A Symbolic Approach to the Design of Robust Cyber-Physical Systems

Matthias Rungger and Paulo Tabuada

Abstract—Robustness plays a major role in the analysis and design of engineering systems. Although robustness is reasonably well understood in control theory, the fundamental tenets of robustness in Cyber-Physical Systems (CPSs) remain to be discovered. In this paper we present a design methodology, based on symbolic models, for robust CPSs. We combine existing notions of robustness, based on input-output stability for physical systems, with a recently developed analogue for cyber systems. Our main result states that robustness for CPS can be achieved through a decomposition of concerns: the combination of robustness of the physical system with respect to continuous disturbances and the robustness of the cyber system with respect to discrete disturbances results in a robust CPS.

I. INTRODUCTION

Cyber-Physical Systems (CPSs) are notoriously difficult to analyze and design due to the intricate interaction between the cyber and the physical components. In particular, they are known to be brittle in the sense that “small” software errors can lead to catastrophic consequences. Although feedback control laws can be designed to be robust with respect to several sources of uncertainty, this robustness is lost in the process of integrating feedback controllers in larger CPSs.

In this paper we take the first steps to develop a theory of robustness for CPSs. We build upon the notion of Input-to-State Dynamic Stability (ISDS) introduced by Grune in [5] as a quantitative version of Input-to-State Stability (ISS) [14].

We have shown in previous work [16] that an input-output version of ISDS provides an adequate notion of robustness for cyber components modeled as transducers. We show in this paper how to combine this notion of robustness for cyber components with an input-output version of ISDS for physical systems to obtain a notion of robustness for CPSs. In addition to propose a notion of robustness for CPSs we show how existing analysis and design techniques based on finite abstractions and symbolic models can be suitably modified to ensure robustness. At the technical level, our results hinge upon showing that robustness is preserved by different notions of simulation relating models at different levels of abstraction.

For a newly introduced notion of simulation — approximate contractive simulation — we can decompose the design of robust CPS into the design of continuous controllers catering to continuous disturbances and discrete controllers catering to discrete disturbances.

One of the intuitive ideas behind robustness is that, bounded disturbances lead to bounded deviations of the system behavior from nominal operation. This is a well-known concept for physical systems and has been studied since the beginning of dynamical system theory, see e.g. [19]. This notion appears in many different flavors, such as operator finite gains, bounded input bounded output stability, input-to-state stability, input-output stability, and many more, see e.g. [14]. Moreover, all kinds of different system models like continuous-time systems [14], discrete-time systems [8], sampled-data and networked control systems [18], [11] as well as general hybrid systems [2] have been analyzed under the light of these concepts.

The mathematical tools used to formalize robustness for physical systems include Euclidean distances and norms to quantify the magnitude of disturbances as well as the deviation of the system behaviors from the nominal behavior. This characterization, although being appropriate for physical systems, is not straightforward to extend to cyber systems since norm and magnitude interpretations for input and output signals have been elusive. Therefore, it is not surprising that, in contrast to physical systems, robustness notions for finite state systems have only recently been developed using different approaches to quantify disturbances and system behaviors, see e.g. [17], [10], [3], [4], [1], [9].

In [5], the author introduces a refined notion of robustness that captures a second intuitive idea underlaying robustness, namely that the effect of sporadic disturbances disappears over time. In [16], this notion has been translated to transducers, i.e., cyber system models, using general cost functions to quantify disturbances and system behaviors. Our notion of robustness for CPS combines those two approaches and we provide a detailed comparison of the different definitions in Sections IV-A and IV-B.

One of our main technical results states that robustness is preserved by simulation relations. A similar result using a stronger notion of simulation relation, so called $(\delta, \rho)$-embedding, is established with respect to ISS in [6].

II. PRELIMINARIES

We denote by $\mathbb{N} = \{0, 1, 2, \ldots\}$ the set of natural numbers and by $B_x(r)$ the closed ball centered at $x \in \mathbb{R}^n$ with radius $r \in \mathbb{R}_{\geq 0}$. For $x = 0$, we identify $B_r(x)$ with $B_0(r)$. Given $x \in \mathbb{R}^n$ and $A \subseteq \mathbb{R}^n$, we use $\|x\|_A := \inf_{x' \in A} \|x - x'\|$ to denote the minimum Euclidean distance between $x$ and $A$. For $a, b \in \mathbb{R}$ with $a \leq b$, we denote the closed, open and half-open intervals in $\mathbb{R}$ by $[a, b]$, $(a, b)$, $[a, b]$ and $(a, b]$, respectively.

M. Rungger and P. Tabuada are with the Electrical Engineering Department at the University of California, Los Angeles, CA 90095 USA.

This work was partially supported by the NSF awards xxx and xxx.
While for \( a, b \in \mathbb{Z}, a \leq b \) we use \([a; b], [a; b[\) and \(]a; b[\), to denote the corresponding intervals in \( \mathbb{Z} \).

Given a function \( f : A \to B \) and \( C \subseteq A \), we use \( f(C) \) to denote the set \( \{ f(a) \mid a \in C \} \). A set-valued function or mapping \( f \) from \( X \) to \( Y \) is denoted by \( f : X \rightrightarrows Y \). A set of finite sequences is denoted by \( A^* \). The set of all infinite sequences in \( A^\omega \). Given a relation \( R \subseteq A \times B \) we use \( \pi_A(R) \) to denote its projection onto the set \( A \). We use \( : A \to B \) to denote the inclusion map from \( A \subseteq B \) to \( B \).

We use the usual formalism of comparison functions:

\[
\begin{align*}
K := \{ \alpha : \mathbb{R}_\geq 0 \to \mathbb{R}_\geq 0 \mid & \alpha \text{ is continuous and strictly increasing with } \alpha(0) = 0 \}, \\
L := \{ \alpha : \mathbb{R}_\geq 0 \to \mathbb{R}_\geq 0 \mid & \alpha \text{ is continuous and strictly decreasing with } \lim_{r \to \infty} \alpha(r) = 0 \}, \\
KL := \{ \beta : \mathbb{R}_\geq 0 \times \mathbb{R}_\geq 0 \to \mathbb{R}_\geq 0 \mid & \forall r, s \in \mathbb{R}_\geq 0 : \beta(r, s) \in K, \beta(r, \cdot) \in L \}, \\
KCD := \{ \beta \in KL \mid & \forall r, s, t \in \mathbb{R}_\geq 0 : \mu(r, 0) = r, \mu(r, s + t) = \mu(\mu(r, s), t) \}.
\end{align*}
\]

III. CYBER-PHYSICAL SYSTEMS

In this paper we consider cyber-physical systems composed of three entities: a physical system to be controlled, a controller, and an environment providing the resources needed by the software implementation of the controller. An illustration of the interplay between these components is given in Figure 1. The inputs of the cyber-physical system are the continuous disturbances that represent actuation errors and sensor noise, respectively. Moreover, we use the discrete disturbance as an input signal modeling deviations of the environment from the nominal environment behavior. The environment could model a scheduler that grants access to resources needed by the controller such as sensors and actuators. The environment might influence how the controller interacts with the physical system as well as the controller itself. The output signal is used to specify the desired behavior of the cyber-physical system. The controller has access to the physical system state only through the noisy measurements.

In the following, we will use a general notion of system to model the dynamics of the CPS that consists of the interplay of the physical system, the computational environment and the dynamics of the controller.

IV. ROBUSTNESS FOR CYBER-PHYSICAL SYSTEMS

We now introduce a notion of robustness for CPSs that describes the consequences of the continuous as well as the discrete disturbances. We first define the notion of system that will be used to describe both physical as well as cyber components. Transition systems have already been successfully used in the analysis and design of CPSs, see e.g. [15]. We then introduce the notion of input-output dynamical stability (IODS) for systems. This notion is closely related to input-to-state dynamical stability (ISDS) introduced by Grün in [5] for control systems.

Definition 1. A system \( S \) is a tuple \( S = (X, X_0, U, r, Y, H) \) consisting of

- a set of states \( X \);
- a set of initial states \( X_0 \subseteq X \);
- a set of inputs \( U \);
- a transition map \( r : X \times U \rightrightarrows X \);
- a set of outputs \( Y \);
- an output map \( H : X \to Y \).

An internal behavior of \( S \) is a pair \((\xi, \nu)\) of sequences \( \xi \in X^\omega, \nu \in U^\omega \) such that \( \xi_0 \in X_0 \) and \( \xi_{t+1} = r(\xi_t, \nu_t) \) for all times \( t \in \mathbb{N} \). Through the output map, every internal behavior \((\xi, \nu)\) defines an external behavior \((\zeta, \nu)\) with \( \zeta \in Y^\omega \) given by \( \zeta_t = H(\xi_t) \) for all \( t \in \mathbb{N} \).

In order to avoid technicalities about possible blocking behavior of \( S \), we simply assume throughout this paper that \( r(x, u) \neq \emptyset \) for all \( x \in X \) and \( u \in U \).

We are interested in the behavior of system \( S \) under disturbances, thus we interpret the set of inputs \( U \) as a set of disturbances.

In order to be able to precisely quantify what it means to deviate from the nominal behavior as well as to quantify the disturbance, we introduce input costs and output costs using the functions

\[
I : U^* \to \mathbb{R}_\geq 0 \text{ and } O : Y^* \to \mathbb{R}_\geq 0,
\]

respectively.

The output cost specifies a preference on the system behaviors: less preferred behaviors have higher costs. In particular, the cost should be zero for the nominal behavior and non-

![Fig. 1. Cyber-physical system composed of the physical control system, the controller and the computational environment.](image-url)
zero for the system behavior under disturbances.

**Definition 2.** Let $S$ be a system and let $I$ and $O$ be input and output costs, respectively, for $S$. We say that $S$ is input-output dynamically stable (IODS) w.r.t. $(I, O)$ if there exist $\mu \in KLD$ and $\gamma \in K_\infty$ such that the following inequality holds for all external behaviors of $S$:
\[
O(\xi_{[0:t+1]}|t) \leq \max_{t' \in [0:t]} \mu(\gamma(I(\nu_{[0:t]})), t - t'), \quad \forall t \in \mathbb{N}. \tag{1}
\]

We take the point of view that the system $S$ can be initialized with the correct behavior, i.e., for all $x_0 \in X_0$ we have $O(H(x_0)) = 0$. Therefore, we don’t account for the deviation from the nominal system behavior due to a perturbed initialization.

The notion of robustness embodied by this inequality captures two intuitive ideas underlying robustness: bounded disturbances lead to bounded deviation from the nominal behavior and the effect of a disturbance disappears over time. For continuous systems this is shown in [5] while for discrete systems it is shown in [16].

We believe that IODS is an adequate notion of robustness for CPS for several reasons. It is based on the quantitative version of ISS, called ISDS, and thus leverages the wealth of existing analysis and design results for continuous systems. It has already been applied to cyber systems, modeled by transducers, for which polynomial time algorithms for verification and synthesis are available [16].

**A. Robustness for physical systems**

The reader who is familiar with the notion of ISDS for physical systems [5], [7] can immediately see the similarities between the two definitions. Consider a discrete-time dynamical system:
\[
\xi_{t+1} = f(\xi_t, \omega_t), \tag{2}
\]

here $\omega_t \in \mathbb{R}^m$ is a disturbance input and let $A \subseteq \mathbb{R}^n$ be a compact set. We call (2) ISDS w.r.t. $A$ if there exist $\mu \in KLD$ and $\sigma, \gamma \in K_\infty$ such that for all $\xi : \mathbb{N} \to \mathbb{R}^n$, $\omega : \mathbb{N} \to \mathbb{R}^m$ satisfying (2) we have
\[
|\xi_{t+1}|_A \leq \max_{t' \in [0:t]} \{ \mu(\sigma(|\xi_0|), t+1), \max \{ \gamma(|\omega_{t'}|), t - t' \} \}
\]

for all times $t \in \mathbb{N}$.

We can model (2) as a system $S(f, A)$ by choosing:

$X := \Theta^*; \quad X_0 := \{ \epsilon \}; \quad U := \Theta^*; \quad \theta' \in r(\theta, \theta') \iff \exists s \in \Theta : \theta = \theta s; \quad Y := \Lambda^*; \quad H := f$.

We now note that the transducer $f$ as a system $S(f)$ by defining:

$X := \Theta^*; \quad X_0 := \{ \epsilon \}; \quad U := \Theta^*; \quad \theta' \in r(\theta, \theta') \iff \exists s \in \Theta : \theta = \theta s; \quad Y := \Lambda^*; \quad H := f$.

We can model the transducer $f$ as a system $S(f)$ by defining:

$X := \Theta^*; \quad X_0 := \{ \epsilon \}; \quad U := \Theta^*; \quad \theta' \in r(\theta, \theta') \iff \exists s \in \Theta : \theta = \theta s; \quad Y := \Lambda^*; \quad H := f$.

We establish the results in this section under the assumption that the cost functions depend on the current output and input only, i.e., $I : U \to \mathbb{R}_{\geq 0}$ and $O : Y \to \mathbb{R}_{\geq 0}$.

**B. Robustness for cyber systems**

A notion of robustness for cyber systems modeled by transducers was introduced in [16]. Given a finite input alphabet $\Theta$ and a finite output alphabet $\Lambda$, a transducer is an input-output mapping $f : \Theta^* \to \Lambda^*$ with the property

$\theta' \leq \theta \Rightarrow f(\theta') \leq f(\theta)$

where $\leq$ denotes the prefix order.

The transducer $f$ is input-output stable (IOS) w.r.t. the cost functions $I : \Theta^* \to \mathbb{N}$, $O : \Lambda^* \to \mathbb{N}$ if there exists $\gamma, \eta \in \mathbb{N}$ such that for all $\theta \in \Theta^*$ the inequality
\[
O(f(\theta)) \leq \max_{\theta' \leq \theta} \{ \gamma(I(\theta)), \eta(|\theta|) - |\theta'| \}
\]

holds. The transducer $f$ is called $(\gamma, \eta)$-IOS if the above inequality holds for $\gamma$ and $\eta$.

We can model the transducer $f$ as a system $S(f)$ by defining:

$X := \Theta^*; \quad X_0 := \{ \epsilon \}; \quad U := \Theta^*; \quad \theta' \in r(\theta, \theta') \iff \exists s \in \Theta : \theta = \theta s; \quad Y := \Lambda^*; \quad H := f$.

We establish the results in this section under the assumption that the cost functions depend on the current output and input only, i.e., $I : U \to \mathbb{R}_{\geq 0}$ and $O : Y \to \mathbb{R}_{\geq 0}$.

**V. PRESERVATION OF IODS BY SIMULATION RELATIONS**

In this section we are concerned with the question:

**Under what conditions is IODS preserved by simulation relations?**

We analyze three different relations: exact simulation relations, approximate simulation relations, and contractive simulation relations. It is well-known that simulation relations play an important role in the representation of control systems as systems, so called abstractions or symbolic models. We will see in Section VI how we can use such symbolic models to design robust cyber-physical systems.

We establish the results in this section under the assumption that the cost functions depend on the current output and input only, i.e., $I : U \to \mathbb{R}_{\geq 0}$ and $O : Y \to \mathbb{R}_{\geq 0}$.

**A. Exact simulation relations**

We consider two systems $S_a = (X_a, X_{a0}, U_a, r_a, Y, H_a)$ and $S_b = (X_b, X_{b0}, U_b, r_b, Y, H_b)$ with the cost functions $(I_a, O_a)$ and $(I_b, O_b)$. Note that $S_a$ and $S_b$ have the same output space.

**Definition 3.** Let $S_a$ and $S_b$ be two systems with output space $Y$. A relation $R \subseteq X_a \times X_b$ is said to be a simulation relation (SR) from $S_a$ to $S_b$ if:

1) $(x_a, x_b) \in R$ implies $H_a(x_a) = H_b(x_b)$;
2) \( \forall x_{a0} \in X_{a0} \exists x_{b0} \in X_{b0} \) such that \((x_{a0}, x_{b0}) \in R; \\
3) for all \((x_a, x_b) \in R, x'_a \in X_a \) and \(u_a \in U_a \) with \(x'_a \in r_a(x_a, u_a)\) there is \(x'_b \in X_b\) and \(u_b \in U_b\) with \(x'_b \in r_b(x_b, u_b)\) such that \((x'_a, x'_b) \in R.\)

We call \( R_e \subseteq X_a \times X_b \times U_a \times U_b \) and extended SR from \( S_a \) to \( S_b \) if \( \pi_{X_a \times X_b}(R_e) \) is a SR from \( S_a \) to \( S_b \) and every \((x_a, x_b, u_a, u_b) \in R_e\) satisfies the third item in the above definition, i.e., for all \(x'_a \in r_a(x_a, u_a)\) there is \(x'_b \in r_b(x_b, u_b)\) such that \((x'_a, x'_b) \in \pi_{X_a \times X_b}(R_e).\)

Simulation relations preserve IODS in the following sense.

**Theorem 1.** Let \( S_a \) and \( S_b \) be two systems with output space \( Y \) and let \((I_a, O_a)\) and \((I_b, O_b)\) be input and output costs for \( S_a \) and \( S_b \), respectively. Suppose there exists an extended SR \( R_e \) from \( S_a \) to \( S_b \) and for all \((x_a, x_b, u_a, u_b) \in R_e\) the inequalities

\[
O_a \circ H_a(x_a) \leq O_b \circ H_b(x_b) \quad \text{and} \quad I_b(u_b) \leq I_a(u_a)
\]

hold. If \( S_a \) is IODS w.r.t. \((I_a, O_a)\) then \( S_a \) is IODS w.r.t. \((I_a, O_a)\).

Note how preservation of IODS is contra-variant, i.e., while the direction of the simulation relation is from system \( S_a \) to system \( S_b \), the propagation of IODS is from system \( S_b \) to system \( S_a \).

**B. Approximate simulation relations**

Exact simulation relations are often too restrictive when one seeks to relate a physical system to a finite-state abstraction or symbolic model. In this case, approximate simulation relations were shown to be adequate in the sense that they can be shown to exist for large classes of physical systems [15].

We consider again two systems \( S_a \) and \( S_b \) with \( Y = Y_a = Y_b \) and assume that the output space \( Y \) is a metric space with the metric \( d_Y \).

**Definition 4.** Let \( S_a \) and \( S_b \) be two systems with output space \( Y \) that is a metric space with metric \( d_Y \) and let \( \varepsilon \in \mathbb{R}_{\geq 0} \). A relation \( R \subseteq X_a \times X_b \) is said to be an \( \varepsilon \)-approximate simulation relation \((\varepsilon-\text{aSR})\) from \( S_a \) to \( S_b \) if:

1) \((x_a, x_b) \in R \) implies \( d_Y(H_a(x_a), H_b(x_b)) \leq \varepsilon; \\
2) \forall x_{a0} \in X_{a0} \exists x_{b0} \in X_{b0} \) such that \((x_{a0}, x_{b0}) \in R; \\
3) \forall (x_a, x_b) \in R, \forall x'_a \in r_a(x_a, u_a), \exists x'_b \in r_b(x_b, u_b) \) such that \((x'_a, x'_b) \in R.\)

We call \( R_e \subseteq X_a \times X_b \times U_a \times U_b \) and extended \( \varepsilon \)-aSR from \( S_a \) to \( S_b \) if \( \pi_{X_a \times X_b}(R_e) \) is an \( \varepsilon \)-aSR from \( S_a \) to \( S_b \) and every \((x_a, x_b, u_a, u_b) \in R_e\) satisfies: for all \(x'_a \in r_a(x_a, u_a)\) there is \(x'_b \in r_b(x_b, u_b)\) such that \((x'_a, x'_b) \in \pi_{X_a \times X_b}(R_e).\)

The analogue of Theorem 1 for approximate simulation relations reads as follows.

**Theorem 2.** Let \( S_a \) and \( S_b \) be two systems with a metric output space \( Y \) and let \((I_a, O_a)\) and \((I_b, O_b)\) be input and output costs for \( S_a \) and \( S_b \), respectively. Suppose there exists an extended \( \varepsilon \)-aSR \( R_e \) from \( S_a \) to \( S_b \) and for all \((x_a, x_b, u_a, u_b) \in R_e\) the following inequalities hold

\[
O_a \circ H_a(x_a) \leq O_b \circ H_b(x_b) + \varepsilon_O \\
I_b(u_b) \leq I_a(u_a) + \varepsilon_I
\]

for some constants \( \varepsilon_O, \varepsilon_I \in \mathbb{R}_{\geq 0} \). If \( S_b \) is IODS w.r.t. \((I_b, O_b)\), then there exist \( \mu_a \in \mathbb{K} \) and \( \gamma_a \in \mathbb{K}_\infty \) such that all external behaviors \((\zeta_a, \nu_a)\) of \( S_a \) satisfy

\[
O_a(\zeta_{a,t+1}) \leq \max_{t' \in [0,t]} (\mu_a(\gamma_a(I_a(\nu_a,t'))), t - t') + \varepsilon_O, \quad \forall t \in \mathbb{N}.
\]

The passage from simulation relations to approximate simulation relation involved trading the requirement that related states have the same outputs by the weaker requirement of having close enough outputs. In the same spirit we relaxed the assumptions (4) on the costs compared to (3), by introducing the constants \( \varepsilon_I \) and \( \varepsilon_O \). This will be useful in the application of the results to CPSs. Moreover, note that (5) represents a “practical” version of the IODS inequality (1).

**C. Contractive Simulation Relations**

The construction of abstractions or symbolic models for physical systems described in [12], [13], [15] results in simulation relations that satisfy a certain contractivity property. In this section we show that such property leads to a strengthened preservation of IODS.

We consider two systems \( S_a \) and \( S_b \) with \( Y = Y_a = Y_b \) and assume that the output space \( Y \) is a metric space with metrics \( d_Y \) and \( d_U \), respectively. We call two systems \( S_a \) and \( S_b \) that satisfy this assumptions compatible metric systems.

**Definition 5.** Let \( S_a \) and \( S_b \) be two compatible metric systems and let \( \kappa, \lambda \in \mathbb{R}_{\geq 0} \) and \( \beta \in [0,1] \) be parameters. We call \( R \subseteq X_a \times X_b \) an \( \kappa \)-approximate \((\beta, \lambda)\)-contractive simulation relation \((\kappa, \beta, \lambda)\)-acSR from \( S_a \) to \( S_b \) if:

1) \( \forall x_{a0} \in X_{a0} \exists x_{b0} \in X_{b0} \) such that \((x_{a0}, x_{b0}) \in R \) and \( d_Y(H_a(x_{a0}), H_b(x_{b0})) \leq \kappa; \\
2) \forall (x_a, x_b) \in R \) and \( u_a \in U_a \) there is \( u_b \in U_b \) such that for all \( x'_a \in r_a(x_a, u_a) \) there is \( x'_b \in r_b(x_b, u_b) \) such that

\[
a) \quad (x'_a, x'_b) \in R \quad \text{and} \quad b) \quad d_Y(y'_a, y'_b) \leq \kappa + \beta d_Y(y_a, y_b) + \lambda d_U(u_a, u_b) \quad \text{with} \quad y_a = H_a(x_a), y'_a = H_a(x'_a), i \in \{a, b\}.\]

We call \( R_e \subseteq X_a \times X_b \times U_a \times U_b \) and extended \((\kappa, \beta, \gamma)\)-acSR from \( S_a \) to \( S_b \) if \( \pi_{X_a \times X_b}(R_e) \) is an \((\kappa, \beta, \lambda)\)-acSR from \( S_a \) to \( S_b \) and for every \((x_a, x_b, u_a, u_b) \in R_e\) and \( x'_a \in r_a(x_a, u_a) \) there is \( x'_b \in r_b(x_b, u_b) \) such that \((x'_a, x'_b) \) satisfy 2.a) and 2.b) in the above definition.

**Remark 1.** Note that an extended \((\kappa, \beta, \lambda)\)-acSR \( R_e \) from \( S_a \) to \( S_b \) is also an extended \( \varepsilon \)-aSR from \( S_a \) to \( S_b \).
$S_b$ for any $\varepsilon \in \mathbb{R}_{\geq 0}$ that satisfies $\varepsilon \geq \beta \varepsilon + \kappa + \lambda \Delta u$ where $\Delta u$ is the maximal distance of related inputs $\Delta u := \sup_{(x_a,x_b,u_a,u_b)} \in \mathcal{R}_e \ d_0(u_a,u_b)$. However, with the inequality in 2.b) we have a defined description of how the approximation accuracy depends on $\kappa$ and the difference of related inputs.

We have the following lemma for the internal behaviors of two related systems.

**Lemma 1.** Let $S_a$ and $S_b$ be two compatible metric systems and let $R_e$ be an extended $(\kappa, \beta, \lambda)$-acSR from $S_a$ to $S_b$. Then there exist $\mu_\Delta \in \mathcal{KLD}$ and $\gamma_\Delta, \kappa_\Delta \in \mathbb{R}_{\geq 0}$ such that every pair of internal behaviors $(\xi_a, \nu_a)$ of $S_a$ and $(\xi_b, \nu_b)$ of $S_b$ with $(\xi_a,t,\xi_b,t,\nu_a,t,\nu_b,t) \in R_e, \forall t \in \mathbb{N}$ and $d_I(H_a(\xi_a,0),H_b(\xi_b,0)) \leq \kappa$, satisfies the inequality

$$\Delta \zeta_{t+1} \leq \kappa_\Delta + \max_{\varepsilon \in [0,\varepsilon]} \mu_\Delta(\gamma_\Delta \Delta \nu_{t'}, t-t') \quad \forall t \in \mathbb{N}$$

with $\Delta \zeta_t := d_I(H_a(\xi_a,t), H_b(\xi_b,t))$ and $\Delta I_t := d_0(\nu_a,t, \nu_b,t).$

Preservation of IODS under contraceptive approximate simulation relations takes the following form.

**Theorem 3.** Let $S_a$ and $S_b$ be two compatible metric systems and let $(I_a,O_a)$ and $(I_b,O_b)$ be input and output costs for $S_a$ and $S_b$, respectively. Suppose there exists an extended $(\kappa, \beta, \lambda)$-acSR $R$ from $S_a$ to $S_b$ and for all $(y_a,y_b,u_a,u_b)$ with $y_a = H_a(x_a)$, $y_b = H_b(x_b)$ and $(x_a,x_b,u_a,u_b) \in R_e$ the inequalities

$$O_a(y_a) \leq O_b(y_b) + \varepsilon_0(d_I(y_a,y_b))$$

$$I_b(u_b) \leq I_a(u_a) + \varepsilon_1(d_I(u_a),u_b))$$

hold for some monotonically increasing functions $\varepsilon_1, \varepsilon_0 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$.

If $S_b$ is IODS w.r.t. $(I_a,O_a)$, then there exist $\mu_\Delta, \mu_\Delta \in \mathcal{KLD}$, $\gamma_\Delta, \kappa, \kappa_\Delta \in \mathbb{R}_{\geq 0}$ such that all external behaviors $(\xi_a, \nu_a)$ of $S_a$ satisfy

$$O_a(\xi_a) \leq \max_{t \in [0,t]} \mu_\Delta(\gamma_\Delta \Delta \nu_{t'}), t-t') + \varepsilon_0'$$

for all $t \in \mathbb{N}$ with $\Delta \nu_t \geq d_0(\nu_a,t,\nu_b,t)$ and $\varepsilon_0'(r) := \varepsilon_0(2r)$.

**VI. APPLICATION TO CYBER-PHYSICAL SYSTEMS**

In this section, we apply the previous results to verify IODS of a CPS by resorting to a symbolic model. In particular, we show that the newly introduced notion of $(\kappa, \beta, \lambda)$-acSR allows us to establish IODS-type inequalities taking into account discrete as well as continuous disturbances, even though we do not explicitly account for the continuous disturbances on the symbolic model.

Note that we do not provide an algorithm to compute the symbolic model $S_M$ along with an $(\kappa, \beta, \lambda)$-acSR from the system $S_\Sigma$ modeling a CPS to $S_M$ but simply assume that we possess such a model and relation. However, we refer the reader to the work described in [12], [13] and [15, Chapter 11] where such algorithms are provided for incrementally stable systems.

Let $S_b$ be a system that models the CPS with output space $Y \subseteq \mathbb{R}^p$ and input space $U_e = W \times D$, where $W = W_a \times W_s$ are the continuous disturbances and $D$ is a finite set of discrete disturbances. We assume that the symbolic model $S_M$ has the same output space $Y$ and input space $U_M = D$. We define the output metric $d_I(y,y') := |y-y'|$ by the Euclidean norm and the input metric by $d_0((w,d),(w',d')) := |w-w'| + d_D(d,d')$

where $d_0$ denotes some metric on $D$. Note that $D$ is isometric to $(0,0) \times D$ and we see that $S_M$ and $S_\Sigma$ are compatible metric systems.

Let $R_e \subseteq X_\Sigma \times X_M \times (W \times D) \times D$ be an extended $(\kappa, \beta, \lambda)$-acSR from $S_\Sigma$ to $S_M$. We assume that we have:

$$(x_\Sigma, x_M, (w,d), d) \in R_e$$

for every $(x_\Sigma, x_M) \in \pi_{X_\Sigma \times X_M}(R_e)$ and $(w,d) \in W \times D$.

This assumption is made without loss of generality since the abstraction based techniques described in [12], [13] and [15, Chapter 11] can be modified to ensure (9). For reasons of space, rather than providing a proof of this fact we provide a characteristic example illustrating that although an abstraction is computed assuming the absence of disturbances, (9) is guaranteed to hold. We consider a system without control inputs. However, the extension to the general case is conceptually clear.

**Example 1.** Let us consider the perturbed system $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ with $W \subseteq \mathbb{R}^n$. Suppose the system is contractive, i.e., for all $x \in \mathbb{R}^n$ and $w \in W$ we have

$$|f(x, w) - f(x', w')| \leq |x-x'| + \lambda |w-w'|$$

with $\beta \in [0,1]$ and $\lambda \in \mathbb{R}_{\geq 0}$. We define the system $S_y$ corresponding to $f$ by $X = X_0 := \mathbb{R}^n; U := W; r(x,w) := \{f(x,w)\}; Y = \mathbb{R}^n$ and $H = \mathbb{I}$.

Let $[\mathbb{R}^n]_2 := \{x \in \mathbb{R}^n \mid \exists k \in \mathbb{Z}^n : x = 2k\eta\}$ denote a uniform grid in $\mathbb{R}^n$. We construct a symbolic model $S_M$ for $S_y$ as follows: $X_M = X_{M_0} := [\mathbb{R}^n]_2; U_M := \{0\}, x' \in r(x_0,0) := \{x' - f(x,0)\} \leq \eta; Y_M = \mathbb{R}^n$ and $H_M = \mathbb{I}$.

It is straightforward to see that $S$ and $S_M$ are two compatible metric systems with metrics $d_I(y,y') := |y-y'|$ and $d_0 = |u-u'|$. Given any bound $\alpha \in \mathbb{R}_{\geq 0}$ on the disturbance $W \subseteq B(\alpha)$ we chose $\varepsilon \in \mathbb{R}_{\geq 0}$ with $\beta \varepsilon + \eta + \lambda \alpha \leq \varepsilon$ and define the relation $R := \{x(x, x_M) \in X \times X_M \mid |x-x_M| \leq \varepsilon\}$.

Now it is easy to verify using the triangle inequality that $R_e := R \times X \times \{0\}$ is an $(\eta, \beta, \lambda)$-acSR from $S$ so $S_M$.

We proceed with the derivation of the cost functions $(I_M, O_M)$ for $S_M$ from the cost functions $(I_\Sigma, O_\Sigma)$. Here, we assume a certain structure on the input cost $I_\Sigma$ that represents the compositional nature of the CPS

$$I_\Sigma(w, d) = |w| + I_D(d).$$
for some cost $I_D : D \to \mathbb{R}_{\geq 0}$ on the discrete disturbances. Since there are no continuous disturbances on the symbolic model, we choose the input costs to be $I_M(d) := I_D(d)$ and it is easy to see that $I_M(d) \leq I_S(w, d)$ holds for all $(w, d)$. Let $Y_M$ denote the set of symbolic outputs. We define the output cost $O^*_M : Y_M \to \mathbb{R}_{\geq 0}$ for $S_M$ as a function of the parameter $\kappa$ of the $(\kappa, \beta, \lambda)$-acSR by

$$O^*_M(y_M) := \inf O_S(B_{y_M}(\kappa)).$$

(11)

The particular choice of output costs $O^*_M$ is motivated by the observation, that it is easier to establish the IODS inequality on $S_M$ for lower output costs.

We define $\varepsilon'_O : Y_M \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by the maximal difference of the output cost in the ball $B_{y_M}(d)$ centered at the abstract output $y_M$ with radius $d$:

$$\varepsilon'_O(y_M, d) = \sup O_S(B_{y_M}(d)) - \inf O_S(B_{y_M}(d))$$

If we define $\Delta y := \partial y(y_M, y_M')$ we see that the inequality

$$O_S(y_M') \leq O^*_M(y_M) + \varepsilon'_O(y_M, \max(\kappa, \Delta y))$$

$$\leq O^*_M(y_M) + \varepsilon_O(y_M, \Delta y) + \varepsilon_O(y_M, \kappa)$$

holds for arbitrary $y_M' \in Y$ and $y_M \in Y_M$. In order to render this inequality independent of $y_M$ we define

$$\varepsilon_O(d) := \sup \varepsilon'_O(Y_M, d)$$

(12)

and obtain $O_S(y_M') \leq O^*_M(y_M) + \varepsilon_O(\Delta y) + \varepsilon_O(\kappa \Delta)$. In the following result, we assume that we were able to successfully verify that the symbolic model $S_M$ is $(\gamma, \eta)$-IOS and we derive an IODS type inequality that holds for the CPS $S_C$. Polynomial time algorithms, for the verification of the IOS inequality for finite systems are given in [16].

**Theorem 4.** Let $S_C$ and $S_M$ be two compatible metric systems and let $(I_S, O_S)$ on $S_C$ and $(I_M, O_M)$ be the input and output costs of $S_M$, where $I_S$ satisfies (10). Let $R_C$ be an $(\kappa, \beta, \lambda)$-acSR from $S_C$ to $S_M$ that satisfies (9). If $S_M$ is IODS wrt. $(I_M, O^*_M)$ with $KLD$ function $\mu_I(r, \gamma) = \max\{0, r - \eta\}$ and constant $\gamma \in \mathbb{R}_{\geq 0}$, then all external behaviors $(\Sigma_C, \omega_C, \delta_C)$ of $S_C$ satisfy

$$O_S(\Sigma_C, t+1) \leq \max_{t' \in [t, t]} \{\gamma I_D(\delta_C, t') - \eta(t - t')\} + \max_{t' \in [t, t]} \varepsilon_O \circ 2 \mu_D(\gamma, |\omega_C, t', t - t'|) + \varepsilon_O(2\kappa \Delta)$$

where $\varepsilon_O$ is given in (12) and $\mu_D \in KLD$ and $\gamma, \kappa, \kappa \in \mathbb{R}_{\geq 0}$ are chosen as in Lemma 1.

The first term in the right hand side of (13) accounts for the perturbed system behavior caused by the discrete disturbances. It follows from the fact that we were able to successfully verify IODS for the symbolic model $S_M$. The second term accounts for the perturbed system behavior due to the continuous disturbances. This term results from the contractivity of the acSR from $S_C$ to $S_M$. The third term, i.e., the constant offset, is a result of the approximate representation of the CPS by the symbolic model and corresponds to the accuracy of $S_M$. Hence, inequality (13) shows how robustness can be guaranteed for the CPS by designing a continuous controller handling the continuous disturbances and a discrete controller handling the discrete disturbances.

**VII. DISCUSSION**

In this paper we proposed IODS of systems as a notion of robustness for CPS. IODS is based on ISDS and thus captures robustness for physical systems. Moreover, a variant of IODS for cyber systems modeled by transducers is shown in [16] to capture robustness for cyber systems. IODS is preserved by different notions of simulation. In particular, it is preserved by contractive simulations, a notion introduced in this paper, arising from the construction of discrete abstractions or symbolic models using the techniques in [12], [13] and [15, Chapter 11]. We use this fact to conclude that the design of CPS using discrete abstractions, can be made robust with respect to both discrete and continuous disturbances by designing a continuous controller focusing on continuous disturbances and a discrete controller focusing on discrete disturbances.

**REFERENCES**


