Online Optimization of LTI Systems Under Persistent Attacks: Stability, Tracking, and Robustness

Felipe Galarza-Jimenez, Gianluca Bianchin, Jorge I. Poveda and Emiliano Dall'Anese

Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, USA.

ARTICLE INFO

ABSTRACT

We study the stability properties of a control system composed of a dynamical plant and a feedback controller, the latter generating control signals that can be compromised by a malicious attacker. We consider two classes of feedback controllers: a static output-feedback controller, and a dynamical gradient-flow controller that seeks to steer the output of the plant towards the solution of a convex optimization problem. In both cases, we analyze the stability properties of the closed-loop system under a class of switching attacks that persistently modify the control inputs generated by the controllers. Our stability analysis leverages the framework of hybrid dynamical systems, Lyapunov-based arguments for switching systems with unstable modes, and singular perturbation theory. Our results reveal that, under a suitable time-scale separation between plant and controllers, the stability of the interconnected system can be preserved when the attack occurs with "sufficiently low frequency" in any bounded time interval. We present simulation results in a power-grid example that corroborate the technical findings.

1. Introduction

This paper studies the stability properties of the interconnection between a Linear Time Invariant (LTI) dynamical system and an output feedback controller, where the control inputs can be maliciously compromised by an attacker. The theoretical and algorithmic endeavors are motivated by a number of applications within the realm of cyber-physical systems (CPSs) – that is, physical systems integrated with computational resources by means of a communication infrastructure. While advances in communication and cyber technologies provide enhanced functionality, efficiency, and autonomy of a CPS, the presence of communication channels as well as the tight integration between cyber and physical components unavoidably introduces new security vulnerabilities. In particular, the modeling adopted in this paper is well suited for a number of applications in power systems [1, 2], transportation systems [3], communication networks [4], and robotics [5], to mention just a few.

As illustrated in Figure 1, our setting considers two types of feedback controllers: (i) a static output-feedback controller that is used as an inner loop to stabilize the LTI system; and, (ii) an output feedback controller based on an appropriate modification of gradient-flow dynamics. For the synthesis of the gradient-flow controller, we start by formulating an unconstrained optimization problem with a composite cost function that captures performance indexes associated with the input and the output of the plant. The cost function is assumed to be smooth and to satisfy the Polyak-Łojasiewicz (PL) inequality condition (where we recall that the PL inequality is a weaker assumption than convexity, and it implies the property of invexity) [6]. We consider adversarial actions of the form of *switching multiplicative attacks against the interconnection between the LTI plant and the controller*, whereby an attacker can persistently modify the inputs of the dynamical system with the objective of destabilizing the equilibrium points. This attack model is rather general and it captures different classes of attacks that can *persistently* modify the sign and/or the magnitude of the control signals, and also jam the communication channels.

With this setting in place, the problem addressed in this paper is that of finding sufficient conditions that guarantee that the equilibrium points of the closed-loop system are asymptotically or exponentially stable. To this end, we adopt the framework of hybrid systems, and we leverage arguments from singular perturbation analysis and input-to-state stability (ISS) for hybrid systems. This framework allows us to establish sufficient conditions in terms of the total
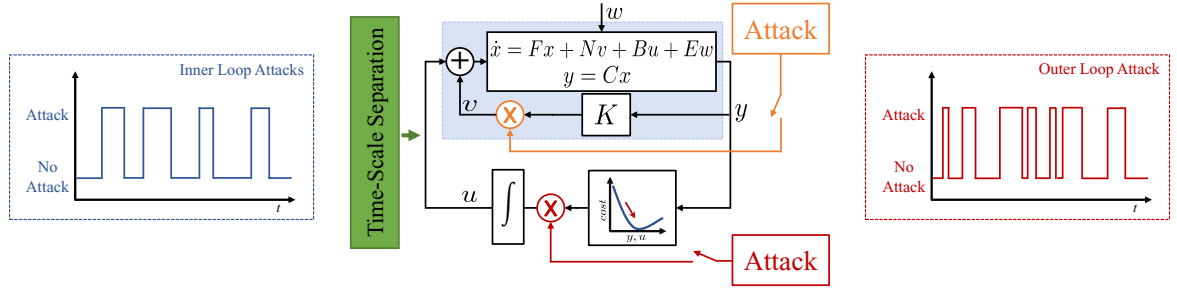
Figure 1: An inner control loop is designed to stabilize the plant, while an outer control loop regulates the plant to an optimizer of (7).

activation time of the attacks acting on the system, as well as the time-scale separation needed between the plant and the gradient-flow controller to preserve the stability properties of the interconnection.

Related Works. The design of feedback-based optimization controllers has recently received significant attention; see, for example [7, 8, 9, 10, 2, 5, 3, 11]. In particular, in [9] sufficient conditions on the time-scale separation between plant and controller were derived to induce asymptotic stability properties. Similarly, the joint stabilization and optimal steady-state regulation of LTI dynamical systems was considered in [10]. LTI dynamical systems with time-varying exogenous inputs were considered in [2], along with the problem of tracking an optimal solution trajectory of a time-varying problem with a strongly convex cost; prediction-correction-type controllers were utilized to track the trajectory of a time-varying problem in [5]. Recently, [3] established exponential stability results for the interconnection of a switched LTI system and a hybrid feedback controller based on accelerated gradient dynamics with resets. Finally, [12] and [13] studied extremum-seeking algorithms for model-free optimization under deception attacks.

In recent years, motivated by the increasing vulnerability of cyber-physical systems operating in adversarial environments, several works have investigated the stability properties of systems under denial-of-service attacks (DoS), whereby an attacker compromises system resources such as sensors or actuators, as well as infrastructure such as communication channels. We refer to [14, 15, 16] (see also references therein) for a comprehensive list of related works, while we present a list of representative references below. The authors in [17] designed a stabilizing controller for communication channels that face malicious random packet losses, while [18] presented a class of stabilizing event-triggered controllers under DoS attacks. The work [19] designed scheduling policies that preserve the stability of an interconnected system when an attacker jams the control inputs that act on the plants. The works [20] and [21] investigated system stability in the presence of deception attacks, namely, attacks where the integrity of control packets or measurements is compromised. Robustness of a sliding mode controller against DoS attacks was investigated in [22]; the effect of DoS attacks in the tracking performance of actuators was investigated in [23]. Attacks to both a state estimator and actuator of an LTI system were considered in [24], where the attacks are assumed to satisfy a given sparse observation condition. A self-triggered consensus network in the presence of communication failures caused by DoS attacks was considered in [15], and a similar problem was studied in [25] and [26]. Event-triggered communication and decentralized control of switched systems under cyber attacks are analyzed in [27]; in particular, conditions on the dwell-time and the gain for the controller are derived to ensure stability. Other lines of work have studied the robustness properties with respect to attacks of different (discrete-time) optimization algorithms; e.g., [28, 29, 30]. In these works, Byzantine attacks in distributed algorithms are modeled by a group of malicious nodes that modify the data transmitted to their neighbors. A similar model was considered in [31] for sub-gradient methods. Furthermore, a model-free moving target defense framework for the detection and mitigation of sensor and/or actuator attacks with discrete-time dynamics was considered in [32].

Contributions. The contribution of this work is threefold. i) We formulate a new class of switching-mode multiplicative attacks against the feedback signals, whereby the attacker transforms the control inputs produced by the controller according to a linear map with the objective of destabilizing the closed-loop system. This class of attacks is novel in the literature and it includes, as a special case, DoS and deception attacks. ii) We present the first stability analysis of the interconnection between an LTI dynamical system and static or dynamic controllers operating under switching multiplicative attacks. Our framework leverages analytical tools from set-valued hybrid dynamical systems theory, optimization, and feedback control theory. We characterize sufficient conditions on the frequency of the attacks to guarantee exponential stability of the closed-loop system under a suitable time-scale separation between the dynam-

ics of the plant and the dynamics of the controller. iii) For systems with time-varying exogenous inputs acting on the plant dynamics, we establish sufficient conditions to guarantee exponential input-to-state stability with respect to the essential supremum of the time derivative of the exogenous signal, similar to the notion of derivative ISS studied in [33]. To derive this result, we leverage Lyapunov theory for switched systems with inputs and unstable modes, where the switching signals are generated by a hybrid dynamical system.

Organization. The rest of this paper is organized as follows. In Section 2 we present the notation used throughout the paper and some preliminaries on hybrid dynamical systems. In Section 3, we formalize the model of the system and the problem under study. Section 4 presents the main results, followed by the analysis presented in Section 5. Section 6 presents some numerical experiments, and Section 7 concludes the paper by summarizing our main findings.

2. Preliminaries and Notation

We begin by introducing the notation that will be used throughout the paper. Given a compact set $\mathcal{A} \subset \mathbb{R}^n$, $n \in \mathbb{Z}_{\geq 0}$, and a vector $z \in \mathbb{R}^n$, we let $|z|_{\mathcal{A}} := \min_{s \in \mathcal{A}} \|z - s\|_2$ denote the minimum distance between z and $\mathcal{A}$. We denote by $\overline{\mathcal{A}}$ the closure of set $\mathcal{A}$. We denote $\mathbb{B} := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. For a bounded function $w : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, we define the sup norm as $\|w\|_t := \sup_{s \in [0,t]} \|w(s)\|$. For a matrix $A \in \mathbb{R}^{n \times d}$, with $d \in \mathbb{Z}_{\geq 0}$, we use $\|A\|$ to denote the largest singular value of A. When $n = d$, and the matrix is symmetric, we use $\bar{\lambda}(A)$ and $\underline{\lambda}(A)$ to denote the largest and smallest eigenvalue of A, respectively. Given a set O, we use $I_O(x)$ to denote the indicator function for the set O, namely $I_O(x) = 1$ if $x \in O$ and $I_O(x) = 0$ if $x \notin O$. Given vectors $p_1 \in \mathbb{R}^d, p_2 \in \mathbb{R}^n$, we denote by $(p_1, p_2) \in \mathbb{R}^{d+n}$ their concatenation. We denote the domain of a function p as $\text{dom}(p)$. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{K}$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. If, in addition, $\lim_{r \to \infty} \alpha(r) = \infty$, it is said to be of class-$\mathcal{K}_\infty$. A function $\sigma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{L}$ if it is continuous, decreasing and $\lim_{r \to \infty} \sigma(r) = 0$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be of class-$\mathcal{KL}$ if it is class-$\mathcal{K}$ in its first argument and class-$\mathcal{L}$ in its second argument.

In this paper, we will consider mathematical models corresponding to set-valued dynamical systems that combine continuous-time dynamics (with inputs) and discrete-time dynamics. These systems are referred to as Hybrid Dynamical Systems (HDS) [34], and they are characterized by the following hybrid inclusions:

$$p \in C, \quad \dot{p} \in F(p, q), \qquad p \in \mathcal{D}, \quad p^+ \in G(p), \tag{1}$$

where $p \in \mathbb{R}^n$ is the state of the system, $q \in \mathbb{R}^m$ is the input, $F : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is a set-valued map (called the *flow map*) that governs the continuous-time dynamics of the system when the state belongs to the flow set $C \subset \mathbb{R}^n$. Similarly, $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued map (called the *jump map*) that describes the discrete-time dynamics of the system when the state belongs to the jump set $\mathcal{D} \subset \mathbb{R}^n$. The tuple $\mathcal{H} = \{C, F, \mathcal{D}, G\}$ completely characterizes the HDS, and it is called the *data* of the system. A HDS is well-posed if it satisfies [34, Def. 6.29] and it is nominally well-posed if it satisfies [34, Def. 6.2]. We recall that a well-posed hybrid system is nominally well-posed. Solutions to (1) are defined on *hybrid time domains*, namely, they are indexed by a parameter $t \in \mathbb{R}_{\geq 0}$ that increases continuously during flow (i.e. while the continuous dynamics are active), and by a parameter $j \in \mathbb{Z}_{\geq 0}$ that increases by one unit during the jumps (i.e., when the discrete dynamics are active). A set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ is called a *compact* hybrid time domain if $E = \cup_{j=0}^{J-1}([t_j, t_{j+1}], j)$ for some $0 = t_0 \leq t_1 \leq \ldots \leq t_J$. The set E is a hybrid time domain if for all $(T, J) \in E$, the set $E \cap ([0, T] \times \{0, \ldots, J\})$ is a compact hybrid time domain. Using the notion of hybrid time domains we can formally introduce the concept of solution to systems of the form (1), which is borrowed from [34, Ch. 2].

Definition 1. *A function $p : dom(p) \to \mathbb{R}^n$ is a hybrid arc if $dom(p)$ is a hybrid time domain and $t \mapsto p(t, j)$ is locally absolutely continuous for each j such that the interval $I_j := \{t : (t, j) \in dom(p)\}$ has nonempty interior. A hybrid arc p is a* solution *to (1) with $dom(q) = dom(p)$ if $p(0, 0) \in \overline{C} \cup \mathcal{D}$, and the following two conditions hold:*

1. *For each $j \in \mathbb{Z}_{\geq 0}$ such that I_j has nonempty interior: $p(t, j) \in C$ for all $t \in int(I_j)$, and $\dot{p}(t, j) \in F(p(t, j), q(t, j))$ for almost all $t \in I_j$.*

2. *For each $(t, j) \in dom(p)$ such that $(t, j + 1) \in dom(p)$: $p(t, j) \in \mathcal{D}$, and $p(t, j + 1) \in G(p(t, j))$.*

In this paper, we will always work with HDS that satisfy the following Basic Conditions whenever q is constant:

a) C and $\mathcal{D}$ are closed subset of $\mathbb{R}^n$.

    b) $F(\cdot, q)$ is outer semicontinuous and locally bounded relative to C, $C \subset \text{dom } F$, and $F(p, q)$ is convex for every $p \in C$.

    c) $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is outer semicontinuous and locally bounded relative to $\mathcal{D}$, and $\mathcal{D} \subset \text{dom } G$.

On the other hand, when q is not constant we will make it a function $\Phi(\cdot)$ of the state p, and in this case the Basic Condition b) should hold for the set-valued map $\hat{F}(\cdot) := F(\cdot, \Phi(\cdot))$. By working with hybrid time domains and nominally well-posed HDS (note that the Basic Conditions are sufficient to obtain well-posedness of the system [34, Thm. 6.8]) we can exploit suitable graphical convergence notions to establish sequential compactness results for the solutions of (1), e.g., the graphical limit of a sub-sequence of solutions is also a solution. Such types of properties will be instrumental for the robustness analysis of the systems studied in this paper.

Definition 2. *A hybrid solution p is maximal if there does not exist another solution ψ to $\mathcal{H}$ such that $dom(p)$ is a proper subset of $dom(\psi)$, and $p(t, j) = \psi(t, j)$ for all $(t, j) \in dom(p)$. A maximal hybrid solution is said to be complete if its domain is unbounded.*

In this paper, we are interested in establishing suitable *convergence and stability* properties for a class of dynamical systems under attack. To this end, the following definitions will be instrumental.

Definition 3. *Given a compact set $\mathcal{A} \subset C \cup \mathcal{D}$, system (1) with $u = 0$ is said to render $\mathcal{A}$ uniformly globally asymptotically stable (UGAS) if there exists a class $\mathcal{KL}$ function β such that every solution of (1) satisfies $|p(t, j)|_{\mathcal{A}} \le \beta(|p(0, 0)|_{\mathcal{A}}, t + j)$ for all $(t, j) \in dom(p)$. If, additionally, $\beta(r, s) = c_1 r e^{-c_2 s}$ for $c_1, c_2 > 0$, the system is said to render the set $\mathcal{A}$ uniformly globally exponentially stable (UGES).*

Note that the stability notion of Definition 3 implies the standard asymptotic stability properties considered in the literature of continuous-time systems ($\mathcal{D} = \emptyset$) and discrete-time systems ($C = \emptyset$). In particular, recall that a continuous-time LTI dynamical system given by $\dot{x} = Ax$ renders the origin $\mathcal{A} = \{0\}$ UGES if and only if A is Hurwitz, i.e., $\text{Re}\{\lambda(A)\} < 0$, for all eigenvalues λ of A.

    When the input u in (1) is not identically zero, the notion of *input-to-state stability* can be used to qualitatively characterize the effect of the input on the stability properties of the system. We refer to [35] for a formal definition of the sup norm $\|q\|_{(t, j)}$ on hybrid time domains.

Definition 4. *For every measurable function $q : dom(q) \to \mathbb{R}^m$ with $dom(q) = dom(p)$, system (1) is said to render the compact set $\mathcal{A}$ input-to-state stable (ISS) with respect to q if there exists a class $\mathcal{KL}$ function β, and a class $\mathcal{K}$ function γ, such that every solution of the system satisfies the bound*

$$|p(t, j)|_{\mathcal{A}} \le \beta(|p(0, 0)|_{\mathcal{A}}, t + j) + \gamma\left(\|q\|_{(t, j)}\right), \tag{2}$$

for all $(t, j) \in dom(p)$. When β has an exponential form, we say that system (1) renders the set $\mathcal{A}$ exponentially input-to-state stable (E-ISS).

The properties of UGAS, UGES, and ISS for hybrid systems can be readily established via suitable Lyapunov-based conditions; see [34, Ch.3] for sufficient conditions that certify UGAS, [36, Thm. 1] for UGES, and [35, Thm. 3.1] for ISS.

3. Problem Formulation

    In this section, we formalize the problem studied in this work. We begin by describing the system and controller model in the absence of attacks. Subsequently, we present our attack model, and we formalize the joint control and optimization problem that is the focus of this work.

3.1. Nominal Model of the Plant and the Controller

    We consider plants modeled as LTI dynamical systems of the form:

$$\dot{x} = Fx + Nv + Bu + Ew, \quad y = Cx, \tag{3}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^{m_v}$ are control inputs, $w \in \mathbb{R}^q$ is a time-varying *unknown* exogenous signal, $y \in \mathbb{R}^p$ is the measurable output, and F, N, B, E, C are matrices of suitable dimensions. We consider general cases where the matrix F may not be Hurwitz-stable. We make the following regularity assumption on the exogenous input w.

Assumption 1. *The exogenous input w is generated by an exosystem of the form $\dot{w} = \Pi(w)$, with $w(0) \in \mathcal{W}$, where $\Pi(\cdot)$ is Lipschitz continuous, and $\mathcal{W} \subset \bar{w}\mathbb{B}$ is a compact set. Moreover, this exosystem renders the set $\mathcal{W}$ forward invariant.*

Throughout this paper, functions w satisfying the conditions of Assumption 1 are said to be *admissible*.

As graphically illustrated in the center panel of Figure 1, we consider two control loops applied to the plant (3). First, an inner control loop, of the form of a static output feedback law, is applied via the control input v:

$$v = Ky, \quad K \in \mathbb{R}^{m_v \times p}. \tag{4}$$

Using (3) and (4), we denote by $A := F + NKC$ the matrix describing the closed-loop dynamics of the plant. We assume that the inner feedback loop is stabilizing, as formalized in the following assumption.

Assumption 2. *For every positive definite matrix $R \in \mathbb{R}^{n \times n}$, there is a unique positive definite matrix $P \in \mathbb{R}^{n \times n}$ such that $A^\top P + PA = -R$.*

Additionally, we consider an outer loop of the form of a low-gain dynamical controller that acts on the input u as follows (see Fig. 1):

$$\dot{u} = -\varepsilon \left(\nabla f_u(u) + G^\top \nabla f_y(y) \right), \quad G := -CA^{-1}B, \tag{5}$$

where $f_u : \mathbb{R}^m \to \mathbb{R}$ and $f_y : \mathbb{R}^p \to \mathbb{R}$ are cost functions (to be characterized below), and $\varepsilon > 0$ is a tunable gain.

By combining the plant model (3) with the controllers (4)-(5), we obtain the following nominal closed-loop system with states $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$\dot{x} = (F + NKC)x + Bu + Ew, \quad y = Cx,$$
$$\dot{u} = -\varepsilon \left(\nabla f_u(u) + G^\top \nabla f_y(y) \right). \tag{6}$$

As shown in [3, 9], under Assumption 2 and suitable conditions on the gradients $(\nabla f_u, \nabla f_y)$, there exists a sufficiently small gain ε such that the controller (5) regulates system (3) towards the solutions of:

$$\min_{u,x,y} \ f_u(u) + f_y(y), \quad \text{s.t.} \quad 0 = Ax + Bu + Ew, \quad y = Cx. \tag{7}$$

where $f_u : \mathbb{R}^m \to \mathbb{R}$ is interpreted as a cost function associated with the steady-state control input, and $f_y : \mathbb{R}^p \to \mathbb{R}$ is interpreted as a cost associated with the steady-state system output. The optimization problem (7) formalizes an optimal equilibrium-selection problem, where the objective is to select an input-state-output triplet (u, x, y) that minimizes the cost function $f_u(u) + f_y(y)$ at steady state. Moreover, since (7) is parametrized by the (deterministic) time-varying signal w, solutions of (7), denoted by (u^*, x^*, y^*) will also be time-varying.

Remark 1. *Steady state optimization problems of the form (7) have received increased attention in the literature during the last years, partially motivated by applications in power systems [2, 37, 10, 9] and transportation networks [38]. Note that when A is Hurwitz the steady-state constraint can be solved for x and substituted in the cost function, leading to an unconstrained optimization problem defined over u.*

Let $G := -CA^{-1}B$ and $H := -CA^{-1}E$, where the inverse of A is well-defined thanks to Assumption 2. Problem (7) can be written as:

$$\min_u f(u, w) := f_u(u) + f_y(Gu + Hw), \tag{8}$$

where w acts as an exogenous signal that parametrizes the optimal solution. Note that a solution of (7) is also a solution of (8), but the reverse implication may not be true, unless the pair (C, A) is observable. Nonetheless, observability is not needed for the subsequent analysis, since we focus on solving problem (8). In the sequel, we use the following expressions for the partial gradients of the costs:

$$\nabla_u f(u, w) = \nabla_u f_u(u) + G^\top \nabla_y f_y(Gu + Hw), \qquad \nabla_w f(u, w) = H^\top \nabla_y f_y(Gu + Hw).$$

We make the following standard regularity assumptions:

Assumption 3. *The functions $u \mapsto f_u(u)$ and $y \mapsto f_y(y)$ are continuously differentiable and globally Lipschitz, namely, there exist ℓ_u, $\ell_y > 0$ such that $\left\| \nabla f_u(u) - \nabla f_u(u') \right\| \leq \ell_u \left\| u - u' \right\|$ and $\left\| \nabla f_y(y) - \nabla f_y(y') \right\| \leq \ell_y \left\| y - y' \right\|$ hold for every $u, u' \in \mathbb{R}^m$ and $y, y' \in \mathbb{R}^p$.*

Assumption 4. *For each w, the function $u \mapsto f(u, w)$ has a unique (bounded) minimizer, denoted by u_w^*. Furthermore, the function $u \mapsto f(u, w)$ satisfies the PL inequality, uniformly in w, namely, $\exists \mu > 0$ such that $\frac{1}{2} \left\| \nabla_u f(u, w) \right\|^2 \geq \mu(f(u, w) - f(u_w^*, w)), \forall u \in \mathbb{R}^m, w \in \mathcal{W}$.*

Assumption 5. *For each $w \in \mathcal{W}$, the solution of (8) satisfies $u_w^* = h(w)$, where $w \mapsto h(w)$ is a continuously differentiable function.*

Remark 2. *Assumption 3 guarantees that the cost functions are sufficiently smooth, which is a standard assumption in the literature of online optimization (see also [2, 10]). Indeed, by Assumption 3, the mapping $u \mapsto f(u, w)$ has a globally Lipschitz-continuous gradient, uniformly in w, with Lipschitz constant $\ell = \ell_v + \ell_y ||G||^2$. On the other hand, Assumption 4 implies that f is invex, and it is one of the weakest assumptions in the optimization literature that ensures exponential convergence to a set of global optimizers [6]. We also note that the PL inequality implies the quadratic growth condition $f(u, w) - f(u_w^*, w) \geq \frac{\mu}{2} \left\| u - u_w^* \right\|^2, \forall u \in \mathbb{R}^m$ for a fixed w; since the minimizer is assumed to be unique and bounded (cf. Assumption 1, 4 and 5), $\mathcal{W}$ is compact, and $h(\cdot)$ is continuous, the quadratic growth condition implies radial unboundedness of f. Finally, note that by Assumption 2, and for fixed u and w, there is a unique solution x_w^* to the constraint $0 = Ax + Bu + Ew$. Thus, by Assumptions 1-5, the solution of (8) exists, it is unique for each time t, and it is a continuous function of w.*

Under Assumptions 3–4, the following chain of inequalities holds:

$$\frac{\mu^2}{2} \left\| u - u_w^* \right\|^2 \leq \mu(f(u, w) - f(u_w^*, w)) \leq \frac{1}{2} \left\| \nabla_u f(u, w) \right\|^2 \leq \frac{\ell^2}{2} \left\| u - u_w^* \right\|^2.$$

The following two lemmas will also be instrumental for our results.

Lemma 1. *Let Assumptions 1 and 5 hold. Then $\exists U_{\bar{w}} \in \mathbb{R}_{>0}$ such that $\left\| \nabla_w h(w) \right\| \leq U_{\bar{w}}$, for all $w \in \mathcal{W}$.*

Lemma 2. *Let Assumptions 3-5 hold. Then, $\nabla_u f(h(w), w)^\mathsf{T} \nabla_w h = 0$*

The result of Lemma 1 follows from the continuity of $\nabla_w h(\cdot)$ (with respect to w) and the compactness of the set $\mathcal{W}$. Regarding Lemma 2, by first-order optimality conditions the gradient of the function f is equal to zero at the set of optimizers; the product is well-defined since $\nabla_u f(u, w)$ and $\nabla_w h$ are continuous by Assumptions 3 and 5, and $\nabla_w h$ is bounded by Lemma 2, hence, the result.

3.2. Attack Model

In this work, we consider attacks that can modify the control signals generated by the controllers. As also illustrated in Figure 1, we focus on attacks that are intermittent and occur in a persistent fashion, but are not necessarily periodic. We account for two types of attacks. First, we consider attacks acting on the gradient-flow controller (5), which can be modeled as:

$$\text{Nominal System: } \dot{u} = -\varepsilon \left(\nabla f_u(u) + G^\mathsf{T} \nabla f_y(y) \right), \tag{9a}$$

$$\text{System Under Attack: } \dot{u} = -\varepsilon M_{\sigma_u} \left(\nabla f_u(u) + G^\mathsf{T} \nabla f_y(y) \right), \tag{9b}$$

where $\sigma_u : \mathbb{R}_{\geq 0} \to \Sigma_{u,a}$ is a switching signal taking values in a finite set of indices $\Sigma_{u,a} \subset \mathbb{Z}_{>0}$, and $M_{\sigma_u} \in \mathbb{R}^{m \times m}$ is a matrix that describes the attack transformation map. The model (9) is general enough to capture different types of attacks, including attacks that modify the sign and/or the magnitude of (some or all components of) the gradient (see e.g. [28, 29, 30, 31, 39, 40, 41] for similar attack models), jamming attacks, and DoS attacks (see e.g. [15, 26]).

Second, we consider attacks targeting the static feedback controller (4), which can be modeled as:

$$\text{Nominal System: } \quad v = Ky, \qquad \dot{x} = Fx + Nv + Bu + Ew, \tag{10a}$$

$$\text{System Under attack:} \quad v = L_{\sigma_v} K y, \quad \dot{x} = Fx + Nv + L_{\sigma_v}^b Bu + L_{\sigma_v}^e Ew, \tag{10b}$$

where $\sigma_v : \mathbb{R}_{\geq 0} \to \Sigma_{v,a}$ is a switching signal taking values in a finite set of indices $\Sigma_{v,a} \subset \mathbb{Z}_{>0}$. For each $\sigma_v \in \Sigma_{v,a}$, $L_{\sigma_v} \in \mathbb{R}^{m_v \times m_v}$ is a matrix that describes the transformation map applied by the attacker to the static feedback loop, and the matrices $L_{\sigma_v}^e, L_{\sigma_v}^b \in \mathbb{R}^{n \times n}$ modify the inputs Bu and Ew, respectively.

For our analysis, we next unify the models of the *Nominal System* and of the *System Under Attack* by combining the two modes into a single switching-system model. Tot his aim, we describe the *Nominal System* in (10) by introducing the index set $\Sigma_{v,s} = \{s\}$ and the identity transformation map $L_s = I \in \mathbb{R}^{m_v \times m_v}$. Similarly, we describe the *Nominal System* in (9) by introducing the index set $\Sigma_{u,s} = \{s\}$ and the identity transformation map $M_s = I \in \mathbb{R}^{m \times m}$. Using these definitions, we can rewrite systems (9)-(10) as switching systems of the form

$$\dot{u} = -\varepsilon M_{\sigma_u} \left(\nabla f_u(u) + G^\top \nabla f_y(y) \right), \tag{11}$$

$$v = L_{\sigma_v} K y, \qquad B_{\sigma_v} := L_{\sigma_v}^b B, \qquad E_{\sigma_v} := L_{\sigma_v}^e E, \tag{12}$$

where now the switching signal σ_v takes values in the extended set $\Sigma_v := \Sigma_{v,s} \cup \Sigma_{v,a}$, and the switching signal σ_u takes values in the extended set $\Sigma_u := \Sigma_{u,s} \cup \Sigma_{u,a}$.

Assumption 6. *The sets Σ_v and Σ_u are finite. Additionally, any attack is destabilizing, namely:*

(i) *For all $\sigma_v \in \Sigma_{v,a}$, the matrix $F + N L_{\sigma_v} KC$ is invertible and admits at least one eigenvalue with positive real part.*

(ii) *For all $\sigma_u \in \Sigma_{u,a}$, the matrix M_{σ_u} has at least one eigenvalue with negative real part. Moreover, we define*
$$\bar{M} := \max_{\sigma_u \in \Sigma_{u,a}} \left\| M_{\sigma_u} \right\|.$$

It follows from Assumption 6 that modes described by the sets $\Sigma_{v,a}$ and $\Sigma_{u,a}$ can make the closed-loop system unstable (6). Moreover, since the sets $\Sigma_{v,a}$ and $\Sigma_{u,a}$ are countable and finite, with a slight abuse of notation, in the remainder we will enumerate these unstable modes as follows: $\Sigma_{v,a} := \{a_1, \dots, a_{|\Sigma_{v,a}|}\}$ and $\Sigma_{u,a} := \{a_1, \dots, a_{|\Sigma_{u,a}|}\}$, where the distinction between the two sets will be made clear by the context.

Lemma 3. *Let Assumption 6-(i) hold. Let $A_{a_i} := F + N L_{a_i} KC$ for $a_i \in \Sigma_{v,a}$, and let $\hat{R}_{a_i} := (A_{a_i}^\top + A_{a_i})$. Then, $\exists \lambda_i > 0$, and $x_i \in \mathbb{C}^n$, $x_i \neq 0$ such that $\hat{R}_{a_i} x_i = \lambda_i x_i$[1].*

Proof: Suppose by contradiction that $\hat{R}_{a_i} \leq 0$, and let $x_a \in \mathbb{C}^n$ be a non-zero vector such that $A_{a_i} x_a = \lambda_a x_a$, then[2]:

$$x_a^H \hat{R}_{a_i} x_a = x_a^H (A_{a_i}^\top + A_{a_i}) x_a = (x_a^H \lambda_a^H x_a + x_a^H \lambda_a x_a) = (\lambda_a^H + \lambda_a) x_a^H x_a = 2 \text{Re}(\lambda_a) x_a^H x_a,$$

and since $\hat{R}_{a_i} \leq 0$ and $x_a \neq 0$, we must have that $\text{Re}(\lambda_a) \leq 0$. Since λ_a is arbitrary, this contradicts Assumption 6-(i). ∎

For any $a_i \in \Sigma_{v,a}$, the matrix $\hat{R}_{a_i}$ is real and symmetric; hence, Rayleigh theorem holds [43, Thm. 4.2.2] and $\underline{\lambda}(\hat{R}_{a_i}) \|x\|^2 \leq x^\top \hat{R}_{a_i} x \leq \bar{\lambda}(\hat{R}_{a_i}) \|x\|^2$. Throughout the rest of this paper, we denote by $\bar{\lambda}(\hat{R}_a) := \max_{a_i \in \Sigma_{v,a}} \bar{\lambda}(\hat{R}_{a_i})$ the largest eigenvalue of all matrices $\hat{R}_{a_i}$.

3.3. Admissible Persistent Attacks

The models (9) and (10) describe a family of persistent multiplicative attacks that intermittently modify the control signal generated by the controllers (4) and (5) via the switching maps L_{σ_v} and M_{σ_u}. As described in Assumption 6, such attacks can induce instability in the closed-loop system (6) by persistently altering the direction of the vector field in the dynamic controller (5), or the gains of the feedback law (4). We note that, because the attack is multiplicative rather than additive, in general, standard approaches used to guarantee closed-loop stability under non-persistent adversarial attacks (such as dynamic filters [20], identification mechanisms [21], dynamic compensators [18], etc) may

---

[1] The existence of eigenvectors and eigenvalues in this case can be proved as in [42].
[2] We denote the conjugate of a matrix $x \in \mathbb{C}^{m \times n}$ as x^H.

not guarantee that the attack is mitigated at all times. For this reason, in this work, we investigate cases where such destabilizing attacks can be tolerated during bounded fractions of time. Specifically, for every time-interval $0 \leq s < t$, we define the total activation time of the attacks (9), and (10), respectively, as follows:

$$T_u(s,t) := \int_s^t I_{\Sigma_{u,a}}(\sigma_u(\tau))d\tau, \qquad \text{and} \qquad T_v(s,t) := \int_s^t I_{\Sigma_{v,a}}(\sigma_v(\tau))d\tau,$$

where we recall that $I_\Sigma(\sigma)$ denotes the indicator function for the set Σ (see Section 2). Notice that, if $\sigma_v = \sigma_u = s$ at all times, then $T_v(s,t) = T_u(s,t) = 0$ for any $0 \leq s < t$ since no attack is active in the closed-loop. The class of admissible attacks is formalized next.

Definition 5. *(Class of Admissible Attacks) Given $\kappa_{v,1}, \kappa_{u,1}, \kappa_{v,2}, \kappa_{u,2} > 0$ and $N_{0,v}, N_{0,u}, T_{0,v}, T_{0,u} \in \mathbb{Z}_{>0}$, the attacks (9) and (10) are said to be admissible for system (6) if the number of discontinuities of σ_v and σ_u in the time-interval $[s,t)$, denoted by $N_v(s,t)$ and $N_u(s,t)$, respectively, satisfy*

$$(a) \qquad N_v(s,t) \leq \kappa_{v,1}(t-s) + N_{0,v}, \quad N_u(s,t) \leq \kappa_{u,1}(t-s) + N_{0,u}, \tag{13}$$

and the activation times $T_v(s,t)$ and $T_u(s,t)$ satisfy:

$$(b) \qquad T_v(s,t) \leq \kappa_{v,2}(t-s) + T_{0,v}, \quad T_u(s,t) \leq \kappa_{u,2}(t-s) + T_{0,u}. \tag{14}$$

Condition (13) imposes an average dwell-time constraint on the switches of σ_u and σ_v, and thus guarantees that any admissible attack cannot generate Zeno behavior. On the other hand, condition (14) says that the total activation time of any admissible persistent attack in a given time interval must grow at most linearly with the length of the time interval. We note that larger values of $\kappa_{v,2}$ and $\kappa_{u,2}$ describe more intrusive attacks, namely, attacks the can remain active for longer periods of time. Finally, we observe that similar attack models have been considered in the literature (see e.g. [13, 15, 26, 12]). Indeed, conditions (13)-(14) are rather general as they capture periodic or aperiodic attacks, and they also allow a finite number of consecutive switches in every interval of time of sufficiently large length.

Remark 3. *In cases where the closed-loop system is monitored by an attack-detection mechanism, the quantities $\{T_{0,v}, \kappa_{v,2}, T_{0,u}, \kappa_{u,2}\}$ can be interpreted as attack design parameters that can be tuned by the attacker to remain undetected from the monitor. Alternatively, in cases where the closed-loop system includes an attack-mitigation mechanism that can reject the effects of attacks for limited periods of time, the quantities $\{T_{0,v}, \kappa_{v,2}, T_{0,u}, \kappa_{u,2}\}$ can be interpreted as parameters that describe the effectiveness of the mitigation mechanism. Both engineering interpretations will be consistent with our theoretical framework.*

3.4. Control Objectives: Approximate Optimal Tracking Under Attacks

Based on the model introduced in the previous sections, we can now formalize the problem under study in this paper. In particular, for the family of closed-loop systems under attacks described by system (6), and satisfying Assumptions 2-6, we are interested in characterizing the admissible persistent attacks, and the time-scale separation between the plant and the controller needed to preserve the stability properties of the nominal closed-loop system. Moreover, when w is constant, we are interested in guaranteeing exponential convergence of the trajectories of the closed-loop system to the set of solutions of the online optimization problem described by (7). Notice that, when w is time-varying, convergence to the solution of the optimization problem can only be achieved up to an error that depends on $\dot{w}$. In this case, we seek to quantify this error using the notion of input-to-state stability.

Based on this, the following problem formalizes the stated objectives

Problem 1. *Determine, when possible, a set of parameters $(\varepsilon, \kappa_{v,1}, \kappa_{v,2})$, or $(\varepsilon, \kappa_{u,1}\kappa_{u,2})$ such that the optimal point $z^* = (x^*, u_w^*)$ of system (6) under attack is exponentially input-to-state stable (E-ISS) with respect to the time-variation (i.e. the derivative) of the exogenous signal $t \mapsto w(t)$.*

Remark 4. *For simplicity and space limitations, we will focus on attacks that are exclusive either to the plant or to the controller, i.e. we do not address simultaneous attacks to the static and dynamic feedback controller. This allows us to simplify our notation since in this case, if $\sigma_v = a_i$, then $\sigma_u = s$, and, if $\sigma_u = a_i$, then $\sigma_v = s$ as defined before. Note, however, that considering simultaneous attacks in the plant and the controller is a natural extension of this work that can be studied following similar steps as in this paper.*

Before presenting our main results, we list three representative examples of applications in cyber-physical systems for which the modeling and control framework considered in this paper is particularly well-suited.

Example 1. *The dynamical model* (3) *and the interconnection* (6) *fit within the context of real-time frequency control and economic optimization of power transmission systems. In particular, the matrix K can be designed based on the so-called automatic generation control (AGC) as well as pertinent droop controllers implemented in the generation units [44]; on the other hand, the gradient-flow controller in* (6) *can be utilized to produce setpoints for the generators (both conventional fossil-fuel and renewable-based units) to solve an economic dispatch (ED) or optimal power flow (OPF) problem in real time [45]; see also the example provided in [2]. The models of attacks considered in this paper capture in this case malicious attacks to the AGC signals, droop control loops, and ED/OPF commands.*

Example 2. *The model* (3) *and* (6) *is also suitable for applications in the context of intelligent transportation systems, where the properties of safety and resilience are critical. For example, for a ramp metering problem in a highway system, our inner control loop can model local controllers such as the ALINEA [46]. On the other hand, gradient-flow controllers have been used to drive the equilibrium point of the traffic flows on the highway system towards the solution of a network-level problem as explained in, e.g., [38] and [47]. It then follows that the attack models consider in* (11) *can be used to capture adversarial jamming of the feedback gains of the ALINEA controller, or attacks acting on the control directions obtained from the gradient flow algorithm.*

Example 3. *The model* (11) *is also suitable for applications in robotics, such as the control of autonomous vehicles using navigation laws based on potential and anti-potential functions; see [48, 49, 50]. Indeed, in this scenario, it is common to approximate the vehicle dynamics as a linear system* (3) *for which an internal feedback loop is designed to stabilize an external reference u; see [51]. In general, the internal controller is designed to guarantee a unitary DC gain, and to operate in a faster time scale (limited by the actuator dynamics of the vehicle), such that the "steady state" of the vehicle satisfies $x^* = u$, where x can model positions or velocities. The input u is then regulated via a gradient flow of the form* (11) *aiming to converge to the minimizer (e.g., the target point) of an artificial potential field. In this case, to prevent the vehicle from converging to the target, an external attacker can persistently alter the sign of the gradient dynamics, i.e., $M_{\sigma_u} \in \{I, -I\}$, effectively moving the vehicle away from the target. Note that this type of attacks will be particularly detrimental for navigation laws that implement anti-potential fields to avoid obstacles, given that in this case, the attacks will effectively push the vehicles towards the obstacles.*

4. Main Results

To address Problem 1, we will use the framework of HDS (1), which allows us to model the switching systems (9)-(10) as time-invariant (hybrid) systems, and also to consider suitable (set-valued) dynamic models of the admissible persistent attacks. For simplicity of exposition, we drop the subscripts from (13) - (14), and we consider a hybrid automaton with state $\varrho := (\tau_1, \tau_2, \sigma) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \Sigma$, and the following hybrid dynamics:

$$\varrho \in C_\varrho := [0, N_0] \times [0, T_0] \times \Sigma, \tag{15a}$$

$$\begin{pmatrix} \dot{\tau}_1 \\ \dot{\tau}_2 \\ \dot{\sigma} \end{pmatrix} \in F(\varrho) := \begin{pmatrix} [0, \kappa_1] \\ [0, \kappa_2] - I_{\Sigma_a}(\sigma) \\ 0 \end{pmatrix}, \tag{15b}$$

$$\varrho \in D_\varrho := [1, N_0] \times [0, T_0] \times \Sigma, \tag{15c}$$

$$\begin{pmatrix} \tau_1^+ \\ \tau_2^+ \\ \sigma^+ \end{pmatrix} \in G(\varrho) := \begin{pmatrix} \tau_1 - 1 \\ \tau_2 \\ \Sigma \backslash \sigma \end{pmatrix}, \tag{15d}$$

where $T_0 \geq 0$, $N_0 \in \mathbb{Z}_{\geq 1}$, $\kappa_1 > 0$, $\kappa_2 \in (0, 1)$. The following Lemma, originally established in [13], will be instrumental for our results.

Lemma 4. *Every complete solution ϱ of* (15) *has a hybrid time-domain dom(ϱ) that satisfies condition* (13) *and* (14) *with $N_0 = N_{0,v} = N_{0,u}$, and $T_0 = T_{0,v} = T_{0,u}$. Moreover, every signal σ modeling an admissible attack can be generated by the dynamical system* (15) *with a suitable initial condition.*

Lemma 4 allows us to study the closed-loop system under persistent switching attacks by studying the interconnection between the dynamics (9), (10) and (15). Similar models have been studied in the literature for the analysis of *static* model-based and model-free optimization problems; see [12, 13]. However, unlike these works, Problem 1 describes an *online optimization problem* with a *dynamic* plant in the loop, which has not been studied before in the literature.

In the ensuing section, we derive sufficient conditions for the stability of the interconnected system when the gradient-flow controller is affected by attacks as in (9). After this, we establish similar results for the case when the static feedback controller is under attack, as in (10).

4.1. Attacks Against Dynamic Feedback

In this section, we analyze the stability properties of the closed-loop system with attacks acting on the dynamic gradient-based controller. To model this scenario, we consider the following HDS with state $\vartheta := (x, u, \tau_{u,1}, \tau_{u,2}, \sigma_u, w)$, and dynamics:

$$\vartheta \in C := \mathbb{R}^n \times \mathbb{R}^m \times [0, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u \times \mathcal{W}, \tag{16a}$$

$$\begin{pmatrix} \dot{x} \\ \dot{u} \\ \dot{\tau}_{u,1} \\ \dot{\tau}_{u,2} \\ \dot{\sigma}_u \\ \dot{w} \end{pmatrix} \in F_{\sigma_u}(\vartheta) := \begin{pmatrix} Ax + Bu + Ew \\ -\varepsilon M_{\sigma_u} \left(\nabla f_u(u) + G^\top \nabla f_y(y) \right) \\ [0, \varepsilon\kappa_{u,1}] \\ [0, \varepsilon\kappa_{u,2}] - \varepsilon I_{\Sigma_{u,a}}(\sigma_u) \\ 0 \\ \Pi(w) \end{pmatrix}, \quad y = Cx, \tag{16b}$$

$$\vartheta \in \mathcal{D} := \mathbb{R}^n \times \mathbb{R}^m \times [1, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u \times \mathcal{W}, \tag{16c}$$

$$\begin{pmatrix} x^+ \\ u^+ \\ \tau_{u,1}^+ \\ \tau_{u,2}^+ \\ \sigma_u^+ \\ w^+ \end{pmatrix} \in G_{\sigma_u}(\vartheta) := \begin{pmatrix} x \\ u \\ \tau_{u,1} - 1 \\ \tau_{u,2} \\ \Sigma_u \backslash \{\sigma_u\} \\ w \end{pmatrix}. \tag{16d}$$

This HDS models the interconnection between the signal generator (15), the closed-loop system (6) with the dynamic controller under attack (9), and the exosystem that generates w, which is treated as an auxiliary state. Note that this is similar to the setting considered in [33, Prop. 5.1] in the context of derivative ISS.

Remark 5. *In* (16b)*, the dynamics of the states* $\tau_{u,1}$ *and* $\tau_{u,2}$ *are multiplied by* ε*. This is done without loss of generality, and only to simplify the modeling framework such that the attacks operate in the same time scale as the gradient controller. In fact, since in our case* $\varepsilon \in (0, 1)$*, the dynamics of* (16b) *could also be written in terms of the scaled time ratios* $\tilde{\kappa}_{u,1} = \varepsilon\kappa_{u,1}$ *and* $\tilde{\kappa}_{u,2} = \varepsilon\kappa_{u,2}$*, and a* ε*-scaled indicator function.*

The following Lemma, which follows directly by construction, will play an important role in the robustness analysis of system (16).

Lemma 5. *For the HDS* (16) *the following holds: (a) The sets* C *and* $\mathcal{D}$ *are closed; (b) The set-valued mapping* $F_{\sigma_u}(\cdot)$ *is outer-semicontinuous, locally bounded, and convex-valued in* C*; (c) The set-valued mapping* G_{σ_u} *is outer-semicontinuous and locally bounded in* $\mathcal{D}$*.*

For system (16), we are interested in characterizing sufficient conditions on the gain ε of the controller, the dwell-time parameter $\kappa_{u,1}$, and the time-ratio parameter $\kappa_{u,2}$, to guarantee the solution of Problem 1. To do this, we first neglect the plant dynamics by assuming they are instantaneous. Namely, we disregard the dynamics $\dot{x}$, and we substitute x by the steady-state mapping $x(u, w) = A^{-1}Bu + Ew$, where w is assumed to be a constant, i.e., $\dot{w} = 0$. For this simplified system, we establish the following proposition, which is the first result of this paper. It establishes a novel characterization of an admissible persistent attack.

Proposition 1. *Suppose that Assumptions 1-6 hold, and that the dwell-time parameter* $\kappa_{u,1}$*, and the time-ratio parameter* $\kappa_{u,2}$ *simultaneously satisfy:*

$$2\mu > \ln(\omega)\kappa_{u,1} + 2\kappa_{u,2}(\mu + \bar{M}\ell), \quad with \quad \omega := \max\left\{ \frac{1}{\mu}, \frac{\ell^2}{\mu} \right\}. \tag{17}$$

Then, for each compact set $K_x \subset \mathbb{R}^n$, the set

$$\mathcal{A}^*_{K_x} = \left\{ (x, u, \tau_{u,1}, \tau_{u,2}, \sigma_u, w) : x \in K_x, \ u = u^*_w, \ \tau_{u,1} \in [0, N_{0,u}], \tau_{u,2} \in [0, T_{0,u}], \sigma_u \in \Sigma_u, w \in \mathcal{W} \right\},$$

is UGES for the HDS (16) *with $\dot{x} = 0$, $\varepsilon = 1$, $y = Gu + Hw$, $\dot{w} = 0$ and with restricted flow and jump sets*

$$C_{K_x} = (\mathbb{R}^n \cap K_x) \times \mathbb{R}^m \times [0, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u \times \mathcal{W}, \ D_{K_x} = (\mathbb{R}^n \cap K_u) \times \mathbb{R}^m \times [1, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u \times \mathcal{W}.$$

The result of Proposition 1 establishes an explicit characterization of parameters $(\kappa_{u,1}, \kappa_{u,2})$ that guarantees exponential stability of the plant dynamics under persistent attacks. In particular, by (17), we obtain that an admissible attack is dependent on the largest value $\left\| M_{\sigma_u} \right\|$ induced by the attacks in the controller, and the function parameters μ and ℓ. Moreover, the parameters $(\kappa_{u,1}, \kappa_{u,2})$ are uniform on the initial conditions of the plant and also on the controller. This uniformity property, established via hybrid Lyapunov functions in the next section, will be instrumental for the analysis of the closed-loop system for the case when $\dot{x} \neq 0$.

Remark 6. *The restriction to a compact set K_x of the x-component of the flow and jump set is imposed for the purpose of regularity –namely, to guarantee that $\mathcal{A}^*_{K_x}$ is compact. Similarly, the stability properties of the signal generator* (15) *and w are asserted with respect to the set $[0, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u$ and $\mathcal{W}$, which are forward pre-invariant and trivially attractive since, by construction of the flow and jump set, no solution can start outside $[0, N_{0,u}] \times [0, T_{0,u}] \times \Sigma_u$ or $\mathcal{W}$ respectively.*

Next, we incorporate the dynamics of the plant into the stability analysis of the closed-loop system. In this case, we provide an upper bound for the gain of the controller, which guarantees exponential stability of the optimal point $z^* = (x^*, u^*_w)$ for the case when $\dot{w} = 0$.

Proposition 2. *Suppose that Assumptions 1-6 hold, and let $\kappa_{u,1}$, and $\kappa_{u,2}$ satisfy the bound* (17)*. Let the gain ε satisfy:*

$$0 < \varepsilon < \varepsilon^* := \frac{\rho \underline{\lambda}(R) \min\{\mu, 1\}}{2\ell_y \max\left\{\bar{M}, 1\right\} \|C\| \|G\| \|PA^{-1}B\| \left(\rho \min\{\mu, 1\} + 2\ell \max\left\{\ell, \bar{M}\right\} e^{\tau_0}\right)}, \tag{18}$$

where $\rho > 0$, and $\tau_0 = \ln(\omega) N_{0,u} + 2T_{0,u}(\mu + \bar{M}\ell)$, and P is defined as in Assumption 2. Then, the set

$$\mathcal{A}^* = \left\{ (x, u, \tau_{u,1}, \tau_{u,2}, \sigma_u, w) : x = x^*, u = u^*_w, \ \tau_{u,1} \in [0, N_{0,u}], \tau_{u,2} \in [0, T_{0,u}], \sigma_u \in \Sigma_u, w \in \mathcal{W} \right\}, \tag{19}$$

is UGES for the HDS (16) *with $\dot{w} = 0$.*

The result of Proposition 2 establishes a solution to Problem 1 under a constant w. Note that the result is actually uniform in w, i.e., the time ratio $\kappa_{u,2}$ and the gain ε are independent of w. Indeed, as we will show below, they are also independent of $\|\dot{w}(t)\|_t$, which is an important property needed to avoid vanishing gains and/or shrinking safety margins under highly oscillating exogenous signals.

Since by Lemma 5 the HDS (16) is well-posed, the following additional robustness result can be asserted for the closed-loop system under attacks *and* small disturbances e acting on the states and dynamics. In general, such type of disturbances is unavoidable in practical applications due to measurements noise, numerical approximations, etc.

Lemma 6. *Consider the HDS* (16) *with $\dot{w} = 0$, and suppose that the conditions of Propositions 1 and 2 hold for the dwell-time parameter $\kappa_{u,1}$, the time-ratio parameter $\kappa_{u,2}$ and the gain $\varepsilon > 0$. Then, there exists $c_1, c_2 > 0$ such that for each compact set of initial conditions $K \subset \mathbb{R}^n \times \mathbb{R}^m$, and for each $\delta > 0$, there exists $\bar{e} > 0$ such that for any measurable disturbance e with $\sup_{t+j \geq 0} |e(t, j)| \leq \bar{e}$ and every initial condition $\vartheta(0, 0) \in K \times [0, N_{0,u}] \times [0, T_{0,u}] \times \Sigma \times \mathcal{W}$, the trajectories of the perturbed HDS*

$$\vartheta + e \in C, \quad \dot{\vartheta} \in F_{\sigma_u}(\vartheta + e) + e, \tag{20a}$$

$$\vartheta + e \in D, \quad \vartheta^+ \in G_{\sigma_u}(\vartheta + e) + e, \tag{20b}$$

satisfy the bound

$$|z(t, j)|_{\mathcal{A}^*} \leq c_1 |z(0, 0)|_{\mathcal{A}^*} \exp(-c_2(t + j)) + \delta, \quad \forall (t, j) \in dom(\vartheta), \tag{21}$$

The robustness result of Lemma 6 is not trivial. Indeed, for general hybrid dynamical systems, it is difficult to guarantee closeness of solutions between nominal and perturbed dynamics, even when the perturbations are arbitrarily small. On the other hand, when the HDS satisfies the regularity conditions of Lemma 5, stability properties turn out to be robust to small disturbances [34, Ch.7]. This idea is at the core of our modeling framework (16), which subsumes the closed-loop system under attack as a well-posed hybrid dynamical system with sufficiently slow control dynamics.

By leveraging the results of Propositions 1 and 2, we can now state the first *main* result of this paper, which establishes uniform global exponential input-to-state stability (E-ISS) for the closed-loop system under attacks and time-varying w. In particular, under the time ratio established in Proposition 1, and the maximal gain ε^* established in Proposition 2, we provide an explicit characterization of the ISS gain γ that maps $\|\dot{w}(t)\|_t$ to the radius of the residual set where the trajectories converge.

Theorem 1. *Let $\mathcal{A}^*$ be given by* (19). *Suppose that Assumptions 1-6 hold, and that conditions* (17) *and* (18) *are also satisfied for* $(\kappa_{u,1}, \kappa_{u,2}, \varepsilon)$. *Then, the HDS* (16) *renders E-ISS the set $\mathcal{A}^*$ with respect to $\dot{w}(t)$, with linear asymptotic gain γ given by*

$$\gamma\left(\|\dot{w}(t)\|_t\right) := \frac{1}{\varepsilon}\left[\left(\frac{\max\left\{\theta\bar{\lambda}(P),(1-\theta)e^{\tau_0}\max\left\{\ell^2/2\mu,1/2\right\}\right\}}{\min\left\{\theta\underline{\lambda}(P),(1-\theta)\min\{\mu/2,1/2\}\right\}}\right)^{1/2}\frac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t, \tag{22}$$

where $k \in (0,1)$, Ξ is a positive definite matrix, and $r := \left(2\theta\left\|PA^{-1}E\right\|,(1-\theta)e^{\tau_0}\max\left\{\ell_y\|H\|\|G\|,U_{\bar{w}}\right\}\right)$ with

$$\theta := \frac{\ell_y\|G\|\|C\|\max\left\{\ell,\bar{M}\right\}}{\ell_y\|G\|\|C\|\max\left\{\ell,\bar{M}\right\}+2\ell e^{-\tau_0}\left\|PA^{-1}B\right\|\max\left\{\bar{M},1\right\}}.$$

Remark 7. *The gain γ in* (22) *reflects explicitly the stability of the plant through the eigenvalues of the matrix P set in Assumption 2, and the qualitative behavior of the cost function given by the Lipschitz and PL constant ℓ and μ, respectively. Additionally, this gain is weighted by $1/\varepsilon$, which implies that a slower controller with respect to the dynamics of the plant leads to larger tracking errors. This behavior is expected since slowing down the controller makes it more difficult to track time-varying optimal points.*

Remark 8. *When $\ell >> 1$ and $\mu << 1$ (i.e., the cost function f is ill-conditioned) the asymptotic gain simplifies to*

$$\gamma\left(\|\dot{w}(t)\|_t\right) := \frac{1}{\varepsilon}\left[cond(f)e^{\tau_0/2}\frac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t, \qquad cond(f) = \frac{\ell}{\mu}, \tag{23}$$

which shows the effect that the condition number of the cost $f(\cdot)$ has on the residual set where the trajectories converge. Note that a cost function f that is ill-conditioned can be identified by examining the eccentricity of its sub-level sets.

Remark 9. *In* (22), *the constant τ_0 comes from the Lyapunov-based analysis used to study the HDS* (16). *This is an upper bound for the weighted time that the closed-loop system* (6) *can flow under attacks. A detailed characterization of these constants is presented in Section 5.*

4.2. Attacks to the Static Feedback

In the previous section, we focused on attacks acting on the dynamic controller. In this section, we now turn our attention to study the effect of persistent attacks acting on the static controller (10). To study this scenario, we consider the following set-valued HDS with state $\vartheta := (x,u,\tau_{v,1},\tau_{v,2},\sigma_v,w)$, which corresponds to the interconnection of the plant dynamics under attacks (10), the nominal dynamic gradient-based controller (9), and the signal generator (15):

$$\vartheta \in C := \mathbb{R}^n \times \mathbb{R}^m \times [0,N_{0,v}] \times [0,T_{0,v}] \times \Sigma_v \times \mathcal{W}, \tag{24a}$$

$$\begin{pmatrix} \dot{x} \\ \dot{u} \\ \dot{\tau}_{v,1} \\ \dot{\tau}_{v,2} \\ \dot{\sigma}_v \\ \dot{w} \end{pmatrix} \in F_{\sigma_v}(\vartheta) := \begin{pmatrix} A_{\sigma_v}x + B_{\sigma_v}u + E_{\sigma_v}w \\ -\varepsilon(\nabla f_u(u)+G^{\top}\nabla f_y(y)) \\ [0,\kappa_{v,1}] \\ [0,\kappa_{v,2}]-I_{\sigma_{v,a}}(\sigma_v) \\ 0 \\ \Pi(w) \end{pmatrix}, \quad y = Cx, \tag{24b}$$

$$\vartheta \in \mathcal{D} := \mathbb{R}^n \times \mathbb{R}^m \times [1, N_{0,v}] \times [0, T_{0,v}] \times \Sigma_v \times \mathcal{W}, \tag{24c}$$

$$\begin{pmatrix} x^+ \\ u^+ \\ \tau_{v,1}^+ \\ \tau_{v,2}^+ \\ \sigma_v^+ \\ w^+ \end{pmatrix} \in G_{\sigma_v}(\vartheta) := \begin{pmatrix} x \\ u \\ \tau_{v,1} - 1 \\ \tau_{v,2} \\ \Sigma_v \backslash \{\sigma_v\} \\ w \end{pmatrix}, \tag{24d}$$

where $T_{0,v} \geq 0$, $\kappa_{v,1} > 0$, $\kappa_{v,2} \in (0, 1)$ is a time-ratio parameter that describes an admissible attack, and $\Pi(w)$ is given by Assumption 1. Note that when $\sigma_v = s$, we have $A_s = A$, hence, Assumption 2 holds for A_s. Moreover, to simplify notation we denote $A_{a_i} := F + N L_{a_i} K C$ for any $a_i \in \Sigma_{v,a}$, thus we will refer to the attacks to the static controller also as attacks to the plant. By construction, system (24) also satisfies the properties of Lemma 5.

For attacks acting on the plant, we introduce an additional technical assumption to guarantee that the switching system has a well-defined unique equilibrium point.

Assumption 7. *For any $w \in \mathcal{W}$ there exists a unique $\bar{x} \in \mathbb{R}^n$ and $\bar{u} \in \mathbb{R}^m$ such that $0 = A_\sigma \bar{x} + B_\sigma \bar{u} + E_\sigma w$, $\forall \sigma \in \Sigma$.*

Remark 10. *The conditions of Assumption 7 assert that the attacks to the plant modify the stability of the equilibria, but they do not alter the set of equilibria [15]. We note that this assumption is aligned with existing works in context [15, 26], and it enables a tractable analysis via Lyapunov-based tools. Without this assumption in place, each attack could generate a different equilibrium point (if any), and to assert stability properties we will need to introduce additional restrictive assumptions on the frequency of the attacks, e.g., condition (13) with a "sufficiently small" constant $\kappa_{v,1}$. The study of switched dynamical systems with multiple equilibria (without attacks) is addressed in [52, 53] to mention a few, and extending our framework to this case is the subject of future research efforts.*

In contrast to the analysis of the previous section, we now first consider the plant under attacks, operating with a constant controller command u (i.e., $\dot{u} = 0$) and a fixed w (i.e., $\dot{w} = 0$). Thus, we study the stability properties of the equilibrium point $x^*(u, w) := A_{\sigma_v}^{-1} B_{\sigma_v} u + A_{\sigma_v}^{-1} E_{\sigma_v} w$. For this reduced system under attacks, we establish the following characterization of the parameters $(\kappa_{v,2}, \kappa_{v,1})$ to guarantee uniform global exponential stability.

Proposition 3. *Suppose that Assumptions 1-7 hold. Given a $\tilde{\beta} > 0$, and the dwell-time parameter $\kappa_{v,1}$, and the time-ratio parameter $\kappa_{v,2}$ simultaneously satisfying:*

$$\underline{\lambda}(R) > \bar{\lambda}(P) \ln(\omega)\kappa_{v,1} + \kappa_{v,2}\left(\underline{\lambda}(R) + \bar{\lambda}(P)\bar{\lambda}(\hat{R}_a)\right), \quad \text{with} \quad \omega := \max\left\{\frac{\bar{\lambda}(P)}{\tilde{\beta}}, \frac{\tilde{\beta}}{\underline{\lambda}(P)}\right\}. \tag{25}$$

where P, R satisfy Assumption 2, and $\bar{\lambda}(\hat{R}_a)$ is a constant derived after Lemma 3. Then, for each each compact set $K_u \subset \mathbb{R}^m$, the set

$$\mathcal{A}^*_{K_u} = \left\{(x, u, \tau_{v,1}, \tau_{v,2}, \sigma_v, w) : x = x^*(u, w), \ u \in K_u, \ \tau_{v,1} \in [0, N_{0,v}], \tau_{v,2} \in [0, T_{0,v}], \sigma_v \in \Sigma_v, w \in \mathcal{W}\right\},$$

is UGES for the HDS (24) with $\varepsilon = 0$, $\dot{w} = 0$, and restricted flow and jump sets given by

$$C_{K_u} = \mathbb{R}^n \times (\mathbb{R}^m \cap K_u) \times [0, N_{0,v}] \times [0, T_{0,v}] \times \Sigma_v \times \mathcal{W}, \ D_{K_u} = \mathbb{R}^n \times (\mathbb{R}^m \cap K_u) \times [1, N_{0,v}] \times [0, T_{0,v}] \times \Sigma_v \times \mathcal{W}.$$

The result of Proposition 3 establishes an explicit characterization of the admissible attack properties (c.f. Definition 5) needed to guarantee exponential stability in the system. For the internal feedback loop controller, these properties are related to the largest and smallest eigenvalues of the matrices P and $\hat{R}_a$.

By leveraging the result of Proposition 3, we can now state a stability and convergence result for the case where the controller is activated, and the state w remains static.

Proposition 4. *Suppose that Assumptions 1-7 hold, and let $\kappa_{v,1}$, $\kappa_{v,2}$ satisfy the bounds in Proposition 3. Let ε satisfy:*

$$0 < \varepsilon < \varepsilon^* := \frac{\rho \min\{\underline{\lambda}(P), \tilde{\beta}\}\mu^2}{2\ell_y e^{\tau_0} \|C\| \|G\| \|A^{-1}B\| \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\}(\mu^2 + \ell^2)}, \tag{26}$$

Attack action on:	Admissible Persistent Attack Condition	Controller Gain: ε^*	ISS Gain: $\gamma(\|\dot{w}\|_t)$
Dynamic Feedback	$2\mu > \ln(\omega)\kappa_{u,1} + 2\kappa_{u,2}(\mu + \bar{M}\ell),$ $\omega := \max\left\{\frac{1}{\mu}, \frac{\ell^2}{\mu}\right\}$	$\dfrac{\rho\underline{\lambda}(R)\min\{\mu,1\}}{2\ell_y\max\{\bar{M},1\}\,\|C\|\,\|G\|\,\|PA^{-1}B\|\,(\rho\min\{\mu,1\}+2\ell\max\{\ell,\bar{M}\}e^{\tau_0})}$	$\dfrac{1}{\varepsilon}\left[\left(\dfrac{\max\{\theta\overline{\lambda}(P),(1-\theta)e^{\tau_0}\max\{\ell^2/2\mu,1/2\}\}}{\min\{\theta\underline{\lambda}(P),(1-\theta)\min\{\mu/2,1/2\}\}}\right)^{1/2}\dfrac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t$
Static Feedback	$\underline{\lambda}(R) > \bar{\lambda}(P)\ln(\omega)\kappa_{v,1}$ $+\kappa_{v,2}\left(\underline{\lambda}(R)+\bar{\lambda}(P)\bar{\lambda}(\hat{R}_a)\right),$ $\omega := \max\left\{\dfrac{\bar{\lambda}(P)}{\tilde{\beta}}, \dfrac{\tilde{\beta}}{\underline{\lambda}(P)}\right\}$	$\dfrac{\rho\min\{\underline{\lambda}(P),\tilde{\beta}\}\mu^2}{2\ell_y e^{\tau_0}\,\|C\|\,\|G\|\,\|A^{-1}B\|\max\{\bar{\lambda}(P),\tilde{\beta}\}(\mu^2+\ell^2)}$	$\dfrac{1}{\varepsilon}\left[\left(\dfrac{\max\{\theta e^{\tau_0}\max\{\bar{\lambda}(P),\tilde{\beta}\},(1-\theta)\ell^2/(2\mu)\}}{\min\{\theta\min\{\underline{\lambda}(P),\tilde{\beta}\},(1-\theta)\mu/2\}}\right)^{1/2}\dfrac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t$

Figure 2: Summary of the results of Propositions 1-4, and Theorems 1-2.

where $\rho > 0$, $\tau_0 = \ln(\omega)N_{0,v} + T_{0,v}\left(\frac{\lambda(R)}{\overline{\lambda}(P)} + \bar{\lambda}(\hat{R}_a)\right)$, and P is defined in Assumption 2. Then, the set

$$\mathcal{A}^* = \left\{(x,u,\tau_{v,1},\tau_{v,2},\sigma_v,w) : x = x^*(u^*,w), u = u^*, \ \tau_{v,1} \in [0, N_{0,v}], \tau_{v,2} \in [0, T_{0,v}], \sigma_v \in \Sigma_v, w \in \mathcal{W}\right\},$$

is UGES for the HDS (24) with $\dot{w} = 0$.

Remark 11. *The result of Proposition 4 establishes a sufficient condition on the bounds of the gain of the controller, the dwell-time parameter $\kappa_{v,1}$, and the time-ratio parameter $\kappa_{v,2}$ to obtain exponential convergence to the solution of Problem (1). Since the HDS is well-posed, the robustness results of Lemma 6 also hold for system (24).*

By leveraging the results of Proposition 3 and Proposition 4, we now establish the second main result of this paper, which asserts exponential input-to-state stability of the closed-loop system under attacks, with respect to the time-variation of w. In conjunction with Theorem 1, this result provides a complete answer to Problem 1.

Theorem 2. *Suppose that Assumptions 1-7 hold, and that conditions (25) and (26) also hold. Then, the HDS (24) renders the set $\mathcal{A}^*$ E-ISS with respect to $\dot{w}(t)$, with linear asymptotic gain γ given by*

$$\gamma\left(\|\dot{w}(t)\|_t\right) := \frac{1}{\varepsilon}\left[\left(\frac{\max\left\{\theta e^{\tau_0}\max\left\{\bar{\lambda}(P),\tilde{\beta}\right\},(1-\theta)\ell^2/(2\mu)\right\}}{\min\left\{\theta\min\left\{\underline{\lambda}(P),\tilde{\beta}\right\},(1-\theta)\mu/2\right\}}\right)^{1/2}\frac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t, \tag{27}$$

where $k \in (0,1)$, Ξ is a positive definite matrix, and $r := \left(2e^{\tau_0}\theta\max\{\bar{\lambda}(P),\tilde{\beta}\}\left\|A^{-1}E\right\|, \ell_y(1-\theta)\|H\|\|G\|\right)$ with

$$\theta := \frac{\ell_y\|G\|\|C\|}{\ell_y\|G\|\|C\| + 2e^{\tau_0}\max\left\{\bar{\lambda}(P),\tilde{\beta}\right\}\|A^{-1}B\|}.$$

Remark 12. *As in Theorem 1, the asymptotic gain (27) relates the magnitude of $\dot{w}$ to the size of the residual set where the trajectories converge. Note that in this case the coefficient e^{τ_0} appears next to the the values and gains related to the plant. Therefore, when $\bar{\lambda}(P) >> 1$ and $\underline{\lambda}(P) << 1$, the asymptotic gain simplifies to*

$$\gamma\left(\|\dot{w}(t)\|_t\right) := \frac{1}{\varepsilon}\left[\sqrt{cond(P)e^{\tau_0}}\frac{\|r\|}{\underline{\lambda}(\Xi)k}\right]\|\dot{w}(t)\|_t, \qquad cond(P) := \frac{\bar{\lambda}(P)}{\underline{\lambda}(P)}, \tag{28}$$

which shows the effect of the condition number of the Lyapunov matrix P on the residual set where the trajectories converge.

We summarize in Table 2 the main results of this section. Each row indicates which subsystem is under attack, and the columns summarize the upper bounds on the time ratios and gains, and also the form of the ISS gain.

5. Analysis

In this section, we present the analysis and the proofs of Theorems 1 and 2. In both cases, our analysis leverages Lyapunov theory for switched systems with unstable subsystems. The structure of the proofs will follow four main steps:

1. For each of the feedback loops under attack, we will introduce an auxiliary Lyapunov-like function for which a strong exponential decrease can be asserted whenever the feedback loop operates in the stable mode, and the dynamics of the other loop are neglected.

2. When the feedback loop is under attacks, we incorporate the switching behavior into the system by extending the auxiliary Lyapunov-like function with a state that depends on the hybrid signal generator.

3. We use the extended auxiliary functions constructed in the previous step, to construct a new Lyapunov function for the complete hybrid system, and we characterize the upper bound ε^* on the gain of the controller needed to guarantee exponential decrease of the function during flows outside of the compact set of interest. We also show that the Lyapunov function does not increase during jumps.

4. The previous argument, in addition to the quadratic upper and lower bounds on the Lyapunov function, and the average dwell-time constraint, allows us to establish uniform global exponential stability and/or uniform global exponential ISS.

5.1. Analysis of System with Dynamic Feedback Controller under Attacks

Consider the closed-loop system with attacks on the gradient-flow controller modeled by the HDS (16). Define the affine map $(u, w) \mapsto \bar{x}(u, w)$ as $\bar{x}(u, w) := -A^{-1}Bu - A^{-1}Ew$, and consider the change of variables $x_e := x - \bar{x}$. In the new variables, the dynamics are given by:

$$\dot{x}_e = Ax_e + A^{-1}B\dot{u} + A^{-1}E\dot{w}, \qquad \dot{u} = -\varepsilon M_{\sigma_u}\varphi(u, w, x_e) = -\varepsilon M_{\sigma_u}\left(\nabla_u f_u(u) + G^\top \nabla_y f_y(Cx_e + Gu + Hw)\right),$$

$$G := -CA^{-1}B, \qquad H := -CA^{-1}E,$$

where we note that $-\varphi(u, w, 0) = -\nabla_u f(u, w) = -\nabla_u f_u(u) - G^\top \nabla_y f_y(Gu + Hw)$. First, we investigate the stability properties of the stable mode of the system by leveraging the stability properties of the gradient-flow. Consider the following auxiliary function:

$$V_s(u, w) = f(u, w) - f(u_w^*, w).$$

For $\vartheta \in \mathcal{D}$, during the jumps it holds that $V_s(u^+, w^+) = V_s(u, w)$. Moreover, when $\vartheta \in \mathcal{C}$, during flows we have:

$$\begin{aligned}
\dot{V}_s(u, w) &= \nabla_u f(u, w)^\top \dot{u} + (H\Pi(w))^\top \nabla_y f_y(Gu + Hw) - \nabla_u f(u_w^*, w)^\top \dot{u}^* - (H\Pi(w))^\top \nabla_y f_y(Gu_w^* + Hw) \\
&= \nabla_u f(u, w)^\top \dot{u} + (H\Pi(w))^\top \nabla_y f_y(Gu + Hw) - (H\Pi(w))^\top \nabla_y f_y(Gu_w^* + Hw) \\
&= -\varepsilon\varphi(u, w, 0)^\top (\varphi(u, w, 0) + \varphi(u, w, x_e) - \varphi(u, w, 0)) + \ell_y \|H\| \|G\| \|u - u_w^*\| \|\Pi(w)\| \\
&\leq -\varepsilon \|\nabla_u f(u, w)\|^2 + \varepsilon\ell_y \|C\| \|G\| \|\nabla_u f(u, w)\| \|x_e\| + \ell_y \|H\| \|G\| \|u - u_w^*\| \|\Pi(w)\| \\
&\leq -2\mu\varepsilon V_s(u, w) + \varepsilon\ell_y \|C\| \|G\| \|\nabla_u f(u, w)\| \|x_e\| + \ell_y \|H\| \|G\| \|u - u_w^*\| \|\Pi(w)\|.
\end{aligned}$$

Next, to study the stability properties of the controller under switching signals with unstable modes, we consider the following auxiliary function:

$$V_a(u, w) = \frac{1}{2} \|u - u_w^*\|^2.$$

When $\vartheta \in \mathcal{D}$, during jumps it holds that $V_a(u^+, w^+) = V_a(u, w)$. For $\vartheta \in \mathcal{C}$, during flows we have

$$\begin{aligned}
\dot{V}_a(u, w) &= (u - u_w^*)^\top (\dot{u} - \dot{u}_w^*) \\
&= -(u - u_w^*)^\top \left(\varepsilon M_{\sigma_u}\varphi(u, w, x_e)\right) - (u - u_w^*)^\top (\nabla_w h\Pi(w))
\end{aligned}$$

$$\leq \varepsilon \left\| u - u_w^* \right\| \left\| M_{\sigma_u} \right\| \left\| \varphi(u, w, x_e) \right\| + \left\| u - u_w^* \right\| \left\| \nabla_w h \right\| \left\| \Pi(w) \right\|$$

$$\leq \varepsilon \left\| u - u_w^* \right\| \left\| M_{\sigma_u} \right\| \left\| \varphi(u, w, 0) - \varphi(u, w, 0) + \varphi(u, w, x_e) \right\| + \left\| u - u_w^* \right\| \left\| \nabla_w h \right\| \left\| \Pi(w) \right\|$$

$$\leq \varepsilon \left\| u - u_w^* \right\| \left\| M_{\sigma_u} \right\| \left(\ell \left\| u - u_w^* \right\| + \ell_y \left\| C \right\| \left\| G \right\| \left\| x_e \right\| \right) + \left\| u - u_w^* \right\| \left\| \nabla_w h \right\| \left\| \Pi(w) \right\|$$

$$\leq \varepsilon \ell \bar{M} \left\| u - u_w^* \right\|^2 + \varepsilon \bar{M} \ell_y \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| + \left\| u - u_w^* \right\| \left\| \nabla_w h \right\| \left\| \Pi(w) \right\|$$

$$= 2\varepsilon \ell \bar{M} V_a(u, w) + \varepsilon \bar{M} \ell_y \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| + U_{\bar{w}} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|.$$

Next, we consider the Lyapunov function:

$$\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) := V_{\sigma_u}(u, w)e^{\tau}, \quad \sigma_u \in \{s, a\}, \quad \tau := \ln(\omega)\tau_{u,1} + \tau_{u,2}(\rho_s + \rho_a), \quad \rho_s := 2\mu, \quad \rho_a := 2\bar{M}\ell, \text{ (29)}$$

where we recall that μ and ℓ are the PL and Lipschitz constants, respectively, $\bar{M}$ is given by Assumption 6, and $\omega = \max\{1/\mu, \ell^2/\mu\}$. Now, during jumps the function $\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u)$ satisfies

$$\bar{V}(u^+, w^+, \tau_{u,1}^+, \tau_{u,2}^+, \sigma_u^+) = V_{\sigma_u^+} e^{\ln(\omega)(\tau_{u,1} - 1) + \tau_{u,2}(\rho_s + \rho_a)} \leq \max_{\sigma_u^+ \in \{s, a\}} V_{\sigma_u^+} e^{-\ln(\omega) + \tau} \leq \omega V_{\sigma_u} e^{-\ln(\omega)} e^{\tau} = \bar{V},$$

hence, we conclude that $\bar{V}$ does not increase during jumps. Let $\tau_0 := \ln(\omega)N_{0,u} + T_{0,u}(\rho_s + \rho_a)$. To analyze the flows, we note that for $\bar{V} = V_s e^{\tau}$ the time derivative satisfies:

$$\dot{\bar{V}}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) = \dot{V}_s(u, w)e^{\tau} + V_s(u, w)e^{\tau}\dot{\tau}$$

$$= \dot{V}_s(u, w)e^{\tau} + V_s(u, w)e^{\tau}(\ln(\omega)\dot{\tau}_{u,1} + (\rho_s + \rho_a)\dot{\tau}_{u,2})$$

$$\leq \dot{V}_s(u, w)e^{\tau} + \varepsilon V(u, w)_1 e^{\tau}(\ln(\omega)\kappa_{u,1} + (\rho_s + \rho_a)\kappa_{u,2})$$

$$\leq -2\mu\varepsilon V_s(u, w)e^{\tau} + \varepsilon\ell_y e^{\tau} \left\| C \right\| \left\| G \right\| \left\| \nabla_u f(u, w) \right\| \left\| x_e \right\|$$
$$\quad + \ell_y e^{\tau} \left\| H \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| \Pi(w) \right\| + \varepsilon V_s(u, w)e^{\tau}(\ln(\omega)\kappa_{u,1} + (\rho_s + \rho_a)\kappa_{u,2})$$

$$= -\varepsilon\rho\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \varepsilon\ell_y e^{\tau} \left\| C \right\| \left\| G \right\| \left\| \nabla_u f(u, w) \right\| \left\| x_e \right\| + \ell_y e^{\tau} \left\| H \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|$$

$$\leq -\varepsilon\rho\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \varepsilon\ell\ell_y e^{\tau_0} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| + \ell_y e^{\tau_0} \left\| H \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|$$

$$\leq -\varepsilon\rho\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \varepsilon\ell\ell_y e^{\tau_0} \max\left\{\ell, \bar{M}\right\} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\|$$
$$\quad + e^{\tau_0} \max\left\{\ell_y \left\| H \right\| \left\| G \right\|, U_{\bar{w}}\right\} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|,$$

where $\rho := \rho_s - \kappa_{u,2}(\rho_s + \rho_a) - \ln(\omega)\kappa_{u,1} > 0$, for $\omega \geq 1$. Similarly, when $\bar{V} = V_a e^{\tau}$ we have that:

$$\dot{\bar{V}}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) = \dot{V}_a(u, w)e^{\tau} + V_a(u, w)e^{\tau}\dot{\tau}$$

$$= \dot{V}_a(u, w)e^{\tau} + V_a(u, w)e^{\tau}(\ln(\omega)\dot{\tau}_{u,1} + (\rho_s + \rho_a)\dot{\tau}_{u,2})$$

$$\leq \dot{V}_a(u, w)e^{\tau} + \varepsilon V_a(u, w)e^{\tau}(\ln(\omega)\kappa_{u,1} + (\rho_s + \rho_a)(\kappa_{u,2} - 1))$$

$$\leq 2\varepsilon\ell\bar{M}V_a(u, w)e^{\tau} + \varepsilon\bar{M}\ell_y e^{\tau} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| + U_{\bar{w}} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\| e^{\tau}$$
$$\quad + \varepsilon V_a(u, w)e^{\tau}(\ln(\omega)\kappa_{u,1} + (\rho_s + \rho_a)(\kappa_{u,2} - 1))$$

$$= -\varepsilon\rho\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \varepsilon\bar{M}\ell_y e^{\tau_0} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| + U_{\bar{w}} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\| e^{\tau_0}$$

$$\leq -\varepsilon\rho\bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \varepsilon\ell\ell_y e^{\tau_0} \max\left\{\ell, \bar{M}\right\} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\|$$
$$\quad + e^{\tau_0} \max\left\{\ell_y \left\| H \right\| \left\| G \right\|, U_{\bar{w}}\right\} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|,$$

Therefore, for any mode we can upper-bound $\dot{\bar{V}}$ as follows:

$$\dot{\bar{V}}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) \leq -\varepsilon\frac{\rho}{2}\min\{\mu, 1\} \left\| u - u_w^* \right\|^2 + \varepsilon\ell_y e^{\tau_0} \max\left\{\ell, \bar{M}\right\} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\|$$
$$\quad + e^{\tau_0} \max\left\{\ell_y \left\| H \right\| \left\| G \right\|, U_{\bar{w}}\right\} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\|.$$

We now analyze the interconnection of the plant dynamics and the switched controller. To do this, we first consider the auxiliary function $W(x_e) = x_e^\top P x_e$, where $P \succ 0$ which satisfies $A^\top P + PA \preceq -R$ with $R \succ 0$ due to Assumption

2. When $\vartheta \in \mathcal{D}$, during jumps we have $W(x_e^+) = W(x_e)$. Moreover, when $\vartheta \in \mathcal{C}$, the time derivative of W along the trajectories of the plant dynamics satisfies the following inequalities:

$$
\begin{aligned}
\dot{W}(x_e) &= \dot{x}_e^\top P x_e + x_e^\top P \dot{x}_e \\
&= 2 x_e^\top P (A x_e + A^{-1} B \dot{u} + A^{-1} E \Pi(w)) \\
&= x_e^\top (A^\top P + PA) x_e + 2 x_e^\top \left(P A^{-1} B \dot{u} \right) + 2 x_e^\top \left(P A^{-1} E \Pi(w) \right) \\
&\leq -x_e^\top R x_e - 2\varepsilon x_e^\top \left(P A^{-1} B M_{\sigma_u} \varphi(u, w, x_e) \right) + 2 x_e^\top \left(P A^{-1} E \Pi(w) \right) \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| \varphi(u, w, x_e) \right\| \left\| x_e \right\| + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| x_e \right\| \left\| \varphi(u, w, 0) + \varphi(u, w, x_e) - \varphi(u, w, 0) \right\| \\
&\quad + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| x_e \right\| \left(\left\| \varphi(u, w, 0) \right\| + \left\| \varphi(u, w, x_e) - \varphi(u, w, 0) \right\| \right) \\
&\quad + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| x_e \right\| \left(\left\| \nabla_u f(u, w) \right\| + \ell_y \left\| C \right\| \left\| G \right\| \left\| x_e \right\| \right) + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&= -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \ell_y \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| C \right\| \left\| G \right\| \left\| x_e \right\|^2 + 2\varepsilon \left\| P A^{-1} B \right\| \left\| M_{\sigma_u} \right\| \left\| x_e \right\| \left\| \nabla_u f(u, w) \right\| \\
&\quad + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \ell_y \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| C \right\| \left\| G \right\| \left\| x_e \right\|^2 + 2\varepsilon \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| x_e \right\| \left\| \nabla_u f(u, w) \right\| \\
&\quad + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| . \\
&\leq -\underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \ell_y \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| C \right\| \left\| G \right\| \left\| x_e \right\|^2 + 2\varepsilon \ell \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| x_e \right\| \left\| u - u_w^* \right\| \\
&\quad + 2 \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| .
\end{aligned}
$$

Next, we consider the following Lyapunov function for the complete HDS (16):

$$
U(\vartheta) = (1 - \theta) \bar{V}(u, w, \tau_{u,1}, \tau_{u,2}, \sigma_u) + \theta W(x_e), \quad \theta \in (0, 1). \tag{30}
$$

This function is bounded from below and above as follows:

$$
\min \left\{ \theta \underline{\lambda}(P), (1 - \theta) \frac{1}{2} \min\{\mu, 1\} \right\} |\vartheta|_{\mathcal{A}^*}^2 \leq U(\vartheta) \leq \max \left\{ \theta \bar{\lambda}(P), (1 - \theta) \frac{1}{2} \max \left\{ \ell^2/\mu, 1 \right\} e^{\tau_0} \right\} |\vartheta|_{\mathcal{A}^*}^2, \tag{31}
$$

$\forall\, \vartheta \in \mathcal{C} \cup \mathcal{D}$, where we used the fact that $\tau_{u,1} \in [0, N_{0,u}]$, $\tau_{u,2} \in [0, T_{0,u}]$, $\sigma_u \in \Sigma_u$, and $w \in \mathcal{W}$ at all times. Moreover, for any $\vartheta \in \mathcal{D}$, it follows that $U(\vartheta^+) \leq (1 - \theta) \bar{V} + \theta W = U(\vartheta)$, hence, during jumps, the Lyapunov function does not increase. For any $\vartheta \in \mathcal{C}$, the time derivative of $\dot{U}$ can be upper bounded as follows:

$$
\begin{aligned}
\dot{U}(\vartheta) &= -\theta \underline{\lambda}(R) \left\| x_e \right\|^2 + 2\varepsilon \theta \ell_y \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| C \right\| \left\| G \right\| \left\| x_e \right\|^2 \\
&\quad + 2\varepsilon \theta \ell \max \left\{ \bar{M}, 1 \right\} \left\| P A^{-1} B \right\| \left\| x_e \right\| \left\| u - u_w^* \right\| + 2\theta \left\| P A^{-1} E \right\| \left\| x_e \right\| \left\| \Pi(w) \right\| \\
&\quad - \varepsilon(1 - \theta) \frac{\rho}{2} \min\{\mu, 1\} \left\| u - u_w^* \right\|^2 + \varepsilon \ell_y e^{\tau_0} (1 - \theta) \max \left\{ \ell, \bar{M} \right\} \left\| C \right\| \left\| G \right\| \left\| u - u_w^* \right\| \left\| x_e \right\| \\
&\quad + e^{\tau_0} (1 - \theta) \max \left\{ \ell_y \left\| H \right\| \left\| G \right\|, U_{\bar{w}} \right\} \left\| u - u_w^* \right\| \left\| \Pi(w) \right\| .
\end{aligned}
$$

Let $\xi := \left(\left\| x_e \right\|, \left\| u - u_w^* \right\| \right)$; using these definitions, we obtain:

$$
\dot{U}(\vartheta) \leq -\varepsilon \xi^\top \Xi \xi + r^\top \xi \left\| \Pi(w) \right\|, \tag{32}
$$

where $r := \left(2\theta \left\| P A^{-1} E \right\|, e^{\tau_0} (1 - \theta) \max \left\{ \ell_y \left\| H \right\| \left\| G \right\|, U_{\bar{w}} \right\} \right)$, and Ξ is a symmetric matrix of the form

$$
\Xi = \begin{bmatrix} \theta \left(\frac{\alpha}{\varepsilon} - \beta \right) & -\frac{1}{2} \left((1 - \theta) \delta + \theta \chi \right) \\ -\frac{1}{2} \left((1 - \theta) \delta + \theta \chi \right) & (1 - \theta) \gamma \end{bmatrix}. \tag{33}
$$

with $\alpha := \underline{\lambda}(R)$, $\beta := 2\ell_y \max\{\bar{M}, 1\} \left\| PA^{-1}B \right\| \|C\| \|G\|$, $\delta := \ell_y e^{\tau_0} \max\{\ell, \bar{M}\} \|G\| \|C\|$,

$\chi := 2\ell \max\{\bar{M}, 1\} \left\| PA^{-1}B \right\|$, $\gamma := \frac{\rho}{2} \min\{\mu, 1\}$, $\theta := \frac{\delta}{\delta + \chi}$, and with ε satisfying (18), hence, Ξ is positive definite according to Lemma 7.

The last step is to show that the quadratic term on the right-hand-side of (32) dominates the linear term whenever ξ is sufficiently large. It follows that for $k \in (0, 1)$, (32) can be further upper bounded as follows:

$$\dot{U}(\vartheta) \le -\varepsilon(1-k)\underline{\lambda}(\Xi) \|\xi\|^2 - \varepsilon k\underline{\lambda}(\Xi) \|\xi\|^2 + r^\top \xi \|\Pi(w)\|_t$$

$$\le -\varepsilon(1-k)\underline{\lambda}(\Xi) \|\xi\|^2 - \varepsilon k\underline{\lambda}(\Xi) \|\xi\|^2 + \|r\| \|\xi\| \|\Pi(w)\|_t$$

$$\le -\varepsilon(1-k)\underline{\lambda}(\Xi) \|\xi\|^2 = -\varepsilon(1-k)\underline{\lambda}(\Xi)|\vartheta|_{\mathcal{A}*}^2, \quad \forall |\vartheta|_{\mathcal{A}*} \ge \frac{\|r\|}{\varepsilon \underline{\lambda}(\Xi)k} \|\Pi(w)\|_t. \tag{34}$$

By applying [35, Thm. 3.1], exploiting the quadratic bounds in (31), and the average dwell-time condition, the result holds. ∎

5.2. Analysis of System with Static Feedback Controller under Attacks

In this section, we study the stability properties of system (24), which models the interconnected system with attacks acting on the inner control loop. We introduce the affine mapping $(u, w) \mapsto \bar{x}(u, w)$ given by $\bar{x}(u, w) = -A_{\sigma_v}^{-1} B_{\sigma_v} u - A_{\sigma_v}^{-1} E_{\sigma_v} w$, which, by Assumption 7, is independent of $\sigma_v \in \Sigma_v$. Using this mapping, we consider the change of variable $x_e := x - \bar{x}$, which leads to the following dynamics:

$$\dot{x}_e = A_{\sigma_v} x_e + A_{\sigma_v}^{-1} B_{\sigma_v} \dot{u} + A_{\sigma_v}^{-1} E_{\sigma_v} \dot{w}, \qquad \dot{u} = -\varepsilon\varphi(u, w, x_e) := -\varepsilon(\nabla f_u(u) + G_{\sigma_v}^\top \nabla f_y(Cx_e + G_{\sigma_v} u + H_{\sigma_v} w)),$$

$$G_{\sigma_v} := -CA_{\sigma_v}^{-1} B_{\sigma_v}, \qquad H_{\sigma_v} := -CA_{\sigma_v}^{-1} E_{\sigma_v},$$

where $-\varphi(u, w, 0) = -\nabla_u f(u, w) = -\nabla f_u(u) - G_s^\top \nabla f_y(G_{\sigma_v} u + H_{\sigma_v} w)$.

We first consider the stable mode, i.e., $\sigma_v = s$. For this mode, we consider the auxiliary function $W_s := x_e^\top P x_e$, where $P \succ 0$ satisfies the algebraic condition $A_s^\top P + PA_s = -R$ for a given matrix $R \succ 0$. When $\vartheta \in \mathcal{D}$, during the jumps we have that $W_s(x_e^+) = W_s(x_e)$. On the other hand, when $\vartheta \in \mathcal{C}$, the time derivative of W_s satisfies:

$$\dot{W}_s(x_e) = \dot{x}_e^\top P x_e + x_e^\top P \dot{x}_e$$

$$= \left(A_s x_e + A_s^{-1} B_s \dot{u} + A_s^{-1} E_s \Pi(w)\right)^\top P x_e + x_e^\top P \left(A_s x_e + A_s^{-1} B_s \dot{u} + A_s^{-1} E_s \Pi(w)\right)$$

$$= x_e^\top (A_s^\top P + PA_s)x_e + 2x_e^\top P(A_s^{-1} B_s \dot{u}) + 2x_e^\top P(A_s^{-1} E_s \Pi(w))$$

$$= -x_e^\top R x_e - 2\varepsilon x_e^\top P A_s^{-1} B_s(\varphi(u, w, x_e) + \varphi(u, w, 0) - \varphi(u, w, 0)) + 2x_e^\top P(A_s^{-1} E_s \Pi(w))$$

$$\le -x_e^\top R x_e + 2\varepsilon \|x_e\| \left\| PA_s^{-1} B_s \right\| \|\varphi(u, w, 0) + (\varphi(u, w, x_e) - \varphi(u, w, 0))\| + 2x_e^\top P(A_s^{-1} E_s \Pi(w))$$

$$\le -x_e^\top R x_e + 2\varepsilon \|x_e\| \left\| PA_s^{-1} B_s \right\| \left(\|\varphi(u, w, 0)\| + \|\varphi(u, w, x_e) - \varphi(u, w, 0)\|\right) + 2x_e^\top P(A_s^{-1} E_s \Pi(w))$$

$$\le -\underline{\lambda}(R) \|x_e\|^2 + 2\varepsilon \left\| PA_s^{-1} B_s \right\| \|x_e\| \|\nabla_u f(u, w)\| + 2\varepsilon\ell_y \left\| PA_s^{-1} B_s \right\| \|C\| \|G_s\| \|x_e\|^2$$

$$\quad + 2 \|x_e\| \left\| PA_s^{-1} E_s \right\| \|\Pi(w)\|$$

$$\le -\rho_s W_s(x_e) + 2\varepsilon \left\| PA_s^{-1} B_s \right\| \|x_e\| \|\nabla_u f(u, w)\| + 2\varepsilon\ell_y \left\| PA_s^{-1} B_s \right\| \|C\| \|G_s\| \|x_e\|^2$$

$$\quad + 2 \|x_e\| \left\| PA_s^{-1} E_s \right\| \|\Pi(w)\|,$$

where we defined $\rho_s := \frac{\underline{\lambda}(R)}{\bar{\lambda}(P)}$.

Next, for the unstable modes, let $\sigma_v = a_i \in \Sigma_{v,a}$. Let $\tilde{\beta} > 0$, and consider the auxiliary function $W_a = \tilde{\beta} x_e^\top x_e$. Recall that $\tilde{\beta}(A_{a_i}^\top + A_{a_i}) = \tilde{\beta}\hat{R}_{a_i}$ as stated in Lemma 3. When $\vartheta \in \mathcal{D}$, during jumps we have that $W_a(x_e^+) = W_a(x_e)$. Moreover, when $\vartheta \in \mathcal{C}$, during flows we have:

$$\dot{W}_a(x_e) = \tilde{\beta}(\dot{x}_e^\top x_e + x_e^\top \dot{x}_e)$$

$$
\begin{aligned}
&= \tilde{\beta}\left(A_{a_i}x_e + A_{a_i}^{-1}B_{a_i}\dot{u} + A_{a_i}^{-1}E_{a_i}\Pi(w)\right)^\top x_e + \tilde{\beta}x_e^\top\left(A_{a_i}x_e + A_{a_i}^{-1}B_{a_i}\dot{u} + A_{a_i}^{-1}E_{a_i}\Pi(w)\right) \\
&= \tilde{\beta}x_e^\top(A_{a_i}^\top + A_{a_i})x_e + 2\tilde{\beta}x_e^\top(A_{a_i}^{-1}B_{a_i}\dot{u}) + 2\tilde{\beta}x_e^\top(A_{a_i}^{-1}E_{a_i}\Pi(w)) \\
&= \tilde{\beta}x_e^\top\hat{R}_{a_i}x_e - 2\varepsilon\tilde{\beta}x_e^\top A_{a_i}^{-1}B_{a_i}(\varphi(u,w,x_e) + \varphi(u,w,0) - \varphi(u,w,0)) + 2\tilde{\beta}x_e^\top(A_{a_i}^{-1}E_{a_i}\Pi(w)) \\
&\leq \tilde{\beta}x_e^\top\hat{R}_{a_i}x_e + 2\varepsilon\tilde{\beta}\left\|x_e\right\|\left\|A_{a_i}^{-1}B_{a_i}\right\|\left\|\varphi(u,w,0) + (\varphi(u,w,x_e) - \varphi(u,0))\right\| + 2\tilde{\beta}x_e^\top(A_{a_i}^{-1}E_{a_i}\Pi(w)) \\
&\leq \tilde{\beta}x_e^\top\hat{R}_{a_i}x_e + 2\varepsilon\tilde{\beta}\left\|x_e\right\|\left\|A_{a_i}^{-1}B_{a_i}\right\|\left(\left\|\varphi(u,w,0)\right\| + \left\|\varphi(u,w,x_e) - \varphi(u,w,0)\right\|\right) + 2\tilde{\beta}x_e^\top(A_{a_i}^{-1}E_s\Pi(w)) \\
&\leq \tilde{\beta}\bar{\lambda}(\hat{R}_{a_i})\left\|x_e\right\|^2 + 2\varepsilon\tilde{\beta}\left\|A_{a_i}^{-1}B_{a_i}\right\|\left\|x_e\right\|\left\|\nabla_u f(u,w)\right\| + 2\varepsilon\ell_y\tilde{\beta}\left\|A_{a_i}^{-1}B_{a_i}\right\|\left\|C\right\|\left\|G_{a_i}\right\|\left\|x_e\right\|^2 \\
&\quad + 2\tilde{\beta}\left\|x_e\right\|\left\|A_{a_i}^{-1}E_{a_i}\right\|\left\|\Pi(w)\right\| \\
&\leq \rho_{a_i}W_a(x_e) + 2\varepsilon\tilde{\beta}\left\|A_{a_i}^{-1}B_{a_i}\right\|\left\|x_e\right\|\left\|\nabla_u f(u,w)\right\| + 2\varepsilon\ell_y\tilde{\beta}\left\|A_{a_i}^{-1}B_{a_i}\right\|\left\|C\right\|\left\|G_{a_i}\right\|\left\|x_e\right\|^2 \\
&\quad + 2\tilde{\beta}\left\|x_e\right\|\left\|A_{a_i}^{-1}E_{a_i}\right\|\left\|\Pi(w)\right\|,
\end{aligned}
\tag{35}
$$

where $\rho_{a_i} := \bar{\lambda}(\hat{R}_{a_i})$ for a fixed a_i. We analyze the stability properties of the switched plant using the following auxiliary function:

$$
\begin{aligned}
&\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) = W_{\sigma_v}(x_e)e^\tau, \quad \sigma_v \in \{s, a\}, \quad \tau = \ln(\omega)\tau_{v,1} + \tau_{v,2}(\rho_s + \rho_a), \\
&\text{where } 0 \leq \tau \leq \tau_0 := \ln(\omega)N_{0,v} + T_{0,v}(\rho_s + \rho_a) \text{ and } \rho_s := \frac{\underline{\lambda}(R)}{\bar{\lambda}(P)} \qquad \rho_a := \bar{\lambda}(\hat{R}_a).
\end{aligned}
\tag{36}
$$

By Assumption 7, $G_s = G_{a_i}$, $\left\|A_s^{-1}B_s\right\| = \left\|A_{a_i}^{-1}B_{a_i}\right\|$, and $\left\|A_s^{-1}E_s\right\| = \left\|A_{a_i}^{-1}E_{a_i}\right\| \;\forall a_i \in \Sigma_{v,a}$, hence, we drop the subindices s and a_i, and we recall that $\bar{\lambda}(\hat{R}_a)$ was defined after Lemma 3. We set $\omega = \max\left\{\frac{\bar{\lambda}(P)}{\tilde{\beta}}, \frac{\tilde{\beta}}{\underline{\lambda}(P)}\right\}$.

The function $\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) = W_{\sigma_v}e^\tau$ satisfies during jumps:

$$
\bar{W}(x_e^+, \tau_{v,1}^+, \tau_{v,2}^+, \sigma_v^+) = W_{\sigma_v^+}e^{\ln(\omega)(\tau_{v,1}-1) + \tau_{v,2}(\rho_s + \rho_a)} \leq \max_{\sigma_v^+ \in \Sigma_v} W_{\sigma_v^+}e^{-\ln(\omega)+\tau} \leq \omega W_{\sigma_v}e^{-\ln(\omega)}e^\tau = \bar{W}
\tag{37}
$$

Similarly, during flows we have that:

$$
\begin{aligned}
\dot{\bar{W}}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) &= \dot{W}_s(x_e)e^\tau + W_s(x_e)e^\tau\dot{\tau} \\
&= \dot{W}_s(x_e)e^\tau + W_s(x_e)e^\tau(\ln(\omega)\dot{\tau}_{v,1} + (\rho_s + \rho_a)\dot{\tau}_{v,2}) \\
&\leq \dot{W}_s(x_e)e^\tau + W_s(x_e)e^\tau(\ln(\omega)\kappa_{v,1} + (\rho_s + \rho_a)\kappa_{v,2}) \\
&\leq -\rho_s W_s(x_e)e^\tau + 2\varepsilon\left\|PA^{-1}B\right\|\left\|x_e\right\|\left\|\nabla_u f(u,w)\right\|e^\tau + 2\varepsilon\ell_y e^\tau\left\|PA^{-1}B\right\|\left\|C\right\|\left\|G\right\|\left\|x_e\right\|^2 \\
&\quad + 2e^\tau\left\|x_e\right\|\left\|PA^{-1}E\right\|\left\|\Pi(w)\right\| + W_s(x_e)e^\tau(\ln(\omega)\kappa_{v,1} + (\rho_s + \rho_a)\kappa_{v,2}) \\
&\leq -\rho\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) + 2\varepsilon\left\|PA^{-1}B\right\|\left\|x_e\right\|\left\|\nabla_u f(u,w)\right\|e^\tau \\
&\quad + 2\varepsilon\ell_y e^\tau\left\|PA^{-1}B\right\|\left\|C\right\|\left\|G\right\|\left\|x_e\right\|^2 + 2e^\tau\left\|x_e\right\|\left\|PA^{-1}E\right\|\left\|\Pi(w)\right\| \\
&\leq -\rho\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) + 2\varepsilon\ell\max\left\{\bar{\lambda}(P), \tilde{\beta}\right\}\left\|A^{-1}B\right\|\left\|x_e\right\|\left\|u - u_w^*\right\|e^{\tau_0} \\
&\quad + 2\varepsilon\ell_y e^{\tau_0}\max\left\{\bar{\lambda}(P), \tilde{\beta}\right\}\left\|A^{-1}B\right\|\left\|C\right\|\left\|G\right\|\left\|x_e\right\|^2 \\
&\quad + 2e^{\tau_0}\max\left\{\bar{\lambda}(P), \tilde{\beta}\right\}\left\|A^{-1}E\right\|\left\|x_e\right\|\left\|\Pi(w)\right\|,
\end{aligned}
$$

where $\rho := \rho_s - \kappa_{v,2}(\rho_s + \rho_a) - \ln(\omega)\kappa_{v,1} > 0$, with $\omega \geq 1$. Similarly, using (35), we can obtain the following bound during unstable modes for the time-derivative of $\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) = W_a e^\tau$:

$$
\dot{\bar{W}}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) = \dot{W}_a(x_e)e^\tau + W_a(x_e)e^\tau\dot{\tau}
$$

$$\begin{aligned}
&= \dot{W}_a(x_e)e^\tau + W_a(x_e)e^\tau(\ln(\omega)\dot{\tau}_{v,1} + (\rho_s + \rho_a)\dot{\tau}_{v,2}) \\
&\leq \dot{W}_a(x_e)e^\tau + W_a(x_e)e^\tau(\ln(\omega)\kappa_{v,1} + (\rho_s + \rho_a)(\kappa_{v,2} - 1)) \\
&\leq \rho_a W_a(x_e)e^\tau + 2\varepsilon\tilde{\beta}e^\tau \left\|A^{-1}B\right\| \|x_e\| \|\nabla_u f(u,w)\| + 2\varepsilon\ell_y\tilde{\beta}e^\tau \left\|A^{-1}B\right\| \|C\| \|G\| \|x_e\|^2 \\
&\quad + 2\tilde{\beta}e^\tau \|x_e\| \left\|A^{-1}E\right\| \|\Pi(w)\| + W_a(x_e)e^\tau(\ln(\omega)\kappa_{v,1} + (\rho_s + \rho_a)(\kappa_{v,2} - 1)) \\
&\leq -\rho\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) + 2\varepsilon\tilde{\beta}e^{\tau_0} \left\|A^{-1}B\right\| \|x_e\| \|\nabla_u f(u,w)\| \\
&\quad + 2\varepsilon\ell_y\tilde{\beta}e^{\tau_0} \left\|A^{-1}B\right\| \|C\| \left\|G_{a_i}\right\| \|x_e\|^2 + 2\tilde{\beta}e^{\tau_0} \|x_e\| \left\|A_{a_i}^{-1}E_{a_i}\right\| \|\Pi(w)\| \\
&\leq -\rho\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) + 2\varepsilon\ell \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\| \|x_e\| \|u - u_w^*\| e^{\tau_0} \\
&\quad + 2\varepsilon\ell_y e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\| \|C\| \|G\| \|x_e\|^2 \\
&\quad + 2e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}E\right\| \|x_e\| \|\Pi(w)\|.
\end{aligned}$$

To analyze the nominal controller (5) we considered the auxiliary function $V(u,w) = f(u,w) - f(u_w^*, w)$, then we get the following upper bound:

$$\begin{aligned}
\dot{V}(u,w) &= -\varepsilon\nabla_u f(u,w)^\top \varphi(x_e, u) + H^\top(\nabla f_y(Gu + Hw) - \nabla f_y(Gu^* + Hw))\Pi(w) \\
&= -\varepsilon\varphi(0,u)^\top(\varphi(0,u) + \varphi(x_e, u) - \varphi(0,u)) + \ell_y \|H\| \|G\| \|u - u^*\| \|\Pi(w)\| \\
&\leq -\varepsilon \|\nabla_u f(u,w)\|^2 + \varepsilon \|\nabla_u f(u,w)\| \|G\| \left\|\nabla f_y(Cx_e + Gu + Hw) - \nabla f_y(Gu + Hw)\right\| \\
&\quad + \ell_y \|H\| \|G\| \|u - u^*\| \|\Pi(w)\| \\
&\leq -\varepsilon \|\nabla_u f(u,w)\|^2 + \varepsilon\ell_y \|G\| \|C\| \|\nabla_u f(u,w)\| \|x_e\| + \ell_y \|H\| \|G\| \|u - u^*\| \|\Pi(w)\| \\
&\leq -\varepsilon \|\nabla_u f(u,w)\|^2 + \varepsilon\ell_y \|G\| \|C\| \|\nabla_u f(u,w)\| \|x_e\| + \ell_y \|H\| \|G\| \|u - u^*\| \|\Pi(w)\| \\
&\leq -\varepsilon\mu^2 \|u - u_w^*\|^2 + \varepsilon\ell\ell_y \|G\| \|C\| \|u - u_w^*\| \|x_e\| + \ell_y \|H\| \|G\| \|u - u^*\| \|\Pi(w)\|.
\end{aligned}$$

Finally, for the HDS (24), we consider the Lyapunov function $U(\vartheta) = \theta\bar{W}(x_e, \tau_{v,1}, \tau_{v,2}, \sigma_v) + (1 - \theta)V(u,w)$, with $\theta \in (0,1)$. This function is bounded as follows:

$$\min\{\theta\underline{\lambda}(P), \theta\tilde{\beta}, (1-\theta)\mu/2\}|\vartheta|_{\mathcal{A}^*}^2 \leq U(\vartheta) \leq \max\left\{\theta\bar{\lambda}(P)e^{\tau_0}, \theta\tilde{\beta}e^{\tau_0}, (1-\theta)\ell^2/(2\mu)\right\} |\vartheta|_{\mathcal{A}^*}^2, \tag{38}$$

$\forall\vartheta \in \mathcal{C} \cup \mathcal{D}$, where we used the facts that $\tau_{v,1} \in [0, N_{0,v}]$, $\tau_{v,2} \in [0, T_{0,v}]$, $\sigma_v \in \Sigma_v$, and $w \in \mathcal{W}$. For every $\vartheta \in \mathcal{D}$ we have that the jumps satisfy $U(\vartheta^+) \leq U(\vartheta)$. Moreover, for $\vartheta \in \mathcal{C}$ the time derivative of U is upper bounded as follows:

$$\begin{aligned}
\dot{U}(\vartheta) &= \theta\dot{\bar{W}}(x_e, u, \tau_{v,1}, \tau_{v,2}, \sigma_v) + (1-\theta)\dot{V}(u,w) \\
&\leq -\rho\theta \min\{\underline{\lambda}(P), \tilde{\beta}\} \|x_e\|^2 + 2\varepsilon\ell\theta \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\| \|x_e\| \|u - u_w^*\| e^{\tau_0} \\
&\quad + 2\varepsilon\ell_y\theta e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\| \|C\| \|G\| \|x_e\|^2 + 2\theta e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}E\right\| \|x_e\| \|\Pi(w)\| \\
&\quad - (1-\theta)\varepsilon\mu^2 \|u - u_w^*\|^2 + \varepsilon\ell\ell_y(1-\theta) \|G\| \|C\| \|u - u_w^*\| \|x_e\| \\
&\quad + \ell_y(1-\theta) \|H\| \|G\| \|u - u^*\| \|\Pi(w)\|.
\end{aligned}$$

Using $\xi := (\|x_e\|, \|u - u_w^*\|)$ we obtain

$$\dot{U}(\vartheta) \leq -\varepsilon\xi^\top \Xi\xi + r^\top\xi \|\Pi(w)\|, \tag{39}$$

where $r := \left(2\theta e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}E\right\|, \ell_y(1-\theta) \|H\| \|G\|\right)$, and Ξ is a symmetric matrix as in Lemma 7 with parameters $\alpha := \rho\min\{\underline{\lambda}(P), \tilde{\beta}\}$, $\beta := 2\ell_y e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\| \|C\| \|G\|$, $\delta := \ell\ell_y \|C\| \|G\|$, $\chi := 2\ell e^{\tau_0} \max\left\{\bar{\lambda}(P), \tilde{\beta}\right\} \left\|A^{-1}B\right\|$, $\gamma := \mu^2$. When $\theta := \frac{\delta}{\delta + \chi}$, and ε satisfies (26) we obtain that Ξ is positive definite via

Lemma 7. Finally, we show that the the quadratic term dominates the linear term in (39). Indeed, for $k \in (0, 1)$, we can rewrite (39) as follows:

$$\dot{U}(\vartheta) \leq -\varepsilon(1-k)\underline{\lambda}(\Xi)\,\|\xi\|^2 - \varepsilon k \underline{\lambda}(\Xi)\,\|\xi\|^2 + r^\top \xi\,\|\Pi(w)\|$$

$$\leq -\varepsilon(1-k)\underline{\lambda}(\Xi)\,\|\xi\|^2 - \varepsilon k \underline{\lambda}(\Xi)\,\|\xi\|^2 + \|r\|\,\|\xi\|\,\|\Pi(w)\|$$

$$\leq -\varepsilon(1-k)\underline{\lambda}(\Xi)\,\|\xi\|^2 = -\varepsilon(1-k)\underline{\lambda}(\Xi)|\vartheta|^2_{\mathcal{A}^*}, \quad \forall\, |\vartheta|_{\mathcal{A}^*} \geq \frac{\|r\|}{\varepsilon \underline{\lambda}(\Xi) k}\,\|\Pi(w)\|_t\,. \tag{40}$$

The result follows by applying [35, Thm. 3.1], exploiting the quadratic bounds in (38), and the average dwell-time condition. ∎

6. Numerical Examples

In this section, we present two numerical experiments to illustrate our theoretical contributions. The first simulation utilizes a synthetic system; in the second experiment, we consider the problem of frequency regulation in power transmission systems.

6.1. Synthetic system

We simulate a linear time-invariant system with state $x \in \mathbb{R}^2$, static control input $v \in \mathbb{R}$, dynamic control input $u \in \mathbb{R}$, and exogenous signal $w \in \mathbb{R}$. The goal is to regulate the solutions of the plant towards the solution of the following optimization problem

$$\min_{u,y} f_u(u) + f_y(y) := \min_{u,y} u^\top R u + (y - y_{\text{ref}})^\top Q(y - y_{\text{ref}}), \tag{41}$$

with $R = 2$, and $Q = [1\ 0; 0\ 2]$.

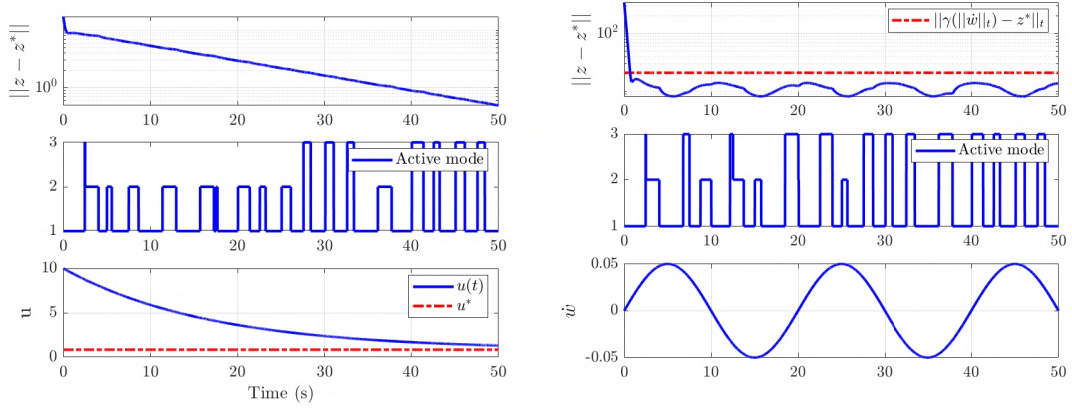#### 6.1.1. Attacks on the static feedback controller

We first consider the scenario where the static controller operates under two switching attacks, and the dynamic controller operates in the nominal stable mode. The plant has the following structure:

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 2 & -1.5 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} v + \begin{bmatrix} -1.06 \\ -0.62 \end{bmatrix} u + \begin{bmatrix} -0.82 \\ -0.79 \end{bmatrix} w, \qquad y = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} x. \tag{42}$$

For this plant, we design an internal static controller that places the poles of the closed loop system $(F + NKC)$ at -2. Hence, $K = [-40 \quad 5]$. The attacks acting on the controller are modeled by the scalars $L_{a_1} = 0$, and $L_{a_2} = -0.1$, which leads to the following matrices:

$$A_{a_1} = \begin{bmatrix} 1 & 0 \\ 2 & -1.5 \end{bmatrix} \quad B_{a_1} = \begin{bmatrix} 0.34 \\ 0.78 \end{bmatrix} \quad E_{a_1} = \begin{bmatrix} 0.30 \\ 0.34 \end{bmatrix},$$

$$A_{a_2} = \begin{bmatrix} 1.4 & -0.5 \\ 2.4 & -1.55 \end{bmatrix} \quad B_{a_2} = \begin{bmatrix} 0.48 \\ 0.92 \end{bmatrix} \quad E_{a_2} = \begin{bmatrix} 0.42 \\ 0.45 \end{bmatrix},$$

which satisfy Assumption 7. The time-ratio parameter is set as $\kappa_{v,2} = 0.365$, which induces $\alpha = 1.71$ in Proposition 3. The gain inducing the time-scale separation is set as $\varepsilon^* = 0.0149$. In our simulation, the nominal stable mode is denoted by the index 1, the unstable mode σ_{a_1} is denoted by the index 2, and σ_{a_2} is denoted by the index 3. Thus $\Sigma_v = \{1, 2, 3\}$ and $\Sigma_{v,a} = \{2, 3\}$. The theoretical value of ε^* is conservative, and to obtain faster convergence we also simulated the system with a gain of $\varepsilon = 20\varepsilon^*$. The left-hand side of Figure 3 shows the performance of the closed-loop system under a constant signal $w = 0.96$. As expected, the trajectories of the system converge to the set of optimal solutions. On the other hand, the right-hand side shows the performance of the system under a time-varying exogenous signal modeled by $\dot{w} = a \sin(\omega t)$, with $a = 0.05$ and $\omega = 2\pi \times 0.05$. Here, we can see that the norm of the tracking error is eventually bounded by 21.

Figure 3: Attack to static feedback controller under constant and time-varying perturbations w.
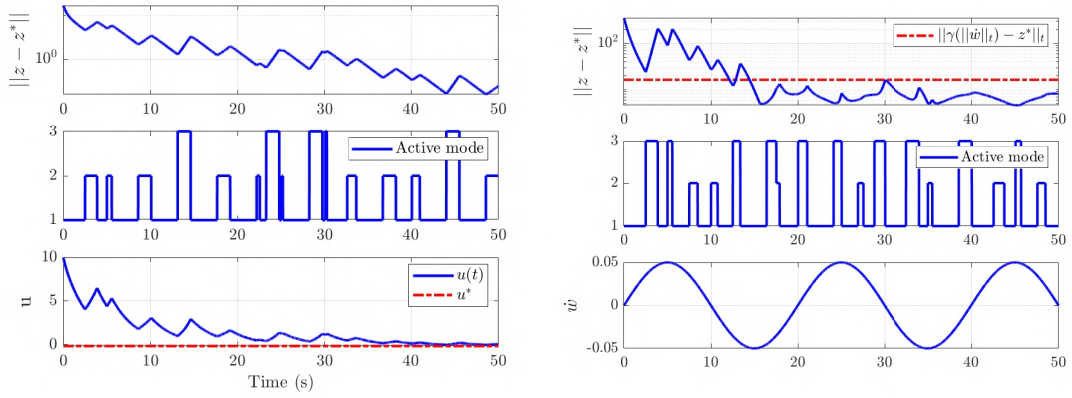
Figure 4: Attack to dynamic feedback controller under constant and time-varying perturbations w.

6.1.2. Attacks on the dynamic feedback controller

We now consider the dynamic controller operating under two switching attacks, while the nominal static controller (stable) remains free from any attack. In this case, we have the following nominal plant

$$\dot{x} = \begin{bmatrix} -3 & 0.5 \\ -2 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u + \begin{bmatrix} 1 \\ 1 \end{bmatrix} w, \qquad y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x. \tag{43}$$

We simulate the system in (43) interconnected with the dynamic controller (9) under attacks. The attacks are modeled by the scalar gains $M_{a_1} = -1$, and $M_{a_2} = -2$. The time-ratio parameter is set to $\kappa_{u,2} = 0.33$. To induce time-scale separation, the gain is set as $\varepsilon^* = 0.0093$. In the simulation, the nominal stable mode is characterized by the index $\{1\}$, the attack σ_{a_1} is characterized by the mode $\{2\}$, and also σ_{a_2} is characterized by $\{3\}$. Since simulations show that the theoretical gain ε^* is very conservative, we set $\varepsilon = 20\varepsilon^*$ to achieve faster convergence in the experiments. In the left plot of Figure 4, we show the decreasing tracking error under a constant w. Similarly, in the right plot, we show the behavior of the system under a time-varying signal $\dot{w} = a\sin(\omega t)$ with $a = 0.05$ and $\omega = 2\pi \times 0.05$. As predicted by Theorem 2, the tracking error is eventually uniformly ultimately bounded.

6.2. Persistent Attack in a Power System

We study the problem of regulating frequency in a power transmission system; in particular, we model the frequency dynamics as an LTI system based on a DC model of the network and a linearized model of the swing equation; see,

e.g., the experimental setup in [2, Section V].[3] Accordingly, we consider the following LTI dynamical system:

$$
\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & I \\ -M^{-1}Y & -M^{-1}D \end{bmatrix}}_{\bar{A}} \begin{bmatrix} \theta \\ \omega \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ B_1 \end{bmatrix}}_{\bar{B}} u + \underbrace{\begin{bmatrix} 0 \\ E_1 \end{bmatrix}}_{\bar{E}} w,
$$

$$
\underbrace{\begin{bmatrix} p_l \\ \bar{\omega} \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} Y & 0 \\ 0 & \frac{1}{n}\mathbb{1}^\top \end{bmatrix}}_{\bar{C}} \begin{bmatrix} \theta \\ \omega \end{bmatrix},
$$

(44)

where the state $z := (\theta, \omega) \in \mathbb{R}^{2n_1}$ collects the phase angles θ_i and frequencies ω_i at n_1 nodes, the vector u collects the powers generated by generator (modeled as synchronous machines), w represents uncontrollable loads, p_l is the vector of power flows on the lines, and $\bar{\omega}$ is the average system frequency. The matrix $Y \in \mathbb{R}^{n_1 \times n_1}$ is the susceptance matrix of the grid whose row sum is equal to zero, i.e. $Y\mathbb{1} = 0$. The matrix $B_1 \in \{0,1\}^{n_1 \times m}$ sets the controllable nodes or generators i.e. $B_1(i,j) = 1$ when i is a generator; and $E_1 \in \{0,1\}^{n_1 \times q}$ sets the uncontrollable loads, i.e. $B_2(i,j) = 1$ when i is a load. Finally, $M = \text{diag}(m_1, \ldots, m_{n_1})$ and $D = \text{diag}(d_1, \ldots, d_{n_1})$ are diagonal matrices respectively containing the rotational inertia coefficients and the damping coefficients in the diagonal.

System (44) is marginally stable since $\bar{A}$ has an eigenvalue at zero with eigenvector $(\mathbb{1}, 0)$. We eliminate the average mode $\bar{\theta}$ with the following coordinate transformation [2, Section V]:

$$
x := \begin{bmatrix} \theta \\ \omega \end{bmatrix} = \underbrace{\begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix}}_{T} x + \begin{bmatrix} \mathbb{1} \\ 0 \end{bmatrix} \bar{\theta},
$$

(45)

where the columns of the matrix $U \in \mathbb{R}^{n_1 \times (n_1-1)}$ form an orthonormal basis orthogonal to span($\mathbb{1}$). One way to select this U is taking the $n_1 - 1$ eigenvectors of Y that correspond to the nonzero eigenvalues. The new system with state $x = T^\top x \in \mathbb{R}^{2n_1-1}$ takes the following form

$$
\begin{aligned}
\dot{x} &= Ax + Bu + Ew, \\
y &= Cx,
\end{aligned}
$$

(46)

where $A := T^\top \bar{A} T$, $B := T^\top \bar{B}$, $E := T^\top \bar{E}$, and $C := \bar{C} T$. System (46) is stable and the output y remains unaltered by the coordinate transformation. Based on this, we consider the following optimization problem associated with the transmission systems:

$$
\begin{aligned}
\min_{u} \quad & u^\top Q u + f_y(y) \\
\text{s.t} \quad & y = Gu + Hw
\end{aligned}
$$

(47)

where $G = -CA^{-1}B$, $H = -CA^{-1}E$, $Q \in \mathbb{R}^{m \times m}$ is positive definite and $f_y(y) = \eta \max\left\{\underline{y}_i - y_i, y_i - \bar{y}_i, 0\right\}^2$. The function $f_y(\cdot)$ imposes soft constraints to the system output y; moreover, $f_y(y)$ is continuously differentiable with a η−Lipschitz gradient. With this formulation, the dynamic feedback controller is given by:
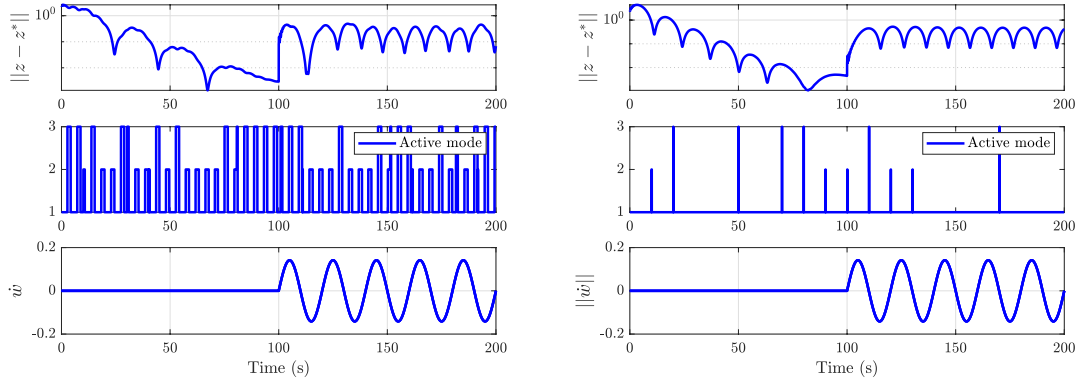
$$
\dot{u} = -\varepsilon(2Hu + \eta G^\top \nabla f_y(y)),
$$

(48)

where $\varepsilon > 0$, and $\nabla f_y(\cdot)$ boils down to the following soft thresholding operator:

$$
\nabla f_y(v) = \begin{cases} v_i - \underline{y}_i, & v_i \leq \underline{y}_i \\ 0 & \underline{y}_i < v_i < \bar{y}_i \\ v_i - \bar{y}_i & v_i \geq \bar{y}_i. \end{cases}
$$

---

[3]In this section, we denote as I_n the identity matrix of dimension n, the vector of ones as $\mathbb{1}$, $0_{n \times m}$ is a matrix of zeros of dimension $n \times m$, and the symbol diag() transforms a vector to a matrix with the elements of the vector on the diagonal.

Figure 5: In the right figure, we present an attack to dynamic feedback controller to system in (46) with $\varepsilon = 1e5\varepsilon^*$. In the left figure, we present an attack to the plant in (46).

The nominal system that we simulate contains a simplified version of the IEEE 9Bus Test case; see [2] for a schematic of the network. In particular, it features $n_1 = 6$ nodes, with generators at 3 nodes, and uncontrollable loads under disturbances in the remaining 3 (with loads modeled as "negative" generators). We build the matrices in system (44) as follows $M = \mathrm{diag}(1, 1, 1, 0.1, 0.1, 0.1)$, $D = 0.1 * I_{n_1}$, $B_1 = \mathrm{diag}(1, 1, 1, 0, 0, 0)$, and $E_1 = (0_{(n_1-q)\times q}, I_q)$, and

$$Y = \begin{bmatrix} 13.5776 & 0 & 0 & -6.9783 & -6.5993 & 0 \\ 0 & 11.7898 & 0 & -4.3844 & 0 & -7.4054 \\ 0 & 0 & 10.4895 & 0 & -4.2507 & -6.2388 \\ -6.9783 & -4.3844 & 0 & 11.3627 & 0 & 0 \\ -6.5993 & 0 & -4.2507 & 0 & 10.8500 & 0 \\ 0 & -7.4054 & -6.2388 & 0 & 0 & 13.6441 \end{bmatrix}.$$

Since $A \in \mathbb{R}^{n \times n}$, where $n := 2n_1 - 1$, we use $R = I_n$ to compute P according to Assumption 2. We consider a constant $w = (-0.9, -1, -1.25)$ for the first 100 s, and $-0.45(\cos(2\pi 0.05t) + 1)$ for the remaining 100 s. For the cost function, we set $Q = I_m$; furthermore, the parameters for f_y are $\eta = 10$, $\bar{y} = 5 * \mathbb{1}$, and $\underline{y} = -5 * \mathbb{1}$.

We simulate attacks to the dynamical controllers with $M_{a_2} = -1$, and $\bar{M}_{a_3} = -2$. The right plots of Figure 5 show the results of this simulation. We implement our algorithm (48) with $\varepsilon = 1e5\varepsilon^*$; we note that the error $||z - z^*||$ decreases whenever $||\dot{w}||$ is constant, and remains bounded whenever w is time-varying. Moreover, we simulate attacks acting on the plant by systematically modifying the matrix M in the following way: $M_2 = \mathrm{diag}(1, 1, -101, 0.1, 0.1, 0.1)$, and $M_3 = \mathrm{diag}(1, 1, -105, 0.1, 0.1, 0.1)$. This modifications lead to an unstable matrix A. On the left side of Figure 5 we present the results. As expected, the error $||z - z^*||$ decreases when w remains constant, but only remains bounded whenever w is time-varying. These results are consistent with our theoretical results.

7. Conclusions

In this paper, we presented the first systematic stability analysis of LTI plants controlled by dynamic gradient-based controllers in subject to persistent attacks. For cost functions that satisfy certain smoothness and gradient-dominated properties, we showed that the static output-feedback controllers guarantee exponential stability of the closed-loop system whenever the total activation time of the attacks in any given window of time is sufficiently bounded, and not too frequent. For the dynamic gradient-flow controllers, we showed that system stability can be further guaranteed by properly tuning the controller gain. To the best knowledge of the authors, our results are the first in the literature of feedback-based automatic optimization that study the convergence and stability properties of the algorithms using the framework of hybrid dynamical systems, and Lyapunov-based tools for switched systems with unstable modes. Overall, our results demonstrate that, under exogenous disturbances, input-to-state stability can be guaranteed for the closed-loop system provided the total activation time of the unstable modes satisfies a particular upper bound, and

enough time-scale separation is induced between the dynamics of the plant and the controller. We also presented explicit characterizations of the ISS gains, the theoretical upper bounds on the controller gains, and the time ratio constraints needed to preserve stability. Our results were validated in two numerical examples. Relevant research directions that require further investigation include the analysis of switching plants with non-common equilibrium points and extensions to plants with nonlinear dynamics.

References

[1] S. Menta, A. Hauswirth, S. Bolognani, G. Hug, F. Dörfler, Stability of dynamic feedback optimization with applications to power systems, in: Annual Conf. on Communication, Control, and Computing, 2018, pp. 136–143.

[2] M. Colombino, E. Dall'Anese, A. Bernstein, Online optimization as a feedback controller: Stability and tracking, IEEE Transactions on Control of Network Systems 7 (2019) 422–432.

[3] G. Bianchin, J. I. Poveda, E. Dall'Anese, Online optimization of switched LTI systems using continuous-time and hybrid accelerated gradient flows, arXiv (2020). ArXiv:2008.03903.

[4] S. H. Low, D. E. Lapsley, Optimization flow control. i. basic algorithm and convergence, IEEE/ACM Transactions on networking 7 (1999) 861–874.

[5] T. Zheng, J. W. Simpson-Porco, E. Mallada, Implicit trajectory planning for feedback linearizable systems: A time-varying optimization approach, arXiv preprint (2019). ArXiv:1910.00678.

[6] H. Karimi, J. Nutini, M. Schmidt, Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition, in: Machine Learning and Knowledge Discovery in Databases, Riva del Garda, Italy, 2016, pp. 795–811.

[7] A. Jokic, M. Lazar, P. P. V. D. Bosch, On constrained steady-state regulation: Dynamic KKT controllers, IEEE Transactions on Automatic Control 54 (2009) 2250–2254.

[8] F. D. Brunner, H.-B. Dürr, C. Ebenbauer, Feedback design for multi-agent systems: A saddle point approach, in: IEEE Conf. on Decision and Control, 2012, pp. 3783–3789.

[9] A. Hauswirth, S. Bolognani, G. Hug, F. Dörfler, Timescale separation in autonomous optimization, IEEE Transactions on Automatic Control 66 (2021) 611–624.

[10] L. S. P. Lawrence, Z. E. Nelson, E. Mallada, J. W. Simpson-Porco, Optimal steady-state control for linear time-invariant systems, in: IEEE Conf. on Decision and Control, Miami Beach, FL, USA, 2018, pp. 3251–3257.

[11] A. Hauswirth, S. Bolognani, G. Hug, F. Dorfler, Optimization algorithms as robust feedback controllers, arXiv:2103.11329 (2021) 1–32.

[12] F. Galarza-Jimenez, J. I. Poveda, G. Bianchin, E. Dall'Anese, Extremum seeking under persistent gradient deception: A switching systems approach, IEEE Control Systems Letters (2021). To appear.

[13] J. I. Poveda, A. R. Teel, A framework for a class of hybrid extremum seeking controllers with dynamic inclusions, Automatica 76 (2017) 113–126.

[14] F. Pasqualetti, F. Dörfler, F. Bullo, Attack detection and identification in cyber-physical systems, IEEE Transactions on Automatic Control 58 (2013) 2715–2729.

[15] D. Senejohnny, P. Tesi, C. De Persis, A jamming-resilient algorithm for self-triggered network coordination, IEEE Transactions on Control of Network Systems 5 (2018) 981–990.

[16] A. A. Cardenas, S. Amin, S. Sastry, Secure control: Towards survivable cyber-physical systems, in: 2008 The 28th International Conference on Distributed Computing Systems Workshops, 2008, pp. 495–500.

[17] A. Cetinkaya, H. Ishii, T. Hayakawa, Networked control under random and malicious packet losses, IEEE Transactions on Automatic Control 62 (2016) 2434–2449.

[18] H. S. Foroush, S. Martinez, On event-triggered control of linear systems under periodic denial-of-service jamming attacks, in: IEEE Conf. on Decision and Control, IEEE, 2012, pp. 2551–2556.

[19] A. Kundu, D. E. Quevedo, Stabilizing scheduling policies for networked control systems, IEEE Transactions on Control of Network Systems 7 (2020) 163–175.

[20] N. Forti, G. Battistelli, L. Chisci, B. Sinopoli, Secure state estimation of cyber-physical systems under switching attacks, IFAC-PapersOnLine 50 (2017) 4979–4986.

[21] S. Z. Yong, M. Zhu, E. Frazzoli, Resilient state estimation against switching attacks on stochastic cyber-physical systems, in: 2015 54th IEEE Conference on Decision and Control (CDC), 2015, pp. 5162–5169.

[22] C. Wu, L. Wu, J. Liu, Z.-P. Jiang, Active defense-based resilient sliding mode control under denial-of-service attacks, IEEE Transactions on Information Forensics and Security 15 (2020) 237–249.

[23] C. Wu, W. Pan, G. Sun, J. Liu, L. Wu, Learning tracking control for cyber–physical systems, IEEE Internet of Things Journal 8 (2021) 9151–9163.

[24] Y. Gao, G. Sun, J. Liu, Y. Shi, L. Wu, State estimation and self-triggered control of cpss against joint sensor and actuator attacks, Automatica 113 (2020) 108687.

[25] Y. Wang, J. Lu, J. Liang, Security control of multiagent systems under denial-of-service attacks, IEEE Transactions on Cybernetics (2020) 1–11.

[26] X.-F. Wang, A. R. Teel, K.-Z. Liu, X.-M. Sun, Stability analysis of distributed convex optimization under persistent attacks: A hybrid systems approach, Automatica 111 (2020) 108607.

[27] Y. Qi, Y. Tang, Z. Ke, Y. Liu, X. Xu, S. Yuan, Dual-terminal decentralized event-triggered control for switched systems with cyber attacks and quantization, ISA Transactions (2020).

[28] S. Sundaram, B. Gharesifard, Distributed optimization under adversarial nodes, IEEE Transactions on Automatic Control 64 (2018) 1063–

1076.

[29] L. Su, N. H. Vaidya, Byzantine-resilient multi-agent optimization, IEEE Transactions on Automatic Control (2020). In press.

[30] Y. Chen, L. Su, J. Xu, Distributed statistical machine learning in adversarial settings: Byzantine gradient descent, ACM on Measur. and Analy. of Computing Syst. 1 (2017) 1–25.

[31] B. Turan, C. A. Uribe, H.-T. Wai, M. Alizadeh, On robustness of the normalized subgradient method with randomly corrupted subgradients, arXiv preprint (2020). ArXiv:2009.13725.

[32] L. Zhai, K. G. Vamvoudakis, Data-based and secure switched cyber–physical systems, Systems and Control Letters 148 (2021) 104826.

[33] D. Angeli, E. D. Sontag, Y. Wang, Input-to-state stability with respect to inputs and their derivatives, International Journal of Robust and Nonlinear Control 13 (2003) 1035–1056.

[34] R. Goebel, R. G. Sanfelice, A. R. Teel, Hybrid Dynamical Systems, Princeton University Press, Princeton, NJ, USA, 2012.

[35] C. Cai, A. R. Teel, Characterizations of input-to-state stability for hybrid systems, Systems and Control Letters 58 (2009) 47 – 53.

[36] A. R. Teel, F. Forni, L. Zaccarian, Lyapunov-based sufficient conditions for exponential stability in hybrid systems, IEEE Transactions on Automatic Control 58 (2013) 1591–1596.

[37] A. Hauswirth, I. Subotić, S. Bolognani, G. Hug, F. Dörfler, Time-varying projected dynamical systems with applications to feedback optimization of power systems, in: IEEE Conf. on Decision and Control, 2018, pp. 3258–3263.

[38] G. Bianchin, J. Cortés, J. I. Poveda, E. Dall'Anese, Time-varying optimization of LTI systems via projected primal-dual gradient flows, arXiv (2021). ArXiv:2101.01799.

[39] L. Gao, F. Li, J. Fu, Output-based decentralised event-triggered dissipative control of ncss under aperiodic dos attacks, International Journal of Systems Science 0 (2021) 1–17.

[40] S. Hu, D. Yue, Z. Cheng, E. Tian, X. Xie, X. Chen, Co-design of dynamic event-triggered communication scheme and resilient observer-based control under aperiodic dos attacks, IEEE Transactions on Cybernetics (2020) 1–11.

[41] D. Zhang, L. Liu, G. Feng, Consensus of heterogeneous linear multiagent systems subject to aperiodic sampled-data and dos attack, IEEE Transactions on Cybernetics 49 (2019) 1501–1511.

[42] J. V. Schaftingen, A direct proof of the existence of eigenvalues and eigenvectors by weierstrass's theorem, The American Mathematical Monthly 120 (2013) 741–746.

[43] R. A. Horn, C. R. Johnson, Matrix Analysis, 2 ed., Cambridge University Press, 2012.

[44] J. W. Simpson-Porco, On area control errors, area injection errors, and textbook automatic generation control, IEEE Transactions on Power Systems (2020).

[45] P. Chakraborty, S. Dhople, Y. C. Chen, M. Parvania, Dynamics-aware continuous-time economic dispatch and optimal automatic generation control, in: 2020 American Control Conference (ACC), IEEE, 2020, pp. 1292–1298.

[46] M. Papageorgiou, A. Kotsialos, Freeway ramp metering: An overview, IEEE Transactions on Intelligent Transportation Systems 3 (2002) 271–281.

[47] P. Grandinetti, C. Canudas-de Wit, F. Garin, Distributed optimal traffic lights design for large-scale urban networks, IEEE Transactions on Control Systems Technology 27 (2018) 950–963.

[48] S. Paternain, D. Koditschek, A. Ribeiro, Navigation functions for convex potentials in a space with convex obstacles, IEEE Transactions on Automatic Control 63 (2018) 2944–2959.

[49] C. Zhang, D. Arnold, N. Ghods, A. Siranosian, M. Krstić, Source seeking with non-holonomic unicycle without position measurement and with tuning of forward velocity, Systems and Control Letters 56 (2007) 245–252.

[50] J. I. Poveda, M. Benosman, A. R. Teel, R. G. Sanfelice, Coordinated hybrid source seeking with robust obstacle avoidance in multi-vehicle autonomous systems, IEEE Transactions on Automatic Control, 10.1109/TAC.2021.3056365 (2021) 1–16.

[51] R. M. Murray, Z. Li, S. S. Sastry, A Mathematical Introduction to Robotic Manipulation, CRC Press, 1994.

[52] T. Alpcan, T. Başar, A stability result for switched systems with multiple equilibria, Dynamics of Continuous, Discrete and Impulsive Systems Series A: Mathematical Analysis 17 (2010) 949–958.

[53] S. Veer, I. Poulakakis, Switched systems with multiple equilibria under disturbances: Boundedness and practical stability, IEEE Transactions on Automatic Control 65 (2020) 2371–2386.

Appendix

Lemma 7. *Let* $\alpha, \beta, \delta, \chi, \gamma$ *be positive scalars, let* $\theta \in (0,1)$ *be tunable parameters, and let*

$$\Xi = \begin{bmatrix} \theta\left(\frac{\alpha}{\varepsilon} - \beta\right) & -\frac{1}{2}((1-\theta)\delta + \theta\chi) \\ -\frac{1}{2}((1-\theta)\delta + \theta\chi) & (1-\theta)\gamma \end{bmatrix}.$$

If $0 < \varepsilon < \alpha\gamma/(\beta\gamma + \delta\chi)$, *then there exists* $\theta < \delta/(\delta + \chi)$, *such that* Ξ *is positive definite.*

Proof: The matrix Ξ is positive definite if and only if the leading principal minors are positive. In this case $(1-\theta)\gamma > 0$ and $\theta(1-\theta)\left(\frac{\alpha}{\varepsilon} - \beta\right)\gamma > \frac{1}{4}((1-\theta)\delta + \theta\chi)^2$. The first inequality is guarantee by the definition of θ and γ. The second inequality can be rewritten as:

$$\varepsilon < \frac{\alpha\gamma}{\beta\gamma + \frac{((1-\theta)\delta + \theta\chi)^2}{4\theta(1-\theta)}} = \hat{\varepsilon}(\theta).$$

The function $\hat{\varepsilon}$ attains its maximum at $\theta = \theta^* := \frac{\delta}{\delta+\chi}$, with $\varepsilon^* := \hat{\varepsilon}(\theta^*) = \frac{\alpha\gamma}{\beta\gamma+\delta\chi}$. ∎