtained by Algorithm 1. To solve the parameter identification problem for system model (14), we suggest to apply IIC $\mathcal{J}_\varepsilon(\theta, N)$ as an objective function for minimization algorithms of various types. If the gradient of the identification criterion is unknown or finding it is computationally expensive, then gradient-free methods such as metaheuristic GA and SA algorithms can be used for minimization.

On the other hand, the gradient-based methods or Newton-like ones can be used for minimization of the IIC. They both take the iterative form of

$$\theta^{j+1} = \theta^j - \beta_j G^{-1}(\theta^j)\nabla_\theta \mathcal{J}_\varepsilon(\theta^j, N) \tag{26}$$

where $\theta^j$ is the parameter vector at the $j$th iteration. In (26), $\nabla_\theta$ denotes the gradient operator $\left[\partial/\partial\theta_1 \,\middle|\, \cdots \,\middle|\, \partial/\partial\theta_p\right]^{\mathrm{T}}$ ($\theta \in \mathbb{R}^p$), which is applied here to the IIC (25) at point $\theta = \theta^j$; $G(\theta^j) = I$, the identity matrix, for the gradient method and $G(\theta^j) = \nabla^2 \mathcal{J}_\varepsilon(\theta^j, N)$, a matrix of second partial derivatives or Hessian matrix of $\mathcal{J}_\varepsilon(\theta^j, N)$ at point $\theta = \theta^j$, for the Newton-like methods. Scalar step size parameter $\beta_j$ is designed to ensure that $\mathcal{J}_\varepsilon(\theta^{j+1}, N) \le \mathcal{J}_\varepsilon(\theta^j, N) + e$ where $e$ is a positive number that can be chosen in a variety of ways [Nocedal and Wright, 2006].

### 3.2 Computing the Gradient of the Proposed Identification Criterion

Let $\theta = [\theta_1, \ldots, \theta_p]^T$ denote the vector of parameters with respect to which the IIC is to be differentiated. From (25) we have

$$\nabla_\theta \mathcal{J}_\varepsilon(\theta, N) = \frac{2}{N}\sum_{k=1}^{N} S_k(\theta)\varepsilon_k(\theta) \tag{27}$$

where $S_k(\theta)$ is the *sensitivity* $(p \times n)$-matrix, and its $(ij)$-th element $s_k^{(ij)}(\theta) = \dfrac{\partial \varepsilon_k^{(j)}(\theta)}{\partial \theta_i}$, $i = 1, \ldots, p$, $j = 1, \ldots, n$.

Let us write the expression for each element of (27):

$$\frac{\partial \mathcal{J}_\varepsilon(\theta, N)}{\partial \theta_i} = \frac{2}{N}\sum_{k=1}^{N} \varepsilon_k^T(\theta)\frac{\partial \varepsilon_k(\theta)}{\partial \theta_i},\ i = 1, \ldots, p. \tag{28}$$

Differentiating $\varepsilon_k(\theta)$ in (18) with respect to $\theta_i$, we obtain

$$\frac{\partial \varepsilon_k(\theta)}{\partial \theta_i} = -\frac{\partial \hat{x}_k^*(\theta)}{\partial \theta_i},\ i = 1, \ldots, p. \tag{29}$$

### 3.3 State Sensitivity Evaluation using Algorithm 1

For simplicity, consider the case $\theta \in \mathbb{R}^1$. To evaluate sensitivities $\dfrac{\partial \hat{x}_k^*(\theta)}{\partial \theta}$ for the estimates of the state $x_k$, we prove the following lemma.

**Lemma 1.** *Let $A(\theta) \in \mathbb{R}^{m\times r}$ be a rectangular matrix parametrized by a scalar parameter $\theta$. Suppose that for a given value of $\theta$ matrix $A = A(\theta)$ has a full column rank and left pseudoinverse matrix $A^+ = (A^TA)^{-1}A^T$. Then*

$$\begin{aligned}\frac{\partial A^+}{\partial \theta} =& (A^TA)^{-1}\frac{\partial A^T}{\partial \theta}(I_m - AA^+) \\ &- A^+\frac{\partial A}{\partial \theta}A^+.\end{aligned} \tag{30}$$

*Proof.*

$$\begin{aligned}\frac{\partial A^+}{\partial \theta} &= \frac{\partial (A^TA)^{-1}A^T}{\partial \theta} \\ &= \frac{\partial (A^TA)^{-1}}{\partial \theta}A^T + (A^TA)^{-1}\frac{\partial A^T}{\partial \theta} \\ &= -(A^TA)^{-1}\frac{\partial (A^TA)}{\partial \theta}(A^TA)^{-1}A^T \\ &\quad + (A^TA)^{-1}\frac{\partial A^T}{\partial \theta} \\ &= -(A^TA)^{-1}\left[\frac{\partial A^T}{\partial \theta}A + A^T\frac{\partial A}{\partial \theta}\right]A^+ \\ &\quad + (A^TA)^{-1}\frac{\partial A^T}{\partial \theta} \\ &= (A^TA)^{-1}\left[\frac{\partial A^T}{\partial \theta} - A^T\frac{\partial A}{\partial \theta}A^+ - \frac{\partial A^T}{\partial \theta}AA^+\right] \\ &= (A^TA)^{-1}\frac{\partial A^T}{\partial \theta}\left(I_m - AA^+\right) \\ &\quad - (A^TA)^{-1}A^T\frac{\partial A}{\partial \theta}A^+ \\ &= (30).\end{aligned}$$

□

Let us construct the sensitivity equations based on the GDM Algorithm 1. Suppose that all filter matrices $F$, $B$, $G$, $Q$, $H$, $R$ can depend on $\theta$, i. e., $F = F(\theta)$, $B = B(\theta)$, etc. Using the rules of matrix differentiation, we obtain

$$\frac{\partial \hat{x}_{k|k-1}}{\partial \theta} = \frac{\partial F}{\partial \theta}\hat{x}_{k-1} + F\frac{\partial \hat{x}_{k-1}}{\partial \theta}, \tag{31}$$

$$\begin{aligned}\frac{\partial P_{k|k-1}}{\partial \theta} =& \frac{\partial F}{\partial \theta}P_{k-1}F^T + F\frac{\partial P_{k-1}}{\partial \theta}F^T \\ &+ FP_{k-1}\frac{\partial F^T}{\partial \theta} + \frac{\partial G}{\partial \theta}QG^T \\ &+ G\frac{\partial Q}{\partial \theta}G^T + GQ\frac{\partial G^T}{\partial \theta},\end{aligned} \tag{32}$$

$$\frac{\partial \widetilde{R}_k}{\partial \theta} = \frac{\partial H}{\partial \theta} P_{k|k-1} H^T + H \frac{\partial P_{k|k-1}}{\partial \theta} H^T + H P_{k|k-1} \frac{\partial H^T}{\partial \theta} + \frac{\partial R}{\partial \theta}. \quad (33)$$

Let us denote $Y_{k-1} = B^T H^T \widetilde{R}_k^{-1} H B$. Then $D_{k-1} = Y_{k-1}^{-1}$ and

$$\frac{\partial D_{k-1}}{\partial \theta} = -Y_{k-1}^{-1} \frac{\partial Y_{k-1}}{\partial \theta} Y_{k-1}^{-1}, \quad (34)$$

$$\frac{\partial Y_{k-1}}{\partial \theta} = \frac{\partial B^T}{\partial \theta} H^T \widetilde{R}_k^{-1} H B + B^T H^T \widetilde{R}_k^{-1} H \frac{\partial B}{\partial \theta} + B^T \left[ \frac{\partial H}{\partial \theta} \widetilde{R}_k^{-1} H - H^T \widetilde{R}_k^{-1} \frac{\partial \widetilde{R}_k^{-1}}{\partial \theta} \widetilde{R}_k^{-1} H + H^T \widetilde{R}_k^{-1} \frac{\partial H}{\partial \theta} \right] B. \quad (35)$$

Let $A = HB$. Then

$$\frac{\partial A}{\partial \theta} = \frac{\partial (HB)}{\partial \theta} = \frac{\partial H}{\partial \theta} B + H \frac{\partial B}{\partial \theta}$$

and $A^+ = M$.

Applying Lemma 1, we obtain

$$\frac{\partial M}{\partial \theta} = (A^T A)^{-1} \frac{\partial A^T}{\partial \theta} (I_m - AM) - M \frac{\partial A}{\partial \theta} M, \quad (36)$$

$$\frac{\partial \hat{u}_{k-1}}{\partial \theta} = \frac{\partial M}{\partial \theta} (z_k - H \hat{x}_{k|k-1}) - M \left[ \frac{\partial H}{\partial \theta} \hat{x}_{k|k-1} + H \frac{\partial \hat{x}_{k|k-1}}{\partial \theta} \right], \quad (37)$$

$$\frac{\partial K_k}{\partial \theta} = \left[ \frac{\partial P_{k|k-1}}{\partial \theta} H^T + P_{k|k-1} \frac{\partial H^T}{\partial \theta} - P_{k|k-1} H^T \widetilde{R}_k^{-1} \frac{\partial \widetilde{R}_k}{\partial \theta} \right] \widetilde{R}_k^{-1}, \quad (38)$$

$$\frac{\partial \hat{x}_k^*}{\partial \theta} = \frac{\partial \hat{x}_{k|k-1}}{\partial \theta} + \frac{\partial B}{\partial \theta} \hat{u}_{k-1} + B \frac{\partial \hat{u}_{k-1}}{\partial \theta}, \quad (39)$$

$$\frac{\partial P_k^*}{\partial \theta} = (I_n - K_k H) \frac{\partial P_{k|k-1}}{\partial \theta} - \left[ \frac{\partial K_k}{\partial \theta} H + K_k \frac{\partial H}{\partial \theta} \right] P_{k|k-1}, \quad (40)$$

$$\frac{\partial \hat{x}_k}{\partial \theta} = \frac{\partial \hat{x}_k^*}{\partial \theta} + \frac{\partial K_k}{\partial \theta} (z_k - H \hat{x}_k^*) - K_k \left[ \frac{\partial H}{\partial \theta} \hat{x}_k^* + H \frac{\partial \hat{x}_k^*}{\partial \theta} \right], \quad (41)$$

$$\frac{\partial P_k}{\partial \theta} = \frac{\partial P_k^*}{\partial \theta} - \left[ \frac{\partial K_k}{\partial \theta} H + K_k \frac{\partial H}{\partial \theta} \right] \times B D_{k-1} B^T (I_n - K_k H)^T + (I_n - K_k H) \times \left[ \frac{\partial B}{\partial \theta} D_{k-1} B^T + B \frac{\partial D_{k-1}}{\partial \theta} B^T + B D_{k-1} \frac{\partial B^T}{\partial \theta} \right] \times (I_n - K_k H)^T - (I_n - K_k H) B D_{k-1} B^T \times \left[ H^T \frac{\partial K_k^T}{\partial \theta} + \frac{\partial H^T}{\partial \theta} K_k^T \right]. \quad (42)$$

Thus, equations (31)–(42) allow us to calculate the partial derivatives of state and input signal estimates for a given value of $\theta$.

## 4 Simulation Results

As a practical application, let us consider the parameter identification problem of an object 2D motion model with unknown exogenuos inputs $u_k$:

$$\begin{cases} x_k = \begin{bmatrix} 1 & \theta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \theta \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} u_{k-1} + \begin{bmatrix} \frac{\theta^2}{2} & 0 \\ \theta & 0 \\ 0 & \frac{\theta^2}{2} \\ 0 & \theta \end{bmatrix} w_k, \\ z_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + v_k, \quad k = 1, \ldots, N \end{cases} \quad (43)$$

where $x_k = [x_1, x_2, x_3, x_4]_k^T$, $x_1 = x$ and $x_3 = y$ are coordinates of the object, $x_2 = v_x$ and $x_4 = v_y$ are its velocity components, initial state $x_0 \sim \mathcal{N}([0, 1, 0, 1]^T, I_2)$, disturbance $w_k \sim \mathcal{N}(0, Q)$ is used to model small accelerations, the turbulence, wind change, and so on, with an appropriate covariance $Q$, measurements noise $v_k \sim \mathcal{N}(0, R)$, and $\theta$ is the model parameter to be identified.

If $u_k \equiv 0$ then the motion model (43) is a nearly constant velocity model [Bar-Shalom et al., 2002].

Let us put the "true" value of the model parameter equal to $\theta^* = 0.2$. Figures 1–3 show estimates of the system state vector $x_k$ obtained with Algorithm 1 for model (43) with $N = 100$, $Q = I_2$, $R = 0.04 I_4$ and input vector

$$u_k = \left[ A \sin \frac{2\pi k}{100}, \; B \cos \frac{2\pi k}{100} \right]^T \quad (44)$$

where $A = B = 1$.

Figures 4, 5 show estimates of the input vector $u_k$ computed by the GDM algorithm.

Let us consider the parameter identification problem for system model (43). We suppose that input vector (44) is unknown.

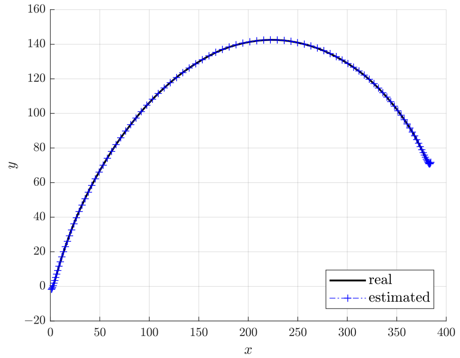


Figure 1. Real and estimated trajectories.

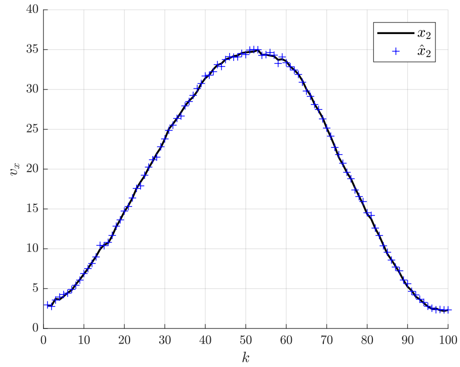


Figure 2. $v_x$ and its estimate

.

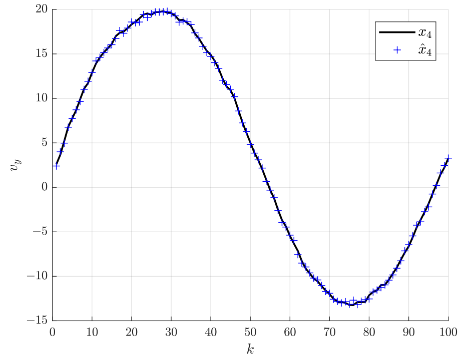


Figure 3. $v_y$ and its estimate.

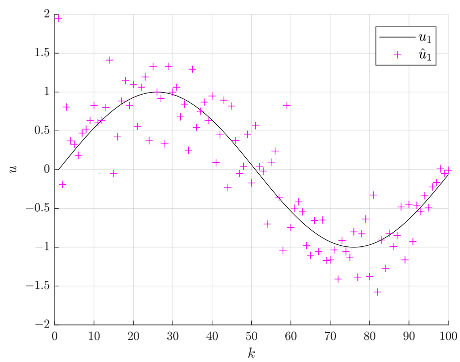


Figure 4. $u_1$ and its estimate.

In order to conduct numerical experiments, we have implemented in MATLAB functions for modeling system dynamics and measurements as well as functions for calculating the identification criterion $\mathcal{J}_\varepsilon(\theta; Z_1^N)$ and its gradient.

Three MATLAB functions were used for numerical minimization of the identification criterion: `simulannealbnd` and `ga` which implement gradient-free metaheuristic algorithms SA (Simulated Annealing) and GA (Genetic Algorithm), respectively; and `fmincon` which was configured to use two different gradient-based algorithms. All experiments were conducted on the following platform: Windows 11, Intel Core i3-1115G4 CPU @ 3.00 GHz, 8 GB of RAM.

Table 1 presents the main settings of the optimizers used in the numerical experiments. The `fmincon` function was used with the following `Algorithm` options: `interior-point` and `trust-region-reflective`.

The `interior-point` algorithm estimates gradients using finite differences and `trust-region-reflective` algorithm uses a user-provided gradient of the objective function. The remaining settings are taken by default.

Table 1. Optimizers settings

| Optimizer | Settings |
|---|---|
| SA | 'Display' = 'off', 'TimeLimit' = Inf, 'MaxIter' = Inf, 'StallIterLimit' = 100, 'ReannealInterval' = 100, 'MaxFunEvals' = Inf |
| GA | 'Display' = 'off', 'TimeLimit' = Inf, 'Generations' = Inf, 'StallGenLimit' = 20, 'PopulationSize' = 10, 'PopInitRange' = [0; 2], 'MutationFcn' = @mutationadaptfeasible |
| FMINCON | 'Display' = 'off', 'MaxFunctionEvaluations' = Inf, 'SpecifyObjectiveGradient' = false, 'Algorithm' = 'interior-point' |
| FMINCON (GRAD) | 'Display' = 'off', 'MaxFunctionEvaluations' = Inf, 'SpecifyObjectiveGradient' = true, 'Algorithm' = 'trust-region-reflective' |

A series of 500 numerical experiments was conducted for different values of noise covariance matrix $R$: $R_1 = \mathrm{diag}(I_2, 4I_2)$, $R_2 = \mathrm{diag}(0.25I_2, I_2)$ and $R_3 = \mathrm{diag}(0.01I_2, 0.04I_2)$. The following settings in (43) and (44) were used: $N = 150$, $Q = I_2$, $A = 0.2$,

Table 3. Identification results.

| | | SA | GA | FMINCON | FMINCON (GRAD) |
|---|---|---|---|---|---|
| $R_1$ | Mean | 0.195693 | 0.195523 | 0.195524 | 0.195528 |
| | RMSE | 0.011436 | 0.006132 | 0.006132 | 0.006129 |
| | MAPE | 4.090858 | 2.603644 | 2.603725 | 2.602772 |
| $R_2$ | Mean | 0.198845 | 0.198653 | 0.198652 | 0.198653 |
| | RMSE | 0.008248 | 0.002596 | 0.002595 | 0.002594 |
| | MAPE | 2.813312 | 1.061867 | 1.061436 | 1.061106 |
| $R_3$ | Mean | 0.199562 | 0.199740 | 0.199740 | 0.199740 |
| | RMSE | 0.008184 | 0.000538 | 0.000536 | 0.000536 |
| | MAPE | 2.469127 | 0.218681 | 0.217678 | 0.217682 |

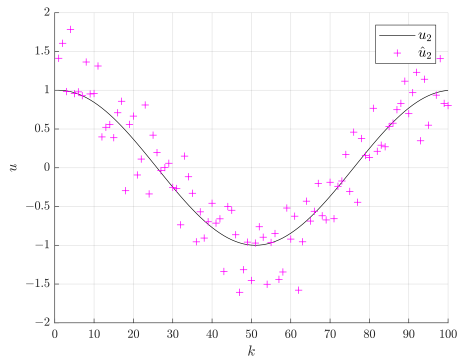


Figure 5. $u_2$ and its estimate.

$B = 0.5$. In each experiment, numerical identification of the parameter $\theta$ was performed based on the results of simulated measurements. The solution $\theta^\dagger$ was searched on the segment $[0; 2]$. The initial point for SA and both gradient-based algorithms was chosen randomly in each experiment.

Table 2 provides average number of iterations and running times for all algorithms. It can be seen that for the problem under consideration both gradient-based algorithms are approximately 10 times faster than GA and 4 times faster than SA, which in turn, on average 2.5 times faster than GA.

Table 2. Average number of iterations and time, sec.

| | | SA | GA | FMINCON | FMINCON (GRAD) |
|---|---|---|---|---|---|
| $R_1$ | Iterations | 170 | 44 | 13 | 9 |
| | Time | 2.439 | 6.245 | 0.630 | 0.609 |
| $R_2$ | Iterations | 172 | 45 | 13 | 9 |
| | Time | 2.632 | 6.843 | 0.684 | 0.638 |
| $R_3$ | Iterations | 174 | 45 | 13 | 9 |
| | Time | 2.582 | 6.591 | 0.604 | 0.682 |

The results of numerical identification of the parameter $\theta$ are summarized in Table 3. They show that with the selected settings, all algorithms demonstrate approximately the same mean accuracy. RMSE and MAPE values decrease with decreasing noise level, but for the SA algorithm they remain significantly larger than for other algorithms.

## 5 Conclusion

In this paper, we have stated and solved the parameter identification problem for a class of discrete-time linear stochastic systems with unknown exogenous inputs. The main results of the paper — the newly constructed instrumental identification criterion (IIC) and the system state sensitivity equations for evaluation of the IIC gradient.

We have proved that the original identification criterion that depends on unobservable estimation error and the new instrumental identification criterion that depends on observable process have one and the same minimizer. Therefore, to solve the parameter identification problem, one can use the IIC instead of the original identification criterion. Moreover, we have shown that the values of the IIC can be calculated using the GDM Algorithm 1, and we have constructed new state sensitivity equations to calculate the values of the IIC gradient.

Numerical experiments have demonstrated how the newly proposed instrumental identification criterion and method for calculation the IIC gradient can be applied for solving the parameter identification problem of the considered stochastic system model with unknown periodic input signal. The comparison results of four different methods (gradient-based and metaheuristic) for minimizing the proposed identification criterion are presented.

It is worth noting that the restriction on the rank of matrix $H$ introduced in Theorem 1 is not principal. Our method can also be used for noises covariances identification as an alternative to the method proposed recently in [Kong et al., 2023].

## Acknowledgements

This work was supported by the Russian Science Foundation, grant no. 23–21–00361, https://rscf.ru/project/23-21-00361/.

## References

Anderson, B.D.O., and Moore, J.B. (1979). *Optimal filtering*. Prentice-Hall.

Åström, K.-J. (1980). Maximum Likelihood and Prediction Error Methods. *Automatica*, **16** (5), pp. 551–574.

Bar-Shalom, Y., Li, X.R., and Kirubarajan, T. (2002) *Estimation with Applications to Tracking and Navigation : Theory, Algorithms and Software*. John Wiley & Sons, Inc., New York, USA.

Darouach, M., Zasadzinski, M., and Xu, S.J. (1994) Full-order observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, **39** (3), pp. 606–609.

Darouach, M., and Zasadzinski, M. (1997). Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, **33** (4), pp. 717–719.

Darouach, M., Zasadzinski, M., and Boutayeb, M. (2003). Extension of minimum variance estimation for systems with unknown inputs. *Automatica*, **39** (5), pp. 867–876.

Friedland, B. (1969). Treatment of bias in recursive filtering. *IEEE Transactions on Automatic Control*, **14**, pp. 359–367.

Furtat, I.B., Zhukov, Y.A., and Matveev, A.M. (2023). Design of optimal-robust control of step motors with partially unknown parameters and disturbances. Cybernetics and Physics, **12** (1), pp. 23–27.

Gillijns, S., and De Moor, B. (2007a). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, **43** (1), pp. 111–116.

Gillijns, S., and De Moor, B. (2007b). Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. *Automatica*, **43** (5), pp. 934–937.

Gillijns, S., Haverbeke, N., and De Moor, B. (2007). Information, covariance and square-root filtering in the presence of unknown inputs. In *Proceedings of the 2007 European Control Conference (ECC)*, Kos, Greece, pp. 2213–2217.

Gillijns, S., and De Moor, B. (2007c). Joint State and Boundary Condition Estimation in Linear Data Assimilation Using Basis Function Expansion. In *Proceedings of the 26th IASTED International Conference on Modelling, Identification, and Control*, Innsbruck, Austria, pp. 458–463.

Grewal, M.S., and Andrews, A.P. (2015). *Kalman filtering: Theory and Practice Using MATLAB* (4th ed.). John Wiley & Sons, Inc., New York, USA.

He, K., and Liu, H. (2023). An adaptive three-stage information filter of simultaneous state and unknown disturbance estimation in linear time-varying systems. *GPS Solut.*, **27** (1).

Hmida, F.B., Khémiri, K., Ragot, J., and Gossa, M. (2010). Robust Filtering for State and Fault Estimation of Linear Stochastic Systems with Unknown Disturbance. *Mathematical Problems in Engineering*, vol. 2010, Article ID 591639, 24 pages.

Hou, M., and Müller, P.C. (1992). Design of observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, **37** (6), pp. 871–875.

Hou, M., and Müller, P.C. (1994). Disturbance decoupled observer design: A unified viewpoint. *IEEE Transactions on Automatic Control*, **39** (6), pp. 1338–1341.

Hou, M., and Patton, R.G. (1998). Optimal filtering for systems with unknown inputs. *IEEE Transactions on Automatic Control*, **43** (3), pp. 445–449.

Hsieh, C.-S. (2000). Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Transactions on Automatic Control*, **45** (12), pp. 2374–2378.

Hua, Y., Wang, N., and Zhao, K. (2021). Simultaneous unknown input and state estimation for the linear system with a rank-deficient distribution matrix. *Mathematical Problems in Engineering*, ID 6693690.

Ignagni, M.B. (1990). Optimal strapdown attitude integration algorithms. *Journal of Guidance*, **13** (2), pp. 363–369.

Kitanidis, P.K. (1987). Unbiased minimum-variance linear state estimation. *Automatica*, **23** (6), pp. 775–778.

Kong, H., Sukkarieh, S., Arnold, T. J., Chen, T., Mu, B., and Zheng, W.X. (2023). Identifiability analysis of noise covariances for LTI stochastic systems with unknown inputs. *IEEE Transactions on Automatic Control*, **68** (7), pp. 4459–4466.

Kudva, P., Viswanadham, N., and Ramakrishna, A. (1980). Observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, **25** (1), pp. 113–115.

Lan, H., Liang, Y., Yang, F., Wang, Z., and Pan, Q. (2013). Joint estimation and identification for stochastic systems with unknown inputs. *IET Control Theory & Applications*, **7** (10), pp. 1377–1386.

Ljung, L. (1999). *System Identification — Theory for the User* (2nd ed). Prentice-Hall, Upper Saddle River, NJ, USA.

Nocedal, J. and Wright, S.J. (2006). *Numerical optimization.* In Springer Series in Operations Research and Financial Engineering. Springer Nature, pp. 1–664.

Palanthandalam-Madapusi, H.J., and Bernstein D.S. (2007). Unbiased Minimum-variance Filtering for Input Reconstruction. In *Proceedings of the 2007 American Control Conference*, New York, NY, USA, pp. 5712–5717.

Semushin, I.V. (2011a). Adaptation in Stochastic Dynamic Systems — Survey and New Results I. *Int. J. Commun. Netw. Syst. Sci.*, **4**, pp. 17–23.

Semushin, I.V. (2011b). Adaptation in Stochastic Dynamic Systems — Survey and new results II. *Int. J. Commun. Netw. Syst. Sci.*, **4**, pp. 266–285.

Semushin, I.V., and Tsyganova, J.V. (2013). Adaptation in Stochastic Dynamic Systems — Survey and New Results IV: Seeking Minimum of API in Parameters of Data. *Int. J. Commun. Netw. Syst. Sci.*, **6**, pp. 513–518.

Semushin, I.V. (2014). The APA based time-variant system identification. In *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC)*, Los Angeles, CA, USA, 15–17 December, pp. 4137–4141.

Semushin, I.V., Tsyganova, J.V., and Tsyganov, A.V. (2018). Numerically Efficient LD-computations for the Auxiliary Performance Index Based Control Optimization under Uncertainties. *IFAC-PapersOnline*, **51**, pp. 568–573.

Tranninger, M., Niederwieser, H., Seeber, R., and Horn, M. (2023). Unknown input observer design for linear time-invariant systems — A unifying framework. *Int.*

*J. Robust Nonlinear Control*, **33** (15), pp. 8911–8934.

Teymouri, D., Sedehi, O., Katafygiotis, L.S., and Papadimitriou, C. (2020). Real-time Bayesian parameter, state and input estimation using output-only vibration measurements. In *Proceedings of the International Conference on Structural Dynamic*, **2**, p. 3590–3598.

Tsyganova, Y.V. (2011). Computing the gradient of the auxiliary quality functional in the parametric identification problem for stochastic systems. *Autom. Remote Control*, **72**, pp. 1925–1940.

Varshney, D., Bhushan, M., and Patwardhan, S.C. (2019). State and parameter estimation using extended Kitanidis Kalman filter. *Journal of Process Control*, **76**, pp. 98–111.

Xiao, M., Zhang, Y., Wang, Z., and Fu, H. (2018). An adaptive three-stage extended Kalman filter for nonlinear discrete-time system in presence of unknown inputs. *ISA Transactions*, **75**, pp. 101–117.

Xiong, Y., and Saif, M. (1994). Unknown disturbance inputs estimation based on a state functional observer design. *Automatica*, **39** (8), pp. 1389–1398.

Yu, D., and Chakravorty, S. (2016). A stochastic unknown input realization and filtering technique. *Automatica*, **63** (1), pp. 26–33.

Yong, S.Z., Zhu, M., and Frazzoli E. (2013). Simultaneous input and state estimation for linear discrete-time stochastic systems with direct feedthrough. In *Proceedings of the 52nd IEEE Conference on Decision and Control*, Firenze, Italy, 10–13 December, pp. 7034–7039.

Yong, S.Z., Zhu, M., and Frazzoli E. (2014). Simultaneous input and state smoothing for linear discrete-time stochastic systems with unknown inputs. In *Proceedings of the 53rd IEEE Conference on Decision and Control*, Los Angeles, CA, USA, 2014, pp. 4204–4209.

Yong, S.Z., Zhu, M., and Frazzoli E. (2016). A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. *Automatica*, **63** (1), pp. 321–329.

Zhirabok, A.N., Zuev, A.V., and Shumsky, A.E. (2023). A sliding mode technique for reconstruction of unknown control input in linear systems. *Cybernetics and Physics*, **12** (1), pp. 82–88.