Abstract—A general notion of bisimulation is defined for linear input-state-output systems, using analogies with the theory of concurrent processes. A characterization of bisimulation and an algorithm for computing the maximal bisimulation relation is derived using geometric control theory. Bisimulation is shown to be a notion which unifies the concepts of state-space equivalence and state-space reduction, and which allows to study equivalence of systems with nonminimal state-space dimension. The notion of bisimulation is especially powerful for “nondeterministic” dynamical systems, and leads in this case to a notion of equivalence which is finer than equality of external behavior. For abstractions of systems it is shown how the results specialize to previously obtained results by other authors. Extensions of the main results to the nonlinear case are provided.

Index Terms—Abstraction, bisimulation, controlled invariance, maximal bisimulation relation, nondeterministic systems, non-linear systems, nonminimal systems, state–space equivalence.

I. INTRODUCTION

A COMMON theme in theoretical computer science (in particular, the theory of distributed processes and computer-aided verification) and in systems and control theory is to characterize systems which are “externally equivalent.” The intuitive idea is that we only want to distinguish between two systems if the distinction can be detected by an external system interacting with these systems. This is a fundamental notion in design, enabling us to take a “divide and rule” strategy, and in analysis, allowing us to switch between externally equivalent representations of the same system and to reduce subsystems to externally equivalent but simpler subsystems.

More specifically, a main issue in the theory of distributed processes and computer-aided verification is to develop a mathematical framework that can handle the state explosion in complex systems. A crucial notion in this endeavor is the concept of bisimulation which expresses when a subprocess can be considered to be externally equivalent to another (hopefully simpler) process. On the other hand, classical notions in systems and control theory are state–space equivalence of dynamical systems, and reduction of a dynamical system to an equivalent system with minimal state–space dimension. These notions have been instrumental in, e.g., linking input–output models to state–space models, and in studying the properties of interconnected systems.

Developments in both areas have been rather independent (“concurrently but without communication”), one of the reasons being that the mathematical formalisms for describing both types of systems (discrete processes on the one hand, and continuous dynamical systems on the other hand) are rather different. However, with the rise of interest in hybrid systems, which are systems with interacting discrete and continuous dynamics, and with the renewed interest in complexity in systems theory, this situation seems not very productive, as has been argued before by several authors.

The aim of this paper is to make another step in this reap-proachment between the theory of concurrent processes and mathematical systems theory by defining and studying in detail a notion of bisimulation for continuous dynamical systems, and to relate it to the common system-theoretic notions of state–space equivalence and state–space reduction.

Such an extension of the notion of bisimulation to continuous dynamical systems has been explored before in a series of innovative papers by Pappas et al. [13], [14], [2], [18]–[20], [9], [29], and this paper is very much inspired by this work. The main difference is that in [18]–[20], and [29], the focus is on characterizing bisimulation of a dynamical system by a “projected” dynamical system with lower state–space dimension (“an abstraction;” see Section V), while in this paper, we deal with a general notion of bisimulation between two continuous dynamical systems and we elaborate on the relations with the system-theoretic notions of state–space equivalence and state–space reduction. (We note that a general definition of bisimulation relations for dynamical systems has been given before in [9].)

In the main part of this paper, we study continuous dynamical systems of the form

$$\begin{align*}
\dot{x} &= Ax + Bu + Gd, & x &\in \mathbb{R}^n, & u &\in \mathbb{R}^m, & d &\in \mathbb{R}^s \\
y &= Cx, & y &\in \mathbb{R}^p
\end{align*}$$

(1)

where $x$ are as usual the state variables, $u$ are the input variables, $y$ are the output variables, while $d$ are additional input variables, which can be thought of as disturbances. The basic problem we address is when two systems of the form (1) can be considered to be externally equivalent, in the sense that for all time instants $t_0$ the solution trajectories for $t \geq t_0$ of one system are mimicked by the other in such a way that the input and output trajectories $u(t)$ and $y(t)$ of both systems are the same for $t \geq t_0$, without imposing any relation between the values of their disturbance variables $d$.

In case the disturbance variables $d$ are absent this problem comes down to the usual system-theoretic notion of state–space equivalence, under the additional assumption that the systems have minimal state–space dimension; see, e.g., [5] and [6]. At the same time, the problem bears much similarity with the notion of bisimulation as introduced by Milner and Park [15], [21]...
in the study of concurrent processes; especially in the presence of disturbance inputs \(d\). Finally, the notion is closely linked to the notion of controlled invariance, as introduced in linear systems theory by Wonham and Morse [31] and Basile and Marro [4]. The last two connections were already explored by Pappas et al. [19], [18], [20], [29] in the more restricted context when a system is bisimilar to an “abstraction” of itself. (Note furthermore that in [18]–[20] and [29], the input term \(Bu\) plays the same role as the disturbance term \(Gd\) in our setting, and thus bisimulation only involves equality of the outputs \(y = Cx\).)

We will make explicit how our results generalize part of the results of [19], [18], [20], [29], and [9] in Section V.

The reason why the aforementioned notion of equivalence is similar to the notion of bisimulation as used in computer science can be briefly explained as follows. The notion of bisimulation has been introduced in order to handle external equivalence of (interactive) concurrent processes. The concept is especially powerful for concurrent processes which are nondeterministic in the sense that branching in the (discrete) state may occur while the traces ("words") generated by the transition system are the same. (Thus, "nondeterministic" is not meant in a stochastic sense but in the sense that the current state together with the label of the transition does not uniquely determine the next state.) In fact, the existence of a bisimulation relation between two deterministic processes is equivalent to equality of their external behaviors (the set of "traces," or the "language" generated by the process), and in this case bisimulation provides an efficient way to check equality of external behavior. For nondeterministic processes, however, bisimulation provides a finer equivalence than equality of external behavior, and, for example, also captures the deadlock behavior of concurrent processes. A similar picture appears to arise for bisimulation of continuous dynamical systems (1). First, a type of "nondeterminism" is present in (1) if we consider \(u\) and \(y\) as the external variables of the system (analogously to the labels of the discrete transitions of a process), while \(d\) denotes a generator for nondeterminism in the evolution of the state \(x\). If \(d\) is absent then (1) reduces to an ordinary "deterministic" system. The interpretation of \(d\) as an internal generator for nondeterminism may be stressed (but for simplicity of notation we will not do so) by suppressing the appearance of the variable \(d\) by rewriting (1) equivalently as the differential inclusion

\[
\begin{align*}
\dot{x} &= Ax - Bu \\
y &= Cx,
\end{align*}
\]

One of the results established in this paper is that for deterministic systems (1) (\(d\) absent) bisimulation is equivalent to equality of external behavior, while generalizing the usual notion of state–space equivalence by allowing for systems with nonminimal state–space dimension. For nondeterministic systems, however, bisimulation will be a stronger (finer) type of equivalence than equality of external behavior.

The structure of the paper is as follows. In Section II, a linear–algebraic characterization of bisimulation is given, based on geometric control theory. The maximal bisimulation relation is computed in Section III. Then, Section IV deals with reduction of dynamical systems using the notion of a bisimulation relation between the system and itself. Simulation and abstraction is treated in Section V, making explicit the relation with previously obtained results in [19]. In Section VI, the theory is specialized to deterministic systems recovering classical results on state–space equivalence and reduction. Although most of the paper is devoted to linear systems (1) we briefly describe in Section VII how the theory developed for linear dynamical systems can be generalized to a nonlinear context by using tools from nonlinear geometric control theory; especially the notions of a controlled invariant submanifold and a controlled invariant distribution or foliation. Finally, Section VIII contains the conclusions and questions for further research.

II. BISIMILAR LINEAR DYNAMICAL SYSTEMS

Consider two dynamical systems of the form (1), as shown in (3) at the bottom of the page, with \(\mathcal{X}_i, \mathcal{D}_i, \mathcal{U}, \mathcal{Y}\) finite-dimensional linear spaces (over \(\mathbb{R}\)). Before defining bisimulation, we need to specify the solution trajectories of the systems (the "semantics"). That is, we have to specify the function classes of admissible input functions \(u : [0, \infty) \rightarrow \mathcal{U}\) and admissible disturbance functions \(d : [0, \infty) \rightarrow \mathcal{D}\), together with compatible function classes of state and output solutions \(x : [0, \infty) \rightarrow \mathcal{X}\) and \(y : [0, \infty) \rightarrow \mathcal{Y}\). For compactness of notation we will usually denote these time–functions, respectively, by \(u(\cdot), d(\cdot), x(\cdot)\) and \(y(\cdot)\). The exact class from which the functions are chosen is for purposes of this paper not really important, as long as the state trajectories \(x(\cdot)\) are at least absolutely continuous. For example, we can take all functions to be \(C^\infty\), or piecewise \(C^\infty\) with the additional requirement that \(x(\cdot)\) is continuous.

**Definition 2.1.** A bisimulation relation between two linear systems \(\Sigma_1\) and \(\Sigma_2\) is a linear subspace

\[\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2\]

with the following property. Take any \((x_{10}, x_{20}) \in \mathcal{R}\) and any joint input function \(u_1(\cdot) = u_2(\cdot)\). Then, for every disturbance function \(d_1(\cdot)\) there should exist a disturbance function \(d_2(\cdot)\) such that the resulting state solution trajectories \(x_1(\cdot)\), with \(x_1(0) = x_{10}\), and \(x_2(\cdot)\), with \(x_2(0) = x_{20}\), satisfy

\[\begin{align*}
&i) \quad (x_1(t), x_2(t)) \in \mathcal{R}, \quad \text{for all } t \geq 0 \quad (4) \\
&ii) \quad C_1 x_1(t) = C_2 x_2(t), \quad \text{for all } t \geq 0 \quad (5)
\end{align*}\]

\[
\Sigma_i : \begin{cases} 
\dot{x}_i = A_i x_i + B_i u_i + G_i d_i, \\
y_i = C_i x_i, \\
x_i \in \mathcal{X}_i, \\
y_i \in \mathcal{Y}, \\
u_i \in \mathcal{U}, \\
d_i \in \mathcal{D}_i 
\end{cases}, \quad i = 1, 2
\]
(or, more precisely, for all \( t \geq 0 \) for which the trajectories are defined), and conversely, for every disturbance function \( d_2(\cdot) \) there should exist a disturbance function \( d_1(\cdot) \) such that again the resulting state trajectories \( x_1(\cdot) \) and \( x_2(\cdot) \) satisfy (4) and (5).

Hence, for every pair \( (x_{10}, x_{20}) \in \mathcal{R} \) all possible trajectories \( x_1(\cdot) \) with \( x_1(0) = x_{10} \) can be “simulated” by a trajectory \( x_2(\cdot) \) with \( x_2(0) = x_{20} \) and conversely all possible trajectories \( x_2(\cdot) \) with \( x_2(0) = x_{20} \) can be “simulated” by a trajectory \( x_1(\cdot) \) with \( x_1(0) = x_{10} \), in the sense of giving the same input-output data for all future times while \( (x_1(t), x_2(t)) \in \mathcal{R} \) for all \( t \geq 0 \).

**Remark 2.2:** For linear systems, it is natural to restrict to bisimulation relations \( \mathcal{R} \) which are linear subspaces. This can be done without much loss of generality. Indeed, the maximal bisimulation relation as produced by Algorithm 3.3 is a linear subspace. Furthermore, if a general subset \( \mathcal{R} \) satisfies the conditions of Definition 2.1, then so does its linear closure.

**Remark 2.3:** A similar definition (in the nonlinear case, cf. Section VII, and for the case that \( u_i \) is absent) was given in [9].

**Remark 2.4:** Note that the existence of a bisimulation relation between systems defines an equivalence relation between systems. Clearly, \( \mathcal{R}_{11} := \{(x_1, x_1), \forall x_1 \in \mathcal{X}_1 \} \) is a bisimulation relation between \( \mathcal{X}_1 \) given by (3) and (5). Furthermore, the existence of a bisimulation relation between \( \mathcal{X}_2 \) and \( \mathcal{X}_2 \) is symmetric. Finally, if \( \mathcal{R}_{12} \subset \mathcal{X}_1 \times \mathcal{X}_2 \) is a bisimulation relation between \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) and \( \mathcal{R}_{22} \subset \mathcal{X}_2 \times \mathcal{X}_2 \) between \( \mathcal{X}_2 \) and \( \mathcal{X}_2 \) then \( \mathcal{R}_{13} := \{(x_1, x_3), \forall x_3 \in \mathcal{X}_3 \} \) is a bisimulation relation between \( \mathcal{X}_1 \) and \( \mathcal{X}_3 \).

**Definition 2.5:** Two systems \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) as in (3) are bisimilar, denoted \( \mathcal{X}_1 \equiv \mathcal{X}_2 \), if there exists a bisimulation relation \( \mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2 \) with the property that

\[
\pi_1(\mathcal{R}) = \mathcal{X}_1 \quad \text{and} \quad \pi_2(\mathcal{R}) = \mathcal{X}_2
\]

where \( \pi_i : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{X}_i, i = 1, 2 \), denote the canonical projections.

**Remark 2.6:** Actually, Definition 2.5 constitutes a slight departure from the definition of bisimulation relation as usually given for discrete processes [15], [21], by imposing the extra requirement (6). The reason is that in computer science discrete processes are usually defined with respect to a fixed initial state (or, a subset of initial conditions). In our setting we consider the behavior of the systems \( \mathcal{X}_i \) for arbitrary initial states. Hence, for every initial condition \( x_{10} \) of \( \mathcal{X}_1 \) there should exist an initial condition \( x_{20} \) of \( \mathcal{X}_2 \) with \( (x_{10}, x_{20}) \in \mathcal{R} \) and vice versa; thus implying (6). The generalization to subsets of initial conditions \( \mathcal{X}_{10} \subset \mathcal{X}_1 \) obviously can be done by relaxing (6) to

\[
\pi_i(\mathcal{R}) = \mathcal{X}_{10}, \quad i = 1, 2
\]

**Remark 2.7:** Clearly (cf. Remark 2.4), the relation \( \equiv \) is an equivalence relation.

**Remark 2.8:** For \( G_1 = G_2 = 0 \), the aforementioned notion of bisimilarity is close to the usual notion of state-space equivalence of two input-state-output systems (cf. [5] and [6])

\[
\begin{align*}
\Sigma_i: & \quad \dot{x}_i = A_i x_i + B_i u_i, \quad x_i \in \mathcal{X}_i, \\
& \quad \forall u_i \in \mathcal{U}, \quad y_i \in \mathcal{Y}, \quad i = 1, 2
\end{align*}
\]

Indeed, in this case one usually starts with a linear equivalence mapping

\[
S : \mathcal{X}_1 \rightarrow \mathcal{X}_2
\]

which is assumed to be invertible (implying that \( \dim \mathcal{X}_1 = \dim \mathcal{X}_2 \) with the property that

\[
S x_1(t) = x_2(t), \quad \text{for all } t \geq 0
\]

\[
C_1 x_1(t) = C_2 x_2(t), \quad \text{for all } t \geq 0
\]

for all state trajectories \( x_1(\cdot) \) and \( x_2(\cdot) \) resulting from initial conditions \( x_{10} \) and \( x_{20} \) related by \( S x_{10} = x_{20} \) and all input-functions \( u_1(\cdot) = u_2(\cdot) \). Defining the linear subspace

\[
\mathcal{R} = \{(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2 : x_2 = S x_1 \} \]

(i.e., the graph of the mapping \( S \)) it is easily seen that \( \mathcal{R} \) is a bisimulation relation which satisfies \( \pi_1(\mathcal{R}) = \mathcal{X}_1 \) trivially and \( \pi_2(\mathcal{R}) = \mathcal{X}_2 \) because of invertibility of \( S \).

Clearly, by allowing \( \mathcal{R} \) to be a relation instead of the graph of a mapping, the notion of bisimilarity even in the case \( G_1 = G_2 = 0 \) is more general than state-space equivalence. In particular, we may allow \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) to be of different dimension.

Furthermore, by doing so we incorporate in the notion of bisimilarity the notion of reduction of an input-state-output system to a lower dimensional input-state-output system, and especially the reduction to a minimal input-state-output system, as we will see later on in Section VI.

Using ideas from the theory of controlled invariance (or \( (A, B) \)-invariance) (see, e.g., [31] and [4]), it is relatively straightforward to derive an algebraic characterization of the notion of a bisimulation relation. We do this in two steps.

**Proposition 2.9:** A subspace \( \mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2 \) is a bisimulation relation between \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) if and only if for all \( (x_1, x_2) \in \mathcal{R} \) and all \( u \in \mathcal{U} \) the following properties hold.

i) For all \( d_1 \in \mathcal{D}_1 \) there should exist a \( d_2 \in \mathcal{D}_2 \) such that

\[
(A_1 x_1 + B_1 u + G_1 d_1, A_2 x_2 + B_2 u + G_2 d_2) \in \mathcal{R}
\]

and, conversely, for every \( d_2 \in \mathcal{D}_2 \) there should exist a \( d_1 \in \mathcal{D}_1 \) such that (12) holds.

ii) \( C_1 x_1 = C_2 d_2. \)

**Proof:** (Sketch) Consider (4). Then, by differentiating \( x_1(t) \) and \( x_2(t) \) with respect to \( t \) and evaluating at any \( t \) we obtain (12), with \( x_1 = x_1(t), x_2 = x_2(t), u = u_1(t) = u_2(t), d_1 = d_1(t), d_2 = d_2(t) \). Conversely, if (12) holds then \( (x_1(t), x_2(t)) \in \mathcal{R} \) for all \( t \geq 0 \) for which the derivative exists, thus implying (4). Equivalence of (5) and (13) is obvious.

**Theorem 2.10:** A subspace \( \mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2 \) is a bisimulation relation between \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) if and only if

\[
\begin{align*}
& \text{a}) \quad \mathcal{R} + \text{im} \begin{bmatrix} G_1 \\ 0 \end{bmatrix} = \mathcal{R} + \text{im} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} : \mathcal{R}_e \\
& \text{b}) \quad \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \mathcal{R} \subset \mathcal{R}_e \\
& \text{c}) \quad \text{im} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \subset \mathcal{R}_e \\
& \text{d}) \quad \mathcal{R} \subset \ker \begin{bmatrix} G_1^T & -C_2 \end{bmatrix} .
\end{align*}
\]
Proof: Clearly, (13) is equivalent with (14d). Let $\mathcal{R}$ satisfy property i) of Proposition 2.9. Take $u = 0$. Then, for all $(x_1, x_2) \in \mathcal{R}$, and for all $d_1$ there exists a $d_2$ such that

$$(A_1x_1 + G_1d_1, A_2x_2 + G_2d_2) \in \mathcal{R}. \tag{15}$$

This is a version of the so-called modified disturbance decoupling problem (with $d_2$ being the “disturbance,” which is assumed to be available for feedforward control action, and $d_2$ being the “control,” [31]). It follows from the solution of this problem (see [31]) that (15) is equivalent to

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \mathcal{R} \subset \mathcal{R} + \mathbb{C} \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \quad \text{im} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \subset \mathcal{R} + \mathbb{C} \begin{bmatrix} 0 \\ G_2 \end{bmatrix}. \tag{16}$$

Analogously, property i) of Proposition 2.9 implies that for all $(x_1, x_2) \in \mathcal{R}$, and all $d_2$ there exists a $d_1$ such that (15) holds, implying similarly

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \mathcal{R} \subset \mathcal{R} + \mathbb{C} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \quad \text{im} \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \subset \mathcal{R} + \mathbb{C} \begin{bmatrix} G_1 \\ 0 \end{bmatrix}. \tag{17}$$

The second lines of (16) and (17) are readily seen to be equivalent to (14a). Consequently, the first lines of (16) and (17) are equivalent with (14b).

Finally, consider property i) for $x_1 = x_2 = 0$ and $d_1 = 0$. Then, (12) implies

$$\text{im} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \subset \mathcal{R} + \mathbb{C} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} = \mathcal{R}_e$$

(the same result follows by taking that $d_2 = 0$ instead of $d_1 = 0$, and thus (14c). By linearity it follows that conversely (14a, b, c) imply property i).

Remark 2.11: Although we have invoked in the proof of Theorem 2.10 the solution of the modified disturbance decoupling problem it is important to stress that the interpretation in the current case is completely different. Indeed, for bisimulation we do not aim to construct a feedback controller in $d_2$ dealing with the disturbance $d_1$, and vice versa. Instead, we only derive algebraic conditions which ensure that the trajectories of $\Sigma_1$ are replicated by $\Sigma_2$ and conversely.

Remark 2.12: Note that a subspace $\mathcal{S} \subset \mathcal{X}_1 \times \mathcal{X}_2$ satisfies properties (14a, b) if and only if the mapping $F$ (from subspaces $\mathcal{S} \subset \mathcal{X}_1 \times \mathcal{X}_2$ to subspaces $F(\mathcal{S}) \subset \mathcal{X}_1 \times \mathcal{X}_2$) defined by

$$\mathcal{S} = \left\{ \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} z + \text{im} \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \right\} \subset \mathcal{S} + \mathbb{C} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \begin{bmatrix} A_1 \\ 0 \\ 0 \\ A_2 \end{bmatrix} z + \text{im} \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \subset \mathcal{S} + \text{im} \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \right\}$$

satisfies $\mathcal{R} \subset F(\mathcal{R})$. This will be instrumental to compute the maximal bisimulation relation; see Algorithm 3.3. In fact, the maximal bisimulation relation turns out to be a fixed-point of this mapping. (This is well known in the theory of bisimulation for concurrent processes [15].)

Remark 2.13: All notions and results regarding bisimulation developed in this paper for continuous-time linear dynamical systems (1) have a direct analogue for discrete-time linear systems $x(k+1) = Ax(k) + Bu(k) + Gd(k), y(k) = Cx(k), k \in \mathbb{Z}$. In particular, bisimulation between two discrete-time linear systems is characterized by the same conditions (14) as in Theorem 2.10.

Bisimilarity is easily seen to imply equality of external behavior. Consider two systems $\Sigma_i, i = 1, 2$, as in (3), with external behavior $B_i$ defined as

$$B_i := \{ (u_i(\cdot), y_i(\cdot)) \left| \exists x_i(\cdot) \text{ such that (3) is satisfied} \right. \} \quad (18)$$

Proposition 2.14: Let $\Sigma_i, i = 1, 2$, be bisimilar. Then, their external behaviors $B_i$ are equal.

However, in the case of nondeterministic systems, that is, $d_i$ is present, systems may have the same external behavior, while not being bisimilar. This is illustrated by the following simple example.

Example 2.15: Consider the two systems

$$\begin{align*}
\Sigma_1 : & \quad x^1 = x^2 \\
& \quad y_1 = x^1 \\
\Sigma_2 : & \quad \dot{z} = d_2 \\
& \quad y_2 = z.
\end{align*} \tag{19}$$

It can be readily seen that there does not exist any bisimulation relation between $\Sigma_1$ and $\Sigma_2$ [consider condition (14a)]. On the other hand, if we restrict to $C^\infty$ external behaviors then obviously $B_1 = B_2$. (Note the different logical quantifiers in the definition of bisimilarity and in equality of external behavior. For bisimilarity, there should exist for every $x^1, x^2$ a $z$ such that for every $d_2$ there exists a $d_1$ with equal external trajectories and conversely, while for equality of external behavior there should exist for every $x^1, x^2, d_1$ a pair $z, d_2$ with equal external trajectories, and conversely.)

An interpretation of the fact that $\Sigma_1$ and $\Sigma_2$ are not bisimilar can be given as follows. Suppose we “test” the system $\Sigma_1$ at some time instant $t = t_0$ in the sense of observing one of its possible external trajectories $y_1(t), t \geq t_0$. At $t = t_0$ the system $\Sigma_1$ is in a given, but unknown, initial state $(x^1(t_0), x^2(t_0))$. Hence, all possible runs $y_1(t), t \geq t_0$, starting from this fixed initial state will have a fixed time-derivative $\dot{y}_1(t_0) = x^1(t_0)$ at $t = t_0$. On the other hand, for $\Sigma_2$ the possible runs $y_2(t), t \geq t_0$, can have arbitrary time-derivative at $t = t_0$. Hence, $\Sigma_1$ and $\Sigma_2$ can be considered to be externally different.

Remark 2.16: Note that in Example 2.15, the solution concept (“semantics”) does play some role. If we choose the solutions different from $C^\infty$, e.g., piecewise $C^\infty$ with $x(\cdot)$ continuous, then the outputs $y_1(\cdot)$ generated by $\Sigma_1$ will have a “higher degree of smoothness” than the outputs $y_2(\cdot)$ of $\Sigma_2$, implying that $B_1 \subset B_2$ while the two behaviors are not exactly equal.

For deterministic systems ($d_i$ absent) it will be shown in Section VI that equality of external behavior does imply bisimilarity.

III. COMPUTING THE MAXIMAL BISIMULATION RELATION

In this section, we show how to compute the maximal bisimulation relation $R \subset X_1 \times X_2$ for two linear dynamical systems
simple observations. The fact also shows up in maximality of the maximal bisimulation relation; a contradiction with the maximality of the subspace dimension. For notational convenience, we define

\[ A^X := \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad G_1^X := \begin{bmatrix} G_1 \\ 0 \end{bmatrix}, \quad G_2^X := \begin{bmatrix} 0 \\ G_2 \end{bmatrix}, \quad C^X := [C_1; -C_2]. \]

Algorithm 3.3: Given two dynamical systems \( \Sigma_1 \) and \( \Sigma_2 \), define the following sequence \( \mathcal{R}^j \), \( j = 0, 1, 2, \ldots \), of subsets of \( A^X \times \Sigma_2 \):

\[ \mathcal{R}^0 = \Sigma_1 \times \Sigma_2 \]

\[ \mathcal{R}^1 = \{ z \in \mathcal{R}^0 | z \in \ker C^X \} \]

\[ \mathcal{R}^2 = \{ z \in \mathcal{R}^1 | A^X z + \im G_1^X \subset \mathcal{R}^1 + \im G_2^X, A^X z + \im G_1^X \subset \mathcal{R}^2 \} \]

\[ \vdots \]

\[ \mathcal{R}^{j+1} = \{ z \in \mathcal{R}^j | A^X z + \im G_1^X \subset \mathcal{R}^j + \im G_2^X, A^X z + \im G_1^X \subset \mathcal{R}^{j+1} \}. \]

Theorem 3.4: The sequence of subsets \( \mathcal{R}^0, \mathcal{R}^1, \ldots, \mathcal{R}^j, \ldots \) satisfies the following properties.

1) \( \mathcal{R}^j, j \geq 0 \), is a linear space or empty. Furthermore \( \mathcal{R}^0 \supset \mathcal{R}^1 \supset \mathcal{R}^2 \supset \cdots \supset \mathcal{R}^j \supset \mathcal{R}^{j+1} \supset \cdots \).

2) There exists a finite \( k \) such that \( \mathcal{R}^k = \mathcal{R}^{k+1} =: \mathcal{R}^* \) and then \( \mathcal{R}^j = \mathcal{R}^* \) for all \( j \geq k \).

3) \( \mathcal{R}^* \) is either empty or equals the maximal subspace of \( X_1 \times X_2 \) satisfying properties (14a, b, d) of Proposition 2.10.

Proof: The proof is very similar to the proof of the corresponding properties of the algorithm for computing the maximal controlled invariant subspace [31] and, hence, only a sketch will be provided.

1) This is easily shown by induction.

2) Follows from Property 1 and finite-dimensionality of the subspaces involved.

3) First, we show that \( \mathcal{R}^* \) obtained in Property 2, whenever nonempty, satisfies properties (14a, b, d). Since \( \mathcal{R}^* = \mathcal{R}^k = \mathcal{R}^{k+1} \) we have

\[ A^X \mathcal{R}^* + \im G_1^X \subset \mathcal{R}^* + \im G_2^X \]

\[ A^X \mathcal{R}^* + \im G_1^X \subset \mathcal{R}^* + \im G_1^X \]

which implies

\[ \im G_1^X \subset \mathcal{R}^* + \im G_2^X \]

\[ \im G_1^X \subset \mathcal{R}^* + \im G_1^X \]

and

\[ A^X \mathcal{R}^* \subset \mathcal{R}^* + \im G_1^X, A^X \mathcal{R}^* \subset \mathcal{R}^* + \im G_1^X. \]

Clearly, (24) is equivalent to (14a), while (25) yields property (14b). Furthermore \( \mathcal{R}^* \subset \mathcal{R}^1 = \ker C^X \), and, thus, \( \mathcal{R}^* \) satisfies property (14d). In order to show that \( \mathcal{R}^* \), if nonempty, is the maximal subspace satisfying properties (14a, b, d), let \( \mathcal{R} \) be any subspace satisfying (14a, b, d). Clearly, \( \mathcal{R} \subset \mathcal{R}^1 \), and by induction we prove \( \mathcal{R} \subset \mathcal{R}^j \) for all \( j \), implying that \( \mathcal{R} \subset \mathcal{R}^* \).

If \( \mathcal{R}^* \) as obtained from Algorithm 3.3 is nonempty and satisfies property (14c), then it follows that \( \mathcal{R}^* \) is the maximal bisimulation relation \( \mathcal{R}^\text{max} \) between \( \Sigma_1 \) and \( \Sigma_2 \), while if \( \mathcal{R}^* \) is empty or does not satisfy property (14c) then there does not exist any bisimulation relation between \( \Sigma_1 \) and \( \Sigma_2 \). With regard to bisimilarity (Definition 2.5), we have the following immediate consequence.

Corollary 3.5: \( \Sigma_1 \) and \( \Sigma_2 \) are bisimilar if and only if \( \mathcal{R}^* \) satisfies Property (14c) and (6).

In the previous section, cf. Proposition 2.14, we have seen that bisimilarity implies equality of external behavior. In fact, a computationally efficient way to check equality of external behavior of two systems \( \Sigma_i, i = 1, 2 \), may be to compute \( \mathcal{R}^* \) as above, and to check bisimilarity by Corollary 3.5. (Recall, however, that nondeterministic systems may have equal behavior while not being bisimilar; cf. Example 2.15.)

IV. REDUCTION OF SYSTEMS BY BISIMULATION

In this section, we study the question how to reduce a linear dynamical system to a system with lower state–space dimension, which is bisimilar to the original system, and in particular how to reduce the system to a bisimilar system with minimal state–space dimension.
It turns out that this can be achieved by considering bisimulation relations between the system and a copy of itself. Furthermore, the reduction to a bisimilar system with minimal state–space dimension can be performed by using the same algorithm as given in the previous section for computing the maximal bisimulation relation. Actually, this idea is well-known in the context of concurrent processes; see, e.g., [11].

So let us consider a linear dynamical system as in (1)

\[ \Sigma: \dot{x} = Ax + Bu + Gd, \quad x \in \mathcal{X}, u \in \mathcal{U}, d \in \mathcal{D} \]

\[ y = Cx, \quad y \in \mathcal{Y} \]

with \( \mathcal{X}, \mathcal{U}, \mathcal{Y}, \) and \( \mathcal{D} \) finite-dimensional linear spaces. Now consider a bisimulation relation between \( \Sigma \) and itself, that is, in view of Theorem 2.10, a subspace \( \mathcal{R} \subset \mathcal{X} \times \mathcal{X} \) satisfying

\[ a) \quad \mathcal{R} + \text{im} \begin{bmatrix} G \\ 0 \end{bmatrix} = \mathcal{R} + \text{im} \begin{bmatrix} 0 \\ G \end{bmatrix} =: \mathcal{R}_e \]

\[ b) \quad \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \mathcal{R} \subset \mathcal{R}_e \]

\[ c) \quad \text{im} \begin{bmatrix} B \\ B \end{bmatrix} \subset \mathcal{R}_e \]

\[ d) \quad \mathcal{R} \subset \ker C. \]

First of all, we note the following obvious fact.

**Proposition 4.1:** The identity relation \( \mathcal{R}_{\text{id}} = \{ (x, x) \mid x \in \mathcal{X} \} \) is a bisimulation between \( \Sigma \) and itself.

Every \( \mathcal{R} \subset \mathcal{X} \times \mathcal{X} \) defines a relation on \( \mathcal{X} \) by saying that \( x_a, x_b \in \mathcal{X} \) are related by \( \mathcal{R} \) if and only if \( (x_a, x_b) \in \mathcal{R} \). For reduction, we should restrict attention to \( \mathcal{R} \subset \mathcal{X} \times \mathcal{X} \) such that the corresponding relation on \( \mathcal{X} \) is an equivalence relation, i.e., \( \mathcal{R} \) is reflexive ((\( x, x \)) \( \in \mathcal{R} \) for all \( x \in \mathcal{X} \)), symmetric ((\( x_a, x_b \)) \( \in \mathcal{R} \) implies \( (x_b, x_a) \in \mathcal{R} \)), and transitive ((\( x_a, x_b \)) \( \in \mathcal{R} \), (\( x_b, x_c \)) \( \in \mathcal{R} \) implies \( (x_a, x_c) \in \mathcal{R} \)). This can be done without loss of generality. Indeed, by Proposition 4.1, we may always add to any bisimulation relation \( \mathcal{R} \), the identity bisimulation relation \( \mathcal{R}_{\text{id}} \), thus enforcing reflexivity. Furthermore, let \( \mathcal{R} \) satisfy (27), then also the inverse relation \( \mathcal{R}^{-1} := \{ (x_a, x_b) \mid (x_b, x_a) \in \mathcal{R} \} \) satisfies (27), implying that the symmetric closure \( \mathcal{R} + \mathcal{R}^{-1} \) satisfies (27). Finally, for linear relations reflexivity and symmetry already implies transitivity: if \( (x_a, x_b), (x_b, x_c) \in \mathcal{R} \), then \( (x_a - x_c, 0) = (x_a, x_b) - (x_b, x_c) \in \mathcal{R} \), and thus \( (x_a, x_c) = (x_a - x_c, 0) + (x_c, x_c) \in \mathcal{R} \). Any equivalence relation \( \mathcal{R} \subset \mathcal{X} \times \mathcal{X} \) can be uniquely associated with a linear subspace \( \mathcal{R} \subset \mathcal{X} \) defined as follows:

\[ \mathcal{R} := \{ x_a - x_b \mid (x_a, x_b) \in \mathcal{R} \} \]

Indeed, \( \mathcal{R} \) defined by (28) is a linear space if and only if \( \mathcal{R} \) is reflexive and symmetric (and therefore an equivalence relation).

Reduction can now be performed by factoring out \( \mathcal{X} \) by the subspace \( \mathcal{R} \), thereby obtaining the linear space of equivalence classes of the equivalence relation \( \mathcal{R} \). In terms of the subspace \( \mathcal{R} \) conditions (27) reduce as follows.

**Theorem 4.2:** Let \( \mathcal{R} \subset \mathcal{X} \times \mathcal{X} \) be an equivalence relation, and define \( \mathcal{R} \subset \mathcal{X} \) as in (28). Conditions (27a, b, c, d) for \( \mathcal{R} \) are equivalent to

\[ A \mathcal{R} \subset \mathcal{R} + \text{im} G \]

\[ \mathcal{R} \subset \ker C. \]
Since $\mathcal{R}$ is a bisimulation relation it immediately follows from (14a, b, d) that
\[
\begin{align*}
A_1\mathcal{R}_1 & \subseteq \mathcal{R}_1 + \text{im } G_1 & \mathcal{R}_1 & \subseteq \ker C_1 \\
A_2\mathcal{R}_2 & \subseteq \mathcal{R}_2 + \text{im } G_2 & \mathcal{R}_2 & \subseteq \ker C_2.
\end{align*}
\] (38)

Hence, similarly as before, the systems $\Sigma_i$ for $i = 1, 2$ can be reduced (via bisimulation) to systems
\[
\Sigma_{\mathcal{R}_i} : \begin{align*}
A_i\mathcal{R}_i & = A_i x_{\mathcal{R}_i} + B_i u + G_i d \\
y & = C_i x_{\mathcal{R}_i}
\end{align*}
\] (39)

living on the state-spaces $\mathcal{X}_i/\mathcal{R}_i$, $i = 1, 2$, with canonical projections $\Pi_{\mathcal{R}_i} : \mathcal{X}_i \rightarrow \mathcal{X}_i/\mathcal{R}_i$, $i = 1, 2$. Furthermore, the relation $\mathcal{R} \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ reduces to a relation $\bar{\mathcal{R}} \subseteq \mathcal{X}_1/\mathcal{R}_1 \times \mathcal{X}_2/\mathcal{R}_2$ defined by
\[
\bar{\mathcal{R}} := \{(x_{\mathcal{R}_1}, \mathcal{R}_2) \mid (x_1, x_2) \in \mathcal{R} \text{ for all }(x_1, x_2)
\text{ with } \Pi_{\mathcal{R}_1}x_1 = x_{\mathcal{R}_1}, \Pi_{\mathcal{R}_2}x_2 = x_{\mathcal{R}_2}\}. \tag{40}
\]

Moreover, it is straightforward to check that $\bar{\mathcal{R}}$ defines a bisimulation between $\Sigma_{\mathcal{R}_1}$ and $\Sigma_{\mathcal{R}_2}$. In fact, we have the following.

Theorem 4.4: Let $\mathcal{R} \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ be bisimulation relation between two systems $\Sigma_i, i = 1, 2$. Then, the systems $\Sigma_i$ can be reduced by bisimulation to the systems $\Sigma_{\mathcal{R}_i}, i = 1, 2$, with $\bar{\mathcal{R}}$ as defined in (40) a bisimulation relation between them. This bisimulation relation $\bar{\mathcal{R}}$ is one-to-one in the sense that $(x_{\mathcal{R}_1}, x_{\mathcal{R}_2}), (x'_1, x'_2) \in \bar{\mathcal{R}}$ implies $x_{\mathcal{R}_1} = x'_1$ and $x_{\mathcal{R}_2} = x'_2$. Furthermore, $\Sigma_i, i = 1, 2$, are bisimilar with bisimulation relation $\mathcal{R}$ if and only if $\Sigma_{\mathcal{R}_i}, i = 1, 2$, are bisimilar with bisimulation relation $\bar{\mathcal{R}}$. In the case of the bisimulation relation $\bar{\mathcal{R}}$, the isomorphism $\mathcal{R}$ corresponds to an invertible map $S : \mathcal{X}_1/\mathcal{R}_1 \rightarrow \mathcal{X}_2/\mathcal{R}_2$.

Theorem 4.14 expresses that every bisimulation relation can be factorized into a one-to-one bisimulation relation between reduced systems, which are themselves bisimilar (by reduction) to the original systems. Furthermore, if $\Sigma_i, i = 1, 2$, are bisimilar then their reduced systems are linked by an invertible mapping (a state-space equivalence mapping).

Corollary 4.5: Suppose the maximal bisimulation relation $\mathcal{R}^{\text{max}}$ between $\Sigma_1$ and $\Sigma_2$ exists and satisfies $\mathcal{R}_1^{\text{max}} = 0, \mathcal{R}_2^{\text{max}} = 0$ and $\pi_1(\mathcal{R}^{\text{max}}) = \mathcal{X}_1, \pi_2(\mathcal{R}^{\text{max}}) = \mathcal{X}_2$. Then, $\mathcal{R}^{\text{max}} = \{(x_1, x_2) = S(x_1) \mid x_1 \in \mathcal{X}_1\}$ for some invertible map $S : \mathcal{X}_1 \rightarrow \mathcal{X}_2$.

V. SIMULATION AND ABSTRACTION

A one-sided version of the notions of bisimulation relation and bisimilarity as provided in Definitions 2.1 and 2.5 can be stated as follows.

Definition 5.1: Consider $\Sigma_1$ and $\Sigma_2$ as given in (3). A simulation relation of $\Sigma_1$ by $\Sigma_2$ is a linear subspace
\[
\mathcal{S} \subseteq \mathcal{X}_1 \times \mathcal{X}_2
\]
with the following property. Take any $(x_{10}, x_{20}) \in \mathcal{S}$ and any joint input function $u_1(\cdot) = u_2(\cdot)$. Then, for every disturbance function $d_1(\cdot)$ there should exist a disturbance function $d_2(\cdot)$ such that the resulting state solution trajectories $x_1(\cdot)$, with $x_1(0) = x_{10}$, and $x_2(\cdot)$, with $x_2(0) = x_{20}$, satisfy
\[
\begin{align*}
i) & \quad (x_1(t), x_2(t)) \in \mathcal{S}, \quad \text{for all } t \geq 0 \quad (41) \\
ii) & \quad C_1x_1(t) = C_2x_2(t), \quad \text{for all } t \geq 0. \quad (42)
\end{align*}
\]

Furthermore, $\Sigma_1$ is simulated by $\Sigma_2$ if the simulation relation $\mathcal{S}$ satisfies $\pi_1(\mathcal{S}) = \mathcal{X}_1$ with $\pi_1$ the canonical projection of $\mathcal{X}_1 \times \mathcal{X}_2$ to $\mathcal{X}_1$.

For several purposes the existence of a simulation relation instead of a bisimulation relation is already sufficient. One may for example think of a system description capturing the required specifications on a system to be designed: If the actual closed-loop system is simulated by the specified system then at least one knows that there is no unwanted behavior in the actual closed-loop system. Or, one may “approximate” a given system by a lower dimensional system in the sense that the given system is simulated by the approximating system. (This is the idea of abstraction as will be discussed later on.) Then checking certain properties for the approximating system will also guarantee these properties for the given system.

By a “one-sided” version of Theorem 2.10, see especially (16), we obtain the following.

Proposition 5.2: $\mathcal{S} \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ is a simulation relation of $\Sigma_1$ by $\Sigma_2$ if and only if
\[
\begin{align*}
a) & \quad \mathcal{S} + \text{im } \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \subseteq \mathcal{S} + \text{im } \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \\
b) & \quad \begin{bmatrix} A_1 \\ 0 \\ A_2 \end{bmatrix} \mathcal{S} \subseteq \mathcal{S} + \text{im } \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \\
c) & \quad \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \subseteq \mathcal{S} + \text{im } \begin{bmatrix} 0 \\ G_2 \end{bmatrix} \\
d) & \quad \mathcal{S} \subseteq \ker C_1 : = \ker C_2. \tag{43}
\end{align*}
\]

For discrete processes, it may occur that there is a simulation relation of $\Sigma_1$ by $\Sigma_2$ and a simulation relation of $\Sigma_2$ by $\Sigma_1$ while there does not exist a bisimulation relation between $\Sigma_1$ and $\Sigma_2$. Compare, e.g., [16] and [7]. In the present context, however, we have the following fact. (Recall the definition of the inverse relation $T^{-1} := \{(x_a, x_b) \mid (x_b, x_a) \in T\}$.)

Proposition 5.3: Let $\mathcal{S} \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ be a simulation relation of $\Sigma_1$ by $\Sigma_2$ and let $T \subseteq \mathcal{X}_2 \times \mathcal{X}_1$ be a simulation relation of $\Sigma_2$ by $\Sigma_1$. Then $\mathcal{R} := \mathcal{S} + T^{-1}$ is a bisimulation relation between $\Sigma_1$ and $\Sigma_2$.

Proof: Let $\mathcal{S}$ satisfy (43) and let $T$ satisfy (43) with index 1 throughout replaced by 2. It follows that
\[
T^{-1} + \mathcal{S} + \text{im } \begin{bmatrix} G_1 \\ 0 \end{bmatrix} \subseteq \mathcal{S} + T^{-1} + \text{im } \begin{bmatrix} 0 \\ G_2 \end{bmatrix}
\]
and thus (14b) results for $\mathcal{R} = \mathcal{S} + T^{-1}$. Similarly, $\mathcal{R}$ satisfies (14b, c, d).

An interesting consequence of Proposition 5.3 is the following. First, we note that if there exists a simulation relation $\mathcal{S}$ of $\Sigma_1$ by $\Sigma_2$ then there also exists the maximal simulation relation $\mathcal{S}^{\text{max}}$ of $\Sigma_1$ by $\Sigma_2$. The argument is the same as in the
case of bisimulation (cf. Section 3): if $\mathcal{S}_1$ and $\mathcal{S}_2$ are simulation relations then also $\mathcal{S}_1 \cup \mathcal{S}_2$ is one. Furthermore, the maximal simulation relation $\mathcal{S}^{\text{max}}$ can be computed by the following simplified version of Algorithm 3.3. Define the decreasing sequence $\mathcal{S}^j$, $j = 0, 1, 2, \ldots$ of linear subspaces of $\mathcal{X}_1 \times \mathcal{X}_2$ as

\[
\begin{align*}
\mathcal{S}^0 &= \mathcal{X}_1 \times \mathcal{X}_2 \\
\mathcal{S}^1 &= \{ z \in \mathcal{S}^0 \mid z \in \ker C^\times \} \\
\mathcal{S}^2 &= \{ z \in \mathcal{S}^1 \mid A^\times z + \im G^\times \subset \mathcal{S}^1 + \im G^\times \} \\
&\quad \vdots \\
\mathcal{S}^{j+1} &= \{ z \in \mathcal{S}^j \mid A^\times z + \im G^\times \subset \mathcal{S}^j + \im G^\times \}. \quad (44)
\end{align*}
\]

Proposition 5.4: Suppose there exists a simulation of $\Sigma_1$ by $\Sigma_2$, and a simulation of $\Sigma_2$ by $\Sigma_1$ (or equivalently, cf. Proposition 5.3, a bisimulation relation between $\Sigma_1$ and $\Sigma_2$). Let $\mathcal{s}^{\text{max}}_1 \subset \mathcal{X}_1 \times \mathcal{X}_2$ denote the maximal simulation relation of $\Sigma_1$ by $\Sigma_2$, and $T^{\text{max}} \subset \mathcal{X}_2 \times \mathcal{X}_1$ the maximal simulation relation of $\Sigma_2$ by $\Sigma_1$. Then, $\mathcal{s}^{\text{max}} = (T^{\text{max}})^{-1} = \mathcal{R}^{\text{max}}$, with $\mathcal{R}^{\text{max}}$ the maximal bisimulation relation.

Proof: By Proposition 5.3 $\mathcal{S}^{\text{max}} + (T^{\text{max}})^{-1}$ is a bisimulation relation, and thus also a simulation relation of $\Sigma_1$ by $\Sigma_2$. By maximality of $\mathcal{S}^{\text{max}}$, it follows that $\mathcal{S}^{\text{max}} + (T^{\text{max}})^{-1} = \mathcal{S}^{\text{max}}$, and thus $(T^{\text{max}})^{-1} \subset \mathcal{S}^{\text{max}}$. The converse inclusion holds by considering $T^{\text{max}}$. Hence, $\mathcal{S}^{\text{max}} = (T^{\text{max}})^{-1}$, which is by Proposition 5.3 a bisimulation relation, and thus contained in $\mathcal{R}^{\text{max}}$. Since $\mathcal{R}^{\text{max}}$ is also a simulation relation, equality follows.

Hence, if there exists a bisimulation relation between $\Sigma_1$ and $\Sigma_2$, then $\mathcal{R}^{\text{max}}$ can also be computed using the simplified algorithm (44).

Now, let us recall the definition of an abstraction as introduced in a general context in [2], [10], [18], and [20], and more specifically in [19] and [29]. Consider a linear system as before, cf. (26)

\[
\begin{align*}
\Sigma : \quad &\dot{x} = Ax + Bu + Gd, \quad x \in \mathcal{X}, u \in \mathcal{U}, d \in \mathcal{D} \\
y &= Cx,
\end{align*}
\]

(45)

together with a surjective linear map $H : \mathcal{X} \to \mathcal{Z}$, $\mathcal{Z}$ being another linear space, satisfying

\[
\ker H \subset \ker C. \quad (46)
\]

Clearly, this implies that there exists a unique linear map $\mathcal{C} : \mathcal{Z} \to \mathcal{Y}$ such that

\[
C = \mathcal{C}H. \quad (47)
\]

Then, define the following dynamical system on $\mathcal{Z}$:

\[
\begin{align*}
\Sigma : \quad &\dot{z} = Az + Bu + Gd, \quad z \in \mathcal{Z}, u \in \mathcal{U}, d \in \mathcal{D} \\
y &= Cz,
\end{align*}
\]

(48)

with

\[
\begin{align*}
\bar{A} &:= HAH^+ \\
\bar{B} &:= HB \\
\mathcal{G} &:= [HGH_1A_1; \ldots ; HGA_v \mathcal{V}]
\end{align*}
\]

(49)

where $H^+$ denotes the Moore–Penrose pseudoinverse of $H$, and $v_1, \ldots, v_k$ span $\ker H$.

We think of $\mathcal{G}$ as an “abstraction” of $\Sigma$ in the sense that we factor out the part of the state variables $x \in \mathcal{X}$ corresponding to $\ker H$. This idea was put forward in a series of innovative papers by Pappas et al. (see especially [19]) in the slightly more restricted case where $B = 0$. (N.B., actually in [19] the “generator for nondeterminism” $d$ is denoted by $u$.)

It has been shown in [19] that $\Sigma$ is simulated by $\mathcal{G}$, as also follows from Proposition 5.2 by taking the simulation relation $\mathcal{R} := \{(x, z) \mid z = Hx\}$. We may think of the transformation of $\Sigma$ into $\mathcal{G}$ as “lumping” the states $x$ corresponding to the same vector $z = Hx$. The underlying idea again comes from the theory of interactive processes, where it is often found convenient to replace a process by a process with a smaller number of states, which however replicates the trajectories of the original system (but in general contains more). Thus, $\mathcal{G}$ can be thought of as a “conservative estimate” of $\Sigma$.

A main result of [19] is that $\mathcal{R} = \{(x, z) \mid z = Hx\}$ is actually a bisimulation relation between $\Sigma$ and $\mathcal{G}$ if and only $\ker H$ is controlled invariant with respect to $A$ and $\im G$. This also follows directly from Theorem 4.2. Generalization of this result to the nonlinear case (see also Section VII) was obtained in [29].

VI. DETERMINISTIC SYSTEMS

In this section, we specialize the results of the previous section to systems without disturbances $d$ (that is, “deterministic systems”), and relate the results to well-known results on state–space equivalence and state–space reduction of standard input-state-output systems [5], [6], [22].

Thus, consider two systems

\[
\begin{align*}
\Sigma_i : \quad &\dot{x}_i = A_i x_i + B_i u_i, \quad x_i \in \mathcal{X}_i, u_i \in \mathcal{U} \\
y_i &= C_i x_i, \quad i = 1, 2
\end{align*}
\]

(50)

with $\mathcal{X}_i, \mathcal{U}, \mathcal{Y}$ finite-dimensional linear spaces. Theorem 2.10 is readily seen to specialize to the following.

Corollary 6.1: A subspace $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ is a bisimulation relation between $\Sigma_1$ and $\Sigma_2$ given in (50) if and only if

\[
\begin{bmatrix}
A_1 & 0 \\
0 & A_2
\end{bmatrix} \subset \mathcal{R} \cap \ker \begin{bmatrix}
C_1^\times \\
C_2^\times
\end{bmatrix}. \quad (51)
\]

For deterministic systems $\Sigma_i$, it turns out that equality of external behaviors $\mathcal{B}_i$, cf. (18), implies bisimilarity (contrary to the case of nondeterministic systems, see Example 2.15). This will be shown using the maximal bisimulation relation. First note that for deterministic systems $\mathcal{R}^*$ as produced by Algorithm 3.3 is nonempty (since $\im G_i^\times$ are zero subspaces).

Theorem 6.2: Consider two systems $\Sigma_1$ and $\Sigma_2$ as given in (50) with equal external behaviors $\mathcal{B}_1 = \mathcal{B}_2$. Then, $\Sigma_1$ and $\Sigma_2$ are bisimilar with bisimulation relation $\mathcal{R}^{\text{max}} = \mathcal{R}^*$ as produced by Algorithm 3.3.
Proof: It is easily seen that $\mathcal{R}^*$ as produced by Algorithm 3.3 is given as

$$
\mathcal{R}^* = \ker \begin{bmatrix} C_1 & \cdots & -C_2 \\ C_1A_1 & \cdots & -C_2A_2 \\ \vdots & \cdots & \vdots \end{bmatrix}. \tag{52}
$$

By considering initial conditions for $\Sigma_1$ and $\Sigma_2$ both equal to zero it is clear that equality of external behaviors implies

$$
\text{im} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \subset \mathcal{R}^*
$$

and, thus, $\mathcal{R}^*$ satisfies (51) and defines a bisimulation relation. What remains to be checked is that $\pi_i(\mathcal{R}^*) = \mathcal{X}_i$, $i = 1, 2$. This follows by considering $u = 0$ and then noting that equality of external behavior of $\Sigma_1$ and $\Sigma_2$ implies that for every $x_1 \in \mathcal{X}_1$ there exists $x_2 \in \mathcal{X}_2$ such that $C_1e^{At_1}x_1 = C_2e^{At_2}x_2$, and conversely. Repeated differentiation of this equality with respect to $t$ and evaluation at $t = 0$ yields in view of (52) that $(x_1, x_2) \in \mathcal{R}^*$.

While equality of external behaviors of two deterministic systems has been shown to be equivalent to their bisimilarity, the existence of a bisimulation relation can be characterized in the following way, relating to classical concepts of system equivalence [5], [6].

Theorem 6.3: There exists a bisimulation relation $\mathcal{R}$ between $\Sigma_1$ and $\Sigma_2$ given in (50) if and only the Markov parameters of $\Sigma_1$ and $\Sigma_2$ are equal, that is

$$
C_1A_k^kB_1 = C_2A_k^kB_2, \quad k = 0, 1, 2, \ldots \tag{53}
$$

or, equivalently, if their transfer matrices $G_i(s) = C_i(sI - A_i)^{-1}B_i$, $i = 1, 2$, are the same.

Moreover, any bisimulation relation $\mathcal{R}$ between $\Sigma_1$ and $\Sigma_2$ satisfies

$$
\pi_i(\mathcal{R}) \supset \text{Reach} (\Sigma_i) \tag{54}
$$

with Reach $(\Sigma_i) = \text{im} [B_1; A_1B_1; \cdots ; A_k^kB_1; \cdots ]$ the reachable subspace of $\Sigma_i$. Hence, if $\Sigma_i, i = 1, 2$, are controllable, then $\Sigma_1$ and $\Sigma_2$ are bisimilar if and only if $G_1(s) = G_2(s)$.

Proof: Condition (53) is equivalent to the property that the subspace

$$
\text{im} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ A_k^kB_1 \\ \cdots \end{bmatrix}
$$

which is the smallest subspace of $\mathcal{X}_1 \times \mathcal{X}_2$ that is invariant under

$$
\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}
$$

and containing $\text{im} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, is contained in $\ker [C_1; - C_2]$. Thus, (53) is equivalent to the existence of $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ satisfying (51), while (54) holds. Controllability is equivalent to Reach $(\Sigma_i) = \mathcal{X}_i$, implying by (54) that $\pi_i(\mathcal{R}) = \mathcal{X}_i$.

With regard to reduction by bisimulation we have the following corollaries of Proposition 4.3 and Theorem 4.4.

Corollary 6.4: Consider

$$
\dot{x} = Ax + Bu, \quad x \in \mathcal{X}, \quad u \in U \quad \text{mod} \quad y = Cx, \quad y \in Y \tag{55}
$$

Then, $\mathcal{R} \subset \mathcal{X} \times \mathcal{X}$ is a bisimulation equivalence relation between $\Sigma$ and itself if and only if the subspace, cf. (28), $\mathcal{R} = \{ x_a - x_b | (x_a, x_b) \in \mathcal{R} \}$ is an unobservable subspace of the system (i.e., $A \mathcal{R} \subset \mathcal{R}, \mathcal{R} \subset \ker C$). The state–space $\mathcal{X}$ of $\Sigma$ can be factored out by $\mathcal{R}$ so as to obtain $\Sigma_R$. Furthermore, $\mathcal{R}^*$ corresponding to the maximal bisimulation relation $\mathcal{R}^\max = \mathcal{R}^*$ between $\Sigma$ and itself equals the maximal unobservability subspace of (55), and can be factored out so as to obtain an observable system $\Sigma_{R^*}$. Conversely, if $\Sigma$ is observable, then $\mathcal{R}^* = \mathcal{R}_v$.

Corollary 6.5: Let $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ be a bisimulation relation between two systems $\Sigma_i$, $i = 1, 2$ given by (50). Then, the systems $\Sigma_i$ can be reduced by bisimulation to the systems $\Sigma_{R_i}, i = 1, 2$, with $\mathcal{R}$ as defined in (40) a bisimulation relation between them. This bisimulation relation $\mathcal{R}$ is one-to-one. Furthermore, $\Sigma_i, i = 1, 2$, are bisimilar with bisimulation relation $\mathcal{R}$ if and only if $\Sigma_{R_i}, i = 1, 2$, are bisimilar with bisimulation relation $\mathcal{R}$. In this case the bisimulation relation $\mathcal{R}$ corresponds to an invertible map. If $\Sigma_1$ and $\Sigma_2$ are bisimilar and both are observable, then the maximal bisimulation relation $\mathcal{R}^*$ is the graph of an invertible map $S : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ satisfying

$$
SA_1S^{-1} = A_2 \quad \text{SB}_1 = B_2, C_1 = C_2S. \tag{56}
$$

Furthermore, by maximality of $\mathcal{R}^*$ the map $S$ satisfying (56) is unique.

Remark 6.6: From the proof of Theorem 6.2 it follows that if $\Sigma_1$ and $\Sigma_2$ are bisimilar then $(x_1, x_2) \in \mathcal{R}^*$ implies that $C_1e^{A_1t_1}x_1 = C_2e^{A_2t_2}x_2$ for all $t \geq 0$. Hence, in this case $\mathcal{R}^*$ is actually given by

$$
\mathcal{R}^* = \ker \begin{bmatrix} C_1 & \cdots & -C_2 \\ C_1A_1 & \cdots & -C_2A_2 \\ \vdots & \cdots & \vdots \end{bmatrix}. \tag{57}
$$

where $n = \max(n_1, n_2), n_i = \dim \mathcal{X}_i$. From here it is directly seen that $\mathcal{R}^* v$ is the maximal unobservable space of $\Sigma_i$. If $\Sigma_i$ and $\Sigma_2$ are bisimilar and at least one of the systems, say $\Sigma_2$, is observable, then $\mathcal{R}^*$ is the graph of the mapping $W_2^{-1}W_1 : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ with

$$
W_1 = \begin{bmatrix} C_1 \\ C_1A_1 \\ \vdots \end{bmatrix}, \quad W_2 = \begin{bmatrix} C_2 \\ C_2A_2 \\ \vdots \end{bmatrix}, \quad W_1^{-1}W_2 \quad \text{where} \ n = n_2 \text{ and } W_2^{-1} \text{ denotes the set-theoretic inverse. (See [21] for a related result in the case of scalar output systems which both are observable.)}

Finally, by combining Theorem 6.3 with Corollary 6.5 we recover the classical state–space equivalence theorem for controllable and observable deterministic systems; see, e.g., [5].

Corollary 6.7: Let the systems $\Sigma_i, i = 1, 2$, given by (50) be observable and controllable. Then, $\Sigma_i, i = 1, 2$, are bisimilar if and only $G_1(s) = G_2(s)$, while the bisimulation relation is given as $\{ (x_1, x_2) | x_1 \in \mathcal{X}_1 \}$ for a uniquely defined invertible map $S : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ (the state–space equivalence map) satisfying (56).
VII. BISIMULATION OF NONLINEAR SYSTEMS

In this section, we briefly indicate the extension of the theory of bisimulation as established for linear dynamical systems (3) to nonlinear systems as shown in (58) at the bottom of the page. Here, as before, $D_i, U, \mathcal{Y}$ are finite-dimensional linear spaces, but the state-space $\mathcal{X}_i$ is a manifold. Furthermore, $a_i(x_i)$ and the columns of the matrices $b_i(x_i)$ and $g_i(x_i)$ denote vector fields on $\mathcal{X}_i$. All data are assumed to be smooth ($C^\infty$). While in the linear case we restricted attention to linear bisimulation relations in this case it is natural to consider nonlinear bisimulation relations $\mathcal{R}$ in Definition 2.1 which are submanifolds of $\mathcal{X}_1 \times \mathcal{X}_2$. (As usual, "submanifold" throughout stands for "imbedded smooth submanifold;" see, e.g., [1] and [17].)

Furthermore, like in Definition 2.5, we say that $\Sigma_1$ and $\Sigma_2$ as in (58) are bisimilar if there exists a nonlinear bisimulation $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ with the property that $\pi_i(\mathcal{R}) = \mathcal{X}_i, i = 1, 2$.

Using the theory of controlled invariance for nonlinear systems; see, e.g., [12] and [17], it is immediate to derive the following generalization of Proposition 2.9.

Proposition 7.1: A submanifold $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ is a bisimulation relation between $\Sigma_1$ and $\Sigma_2$ if and only if for all $(x_1, x_2) \in \mathcal{R}$ and all $u \in U$ the following properties hold.

- For all $d_1 \in D_1$, there should exist a $d_2 \in D_2$ such that

$$ (a_1(x_1) + b_1(x_1)u + g_1(x_1)d_1, a_2(x_2) + b_2(x_2)u + g_2(x_2)d_2) \in \mathcal{T}(x_1, x_2) \mathcal{R} \tag{59} $$

and conversely for every $d_2 \in D_2$ there should exist a $d_1 \in D_1$ such that (59) holds.

- $c_1(x_1) = c_2(x_2). \tag{60}$

Remark 7.3: Note that $\mathcal{R}_e(x_1, x_2)$ is a linear subspace of the tangent space $\mathcal{T}_{x_1} \mathcal{X}_1 \times \mathcal{T}_{x_2} \mathcal{X}_2$, which is defined for all $(x_1, x_2) \in \mathcal{R}$.

Remark 7.4: For simplicity of presentation, we have restricted ourselves to nonlinear systems (58), instead of the more general form $\dot{x}_i = f_i(x_i, u_i, d_i), y_i = h_i(x_i, u_i)$. Proposition 7.1 easily generalizes by considering instead of (59) the conditions $(f_1(x_1, u_1, d_1), f_2(x_2, u_2, d_2)) \in \mathcal{T}(x_1, x_2) \mathcal{R}$ and $h_1(x_1, u) = h_2(x_2, u)$. An appropriate generalization of Theorem 2.7 can be obtained along the lines of [17, Ch. 13].

For computing the maximal nonlinear bisimulation relation one may use the following version of the algorithm to compute the maximal controlled invariant submanifold, cf. [17].

Algorithm 7.5: Given two nonlinear dynamical systems $\Sigma_1$ and $\Sigma_2$ as in (58). Define the following sequence $\mathcal{R}^j, j = 0, 1, 2, \ldots$, of subsets of $\mathcal{X}_1 \times \mathcal{X}_2$, and assume that they are all submanifolds (possibly empty)

$$ \mathcal{R}^0 = \mathcal{X}_1 \times \mathcal{X}_2 $$

$$ \mathcal{R}^1 = \{(x_1, x_2) \in \mathcal{R}^0 | c_1(x_1) = c_2(x_2)\} $$

$$ \mathcal{R}^2 = \{(x_1, x_2) \in \mathcal{R}^1 | \begin{bmatrix} a_1(x_1) \\ a_2(x_2) \end{bmatrix} + \begin{bmatrix} g_1(x_1) \\ g_2(x_2) \end{bmatrix} \} \subset \mathcal{T}(x_1, x_2) \mathcal{R}^1 + \begin{bmatrix} 0 \\ g_1(x_1) \end{bmatrix} \tag{59} $$

$$ \vdots $$

$$ \mathcal{R}^j = \mathcal{R}^{j-1} + \begin{bmatrix} 0 \\ g_1(x_1) \end{bmatrix} \tag{62} $$

Remark 7.6: For a discussion of the condition that $\mathcal{R}^j$ at every step of the algorithm is a submanifold we refer to [17, Ch. 11]. For the nonsmooth case a similar type of algorithm has been provided in the context of viability theory; see, e.g., [3] and the references therein.

Proposition 7.7:

1. $\mathcal{R}^0 \supset \mathcal{R}^1 \supset \mathcal{R}^2 \supset \cdots \supset \mathcal{R}^j \supset \mathcal{R}^{j+1} \supset \cdots$
2. There exists a finite $k$ such that $\mathcal{R}^k = \mathcal{R}^{k+1} = \mathcal{R}^*$. Then $\mathcal{R}^* = \mathcal{R}^j$ for all $j \geq k$.
3. $\mathcal{R}^*$ is the maximal submanifold of $\mathcal{X}_1 \times \mathcal{X}_2$ (possibly empty) satisfying properties (61a, b, d).
Corollary 7.8: Suppose that Algorithm 7.5 is well defined, that is, the subsets $\mathcal{R}_i$ are all submanifolds. Then, $\Sigma_1$ and $\Sigma_2$ are bisimilar if and only if $\mathcal{R}^e$ satisfies (61c) and $\pi_i(\mathcal{R}^e) = \mathcal{X}_i$, $i = 1, 2$.

Reduction of a nonlinear system

$$\Sigma : \dot{x} = a(x) + b(x)u + g(x)d, \quad x \in \mathcal{X}, u \in \mathcal{U}, \quad d \in \mathcal{D} \quad y = c(x)$$

(63)

to a nonlinear system with lower-dimensional state-space $\mathcal{X}_R$ can be achieved, similarly to the linear theory exposed in Section IV, by considering a nonlinear bisimulation relation $\mathcal{R}$ between $\Sigma$ and itself. Such a bisimulation relation is given by a submanifold $\mathcal{R} \subset \mathcal{X} \times \mathcal{X}$ satisfying for all $(x_1, x_2) \in \mathcal{R}$

$$\begin{align*}
a) & \quad (x_1, x_2) \in \mathcal{R} \implies T(x_1, x_2)\mathcal{R} + \text{im} \left[ \begin{bmatrix} g(x_1) \\ 0 \end{bmatrix} \right] = \mathcal{R}e(x_1, x_2) \\
b) & \quad a(x_1) = a(x_2) \\
c) & \quad (b(x_1), b(x_2)) \in \mathcal{R}e(x_1, x_2) \\
d) & \quad c(x_1) = c(x_2),
\end{align*}$$

(64)

Similar to the linear theory in Section IV, we assume that $\Sigma$ defines an equivalence relation on $\mathcal{X}$. Then, for all $x \in \mathcal{X}$, we may define the equivalence classes

$$\mathcal{F}_R = \{x' \in \mathcal{X} \mid (x', x) \in \mathcal{R} \} = \pi(\mathcal{R} \cap \{(x', x) \mid x' \in \mathcal{X}\})$$

(65)

where $\pi$ denotes the projection on the first component. Throughout we assume that the subsets $\mathcal{F}_R$ of $\mathcal{X}$ are submanifolds of $\mathcal{X}$ for every $x \in \mathcal{X}$. (Under generic transversality conditions this will be the case since $F_R$ is defined as the intersection of submanifolds.)

Since $\mathcal{R}$ is assumed to be an equivalence relation it follows that $\mathcal{F}_R, x \in \mathcal{X}$, defines a foliation $\mathcal{F}_R$ of $\mathcal{X}$ in the following sense:

i) $x \in \mathcal{F}_R$

ii) $y \in \mathcal{F}_R \iff x \in \mathcal{F}_R^u$

iii) $\mathcal{F}_R \cap \mathcal{F}_R^u \neq \emptyset \Rightarrow \mathcal{F}_R = \mathcal{F}_R^u$.

(66)

Now, let us consider conditions (64) in more detail. First of all, condition (64b) implies that there exists a vector $\alpha(x_1, x_2)$ such that

$$\begin{bmatrix} a'(x_1) \\ a'(x_2) \end{bmatrix} = \begin{bmatrix} a(x_1) \\ a(x_2) \end{bmatrix} + \begin{bmatrix} g(x_1) \\ 0 \end{bmatrix} \alpha(x_1, x_2) \in T(x_1, x_2)\mathcal{R}.$$  

(67)

Hence, the vector field $\begin{bmatrix} a'(x_1) \\ a'(x_2) \end{bmatrix}$ is tangent to $\mathcal{R}$ for every $(x_1, x_2) \in \mathcal{R}$. This means that the flow of this vector field when starting at $(x_1, x_2) \in \mathcal{R}$ will remain in $\mathcal{R}$ for all $t > 0$ for which it is defined. This implies that the flow of this vector field when projected to the first factor of $\mathcal{X} \times \mathcal{X}$ leaves the foliation $\mathcal{F}_R$ invariant. Using the theory of invariant distributions (see, e.g., [12] and [17]) this is known to be equivalent to the condition

$$[d', D_R] \subset D_R$$

(68)

where $D_R$ is the integrable distribution on $\mathcal{X}$ corresponding to $\mathcal{F}_R$, i.e.,

$$D_R(x) := T_x\mathcal{F}_R,$$  

(69)

Similar statements hold with respect to invariance of $D_R$ with respect to the columns of $b$ and $g$. Using the theory of controlled invariant distributions [12], [17] one then arrives at the following nonlinear analogue of Theorem 4.2.

Proposition 7.9: Let the submanifold $\mathcal{R} \subset \mathcal{X} \times \mathcal{X}$ define an equivalence relation on $\mathcal{X}$. Define the integrable distribution $D_R$ on $\mathcal{X}$ by (69). Then, $\mathcal{R}$ is a bisimulation relation between $\Sigma$ and itself if and only if $D_R$ satisfies

$$[a, D_R] \subset D_R + \text{im} g$$

$$[b, D_R] \subset D_R + \text{im} g$$

$$[g, D_R] \subset D_R + \text{im} g$$

$$D_R \subset \ker d.$$  

(70)

Conditions (70) express that the distribution $D_R$ is controlled invariant (for $a$ and the columns of $b$ with respect to the “input” distribution $\text{im} g(x)$). It follows that the foliation $\mathcal{F}_R$ (or equivalently the distribution $D_R$) can be factored out to obtain a reduced nonlinear system $\Sigma_R$ (assuming that the quotient space is again a manifold; for a discussion of this, see [8]). We leave the details to the reader.

Example 7.10: Consider the nonlinear system (cf. [28])

$$\begin{align*}
\dot{x}_1 &= x_2^2 + d^2 \\
\dot{x}_2 &= d^2 \\
y &= x_1.
\end{align*}$$

(71)

This system is not reducible by bisimulation to the system

$$\begin{align*}
\dot{z} &= d \\
y &= z
\end{align*}$$

(72)

(although their $C^\infty$ or piecewise $C^\infty$ external behaviors are equal!). Indeed, the distribution $D_R := \text{span} \{(\partial/\partial x^2)\}$ does not satisfy (70). Reason is that $\Sigma$ given by (71) has a “deadlock” behavior for $x^2 = 0$, which is not present in the system given by (72). Note that if we replace the first equation of (71) by $\dot{x}_1 = x^2 + d^2$ then $\Sigma$ is (linearly) bisimilar to (72).

Finally, we obtain the following nonlinear (partial) analogue of Theorem 6, 3 for the case that $g_i \equiv 0$ (the “deterministic” case).

Proposition 7.11: A submanifold $\mathcal{R} \subset \mathcal{X}_1 \times \mathcal{X}_2$ is a bisimulation relation between $\Sigma_1$ and $\Sigma_2$ for $g_i \equiv 0$ if and only if for all $(x_1, x_2) \in \mathcal{R}$

$$\begin{align*}
a) & \quad \begin{bmatrix} a_1(x_1) \\ a_2(x_2) \end{bmatrix} \in T(x_1, x_2)\mathcal{R} \\
b) & \quad \text{im} \begin{bmatrix} b_1(x_1) \\ b_2(x_2) \end{bmatrix} \subset T(x_1, x_2)\mathcal{R} \\
c) & \quad c_1(x_1) = c_2(x_2),
\end{align*}$$

(73)
Furthermore, $\mathcal{R}$ is a bisimulation relation if and only if for all $(x_1, x_2) \in \mathcal{R}$:

$$H_1(x_1) = H_2(x_2)$$

(74)

for all corresponding functions $H_i := L_{X_1} L_{X_2} \ldots L_{X_j} c_i$, where $X_j$ are the vector fields given by $a_i$ or the columns of $b_i$, in the observation space of $\Sigma_i$. Finally, the maximal bisimulation relation between $\Sigma_1$ and $\Sigma_2$ equals the set of points $(x_1, x_2)$ satisfying (74).

**Proof:** Obviously, (73) is the specialization of (64) to the present case. By invariance of the observation space of $\Sigma_2$ under $a_i$ and the columns of $b_i$ it follows that satisfaction of (74) implies (73).

### VIII. Conclusion and Outlook

In this paper, we have defined and studied a notion of bisimulation for continuous dynamical systems, motivated by the theory of bisimulation for concurrent processes and by previously obtained results by Pappas et al. Furthermore, we have made explicit the relations with well-known system-theoretic notions of state-space equivalence and reduction. The notion of bisimulation appears to be a notion which unifies the concepts of state-space equivalence and state-space reduction, and which allows to study equivalence of systems with nonminimal state-space dimension. Classical results concerning reduction and state-space equivalence for minimal deterministic systems are readily and elegantly recovered. A crucial tool in all this is the notion of controlled invariance as developed in linear and nonlinear geometric control theory.

Compared with classical systems theory a new twist to the problem is given by the idea of considering nondeterministic continuous dynamical systems. For concurrent discrete processes the advantages of allowing nondeterminism are clear [15], [11], but for continuous dynamical systems such arguments have not been explicitized except in the context of abstraction. Indeed, in the area of verification of hybrid systems it is natural to look for notions of abstraction which allow to extend methods for verification of discrete processes to the hybrid case, see [10], [2], [13], [14], and [30]. This naturally leads to considering ways of abstracting continuous dynamical systems. Apart from abstraction we believe that there are other compelling reasons to study some type of systems. Apart from abstraction we believe that there are other leads to considering ways of abstracting continuous dynamical systems.

Finally, a generalization to the bisimulation of continuous dynamical systems as exposed in the present paper is to suppress the a priori distinction of the external variables into inputs $u$ and outputs $y$. This can be achieved by considering a behavioral version of (1), e.g., given in pencil representation as

$$Ez = Az$$

$$w = Hz$$

(75)

where now $w$ denotes the whole vector of external variables (not split into $u$ and $y$). Note that (1) may be put into the form (75) by first premultiplying the first line of (1) by an annihilating matrix $G^\perp (G^\perp G = 0)$ of maximal rank, to obtain $G^\perp \dot{z} = G^\perp Ax + G^\perp Bu$, and then considering the extended state vector $z = (x, \dot{x})$. Generalization of the bisimulation theory to this class of systems is pursued in [25]. An additional and important advantage of (75) is that it may also include algebraic constraints on the state variables, and thus is eminently suited to model systems as resulting from interconnection of systems, or multimodal systems arising from varying (event-driven) state constraints [27].

### Acknowledgment

Initial discussions with G. Pappas and P. Tabuada (University of Pennsylvania, Philadelphia, PA) and ongoing conversations with J. W. Polderman, R. Langerk, A. Julius, and S. Strubbe (University of Twente, Enschede, The Netherlands) are highly appreciated.

### References


References


