Proc. IFAC Symp. on Robust Control Design, ROCOND '97 Budpapest, Hungary, June 25-27, 1997

ROBUSTNESS-ORIENTED CONTROLLER IDENTIFICATION

Michael G. Safonov [*,1] **Fabricio B. Cabral** [*,2]

[*] *University of Southern California*
Los Angeles, CA 90089-2563, USA

Abstract: The problem of real-time robust controller identification is posed and solved within the framework of Willems' behavioural system theory. The theory is augmented with projection operators to accommodate partial information about signals (e.g., past only and sampling), as well as constraints introduced by feedback control. With this refinement, the Willems theory provides a precise characterization of the increment in control-relevant knowledge that results from the addition of each experimental data point to the knowledge base. Software simulations demonstrate that the theory can significantly enhance the reliability of robust controller designs by reducing reliance on a priori estimates of uncertainty size and structure.

Keywords: Robust control; identification; uncertain dynamic systems; validation; learning control; adaptive control; behavioural science

1. INTRODUCTION

The philosophy of behaviourism holds that science is the art of constructing models which, though purely fictitious creations of the mind, are capable of reproducing experimentally observed behaviours. That is, science seeks to discover patterns amongst experimentally observed signals without pretending that the hypothesized representations of these patterns (i.e., the models and their internal variables) *necessarily* represent physical truth — though the hope is that they do. In (Willems, 1991), a behavioural approach to the theory of dynamical systems was introduced. In this approach, the basic variables considered were the external or manifest variables of the system without making any distinction between them. The collection of trajectories describing the evolution of the manifest variables over time defined a dynamical system (Antoulas and Willems, 1993).

Despite of the fact that one of the central points of the behavioural framework was to have a representation-independent definition of a dynamical system, when control problems were considered in the behavioural framework, plant representations were usually assumed. For instance, a state space representation was assumed in (Willems, 1993) and an image representation was assumed in (Trentelman and Willems, 1994).

Let us notice that control problems without assumed plant representations have been formulated in the past few years in the so called unfalsified control approach (Cabral, 1996; Cabral and Safonov, 1996; Brozenec and Safonov, 1997; Safonov and Tsao, 1995, 1997). The main idea behind this approach is that we can eliminate undesirable controllers through a check of the consistency between (i) the data, (ii) the set of candidate control laws, and (iii) the performance specification.

In this work, we formulate, in the behavioural framework, control problems which do not assume any plant representation.

2. MATHEMATICAL MODELS

The fundamental piece of the "behavioural" framework of Willems (1991) is the definition of a mathematical model. This definition is formulated according to the black box point of view, "in which we focus on how a system behaves, on the way it

[1] Research supported in part by AFOSR Grant F49620-95-I-0095.
[2] Supported by a grant from CAPES and CNPq.

interacts with its environment, instead of trying to understand, in the tradition of physics, how it is put together and how its components work (Willems, 1991)". This definition of a mathematical model formalizes, in its ultimate generality, the black box point of view. This ultimate generality can be more strongly evidenced by the fact that Willems backs off "from the usual input/output setting, from the processor point of view, in which systems are seen as influenced by inputs, acting as causes, and producing outputs through these inputs, the internal conditions, and the system dynamics."

Willems (1991) starts by assuming that there is a phenomenon to be modelled. Then he "casts the situation in the language of mathematics by assuming that the phenomenon produces elements in a set $\mathbf{U}$", called the universum. The elements of $\mathbf{U}$ are called the outcomes of the phenomenon. "A (deterministic) mathematical model for the phenomenon (viewed purely from the behavioural, the black box point of view) claims that certain outcomes are possible, while others are not. Hence a model recognizes a certain subset $\mathcal{B}$ of $\mathbf{U}$. This subset will be called the behaviour (of the model)." And the formal definition is given as follows.

Definition 2.1. A mathematical model is a pair $(\mathbf{U}, \mathcal{B})$, with $\mathbf{U}$ the universum — its elements are called outcomes — and $\mathcal{B} \subset \mathbf{U}$ the behaviour. □

With respect to data and measurements, we cite the following excerpts from Willems (1991): "We will now cast measurements in this setting. We will assume that we make certain measurements which we will call the data." "... we ... assume that the data consists of observed realizations of the phenomenon itself. Thus, a data set will be a nonempty subset $\mathcal{D}$ of $\mathbf{U}$." The formal definition of a data set is, then, given as follows.

Definition 2.2. A data set is a nonempty subset $\mathcal{D}$ of $\mathbf{U}$. □

2.1 *Unfalsified Models*

Willems (1991) also introduces the concept of an unfalsified model. This concept is related to earlier works on model validation and deterministic hypothesis testing. It was introduced by Willems as a reminiscence of Popper's philosophy of science in his language for modelling from data. Implicit in both Willems work and Popper's falsification concept is the existence of a unique, but unknown model $(\mathbf{U}, \mathcal{B}_{true})$ representing the undiscovered 'truth'. In this context experimentally observed data is regarded as a source of information about some of the elements z_{true} of the set $\mathcal{B}_{true}$. A model $(\mathbf{U}, \mathcal{B})$ of $(\mathbf{U}, \mathcal{B}_{true})$ is regarded as *valid* if it predicts outcomes in the sense $\mathcal{B}_{true} \subset \mathcal{B}$, otherwise it is *invalid.* Experimental observation may provide information about some of the elements of $\mathcal{B}_{true}$ and hence may sometimes *falsify* a model, i.e., prove it to be invalid.

For the case in which data consists of a set $\mathcal{D}$ of direct measurements of outcomes $z \in \mathcal{B}$, Willems (1991) offers the following definition: "A mathematical model $(\mathbf{U}, \mathcal{B})$ is said to be *unfalsified* by the data $\mathcal{D} \subset \mathcal{B}_{true}$ if $\mathcal{D} \subset \mathcal{B}$". However, Willems (1991) noted that this definition may be inadequate when only partial information of each outcome is available, as with "the observations ... generated by a function of the outcomes", but chose "not to pursue these complications here." Yet partial information is typically all that we may ever have in the case of time signals z whose past may be seen but whose future values at any time necessarily remain beyond view. In dealing with dynamical systems, it is essential to have a more flexible definition of unfalsification that can accommodate partial knowledge of outcomes z.

2.1.1. *Partial Information*

Following Safonov and Tsao (1995; 1997), let us consider the case in which universum $\mathbf{U}$ is a vector-space of time-signals — e.g., $\mathbf{U} = \mathcal{L}_{2e}$. Suppose that measurement information evolves as time $\tau \in \mathsf{R}_+$ increases. Suppose further that at each time τ, we have available measurement data $\mathcal{D} \subset P_\tau \mathcal{B}_{true}$ where $P_\tau : \mathbf{U} \to \mathbf{U}$ is a non-invertible mapping, so the data $\mathcal{D}$ provides only partial information about signal vectors $z_{true} \in \mathcal{B}_{true}$, viz., there exists at least one

$$z_{true} \in P_\tau^{-1}(\mathcal{D}) \triangleq \{z \mid (P_\tau z) \in \mathcal{D}\}. \quad (1)$$

Here $P_\tau^{-1}(\mathcal{D})$ denotes the *inverse image* of a set $\mathcal{D}$ under the map P_τ. It is a set-valued map of subsets of $\mathbf{U}$ into subsets of $\mathbf{U}$; it is always well-defined, even though P_τ itself is non-invertible.

For example, measurements of the past history of $z(t)$ would correspond to the case in which P_τ is the familiar time-truncation projection operator of input-output stability theory

$$[P_\tau z](t) = \begin{cases} z(t), & \text{if } t \in [0, \tau] \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Because it contains no information about the future $z(t)$, $(t > \tau)$, time-truncated data $z_{data} = P_\tau z_{true}$ provides only partial information about the signal z_{true}.

Definition 2.3. Given a model $(\mathbf{U}, \mathcal{B})$, a mapping $P_\tau : \mathbf{U} \to \mathbf{U}$ and a data set $\mathcal{D} \subset P_\tau \mathbf{U}$, we say that the model $(\mathbf{U}, \mathcal{B})$ is *falsified* if knowledge of P_τ and $\mathcal{D}$ is sufficient to deduce that the model $(\mathbf{U}, \mathcal{B})$ is invalid; otherwise, the model $(\mathbf{U}, \mathcal{B})$ is said to be *unfalsified.*

Thus a model $(\mathbf{U}, \mathcal{B})$ is unfalsified if, for each $z_{data} \in \mathcal{D}$, there exists an outcome $z \in \mathcal{B}$

associated with the model such that $z_{data} = P_\tau z$. For the special case considered by Willems (1991), P_τ is the identity map and $z_{data} = z$.

This definition leads to the following result.

Theorem 2.1. Let the set $\mathcal{D}$ be as in (1) above. A behavioural model $(\mathbf{U}, \mathcal{B})$ of $(\mathbf{U}, \mathcal{B}_{true})$ is falsified by data $\mathcal{D} \subset P_\tau \mathcal{B}_{true}$ if and only if for some $z_{data} \in \mathcal{D}$

$$P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B} = \emptyset$$

Proof. From Definition 2.3, it follows that falsification is equivalent to $z_{data} \notin P_\tau \mathcal{B}$ for some $z_{data} \in \mathcal{D}$ or, equivalently, $P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B} = \emptyset$. □

The following result is immediate.

Corollary 2.1. A behavioural model $(\mathbf{U}, \mathcal{B})$ of $(\mathbf{U}, \mathcal{B}_{true})$ is unfalsified by data $\mathcal{D} \subset P_\tau \mathcal{B}_{true}$ if, and only if, for each $z_{data} \in \mathcal{D}$, the set $P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B}$ is non-empty. □

3. UNFALSIFIED CONTROLLERS

As stated in (Willems, 1991), the intersection of behaviours is "a way of formalizing that additional laws are imposed on a system." Thus, the role of a controller is to impose constraints on the plant behaviour so that the plant behaves as desired. The term unfalsified, on the other hand was first associated with models in the Willems's (1991) definition of an unfalsified model. The concept of an unfalsified model is related to earlier works on model validation and deterministic hypothesis testing. It was introduced by Willems as a reminiscence of Popper's philosophy of science in his language for modelling from data. In an analogous way, Safonov and Tsao (1995) introduced the concept of an unfalsified controller as a reminiscence of Popper's philosophy of science in their paradigm for identifying controllers directly from data.

It is possible to dispense with the references to Popper and explain these concepts in engineering terms. Basically, an unfalsified controller is a controller that fits the data and performance constraints in the same sense that a curve fits the data in a curve fitting process. Thus, unfalsified control deals with getting to know if a controller fits or does not fit the data and performance constraints.

Mathematically, the plant may be regarded as a true but unknown system $(\mathbf{U}, \mathcal{B}_p)$ and the controller as a known system $(\mathbf{U}, \mathcal{B}_c)$. Willems observed that the feedback control results in 'closed-loop' behaviour $(\mathbf{U}, \mathcal{B}_{cl})$ given by $\mathcal{B}_{cl} \triangleq \mathcal{B}_p \cap \mathcal{B}_c$. That is, feedback interconnection corresponds to intersecting the behaviours of plant and controller and, if $\mathcal{B}_p$ is the true but unknown open-loop plant, then the true closed-loop behaviours that are possible when the controller is in place must satisfy $z_{true} \in \mathcal{B}_p \cap \mathcal{B}_c$. Given a desired closed-loop behaviour $(\mathbf{U}, \mathcal{B}_d)$, a controller $(\mathbf{U}, \mathcal{B}_c)$ is said to be *valid* if $\mathcal{B}_p \cap \mathcal{B}_c \subset \mathcal{B}_d$ or, equivalently, $\mathcal{B}_p \cap \mathcal{B}_c \cap \bar{\mathcal{B}}_d = \emptyset$; otherwise, it is said to be *invalid*. Measurements of open-loop plant data can be used to demonstrate that a controller is invalid.

Definition 3.1. Given plant data $\mathcal{D} \subset P_\tau \mathcal{B}_p$ and a desired closed behaviour $(\mathbf{U}, \mathcal{B}_d)$, we say that a controller $(\mathbf{U}, B_c)$ is *falsified* if knowledge of P_τ and $\mathcal{D}$ is sufficient to deduce that the controller $(\mathbf{U}, \mathcal{B}_c)$ is invalid; otherwise, the controller $(\mathbf{U}, \mathcal{B}_c)$ is said to be *unfalsified.*

Note that the controller need not be in the loop when the data $\mathcal{D}$ is collected. The data that falsifies a controller may be open-loop plant data.

In order for information in open-loop data $\mathcal{D}$ to be useful in falsifying controllers, it is necessary that one be able to ensure that that data is sufficiently rich to predict at least one true closed-loop behaviour. That is, the set $P_\tau^{-1}(\mathcal{D}) \cap \mathcal{B}_c \cap \mathcal{B}_p$ must known to be non-empty. For example, this is assured if the controller was actually in place when the data was collected or, more generally, if the controller could have produced this data if it were in place. More precisely, the following well-posedness condition is required.

Definition 3.2. The quadruple $(\mathbf{U}, B_p, B_c, P_\tau)$ is said to define a *well-posed* unfalsified control problem if, for each $z_{data} \in P_\tau \mathcal{B}_p$, it holds that $P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B}_c \cap \mathcal{B}_p \neq \emptyset$ whenever $P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B}_c \neq \emptyset$. □

Evidently (cf. Safonov and Tsao, 1997), for causal plants a sufficient condition for well-posedness is that the data include all past values of all signals (viz., sensor and actuator signals) through which the controller could have interacted with the plant had it been in place up to the current time. The latter is consistent with the logical premise that control actions at any time τ ought not to directly depend on unseen internal or future values of plant variables.

For a well-posed problem the information provided by open-loop plant data $\mathcal{D} \subset P_\tau \mathcal{B}_p$ may be combined with knowledge of the controller behaviour $\mathcal{B}_c$ to obtain partial information about true closed-loop behaviours $z_{true} \in \mathcal{B}_{cl} = \mathcal{B}_p \cap \mathcal{B}_c$, viz. there exists at least one such $z_{true} \in P_\tau^{-1}(\{z_{data}\}) \cap \mathcal{B}_c \cap \mathcal{B}_p$. Based upon this observation, one obtains the the following feedback control counterpart of Theorem 2.1.

Theorem 3.1. Suppose that $(\mathbf{U}, B_p, B_c, P_\tau)$ is well-posed. The controller $(\mathbf{U}, \mathcal{B}_c)$ is unfalsified by the measurement data $z_{data} \in \mathcal{D} \subset P_\tau \mathcal{B}_p$ if and only if

$$P_\tau^{-1}(\{z_{data}\}) \cap B_c \cap \mathcal{B}_d \neq \emptyset \qquad (3)$$

whenever

$$P_\tau^{-1}(\{z_{data}\}) \cap B_c \neq \emptyset.$$

Proof. By well-posedness, we have that for each $z_{data} \in P_\tau^{-1}(\{z_{data}\}) \cap B_c$ there exists at least one z_{true} in $P_\tau^{-1}(\{z_{data}\}) \cap B_c \cap \mathcal{B}_p$ and, if the controller is to be valid, this z_{true} must also be in $\mathcal{B}_d$. Whence, the controller is falsified by the data $\mathcal{D}$ if and only if every element of $P_\tau^{-1}(\{z_{data}\}) \cap B_c$ is not an element of $\mathcal{B}_d$. □

Comparing the above definitions and results for controller unfalsification, we see that controller unfalsification is essentially the same as model unfalsification, except that the open-loop information set $P_\tau^{-1}(\mathcal{D})$ is replaced by the smaller closed-loop information set $\mathcal{D}_{cl} \triangleq P_\tau^{-1}(\mathcal{D}) \cap \mathcal{B}_c$. Roughly, the foregoing may be interpreted as saying that the feedback controller shrinks the universum to $\mathbf{U} \leftarrow \mathcal{B}_c \cap \mathbf{U}$ — exactly like the effect of prior knowledge for unfalsified models. In this sense, controller unfalsification and model unfalsification are mathematically indistinguishable — system identification and controller identification are one and the same.

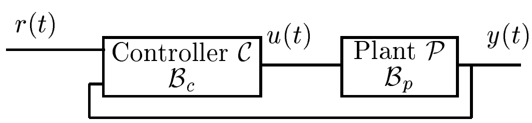

Fig. 1. Feedback control system.

Example 3.1. Consider the system in Figure 1. Let $z = \mathrm{col}(r, y, u) \in \mathbf{U}$ where $\mathbf{U} = \mathcal{R} \times \mathcal{Y} \times \mathcal{U} = \mathcal{L}_{2e}^{n_z}$. Here $\mathcal{R} = \mathcal{L}_{2e}^{n_r}$ is the set of reference signals, $\mathcal{Y} = \mathcal{L}_{2e}^{n_y}$ and $\mathcal{U} = \mathcal{L}_{2e}^{n_u}$ are sets of plant signals, and $n_z = n_r + n_y + n_u$. In keeping with Willems (1991), we omit any arrows on the block diagram in Figure 1, thereby emphasizing the fact that one need not suppose that u is input or y is output.

The plant characteristic is nominally contained in $\mathcal{U} \times \mathcal{Y}$. However, since we are interested in the control system's behaviour which is contained in $\mathcal{R} \times \mathcal{Y} \times \mathcal{U}$, we must consider the 'true' plant behaviour $(\mathbf{U}, \mathcal{B}_p)$ as a trivial extension of its characteristic $\mathcal{P} \subset \mathcal{U} \times \mathcal{Y}$, i.e.

$$\mathcal{B}_p = \{(r, y, u) \mid (u, y) \in \mathcal{P},\ r \in \mathcal{R}\}.$$

Let us suppose that $\mathcal{P}$ is unknown, save that we a measurement data $(u_{data}(t), y_{data}(t)),\ t \in [0, \tau]$. Taking into account the fact that, in the absence of feedback control, this data would be observed for any $r \in \mathcal{R}$, we have that each single measurement (u_{data}, y_{data}) results in a data set

$$\mathcal{D} = \{(r, y, u) \mid \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} u_{data}(t) \\ y_{data}(t) \end{bmatrix} \forall t \in [0, \tau]\}.$$

where $P_\tau : \mathbf{U} \to \mathbf{U}$ is the projection operator

$$[P_\tau(r, y, u)](t) = \begin{cases} (0, y_{data}(t), u_{data}(t)), \forall t \in [0, \tau] \\ (0, 0, 0),\ \text{otherwise} \end{cases}$$

This quadruple $(\mathbf{U}, B_p, B_c, P_\tau)$ is well-posed (cf. Definition 3.2) since the data includes measurements of all past values of the signals (u, y) through which the hypothetical controllers $\mathcal{B}_c(\theta)$ interact with the plant.

Consider now a parametrized class of controllers $(\mathbf{U}, \mathcal{B}_c(\theta))$ with

$$\mathcal{B}_c(\theta) = \{(r, y, u) \mid u = r + \theta y\},$$

where $\theta \in \mathsf{R}$ is a constant real parameter. It follows that $(r, y, u) \in P_\tau^{-1}(\mathcal{D}) \cap \mathcal{B}_c(\theta)$ if and only if for all $t \in [0, \tau]$

$$r = u_{data} - \theta y_{data}, y = y_{data}, u = u_{data}.$$

For instance, let the desired behaviour be given by

$$B_d = \{(r, y, u) \mid \mathcal{J}(r, y, u, t) \geq 0\ \forall t \in \mathcal{T}\},$$

where $\mathcal{J}(r, y, u, t) \triangleq \delta(t) - \int_0^t (y - w_m * r)^2 d\tau$, w_m is a time domain reference model transfer function, δ is a function $\delta : \mathcal{T} \to \mathsf{R}_+$ and $\mathcal{T} \subset [0, \tau]$ is a set of time instants. Then, by Theorem 3.1 $(\mathbf{U}, \mathcal{B}_c(\theta))$ is unfalsified by the data $(u_{data}(t), y_{data}(t)),\ t \in [0, \tau]$ if and only if for all $t \in \mathcal{T}$ with $t \leq \tau$

$$\mathcal{J}(u_{data} - \theta y_{data}, y_{data}, u_{data}, t) \geq 0.$$

Thus, at each time τ, the set of unfalsified controllers is given by the set of parameters

$$\{\theta \mid \theta^2 A_t - 2\theta B_t + C_t \leq 0,\ \forall t \in \mathcal{T} \cap [0, \tau]\},$$

where

$$A_t = \int_0^t (w_m * y_{data})^2 d\tau \qquad (4)$$

$$B_t = \int_0^t (w_m * u_{data} - y_{data}) w_m * y_{data} d\tau \qquad (5)$$

$$C_t = \int_0^t (y_{data} - w_m * u_{data})^2 d\tau - \delta(t). \qquad (6)$$

Notice that from the expressions (4) and (5), we have that $A_t = 0$ implies that $B_t = 0$. Noticing, also, that the expression (4) implies that $A_t \geq 0$, the set of unfalsified controllers $\Theta^*(\tau)$ can be given in an explicit form as follows.

If $A_t = 0$ and $C_t > 0$ for some $t \in \mathcal{T}$, then $\Theta^*(\tau)$ is empty, and all candidate controllers are falsified. Otherwise, $\Theta^*(\tau),\ \tau \in \mathcal{T}$ is given by

$$\Theta^*(\tau) = \bigcap_{\substack{t \in \mathcal{T} \cap [0,\tau] \\ A_t \neq 0}} \Theta_t$$

where Θ_t denotes the set of values of $\theta \in \mathsf{R}$ satisfying $(B_t - \gamma_t)/A_t \leq \theta \leq (B_t + \gamma_t)/A_t$, and $\gamma_t = \sqrt{B_t^2 - A_t C_t}$.

Summarizing, from an initial data point (u_{data}, y_{data}), which can correspond to open- or closed-loop data, we are able to make a statement, based on the data information only, about each controller as to whether it constrains the plant in a desired way or not.

4. CURVE FITTING

When we introduced the term unfalsified in this paper, we said that we could dispense with the references to Popper and explain it in engineering terms. We said that an unfalsified controller is a controller that fits the data and performance constraints in the same sense that a curve fits the data in a curve fitting process. In this section we illustrate with examples what we mean by a curve fitting process.

4.1 *A Curve Fitting Example*

Example 4.1. Let $\mathcal{S}$ be a finite set of points in R^2. Let us say that we are interested in finding a linear curve that passes through the origin and fits the data points in $\mathcal{S}$ in some sense. It is clear that in order to proceed we need to formalize what we mean by a fit in some sense. Since the problem is restricted to linear curves that pass through the origin, the candidate curves can be characterized by

$$\mathcal{C}_\theta = \{(u, y) \mid y = \theta u.\}$$

Let us say then that a curve fits the data if $\sum_{(u,y)\in\mathcal{S}} (y - \theta u)^2 \leq \epsilon n_{\mathcal{S}}$, where $n_{\mathcal{S}}$ is the number of elements of $\mathcal{S}$ and $\epsilon \in \mathsf{R}_+$ indicates our tolerance in the fitting process. An alternative way of specifying this fitness criterion is, first, to define a performance index

$$I(\mathcal{S}, \theta) = \sum_{(u,y)\in\mathcal{S}} (y - \theta u)^2 - \epsilon n_{\mathcal{S}},$$

and, then, to say that the curve $\mathcal{C}_\theta$ fits the data if $I(\mathcal{S}, \theta) \leq 0$.

Thus, in the same sense that we say that a curve fits the specification, we can say that it does not violate the fitness specification. Or that it is unfalsified by the data and fitness specification.

4.2 *A More Difficult Curve Fitting Example*

Example 4.2. This example is particularly related to our control examples, in which the signals belong to $\mathcal{L}_{2,[0,T]}$. A more difficult curve fitting problem than the previous one is to consider, instead of the set $\mathcal{S}$ with $n_{\mathcal{S}}$ elements, the unitary set

$$\{(u_{data}, y_{data})\} \subset \mathsf{R}^{n_S} \times \mathsf{R}^{n_S},$$

i.e.,

$$u_{data} = (u_1, u_2, \ldots, u_{n_{\mathcal{S}}}) \text{ with } u_i \in \mathsf{R},\ 1 \leq i \leq n_{\mathcal{S}} \text{ and } y_{data} = (y_1, y_2, \ldots, y_{n_{\mathcal{S}}}) \text{ with } y_i \in \mathsf{R},\ 1 \leq i \leq n_{\mathcal{S}};$$

and to have, instead of one performance index, $n_{\mathcal{S}}$ performance indexes, one for each element of the set $\{((u_1, \ldots, u_j), (y_1, \ldots, y_j)) \mid 1 \leq j \leq n_{\mathcal{S}}\}$; and, as a criterion of fitness, to require that all the $n_{\mathcal{S}}$ performance indexes be non positive.

5. THE OPEN-LOOP VS. CLOSED-LOOP DATA PUZZLE

Having clarified that the term unfalsified can be understood in the same way that we understand the term fit in a curve fitting process, the only remaining puzzle is related to the question of open or closed loop data.

Prior to the introduction of the unfalsified control paradigm (Safonov and Tsao, 1995), the idea of falsifying a controller had always been associated with plugging one controller in closed-loop and observing the behaviour of the resulting system (Martensson, 1985; Fu and Barmish, 1986). For instance, in (Martensson, 1985), a search over a dense set of controllers is performed while the observed behaviours lead to a violation of a performance specification. In each step of this search, only the data corresponding to the actual controller is considered. A similar technique is used in (Fu and Barmish, 1986) in which a search over a finite set of controllers is considered. In our case, we are not concerned with how the plant data was obtained. The plant data can be open or closed-loop data, and, in this last case, it does not matter which controller was used to close the loop.

For us, as for Willems (1991), a control law is a constraint on the plant behaviour. Moreover, Willems says that this is new: "Our concept of a dynamical system leads to a new view ..., in particular, to what constitutes a feedback control law (Willems, 1991)". Thus it seems that the key to the open-loop vs. closed-loop data puzzle lies in the understanding what is meant by a control law *from the behavioural perspective.*

6. SIMULATIONS

6.1 *Problem Formulation*

Consider the framework of the example 3.1. The goal is to determine at each time τ an unfalsified $\hat{\theta}(\tau) \in \Theta^*(\tau)$. The unfalsified controller $\mathcal{B}_c(\hat{\theta}(\tau))$ is employed and the reference signal is a unit step $r(t) = 1\ \forall t \geq 0$. We choose $w_m = \mathcal{L}^{-1}(\frac{1}{s+1})$, $\delta(t)$ be defined by $\delta(t) = exp(-2t)$, and $\mathcal{T}$ be the set of time instants determined by an initial time=

0 seconds, a sampling frequency= 100 Hz, and a final time= T seconds.

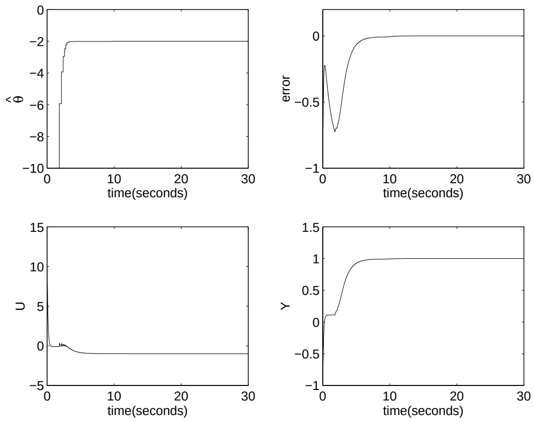

Fig. 2. Initial $\hat{\theta}(0) = -10$ and $x_0 = -1$

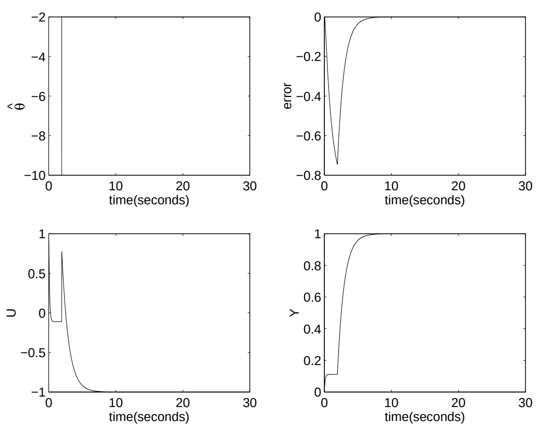

Fig. 3. Initial $\hat{\theta}(0) = -10$ and $x_0 = 0$

6.2 *Simulation Results*

We present two simulations. For purposes of this simulation the 'true, but unknown plant' $\mathcal{B}_p$ has open-loop transfer function $Y(s)/U(s) = \frac{1}{s-1}$ and the initial value for $\hat{\theta}(t)$ is $\hat{\theta}(0) = -10$. We update the controller parameter $\hat{\theta}(\tau)$ only when a falsification occurs and, in this case, we use as $\hat{\theta}(\tau)$ the point at the center of the set $\Theta^*(\tau)$ of unfalsified controllers. In the first simulation (see figure 2), we use $x_0 = -1$ as the value of the plant initial condition. In the second simulation, on the other hand (see figure 3), we use $x_0 = 0$.

7. CONCLUDING REMARKS

In this paper we laid the unfalsified control paradigm (Safonov and Tsao, 1995) for robust controller identification in the general framework for behavioural models of Willems (1991). The behavioural theory was expanded to give more careful attention to the implications of partial information induced when measurements consist of a non-invertible 'projection' $P_\tau z_{true}$ giving us partial information about only the past of z_{true}. The definition of an unfalsified controller was formulated and the term *unfalsified* was explained not only in terms of a reminiscence of Popper's philosophy of science but also in terms of a curve fitting process. Feedback control was seen to work by introducing constraints that effectively 'shrink' the universum, much as prior knowledge shrinks the class of models in system identification. The central concepts in this framework were the Willems' definition of a mathematical model, the understanding of control laws as constraints, and the fact that constraints imposed on a system can be formalized as the intersection of behaviours.

8. REFERENCES

Antoulas, A. C. and J. C. Willems (1993). "A Behavioral Approach to Linear Exact Modeling". *IEEE Trans. on Automatic Control* **38**(12), 1776–1802.

Brozenec, T. F. and M. G. Safonov (1997). Controller identification. In: *Proceedings of the American Control Conference*. Albuquerque, NM.

Cabral, F. B. (1996). Data-Based Control. PhD thesis. University of Southern California. Los Angeles, CA.

Cabral, F. B. and M. G. Safonov (1996). "A Falsification Perspective on Model Reference Adaptive Control ". In: *to appear on Proc. IEEE Conf. on Decision and Control*. Kobe, Japan.

Fu, M. and B. R. Barmish (1986). "Adaptive Stabilization of Linear Systems Via Switching Control". *IEEE Trans. on Automatic Control* **31**, 1097–1103.

Martensson, B. (1985). "The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization". *System Control Letters* **6**, 87–91.

Safonov, M. G. and T. C. Tsao (1995). "The Unfalsified Control Concept: A Direct path from Experiment to Controller". In: *Feedback Control, Nonlinear Systems and Complexity* (B. A. Francis and A. R. Tannenbaum, Eds.). pp. 196–214. Springer-Verlag. Berlin.

Safonov, M. G. and T. C. Tsao (1997). "The Unfalsified Control Concept and Learning". *IEEE Trans. on Automatic Control.*

Trentelman, H. L. and J. C. Willems (1994). "H_∞ Control in a Behavioral Context". In: *Proc. IEEE Conf. on Decision and Control*. Lake Buena Vista, FL. pp. 3695–3700.

Willems, J. C. (1991). "Paradigms and Puzzles in the Theory of Dynamical Systems". *IEEE Trans. on Automatic Control* **36**(3), 259–294.

Willems, J. C. (1993). "LQ-Control: A Behavioral Approach". In: *Proc. IEEE Conf. on Decision and Control*. San Antonio, TX. pp. 3664–3665.