ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 2, pp. 97–140.

REVIEWS

Overview on Observer Design Principles for Systems with Unknown Parameters

A. I. Glushchenko[*,a] **and K. A. Lastochkin**[*,b]

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

Received September 7, 2024
Revised July 11, 2025
Accepted July 15, 2025

Abstract—The overview provides a comparative analysis of the feasibility conditions for key methods to design state observers for linear dynamical SISO systems with unknown parameters. A distinctive feature of this paper is the comparison between state observers that are based on the robust, invariant, and adaptive methods. The overview is written without excessive mathematical details and aims to familiarize a wide range of readers with the basic principles of design and functioning of state observers for linear systems with unknown parameters.

Keywords: state observers, parametric uncertainty, robustness, invariance, adaptivity

DOI: 10.7868/S1608303226020012

1. INTRODUCTION

State observers are special software implementable algorithms that reconstructs unmeasurable system states at any given time instant on the basis of available *a priori* information and online measurements of system input and output. The invention and development of state observers is a response of the mathematical community to the desire of engineers to increase reliability and simplify the design of measurement systems as an essential part of automated control systems. Observers have been most widely applied in the field of automated electric drives design, where today most control systems for synchronous and asynchronous drives are implemented by means of sensorless control [1]. Such success could not have been achieved without a deep theoretical investigation of the problem of state estimation.

The mathematical theory of observer design originated from the studies of D. Luenberger [2] and R. Kalman [3], who respectively stated deterministic and stochastic estimation problems. By now, these two parallel tasks have received an incredible number of extensions and generalizations in both continuous and discrete time. This circumstance forces us in this paper to limit ourselves to consideration and analysis of only a small part of the available results and developments. For interested readers, who would like to engage deeply into the topic and cover the problem more broadly, we recommend to refer to the reviews [4, 5] and books [6–12].

In this paper, the simplest case of a linear dynamical system with one input and one output is considered, and the task is to reconstruct the state of such a system in its original form (physical state) rather than the virtual state of its canonical forms. If the matrices of such a system are precisely known in advance, and the observability condition is met, then the Luenberger observer [2] is a solution to the estimation problem and allows one to reconstruct unmeasurable state with a given exponential convergence rate. However, how to solve the problem if the matrices of the system are unknown or not known precisely enough? The theories of robust [8–10], invariant [7, 8, 11] and adaptive [12] control are simultaneously looking for an answer to this question.[1]

Robust observers. This group mainly concerns the improvement of (and development of new) offline methods to calculate a correction gain for the basic structure of the Luenberger observer. Various design methods are being developed and improved, which (*i*) at the stage of observer parameters computation, provide guarantees of state estimation quality for the parametrically uncertain systems and additionally (*ii*) solve some optimization problems ($\mathcal{H}_2$, $\mathcal{H}_\infty$ norms of the transfer function, the quadratic functional for a linear-quadratic problem, the "size" of an invariant (or attractive) ellipsoid for problems related to the external disturbances attenuation, etc.). The general limitations of robust methods are that *i*) the amplitude of the disturbance and the range of parametric uncertainty are required to be *a priori* known, *ii*) the parametric uncertainty of the system is to satisfy special parameterizations. The approaches of this group have been developed at different times by following foreign and Russian researchers: S. Bhattacharyya, J. Doyle, K. Zhou, H. Kwakernaak, I. Petersen, B. Barmish, R. Tempo, W. Schmitendorf, F. Jabbari, D. Bernstein, M. Corless, C.H. Lien, C.E. de Souza, J.C. Geromel, D. Peaucelle, B.T. Polyak, M.V. Khlebnikov, P.S. Shcherbakov, I.G. Vladimirov, A.P. Kurdyukov, D.V. Balandin, M.M. Kogan, A.G. Alexandrov, V.N. Chestnov, S.K. Korovin, V.V. Fomichev, A.V. Ushakov, R.O. Omorov, etc.

Estimation on the basis of invariance theory. The essence of the methods from this group is the idea to ensure invariance of state reconstruction with respect to the parametric and/or additive disturbances. Invariance is achieved by algebraic elimination of a generalized disturbance from the state reconstruction error equation, or by its high-gain attenuation via choice of correction gain from the class of functions with high-gain coefficients or discontinuous signals. The feasibility conditions of observers based on the invariance theory are, firstly, knowledge of the system uncertainty range (to get to a sliding surface), and secondly, the fulfillment of strict output matching conditions (to ensure existence of such sliding surface). The application of invariance theory methods to the problem of state reconstruction is primarily related to C. Edwards, S. K. Spurgeon, B. L. Walcott, T. Floquet, J.-P. Barbot, M. Darouach, P. Kudva, S. Zak, A. Levant, L. Fridman, Y. Shtessel, K. Khalil, J. Slotine, V.I. Utkin, A.S. Poznyak, S.A. Krasnova, V.A. Utkin, A.N. Zhirabok, etc.

Adaptive observers. Observers from this group simultaneously estimate unmeasurable state and system unknown parameters. Owing to this essentially simple idea, the *a priori* known range of parametric uncertainty is not required for the observer design. Instead, classical adaptive observers require to meet the condition of persistent excitation of some function from the measured signals (control and output) for the asymptotic convergence of the state estimates to their true value. For a long period of time, this condition stymied practical interest in adaptive state reconstruction methods, since, according to various interpretations, such condition is equivalent to (*i*) the global sufficient richness of a control signal of such order that is equal to the dynamic order of the system, (*ii*) the complete observability of the extended system (states + unknown parameters). By the efforts of G. Chowdharry, R. Ortega, Y. Pan, S.B. Roy, A.A. Bobtsov, S.V. Aranovskiy and many other researchers, this condition has been relaxed to the requirement of a regressor finite excitation, which is equivalent to the sufficient richness of the control signal over some finite time interval or the local complete observability of the extended system. The development of methods to design adaptive state observers is closely related to studies by K. Narendra, G. Luders, R. Carroll, D. Lindorf, G. Kreisselmeier, A. Annaswamy, P. Ioannou, R. Marino, P. Tomei, A. Isidori, R. Ortega, S.B. Roy, D. Efimov, A.A. Bobtsov, V.O. Nikiforov, A.A. Pyrkin, S.V. Aranovskiy, N.N. Karabutov, etc.

The aim of this study is to compare the feasibility conditions of state observers designed on the basis of these three theories. In authors' opinion, the relevance of such a comparative analysis is caused by two circumstances. First, the application of observers for practical scenarios requires an

[1] Here the books are cited that summarize and systematize the main results in the mentioned fields of control theory.

engineer to choose the algorithm that best suits the conditions of the applied problem being solved. This, in turn, requires such engineer to know the basic methods to design state observers and the conditions of their applicability. This overview provides such information in a compact form, which makes it possible to familiarize oneself with the main existing approaches. Secondly, researchers interested in one theory (robust, invariant, or adaptive) do not always, but very often speak different languages with colleagues who solve the same problem, but use methods from another theory. The authors hope that this overview will help to improve the situation and achieve greater mutual understanding between these groups, which in the future will result in the enrichment of all three theories with new approaches.

Of course, it is impossible to provide an exhaustive overview, so we would like to note that this paper discusses the results that seem to be the most interesting from the subjective point of view of the authors and mainly relate to the task of physical rather than virtual state reconstruction. At the same time, not only the authors do not disclaim responsibility to *i*) the readers for subjectivity and incompleteness of information, *ii*) the researchers for the possible lack of references to their papers on the subject under consideration, but also hope to receive feedback, which will undoubtedly be of use.

The overview has the following structure. The second section show the relevance of the problem of the state vector reconstruction from the point of view of the stabilization task for a linear time-invariant system. Sections 3–5 discuss in detail the existing approaches to design observers using robust, invariant, and adaptive control methods, respectively. The sixth section is devoted to the analysis of the applicability of the state estimate obtained with the help of the overviewed observers for the purposes of feedback control. The review is wrapped up with the general conclusions presented in Section 7. The research interest of the authors is related to adaptive control theory, therefore, in the overview, more attention is paid to the methods of adaptive observers design.

Further the following notation is used: $|.|$ is the absolute value, $\|.\|$ is the suitable vector or matrix norm of $(.)$, $I_{n\times n} = I_n$ is an identity $n\times n$ matrix, $0_{n\times n}$ is a zero $n\times n$ matrix, 0_n stands for a zero vector of length n, $\det\{.\}$ stands for a matrix determinant, $\mathrm{adj}\{.\}$ represents an adjoint matrix, $\mathrm{tr}\{.\}$ is a matrix trace, $\mathrm{rank}(.)$ denotes a matrix rank, $\dim(.)$ stands for a dimension of a vector or matrix, $\sigma\{.\}$ is a matrix spectrum, $\mathrm{sgn}(.)$ denotes a sign function, $\ln(.)$ stands for a natural logarithm, $\mathrm{vec}\,(.)$ is the operation of a matrix vectorization, $\mathrm{mat}(.)$ stands for an operation that is inverse to the vectorization, $\mathrm{diag}\{.\}$ denotes a diagonal matrix. The matrix $A \in \mathbb{R}^{n\times n}$ minimum and maximum eigenvalues are denoted as $\lambda_{\min}(A) = \min\limits_i (\mathrm{Re}\{\lambda_i(A)\})$ and $\lambda_{\max}(A) = \max\limits_i (\mathrm{Re}\{\lambda_i(A)\})$, respectively, where $i = 1, 2, \ldots, n$. The Latin abbreviation exp is used to denote the exponential convergence rate. For a mapping $\mathcal{F} : \mathbb{R}^n \mapsto \mathbb{R}^n$ we denote its Jacobian by $\nabla_x \mathcal{F}(x) = \frac{\partial \mathcal{F}}{\partial x}(x)$. We also use the fact that for all (possibly singular) matrices M of dimension $n\times n$ the following holds: $\mathrm{adj}\{M\}M = \det\{M\}I_{n\times n}$, and $f \in L_q$ means $\sqrt[q]{\int_{t_0}^{t} \|f(s)\|^q ds} < \infty$ for all $t \geq t_0$.

2. STATE ESTIMATION TASK FROM PERSPECTIVE OF CONTROL PROBLEM

Many real-world technical systems can be described with sufficient accuracy by linear models with time-invariant parameters (single input–single output systems are considered hereafter):

$$\begin{gathered} \dot{x}(t) = Ax(t) + Bu(t), \quad x(t_0) = x_0, \\ y(t) = C^{\mathrm{T}} x(t), \end{gathered} \tag{2.1}$$

where $t \geq t_0 \geq 0$ stands for time, $x(t) \in \mathbb{R}^n$ is a vector of unmeasurable physical states with unknown initial conditions x_0, $u(t) \in \mathbb{R}$ denotes a measurable signal or controller output, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ are unknown matrix and vector, $C \in \mathbb{R}^n$ stands for a known vector that forms a measurable output $y(t) \in \mathbb{R}$. The pairs $\left(C^{\mathrm{T}}, A\right)$ and (A, B) are completely observable and controllable, respectively (definitions and criteria of controllability and observability can be found in [6, 23]).

One of the basic problems of control theory is to solve a stabilization task, that is, to ensure convergence of the linear system states from arbitrary bounded initial condition x_0 to the origin. If the pair of matrices (A, B) is known and controllable, then this task can be solved using a state-feedback control law:

$$u(t) = -Kx(t), \tag{2.2}$$

where the row-vector $K \in \mathbb{R}^{1\times n}$ is chosen so that to satisfy $\lambda_{\max}(A - BK) < 0$ and can be calculated, for example, via application of a pole placement method [13, 14]:

$$\begin{gathered} AM - MA_{ref} = Bh, \\ K = hM^{-1}, \end{gathered} \tag{2.3}$$

where the matrix A_{ref} and vector h are chosen so that $\sigma\{A\} \cap \sigma\{A_{ref}\} = \emptyset$, and the pair (h, A_{ref}) is observable.

However, since, according to the problem statement, the states of the system are unmeasurable, and the parameters of the matrices A and B are unknown, then the feedback (2.2) cannot be implemented, and the task to design dynamic feedback naturally takes the following forms:

$$u(t) = -K_r\hat{x}(t), \tag{2.4a}$$

$$u(t) = -\hat{K}(t)\hat{x}(t), \tag{2.4b}$$

where:

(2.4a) is a robust dynamic feedback (K_r ensures $\lambda_{\max}(A - BK_r) < 0$ for all matrices A and B that belong to some known sets);

(2.4b) denotes an adaptive dynamic feedback (the adaptive laws for $\hat{K}(t)$ and $\hat{x}(t)$ jointly ensure asymptotic stability of the system (2.1) origin). In a particular case, $\hat{K}(t)$ can be an estimate of the parameters K of a control law (2.2), (2.3).

In order to implement the feedback laws (2.4a) and (2.4b), the estimate of the system (2.1) state is required to satisfy the following:

$$\lim_{t\to\infty}\|\tilde{x}(t)\| = \lim_{t\to\infty}\|\hat{x}(t) - x(t)\| = 0, \tag{2.5}$$

where $\hat{x}(t)$ is a system state estimate, $\tilde{x}(t)$ is a error of state reconstruction/observation/estimation.

Further, the overview focuses on methods to solve the problem (2.5) using the results of the theories of robust, invariant and adaptive control, while the issues of adaptive law design for $\hat{K}(t)$ and equations (procedures) to calculate the parameters K_r are almost everywhere stay beyond the scope of the overview.

In this regard, in order to correctly analyze methods from literature, the following classical assumption is adopted.

Assumption 1. The control signal $u(t)$ is bounded and ensures boundedness of the states $x(t)$ for all time instants.

Remark 1. In addition to the problem of system stabilization via dynamic state-feedback control law discussed in this section, for which, as mentioned earlier, estimates of virtual states of the observer canonical form may be sufficient instead of $\hat{x}(t)$, practice is replete with other tasks that require to estimate the vector of unmeasurable physical states $x(t)$ of the system. For example, it is fault detection, technological processes monitoring and logging of unmeasurable signals, design of digital twins and other applied tasks. Therefore, the goal of state observer design is defined in the form (2.5).

3. ROBUST OBSERVERS

A conventional Luenberger observer is described as follows:

$$\dot{\hat{x}}(t) = A_n\hat{x}(t) + B_n u(t) + L\left(y(t) - C^{\mathrm{T}}\hat{x}(t)\right), \quad \hat{x}(t_0) = \hat{x}_0, \tag{3.1}$$

where $A_n \in \mathbb{R}^{n\times n}$ is a matrix of nominal system parameters, $B_n \in \mathbb{R}^{n\times 1}$ stands for a nominal input matrix, $L \in \mathbb{R}^{n\times 1}$ denotes a gain vector of Luenberger correction term, the pair (C^{T}, A_n) is observable.

If it is true for the system (2.1) that $A = A_n$, $B = B_n$, then, as it results from the following equation (in this section we redefine $e(t) = -\tilde{x}(t)$):

$$\dot{e}(t) = \left(A_n - LC^{\mathrm{T}}\right)e(t), \quad e(t_0) = x_0 - \hat{x}_0, \tag{3.2}$$

the observer (3.1) is a solution of the stated problem (2.5) in case $A_n - LC^{\mathrm{T}}$ is a Hurwitz matrix.

If $A \neq A_n$ and/or $B \neq B_n$, then, instead of (3.2), we have:

$$\begin{aligned} \dot{e}(t) &= Ax(t) + Bu(t) - A_n\hat{x}(t) - B_n u(t) - L\left(y(t) - C^{\mathrm{T}}\hat{x}(t)\right) \pm A_n x(t) \\ &= \left(A_n - LC^{\mathrm{T}}\right)e(t) + (B - B_n)u(t) + (A - A_n)x(t). \end{aligned} \tag{3.3}$$

Robust state observers are based on the idea of ensuring that the goal (2.5) is met not for a single pair of known matrices A_n, B_n, but for some predefined and known set of matrices. From the point of view of the system (3.3) stability, two different situations can be singled out.

Filtering problem. Assumption 1 is met, the signal $u(t)$ is measurable, but not allowed to be chosen by designer/engineer. Under these conditions, the observer (3.1) does not allow one to solve the problem (2.5) without additional restrictions imposed on the class of signal $u(t)$. If $u \in L_2$, then the goal (2.5) can be achieved and additionally we can state the task of $\mathcal{H}_\infty$ filtering, *i.e.*, to ensure a given $\gamma > 0$ ratio of L_2 norm of the error $e(t)$ to L_2 norm of the signal $u(t)$:

$$\sqrt{\int_{t_0}^{\infty} e^2(s)\,ds} \leq \gamma\sqrt{\int_{t_0}^{\infty} u^2(s)\,ds}. \tag{3.4}$$

If $u \in L_\infty$ and A is a Hurwitz matrix, then (3.1) is able to ensure only convergence of the estimation error into a certain bounded set of the phase space, and the best we can do is to state and solve a problem of such set minimization in a certain sense.

Linear dynamic controller design problem. Control signal is a linear function of the observer states, that is, $u(t) = -K_r\hat{x}(t)$. The feedback parameters K_r, along with the correction gain vector L can be chosen in the course of the observer design. In this situation, by joint calculation of K_r and L, the goal (2.5) can be achieved.

In order to use methods of robust theory to solve these problems, the parametric uncertainties in A and B matrices are parametrized into special forms. For example, a structured form is often used (such type of perturbations are also called norm-bounded) [15]:

$$A = A_n + F_A\Delta_A H_A, \quad B = B_n + F_B\Delta_B H_B, \tag{3.5}$$

where matrix uncertainties $\Delta_A \in \mathbb{R}^{p_A\times q_A}$ and $\Delta_B \in \mathbb{R}^{p_B\times q_B}$ satisfy the inequalities

$$\|\Delta_A\| \leq 1, \quad \|\Delta_B\| \leq 1,$$

and weight matrices $F_A \in \mathbb{R}^{n\times p_A}$, $H_A \in \mathbb{R}^{q_A\times n}$, $F_B \in \mathbb{R}^{n\times p_B}$, $H_B \in \mathbb{R}^{q_B\times n}$ are known.

An alernative parametrizaion is an affine one [16]:

$$R(q) = (A(q), B(q)) \in \mathcal{R},$$
$$\mathcal{R}{:} = \left\{ R(q) : R(q) = \sum_{i=1}^{s} q_i R_i;\ \sum_{i=1}^{s} q_i = 1,\ q_i \geq 0 \right\}. \tag{3.6}$$

Regardless of the chosen uncertainty parameterization, it is usually additionally assumed that A_n is a Hurwitz matrix (this can always be achieved by *a priori* selection of the nominal part of the system).

In the following subsections, possible solutions are considered to the problems of filtering and linear dynamic controller design in the presence of a structured (3.5) or affine (3.6) uncertainty.

3.1. Robust Filtering Problem

As, considering the filtering problem, we do not design the input signal $u(t)$, then the success of its solution depends entirely on *a priori* assumptions about the class of this signal. There is a distinction between stochastic filtering tasks (for example, $u(t)$ is a random white noise with zero mean) and deterministic ones (for example, $u \in L_2$ or $u \in L_\infty$). In this overview, we consider some solutions to the filtering problem only for a deterministic case.

3.1.1. $\mathcal{H}_\infty$ filtering. The aim of $\mathcal{H}_\infty$ filtering is to obtain state estimate $\hat{x}(t)$ such that, if $u \in L_2$, then (2.5) holds, and, if $x_0 = 0$, then (3.4) holds for a given $\gamma > 0$ for all uncertainties (3.5) or (3.6).

Up to date, a solution to this problem has been obtained for both structured (3.5) and affine (3.6) uncertainties (see [17, 18] and [19–22], respectively). Considering structured uncertainty, the solution is obtained in terms of solutions of two related Riccati equations (2-Riccati approach). As for the case of affine uncertainty, solutions are obtained in terms of linear matrix inequalities derived on the basis of quadratic Lyapunov functions, both independent [19, 20] and dependent [21, 22] from a parameter. For a detailed comparative overview of existing solutions, we would like to address an interested reader to the treatise [9]. In this paper, we restrict our consideration to one algorithm that ensures (2.5) and (3.4) for affine uncertainty.

The solution to the $\mathcal{H}_\infty$ filtering problem is usually designed using an observer defined by the following equation:

$$\dot{\hat{x}}(t) = G\hat{x}(t) + Ly(t), \quad \hat{x}(t) = \hat{x}_0. \tag{3.7}$$

Unlike (3.1), here the input signal $u(t)$ is not used to obtain state estimate, and not only correction gain vector L is a design parameter, but also the matrix G, which is no longer fixed equal to the nominal matrix of the system A_n. The procedure to calculate these parameters is given by the following theorem.

Theorem 1. *Let $\hat{Y}_1$, $\hat{Y}_2$ and $\hat{P}_1$, $\hat{P}_2$ be solutions of the following linear matrix inequalities (LMI)*

$$\begin{bmatrix} P_1 & P_2 \\ P_2 & P_2 \end{bmatrix} > 0, \quad \begin{bmatrix} \Omega_{1,i} + \Omega_{1,i}^{\mathrm{T}} & Y_1 + \Omega_{2,i}^{\mathrm{T}} & P_1 B_i & I \\ * & Y_1 + Y_1^{\mathrm{T}} & P_2 B_i & 0 \\ * & * & -\gamma^2 I & 0 \\ * & * & * & -I \end{bmatrix} < 0, \quad \forall i = 1, \ldots s, \tag{3.8}$$

where

$$\Omega_{1,i} = P_1 A_i + Y_2 C^{\mathrm{T}}, \quad \Omega_{2,i} = P_2 A_i + Y_2 C^{\mathrm{T}},$$

with respect to $Y_1 \in \mathbb{R}^{n\times n}$, $Y_2 \in \mathbb{R}^n$ and $P_1 \in \mathbb{R}^{n\times n}$, $P_2 \in \mathbb{R}^{n\times n}$ for a given $\gamma > 0$.

Then the parameters $G = \hat{P}_2^{-1}\hat{Y}_1$, $L = \hat{P}_2^{-1}\hat{Y}_2$ ensure that the goals (2.5) *and* (3.4) *are achieved.*

Proof is given in [9].

The observer (3.7) designed via solution of LMI (3.8) has three main problems. First, the goal (2.5) can be achieved only when the uncertainty satisfies the predetermined parameterization (3.6), which is, in fact, typical of all robust solutions. Second, the excessive conservatism is imposed by the design procedure and related to the application of a common Lyapunov function for s vertices of the polytope (3.6). Third, the goal (2.5) is achieved only if the restrictive condition $u \in L_2$ is satisfied. As for the second problem, there exist procedures [21, 22] that solve the problem (3.4) based on a parameter-dependent Lyapunov function.

3.1.2. Invariant ellipsoid method. Considering a more practically common case $u \notin L_2$, $u \in L_\infty$ and when the matrix $A_n - LC^{\mathrm{T}}$ is Hurwitz one, we can only conclude from (3.3) that the norm of the estimation error converges asymptotically to some set:

$$\lim_{t\to\infty} \|e(t)\| \le e_{\max},$$

where $e_{\max}$ depends from both the upper bound of the signal $(B - B_n)\,u(t) + (A - A_n)\,x(t)$, and the chosen correction gain vector L.

For this situation, a more precise description of the asymptotic behavior of the error can be obtained using the notion of an invariant ellipsoid of the system.

Definition 1. An ellipsoid centered at the origin

$$\mathcal{E}_e = \left\{ e \in \mathbb{R}^n \colon e^{\mathrm{T}} P_e^{-1} e \le 1 \right\} \quad P_e > 0$$

is said to be positively invariant for dynamical system (3.3) + (3.5) if for all uncertainties $\|\Delta_A\| \le 1$, $\|\Delta_B\| \le 1$ and all input signals $|u(t)| \le 1$ it holds that:

1) $e_0 \in \mathcal{E}_e \Rightarrow e(t) \in \mathcal{E}_e$ for all $t \ge t_0$;
2) $e_0 \notin \mathcal{E}_e \Rightarrow e(t) \to \mathcal{E}_e$ for $t \to \infty$.

The matrix $P_e \in \mathbb{R}^{n\times n}$ is a configuration matrix of the ellipsoid $\mathcal{E}_e$.

Obviously, there exists an infinite number of invariant ellipsoids for the considered system (3.3). Since, in case $u \in L_\infty$, $A \ne A_n$ and/or $B \ne B_n$, the goal (2.5) for the filtering problem cannot be achieved by means of robust control, it remains only to try to provide convergence of the error $e(t)$ into the minimal ellipsoid. Minimal ellipsoid can be understood in different ways. Most commonly, in the literature, the minimality criterion is chosen as the minimization of the sum of squared semi-axes of the ellipsoid or, equivalently, the minimization of the trace of the ellipsoid matrix P_e [23]. Strictly, the problem of minimization of the sum of squared semi-axes of the ellipsoid is stated as follows.

Assume that the signal $u(t)$ is bounded by $|u(t)| \le 1$ (this constraint can always be satisfied by normalization). Then we need to compute the correction gain vector $L \in \mathbb{R}^{n\times 1}$ that ensures that the system (3.3) trajectories meet the conditions (1) and (2) from Definition 1, where the ellipsoid $\mathcal{E}_e$ matrix has minimal trace $\operatorname{tr} P_e$.

Such problem solution is designed in the following way. First of all, an extended system is considered:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{e}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} A_n + F_A\Delta_A H_A & 0 \\ F_A\Delta_A H_A & A_n - LC^{\mathrm{T}} \end{bmatrix}}_{\tilde{A}} \underbrace{\begin{bmatrix} x(t) \\ e(t) \end{bmatrix}}_{g(t)} + \underbrace{\begin{bmatrix} B_n + F_B\Delta_B H_B \\ F_B\Delta_B H_B \end{bmatrix}}_{\tilde{B}} u(t)\,.$$

Secondarily, its state $g(t)$ is put into an ellipsiod $\mathcal{E}_g$, which is defined by the matrix

$$P = \begin{bmatrix} P_x & 0 \\ 0 & P_e \end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \quad P > 0$$

and the ellipsoid defined by the matrix P_e is minimized. As a result, using techniques based on the Lyapunov second method, Petersen lemma [24], S-procedure and Schur lemma, the following result is obtained.

Theorem 2. *Let $\hat{Q}_e$ and $\hat{Y}$ be the solutions of the minimization problem:*

$$\mathrm{tr}\, H \to \min$$

subjected to the constraints

$$\begin{bmatrix} \Omega_1 & 0 & Q_x B_n & Q_x F_A & Q_x F_B \\ * & \Omega_2 & 0 & Q_e F_A & Q_e F_B \\ * & * & -\alpha I + \varepsilon_2 H_B^{\mathrm{T}} H_B & 0 & 0 \\ * & * & * & -\varepsilon_1 I & 0 \\ * & * & * & * & -\varepsilon_2 I \end{bmatrix} \leq 0, \quad \begin{bmatrix} H & I \\ I & Q_e \end{bmatrix} \geq 0,$$

where

$$\Omega_1 = A_n^{\mathrm{T}} Q_x + Q_x A_n + \alpha Q_x + \varepsilon_1 H_A^{\mathrm{T}} H_A,$$
$$\Omega_2 = A_n^{\mathrm{T}} Q_e + Q_e A_n - Y C^{\mathrm{T}} - C Y^{\mathrm{T}} + \alpha Q_e,$$

with respect to the variables $Q_x = Q_x^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, $Q_e = Q_e^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, $Y \in \mathbb{R}^n$, ε_1, $\varepsilon_2 \in \mathbb{R}$ and scalar parameter $\alpha > 0$.

Then the minimal invariant ellipsoid is defined by the matrix $\hat{P}_e = \hat{Q}_e^{-1}$, and the respective correction gain of the observer is obtained as $L = \hat{Q}_e^{-1}\hat{Y}$.

Proof is given in [25].

According to Theorem 2, the computation of the correction gain vector L for fixed $\alpha > 0$ is reduced to a semi-definite programming problem, which is solved numerically using various software packages.

The open problems of the observers designed by the invariant ellipsoid method include the fundamental inability to meet the goal (2.5) without additional assumptions about the control signal $u(t)$. This problem is caused by the fact that regardless of the choice of the correction gain vector L, in error equation (3.3) there exists a non-vanishing component $(B - B_n)\, u(t) + (A - A_n)\, x(t)$ considered as a perturbation. The second problem is that, in order to minimize the ellipsoid trace, we need to solve optimization problem w.r.t. the parameter $\alpha > 0$ (convexity of the objective function w.r.t. this parameter has not been proved). The third and more significant problem is that the stated problem is solved only when the real uncertainty of the parameters A and B satisfies the chosen parameterization (3.5). If the real uncertainty does not satisfy this parameterization (*i.e.*, the chosen weight matrices F_A, H_A and F_B, H_B do not match the reality), then the observer correction gain no longer provides optimality in the sense of the trace of the invariant ellipsoid. The conservative choice of weight matrices, of course, can solve this problem, but leads to excessive conservatism of the obtained state estimate.

3.2. Dynamic Feedback Based Control

As it was mentioned earlier, the feedback design is almost always out of the scope of this overview. However, in this case the choice of feedback allows one to improve the properties of the solution of state reconstruction problem.

Using dynamic feedback control, it is possible to overcome the first drawback of the observer (3.1), roughly speaking, by ensuring that the summand $(B - B_n)\, u(t) + (A - A_n)\, x(t)$ converges to zero by choosing a feedback control law as $u(t) = -K_r \hat{x}(t)$ and stabilizing the system (2.1). In the presence of parametric uncertainty, the separation principle commonly used for linear systems is no longer valid [6]. This means that, in order to stabilize the system (2.1) via feedback based on the observer state, a joint calculation of the parameters of the observer and the controller is necessary to ensure the stability of the extended system composed of the equations of both the system and the observer. Let us briefly review the application of the Riccati approach and the linear matrix inequalities technique for the joint calculation of the parameters L and K_r.

3.2.1. Riccati approach. The control law for the system (2.1) with uncertainty (3.5) is chosen as a linear dynamic controller (2.4a) that uses the states of the observer (3.1). Then we need to calculate the parameters $L \in \mathbb{R}^{n\times 1}$ and $K_r \in \mathbb{R}^{1\times n}$ in such a way that

$$\lim_{t\to\infty} \|e(t)\| = \lim_{t\to\infty} \|x(t) - \hat{x}(t)\| = 0, \quad \lim_{t\to\infty} \|x(t)\| = 0. \tag{3.9}$$

The solution of this problem on the basis of Riccati approach is obtained in [26] and given below in the form of the following theorem.

Theorem 3. *Let there exist constants $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ such that the following conditions hold:*

i) there exists $P_c = P_c^{\mathrm{T}} > 0$, which satisfies Riccati equation

$$\begin{gathered} A_n^{\mathrm{T}} P_c + P_c A_n + 2H_A^{\mathrm{T}} H_A + \varepsilon_1 Q_1 \\ - P_c \left[\frac{1}{\varepsilon_1} \left(B_n \left(R_1^{-1} - 2R_1^{-1} H_B^{\mathrm{T}} H_B R_1^{-1} \right) B_n^{\mathrm{T}} - 2F_B F_B^{\mathrm{T}} \right) - F_A F_A^{\mathrm{T}} \right] P_c = 0, \end{gathered} \tag{3.10a}$$

where $R_1 = R_1^{\mathrm{T}} > 0$, $Q_1 = Q_1^{\mathrm{T}} > 0$.

ii) there exists $P_o = P_o^{\mathrm{T}} > 0$, which satisfies Riccati equation

$$\begin{gathered} A_n^{\mathrm{T}} P_o + P_o A_n - \frac{1}{\varepsilon_2} C R_2^{-1} C^{\mathrm{T}} + P_o \left(F_A F_A^{\mathrm{T}} + \frac{2}{\varepsilon_1} F_B F_B^{\mathrm{T}} + \varepsilon_2 Q_2 \right) P_o \\ + \frac{2}{\varepsilon_2} P_c B_n R_1^{-1} H_B^{\mathrm{T}} H_B R_1^{-1} B_n^{\mathrm{T}} P_c = 0, \end{gathered} \tag{3.10b}$$

where $R_2 = R_2^{\mathrm{T}} > 0$, $Q_2 = Q_2^{\mathrm{T}} > 0$.

iii) the matrices P_c and P_o satisfy matrix inequality

$$\begin{gathered} \varepsilon_1 \left(Q_1 + \frac{1}{\varepsilon_1^2} P_c B_n R_1^{-1} B_n^{\mathrm{T}} P_c \right) \\ - \frac{1}{\varepsilon_1^2 \varepsilon_2} P_c B_n R_1^{-1} B_n^{\mathrm{T}} P_c \left[P_o Q_2 P_o + \frac{1}{\varepsilon_2^2} C R_2^{-1} C^{\mathrm{T}} \right]^{-1} P_c B_n R_1^{-1} B_n^{\mathrm{T}} P_c > 0. \end{gathered} \tag{3.10c}$$

Then the following choice of $K_r = \frac{1}{\varepsilon_1} R_1^{-1} B_n^{\mathrm{T}} P_c$ and $L = \frac{1}{\varepsilon_2} P_o^{-1} C R_2^{-1}$ ensures that the stated goal (3.9) *is achieved.*

Proof is given in [26].

Equations (3.10a) and (3.10b) are obtained via finding majorants of the uncertainty terms Δ_A, Δ_B, which occur in the derivative of the function $V = x^{\mathrm{T}} P_c x + e^{\mathrm{T}} P_o e$, with the help of the following inequality

$$a^{\mathrm{T}} M F N b \le a^{\mathrm{T}} M M^{\mathrm{T}} a + b^{\mathrm{T}} N^{\mathrm{T}} N b,$$

which holds for such F that $F^{\mathrm{T}} F \le I$, and any $a, b \in \mathbb{R}^n$, $M \in \mathbb{R}^{n\times p}$, $N \in \mathbb{R}^{p\times n}$.

Inequality (3.10c) is obtained via application of Schur lemma and means positive definiteness of the matrix Ω that defines the derivative $\dot{V} \le -\begin{bmatrix} x \\ e \end{bmatrix}^{\mathrm{T}} \Omega \begin{bmatrix} x & e \end{bmatrix}$.

3.2.2. Linear matrix inequalities technique. Application of the LMI [27] technique to solve the dynamic feedback control problem requires, in addition to L and K_r calculation, to compute the observer state matrix. Therefore, equation (3.1) is rewritten in the form of

$$\dot{\hat{x}}(t) = G\hat{x}(t) + B_n u(t) + L\left(y(t) - C^{\mathrm{T}}\hat{x}(t)\right), \quad \hat{x}(t) = \hat{x}_0.$$

It is required to calculate the parameters $G \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times 1}$ and $K_r \in \mathbb{R}^{1\times n}$ such that the stated goal (3.9) is achieved. Using techniques based on the Lyapunov second method, Petersen and Schur lemmas, the following result is presented in [28].

Theorem 4. *Let $\hat{Q}_{\hat{x}}$, $\hat{Y}_1$, $\hat{Q}_e$, $\hat{Y}_2$ and $\hat{\alpha}$, $\hat{Y}_3$ be a solution of the following LMI:*

$$\begin{bmatrix} \Omega_{11} & \Omega_{12} & Q_{\hat{x}}H_A^{\mathrm{T}} & -Y_1^{\mathrm{T}}H_B^{\mathrm{T}} \\ * & \Omega_{22} & Q_e H_A^{\mathrm{T}} & 0 \\ * & * & -\varepsilon_1 I & 0 \\ * & * & * & -\varepsilon_2 I \end{bmatrix} < 0$$

under the condition

$$C^{\mathrm{T}}Q_e = \alpha C^{\mathrm{T}},$$

where

$$\begin{aligned} \Omega_{11} &= Y_3 + Y_3^{\mathrm{T}} - B_n Y_1 - Y_1^{\mathrm{T}}B_n^{\mathrm{T}}, \\ \Omega_{12} &= Y_2 C^{\mathrm{T}} + Q_{\hat{x}}A_n^{\mathrm{T}} - Y_3^{\mathrm{T}}, \\ \Omega_{22} &= A_n Q_e + Q_e A_n^{\mathrm{T}} - Y_2 C^{\mathrm{T}} - C Y_2^{\mathrm{T}} + \varepsilon_1 F_A F_A^{\mathrm{T}} + \varepsilon_2 F_B F_B^{\mathrm{T}} \end{aligned}$$

with respect to $Q_{\hat{x}} = Q_{\hat{x}}^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, $Q_e = Q_e^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, $Y_1 \in R^{1\times n}$, $Y_2 \in \mathbb{R}^n$, $Y_3 \in \mathbb{R}^{n\times n}$, $\varepsilon_1, \varepsilon_2, \alpha > 0$.

Then the parameters $K_r = \hat{Y}_1\hat{Q}_{\hat{x}}^{-1}$, $G = \hat{Y}_3\hat{Q}_{\hat{x}}^{-1}$ and $L = \hat{Y}_2\hat{\alpha}^{-1}$ ensure that the stated goal (3.9) *is achieved.*

Proof is given in [28].

The common open problems of the Riccati and LMI approaches include the necessity to choose a control law in the form of $u(t) = -K_r\hat{x}(t)$ and the need of joint calculation of the observer and controller parameters.

3.3. Conclusions on Overview of Robust Observers

A distinctive feature of all considered robust approaches to the state observers design is the assumption that the parametric uncertainty of the system meets the parameterizations (3.5) or (3.6). In practice, this assumption requires the range of variation of all unknown parameters to be known. Excessive conservatism in such ranges choice can lead to high values of the feedback parameters, and, as a consequence, to non-implementable solutions.

4. STATE ESTIMATION BASED ON INVARIANCE THEORY

Invariant observers are designed on the basis of idea of elimination (whether algebraic or high-gain one) of the parametric uncertainty effect on the state estimate [7, 29–31]. In the first step of such solutions design, equation (2.1) is rewritten in the following form:

$$\begin{gathered} \dot{x}(t) = (A_n + \Delta A)x(t) + (B_n + \Delta B)u(t) = A_n x(t) + B_n u(t) + Dw(x,u), \\ Dw(x,u) = \Delta A x(t) + \Delta B u(t), \end{gathered} \tag{4.1}$$

where ΔA, ΔB are the aggregated parametric uncertainties of the corresponding dimensions, $D \in \mathbb{R}^{n\times m}$ is a known vector that assigns gains to the uncertainties and allocates them to the equations, $w(x,\ u) \in \mathbb{R}^m$ is the unknown bounded (by Assumption 1) perturbation $\|w(x,\ u)\| \le w_{\max}$.

The stated goal (2.5) is now interpreted as the reconstruction of the state of the system (2.1) that has an unmeasured (unknown) input $w(x,\ u)$. Let us consider the solutions of this problem based on the Luenberger observer for systems with unknown input and various sliding mode observers.

4.1. Luenberger Observer for Systems with Unknown Input

In accordance with the results [32, 33], it is proposed to obtain the estimate of unmeasurable states by means of the following set of equations:

$$\begin{aligned} \dot{z}(t) &= Nz(t) + Gu(t) + Ly(t), \\ \hat{x}(t) &= z(t) - Ey(t), \end{aligned} \tag{4.2}$$

where $z(t) \in \mathbb{R}^n$ and matrices of corresponding dimensions are chosen as follows:

$$\begin{gathered} \left(I + EC^{\mathrm{T}}\right) D = 0, \\ N = \left(I + EC^{\mathrm{T}}\right) A_n - (L + NE)\, C^{\mathrm{T}}, \\ \left(I + EC^{\mathrm{T}}\right) B_n = G, \end{gathered} \tag{4.3}$$

and the vector L makes the matrix N be Hurwitz one.

Owing to (4.2), the following holds for the state reconstruction error $\tilde{x}(t) = \hat{x}(t) - x(t)$:

$$\tilde{x}(t) = z(t) - EC^{\mathrm{T}}x(t) - x(t) = z(t) - \left(I + EC^{\mathrm{T}}\right) x(t),$$

and then the error equation is written as:

$$\begin{gathered} \dot{\tilde{x}}(t) = \dot{z}(t) - \left(I + EC^{\mathrm{T}}\right)\dot{x}(t) \\ = Nz(t) + Gu(t) + Ly(t) - \left(I + EC^{\mathrm{T}}\right)\left(A_n x(t) + B_n u(t) + Dw(x,u)\right) \\ = Nz(t) + Ly(t) - \left(I + EC^{\mathrm{T}}\right) A_n x(t) \\ = Nz(t) + LC^{\mathrm{T}}x(t) - Nx(t) - (L + NE)\, C^{\mathrm{T}}x(t) \\ = Nz(t) - Nx(t) - NEC^{\mathrm{T}}x(t) = N\tilde{x}(t). \end{gathered} \tag{4.4}$$

Considering equation (4.4), it is easy to conclude that the invariance of the estimation error $\tilde{x}(t)$ with respect to the parametric uncertainty $w(x,\ u)$ is achieved because of its algebraic elimination from the error equation. Necessary and sufficient conditions to make equations (4.3) solvable and the matrix N be Hurwitz one have been derived and revised in [32] and [33], respectively. For the considered case of the class of single-input-single-output systems, these conditions are formulated as follows.

Theorem 5. *Set of equations* (4.3) *has a solution if and only if the following conditions hold:*
1) $m = \dim(y(t)) = 1$,
2) $\operatorname{rank}\left(C^{\mathrm{T}}D\right) = \operatorname{rank}(D)$,
3) *the invariant zeros of the triple* $\left(A_n,\ D,\ C^{\mathrm{T}}\right)$ *are stable (have a negative real part).*

Proof is given in [32, 33].

The first two premises of Theorem 5 are restrictive for applications. In particular, the first condition is necessary (but not sufficient [34]) to meet the second one and does not allow two functionally different uncertainties to be in different equations of the system. For example, the observer (4.2) cannot be applied to the following second-order system:

$$\begin{aligned}\dot{x}(t) &= \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}x(t) + \begin{bmatrix}0\\ 1\end{bmatrix}u(t) + \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}a_1x_1(t)\\ a_2x_1(t) + a_3x_2(t)\end{bmatrix},\\ y(t) &= \begin{bmatrix}1 & 0\end{bmatrix}x(t),\end{aligned} \tag{4.5}$$

where $m = 2 > \dim(y(t)) = 1$, which contradicts the first condition.

In its turn, the second condition or output matching condition requires a perturbation to affect the equation, which state is measurable. For example, the observer (4.2) is not applicable to the following second-order system:

$$\begin{aligned}\dot{x}(t) &= \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}x(t) + \begin{bmatrix}0\\ 1\end{bmatrix}u(t) + \begin{bmatrix}0\\ 1\end{bmatrix}(a_2x_1(t) + a_3x(t) + b_1u),\\ y(t) &= \begin{bmatrix}1 & 0\end{bmatrix}x(t),\end{aligned} \tag{4.6}$$

where $\operatorname{rank}\left(C^{\mathrm{T}}D\right) = 0 \neq \operatorname{rank}(D) = 1$, which contradicts the second condition.

The independence of the estimation error from an arbitrary external perturbation is the main competitive advantage of the considered Luenberger observer in comparison with the robust alternatives (3.1) and (3.7). However, strict structural constraints do not allow one to use the considered observer for state reconstruction of a wide class of real technical systems with output unmatched uncertainty.

4.2. Sliding Mode Observers. Output Matching Conditions

The invariance of the state reconstruction error with respect to parametric uncertainty can be achieved not only by its algebraic elimination, but also using special sliding modes [7, 29–31]. The first observer based on sliding modes was proposed by V.I. Utkin [35], later this solution was improved in [36] with the help of additional Luenberger correction term. In [37], an alternative way to choose the sliding surface was proposed, and necessary and sufficient conditions for the existence of all three types of sliding mode observers [35–37] were obtained.

According to an algorithm from [37], the following observer is designed for a system represented as (4.1):

$$\begin{aligned}\dot{\hat{x}}(t) &= A_n\hat{x}(t) + B_nu(t) - G_l(\hat{y}(t) - y(t)) + G_nv(t),\\ \hat{y}(t) &= C^{\mathrm{T}}\hat{x}(t),\end{aligned} \tag{4.7}$$

where

$$G_l = T^{-1}\begin{bmatrix}F_{12}\\ F_{22} - F_m\end{bmatrix} \in \mathbb{R}^n, \quad G_n = T^{-1}\begin{bmatrix}0_{n-1}\\ 1\end{bmatrix} \in \mathbb{R}^n,$$

$$v(t) := \begin{cases} -\rho\dfrac{P(\hat{y}(t) - y(t))}{\|P(\hat{y}(t) - y(t))\|}, & \text{if} \quad \hat{y}(t) - y(t) \neq 0,\\ 0 & \text{otherwise},\end{cases}$$

and F_m is a Hurwitz matrix, P stands for a solution of Luapunov equation for F_m, $\rho > 0$ denotes a sufficiently large scalar, $T \in \mathbb{R}^{n\times n}$ is full-rank similarity transformation matrix, F_{12}, F_{22} stand for known matrices, which definition algorithm will become clear from further explanations.

According to [37, 38], the observer (4.7) is implementable and ensures convergence of the state reconstruction error (2.5) when the existence conditions of the following structural transformations are satisfied.

Lemma 1. *Let the following conditions be met:*

1) $m = \dim(y(t)) = 1$,
2) $\operatorname{rank}\left(C^{\mathrm{T}} D\right) = \operatorname{rank}(D)$,
3) *the invariant zeros of the triple* $\left(A_n, D, C^{\mathrm{T}}\right)$ *are stable.*

Then there exists a nonsingular coordinate transformation $\begin{pmatrix} z_1(t) \\ y(t) \end{pmatrix} = Tx(t)$, *which represents the system* (4.1) *and observer* (4.7) *in a block form:*

$$\begin{aligned} \dot{z}_1(t) &= F_{11} z_1(t) + F_{12} y(t) + G_1 u(t), \\ \dot{y}(t) &= F_{21} z_1(t) + F_{22} y(t) + G_2 u(t) + D_2 w(x,u), \\ \dot{\hat{z}}_1(t) &= F_{11} \hat{z}_1(t) + F_{12} \hat{y}(t) + G_1 u(t) - F_{12} e_y(t), \\ \dot{\hat{y}}(t) &= F_{21} \hat{z}_1(t) + F_{22} \hat{y}(t) + G_2 u(t) - (F_{22} - F_m) e_y(t) + v(t), \end{aligned} \tag{4.8}$$

where $D_2 > 0$, F_{11} *is a Hurwitz matrix according to the design,* $e_y(t) = \hat{y}(t) - y(t)$ *is a error of system output observation.*

Proof is given in [37, 38].

The detailed algorithm to obtain the transformation matrix T is given in [7, 29–31, 35, 37] and is reduced to elimination of the external perturbation from the equation related to the unmeasured variable. The error equation between the system and observer dynamics represented via new coordinates (4.8) has the form:

$$\begin{aligned} \dot{\tilde{z}}_1(t) &= F_{11} \tilde{z}_1(t), \\ \dot{e}_y(t) &= F_{21} \tilde{z}_1(t) + F_m e_y(t) + v(t) - D_2 w(x,u). \end{aligned} \tag{4.9}$$

The fact that F_{11} is a Hurwitz matrix leads to the exponential convergence of the error $\tilde{z}_1(t)$ to zero. Using the Lyapunov second method, the error $e_y(t)$ is shown to reach the line $e_y(t) = 0$ in finite time. By the method of equivalent control $\mu^{-1}\dot{v}_{eq}(t) = v(t) - v_{eq}(t)$, $\mu > 0$, it is also shown that $D_2^{-1} v_{eq}(t) \to w(x, u)$. Since T is a nonsingular transformation, then, as the errors (4.9) are exponentially stable, we immediately have that the goal (2.5) is achieved.

Thus, the sliding mode observer (4.7) requires to meet the same strict structural constraints as the perturbation invariant Luenberger observer (4.2). However, in contrast to the solution (4.2), the observer (4.7) uses another technique to eliminate the external perturbation, which is based on the robust properties of the motion on the sliding surface. A more detailed comparison of the properties of these two types of invariant observers are given in [39, 40].

4.3. Sliding Mode Observers. Extended Matching Conditions

The output matching conditions $\operatorname{rank}\left(C^{\mathrm{T}} D\right) = \operatorname{rank}(D)$ significantly restrict the class of systems, which state can be reconstructed with the help of (4.2) and (4.7) observers. It is possible to relax this requirement by increasing the dimensionality of the output $y(t)$ via addition of new physical or virtual measurements to the existing ones. New physical measurements requires to install additional sensors, which contradicts the theoretical formulation of the problem, and, considering practical scenarios, is often impractical or impossible. So, the number of measurements can be increased only by means of virtual states. In [41], it is proposed to introduce into consideration a virtual output variable

$$\xi(t) = \mathcal{O}^{-1} x(t) = \begin{bmatrix} C^{\mathrm{T}} \\ \vdots \\ C^{\mathrm{T}} A_n^{n-1} \end{bmatrix} x(t) \tag{4.10}$$

and design the following state observer:

$$\begin{aligned} \dot{\hat{x}}(t) &= A_n \hat{x}(t) + B_n u(t) - G_l \left(\hat{y}_a(t) - \hat{\xi}(t) \right) + G_n v(t), \\ \hat{y}_a(t) &= \mathcal{O}^{-1} \hat{x}(t), \end{aligned} \tag{4.11}$$

where $\hat{y}_a(t)$ is an estimate of the output (4.10) with the help of (4.11), $\hat{\xi}(t)$ denotes the estimate of the output (4.10) with the help of an observer, which will be defined further, and the matrices G_l, G_n and function $v(t)$ are chosen like in (4.7).

According to the results of [41, 42], the observer (4.11) is implementable and ensures convergence of the state reconstruction error in case the following conditions for the system that is reducible to block and triangular forms are satisfied.

Lemma 2. *Let the following conditions be met:*

1) $m = \dim(y(t)) = 1$,
2) $\operatorname{rank}(\mathcal{O}^{-1} D) = \operatorname{rank}(D)$,
3) $\begin{bmatrix} C^{\mathrm{T}} \\ \vdots \\ C^{\mathrm{T}} A_n^{n-2} \end{bmatrix} D = 0,$
4) *the invariant zeros of the triple* $\left(A_n, D, C^{\mathrm{T}}\right)$ *are stable.*

Then:

a) *there exists* [41] *a nonsingular transformation* $\begin{pmatrix} z_1(t) \\ y_a(t) \end{pmatrix} = T x(t)$, *which represents the system* (4.1) *with the output* $\xi(t)$ *and the observer* (4.11) *in a block form* (4.8) *up to a change of* $y(t)$ *and* $\hat{y}(t)$ *by* $\xi(t)$ *and* $\hat{y}_a(t)$, *respectively;*

b) *there exists* [42] *a nonsingular transformation* (4.10), *which represents the system* (4.1) *in a triangular form:*

$$\begin{aligned} \dot{\xi}(t) &= A_0 \xi(t) + B_0 \delta(t) + B_e u(t), \\ y(t) &= C_0^{\mathrm{T}} \xi(t), \end{aligned} \tag{4.12}$$

where

$$A_0 = \begin{bmatrix} 0_n & \begin{matrix} I_{n-1} \\ 0_{1\times(n-1)} \end{matrix} \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0_{n-1} \\ 1 \end{bmatrix}, \quad B_e = \begin{bmatrix} C^{\mathrm{T}} B_n \\ \vdots \\ C^{\mathrm{T}} A_n^{n-2} B_n \\ C^{\mathrm{T}} A_n^{n-1} B_n \end{bmatrix},$$

$$\delta(t) = C^{\mathrm{T}} A_n^n x(t) + C^{\mathrm{T}} A_n^{n-1} D w(t), \quad C_0^{\mathrm{T}} = \begin{bmatrix} 1 & 0_{n-1}^{\mathrm{T}} \end{bmatrix}.$$

Proof is given in [41, 42].

The implementation of the observer (4.11) requires to obtain an estimate $\hat{\xi}(t)$ of the output (4.10), which is simultaneously the state of the triangular system (4.12). In [41], a cascade observer is proposed to estimate the states of the system (4.12) in finite time:

$$\left\{\begin{array}{l} \dot{\hat{\zeta}}_1(t) = v\left(y(t) - \hat{\zeta}_1(t)\right) \\ \dot{\hat{\zeta}}_2(t) = E_1 v\left(\tilde{\zeta}_2(t) - \hat{\zeta}_2(t)\right) \\ \vdots \\ \dot{\hat{\zeta}}_{n-1}(t) = E_{n-2} v\left(\tilde{\zeta}_{n-1}(t) - \hat{\zeta}_{n-1}(t)\right) \\ \dot{\hat{\zeta}}_n(t) = E_{n-1} v\left(\tilde{\zeta}_n(t) - \hat{\zeta}_n(t)\right) \end{array}\right. + B_e u(t), \tag{4.13}$$

$$\hat{\xi}(t) = \begin{bmatrix} y(t) \\ v\left(\tilde{\zeta}_2(t) - \hat{\zeta}_2(t)\right) \\ \vdots \\ v\left(\tilde{\zeta}_{n-1}(t) - \hat{\zeta}_{n-1}(t)\right) \end{bmatrix},$$

where $\tilde{\zeta}_1(t) = y(t)$ and

$$\begin{gathered} \tilde{\zeta}_i(t) = v\left(\tilde{\zeta}_{i-1}(t) - \hat{\zeta}_{i-1}(t)\right), \quad 2 \le i \le n, \\ E_j = \begin{cases} 1, & \text{if } \left|\tilde{\zeta}_i(t) - \hat{\zeta}_i(t)\right| \le \varepsilon \\ 0 & \text{otherwise}, \end{cases} \qquad j \le i, \\ v(.) = \varphi(t) + \lambda_s |.|^{\frac{1}{2}} \operatorname{sgn}(.), \quad \lambda_s > 0, \\ \dot{\varphi}(t) = \alpha_s \operatorname{sgn}(.), \quad \alpha_s > 0. \end{gathered}$$

The error equation between the systems (4.12) and (4.13) is written as:

$$\left\{\begin{array}{l} \dot{\bar{\zeta}}_1(t) = \xi_2(t) - v\left(y(t) - \hat{\zeta}_1(t)\right) \\ \dot{\bar{\zeta}}_2(t) = \xi_3(t) - E_1 v\left(\tilde{\zeta}_2(t) - \hat{\zeta}_2(t)\right) \\ \vdots \\ \dot{\bar{\zeta}}_{n-1}(t) = \xi_n(t) - E_{n-2} v\left(\tilde{\zeta}_{n-1}(t) - \hat{\zeta}_{n-1}(t)\right) \\ \dot{\bar{\zeta}}_n(t) = \delta(t) - E_{n-1} v\left(\tilde{\zeta}_n(t) - \hat{\zeta}_n(t)\right), \end{array}\right. \tag{4.14}$$

and, in case of sufficiently large values of λ_s, α_s, the sliding modes emerge one by one in (4.14) at $\tilde{\zeta}_i(t) - \hat{\zeta}_i(t) = 0$, $i = \overline{1, n}$, which results in convergence of the error $\hat{\xi}(t) - \xi(t)$ in finite time T_ξ. When the estimation process of the output (4.10) of the system (4.1) is over, a sliding mode occurs at $\hat{y}_a(t) - \hat{\xi}(t) = 0$ in finite time $T_x \ge T_\xi$, which, in accordance with the analysis (4.9), ensures that the stated goal (2.5) is achieved.

According to the results of [41], when the new output (4.10) is measured, even if the matching condition $\operatorname{rank}\left(C^{\mathrm{T}} D\right) = \operatorname{rank}(D)$ is not satisfied, the extended matching condition 2)–3) of Lemma 2 can be met, and the observer (4.11) is implementable and ensures the convergence of the state reconstruction error.

For example, we have for (4.6) that:

$$\begin{bmatrix} C^{\mathrm{T}} \\ C^{\mathrm{T}}A \end{bmatrix} D = D \Rightarrow \operatorname{rank}\left(C_a^{\mathrm{T}} D\right) = \operatorname{rank}\left(D\right),$$

which allows one to reconstruct the state of the system (4.6) with the help of (4.11) + (4.13).[2]

Alternative ways to reconstruct the states of triangular systems were discussed in detail in [43, 44] and, by analogy with (4.13), they can be used to obtain an estimate of the output variable $\xi\left(t\right)$ under the same structural constraints from Lemma 2. Also it should be noted that, as the pair $\left(A,\, C^{\mathrm{T}}\right)$ is observable, the states $x\left(t\right)$ can be estimated in finite time without implementation of (4.11) but using $\hat{x}\left(t\right) = \mathcal{O}\hat{\xi}\left(t\right)$.

If the extended matching conditions 2)-3) from Lemma 2 are satisfied, the system is reduced to a triangular form just after the first choice of the virtual output (4.10), which allows one to use observers of the type (4.11) + (4.13) or (4.11) + $\hat{x}\left(t\right) = \mathcal{O}\hat{\xi}\left(t\right)$ to reconstruct the states. For high-dimensional systems with multiple inputs/outputs, an introduction of a single virtual output may be not enough to reduce the whole system to a triangular form [11, 45]. In this case, the transformations are continued until the whole system is divided into triangularly shaped blocks. The system equations split obtained as a result of such iterative procedure is called the quasi block triangular observable form (QBTOF), about which, strictly speaking, two facts are known:

(i) states and perturbations of the system are observable by measurements if and only if the system is reducible to QBTOF [11],

(ii) the system is reducible to QBTOF in case the matching conditions 2) from Theorem 5 and 3) from Lemma 2 are not satisfied [45].

Therefore, the sliding mode observers from [11, 45], designed on the basis of the QBTOF, have the weakest structural constraints in comparison with other invariant state observers (e.g., (4.7), (4.11) + (4.13)). However, reduction of the system to the QBTOF is a rather laborious iterative process, and some of the structural constraints of this form of system representation seem not to be relaxable either. For example, following step 1(a) of the constructive observability criterion [11, Section 3.2], let us write down the previously considered system (4.5) in the following form (using the authors' notation):

$$\begin{aligned} \dot{y}\left(t\right) &= A_{11}y\left(t\right) + D_1x_1\left(t\right) + Q_1\psi\left(t\right) + B_1u\left(t\right), \\ \dot{x}_1\left(t\right) &= A_{x_{11}}y\left(t\right) + A_{x_1}x_1\left(t\right) + Q_{x_1}\psi\left(t\right) + B_{x_1}u\left(t\right), \end{aligned} \tag{4.15}$$

where in terms of the system (4.5) we adopt the following re-definitions:

$$y_1\left(t\right) := y\left(t\right) = x_1\left(t\right), \quad x_1\left(t\right) := x_2\left(t\right), \quad \psi\left(t\right) := \begin{bmatrix} \psi_1\left(t\right) \\ \psi_2\left(t\right) \end{bmatrix} = w\left(x, u\right),$$
$$A_{11} = 0, \quad D_1 = 1, \quad Q_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad B_1 = 0,$$
$$A_{x_{11}} = A_{x_1} = 0, \quad B_{x_1} = 1, \quad Q_{x_1} = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

In order to reduce the system (4.5), (4.15) into QBTOF, the following condition (equation (**) from [11]) is to be met:

$$\operatorname{rank}\left(Q_1\right) - \operatorname{rank}\left(\begin{bmatrix} Q_1 & D_1 \end{bmatrix}\right) \neq 0.$$

As for the considered case (4.5), we have that $\operatorname{rank}\left(Q_1\right) - \operatorname{rank}\left(\begin{bmatrix} Q_1 & D_1 \end{bmatrix}\right) = 0$, which does not allow one to reduce the system to QBTOF, and hence, owing to the design criterion from [11],

[2] It holds for (4.6) that $\xi\left(t\right) = x\left(t\right)$, so the implementation of (4.11) is not really required.

the states and perturbations of the system (4.5) are unobservable. In order to reduce the system to QBTOF, the inequality $m \leq \dim(y(t))$ is required to be satisfied, which for the class of linear systems (2.1) with one output that is considered in this paper is almost always not met.

Thus, all existing invariant state observers based on sliding mode technique [11, 35–41, 43–45], including recent results [75, 76], which allow one to improve the quality of transients for the state estimates, impose strict structural requirements on a dynamical system with a single output. Moreover, as shown by the constructive observability criterion [11, 45], some of these conditions for single-output systems cannot even potentially be relaxed for a class of sliding mode observers. Indeed, let us introduce an observer of the following form for the system (4.5), (4.15):

$$\begin{aligned}\dot{\hat{y}}_1(t) &= A_{11}\hat{y}_1(t) + D_1\hat{x}_1 + B_1 u + L(\hat{y}_1(t) - y_1(t)) - \rho\,\mathrm{sgn}(\hat{y}_1(t) - y_1(t)),\\ \dot{\hat{x}}_1(t) &= A_{x_{11}}\hat{y}_1(t) + A_{x_1}\hat{x}_1 + B_{x_1}u - K_{x_1}\mathrm{sgn}(\hat{y}_1(t) - y_1(t)),\end{aligned}$$

which allows one to obtain the following error equation:

$$\begin{aligned}\dot{\tilde{y}}_1(t) &= (A_{11} + L)\tilde{y}_1(t) + D_1\tilde{x}_1(t) - Q_1\psi(t) - v(t),\\ \dot{\tilde{x}}_1(t) &= A_{x_{11}}\tilde{y}_1(t) + A_{x_1}\tilde{x}_1(t) - Q_{x_1}\psi(t) - \frac{K_{x_1}}{\rho}v(t),\\ v(t) &= \rho\,\mathrm{sgn}(\tilde{y}_1(t)), \quad \tilde{y}_1(t) = \hat{y}_1(t) - y_1(t).\end{aligned}$$

In case $\rho > 0$ is sufficiently large, a sliding mode occurs in the first subsystem at $\tilde{y}_1(t) = 0$, which, using the method of equivalent control, allows one to obtain:

$$\begin{aligned}0 &= D_1\tilde{x}_1(t) - Q_1\psi(t) - v_{eq}(t),\\ \dot{\tilde{x}}_1(t) &= \left(A_{x_1} - \frac{K_{x_1}}{\rho}D_1\right)\tilde{x}_1(t) - Q_{x_1}\psi(t) + \frac{K_{x_1}}{\rho}Q_1\psi(t)\\ &= \left(A_{x_1} - \frac{K_{x_1}}{\rho}D_1\right)\tilde{x}_1(t) - \psi_2(t) + \frac{K_{x_1}}{\rho}\psi_1(t),\end{aligned}$$

from which, in general case, we have only boundedness of the error $\tilde{x}_1(t)$, while the stated goal (2.5) can be achieved only if the condition $\psi_2(t) - \frac{K_{x_1}}{\rho}\psi_1(t) = 0$ holds, which is not satisfied for the example under consideration (4.5), (4.15), as for all $K_{x1} \in \mathbb{R}$ it holds that $a_2 x_1(t) + a_3 x_2(t) - \frac{K_{x1}}{\rho}a_1 x_1(t) \neq 0$.

Considering real technical systems, different parametric perturbations can potentially affect equations of the system, *i.e.*, the inequality $m > 1$ is often satisfied for the parameterization (4.1), which restricts the potential of sliding mode observers application to the problem of state reconstruction of linear systems with parametric uncertainty.

4.4. Conclusions on Overview of Observers Based on Invariance Theory

Owing to algebraic uncertainty elimination or sliding modes, the invariant observers provide insensitivity of the state estimation error to the parametric uncertainty of the system. This property allows one to achieve the goal (2.5) without additional assumptions regarding the parameterization of uncertainty and the properties of the signal $u(t)$.

The open problems for invariant observers are the need to use discontinuous correction signals that are sensitive to measurement noise and to fulfill strict output matching conditions.

5. ADAPTIVE OBSERVERS

In contrast to robust state reconstruction methods, adaptive algorithms theoretically do not require any *a priori* information about the system (2.1) parameters, because they simultaneously estimate unknown parameters and reconstruct state vector. Compared to observers that are invariant to perturbations, adaptive algorithms, by means of special parameterizations, are able to take into account the structure of the perturbation that affects the system, which, in many cases, makes it possible not to require to meet output matching conditions.

It is logical (but naive) to choose the structure of an adaptive state observer as follows:

$$\dot{\hat{x}}(t) = \hat{A}(t)\hat{x}(t) + \hat{B}(t)u(t) + \hat{L}(t)(\hat{y}(t) - y(t)), \quad \hat{x}(t_0) = \hat{x}_0, \tag{5.1}$$

where all unknown matrices of the system are substituted by their dynamic estimates.

Then the error equation between (2.1) and (5.1) is written as:

$$\begin{aligned} \dot{\tilde{x}}(t) &= \hat{A}(t)\hat{x}(t) + \hat{B}(t)u(t) + \hat{L}(\hat{y}(t) - y(t)) - Ax(t) - Bu(t) \pm L(\hat{y}(t) - y(t)) \pm A\hat{x}(t) \\ &= A\hat{x}(t) + \tilde{B}(t)u(t) + \tilde{L}\tilde{y}(t) + \tilde{A}(t)\hat{x}(t) - Ax(t) + LC^{\mathrm{T}}\tilde{x}(t) \\ &= \left(A + LC^{\mathrm{T}}\right)\tilde{x}(t) + \tilde{B}(t)u(t) + \tilde{L}(t)\tilde{y}(t) + \tilde{A}(t)\hat{x}(t), \\ \tilde{y}(t) &= C^{\mathrm{T}}\tilde{x}(t), \end{aligned} \tag{5.2}$$

where $\tilde{A}(t)$, $\tilde{B}(t)$ are the estimation errors of the system (2.1) parameters, $\tilde{L}(t)$ stands for the estimation error of L, which ensures that $\lambda_{\max}\left(A + LC^{\mathrm{T}}\right) < 0$.

As it follows from the differential equation (5.2), there are two principal mechanisms to achieve the goal (2.5). First, one can try to obtain estimates in such a way that the sum $\tilde{B}(t)u(t) + \tilde{L}(t)\tilde{y}(t) + \tilde{A}(t)\hat{x}(t)$ converges to zero. The second way is to ensure convergence to zero of the errors $\tilde{A}(t)$, $\tilde{B}(t)$ and $\tilde{L}(t)$.

Traditionally, choosing the first mechanism, adaptive laws are designed using the Lyapunov second method. According to this approach, a quadratic form is introduced into consideration:

$$V = \tilde{x}^{\mathrm{T}}P\tilde{x} + tr\left(\tilde{A}^{\mathrm{T}}\tilde{A}\right) + \tilde{B}^{\mathrm{T}}\tilde{B} + \tilde{L}^{\mathrm{T}}\tilde{L}, \tag{5.3}$$

where $P = P^{\mathrm{T}} > 0$ is a solution of the Lyapunov equation for $Q = Q^{\mathrm{T}} > 0$.

By easy but lengthy transformations, the derivative of the function (5.3) is obtained as:

$$\dot{V} = -\tilde{x}^{\mathrm{T}}Q\tilde{x} + 2\mathrm{tr}\left(\tilde{A}^{\mathrm{T}}P\tilde{x}\hat{x}^{\mathrm{T}} + \tilde{A}^{\mathrm{T}}\dot{\hat{A}}\right) + 2\tilde{B}^{\mathrm{T}}P\tilde{x}u + 2\tilde{B}^{\mathrm{T}}\dot{\hat{B}} + 2\tilde{L}^{\mathrm{T}}P\tilde{x}\tilde{y} + 2\tilde{L}^{\mathrm{T}}\dot{\hat{L}}. \tag{5.4}$$

Then the following adaptive laws:

$$\dot{\hat{L}}(t) = -P\tilde{x}(t)\tilde{y}(t), \quad \dot{\hat{B}}(t) = -P\tilde{x}(t)u(t), \quad \dot{\hat{A}}(t) = -P\tilde{x}(t)\hat{x}^{\mathrm{T}}(t) \tag{5.5}$$

ensure $\dot{V} = -\tilde{x}^{\mathrm{T}}Q\tilde{x}$, which, with the help of Barbalat lemma, allows one to prove the asymptotic convergence of the estimation error $\tilde{x}(t)$. However, the adaptive laws (5.5) use an unmeasurable estimation error and, hence, cannot be implemented.

The adaptive laws that use the second mechanism of error equation stabilization are designed with the help of the gradient method, and therefore, they require parameterization of the regression equation based on the measured signals $u(t)$ and $y(t)$ with respect to $n^2 + n$ parameters of the matrices A, B. Let us show that such a parameterization cannot be obtained in the general case.

A transfer function from control $u(t)$ to system output $y(t)$ has the following form ($m \leq n - 1$):

$$W(s) = C^{\mathrm{T}}(sI_n - A)^{-1}B = \frac{b_m s^m + b_{m-1}s^{m-1} + \ldots + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_0}, \tag{5.6}$$

that is, only $2n \le n^2 + n$ parameters are associated with the measured signals $u\,(t)$ and $y\,(t)$ (and their derivatives).

Let us rewrite equation (5.6) in the following form (here s is the differentiation operator):

$$\left(b_m s^m + b_{m-1} s^{m-1} + \ldots + b_0\right) u\,(t) = \left(s^n + a_{n-1} s^{n-1} + \ldots + a_0\right) y\,(t)\,. \tag{5.7}$$

Having applied the operator $\frac{1}{\Lambda(s)}$ ($\Lambda\,(s)$ is a monic Hurwitz polynomial of order n) to the left- and right-hand sides of equation (5.7) and expressed $\frac{s^n}{\Lambda(s)} y\,(t)$ from the obtained equation, the following parameterization is written:

$$\begin{gathered} z\,(t) = \frac{s^n}{\Lambda\,(s)} y\,(t) = \varphi^{\mathrm{T}}\,(t)\,\psi, \\ \varphi\,(t) = \begin{bmatrix} -\dfrac{\alpha_{n-1}^{\mathrm{T}}(s)}{\Lambda\,(s)}\, y\,(t) & \dfrac{\alpha_{n-1}^{\mathrm{T}}(s)}{\Lambda\,(s)}\, u\,(t) \end{bmatrix}^{\mathrm{T}}, \quad \alpha_{n-1}^{\mathrm{T}}\,(s) = \begin{bmatrix} s^{n-1} & \cdots & s & 1 \end{bmatrix}, \\ \psi = \begin{bmatrix} a_{n-1} & a_{n-2} & \cdots & b_0 \end{bmatrix}, \end{gathered} \tag{5.8}$$

which relates the signals $u\,(t)$ and $y\,(t)$ to the parameters of the transfer function (5.6).

The signals $\varphi\,(t)$ and $z\,(t)$ are known (can be calculated via $u\,(t)$ and $y\,(t)$), and thus equation (5.8) can be used to design the estimation laws for the parameters ψ using the gradient method. However, since in case $n > 1$ an infinite number of triplets $(A,\, B,\, C)$ is associated with one transfer function (5.6), then in general case the parameters of matrices $(A,\, B)$ of arbitrary state space representation (2.1) cannot be uniquely estimated using information only about $u\,(t)$ and $y\,(t)$. We can obtain only parameters of polynomials of the transfer function (5.8) numerator and denominator, which are directly related to the measured signals and their derivatives. Therefore, in the general case, it is also impossible to design adaptive laws using the second mechanism of stabilization of the reconstruction error $\tilde{x}\,(t)$ equation.

Thus, regardless of the chosen mechanism to stabilize the error equation (5.2), in general case the adaptive observer (5.1) cannot be implemented, and it is required to put forward some additional conditions that allow one to consider narrower classes of systems, for which the problem of state estimation under parametric uncertainty can be solved.

Two types of conditions are distinguished in the existing literature. The first one describes the class of systems, for which the application of the Lyapunov second method (similar to (5.3)–(5.5)) allows obtaining implementable adaptive laws. The second type of conditions describes the class of systems, for which estimation of the transfer function (5.2) parameters is sufficient to reconstruct the states of the system (2.1). Let us proceed to a detailed analysis of these requirements and the adaptive observers that can be designed when they are satisfied.

5.1. Adaptive Observers for Systems with Strictly Positive Real Transfer Function from Perturbation to Output

The idea of this approach is to obtain the conditions, under which the output estimation error $\tilde{y}\,(t)$ can be used in the adaptive laws instead of the unmeasured state estimation error $\tilde{x}\,(t)$. Namely, it turns out that such a substitution is valid for systems with strictly positive real (SPR) transfer function from perturbation to output.

The main steps of observer design for systems with strictly positive real transfer function from perturbation to output are presented in Fig. 1.

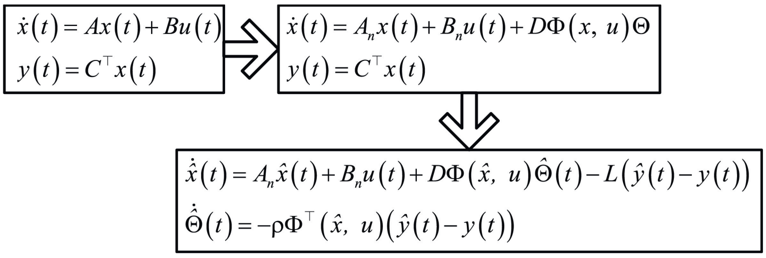

Fig. 1. Procedure of adaptive observers design for systems with strictly positive real transfer function.

As follows from Fig. 1, in the first step, the parametric perturbations of the system are represented as a new unknown input:

$$\begin{gathered}\dot{x}(t) = (A_n + \Delta A)\,x(t) + (B_n + \Delta B)\,u(t) = A_n x(t) + B_n u(t) + D\Phi(x, u)\,\Theta,\\ D\Phi(x, u)\,\Theta = \Delta A x(t) + \Delta B u(t),\end{gathered} \tag{5.9}$$

where $\Phi\colon \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{m \times p}$ is a linear function, $\Theta \in \mathbb{R}^p$ stands for a vector of unknown parameters.

In contrast to the parameterization (4.1) used for the invariant observers design, here the generalized perturbation is parametrized as a linear regression with unknown parameters and an unmeasured regressor.

The adaptive observer for the system (5.9) is proposed [46] as follows:

$$\begin{aligned}&\dot{\hat{x}}(t) = A_n\hat{x}(t) + B_n u(t) + D\Phi(\hat{x},\ u)\,\hat{\Theta}(t) - L(\hat{y}(t) - y(t)),\\ &\dot{\hat{\Theta}}(t) = -\rho\Phi^{\mathrm{T}}(\hat{x},\ u)(\hat{y}(t) - y(t)),\end{aligned} \tag{5.10}$$

where $\rho > 0$ is an adaptive gain, $L \in \mathbb{R}^n$ stands for a known vector that is chosen in such a way that $A_n + LC^{\mathrm{T}}$ is a Hurwitz matrix, and the adaptive law for $\hat{\Theta}(t) \in \mathbb{R}^p$ is derived using the Lyapunov second method.

The conditions of state estimates convergence using (5.10) are described in the following lemma:

Lemma 3. *Let the following conditions hold:*

1) $\|\Phi(x,\ u) - \Phi(\hat{x},\ u)\| \leq \gamma\|x - \hat{x}\|,\ \gamma > 0$;
2) *there exists a solution of the following euqations for* $Q = Q^{\mathrm{T}} > 0$:

$$\begin{aligned}&\left(A_n + LC^{\mathrm{T}}\right)^{\mathrm{T}}P + P\left(A_n + LC^{\mathrm{T}}\right) = -Q,\\ &PD = C,\end{aligned}$$

and it holds that $\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)} > \gamma$.

Then the adaptive observer (5.10) *ensures that:*

$$\lim_{t\to\infty}\|\tilde{x}(t)\| = 0, \quad \lim_{t\to\infty}\left\|D\Phi(\hat{x},\ u)\,\hat{\Theta}(t) - D\Phi(x,\ u)\,\Theta\right\| = 0.$$

Proof is given in [46].

The fact that the second premise of Lemma 3 is met allows one to use substitution of $\tilde{x}^{\mathrm{T}}(t)\,PD$ by $\tilde{y}(t) = C^{\mathrm{T}}\tilde{x}(t)$ in the estimation law and obtain an implementable adaptive law from (5.10), which is easy to check by consideration $V = \tilde{x}^{\mathrm{T}}P\tilde{x} + \rho^{-1}\tilde{\Theta}^{\mathrm{T}}\tilde{\Theta}$.

The first premise of Lemma 3 is always satisfied when a control law is linear w.r.t. states, because, owing to (5.9), the function $\Phi(x, u)$ is linear w.r.t. both inputs. According to the Kalman–Yakubovich–Popov lemma, the set of equations from the second premise has a solution if and only if the transfer function from the perturbation $\Phi(x,\ u)\,\Theta$ to the output $y(t)$ is strictly positively real. Moreover, the equation $PD = C^{\mathrm{T}}$ is solvable if PD lies in the linear span of C^{T}, which requires the output matching condition $\mathrm{rank}\left(C^{\mathrm{T}}D\right) = \mathrm{rank}\,(D)$ to be met, and hence also the equality $m = \dim\left(y\left(t\right)\right) = 1$.

Thus, the convergence conditions of state estimates obtained by the adaptive observer (5.10) coincide with the ones of the basic invariant Luenberger observer (4.2) and the sliding mode observer (4.7). In [47], an approach to relax the output matching condition is developed. By analogy with the solution [41], in [47], it is proposed to first obtain a new output (4.10), for which the premises of Lemma 3 are assumed to be satisfied, and then, using the new output (4.10), to implement the adaptive observer (5.5). The disadvantages of the solution [47] coincide with the ones of [41], *i.e.* the observer [47] is applicable only to the systems, for which the extended matching condition is met.

The extended and standard matching conditions are restrictive and do not allow one to design adaptive observers for systems, in which parametric uncertainties affect several equations of the system (e.g., for (4.5)).

5.2. Adaptive Observers Based on System Transformation into Observer Canonical Form

As noted earlier, the numerator and denominator parameters of the transfer function (5.3) are related to the control and output signals. This observation motivates the transformation of the system (2.1) into the observer canonical form of state space, the unknown parameters of which are precisely such transfer function parameters. For this purpose, a nonsingular (for completely observable systems) transformation $\xi\left(t\right) = Tx\left(t\right)$ is introduced as follows:

$$\begin{gathered} T^{-1} = \begin{bmatrix} A^{n-1}\mathcal{O}_n & A^{n-2}\mathcal{O}_n & \cdots & \mathcal{O}_n \end{bmatrix}, \quad \mathcal{O}_n = \mathcal{O}\begin{bmatrix} 0_{1\times(n-1)} & 1 \end{bmatrix}^{\mathrm{T}}, \\ \mathcal{O}^{-1} = \begin{bmatrix} C & A^{\mathrm{T}}C & \cdots & \left(A^{n-1}\right)^{\mathrm{T}}C \end{bmatrix}^{\mathrm{T}}, \end{gathered} \tag{5.11}$$

which, in accordance with [48, 49], allows one to rewrite equations of the system (2.1):

$$\dot{\xi}\left(t\right) = A_0\xi\left(t\right) + \psi_a y\left(t\right) + \psi_b u\left(t\right), \tag{5.12}$$

$$y\left(t\right) = C_0^{\mathrm{T}}\xi\left(t\right), \quad \xi\left(t_0\right) = Tx_0, \tag{5.13}$$

where

$$\begin{gathered} \psi_a = TAT^{-1}C_0 = \begin{bmatrix} a_{n-1} & a_{n-2} & \ldots & a_0 \end{bmatrix}^{\mathrm{T}}, \\ \psi_b = TB = \begin{bmatrix} b_{n-1} & b_{n-2} & \ldots & b_0 \end{bmatrix}^{\mathrm{T}}, \\ A_0 = \begin{bmatrix} 0_n & \begin{matrix} I_{n-1} \\ 0_{1\times(n-1)} \end{matrix} \end{bmatrix}, \quad C_0^{\mathrm{T}} = C^{\mathrm{T}}T^{-1} = \begin{bmatrix} 1 & 0_{n-1}^{\mathrm{T}} \end{bmatrix}. \end{gathered}$$

Then the problem of physical states estimation (2.5) is transformed into the one for the virtual states:

$$\lim_{t\to\infty}\left\|\tilde{\xi}\left(t\right)\right\| = \lim_{t\to\infty}\left\|\hat{\xi}\left(t\right) - \xi\left(t\right)\right\| = 0 \ (\exp), \tag{5.14}$$

where $\hat{\xi}\left(t\right)$ is a virtual state estimate, $\tilde{\xi}\left(t\right)$ denotes error of reconstruction/observation/estimation.

Unlike (2.5), the goal (5.14) is achievable using adaptive control techniques. For example, the estimation laws for the unknown parameters ψ_a and ψ_b can be designed on the basis of the regression equation (5.8). However, a contradiction arises between the objective (5.14) and the practical need

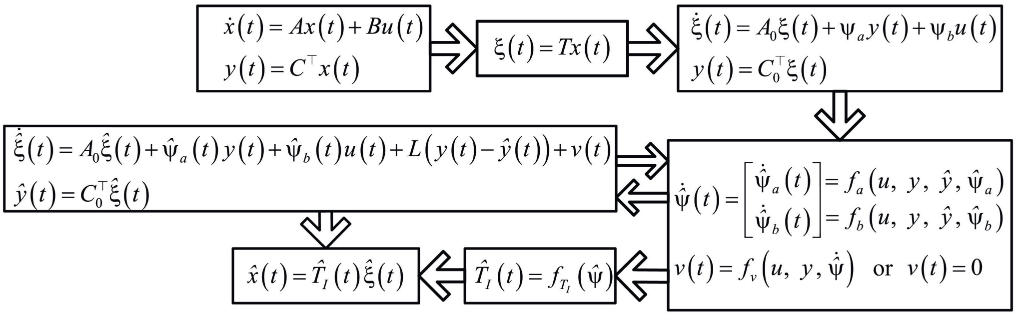

Fig. 2. Procedure of adaptive observers design on basis of system transformation into observer canonical form.

to estimate physical states to be used in the dynamic feedback (2.4a) and (2.4b). In other words, the state estimates $\hat{\xi}(t)$ are useless for *a priori* chosen feedback (2.4a) and (2.4b). Moreover, in practice, estimates of physical states $\hat{x}(t)$ are required not only for control problems, but also for other purposes (e.g., fault detection, predictive control, process monitoring, online tuning of digital twins, etc.). In order to resolve this contradiction, the physical state estimate is to be obtained as follows:

$$\hat{x}(t) = \hat{T}_I(t)\,\hat{\xi}(t), \tag{5.15}$$

where, for convenience of notation, we use the re-definition $T_I := T^{-1}$, and the estimate $\hat{T}_I(t)$ of the transformation matrix T_I is obtained on the basis of the parameter estimates of the system (5.12) via the following equation:

$$\hat{T}_I(t) = f_{T_I}\left(\hat{\psi}\right), \tag{5.16}$$

where $f_{T_I}\colon \mathbb{R}^{2n} \mapsto \mathbb{R}^n \times \mathbb{R}^n$ is a function to recalculate the transfer function parameters into the transformation matrix, $\hat{\psi}(t) = \left[\hat{\psi}_a^{\mathrm{T}}(t) \quad \hat{\psi}_b^{\mathrm{T}}(t)\right]^{\mathrm{T}} \in \mathbb{R}^{2n}$, and $\hat{\psi}_a(t)$, $\hat{\psi}_b(t)$ stand for the estimates of the parameters of the denominator and numerator of the transfer function (5.3), respectively.

Then the original goal (2.5) is equivalent to the following equalities:

$$\begin{gathered} \lim_{t\to\infty}\left\|\tilde{\xi}(t)\right\| = \lim_{t\to\infty}\left\|\hat{\xi}(t) - \xi(t)\right\| = 0 \ (\exp), \\ \lim_{t\to\infty}\left\|\tilde{\psi}(t)\right\| = \lim_{t\to\infty}\left\|\hat{\psi}(t) - \psi\right\| = 0 \ (\exp). \end{gathered} \tag{5.17}$$

In order to achieve (5.17) using the observer (5.15) + (5.16), it is required, first, the existence of a Lipschitz continuous recalculation function $f_{T_I}\colon \mathbb{R}^{2n} \mapsto \mathbb{R}^n{\times}\mathbb{R}^n \times \mathbb{R}^n$, and secondly, to design estimation laws for unknown parameters and unmeasured states, which ensure achievement of (5.17).

Block diagram of the observer (5.15) design is depicted in Fig. 2.

We first review the existing procedures to design estimation laws (functions $f_a(.)$ and $f_b(.)$ in Fig. 2). Then we show the necessity to satisfy the Lipschitz condition to achieve (2.5) when (5.17) is met and (5.16) is used.

5.2.1. Adaptive Luenberger observer. Lyapunov approach. The problem of adaptive state estimation of linear systems represented in the observer canonical form was first considered in the studies by R. Carroll and D. Lindorff [49] and G. Luders and K. Narendra [50]. It is noteworthy that these two papers, which are very close in terms of problem statement and content, were published 10 years after D. Luenberger's seminal paper [2] in the same issue of IEEE Transactions on Automatic Control with a difference of 61 pages. Later in [51] the early results of [49, 50] were revised and derived from unified perspective. Using modern notation, the observers of type [49–51] are given in [52]. In accordance with the approach under consideration, the structure of the state observer for the system (5.5) is defined in the following form:

$$\begin{aligned} \dot{\hat{\xi}}(t) &= A_0\hat{\xi}(t) + \hat{\psi}_a(t)\, y(t) + \hat{\psi}_b(t)\, u(t) + L\left(\hat{y}(t) - y(t)\right) + v(t), \\ \hat{y}(t) &= C_0^{\mathrm{T}}\hat{\xi}(t), \end{aligned} \tag{5.18}$$

where $v(t) \in \mathbb{R}^n$ is an additional signal, and the correction gain L is chosen so that to meet the condition $\lambda_{\max}\left(A_0 + LC_0^{\mathrm{T}}\right) < 0$.

The error equation between (5.18) and (5.12) is written as:

$$\begin{aligned} \dot{\tilde{\xi}}(t) &= A_m\tilde{\xi}(t) + \tilde{\psi}_a(t)\, y(t) + \tilde{\psi}_b(t)\, u(t) + v(t), \\ \tilde{y}(t) &= C_0^{\mathrm{T}}\tilde{\xi}(t), \end{aligned} \tag{5.19}$$

where $A_m = A_0 + LC_0^{\mathrm{T}}$.

The basic idea of approach from [49–51] is to transform, by means of a special choice of signal $v(t)$, the error equation (5.19) into a form such that the so-called adaptive control Lyapunov function exists.

Lemma 4. *There exists a signal vector $v(t) \in \mathbb{R}^n$ generated from measurements of known signals, for which the system* (5.19) *becomes:*

$$\begin{aligned} \dot{e}(t) &= A_m e(t) + B_c\varphi^{\mathrm{T}}(t)\,\tilde{\psi}(t), \\ \tilde{y}(t) &= C_0^{\mathrm{T}} e(t), \end{aligned}$$

where $C_0^{\mathrm{T}}\left(sI_n - A_m\right)^{-1} B_c$ is a strictly positive real (SPR) transfer function, $\varphi(t) \in \mathbb{R}^{2n}$ stands for a regressor generated on the basis of the control and output signals, $e(t) \in \mathbb{R}^n$ denotes a new state vector.

Proof is given in [52, p. 280].

By the Kalman–Yakubovich–Popov Lemma [52], the following set of equations is solvable for systems with SPR transfer function:

$$\begin{aligned} &\left(A_0 + LC_0^{\mathrm{T}}\right)^{\mathrm{T}} P + P\left(A_0 + LC_0^{\mathrm{T}}\right) = -Q, \quad Q = Q^{\mathrm{T}} > 0, \\ &PB_c = C_0, \end{aligned}$$

and, therefore, consideration of the following quadratic control Lyapunov function

$$V = e^{\mathrm{T}}Pe + \tilde{\psi}^{\mathrm{T}}\Gamma^{-1}\tilde{\psi}$$

allows one to obtain the below-given result.

Theorem 6. *The adaptive observer* (5.18) *with an adaptive law:*

$$\dot{\hat{\psi}}(t) = \dot{\tilde{\psi}}(t) = -\Gamma\tilde{y}(t)\,\varphi(t), \quad \Gamma = \Gamma^{\mathrm{T}} > 0, \tag{5.20}$$

in case $\xi \in L_\infty$ and $u \in L_\infty$, ensures that:

(i) *all signals are bounded,*

(ii) $\lim\limits_{t\to\infty} |\tilde{y}(t)| = 0$,

(iii) $\dot{\hat{\psi}} \in L_2 \cap L_\infty$ *and* $\lim\limits_{t\to\infty} \left\|\dot{\hat{\psi}}(t)\right\| = 0.$

In addition, if for all $t \geq t_0$ there exist $T > 0$ and $\alpha > 0$, such that[3]

$$\int\limits_{t}^{t+T} \varphi(\tau)\varphi^{\mathrm{T}}(\tau)\,d\tau \geq \alpha > 0,$$

then the following equalities hold:

$$\lim_{t\to\infty}\left\|\tilde{\psi}(t)\right\| = 0 \ (\exp), \quad \lim_{t\to\infty}\left\|\tilde{\xi}(t)\right\| = 0 \ (\exp).$$

Proof is given in [52, p. 282].

As for the considered solution, the equation of the observer (5.18) dynamics differs in the course of transients from the dynamics of the system (5.12) due to the additional signal $v(t)$, which is asymptotically decaying but imposes significant distortion into the estimates $\hat{\xi}(t)$. In addition, the adaptive law (5.18) for the observer parameters cannot be chosen arbitrarily but is derived using the control Lyapunov function, and therefore, it does not provide any guarantees of the transient quality for the parametric error $\tilde{\psi}(t)$ (e.g., it does not guarantee monotonicity of the error vector by norm or elements).

5.2.2. Adaptive Luenberger observer. Estimation-based approach. To eliminate the coupling between the dynamics of the estimator (5.20) and the observer (5.18) through $\tilde{y}(t)$, and to improve the transient quality of error $\tilde{\psi}(t)$, an estimation-based approach to design adaptive state observers [52] was proposed. The observer structure is chosen as (5.18), but the additional signal is eliminated from the equations by choosing $v(t) \stackrel{\Delta}{\equiv} 0$, and the estimation laws are derived without Lyapunov functions. The basis for the design of the estimation law is the regression equation (5.8) obtained using dynamic filters.

The regression equation (5.8) allows one to use a wide class of methods from the parameter estimation theory (gradient method, variations of the least squares method, etc. [52]) to estimate the parameters of the transfer function (5.6). In particular, in case the gradient method is applied, the following theorem holds.

Theorem 7. *The adaptive observer* (5.18) *with* $v(t) \stackrel{\Delta}{\equiv} 0$ *and the adaptive law:*

$$\begin{aligned}\dot{\tilde{\psi}}(t) = \dot{\hat{\psi}}(t) &= -\Gamma\varphi(t)\left(\varphi^{\mathrm{T}}(t)\hat{\psi}(t) - z(t)\right) \\ &= -\Gamma\varphi(t)\left(\varphi^{\mathrm{T}}(t)\hat{\psi}(t) - \varphi^{\mathrm{T}}(t)\psi\right), \quad \Gamma = \Gamma^{\mathrm{T}} > 0\end{aligned} \tag{5.21}$$

in case $\xi \in L_\infty$ and $u \in L_\infty$, ensures that:

(i) *all signals are bounded,*

(ii) $\lim\limits_{t\to\infty}|\tilde{y}(t)| = 0$,

(iii) $\dot{\tilde{\psi}} \in L_2 \cap L_\infty$ *and* $\lim\limits_{t\to\infty}\left\|\dot{\tilde{\psi}}(t)\right\| = 0$.

In addition, if the condition $\varphi \in \mathrm{PE}$ is met, then it holds that

$$\lim_{t\to\infty}\left\|\tilde{\psi}(t)\right\| = 0 \ (\exp), \quad \lim_{t\to\infty}\left\|\tilde{\xi}(t)\right\| = 0 \ (\exp).$$

Proof is given in [52, p. 271].

As it follows from comparison of the premises and statements of Theorems 6 and 7, the properties of the observer (5.18) + (5.20) derived on the basis of adaptive control Lyapunov function, and the one (5.18) + (5.21) based on estimation approach coincide to each other. However, estimation laws

[3] This requirement is a regressor persistent excitation condition, which is further denoted as $\varphi \in \mathrm{PE}$.

for $\hat{\psi}(t)$ and $\hat{\xi}(t)$ are decoulped in (5.18) + (5.21), and additional signal $v(t)$, which distorts the estimate $\hat{\xi}(t)$, is not used. Moreover, as it can be easily checked by consideration of the quadratic form $V = \tilde{\psi}^{\mathrm{T}}\Gamma^{-1}\tilde{\psi}$, unlike (5.20), in case we choose $\Gamma = \gamma I_{2n} > 0$, the estimation law (5.21) ensures monotonicity of the norm of the error $\tilde{\psi}(t)$:

$$\left\|\tilde{\psi}(t_a)\right\| \leq \left\|\tilde{\psi}(t_b)\right\| \;\forall t_a \geq t_b,$$

which, in comparison with (5.20), allows one to improve the transient quality of the estimates $\hat{\psi}(t)$ and $\hat{\xi}(t)$.

In addition, the considered estimation-based approach and parameterization (5.8) open a wide range of possibilities for the design of adaptive state observers with relaxed regressor excitation requirements and accelerated/improved convergence. These solutions will be discussed later in Section 5.2.6.

5.2.3. Extended Kalman filter. Joint estimation of the states and unknown parameters of the system (5.12) can also be achieved using the extended Kalman filter [3]. To apply it, the system is rewritten in the following form:

$$\begin{gathered}\dot{\xi}_e(t) = \begin{bmatrix} A_0 & \Psi(t) \\ 0_{2n\times n} & 0_{2n\times 2n}\end{bmatrix}\xi_e(t) = A_e(t)\,\xi_e(t), \\ y(t) = C_e^{\mathrm{T}}\xi_e(t) = \begin{bmatrix} 1 & 0_{3n-1}\end{bmatrix}\xi_e(t),\end{gathered} \tag{5.22}$$

where $\xi_e(t) = \begin{bmatrix}\xi(t) & \psi\end{bmatrix}^{\mathrm{T}}$ is an extended state vector, $\Psi(t) = \begin{bmatrix} y(t)\,I_n & u(t)\,I_n\end{bmatrix} \in \mathbb{R}^{n\times 2n}$ stands for a measurable signal, $A_e(t) \in \mathbb{R}^{3n\times 3n}$ denotes a known time-varying matrix.

An extended Kalman-Bucy filter [3] for the system (5.22) is defined as follows:

$$\begin{aligned}&\dot{\hat{\xi}}_e(t) = A_e(t)\,\hat{\xi}_e(t) + L(t)\left(y(t) - C_e^{\mathrm{T}}\hat{\xi}_e(t)\right), \\ &L(t) = P(t)\,C_e R^{-1}, \\ &\dot{P}(t) = A_e(t)\,P(t) + P(t)\,A_e^{\mathrm{T}}(t) + Q - P(t)\,C_e R^{-1} C_e^{\mathrm{T}} P(t),\end{aligned} \tag{5.23}$$

where $R^{-1} \in \mathbb{R}$, $Q \in \mathbb{R}^{3n\times 3n}$ stand for covariances of disturbances that affect the state and output equations, respectively.

Using the Kalman filtering algorithm (5.23), the condition for exponential convergence of the estimates $\hat{\xi}_e(t)$ to their true values is [10, 52] a uniform complete observability of the extended system (5.22).

Definition 2. A pair $(C_e, A_e(t))$ is uniformly completely observable, if for all $t \geq t_0$ there exist scalars α_1, α_2, $T > 0$ such that the following inequality holds:

$$\alpha_2 I_{3n} \geq \int\limits_t^{t+T} \Phi(\tau,\ t) C_e C_e^{\mathrm{T}} \Phi(\tau,\ t)\, d\tau \geq \alpha_1 I_{3n}, \tag{5.24}$$

where $\Phi(t, t_0)$ is a fundamental matrix of the system (5.22), which is defined as:

$$\dot{\Phi}(t,\ t_0) = A_e(t)\,\Phi(t,\ t_0), \quad \Phi(t_0,\ t_0) = I_{3n}.$$

As follows from the definition, the Kalman filter (5.23) requires uniform nonsingularity of the grammian (5.24) with dimension $3n \times 3n$, while the previously considered adaptive observers require nonsingularity of some integral of dimension $2n \times 2n$. So, in general, the convergence conditions of the adaptive observers (5.18) + (5.20) and (5.18) + (5.21) are (if not the same) weaker than the ones of the extended Kalman filter (5.23). Following the review [54], it can also be shown by means of some extensive calculations that the nonsingularity of (5.24) always requires to meet the persistent excitation condition for the regressor $\varphi(t)$ from the parameterization (5.8).

5.2.4. Algebraic observers. Kreisselmeier parametrization. The above-considered adaptive obser-vers form the state estimates using differential equations, the right-hand side of which depends on the estimates of unknown parameters and contains the Luenberger correction term. Therefore, transients of the system parameters estimates affect significantly the transient quality of the state estimates, and the correction term, which is a feedback for the observer, worsens the situation even more, causing the peak effect. In order to eliminate the noted problems, G. Kreisselmeier proposed to obtain the state estimates with the help of special parameterizations using not differential but algebraic equation, which does not require to use the correction term.

The essence of the approach is as follows. A set of differential filters for control and output signals is introduced:

$$\begin{aligned} \dot{\Omega}(t) &= A_K^{\mathrm{T}}\Omega(t) + C_0 u(t), \quad \Omega(t_0) = 0_n, \\ \dot{P}(t) &= A_K^{\mathrm{T}}P(t) + C_0 y(t), \quad P(t_0) = 0_n, \end{aligned} \tag{5.25}$$

where $A_K = A_0 + KC_0^{\mathrm{T}}$ is a Hurwitz matrix.

According to the results from [55], the following holds:

$$\begin{gathered} \xi(t) = H^{\mathrm{T}}(t)\psi + e^{A_K(t-t_0)}\xi(t_0), \\ \psi^{\mathrm{T}} = \begin{bmatrix}\psi_a^{\mathrm{T}} - K^{\mathrm{T}} & \psi_b^{\mathrm{T}}\end{bmatrix}, \quad H^{\mathrm{T}}(t) = [h_1(t)\ldots h_{2n}(t)], \\ h_i(t) = F_i P(t), \quad h_{i+n}(t) = F_i\Omega(t), \quad i = \overline{1,\ n}, \end{gathered} \tag{5.26}$$

where $H(t) \in \mathbb{R}^{2n\times n}$ is a measurable regressor, $F_i \in \mathbb{R}^{n\times n}$ stands for a transformation matrix, which is composed of the parameters of the numerator polynomial of the vector function $(sI - A_K)^{-1}e_i$, $e_i \in \mathbb{R}^n$ denotes a vector, which i^{th} element equals to one.

The parametrization (5.26) motivates to estimate the unmeasurable state of the system (5.12) with the help of an algebraic equation:

$$\hat{\xi}(t) = H^{\mathrm{T}}(t)\hat{\psi}(t), \tag{5.27}$$

where the estimates $\hat{\psi}(t)$ are obtained from the regression equation:

$$\begin{aligned} y(t) &= C_0^{\mathrm{T}}\xi(t) = \varphi(t)\psi + C_0^{\mathrm{T}}e^{A_K(t-t_0)}\xi(t_0), \\ \varphi(t) &= C_0^{\mathrm{T}}H^{\mathrm{T}}(t) \end{aligned} \tag{5.28}$$

with the help of a broad range of the parameter estimation laws.

For example, in case $\xi \in L_\infty$, $u \in L_\infty$ and the condition $\varphi \in \mathrm{PE}$ is met, a gradient estimation law

$$\dot{\hat{\psi}}(t) = \dot{\tilde{\psi}}(t) = -\Gamma\varphi(t)\left(\varphi^{\mathrm{T}}(t)\hat{\psi}(t) - y(t)\right), \quad \Gamma = \Gamma^{\mathrm{T}} > 0 \tag{5.29}$$

ensures that

$$\lim_{t\to\infty}\left\|\tilde{\psi}(t)\right\| = 0 \ (\exp), \quad \lim_{t\to\infty}\left\|\tilde{\xi}(t)\right\| = 0 \ (\exp).$$

Here it should be noted that, owing to the fact that A_K is a Hurwitz matrix, an exponentially decaying term in (5.26) and (5.28) does not affect the asymptotic properties of the estimation errors, but limits the maximal convergence rate [52]. In contrast to the previously considered observers, in case of the algebraic state reconstruction (5.27) + (5.29), the transient quality improvement for $\tilde{\psi}(t)$ results in uniform transient quality improvement for estimates $\hat{\xi}(t)$ (parameter estimates are not integrated in the observer equation, and there is no correction term).

5.2.5. Algebraic observers. Alternative parametrization. The computation of the regressor $H(t)$ to implement the parameterization (5.26) requires a rather laborious calculation of the matrices F_i, especially in case if n grows. A simplified parameterization is proposed in [56], [57]. Instead of two filters with vector states (5.25), it is proposed to use three filters—two with matrix state and one with vector state:

$$\begin{aligned} \dot{\chi}(t) &= A_K \chi(t) + K y(t), \quad \chi(t_0) = 0_{n\times 1}, \\ \dot{P}(t) &= A_K P(t) + I_n y(t), \quad P(t_0) = 0_{n\times n}, \\ \dot{\Omega}(t) &= A_K \Omega(t) + I_n u(t), \quad \Omega(t_0) = 0_{n\times n}. \end{aligned} \tag{5.30}$$

Then, according to [56, 57], the parameters ψ satisfy the following linear regression equation:

$$\begin{gathered} \xi(t) = \chi(t) + H^{\mathrm{T}}(t)\psi + e^{A_K(t-t_0)}\xi(t_0), \\ H^{\mathrm{T}}(t) = \begin{bmatrix} \Omega(t) & P(t) \end{bmatrix} \in \mathbb{R}^{n\times 2n}. \end{gathered} \tag{5.31}$$

Equation (5.31) can be easily verified by direct differentiation of the error $\xi(t) - \chi(t) + H^{\mathrm{T}}(t)\psi$. Similar to (5.28), the multiplication of (5.31) by C_0^{T} yields:

$$\begin{aligned} & z(t) = y(t) - C_0^{\mathrm{T}}\chi(t) = \varphi^{\mathrm{T}}(t)\psi + C_0^{\mathrm{T}} e^{A_K(t-t_0)}\xi(t_0), \\ & \varphi^{\mathrm{T}}(t) = C_0^{\mathrm{T}} H^{\mathrm{T}}(t), \end{aligned} \tag{5.32}$$

which allows one to implement the following observer:

$$\hat{\xi}(t) = \chi(t) + H^{\mathrm{T}}(t)\hat{\psi}(t), \tag{5.33}$$

where $\hat{\psi}(t)$ can be obtained on the basis of (5.32) with the help of a broad class of parameter estimation laws.

5.2.6. Observers with accelerated/improved convergence. One of the main advantages of algebraic adaptive observers (5.27) and (5.33) over differential type adaptive observers (5.18) + (5.20), (5.18) + (5.21) is the direct dependence of the convergence rate of the observation error $\xi(t)$ on the convergence rate of the parametric error and the filter parameters (5.25) or (5.30).

Indeed, when $\xi \in L_\infty$ and $u \in L_\infty$, it holds that

$$\left\|\tilde{\xi}(t)\right\| \le H_{\max}\left\|\tilde{\psi}(t)\right\| + e^{A_K(t-t_0)}\xi_{\max},$$

which allows one to adjust the convergence rate of $\tilde{\xi}(t)$ by choosing A_K and increasing the rate of convergence of $\tilde{\psi}(t)$.

At the same time, the rate of convergence of the error from the error equation (5.19) is known to be defined not only by the rate of convergence of its input, but also by the eigenvalues of the matrix of the closed-loop system $A_0 + L_0 C^{\mathrm{T}}$. In this case, the increase (in terms of absolute value) of the eigenvalues real part can lead to significant peaks in the course of transient process.

When $\Gamma = \gamma I_{2n}$ and the condition of regressor persistent excitation is met, the convergence rate of the previously discussed gradient estimation law of type (5.21) or (5.29), according to the results of [58], can be written in the following form:

$$\begin{gathered} \left\|\tilde{\psi}(t)\right\| \le \sqrt{a}\,\|\varphi\|\, e^{-\frac{1}{2}\gamma a^{-1}(t-t_0)}\left\|\tilde{\psi}(t_0)\right\|, \\ a = \gamma b^{-1} e^{2bT}, \quad b = -\frac{1}{2T}\ln\left(1 - \frac{\gamma\alpha}{1+\gamma^2 T^2\|\varphi\|_\infty^4}\right). \end{gathered}$$

Thus, using conventional gradient estimation laws, the convergence rate of the parametric error cannot be arbitrarily increased by raising the gain $\gamma > 0$, since the product γa^{-1} may decrease as γ increases. Therefore, in fact, using algebraic observers (5.27) and (5.33) with conventional gradient estimation laws of the form (5.21), (5.29), as well as observers (5.18) + (5.20) or (5.18) + (5.21), it is not easy to increase the convergence rate of the unmeasured state estimation without deterioration of the transient quality. To solve this problem, modified parameter estimation laws with accelerated convergence have been proposed. Without loss of generality, and for simplicity, we illustrate the properties of these improved estimation laws using the estimation problem with parametrization (5.32). Here, we temporarily omit the term $C_0^{\mathrm{T}} e^{A_K (t-t_0)} \xi(t_0)$ but acknowledge its constraints on the achievable convergence rate.

1) Kreisselmeier's regressor extension scheme. G. Kreisselmeier proposed [55] to transform the vector regressor $\varphi(t) \in \mathbb{R}^{2n}$ from the regression equation (5.32) into a symmetric positive-semidefinite matrix $\Phi(t) \in \mathbb{R}^{2n\times 2n}$, $\Phi(t) = \Phi^{\mathrm{T}}(t) \geq 0$ with the help of the following filtering ($l > 0$):

$$\begin{aligned} \dot{Y}(t) &= -lY(t) + \varphi(t) z(t), \quad Y(t_0) = 0_{2n}, \\ \dot{\Phi}(t) &= -l\Phi(t) + \varphi(t)\varphi^{\mathrm{T}}(t), \quad \Phi(t_0) = 0_{2n\times 2n}, \end{aligned} \tag{5.34}$$

which, owing to $\psi = \text{const}$, allows one to obtain a new regression equation:

$$Y(t) = \Phi(t)\psi. \tag{5.35}$$

Based on equation (5.35), a new estimation law is designed:

$$\dot{\tilde{\psi}}(t) = \dot{\hat{\psi}}(t) = -\Gamma\left(\Phi(t)\hat{\psi}(t) - Y(t)\right), \quad \Gamma = \Gamma^{\mathrm{T}} > 0, \tag{5.36}$$

which has significantly different properties in comparison with the previously considered laws [59]:

1) $\lambda_{\min}(\Phi(t)) \notin L_1 \Leftrightarrow \lim\limits_{t\to\infty} \left\|\tilde{\psi}(t)\right\| = 0;$
   $\varphi \in \mathrm{PE} \Leftrightarrow \lim\limits_{t\to\infty} \left\|\tilde{\psi}(t)\right\| = 0 \ (\exp);$
2) $\left\|\tilde{\psi}(t_a)\right\| \leq \left\|\tilde{\psi}(t_b)\right\| \ \forall t_a \geq t_b,$
3) if $\varphi \in \mathrm{PE}$ and $\Gamma = \gamma I_{2n}$, then the rate of exponential convergence of the parametric error $\tilde{\psi}(t)$ can be made arbitrarily fast by increasing γ.

The most important one is the third property of the law (5.36), which is explained by the property

$$\varphi \in \mathrm{PE} \Leftrightarrow \forall t \geq kT \ \lambda_{\min}(\Phi(t)) > \mu > 0,$$

proved in [60] and allowed one to obtain the following bound of the parametric error norm:

$$\frac{d}{dt}\left\|\tilde{\psi}\right\|^2 = -2\gamma\tilde{\psi}^{\mathrm{T}}\Phi\tilde{\psi} \leq -2\gamma\lambda_{\min}(\Phi)\left\|\tilde{\psi}\right\|^2$$
$$\Downarrow$$
$$\left\|\tilde{\psi}(t)\right\| \leq e^{-\gamma\int\limits_{t_0}^{t}\lambda_{\min}(\Phi)d\tau}\left\|\tilde{\psi}(t_0)\right\| \leq e^{-\gamma\mu(t-t_0)}\left\|\tilde{\psi}(t_0)\right\|,$$

from which we immediately have the third property.

Considering estimation law (5.36), the filtering operation (5.34) allows, in addition to current indirect information about system parameters, to use a certain amount of historical data (determined by parameter l). This enables the system to have indirect information (via $Y(t)$ and $\Phi(t)$) about all unknown parameters at each time instant, which in turn allows one to increase the convergence rate by means of increasing γ. Thus, when using the algebraic state observer (5.33) augmented with the accelerated convergence estimator (5.36), it becomes possible to enhance the exponential convergence rate of state reconstruction errors by increasing γ. It should also be noted that, unlike previously considered observers, the observer (5.33) + (5.36) does not require the persistent excitation condition to be satisfied to achieve asymptotic (though not exponential) convergence of the estimation error.

2) Dynamic regressor extension and mixing procedure. In [61], it is proposed to use a time-varying adaptive gain for the law (5.36):[4]

$$\Gamma = \gamma \operatorname{adj}\{\Phi(t)\}, \quad \gamma > 0, \tag{5.37}$$

which, using the property $\operatorname{adj}\{\Phi(t)\}\Phi(t) = \det\{\Phi(t)\} I_{2n\times 2n}$, allows one to obtain:

$$\begin{gathered}
\dot{\hat{\psi}}(t) = \dot{\tilde{\psi}}(t) = -\gamma \operatorname{adj}\{\Phi(t)\}\left(\Phi(t)\hat{\psi}(t) - Y(t)\right) \\
= -\gamma\left(\Delta(t)\hat{\psi}(t) - \mathcal{Y}(t)\right) \\
\Downarrow \\
\dot{\hat{\psi}}_i(t) = \dot{\tilde{\psi}}_i(t) = -\gamma\Delta(t)\tilde{\psi}_i(t), \\
\Delta(t) = \det\{\Phi(t)\}, \quad \mathcal{Y}(t) = \operatorname{adj}\{\Phi(t)\}Y(t) = \Delta(t)\psi,
\end{gathered} \tag{5.38}$$

as a result, the new law (5.38) has the properties 1)–3) of (5.36) and additionaly ensures element-wise monotonicity of the parametric error:

$$\left|\tilde{\psi}_i(t_a)\right| \leq \left|\tilde{\psi}_i(t_b)\right| \; \forall t_a \geq t_b, \quad i \in \{1, \ldots, n\}. \tag{5.39}$$

As it will be shown further, the property (5.39) is essential for recalculation via (5.16) of the estimates $\hat{\psi}(t) = \begin{bmatrix}\hat{\psi}_a(t) & \hat{\psi}_b(t)\end{bmatrix}^{\mathrm{T}}$ into the estimate of the transformation matrix $\hat{T}(t)$. Indeed, having (5.39) at hand, it holds that:

$$\left.\begin{array}{l} \hat{\psi}_i(t_0) \in \Xi \\ \psi_i \in \Xi \end{array}\right\} \Rightarrow \hat{\psi}_i(t) \in \Xi, \tag{5.40}$$

which, in some cases, given certain *a priori* information about the system parameters (knowledge that $\psi_i \in \Xi$), allows one to avoid discontinuities in the transformation (5.15). The role of property (5.40) will be discussed in more detail further.

3) Relaxation of Exponential Convergence Condition. Integral Dynamic Regressor Extension and Mixing Procedure. The considered estimation laws (5.20), (5.21), (5.29), (5.36), or (5.38) ensure exponential convergence of the parametric error and estimation error under the condition of regressor persistent excitation (nonsingularity of integral of dimension $2n \times 2n$), which typically requires the control input to contain at least n distinct frequencies [52, 63]. In practice, in order to meet such condition, non-decaying harmonic dither signals are required to be added to the control signal, which may contradict the original control objective.

Thus, to relax the convergence conditions for state and parameter estimates to their true values, estimation laws with weaker regressor excitation requirements have been proposed in the literature. One such condition is the finite excitation requirement ($\varphi \in \mathrm{FE}$).[5] Unlike regressor persistent

[4] It should be noted that a similar regressor scalarization procedure was proposed ten years earlier in a less cited paper [62].

[5] The condition $\varphi \in \mathrm{FE}$ means that there exist $t_e > t_r \geq t_0$ and $\alpha > 0$ such that $\int\limits_{t_r}^{t_e} \varphi(\tau)\varphi^{\mathrm{T}}(\tau)\,d\tau \geq \alpha > 0$.

excitation, the finite excitation condition only requires the nonsingularity of a certain integral over a finite time interval rather than the entire time axis. In terms of convergence conditions of the Kalman–Bucy filter (5.23), this relaxed condition is equivalent to observability of the extended system (5.22) only over a specific time interval [3]. The finite excitation condition can be satisfied by addition of an exponentially decaying harmonic signal with n distinct frequencies to the control input, which does not prevent the achievement of the asymptotic control objective. For many systems, this condition is met even in case of complete absence of dither signals.

To relax the persistent excitation requirement, numerous approaches have been proposed to handle the regressor and regressand. A detailed overview of such methods is provided in [64–67]. The core idea of most approaches is to "memorize" in a special way the indirect information about unknown parameters obtained during the finite excitation interval.

In [68], instead of (5.34), it is proposed to use the following filtering ($\sigma > 0$):

$$\begin{gathered}\dot{Y}(t) = e^{-\sigma(t-t_0)}\varphi(t)\,z(t), \quad Y(t_0) = 0_{2n},\\ \dot{\Phi}(t) = e^{-\sigma(t-t_0)}\varphi(t)\,\varphi^{\mathrm{T}}(t), \quad \Phi(t_0) = 0_{2n\times 2n},\end{gathered} \tag{5.41}$$

which allows one to obtain a new regression equation of the form (5.35), for which regressor, in case $\varphi \in \mathrm{FE}$, it holds that:

$$\begin{gathered}\Phi(t) = \int\limits_{t_0}^{t} e^{-\sigma(\tau-t_0)}\varphi(\tau)\,\varphi^{\mathrm{T}}(\tau)\,d\tau \geq \int\limits_{t_0}^{t_e} e^{-\sigma(\tau-t_0)}\varphi(\tau)\,\varphi^{\mathrm{T}}(\tau)\,d\tau\\ \geq e^{-\sigma(t_e-t_0)}\int\limits_{t_0}^{t_e}\varphi(\tau)\,\varphi^{\mathrm{T}}(\tau)\,d\tau \geq \alpha e^{-\sigma(t_e-t_0)}I_{2n} > 0.\end{gathered} \tag{5.42}$$

Based on equation (5.35) obtained by filtering (5.41), an estimation law with matrix (5.36) or scalar (5.38) regressor can be implemented. The properties of these laws differ only by the type of monotonicity of the parametric error (by norm or element-wise):

1) $\varphi \in \mathrm{FE} \Leftrightarrow \lim\limits_{t\to\infty}\left\|\tilde{\psi}(t)\right\| = 0 \ (\exp)$;
2) $(5.41) + (5.36) \Rightarrow \left\|\tilde{\psi}(t_a)\right\| \leq \left\|\tilde{\psi}(t_b)\right\| \ \forall t_a \geq t_b$
   $(5.41) + (5.38) \Rightarrow \left|\tilde{\psi}_i(t_a)\right| \leq \left|\tilde{\psi}_i(t_b)\right| \ \forall t_a \geq t_b, \ i \in \{1,\ldots,n\}$,
3) if $\varphi \in \mathrm{FE}$ and $\Gamma = \gamma I_{2n}$, then the rate of exponential convergence of the parametric error $\tilde{\psi}(t)$ can be made arbitrarily fast by increasing γ.

As it follows from the comparison of properties of the laws (5.34) + (5.36), (5.34) + (5.38) and (5.41) + (5.36), (5.41) + (5.38), the second group inherits the advantages of the first one but relaxes the requirement necessary to achieve the goals (2.5) and (5.17).

Relaxation of the requirement of the regressor persistent excitation allows one to extend the estimation problem for parameterizations (5.26) and (5.31) by taking into consideration an exponentially decaying term:[6]

$$\begin{gathered}\xi(t) = H^{\mathrm{T}}(t)\,\psi_e,\\ \psi_e^{\mathrm{T}} = \begin{bmatrix}\psi_a^{\mathrm{T}} - K^{\mathrm{T}} & \psi_b^{\mathrm{T}} & \xi^{\mathrm{T}}(t_0)\end{bmatrix}, \quad H^{\mathrm{T}}(t) = \begin{bmatrix}h_1(t)\ldots h_{2n}(t) & e^{A_K(t-t_0)}\end{bmatrix},\\ \xi(t) = \chi(t) + H^{\mathrm{T}}(t)\,\psi_e,\\ H^{\mathrm{T}}(t) = \begin{bmatrix}\Omega(t) & P(t) & e^{A_K(t-t_0)}\end{bmatrix},\end{gathered}$$

[6] Extension of the estimation problem without relaxation of the persistent excitation requirement is impossible, as regressors containing a decaying term do not satisfy this condition.

which, in case of application of the corresponding parameter estimation law for ψ_e, results in the following state reconstruction error bound:

$$\left\|\tilde{\xi}\left(t\right)\right\| \leq H_{\max}\left\|\tilde{\psi}_e\left(t\right)\right\|$$

and eliminates the constraints that the eigenvalues of matrix A_K impose on the error $\tilde{\xi}\left(t\right)$ convergence rate, allowing its adjustment using only a single scalar parameter γ.

4) Relaxation of the regressor persistent excitation requirement. Estimation algorithm with finite-time convergence. An alternative estimation law [69] with relaxed regressor excitation requirements ensures finite-time convergence of parameter/state estimation errors and have the form:

$$\begin{gathered}\hat{\psi}_i^{\mathrm{FTC}}\left(t\right)=\frac{1}{1-\phi_c\left(t\right)}\left[\hat{\psi}_i\left(t\right)-\phi_c\left(t\right)\hat{\psi}_i\left(t_0\right)\right], \\ \phi_c\left(t\right)=\begin{cases}\sigma, & \text{if} \quad \phi\geq\sigma \\ \phi\left(t\right), & \text{if} \quad \phi<\sigma,\end{cases} \qquad \dot{\phi}\left(t\right)=-\gamma\Delta^2\left(t\right)\phi\left(t\right), \quad \phi\left(t_0\right)=1.\end{gathered} \tag{5.43}$$

In [69], the convergence of $\hat{\psi}_i^{\mathrm{FTC}}\left(t\right)$ to ψ_i in finite time t_e is proved in case the following condition holds:

$$\int\limits_{t_0}^{t_e}\Delta^2\left(\tau\right)d\tau\geq-\frac{1}{\gamma_i}\ln\left(\sigma\right). \tag{5.44}$$

In order to explain the relaxation method, a solution of equation (5.38) is written:

$$\begin{gathered}\tilde{\psi}_i\left(t\right)=\phi\left(t\right)\tilde{\psi}_i\left(t_0\right) \\ \Downarrow \\ \hat{\psi}_i\left(t\right)-\phi\left(t\right)\hat{\psi}_i\left(t_0\right)=\left[1-\phi\left(t\right)\right]\psi_i.\end{gathered} \tag{5.45}$$

Since, owing to $\Delta^2\left(t\right)\geq 0$, $\phi\left(t\right)$ is a non-increasing function of time, then, when the condition (5.44) is met, the logical operator (5.43) switches at time instant t_e, and, following (5.45), the unknown parameters values are determined analytically. In parameterization (5.45), the regressor $1-\phi\left(t\right)$ and regressand $\hat{\psi}_i\left(t\right)-\phi\left(t\right)\hat{\psi}_i\left(t_0\right)$ play the roles of $\Phi\left(t\right)$ and $Y\left(t\right)$ from parametrization (5.35) + (5.41). When the condition (5.44) is met, the regressor $1-\phi\left(t\right)$ is globally bounded away from zero starting from some time instant, *i.e.*, indirect information about all unknown parameters has been obtained and preserved, enabling their analytical calculation. Consequently, the counterpart to law (5.43) in terms of parameterization (5.41) is the following one:

$$\hat{\psi}^{\mathrm{FTC}}\left(t\right)=\left\{\begin{array}{ll}\hat{\psi}\left(t_0\right), & \text{if} \quad \det\left\{\Phi\left(t\right)\right\}<\rho\leq\det\left\{\alpha e^{-\sigma t_e}I_{2n}\right\} \\ \Phi^{-1}\left(t\right)Y\left(t\right), & \text{otherwise}\end{array}\right. \tag{5.46}$$

which, owing to the fact that inequality (5.42) holds, also allows one to estimate the parameters in finite-time in case of proper choice of ρ. An open problem of finite-time convergence laws (5.43) or (5.46) is their high sensitivity to parameters σ, γ_i, ρ—parametric convergence (or its absence in case of poor choice) depends entirely on the selection of these values.

5.2.7. Problem of singularities of recalculation (5.16). Having considered in detail most of the existing algorithms and procedures to estimate the virtual states $\xi\left(t\right)$ of the observer canonical form (5.12) and the parameters ψ_a, ψ_b of the numerator and denominator of the transfer function (5.6), we return to the design problem of the observer (5.15) + (5.16) of the physical states of the original system (2.1).

First, in order to obtain estimates (5.15), it is necessary to have an equation (5.16) to recalculate parameters of the transfer function (5.6) into elements of the matrix T_I. Generally speaking, complete observability of the system is a necessary, but not a sufficient condition for the existence of such an equation.

Example 1. Consider a completely observable (in case $\theta_2 \neq 0$) system:

$$A = \begin{bmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \theta_5 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In accordance with the transformation (5.11), the parameters of the numerator and denominator of the transfer function (5.6) are defined as follows:

$$\psi_a = \begin{bmatrix} \theta_1 + \theta_4 \\ \theta_2\theta_3 - \theta_1\theta_4 \end{bmatrix}, \quad \psi_b = \begin{bmatrix} 0 \\ \theta_2\theta_5 \end{bmatrix}, \tag{5.47}$$

and the transformation T_I is written as:

$$T_I = \begin{bmatrix} 1 & 0 \\ \theta_4\theta_2^{-1} & \theta_2^{-1} \end{bmatrix}.$$

A set of algebraic equations (5.47) has no solution, and therefore, it is impossible to implement (5.16) despite complete observability of the system. ■

Secondly, even if the recalculation equation (5.16) exists, then exponential convergence of the physical state reconstruction error $\tilde{x}(t)$ follows from exponential convergence of $\tilde{\psi}(t)$ and $\tilde{\xi}(t)$ if the function $f_{T_I}\colon \mathbb{R}^{2n} \mapsto \mathbb{R}^n \times \mathbb{R}^n$ is Lipschitz continuous.

Lemma 5. *Let $\tilde{\psi}(t)$ and $\tilde{\xi}(t)$ converge exponentially to zero, then the limit* (2.5) *exists, if for all $a,\ b \in \mathbb{R}^{2n}$ there exists a scalar $\mu > 0$ such that:*

$$\left\| f_{T_I}(a) - f_{T_I}(b) \right\| \leq \mu \left\| a - b \right\|. \tag{5.48}$$

Proof. The following error equation is written:

$$\begin{aligned} \tilde{x}(t) &= \hat{T}_I(t)\hat{\xi}(t) - T_I\xi(t) \pm \hat{T}_I(t)\xi(t) = \hat{T}_I(t)\tilde{\xi}(t) + \tilde{T}_I(t)\xi(t) \pm T_I\tilde{\xi}(t) \\ &= \tilde{T}_I(t)\tilde{\xi}(t) + T_I\tilde{\xi}(t) + \tilde{T}_I\xi(t), \end{aligned}$$

Using the condition (5.48), an upper bound is obtained:

$$\begin{aligned} \left\| \tilde{x}(t) \right\| &\leq \left\| \tilde{T}_I \right\| \left\| \tilde{\xi}(t) \right\| + \left\| T_I \right\| \left\| \tilde{\xi}(t) \right\| + \left\| \tilde{T}_I(t) \right\| \left\| \xi(t) \right\| \\ &\leq \left(\left\| T_I \right\| + \mu \left\| \tilde{\psi}(t) \right\| \right) \left\| \tilde{\xi}(t) \right\| + \mu \left\| \tilde{\psi}(t) \right\| \left\| \xi(t) \right\|, \end{aligned}$$

from which we have exponential convergence of the error $\tilde{x}(t)$ in case of exponential convergence of $\tilde{\psi}(t)$ and $\tilde{\xi}(t)$. ■

For mathematical models of physical systems, recalculation equations (5.16) typically include division operation. This violates condition (5.48) and, regardless of the parameter estimation law used for $\hat{\psi}(t)$, may cause singularity of estimates $\hat{T}_I(t)$, and consequently $\hat{x}(t)$. Let us validate this claim through the following examples.

Example 2. Consider the system (2.1) with the following matrices:

$$A = \begin{bmatrix} 0 & \theta_1 \\ \theta_1 & \theta_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

According to the transformation (5.11), the parameters of the numerator and denominator of the transfer function (5.6) are defined as follows:

$$\psi_a = \begin{bmatrix} \theta_2 \\ \theta_1^2 \end{bmatrix}, \quad \psi_b = \begin{bmatrix} 0 \\ \theta_1\theta_3 \end{bmatrix},$$

and the transformation T_I is written as:

$$T_I = \begin{bmatrix} 1 & 0 \\ \theta_2\theta_1^{-1} & \theta_1^{-1} \end{bmatrix}.$$

Then the recalculation equation (5.16) is obtained as:

$$\hat{T}_I(t) = \begin{bmatrix} 1 & 0 \\ \dfrac{\hat{\psi}_{1a}(t)}{\sqrt{\hat{\psi}_{2a}(t)}} & \dfrac{1}{\sqrt{\hat{\psi}_{2a}(t)}} \end{bmatrix}.$$

The recalculation function does not satisfy the Lipschitz condition and is singular in case $\hat{\psi}_{2a}(t) = 0$. Moreover, the negative values of $\hat{\psi}_{2a}(t)$ are out of admissible range. ■

On the basis of the dynamic properties of the parametric error, the previously discussed estimation laws of the parameters ψ_a, ψ_b can be divided into three groups:

1) do not ensure monotonicity of the error $\tilde{\psi}(t)$ neither by norm, nor by elements (5.20), (5.23);
2) guarantee monotonicity by norm (5.21), (5.29), (5.36);
3) ensure element-wise monotonicity (5.34) + (5.38), (5.41) + (5.38), (5.34) + (5.43).

Considering the estimation laws from the first two groups, it is not possible to guarantee that the estimate $\hat{\psi}_{2a}(t)$ does not enter the forbidden region $\hat{\psi}_{2a}(t) \leq 0$ in the course of the transient even if the initial conditions are chosen so that $\hat{\psi}_{2a}(t_0) > 0$. Hence, the law of physical state reconstruction (5.15) + (5.16) cannot be implemented when using observers and estimators (5.18) + (5.20), (5.18) + (5.21), (5.23), (5.26) + (5.29), (5.33) + (5.34) + (5.36) and (5.33) + (5.41) + (5.36).

The estimation laws with a scalar regressor (5.34) + (5.38), (5.41) + (5.38), (5.44) + (5.43) ensure element-wise monotonicity of the parametric error, and therefore, owing to (5.39) and (5.40), it holds for the example under consideration that

$$\left.\begin{array}{l} \hat{\psi}_2(t_0) > 0 \\ \psi_2 > 0 \end{array}\right\} \Rightarrow \hat{\psi}_2(t) > 0,$$

as it is *a priori* known from the definition of ψ_a that $\psi_2 = \psi_{2a} > 0$.

The property (5.40) allows one to apply estimators and observers (5.33) + (5.34) + (5.38), (5.33) + (5.41) + (5.38), (5.33) + (5.34) + (5.43) to implement the observer of the physical state (5.15) + (5.16). However, the properties (5.39) and (5.40) are fragile with respect to external perturbations, and more realistic practical problems may require unavailable *a priori* information to implement these estimators. Let us confirm the latter conclusion with an example.

Example 3. Consider a system (2.1) with the following matrices:

$$A = \begin{bmatrix} 0 & \theta_1 - \theta_2 \\ \theta_1 & \theta_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{5.49}$$

Following the transformation (5.11), the parameters of the numerator and denominator of the transfer function (5.6) are defined as follows:

$$\psi_a = \begin{bmatrix} \theta_1 \\ \theta_1 (\theta_1 - \theta_2) \end{bmatrix}, \quad \psi_b = \begin{bmatrix} 0 \\ \theta_3 (\theta_1 - \theta_2) \end{bmatrix}, \tag{5.50}$$

and the transformation T_I is written as:

$$T_I = \begin{bmatrix} 1 & 0 \\ \dfrac{\theta_1}{\theta_1 - \theta_2} & \dfrac{1}{\theta_1 - \theta_2} \end{bmatrix}.$$

Then the recalculation equation (5.16) is obtained as:

$$\hat{T}_I (t) = \begin{bmatrix} 1 & 0 \\ \dfrac{\hat{\psi}_{1a}^2 (t)}{\hat{\psi}_{2a} (t)} & \dfrac{\hat{\psi}_{1a} (t)}{\hat{\psi}_{2a} (t)} \end{bmatrix},$$

from which we have a singularity of the function (5.16) for $\hat{\psi}_{2a} (t) = 0$.

The estimation laws from the first two groups still cannot be used, and the implementation of the laws from the third group requires knowledge of the sign of the product $\theta_1 (\theta_1 - \theta_2)$, since the following holds:

$$\begin{gathered} \operatorname{sgn} \left(\hat{\psi}_2 (t_0) \right) = \operatorname{sgn} (\psi_2) \\ \Downarrow \\ \operatorname{sgn} \left(\hat{\psi}_2 (t) \right) = \operatorname{sgn} (\psi_2), \end{gathered}$$

where $\psi_2 = \psi_{2a} = \theta_1 (\theta_1 - \theta_2)$.

Unlike the previous example, in the considered case (5.49), (5.50), the information about the sign of ψ_2 cannot be obtained from (5.50), and hence without additional *a priori* information about the system parameters the estimation laws of the third group also cannot be used. So, for the system (5.49), the observer (5.15) + (5.16) is not implementable for any existing laws to estimate the parameters of the transfer function (5.6). ■

Examples 1–3 show that, in general case, if the system (2.1) has parametric uncertainty, then, using the observers on the basis of transformation of such system into the observer canonical form, the only solvable problem is the estimation of its virtual states $\xi (t)$, which are useless to implement the dynamic feedback of the form (2.4a) and (2.4b).

However, for few particular cases, in which the transformation equation (5.16) exists and satisfies (possibly locally) the Lipschitz condition, these solutions, unlike other adaptive and invariant observers, do not require the restrictive output matching conditions to be met. Indeed, if $\psi \in \mathcal{D}_\psi$, and the function $f_{T_I} \colon \mathbb{R}^{2n} \mapsto \mathbb{R}^n \times \mathbb{R}^n$ exists and meets the local Lipschitz condition inside $\mathcal{D}_\psi$, and, owing to (5.40), it is ensured that $\hat{\psi} (t) \in \mathcal{D}_\psi$, then, the observer (5.15) + (5.16) guarantees state reconstruction of the system, which does not meet the output matching conditions. For example, the system from Example 2 does not satisfy such condition:

$$\dot{x} = \begin{bmatrix} 0 & \theta_1 \\ \theta_1 & \theta_2 \end{bmatrix} x + \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix} u = A_n x + B_n u + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 & 0 & 0 & 0 \\ 0 & x_1 & x_2 & u \end{bmatrix} \begin{bmatrix} \theta_1 - 1 \\ \theta_1 \\ \theta_2 \\ \theta_3 - 1 \end{bmatrix}}_{D\Phi(x,\, u)\Theta},$$

but, nevertheless, owing to the property (5.40), for some choice of the estimates initial conditions, such system state can be reconstructed with the help of the observers (5.33) + (5.34) + (5.38), (5.33) + (5.41) + (5.38), (5.33) + (5.34) + (5.43). As it is shown in Example 3, unavailable information about the system parameters may be needed to meet the condition $\hat{\psi}(t) \in \mathcal{D}_\psi$.

Thus, the challenges of implementation of state observers based on system transformation to observer canonical form are less about to obtain estimation algorithms for $\hat{\psi}_a(t)$, $\hat{\psi}_b(t)$ and $\hat{\xi}(t)$ that ensure (5.17), but more about implementation of the recalculation equation (5.16), which may:

— not exist even for completely observable systems;

— exist but does not meet the Lipschitz condition (it can include operations of division, raising to a power, square root, their combinations, etc.).

Several different approaches have been proposed in the literature to overcome the problem of recalculation (5.16) singularity. Before proceeding to their consideration, let us study in detail the conditions for the existence of the recalculation equation (5.16). Note that in Examples 1–3 we considered systems with overparameterization, *i.e.*, with a number of unknown parameters that is larger than necessary for a minimal system representation. Such a class of systems in the state space is defined as follows:

$$\begin{aligned} \dot{x}(t) &= A(\theta)\, x(t) + B(\theta)\, u(t), \\ y(t) &= C^{\mathrm{T}} x(t), \quad x(t_0) = x_0, \end{aligned} \tag{5.51}$$

where $A : \mathbb{R}^{n_\theta} \mapsto \mathbb{R}^{n \times n}$, $B : \mathbb{R}^{n_\theta} \mapsto \mathbb{R}^n$ are mappings from the physical parameters $\theta \in \mathbb{R}^{n_\theta}$ to system matrices, while other notation coincides with the one from (2.1).

Then, using re-definition $\psi_a := \psi_a(\theta)$, $\psi_b := \psi_b(\theta)$, $T := T(\theta)$, we take into consideration the dependence of the recalculation matrix and parameters of the observer canonical form (5.12), (5.13) from the unknown parameters θ. Now the existence conditions of the recalculation equation (5.16) can be represented in terms of sufficient conditions for the existence of the inverse function that allows one to recalculate the parameters of the observer canonical form $\psi_a(\theta)$, $\psi_b(\theta)$ into physical parameters:

$$\begin{gathered} \det{}^2 \{\nabla_\theta \psi_{ab}(\theta)\} > 0, \\ \psi_{ab}(\theta) = \mathcal{L}_{ab} \begin{bmatrix} \psi_a(\theta) \\ \psi_b(\theta) \end{bmatrix} = \mathcal{L}_{ab} \psi(\theta), \end{gathered} \tag{5.52}$$

where $\psi(\theta) = \begin{bmatrix} \psi_a^\top(\theta) & \psi_b^\top(\theta) \end{bmatrix}^\top$.

If the condition (5.52) is met, then there exists an inverse function such that $\mathcal{F}(\psi_{ab}) = \theta$, and therefore, the mapping $f_{T_I} \colon \mathbb{R}^{2n} \mapsto \mathbb{R}^n \times \mathbb{R}^n$ from (5.16) is defined as:

$$f_{T_I}\left(\hat{\psi}\right) = (T_I \circ \mathcal{F})\left(\mathcal{L}_{ab}\hat{\psi}\right). \tag{5.53}$$

The motivation behind the condition (5.52) is as follows. As previously shown, using the input $u(t)$ and the output $y(t)$ signals, if certain convergence conditions are met, the parameters of the observer canonical form can be estimated. The inverse coordinate transformation matrix is calculated through the physical parameters of the system. So, it is possible to calculate it only when the parameters of the observer canonical form can be recalculated into θ parameters. Considering the system with overparameterization (5.51), if the existence conditions of the inverse function are satisfied, we can describe the existing ways to avoid singularity of the recalculation (5.16).

1) Gradient projection method. Let us introduce into consideration the following set

$$\mathcal{D}_\psi := \left\{\psi \in \mathbb{R}^{2n} : \overline{\psi_i} \le \psi_i \le \underline{\psi_i}, \quad i = 1, \dots 2n\right\},$$

where $\underline{\psi_i}$ and $\overline{\psi_i}$ are known scalars.

Suppose condition (5.48) holds for all $\psi \in \mathcal{D}_\psi \subset \mathbb{R}^{2n}$. To resolve the singularity issue, it then remains to ensure that the implication $\hat{\psi}(t_0) \in \mathcal{D}_\psi \Rightarrow \hat{\psi}(t) \in \mathcal{D}_\psi, \forall t \ge t_0$ holds. This is achieved by augmenting the estimation law with a projection operator [52, p. 788], [53, p. 133], which projects estimates onto $\mathcal{D}_\psi$. For example, the estimation law (5.21) with projection operator becomes:

$$\begin{gathered}
\dot{\hat{\psi}}(t) = \lambda(t) + \mu(t), \\
\lambda(t) = -\Gamma\varphi(t)\left(\varphi^{\mathrm{T}}(t)\hat{\psi}(t) - z(t)\right), \quad \hat{\psi}(t_0) \in \mathcal{D}_\psi \\
\mu(t) = \begin{cases} 0, & \text{if } \hat{\psi}_i(t) \in \left(\underline{\psi_i}, \overline{\psi_i}\right), \text{ or} \\ & \text{if } \hat{\psi}_i(t) = \overline{\psi_i} \text{ and } \lambda(t) \le 0, \text{ or} \\ & \text{if } \hat{\psi}_i(t) = \underline{\psi_i} \text{ and } \lambda(t) \ge 0, \\ -\lambda(t), & \text{otherwize.} \end{cases}
\end{gathered} \tag{5.54}$$

The law (5.54) have all properties described in Theorem 7 and additionally guarantees $\hat{\psi}(t) \in \mathcal{D}_\psi$, $\forall t \ge t_0$, which, if Lipschitz continuity inside $\mathcal{D}_\psi$ holds, allows one to obtain singularity-free estimate $\hat{T}_I(t)$. Similarly, all previously discussed estimation laws can be augmented with projection operator. The limitations of the modified law (5.54) is the need to know the lower/upper bounds $\underline{\psi_i}$, $\overline{\psi_i}$ and represent the singularity-avoidance conditions via constraints for ψ_i. Generally speaking, the obtained estimates can be projected not only onto the considered set, but any convex set [52, p. 788] defined as follows:

$$\mathcal{D}_\psi := \left\{\psi \in \mathbb{R}^{2n} : g(\psi) \le 0\right\}.$$

In such case, the estimation law (5.54) is written as:

$$\begin{gathered}
\dot{\hat{\psi}}(t) = \begin{cases} \lambda(t), \text{ if } \; g\left(\hat{\psi}\right) < 0 \text{ or } g\left(\hat{\psi}\right) = 0 \text{ and } \lambda^{\mathrm{T}}(t)\nabla g \le 0, \\ \lambda(t) - \Gamma\dfrac{\nabla g \nabla g^{\mathrm{T}}}{\nabla g^{\mathrm{T}}\Gamma\nabla g}\lambda(t), \text{ otherwise,} \end{cases} \\
\lambda(t) = -\Gamma\varphi(t)\left(\varphi^{\mathrm{T}}(t)\hat{\psi}(t) - z(t)\right), \; \hat{\psi}(t_0) \in \mathcal{D}_\psi
\end{gathered}$$

and ensures $\hat{\psi}(t) \in \mathcal{D}_\psi$, $\forall t \ge t_0$. Unfortunately, considering practical scenarios, the set of $\psi(\theta)$, for which (5.16) is singular, is often non-convex, which motivated the development of other approaches.

2) Asymptotic cascade recalculation. An approach proposed in [70] is considered to avoid singularity of the recalculation (5.16), (5.53). Assume that the singularity is caused solely by operations of division by functions from $\psi_{ab}(\theta)$ and θ. In this case, the mappings $\mathcal{F}(\theta)$ and $T_I(\theta)$

can be decomposed into a matrix numerator and denominator as follows:

$$\mathcal{S}\left(\psi_{ab}\right)=\mathcal{G}\left(\psi_{ab}\right)\mathcal{F}\left(\psi_{ab}\right)=\mathcal{G}\left(\psi_{ab}\right)\theta, \tag{5.55a}$$

$$\mathcal{P}\left(\theta\right)=\mathcal{Q}\left(\theta\right)vec\left(T_I\left(\theta\right)\right), \tag{5.55b}$$

where $\mathcal{S}:\mathbb{R}^{n_\theta}\mapsto\mathbb{R}^{n_\theta}$, $\mathcal{G}:\mathbb{R}^{n_\theta}\mapsto\mathbb{R}^{n_\theta\times n_\theta}$, $\mathcal{P}:\mathbb{R}^{n_\theta}\mapsto\mathbb{R}^{n^2}$, $\mathcal{Q}:\mathbb{R}^{n_\theta}\mapsto\mathbb{R}^{n^2\times n^2}$ stand for known mappings.

Then, having substituted $\hat{\psi}_{ab}\left(t\right)=\mathcal{L}_{ab}\hat{\psi}\left(t\right)$ into (5.55a), a regression equation is written (the estimates $\hat{\psi}\left(t\right)$ can be obtained with the help of any previously considered estimation algorithm):

$$\mathcal{Y}_\theta\left(t\right)=\mathcal{M}_\theta\left(t\right)\theta, \tag{5.56}$$

where

$$\mathcal{Y}_\theta\left(t\right):=\mathcal{S}\left(\hat{\psi}_{ab}\right),\ \mathcal{M}_\theta\left(t\right):=\mathcal{G}\left(\hat{\psi}_{ab}\right).$$

Based on (5.56), we can design, for example, the following estimation law ($\gamma>0$):

$$\dot{\hat{\theta}}\left(t\right)=-\gamma\mathcal{M}_\theta^{\mathrm{T}}\left(t\right)\left(\mathcal{M}_\theta\left(t\right)\hat{\theta}\left(t\right)-\mathcal{Y}_\theta\left(t\right)\right),\quad \hat{\theta}\left(t_0\right)=\hat{\theta}_0. \tag{5.57}$$

Now, having the estimate $\hat{\theta}\left(t\right)$ at hand and following (5.56), (5.57), the required estimate of the recalculation matrix is obtained:

$$\dot{\hat{T}}_I\left(t\right)=-\mathrm{mat}\left[\gamma\mathcal{M}_{T_I}^{\mathrm{T}}\left(t\right)\left(\mathcal{M}_{T_I}\left(t\right)\mathrm{vec}\left(\hat{T}_I\left(t\right)\right)-\mathcal{Y}_{T_I}\left(t\right)\right)\right],\quad \hat{T}_I\left(t_0\right)=\hat{T}_{I0}, \tag{5.58}$$

where

$$\mathcal{M}_{T_I}\left(t\right):=\mathcal{Q}\left(\hat{\theta}\right),\quad \mathcal{Y}_{T_I}\left(t\right):=\mathcal{P}\left(\hat{\theta}\right),$$

and mat $[.]$ denotes an operation, which is inverse to vectorization.

As, owing to (5.55a), (5.55b), $\mathcal{S}\left(\psi_{ab}\right)$, $\mathcal{G}\left(\psi_{ab}\right)$, $\mathcal{P}\left(\theta\right)$, $\mathcal{Q}\left(\theta\right)$ have no singularities for all $\psi_{ab}\in\mathbb{R}^{n_\theta}$ and $\theta\in\mathbb{R}^{n_\theta}$, then the asymtotic cascade recalculation (5.57), (5.58) ensures $\hat{T}_I\left(t\right)\to T_I\left(\theta\right)$ in case $\hat{\psi}\left(t\right)\to\psi\left(\theta\right)$ and any of the previously considered estimation laws for $\psi\left(\theta\right)$ is used.

The disadvantages of the asymptotic cascade approach include, first, the need of additional identification of parameters θ, and, second, the slow rate of convergence of $\hat{T}_I\left(t\right)$ to $T_I\left(\theta\right)$ due to the necessity of convergence, firstly, of $\hat{\psi}\left(t\right)$ to $\psi\left(\theta\right)$, then, of $\hat{\theta}\left(t\right)$ to θ, and only after that, of $\hat{T}_I\left(t\right)$ to $T_I\left(\theta\right)$, thirdly, the cascade approach provides no protection against singularities, particularly if mappings $\mathcal{S}\left(\psi_{ab}\right),\ \mathcal{G}\left(\psi_{ab}\right),\ \mathcal{P}\left(\theta\right),\ \mathcal{Q}\left(\theta\right)$ include, for example, arithmetical root operation. Unlike gradient projection method, the cascade procedure (5.57), (5.58) requires no prior knowledge of upper/lower bounds for elements of the vector $\psi\left(\theta\right)$ or convexity of $\mathcal{D}_\psi$.

3) Heterogeneous mappings. Another procedure to obtain the estimate $\hat{T}_I\left(t\right)$ of the recalculation matrix without singularities is based on the definiton and properties of heterogeneous mappings.

Definition 3. A mapping $\mathcal{F}\colon\mathbb{R}^{n_x}\mapsto\mathbb{R}^{n_\mathcal{F}\times m_\mathcal{F}}$ is heterogeneous of degree $\ell_\mathcal{F}\geq 1$ if there exist $\Pi_\mathcal{F}\left(\omega\right)\in\mathbb{R}^{n_\mathcal{F}\times n_\mathcal{F}}$, $\Xi_\mathcal{F}\left(\omega\right)=\overline{\Xi}_\mathcal{F}\left(\omega\right)\omega\in\mathbb{R}^{\Delta_\mathcal{F}\times n_x}$, and a mapping $\mathcal{T}_\mathcal{F}:\mathbb{R}^{\Delta_\mathcal{F}}\mapsto\mathbb{R}^{n_\mathcal{F}\times m_\mathcal{F}}$ such that for all $\omega\in\mathbb{R}$ and $x\in\mathbb{R}^{n_x}$ the following functional equation has a solution:

$$\Pi_\mathcal{F}\left(\omega\right)\mathcal{F}\left(x\right)=\mathcal{T}_\mathcal{F}\left(\Xi_\mathcal{F}\left(\omega\right)x\right), \tag{5.59a}$$

in such a way that:

$$\begin{gathered}\det\left\{\Pi_\mathcal{F}\left(\omega\right)\right\}\geq\omega^{\ell_\mathcal{F}}\left(t\right),\\ \Xi_{\mathcal{F}ij}\left(\omega\right)=c_{ij}\omega^{\ell_{ij}},\quad \overline{\Xi}_{\mathcal{F}ij}\left(\omega\right)=c_{ij}\omega^{\ell_{ij}-1},\\ c_{ij}\in\left\{0,1\right\},\quad \ell_{ij}\geq 1.\end{gathered} \tag{5.59b}$$

Example 4. A mapping $\mathcal{F}(\theta) = \mathrm{col}\{\theta_1\theta_2 + \theta_1, \theta_1\}$ with $\Pi_{\mathcal{F}}(\omega) = \mathrm{diag}\{\omega^2, \omega\}$, $\Xi_{\mathcal{F}}(\omega) = \mathrm{diag}\{\omega, \omega\}$ is heterogeneous of degree $\ell_{\mathcal{F}} = 3$. Thus, using a measurable signal $\mathcal{Y}_\theta(t) = \omega(t)\theta$, the main property $\Xi_{\mathcal{F}}(\omega)\theta = \overline{\Xi}_{\mathcal{F}}(\omega)\omega(t)\theta$ from Definition 3 allows one to parametrize a linear regression equation with respect to $\mathcal{F}(\theta)$ in the following way:

$$\Pi_{\mathcal{F}}(\omega)\mathcal{F}(\theta) = \mathcal{T}_{\mathcal{F}}\left(\overline{\Xi}_{\mathcal{F}}(\omega)\mathcal{Y}_\theta\right),$$
$$\begin{bmatrix}\omega^2(t) & 0\\ 0 & \omega(t)\end{bmatrix}\mathcal{F}(\theta) = \begin{bmatrix}\mathcal{Y}_{1\theta}(t)\mathcal{Y}_{2\theta}(t) + \mathcal{Y}_{1\theta}(t)\\ \mathcal{Y}_{1\theta}(t)\end{bmatrix}. \qquad \blacksquare$$

A particular case of heterogeneity conditions are homogeneity conditions, since for homogeneous mappings, in terms of equation (5.59a), it holds that

$$\Pi_{\mathcal{F}}(\omega)\mathcal{F}(x) = \mathcal{F}(\Xi_{\mathcal{F}}(\omega)x),$$
$$\Xi_{\mathcal{F}}(\omega) = \omega^\ell I_{n_x}, \quad \ell_{ij} = \ell,$$

that is, in this case, the product $\Pi_{\mathcal{F}}(\omega)\mathcal{F}(x)$ can be calculated via the function $\mathcal{F}(x)$ itself by the multiplication of its argument by $\Xi_{\mathcal{F}}(\omega)$.

In [71], it is shown that the functional equation (5.59a) has a solution for a sufficiently general algebraic function that covers all polynomials and monomials, as well as some of the irrational functions.

Proposition 1. *Let $n_{\mathcal{F}} = m_{\mathcal{F}} = 1$, then the mapping $\mathcal{F}(x) = \sum\limits_j a_j \prod\limits_{i=1}^{n_x} x_i^{k_{ji}}$, $k_{ji} \geq 0$ is heterogeneous in the sense of* (5.59a), (5.59b).

Proof of proposition is given in [71].

Let us show how the properties of heterogeneous mappings can be used to obtain an estimate of $\hat{T}_I(t)$. It should be recalled that using the dynamic regressor extension and mixing procedure (5.38), the signals $\mathcal{Y}(t)$ and $\Delta(t)$ from the following regression equation are available for measurement:

$$\mathcal{Y}(t) = \Delta(t)\psi(\theta).$$

Then, the multiplication of $\mathcal{Y}(t)$ by $\mathcal{L}_{ab}$ yields:

$$\mathcal{Y}_{ab}(t) = \Delta(t)\psi_{ab}(\theta),$$

where $\mathcal{Y}_{ab}(t) := \mathcal{L}_{ab}\mathcal{Y}(t)$.

Assume that the mappings $\mathcal{S}(\psi_{ab})$, $\mathcal{G}(\psi_{ab})$ from (5.55a) are heterogeneous in the sense of Definition 3. Then, having multiplied the left- and right-hand sides of (5.55a) by $\Pi(\Delta)$ and used the property (5.59b), it is obtained that:

$$\begin{gathered}\mathcal{T}_{\mathcal{S}}(\Xi_{\mathcal{S}}(\Delta)\psi_{ab}) = \mathcal{T}_{\mathcal{G}}(\Xi_{\mathcal{G}}(\Delta)\psi_{ab})\theta,\\ \Downarrow\\ \mathcal{T}_{\mathcal{S}}\left(\overline{\Xi}_{\mathcal{S}}(\Delta)\mathcal{Y}_{ab}\right) = \mathcal{T}_{\mathcal{G}}\left(\overline{\Xi}_{\mathcal{G}}(\Delta)\mathcal{Y}_{ab}\right)\theta,\end{gathered} \tag{5.60}$$

where the signals $\mathcal{T}_{\mathcal{S}}\left(\overline{\Xi}_{\mathcal{S}}(\Delta)\mathcal{Y}_{ab}\right)$, $\mathcal{T}_{\mathcal{G}}\left(\overline{\Xi}_{\mathcal{G}}(\Delta)\mathcal{Y}_{ab}\right)$ are measurable as $\mathcal{Y}_{ab}(t)$ and $\Delta(t)$ are measurable, and the mappings $\mathcal{T}_{\mathcal{S}}(.)$, $\mathcal{T}_{\mathcal{G}}(.)$, $\overline{\Xi}_{\mathcal{S}}(.)$, $\overline{\Xi}_{\mathcal{G}}(.)$ are known.

Having multiplied (5.60) by $\mathrm{adj}\{\mathcal{T}_{\mathcal{G}}\left(\overline{\Xi}_{\mathcal{G}}(\Delta)\mathcal{Y}_{ab}\right)\}$, a regression equation with a scalar regressor is obtained:

$$\mathcal{Y}_\theta(t) = \mathcal{M}_\theta(t)\theta, \tag{5.61}$$

where

$$\mathcal{Y}_\theta(t) := \operatorname{adj}\left\{\mathcal{T}_\mathcal{G}\left(\overline{\Xi}_\mathcal{G}(\Delta)\,\mathcal{Y}_{ab}\right)\right\}\mathcal{T}_\mathcal{S}\left(\overline{\Xi}_\mathcal{S}(\Delta)\,\mathcal{Y}_{ab}\right), \quad \mathcal{M}_\theta(t) := \det\left\{\mathcal{T}_\mathcal{G}\left(\overline{\Xi}_\mathcal{G}(\Delta)\,\mathcal{Y}_{ab}\right)\right\}.$$

Then, having assumed the heterogeneity of $\mathcal{P}(\theta)$, $\mathcal{Q}(\theta)$ and used the property (5.59b), the same derivations as in (5.60), (5.61) yields

$$\mathcal{Y}_{T_I}(t) = \mathcal{M}_{T_I}(t)\operatorname{vec}(T_I(\theta)), \tag{5.62}$$

where

$$\mathcal{Y}_{T_I}(t) := \operatorname{adj}\left\{\mathcal{T}_\mathcal{Q}\left(\overline{\Xi}_\mathcal{Q}(\mathcal{M}_\theta)\,\mathcal{Y}_\theta\right)\right\}\mathcal{T}_\mathcal{P}\left(\overline{\Xi}_\mathcal{P}(\mathcal{M}_\theta)\,\mathcal{Y}_\theta\right), \quad \mathcal{M}_{T_I}(t) := \det\left\{\mathcal{T}_\mathcal{Q}\left(\overline{\Xi}_\mathcal{Q}(\mathcal{M}_\theta)\,\mathcal{Y}_\theta\right)\right\}.$$

Based on equation (5.62), the following estimation algorithm is introduced:

$$\dot{\hat{T}}_I(t) = -\operatorname{mat}\left[\gamma\mathcal{M}_{T_I}(t)\left(\mathcal{M}_{T_I}(t)\operatorname{vec}\left(\hat{T}_I(t)\right) - \mathcal{Y}_{T_I}(t)\right)\right], \quad \hat{T}_I(t_0) = \hat{T}_{I0}. \tag{5.63}$$

Unlike the gradient projection method, the approach based on heterogeneous mappings does not require the convexity of the set $\mathcal{D}_\psi$ or knowledge of any *a priori* information about the unknown parameters $\psi(\theta)$. Unlike the cascade recalculation (5.57), (5.56), only the inverse coordinate transformation matrix is estimated without identification of θ, which ensures faster convergence.

5.3. Conclusions on Adaptive Observers Overview

Compared to robust observers, adaptive observers *i*) do not require *a priori* information about system parameters specified by parameterization (3.5) or (3.6), *ii*) to achieve the goal (2.5), instead of the restrictive condition $u \in L_2$ or the equality $u(t) = -K_r\hat{x}(t)$, they only require the boundedness of control signal and states (for a discussion on the relaxation of this condition, see Section 6.1). In comparison with the invariant observers, the adaptive ones do not require to meet the restrictive extended and standard output matching conditions, and therefore, are applicable to a wider class of systems.

The open problems of adaptive observers are *i*) that the goal (2.5) is achieved only when the finite excitation condition $\varphi \in \mathrm{FE}$ is satisfied, *i.e.*, the necessary and sufficient condition for the identifiability of unknown parameters [72], *ii*) the applicability of solutions only to systems with time-invariant parameters (for comparison, all observers considered earlier, both robust and invariant, generally do not assume the system unknown parameters to be time-invariant). Note that adaptive observers, in which parameter changes are described by a linear dynamic exosystems with known parameters and unknown initial conditions [73], do not allow the second problem to be overcome: in practice, system parameter changes do not usually fit such exosystem models.

6. SOME REMARKS ON STATE-ESTIMATE-BASED CONTROL

Having compared robust, invariant, and adaptive observers, it may seem that only robust observers do not require *a priori* boundedness of $u(t)$ and $x(t)$ (see Assumption 1) and allow one to use of the obtained estimates $\hat{x}(t)$ for feedback control purposes. Indeed, in most studies on adaptive and invariant observers, the problem of state estimation is considered separately from the control task. As an exception, we note the paper [70], which is devoted to the design of adaptive control based on an adaptive observer. Naturally, an interested reader may have the following questions. To satisfy Assumption 1, has the stabilizing controller already been designed? How is it designed if the system parameters are unknown and the state vector is unmeasurable?

On the one hand, a stabilizing controller can be designed based on the system output using, for example, robust control techniques. In this situation, as noted in Remark 1, the state observer is not aimed at control task solution, but rather at fault detection, process monitoring and logging of unmeasured variables, digital twin design, etc. On the other hand, following the principle of adaptive modular design, the estimation and control tasks in linear systems can be considered independently of each other. Let us focus on the latter point in more detail using the example of an adaptive observer consisting of equations (5.15), (5.18) at $v(t)=0$, *i.e.*

$$\begin{gathered}\hat{x}(t)=\hat{T}_I(t)\hat{\xi}(t),\\ \dot{\hat{\xi}}(t)=A_0\hat{\xi}(t)+\hat{\psi}_a(t)y(t)+\hat{\psi}_b(t)u(t)+L\left(C_0^{\mathrm{T}}\hat{\xi}(t)-y(t)\right).\end{gathered}\tag{6.1}$$

For simplicity and without loss of generality, we assume that the control signal is a robust law (2.4a). We define the parameter estimation laws $\psi_{ab}(\theta)$ and $T_I(\theta)$ by (5.38) + (5.41) and (5.63), *i.e.*:

$$\begin{aligned}\dot{\tilde{\psi}}(t)=\dot{\hat{\psi}}(t)&=-\gamma\left(\Delta(t)\hat{\psi}(t)-\mathcal{Y}(t)\right)=-\gamma\Delta(t)\tilde{\psi}(t),\\ \dot{\tilde{T}}_I(t)=\dot{\hat{T}}_I(t)&=-\mathrm{mat}\left[\gamma_{T_I}\mathcal{M}_{T_I}(t)\left(\mathcal{M}_{T_I}(t)\,vec\left(\hat{T}_I(t)\right)-\mathcal{Y}_{T_I}(t)\right)\right]\\ &=-\mathrm{mat}\left[\gamma_{T_I}\mathcal{M}_{T_I}^2(t)\,vec\left(\tilde{T}_I(t)\right)\right].\end{aligned}\tag{6.2}$$

When the condition $\varphi\in\mathrm{FE}$ is met, regardless of the boundedness of $x(t)$ and $u(t)$, it is easy to show (5.42) the existence of following lower bounds for all $t\geq t_e$:

$$|\Delta(t)|\geq\Delta_{LB}>0,\quad|\mathcal{M}_{T_I}(t)|\geq\underline{\mathcal{M}}_{T_I}>0,$$

therefore, the errors $\tilde{\psi}(t)$ and $\tilde{T}_I(t)$ converge exponentially to zero with the rate defined by parameters $\gamma>0$ and $\gamma_{T_I}>0$, *i.e.*:

$$\left\|\tilde{\psi}(t)\right\|=e^{-\gamma\Delta_{LB}(t-t_e)}\left\|\tilde{\psi}(t_e)\right\|,\quad\left\|\tilde{T}_I(t)\right\|=e^{-\gamma_{T_I}\underline{\mathcal{M}}_{T_I}^2(t-t_e)}\left\|\tilde{T}_I(t_e)\right\|.\tag{6.3}$$

The differential equations for $x(t)$ and $\tilde{\xi}(t)$ are written as follows:

$$\begin{gathered}\begin{bmatrix}\dot{x}(t)\\ \dot{\tilde{\xi}}(t)\end{bmatrix}=(\mathcal{A}+\mathcal{B}(t))\begin{bmatrix}x(t)\\ \tilde{\xi}(t)\end{bmatrix}\\ =\begin{bmatrix}A-BK_r\left(I+\tilde{T}_I(t)T(\theta)\right) & -BK_r\left(\tilde{T}_I(t)+T_I(\theta)\right)\\ \tilde{\psi}_a(t)C^{\mathrm{T}}-\tilde{\psi}_b(t)K_r\left(I+\tilde{T}_I(t)T(\theta)\right) & A_m-\tilde{\psi}_b(t)K_r\left(\tilde{T}_I(t)+T_I(\theta)\right)\end{bmatrix}\begin{bmatrix}x(t)\\ \tilde{\xi}(t)\end{bmatrix},\end{gathered}\tag{6.4}$$

where

$$\mathcal{A}=\begin{bmatrix}A-BK_r & -BK_rT_I(\theta)\\ 0 & A_m\end{bmatrix},$$

$$\mathcal{B}(t)=\begin{bmatrix}-BK_r\left(\tilde{T}_I(t)T(\theta)\right) & -BK_r\left(\tilde{T}_I(t)\right)\\ \tilde{\psi}_a(t)C^{\mathrm{T}}-\tilde{\psi}_b(t)K_r\left(I+\tilde{T}_I(t)T(\theta)\right) & -\tilde{\psi}_b(t)K_r\left(\tilde{T}_I(t)+T_I(\theta)\right)\end{bmatrix}.$$

Owing to the estimates (6.3), we have $||\mathcal{B}(t)||\to 0$ as $t\to\infty$, then, if robust control techniques were used to calculate K_r such that the matrix $A-BK_r$ is Hurwitz, then, by Lemma 9.5 from [74], we obtain global exponential convergence of states $x(t)$ and estimation errors $\tilde{\xi}(t)$, $\tilde{x}(t)$ to zero. Using adaptive feedback (2.4b), the situation is similar, except that the control scheme includes the parameter estimation law for K (e.g., for the controller (2.3)), and the matrix $\mathcal{B}(t)$ now additionally depends on the error $\tilde{K}(t)=\hat{K}(t)-K$. Thus, the adaptive observers considered in the overview, which achieve the goal (2.5), can be fully justified for application to design the control law of the type (2.4a) or (2.4b) when $\varphi\in\mathrm{FE}$.

7. CONCLUSION

Based on the stated above, the following conclusions can be made regarding the feasibility conditions of existing state observers for linear dynamical systems based on robust, invariant, and adaptive control methods:

— robust state observers require the system uncertainty to be parameterized in the form of (3.5) or (3.6). This, in turn, demands the bounds of the system parameter variation range to be known. Such observers achieve the goal if $u \in L_2$ or $u(t) = -K_r \hat{x}(t)$ and the parameter K_r is calculated jointly with the observer parameters. In the general case $u \in L_\infty$, robust observers ensure that the estimation error belongs to some invariant ellipsoid that can be minimized.

— invariant state observers solve the problem of state estimation in case of arbitrary parametric uncertainty, but they require strict structural constraints to be met—standard or extended output matching conditions. As the analysis shows, some of these conditions for systems with scalar output cannot even potentially be relaxed (in particular, the requirement that the dimension of the measured output exceed the dimension of the disturbance). Therefore, the practical application of invariant state observers is restricted to the class of systems, for which the specified strict structural constraints are satisfied.

— adaptive state observers can be divided into two groups. Methods from the first group (see Section 5.1) allow one to reconstruct the states of systems with a SPR transfer function from a parametric disturbance to a measured output, which is equivalent to the restrictive output matching conditions required for the application of invariant observers. Algorithms from the second group (see Section 5.2) reconstruct the states of linear systems with completely unknown parameters when the following requirements are met: *i*) the necessary and sufficient condition for identifiability [72] of the parameters of the observer canonical form $\psi(\theta)$, *i.e.*, $\varphi \in \mathrm{FE}$, *ii*) the sufficient condition (5.52) for the existence of the inverse function $\mathcal{F}(\psi_{ab})$, which allows the parameters $\psi(\theta)$ to be converted into physical parameters θ, *iii*) the singularity problem is avoidable using one of the techniques discussed in Section 5.2.7.

In general, research into the development of new methods to design observers for linear dynamical systems continues within all three branches of control theory discussed here, and their direction is largely driven by the existing problems mentioned in this overview, particularly in Sections 3.3, 4.4, and 5.3.

REFERENCES

1. Pacas, M., Sensorless drives in industrial applications, *IEEE Ind. Electron. Mag.*, 2011, vol. 5, no. 2, pp. 16–23.

2. Luenberger, D.G., Observing the state of a linear system, *IEEE Trans. Mil. Electron.*, 1964, vol. 8, no. 2, pp. 74–80.

3. Kalman, R.E. and Bucy, R.S., New results in linear filtering and prediction theory, *J. Basic Eng.*, 1961, vol. 83, pp. 95–108.

4. Bernard, P., Andrieu, V., and Astolfi, D., Observer design for continuous-time dynamical systems, *Annu. Rev. Control*, 2022, vol. 53, pp. 224–248.

5. Andrievsky, B. and Furtat, I., Disturbance observers: Methods and applications. I, *Autom. Remote Control*, 2020, vol. 81, no. 9, pp. 1563–1610.

6. O'Reilly, J., *Observers for Linear Systems*, Academic Press, 1983.

7. Shtessel, Y., Edwards, C., Fridman, L., and Levant, A., *Sliding Mode Control and Observation*, Springer, 2011.

8. Korovin, S.K. and Fomichev, V.V., *State Observers for Linear Systems with Uncertainty*, Walter de Gruyter, 2009.

9. Gao, H. and Li, X., *Robust Filtering for Uncertain Systems*, Springer, 2014.

10. Lewis, F.L., Xie, L., and Popa, D., *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*, CRC Press, 2017.

11. Krasnova, S.A. and Utkin, V.A., *Kaskadnoe proektirovanie nablyudatelei sostoyaniya dinamicheskikh sistem* (Cascade Design of Dynamical Systems State Observers), Moscow: Nauka, 2006.

12. Narendra, K.S. and Annaswamy, A.M., *Stable Adaptive Systems*, Courier Corporation, 2012.

13. Bhattacharyya, S.P. and de Souza, E., Pole assignment via Sylvester's equation, *Syst. Control Lett.*, 1982, vol. 1, no. 4, pp. 261–263.

14. Dudarenko, N.A., Slita, O.V., and Ushakov, A.V., Algebraic conditions of generalized modal control, *IFAC Proc. Vol.*, 2012, vol. 45, no. 13, pp. 150–155.

15. Petersen, I.R., A stabilization algorithm for a class of uncertain linear systems, *Syst. Control Lett.*, 1987, vol. 8, no. 4, pp. 351–357.

16. Geromel, J.C., Peres, P.L.D., and Bernussou, J., On a convex parameter space method for linear control design of uncertain systems, *SIAM J. Control Optim.*, 1991, vol. 29, no. 2, pp. 381–402.

17. Fu, M., de Souza, C.E., and Xie, L., $\mathcal{H}_\infty$ estimation for uncertain systems, *Int. J. Robust Nonlinear Control*, 1992, vol. 2, no. 2, pp. 87–105.

18. Xie, L. and de Souza, C.E., On robust filtering for linear systems with parameter uncertainty, *Proc. 34th IEEE Conf. Decis. Control*, 1995, vol. 2, pp. 2087–2092.

19. Geromel, J.C. and De Oliveira, M.C., $\mathcal{H}2$ and $\mathcal{H}\infty$ robust filtering for convex bounded uncertain systems, *IEEE Trans. Autom. Control*, 2001, vol. 46, no. 1, pp. 100–107.

20. Jin, S.H. and Park, J.B., Robust $\mathcal{H}2$, $\mathcal{H}\infty$ filtering for polytopic uncertain systems via convex optimisation, *IEEE Proc., Control Theory Appl.*, 2001, vol. 148, no. 1, pp. 55–59.

21. Duan, Z., Zhang, J., Zhang, C., and Mosca, E., Robust $\mathcal{H}2$ and $\mathcal{H}\infty$ filtering for uncertain linear systems, *Automatica*, 2006, vol. 42, no. 11, pp. 1919–1926.

22. Lacerda, M.J., Oliveira, R.C.L.F., and Peres, P.L.D., Robust $\mathcal{H}2$ and $\mathcal{H}\infty$ filter design for uncertain linear systems via LMIs and polynomial matrices, *Signal Process.*, 2011, vol. 91, no. 5, pp. 1115–1122.

23. Polyak, B.T., Khlebnikov, M.V., and Rappoport, L.B., *Matematicheskaya teoriya avtomaticheskogo upravleniya* (Mathematical Theory of Automatic Control), Moscow: Lenand, 2019.

24. Petersen, I.R., A stabilization algorithm for a class of uncertain linear systems, *Syst. Control Lett.*, 1987, vol. 8, no. 4, pp. 351–357.

25. Khlebnikov, M.V., Robust filtering under nonrandom disturbances: The invariant ellipsoid approach, *Autom. Remote Control*, 2009, vol. 70, no. 1, pp. 133–146.

26. Petersen, I.R., A Riccati equation approach to the design of stabilizing controllers and observers for a class of uncertain linear systems, *Proc. Am. Control Conf.*, 1985, pp. 772–777.

27. Boyd, S., El Ghaoui, L., Feron, E., et al., *Linear Matrix Inequalities in System and Control Theory*, SIAM, 1994.

28. Lien, C.H. and Yu, K.W., LMI optimization approach on robustness and $\mathcal{H}_\infty$ control analysis for observer-based control of uncertain systems, *Chaos Solitons Fractals*, 2008, vol. 36, no. 3, pp. 617–627.

29. Perruquetti, W. and Barbot, J.P., *Sliding Mode Control in Engineering*, Marcel Dekker, 2002.

30. Alwi, H., Edwards, C., and Tan, C.P., *Fault Detection and Fault-Tolerant Control Using Sliding Modes*, Springer, 2011.

31. Utkin, V., Poznyak, A., Orlov, Y., and Polyakov, A., *Road Map for Sliding Mode Control Design*, Springer, 2020.

32. Yang, F. and Wilde, R.W., Observers for linear systems with unknown inputs, *IEEE Trans. Autom. Control*, 1988, vol. 33, no. 7, pp. 677–681.

33. Darouach, M., Zasadzinski, M., and Xu, S.J., Full-order observers for linear systems with unknown inputs, *IEEE Trans. Autom. Control*, 1994, vol. 39, no. 3, pp. 606–609.

34. Dimassi, H., Loria, A., and Belghith, S., Continuously-implemented sliding mode adaptive unknown-input observers under noisy measurements, *Syst. Control Lett.*, 2012, vol. 61, no. 12, pp. 1194–1202.

35. Utkin, V.I., *Sliding Modes in Control and Optimization*, Springer, 2013.

36. Slotine, J.J.E., Hedrick, J.K., and Misawa, E.A., On sliding observers for nonlinear systems, *J. Dyn. Syst. Meas. Control*, 1987, vol. 109, pp. 245–252.

37. Edwards, C. and Spurgeon, S.K., On the development of discontinuous observers, *Int. J. Control*, 1994, vol. 59, no. 5, pp. 1211–1229.

38. Edwards, C., Spurgeon, S.K., and Patton, R.J., Sliding mode observers for fault detection and isolation, *Automatica*, 2000, vol. 36, no. 4, pp. 541–553.

39. Edwards, C. and Tan, C.P., A comparison of sliding mode and unknown input observers for fault reconstruction, *Eur. J. Control*, 2006, vol. 12, no. 3, pp. 245–260.

40. Darouach, M., Zasadzinski, M., and Boutat-Baddas, L., Discussion on "A comparison of sliding mode and unknown input observer for fault reconstruction", *Eur. J. Control*, 2006, vol. 12, no. 3, pp. 261–266.

41. Floquet, T., Edwards, C., and Spurgeon, S.K., On sliding mode observers for systems with unknown inputs, *J. Adapt. Control Signal Process.*, 2007, vol. 21, no. 8–9, pp. 638–656.

42. Isidori, A., *Lectures in Feedback Design for Multivariable Systems*, Springer, 2017.

43. Kalsi, K., Lian, J., Hui, S., and Zak, S.H., Sliding-mode observers for systems with unknown inputs: A high-gain approach, *Automatica*, 2010, vol. 46, no. 2, pp. 347–353.

44. Zhu, F., State estimation and unknown input reconstruction via both reduced-order and high-order sliding mode observers, *J. Process Control*, 2012, vol. 22, no. 1, pp. 296–302.

45. Floquet, T. and Barbot, J.-P., A canonical form for the design of unknown input sliding mode observers, in *Advances in Variable Structure and Sliding Mode Control*, 2006, vol. 334, pp. 271–292.

46. Cho, Y.M. and Rajamani, R., A systematic approach to adaptive observer synthesis for nonlinear systems, *IEEE Trans. Autom. Control*, 1997, vol. 42, no. 4, pp. 534–537.

47. Cecilia, A., and Costa-Castell, R., Addressing the relative degree restriction in nonlinear adaptive observers: A high-gain observer approach, *J. Franklin Inst.*, 2022, vol. 359, no. 8, pp. 3857–3882.

48. Antsaklis, P.J. and Michel, A.N., *Linear Systems*, McGraw-Hill, 1997.

49. Carroll, R. and Lindorff, D., An adaptive observer for single-input single-output linear systems, *IEEE Trans. Autom. Control*, 1973, vol. 18, no. 5, pp. 428–435.

50. Luders, G. and Narendra, K.S., An adaptive observer and identifier for a linear system, *IEEE Trans. Autom. Control*, 1973, vol. 18, no. 5, pp. 496–499.

51. Narendra, K.S. and Kudva, P., Stable adaptive schemes for system identification and control-Part I, *IEEE Trans. Syst. Man Cybern.*, 1974, vol. 6, pp. 542–551.

52. Ioannou, P.A. and Sun, J., *Robust Adaptive Control*, McGraw-Hill, 1996.

53. Tao, G., *Adaptive Control Design and Analysis*, Wiley, 2003.

54. Zhang, Q., Revisiting different adaptive observers through a unified formulation, *Proc. 44th IEEE Conf. Decis. Control*, 2005, pp. 3067–3072.

55. Kreisselmeier, G., Adaptive observers with exponential rate of convergence, *IEEE Trans. Syst. Man Cybern.*, 1977, vol. 22, no. 1, pp. 2–8.

56. Prokopov, B.I., On design of adaptive observers, *Autom. Remote Control*, 1981, vol. 42, no. 5, pp. 633–637.

57. Bobtsov, A., Pyrkin, A., Vedyakov, A., Vediakova, A., and Aranovskiy, S., A modification of generalized parameter-based adaptive observer for linear systems with relaxed excitation conditions, *IFAC-PapersOnLine*, 2022, vol. 55, no. 12, pp. 324–329.

58. Efimov, D. and Fradkov, A., Design of impulsive adaptive observers for improvement of persistency of excitation, *Int. J. Adapt. Control Signal Process.*, 2015, vol. 29, no. 6, pp. 765–782.

59. Gerasimov, D.N. and Nikiforov, V.O., On key properties of the Lion's and Kreisselmeier's adaptation algorithms, *Int. J. Adapt. Control Signal Process.*, 2022, vol. 36, no. 6, pp. 1285–1304.

60. Aranovskiy, S., Ushirobira, R., Korotina, M., and Vedyakov, A., On preserving excitation properties of Kreisselmeier's regressor extension scheme, *IEEE Trans. Autom. Control*, 2023, vol. 68, no. 2, pp. 1296–1302.

61. Aranovskiy, S., Bobtsov, A., Ortega, R., and Pyrkin, A., Performance enhancement of parameter estimators via dynamic regressor extension and mixing, *IEEE Trans. Autom. Control*, 2017, vol. 62, no. 7, pp. 3546–3550.

62. Baev, S., Shkolnikov, I., Shtessel, Y., and Poznyak, A., Parameter identification of non-linear system using traditional and high order sliding modes, *Proc. Am. Control Conf.*, 2006, pp. 1–6.

63. Boyd, S. and Sastry, S.S., Necessary and sufficient conditions for parameter convergence in adaptive control, *Automatica*, 1986, vol. 22, no. 6, pp. 629–639.

64. Ortega, R., Nikiforov, V., and Gerasimov, D., On modified parameter estimators for identification and adaptive control. A unified framework and some new schemes, *Annu. Rev. Control*, 2020, vol. 50, pp. 278–293.

65. Pan, Y., Shi, T., and Ortega, R., Comparative analysis of parameter convergence for several least-squares estimation schemes, *IEEE Trans. Autom. Control*, 2024, vol. 69, no. 5, pp. 3341–3348.

66. Pan, Y. and Shi, T., Adaptive estimation and control with online data memory: A historical perspective, *IEEE Control Syst. Lett.*, 2024, vol. 8, pp. 267–278.

67. Glushchenko, A., Petrov, V., and Lastochkin, K., Regression filtration with resetting to provide exponential convergence of MRAC for plants with jump change of unknown parameters, *IEEE Trans. Autom. Control*, 2023, vol. 68, no. 8, pp. 5127–5134.

68. Glushchenko, A. and Lastochkin, K., Unknown piecewise constant parameters identification with exponential rate of convergence, *Int. J. Adapt. Control Signal Process.*, 2022, vol. 37, no. 1, pp. 315–346.

69. Ortega, R., Aranovskiy, S., Pyrkin, A.A., Astolfi, A., and Bobtsov, A.A., New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases, *IEEE Trans. Autom. Control*, 2021, vol. 66, no. 5, pp. 2265–2272.

70. Kreisselmeier, G., Adaptive control via adaptive observation and asymptotic feedback matrix synthesis, *IEEE Trans. Autom. Control*, 1980, vol. 25, no. 4, pp. 717–722.

71. Glushchenko, A. and Lastochkin, K., Exponentially stable adaptive observation for systems parameterized by unknown physical parameters, *IEEE Trans. Autom. Control*, 2024, vol. 69, no. 8, pp. 5635–5642.

72. Wang, L., Ortega, R., Bobtsov, A., Romero, J.G., and Yi, B., Identifiability implies robust, globally exponentially convergent on-line parameter estimation, *Int. J. Control*, 2024, vol. 97, no. 9, pp. 1967–1983.

73. Pyrkin, A., Bobtsov, A., Ortega, R., and Isidori, A., An adaptive observer for uncertain linear time-varying systems with unknown additive perturbations, *Automatica*, 2023, vol. 147, p. 110677.

74. Khalil, H.K. and Grizzle, J.W., *Nonlinear Systems*, Prentice Hall, 2002.

75. Krasnova, S.A. and Utkin, A.V., Sigma-function for design of state and disturbance observers, *Probl. Control*, 2015, no. 5, pp. 27–36.

76. Krasnova, S.A. and Mysik, N.S., Cascade synthesis of a state observer with nonlinear correcting influences, *Autom. Remote Control*, 2014, vol. 75, no. 2, pp. 263–280.

This paper was recommended for publication by M.V. Khlebnikov, a member of the Editorial Board