Accelerating Adaptive Systems via Normalized Parameter Estimation Laws

Mohammad Boveiri[a], Mohammad Khosravi[a], Peyman Mohajerin Esfahani[a,b]

Abstract. In this paper, we propose a new class of parameter estimation laws for adaptive systems, called *normalized parameter estimation laws*. A key feature of these estimation laws is that they accelerate the convergence of the system state, $x(t)$, to the origin. We quantify this improvement by showing that our estimation laws guarantee finite integrability of the r-th root of the squared norm of the system state, i.e., $\|x(t)\|_2^{2/r} \in \mathcal{L}_1$, where $r \geq 1$ is a pre-specified parameter that, for a broad class of systems, can be chosen arbitrarily large. In contrast, standard Lyapunov-based estimation laws only guarantee integrability of $\|x(t)\|_2^2$ (i.e., $r = 1$). We motivate our method by showing that, for large values of r, this guarantee serves as a sparsity-promoting mechanism in the time domain, meaning that it penalizes prolonged signal duration and slow decay, thereby promoting faster convergence of $x(t)$. The proposed estimation laws do not rely on time-varying or high adaptation gains and do not require persistent excitation. Moreover, they can be applied to systems with matched and unmatched uncertainties, regardless of their dynamic structure, as long as a control Lyapunov function (CLF) exists. Finally, they are compatible with any CLF-based certainty equivalence controllers. We further develop higher-order extensions of our estimation laws by incorporating momentum into the estimation dynamics. We illustrate the performance improvements achieved with the proposed scheme through various numerical experiments.

1. Introduction

Adaptive control is a powerful methodology that enables real-time learning, i.e., model parameter estimation, and performance improvement in dynamic systems while ensuring stability. Due to this key feature, it has been extensively studied for decades [3, 19, 22, 43], with applications across diverse fields, including robotics [45, 50, 38, 49], aerospace engineering [16, 8, 58], automated driving [53, 25], and energy systems [40, 38].

The primary objective in adaptive control is the convergence of the system state to a desired value or trajectory, whereas the parameter convergence, while desirable, is of secondary importance. Parameter convergence occurs only under a strong condition known as persistent excitation (PE) [33, 43], which typically arises from sufficiently rich reference inputs.

In the presence of PE, the system state and parameter estimate converge respectively to their desired and actual values with a fast rate, typically exponential [19, 5, 44, 30]. Nonetheless, in most control tasks, such as stabilization and set-point regulation, the PE property is generally absent. In standard Lyapunov-based adaptive control design, without the PE condition, the parameter estimate does not converge to the true value, and the convergence of the system state can be slow. The primary reason for slow convergence in the absence of PE is that parameter adaptation decelerates as the system state approaches the origin. Typically, in this scenario, it can only be shown that the squared

1

norm of the system state asymptotically converges to zero and belongs to the space of integrable functions. The reader is referred to Section 2 for a detailed discussion of this topic.

Related works. Research on accelerating the convergence of adaptive systems can be broadly divided into the categories discussed below.

(i) **Convergence via altering the reference signal.** A common approach to accelerate convergence is to enrich the reference signal, for instance by adding time-varying noise, thereby trying to enforce PE and accelerate the convergence of both the parameter estimate and the system state [12, 11, 32]. However, this comes at the cost of compromising the transient behavior of the system. Moreover, if the PE of the reference signal does not imply the PE of the regressor vector, this approach does not necessarily accelerate convergence [20].

(ii) **Parameter estimator design for a given controller.** This approach, which is conceptually closer to our work, focuses on parameter estimation algorithms without modifying the control law. In line with recent advances in optimization and machine learning [4, 36, 34], this line of research incorporates ideas such as natural gradient and mirror descent-like algorithms [37, 24, 57], as well as momentum-based methods [18], to exploit the geometry of the problem and improve convergence. However, to the best of our knowledge, the improvements are primarily empirical. More precisely, the theoretical guarantees remain similar to those of standard Lyapunov-based parameter estimation algorithms; namely, asymptotic convergence of the error to zero, with the squared error signal belonging to the space of integrable functions.

(iii) **Joint controller and parameter estimator design.** This category includes works that design new adaptive controllers by modifying both the parameter estimation law and the control law, mainly for specific classes of nonlinear systems. For instance, the method in [47, 46] applies a speed transformation to construct an accelerated (virtual) system, whose stabilization enables arbitrary adjustment of the convergence rate in the original system. This approach ensures exponential convergence without requiring persistent excitation. However, it relies on time-varying adaptation and control gains that grow exponentially over time, and is only applicable to certain system classes, such as those in strict-feedback form. Similarly, the works in [7, 6] propose adaptive controllers for feedback-linearizable and strict-feedback systems, guaranteeing convergence to a predefined residual set at a prescribed rate. However, this guarantee is non-asymptotic, i.e., the specified convergence rate does not apply when the error signal is close to zero, namely within the residual set. Adaptive fixed-time stabilization [31, 56] is also proposed for a specific class of nonlinear systems, such as nonlinear systems with linear growth conditions.

Contributions. This study conceptually falls into category (ii), where we propose a new class of parameter estimation laws, called *normalized parameter estimation laws*, designed to accelerate convergence in adaptive systems. The introduced estimation laws incorporate a form of normalization based on the Lyapunov function, enabling faster adaptation when the system state is close to the origin. We note that our parameter update laws do not rely on persistent excitation, time-varying adaptation gains, or side information. Additionally, they have the following distinct features:

(i) **Promoting sparsity in time-domain: Penalizing signal duration and slow decay.** We introduce the notion of the *vanishing degree vd* (Definition 2) for a given nonlinear system, an intrinsic property of a system that characterizes its behavior around the origin. We show that for any $r \leq vd$, our estimation laws ensure rth-root integrability of the squared norm of the system state, i.e., $\|x(t)\|_2^{2/r} \in \mathcal{L}_1$ (Theorem 3). We demonstrate that, for large values of r, this guarantee serves as a form of sparsity-promoting mechanism in the time domain, meaning it penalizes signal duration and

slow decay of $x(t)$, thereby accelerating convergence of the system state. This effect becomes more pronounced as r increases.

Further, we show that for a considerably large class of nonlinear systems, namely systems that are exponentially stabilizable and have a regressor that is zero at the origin, the vanishing degree vd is unbounded (Theorem 4), allowing the parameter r to be chosen arbitrarily large.

(ii) **Compatibility with certainty-equivalence CLF-based controllers.** Our parameter estimation laws are applicable to systems with both matched and unmatched uncertainties. Moreover, they can be seamlessly integrated into any certainty-equivalence Lyapunov-based control scheme..

(iii) **Extension to higher-order estimation laws.** We further develop higher-order extensions of our estimation laws by incorporating momentum into the update dynamics (Theorem 5). We prove that these momentum-based algorithms are stable and globally convergent.

Structure of the paper. The rest of this paper is organized as follows. Section 2 presents a motivating example of a scalar system, illustrating how our proposed parameter estimation law accelerates convergence. Section 3 generalizes this result to a multi-dimensional nonlinear system with unmatched uncertainty. In Section 4 we propose a higher-order extension of our estimation law incorporating momentum. Section 5 provides the technical proofs of theorems and propositions presented in the paper. Simulation results in Section 6 demonstrate the performance improvements achieved with the proposed method. Finally, in Section 7, we discuss some conclusions and future directions.

Notation. Throughout the paper, $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ denote the set of real numbers and non-negative real numbers, respectively. The symbol $\mathcal{L}_p$ denotes the set of functions $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that $\int_0^\infty \|z(t)\|_2^p \, dt < \infty$. Given a real-valued function $V : \mathbb{R}^n \to \mathbb{R}$, we denote the vector of partial derivatives by $\frac{\partial V(x)}{\partial x} = \left[\frac{\partial V}{\partial x_1}, \ldots, \frac{\partial V}{\partial x_s} \right]$. A function V is said to be of class C^r if all its partial derivatives up to order r exist and are continuous. Furthermore, a function is called smooth, or of class C^∞, if its partial derivatives of all orders exist and are continuous. A function $V(x)$ is positive definite if $V(0) = 0$ and $V(x) > 0$ for all $x \neq \mathbf{0}$. Moreover, $V(x)$ is radially unbounded if $V(x) \to \infty$ as $\|x\| \to \infty$. The span of a set of vectors $\{v_1, \ldots, v_m\}$ is given by

$$\text{span}\{v_1, \ldots, v_m\} = \{ v \mid \exists \lambda_1, \ldots, \lambda_m \in \mathbb{R}, \quad \text{s.t.} \quad v = \lambda_1 v_1 + \cdots + \lambda_m v_m \}.$$

For symmetric matrices A and B, the notation $A < B$ (resp. $A \leq B$) means that $A - B$ is negative definite (resp. negative semi-definite). Furthermore, I_n denotes the identity matrix of size $n \in \mathbb{N}$.

2. A Motivating Example and Preliminary Result

In this section, a motivating example is provided to briefly demonstrate the core idea of this article, namely, how suitable Lyapunov-based normalization of the parameter estimation law accelerates the convergence of the system state to the origin. To this end, we consider the scalar nonlinear system described by the differential equation

$$\dot{x} = \theta|x| + u, \tag{1}$$

where x is the state, u is the input, and θ is an unknown parameter. The goal is to design an adaptive controller to regulate the state to the origin. To better understand our proposed method and enable direct comparison, we begin by reviewing the standard direct adaptive control design procedure based on Lyapunov theory and elaborate on how it forms the foundation of our approach.

2.1. Standard Lyapunov-based estimation law.

To design an adaptive controller for system (1), we use an estimate of the unknown parameter, denoted as $\hat{\theta}(t) \in \mathbb{R}$. Adding and subtracting $\hat{\theta}x^2$ to the right-hand side of the above equation, we can rewrite the system as

$$\dot{x} = \hat{\theta}|x| + u - (\hat{\theta} - \theta)|x|. \tag{2}$$

Consider the following augmented Lyapunov function

$$Q(x, \tilde{\theta}) := V(x) + \frac{1}{2\gamma}\tilde{\theta}^2,$$

where $\gamma > 0$, $\tilde{\theta} = \hat{\theta}(t) - \theta$ represents the parameter estimation error, and $V(x) := \frac{1}{2}x^2$ is a control Lyapunov function (CLF) for system (2), as detailed in Definition 1. Taking the time derivative of $Q(x, \hat{\theta})$ along (2) yields

$$\dot{Q}(x, \tilde{\theta}) = \frac{\partial V(x)}{\partial x}\left(\hat{\theta}|x| + u - \tilde{\theta}|x|\right) + \frac{1}{\gamma}\tilde{\theta}\dot{\hat{\theta}}.$$

By selecting the control law

$$u = -\hat{\theta}|x| - \frac{1}{2}x, \tag{3}$$

and estimation law

$$\dot{\hat{\theta}} = \frac{\partial V(x)}{\partial x}|x| = \gamma\,\text{sign}(x)\,x^2, \tag{4}$$

we obtain

$$\dot{Q} \leq -V(x) \leq 0. \tag{5}$$

Thus, $Q(x, \hat{\theta})$ is non-increasing, which implies that $x(t)$ and $\hat{\theta}(t)$ are bounded. Furthermore, we have

$$\dot{V}(x) = -\frac{1}{2}x^2 - \tilde{\theta}\,\text{sign}(x)\,x^2,$$

which shows that $\dot{V}(x)$ is also bounded. As a result, $V(x)$ is uniformly continuous. Moreover, by integrating both sides of inequality (5) with respect to time, we obtain

$$\int_0^\infty V(x)\mathrm{d}t \leq Q\big(x(0), \hat{\theta}(0)\big) - Q\big(x(\infty), \hat{\theta}(\infty)\big) \leq Q\big(x(0), \hat{\theta}(0)\big) < \infty,$$

implying that $V(x)$ is integrable, i.e., $V(x) \in \mathcal{L}_1$. Thus, $x(t) \in \mathcal{L}_2$. Barbalat's Lemma, presented below, then guarantees that $V(x)$ convergence to zero as t goes to infinity, given its uniform continuity and integrability.

Lemma 1 (Barbalat's Lemma [22, 43]). *Let $v(t) \in \mathcal{L}_1$ be a uniformly continuous function defined on $[0, \infty)$. Then, $\lim_{t \to \infty} v(t) = 0$.*

Firstly, note that if the system parameter θ is exactly known and we set $\hat{\theta} = \theta$, then the controller (3) guarantees exponential convergence of the system state since $\tilde{\theta} = 0$. However, when θ is unknown, as discussed above, we can only establish the asymptotic convergence of $x(t)$ to zero and its square integrability, i.e., $x(t) \in \mathcal{L}_2$. The main reason for this slower convergence in the adaptive case is that as $x(t)$ approaches zero, the adaptation law (4) slows down significantly, which results in slower parameter adjustment and consequently slower convergence of the system state.

(a) Evolution of $\log V(x)$ (b) $x(t)$ for different values of r.   (c) $u(t)$ for different values of r.
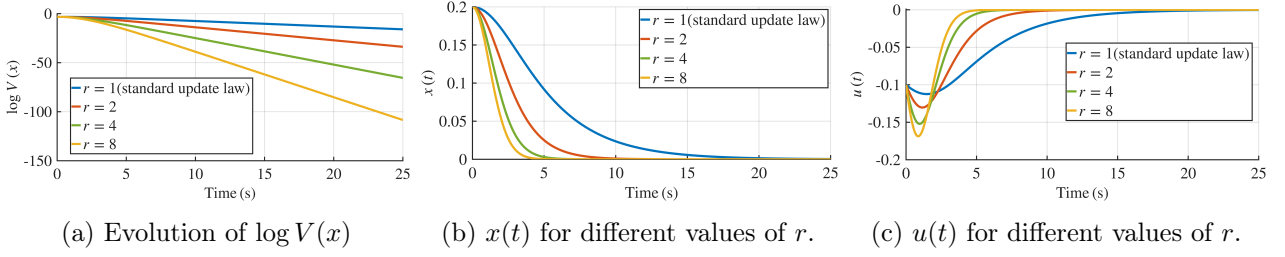
Figure 1. (a) Plots of the logarithm of the Lyapunov function $V(x) = \frac{1}{2}x^2$ versus time for system (1) under control law (3) and adaptive law (6) with $\gamma = 1$ and different values of r. (b) Closed-loop system state variable. (c) Control signal $u(t)$ for different values of r.

2.2. Normalized parameter estimation law

In this subsection, we provide an alternative parameter estimation law for system (1) that accelerates the convergence of the system state to zero.

Proposition 2 (Normalized estimation law for (1)). *Consider the nonlinear system* (1) *with the control law* (3) *and the estimation law*

$$\dot{\hat{\theta}} = \gamma \operatorname{sign}(x)\big(x^2 + |x|^{2/r}\big), \tag{6}$$

where $r \geq 1$. Then $x(t)$ asymptotically converges to zero. Moreover, we have $V(x)^{1/r} \in \mathcal{L}_1$.

The proof of Proposition 2 is provided in Section 5.1.

Remark 1 (Standard vs. normalized estimation law). A direct comparison between the standard estimation law (4) and the estimation law (6) reveals the presence of an additional term of the form $|x|^{2/r}$ in (6). This term prevents the adaptation from slowing down too quickly as the system state converges to zero, resulting in faster parameter adjustment and improved convergence of the system state. This effect becomes more pronounced as the parameter r increases. We illustrate this behavior in Figure 2, which corresponds to simulations where the control law (3) and estimation law (6) are applied to system (1). Figure 1a shows the logarithm of $V(x) = \frac{1}{2}x^2$ as a function of time, Figure 1b shows the evolution of the system state, and Figure 1c shows the control signal $u(t)$ applied to the system. We observe that increasing r accelerates the convergence of the system state to the origin.

A key feature of the update law (6) is that it ensures r-th root integrability of $V(x)$, that is $V(x)^{1/r} \in \mathcal{L}_1$, rather than absolute integrability, i.e., $V(x) \in \mathcal{L}_1$, established in the previous section. We next outline the implications and properties of this guarantee.

Monotonicity. Let $\nu(t)$ be a nonnegative function defined on $[0, \infty)$. For any $r_1 \geq r_2 \geq 1$, if $\nu(t)^{1/r_1} \in \mathcal{L}_1$, then $\nu(t)^{1/r_2} \in \mathcal{L}_1$. In other words, the condition $\nu(t)^{1/r} \in \mathcal{L}_1$ becomes stronger as r increases. To build intuition, consider $\nu(t) = \frac{1}{(t+1)^q}$ with $q \geq 1$. We have $\nu(t)^{1/r} \in \mathcal{L}_1$ if and only if $q > r$. When $q \leq r$, the decay of $\nu(t)$ is too slow and the integral diverges.

Promoting sparsity: Penalizing signal duration for large r. As $r \to \infty$, we have $\nu(t)^{1/r} \to \mathbf{1}_{\{\nu(t) \neq 0\}}$ where $\mathbf{1}_{\{\nu(t) \neq 0\}}$ is the indicator function defined as

$$\mathbf{1}_{\{\nu(t) \neq 0\}} = \begin{cases} 1, & \nu(t) \neq 0, \\ 0, & \nu(t) = 0. \end{cases}$$

Hence, the integral

$$\int_0^\infty \mathbf{1}_{\{\nu(t)\neq 0\}}\mathrm{d}t = \int_0^\infty \lim_{r\to\infty} \nu(t)^{1/r}\mathrm{d}t$$

measures the length of the time interval where $\nu(t)$ is nonzero. Therefore, if this integral is bounded, it means that $\nu(t)$ is nonzero only for a finite amount of time.

Now consider the case where $\nu(t) = V(x(t)) = \frac{1}{2}x(t)^2$. While Proposition 2 does not guarantee the integrability of $V(x)^{1/r}$ as $r \to \infty$, it does guarantee the integrability of $V(x)^{1/r}$ for any finite r. Hence, increasing r can be viewed as a sparsity-promoting mechanism as it penalizes signals with long duration and slow decay. It is interesting to note that, as r increases, the update law (6) approaches the limiting form

$$\lim_{r\to\infty} \dot{\hat{\theta}} \;=\; \gamma \operatorname{sign}(x)\Big(x^2 + \mathbf{1}_{\{x(t)\neq 0\}}\Big).$$

This estimation law maintains a constant adaptation rate even when the system state is close to zero, and switches off adaptation only when the state is exactly zero.

Example 1 (Estimation laws with smooth limits)**.** The limiting forms of the normalized estimation laws (i.e., as $r \to \infty$) depend on the system dynamics and can be smooth. For instance, consider the system

$$\dot{x} = \theta x^2 + u,$$

with the control and estimation laws given by

$$u = -\hat{\theta}x^2 - \tfrac{1}{2}x, \quad \dot{\hat{\theta}} = \gamma\big(x^3 + x\,|x|^{2/r}\big),$$

respectively. Following a similar line of reasoning as in the proof of Proposition 2, one can show that $x(t) \to 0$ as $t \to \infty$. Moreover, $\|x(t)\|_2^{2/r} \in \mathcal{L}_1$. For this system, the limiting form of the estimation law is

$$\lim_{r\to\infty} \dot{\hat{\theta}} = \gamma(x^3 + x),$$

which is smooth.

3. Accelerating Adaptation via Normalized Parameter Estimation Laws

In this section, we generalize the results and ideas from the previous section to the case of multidimensional nonlinear dynamical systems. In our analysis, we consider both matched and unmatched uncertainties.

Consider an uncertain dynamical system of the form

$$\dot{x} = f(x) + \Delta(x)^\top\theta + B(x)u, \tag{7}$$

where $x \in \mathbb{R}^n$ is the system state, $u \in \mathbb{R}^m$ is the control input, $\theta \in \Theta \subseteq \mathbb{R}^p$ is an unknown vector of fixed parameters characterizing the uncertainty in the dynamics of the system, Θ is a convex set, and $f : \mathbb{R}^n \to \mathbb{R}^n$, $B : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, and $\Delta : \mathbb{R}^n \to \mathbb{R}^{p\times n}$ are known functions. For ease of discussion, we focus here on the stabilization problem of the above system, i.e., driving the system state to the origin, $x = 0$. In such problems, persistent excitation is typically absent due to the lack of richness of the input signal.

We begin by recalling the definition of control Lyapunov functions for the introduced system (see, e.g., [21, 22])

Definition 1 (Control Lyapunov function). *A smooth and positive definite function $V(x, \theta)$ is an exponential control Lyapunov function (CLF) for* (7) *if there exist positive constants c_1, c_2, and λ such that*

$$c_1 \|x\|_2^2 \leq V(x, \theta) \leq c_2 \|x\|_2^2, \tag{8}$$

$$\inf_{u \in \mathbb{R}^m} \Big[\frac{\partial V(x, \theta)}{\partial x} \Big(f(x) + \Delta(x)^\top \theta + B(x) u \Big) \Big] \leq -\lambda V(x, \theta), \tag{9}$$

for all $x \in \mathbb{R}^n$ and $\theta \in \Theta$.

We note that the existence of a CLF is a necessary and sufficient condition for the exponential stabilizability of a dynamical system [21, Theorem 4.14] [2, 48]. Depending on the structure of the system, CLFs can be constructed using various methods, including physics-based approaches [45, 50, 38], numerical techniques, such as semidefinite and sum-of-squares programming [13, 14, 51, 29], and control design tools, like back-stepping and feedback linearization [22, 42, 17].

CLF-based Controllers. Let $V(x, \theta)$ be a CLF for the nonlinear system described by (7), for example, obtained through the aforementioned methodologies. Let $\hat{\theta}(t)$ be an estimate of the unknown parameter θ at time t, and assume that a controller $u(t)$ is in place such that the inequality

$$\frac{\partial V(x, \hat{\theta})}{\partial x} \Big(f(x) + \Delta(x)^\top \hat{\theta} + B(x) u \Big) \leq -\lambda V(x, \hat{\theta}) \tag{10}$$

holds. Note that if $\hat{\theta}(t) = \theta$, i.e., when the estimated parameter vector exactly matches the true system parameter vector, then the above controller ensures that $V(x, \hat{\theta})$ converges to zero exponentially. Such controllers are referred to as *certainty-equivalence controllers* [22], as they are derived by simply replacing θ with $\hat{\theta}$. There are several well-known controllers that satisfy the inequality (10). One example is the point-wise minimum norm controller [41], defined by

$$u(x, \hat{\theta}) = \underset{\bar{u} \in \mathbb{R}^m}{\operatorname{argmin}} \Big\{ \bar{u}^\top \bar{u} \ \Big| \ \text{subject to (10)} \Big\}. \tag{11}$$

Parameter estimation law. Similar to Section 2, our main objective is to develop an update law for the parameter estimate $\hat{\theta}$ in (10) that guarantees finite integrability of $\|x\|_2^{2/r}$ for a prespecified constant $r \in \mathbb{R}_{>0}$. Achieving this requires characterizing how the regressor $\Delta(x)$ converges to zero with respect to a given CLF. To this end, we introduce the notion of the *vanishing degree* of the regressor $\Delta(x)$, which is an intrinsic attribute of the system.

Definition 2 (Vanishing degree). *Given the CLF $V(x, \theta)$, we define* vanishing degree, *denoted by vd, of the regressor $\Delta(x)$ with respect to V as*

$$vd(\Delta, V) := \sup \Big\{ r \in \mathbb{R}_{>0} \ \Big| \ \lim_{x \to 0} \Big\| V(x, \theta)^{1/r - 1} \frac{\partial V(x, \theta)}{\partial x} \Delta(x)^\top \Big\|_\infty < \infty, \ \ \forall \theta \in \Theta \Big\}.$$

The following example further illustrates the notion of vanishing degree discussed above.

Example 2 (Vanishing degree). Consider the scalar system

$$\dot{x} = \Delta(x)^\top \theta + u,$$

with the regressor term defined as

$$\Delta(x) = [\sin(x), \ x, \ x^2]^\top,$$

and the CLF $V(x) = x^2$. For the introduced system, we have

$$\lim_{x \to 0} V(x)^{1/r-1} \frac{\partial V(x)}{\partial x} \Delta(x) = \lim_{x \to 0} |x|^{2/r} \begin{bmatrix} \frac{\sin x}{x} \\ 1 \\ x \end{bmatrix} = 0,$$

for all $r \in \mathbb{R}_{>0}$. Therefore, the vanishing degree of this system is unbounded, i.e., $vd(\Delta, V) = \infty$. We have a similar property for higher-dimensional systems, provided that $\Delta(x)$ approaches zero at least linearly as x goes to zero (see Theorem 4 for more details). On the other hand, if $\Delta(x)$ includes any component that is nonzero at the origin, such as $\cos(x)$ or a nonzero constant, then the vanishing degree is $vd(\Delta, V) = 2$.

The following theorem establishes an update law for $\hat{\theta}(t)$ that ensures $\|x(t)\|^{2/r} \in \mathcal{L}_1$ for $r \leq vd(\Delta, V)$.

Theorem 3 (Normalized parameter estimation law). *Let $\varepsilon > 0$ be a scalar, $\Gamma \in \mathbb{R}^{p \times p}$ a positive definite matrix, and $\omega : \mathbb{R} \to \mathbb{R}_{>0}$ a strictly positive and strictly increasing function, with derivative denoted by $\omega_\rho(\rho) := \frac{\partial \omega(\rho)}{\partial \rho}$. Let $u(t)$ be any control signal satisfying (10), with the parameter estimation law defined as*

$$\dot{\hat{\theta}} = \Gamma \, \omega(\rho) \, V(x, \hat{\theta})^{1/r-1} \, \Delta(x) \frac{\partial V(x, \hat{\theta})}{\partial x}^\top, \tag{12a}$$

$$\dot{\rho} = -\frac{\omega(\rho) \, V(x, \hat{\theta})^{1/r-1}}{\omega_\rho(\rho) \left(rV(x, \hat{\theta})^{1/r} + \varepsilon \right)} \frac{\partial V(x, \hat{\theta})}{\partial \hat{\theta}} \dot{\hat{\theta}}, \tag{12b}$$

where r is a positive real scalar such that $r \leq vd(\Delta, V)$. Then the following properties hold:

(1) *All closed-loop signals remain bounded.*
(2) *$x(t) \to 0$ as $t \to \infty$. Moreover, $\|x(t)\|_2^{2/r} \in \mathcal{L}_1$.*

The proof of Theorem 3 is provided in Section 5.2.

Remark 2 (Simpler update law for matched uncertainty). When the uncertainty satisfies the *matching condition*

$$\Delta(x)^\top \theta \in \text{span}\{b_1(x), \dots, b_m(x)\}, \tag{13}$$

where $b_1(x), \dots, b_m(x)$ are the columns of input matrix $B(x)$, parameter-independent CLFs can be constructed for (7). Specifically, any CLF $V(x)$ for the nominal system

$$\dot{x} = f(x) + B(x)u,$$

remains valid for the uncertain system (7), since the term $\Delta(x)^\top \theta$ can be canceled by the control input. As a result, in the matched uncertainty case, the update law (12b) becomes zero, i.e., the update law (12b) is only active when the CLF depends on the system parameters. The main idea behind the update law (12b), inspired by [27], is to modulate the adaptation rate so that transients in parameter estimation do not compromise stability.

Note that the parameter r in Theorem 3 is bounded above by the vanishing degree. In the next theorem, we show that for a broad class of dynamical systems, the vanishing degree is infinite, i.e., $vd(\Delta, V) = \infty$. Consequently, for these systems, r can be chosen arbitrarily large.

Theorem 4 (Computing vanishing degree). *Let $V(x, \theta)$ be a CLF for system (7), and the regressor $\Delta(x)$ be twice continuously differentiable, i.e., $\Delta(x) \in C^2$, with $\Delta(0) = 0$. Then, the corresponding vanishing degree is infinite, i.e., $vd(\Delta, V) = \infty$.*

The proof of Theorem 5 is provided in Section 5.3. We note that the condition $\Delta(0) = 0$ mentioned in the above theorem is common in adaptive stabilization of nonlinear systems, e.g., see [23, 15, 22].

4. Normalized estimation laws with momentum

In this section, we propose a new parameter estimation law for nonlinear systems with unmatched uncertainty, incorporating the normalization technique from the previous section along with a momentum term. The developed estimation law introduces second-order dynamics for the parameter estimate $\hat{\theta}(t)$, similar to the momentum-based algorithms commonly used in the optimization literature [35, 13]. We prove general conditions under which these momentum algorithms provide guarantees of the form $\|x(t)\|^{2/r} \in \mathcal{L}_1$ for any prespecified $r \in \mathbb{R}_{>0}$.

While the results in this section can be stated more generally, for ease of discussion, we focus on the case where a quadratic CLF can be found for system (7). This assumption is closely related to the concept of quadratic stabilizability discussed in [14, 39]. The following theorem introduces a higher-order update law for $\hat{\theta}(t)$, which ensures $\|x(t)\|^{2/r} \in \mathcal{L}_1$ for any $r \in \mathbb{R}_{>0}$.

Theorem 5 (Normalized estimation law with momentum). *Suppose that there exists a quadratic CLF of the form $V(x, \theta) = \frac{1}{2} x^\top P(\theta) x$ for system (7) and The regressor $\Delta(x)$ satisfies the conditions stated in Theorem 4. Also, let ε, Γ, and ω be specified as in Theorem 3. Then, any control signal $u(t)$ satisfying (10) and the estimation laws defined as*

$$\dot{\hat{a}} = \Gamma \, \omega(\rho) V(x, \hat{\theta})^{1/r-1} \Delta(x) \frac{\partial V(x, \hat{\theta})}{\partial x}^\top, \tag{14a}$$

$$\dot{\hat{\theta}} = \frac{1}{\lambda} \Gamma \, \omega(\rho) \, V(x, \hat{\theta})^{1/r-1} \Delta(x) P(\hat{\theta}) \Delta(x)^\top (\hat{a} - \hat{\theta}), \tag{14b}$$

$$\dot{\rho} = -\frac{\omega(\rho) V(x, \hat{\theta})^{1/r-1}}{\omega_\rho(\rho) \left(r V(x, \hat{\theta})^{1/r} + \varepsilon \right)} \frac{\partial V(x, \hat{\theta})}{\partial \hat{\theta}} \dot{\hat{\theta}}, \tag{14c}$$

The resulting closed-loop signals remain bounded and $x(t) \to 0$ as $t \to \infty$. Moreover, $\|x\|_2^{2/r} \in \mathcal{L}_1$.

The proof of Theorem 5 is provided in Section 5.4. Similar to the update laws presented in Theorem 3, when the uncertainty is matched, the update law (14c) becomes inactive, resulting in simpler update dynamics.

Remark 3 (Unmatched uncertainty and fast convergence). Compared to existing higher-order parameter estimation laws [37, 18], the estimation law proposed in Theorem 5 offers two key advantages. First, it can handle unmatched uncertainties. Second, it guarantees that $\|x(t)\|_2^{2/r} \in \mathcal{L}_1$ for any $r \in \mathbb{R}_{>0}$, whereas the methods in [37, 18] achieve only $\|x(t)\|_2^2 \in \mathcal{L}_1$ (i.e., the case $r = 1$).

5. Technical Proofs

In this section, we provide the proofs of the theorems and propositions presented in the preceding sections.

5.1. Proof of Proposition 2

Consider the nonnegative function $Q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as

$$Q(x, \hat{\theta}) = x^2 + r|x|^{2/r} + \frac{1}{\gamma}\tilde{\theta}^2. \tag{15}$$

By taking the time derivative of Q along the dynamics (1), we find that

$$\dot{Q} = 2(x + \text{sign}(x)|x|^{2/r-1})(\hat{\theta}|x| + u) - 2\tilde{\theta}\,\text{sign}(x)(x^2 + |x|^{2/r}) + \frac{2}{\gamma}\tilde{\theta}\dot{\hat{\theta}}.$$

Substituting the control law (3) and update law (4) into the above inequality gives

$$\dot{Q} \leq -x^2 - |x|^{2/r} \leq 0 \tag{16}$$

which means that $x(t)$ and $\hat{\theta}(t)$ are bounded. By integrating both sides of the above inequality with respect to time, we obtain

$$\int_0^{\infty} x^2(t) + |x(t)|^{2/r}\,\mathrm{d}t \leq Q\big(x(0), \theta(\hat{0})\big) - Q\big(x(\infty), \theta(\hat{\infty})\big)$$

$$\leq Q\big(x(0), \theta(\hat{0})\big) < \infty,$$

meaning that $|x(t)|^{2/r}$ is integrable, i.e., $|x(t)|^{2/r} \in \mathcal{L}_1$. Moreover, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}r|x|^{2/r} &= 2\text{sign}(x)|x|^{2/r-1}(\hat{\theta}|x| + u) - 2\tilde{\theta}\text{sign}(x)|x|^{2/r} \\
&= -|x|^{2/r} - 2\tilde{\theta}\text{sign}(x)|x|^{2/r}.
\end{aligned} \tag{17}$$

Since $x(t)$ and $\hat{\theta}(t)$ are bounded, the right-hand side of (17) is bounded. As a result, $|x(t)|^{2/r}$ is uniformly continuous. Hence, by Barbalat's lemma, $x(t)$ converges to zero as $t \to \infty$.

5.2. Proof of Theorem 3

Let the non-negative function $Q : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$ be defined as

$$Q\big(x, \hat{\theta}, \rho\big) = \omega(\rho)\big(rV(x, \hat{\theta})^{1/r} + \varepsilon\big) + \frac{1}{2}\tilde{\theta}^{\top}\Gamma^{-1}\tilde{\theta},$$

Taking the time derivative of Q along the dynamics (7) gives

$$\begin{aligned}
\dot{Q} &= \omega(\rho)V(x, \hat{\theta})^{1/r-1}\frac{\partial V(x, \hat{\theta})}{\partial x}\Big(f(x) + \Delta(x)^{\top}\theta + B(x)u\Big) + \omega(\rho)V(x, \hat{\theta})^{1/r-1}\frac{\partial V(x, \hat{\theta})}{\partial \theta}\dot{\hat{\theta}} \\
&\quad + \frac{\partial \omega(\rho)}{\partial \rho}\dot{\rho}\big(rV(x, \hat{\theta})^{1/r} + \varepsilon\big) + \tilde{\theta}^{\top}\Gamma^{-1}\dot{\hat{\theta}}.
\end{aligned}$$

By rearranging the terms, it follows that

$$\begin{aligned}
\dot{Q} &= \omega(\rho)V(x, \hat{\theta})^{1/r-1}\frac{\partial V(x, \hat{\theta})}{\partial x}\Big(f(x) + \Delta(x)^{\top}\hat{\theta} + B(x)u\Big) - \omega(\rho)V(x, \hat{\theta})^{1/r-1}\frac{\partial V(x, \hat{\theta})}{\partial x}\Delta(x)^{\top}\tilde{\theta} \\
&\quad + \omega(\rho)V(x, \hat{\theta})^{1/r-1}\frac{\partial V(x, \hat{\theta})}{\partial \theta}\dot{\hat{\theta}} + \tilde{\theta}^{\top}\Gamma^{-1}\dot{\hat{\theta}} + \frac{\partial \omega(\rho)}{\partial \rho}\dot{\rho}\big(rV(x, \hat{\theta})^{1/r} + \varepsilon\big).
\end{aligned} \tag{18}$$

Substituting u in the right hand sight of (18) with a control law satisfying (10) gives

$$
\begin{aligned}
\dot{Q} \leq & -\omega(\rho)\lambda V(x,\hat{\theta})^{1/r} - \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\Delta(x)^\top\tilde{\theta} \\
& + \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial\theta}\dot{\hat{\theta}} + \tilde{\theta}^\top\Gamma^{-1}\dot{\hat{\theta}} + \frac{\partial\omega(\rho)}{\partial\rho}\dot{\rho}\big(rV(x,\hat{\theta})^{1/r} + \varepsilon\big).
\end{aligned}
\tag{19}
$$

Furthermore, utilizing the update law (12b) in the right-hand side of (19) results in the cancellation of the third and fifth terms. Thus, from the update law (12a), one can see that

$$
\begin{aligned}
\dot{Q} \leq & -\omega(\rho)\lambda V(x,\hat{\theta})^{1/r} - \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\Delta(x)^\top\tilde{\theta} + \tilde{\theta}^\top\Gamma^{-1}\dot{\hat{\theta}} \\
\leq & -\omega(\rho)\lambda V(x,\hat{\theta})^{1/r} - \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\Delta(x)^\top\tilde{\theta} + \tilde{\theta}^\top\omega(\rho)V(x,\hat{\theta})^{1/r-1}\Delta(x)\frac{\partial V(x,\hat{\theta})}{\partial x}^\top.
\end{aligned}
$$

Consequently, with the last two terms canceled, it follows that

$$
\dot{Q} \leq -\omega(\rho)\lambda V(x,\hat{\theta})^{1/r} \leq 0,
\tag{20}
$$

which implies that $Q\big(x,\hat{\theta},\omega(\rho)\big)$ is non-increasing. As a result, the value of $\omega(\rho)$, $V(x,\hat{\theta})$, $x(t)$ and $\hat{\theta}(t)$ remain bounded. Furthermore, by taking the time-integral from both sides of the inequality (20), we have

$$
\int_0^\infty \lambda\omega(\rho)V(x,\hat{\theta})^{1/r}\mathrm{d}t \leq Q\big(x(0),\hat{\theta}(0),\omega(\rho(0))\big) - \lim_{t\to\infty} Q\big(x(t),\hat{\theta}(t),\omega(\rho(t))\big)
$$

$$
\leq Q\big(x(0),\hat{\theta}(0),\omega(\rho(0))\big) < \infty,
$$

where the second inequality follows from the fact that $Q(x,\hat{\theta},\omega(\rho))$ is a non-negative function. Since $\omega(\rho)$ is strictly positive, the inequality above implies that $V(x,\hat{\theta})^{1/r} \in \mathcal{L}_1$. Moreover, from the inequality (8), we have

$$
\int_0^\infty ||x||_2^{2/r}\mathrm{d}t \leq \int_0^\infty \frac{1}{c_1^{1/r}}V(x,\hat{\theta})^{1/r}\mathrm{d}t < \infty,
$$

meaning that $||x||_2^{2/r} \in \mathcal{L}_1$. Furthermore, we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}V(x,\hat{\theta}) &= \frac{\partial V(x,\hat{\theta})}{\partial x}\dot{x} + \frac{\partial V(x,\hat{\theta})}{\partial x}\dot{\hat{\theta}} \\
&= \frac{\partial V(x,\hat{\theta})}{\partial x}\big(f(x) + \Delta(x)^\top\theta + B(x)u\big) + \frac{\partial V(x,\hat{\theta})}{\partial\hat{\theta}}\Gamma\omega(\rho)V(x,\hat{\theta})^{1/r-1}\Delta(x)\frac{\partial V(x,\hat{\theta})}{\partial x}^\top.
\end{aligned}
$$

Note that for $r \leq vd(\Delta,V)$, the last term on the right-hand side of the above expression remains bounded. This implies that $\dot{V}(x,\hat{\theta})$ is bounded. Consequently, $V(x,\hat{\theta})$ is uniformly continuous. Since $V(x,\hat{\theta})$ is both integrable and uniformly continuous, Barbalat's lemma ensures that $V(x,\hat{\theta})$ converges to zero as $t\to\infty$, which in turn implies that $x(t)\to 0$ as $t\to\infty$.

5.3. **Proof of Theorem 4**

Since $V(x,\theta)$ is a smooth and positive definite function with $V(0,\theta) = 0$, it has a Taylor expansion of the form

$$
V(x,\theta) = \frac{1}{2}x^\top H(\theta)x + R_V(x,\theta),
\tag{21}
$$

where $R_V : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ is a smooth function satisfying

$$
\lim_{x\to 0} \frac{R_V(x,\theta)}{||x||_2^2} = 0,
\tag{22}
$$

and $H(\theta)$ is a positive definite matrix, that satisfies $c_1 I_n \leq H(\theta) \leq c_2 I_n$, for all $\theta \in \Theta$. Note that, due to positivity and smoothness, the gradient of $V(x, \theta)$ vanishes at zero, and therefore, the expansion (21) does not have a linear term. Furthermore, since $\Delta(x)^\top$ is twice continuously differentiable, each column of it has a first-order Taylor expansion of the form

$$[\Delta(x)^\top]_i = F_i x + R_i(x),$$

where $F_i \in \mathbb{R}^{r \times r}$ and $R_i : \mathbb{R}^n \to \mathbb{R}^n$ is a function satisfying

$$\lim_{x \to 0} \frac{R_i(x)}{\|x\|_2} = 0.$$

Thus, using the above expressions, for all $r \in \mathbb{R}_{>0}$ and $i \in \{1, \ldots, p\}$, we have

$$\limsup_{x \to 0} \left| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} [\Delta(x)^\top]_i \right| = \limsup_{x \to 0} \left| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} (F_i x + R_i(x)) \right|$$

$$= \limsup_{x \to 0} \left| V(x, \theta)^{1/r} \frac{x^\top H F_i x}{\frac{1}{2} x^\top H x + R_V(x, \theta)} + \eta_i(x, \theta) \right|,$$

where $\eta_i(x, \theta)$ corresponds the higher order terms in the resulting expansion. Accordingly, it follows that

$$\limsup_{x \to 0} \left| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} [\Delta(x)^\top]_i \right| \leq \limsup_{x \to 0} \left| V(x, \theta)^{1/r} \frac{x^\top H F_i x}{x^\top H x} \frac{x^\top H x}{\frac{1}{2} x^\top H x + R_V(x, \theta)} \right|. \quad (23)$$

From the positive-definiteness of H, one can easily see that

$$\frac{1}{\lambda_{\max}(H)} \frac{|R_V(x, \theta)|}{\|x\|_2^2} \leq \left| \frac{R_V(x, \theta)}{x^\top H x} \right| \leq \frac{1}{\lambda_{\min}(H)} \frac{|R_V(x, \theta)|}{\|x\|_2^2}.$$

Consequently, (22) implies that

$$\lim_{x \to 0} \left| \frac{x^\top H x}{\frac{1}{2} x^\top H x + R_V(x, \theta)} \right| = 2. \quad (24)$$

Also, from the definition of matrix operator norm, we have

$$\limsup_{x \to 0} \left| \frac{x^\top H F_i x}{x^\top H x} \right| \leq \|H^{\frac{1}{2}} F_i H^{-\frac{1}{2}}\|. \quad (25)$$

Subsequently, from (23), (24), and (25), it follows that

$$\limsup_{x \to 0} \left| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} [\Delta(x)^\top]_i \right| \leq 2\|H^{\frac{1}{2}} F_i H^{-\frac{1}{2}}\| \limsup_{x \to 0} V(x, \theta)^{1/r} = 0.$$

As a result, one can see that

$$\lim_{x \to 0} \left\| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} \Delta(x)^\top \right\|_\infty = \lim_{x \to 0} \max_i \left| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} [\Delta(x)^\top]_i \right| = 0,$$

for any $r \in \mathbb{R}_{>0}$. Therefore, following the definition of the vanishing degree, we have

$$vd(\Delta, V) = \sup \left\{ r \in \mathbb{R}_{>0} \ \middle| \ \lim_{x \to 0} \left\| V(x, \theta)^{1/r-1} \frac{\partial V(x, \theta)}{\partial x} \Delta(x)^\top \right\|_\infty < \infty, \ \forall \theta \in \Theta \right\} = \infty,$$

as claimed.

5.4. **Proof of Theorem 5**

Consider the non-negative function

$$Q\big(x,\hat{a},\hat{\theta},\omega\big) = \omega(\rho)\big(rV(x,\hat{\theta})^{1/r} + \varepsilon\big) + \frac{1}{2}(\hat{a}-\theta)^\top\Gamma^{-1}(\hat{a}-\theta) + \frac{1}{2}(\hat{\theta}-\hat{a})^\top\Gamma^{-1}(\hat{\theta}-\hat{a}).$$

Taking the time derivative of the above function along the dynamic (7) gives

$$\dot{Q} = \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\big(f(x) + \Delta(x)^\top\theta + B(x)u\big) + \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial\theta}\dot{\hat{\theta}}$$

$$+ \frac{\partial\omega(\rho)}{\partial\rho}\dot{\rho}\big(rV(x,\hat{\theta})^{1/r} + \varepsilon\big) + (\hat{a}-\theta)^\top\Gamma^{-1}\dot{\hat{a}} + (\hat{\theta}-\hat{a})^\top\Gamma^{-1}(\dot{\hat{\theta}} - \dot{\hat{a}}).$$

The above equation can be written as

$$\dot{Q} = \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\big(f(x) + \Delta(x)^\top\hat{\theta} + B(x)u\big) + \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial\theta}\dot{\hat{\theta}}$$

$$+ \frac{\partial\omega(\rho)}{\partial\rho}\dot{\rho}\big(rV(x,\hat{\theta})^{1/r} + \varepsilon\big) - \omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\hat{\theta})}{\partial x}\Delta(x)^\top\tilde{\theta} + (2\hat{a}-\theta-\hat{\theta})^\top\Gamma^{-1}\dot{\hat{a}}$$

$$+ (\hat{\theta}-\hat{a})^\top\Gamma^{-1}\dot{\hat{\theta}}.$$

By substituting the control law (10) and the update law (14a), we find that

$$\dot{Q} \leq -\lambda\omega(\rho)V(x,\hat{\theta})^{1/r} + 2\omega(\rho)V(x,\hat{\theta})^{1/r-1}\frac{\partial V(x,\theta)}{\partial x}\Delta(x)^\top(\hat{a}-\hat{\theta}) + (\hat{\theta}-\hat{a})^\top\Gamma^{-1}\dot{\hat{\theta}}.$$

By setting $\dot{\hat{\theta}}$ according to the update law (14b), we have

$$\dot{Q} \leq -\frac{\lambda}{2}\omega(\rho)V(x,\hat{\theta})^{1/r} - \frac{1}{2}\omega(\rho)V(x,\hat{\theta})^{1/r-1}\bigg(\lambda V(x,\hat{\theta})$$

$$+ 2\frac{\partial V(x,\theta)}{\partial x}\Delta(x)^\top(\hat{a}-\hat{\theta}) - \frac{1}{\lambda}(\hat{\theta}-\hat{a})^\top\Delta(x)P(\theta)\Delta(x)^\top(\hat{\theta}-\hat{a})\bigg). \qquad (26)$$

Note that, in Theorem 5, we have $V(x,\theta) = \frac{1}{2}x^\top P(\theta)x$ and $\frac{\partial}{\partial x}V(x,\theta) = x^\top P(\theta)$. Substituting these terms into (26) makes the expression inside the parentheses a perfect square, as follows

$$\dot{Q} \leq -\frac{\lambda}{2}\omega(\rho)V(x,\hat{\theta})^{1/r} - \frac{\lambda}{2}\omega(\rho)V(x,\hat{\theta})^{1/r-1}\big(x - \frac{(\hat{\theta}-\hat{a})}{\lambda}\big)^\top P(\theta)\big(x - \frac{(\hat{\theta}-\hat{a})}{\lambda}\big) \leq 0.$$

Therefore, we conclude that the function $Q\big(x,\hat{a},\hat{\theta},\omega\big)$ is non-increasing. Hence, the signals $x(t)$, $\hat{\theta}(t)$ and $\hat{a}(t)$ are bounded. Furthermore, by taking the time integral from both sides of the above inequality, we find that

$$\int_0^\infty \lambda\omega(\rho)V(x,\hat{\theta})^{1/r}\mathrm{d}t \leq 2Q\big(x(0),\hat{\theta}(0),\omega\big(\rho(0)\big)\big) - 2\lim_{t\to\infty}Q\big(x(t),\hat{\theta}(t),\omega(\rho(t))\big)$$

$$- \int_0^\infty \lambda\omega(\rho)V(x,\hat{\theta})^{1/r-1}\big(x - \frac{(\hat{\theta}-\hat{a})}{\lambda}\big)^\top P(\theta)\big(x - \frac{(\hat{\theta}-\hat{a})}{\lambda}\big)\,\mathrm{d}t$$

$$\leq 2Q\big(x(0),\hat{\theta}(0),\omega\big(\rho(0)\big)\big) < \infty,$$

which implies that $V(x,\hat{\theta})^{1/r}$ belongs to $\mathcal{L}_1$. As a result, from the inequality (8), we conclude that $\|x\|_2^{2/r} \in \mathcal{L}_1$. Moreover, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}V(x,\hat{\theta}) = \frac{\partial V(x,\hat{\theta})}{\partial x}\dot{x} + \frac{\partial V(x,\hat{\theta})}{\partial x}\dot{\hat{\theta}} = \frac{\partial V(x,\hat{\theta})}{\partial x}\big(f(x) + \Delta(x)^\top\theta + B(x)u\big) + \frac{\partial V(x,\hat{\theta})}{\partial\hat{\theta}}\dot{\hat{\theta}}.$$

(a) Evolution of $\log V(x, \hat{\theta})$ (b) Evolution of state for $r = 1$ (c) Evolution of state for $r = 8$
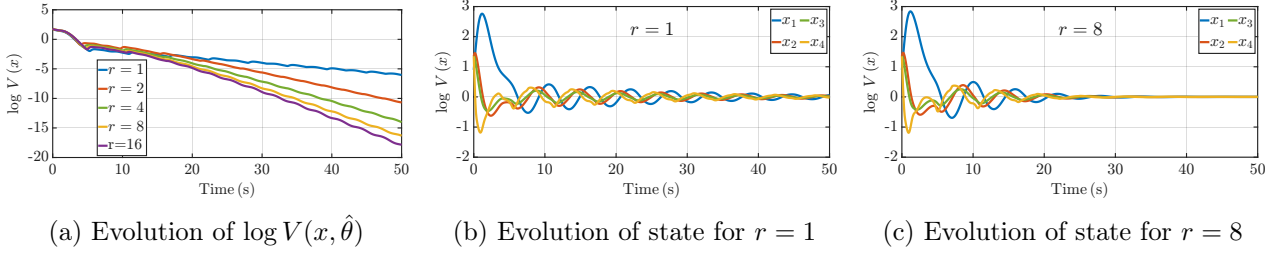
Figure 2. (a) Plot of the log Lyapunov function versus time for Example 3, where the control law (10) and the update laws (12a) are employed for different values of r. (b)–(c) Closed-loop state variables in Example 3, where the update law (12a) is employed for $r = 1$ and $r = 8$, respectively.

(a) Evolution of $\log V(x, \hat{\theta})$ (b) Evolution of state for $r = 1$ (c) Evolution of state for $r = 8$
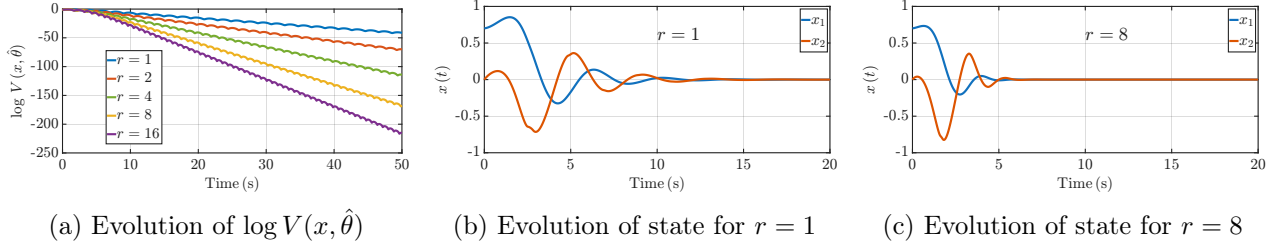
Figure 3. (a) Plot of the log Lyapunov function versus time for Example 2, where the control law (10) and the update laws (12a) are employed for different values of r. (b)–(c) Closed-loop state variables in Example 4, where the update laws (12a) and (12b) are employed for $r = 1$ and $r = 8$, respectively.

The assumptions stated in Theorem 5 ensure that the update law $\dot{\hat{\theta}}$ remains bounded. As a result, the right-hand side of the above expression is bounded, which implies $V(x, \hat{\theta})$ is uniformly continuous. Hence, by Barbalat's lemma, $V(x, \hat{\theta})$ and, consequently, $x(t)$ converge to zero as $t \to \infty$.

6. NUMERICAL RESULTS

In this section, we demonstrate the performance improvements achieved by the proposed parameter estimation laws using three numerical examples, each representing a different class of dynamics and uncertainty structure. The first example considers a system with a linear nominal term and a nonlinear matched uncertainty; the second addresses systems with unmatched uncertainty in parametric strict-feedback form; and the third investigates a robot manipulator with flexible joints.

Example 3 (Matched uncertainty). Consider a dynamical system of the form

$$\dot{x} = Ax + B(u + \varphi(x)^\top \theta). \tag{27}$$

Our goal is to regulate the system state to zero. Note that the uncertainty term, defined here as $\Delta(x)^\top \theta = B(x)\varphi(x)^\top \theta$, satisfies the matching condition (13). By applying Finsler's lemma [55] to (9), we obtain that any quadratic function of the form $V(x) = x^\top P^{-1} x$ serves as a CLF for system (27) if and only if the positive definite matrix P satisfies the linear matrix inequality

$$B^\perp (AP + PA^\top + \lambda P)B^\perp \leq 0, \tag{28}$$

where $B^\perp$ is the annihilator matrix of B, i.e., $B^\perp B = 0$.

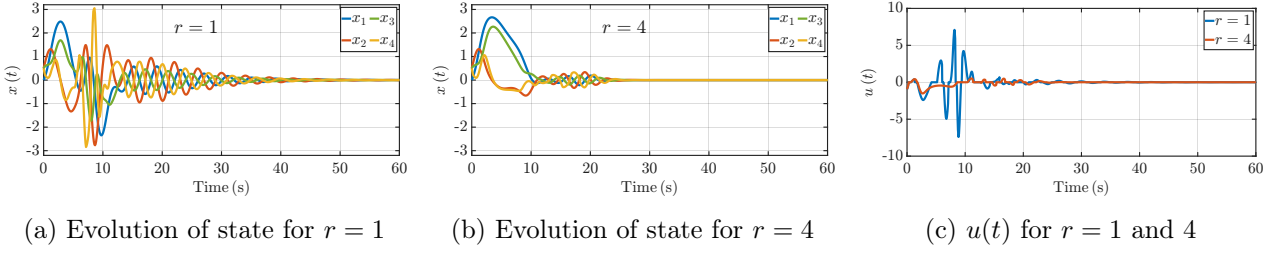(a) Evolution of state for $r = 1$ (b) Evolution of state for $r = 4$ (c) $u(t)$ for $r = 1$ and 4

Figure 4. (a)–(b) Closed-loop state variables in Example 5, using the update laws from Theorem 3 with $r = 1$ and $r = 4$, respectively. (c) Control signal $u(t)$ in Example 5 using the update laws from Theorem 3 with $r = 1$ and $r = 4$.

For our numerical demonstration, we set A, B, and $\varphi(x)$ respectively as

$$
A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \varphi(x) = \begin{bmatrix} x_1^2 \\ \arctan(x_2) \\ x_3 \\ \arctan(x_4) \end{bmatrix}.
$$

Following Theorem 4, we have that the vanishing degree $vd(\Delta, V)$ is infinite, implying that the parameter r in Theorem 3 can be chosen arbitrarily large. In the rest of this example, we set the true system parameter vector to

$$
\theta = \begin{bmatrix} 0.5 & 1 & 1 & 1 \end{bmatrix}^\top.
$$

The system state and the parameter estimate are initialized as

$$
x(0) = \begin{bmatrix} 1 & 1.5 & 1.5 & 1.5 \end{bmatrix}^\top, \quad \hat{\theta}(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^\top.
$$

The adaptation gain is set to $\Gamma = I_4$. Figure 2a illustrates the log of $V(x)$ as a function of time. Each curve is obtained by applying the controller (11) and the parameter update law (12a) for a fixed value of r, namely $r \in \{1, 2, 4, 8, 16\}$. As we can observe, increasing r accelerates the convergence of $V(x)$. Figure 2b and Figure 2c present the system state trajectories for $r = 1$ and $r = 8$, respectively. It can be seen that the states converge to zero significantly faster for $r = 8$ compared to $r = 1$.

Example 4 (Unmatched uncertainty). Consider the following system with the parametric strict-feedback form as

$$
\begin{aligned}
\dot{x}_1 &= x_2 + \theta^\top \varphi_1(x_1), \\
\dot{x}_2 &= \theta^\top \varphi_2(x_1, x_2) + u,
\end{aligned}
\tag{29}
$$

where $x = [x_1, x_2]^\top$ is the system state, u is the control input, $\theta \in \mathbb{R}^p$ is an unknown parameter vector, and $\varphi_1(x_1)$, $\varphi_2(x_1, x_2)$ are known functions. Using the back-stepping technique [22], we construct the following CLF

$$
V(x, \theta) = \frac{1}{2} x_1^2 + \frac{1}{2} \left(x_2 + \theta^\top \varphi_1(x_1) + \lambda x_1 \right)^2,
$$

for a given $\lambda > 0$. For our numerical simulations, we assume

$$
\varphi_1(x_1) = \begin{bmatrix} \sin(x_1) \\ \arctan(x_1) \\ 0 \\ 0 \end{bmatrix}, \quad \varphi_2(x_1, x_2) = \begin{bmatrix} 0 \\ 0 \\ \sin(x_1) \\ \arctan(x_2) \end{bmatrix}.
$$

Note that in this example, $V(x,\theta)$ is a function of the system parameters. Moreover, according to Theorem 4, one can see that the vanishing degree $vd(\Delta, V)$ is infinite.

We set the design parameters in our estimation law as

$$\Gamma = 0.2I_4, \quad \omega(\rho) = \frac{\pi}{2} + \arctan(0.5\,\rho), \quad \varepsilon = 0.5.$$

Figure 3a presents the logarithm of $V(x,\theta)$ over time. Similar to the previous example, each curve corresponds to a fixed r, for $r \in \{1, 2, 4, 8, 16\}$, and is obtained using the controller in (11) together with the parameter estimation law from Theorem 3. Furthermore, Figure 3b and Figure 3c present the system state trajectories for $r = 1$ and $r = 8$, respectively. We can observe that increasing r accelerates convergence.

Example 5 (Manipulator with flexible joints)**.** The dynamics of a single link manipulator with flexible joints can be described as [21]

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \theta_1 \sin x_1 + \theta_2(x_3 - x_1), \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= \theta_3(x_1 - x_3) + \theta_4 x_4 + u,
\end{aligned} \tag{30}$$

where x_1 and x_3 are angular positions, x_2 and x_4 are angular velocities, and $\theta = [\theta_1 \; \theta_2 \; \theta_3 \; \theta_4]^\top$ is an unknown parameter vector, with $\theta_2 > 0$. We note that system (30) is feedback linearizable, and therefore, a CLF can be obtained using feedback linearization techniques [17]. In particular, one can verify that the function

$$V(x,\theta) = \Psi(x,\theta)^\top P^{-1} \Psi(x,\theta),$$

where

$$\Psi(x,\theta) = \begin{bmatrix} x_1 \\ x_2 \\ \theta_1 \sin x_1 + \theta_2(x_3 - x_1) \\ \theta_1 x_2 \cos x_1 + \theta_2(x_4 - x_2) \end{bmatrix},$$

and P is a positive-definite matrix satisfying the linear matrix inequality (28) for

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

is a CLF for system (30). According to Theorem 4, the vanishing degree $vd(\Delta, V)$ associated with system (30) is infinite. Hence, the parameter r in Theorem 3 can be chosen arbitrarily large.

Figure 4a and Figure 4b show the system states when the parameter estimation law in Theorem 3 are employed, for $r = 1$ and $r = 4$, together with the control (11). It can be observed that the system states converge significantly faster to zero for $r = 4$ compared to $r = 1$. Figure 4c shows the input signal $u(t)$ for the mentioned scenarios. One can see that the control effort for $r = 4$ has a smaller magnitude and fewer oscillations.

7. Conclusions And Future Directions

In this paper, we proposed a new family of parameter estimation laws and their higher-order extensions for adaptive systems. We demonstrated that these estimation laws can significantly accelerate the convergence of the system states by promoting signal sparsity in the time domain, i.e., penalizing prolonged signal duration and slow decay. The developed approach does not require persistent excitation, high adaptation gains, or prior knowledge of the system parameters. Furthermore, it can be integrated with any certainty-equivalence CLF-based controller. Moreover, it is applicable to uncertain systems with both matched and unmatched uncertainties. Numerical experiments and simulation results confirm the performance improvements achieved using the proposed method.

Future work could extend these results to accelerate convergence in adaptive safety and control barrier functions [1, 28], fault compensation in dynamical systems [52, 9], adaptive observer design [22, 26, 10], adaptive control for systems with time-varying parameters [15], and adaptive control for nonlinearly parameterized systems [54].

References

[1] Aaron D. Ames et al. "Control barrier function based quadratic programs for safety-critical systems". In: *IEEE Transactions on Automatic Control* 62.8 (2017), pp. 3861–3876.

[2] Zvi Artstein. "Stabilization with relaxed controls". In: *Nonlinear Analysis-Theory Methods & Applications* 7.11 (1983), pp. 1163–1173.

[3] Karl Johan Åström and Björn Wittenmark. *Adaptive control.* 2nd. Mineola, NY: Dover Publications, 2008.

[4] Francis Bach. *Learning theory from first principles.* Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press, 2024.

[5] Georges Bastin. *On-line estimation and adaptive control of bioreactors.* Vol. 1. Elsevier, 2013.

[6] Charalampos P. Bechlioulis and George A. Rovithakis. "Adaptive control with guaranteed transient and steady state tracking error bounds for strict feedback systems". In: *Automatica* 45.2 (2009), pp. 532–538.

[7] Charalampos P. Bechlioulis and George A. Rovithakis. "Robust Adaptive Control of Feedback Linearizable MIMO Nonlinear Systems With Prescribed Performance". In: *IEEE Transactions on Automatic Control* 53.9 (2008), pp. 2090–2099.

[8] Marc Bodson and Joseph E. Groszkiewicz. "Multivariable adaptive algorithms for reconfigurable flight control". In: *IEEE Transactions on Control Systems Technology* 5.2 (1997), pp. 217–229.

[9] Mohammad Boveiri and Mohammad Saleh Tavazoei. "Adaptive Actuator Failure Compensation on the Basis of Contraction Metrics". In: *IEEE Control Systems Letters* 6 (2022), pp. 1376–1381.

[10] Mohammad Boveiri and Mohammad Saleh Tavazoei. "Output feedback and adaptive output feedback control design for input-affine nonlinear systems based on Riemannian metrics". In: *European Journal of Control* 75 (2024), p. 100938.

[11] Stephen Boyd and Shankar Sastry. "Necessary and sufficient conditions for parameter convergence in adaptive control". In: *Automatica* 22.6 (1986), pp. 629–639.

[12] Stephen Boyd and Shankar Sastry. "On parameter convergence in adaptive control". In: *Systems & Control Letters* 3.6 (1983), pp. 311–319.

[13] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[14] Stephen Boyd and Lieven Vandenberghe. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[15] Kaiwen Chen and Alessandro Astolfi. "Adaptive control for systems with time-varying parameters". In: *IEEE Transactions on Automatic Control* 66.5 (2021), pp. 1986–2001.

[16] Zachary T. Dydek, Anuradha M. Annaswamy, and Eugene Lavretsky. "Adaptive control and the NASA X-15-3 flight revisited". In: *IEEE Control Systems Magazine* 30.3 (2010), pp. 32–48.

[17] Randy Freeman and Petar V. Kokotovic. *Robust nonlinear control design: State-Space and Lyapunov techniques*. Birkhäuser, 2008.

[18] Joseph E. Gaudio et al. "A class of high-order tuners for adaptive systems". In: *IEEE Control Systems Letters* 5.2 (2021), pp. 391–396.

[19] Petros A. Ioannou and Jing Sun. *Robust adaptive control*. Prentice Hall, 1996.

[20] Benjamin M. Jenkins et al. "Convergence properties of adaptive systems and the definition of exponential stability". In: *SIAM Journal on Control and Optimization* 56.4 (2018), pp. 2463–2484.

[21] Hassan K. Khalil. *Nonlinear systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[22] Miroslav Krstic, Ioannis Kanellakopoulos, and Petar Kokotovic. *Nonlinear and Adaptive Control Design*. John Willey, 1995.

[23] Miroslav Krstić and Petar V. Kokotović. "Modular approach to adaptive nonlinear stabilization". In: *Automatica* 31.1 (1995), pp. 95–100.

[24] Taeyoon Lee, Jaewoon Kwon, and Frank C. Park. "A natural adaptive control law for robot manipulators". In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2018, pp. 8628–8635.

[25] Xuefang Li et al. "Robust adaptive learning-based path tracking control of autonomous vehicles under uncertain driving environments". In: *IEEE Transactions on Intelligent Transportation Systems* 23.11 (2022), pp. 20798–20809.

[26] Yusheng Liu. "Robust adaptive observer for nonlinear systems with unmodeled dynamics". In: *Automatica* 45.8 (2009), pp. 1891–1895. ISSN: 0005-1098.

[27] Brett T. Lopez and Jean-Jacques E. Slotine. "Universal adaptive control of nonlinear systems". In: *IEEE Control Systems Letters* (2021).

[28] Brett T. Lopez, Jean-Jacques E. Slotine, and Jonathan P. How. "Robust adaptive control barrier functions: An adaptive and data-driven approach to safety". In: *IEEE Control Systems Letters* 5.3 (2021), pp. 1031–1036.

[29] Ian R Manchester and Jean-Jacques E. Slotine. "Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design". In: *IEEE Transactions on Automatic Control* 62.6 (2017), pp. 3046–3053.

[30] Frederic Mazenc, Marcio de Queiroz, and Michael Malisoff. "Uniform Global Asymptotic Stability of a Class of Adaptively Controlled Nonlinear Systems". In: *IEEE Transactions on Automatic Control* 54.5 (2009), pp. 1152–1158.

[31] Qingtan Meng, Qian Ma, and Yang Shi. "Adaptive fixed-time stabilization for a class of uncertain nonlinear systems". In: *IEEE Transactions on Automatic Control* 68.11 (2023), pp. 6929–6936.

[32] Kumpati S Narendra and Anuradha M Annaswamy. "Persistent excitation in adaptive systems". In: *International Journal of Control* 45.1 (1987), pp. 127–160.

[33] Kumpati S. Narendra and Anuradha M. Annaswamy. *Stable adaptive systems*. Dover Publications, 2012.

[34] Arkadi Nemirovski et al. "Robust stochastic approximation approach to stochastic programming". In: *SIAM Journal on Optimization* 19.4 (2009), pp. 1574–1609.

[35] Yurii Nesterov. "A method for solving the convex programming problem with convergence rate $O(1/k^2)$". In: *Soviet Mathematics Doklady* 27.2 (1983), pp. 372–376.

[36] Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course.* Vol. 87. Applied Optimization. Springer, 2004.

[37] Jean-Jacques E. Slotine Nicholas M. Boffi. "Implicit regularization and momentum algorithms in nonlinearly parameterized adaptive control and prediction". In: *Neural Computation* 33 (2021), pp. 590–673.

[38] Romeo Ortega et al. *Passivity-based Control of Euler-Lagrange Systems: Mechanical, Electrical and Electromechanical Applications.* Springer, 1998.

[39] Alexey Pavlov, Nathan Wouw, and Henk Nijmeijer. *Uniform output regulation of nonlinear systems: A convergent dynamics approach.* Springer Science & Business Media, 2006.

[40] Donald A. Pierre. "A perspective on adaptive control of power systems". In: *IEEE Transactions on Power Systems* 2.2 (1987), pp. 387–395.

[41] James A. Primbs, Vesna Nevistic, and John C. Doyle. "A receding horizon generalization of pointwise min-norm controllers". In: *IEEE Transactions on Automatic Control* 45.5 (2000), pp. 898–909.

[42] Rodolphe Sepulchre, Mrdjan Jankovic, and Petar V. Kokotovic. *Constructive nonlinear control.* Berlin, Heidelberg: Springer, 1997.

[43] Jean-Jacques E. Slotine and Weiping Li. *Applied nonlinear control.* Englewood Cliffs, NJ: Prentice Hall, 1991.

[44] Jean-Jacques E. Slotine and Weiping Li. "Composite adaptive control of robot manipulators". In: *Automatica* 25.4 (1989), pp. 509–519. ISSN: 0005-1098.

[45] Jean-Jacques E. Slotine and Weiping Li. "On the adaptive control of robot manipulators". In: *International Journal of Robotics Research* 6.3 (1987), pp. 49–63.

[46] Yong-Duan Song, Richard H. Middleton, and Joe N. Anderson. "Study on the exponential path tracking control of robot manipulators via direct adaptive methods". In: *Robotics and Autonomous Systems* 9.4 (1992), pp. 271–282.

[47] Yongduan Song, Kai Zhao, and Miroslav Krstic. "Adaptive control with exponential regulation in the absence of persistent excitation". In: *IEEE Transactions on Automatic Control* 62.5 (2017), pp. 2589–2596.

[48] Eduardo D. Sontag. "A 'universal' construction of Artstein's theorem on nonlinear stabilization". In: *Systems & Control Letters* 13.2 (1989), pp. 117–123.

[49] Mark W. Spong. "An historical perspective on the control of robotic Manipulators". In: *Annual Review of Control, Robotics, and Autonomous Systems* 5 (2022), pp. 1–31.

[50] Mark W. Spong, Seth Hutchinson, and Mathukumalli Vidyasagar. *Robot Modeling and Control.* 2nd. Wiley, 2020.

[51] Weehong Tan and Andrew Packard. "Stability region analysis using polynomial and composite polynomial Lyapunov functions and sum-of-squares programming". In: *IEEE Transactions on Automatic Control* 53.2 (2008), pp. 565–571.

[52] Gang Tao, Suresh M. Joshi, and Xiaoli Ma. "Adaptive state feedback and tracking control of systems with actuator failures". In: *IEEE Transactions on Automatic Control* 46.1 (2001), pp. 78–95.

[53] Andrew J. Taylor and Aaron D. Ames. "Adaptive safety with control barrier functions". In: *2020 American Control Conference (ACC)*. 2020, pp. 1399–1405.

[54] Ivan Yu. Tyukin, Danil V. Prokhorov, and Cees van Leeuwen. "Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization". In: *IEEE Transactions on Automatic Control* 52.9 (2007), pp. 1543–1559.

[55] Frank Uhlig. "A recurring theorem about pairs of quadratic forms and extensions: a survey". In: *Linear Algebra and its Applications* 25 (1979), pp. 219–237.

[56] Fang Wang and Guanyu Lai. "Fixed-time control design for nonlinear uncertain systems via adaptive method". In: *Systems & Control Letters* 140 (2020), p. 104704.

[57] Patrick M. Wensing and Jean-Jacques E. Slotine. *Beyond convexity – contraction and global convergence of gradient descent*. 2018. arXiv: 1806.06655.

[58] Zongyu Zuo and Pengkai Ru. "Augmented L_1 adaptive tracking control of quad-rotor unmanned aircrafts". In: *IEEE Transactions on Aerospace and Electronic Systems* 50.4 (2014), pp. 3090–3101.