Tutorial on Supervisory Control*

João P. Hespanha[†]
hespanha@ece.ucsb.edu

Dept. Electrical & Computer Engineering, Univ. of California
Santa Barbara, CA 93106-9560

April 1, 2002[‡]

Abstract

This paper provides an overview of the results available on supervisory control in an adaptive context, with special emphasis on estimator-based supervisory architectures. We cover both the linear and nonlinear cases in a unified manner, highlighting the common properties and analysis techniques. The style of the paper is tutorial and provides the reader with enough detail to apply these techniques to general classes of linear and nonlinear systems.

Contents

*Lecture Notes for the workshop *Control using Logic and Switching* for the 40th Conf. on Decision and Contr., Orlando, Florida

[†]This material is based upon work partially supported by the National Science Foundation under Grant No. ECS-0093762.

Background

This paper is mostly self-contained. However, the reader may want to also get familiar with the following material:

1. Design of linear multi-estimators—This material is briefly summarized in Appendix A.
2. Linear Certainty Equivalence Stabilization and Output Stabilization Theorems [55].
3. (Average) Dwell-time Switching Theorem [32].

1 Overview

Section Summary

In this section we present the basic supervisory control architecture and put it in perspective with respect to related work. We also introduce key properties of the overall system and explain how they are used to analyze it. This is done in a in a qualitative fashion appealing to the intuition of the reader.

1.1 Supervisory control

The idea of integrating logic with continuous dynamics in the control of complex systems is certainly not new. Consider for example an industrial setting where a human operator periodically adjusts the set points of an array of PID controllers to account for changes in the environment. One can recognize the human operator as a component of the feedback loop that adjusts the continuous dynamics using logic-based decision rules.

The basic problem considered here is the control of complex systems for which traditional control methodologies based on a single controller do not provide satisfactory performance. In switching control, one builds a bank of alternative candidate controllers and switches among them based on measurements collected online. The switching is orchestrated by a specially designed logic that uses the measurements to assess the performance of the candidate controller currently in use and also the potential performance of alternative controllers. Figure 1 shows the basic architecture employed by switching control. In this figure u represents the control signal, w an exogenous disturbance and/or

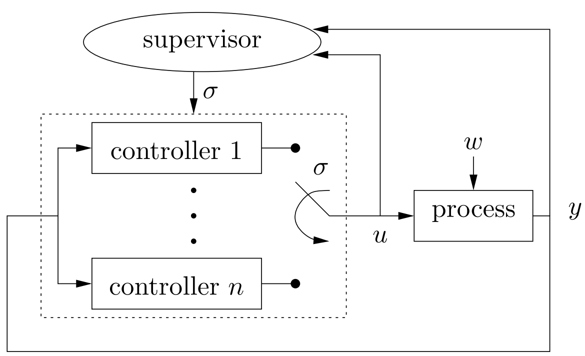

Figure 1: Switching control

measurement noise, and y the measured output. The dashed box is as a conceptual representation of a switching controller. In practice, switching controllers are implemented differently. Suppose that we desire to switch among a family $\mathcal{C}$ of controllers parameterized by some variable $q \in \mathcal{Q}$. For example, we could have

$$\mathcal{C} := \big\{\dot{z}_q = F_q(z_q, y),\ u = G_q(z_q, y) : q \in \mathcal{Q}\big\},$$

with the parameterizing set $\mathcal{Q}$ finite, infinite but countable, or not even countable (e.g., a ball in $\mathbb{R}^m$); and all the z_q with the same dimension. Switching among the controllers in $\mathcal{C}$ can be accomplished using the following *multi-controller*:

$$\dot{x}_C = F_\sigma(x_C, y), \qquad u = G_\sigma(x_C, y), \tag{1}$$

where the signal $\sigma : [0, \infty) \to \mathcal{Q}$—called the *switching signal*—effectively determines which controller is in the loop at each instant of time. The points of discontinuity of σ correspond to a change in candidate controller and are therefore called *switching times*. The multi-controller in (1) is far more efficient than the conceptual structure in Figure 1 as its dimension is independent of the number of candidate controllers. Moreover, it avoids the difficulties that arise in Figure 1 when some of the out-of-loop controllers are unstable [56]. In this section, we use a continuous-time multi-controller such as (1) to keep the exposition concrete. However, the concepts presented generalize to other types of candidate control laws, such as discrete-time [8] or hybrid controllers [25].

The top element in Figure 1 is the logic that controls the switch, or more precisely, that generates the switching signal in (1). This logic is called the *supervisor* and its purpose is to monitor the signals that can be measured (in this case u and y) and decide, at each instant of time, which candidate controller should be put in the feedback loop with the process. The key difference between the type of switching control proposed here and adaptive algorithms based on continuous tuning is the use of logic within the supervisor to control the learning process. In fact, traditional adaptive control could be viewed as a form of switching control where the "switching signal" σ is generated by means of an ordinary differential equation of the form

$$\dot{\varphi} = \Phi(\sigma, \varphi, u, y), \qquad \dot{\sigma} = \Psi(\sigma, \varphi, u, y). \tag{2}$$

In supervisory control, the supervisor combines continuous dynamics with discrete logic and is therefore an *hybrid system*. The modeling of such systems has received considerable attention in the literature in the last few years [77]. A typical hybrid supervisor can be defined by a an ordinary differential equation coupled with a recursive equation such as

$$\dot{\varphi} = f_\delta(\varphi, u, y), \qquad \delta = \phi(\varphi, \delta^-), \qquad t \geq t_0, \tag{3}$$

where, for each $t > t_0$, $\delta^-(t)$ denotes the limit from the left of $\delta(\tau)$ as $\tau \uparrow t$. The signal φ is called the *continuous state* of the supervisor and δ the *discrete state*. The output $\sigma \in \mathcal{Q}$ of such a hybrid system is defined by an output equation of the form

$$\sigma = \psi_\delta(\varphi). \tag{4}$$

More general models for hybrid systems can be found, e.g., in [77]. Algorithms that use supervisors such as (3)–(4) are called *supervisory controllers*, whereas algorithms that use continuous tuning laws such as (2) are called *adaptive controllers*. The use of supervisory control has several potential advantages over traditional forms of adaptive control. These are discussed with some depth in [27] and summarized below:

Rapid Adaptation. With supervisory control, online adaptation to sudden changes in the process or the control objectives can occur very fast because the signal σ that controls which controller is placed in the feedback loop is not restricted to vary in a continuous fashion. Rapid adaptation is critical to deal with processes likely to undergo sudden changes in its dynamics (e.g., due to faults or external interferences) that may render the system unstable or degrade its performance significantly. We will see below that this is the case in some of our application areas.

Flexibility and Modularity. Supervisory control is based on a modular architecture that separates the candidate controllers (implemented as a multi-controller) from the learning mechanism (i.e., the supervisor). This allows for the integration into supervisory control of off-the-shelf candidate controllers, designed using existing theories for nonadaptive systems. In adaptive control, on the other hand, the candidate controllers usually have to be specially tailored to the tuning mechanism, leaving little freedom to the controller designer.

The flexibility and modularity of supervisory control is particularly important for processes that are difficult to control or for which the performance requirements call for advanced techniques to design the candidate controllers. Supervisory control allows the use of candidate controllers designed using techniques such as LQG/LQR [45], $\mathcal{H}_\infty$ [6], μ-synthesis [91], feedback linearization [35], backstepping [42], inverse optimality [19], etc. Flexibility in design is also important when there are constrains on the structure of the candidate controllers, e.g., in many applications in process control it is desirable to use PID controllers that are already in place. This feature of supervisory control is crucial for the successful industrial integration of algorithms capable of sophisticated forms of adaptation and learning.

Decoupling between Supervision and Control. Between switchings times the process is connected to one of the candidate controllers and the dynamics of the supervisor play not role in the evolution of the resulting closed-loop system. This simplifies the analysis of the overall algorithm considerably. We will see below that, for analysis purposes, we can abstract the detailed behaviors of the multi-controller and supervisor and concentrate on a small set of properties that these systems exhibit. These properties can, in turn, be used to infer properties of the overall system. This type of analysis provides an intuitive understanding of the behavior of the complex system and guides the controller designer in the pursuit of higher performance.

This form of decoupling also adds robustness to the design because nonlinearities in the supervisor do not affect directly the dynamics of the system. For example, when the process and each candidate controller are linear, the overall system will be linear between switching times. This is significantly different from adaptive control, in which the tuning equation (2) always renders the closed-loop system nonlinear. Because of this, unmodeled dynamics may cause finite escape time (i.e., signals that grow to infinity in finite time) in adaptive control, even when both the process and the candidate controllers are linear. This is excluded from approaches based on supervisory control.

1.2 Adaptive supervisory control

The introduction of adaptation and learning in the field of automatic control dates, at least, as far back as the 1950s. However, it was not until the 1980s that the field was sufficiently well understood to find application in the industry. Instrumental to this were the development of theoretical tools to analyze the stability of adaptive systems. For an historical perspective on the early work on adaptive control see [5]. The use of switching and logic to overcome the difficulties that arise in the control of a poorly modeled processes has its roots in the pioneer work of Mårtenson [50]. Mårtenson showed that it is possible to build a switching controller capable of stabilizing *every* process that can be stabilized by some controller with order no larger than a given integer n. The controller proposed effectively "searched" for a stabilizing controller by switching among the elements of an ordered set of candidate controllers that was dense on the set of all controllers of order up to n. Mårtenson's work was of a theoretical nature and the controller proposed has no practical application due to its very poor performance.

Since [50], a sequential or "pre-routed" search among a set of controllers has been explored in a several control algorithms proposed in the literature [21, 14, 52, 53, 13, 78, 20]. Many of these algorithms attempt to address the issue of performance but usually in an asymptotic sense, ignoring the transient behavior. In fact, when the number of candidate controllers is large, these algorithms tend to take a fair amount of time to find an acceptable controller because the search is essentially blind. This often leads to long periods in which the closed-loop system is unstable and, consequentially, to unacceptable transients. In practice, pre-routed supervision is usually restricted to a small number of candidate controllers.

The switching algorithms that seem to be the most promising are those that evaluate online the potential performance of each candidate controller and use this to direct their search, which no longer follows a pre-computed order. These algorithms can roughly be divided into two categories: those based in process estimation, using either Certainty Equivalence [7, 57, 63, 59, 44, 67, 8, 34, 31, 33, 25, 1, 30, 68, 12, 36, 64, 28] or Model Validation [38, 15, 39, 89]; and those based on a direct performance evaluation of each candidate controller [74, 75, 73, 88, 40, 62]. Although these algorithms originate from fundamentally different approaches, they share key structures and exhibit important common properties. We proceed by presenting the basic types of supervisors that appear in the literature. We shall use the formalism introduced in Section 1.1, which is not always the one found in the papers referenced.

1.2.1 Estimator-based supervision

Estimator-based supervision was developed to overcome the poor performance that afflicts most pre-routed supervisors. Estimator-based supervisors continuously compare the behavior of the process with the behavior a several admissible process models to determine which model is more likely to describe the actual process. This model is regarded as an "estimate" of the actual process. From time to time, the supervisor then places in the loop that candidate controller that is more adequate for the estimated model. This is a manifestation of the well known Certainty Equivalence Principle. Some estimator-based supervisors follow a Model Validation paradigm towards identification [38, 15, 39, 89], instead of the "best-fit" type of estimation usually used in Certainty Equivalence. These identification algorithms are usually able to take unmodeled dynamics directly into account but their applicability is sometimes compromised by computational difficulties.

Typically, the family of admissible process models considered in estimator-based supervision is of the form

$$\mathcal{M} := \bigcup_{p \in \mathcal{P}} \mathcal{M}_p \tag{5}$$

where each $\mathcal{M}_p$ is a small set of admissible processes centered around a nominal model N_p. Typically, each N_p is a finite dimensional dynamical system modeled by an ordinary differential equation and each set $\mathcal{M}_p$ is of the general form

$$\mathcal{M}_p := \big\{ \mathrm{M}_p : \mathrm{d}(\mathrm{M}_p, \mathrm{N}_p) \leq \epsilon_p \big\}, \tag{6}$$

where d represents some metric defined on the set of process models, e.g., d could be defined using the $\mathcal{H}_\infty$ norm, the gap-metric [90, 16, 23], the Vinnicombe metric [86], or their generalizations to nonlinear systems. Families of admissible models of the type described by (5)–(6) allow one to take into account both parametric uncertainty and unmodeled dynamics.

For each admissible process model in $\mathcal{M}$, there must exist at least one candidate controller in $\mathcal{C} := \{\mathrm{C}_q : q \in \mathcal{Q}\}$ capable of providing satisfactory performance for that model. Usually, the same candidate controller provides satisfactory performance when connected to any process model in one of the sets $\mathcal{M}_p$. We can then define a *controller selection function* $\chi : \mathcal{P} \to \mathcal{Q}$ that maps each parameter value $p \in \mathcal{P}$ with the index $q = \chi(p) \in \mathcal{Q}$ of the controller C_q that provides satisfactory performance when connected to any process model in $\mathcal{M}_p$. In accordance with certainty equivalence, if at some point in time the process is believed to be in the family $\mathcal{M}_p$ for some $p \in \mathcal{P}$ then the controller C_q with $q := \chi(p)$ should be used.

An estimator-based supervisor can be represented by the diagram in Figure 2. This type of supervisor consists of a *multi-estimator* responsible for evaluating which admissible model best matches the process and a *decision logic* that generates σ and therefore effectively selects which candidate controller should be used. Typically, the *multi-estimator* is a dynamical system of the

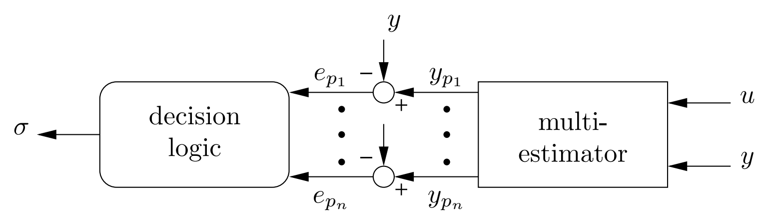

Figure 2: Estimator-based Supervisor

form

$$\dot{x}_E = A_E(x_E, u, y), \qquad y_p = C_E(p, x_E, u, y), \qquad e_p = y_p - y, \qquad p \in \mathcal{P}, \tag{7}$$

whose inputs are the signals that can be measured (in this case u and y) and whose outputs are the *estimation errors*, e_p, $p \in \mathcal{P}$. A multi-estimator is designed according to the general principle that if the actual process belongs to the ball $\mathcal{M}_p$, $p \in \mathcal{P}$, then the corresponding output estimate y_p should match the process output y and therefore the estimation error e_p should be small. Each e_p can therefore be regarded as a measure of the likelihood that the actual process is inside the ball $\mathcal{M}_p$. Multi-estimators can be designed using Observer Theory [72] or identification filters from Adaptive Control [70].

The *decision logic*, essentially compares the several estimation errors and, when a particular error e_p, $p \in \mathcal{P}$, is small, it places in the feedback loop the corresponding candidate controller C_q, $q := \chi(p)$. This is motivated by the fact that if e_p is small then the actual process model is likely to be in $\mathcal{M}_p$ and therefore the candidate controller C_q, $q := \chi(p)$ should perform satisfactory. Although intuitive, this reasoning cannot be used to carry out a formal analysis. This is because, smallness of e_p is generally not sufficient to guarantee that the actual process is in $\mathcal{M}_p$ [1]. It turns out that smallness of e_p is sufficient to guarantee that the candidate controller C_q, $q := \chi(p)$ will do a good job at controlling the actual process. It was showed by Morse [55] for the linear case that, while the controller C_q, $q := \chi(p)$ is in the feedback loop (i.e., while $\sigma = \chi(p)$), the system formed by the process, multi-controller, and multi-estimator is detectable through the estimation error e_p. This means that smallness of e_p is indeed sufficient to guarantee that the state of the overall system remains well behaved. Hespanha [24], Hespanha and Morse [31] extended this result to the

[1] Using persistence of excitation [70], it is possible to conclude from smallness of e_p that the actual process model is close to N_p. However, this usually requires tracking of specific reference signals or the addition of dither noise, which make this approach unattractive.

nonlinear case. This was achieved by making use of the recently introduced notion of detectability for nonlinear systems [56, 84].

The decision logic in Figure 2 will generally only need to compare the estimation errors. This can be done directly using the state x_E of the multi-estimator and therefore the input to the decision logic only needs to contain the three signals x_E, u, and y, *regardless of how large* $\mathcal{P}$ *is.* In fact, estimator-based supervision can be used even when $\mathcal{P}$ is an infinite set, in which case the diagram in Figure 2 is purely conceptual. This issue will be discussed in detail later.

1.2.2 Performance-based supervision

Performance-based supervision is characterized by the fact that the supervisor attempts to assess directly the potential performance of every candidate controller, without trying to estimate the model of the process [74, 75, 11, 73, 76, 62]. To achieve this, the supervisor computes *performance signals* π_q, $q \in \mathcal{Q}$, that provide a measure of how well the controller C_q would perform in a conceptual experiment in which the actual control signal u would be generated by C_q as a response to the measured process output y. This conceptual experiment is usually formulated imagining that the controller C_q is being used to achieve a control objective that would make u the response to y (e.g., trying to track a particular reference signal). This type of supervision is inspired by the idea of controller falsification introduced in [74]. When a particular performance signal π_q, $q \in \mathcal{Q}$, is large we know that the controller C_q would behave poorly for a particular pair of signals u, y. A supervisor should then avoid using such a controller because it has demonstrated poor performance under the hypothetical conditions of the virtual experiment. In performance-based supervision, the supervisor then only keeps in the feedback loop candidate controllers for which the corresponding performance signals are small. We refer the reader to [73] for an in-depth review of the controller falsification paradigm.

Figure 3 shows the block diagram of a performance-based supervisor. This type of supervisor consists of a *performance monitor* that generates the performance signals π_q, $q \in \mathcal{Q}$, together with a *decision logic* that generates the switching signal σ. The performance monitor is generally a dynamical system that resembles the multi-estimator in (7).

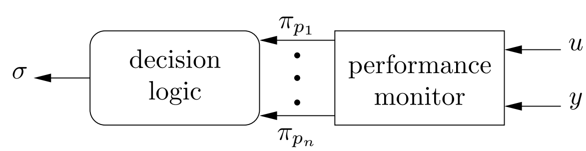

Figure 3: Performance-based Supervisor

1.3 Abstract supervision

Both estimator-based and performance-based supervision share the same basic control architecture depicted in Figure 4. We will see next that it is convenient to abstract from the detailed implementation of each individual block in Figure 4 and take instead the diagram in this figure as an *abstract supervision problem*, without regard to whether the top/right block is a multi-estimator or a performance signal generator. In fact, the diagram in Figure 4 also generalizes to other types of supervision, e.g., pre-routed algorithms. However, and to avoid introducing additional notation, for now we assume that we have an estimator-based supervisor.

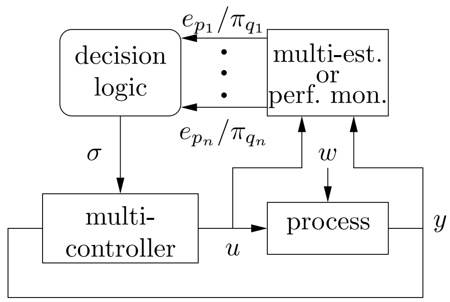

Figure 4: Common architecture to Estimator/Performance-based Supervision

1.3.1 The switched system

We start by focusing our attention on the aggregate dynamics of the process, multi-controller, and multi-estimator. The resulting system, called the *switched system*, can be modeled by a differential equation of the form

$$\dot{x} = A_\sigma(x, w), \tag{8}$$

where x denotes the aggregate state of the process, multi-controller, and multi-estimator; and w the exogenous disturbance/measurement noise. From the perspective of the decision logic, the outputs of this system are the estimation errors that can be generated by

$$e_p = C_p(x, w), \qquad p \in \mathcal{P}. \tag{9}$$

The family of functions $\{A_q : q \in \mathcal{Q}\}$ that define the dynamics of the switched system in (8) and the output functions $\{C_p : p \in \mathcal{P}\}$ in (9), can be easily constructed using the models of each subsystem. The switched system has a few basic properties that are crucial for the understanding of the overall system: the Matching Property and the Detectability Property. The former is essentially a property of the multi-estimator, whereas the latter is a property of the multi-controller. We proceed to qualitatively describe these properties and defer a formal presentation for later.

Matching Property The *Matching Property* refers to the fact that the multi-estimator should be designed so that each particular y_p provides a “good” approximation to the process output y—and therefore e_p is “small”—whenever the actual process model is inside the corresponding $\mathcal{M}_p$. Since the process is assumed to match one of the models in the set (5), we should then expect at least one of the estimation errors, say e_{p^*}, to be small in some sense. For example, we may require that in the absence of unmodeled dynamics, noise, and disturbances e_{p^*} converge to zero exponentially fast for every control input u. It is also desirable to have an explicit characterization of e_{p^*} in the presence of unmodeled dynamics, noise, and disturbances. For linear systems, a multi-estimator satisfying such requirements can be obtained as explained in [57]. In that paper it is also shown how the multi-estimator can be constructed in a *state-shared* fashion (so that it is finite-dimensional even if $\mathcal{P}$ is infinite), using standard results from realization theory. Multi-estimators with similar properties can also be designed for some useful classes of nonlinear systems, as discussed in [31]. State-sharing is always possible if the parameters enter the process model “separably” (but not necessarily linearly).

Detectability Property The *Detectability Property* that we impose on the candidate controllers is that for every fixed $p \in \mathcal{P}$, the switched system (8)–(9) must be detectable with respect to the estimation error e_p when the value of the switching signal is frozen at $\chi(p) \in \mathcal{Q}$. Adopted to the present context, the results proved in [55] imply that in the linear case the Detectability Property holds if the controller asymptotically stabilizes the multi-estimator and the process is detectable ("Certainty Equivalence Stabilization Theorem"), or if the controller asymptotically output-stabilizes the multi-estimator and the process is minimum-phase ("Certainty Equivalence *Output* Stabilization Theorem"). These conditions are useful because they decouple the properties that need to be satisfied by the parts of the system constructed by the designer from the properties of the unknown process. Extensions of these results to nonlinear systems are discussed in [31, 28, 48]. In particular, it is shown in [31] that detectability, defined in a suitable way for nonlinear systems, is guaranteed if the process is detectable and the controller *input-to-state stabilizes* the multi-estimator with respect to the estimation error (in the sense of Sontag [79]). The design of candidate controllers is thereby reduced to a disturbance attenuation problem well studied in the nonlinear control literature. The paper [28] develops an integral variant of this result, and the recent work [48] contains a nonlinear version of the certainty equivalence output stabilization theorem.

1.3.2 The switching logic

The index σ of the controller in the feedback loop is determined by the switching logic, whose inputs are the estimation errors e_p, $p \in \mathcal{P}$. In accordance to certainty equivalence, when a particular output estimation error e_p, $p \in \mathcal{P}$ is the smallest—and therefore p seems to be the most likely value for the parameter—the logic should select $\sigma = \chi(p) \in \mathcal{Q}$. To prevent chattering, one approximates this mechanism by introducing a *dwell-time* [57] or *hysteresis* [24, 31, 30, 46]. Since the value p that corresponds to the smallest e_p varies, it is convenient to introduce a *process switching signal* $\rho : [0, \infty) \to \mathcal{P}$ that for each time t indicates the current estimate $\rho(t) \in \mathcal{P}$ of the index p of the family $\mathcal{M}_p$ where the actual process lies. Typically, $\sigma = \chi(\rho)$. Note that the actual output of the switching logic is the switching signal σ that determines which candidate controller should be used. Process switching signals are often just internal variables to the logic or simply "virtual" signals used in the analysis[2]. Two properties need to be satisfied by the switching logic and the monitoring signal generator: the Non-Destabilization Property and the Small Error Property.

Small Error Property The *Small Error Property* calls for a bound on e_ρ in terms of the smallest of the signals e_p, $p \in \mathcal{P}$ for a process switching signal ρ for which $\sigma = \chi(\rho)$. For example, if $\mathcal{P}$ is a finite set and the monitoring signals are defined as $\mu_p(t) = \int_0^t e_p^2(s)ds$, then the scale-independent hysteresis switching logic of [24] guarantees that for every $p \in \mathcal{P}$,

$$\int_0^t e_\rho^2(s)ds \le C \int_0^t e_p^2(s)ds \tag{10}$$

where C is a constant (which depends on the number of controllers and the hysteresis parameter) and the integral on the left is to be interpreted as the sum of integrals over intervals on which ρ is constant. If e_{p^*} decays exponentially as discussed earlier, then (10) guarantees that the signal e_ρ is in $\mathcal{L}_2$. At the heart of the switching logic, there is a conflict between the desire to switch to the smallest estimation error to satisfy the Small Error Property and the concern that too much switching may violate the Non-Destabilization Property.

[2]We see later shortly that in some cases $\sigma = \chi(\rho)$ does not uniquely define ρ and we can utilize this degree of freedom in the definition of ρ to simplify the analysis.

Non-Destabilization Property Recall that, in view of the Detectability Property, for every *fixed* value of $\sigma = \chi(p)$, $p \in \mathcal{P}$ the system (8)–(9) is detectable with respect to the corresponding estimation error e_p. The switching signal σ is said to have the *Non-Destabilization Property* if it preserves the detectability in a time-varying sense, i.e., if the switched system (8)–(9) is detectable with respect to the switched output e_ρ, for a process switching signal ρ for which $\sigma = \chi(\rho)$. The Non-Destabilization Property trivially holds if the switching stops in finite time (which is the case if the scale-independent hysteresis switching logic of [24] or its variants proposed in [30, 46] are applied in the absence of noise, disturbances, and unmodeled dynamics). In the linear case, a standard output injection argument shows that detectability is not destroyed by switching if the switching is sufficiently slow (so as not to destabilize the injected switched system). According to the results of [32], it actually suffices to require that the switching be *slow on the average*. However, it should be noted that the Non-Destabilization Property does not necessarily amount to a slow switching condition; for example, the switching can be fast if the systems being switched are in some sense "close" to each other. In [57, Section VIII] one can find another fast switching result that exploits the structure of linear multi-controllers and multi-estimators.

1.3.3 Putting it all together

We now briefly explain how the above properties of the various blocks of the supervisory control system can be put together to analyze its behavior. For simplicity, let us assume that the process, as well as the multi-controller and the multi-estimator, are linear systems and that there are no disturbances or noise. Then the switched closed-loop system can be represented in the form

$$\dot{x} = A_\sigma x, \qquad e_p = C_p x, \qquad p \in \mathcal{P}. \tag{11}$$

In view of the Detectability Property, (C_p, A_q), $q := \chi(p)$ is a detectable pair for each $p \in \mathcal{P}$. Choosing output injection matrices K_p, $p \in \mathcal{P}$ such that $A_q - K_p C_p$ is stable for all p, we can rewrite the dynamics of (11) as

$$\dot{x} = (A_\sigma - K_\rho C_\rho)x + K_\rho e_\rho,$$

for any process switching signal ρ. Because of the Matching Property, there exists some $p^* \in \mathcal{P}$ for which e_{p^*} is small (e.g., converges to zero exponentially fast as above). This together with the Small error Property guarantees that e_ρ is small in an appropriate sense (e.g., an $\mathcal{L}_2$ signal) for some process switching signal ρ for which $\sigma = \chi(\rho)$. To establish boundedness of the overall system all that remains to be verified is that the switched system

$$\dot{x} = (A_\sigma - K_\rho C_\rho)x$$

is asymptotically stable. In view of stability of the individual matrices $A_q - K_p C_p$ with $q := \chi(p)$, this is guaranteed if σ has the additional Non-destabilization Property: For example, if the switching stops in finite time or is *slow on the average* in the sense of [32]. Switching signals produced by the dwell-time switching logic [57, 59] or by the scale-independent hysteresis switching logic of [24] and its variants proposed in [30, 26] are known to possess these desired properties. Proceeding in this fashion, it is possible to analyze stability and robustness of supervisory control algorithms for quite general classes of uncertain systems [57, 59, 32, 30, 26]

Not surprisingly, the four properties that were just introduced for supervisory control have direct counterparts in classical adaptive control. The Detectability Property was first recognized in the context of adaptive control in [54], where it was called *tunability*. The Matching Property is usually implicit in the derivation of the error model equations, where one assumes that, for a specific value of the parameter, the output estimate matches the true output. Both the Small Error Property and the Non-Destabilization Property are pertinent to the tuning algorithms, being typically stated in terms of the smallness (most often in the $\mathcal{L}_2$ sense) of the estimation error and the derivative of the parameters estimate, respectively.

2 Estimator-based linear supervisory control

Section Summary

In this section we specialize the estimator-based architecture of Section 1 to the case of a linear process, linear candidate controllers, and linear multi-estimator. In this context we summarize the results available and go through the main steps of the stability argument.

2.1 Class of admissible processes and candidate controllers

We assume that the uncertain process to be controlled admits the model of a finite-dimensional stabilizable and detectable linear system with control input u and measured output y, perturbed by a bounded disturbance input d and a bounded output noise signal n (cf. Figure 5). The disturbance/noise vector is then defined by $w := \begin{bmatrix} d & n \end{bmatrix}'$. It is assumed known that the process transfer

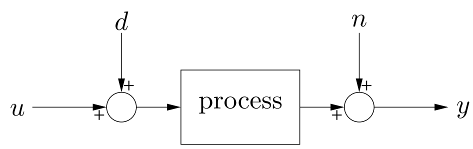

Figure 5: Process

function from u to y belongs to a family of admissible process model transfer functions

$$\mathcal{M} := \bigcup_{p \in \mathcal{P}} \mathcal{M}_p$$

where p is a parameter taking values in some index set $\mathcal{P}$. Here, for each p, $\mathcal{M}_p$ denotes a family of transfer functions "centered" around some known *nominal* process model transfer function ν_p (cf. below). Throughout the paper, we will take $\mathcal{P}$ to be a compact subset of a finite-dimensional normed linear vector space.

The problem of interest is to design a feedback controller that achieves output regulation, i.e., drives the process output y to zero, whenever the noise and disturbance signals are zero. Moreover, all system signals must remain (uniformly) bounded in response to arbitrary bounded noise and disturbance inputs. Everything that follows can be readily extended to the more general problem of set-point control (i.e., tracking an arbitrary constant reference r) with the help of adding an integrator in the feedback loop, as in [57, 59]. Such a modification would not introduce any significant changes as far as the principal developments of this paper are concerned. Control algorithms of the type described here can also be applied to the problem of disturbance suppression [22].

The set $\mathcal{P}$ can be thought of as representing the range of parametric uncertainty, while for each fixed $p \in \mathcal{P}$ the subfamily $\mathcal{M}_p$ accounts for unmodeled dynamics. There are several ways of specifying allowable unmodeled dynamics around the nominal process model transfer functions ν_p. For example, take two arbitrary numbers $\epsilon > 0$ and $\lambda \geq 0$. Then we can define

$$\mathcal{M}_p := \{\nu_p(1 + \delta_m) + \delta_a : \|\delta_m\|_{\infty,\lambda} \leq \epsilon,\ \|\delta_a\|_{\infty,\lambda} \leq \epsilon\}, \qquad p \in \mathcal{P} \tag{12}$$

where $\|\cdot\|_{\infty,\lambda}$ denotes the $e^{\lambda t}$-weighted $\mathcal{H}_\infty$ norm of a transfer function: $\|\nu\|_{\infty,\lambda} = \sup_{\Re[s]\geq 0} |\nu(s - \lambda)|$. This yields the class of admissible process models treated in [57, 59] for the SISO case.

Alternatively, one can define $\mathcal{M}_p$ to be the ball of radius ϵ around ν_p with respect to the Vinnicombe metric [86]. Another possible definition for SISO processes is

$$\mathcal{M}_p := \Big\{\frac{n_p+\delta_n}{d_p+\delta_d} : \|\delta_n\|_{\infty,\lambda} \le \epsilon,\ \|\delta_d\|_{\infty,\lambda} \le \epsilon\Big\}, \qquad p \in \mathcal{P} \tag{13}$$

where $\nu_p = n_p/d_p$ is the normalized coprime factorization of ν_p (see, e.g., [91]). This is more general than (12) in that it allows for uncertainty about the pole locations of the nominal process model transfer functions. In the sequel, allowable unmodeled dynamics are assumed to be specified in either one of the aforementioned ways. We will refer to the positive parameter ϵ as the *unmodeled dynamics bound*.

Modeling uncertainty of the kind described above may be associated with unpredictable changes in operating environment, component failure, or various external influences. Typically, no single controller is capable of solving the regulation problem for the entire family of admissible process models. Therefore, one needs to develop a controller whose dynamics can change on the basis of available real-time data. Within the framework of supervisory control discussed here, this task is carried out by the *supervisor*, whose purpose is to orchestrate the switching among a parameterized family of *candidate controller* transfer functions

$$\mathcal{C} := \{\kappa_q : q \in \mathcal{Q}\}, \tag{14}$$

where $\mathcal{Q}$ is an index set. We require this controller family to be sufficiently rich so that every admissible process model can be stabilized by placing it in the feedback loop with some controller in $\mathcal{C}$. In particular, we assume that there exists a *controller selection function* $\chi : \mathcal{P} \to \mathcal{Q}$ that maps each parameter value $p \in \mathcal{P}$ with the index $q = \chi(p) \in \mathcal{Q}$ of the controller κ_q that stabilizes the nominal process model transfer function ν_p as well as all transfer functions in the family $\mathcal{M}_p$ "centered" at ν_p. In accordance with certainty equivalence, if at some point in time the process is believed to be in the family $\mathcal{M}_p$ for some $p \in \mathcal{P}$ then the controller κ_q with $q := \chi(p)$ should be used.

2.2 Multi-estimator and multi-controller

We utilize here a state-shared multi-estimator of the form

$$\dot{x}_E = A_E x_E + D_E y + B_E u, \qquad y_p = C_p x_E, \qquad e_p = y_p - y, \qquad p \in \mathcal{P} \tag{15}$$

with A_E an asymptotically stable matrix. This type of structure is quite common in adaptive control. Note that even if $\mathcal{P}$ is an infinite set, the above dynamical system is finite-dimensional. In this case the multi-estimator formally has an infinite number of outputs, however they can all be computed from x_E.

The key property of the of multi-estimator is the *Matching Property*, which refers to the fact that if the process transfer function is within a particular family $\mathcal{M}_{p^*}$, $p^* \in \mathcal{P}$ then the corresponding output estimate y_{p^*} should be close to the process out y and therefore e_{p^*} should be small. It turns out that it is always possible to design state-shared multi-estimators for linear systems with such a property (cf. Appendix A). Formally, the Matching property can be stated as follows:

Property 1 (Matching). There exist positive constants $c_0, c_w, c_\epsilon, \lambda$ and some $p^* \in \mathcal{P}$ such that

$$\|e_{p^*}\|_{\lambda,[0,t)} \le c_0 + c_w\|w\|_{\lambda,[0,t)} + \epsilon\, c_\epsilon \|u\|_{\lambda,[0,t)}, \qquad \forall t \ge 0.$$

In the above property, $\|\cdot\|_{\lambda,[0,t)}$ denotes the $e^{\lambda t}$-weighted $\mathcal{L}_2$-norm truncated to the interval $[0,t)$, i.e., given a signal v $\|v\|_{\lambda,[0,t)} = \left(\int_0^t e^{2\lambda\tau} v(\tau)^2 d\tau\right)^{\frac{1}{2}}$. We will denote by $\mathcal{L}_2(\lambda)$ the set of all signals that have finite $e^{\lambda t}$-weighted $\mathcal{L}_2$-norm on $[0,\infty)$.

To build the multi-controller, we start by constructing a family $\{(F_q, G_q, H_q, J_q) : q \in \mathcal{Q}\}$ of n_C-dimensional stabilizable and detectable realizations for the candidate controllers, with the understanding that (F_q, G_q, H_q, J_q) is a realization for the controller κ_q. The multi-controller is then defined by

$$\dot{x}_C = F_\sigma x_C + G_\sigma y, \qquad\qquad u = H_\sigma x_C + J_\sigma y.$$

And therefore, when $\sigma = q$ we effectively have the controller κ_q in feedback with the process. As mentioned before the main requirement on the multi-controller is that it satisfy the Detectability property. The best way to understand what this amounts to passes through the introduction of what is called the "injected systems" that we introduce next.

2.3 The injected system

We recall that we called the aggregate system consisting of the process, multi-estimator, and multi-controller the *switched system*. We propose now to actually regard the switched system as the feedback interconnection of two subsystem: the process and the "injected system." Formally, this can be done as follows:

1. Take a piecewise constant *process switching signal* $\rho : [0,\infty) \to \mathcal{P}$. It is useful to think of $\rho(t)$ as the estimate (at time t) of the parameter value $p^* \in \mathcal{P}$ that indexes the family $\mathcal{M}_{p^*}$ where the process lies.

2. Define the signal

$$v(t) := e_{\rho(t)}(t) = y_{\rho(t)}(t) - y(t), \qquad\qquad t \geq 0. \tag{16}$$

3. Replace y in the equations of the multi-estimator and multi-controller by $y_\rho - v$. The resulting system (with state $x := \begin{bmatrix} x_E' & x_C' \end{bmatrix}'$) is called the *injected system* and has input v and outputs u and all the y_p, $p \in \mathcal{P}$. The name "injected" comes from the fact that to construct it we inject the output y_ρ of the multi-estimator back into its input y.

We can then regard the switched system as the interconnection of two system: the process and the injected system, with the interconnection defined by (16). (cf. Figure 6).

The state-space model of the injected system is of the form

$$\dot{x} = A_{\rho\sigma} x + B_\sigma v, \qquad u = F_{\rho\sigma} x + G_\sigma v, \qquad y_p = C_p x, \qquad p \in \mathcal{P}, \tag{17}$$

for appropriately defined matrices $A_{pq}, B_q, F_{pq}, G_q, C_p$, $p \in \mathcal{P}$, $q \in \mathcal{Q}$. By writing A_{pq} explicitly, one can see by inspection that the eigenvalues of this matrix are precisely the poles of the feedback interconnection of the nominal process ν_p with the controller κ_q, together with some of the (stable) eigenvalues of A_E and any (stable) eigenvalues of κ_q's realization (F_q, G_q, H_q, J_q) that are not observable or controllable. To verify this, one uses the fact that $(A_E + D_E C_p, B_E, C_p)$ is a stabilizable and detectable realization of ν_p, which is a necessary condition for the matching property to hold.

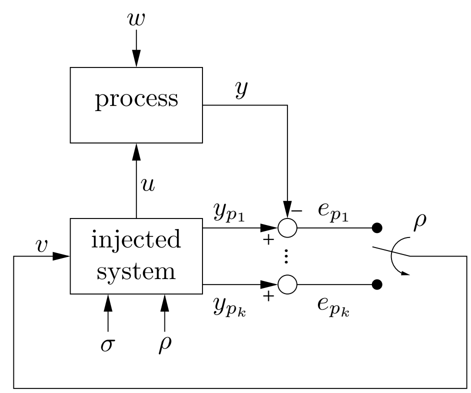

Figure 6: The switched system as the interconnection of the process with the injected system.

An immediate consequence of the above is that when $\sigma = q \in \mathcal{Q}$, $\rho = p \in \mathcal{P}$, and the candidate controller κ_q stabilizes the nominal process model ν_p, the injected system is asymptotically stable. Therefore, if $v := e_p$ converges to zero, then so does the state and all the outputs of the injected system. In particular, u and y_p. But then $y = y_p - v$ also converges to zero. Since it has been established that both the input u and output y of the process converge to zero, its internal state must also converge to zero (assuming the process is detectable). This argument proves that the switched system is detectable. This is because only for detectable (linear) systems it is possible to argue that its state converge must converge to zero when it is observed that its output converges to zero.

Property 2 (Detectability). Let $p \in \mathcal{P}$ and $q \in \mathcal{Q}$ be such that the candidate controller κ_q stabilizes the nominal process model ν_p (e.g., $q = \chi(p)$ where χ denotes the controller selection function defined before). Then, for $\sigma = q \in \mathcal{Q}$ and $\rho = p \in \mathcal{P}$, the injected system is asymptotically stable and the switched system is detectable through the output e_p.

This result is also known as the Certainty Equivalence Stabilization Theorem [54]. The connection with certainty equivalence stems from the fact that if the estimation error e_p is small—and therefore it seems reasonable to assume that the process is in $\mathcal{M}_p$—we actually achieve detectability through e_p by using the controller κ_q that stabilizes ν_p. Because of detectability, smallness of e_p will then results in smallness of the overall state of the switched system *whether or not the process is in* $\mathcal{M}_p$.

Although stability of the injected system is a simple mechanism to obtain detectability of the switched system, it is not the only mechanism. For example, it is possible to show that output-stability of the injected system together with minimum-phase of the process are also sufficient to prove that the switched system is detectable. This is known as the Certainty Equivalence *Output* Stabilization Theorem [54]. Cyclic switching is another mechanism to achieve detectability [66, 69].

2.4 Dwell-time switching logic

The decomposition of the switched system shown in Figure 6 provides insight into the challenges in designing the switching logic that generates σ:

1. To achieve stability one wants $v := e_\rho$ to be small. The simplest way to achieve this is to select ρ to be the index in $\mathcal{P}$ for which e_p is smallest.

2. However, smallness of v is only useful if the injected system is stable. To achieve this one wants $\sigma = \chi(\rho)$ to make sure that the injected system is stable. However, this only guarantees that $A_{\rho(t)\sigma(t)}$ is a stability matrix for every time T and not that the time-varying system is exponentially stable. In fact, it is well known that switching among stable matrices can easily result in an unstable system [9, 47]. To avoid a possible loss of stability caused by switching one should then require the switching logic to prevent "too much" switching. Unfortunately, this may conflict with the requirement that $v := e_\rho$.

The items (1) and (2) above directly motivate the Small Error Property and the Non-destabilization Property, respectively.

The *dwell-time switching logic* resolved the previous conflict by select ρ to be the index in $\mathcal{P}$ for which e_p is smallest, but "dwelling" on this particular choice for ρ and $\sigma = \chi(\rho)$ for at least a pre-specified amount of time τ_D called the *dwell-time* [57, 59]. Figure 7 shows a simplified version of this logic that we will utilize here. In this figure, the signals μ_p, $p \in \mathcal{P}$ are called the *monitoring*

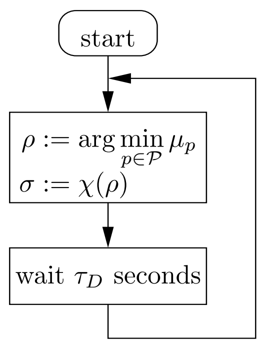

Figure 7: Dwell-time switching logic

signals and are defined by

$$\mu_p(t) := \int_0^t e^{-2\lambda(t-\tau)} \|e_p(\tau)\|^2 d\tau, \qquad p \in \mathcal{P}, \tag{18}$$

where λ denotes a non-negative constant. For convenience, we make it the same as in (12) but we could also take a strictly smaller value than that in (12). The reason for this will become clear later. The monitoring signals should be viewed as measures of the size of the estimation errors over a window whose length is defined by the forgetting factor λ. Thus smallness of a monitoring signal μ_p, $p \in \mathcal{P}$ means that the corresponding estimation error e_p has been small for some time interval (on the order of $1/\lambda$ seconds). We opted here for an $\mathcal{L}_2$-type norm to measure the estimation errors, but other norms would be possible. In fact, this extra flexibility will be needed for nonlinear systems.

Note that the process switching signal ρ defined by the logic is actually not used by the supervision algorithm, as only σ is used by the multi-controller. The signal ρ is only used in the analysis to define the injected system and does not actually need to be generated explicitly by the logic. In fact, one can viewed ρ as a degree-of-freedom to be used in constructing a stability prove. However, in the definition of the dwell-time switching logic in Figure 7 we are somewhat getting ahead of ourselves and already specifying the process switching signal ρ that "makes sense" to use in light of the Detectability Property 2. However, this is certainly not the only choice and in fact we will shortly be forced to use a slightly different process switching signal.

2.4.1 Dwell-time switching properties

By construction, the dwell-time switching logic guarantees that the interval between two consecutive discontinuities of σ. We state this formally for later reference.

Property 3 (Non-destabilization). The minimum interval between consecutive discontinuities of any switching signal σ generated by the dwell-time switching logic is equal to $\tau_D > 0$.

The Small Error Property for the dwell-time switching logic is less trivial and we will state two version of it. For simplicity, we start by considering the case of a finite set $\mathcal{P}$. Let us start by assuming that one of the estimation errors $e_{p^*} \in \mathcal{L}_2(\lambda)$, i.e.,

$$\|e_{p^*}\|^2_{[0,\infty)} = \int_0^\infty e^{2\lambda\tau}\|e_{p^*}(\tau)\|^2 d\tau \leq C^* < \infty, \tag{19}$$

(for example because we are ignoring noise and unmodeled dynamics, c.f the Matching Property 1). This means that $e^{2\lambda t}\mu_{p^*}(t) \leq C^*$, $t \geq 0$ and therefore, whenever ρ is selected to be equal to some $p \in \mathcal{P}$ at time t, we must have

$$\int_0^t e^{2\lambda\tau}\|e_p(\tau)\|^2 d\tau = e^{\lambda t}\mu_p(t) \leq e^{\lambda t}\mu_{p^*}(t) \leq C^*. \tag{20}$$

Two options are then possible:

1. Switching will stop in finite time T at some value $p \in \mathcal{P}$ for which (20) holds for $t \geq T$. In this case,

$$\int_0^\infty e^{2\lambda\tau}\|e_{\rho(\tau)}(\tau)\|^2 d\tau = \int_0^T e^{2\lambda\tau}\|e_{\rho(\tau)}(\tau)\|^2 d\tau + \int_T^\infty e^{2\lambda\tau}\|e_p(\tau)\|^2 d\tau < \infty.$$

2. Switching will not stop, but after some finite time T it must only occur among elements of a subset $\mathcal{P}^*$ of $\mathcal{P}$, each appearing in ρ infinitely many times. Therefore (20) holds for all elements of $\mathcal{P}^*$ and we have

$$\int_0^\infty e^{2\lambda\tau}\|e_{\rho(\tau)}(\tau)\|^2 d\tau = \int_0^T e^{2\lambda\tau}\|e_{\rho(\tau)}(\tau)\|^2 d\tau + \sum_{p\in\mathcal{P}^*}\int_T^\infty e^{2\lambda\tau}\|e_p(\tau)\|^2 d\tau < \infty.$$

   The last inequality requires finiteness of $\mathcal{P}^*$.

The following can then be stated:

Property 4 (Small Error—$\mathcal{L}_2$ case). Assume that $\mathcal{P}$ is a finite set. If there exists some $p^* \in \mathcal{P}$ for which $e_{p^*} \in \mathcal{L}_2(\lambda)$ then $e_\rho \in \mathcal{L}_2(\lambda)$, i.e.,

$$\int_0^\infty e^{2\lambda\tau}\|e_{\rho(\tau)}(\tau)\|^2 d\tau < \infty. \tag{21}$$

It turns out that, even when no $p^* \in \mathcal{P}$ is necessarily $\mathcal{L}_2(\lambda)$, it is possible to prove that a suitable Small Error Property holds. The finiteness of $\mathcal{P}$ is also not needed and can be replaced by finiteness of the set of candidate controllers:

Property 5 (Small Error—general case). Assume that $\mathcal{Q}$ is a finite set with m elements. For every $t \geq 0$ there exists a process switching signal $\rho_t : [0, t) \to \mathcal{P}$, such that $\sigma = \chi(\rho_t)$ except for at most m time intervals of length τ_D, such that

$$\int_0^t e^{2\lambda\tau} \|e_{\rho_t(\tau)}(\tau)\|^2 d\tau \leq m \int_0^t e^{2\lambda\tau} \|e_p(\tau)\|^2 d\tau, \qquad \forall p \in \mathcal{P}. \tag{22}$$

The process switching signal $\rho_t : [0, t) \to \mathcal{P}$ can be constructed as follows: For each $q \in \mathcal{Q}$, let $[\tau_q, \tau_q + \tau_D]$ be the last interval on $[0, t)$ on which σ is equal to q and set $\rho_t(\tau) = \rho(\tau_q)$ for any time $\tau < \tau_q$ for which $\sigma(\tau) = q = \chi(\rho(\tau_q)) = \chi(\rho_t(\tau))$ and $\rho_t(\tau) = p$ for $\tau \in [\tau_q, \tau_q + \tau_D]$. With this construction $\sigma = \chi(\rho_t)$ except for at most m intervals of length τ_D. We leave to the reader the proof that (22) holds with ρ_t defined in this manner. This proof is essentially done in [60].

The Small Error Property can still be generalized to infinite sets of candidate controllers, under suitable compactness assumptions [59, Lemma 5].

2.4.2 Implementation Issues

Before proceeding some discussion is needed regarding the implementation of the switching logic, especially when $\mathcal{P}$ is an infinite set and therefore the generation of the monitoring signals in (18) seems to require an infinite dimensional system. It turns out that the μ_p, $p \in \mathcal{P}$ can be efficiently generated by a finite-dimensional system. To see why, note that it is always possible to write

$$\|e_p\|^2 = \|C_p x_E - y\|^2 = k(p)' h(y, x_E), \qquad \forall p \in \mathcal{P},\ y,\ x_E$$

where $k(p)$ and $h(y, x_E)$ are appropriately defined vector functions. The monitoring signals can then be generated by

$$\dot{x}_\mu = -2\lambda x_\mu + h(y, x_E), \quad x_\mu(0) = 0, \qquad \mu_p = k(p) x_\mu, \quad p \in \mathcal{P}.$$

This can be checked by verifying that the μ_p so defined satisfy the differential equation $\dot{\mu}_p = -2\lambda\mu_p + \|e_p\|^2$, $\mu_p(0) = 0$, whose solution is given by (18) with e_p as in (15). This also means that the generation of ρ in the middle box of diagram in Figure 7 can be written as

$$\rho := \arg\min_{p \in \mathcal{P}} k(p) x_\mu \tag{23}$$

and is, in fact, an optimization over the elements of $\mathcal{P}$. This means that we never actually need to explicitly compute all the estimation errors e_p or the monitoring signals μ_p, as long as we know how to solve the optimization problem (23). This is certainly true when C_p is linear on the parameter p and therefore $k(p)$ is quadratic on p, in which case closed form solutions can often be found. It is interesting to point-out that most traditional adaptive control algorithms can only address this case. The dwell-time logic, however, can still be efficiently implemented when this is not the case but the optimization (23) is tractable. This happens when there is a closed-form solution or when there are efficient numerical solution (e.g., due to convexity). These issues are further discussed in [27].

It is also worth noticing that the Small Error Properties above still hold if the optimization in (23) is not instantaneous and takes some computation time $\tau_C > 0$, i.e., if

$$\rho(t) := \arg\min_{p\in\mathcal{P}} k(p) x_\mu(t-\tau_C).$$

The only change needed in Property (5) is that now $\sigma = \chi(\rho_t)$ except for at most m intervals of length $\tau_D + \tau_C$.

2.4.3 Slow switching

In this section we provide the necessary details to prove the stability of the supervisory control closed-loop system. We will do this under the simplifying assumption that the dwell-time constant τ_D is large:

Assumption 1 (Slow switching). The dwell-time τ_D and the forgetting factor λ are chosen so that there exist constants $c > 0$ and $\bar{\lambda} > \lambda$ for which, for every process switching signal $\bar{\rho}$ with interval between consecutive discontinuities no smaller than τ_D,

$$\|\Phi_{\bar{\rho}}(t,\tau)\| \leq c e^{-\bar{\lambda}(t-\tau)}, \qquad t \geq \tau \geq 0, \tag{24}$$

where $\Phi_{\bar{\rho}}$ denotes the state transition matrix of the time varying system $\dot{z} = A_{\bar{\rho}\bar{\sigma}} z$, $\bar{\sigma} := \chi(\bar{\rho})$.

Since for every fixed time $t \geq 0$ the controller $\kappa_{\bar{\sigma}(t)}$ stabilizes the nominal process model $\nu_{\bar{\rho}(t)}$ and therefore the matrix $A_{\bar{\rho}(t)\bar{\sigma}(t)}$ is asymptotically stable. By choosing λ sufficiently small so that all matrices $A_{p\chi(p)} + \lambda I$, $p \in \mathcal{P}$ are asymptotically stable, it is then possible to make sure that (24) holds by selecting τ_D sufficiently large (cf. [32]). We will see later that we can actually prove stability for *any* arbitrarily small value of τ_D.

We start by considering the case in which there is no unmodeled dynamics and no noise, i.e., when $\epsilon = 0$ and $w(t) = 0$, $t \geq 0$. From the Matching Property 1 and the Small Error Property 4, we then conclude that there is some $p^* \in \mathcal{P}$ for which e_{p^*} is $e^{\lambda t}$-weighted $\mathcal{L}_2$ in the sense of (19) and e_ρ is also $e^{\lambda t}$-weighted $\mathcal{L}_2$, now in the sense of (21). This means that the state transition matrix of the injected system (17) decays to zero faster than $e^{-\lambda t}$ (cf. Assumption 1) and its input $v := e_\rho$ is $e^{\lambda t}$-weighted $\mathcal{L}_2$. From this we conclude immediately that the state x all the outputs of the injected system are also $e^{\lambda t}$-weighted $\mathcal{L}_2$ and even converge to zero. This is true, in particular, for u and y_ρ. But then $y := y_\rho - e_\rho$ is also $e^{\lambda t}$-weighted $\mathcal{L}_2$. Since the input and output of the process are $\mathcal{L}_2$ then its state must converge to zero (assuming the process is detectable). The following was proved:

Theorem 1. *Assuming that $\mathcal{P}$ is finite, that the process is detectable, and in the absence of noise and unmodeled dynamics (i.e., when $\epsilon = 0$ and $w(t) = 0$, $t \geq 0$), the states of the process, the multi-estimator, and the multi-controller are all $e^{\lambda t}$-weighted $\mathcal{L}_2$ and converge to zero as $t \to \infty$.*

We proceed now to consider the general case in which e_{p^*} is not known to be $\mathcal{L}_2$. In this case, we have to use the Small Error Property 5 instead of 4. To do this, let us focus out attention on an interval $[0,t)$, $t > 0$. We start by "cheating" and pretending that $\sigma = \chi(\rho_t)$ on $[0,t)$. In this case, Assumption 1 guarantees that the injected system (17) obtained using the process switching signal ρ_t has finite induced $\|\cdot\|_{\lambda,[0,t)}$-norm, i.e., that there exists a finite constant γ such that

$$\|u\|_{\lambda,[0,t)} \leq \gamma \|v\|_{\lambda,[0,t)} + \bar{c}_0, \tag{25}$$

where c_0 only depends on the initial conditions $x(0)$. Moreover, because of the Small Error Property 5 and the Matching Property 1 we also have that

$$\|e_{\rho_t}\|_{\lambda,[0,t)} \leq \sqrt{m}\|e_{p^*}\|_{\lambda,[0,t)} \leq c_0\sqrt{m} + c_w\sqrt{m}\|w\|_{\lambda,[0,t)} + \epsilon\, c_\epsilon\sqrt{m}\|u\|_{\lambda,[0,t)}, \tag{26}$$

But, since we constructed the injected system using the process switching signal ρ_t, $v := e_{\rho_t}$ and therefore we conclude from (26) and (25) that

$$\|v\|_{\lambda,[0,t)} \leq \frac{(c_0 + \epsilon\, c_\epsilon \bar{c}_0)\sqrt{m}}{1 - \epsilon\,\gamma\, c_\epsilon\sqrt{m}} + \frac{c_w\sqrt{m}}{1 - \epsilon\,\gamma\, c_\epsilon\sqrt{m}}\|w\|_{\lambda,[0,t)}, \tag{27}$$

assuming that

$$\epsilon < \frac{1}{\gamma\, c_\epsilon\sqrt{m}}.$$

From this bound it is then straightforward to conclude that there is a finite induced $\|\cdot\|_{\lambda,[0,t)}$-norm from w to any other signal; that all signals remain bounded, provided that $w(t)$ is uniformly bounded for $t \in [0,\infty)$; and that all signals converge to zero when $w(t) = 0$, $t \geq 0$.

The key insight to be taken from the reasoning above is that one can apply a small-gain argument to the switched system in Figure 6 by regarding the switch as a system with a finite induced $\|\cdot\|_{\lambda,[0,t)}$-norm specified by Small Error Property. It turns out that a similar argument can be made even if σ is not equal to $\chi(\rho_t)$ all over $[0,t)$.

For time instants $\tau \in [0,t)$ on which $\sigma(\tau) = \chi\big(\rho_t(\tau)\big)$, the matrix $A_{\rho_t(\tau)\sigma(\tau)}$ is asymptotically stable. However, when $\sigma(\tau) \neq \chi\big(\rho_t(\tau)\big)$, $A_{\rho_t(\tau)\sigma(\tau)}$ may be unstable. Fortunately, this will only occur for a union of time intervals with total length no larger than $m\tau_D$. Therefore the time-varying system $\dot{z} = A_{\rho_t\sigma}z$ is still exponentially stable and its state transition matrix $\Phi_{\rho_t\sigma}$ can still be bounded by an equation like (24), in fact it is straightforward to show that

$$\|\Phi_{\rho_t\sigma}(t,\tau)\| \leq c e^{a\,m\,\tau_D} e^{-\bar{\lambda}(t-\tau)}, \qquad t \geq \tau \geq 0, \tag{28}$$

where $a := \max_{p\in\mathcal{P},q\in\mathcal{Q}}\|A_{pq}\|$. We can therefore still use the argument above to establish the bound (27) but keeping in mind that γ must now be replaced by an upper-bound $\bar{\gamma}$ on the induced $\|\cdot\|_{\lambda,[0,t)}$-norm from v to u of the injected system obtained using the process switching signal ρ_t. Such an upper bound can be easily derived from (28) and (17):

$$\bar{\gamma} := \frac{c e^{a\,m\,\tau_D}}{\bar{\lambda} - \lambda}\Big(\max_{p\in\mathcal{P},q\in\mathcal{Q}}\|F_{pq}\|.\|B_q\|\Big) + \max_{q\in\mathcal{Q}}\|G_q\| \tag{29}$$

[32]. This leads to the following result:

Theorem 2. *Assuming that the set $\mathcal{Q}$ is finite, that the process is detectable, that assumption 1 holds, and that*

$$\epsilon < \frac{1}{\bar{\gamma}\, c_\epsilon\sqrt{m}},$$

the $\|\cdot\|_{\lambda,[0,t)}$-norm of the state of the multi-estimator, the state of the multi-controller, and of the input and output of the process can all be bounded by expressions of the form

$$\bar{c}_0 + \bar{c}_w\|w\|_{\lambda,[0,t)},$$

where $\bar{c}_0$ and $\bar{c}_w$ are finite constants, with $\bar{c}_0$ depending on initial conditions and $\bar{c}_w$ not. Moreover, all signals remain bounded, provided that $w(t)$ is uniformly bounded for $t \in [0,\infty)$; and that all signals converge to zero as $t \to \infty$ when $w(t) = 0$, $t \geq 0$.

The reader is referred to [59] for a generalization of Theorem 2 to infinite sets of candidate controllers.

2.4.4 Fast switching

As mentioned before Assumption 1 can be made significantly less restrictive and, in particular the dwell-time τ_D can be made arbitrarily small without compromising stability. In fact, we can replace this assumption simply the following one.

Assumption 2 (Fast switching). The forgetting factor λ is chosen so that all matrices $A_{p\chi(p)}+\lambda I$, $p \in \mathcal{P}$ are asymptotically stable.

The following was then proved in [59]:

Theorem 3. *Assuming that the process is SISO, that assumption 2 holds, and that the multi-controller is of the form*

$$\dot{x}_C = (A_C + d_C f_\sigma)x_C + b_C y, \qquad u = f_\sigma x_C + j_\sigma y,$$

there exists a constant ϵ^* such that when

$$\epsilon < \epsilon^*,$$

the $\|\cdot\|_{\lambda,[0,t)}$-norm of the state of the multi-estimator, the state of the multi-controller, and of the input and output of the process can all be bounded by expressions of the form

$$\bar{c}_0 + \bar{c}_w \|w\|_{\lambda,[0,t)},$$

where $\bar{c}_0$ and $\bar{c}_w$ are finite constants, with $\bar{c}_0$ depending on initial conditions and $\bar{c}_w$ not. Moreover, all signals remain bounded, provided that $w(t)$ is uniformly bounded for $t \in [0,\infty)$; and that all signals converge to zero as $t \to \infty$ when $w(t) = 0$, $t \geq 0$.

We do not prove this result here for lack of space.

2.5 Other switching logics

In the sequel we describe a few other switching logics that can be used to generate the switching signal in estimator-based supervision.

2.5.1 Scale-independent hysteresis switching logic

The idea behind hysteresis-based switching logics is to slowdown switching based on the observed growth of the estimation errors instead of forcing a fixed dwell-time. Although hysteresis logics do not enforce a minimum interval between consecutive switchings, they can still be used to achieve non-destabilization of the switched system. The *Scale-independence hysteresis switching logic* [24, 26, 28] presented here is inspired by its non-scale-independent counter part introduced in [51, 61]. Figure 8 shows a graphical representation of this logic, where h is a positive *hysteresis constant*; the signals μ_p, $p \in \mathcal{P}$ are called the *monitoring signals* and are defined by

$$\mu_p(t) := \epsilon + e^{-\lambda t}\epsilon_0 + \int_0^t e^{-2\lambda(t-\tau)} \|e_p(\tau)\|^2 d\tau, \qquad p \in \mathcal{P};$$

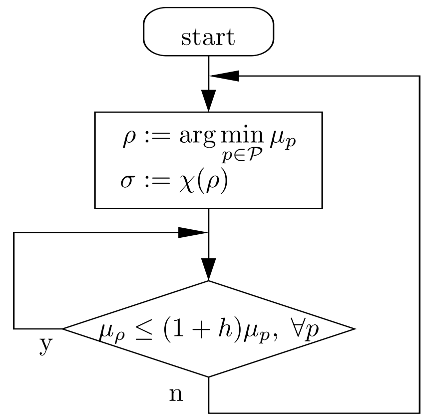

Figure 8: Scale-independent hysteresis switching logic

λ is a constant non-negative *forgetting factor*; and ϵ, ϵ_0 nonnegative constants, with at least one of them strictly positive. positive.

The denomination "scale-independent" comes from the fact that the switching signal σ generated by the logic would not change if all the monitoring signals we simultaneously scaled, i.e., if all $\mu_p(t)$, $p \in \mathcal{P}$ were replaced by $\vartheta(t)\mu_p(t)$, $p \in \mathcal{P}$ for some positive signal $\vartheta(t)$. This property is crucial to the proof of the Scale-Independent Hysteresis Switching Theorem [28] that provides the type of bounds needed to establish the Non-destabilization and Small Error Properties. We state below a version of this theorem adapted to the monitoring signal defined in (53). The following notation is needed: given a switching signal σ, we denote by $N_\sigma(\tau, t)$, $t > \tau \geq 0$ the number of discontinuities of σ in the open interval (τ, t).

Theorem 4 (Scale-Independent Hysteresis Switching). *Let $\mathcal{P}$ be a finite set with m elements. For any $p \in \mathcal{P}$ we have that*

$$N_\sigma(\tau, t) \leq 1 + m + \frac{m \log\left(\frac{\mu_p(t)}{\epsilon + e^{-\lambda t}\epsilon_0}\right)}{\log(1+h)} + \frac{m\lambda(t-\tau)}{\log(1+h)}, \qquad \forall t > \tau \geq 0, \tag{30}$$

and

$$\int_0^t e^{-\lambda(t-\tau)} \|e_\rho(\tau)\|^2 d\tau \leq (1+h)m\mu_p(t), \qquad \forall t > 0. \tag{31}$$

Equations (56) and (31) can be used to establish suitable Non-destabilization and Small Error Properties, respectively. We refer the reader to [26] for details of the stability analysis .

2.5.2 Hierarchical hysteresis switching logic

A key assumption of the Scale-Independent Hysteresis Switching Theorem 4 was the finiteness of the parameter set $\mathcal{P}$. In fact, when $\mathcal{P}$ has infinitely many elements the scale-independent hysteresis switching logic could, in principle, produce an arbitrarily large number of switchings in a finite interval $(\tau, t) \subset [0, \infty)$. This difficulty is avoided by the *hierarchical hysteresis switching logic* introduced in [46, 29]. Figure 8 shows a graphical representation of this logic, where h is a positive

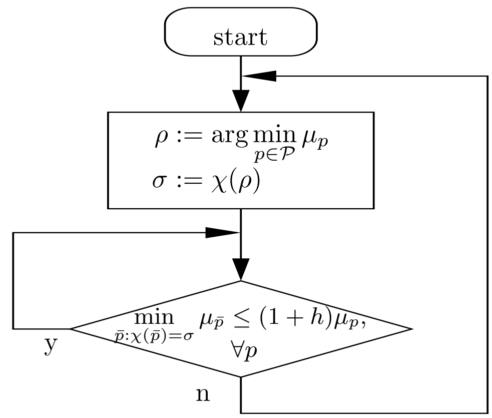

Figure 9: Hierarchical hysteresis switching logic

hysteresis constant; the signals μ_p, $p \in \mathcal{P}$ are called the *monitoring signals* and are defined by

$$\mu_p(t) := \epsilon + e^{-\lambda t}\epsilon_0 + \int_0^t e^{-2\lambda(t-\tau)}\|e_p(\tau)\|^2 d\tau, \qquad p \in \mathcal{P};$$

λ is a constant non-negative *forgetting factor*; and ϵ, ϵ_0 nonnegative constants, with at least one of them strictly positive.

The hierarchical hysteresis switching logic guarantees bounds like the ones in the Scale-Independent Hysteresis Switching Theorem 4 even when the parameter set $\mathcal{P}$ is infinite, provided that the set of candidate controllers is finite.

Theorem 5 (Hierarchical Hysteresis Switching). *Let $\mathcal{Q}$ be a finite set with m elements. For any $p \in \mathcal{P}$ we have that*

$$N_\sigma(\tau,t) \le 1 + m + \frac{m\log\left(\frac{\mu_p(t)}{\epsilon+e^{-\lambda t}\epsilon_0}\right)}{\log(1+h)} + \frac{m\lambda(t-\tau)}{\log(1+h)}, \qquad \forall t > \tau \ge 0, \tag{32}$$

and for every $t \ge 0$, there exists a process switching signal $\rho_t : [0,t) \to \mathcal{P}$ such that $\sigma = \chi(\rho_t)$ on $[0,t)$ and

$$\int_0^t e^{-\lambda(t-\tau)}\|e_{\rho_t}(\tau)\|^2 d\tau \le (1+h)m\mu_p(t). \tag{33}$$

A noticeable difference between the Scale-Independent Hysteresis Switching Theorem 4 and the Hierarchical Hysteresis Switching Theorem 5 is that the process switching signal ρ_t that appear in the integral bound on the estimation error is not the signal ρ defined by the logic (cf. middle box in Figure 9). However, for each $t \ge 0$, the process switching signal ρ_t still satisfies $\sigma = \chi(\rho_t)$ on $[0,t)$ and therefore each candidate controller $\kappa_{\sigma(\tau)}$, $\tau \in [0,t)$ stabilizes the nominal process model $\nu_{\rho_t(\tau)}$. The use of process switching signal ρ_t that depends on the interval $[0,t)$ on which we want to establish boundedness does not introduce any particular difficulty (in fact, this was already done in Section 2.4.3).

To use the Hierarchical Hysteresis Switching Theorem 5 one needs to work with a finite family of candidate controllers, even if the process parameter set $\mathcal{P}$ has infinitely many elements. This raises the question of whether or not it is possible to stabilize an infinite family of process models $\mathcal{M} := \bigcup_{p \in \mathcal{P}} \mathcal{M}_p$ with a finite set of controllers $\{\nu_q : q \in \mathcal{Q}\}$. It turns out that the answer to this question is affirmative provided that $\mathcal{P}$ is compact and under mild continuity assumptions hold [1].

The reader is referred to [44, 34, 67, 68, 71] for several alternative switching logics.

3 Estimator-based nonlinear supervisory control

Section Summary

In this section we consider the general nonlinear case. We introduce several classes of systems for which it is known how to build multi-estimators and multi-controller. We also present the main stability results available and go through the arguments of the proof.

3.1 Class of admissible processes and candidate controllers

We assume that the uncertain process to be controlled admits a state-space model of the form

$$\dot{x}_P = A(x_P, w, u), \qquad y = C(x_P, w), \tag{34}$$

where u denotes the control input, w an exogenous disturbance and/or measurement noise, and y the measured output. The process model is assumed to belong to a family of the form

$$\mathcal{M} := \bigcup_{p \in \mathcal{P}} \mathcal{M}_p,$$

where p is a parameter taking values on the set $\mathcal{P}$ and each $\mathcal{M}_p$ denotes a family of models centered around a nominal state-space model N_p of the form

$$\dot{z} = A_p(z, w, u), \qquad y = C_p(z, w), \qquad p \in \mathcal{P}.$$

Typically,

$$\mathcal{M}_p := \left\{ \mathrm{M}_p : \mathrm{d}(\mathrm{M}_p, \mathrm{N}_p) \le \epsilon_p \right\},$$

where d represents some metric defined on the set of state-space models. Most of the results resented here are either independent of the metric d used (e.g., those related to the Detectability property) or just consider the case $\epsilon_p = 0$, $p \in \mathcal{P}$ (e.g., those related to the Matching property).

The problem of interest is to stably design a feedback controller that drives the output y to zero. All that follows could be easily extended to the more general set-point control problem, in which one attempts to track an arbitrary constant reference r. Within the framework of supervisory control, this will be achieved by switching among a parameterized family of candidate feedback-controllers

$$\mathcal{C} := \left\{ \dot{z}_q = F_q(z_q, y),\ u = G_q(z_q, y) : q \in \mathcal{Q} \right\},$$

Without loss of generality, we assume that all the state-space models in $\mathcal{C}$ have the same dimension and therefore switching among the controllers in $\mathcal{C}$ can be accomplished using the multi-controller:

$$\dot{x}_C = F_\sigma(x_C, y), \qquad u = G_\sigma(x_C, y),$$

where $\sigma : [0, \infty) \to \mathcal{Q}$ denotes the switching signal.

3.2 Multi-estimator

Currently, a general methodology to design multi-estimators for any class of admissible nonlinear processes does not seem to exist. However, we can design multi-estimators for important specific classes of nonlinear processes. We present some of these next.

3.2.1 State accessible and no exogenous disturbances

Suppose that the nominal state-space models N_p, $p \in \mathcal{P}$ are of the form

$$\dot{z} = A_p(z,u), \qquad y = z, \qquad p \in \mathcal{P}, \tag{35}$$

and therefore that the state is accessible and there is no exogenous disturbance w. One simple multi-estimator for this family of processes is given by

$$\dot{z}_p = A(z_p - y) + A_p(y,u), \qquad y_p = z_p, \qquad p \in \mathcal{P}, \tag{36}$$

where A can be any asymptotically stable matrix. In principle, the state of this multi-estimator would then be $x_E := \{z_p : p \in \mathcal{P}\}$. However, we shall see shortly that it is often possible to implement this type of multi-estimator using "state-sharing," which results in multi-estimators with much smaller dimension.

Using the multi-estimator in (36), when the process model is given by N_{p^*} for some $p^* \in \mathcal{P}$, $e_{p^*} := y_{p^*} - y$ converges to zero exponentially fast at a rate determined by the eigenvalues of A. This is because $e_{p^*} = z_{p^*} - z$ and therefore $\dot{e}_{p^*} = Ae_{p*}$. The Matching Property can then be stated as follows:

Property 6 (Matching). Assume that $\mathcal{M} := \{\mathrm{N}_p : p \in \mathcal{P}\}$, with N_p as in (35). There exist positive constants c_0, λ^* and some $p^* \in \mathcal{P}$ such that

$$\|e_{p^*}(t)\| \leq c_0 e^{-\lambda^* t}, \qquad t \geq 0. \tag{37}$$

Also with nonlinear systems, it is often possible to state-share the multi-estimator, i.e., generate a large number of estimation errors using a state with small dimension. The condition needed here is *separability* of $A_p(\cdot,\cdot)$, in the sense that this function can be written as

$$A_p(y,u) = M(y,u)k(p), \qquad \forall p \in \mathcal{P},\ u,\ y$$

for an appropriately defined matrix-valued function $M(y,u)$ and a vector-valued function $k(p)$. In this case the multi-estimator (36) can be realized as

$$\dot{X}_E = A(X_E - Y) + M(y,u), \qquad y_p = X_E k(p), \quad p \in \mathcal{P} \tag{38}$$

where Y is a matrix with the same size as $M(y,u)$ and all columns equal and y. Note that the separability conditions holds trivially when the unknown parameters enter linearly in the nominal models (35). This is usually required by adaptive control algorithms based on continuous tuning.

3.2.2 Output-injection away from a stable linear system

Suppose that the nominal state-space models N_p, $p \in \mathcal{P}$ are of the form

$$\dot{z} = A_p z + B_p w + H_p(y,u), \qquad y = C_p z + D_p w, \qquad p \in \mathcal{P}, \tag{39}$$

where each A_p is an asymptotically stable matrix. This is actually a generalization (35). One simple multi-estimator for this family of processes is given by

$$\dot{z}_p = A_p z_p + H_p(y,u), \qquad y_p = C_p z_p, \qquad p \in \mathcal{P}. \tag{40}$$

In this case, when the process model is given by N_{p^*} for some $p^* \in \mathcal{P}$, defining $\tilde{z}_{p^*} := z_{p^*} - z$ we have

$$\dot{\tilde{z}}_{p^*} = A_p \tilde{z}_{p^*} - B_p w, \qquad e_{p^*} = C_p \tilde{z}_{p^*} - D_p w.$$

The Matching Property can then be stated as follows:

Property 7 (Matching). Assume that $\mathcal{M} := \{\mathrm{N}_p : p \in \mathcal{P}\}$, with N_p as in (39) and all $A_p + \lambda^* I$ asymptotically stable for some $\lambda^* > 0$. There exist positive constants c_0, c_w and some $p^* \in \mathcal{P}$ such that

$$\|e_{p^*}(t)\| \leq c_0 e^{-\lambda^* t} + c_w, \qquad t \geq 0.$$

In case $w = 0$, c_w can be chosen equal to zero.

The multi-estimator (40) can be state-shared if the matrices A_p are independent of p, i.e., $A_p = A$, $\forall p \in \mathcal{P}$ and the function $H_p(y, u)$ is separable in the sense that is can be written as

$$H_p(y, u) = M(y, u) k(p), \qquad \forall p \in \mathcal{P},\ u,\ y$$

for an appropriately defined matrix-valued function $M(y, u)$ and a vector-valued function $k(p)$. In this case the multi-estimator (40) can be realized as

$$\dot{X}_E = A(X_E - Y) + M(y, u), \qquad y_p = C_p X_E k_p, \quad p \in \mathcal{P}$$

where Y is a matrix with the same size as $M(y, u)$ and all columns equal and y.

3.2.3 Output-injection and coordinate transformation away from a stable linear system

The nominal state-space models N_p in (39) can still be generalized by considering coordinate transformations $\bar{z} = \xi_p(z)$, possibly dependent of the unknown parameter $p \in \mathcal{P}$. In particular, the nominal state-space models N_p, $p \in \mathcal{P}$ can also be of the form

$$\dot{\bar{z}} = \zeta_p(\bar{z}) \Big(A_p \xi_p^{-1}(\bar{z}) + B_p w + H_p\big(C_p \xi_p^{-1}(\bar{z}) + D_p w, u\big) \Big), \qquad y = C_p \xi_p^{-1}(\bar{z}) + D_p w, \qquad p \in \mathcal{P},$$

where each A_p is an asymptotically stable matrix, each ξ_p is a continuously differentiable function with continuous inverse ξ_p^{-1}, and $\zeta_p := \xi_p' \circ \xi_p^{-1}$. Since from an input-output perspective these models are similar to those in (39), the multi-estimator (40) can also be used here and the Matching Property 7 also holds.

The above classes of nominal state-space models are not the only ones for which multi-estimators can be designed. For example in [12], we can find a multi-estimator that does not fall in any of the above classes. It is also possible to design multi-estimators for any class of processes for which it is known how to design state observers. The reader is referred to [65, 4, 2] for work on this area that is particularly relevant to our purposes.

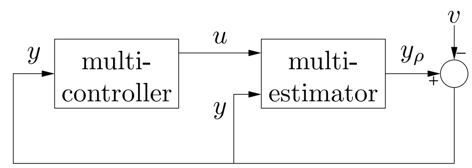

Figure 10: The injected system obtained by replacing y in the multi-estimator and multi-controller by $y_\rho - v$.

3.3 The injected system

Also for nonlinear systems it is useful to regard the *switched system*— defined to be the interconnection of process, multi-estimator, and multi-controller—as the interconnection in Figure 6 of the process and the injected system. As in the linear case, we define the *injected system* by replacing y in the equations of the multi-estimator and multi-controller by $y_\rho - v$, where $\rho : [-0, \infty) \to \mathcal{P}$ is a *process switching signal* and $v := y_\rho - y$. A block diagram of the injected system is shown in Figure 10 and its state-space model is of the form

$$\dot{x} = A_{\rho\sigma}(x, v), \qquad u = F_{\rho\sigma}(x, v), \qquad y_p = C_p(x), \qquad p \in \mathcal{P}. \tag{41}$$

Also here, the simplest mechanism to achieve detectability of the switched system is make sure that the injected system is stable. However, to prove this we need appropriate extensions of these concepts to nonlinear systems. Before proceeding we make a brief detour to recall the relevant definitions of stability and detectability for a nonlinear systems.

3.3.1 Input-to-state stability and detectability

We start be recalling a few definition: We say that a function $\alpha : [0, \infty) \to [0, \infty)$ is of *class* $\mathcal{K}$, and write $\alpha \in \mathcal{K}$, if it is continuous, strictly increasing, and $\alpha(0) = 0$. If α is also unbounded, then we say it is of *class* $\mathcal{K}_\infty$ and write $\alpha \in \mathcal{K}_\infty$. We say that a function $\beta : [0, \infty) \times [0, \infty) \to [0, \infty)$ is of *class* $\mathcal{KL}$, and write $\beta \in \mathcal{KL}$ if $\beta(\cdot, t)$ is of class $\mathcal{K}$ for each fixed $t \geq 0$ and $\beta(s, t)$ decreases to 0 as $t \to \infty$ for each fixed $s \geq 0$.

Consider a general nonlinear system

$$\dot{x} = A(x, u), \qquad y = C(x, u), \tag{42}$$

with state x, piecewise-continuous input u, and output y, where A is assumed locally Lipschitz and $A(0, 0) = 0$, $C(0, 0) = 0$. Although in this paper we take the equilibrium state of the system (42) to be the origin, the subsequent definitions and results can be extended to the case of nonzero equilibrium states [31, 33].

Following [79], we say that (42) is *input-to-state stable* (ISS) if there exist functions $\beta \in \mathcal{KL}$, $\gamma \in \mathcal{K}$ such that

$$\|x(t)\| \leq \beta(\|x(0)\|, t) + \sup_{\tau \in [0,t)} \gamma(\|u(\tau)\|), \qquad t \geq 0,$$

along solutions to (42). The system (42) is said to be *integral input-to-state stable* (iISS) if there exist functions $\alpha \in \mathcal{K}_\infty$, $\beta \in \mathcal{KL}$, $\gamma \in \mathcal{K}$ such that

$$\alpha(\|x(t)\|) \le \beta(\|x(0)\|, t) + \int_{\tau \in [0,t)} \gamma(\|u(\tau)\|), \qquad t \ge 0,$$

along solutions to (42). Integral input-to-state stability was introduced in [80] and is a weaker version of stability than input-to-state stability. In fact, every ISS system is iISS. This is immediate from the characterizations of input-to-state stability and integral input-to-state stability in terms of dissipation inequalities, which can be found in [82] and [3], respectively. It turns out that iISS is a strictly weaker property as there are iISS systems that are not ISS [80].

The system (42) is said to be *detectable* (or input/output-to-state stable IOSS [85]) if there exist functions $\beta \in \mathcal{KL}$, $\gamma_u, \gamma_y \in \mathcal{K}$ such that

$$\|x(t)\| \le \beta(\|x(0)\|, t) + \sup_{\tau \in [0,t)} \gamma_u(\|u(\tau)\|) + \sup_{\tau \in [0,t)} \gamma_y(\|y(\tau)\|), \qquad t \ge 0,$$

along solutions to (42). An equivalent characterization of detectability in terms of dissipation inequalities can be found in [41]. Similarly, (42) is said to be *integrable detectable* (or integral input/output-to-state stable iIOSS) if there exist functions $\alpha \in \mathcal{K}_\infty$, $\beta \in \mathcal{KL}$, $\gamma_u, \gamma_y \in \mathcal{K}$ and such that

$$\alpha(\|x(t)\|) \le \beta(\|x(0)\|, t) + \int_{\tau \in [0,t)} \gamma_u(\|u(\tau)\|) + \int_{\tau \in [0,t)} \gamma_y(\|y(\tau)\|), \qquad t \ge 0,$$

along solutions to (42). Both notions of detectability are consistent with the usual definition of detectability for linear systems, which basically says that the state eventually becomes small if the inputs and outputs are small. In fact, a detectable linear system (in the usual sense) is both detectable and integral detectable in the above senses. However, in general integral detectability is a weaker property than detectability. Indeed, every detectable system is integral detectable. This can be shown using the characterization of detectability in terms of the exponentially decaying dissipation inequality in [41, Section 5.1]. The following result will be needed [80, Proposition 6].

Lemma 1. *Suppose that the system* (42) *is integral detectable and that the initial state* $x(0)$ *and the input* u *are such that the corresponding solution of* (42) *is globally defined and*

$$\int_0^\infty \gamma_u(\|u(\tau)\|)d\tau < \infty, \qquad \int_0^\infty \gamma_y(\|y(\tau)\|)d\tau < \infty.$$

Then $x(t) \to 0$ *as* $t \to \infty$.

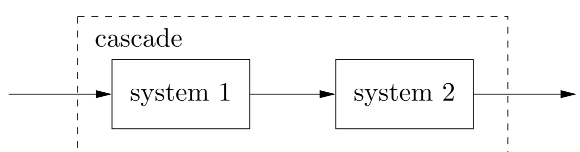

Figure 11: Cascade connection

The following are well-known properties of the interconnections between input-to-state stable and detectable systems.

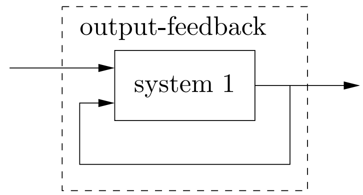

Figure 12: Output-feedback connection

Lemma 2 (Cascade). *Consider the cascade connection in Figure 11:*

1. *If system Σ_1 is input-to-state stable and system Σ_2 is detectable, then the cascade is detectable.*
2. *If system Σ_1 is integral input-to-state stable and system Σ_2 is detectable, then the cascade is integral detectable.*

Lemma 3 (Feedback). *Consider the output-feedback connection in Figure 12.*

1. *If system Σ_1 is detectable, then the output-feedback connection is detectable.*
2. *If system Σ_1 is integral detectable, then the output-feedback connection is integral detectable.*

The proof that the cascade of an input-to-state stable with a detectable system is detectable can be found in [31, Appendix] and its integral version in [28, Proof of Theorem 3]. These proofs use fairly standard techniques that were developed to analyze the cascade of input-to-state stable systems [79, 81] (see also [37, Section 5.3]). Lemma 3 is a straightforward consequence of the detectability definitions.

3.3.2 Nonlinear Certainty Equivalence Stabilization Theorems

We are now ready to state the nonlinear version of the Certainty Equivalence Stabilization Theorem that establishes the detectability of the switched system from the input-to-state stability of the injected system:

Theorem 6 (Certainty Equivalence Stabilization). *Suppose that the process* (34) *is detectable and take a fixed $\rho = p \in \mathcal{P}$ and $\sigma = q \in \mathcal{Q}$. Then*

1. *If the injected system is input-to-state stable then the switched system is detectable [31].*
2. *If the injected system is integral input-to-state stable then the switched system is integral detectable [28].*

To prove this result we redraw in Figure 13 the diagram from Figure 6 that shows the switched system as the interconnection of the process and the injected system. Since we are setting $\rho = p \in \mathcal{P}$, we only show the output estimate y_p.

Figure 13 shows that the switched system can be obtained by first cascading the injected system Σ_1 with input v and output (u, y_p) with a system Σ_2 with input (w, u, y_p) and output e_p (whose dynamics are essentially those of the process) and then closing the loop with the output-feedback $v := e_p$.

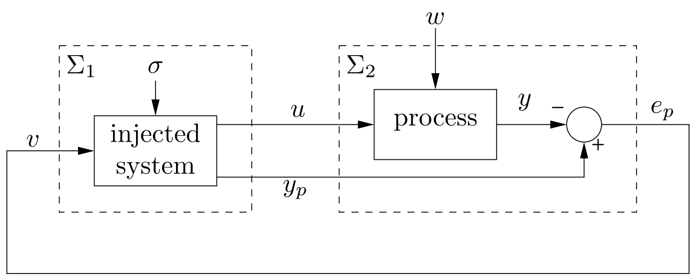

Figure 13: The switched system as the interconnection of the process with the injected system (version 2).

From the definition of detectability it is immediate that Σ_2 is detectable because the process is detectable. Suppose now that the injected system Σ_1 is input-to-state stable. It turns out that the cascade of an input-to-state stable with a detectable system is detectable, therefore the cascade of Σ_1 and Σ_2 (with input (w, v) and output e_p) is detectable (cf. Lemma 2). Moreover, output-feedback preserves detectability so the overall switched system is detectable (cf. Lemma 3). This proves part 1 of Theorem 6. As for part 2 we can use the same argument but keeping in mind that the cascade of an integral input-to-state stable with a detectable system is integral detectable.

As for the linear case, (integral) input-to-state stability of the injected system is not the only mechanism to obtain detectability of the switched system. For example, a nonlinear version of the Certainty Equivalence Output Stabilization Theorem is given in [48], where it is shown that input-to-output stability of the injected system together with "minimum-phase" of the process and detectability of both the multi-controller and multi-estimator are also sufficient to prove that the switched system is detectable. Minimum-phase here should be understood in the sense defined in [48].

3.3.3 Achieving detectability

To use the Certainty Equivalence Stabilization Theorem 6 to prove detectability one needs to design candidate controllers that (integral) input-to-state stabilize the injected system, i.e., that make the feedback connection in Figure (10) at least integral input-to-state stable with respect to the input v. In particular, one would like to determine a family of candidate controllers

$$\mathcal{C} := \left\{ \dot{z}_q = F_q(z_q, y),\ u = G_q(z_q, y) : q \in \mathcal{Q} \right\},$$

and a controller selection function $\chi : \mathcal{P} \to \mathcal{Q}$ such that for each $p \in \mathcal{P}$ the injected system in Figure (10) with $\rho = p$ and $\sigma = \chi(p)$ is integral input-to-state stable.

The input v in Figure (10) can be viewed as a form of measurement noise. Therefore the design of candidate controllers for supervisory control fall into the general problem of designing "robust" controllers for nonlinear systems. There are however, a few simplifying assumptions:

1. The disturbance v can be measured and used for control (recall that v is noting more than the output estimation error $e_\rho := y_\rho - y$.
2. The state of the injected system can also be measured and used for control (recall that the state of the injected system is nothing but the combined state of the multi-estimator and multi-controller).

We present next a few cases for which it is known how to solve the candidate controller design problem.

Feedback fully linearizable multi-estimator Suppose that the multi-estimator is of the form

$$\dot{x}_E = Ax_E + B(\Psi(x_E, y) + u) + Dy, \qquad y_p = C_p x_E, \qquad p \in \mathcal{P}, \tag{43}$$

with each pair $(A + DC_p, B)$ stabilizable. By setting $y := y_\rho - v$, $\rho = p \in \mathcal{P}$ we obtain the system that needs to be stabilized by the candidate controller $q := \chi(p)$:

$$\dot{x}_E = (A + DC_p)x_E + B\big(\Psi(x_E, C_p x_E - v) + u\big) - Dv.$$

To achieve this we can simply set $\chi(p) := p$, $p \in \mathcal{P} := \mathcal{Q}$ and define the candidate controller $q := \chi(p) = p$ to be

$$u = -\Psi(x_E, C_p x_E - v) + F_p x_E, \tag{44}$$

where F_p is any matrix for which $A + DC_p + BF_p$ is asymptotically stable. The corresponding injected system is given by

$$\dot{x}_E = (A + DC_p + BF_p)x_E - Dv,$$

and is therefore input-to-state stable.

Even when the multi-estimator is feedback linearizable, more often than not is will not be in the "canonical form" (43)—for example, it may be in the strict-feedback or the pure-feedback forms [42]—, in which case a coordinate transformation is needed to put the system in the form (43). Suppose then that the multi-estimator is of the form

$$\dot{\bar{x}}_E = \zeta(\bar{x}_E)\Big(A\xi^{-1}(\bar{x}_E) + B\big(\Psi(\xi^{-1}(\bar{x}_E), y) + u\big) + Dy\Big), \qquad y_p = C_p\xi^{-1}(\bar{x}_E), \qquad p \in \mathcal{P}, \tag{45}$$

where each the pair $(A + DC_p, B)$ is stabilizable, ξ is a continuously differentiable function with continuous inverse ξ^{-1}, and $\zeta := \xi' \circ \xi^{-1}$. The multi-estimator (45) results from applying the coordinate transformation $\bar{x}_E = \xi(x_E)$ to (43). By setting $y := y_\rho - v$, $\rho = p \in \mathcal{P}$ we obtain the system that needs to be stabilized by the candidate controller $q := \chi(p)$:

$$\dot{\bar{x}}_E = \zeta(\bar{x}_E)\Big((A + DC_p)\xi^{-1}(\bar{x}_E) + B\big(\Psi(\xi^{-1}(\bar{x}_E), C_p\xi^{-1}(\bar{x}_E) - v) + u\big) - Dv\Big)$$

In this case, we could set $\chi(p) := p$, $p \in \mathcal{P} := \mathcal{Q}$ and define the candidate controller $q := \chi(p) = p$ to be

$$u = -\Psi(\xi^{-1}(\bar{x}_E), C_p\xi^{-1}(\bar{x}_E) - v) + F_p\xi^{-1}(\bar{x}_E), \tag{46}$$

where F_p is any matrix for which $A + DC_p + BF_p$ is asymptotically stable. The corresponding injected system is given by

$$\dot{\bar{x}}_E = \zeta(\bar{x}_E)\Big((A + DC_p + BF_p)\xi^{-1}(\bar{x}_E) - Dv\Big). \tag{47}$$

To check the stability of this system, we do the change of coordinates $x_E := \xi^{-1}(\bar{x}_E)$ and obtain

$$\dot{x}_E = (A + DC_p + BF_p)x_E - Dv.$$

Since this system is input-to-state stable, so is the injected system (47).

Input-output feedback linearizable multi-estimator Suppose that for each $p \in \mathcal{P}$ we can partition the state x_E of the multi-estimator as $x_E = (x_p, \bar{x}_p)$ and that its dynamics are of the form

$$\dot{x}_p = A_p x_p + B_p(\Psi_p(x_E, y) + u) + D_p y, \qquad y_p = C_p x_p, \tag{48}$$
$$\dot{\bar{x}}_p = \bar{A}_p(\bar{x}_p, x_p, u, y), \tag{49}$$

where each pair $(A_p + D_p C_p, B_p)$ is stabilizable and (49) is input-to-state stable when (x_p, u, y) is viewed as a disturbance input. In this case, we could set $\chi(p) := p$, $p \in \mathcal{P} := \mathcal{Q}$ and define the candidate controller $q := \chi(p) = p$ to be

$$u = -\Psi(x_E, C_p x_p - v) + F_p x_p, \tag{50}$$

where F_p is any matrix for which $A_p + D_p C_p + B_p F_p$ is asymptotically stable. The corresponding injected system is given by

$$\dot{x}_p = (A_p + D_p C_p + B_p F_p) x_p - D_p v,$$
$$\dot{\bar{x}}_p = \bar{A}_p(\bar{x}_p, x_p, -\Psi(x_E, C_p x_p - v) + F_p x_p, C_p x_p - v),$$

which is input-to-state stable because it can be viewed as the cascade of two input-to-state stable systems [79, 81].

This generalization is important because typically the multi-estimator is "non-minimal" (due to state-sharing) and it is often possible to linearize its dynamics from its inputs to one particular output y_p but not to all its outputs simultaneously. This typically happens when every nominal process model is feedback linearizable but the coordinate transformation needed to achieve the linearization depends on the unknown parameter and is therefore not the same for every nominal model. A typical example is the family of nominal models:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} p_1 x_1^3 + p_2 x_2 \\ u \end{bmatrix}, \qquad y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad (p_1, p_2) \in \mathcal{P}.$$

Also here we could consider (possibly p-dependent) coordinate transformation from the canonical form (48)–(49) to more general forms. We leave this as an exercise to the reader.

Using the candidate controllers defined above and making use of the Certainty Equivalence Stabilization Theorem 6 we obtain the desired Detectability Property:

Property 8 (Detectability). Suppose that the process is detectable, the multi-estimator is of the form (48)–(49) and $\mathcal{P} := \mathcal{Q}$ with the candidate controller $q = \chi(p) := p$ given by (50). Then, for $\rho = p \in \mathcal{P}$ and $\sigma = \chi(p) = p$ the injected system (41) is input-to-state stable (and consequently integral input-to-state stable), therefore there exist functions $\bar{\beta}_p \in \mathcal{KL}$, $\bar{\alpha}_p, \bar{\gamma}_p \in \mathcal{K}$ such that

$$\bar{\alpha}_p(\|x(t)\|) \leq \bar{\beta}_p(\|x(0)\|, t) + \int_{\tau \in [0,t)} \bar{\gamma}_p(\|v(\tau)\|), \qquad t \geq 0. \tag{51}$$

Moreover, the switched system is detectable (and consequently integral detectable) through the output e_p, therefore there exist functions $\beta \in \mathcal{KL}$, $\alpha_p, \gamma_p, \varphi_p \in \mathcal{K}$ such that

$$\alpha_p(\|\bar{x}(t)\|) \leq \beta_p(\|\bar{x}(0)\|, t) + \int_{\tau \in [0,t)} \gamma_p(\|e_p(\tau)\|) + \int_{\tau \in [0,t)} \varphi_p(\|w(\tau)\|), \qquad t \geq 0, \tag{52}$$

where $\bar{x}$ denotes the aggregate state of the process, multi-estimator, and multi-controller.

Note that the candidate controllers (44), (46), (50) are not the only ones that would lead to detectability for the given multi-estimators. Another option would be to used these feedback linearizing control laws to determine a ISS control Lyapunov function and then use it to construct more suitable controllers that still make the injected system input-to-state stable [18, 83, 87, 43, 49]. In [18, 43] it is shown that these control laws often correspond to meaningful inverse optimal robust stabilization problems for nonlinear systems with disturbances. The pointwise min-norm control laws proposed in [18] were used in the context of supervisory control in [24, 27].

Feedback linearizability in the above senses is a property of a multi-estimator that can be used to achieve detectability. However, it is not the only one and, e.g., [25, 12] one can find candidate controllers for non-feedback linearizable multi-estimators. The [25] is particularly interesting because the nominal process models (and by extension the multi-estimator) cannot be stabilized by any continuous time-invariant feedback control law due to their nonholonomic nature [10]. In [25], the candidate controllers are then themselves hybrid systems, leading to a two-layer hierarchical supervisory control structure.

3.4 Scale-independent hysteresis switching logic

Although dwell-time switching logics are extremely successful in linear supervisory, they are generally not adequate in the context of nonlinear supervision because of finite escape. Indeed, when the switched system is not globally Lipschitz, "dwelling" on a non-stabilizing controller for a fixed interval of time τ_D may lead to state unboundedness in finite time.

Hysteresis-based switching logics have been proposed to address this difficulty. The idea behind these logics is to slowdown switching based on the observed growth of the estimation errors instead of forcing a fixed dwell-time. Although hysteresis logics do not enforce a minimum interval between consecutive switchings, they can still be used to achieve non-destabilization of the switched system.

The *Scale-independence hysteresis switching logic* [24, 26, 28] presented here is inspired by its non-scale-independent counter part introduced in [51, 61]. Figure 8 shows a graphical representation of this logic, where h is a positive *hysteresis constant* and the signals μ_p, $p \in \mathcal{P}$ are called the *monitoring signals* and are defined by

$$\mu_p(t) := \epsilon + e^{-\lambda t}\epsilon_0 + \int_0^t e^{-\lambda(t-\tau)}\gamma_p(\|e_p(\tau)\|)d\tau, \qquad p \in \mathcal{P}, \tag{53}$$

where λ denotes a constant non-negative *forgetting factor*; ϵ, ϵ_0 nonnegative constants, with at least one of them strictly positive; and the γ_p, $p \in \mathcal{P}$ class $\mathcal{K}$ functions.

When the maps $p \mapsto \gamma_p(\|e_p(\tau)\|)$ have adequate separability properties, it is possible it is possible to efficiently generate a large number (or even a continuum) of monitoring signals with a low-dimensional dynamical system. This happens when we can write

$$\gamma_p(\|e_p\|) = k(p)'h(y, u, x_E), \qquad \forall p \in \mathcal{P},\ u,\ y,\ x_E,$$

where x_E is the state of the multi-estimator and $k(p)$, $h(y, u, x_E)$ are appropriately defined vector-valued functions. In this case, the monitoring signals could be generated by

$$\dot{x}_\mu = -\lambda x_\mu + h(y, u, x_E), \qquad \mu_p = k(p)'x_\mu, \qquad p \in \mathcal{P}.$$

This means that the generation of ρ in the middle box of diagram in Figure 8 can be written as

$$\rho := \arg\min_{p\in\mathcal{P}} k(p)x_\mu, \tag{54}$$

and the condition in the lower box as

$$k(\rho)x_\mu \le (1+h)k(p)x_\mu, \qquad \forall p \in \mathcal{P}. \tag{55}$$

This means that we never actually need to explicitly compute all the estimation errors e_p or the monitoring signals μ_p, as long as we know how to solve the optimization problem (54) and the feasibility problem (55).

3.4.1 Scale-independent hysteresis switching properties

The denomination "scale-independent" comes from the fact that the switching signal σ generated by the logic would not change if all the monitoring signals we simultaneously scaled, i.e., if all $\mu_p(t)$, $p \in \mathcal{P}$ were replaced by $\vartheta(t)\mu_p(t)$, $p \in \mathcal{P}$ for some positive signal $\vartheta(t)$. This property is crucial to the proof of the Scale-Independent Hysteresis Switching Theorem [28] that provides the type of bounds needed to establish the Non-destabilization and Small Error Properties. We state below a version of this theorem adapted to the monitoring signal defined in (53). The following notation is needed: given a switching signal σ, we denote by $N_\sigma(\tau, t)$, $t > \tau \ge 0$ the number of discontinuities of σ in the open interval (τ, t).

Theorem 7 (Scale-Independent Hysteresis Switching). *Let $\mathcal{P}$ be a finite set with m elements. For any $p \in \mathcal{P}$ we have that*

$$N_\sigma(\tau,t) \le 1 + m + \frac{m \log\left(\frac{\mu_p(t)}{\epsilon + e^{-\lambda t}\epsilon_0}\right)}{\log(1+h)} + \frac{m\lambda(t-\tau)}{\log(1+h)}, \qquad \forall t > \tau \ge 0, \tag{56}$$

and

$$\int_0^t e^{-\lambda(t-\tau)}\gamma_\rho(\|e_\rho(\tau)\|)d\tau \le (1+h)m\mu_p(t), \qquad \forall t > 0. \tag{57}$$

To derive the Non-destabilization and the Small Error Properties, let $[0, T)$ be the maximum interval on which the solution to the switched system supervised by the scale-independent hysteresis logic is defined. Since we are dealing with systems that may not be globally Lipschitz, we must consider the possibility that $T_{\max} < \infty$.

Equation (56) provides an upper bound on the maximum number of switchings that can occur in an arbitrary interval (τ, t), $0 \le \tau < t < T_{\max}$. Not surprisingly, this number decreases as the hysteresis constant h increases or the forgetting factor λ decreases. From (56) we can obtain the desired Non-destabilization Property as follows: Set $\epsilon = 0$, $\epsilon_0 > 0$ and assume that the norm one of the estimation errors $e_{p^*}(t)$, $p^* \in \mathcal{P}$ can be bounded by $c_0 e^{-\lambda^* t}$, $t \in [0, T_{\max})$ with $\lambda^* > \lambda > 0$, $c_0 > 0$. Then, assuming that γ_{p^*} in (53) is locally Lipschitz, $\gamma_{p^*}(\|e_{p^*}(\tau)\|)$ can also be bounded by a similar expression. From this it is straightforward to conclude that $\mu_{p^*}(t)$ can be bounded by $\bar{c}_0 e^{-\lambda t}$, $t \in [0, T_{\max})$ with $\lambda > 0$, $\bar{c}_0 > 0$. This means that $N_\sigma(0, t)$ is uniformly bounded for $t \in [0, T_{\max})$, or equivalently, that there is only a finite number of switching times on $[0, T_{\max})$ (even if $T_{\max} = \infty$). The following can then be stated:

Property 9 (Non-destabilization). Assume that $\mathcal{P}$ is finite, that all the γ_p, $p \in \mathcal{P}$ in (53) are locally Lipschitz, and that $\epsilon = 0$, $\epsilon_0 > 0$. If one of the norm one of the estimation errors $e_{p^*}(t)$, $p^* \in \mathcal{P}$ can be bounded by $c_0 e^{-\lambda^* t}$, $t \in [0, T_{\max})$ with $\lambda^* > \lambda > 0$, $c_0 > 0$ then switching will stop after some finite time $T^* < T_{\max}$.

Equation (57), on the other hand, provides and upper bound on a integral-norm of the switched estimation error e_ρ that can be used to establish the desired Small Error Property. Indeed, under the assumptions of the Non-destabilization Property 9, $e^{\lambda t}\mu_{p^*}(t)$ is uniformly bounded. From this and (57) we conclude that the following property holds:

Property 10 (Small Error). Assume that $\mathcal{P}$ is finite, that all the γ_p, $p \in \mathcal{P}$ in (53) are locally Lipschitz, and that $\epsilon = 0$, $\epsilon_0 > 0$. If one of the norm one of the estimation errors $e_{p^*}(t)$, $p^* \in \mathcal{P}$ can be bounded by $c_0 e^{-\lambda^* t}$, $t \in [0, T_{\max})$ with $\lambda^* > \lambda > 0$, $c_0 > 0$ then the signal

$$\int_0^t e^{\lambda\tau}\gamma_\rho(\|e_\rho(\tau)\|)d\tau$$

is uniformly bounded for $t \in [0, T_{\max})$.

3.4.2 Analysis

With the four Properties—Matching 6, Detectability 8, Non-destabilization 9, and Small error 10—at hand it is then straightforward to prove the following:

Theorem 8. *Consider a detectable process, a multi-estimator as in the Matching Property 6 (or the Matching Property 7 with $w = 0$); candidate controllers as in the Detectability Property 8; and a scale-independent hysteresis supervisor with $\epsilon = 0$, $\epsilon_0 > 0$, $\lambda < \lambda^*$ (with λ^* as in the Matching Property), and the γ_p, $p \in \mathcal{P}$ from (52). Assuming that $\mathcal{P}$ is finite and that all the γ_p are locally Lipschitz, then the state of the process, multi-estimator, multi-controller, and all other signals converge to zero as $t \to \infty$.*

To prove this result we simply have to note that the Matching Property guarantees that the norm one of the estimation errors $e_{p^*}(t)$, $p^* \in \mathcal{P}$ can be bounded by $c_0 e^{-\lambda^* t}$, $t \in [0, T_{\max})$ with $\lambda^* > \lambda > 0$, $c_0 > 0$. Therefore, because of the Non-destabilization Property, switching will stop after some finite time $T^* < T_{\max}$. Let then $p \in \mathcal{P}$ denote the (constant) value of $\rho(t)$ for $t \in [T^*, T_{\max})$. Because of the Small error Property, we must have

$$\int_{T^*}^{T_{\max}} \gamma_p(\|e_p(\tau)\|)d\tau < \infty.$$

From this and the Detectability Property (recall that $w = 0$), we conclude that the state $\bar{x}$ of the switched system with $\sigma = \chi(\rho) = p$, $t \geq T^*$ must by bounded on $[T^*, T_{\max})$ and therefore $T_{\max} = \infty$. Moreover, because of Lemma 1, we actually have that $\bar{x}(t)$ converges to 0 as $t \to \infty$.

It is often convenient to use the class $\mathcal{K}$ functions $\bar{\gamma}_p$, $p \in \mathcal{P}$ in (51) instead of the functions γ_p in (52) to define the monitoring signals μ_p. This is because the $\bar{\gamma}_p$ do not depend on the unknown process model. Reasoning as above, in this case we conclude that after some finite time $T^* < T_{\max}$, $\rho(t)$ will become constant and equal to some $p \in \mathcal{P}$ for which

$$\int_{T^*}^{T_{\max}} \bar{\gamma}_p(\|e_p(\tau)\|)d\tau < \infty.$$

From this and the Detectability Property, we now conclude that the state x of the injected system with $\sigma = \chi(\rho) = p$, $t \geq T^*$ must by bounded on $[T^*, T_{\max})$. Therefore u and y_{p^*}. From this and the boundedness of e_{p^*}, we conclude that both the input u and the output $y = y_{p^*} - e_{p^*}$ of the process are bounded. Since the process is detectable, its state must also be bounded. Once boundedness of all signals has been established we conclude that $T_{\max} = \infty$. Also here, using Lemma 1, we could conclude that all signals actually converges to 0 as $t \to \infty$.

Corollary 1. *The results in Theorem 8 also holds if one uses in* (53) *the class* $\mathcal{K}$ *functions* $\bar{\gamma}_p$ *from* (51) *instead of the* γ_p *from* (52)*.*

Acknowledgements

The author would like to thanks Daniel Liberzon and Stephen Morse for very helpful discussions and suggestions on the contents of this paper.

Appendix

A Linear multi-estimator design

The basic requirement behind the multi-estimator is that a particular y_p, $p \in \mathcal{P}$ be "small" when the process model matches one element of the family $\mathcal{M}_p$. Both linear and nonlinear control theory provide numerous procedures to design systems with this property that can be used in the context of supervisory control.

Let us start by considering families $\mathcal{M}_p$, $p \in \mathcal{P}$ each consisting of a single linear time-invariant process with transfer matrix ν_p and minimal realization

$$\dot{z} = A_p z + B_p(u + d), \qquad y = C_p z + D_p u + n,$$

where n and d denote measurement noise and an input disturbance, respectively. We are assuming here that $\mathcal{P}$ is a finite set. The extension of these ideas to sets $\mathcal{P}$ that have infinitely many elements will be addressed shortly.

One of the simplest multi-estimator for this process utilizes the following family of Luenberger observers

$$\dot{z}_p = (A_p - K_p C_p) z_p + B_p u + K_p y, \qquad y_p = C_p z_p + D_p u, \qquad p \in \mathcal{P}, \tag{58}$$

where the K_p are chosen so that $A_p - K_p C_p$ are asymptotically stable matrices. From observer theory we know that if the process transfer matrix is ν_{p^*}, $p^* \in \mathcal{P}$ and in the absence of noise and disturbances ($n = d = 0$) then y_{p^*} converges exponentially fast to y. In case n and/or d are non zero then the estimation error $e_{p^*} := y_{p^*} - y$ is simply bounded. The Matching Property can then be formally stated as

Property 11 (Matching–1). There exist positive constants c_0, c_w, λ and some $p^* \in \mathcal{P}$ such that

$$\|e_{p^*}(t)\| \leq c_0 e^{-\lambda t} + c_w, \qquad t \geq 0.$$

In case $n = d = 0$, c_w can be chosen equal to zero.

Another method to build multi-estimators is based on coprime factorizations over the ring RH_∞ of stable transfer matrices. For a given $p \in \mathcal{P}$, consider the following factorization

$$\nu_p = D_p^{-1} N_p,$$

with $D_p, N_p \in \mathrm{RH}_\infty$ coprime over RH_∞. Without loss of generality we can write

$$D_p = \frac{1}{\omega_p}\eta_p, \qquad N_p = \frac{1}{\omega_p}\mu_p,$$

where ω_p is a stable polynomial and η_p, μ_p matrices of polynomials. We therefore conclude that when the process transfer matrix is equal to μ_{p^*}, $p^* \in \mathcal{P}$ [3]

$$\eta_{p^*}(y-n) = \mu_{p^*}(u+d).$$

Therefore, generating the y_p by the family of differential equations

$$\omega_p y_p = (\omega_p I - \eta_p)y + \mu_p u, \qquad p \in \mathcal{P} \tag{59}$$

we conclude that, when the process transfer matrix is equal to ν_{p^*},

$$\omega_p y_{p^*} = \omega_{p^*} y - \eta_{p^*} n - \mu_{p^*} d$$

and therefore, in the absence of noise and disturbances, y_{p^*} converges exponentially fast to y, at a rate determined by the roots of ω_{p^*}. Also here, when either d or n are non zero then e_{p^*} is uniformly bounded and the Matching Property 11 holds.

An interesting feature of the multi-estimator (59) is that if all the ν_p have McMillan degree smaller than m, then we can choose all the ω_p to be equal to some polynomial ω with degree m and (59) can be realized as

$$\dot{w}_1 = A_1 w_1 + B_1 u, \qquad \dot{w}_2 = A_2 w_2 + B_2 y, \qquad y_p = C_p \begin{bmatrix} w_1' & w_2' \end{bmatrix}', \quad p \in \mathcal{P}, \tag{60}$$

where A_1 a $km \times km$ matrix with characteristic polynomial ω^k and A_2 a $lm \times lm$ matrix with characteristic polynomial ω^l. Here, k and l are the number of process inputs and outputs, respectively. What makes the multi-estimator (60) particularly attractive is the fact that a large number of output estimates y_p (possibly infinitely many) can be generated by a low-dimensional system. Moreover, since all the estimation errors e_p, $p \in \mathcal{P}$ can be computed directly from w and y, we can actually take these two signals to be the output of the multi-estimator. When this happens we call the multi-estimator *state-shared*. It is worth mentioning that structures like (60) appear in traditional adaptive control, where w_1 and w_2 are often called *regression vectors*.

Let us consider now the more general case in which the $\mathcal{M}_p$, $p \in \mathcal{P}$ are infinite dimensional balls of unmodeled dynamics around a nominal process ν_p, e.g.,

$$\mathcal{M}_p := \{\nu_p(1+\delta_m) + \delta_a : \|\delta_m\|_{\infty,\lambda} \le \epsilon,\ \|\delta_a\|_{\infty,\lambda} \le \epsilon\}, \qquad p \in \mathcal{P}$$

where ϵ and λ are nonnegative constants and $\|\cdot\|_{\infty,\lambda}$ denotes the $e^{\lambda t}$-weighted $\mathcal{H}_\infty$ norm of a transfer function: $\|\nu\|_{\infty,\lambda} = \sup_{\omega\in\mathbb{R}} |\nu(j\omega - \lambda)|$. If we now still use the multi-estimator (59)—or its state-shared version (60)—it is straightforward to show that is the process transfer matrix is equal to $\nu_{p^*}(1+\delta_m) + \delta_a$, then

$$\omega_p y_{p^*} = \omega_{p^*} y - \eta_{p^*} n - \mu_{p^*} d - (\mu_{p^*}\delta_m + \eta_{p^*}\delta_a) \circ (u+d), \tag{61}$$

where $\delta_m \circ u$ and $\delta_a \circ u$ denote the convolution of u with the impulse responses of δ_m and δ_a, respectively. In this case the Matching Property can be stated as

Property 12 (Matching–2). There exist positive constants $c_0, c_d, c_n, c_\epsilon, \lambda$ and some $p^* \in \mathcal{P}$ such that

$$\|e_{p^*}\|_{\lambda,[0,t)} \le c_0 + c_d\|d\|_{\lambda,[0,t)} + c_n\|n\|_{\lambda,[0,t)} + \epsilon\, c_\epsilon \|u\|_{\lambda,[0,t)}, \qquad \forall t \ge 0,$$

where $\|\cdot\|_{\lambda,[0,t)}$ denotes the $e^{\lambda t}$-weighted $\mathcal{L}_2$-norm truncated to the interval $[0,t)$: $\|u\|_{\lambda,[0,t)} = \int_0^t e^{\lambda\tau} u(\tau) 2 d\tau$.

[3] Given a $l \times k$ polynomial matrix ρ and a signal $a : [0,\infty) \to \mathbb{R}^k$, we denote by ρw the signal $b : [0,\infty) \to \mathbb{R}^l$ defined by applying the differential operator $\rho\left(\frac{d}{dt}\right)$ to a.

References

[1] B. D. O. Anderson, T. S. Brinsmead, F. de Bruyne, J. P. Hespanha, D. Liberzon, and A. S. Morse. Multiple model adaptive control. i. finite controller coverings. George Zames Special Issue of the *Int. J. of Robust and Nonlinear Control*, 10(11-12):909–929, Sept. 2000.

[2] D. Angeli. A Lyapunov approach to incremental stability properties. Submitted to IEEE Trans. on Automat. Contr., 2001.

[3] D. Angeli, E. D. Sontag, and Y. Wang. A characterization of integral input-to-state stability. *IEEE Trans. on Automat. Contr.*, 45(6):1082–1097, June 2000.

[4] M. Arcak and P. V. Kokotović. A new nonlinear observer: Design and robustness analysis. Submitted to the IEEE Trans. on Autom. Control., Nov. 1999.

[5] K. J. Åström. Adaptive control around 1960. *IEEE Contr. Syst. Mag.*, 16(3):44–49, June 1996.

[6] T. Başar and P. Bernhard. *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Systems & Control: Foundations & Applications. Bikhäuser, Boston, 2nd edition, 1995.

[7] J. Balakrishnan. *Control Systems Design Using Multiple Models, Switching, and Tuning*. PhD thesis, Yale University, 1995.

[8] D. Borrelli, A. S. Morse, and E. Mosca. Discrete-time supervisory control of families of two-degrees-of-freedom linear set-point controllers. *IEEE Trans. on Automat. Contr.*, 44(1):178–181, Jan. 1998.

[9] M. S. Branicky. Multiple Lyapunov functios and other analysis tools for switched and hybrid systems. *IEEE Trans. on Automat. Contr.*, 43(4):475–482, Apr. 1998.

[10] R. W. Brockett. Asymptotic stability and feedback stabilization. In R. W. Brockett, R. S. Millman, and H. J. Sussmann, editors, *Differential Geometric Control Theory*, pages 181–191. Birkhäuser, Boston, 1983.

[11] F. B. Cabral and M. G. Safonov. A falsification perspective on model reference adaptive control. In *Proc. of the 35th Conf. on Decision and Contr.*, pages 2982–2985, Dec. 1996.

[12] G. Chang, J. P. Hespanha, A. S. Morse, M. Netto, and R. Ortega. Supervisory field-oriented control of induction motors with uncertain rotor resistance. *Int. J. of Adaptive Control and Signal Processing* Special Issue on Switching and Logic, 15(3):353–375, May 2001.

[13] M. H. Chang and E. J. Davison. Switching control of a family of plants. In *Proc. of the 1995 Amer. Contr. Conf.*, volume 1, pages 1015–1020, June 1995.

[14] S. J. Cusumano and K. Poolla. Adaptive control of uncertain systems: A new approach. In *Proc. of the AACC*, pages 355–359, 1988.

[15] M. A. Dahleh and M. M. Livstone. A unified framework for identification and control. In Francis and Tannenbaum [17], pages 196–214.

[16] A. K. El-Sakkary. The gap metric: Robustness of stabilization of feedback systems. *IEEE Trans. Auto. Control*, AC-30(3):240–247, Mar. 1985.

[17] B. A. Francis and A. R. Tannenbaum, editors. *Feedback Control, Nonlinear Systems and Complexity.* Number 202 in Lecture Notes in Control and Information Sciences. Springer-Verlag, London, 1995.

[18] R. A. Freeman and P. V. Kokotović. Inverse optimality in robust stabilization. *SIAM J. Contr. Optimization*, 34(4):1365–1391, July 1996.

[19] R. A. Freeman and P. V. Kokotović. *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques.* Systems & Control: Foundations & Applications. Birkhäuser, Boston 1996, 1996.

[20] M. Fu. Minimum switching control for adaptive tracking. In *Proc. of the 35th Conf. on Decision and Contr.*, pages 3749–3754, Dec. 1996.

[21] M. Fu and B. Barmish. Adaptive stabilization of linear systems via switching control. *IEEE Trans. on Automat. Contr.*, 31:1097–1103, Dec. 1986.

[22] S. Fujii, J. P. Hespanha, and A. S. Morse. Supervisory control of families of noise suppressing controllers. In *Proc. of the 37th Conf. on Decision and Contr.*, pages 1641–1646, Dec. 1998.

[23] T. T. Georgiou, A. Pascoal, and P. P. Khargonekar. On the robust stabilizability of uncertain linear time-invariant plants using nonlinear time-varying controllers. *Automatica*, 35(5):617–624, Sept. 1987.

[24] J. P. Hespanha. *Logic-Based Switching Algorithms in Control.* PhD thesis, Yale University, New Haven, CT, 1998.

[25] J. P. Hespanha, D. Liberzon, and A. S. Morse. Logic-based switching control of a nonholonomic system with parametric modeling uncertainty. *Syst. & Contr. Lett.* Special Issue on Hybrid Systems, 38(3):167–177, Nov. 1999.

[26] J. P. Hespanha, D. Liberzon, and A. S. Morse. Bounds on the number of switchings with scale-independent hysteresis: Applications to supervisory control. In *Proc. of the 39th Conf. on Decision and Contr.*, volume 4, pages 3622–3627, Dec. 2000.

[27] J. P. Hespanha, D. Liberzon, and A. S. Morse. Overcoming the limitations of adaptive control by means of logic-based switching. To appear in *Syst. & Contr. Lett.*, 2002.

[28] J. P. Hespanha, D. Liberzon, and A. S. Morse. Supervision of integral-input-to-state stabilizing controllers. *Automatica*, 38(8):1327–1335, Aug. 2002.

[29] J. P. Hespanha, D. Liberzon, and A. S. Morse. Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39(2), Feb. 2003. To appear.

[30] J. P. Hespanha, D. Liberzon, A. S. Morse, B. D. O. Anderson, T. S. Brinsmead, and F. de Bruyne. Multiple model adaptive control, part 2: switching. *Int. J. of Robust and Nonlinear Control* Special Issue on Hybrid Systems in Control, 11(5):479–496, Apr. 2001.

[31] J. P. Hespanha and A. S. Morse. Certainty equivalence implies detectability. *Syst. & Contr. Lett.*, 36(1):1–13, Jan. 1999.

[32] J. P. Hespanha and A. S. Morse. Stability of switched systems with average dwell-time. In *Proc. of the 38th Conf. on Decision and Contr.*, pages 2655–2660, Dec. 1999.

[33] J. P. Hespanha and A. S. Morse. Supervisory control of integral-input-to-state stabilizing controllers. In *Proc. of the 1999 European Contr. Conf.*, Aug. 1999.

[34] J. Hockerman-Frommer, S. R. Kulkarni, and P. J. Ramadge. Controller switching based on output prediction errors. *IEEE Trans. on Automat. Contr.*, 43(5):596–607, May 1998.

[35] A. Isidori. *Nonlinear Control Systems.* Communications and Control Engineering Series. Springer-Verlag, Berlin, 3rd edition, 1995.

[36] A. Karimi, I. D. Landau, and N. Motee. Effects of the design paramters of multimodel adaptive control on the performance of a flexible transmission system. *Int. J. of Adaptive Control and Signal Processing* Special Issue on Switching and Logic, 15(3):335–352, May 2001.

[37] H. K. Khalil. *Nonlinear Systems.* Prentice-Hall, New York, 2nd edition, 1996.

[38] R. Kosut, M. Lau, and S. Boyd. Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Trans. on Automat. Contr.*, 37(7):929–941, July 1992.

[39] R. L. Kosut. Uncertainty model unfalsification: A system identification paradigm compatible with robust control design. In *Proc. of the 34th Conf. on Decision and Contr.*, pages 3492–3497, Dec. 1995.

[40] R. L. Kosut. Iterative unfalsified adaptive control: Analysis of the disturbance-free case. In *Proc. of the 1999 Amer. Contr. Conf.*, pages 566–570, June 1999.

[41] M. Krichman, E. D. Sontag, and Y. Wang. Input-output-to-state stability. *SIAM J. Contr. Optimization*, 39:1874–1928, 2001.

[42] M. Krstić, I. Kanellakopoulos, and P. Kokotović. *Nonlinear and Adaptive Control Design.* Adaptive and Learning Systems for Signal Processing, Communications, and Control. John Wiley & Sons, New York, 1995.

[43] M. Krstic and Z. Li. Inverse optimal design of input-to-state stabilizing nonlinear controllers. *IEEE Trans. on Automat. Contr.*, 43:336–350, 1998.

[44] S. R. Kulkarni and P. J. Ramadge. Model and controller selection policies based on output prediction errors. *IEEE Trans. on Automat. Contr.*, 41(11):1594–1604, Nov. 1996.

[45] H. Kwakernaak and R. Sivan. *Linear optimal control systems.* Wiley Interscience, New York, 1972.

[46] D. Liberzon, J. P. Hespanha, and A. S. Morse. Hierarchical hysteresis switching. In *Proc. of the 39th Conf. on Decision and Contr.*, volume 1, pages 484–489, Dec. 2000.

[47] D. Liberzon and A. S. Morse. Basic problems in stability and design of switched systems. *IEEE Contr. Syst. Mag.*, 19(5):59–70, Oct. 1999.

[48] D. Liberzon, A. S. Morse, and E. D. Sontag. Output-input stability and minimum-phase nonlinear systems. To appear in IEEE Trans. on Automat. Control, 2001.

[49] D. Liberzon, E. D. Sontag, and Y. Wang. Universal construction of feedback laws achieving iss and integral-iss disturbance attenuation. Submitted to Syst. & Contr. Lett., May 2001.

[50] B. Mårtenson. The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization. *Syst. & Contr. Lett.*, 6(2):87–91, Feb. 1985.

[51] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne. Design issues in adaptive control. *IEEE Trans. on Automat. Contr.*, 33(1):50–58, Jan. 1988.

[52] D. Miller. *Adaptive Control of Uncertain Systems*. PhD thesis, 1989.

[53] D. Miller and E. J. Davison. An adaptive controller which provides an arbitrary good transient and steady-state response. *IEEE Trans. on Automat. Contr.*, 36(1):68–81, Jan. 1991.

[54] A. S. Morse. Towards a unified theory of parameter adaptive control: tunability. *IEEE Trans. on Automat. Contr.*, 35(9):1002–1012, Sept. 1990.

[55] A. S. Morse. Towards a unified theory of parameter adaptive control—part II: Certainty equivalence and implicit tuning. *IEEE Trans. on Automat. Contr.*, 37(1):15–29, Jan. 1992.

[56] A. S. Morse. Control using logic-based switching. In A. Isidori, editor, *Trends in Control: An European Perspective*, pages 69–113. Springer-Verlag, London, 1995.

[57] A. S. Morse. Supervisory control of families of linear set-point controllers—part 1: exact matching. *IEEE Trans. on Automat. Contr.*, 41(10):1413–1431, Oct. 1996.

[58] A. S. Morse, editor. *Control Using Logic-Based Switching*. Number 222 in Lecture Notes in Control and Information Sciences. Springer-Verlag, London, 1997.

[59] A. S. Morse. Supervisory control of families of linear set-point controllers—part 2: robustness. *IEEE Trans. on Automat. Contr.*, 42(11):1500–1515, Nov. 1997.

[60] A. S. Morse. A bound for the disturbance-to-tracking-error gain of a supervised set-point control system. In D. N. Cyrot, editor, *Perspectives in Control—Theory and Application*, pages 23–41. Springer-Verlag, 1998.

[61] A. S. Morse, D. Q. Mayne, and G. C. Goodwin. Applications of hysteresis switching in parameter adaptive control. *IEEE Trans. on Automat. Contr.*, 37(9):1343–1354, Sept. 1992.

[62] E. Mosca, F. Capecchi, and A. Casavola. Designing predictors for MIMO switching supervisory control. *Int. J. of Adaptive Control and Signal Processing* Special Issue on Switching and Logic, 15(3):265–286, May 2001.

[63] K. S. Narendra and J. Balakrishnan. Adaptive control using multiple models. *IEEE Trans. on Automat. Contr.*, 42(2):171–187, Feb. 1997.

[64] K. S. Narendra and O. A. Driollet. Stochastic adaptive control using multiple models for improved performance in the presence of random disturbances. *Int. J. of Adaptive Control and Signal Processing* Special Issue on Switching and Logic, 15(3):297–317, May 2001.

[65] A. D. B. Paice and A. J. van der Schaft. The class of stabilizing nonlinear plant controller pairs. *IEEE Trans. on Automat. Contr.*, 41(5):634–645, May 1996.

[66] F. M. Pait. *Achieving Tunability in Parameter-Adaptive Control*. PhD thesis, Yale University, May 1993.

[67] F. M. Pait and F. Kassab, Jr. Parallel algorithms for adaptive control: Robust stability. In Morse [58], pages 262–276.

[68] F. M. Pait and F. Kassab, Jr. On a class of switched, robustly stable, adaptive systems. *Int. J. of Adaptive Control and Signal Processing* Special Issue on Switching and Logic, 15(3): 213–238, May 2001.

[69] F. M. Pait and A. S. Morse. A cyclic switching strategy for parameter-adaptive control. *IEEE Trans. on Automat. Contr.*, 39(6):1172–1183, June 1994.

[70] J. W. Polderman and I. Mareels. *Adaptive Systems—An Introduction*. Birkäuser, 1996.

[71] M. Prandini and M. C. Campi. Logic-based switching for the stabilization of stochastic systems in the presence of unmodeled dynamics. In *Proc. of the 40th Conf. on Decision and Contr.*, volume 1, pages 393–398, Dec. 2001.

[72] W. J. Rugh. *Linear Systems Theory*. Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[73] M. G. Safonov. Focusing on the knowable: Controller invalidation and learning. In Morse [58], pages 224–233.

[74] M. G. Safonov and T.-C. Tsao. The unfalsified control concept and learning. In *Proc. of the 33rd Conf. on Decision and Contr.*, pages 2819–2824, Dec. 1994.

[75] M. G. Safonov and T.-C. Tsao. The unfalsified control concept: A direct path from experiment to controller. In Francis and Tannenbaum [17], pages 196–214.

[76] M. G. Safonov and T.-C. Tsao. The unfalsified control concept and learning. *IEEE Trans. on Automat. Contr.*, 42(6):843–847, June 1997.

[77] A. V. D. Schaft and H. Schumacher. *An Introduction to Hybrid Dynamical Systems*. Number 251 in Lecture Notes in Control and Information Sciences. Springer-Verlag, London, 2000.

[78] J. S. Shamma and K. Poolla. Optimal asymptotic robust performance through logic-based switching. In Morse [58], pages 201–207.

[79] E. D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Trans. on Automat. Contr.*, 34(4):435–443, Apr. 1989.

[80] E. D. Sontag. Comments on integral variants of ISS. *Syst. & Contr. Lett.*, 34:93–100, 1998.

[81] E. D. Sontag and A. Teel. Changing supply functions in input/state stable systems. *IEEE Trans. on Automat. Contr.*, 40(8):1476–1478, Aug. 1995.

[82] E. D. Sontag and Y. Wang. On characterizations of the input-to-state stability property. *Syst. & Contr. Lett.*, 24:351–359, 1995.

[83] E. D. Sontag and Y. Wang. On characterizations of the input-to-state stability property with respect to compact sets. In *Proc. IFAC Non-linear Control Systems Design Symposium (NOLCOS'95)*, pages 226–231, June 1995.

[84] E. D. Sontag and Y. Wang. Detectability of nonlinear systems. In *Proc. of the 30th Annual Conf. on Information Sciences and Syst.*, volume 2, pages 1031–1036, Mar. 1996.

[85] E. D. Sontag and Y. Wang. Output-to-state stability and detectability of nonlinear systems. *Syst. & Contr. Lett.*, 29(5):279–290, 1997.

[86] G. Vinnicombe. Frequency domain uncertainty and the graph topology. *IEEE Trans. on Automat. Contr.*, 38(9):1371–1383, Sept. 1993.

[87] Y. Wang. A converse Lyapunov theorem with applications to ISS-disturbance attenuation. In *Proc. 13th IFAC World Congress*, volume E, pages 79–84, 1996.

[88] B. R. Woodley, J. P. How, and R. L. Kosut. Direct unfalsified controller design – solution via convex optimization. In *Proc. of the 1999 Amer. Contr. Conf.*, pages 3302–3306, June 1999.

[89] G. Zames. Towards a general complexity-based theory of identification and adaptive control. In Morse [58], pages 208–223.

[90] G. Zames and A. K. El-Sakkary. Unstable systems and feedback: The gap metric. *Proc. Allerton Conf.*, pages 380–385, 1980.

[91] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice Hall, New Jersey, 1996.