On a Control-Oriented Notion of Finite State Approximation

Danielle C. Tarraf

Abstract—We consider the problem of approximating discrete-time plants with finite-valued sensors and actuators by deterministic finite memory systems for the purpose of certified-by-design controller synthesis. We propose a control-oriented notion of input/output approximation for these systems, building on ideas from classical robust control theory, and we explain the relevance of the proposed notion of approximation to the control synthesis problem.

I. INTRODUCTION

This paper has two objectives: First, to formalize a control-oriented notion of finite state approximation for plants with discrete actuation and sensing, previously demonstrated for a class of systems in [1], [2] and proposed in a more primitive form in [3]. Second, to demonstrate the relevance of the proposed notion to the problem of certified-by-design control synthesis, drawing on ideas from classical robust control theory and behavioral systems theory.

High fidelity models that accurately describe a dynamical system are often too complex to use for controller design. The problem of finding a lower complexity model that approximates a given plant has thus been extensively studied and continues to receive much deserved attention (see [4], [5] and the references therein). In order to be useful in this setting, a model complexity reduction approach should provide both a lower complexity model and a rigorous assessment of the quality of approximation, allowing one to quantify the performance of a controller designed based on the lower complexity model and implemented in the actual system (assumed to be faithfully captured by the high complexity model).

One of the big success stories has been for the class of linear time-invariant systems, where model complexity is captured by the order of the system. More recently, the problem of approximating hybrid systems, that is systems involving interacting analog and discrete dynamics, by simpler systems has been considered [6]–[8]. In particular, the problem of constructing finite state (i.e. finite memory) approximations of hybrid systems has been the object of intense study over the past two decades, due to the amenability of finite state models to efficient control synthesis. Viewed at a high level, the notions of approximation that have been systematically explored can be placed under two headings: ‘Behavioral’ abstractions, and ‘simulation/bisimulation’ based abstractions.

While perhaps not standard terminology, ‘behavioral abstractions’ is used here to refer to non-deterministic finite state automata constructed so that their input/output behavior (or simply their output behavior in the case of autonomous systems) contains that of the original model. The problem of controller synthesis in this case reduces to a standard supervisory control problem for discrete-event systems, which can be addressed using the Ramadge-Wonham framework [9], [10]. The results on qualitative models [11], [12], qualitative state reconstruction from quantized input and output observations [13]–[15] and l-complete approximations [16], [17] can all be considered to fall in this category.

Inspired by the theory of bisimulation in concurrent processes [18], [19], finite bisimulation abstractions of hybrid systems have been explored. Having discovered that the classes of systems admitting finite bisimulations are limited [20], [21], recent focus has turned to extending this approach by proposing weaker notions of abstraction and less stringent metrics [22]–[25]. These notions of abstraction effectively ensure that the set of state trajectories of the original model is exactly matched by (bisimulation), contained in (simulation), matched to within some distance $\epsilon$ by (approximate bisimulation), or contained to within some distance $\epsilon$ in (approximate simulation), the set of state trajectories of the finite state abstraction. The specifications are typically constraints, possibly linear temporal formulas, on the state trajectories of the original hybrid system. The corresponding controller design problem reduces to a two step procedure: A finite state supervisory controller is first designed using the Ramadge-Wonham framework. This controller is subsequently refined to yield a certified hybrid controller for the original plant [26].

In [1], [2], we demonstrated the use of specially constructed finite state approximations for a class of systems, namely switched second order homogeneous systems with binary sensors. These finite approximations were constructed to approximately match the input/output behavior of the original model in a specific sense, quantified in terms of a “gain” bound on an appropriately defined approximation error. The finite models and the corresponding error bounds were then used to systematically and efficiently synthesize certified-by-design stabilizing finite state controllers. In this paper, we take a step back to basics, by formalizing a control oriented notion of input/output finite state approximation for discrete-time plants that are allowed to interact with their controllers via fixed finite alphabets, of which the demonstrations in [1], [2] are particular instances. While the proposed notion is primarily inspired from classical robust control theory, the class of problems considered here poses unique challenges for two main reasons:

1) The lack of algebraic structure: In general, the input and output signals of the system are assumed to take
their values in arbitrary sets of symbols.

2) The need to approximate both the model and the performance objective, while simultaneously quantifying the approximation error.

The paper is organized as follows: We begin by reviewing basic concepts that will be used in our development in Section II. We present the proposed notion of approximation in Section III and explain its relevance to the control synthesis problem in Section IV. We describe directions for future work in Section V.

II. PRELIMINARIES

We begin by briefly reviewing basic definitions that will be used in our development. Readers are referred to [27] for a more detailed treatment. We begin with a word on notation: \( \mathbb{Z}_+ \) and \( \mathbb{R}_+ \) denote the set of non-negative integers and non-negative reals, respectively. Given a set \( A \), \( A^{\mathbb{Z}_+} \) is the set of all infinite sequences over \( A \). An element of \( A \) is denoted by \( a \) while an element of \( A^{\mathbb{Z}_+} \) is denoted by boldface \( a \). For \( a \in A^{\mathbb{Z}_+} \), \( a(i) \) denotes the \( i \)th component of \( a \).

A discrete-time signal is understood to be an infinite sequence over some prescribed set, referred to as an ‘alphabet’.

**Definition 1.** A discrete-time system \( S \) is a set of pairs of signals, \( S \subseteq U^{\mathbb{Z}_+} \times Y^{\mathbb{Z}_+} \), where \( U \) and \( Y \) are given alphabets.

A discrete-time system is thus a process characterized by its feasible signals set, which is simply the set of ordered pairs of all the signals that can be applied as an input to this process, and all the output signals that can be potentially exhibited by the process in response to each of the input signals. This view of systems can be considered an extension of the graph theoretic approach [28] to include the finite alphabet setting. It also shares some similarities with the behavioral approach [29], though we insist on differentiating between input and output signals upfront. In this setting, system properties of interest are captured by means of ‘integral constraints’ on the feasible signals.

**Definition 2.** Consider a system \( S \subseteq U^{\mathbb{Z}_+} \times Y^{\mathbb{Z}_+} \) and let \( \rho : U \to \mathbb{R} \) and \( \mu : Y \to \mathbb{R} \) be given functions. \( S \) is \( \rho/\mu \) gain stable if there exists a finite non-negative constant \( \gamma \) such that the following inequality is satisfied for all \((u, y)\) in \( S\):

\[
\inf_{T \geq 0} \sum_{t=0}^{T} \gamma \rho(u(t)) - \mu(y(t)) > -\infty.
\]

In particular, when \( \rho \) and \( \mu \) are non-negative (and not identically zero), a notion of ‘gain’ can be defined.

**Definition 3.** Consider a system \( S \subseteq U^{\mathbb{Z}_+} \times Y^{\mathbb{Z}_+} \). Assume that \( S \) is \( \rho/\mu \) gain stable for \( \rho : U \to \mathbb{R}_+ \) and \( \mu : Y \to \mathbb{R}_+ \), and that neither function is identically zero. The \( \rho/\mu \) gain of \( S \) is then the infimum of \( \gamma \) such that (1) is satisfied.

**Remark 1.** These notions of ‘gain stability’ and ‘gain’ can be considered an extension of the classical definitions to include the finite alphabet setting. When \( U \) and \( Y \) are Euclidean vector spaces and \( \rho \) and \( \mu \) are the corresponding Euclidean norms, we simply recover the standard definitions of \( l_2 \) stability and \( l_2 \) gain.

We are specifically interested in discrete-time plants that interact with their controllers through fixed discrete alphabets (i.e. plants with finite-valued actuators and sensors). We will refer to such plants as ‘systems over finite alphabets’.

**Definition 4.** A system over finite alphabets \( S \) is a discrete-time system \( S \subseteq (U \times \mathbb{R})^{\mathbb{Z}_+} \times (Y \times \mathbb{R})^{\mathbb{Z}_+} \) whose alphabets \( U \) and \( Y \) are finite.

In this setting, \( r \in \mathbb{R}^{\mathbb{Z}_+} \) represents an exogenous input to the plant, \( u \in U^{\mathbb{Z}_+} \) represents the control input to the plant, \( v \in Y^{\mathbb{Z}_+} \) represents the performance output of the plant, and \( y \in Y^{\mathbb{Z}_+} \) represents the sensor output of the plant. No restrictions are assumed on the internal dynamics of the plant: The underlying state-space may be analog, discrete or hybrid. Likewise, alphabets \( R \) and \( V \) may be finite, countably infinite or simply infinite.

The models that will be used to approximate systems over finite alphabets will be drawn from a specific class of models, namely the class of deterministic finite state machines (DFM).

**Definition 5.** A deterministic finite state machine (DFM) is a discrete-time system \( S \subseteq U^{\mathbb{Z}_+} \times Y^{\mathbb{Z}_+} \) with finite alphabets \( U \) and \( Y \), whose feasible input and output signals \((u, y)\) are related by

\[
q(t+1) = f(q(t), u(t)) \quad y(t) = g(q(t), u(t))
\]

where \( t \in \mathbb{Z}_+ \), \( q(t) \in Q \) for some finite set \( Q \) and functions \( f : Q \times U \to Q \) and \( g : Q \times U \to Y \).

In this setting, it is understood that finite set \( Q \) and functions \( f \) and \( g \) represent the set of states of the DFM, the state transition map, and the output map, respectively, in the traditional state-space sense.

III. PROPOSED NOTION: INPUT/OUTPUT APPROXIMATION

In this section we formalize a control-oriented notion of finite state approximation for systems over finite alphabets. Our assumption is that the purpose of deriving a DFM approximation of a system \( P \) over finite alphabets is to simplify the process of synthesizing a controller \( K \) such that the closed loop system \((P, K)\) is \( \rho/\mu \) gain stable with \( \gamma = 1 \) for some given functions \( \rho \) and \( \mu \).

We begin by introducing some further notation: Given a system \( P \subseteq (U \times \mathbb{R})^{\mathbb{Z}_+} \times (Y \times \mathbb{R})^{\mathbb{Z}_+} \) and a choice of signals \( u_o \in U^{\mathbb{Z}_+} \) and \( y_o \in Y^{\mathbb{Z}_+} \), we use the notation \( P|_{u_o, y_o} \) to denote the subset of feasible signals of \( P \) whose first component is \( u_o \) and whose third component is \( y_o \). That is

\[
P|_{u_o, y_o} = \left\{ (u, r, y, v) \in P | u = u_o \text{ and } y = y_o \right\}.
\]

Note that \( P|_{u_o, y_o} \) may be an empty set for specific choices of \( u_o \) and \( y_o \).
Definition 6 (Notion of DFM Approximation). Consider a system over finite alphabets $P \subset (\mathcal{U} \times \mathcal{R})^Z_+ \times (\mathcal{Y} \times \mathcal{V})^Z_+$ and a desired closed loop performance objective
\[
\inf_{T \geq 0} \sum_{t=0}^T \rho(r(t)) - \mu(v(t)) > -\infty
\]
for given functions $\rho : \mathcal{R} \to \mathbb{R}$ and $\mu : \mathcal{V} \to \mathbb{R}$. A sequence \( \{M_i\} \) of deterministic finite state machines
\[
\hat{M}_i \subset (\mathcal{U} \times \hat{\mathcal{R}}_i \times \mathcal{W})^Z_+ \times (\mathcal{Y} \times \hat{\mathcal{V}}_i \times \mathcal{Z})^Z_+
\]
with $\hat{\mathcal{R}}_i \subset \mathcal{R}$ and $\hat{\mathcal{V}}_i \subset \mathcal{V}$ is a $\rho/\mu$ approximation of $P$ if there exists a corresponding sequence of systems \( \{\Delta_i\} \), \( \Delta_i \subset \mathcal{Z}^Z_+ \times \mathcal{W}^Z_+ \), and non-zero functions $\rho_\Delta : \mathcal{Z}^Z_+ \to \mathbb{R}_+$ and $\mu_\Delta : \mathcal{W}^Z_+ \to \mathbb{R}_+$, such that for every $i$:

1) There exists a surjective map $\psi_i : P \to \hat{P}_i$ satisfying
\[
\psi_i(P_{\mid u,y}) \subseteq \hat{P}_{\mid u,y}
\]
for all $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$, where
\[
\hat{P}_i \subset (\mathcal{U} \times \hat{\mathcal{R}}_i)^Z_+ \times (\mathcal{Y} \times \hat{\mathcal{V}}_i)^Z_+
\]
is the feedback interconnection of $\hat{M}_i$ and $\Delta_i$ as shown in Figure 1.

2) For every feasible signal $((u, r), (y, v)) \in P$, we have
\[
\rho(r(t)) - \mu(v(t)) \geq \rho(\hat{r}_i(t)) - \mu(\hat{v}_i(t)), \quad \forall t \in \mathbb{Z}_+
\]
where
\[
((u, \hat{r}_i), (y, \hat{v}_i)) = \psi_i((u, r), (y, v))
\]

3) $\Delta_i$ is $\rho_\Delta/\mu_\Delta$ gain stable, and moreover, the corresponding $\rho_\Delta/\mu_\Delta$ gains satisfy $\gamma_i \geq \gamma_{i+1}$.

IV. RELEVANCE OF THE PROPOSED NOTION

In this section, we demonstrate the relevance of the proposed notion of approximation to the problem of certified-by-design controller synthesis for the plant. We begin by establishing several facts that will aid in our understanding of the proposed notion of approximation.

Proposition 1. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. The sets $P_{\mid u,y}$, $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$, partition $P$ into equivalence classes. For each $i$, the sets $\hat{P}_{\mid u,y}$, $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$, partition $\hat{P}_i$ into equivalence classes.

Proof: It immediately follows from the definition that $P_{\mid u_1,y_1} \cap P_{\mid u_2,y_2} = \emptyset$ whenever $(u_1, y_1) \neq (u_2, y_2)$. It also follows from the definition that every $((u, r), (y, v)))$ in $P$ belongs to some $P_{\mid u,y}$, hence $\bigcup_{u,y} P_{\mid u,y} = P$. The proof for each $\hat{P}_i$ is similar and is thus omitted.

Proposition 2. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. For every $i$, $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$, we have $\psi_i(P_{\mid u,y}) = \hat{P}_{\mid u,y}$.

Proof: By condition (1) of Definition 6, for each $i$ there exists a $\psi_i : P \to \hat{P}_i$ with $\psi_i(P_{\mid u,y}) \subseteq \hat{P}_{\mid u,y}$ for all $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$. What remains is to show equality. Fix index $i$. For a given choice of $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$:

If $\hat{P}_{\mid u,y} = \emptyset$, we have $\psi_i(P_{\mid u,y}) \subseteq \hat{P}_{\mid u,y} = \emptyset$, and equality holds. Otherwise, assume there exists an $x \in \hat{P}_{\mid u,y}$, $x \notin \psi_i(P_{\mid u,y})$. Since $\psi_i$ is surjective, $x \in \psi_i(P_{\mid u,y})$ for some $(u_1, y_1) \neq (u, y)$. We then have $x \not\in \hat{P}_{\mid u,y} \cap \hat{P}_{\mid u,y}$, leading to a contradiction by Proposition 1. Thus, such an $x$ cannot exist, and equality holds. Finally, note that the proof is independent of the choice of index $i$.

Corollary 1. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. For every $i$, $(u, y) \in \mathcal{U}^Z_+ \times \mathcal{Y}^Z_+$, we have $P_{\mid u,y} = \emptyset$ iff $\hat{P}_{\mid u,y} = \emptyset$.

Proof: For any index $i$, we have
\[
\hat{P}_{\mid u,y} = \emptyset \iff \psi_i(P_{\mid u,y}) = \emptyset \iff P_{\mid u,y} = \emptyset
\]
where the first equivalence follows from Proposition 2.

The consequence of these simple facts is as follows: If we were to partition each of $P$ and $\hat{P}_i$ into equivalence classes of feasible signals having identical first and third components.
Corollary 2. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. For each $i$, there exists a 1-1 correspondence between the equivalence classes \( \{P_{i,u,y}\} \) of $P$ and the equivalence classes \( \{\hat{P}_{i,u,y}\} \) of $\hat{P}_i$.

Proof: For any index $i$, consider the map $\psi_i : \{\hat{P}_{i,u,y}\} \to \{P_{i,u,y}\}$ (with a slight abuse of notation). It follows from Proposition 2 and Corollary 1 that $\psi_i$ is both surjective and injective.

Lemma 1. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. For any given choice of $i$, $(u, y) \in U^Z_+ \times Y^Z_+$, if every $(u, \hat{r}, (y, \hat{v})) \in \hat{P}_{i,u,y}$ satisfies

$$\inf_{T \geq 0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)) > -\infty$$

then every $(u, r, (y, v)) \in P_{i,u,y}$ satisfies

$$\inf_{T \geq 0} \sum_{t=0}^{T} \rho(r(t)) - \mu(v(t)) > -\infty$$

Proof: Fix $i$ and consider any $(u, y) \in U^Z_+ \times Y^Z_+$. If $\hat{P}_{i,u,y} = \emptyset$, then $P_{i,u,y} = \emptyset$ by Corollary 1 and the statement holds vacuously. Now suppose that $\hat{P}_{i,u,y} \neq \emptyset$ and every $(u, \hat{r}, (y, \hat{v})) \in \hat{P}_{i,u,y}$ satisfies (4). Pick any $(u, r, (y, v)) \in P_{i,u,y}$ and consider its image $\psi_i((u, r, (y, v))) = ((u, \hat{r}), (y, \hat{v}))$. By condition (2) of Definition 6, we have

$$\rho(r(t)) - \mu(v(t)) \geq \rho(\hat{r}(t)) - \mu(\hat{v}(t)), \quad \forall t \in Z_+$$

$$\Rightarrow \sum_{t=0}^{T} \rho(r(t)) - \mu(v(t)) \geq \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)), \quad \forall T \in Z_+$$

$$\Rightarrow \sum_{t=0}^{T} \rho(r(t)) - \mu(v(t)) \geq \inf_{T \geq 0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)),$$

$$\forall T \in Z_+$$

$$\Rightarrow \inf_{T \geq 0} \sum_{t=0}^{T} \rho(r(t)) - \mu(v(t)) \geq \inf_{T \geq 0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t))$$

Thus it follows that if every element of $\hat{P}_{i,u,y}$ is such that (4) is satisfied, then every element of $P_{i,u,y}$ is such that (2) is satisfied.

We are now ready to turn our attention to the problem of control synthesis.

Theorem 1. Consider a plant $P$ and a $\rho/\mu$ approximation \( \{M_i\} \) as in Definition 6. Let $K \subset \hat{Y}^Z_+ \times U^Z_+$ be such that the feedback interconnection of $\hat{P}_i$ and $K$, $(\hat{P}_i, K) \subset \hat{R}_i^Z \times \hat{V}_i^Z$, satisfies

$$\inf_{T \geq 0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)) > -\infty$$

for some index $i$. Then the feedback interconnection of $P$ and $K$, $(P, K) \subset \hat{R}_i^Z \times \hat{V}_i^Z$, satisfies

$$\inf_{T \geq 0} \sum_{t=0}^{T} \rho(r(t)) - \mu(v(t)) > -\infty$$

Proof: Let

$$P_{i,K} = \left\{ ((u, r), (y, v)) \in P \mid (y, u) \in K \right\},$$

$$\hat{P}_{i,K} = \left\{ ((u, \hat{r}), (y, \hat{v})) \in \hat{P}_i \mid (y, u) \in K \right\}.$$
robust control setting, the idea is to design \( K \) such that the interconnection of \( M_i, K \) and any \( \Delta \) in the class \( \Delta_i \)

\[
\Delta_i = \{ \Delta \in \mathbb{Z}^+ \times \mathbb{W}^+ | \inf_{T \geq T_0} \sum_{t=0}^{T} \gamma_i \rho_{\Delta}(z(t)) - \mu_{\Delta}(w(t)) > -\infty \}
\]

satisfies the auxiliary performance objective (4). This synthesis problem can be elegantly formulated using the ‘Small Gain Theorem’ proposed in [27].

**Theorem 2 (Small Gain Theorem - Adapted from [27]).** Consider the feedback interconnection of two systems \( S \) and \( \Delta \) as in Figure 2. If \( S \) satisfies

\[
\inf_{T \geq T_0} \sum_{t=0}^{T} \rho_S(\hat{r}(t), w(t)) - \mu_S(\hat{v}(t), z(t)) > -\infty \quad (5)
\]

for some \( \rho_S : \hat{\mathcal{R}} \times \mathcal{W} \to \mathbb{R}, \mu_S : \hat{\mathcal{V}} \times \mathcal{Z} \to \mathbb{R} \) (\( \hat{\mathcal{R}}, \mathcal{W}, \hat{\mathcal{V}} \) and \( \mathcal{Z} \) are finite alphabets), and \( \Delta \) satisfies

\[
\inf_{T \geq T_0} \sum_{t=0}^{T} \gamma_{\Delta} \rho_{\Delta}(u(t)) - \mu_{\Delta}(w(t)) > -\infty
\]

for some scalar \( \gamma_{\Delta}, \rho_{\Delta} : \mathcal{Z} \to \mathbb{R}, \mu_{\Delta} : \mathcal{W} \to \mathbb{R} \), then \((S, \Delta)\) satisfies

\[
\inf_{T \geq T_0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)) > -\infty
\]

for \( \rho : \hat{\mathcal{R}} \to \mathbb{R}, \mu : \hat{\mathcal{V}} \to \mathbb{R} \) defined by

\[
\rho(\hat{r}) = \max_{w \in \mathcal{W}} \{ \rho_S(\hat{r}, w) - \tau \mu_{\Delta}(w) \}
\]

\[
\mu(\hat{v}) = \min_{z \in \mathcal{Z}} \{ \mu_S(\hat{v}, z) - \tau \gamma_{\Delta} \rho_{\Delta}(z) \}
\]

for any \( \tau > 0 \). \( \square \)

Interpreting Theorem 2 in the setting where “\( S \)” represents the feedback interconnection of \( \hat{M}_i \) and controller \( K \), and where “\( \Delta \)” represents the corresponding approximation error \( \Delta_i \), we can formulate the following statement.

**Theorem 3.** Consider a plant \( P \) and a \( \rho/\mu \) approximation \( \{ \hat{M}_i \} \) as in Definition 6. If for some index \( i \), there exists a controller \( K \subset \hat{\mathcal{Y}}^+ \times \mathcal{U}^+ \) such that the feedback interconnection \( \langle \hat{M}_i, K \rangle \subset (\hat{\mathcal{R}}_i \times \mathcal{W})^{\mathbb{Z}^+} \times (\hat{\mathcal{V}}_i \times \mathcal{Z})^{\mathbb{Z}^+} \) satisfies

\[
\inf_{T \geq T_0} \sum_{t=0}^{T} \rho(\hat{r}(t)) + \tau \mu_{\Delta}(w(t)) - \mu(\hat{v}(t)) > -\infty
\]

for some \( \tau > 0 \), then the feedback interconnection of \( \hat{P}_i \) and \( K, \langle \hat{P}_i, K \rangle \subset \hat{\mathcal{R}}_i^{\mathbb{Z}^+} \times \hat{\mathcal{V}}_i^{\mathbb{Z}^+} \), satisfies

\[
\inf_{T \geq T_0} \sum_{t=0}^{T} \rho(\hat{r}(t)) - \mu(\hat{v}(t)) > -\infty \quad (4)
\]

**Proof:** Letting \( S = \langle \hat{M}_i, K \rangle, \Delta = \Delta_i, \rho_S(\hat{r}, w) = \rho(\hat{r}) + \tau \mu_{\Delta}(w), \mu_S(\hat{v}, z) = \mu(\hat{v}) - \tau \gamma_{\Delta} \rho_{\Delta}(z), \) and \( \gamma_{\Delta} = \gamma_i \), we have by Theorem 2 that the interconnection of \( K, \hat{M}_i \) and \( \Delta_i \) satisfies (4). Equivalently, the feedback interconnection of \( \langle \hat{P}_i, K \rangle \) satisfies (4).

The problem of designing a controller \( K \) for a DFM \( M_i \) so that the closed loop system satisfies a gain condition such as (6) can be systematically addressed by solving a corresponding discrete minimax problem using techniques inspired from dynamic programming. Interested readers are referred to [2] for the details of the approach.

Intuitively, the availability of finite state approximations of a plant in the sense of Definition 6 allows one to successively replace the original synthesis problem by two problems: The first, proposed in Theorem 1, allows us to approximate the performance objectives in the general setting where the exogenous input and performance output of the plant are not finite valued. The second, proposed in Theorem 3, allows us to simplify the design problem at the expense of additional conservatism due to the introduction of a set based description of the approximate model, along the lines of what is traditionally done in robust control. In practice, since the approximation error \( \Delta_i \) may be an arbitrarily complex system, exact computation of its gain may be computationally prohibitive, if not impossible. Rather, gain bounds would be typically obtained, leading to a hierarchy of synthesis problems and corresponding controllers.

**V. CONCLUSIONS AND FUTURE WORKS**

We considered the problem of approximating plants with discrete sensors and actuators (termed ‘systems over finite alphabets’) by deterministic finite memory systems for the purpose of certified-by-design controller synthesis. We formalized a new control-oriented notion of input/output approximation for these systems, and we explained its relevance to the problem of certified-by-design control synthesis.

Future work will focus on:

- Exploring the limitations imposed by the use of gain conditions to encode performance objectives. A question of particular interest is whether linear temporal logic constraints can be alternatively described by gain conditions.
- Developing algorithmic approaches for constructing approximate models for broader classes of systems than those already addressed in [1], [2].
VI. ACKNOWLEDGMENTS

The author is indebted to A. Megretski for many stimulating discussions. The author thanks M. A. Dahleh for feedback on early versions of some of the ideas presented here, and one of the anonymous reviewers for their comments that helped to streamline the proofs of Proposition 2 and Corollary 1.

REFERENCES


