Robust Stability for Multiple Model Adaptive Control: Part I—The Framework

Dominic Buchstaller and Mark French, *Member, IEEE*

***Abstract*—An axiomatic framework providing robust stability and performance bounds for a wide class of Estimation based Multiple Model Switched Adaptive Control (EMMSAC) algorithms is developed. The approach decouples development of both the atomic control designs and the estimation processes thus permitting the usage of standard controller design and optimization approaches for these components. The framework is shown to give tractable algorithms for MIMO LTI plants, and also for some classes of nonlinear systems (for example, an integrator with input saturation). The gain bounds obtained have the key feature that they are functions of the complexity of the underlying uncertainty as described by metric entropy measures. For certain important geometries, such as a compact parametric uncertainties, the gain bounds are independent of the number of plant models (above a certain threshold) which are utilized in the implementation. Design processes are described for achieving a suitable sampling of the plant uncertainty set to create a finite candidate plant model set (whose size is also determined by a metric entropy measure) which achieves a guaranteed robustness/performance.**

***Index Terms*—Estimation based multiple model switched adaptive control (EMMSAC).**

I. Introduction

A multiple model adaptive control scheme consists of a set of candidate plant models, each with an associated controller, coupled with an on-line process for ranking the ability of each model to explain the observed signals. An on-line switching logic selects an appropriate controller based on this ranking. Typically the ranking process is realised via monitoring the output errors of a bank of observers or Kalman Filters. Despite strong advances, key challenges for this approach include the development of a strong robust stability framework and the development of a principled design theory. This paper provides a framework for both robust stability and a principled approach to synthesis, as a step towards addressing these challenges.

The outcome of any design process in MMAC must include the construction of a candidate plant model set. A designer is necessarily confronted with the following design questions:

1) How many plant models are needed?
2) How should the plant models be (geometrically) distributed over the uncertainty set?

Manuscript received October 17, 2012; revised July 31, 2014; accepted January 5, 2015. Date of publication October 19, 2015; date of current version February 25, 2016. Recommended by Associate Editor A. Astolfi.

D. Buchstalker is with Siemens Corporate Technology, Erlangen 91058, Germany (e-mail: dbuchstaller@web.de).

M. French is with the School of Electronics and Computer Science University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: mcf@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TAC.2015.2492518

Furthermore, these design questions should be addressed in a framework that addresses the following questions:

3) What are the robustness guarantees?
4) How can a conservative design be avoided?

These four questions are central to both the development of the EMMSAC framework given in this paper and the proof of the gain bounds given in the companion paper [7]. Questions 1)–3) are addressed in this paper and Question 4) is addressed the second paper [7]. An axiomatic approach (based on the initial work [11]) is used to describe and prove properties of a wide family of MMAC algorithms; the analysis differs strongly from all preceding contributions. In particular, a novel optimization viewpoint of the model ranking process due to [9], [25], is the key insight which underpins the analysis.

The first two questions are considered to be key outstanding issues in the field of multiple model control, e.g., see [1], [2], [8]. For example, in [8] the authors ask: "How to divide the initial parameter uncertainty set into N smaller subsets, how large should N be, etc." and in [1]: "How many plants (models) should be chosen, how does one choose a representative set of plants (plant model set), etc.". A major issue with previous performance bounds developed for MMAC is their exponential scaling with the number of plant models, irrespective of the geometry [14], with the single important exception of the structured switching mechanism [17] which avoids this scaling problem when the uncertainty is a compact continuum. A key contribution of this paper is to give gain bounds, which, for certain important geometries, are independent of the size of the candidate plant model set, depending instead on the complexity of the uncertainty set. In turn the characterization of this complexity, together with optimization of the gain bounds, leads to a principled selection of the distribution of plant models over the uncertainty set. An important consequence is that a structured uncertainty described by a compact continuum can always be robustly stabilised by a EMMSAC design with a sufficiently large number of plant models, and further refinements of the candidate plant model set do not degrade the gain bounds. A pragmatic conclusion is that beyond a certain threshold it does not matter how many plant models are utilized in an implementation: there is no loss of performance guarantees through using "too many" plant models; hence the control designer can use as many plant models as the real-time implementation can support. Foruncertainty sets described by continua, these issues have been central goals of the MMAC literature, see e.g., [8], [14], [17], [20].

The third question represents a goal within the field of adaptive control which has been elusive for decades. Ever since the

publication of the Rohrs example [21] it has been known that adaptive controllers can induce severe instabilities in practice despite ideal nominal behaviour. Much of the effort of the 1980's was concerned with modification of classical adaptive control schemes to achieve limited robustness guarantees (typically restricted to unmodelled dynamics of an additive or multiplicative type). More recently techniques introduced from nonlinear input-output stability theory (involving the nonlinear gap metric) have been utilized to revisit the classical schemes to provide more satisfactory robustness guarantees (e.g., for gap or co-prime factor uncertainty models) [10]; and this approach to robust stability analysis forms the approach taken in this paper. It is of particular importance for MMAC that robustness results incorporate uncertainty at both low and high frequencies: uncertainties arise naturally both in the low frequency range (due to parametric mismatch between the true plant and the finite collection of candidate plant models) and in the high frequency range (due to unmodelled dynamics). Addressing question 3 pervades the entire approach, and Theorem 18 guarantee stabilization of both structured uncertainty sets (for example a parametric uncertainty) together with gap perturbations around these sets, (the size of the perturbation being determined by the robust stability margin).

The axiomatic framework considered leads to a unified treatment of large classes of algorithms, with the important feature that the estimation component of the design is completely decoupled from the underlying control design process for each candidate plant model. In particular the control design process inherits no structural constraints and can be implemented via any standard controller structure (PID, H_∞, etc.). The estimation process, which determines the candidate plant model ranking, encompasses both the Kalman Filter bank of the historical algorithms, but also a variety of "optimization based" processes, where finite horizon optimizations are aggregated to realise an residual which ranks the quality of each model [9], [25]. The optimizations and controller designs can be cast in a variety of different signal spaces yielding a wide class of algorithms. Although oriented towards the case of MIMO LTI systems, the analysis can also apply to broad classes of nonlinear controllers which achieve closed loop gain stability. Whilst the nonlinear optimizations within the estimator are not tractable in real-time in general, there are important classes of nonlinear systems which do result in tractable algorithms: here we illustrate the case of an integrator with input saturation.

Arguments for the benefits of MMAC compared to other adaptive approaches have been made previously, see, e.g., [6], [8] and [13]; these apply equally to the EMMSAC class of algorithms. Additionally, the ability to give stabilization guarantees over uncertainty sets described by continua, and by unbounded sets [7], means that the domain of MMAC now encompasses all the uncertainty sets considered in classical (linear) adaptive control. Furthermore, MMAC inherits none of the standard adaptive control requirements of convex uncertainty sets, or parameterizations limited to particular forms, which limit the problem domain of classical adaptive control. MMAC can deal with plants where the sign of the high frequency gain is unknown; such plant pairs are not simultaneously stabilizable by LTI design compensators, nor do classical adaptive algorithms have satisfactory performance (e.g., the Nussbaum universal controller). Unfalsified control [3], [23], [26], [29] is an alternative approach to switching between compensators wherein robust stability follows from a weak feasibility assumption. It is likely that schemes incorporating both MMAC and unfalsified concepts will prevail: see for example [3] for work in this direction; on the other hand, a version of dynamic EMMSAC considered in [7] has unfalsified characteristics.

Since adaptive control is necessarily a theory of controlling processes with large uncertainties, the framework is cast in a setting in which closed loops are shown to be robust w.r.t. a combination of large structured uncertainty sets and small unstructured uncertainties. A description of the complexity of a large uncertainty set is given, in terms of metric covers and entropy. The resulting gain bounds are shown to be functions of the complexity of the uncertainty set. Key to the development of the theory is to analyze not the actual realization of the algorithm, but rather a potentially infinite dimensional object which comprises of (typically) a continuum of estimators and potentially also a continuum of controllers (e.g., a MMAC controller based on an infinite number of candidate plant models in one to one correspondence with the structured uncertainty set and with one estimator and one controller associated to each candidate plant). A reduction theory is given to approximate this by a finite dimensional realization (e.g., a MMAC controller based on a finite number of controllers and estimators), and bounds are given to relate the performance of the finite dimensional realization to that of the infinite dimensional object. The necessary complexity of the finite dimensional controller (i.e., the size of the candidate plant set) is bounded in terms of the metric entropy of the uncertainty, and the resulting realizable algorithms are then proven to be robust to both the uncertainty set sampling error and the underlying unstructured uncertainty.

The objective of design therefore arises as the question of how to achieve the finite dimensional realization via an appropriate sampling of the structured uncertainty sets to arrive at a finite number of nominal candidate plant models and hence an implementable controller. Here the construction is similar to the explicit, albeit heuristic, design procedures of [2], [8], [14], [20] which construct a candidate plant model set based on covers generated from the atomic closed loop performance of matching plant and controller pairs.

The paper is structured as follows. In Section II we introduce the setting and notation. The structure of the EMMSAC algorithm is given in Section III, and the axiomatic requirements of the estimation process are given in Section IV together with important examples of estimators. The main result which establishes complexity dependent gain bounds and robust stability is given in Section V. The proof of this result, which is long and involved, is given in the sequel [7]. Section VI presents a number of consequences of the theorem as well as robust stability certificates.

II. Preliminaries

For $0 \leq a \leq b,\ a, b \in \mathbb{Z}$ let $[a,b] := \{x \in \mathbb{Z} \mid a \leq x \leq b\}$, $[a,b) := \{x \in \mathbb{Z} \,|\, a \leq x < b\}$ and define $|[a,b]| := b-a+1$ and $|[a,b)| := b-a$. For a signal $v \in \mathcal{S}$ we then define

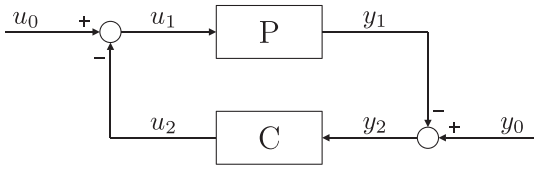

Fig. 1. Closed loop $[P, C]$.

the restriction of v over the interval $I = [c, d]$ by $v|_I := (v(c), \ldots, v(d))$ where $c \le d,\ c, d \in \mathbb{Z}$, and similarly for $I = [c, d)$. Denote the collection of all maps $\mathcal{S} := \mathrm{map}(\mathbb{Z}, \mathbb{R}^h)$ and let $\mathcal{S}|_{[a,b]} := \mathrm{map}([a, b], \mathbb{R}^h)$. Let $\mathscr{T}_t : \mathcal{S} \cup_{b\in\mathbb{Z}} \mathcal{S}|_{[0,b]} \to \mathcal{S},\ t \in \mathbb{Z}$ denote the truncation operator defined by

$$(\mathscr{T}_t v)(\tau) = \begin{cases} v(\tau) & \text{if } \tau \in \mathrm{dom}(v),\ \tau \le t \\ 0 & \text{otherwise.} \end{cases}$$

For $x \in \mathcal{S}$ define the norms $\|x\| = \|x\|_r = (\sum_{i\in\mathrm{dom}(x)} |x(i)|^r)^{1/r},\ \ 1 \le r < \infty,\ \ \|x\| = \|x\|_\infty = \sup_{i\in\mathrm{dom}(x)} |x(i)|$. We consider signal spaces $\mathcal{V} \subset \mathcal{S}$, interval spaces $\mathcal{V}|_{[a,b]}$ and extended signal spaces $\mathcal{V}_e \subset \mathcal{S}$

$$\begin{aligned} \mathcal{V} &:= \{v \in \mathcal{S} \mid v(-t) = 0,\ \forall t \in \mathbb{N};\ \|v\| < \infty\} \\ \mathcal{V}_{[a,b]} &:= \left\{v \in \mathcal{S}|_{[a,b]} \,\middle|\, \exists x \in \mathcal{V}\ s.t.\ v = x|_{[a,b]}\right\}. \\ \mathcal{V}_e &:= \{v \in \mathcal{S} \mid \forall t \in \mathbb{Z} : \mathscr{T}_t v \in \mathcal{V}\}. \end{aligned} \tag{2.1}$$

We take $\mathcal{V} = l_r$ to be defined by (2.1) with $\|\cdot\| = \|\cdot\|_r$. The input and output signal spaces are defined as: $\mathcal{U} := \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{m} = \mathcal{V}^m, \mathcal{Y} := \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{o} = \mathcal{V}^o$, and let $\mathcal{W} := \mathcal{U} \times \mathcal{Y}$. Given a plant $P : \mathcal{U}_e \to \mathcal{Y}_e$ satisfying $P(0) = 0$ and a controller $C : \mathcal{Y}_e \to \mathcal{U}_e$ satisfying $C(0) = 0$ the closed loop system $[P, C]$ in Fig. 1 is defined by

$$y_1 = Pu_1 \tag{2.2}$$
$$u_0 = u_1 + u_2, \quad y_0 = y_1 + y_2 \tag{2.3}$$
$$u_2 = Cy_2. \tag{2.4}$$

Here $w_i = (u_i, y_i)^\top \in \mathcal{W}_e$ represents the plant input and output $(i = 1)$, disturbances $(i = 0)$ and observations $(i = 2)$.

$[P, C]$ is said to be well-posed if for all $w_0 \in \mathcal{W}$ there exists a unique solution $(w_1, w_2) \in \mathcal{W}_e \times \mathcal{W}_e$. Note that linear switched systems are well-posed. For a well-posed system $[P, C]$ we define the closed loop operator

$$\Pi_{P//C} : \mathcal{W} \to \mathcal{W}_e \times \mathcal{W}_e : w_0 \mapsto (w_1, w_2).$$

$[P, C]$ is said to be gain stable if there exists a $M > 0$ s.t.:

$$\sup_{w_0\in\mathcal{W},\ w_0\neq 0} \frac{\|\Pi_{P//C} w_0\|}{\|w_0\|} = \|\Pi_{P//C}\| < M < \infty.$$

Define $\mathcal{P}_{\mathrm{LTI}}$ to be the set of all $p = (A, B, C, D) \in \cup_{n\ge 1} \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{o\times n} \times \mathbb{R}^{o\times m}$ such that p is minimal. Let

$$P_p : \mathcal{U}_e \to \mathcal{Y}_e,\ u_1^p \mapsto y_1^p,\ p \in \mathcal{P}_{\mathrm{LTI}} \tag{2.5}$$

be defined by

$$x_p(k+1) = Ax_p(k) + Bu_1^p(k) \tag{2.6}$$
$$y_1^p(k) = Cx_p(k) + Du_1^p(k) \tag{2.7}$$
$$x_p(-k) = 0, \quad k \in \mathbb{N}. \tag{2.8}$$

Note that since $x_p(-k) = 0$ for all $k \in \mathbb{N}$ it follows that $y_1^p(-k) = (P_p u_1^p)(-k) = 0$ for all $k \in \mathbb{N}$. Also define

$$\bar{\mathcal{P}}_{\mathrm{LTI}} := \{(A, B, C, D) \in \mathcal{P}_{\mathrm{LTI}} \mid D = 0\}. \tag{2.9}$$

Similarly, let $\mathcal{C}_{\mathrm{LTI}}$ to be the set of all $c = (A, B, C, D) \in \cup_{n\ge 1} \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times o} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times o}$ such that c is minimal, and define the control operator

$$C_c : \mathcal{Y}_e \to \mathcal{U}_e : y_2^c \mapsto u_2^c,\ c \in \mathcal{C}_{\mathrm{LTI}} \tag{2.10}$$

analogously to equations (2.6)–(2.8) and let

$$\bar{\mathcal{C}}_{\mathrm{LTI}} := \{(A, B, C, D) \in \mathcal{C}_{\mathrm{LTI}} \mid D = 0\}. \tag{2.11}$$

The collection of bounded pairs $(u_1^p, y_1^p)^\top \in \mathcal{W}$ compatible with the plant $P_p,\ p \in \mathcal{P}$ where $\mathcal{P}$ is an indexing set (for example $\mathcal{P} = \mathcal{P}_{\mathrm{LTI}}$) forms the graph $\mathcal{M}_p \subset \mathcal{W}$

$$\mathcal{M}_p = \left\{ v \in \mathcal{W} \,\middle|\, \begin{array}{c} \exists (u_1^p, y_1^p)^\top \in \mathcal{W} \text{ s.t. } P_p u_1^p = y_1^p, \\ v = (u_1^p, y_1^p)^\top \end{array} \right\}.$$

An appropriate measure of the distance between graphs defines the nonlinear gap as follows. Let $\mathcal{O}_{p_1,p_2} = \{\Phi : \mathcal{M}_{p_1} \to \mathcal{M}_{p_2} \mid \Phi \text{ is causal, bijective, and } \Phi(0) = 0\}$. Define the non-linear directed gap between $p_1, p_2 \in \mathcal{P}$ by

$$\vec{\delta}(p_1, p_2) := \inf_{\Phi\in\mathcal{O}_{p_1,p_2}} \sup_{x\in\mathcal{M}_{p_1}\setminus 0,\ k>0} \left(\frac{\|\mathscr{T}_k(\Phi - I)x\|}{\|\mathscr{T}_k x\|} \right)$$

if $\mathcal{O}_{p_1,p_2} \neq \emptyset$, and $\vec{\delta}(p_1, p_2) := \infty$ if $\mathcal{O}_{p_1,p_2} = \emptyset$. Define $\delta(p_1, p_2) = \max\{\vec{\delta}(p_1, p_2), \vec{\delta}(p_2, p_1)\}$. Note that the non-linear gap is a generalization of the standard linear definition via coprime factors [12, Appendix]. In the linear setting, small time delays, multiplicative, inverse multiplicative, parametric and co-prime factor perturbations are all small in the gap. For nonlinear systems, similar relationships hold. The central robust stability theorem is as follows:

Theorem 1: Let $\mathcal{U} = \mathcal{Y} = l_r, 1 \le r \le \infty$. Let $P_{p_1} : \mathcal{U}_e \to \mathcal{Y}_e$, $P_{p_2} : \mathcal{U}_e, \to \mathcal{Y}_e,\ C : \mathcal{Y}_e \to \mathcal{U}_e$ and suppose that the closed loops $[P_{p_i}, C],\ i = \{1, 2\}$ are well-posed. Let the closed loop $[P_{p_1}, C]$ be gain stable. If

$$\vec{\delta}(p_1, p_2) < \|\Pi_{P_{p_1}//C}\|^{-1} = b_{P_{p_1},C}$$

then the closed loop system $[P_{p_2}, C]$ is gain stable and

$$\|\Pi_{P_{p_2}//C}\| \le \|\Pi_{P_{p_1}//C}\| \frac{1 + \vec{\delta}(p_1, p_2)}{1 - \|\Pi_{P_{p_1}//C}\| \vec{\delta}(p_1, p_2)}.$$

Proof: The proof can be found in [12]. ■

Throughout the paper we consider $\mathcal{P}$ to be a topological space, with topology determined by the gap δ.

As $v \in \mathcal{V}$ has the property that $v(-t) = 0$ for $t \in \mathbb{N}$ and since we are requiring $P(0) = C(0) = 0$, it follows that LTI state space models for P and C are required to have their initial conditions set to zero $(x(0) = 0)$. This is a standard assumption in the input/output setting. However, it should be noted that the discrete-time analogue of the approach via input, output injection of [10, Th. 5.3] means that for linear plants (but with potentially nonlinear controllers, as here) the zero initial condition results also imply stability results for non-zero initial conditions.

III. Estimation-Based Multiple Model Switched Adaptive Control

In this Section we develop the structure of EMMSAC. We introduce the controller design procedure K that assigns a stabilizing controller to every plant model and then describe the structure for switching between these atomic controllers.

A. Finite Horizon Behaviour of the Atomic Closed Loop

The controller design procedure is specified by a map $K : \mathcal{P} \to \mathcal{C}$ where $\mathcal{C}$ is a set parametrizing a collection of controller operators

$$u_2^c = C_c y_2^c \tag{3.12}$$

for $c \in \mathcal{C}$, for example $\mathcal{C} = \mathcal{C}_{\text{LTI}}$. $K : \mathcal{P} \to \mathcal{C}$ is said to be a stabilizing design if $[P_p, C_{K(p)}]$ is gain stable for all $p \in \mathcal{P}$.

Let $\sigma(c),\ c \in \mathcal{C}$ denote the minimum length of the interval that the signal $(u_2^c, y_2^c)^\top$ needs to be observed to uniquely determine the initial condition of C_c, i.e.

$$\sigma(c) = \min \left\{ k \geq 0 : \begin{array}{c} \forall\, l \geq 0, \\ u_2^c = C_c y_2^c,\ \hat{u}_2^c = C_c \hat{y}_2^c, \\ (u_2^c, y_2^c)^\top|_{[l,l+k]} = (\hat{u}_2^c, \hat{y}_2^c)^\top|_{[l,l+k]}, \\ y_2^c = \hat{y}_2^c \Rightarrow u_2^c = \hat{u}_2^c \end{array} \right\} \tag{3.13}$$

Similarly let $\sigma(p),\ p \in \mathcal{P}$ denote the minimum length of the interval that the signal $(u_1^p, y_1^p)^\top$ needs to be observed to uniquely determine the initial condition of P_p. For $p \in \mathcal{P}_{\text{LTI}},\ c \in \mathcal{C}_{\text{LTI}}$ note that $\sigma(p) = n_p - 1$, $\sigma(c) = n_c - 1$ where n_p, n_c are the McMillan degrees of p and c respectively.

We now state two general requirements imposed upon the atomic closed loop systems $[P_p, C_c]$ and $[P_p, C_{K(p)}]$.

Assumption 2: There exist functions $\alpha, \beta : \mathcal{P} \times \mathcal{C} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that the following holds:

1) (Linear growth of $[P_p, C_c]$): Let $p \in \mathcal{P},\ c \in \mathcal{C}$ and suppose $[P_p, C_c]$ is well-posed. Let $t_1, t_2, t_3, t_4 \in \mathbb{N},\ t_1 < t_2 \leq t_3 < t_4$ and $I_1 = [t_1, t_2), I_2 = [t_2, t_3), I_3 = [t_3, t_4)$. Suppose $w_2, w_2^c, w_1^p \in \mathcal{W}_e,\ w_0^p \in \mathcal{W}$ satisfy the equations

$$y_1^p = P_p u_1^p,\ u_2^c = C_c y_2^c,\ u_0^p = u_1^p + u_2,\ y_0^p = y_1^p + y_2$$

on $I_1 \cup I_2 \cup I_3$. Suppose that either

$$\begin{array}{rl} w_2^c|_{I_1} = 0,\ w_2^c|_{I_2 \cup I_3} = w_2|_{I_2 \cup I_3}, & \text{or} \\ w_2^c|_{I_1 \cup I_2 \cup I_3} = w_2|_{I_1 \cup I_2 \cup I_3} & \end{array}$$

where

$$|I_1| = t_2 - t_1 \geq \max\{\sigma(p), \sigma(c)\}. \tag{3.14}$$

Then, in both cases

$$\begin{aligned} \|w_2|_{I_3}\| \leq\ & \alpha(p, c, |I_2|, |I_3|)\|w_2|_{I_1}\| \\ & + \beta(p, c, |I_2|, |I_3|)\|w_0^p|_{I_1 \cup I_2 \cup I_3}\|. \end{aligned} \tag{3.15}$$

2) (Stability of $[P_p, C_{K(p)}]$): Let $p \in \mathcal{P}$ and $x \in \mathbb{N}$. Then

$$\alpha(p, K(p), a, x) \to 0 \text{ as } a \to \infty \tag{3.16}$$

and α is monotonic in a.

Note that the monotonicity requirement in the second assumption follows without loss of generality since any function α satisfying (3.16) can be dominated point-wise by a monotonic function $\hat{\alpha}$ satisfying (3.16). Assumption 2 is interpreted as follows. The choice $w_2^c|_{I_1} = 0$ corresponds to an initialization of the controller to zero at time t_2 and the choice $w_2^c|_{I_1} = w_2|_{I_1}$ corresponds to continued closed loop operation of the same controller. We expect to be able to bound future signals $\|w_2|_{I_3}\|$ by some function of the size of the system's initial conditions, determined by $\|w_1|_{I_1}\|$, $\|w_2|_{I_1}\|$, and the system's input $w_0^p|_{I_1 \cup I_2 \cup I_3}$ for any well-posed closed loop system $[P_p, C_c]$. This is reflected by (3.15). However $w_1|_{I_1},\ w_2|_{I_1}$ can only be interpreted as an initial condition if the interval I_1 is sufficiently long. This is reflected by (3.13), (3.14). For $K : \mathcal{P} \to \mathcal{C}$ where $(\mathcal{P}, \mathcal{C}) \subset (\bar{\mathcal{P}}_{\text{LTI}}, \mathcal{C}_{\text{LTI}}) \cap (\mathcal{P}_{\text{LTI}}, \bar{\mathcal{C}}_{\text{LTI}})$ it can be shown that Assumption 2(1) holds as follows. $\Pi_{C_c//P_p}$ is linear and hence a state space representation (A, B, C, D), with state x, so $w_2(t) = CA^t x_0 + C(\Pi_{C_c//P_p} w_0)(t) + Dw_0(t)$, Observability matrices M, N can be constructed from (A, B, C, D) such that $x(0) = Mw_0|_{I_1} + Nw_2|_{I_1}$ since w_0 and w_2 are the inputs and outputs of $\Pi_{P//C}$. Hence: $\|w_2|_{I_3}\| \leq \|C(A^{t_3}, \ldots, A^{t_4})x_0 + C(\Pi_{C_c//P_p} w_0)|_{I_3} + (Dw_0)|_{I_3}\|$, and consequently inequality (3.15) holds with

$$\alpha = \|M, N\|\|A^{t_3}, \ldots, A^{t_4}\| \tag{3.17}$$

$$\beta = \|\Pi_{C_c//P_p}\| + \|M, N\|\|A^{t_3}, \ldots, A^{t_4}\|. \tag{3.18}$$

Tighter expressions for α and β can be found in [6]. If additionally $K : \mathcal{P} \to \mathcal{C}$ is an (asymptotic) stabilizing design, it can be shown that Assumption 2(2) holds since asymptotic stability implies l_r stability $1 \leq r \leq \infty$.

B. The Switching Algorithm

We now formally introduce the estimation-based switching operator $S = DM(X, G)$ where G is the plant-generating operator which specifies which candidate plants can be considered at each step, X is the residual operator which returns a scalar for each plant which assesses the quality of the model, M is the minimization operator which returns the plant model with the smallest residual and D is the delay operator whose role is to prevent rapid destabilizing switches. The dynamic versions of EMMSAC, as motivated in the introduction, are characterised by a time varying set of candidate plant models. These are specified by the notion of a plant generating operator defined as follows. Let $\mathcal{P}^*$ be the powerset of $\mathcal{P}$. Then:

Definition 3: A causal map $Q : \mathcal{W}_e \to \text{map}(\mathbb{N}, \mathcal{P}^* \setminus \emptyset)$ is said to be a plant-generating operator. We define $\mathcal{P}^Q$ is the union of all plant model sets represented by Q

$$\mathcal{P}^Q := \bigcup_{w_2 \in \mathcal{W}_e} \bigcup_{k \in \mathbb{N}} Q(w_2)(k) \subset \mathcal{P}.$$

Q is said to be finite if $Q(w_2)(k)$ is a finite set for all $w_2 \in \mathcal{W}$, $k \in \mathbb{N}$, constant if $Q(w_2)(i) = Q(w_2)(j)$, for all $w_2 \in \mathcal{W}$, $i, j \in \mathbb{N}$, monotonic if $Q(w_2)(k) \subset Q(w_2)(k+1)$ for all $w_2 \in \mathcal{W}$, $k \in \mathbb{N}$ and compact if $Q(w_2)(k)$ is compact as a subset of $\mathcal{P}$ for all $w_2 \in \mathcal{W},\ k \in \mathbb{N}$. For notational economy we often write $Q(k) := Q(w_2)(k),\ k \in \mathbb{N}$.

Within an EMMSAC algorithm, the candidate plant models $G(w_2)(k)$ which are available for consideration at any time k are determined by an underlying plant-generating operator

$$G : \mathcal{W}_e \to \mathrm{map}(\mathbb{N}, \mathcal{P}^*) \tag{3.19}$$

The dependence on w_2 allows this set of candidate plants to be determined adaptively.

The residual operator is of the form

$$X : \mathcal{W}_e \to \mathrm{map}\left(\mathbb{N}, \mathrm{map}(\mathcal{P}, \mathbb{R}^+)\right) : w_2 \mapsto [k \to (p \mapsto r_p[k])] \tag{3.20}$$

where $r_p[k]$ is said to be the residual of a plant $P_p,\ p \in \mathcal{P}$ at time $k \in \mathbb{N}$. The residual is a scalar that represents the quality assessment of the associated plant model. In classical MMAC it is the residual of the Kalman filter, or the weighted cumulative output error of the associated observer. Alternative residuals based on optimizations are considered in Section IV. Note that all residuals considered will necessarily measure the performance of the models over the full period $[0, k]$.

At time k, the minimizing operator M selects the plant with the smallest residual which is available for switching (i.e., which lies in $G(w_2)(k)$)

$$M : \left(\mathrm{map}\left(\mathbb{N}, \mathrm{map}(\mathcal{P}, \mathbb{R}^+)\right), \mathrm{map}(\mathbb{N}, \mathcal{P}^*)\right) \to \mathrm{map}(\mathbb{N}, \mathcal{P}^*) \tag{3.21}$$

$$[k \mapsto (p \mapsto r_p[k]), k \mapsto G(k)] \mapsto [k \mapsto q_f(k)] \tag{3.22}$$

where

$$q_f(k) := \arg \min_{p \in G(k)} r_p[k], \quad \forall\, k \in \mathbb{N}.3. \tag{3.23}$$

If there are multiple minimizing residuals, an arbitrary ordering on $G(k)$ is imposed *a priori*, i.e., $G(k) = \{p_1, p_2, \ldots, p_n\}$, and $\arg\min_{p \in G(k)} r_p[k]$ is defined to return the parameter $p_i \in G(k)$ with the smallest index i such that $r_{p_i}[k]$ is minimal. Equation (3.23) also includes the implicit assumption that a minimiser exists. In the scenario considered in this paper, whereby G is finite or G is compact and $p \mapsto r_p[k]$ is continuous, this holds.

It would be natural at time k to utilize the controller specified by plant $q_f(k)$. However, there is the potential for instability to occur if the switches are too fast [16], hence the purpose of the delay operator D is to slow down the free switching signal q_f for long enough to prevent the potential for these instabilities. We encode this information into the transition delay function $\Delta : \mathcal{P} \to \mathbb{N}$: to every plant $P_p,\ p \in \mathcal{P}$, we associate a minimum delay $\Delta(p)$ which must elapse before another switch is permitted; the analogue of the transition delay is taken by dwell time switching in other versions of MMAC e.g., [18]. Here the transition delay is plant dependent; this reduces the bounds, for if a uniform delay is utilized, then the delay would be determined by the time-scale of the response of the slowest candidate closed loop $[P_p, C_p]$, which can produce larger than necessary transients when a mismatched controller is switched into the loop. The transition delays are design parameters and the required lower limits on these delays will be determined by Assumption 13.

This leads to the following structure for the delay operator:

$$D : \mathrm{map}(\mathbb{N}, \mathcal{P}) \to \mathrm{map}(\mathbb{N}, \mathcal{P}) \tag{3.24}$$

$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{3.25}$$

where $q(k)$ is defined recursively

$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{3.26}$$

and where $k_s : \mathbb{N} \to \mathbb{N}$ is given by

$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k,\ q(i) \neq q(i-1)\} \tag{3.27}$$

Note that $k_s(k)$ returns the last time up to time $k \in \mathbb{N}$ where the algorithm switches from one plant to another. The switching operator is now given as follows:

$$S = \mathrm{DM}(X, G) : \mathcal{W}_e \to \mathrm{map}(\mathbb{N}, \mathcal{P}^*) : w_2 \mapsto q.$$

Given a control design procedure $K : \mathcal{P} \to \mathcal{C}$, the switching controller

$$C : \mathcal{Y}_e \to \mathcal{U}_e : y_2 \mapsto u_2 \tag{3.28}$$

is then defined via the switching signal q as follows:

$$u_2(k) = C_{K(q(k))}(y_2 - \mathscr{T}_{k_s(k)-1} y_2)(k). \tag{3.29}$$

By the definition of the truncation operator and (3.26), note that $y_2(s) - \mathscr{T}_{k_s(k)} y_2(s)$ is zero for times $s < k_s(k)$. Hence (3.29) ensures a zero initial condition for the atomic controller $C_{K(q(k))}$ when it is switched into closed loop at time $k_s(k)$. Note that if X is causal and G is causal, then S is causal. We therefore have arrived at the closed-loop given by Fig. 2 where all the involved sub systems have been defined. Note that further structure on the residual operator X has also been illustrated (including operators N, E and the signals $d_{p_1}, \ldots d_{p_n}$): see Section IV.

IV. Disturbance Estimation

In this section we will impose conditions on the residual operator X which permits the residual $r_p[k]$ to have the interpretation of being a measure of the size of the disturbance signals $w_0^p = (u_0^p, y_0^p)^\top$ required to explain the observation $w_2 = (u_2, y_2)^\top$ in a manner consistent with the candidate plant P_p on the interval $[0, k]$.

We first formally define the notion of disturbances which are consistent with a plant P_p and an observation on a specified interval $[a, b]$:

Definition 4: Let $a \leq b,\ a, b \in \mathbb{Z}$. The set of weakly consistent disturbance signals $\mathcal{N}_p^{[a,b]}(w_2)$ for a plant $P_p,\ p \in \mathcal{P}$ and the observation $w_2 = (u_2, y_2)^\top$ is defined by

$$\mathcal{N}_p^{[a,b]}(w_2) := \left\{ v \in \mathcal{W}|_{[a,b]} \;\middle|\; \exists (u_0^p, y_0^p)^\top \in \mathcal{W}_e \text{ s.t.} \right.$$
$$\left. \begin{array}{l} \mathscr{R}_{b-a,b} P_p \left(u_0^p - u_2\right) = \mathscr{R}_{b-a,b}(y_0^p - y_2), \\ v = (\mathscr{R}_{b-a,b} u_0^p, \mathscr{R}_{b-a,b} y_0^p) \end{array} \right\}.$$

where the restriction operator $\mathscr{R}_{\sigma,t} : \mathcal{S} \to \mathbb{R}^{h(\sigma+1)}$ extracts a finite window of a signal, i.e., for $\sigma, t \in \mathbb{Z}$

$$\mathscr{R}_{\sigma,t} v := (v(t-\sigma), \ldots, v(t)),\ v \in \mathrm{map}(\mathbb{Z}, \mathbb{R}^h).$$

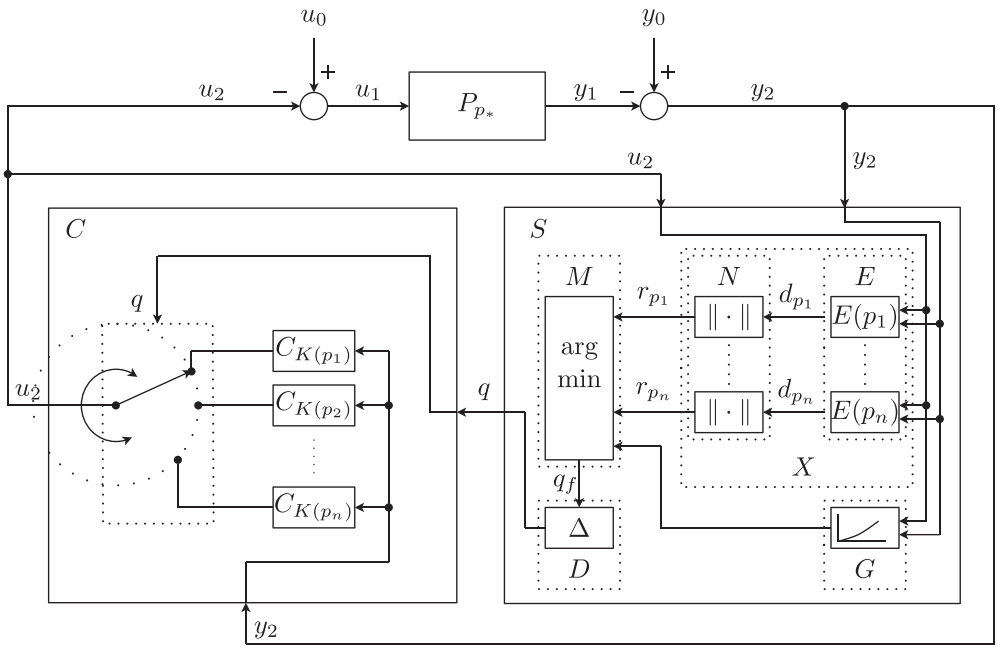

Fig. 2. The EMMSAC structure. The switch S outputs the switching signal q which determines the atomic controller choice $C_{K(q(k))}$. q is generated via the delay operator D from the free switching signal q_f, which in turn is generated from the minimization operator M acting on the residuals $r_p[k]$ from the estimator X, where the admissible plants are detemined by the plant generating operator G. We also illustrate the internal structure of X as in Section IV.

For the remainder of this paper we assume $\mathcal{N}_p^{[a,b]}(w_2)$ is closed and convex for all $a \leq b \in \mathbb{Z},\ w_2 \in \mathcal{W}_e$, noting that if P_p is linear, then this holds.

We now give two examples of residual operators. Let $k, \lambda \in \mathbb{N}$ and $w_2 \in \mathcal{W}_e$, and define the infinite horizon operator

$$X_A(w_2)(k)(p) = r_p^A[k] = \inf\{r \geq 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[0,k]}(w_2)\}. \quad (4.30)$$

Similarly, define the finite horizon operator

$$X_B(w_2)(k)(p) = r_p^B[k] = \left\|r_p^B[k-1], i_p[k]\right\| \quad (4.31)$$

$$i_p[k] = \inf\{r \geq 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\} \quad (4.32)$$

where note that for l_r: $\|a, b\| = (|a|^r + |b|^r)^{1/r}$ if $1 \leq r < \infty$ and $\|a, b\| = \max\{|a|, |b|\}$ if $r = \infty$.

These examples illustrate the EMMSAC approach: control selection is done via assessing the quality of the associated models thorough an identification based procedure: models are assessed on their ability to explain the observed signals (w_2) with the smallest disturbances (w_0). Such a quality assessment lends itself naturally to finite dimensional optimizations, and contrasts to the standard approach of MMAC where model assessment is achieved via monitoring the output of associated observers. Nevertheless, as we will see, estimator A can be implemented by monitoring the output of Kalman Filters, hence providing the linkage to standard MMAC.

The finite horizon estimator X_B is recursive by construction, therefore the computational complexity of the direct optimization does not depend on $k \in \mathbb{N}$ but only on the complexity of the involved optimization at each time step. The direct optimization is the computation of the optimal v_0 to determine $i_p[k]$, and this computation is bounded independent of $k \in \mathbb{N}$. The norm in (4.31) (also in (4.32)) can be taken to be $l_r,\ 1 \leq r \leq \infty$, giving rise to different optimizations. Such standard optimization problems can be solved by many possible implementations, i.e., in the linear case via computing a suitable pseudo inverse in l_2 or via linear programming in l_1 or l_∞, or convex programming in other norm settings. The implementation of the estimators in the nonlinear setting is discussed further in Section VI-C where it is shown that linear systems with input saturation have estimator optimizations which can be solved by linear or quadratic programming.

The infinite horizon estimator X_A has the direct interpretation as generating the size of the smallest disturbances compatible with the plant P_p and the observation w_2 up to the current time k. However, any direct implementation of the optimization defining the infinite horizon estimator X_A is not realizable (e.g., by using any of the optimizations methods described above for estimator B, but over the horizon $[0, k]$), since the computational complexity of these optimization algorithms grows with $k \in \mathbb{N}$. But, importantly, in the l_2 setting with linear plants, the residuals $r_p^A[k]$ for $p = (A, B, C, 0) \in \bar{\mathcal{P}}_{\text{LTI}}$ can be determined indirectly from the residuals in a Kalman filter bank (see also [9]). With $\hat{x} : [0, \tau] \mapsto \mathbb{R}^n,\ \tau \in \mathbb{N},\ \Sigma : \mathbb{N} \mapsto \mathbb{R}^{n \times n}$, the discrete-time Kalman filter equations are given as follows:

$$\hat{x}(k+1/2) = \hat{x}(k) - \Sigma(k)C^\top \left[C\Sigma(k)C^\top + I\right]^{-1} \cdot [y_2(k) + C\hat{x}(k)] \quad (4.33)$$

$$\Sigma(k+1/2) = \Sigma(k) - \Sigma(k)C^\top \cdot \left[C\Sigma(k)C^\top + I\right]^{-1} C\Sigma(k) \quad (4.34)$$

$$\hat{x}(k+1) = A\hat{x}(k+1/2) - Bu_2(k) \quad (4.35)$$

$$\Sigma(k+1) = A\Sigma(k+1/2)A^\top + BB^\top \quad (4.36)$$

where the initial conditions are specified by $\Sigma(0), \hat{x}(0)$. As a notion of the output error between the observation y_2 and the estimation of the Kalman filter, define the (scaled) residual $r : \mathbb{N} \to \mathbb{R}^+$ for $\tau \geq 0$ by

$$r_{\mathrm{KF}(\Sigma)}(\tau) = \left[\sum_{k=0}^{\tau} \|y_2(k) + C\hat{x}(k)\|^2_{[C\Sigma(k)C^\top + I]^{-1}}\right]^{\frac{1}{2}}.$$

Note that $[C\Sigma(k)C^\top + I]^{-1}$ is defined since it can be shown that $\Sigma(k)$ is positive semi-definite for all $k \in \mathbb{N}$ provided $\Sigma(0) = \Sigma(0)^\top \geq 0$. The key result establishing the equality between the Kalman Filter residual and the infinite horizon estimator is as follows:

Theorem 5: Let $p = (A, B, C, 0) \in \bar{\mathcal{P}}_{\mathrm{LTI}}$ and suppose that C is full row rank. Let the Kalman filter be described by equations (4.33)–(4.36) with the interconnection specified by (2.2), (2.3). Let $\hat{x}(0) = 0$ and $\Sigma(0) = 0$. Then

$$r_{\mathrm{KF}(\Sigma)}(\tau) = r_p^A[\tau] = X_A(w_2)(\tau)(p), \quad \forall\, w_2 \in \mathcal{W}_e,\ \tau \in \mathbb{N}.$$

Proof: The proof can be found in [6] and is related to previous work on the deterministic interpretation of the Kalman Filter, see e.g., [22], [27]. ■

This makes the realization of $X_A(\cdot)(\cdot)(p)$ finite dimensional as the Kalman filter algorithm is recursive—the computational complexity is invariant to $k \in \mathbb{N}$ and is dependent only on the order of the corresponding plant model $p \in \mathcal{P}_{\mathrm{LTI}}$.

Finally we observe that the switching algorithm requires the computation of the estimator $X(w_2)(k)(p)$ for all candidate plant models $p \in G(k)$. This is the limiting real-time computational requirement of EMMSAC: bounds on the number of candidate plant models required are the focus of Section VI. Note also that the computation involved in realizing an estimator bank is ideally suited to parallel computing, and may be realised e.g., on GPU or FPGA architectures.

We now state five abstract estimator assumptions that the residual operator is required to satisfy and on which the subsequent analysis will rest, and show that both the infinite and finite horizon residual operators, X_A and X_B, satisfy these axioms. These axioms ensure that the residuals have an interpretation as capturing the size of the smallest disturbances compatible with the plant model and the observed signals. The key to this interpretation is the requirement that X can be factorised, $X = NE$, where N and E are norm and estimation operators as defined next. For $k \in \mathbb{N},\ p \in \mathcal{P}$ the estimation operator has the structure

$$E : \mathcal{W}_e \to \mathrm{map}\left(\mathbb{N}, \mathrm{map}\left(\mathcal{P}, \mathrm{map}(\mathbb{N}, \mathbb{R}^h)\right)\right) \tag{4.37}$$

$$w_2 \mapsto [k \mapsto (p \mapsto d_p[k])] \tag{4.38}$$

where $d_p[k] : \mathbb{N} \to \mathrm{map}(\mathbb{N}, \mathbb{R}^h)$ represents the time series of the disturbance estimates at time $k \in \mathbb{N}$ corresponding to a plant $p \in \mathcal{P}$

$$d_p[k] = (d_p[k](0), d_p[k](1), \ldots, d_p[k](k), 0, \cdots)$$

where $h \in \mathbb{N} \cup \{\infty\}$ depends on the plant. Note that, in general, this estimate will not be recursive, i.e., $\mathscr{T}_k d_p[l] \neq \mathscr{T}_k d_p[k], l > k$.

The norm operator is defined

$$N : \mathrm{map}(\mathbb{N}, \mathrm{map}(\mathcal{P}, \mathrm{map}(\mathbb{N}, \mathbb{R}^h))) \to \mathrm{map}(\mathbb{N}, \mathrm{map}(\mathcal{P}, \mathbb{R}^+)) \tag{4.39}$$

$$[k \mapsto (p \mapsto d_p[k])] \mapsto [k \mapsto (p \mapsto \|d_p[k]\| = r_p[k])]. \tag{4.40}$$

Assumption 6: Let $\lambda \in \mathbb{R}$ be given. The residual operator X factorises $X = NE$ where N is the norm operator, E is an estimation operator, and:

1) (Causality): E is causal.
2) (Weak consistency): For all $p \in \mathcal{P}$ there exists a map $\Phi_\lambda : \mathrm{map}(\mathbb{N}, \mathbb{R}^h) \to \mathbb{R}^{m(\lambda+1)} \times \mathbb{R}^{o(\lambda+1)}$, such that for all $w_2 \in \mathcal{W}_e$ and for all $k \in \mathbb{N}$
$$\Phi_\lambda E(w_2)(k)(p) \in \mathcal{N}_p^{[k-\lambda,k]}(w_2), \quad \text{and,}$$
$$\|\Phi_\lambda E(w_2)(k)(p)\| \leq \|\mathscr{R}_{\lambda,k} E(w_2)(k)(p)\|.$$
3) (Monotonicity): For all $p \in \mathcal{P}$, for all $k, l \in \mathbb{N}$ with $0 \leq k \leq l$ and for all $w_2 \in \mathcal{W}_e$
$$\|E(w_2)(k)(p)\| \leq \|\mathscr{T}_k E(w_2)(l)(p)\|.$$
4) (Continuity): There exists a function $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}^+$, $\chi(p, p) = 0$ for all $p \in \mathcal{P}$, such that for all $k \in \mathbb{N}$, $p_1, p_2 \in \mathcal{P}$ and $w_2 \in \mathcal{W}_e$
$$\|E(w_2)(k)(p_1) - E(w_2)(k)(p_2)\| \leq \chi(p_1, p_2)\|\mathscr{T}_k w_2\|.$$
5) (Minimality): There exists $\mu > 0$ such that for all $k \geq 0$, for $p \in \mathcal{P}$ and for all $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying (2.2), (2.3) for $P = P_p$
$$\|E(w_2)(k)(p)\| \leq \mu\|\mathscr{T}_k w_0\|.$$

The treatment of the finite and infinite horizon case in a unified framework is possible since:

Proposition 7: Both X_A and X_B satisfy assumption 6.

Proof: See Appendix. ■

In the case of estimator A, we can take $\chi = \chi_A$, where

$$\chi_A(p_1, p_2) = \sup_{k \geq 0} \left\| \Pi_{p_1}^{[0,k]} - \Pi_{p_2}^{[0,k]} \right\|. \tag{4.41}$$

In the case of estimator B, we can take $\chi = \chi_B$, where

$$\chi_B(p_1, p_2) = (\lambda + 1)^{\frac{1}{r}} \max_{0 \leq k \leq \lambda+\sigma+1} \left\| \Pi_{p_1}^{[k-\lambda,k]} - \Pi_{p_2}^{[k-\lambda,k]} \right\|. \tag{4.42}$$

In the important case of l_2, where $r = 2$, we can alternatively take χ to be the l_2 gap δ by the following bound:

Proposition 8: Let $r = 2$. Then $\chi_A(p_1, p_2) \leq \delta(p_1, p_2)$.

Proof: See Appendix. ■

The continuity of χ (with respect to (w.r.t.) the gap topology) plays an important role in establishing the existence of finite dimensional EMMSAC controllers for uncertainty sets described by compact continua (Proposition 17, Theorem 18 below). In the case of Estimator A in l_2, continuity follows from Proposition 8. For Estimator B in general l_r, we have:

Proposition 9: Let $1 \leq r \leq \infty$. Suppose $\Omega \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{o \times n} \times \mathbb{R}^{o \times m} \subset \overline{\mathcal{P}}_{\mathrm{LTI}}$ is compact. Then $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_+ \cup \{\infty\}$ as given by (4.42) is continuous on $\Omega \times \Omega$.

Proof: See Appendix. ■

V. Stability and Gain Bound Analysis

In this section we establish the underlying gain bounds for EMMSAC algorithms. A key feature of the bounds is the explicit appearance of terms related to the metric complexity (or entropy) of the underlying uncertainty set rather than on the complexity of the controller (as measured, for example, by the number of candidate plant models). We will show in Section VI that the results of this section lead to gain bounds for realizable algorithms where it is the geometry of the plant model set that influences the gain bound, rather than the absolute size of the plant model set. In particular in some geometries the performance is independent of the candidate plant set size (e.g., beyond a certain threshold, the gain bounds are independent of the plant model density within a fixed uncertainty set).

The complexity of the underlying uncertainty set will be captured through the notion of a cover of the uncertainty set. Let $U : \mathcal{W}_e \to \mathrm{map}(\mathbb{N}, \mathcal{P}^*)$ be a monotonic plant-generating operator. U has the role of specifying an uncertainty set we seek to control at a given time $k \in \mathbb{N}$. Let $\chi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}^+$ be as in Assumption 6(4). Let

$$H : \mathcal{W}_e \to \mathrm{map}(\mathbb{N}, \mathcal{P}^*) \tag{5.43}$$

be a plant-generating operator. Let $\nu : \mathcal{W}_e \to \mathrm{map}(\mathbb{N}, \mathrm{map}(\mathcal{P}, \mathbb{R}^+))$ be given. As in Section III we write $U(k), H(k), \nu(k)$ for $U(w_2)(k), H(w_2)(k), \nu(w_2)(k)$ respectively. Now define

$$B_\chi(p, \nu(k)(p)) := \{p\} \cup \{p_1 \in \mathcal{P} \mid \chi(p, p_1) < \nu(k)(p)\} \cap U(k),\ p \in \mathcal{P},\ k \in \mathbb{N}. \tag{5.44}$$

For an appropriate choice of H, ν, the union of the corresponding neighbourhoods in U then leads to a cover for U:

Definition 10: (H, ν) is said to be a monotonic cover for a plant-generating operator U if $\forall\, k \in \mathbb{N}$, $w_2 \in \mathcal{W}_e$:

1) H and ν define a cover for U

$$U(k) \subset R(k) := \cup_{p \in H(k)} B_\chi(p, \nu(k)(p)).$$

2) The cover is monotonic: $R(k) \subset R(k+1)$, $\forall k \in \mathbb{N}$. (H, ν) is said to be a finite cover if $H(k)$ is a finite set for all $k \in \mathbb{N},\ w_2 \in \mathcal{W}_e$.

Sufficient conditions for the existence of a finite cover (H, ν) for U will be established in Section VI. We will utllize the objects U, H, ν, G in the following way:

1) U is the uncertainty, or more precisely U specifies the uncertainty set.
2) The cover (H, ν) for U is the device by which we assess the complexity of the uncertainty U.
3) G is an appropriate sampling of the uncertainty set U and determines the candidate plant set used by the controller.

See Fig. 3. (H, ν) is the device by which we are able to express gain bounds which scale in terms of the number of elements of $|H(k)|$ rather than the size of the set $G(k)$.

The static version of EMMSAC has G as a constant operator, which is well suited to the case where U is also constant and U is used to directly describe the structured uncertainty set; for example $U(k) = \mathcal{P}_{a_{\max}}$ for all $k \in \mathbb{N}$, where $\mathcal{P}_{a_{\max}} = \{(a, 1, 1, 0) \in \bar{\mathcal{P}}_{\mathrm{LTI}} \subset \mathbb{R}^4 \mid a \in \mathbb{R},\ |a| \le a_{\max}\}$. G would then represent a suitable sampling of the uncertainty set $\mathcal{P}_{a_{\max}}$, for example with $\epsilon > 0$, $G(k) = \{(n\epsilon, 1, 1, 0) \in \bar{\mathcal{P}}_{\mathrm{LTI}} \subset \mathbb{R}^4 \mid n \in \mathbb{Z},\ n\epsilon \le a_{\max}\}$.

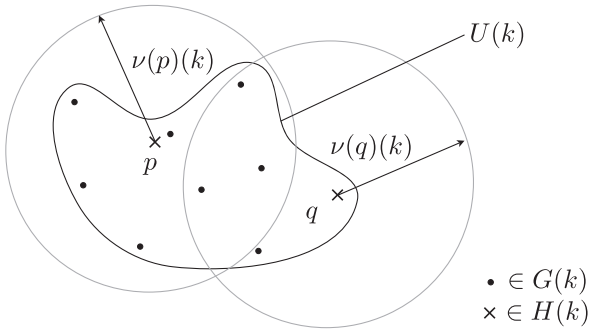

Fig. 3. Uncertainty set $U(k)$, cover $(H(k), \nu(k))$, candidate plant set $G(k)$.

The time varying nature of the operators is motivated by the requirements of dynamic EMMSAC. A typical dynamic EMMSAC algorithm varies the candidate plant set available at time K specified by $G(k)$ until some performance requirement is met. For example, if the uncertainty was unbounded, e.g., it is only known that the true plant p_* lies in the set $\mathcal{P}_\infty = \{(a, 1, 1, 0) \in \bar{\mathcal{P}}_{\mathrm{LTI}} \subset \mathbb{R}^4 \mid a \in \mathbb{R}\}$, then we could choose $G(k) = P_{i(k)}$ where $i(0) = \mathcal{P}_0$ and $i(k) = i(k-1)$ if performance is satisfactory, else $i(k) = i(k-1) + 1$. Dynamic EMMSAC formalises this process of expansion of the uncertainty set by choosing G dependent on the magnitude of the observations w_2, indeed in [7] we give an explicit example of such an algorithm.

We impose the following constraint on the set $\mathcal{P}^U$:

Assumption 11: Let $\mathcal{P}^U \subset \mathcal{P}$ have the property

$$\sigma := \max_{p_1, p_2 \in \mathcal{P}^U} \max\{\sigma(p_1), \sigma(K(p_2))\} < \infty.$$

$[P_{p_1}, C_{K(p_2)}],\ p_1, p_2 \in \mathcal{P}^U$ is uniquely determined. Further, we assume:

Assumption 12: The delay transition function $\Delta : \mathcal{P} \to \mathbb{N}$ satisfies $\Delta(p) > \sigma,\ \forall p \in \mathcal{P}^U$.

Assumptions 11 and 12 combined therefore ensure that there is sufficient time between controller switches to determine the initial conditions associated to any of the possible closed loops.

In the standard EMMSAC design, the delay transition function represents a sufficient period of time to ensure that each atomic controller provides enough of a stabilizing effect on it's associated plant before it can be switched away. This is formalised as follows:

Assumption 13: The control design K and delay transition function Δ satisfies

$$J(\xi) \sup_{p_1 \in \mathcal{P}^U} \alpha^\xi(p_1, K(p_1), \Delta(p_1) - \sigma, \sigma) < 1$$

where Δ satisfies Assumption 12, α is defined in Assumptions 2 and for $1 \le x, y < \infty$ we define $\lfloor c \rfloor := \max\{n \in \mathbb{Z} \mid n \le c\}$, $\binom{x}{y} := x!/y!(x-y)!$ and $J(\xi) = \xi\binom{\xi}{\lfloor \xi/2 \rfloor}$.

We note that Assumption 13 is achievable by design. Given K satisfying Assumption 2(2) a transition delay function Δ exists which meets Assumption 13. For linear systems, by considering the expression for α in (3.17), it can be seen that powers of the closed loop matrix A associated to $[P_p, C_K(p)]$ will determine the necessary lower bound on Δ_p. For a given control design, this can be bounded by considering the pseudo-spectra of A [24], or simply assessed numerically. We also remark that the minimization of the transient generated by the powers of A may be a legitimate target of the control design K e.g., through pole placement (at the extreme, a dead-beat design sets $A = 0$). Alternatively the complete design freedom available to the atomic design can be exploited and good transient performance can be indirectly ensured via designs such as mixed H_2/H_∞ [15], LQR etc. and suitable values for Δ_p can be determined post-design by the preceding comments. In the examples in this contribution the designs are: LQR ([7, Sec. 7]) and proportional (Section VI-B). Further examples include an EMMSAC design for a pendulum system using pole placement [9], and LQR controllers with nonlinear inversion for an application of EMMSAC to electrical stimulation of human muscle [5].

The construction of each $K(p),\ \Delta(p),\ p \in \mathcal{P}^G$ by hand is possible for small uncertainty sets, although this will not be feasible in many situations, i.e., if $\mathcal{P}^G$ is large or unknown. Automated design procedures for K and Δ can for example be implemented by using (the code from) suitable MATLAB toolboxes, e.g., to automatically construct stabilizing H_∞, LQR, PID controllers. The design challenge is then to set suitable parameters for the automation (e.g., plant parameterised weights in H^∞ design process); this may be achieved, for example, by extensive off-line simulations. For constant G, these designs would be typically computed a-priori, however, in many dynamic schemes, the controllers can be constructed on-line: only one controller is active at a time, hence only a single controller and corresponding delay needs to be calculated every time the algorithm performs a switch. Hence determining the controller and delay on-line reduces the (possibly infinitely large) computational complexity of determining K and Δ off-line to a single computational operation every time a switch occurs. We can therefore trade off memory size and computational off-line resource versus computational on-line resource, or have a hybrid of both.

The following definition now defines the general class of EMMSAC controllers considered (see Table I).

Definition 14: An EMMSAC controller $C(U, K, \Delta, G, X)$ is said to be standard if it satisfies:

- $K : \mathcal{P} \to \mathcal{C}$ is a given control design satisfying Assumption 2(1), (2).
- U satisfies assumption 11 for $\sigma < \infty$.
- $\Delta : \mathcal{P} \to \mathbb{N}$ is a delay transition function satisfying Assumption 12.
- K, Δ satisfy Assumption 13.
- E satisfies Assumptions 6(1)–(5) where

$$\lambda = \max_{p \in \mathcal{P}^U} (2\Delta(p) + \sigma). \tag{5.45}$$

- The switching operator $S = \mathrm{DM}(X, G)$ is given by (3.19)–(3.23) and (3.24)–(3.27).
- The switching controller C is defined by (3.28), (3.29).

TABLE I
RELEVANT EQUATIONS FOR DEFINTION 14 AND THEOREMS 15, 18

$$y_1 = Pu_1 \tag{2.2}$$
$$u_0 = u_1 + u_2, \quad y_0 = y_1 + y_2 \tag{2.3}$$
$$u_2 = Cy_2 \tag{2.4}$$
$$X : \mathcal{W}_e \to map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)) : w_2 \mapsto [k \to (p \mapsto r_p[k])] \tag{3.20}$$
$$G : \mathcal{W}_e \to map(\mathbb{N}, \mathcal{P}^*) \tag{3.19}$$
$$M : (map(\mathbb{N}, map(\mathcal{P}, \mathbb{R}^+)), map(\mathbb{N}, \mathcal{P}^*)) \to map(\mathbb{N}, \mathcal{P}^*) \tag{3.21}$$
$$\big[k \mapsto (p \mapsto r_p[k]), k \mapsto G(k)\big] \mapsto \big[k \mapsto q_f(k)\big] \tag{3.22}$$
$$q_f(k) := \operatorname*{argmin}_{p \in G(k)} r_p[k],\ \forall k \in \mathbb{N} \tag{3.23}$$
$$D : map(\mathbb{N}, \mathcal{P}) \to map(\mathbb{N}, \mathcal{P}) \tag{3.24}$$
$$[k \mapsto q_f(k)] \mapsto [k \mapsto q(k)] \tag{3.25}$$
$$q(k) := \begin{cases} q_f(k) & \text{if} \quad k - k_s(k) \geq \Delta(q(k_s(k))) \\ q(k_s(k)) & \text{else} \end{cases} \tag{3.26}$$
$$k_s(k) := \max\{i \in \mathbb{N} \mid 0 \leq i \leq k,\ q(i) \neq q(i-1)\} \tag{3.27}$$
$$C : \mathcal{Y}_e \to \mathcal{U}_e : y_2 \mapsto u_2 \tag{3.28}$$
$$u_2(k) = C_{K(q(k))}(y_2 - \mathscr{T}_{k_s(k)-1} y_2)(k) \tag{3.29}$$

The following Theorem establishes a gain bound where a bias term arises from the behaviour of the system on an interval $[0, k_*)$ and a gain term depends on the behaviour on $[k_*, \infty]$. Here $k_* \in \mathbb{N} \cup \infty$ is defined by (5.47), and is the first time at which a matching condition w.r.t. p_* and $G(k)$ is satisfied (either exactly ($\varepsilon = 0$) or approximately ($\varepsilon > 0$)). The interval $[0, k_*)$ is analyzed under no assumption of the (approximate) presence of p_* in the candidate plant set as specified by $G(k)$, and the period $[k_*, \infty]$ captures the behaviour of the closed loop once (an approximation of) the plant p_* is available to the switching mechanism. In the classical (static) setup (e.g., [9], [11], [13], [18], [19]) we have $p_* \in G(j) = G(k), \forall j, k \in \mathbb{N}$ so $k_* = 0$. The case $k^* > 0$ arises in dynamic EMMSAC.

In order to define the time $k^* \geq 0$ by (5.47), we let $0 = l_0 \leq l_i < l_{i+1}$ form the ordered sequence of times when the switching sequence changes value, i.e., when $q(l_i - 1) \neq q(l)_i$ and define

$$Q_\infty = \cup_{i \geq 0} [l_i, l_{i+1}] \cap \{l_i + b\Delta(q(l_i)) \mid b \in \mathbb{N}\}. \tag{5.46}$$

The main result establishes closed loop gain bounds for both dynamic and static EMMSAC.

Theorem 15: Let $1 \leq r \leq \infty$. Let $P = P_{p_*}$, where $p_* \in \mathcal{P}^U \subset \mathcal{P}$. Let U be a monotonic plant generating operator and suppose (H, ν) defines a monotonic finite cover for U. Let $k \in \mathbb{N}$. Suppose the EMMSAC controller $C(U, K, \Delta, G, X)$ is standard, and $G(j) \subset U(j),\ j \leq k$. Suppose $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfy the closed loop (2.2), (2.3). Let $\varepsilon > 0$. Let

$$k_* := \begin{cases} \min\{i \in Q_\infty \mid \exists\, p \in G(i),\ \chi(p, p_*) \leq \varepsilon \chi_\nu(H, \nu)\} \\ \text{if } \exists i \text{ s.t. } \exists\, p \in G(j), \chi(p, p_*) \leq \varepsilon \chi_\nu(H, \nu),\ \forall j \geq i, \\ \infty \text{ if not} \end{cases} \tag{5.47}$$

and suppose $k_* < \infty$. If

$$\pi\left(U(j), H(j), \nu(j), \varepsilon, p^*\right) > 0, \quad \forall\, j \leq k \tag{5.48}$$

TABLE II
FUNCTIONS SPECIFYING THE GAIN BOUND

For $\mathcal{Q}_1 \subset \mathcal{P}^U$ and $\xi = \begin{cases} r & \text{for} \quad 1 \le r < \infty \\ 1 & \text{for} \quad r = \infty \end{cases}$ let:

$$J(\xi) = \xi \begin{pmatrix} \xi \\ \max\{n \in \mathbb{Z} \mid n \le \xi/2\} \end{pmatrix} \text{ where } \begin{pmatrix} x \\ y \end{pmatrix} := \frac{x!}{y!(x-y)!},$$

$$\begin{aligned}
\alpha_{OP}(\mathcal{Q}_1) &= J(\xi) \sup_{p_1 \in \mathcal{Q}_1} \alpha^{\xi}(p_1, K(p_1), \Delta(p_1) - \sigma, \sigma) \\
\beta_{OP}(\mathcal{Q}_1) &= J(\xi) \sup_{\substack{\Delta(p_1) \le x \le 2\Delta(p_1) \\ p_1 \in \mathcal{Q}_1}} \beta^{\xi}(p_1, K(p_1), x - \sigma, \sigma) \\
\alpha_{OS}(\mathcal{Q}_1) &= J(\xi) \sup_{\substack{\Delta(p_1) \le x \le 2\Delta(p_1) \\ p_1 \in \mathcal{Q}_1}} \alpha^{\xi}(p_1, K(p_1), 0, x - \sigma) \\
\beta_{OS}(\mathcal{Q}_1) &= J(\xi) \sup_{\substack{\Delta(p_1) \le x \le 2\Delta(p_1) \\ p_1 \in \mathcal{Q}_1}} \beta^{\xi}(p_1, K(p_1), 0, x - \sigma) \\
\gamma_1(p, p_*) &= 1 + \sup_{\Delta(p) \le x \le 2\Delta(p)} \alpha(p_*, K(p), 0, x) \\
\gamma_2(p, p_*) &= \sup_{\Delta(p) \le x \le 2\Delta(p)} \beta(p_*, K(p), 0, x), \\
\bar{\gamma}_i(\mathcal{Q}_2, \mathcal{Q}_1) &= \sup_{p_2 \in \mathcal{Q}_2} \sup_{p_1 \in \mathcal{Q}_1} \gamma_1(p_2, p_1), \quad i = 1, 2,
\end{aligned}$$

If $1 \le r < \infty$ let:

$$\begin{aligned}
\gamma_3(\mathcal{Q}_1) &= (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1)) \left(\frac{\alpha_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)} \right)^{1/r} + \alpha_{OS}^{1/r}(\mathcal{Q}_1) \\
\gamma_4(\mathcal{Q}_1) &= (1 + \alpha_{OS}^{1/r}(\mathcal{Q}_1)) \left(\frac{\beta_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)} \right)^{1/r} \\
\gamma_5(\mathcal{Q}_1) &= \beta_{OS}^{1/r}(\mathcal{Q}_1),
\end{aligned}$$

and if $r = \infty$ let:

$$\begin{aligned}
\gamma_3(\mathcal{Q}_1) &= \max\{1, \alpha_{OS}(\mathcal{Q}_1))\}\alpha_{OP}(\mathcal{Q}_1) + \alpha_{OS}(\mathcal{Q}_1) \\
\gamma_4(\mathcal{Q}_1) &= \max\{1, \alpha_{OS}(\mathcal{Q}_1)\}) \frac{\beta_{OP}(\mathcal{Q}_1)}{1 - \alpha_{OP}(\mathcal{Q}_1)} \\
\gamma_5(\mathcal{Q}_1) &= \beta_{OS}(\mathcal{Q}_1).
\end{aligned}$$

For $\mathcal{Q}_2 \subset \mathcal{P}^H$, $v : \mathcal{P} \to \mathbb{R}^+$, $\mu > 0$, $\varepsilon > 0$ let:

$$\begin{aligned}
\chi_\nu(\mathcal{Q}_2, v) &= 2 \sup_{p \in \mathcal{Q}_2} v(p) \\
\pi(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p^*) &= 1 - 2^{1/r} \varepsilon \chi_\nu(\mathcal{Q}_2, v)(1 + \bar{\gamma}_1^2(\mathcal{Q}_2, \mathcal{Q}_1)) \cdot \\
&\qquad \cdot \big(\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)\big)
\end{aligned}$$

$$\eta(\mathcal{Q}_2, v, \varepsilon, p_*) = 2^{1/r}(\mu + \varepsilon \chi_\nu(\mathcal{O}_2, v)\bar{\gamma}_2(\mathcal{O}_2, \{p_*\})(1 + \bar{\gamma}_1(\mathcal{Q}_2, \{p_*\})))$$

$$\beta(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p_*) = \left(\frac{1 + \gamma_3(\mathcal{Q}_1)}{\pi(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p_*)} \right)^{|\mathcal{Q}_2|} \prod_{p \in \mathcal{Q}_2} \gamma_1(p, p_*)$$

$$\begin{aligned}
\hat{\gamma}(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p_*) &= \beta(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p_*) \left(\sum_{p \in \mathcal{Q}_2} \gamma_2(p, p_*) + \right. \\
&\qquad \left. \eta(\mathcal{Q}_2, v, \varepsilon, p_*) \frac{\gamma_4(\mathcal{Q}_1) + \gamma_5(\mathcal{Q}_1)}{\pi(\mathcal{Q}_1, \mathcal{Q}_2, v, \varepsilon, p_*)} \right).
\end{aligned}$$

then

$$\begin{aligned}
\|\mathscr{T}_k w_2\| \le \beta(U(k), H(k), \nu(k), \varepsilon, p_*) \|\mathscr{T}_{k_*-1} w_2\| \\
+ \hat{\gamma}(U(k), H(k), \nu(k), \varepsilon, p_*) \|w_0\| \quad (5.49)
\end{aligned}$$

where $\pi, \beta, \hat{\gamma}$ are given in Table II.

Proof: See [7] ■

A detailed discussion of the interpretation of the terms in the above bounds can be found in the companion paper [7]. Prior to fully exploiting the implications of this result in Section VI, we make some observations. There are two principal conditions under which Theorem 15 holds. The requirement that the design is standard, incorporates Assumption 13 which requires the condition that $\alpha_{\mathrm{OP}}(U(j)) < 1$. As discussed previously, this is a condition on the atomic controllers which is achievable by design. The second condition (inequality (5.48)) relates the allowable cover H to the underlying uncertainty set U. Proposition 17 below shows how a construction of a finite cover H meeting inequality (5.48) can be achieved for the case of compact U.

We discuss three special cases. First if the structured uncertainty set is finite, as for example in the case of seeking to stabilize a discrete integrator with unknown sign: $p_* = (A, B, C, D) \in \{(1, 1, 1, 0), (1, -1, 1, 0)\} \subset \bar{\mathcal{P}}_{\mathrm{LTI}}$, then we can choose $G = U$ to be constant, $G(k) = U(k) = \{(1, 1, 1, 0),\ (1, -1, 1, 0)\} \subset \bar{\mathcal{P}}_{\mathrm{LTI}}$. Then taking $H = G$, $\nu = 0$, yields gain bounds for any stabilizing atomic controllers, where note that we can take $\varepsilon = 0$, and hence $k_* = 0$. Consequently, Theorem 1 provides robustness margins.

Secondly, in the case where the underlying uncertainty set is a continuum, for example, if $p_* \in \mathcal{P}_{a_{\max}} = \{(a, 1, 1, 0) \in \bar{\mathcal{P}}_{\mathrm{LTI}} \subset \mathbb{R}^4 \mid a \in \mathbb{R},\ |a| \le a_{\max}\}$, then choosing $G = U$ to be constant, $G(k) = U(k) = \mathcal{P}_{a_{\max}}$, together with a continuum of stabilizing controllers, yields finite gain bounds (where again we can take $\varepsilon = 0$ and hence $k_* = 0$) provided H is a finite cover (we provide sufficient conditions for this in Proposition 17 below). Such an infinite dimensional controller has a robustness margin provided by Theorem 1 but will not be directly implementable. However, in the following Section VI we will show how the underlying infinite dimensional controller can be sampled to produce a realizable design with guaranteed robustness margins, based on a finite candidate plant model set of appropriate geometry. This realization will be based on an application of Theorem 15 with $G \ne U$.

The third case is also in the setting where the underlying constant uncertainty set U is a continuum. Suppose K is a stabilizing control design where $K(U) \subseteq \{C_1, \ldots, C_n\}$, that is, such that each $p \in U$ is stabilised by $K(p)$ which is one of $\{C_1, \ldots, C_n\}$. In Section VI it is shown that by taking $\varepsilon > 0$ to be sufficiently small we can determine a suitable cover G thus determining a suitable estimator structure for the given controller bank (again with $k_* = 0$). This procedure typically yields multiple plant estimators per atomic controller.

The gain bound in Theorem 15 is a function of the complexity of U, as measured by (H, ν), and U itself, however is invariant to the number of elements in either U or G. Instead the gain bound scales (exponentially) with the number of elements in H. This is a substantial improvement on previous bounds, which scaled exponentially with the number of elements in the candidate plant set (i.e., with G). For example, this invariance suggests that there is no disadvantage in having a large number of high fidelity models in the candidate plant set (we will return to this point in Section VI). In Proposition 17 below we show that for compact operators U, suitable covers H exist and we bound the size of H (and hence the exponential exponents in the gain bound) in terms of the metric entropy of $U(j)$, $j \in \mathbb{N}$ where the metric entropy of a set Ω with distance χ and $\zeta > 0$ is given by

$$\begin{aligned}
\mathcal{C}_E(\chi, \Omega, \zeta) = \min\{n \in \mathbb{N} \mid h = \{p_1, \ldots, p_n\} \subset \mathcal{P} \\
\Omega \subset \cup_{p \in h} \{q \in \mathcal{P} \mid \chi(p, q) < \zeta\}\ \}.
\end{aligned}$$

A higher complexity implies less prior information. This concept of interlinking information with complexity is due to [28], where it is utilised to seek to define the term 'adaptive' in a control context.

Definition 16: Let $\sigma \in \mathbb{N}$. Let U be a plant-generating operator. Let α, β be defined by Assumptions 2 and let $\Delta : \mathcal{P} \to \mathbb{N}$ be the transition delay function. A control design $K : \mathcal{P} \to \mathcal{C}$ is said to be U regular if for all $\Delta(p) \leq x \leq 2\Delta(p)$, the functions $\alpha(p, K(p), \Delta(p) - \sigma, \sigma)$, $\beta(p_1, K(p), x - \sigma, \sigma)$, $\alpha(p_1, K(p), 0, x - \sigma)$, $\beta(p_1, K(p), 0, x - \sigma)$, $x \in \mathbb{N}$ are continuous with respect to all $p_1, p \in \mathcal{P}^U$.

The key result establishing the existence and complexity of a finite cover is given next. It is dependent on the continuity of χ, as for example established by Propositions 8, 9 for X_A $(r = 2)$ and X_B $(1 \leq r \leq \infty)$ respectively. Note also that the continuity and compactness requirements are w.r.t. the topology on $\mathcal{P}$ induced by the nonlinear gap $\delta(\cdot, \cdot)$.

Proposition 17: Let U be a compact plant-generating operator and suppose K is U regular. Suppose $\mathcal{P}^U$ is bounded and χ is continuous on $\mathcal{P}^U \times \mathcal{P}^U$. Suppose $\alpha_{OP}(U(j)) < 1$. Let $k \in \mathbb{N}$ and $j \leq k$. Let $\varepsilon > 0$. Then there exists a finite cover (H, ν) of U which satisfies inequality (5.48). The size of the cover is bounded

$$\#H(j) \geq \mathcal{C}_E \Big(\chi, U(j), 2^{-(r+1)/r}\varepsilon^{-1}(1 + \bar{\gamma}_1^2(\mathcal{P}^U, \mathcal{P}^U))^{-1} \times (\gamma_4(U(j)) + \gamma_5(U(j)))^{-1}\Big). \quad (5.50)$$

Proof: Let $j \leq k \in \mathbb{N}$. Since U is compact and K is U regular it follows that $\alpha_{\mathrm{OP}}(\mathcal{Q}) = J(\xi) \sup_{p_1 \in \mathcal{Q}} \alpha^{\xi}(p_1, K(p_1), \Delta(p_1) - \sigma, \sigma) < \infty$, for $\mathcal{Q} \subset \mathcal{P}^U$. Similarly, $\alpha_{OS}(U(j)) < \infty$ and $\beta_{OP}(U(j)) < \infty$. By assumption $\alpha_{OP}(U(j)) < 1$ and hence $\gamma_4(U(j))$, $\gamma_5(U(j))$ are defined. Let $0 < \zeta_j < (2^{(r+1)/r}\varepsilon(1 + \bar{\gamma}_1^2(\mathcal{P}^U, \mathcal{P}^U))(\gamma_4(U(j)) + \gamma_5(U(j)))^{-1}$. Since $\chi|_{\mathcal{P}^U \times \mathcal{P}^U}$ is continuous, it follows from (5.44) that $B_\chi(p, \zeta_j)$ is open for $p \in \mathcal{P}^U$ and hence $\{B_\chi(p, \zeta_j)\}_{p \in U(j)}$ is an open cover of $U(j)$ with respect to the subspace topology of $U(j)$. Since $U(j)$ is compact, there exists a finite set $h_j \subset U(j)$ such that $\{B_\chi(p, \zeta_j)\}_{p \in h_j}$ covers $U(j)$. Let $\nu_j(p) = \zeta_j, \ \forall p \in \mathcal{P}$ hence $(h_j, \nu_j) \in (\mathcal{P}^U, \mathrm{map}(\mathcal{P}, \mathbb{R}^+))$ is a finite cover of $U(j)$. We construct a monotonic cover (H, ν) by letting $H(k) = \cup_{j \leq k} h_j$, $\nu(k)(p) = \min_{j \leq k} \zeta_j, \ \forall p \in \mathcal{P}^H$. Since $p_* \in \mathcal{P}^U$, $H(j) \subset U(j)$ it follows that $\bar{\gamma}_1(H(j), \{p_*\}) \leq \bar{\gamma}_1(\mathcal{P}^U, \mathcal{P}^U)$, and since ν_j is constant it follows that $\zeta_j = (1/2)\chi_{\nu_j}(h_j, \nu_j)$ and hence (5.48) holds. By construction, the size of $H(j)$ is given by the right hand side of (5.50) ■

VI. Design

We consider the case of design for a compact structured uncertainty set, for example as specified by a closed and bounded parametric uncertainty $\Omega \subset \mathcal{P}$. U is taken to be a constant, compact plant generating operator representing the uncertainty: $U(j) = \Omega$, $j \in \mathbb{N}$. Determining stability guarantees for feasible MMAC controllers where the uncertainty is given by a continuum has been a central topic in the literature [14], and is unresolved in general, with the exception of [17] where a structured switching mechanism achieves the requisite stability. Section V has established a complexity dependent gain bound, applicable to EMMSAC controllers $C(U, K, \Delta, G, X)$ which can be applied with $G = U$, i.e., to $C(U, K, \Delta, U, X)$, however, this typically yields an unrealizable infinite dimensional controller (for example if the uncertainty U represents a continuum, then a continuum of estimators are required, and the atomic control design K generally represents a continuum of distinct controllers). In this section we give a principled route to constructing a finite dimensional controller (based on a finite number of estimators) which robustly stabilises all plants in the uncertainty set and inherits a gain bound which is quantifiably close to the original bound. Additionally, and as an alternative route, we show that by starting from a stabilizing atomic control design based on only a finite number of controllers, we can construct a corresponding finite dimensional estimator structure yielding a stabilizing EMMSAC controller. As the optimal χ cover gives a measure of the complexity of the underlying uncertainty, we can thus interpret both routes as design processes which take the uncertainty description (U) and yield a concrete algorithm $C(U, K, \Delta, G, X)$, together with associated complexity dependent gain bounds.

A. Candidate Plant Sampling of Compact Uncertainty Sets

Suppose that for a given constant and compact uncertainty set U, the cover (H, ν) satisfies inequality (5.48). The final design step is then to construct a suitable finite plant model set G whose associated EMMSAC controller $C(U, K, \Delta, G, X)$ has guaranteed robust stability. Since the gain bounds of Theorem 15 and hence the robust stability margins (determined by Theorem 1) are independent of G, all that is now required is to ensure that the true plant either lies inside the candidate plant set, or lies sufficiently close to an element within the candidate plant set, where the maximum distance is determined by the robust stability margin given by Theorem 15. Here it is critical that the gain bound (and hence the robust stability margin) is independent of G—in the previous literature the gain bounds scaled with the size of the candidate plant set, hence the margins decreased as the plant set grew, and so it was unclear whether it was possible to give a stabilization guarantee for all plants within a continuum—for large candidate plant sets the margins were smaller, so it may not have been possible to find a plant set whereby the whole of the uncertainty was included within the the union of the plants guaranteed to be stable via the robustness margins, see [14] for a similar discussion.

To specify the number of candidate plants required, we introduce a notion of metric entropy where non-uniform neighborhoods are considered. Given $\zeta : \Omega \to \mathbb{R}_+$, the functional metric entropy $\mathcal{C}_{fE}$ is given by

$$\mathcal{C}_{fE}(\vec{\delta}, \Omega, \zeta) = \min\{n \in \mathbb{N} \mid h = \{p_1, \ldots, p_n\} \subset \mathcal{P} \quad \Omega \subset \cup_{p \in h}\{q \in \mathcal{P} \mid \vec{\delta}(p, q) < \zeta(p)\}\}.$$

In turn this is bounded by the standard metric entropy of $\mathcal{P}^U$: $\mathcal{C}_{fE}(\vec{\delta}, \mathcal{P}^U, \zeta(p)) \leq \mathcal{C}_E(\vec{\delta}, \mathcal{P}^U, \bar{\zeta})$, where $\bar{\zeta} = \inf_{p \in \mathcal{P}^U} \zeta(p)$. We note that constructing candidate plant sets via covers of the uncertainty sets has its antecedents in, for example, [2], [8], [14], [20].

Theorem 18: Let U be a constant, compact plant-generating operator. Suppose the EMMSAC controller $C(U,K,\Delta,U,X)$ is standard. Suppose there exists a function $\hat{\gamma}:\mathcal{P}\to\mathbb{R}$ such that $\|\Pi_{C(U,K,\Delta,U,X)//P_p}\|\leq\hat{\gamma}(p)$, for all $p\in\mathcal{P}^U$. Let $1>\xi>0$. Then there exists a constant plant generating operator $G(k)=\{p_1,\dots p_n\}$, $\forall k\in\mathbb{N}$, where $n=\mathcal{C}_{fE}(\vec{\delta},\mathcal{P}^U,\xi\hat{\gamma}(\cdot)^{-1})<\infty$, and

$$d(p):=\inf_{p_i\in G(0)}\vec{\delta}(p_i,p)<\xi\hat{\gamma}(p_i)^{-1},\qquad\forall p\in\mathcal{P}^U. \quad (6.51)$$

For all $p_*\in\cup_{p_1\in\mathcal{P}^G}\{p\in\mathcal{P}\mid\vec{\delta}(p_1,p)<d(p)\}$

$$\|\Pi_{P_{C(U,K,\Delta,G,X)//P_{p_*}}}\|\leq\frac{\xi(1+d(p_*))}{(1-\xi)d(p_*)}. \quad (6.52)$$

Furthermore, any monotonic plant generating operator G satisfying (6.51) also yields the gain bound (6.52).

Since $\Pi_{P_{p_*}//C(U,K,\Delta,G,X)}=I-\Pi_{C(U,K,\Delta,G,X)//P_{p_*}}$, it follows that under the conditions of Theorem 18, that $[P_{p_*},C(U,K,\Delta,G,X)]$ has the property that

$$\|\mathscr{T}_k w_1\|\leq\frac{1+\xi}{(1-\xi)d(p_*)}\|\mathscr{T}_k w_0\|\qquad\forall k\geq 0.$$

Proof: Since Ω is compact, and by the definition of $\mathcal{C}_{fE}$, there exists $p_1,\dots,p_n\subset\mathcal{P}$, $n<\infty$ such that for all $p\in\mathcal{P}^U$, $\vec{\delta}(p_i,p)<\xi\hat{\gamma}(p_i)^{-1}$ for some $1\leq i\leq n$. Hence by choosing G to be the constant plant generating operator $G(k)=\{p_1,\dots,p_n\}$, it follows that inequality (6.51) holds.

Let $p_*\in\mathcal{P}^U$ and let $1\leq i\leq n$ be such that $\vec{\delta}(p_i,p_*)=d(p_*)<\xi\hat{\gamma}(p_i)^{-1}$. Then by Theorem 1

$$\|\Pi_{P_{p_*}//C(U,K,\Delta,G,X)}\|\leq\hat{\gamma}(p_i)\frac{1+\vec{\delta}(p_i,p_*)}{1-\vec{\delta}(p_i,p_*)\hat{\gamma}(p_i)}$$

thus yielding inequality (6.52). This holds for any G satisfying inequality (6.51) as required. ■

The condition of the theorem, i.e., the existence of $\hat{\gamma}:\mathcal{P}\to\mathbb{R}$ such that $\|\Pi_{C(U,K,\Delta,U,X)//P_p}\|\leq\hat{\gamma}(p)$, for all $p\in\mathcal{P}^U$, is exactly the form of the bound supplied by Theorem 15 (with $G=U$) noting that $k^*=0$ since $p\in U(0)$. Additionally it is important to note that by taking $\hat{\gamma}(p)=\|\Pi_{C(U,K,\Delta,U,X)//P_p}\|$, the result shows how the true gains $\|\Pi_{C(U,K,\Delta,U,X)//P_{p_i}}\|$ and $\|\Pi_{P_{C(U,K,\Delta,G,X)//P_{p_*}}}\|$ are related.

If G is constant then this theorem shows that a system with a compact uncertainty can always be robustly stabilised by a MMAC algorithm with a fixed, finite candidate model set. To determine an appropriate geometry for the candidate plant set, we have to construct a covering of U by gap balls as determined from Theorem 15, and we can then take G to comprise the centres of these neighbourhoods. Theorem 18 also caters for the case where G is time varying. For a compact uncertainty set, this may arise by a dynamic mechanism which refines the candidate plant model set over time, introducing new higher fidelity models and controllers as required, see [7].

If computational resource is unlimited we may include as many plant models in $G\subset U$ as we like without weakening the gain bounds from Theorem 18. Furthermore when $G=U$ the bounds are minimised and collapse to the one in Theorem 15. This leads to the pragmatic guideline: populate the candidate plant set at as high a resolution as the implementation hardware constraints permit. An apparent over-population of plant models may arise also from a lack of tightness in the bounds utilized in the design process. So this is a rather unusual situation: bounds which may not be tight are required to be used in a design process; but performance does not degrade with the conservatism of these bounds.

B. Determination of an Estimator for a Fixed Controller Bank

By exploiting the fact that the control design K is not required to be injective, we can choose a stabilizing control design s.t. $K(U)\subseteq\{C_1,\dots,C_n\}$, that is, such that each $p\in U$ is stabilised by $K(p)$ which is one of a fixed a-priori chosen bank of controllers $\{C_1,\dots,C_n\}$. There are a variety of means to obtain such designs, for example the mixed μ and % FNARC technique of [8]. By now taking $\varepsilon>0$ to be sufficiently small, we can construct a constant cover (G,ρ) which satisfies the dual requirement that $\pi(U(j),H(j),\nu(j),\varepsilon,p^*)>0$ (inequality (5.48)) where $\rho=\varepsilon\chi_\nu(H,\nu)$ and that $\chi(p,p_*)<\rho$ for all $p\in G$, $p_*\in U$ ((5.47)). Note that the construction of G does not require the construction of H. H is only required to determine performance bounds, and the trade-off between ε and $\chi_\nu(H,\nu)$ is not relevant at the design stage. Typically, there are multiple estimators corresponding to each controller.

We illustrate this process in the following academic example which is chosen to enable exact computation, although this procedure remains tractable on more complex examples with numerical computation. We consider SISO LTI plants defined by the uncertainty set: $U=U_I=\{x\in\mathbb{R}^4:x=(a_1,1,1,0),\ a_1\in I\}$, $M>0$; our final numerical results will be for $I=[N,M]$ where $N=0.7$, $M=1.2$. We consider proportional controllers, with the design requirement that each controller C_i achieves closed loop pole placement within the disk of radius $d=0.14$ centred at 0 for each associated plant, i.e., for every plant in $U\cap K^{-1}(C_i)$. A straightforward calculation shows that 2 controllers suffice with gains 0.84, 1.12 corresponding to C_1,C_2 respectively, with $K^{-1}(C_1)\cap U=U_{[0.7,0.98]}$, $K^{-1}(C_2)\cap U=U_{(0.98,1.2]}$.

Since $\sigma(p)=1$, $\sigma(K(p))=0$, it follows that we can take $\Delta(p)=\sigma=1$ for all $p\in U$. EMMSAC will be implemented by Estimator B, with $\lambda=1$. It follows that $\mu=(\lambda+1)^{1/r}=\sqrt{2}$, and $\|c\|=\sqrt{2}$.

Some lengthy (but elementary) calculations establish

$$\begin{aligned}\alpha_{OS}(U)=\alpha_{OP}(U)&\leq 4(1+M^2)d^2(1+d^2),\\ \beta_{OP}(U)&\leq 4(1+M^2)((4+2M^2),\\ &\quad+d^2(2+2M^2)+d^4(1+M^2)),\\ \beta_{OS}(U)&\leq 8(1+M^2)^2(1+d^2).\end{aligned}$$

Hence $\alpha_{\mathrm{OP}}(U)=\alpha_{OS}(U)\leq 0.4$, $\beta_{OP}(U)\leq 67.81$ and $\beta_{OS}(U)\leq 48.56$. Consequently, $\gamma_4(U)\leq 10.85$ and $\gamma_5(U)\leq 6.99$. Similar computations show

$$\bar{\gamma}_1(U,U)\leq 1+(1+M^2)\sqrt{D^2+D^4+D^6}=1.89$$

where $D=N-M=0.5$. We next compute $\chi(p_1,p_2)$. Let $x_1=(1,-a_1,-1)^\top$, $x_2=(1,-a_2,-1)^\top$, $\hat{x}_1=x_1/\|x_1\|$, $\hat{x}_2=x_2/\|x_2\|$, $w_k=(y_2(k),y_2(k-1),u_2(k-1)^\top$. The projection $\Pi^k_{p_i}:=\Pi^{[k-\lambda,k]}_{p_i}$, $i=1,2$ is given by $\Pi^k_{p_i}(w_2)=\hat{x}_i^\top w_k\hat{x}_i$, so

$\chi(p_1,p_2)=\max_{k\geq 0}\|\Pi^k_{p_1}-\Pi^k_{p_2}\|=\max_{k\geq 0}\max\{\|\Pi^k_{p_2}(I-\Pi^k_{p_1})\|,\|(I-\Pi^k_{p_2})\Pi^k_{p_1}\|\}$. Since $\|\Pi^k_{p_2}(I-\Pi^k_{p_1})\|=\sup_{\|w_k\|=1}\|\hat{x}_2^\top(w_k-\hat{x}_1^\top w_k\hat{x}_1)\hat{x}_2\|=\sup_{\|w_k\|=1}\|(\hat{x}_2-\hat{x}_1^\top\hat{x}_2\hat{x}_1)^\top w_k\hat{x}_2\|=\|(\hat{x}_2-\hat{x}_1^\top\hat{x}_2\hat{x}_1)\|$, and similarly, $\|(I-\Pi^k_{p_2})\Pi^k_{p_1}\|=\|\hat{x}_1-\hat{x}_2^\top\hat{x}_1\hat{x}_2\|=\|(\hat{x}_2-\hat{x}_1^\top\hat{x}_2\hat{x}_1)\|$, it follows that:

$$\chi(p_1,p_2)=\|\hat{x}_2-\hat{x}_1^\top\hat{x}_2\hat{x}_1\|$$
$$=(1-(\hat{x}_1^\top\hat{x}_2)^2)^{\frac{1}{2}}=\sqrt{\frac{2(a_1-a_2)^2}{(2+a_1^2)(2+a_2^2)}} \quad (6.53)$$

By the definition of $\pi(U,\mathcal{Q}_2,\epsilon)$ and by the bounds on $\gamma_4(U)$, $\gamma_5(U)$ and $\bar{\gamma}_1(U,U)$ we can obtain the requirement $\rho:=\varepsilon\chi_\nu(\mathcal{Q}_2,\epsilon)\leq 0.00864$. To determine the estimator structure, when $G=H$, we need to find a constant (G,ρ) cover, where $\rho=0.00864/2$. We solve this by finding points a_n such that $\chi(a_{n+1},a_n)=0.00864=\delta$. That is, solving the iteration (from (6.53)): $a_{n+1}=(-b+\sqrt{b^2-4ac})/2a$ where $a=2-2\delta^2-\delta^2a_n^2$, $b=4a_n$ and $c=2a_n^2-2\delta^2a_n^2-4\delta^2$, initialized with $a_0=0.7$. This yields 17 estimators $a_0=0.7,\ldots,a_{16}=0.963$ corresponding to the controller C_1. Initializing the iteration again at $a_{17}=0.98$ yields 12 estimators $a_{17}=0.980,\ldots,a_{28}=1.194$ corresponding to controller C_2.

Due to the high degree of correlation between the estimator residuals of nearby estimators, a far courser grid of estimators is likely to remain effective. This is supported by indicative simulations, and remains a rich area for further theoretical studies. It is relevant to contrast this design procedure to results to [3], [23], [26], [29], where stabilizing schemes based on model falsification require only a feasibility assumption, that is they can be built with the minimum number of (atomic) controllers required to stabilize the uncertainty set U. The above design process shows that MMAC can also be based on a limited number of stabilizing controllers: it is the estimator structure which may require a larger number of candidate plants: note that this has no analogue in falsification schemes.

C. Nonlinear Systems: Input Saturation

Although our presentation has in the most part been oriented towards linear systems, a key feature of the axiomatic framework is that no assumption of linearity is made. The tractability of the EMMSAC in the nonlinear setting rests on a) the ability to achieve the controller assumptions (e.g., gain stability for the atomic closed loops) and b) the feasibility of the implementation of the optimization required by the estimator. For linear systems and for l_r, $1\leq r\leq\infty$ signal spaces, estimator B reduces to standard convex optimizations (linear, quadratic programming etc). In the more general setting, the direct estimator optimizations may not be tractable, unless restrictive convexity assumptions are imposed; however, the following example shows that there are important nonlinear cases for which EMMSAC is implementable.

We first establish a result that shows that the optimization required within the finite horizon estimator for a linear system with saturation (i.e., the computation of $i_p[k]$ ((4.32)) is equivalent to a constrained optimization problem, which in turn is solvable by standard convex optimizations:

Proposition 19: Let $1\leq r<\infty$. Let $p=(A,B,C,0)\in\bar{\mathcal{P}}_{\text{LTI}}$, $\tilde{p}=(p,S)$ and define $P_{\tilde{p}}=P_p\circ\text{SAT}_S$, where

$$\text{SAT}_S(u)(k)=\begin{cases}u(k)= & \text{if}|u(k)|\leq S\\ \frac{u(k)}{|u(k)|}= & \text{if}|u(k)|>S.\end{cases}$$

Then, $i_{(p,S)}[k]$ is determined by the following convex optimization:

$$i_{(p,S)}[k]=\min_{v_0\in N_1}\|v_0\|$$

where $N_1=\{(u_0,y_0)^\top\in\mathcal{W}\mid(u_0,y_0)^\top\in\mathcal{N}_p^{[k-\lambda,k]}(u_2,y_2)^\top,\ |u_0(t)-u_2(t)|\leq S,\ t\in[k-\lambda,k]\}$.

Proof: Let $N_2=\mathcal{N}_{(p,S)}^{[k-\lambda,k]}(u_2,y_2)^\top$. Clearly $N_1\subset N_2$, since if $v=u_0(t)-u_2(t)$ and $|v|\leq S$, then $v=\text{SAT}_S(u_0-u_2)$. Let $r_1=\min_{v_0\in N_1}\|v_0\|$. Suppose $w_0=\arg\min_{v_0\in N_2}\|v_0\|$, and by (4.32), $r_2=\|w_0\|=i_{(p,S)}[k]$. Since $N_1\subset N_2$, it follows that $r_1\geq r_2$. It thus suffices to show $w_0\in N_1$, for then $r_2\geq r_1$, and hence $i_{(p,S)}[k]=r_1$ as required.

For a contradiction, suppose $w_0=(u_0,y_0)^\top\notin N_1$ and let k be a time at which $|u_0(k)-u_2(k)|>S$. Let

$$\tilde{u}_0(t)=\begin{cases}u_0(t) & \text{if}\quad t\neq k\\ u_2(k)+S & \text{if}\quad t=k\text{ and }u_0(k)-u_2(k)>S\\ u_2(k)-S & \text{if}\quad t=k\text{ and }u_0(k)-u_2(k)<-S.\end{cases}$$

Then $\text{SAT}_S(\tilde{u}_0(t)-u_2(t))=\text{SAT}_S(u_0(t)-u_2(t))$ for all $t\geq 0$, hence $\tilde{w}_0=(\tilde{u}_0,y_0)^\top\in N_2$. But, by construction, $\|\tilde{w}_0\|<\|w_0\|$ for $1\leq r<\infty$, hence w_0 is not the minimizer in N_2. This is a contradiction, and $w_0\in N_1$ as required. ■

Hence the difference between the optimization required in the linear case to that of the case with input saturation is simply the addition of an inequality constraint. The optimization remains convex, and for example is solvable e.g., via linear programming $(r=1)$ and quadratic programming $(r=2)$. This fully addresses point b) for this class of systems.

For point b), since our analysis is global, the system class is further restricted to the class neutrally stable LTI plants with saturation, i.e., those which have the eigenvalues of the state space matrix A inside or on the unit circle, with those on the unit circle having all Jordan blocks of size one, see [4]. As a concrete example, we note the pair of saturated stable first order systems of unknown input sign, i.e., the plant model set $\{P_{(a,a,1,0)}\circ\text{SAT}_1,P_{(a,-a,1,0)}\circ\text{SAT}_1\}$, $a<1$ can be usefully controlled by EMMSAC, since the atomic pole-placement controllers can gain stabilize and meet the controller assumptions (e.g., $Cy_2=\pm(a+\gamma)y_2$, $|\gamma|<1$).

Alternatively, we can consider a saturation occurring at the output of the controller, e.g., the above equation becomes $u_1=u_0-\text{SAT}_S(u_2)$ (this is a common scenario, for example corresponding to a mechanical force actuator with limited authority and u_0 representing an external force). Then given LTI plant dynamics, the resulting estimator optimizations remain as in the fully linear setting. This underlines further that the estimator optimizations for a linear plant and a nonlinear controller are linear estimation problems only, and the nonlinearity does not complicate the estimation part of EMMSAC; thus the implementation of EMMSAC remains tractable. This idea is taken further in [5], which considers an application example of a

Hammerstein system with uncertainty in both the nonlinearity and the linear dynamics, and a tractable (and exact) estimator based on Kalman filtering is constructed.

VII. Conclusion

This paper presents comprehensive robustness and performance guarantees for Estimation-based Multiple Model Switched Adaptive Control (EMMSAC) algorithms in terms of $l_r,\ 1 \le r \le \infty$ gain (function) bounds on the gain from the external disturbances w_0 to the internal signals w_2. The axiomatic style of the analysis leads to the generality of the results: they apply to the class of minimal MIMO LTI plants but also to non-linear plants which can be gain stabilized. The axiomatic approach utiliized makes future generalizations appear inevitable, e.g., to time-varying plants and to non-linear plants with super-linear growth. Investigating the viability of the resulting estimator (sub-)optimizations in the nonlinear setting is an interesting open area; identifying tractable classes of nonlinear systems is very worthwhile. The EMMSAC approach is completely modularised: allowing for the integration of standard control designs for the atomic controllers and standard optimization approaches such as Kalman filters or convex programming methods for the estimators.

The robustness analysis leads naturally to a principled route to design, and we have shown how the complexity of the underlying uncertainty set leads to complexity dependent gain bounds for infinite dimensional controllers which can then be systematically reduced to finite dimensional realizations with guaranteed performance and robustness. The resulting complexity of the controller has also been related to metric entropy measures of the underlying uncertainty. A key consequence is that, for many geometries, the bounds are independent of the size of the candidate plant model set above a certain threshold, hence the designer can maintain (even improve) performance bounds whilst increasing the number of plant models to the maximum which are supportable in real-time. A pragmatic design guideline is therefore to populate the candidate plant set at as high a resolution as the implementation hardware constraints permit. We have thus provided an integrated, conceptual approach to address the first three questions stated in the Introduction. The fourth question concerning the construction of non-conservative designs is considered in the sequel [7].

Appendix

Proof of Proposition 7: We first consider estimator A. Let $1 \le r \le \infty$ and let $\lambda = h = \infty$. For $k \in \mathbb{N}$, $w_2 \in \mathcal{W}$, let E_A be given by

$$E_A(w_2)(k)(p) = d_p^A[k] \in \mathrm{map}(\mathbb{N}, \mathbb{R}^h) \tag{7.54}$$

$$d_p^A[k] = \mathscr{T}_k \arg \min_{w_0 \in \mathcal{N}_p^{[0,k]}(w_2)} \|w_0\| \tag{7.55}$$

if there exists a unique minimum, or any $d_p^A[k] \in \mathcal{N}_p^{[0,k]}(w_2)$ such that $\|d_p^A[k]\| \le \|v\|$ for all $v \in \mathcal{N}_p^{[0,k]}(w_2)$ if the minimum is not unique. To see that X_A factorises as $X_A = \mathrm{NE}_A$, observe that for all $w_2 \in \mathcal{W}_e$, $k \in \mathbb{N}$ and $p \in \mathcal{P}$

$$\mathrm{NE}_A(w_2)(k)(p) = \|d_p[k]\| = r_p[k] = X_A(w_2)(k)(p)$$

as required. We now verify 1–5.

1) Causality: The disturbance estimate at time $k \in \mathbb{N}$ does not depend on future information $w_2|_{(k,\infty)}$ and is therefore causal.
2) Weak consistency: Let $p \in \mathcal{P},\ w_2 \in \mathcal{W}_e$. Let Φ_λ be defined by $\Phi_\lambda x = \mathscr{R}_{\lambda,k} x,\ \ x \in \mathcal{S}$, and clearly $\|\Phi_\lambda E_A(w_2)(k)(p)\| \le \|\mathscr{R}_{\lambda,k} E_A(w_2)(k)(p)\|$. We then have
$$\begin{aligned}\Phi_\lambda E_A(w_2)(k)(p) &= \mathscr{R}_{\lambda,k} E_A(w_2)(k)(p) \\ &\in \mathscr{R}_{\lambda,k} \mathcal{N}_p^{[0,k]}(w_2) \subset \mathcal{N}_p^{[k-\lambda,k]}(w_2).\end{aligned}$$
3) Monotonicity: Let $p \in \mathcal{P}$, let $0 \le k \le l,\ \ k, l \in \mathbb{N}$. Observe that $\mathscr{T}_k E_A(w_2)(l)(p) \in \mathscr{T}_k \mathcal{N}_p^{[0,k]}(w_2)$. Since
$$\|E_A(w_2)(k)(p)\| = \inf\{r \ge 0 | r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[0,k]}(w_2)\}$$
it follows that $\|E_A(w_2)(k)(p)\| \le \|\mathscr{T}_k E_A(w_2)(l)(p)\|$ as required.
4) Continuity: For $p_1, p_2 \in \mathcal{P}$ let $\chi(p_1, p_2)$ be given by (4.41). Then $\chi(p,p) = \delta(p,p) = 0$ for all $p \in \mathcal{P}$. Define $\Pi_p^{[0,k]} : \mathcal{W}|_{[0,k]} \to \mathcal{W}|_{[0,k]}$ by the projection: $\Pi_p^{[0,k]} \mathscr{R}_{k,k} w_2 = d_p^A[k]$. Then
$$\begin{aligned}&\|E(w_2)(k)(p_1) - E(w_2)(k)(p_2)\| \\ &\quad \le \|\Pi_{p_1}^{[0,k]} \mathscr{R}_{k,k} w_2 - \Pi_{p_2}^{[0,k]} \mathscr{R}_{k,k} w_2\| \\ &\quad \le \|\Pi_{p_1}^{[0,k]} - \Pi_{p_2}^{[0,k]}\| \|\mathscr{R}_{k,k} w_2\| \\ &\quad = \chi(p_1, p_2) \|\mathscr{T}_k w_2\|.\end{aligned}$$
It remains to show $\chi(p_1, p_2) < \infty$. Define $L_k : \mathcal{W} \to \mathcal{W}$ by: $L_k = \mathscr{T}_k(\Pi_{p_1}^{[0,k]} - \Pi_{p_2}^{[0,k]}) \mathscr{R}_{k,k}$. It is easily follows that $\|L_k\| = \|\Pi_{p_1}^{[0,k]} - \Pi_{p_2}^{[0,k]}\|$. Now, for all $w_2 \in \mathcal{W}$
$$\begin{aligned}\sup_{k \ge 0} \|L_k w_2\| &= \|\mathscr{T}_k(\Pi_{p_1}^{[0,k]} - \Pi_{p_2}^{[0,k]}) \mathscr{R}_{k,k} w_2\| \\ &\le (\|(\Pi_{p_1}^{[0,k]}\| + \|\Pi_{p_2}^{[0,k]}\|) \|\mathscr{R}_{k,k} w_2\| \\ &\le 2\|w_2\| < \infty.\end{aligned}$$
Hence by the Banach-Steinhaus Theorem, $\chi(p_1, p_2) = \sup_{k \ge 0} \|L_k\| < \infty$.
5) Minimality: Observe that for any $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying (2.2), (2.3) for $P = P_p$ and for $k \in \mathbb{N}$ we have $\mathscr{T}_k w_0 \in \mathscr{T}_k \mathcal{N}_p^{[0,k]}(w_2)$. Hence by the definition of E_A, $\|E_A(w_2)(k)(p)\| \le \|\mathscr{T}_k w_0\|$.

We now consider estimator B. Let $1 \le r \le \infty$ and let $\lambda \in \mathbb{N}$, $h = (m+o)(\lambda+1)$. For $0 \le i \le k$, $w_2 \in \mathcal{W}$, let estimator B be given by

$$E_B(w_2)(k)(p) = d_p^B[k] \in \mathrm{map}(\mathbb{N}, \mathbb{R}^h) \tag{7.56}$$

$$d_p^B[k](i) = \arg \min_{w_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)} \|w_0\| \tag{7.57}$$

if there exists a unique minimum, or any $d_p^B[k](i) \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)$ satisfying $\|d_p^B[k](i)\| \le \|v\|$ for all $v \in \mathcal{N}_p^{[i-\lambda,i]}(w_2)$ if the minimum is not unique. To see that X_B does indeed factorise as $X_B = \mathrm{NE}_B$, we argue as follows. Since

$d_p^B[k](i) = d_p^B[i](i)$ for $0 \le i \le k$ and $\|a,b\|_r = \|\|a\|_r, \|b\|_r\|_r$, $a, b \in l_r,\ 1 \le r \le \infty$, we have for all $w_2 \in \mathcal{W}_2$ that

$$\begin{aligned}\mathrm{NE}_B(w_2)(k)(p) &= \|d_p[k]\| \\ &= \|d_p[k](0), d_p[k](1), \cdots, d_p[k-1](k), d_p[k](k)\| \\ &= \|d_p[k-1](0), d_p[k-1](1), \cdots, \\ &\qquad d_p[k-1](k-1), d_p[k](k)\| \\ &= \|\|d_p[k-1]\|, \|d_p[k](k)\|\|\end{aligned}$$

where $d_p[k](k) \in \{w_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2) \mid \|w_0\| = \inf\{r \ge 0 \mid r = \|v_0\|,\ v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\}$. Since $i_p[k] = \|d_p[k](k)\| = \inf\{r \ge 0 \mid r = \|v_0\|, v_0 \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)\}$, and $NE_B(w_2)(k)(p) = \|d_p[k]\|$, we obtain

$$\mathrm{NE}_B(w_2)(k)(p) = \|r_p[k-1], i_p[k]\| = X_B(w_2)(k)(p).$$

Hence $X_B = \mathrm{NE}_B$ as required. We now verify 1–5.

1) Causality: E_B is invariant to $w_2|_{(k,\infty)}$. 2) Weak consistency: Let $p \in \mathcal{P}$. Let Φ_λ be defined by $\Phi_\lambda d_p^B[k] = \mathscr{R}_{\lambda,\lambda} d_p^B[k](k)$. Since $\mathscr{R}_{\lambda,\lambda} d_p^B[k](k) \subset \mathscr{R}_{\lambda,k} d_p^B[k]$, it follows that $\|\Phi_\lambda E_B(w_2)(k)(p)\| = \|\mathscr{R}_{\lambda,\lambda} d_p^B[k](k)\| \le \|\mathscr{R}_{\lambda,k} d_p^B[k]\| = \|\mathscr{R}_{\lambda,k} E_B(w_2)(k)(p)\|$. Furthermore, $\Phi_\lambda d_p^B[k] = \mathscr{R}_{\lambda,\lambda} d_p^B[k](k) \in \mathcal{N}_p^{[k-\lambda,k]}(w_2)$. 3) Monotonicity: Let $p \in \mathcal{P}$, let $0 \le k \le l$, $k, l \in \mathbb{N}$. Since $\mathscr{T}_k d_p^B[l] = d_p^B[k]$ it follows that $\|E_p^B(w_2)(k)(p)\| = \|\mathscr{T}_k E_p^B(w_2)(l)(p)\|$ as required. 4) Continuity: Let $k \in \mathbb{N}$, $p \in \mathcal{P}$. From Assumption 2 let Φ_λ be defined by $\Phi_\lambda d_p^B[k] = \mathscr{R}_{\lambda,\lambda} d_p^B[k](k)$. Define $\Pi_p^{[k-\lambda,k]} : \mathcal{W}|_{[k-\lambda,k]} \to \mathcal{W}|_{[k-\lambda,k]}$ by the projection: $\Pi_p^{[k-\lambda,k]} \mathscr{R}_{\lambda,k} w_2 = d_p^B[k](k)$. For $p_1, p_2 \in \mathcal{P}$ let $\chi(p_1, p_2)$ be given by (4.42). It follows trivially that $\chi(p,p) = 0, p \in \mathcal{P}$. Since $\Phi_\lambda d_p^B[k] = \Pi_p^{[k-\lambda,k]} \mathscr{R}_{\lambda,k} w_2$, it follows that:

$$\begin{aligned}&\|E(w_2)(k)(p_1) - E(w_2)(k)(p_2)\| \\ &= \|d_{p_1}[0](0) - d_{p_2}[0](0), \ldots, d_{p_1}[k](k) - d_{p_2}[k](k)\|\| \\ &= \|\Pi_{p_1}^{[-\lambda,0]} \mathscr{R}_{\lambda,0} w_2 - \Pi_{p_2}^{[-\lambda,0]} \mathscr{R}_{\lambda,0} w_2, \\ &\qquad \ldots, \Pi_{p_1}^{[k-\lambda,k]} \mathscr{R}_{\lambda,k} w_2 - \Pi_{p_2}^{[k-\lambda,k]} \mathscr{R}_{\lambda,k} w_2\| \\ &\le \max_{k \ge 0} \|\Pi_{p_1}^{[k-\lambda,k]} - \Pi_{p_2}^{[k-\lambda,k]}\| \|\|\mathscr{R}_{\lambda,0} w_0\|, \ldots, \|\mathscr{R}_{\lambda,k} w_0\|\| \\ &\le \chi(p_1,p_2) \|\mathscr{T}_k w_0\|\end{aligned}$$

5) Minimality: Observe that for any $(w_0, w_1, w_2) \in \mathcal{W} \times \mathcal{W}_e \times \mathcal{W}_e$ satisfying (2.2), (2.3) for $P = P_p$ and for $k \in \mathbb{N}$ we have $\mathscr{R}_{\lambda,i} w_0 \in \mathcal{N}_p^{[i-\lambda,i]}(w_2),\ 0 \le i \le k$. Hence by the definition of $d_p^B[k](i)$, $\|d_p^B[k](i)\| \le \|\mathscr{R}_{\lambda,i} w_0\|,\ 0 \le i \le k,\ k \in \mathbb{N}$. Hence we obtain

$$\begin{aligned}\|E_B(w_2)(k)(p)\| &= \|\|d_p^B[k](0)\|, \ldots, \|d_p^B[k](k)\|\| \\ &\le \|\|\mathscr{R}_{\lambda,0} w_0\|, \ldots, \|\mathscr{R}_{\lambda,k} w_0\|\| \\ &\le (\lambda+1)^{\frac{1}{r}} \|\mathscr{T}_k w_0\|\end{aligned}$$

which is the required inequality (with $\mu = (\lambda+1)^{1/r}$). ■

Proof of Proposition 8: Let $k \ge 0$ and $\Pi_1^k = I - \Pi_{p_1}^{[0,k]}$, $\Pi_2^k = I - \Pi_{p_2}^{[0,k]}$, $\Pi_1 = I - \Pi_1^\infty$, $\Pi_2 = I - \Pi_2^\infty$. The following identity holds for any projections Π_A, Π_B:

$$\begin{pmatrix} \Pi_B \\ \Pi_B^\perp \end{pmatrix} (\Pi_A - \Pi_B) \begin{pmatrix} \Pi_A^\perp & \Pi_A \end{pmatrix} = \begin{pmatrix} -\Pi_B \Pi_A^\perp & 0 \\ 0 & \Pi_B^\perp \Pi_A \end{pmatrix}. \quad (7.58)$$

Let $\Pi_A = \Pi_1^k$, $\Pi_B = \Pi_2^k$. Then since $(\Pi_B\ \Pi_B^\perp)^\top$ and $(\Pi_A^\perp\ \Pi_A)$ are isometric isomorphisms, it follows that: $\|\Pi_1^k - \Pi_2^k\| = \max\{\|\Pi_2^k(\Pi_1^k))^\perp\ \|, \|(\Pi_2^k)^\perp \Pi_1^k\|\}$. Since $(\Pi_1^k)^\perp \Pi_2^k$ has the adjoint: $((\Pi_1^k)^\perp \Pi_2^k)^* = \Pi_2^k (\Pi_1^k)^\perp$, it follows that: $\|\Pi_1^{[0,k]} - \Pi_2^{[0,k]}\| = \max\{\|(\Pi_1^k)^\perp \Pi_2^k\|, \|(\Pi_2^k)^\perp \Pi_1^k\|\}$.

Let $L = (M\ N)$ be a normalized right co-prime factorization of P_1 over H_∞. Let $L^* = (M^*\ N^*)$, so we have the Bezout identity $L^* L = I$. For every $z_k \in \mathscr{R}_{k,k} \mathcal{G}_{p_1}$, there exists $\tilde{z} \in \mathcal{G}_{p_1}$ such that $\mathscr{R}_{k,k} \tilde{z} = z_k$. Since $\tilde{z} \in \mathcal{G}_{p_1}$ it follows that there exists $v \in \mathcal{W}$ such that $\tilde{z} = Lv$. Then by the causality of L, we have $\mathscr{T}_k z_k = \mathscr{T}_k \tilde{z}_k = \mathscr{T}_k L \mathscr{T}_k v$. Define $\tilde{z}_k = L \mathscr{T}_k v$. Since L is bounded, and $\mathscr{T}_k v \in \mathcal{W}$, it follows that $\tilde{z}_k \in \mathcal{G}_{p_1}$. Hence by the causality of L, $\mathscr{R}_{k,k} \tilde{z}_k = \mathscr{R}_{k,k} L \mathscr{T}_k v = \mathscr{R}_{k,k} \mathscr{T}_k L \mathscr{T}_k v = \mathscr{R}_{k,k} \mathscr{T}_k z_k = z_k$.

Observe that since $\tilde{z}_k = L \mathscr{T}_k v_k$, it follows that $L^* \tilde{z}_k = \mathscr{T}_k v$ and $\mathscr{T}_k L^* \tilde{z}_k = \mathscr{T}_k v$. Furthermore, since L is normalized and L^* is causal, we have: $\|\tilde{z}_k\| = \|L \mathscr{T}_k v\| = \|\mathscr{T}_k v\| = \|\mathscr{T}_k L^* \tilde{z}_k\| = \|\mathscr{T}_k L^* \mathscr{T}_k \tilde{z}_k\| = \|\mathscr{T}_k L^* \mathscr{T}_k z_k\|$. It is straightforward to see that $\|\mathscr{T}_k L^* \mathscr{T}_k\| \le \|L^*\| \le 1$, hence $\|\tilde{z}_k\| \le \|z_k\|$.

Since $\Pi_2^k : \mathcal{W}_{[0,k]} \to \mathscr{R}_{k,k} \mathcal{G}_{p_2}$ is a projection, $\tilde{z}_k \in \mathcal{G}_{p_1}$ and hence $\mathscr{R}_{k,k} \Pi_2 \tilde{z}_k \in \mathscr{R}_{k,k} \mathcal{G}_{p_2}$, it follows that for all $z_k \in \mathscr{R}_{k,k} \mathcal{G}_{p_1}$: $\|(\Pi_2^k)^\perp z_k\| = \|z_k - \Pi_2^k z_k\| \le \|z_k - \mathscr{R}_{k,k} \Pi_2 \tilde{z}_k\| \le \|\mathscr{R}_{k,k} \Pi_2^\perp \tilde{z}_k\|$, hence it follows that:

$$\begin{aligned}\|(\Pi_2^k)^\perp \Pi_1^k\| &= \|(\Pi_2^k)^\perp|_{\mathscr{R}_{k,k} \mathcal{G}_{p_1}}\| \\ &\le \sup_{z_k \in \mathscr{R}_{k,k} \mathcal{G}_{p_1}} \frac{\|\mathscr{R}_{k,k} \Pi_2^\perp \tilde{z}_k\|}{\|\tilde{z}_k\|} \frac{\|\tilde{z}_k\|}{\|z_k\|} \\ &\le \|\Pi_2^\perp|_{\mathcal{G}_{p_1}}\| \sup_{z_k \in \mathscr{R}_{k,k} \mathcal{G}_{p_1}} \frac{\|\tilde{z}_k\|}{\|z_k\|} \le \|\Pi_2^\perp \Pi_1\|\end{aligned}$$

Hence $\|(\Pi_2^k)^\perp \Pi_1^k\| \le \|\Pi_2^\perp \Pi_1\| = \|(I - \Pi_2)\Pi_1\| = \vec{\delta}(p_1, p_2)$, where the equality with the directed gap follows from [12]. Similarly $\|(\Pi_1^k)^\perp \Pi_2^k\| \le \vec{\delta}(p_2, p_1)$. Hence $\|\Pi_1^{[0,k]} - \Pi_2^{[0,k]}\| \le \delta(p_1, p_2)$. Since this holds for all k, the proof is complete. ■

Proof of Proposition 9: Let $k \ge 0$ and $\Pi_A = \Pi_{p_1}^{[k-\lambda,k]}$, $\Pi_B = \Pi_{p_2}^{[k-\lambda,k]}$. Since Π_A, Π_B are projections, by an analogous argument to the proof of Proposition 8, we have $\|\Pi_A - \Pi_B\| \le c_0 \max\{\|\Pi_B \Pi_A^\perp\ \|, \|\Pi_B^\perp \Pi_A\|\}$ for some $c_0 > 0$, since the identity (7.58) holds and where $(\Pi_B^\perp\ \Pi_B)^\top$ and $(\Pi_A^\perp\ \Pi_A)$ are isomorphisms (not isometric for $r \ne 21$). It therefore suffices to check the continuity of $\|\Pi_B \Pi_A^\perp\|$, $\|\Pi_B^\perp \Pi_A\|$ w.r.t. p_1, p_2. Let $w_2 = (u_2, y_2)^\top \in \mathcal{W}|_{[k-\lambda,k]} \setminus \{0\}$, $w_0^{p_1} = (u_0^{p_1}, y_0^{p_1})^\top = \Pi_A w_2$ and $w_1^{p_1} = -\Pi_A^\perp w_2$. Let $x_0 \in \mathbb{R}^n$ be such that (2.6)–(2.8) hold on $[k-\lambda, k]$ with $x(k-\lambda) = x^{p_1}(k-\lambda) = x_0$. Let $y_1^{p_2} \in \mathcal{Y}|_{[k-\lambda,k]}$ be the output of the system p_2 when $x^{p_2}(k-\lambda) = x^{p_1}(k-\lambda)$, and $u_1^{p_2} = u_1^{p_1} \in \mathcal{U}|_{[k-\lambda,k]}$.

We first consider $\|\Pi_B^\perp \Pi_A\|$. Let $w_0^{p_2} = (u_2 + u_1^{p_2}, y_2 + y_1^{p_2})^\top$. Since Π_B is a projection, it follows that $\|\Pi_B^\perp \Pi_A w_2\| = \|w_0^{p_1} - \Pi_B w_0^{p_1}\| \le \|w_0^{p_1} - w_0^{p_2}\| = \|y_0^{p_1} - y_0^{p_2}\| = \|y_1^{p_1} - y_1^{p_2}\|$. We now consider $\|\Pi_B \Pi_A^\perp\|$. Let $u_2^{p_2} = -u_1^{p_1} = u_1^{p_2}$ so $u_0^{p_2} = 0$. We let $y_0^{p_2} = y_1^{p_2} - y_1^{p_1}$, so $y_2^{p_2} = -y_1^{p_1}$. By definition of Π_B, it follows that $\|\Pi_B \Pi_A^\perp w_2\| = \|\Pi_B(-w_1^{p_1})\| = \|\Pi_B(w_2^{p_2})\| \le \|w_0^{p_2}\| = \|y_0^{p_2}\| = \|y_1^{p_1} - y_1^{p_2}\|$.

Let $\epsilon > 0$. In both cases, by the continuity of the solution of a discrete time system with respect to the entries of (A, B, C) over a finite time interval, it follows that there exists $\delta > 0$ such that if $|p_1 - p_2| \leq \delta$, then $\|y_1^{p_1} - y_1^{p_2}\| \leq \epsilon$. Further, we have the bound: $\|y_1^{p_1} - y_1^{p_2}\| \leq c_1\epsilon(\|\mathscr{R}_{\lambda,k}u_1^{p_1}\| + |x_0^{p_1}|)$ for some $c_1 > 0$. By observability and $\lambda \geq \sigma$, it follows that $|x_0^{p_1}| \leq c_2\|w_1^{p_1}\| = c_2\|(I - \Pi_A)w_2\|$. Hence $\max\{\|\Pi_B\Pi_A^{\perp}w_2\|, \|\Pi_B^{\perp}\Pi_A w_2\|\} \leq c_1c_2\epsilon\|(I - \Pi_A)w_2\|$ and

$$\begin{aligned}\|\Pi_A - \Pi_B\| &\leq c_0 \max\{\|\Pi_B\Pi_A^{\perp}\|, \|\Pi_B^{\perp}\Pi_A\|\} \\ &= c_0 \sup_{w_2 \neq 0} \max\{\|\Pi_B\Pi_A^{\perp}w_2\|, \|\Pi_B^{\perp}\Pi_A w_2\|\}\|w_2\| \\ &\leq c_0c_1c_2\epsilon \sup_{w_2 \neq 0} \frac{\|\Pi_A^{\perp}w_2\|}{\|w_2\|} \leq c_0c_1c_2\|\Pi_A^{\perp}\|\epsilon.\end{aligned}$$

For all $p \in \Omega$, $\Pi_p^{[i-\lambda,i]} = \Pi_p^{[j-\lambda,j]}$ for all $i, j > \lambda + \sigma$, $i, j \in \mathbb{N}$ since $\mathcal{N}_p^{[i-\lambda,i]}(w_2) = \mathcal{N}_p^{[j-\lambda,j]}(w_2)$ for all $i, j > \lambda + \sigma$. As $\|\Pi_{p_1}^{[k-\lambda,k]} - \Pi_{p_2}^{[k-\lambda,k]}\|$ and hence $\chi_k(p_1, p_2)$ is continuous on $\Omega \times \Omega$, and $\chi(p_1, p_2) = \max_{k \geq 0} \chi_k(p_1, p_2) = \max_{0 \leq k \leq \lambda+\sigma+1} \chi_k(p_1, p_2)$, χ is continuous on $\Omega \times \Omega$. ■

ACKNOWLEDGMENT

The authors wish to thank G. Vinnicombe, for his profound perspective, and S. Trenn (TU Kaiserslautern), for many useful discussions and initial input to this research and for his key input in [11].

REFERENCES

[1] B. D. O. Anderson, "Failures of adaptive control theory and their resolution," *Commun. Inform. Syst.*, vol. 5, no. 1, pp. 1–20, 2005.

[2] B. D. O. Anderson, T. S. Brinsmead, F. De Bruyne, J. P. Hespanha, D. Liberzon, and A. S. Morse, "Multiple model adaptive control. Part 1: Finite controller coverings," *Int. J. Robust Nonlin. Control*, vol. 10, pp. 909–929, 2000.

[3] S. Baldi, G. Battistelli, E. Mosca, and P. Tesi, "Multi-model unfalsified adaptive switching supervisory control," *Automatica*, vol. 46, pp. 249–259, 2010.

[4] X. Bao, E. LIn, and E. D. Sontag, "Finite gain stabilization of discrete-time linear systems subject to actuator saturation," *Automatica*, vol. 36, pp. 269–277, 2000.

[5] O. W. Brend, C. Freeman, and M. French, "Application of multiple model adaptive control to upper limb stroke rehabilitation," in *Proc. IEEE Multi-Conf. Syst. Control*, pp. 69–74.

[6] D. Buchstaller, "Robust Stability and Performance for Multiple Model Switched Adaptive Control," Ph.D. dissertation, School of Electron. Comp. Sci., University of Southampton, Southampton, U.K., 2010. [Online]. Available: http://eprints.ecs.soton.ac.uk/18435/

[7] D. Buchstaller and M. French, "Robust stability and performance for multiple model adaptive control: Part II—Gain bounds," *IEEE Trans. Autom. Control*, vol. 61, no. 3, Mar. 2016.

[8] S. Fekri, M. Athans, and A. Pascoal, "Issues, progress and new results in robust adaptive control," *Int. J. Adaptive Control Signal Processing*, vol. 20, no. 10, pp. 519–579, 2006.

[9] T. P. Fisher-Jeffes, "Multiple-Model Switching Control to Achieve Asymptotic Robust Performance," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., 2003.

[10] M. French, "Adaptive control and robustness in the gap metric," *IEEE Trans. Autom. Control*, vol. 53, no. 2, pp. 461–478, 2008.

[11] M. French and S. Trenn, "l^p Gain bounds for switched adaptive controllers," in *Proc. 44th IEEE Conf. Decision Control Eur. Control Conf.*, Seville, Spain, 2005, pp. 2865–2870.

[12] T. T. Georgiou and M. C. Smith, "Robustness analysis of nonlinear feedback systems: An input-output approach," *IEEE Trans. Autom. Control*, vol. 42, no. 9, pp. 1200–1220, Sep. 1997.

[13] J. P. Hespanha, D. Liberzon, and A. S. Morse, "Overcoming the limitations of adaptive control by means of logic-based switching," *Syst. Control Lett.*, vol. 49, no. 1, pp. 49–65, 2003.

[14] 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. Robust Nonlin. Control*, vol. 11, pp. 479–496, 2001.

[15] P. M. Khargonekar and M. A. Rotea, "Mixed H_2/H_∞ control: A convex optimization approach," *IEEE Trans. Autom. Control*, vol. 36, pp. 824–837, 1991.

[16] D. Liberzon, *Switching in Systems and Control*. Boston, MA, USA: Birkhauser, 2003.

[17] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Hysteresis-based switching algorithms for supervisory control of uncertain systems," *Automatica*, vol. 39, pp. 263–272, 2003.

[18] A. S. Morse, "Supervisory control of families of linear set-point controllers—Part 1: Exact matching," *IEEE Trans. Autom. Control*, vol. 41, no. 10, pp. 1413–1431, Oct. 1996.

[19] A. S. Morse, "Supervisory control of families of linear set-point controllers—Part 2: Robustness," *IEEE Trans. Autom. Control*, vol. 42, no. 11, pp. 1500–1515, Nov. 1997.

[20] F. M. Pait and F. Kassab, Jr., "On a class of switched, robustly stable, adaptive systems," *Int. J. Adaptive Contr. Signal Processing*, vol. 15, pp. 213–238, 2000.

[21] C. E. Rohrs, L. Valavani, M. Athans, and G. Stein, "Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics," *IEEE Trans. Autom. Control*, vol. AC-30, no. 9, pp. 881–889, 1985.

[22] E. D. Sontag, *Mathematical Control Theory*. New York, NY, USA: Springer, 1998.

[23] M. Stefanovic and M. G. Safonov, "Safe adaptive switching control: Stability and convergence," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2012–2021, 2008.

[24] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton, NJ, USA: Princeton Univ. Press, 2005.

[25] G. Vinnicombe, "Examples and counterexamples in finite $\mathcal{L}_2$-gain adaptive control," in *Proc. 16th Int. Symp. Math. Theory Netw. Syst.*, Leuven, Belgium, 2004, pp. 1–7.

[26] R. Wang, A. Paul, M. Stefanovic, and M. G. Safonov, "Cost detectability and stability of adaptive control systems," *Int. J. Rob. Nonlin. Control*, vol. 17, pp. 549–561, 2007.

[27] J. C. Willems, "Deterministic least squares filtering," *J. Econometrics*, vol. 118, pp. 341–370, 2004.

[28] G. Zames, "Adaptive control: Towards a complexity-based general theory," *Automatica*, vol. 34, no. 10, 1998.

[29] P. V. Zhivoglyadov, R. H. Middleton, and M. Fu, "Localization based switching adaptive control for time-varying discrete-time systems," *IEEE Trans. Autom. Control*, vol. 45, no. 4, pp. 752–755, 2000.

See Part II for Author Photos and Biographies.