2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Design of Experiments for Guaranteed Parameter Estimation in Membership Setting

S. Borchers and R. Findeisen

Abstract— **We address the problem of designing experiments to obtain guaranteed and as good as possible parameter estimates for linear systems subject to bounded disturbances. First, we review some existing results relevant for the set-membership parameter estimation and outer-bounding. Based on these results, we approach the a priori experimental design problem. By considering a min–max setup, a selection approach is proposed to choose experiments which provide maximum information in worst-case. The proposed approach allows us furthermore to study identifiability from a practical perspective, to investigate the role of initial conditions for identification, and to analyze how disturbances affect the desired estimates.**

I. INTRODUCTION

Obtaining or refining the parameters of a mathematical model describing a dynamic process is an ubiquitous problem and required for prediction or control synthesis. To this end, experiments have to be performed with the process, to obtain measurements for parameter estimation. The data however is typically affected by some noise or disturbances, which has to be considered at stage of experimental design and for parameter estimation.

The parameter estimation and experimental design problem for dynamical systems has been studied extensively for the case the data uncertainty is caused by (random and additive) noise, see e.g. [14], [19] and references therein. An alternative approach, known as set–membership or bounded error description, is to assume the uncertainty to be bounded, but otherwise unknown.

Early references of this approach are [27] and [22] in the domain of state estimation, and for parameter estimation of linear (output) systems see e.g. [26], [17], [4] and the references therein; for an application of the set–membership approach to nonlinear systems see e.g. [9], [10] and [20]. For linear systems, the membership setting allows to derive a polytopic set of the feasible parameters, and various approaches have been derived to determine simple–shaped sets which are guaranteed to contain the set of feasible parameters. For example, ellipsoids [22], [23] and [12] have been considered, as well as orthotopes ([16]), zonotopes ([25], [18]), or homothety ([6]).

In this contribution, we outline a novel min–max experimental design and identifiability analysis approach for linear, discrete time systems, in membership setting. We extend our previous one–step ahead approach [6] to the multi–step case. The robust experimental design problem is approached in a min-max setting, where the volume of the consistent parameter set is considered as selection criterion. The methods are illustrated by several examples.

Paper Structure: We first outline the considered setup in Section II. In Section III, we review shortly the set–membership parameter identification approach [6] following the ideas of set–dynamics employed in [2], [1]. In Section IV, we focus on the min–max experimental design problem. In Section V, we relate the results to the (N–step) identifiability problem. In Section VI, we present an learning strategy based on one–step experiments.

Basic Nomenclature: The sets of non–negative and non–negative real numbers are denoted, respectively, by $\mathbb{N}, \mathbb{R}_+$. All sets considered in the remainder are compact and convex sets (unless otherwise stated). The collection of non–empty compact sets in $\mathbb{R}^n$ is denoted by $Com(\mathbb{R}^n)$. For shorthand of notation, we denote $z_k \doteq (z_{1,k}, z_{2,k}, \ldots, z_{n_x,k})^T$ and $u_k \doteq (u_{1,k}, u_{2,k}, \ldots, u_{n_u,k})^T$ the state and input vectors at time k. The integer sequence is denoted by $\mathbb{N}_{[a:b]} \doteq \{a, a+1, \ldots, b\}$ with $a \in \mathbb{N}, b \in \mathbb{N}, a < b$.

II. SETUP

We consider linear systems of the form:

$$x_{k+1} = A(\lambda)x_k + B(\lambda)u_k + w_k, \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ and $w_k \in \mathbb{R}^{n_x}$ are the current state, control and the unknown disturbance respectively, x_{k+1} is the successor state, and $\lambda \in \mathbb{R}^{n_\lambda}$ denotes the (unknown) system parameters. The system structure is known, i.e. the matrices $A(\lambda)$, $B(\lambda)$ are given by:

$$A(\lambda) = \sum_{i=1}^{n_\lambda} A_i\lambda_i,\ B(\lambda) = \sum_{i=1}^{n_\lambda} B_i\lambda_i, \tag{2}$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{n_\lambda})$, and for all $i \in \{1, 2, \ldots, n_\lambda\}$, the matrix pairs (A_i, B_i) are known and are of compatible dimension, i.e. $(A_i, B_i) \in \mathbb{R}^{n_x \times n_x} \times \mathbb{R}^{n_x \times n_u}$.

We furthermore assume some (limited) prior knowledge on the parameters and the disturbances to be available, i.e. prior bounding sets of the parameters and the disturbance. We denote the sets by Λ and W respectively, and assume for simplicity that both sets are polytopic (compact and convex) sets in $\mathbb{R}^{n_\lambda}$ and $\mathbb{R}^{n_x}$ respectively,

$$\begin{aligned} \Lambda : &= \{\lambda \in \mathbb{R}^{n_\lambda} : M_0\lambda \le l_0\}, \\ W : &= \{w \in \mathbb{R}^{n_x} : M_w w \le l_w\}, \end{aligned} \tag{3}$$

The authors are with the Institute for Automation Engineering, Otto-von-Guericke-Universität Magdeburg, Germany. S. Borchers is also a member of the International Max Planck research school (IMPRS) Magdeburg. Thanks to S.V. Rakovic for helpful discussions and ideas.

with known matrix–vector pairs $(M_0, l_0) \in \mathbb{R}^{r_i \times n_\lambda} \times \mathbb{R}^{r_i}$ and $(M_w, l_w) \in \mathbb{R}^{r_w \times n_\lambda} \times \mathbb{R}^{r_w}$.

Remark 1: The parameters λ are not known apart from being bounded, though they do not change with time. In contrast, the disturbances w can take different values in time, known only to be bounded.

For ease of notation, we denote for any state/control pair $(x_k, u_k) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$ and for any $i \in \{1, 2, \ldots, n_\lambda\}$,

$$\begin{aligned} y_i(x, u) &\doteq A_i x_k + B_i u_k, \qquad (4)\\ Y(x_k, u_k) &\doteq (y_1(x_k, u_k)\ y_2(x_k, u_k)\ \cdots\ y_{n_\lambda}(x_k, u_k)), \end{aligned}$$

where $y_i(x_k, u_k) \in \mathbb{R}^{n_x}$, $Y(x_k, u_k) \in \mathbb{R}^{n_x \times n_\lambda}$. Notice that, under the construction above, for any (x_k, u_k), $Y(x_k, u_k)\lambda = A(\lambda)x_k + B(\lambda)u_k$.

Finally, when referring to an N–step experiment, we mean an instance $E(x_0, \boldsymbol{u})$ with feasible initial condition $x_0 \in \mathcal{X}_0 \subset \mathbb{R}^{n_x}$ and feasible N–step input sequence $\boldsymbol{u} = \{u_k \in \mathcal{U}\}_{k=0}^{N-1}$, where $\mathcal{X}_0 \subset \mathbb{R}^{n_x}$ denotes an initial condition, and $\mathcal{U}$ an input domain. *Performing* such an experiment yields, typically disturbed, state sequences $\{x_k \in \mathbb{R}^{n_x}\}_{k=1}^{N}$.

III. Set–Membership Parameter Estimation

Parameter estimation is the task of obtaining as good as possible parameter estimates considering the available measurements.

We assume given, besides prior knowledge on the initial parameter and disturbance bounds (3), a possibly disturbed state sequence $\{x_k\}_{k=1}^{N}$ obtained from an (N–step) experiment $E(x_0, \boldsymbol{u})$. For simplicity, we consider all states to be measured; the more general case can be found in [7], [20]. The set–membership parameter estimation problem takes then the following form:

Problem 1 (Parameter identification): Estimate the set $\Theta_N \subseteq \Lambda$ of parameters that is *consistent* with the available experimental data $\{x_k\}_{k=0}^{N}, \{u_k\}_{k=0}^{N-1}$, i.e. estimate the *consistent parameter set*

$$\begin{aligned} \Theta_N \doteq\ & \{\lambda \in \Lambda:\ \forall k \in \mathbb{N}_{[0:N-1]}, \qquad (5)\\ & x_{k+1} = A(\lambda)x_k + B(\lambda)u_k + w_k,\\ & w_k \in W\}. \end{aligned}$$

A. Exact Description

Recall that the model parameters λ are known only to the extend that $\lambda \in \Lambda$ and that they do not change over time (i.e. the values of λ are, at any time instance $k \in \mathbb{N}$, equal to its values at the beginning of the process). However, the disturbance w is not known and it can take, at any point in time, any arbitrary value in the set W. Following the set–dynamics ideas presented in [2], [1], we have:

Proposition 3.1 (Parameter set dynamics): The consistent parameter set (5) is described by the dynamic map

$$\Theta_{k+1} = F(\Theta_k, x_{k+1}, x_k, u_k), \qquad (6)$$

where $F(\cdot,\cdot,\cdot,\cdot) : Com(\mathbb{R}^{n_\lambda}) \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_\lambda}$ is given by:

$$F(\Theta_k, x_{k+1}, x_k, u_k) = \{\lambda \in \Theta_k : x_{k+1} - Y(x_k, u_k)\lambda \in W\}. \qquad (7)$$

The proof can be found in the appendix. Hence, parameter identification reduces to the determination of the sequence $\{\Theta_k\}_{k=1}^{N}$ of consistent parameter sets, for the given initial parameter set $\Theta_0 = \Lambda$, and the available data $\{x_k\}_{k=0}^{N}$ and $\{u_k\}_{k=0}^{N-1}$. In the considered linear–polytopic setting, the computation of the sequence $\{\Theta_k\}_{k=1}^{N}$ simplifies then:

Proposition 3.2 (Consistent parameter set): The consistent parameter sets Θ_k, $k \in \{1, 2, \ldots, N\}$ are given by:

$$\Theta_k = \{\lambda \in \Lambda :\ M_k\lambda \le l_k\}, \qquad (8)$$

with $\Theta_k = F(\Lambda, \{x_i\}_1^k, x_0, \{u_i\}_0^{k-1})$, where for all $j \in \{1, 2, \ldots, k\}$:

$$M_j = \begin{pmatrix} M_{j-1} \\ -M_w Y(x_{j-1}, u_{j-1}) \end{pmatrix},\ l_j = \begin{pmatrix} l_{j-1} \\ l_w - M_w x_j \end{pmatrix}. \qquad (9)$$

The proof is provided in the appendix. The exact consistent parameter set (8) is constructed recursively. In the case of parameter estimation, usually only few inequalities of (8) contribute to the boundary of the consistent parameter set. Redundant constraints can be neglected, e.g. following [15], to obtain a minimal representation of the consistent parameter set.

B. Outer–Bounding

For the considered system class, the consistent parameter sets Θ_k (8) are polytopic, see Prop. 3.2. In practice, one is often interested in the uncertainty interval associated with a parameter λ_i, i.e. the axis–aligned projection of the consistent parameter set Θ_N (8) onto the respective coordinate axis. Its length provides e.g. a measure of the quality of the estimate, analogously to the confidence intervals considered in a statistical setting.

The lower and upper bound which define the (compact) uncertainty interval of the i–th parameter, $i \in \{1, \ldots, n_\lambda\}$, are given by:

$$\begin{aligned} \mathcal{O}_i(\Theta_N) &\doteq [\underline{\lambda}_i, \overline{\lambda}_i], \qquad (10)\\ & \underline{\lambda_i} = \min_\lambda \{\lambda_i\},\ \overline{\lambda_i} = \max_\lambda \{\lambda_i\},\\ s.t. \quad & \lambda \in \Theta_N. \end{aligned}$$

The length of the (inner and outer) bounding interval of a parameter $\lambda_i \in \Theta_N$ is denoted by

$$\ell_i^N = \overline{\lambda}_i - \underline{\lambda}_i. \qquad (11)$$

For ease of presentation, we define the bounding orthotope as the Cartesian product of all n_λ bounding intervals, i.e.

$$\mathcal{O}(\Theta_N) \doteq \mathcal{O}_1(\Theta_N) \times \mathcal{O}_2(\Theta_N) \times \ldots \times \mathcal{O}_{n_\lambda}(\Theta_N).$$

By definition, $\mathcal{O}(\Theta_N)$ is Lebesgue measurable (see e.g. [21]), and its volume $Vol(\cdot) : \; Com(\mathbb{R}^{n_\lambda}) \rightarrow \mathbb{R}_+$ takes the form

$$Vol\left(\mathcal{O}(\Theta_N)\right) = \prod_{i=1}^{n_\lambda} \ell_i^N.$$

The bounding orthotope and in particular its volume are used later on as selection criterion for experimental design, obtained by (10) solving $2n_\lambda$ linear programs. Alternatively, the bounding orthotope can be obtained via a single geometric optimization, which is required later on for experimental design, as follows:

Proposition 3.3 (Bounding orthotope): The collection of bounds of Θ_N and respective volume are obtained by:

$$\begin{aligned} \mathcal{O}(\Theta_N) \quad &= \quad \arg\max_{\underline{\lambda},\overline{\lambda}} \{ \prod_{i=1}^{i=n_\lambda} (\overline{\lambda}_i^{(i)} - \underline{\lambda}_i^{(i)}) \} \qquad (12) \\ s.t. \quad & \qquad \forall i \in \mathbb{N}_{[1:n_\lambda]},\ \overline{\lambda}_i^{(i)} \geq \underline{\lambda}_i^{(i)}, \\ & \qquad \overline{\lambda}^{(i)} \in \Theta_N,\ \underline{\lambda}^{(i)} \in \Theta_N. \end{aligned}$$

The volume is simply obtained by replacing "argument" with "max" in (12). The proof immediately follows from construction.

Remark 2: Notice that $2n_\lambda$ (independent) variables are introduced, denoted by $\underline{\lambda}^{(i)}$ and $\overline{\lambda}^{(i)}$ for $i \in \{1, 2, \ldots, n_\lambda\}$, and that $\underline{\lambda} = (\underline{\lambda}_1^{(1)}, \ldots, \underline{\lambda}_{n_\lambda}^{(n_\lambda)})$, $\overline{\lambda} = (\overline{\lambda}_1^{(1)}, \ldots, \overline{\lambda}_{n_\lambda}^{(n_\lambda)})$.

Remark 3: Note that by construction it holds that $\Theta_{k+1} \subseteq \Theta_k$, hence $\mathcal{O}(\Theta_{k+1}) \subseteq \mathcal{O}(\Theta_k)$ and $\ell_i^{k+1} \leq \ell_i^k$, i.e. the uncertainty intervals sequences are monotonically non–increasing (compare Ex. 1). Also, whenever $Vol\left(\mathcal{O}(\Theta_k)\right) = \{\emptyset\}$ 12, $\Theta_k = \{\emptyset\}$, thus providing fact that the model (1) is invalid (inconsistent with the measurements).

Illustrative Example 1

As example we consider the following uncertain linear system

$$x_{k+1} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} x_k + \begin{pmatrix} \lambda_5 \\ \lambda_6 \end{pmatrix} u_k + w_k \qquad (13)$$

with $n_x = 2$, $n_u = 1$, and $n_w = 2$. The disturbances $w_k = (w_{1,k}, w_{2,k})^T$ are bounded, $0 \leq w_{1,k} \leq 0.2$, $0 \leq w_{2,k} \leq 0.2$, and the six parameters are unknown to the extend

$$\Lambda = \Theta_0 = \{\lambda \in \mathbb{R}^6 : \forall i \in \mathbb{N}_{[1:6]}, 0 \leq \lambda_i \leq 1\}.$$

We generate artificial measurements ($N = 30$) using the reference parameters $\lambda^* = (0.1, 0.2, 0.1, 0.3, 0.2, 0.1)^T$. We consider two experiments with same initials $x_0 = (0, 0)^T$, same input sequence $u_0 = 1$, $\{u_k \sim \{0, 1\}\}_1^{29}$. Two different realizations are obtained by considering two independent random disturbance sets $\{w_{i,k}^{(i)} \sim [0, 0.2]\}_1^{29}$, $i = \{1, 2\}$, by which two sequences $\{x_k^{(i)}\}_{k=0}^{30}$, $i = \{1, 2\}$ are obtained.

For this two measurement sequences, we estimate the dynamics of bounding intervals for the six parameters according to Prop. 3.3. The results are depicted in Fig. 1. The example demonstrates that although parameters intervals can be narrowed, the estimates quality strongly depends on the actual disturbances.

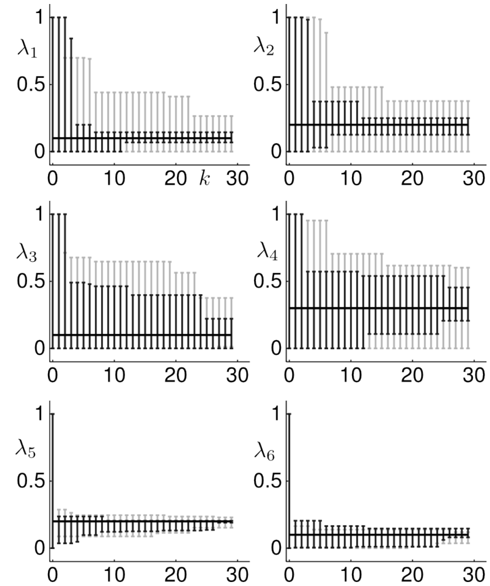

Fig. 1. Orthotopic outer bounding. Evolution of the bounding intervals $\mathcal{O}_i(\Theta_k)$ for two realizations of the same experiment, shown in different colors. Reference values are indicated by the black lines.

IV. Experimental Design

We now turn on the problem of designing *optimal* experiments in membership setting. Particularly, we aim to plan experiments which lead to a *minimal volume consistent parameter set* in worst–case, ideally a singleton set, thereby providing a maximum of information. Since the actual parameters are unknown, "worst–case" here means the most unfavorable disturbances *and* parameters (in Λ).

Obviously, this problem is much more challenging then parameter estimation, since, apart from prior knowledge (3), little further information is available. Actual measurements are not known and can take any feasible value. However, we can exploit the information that N consecutive, singleton, and feasible measurements will be available.

We denote by $\boldsymbol{z} \doteq \{z_k \in \mathbb{R}^{n_x}\}_{k=1}^N$ a feasible state sequence, and consider again the system (1) with prior bounds on parameters and disturbance (3). Controls of the domain $U = \{u : u \in \mathbb{R}^{n_u}\}$ can be applied, and the initial condition can be chosen from $x_0 \in \mathcal{X}_0 \subset \mathbb{R}^{n_x}$. The experimental design problem in min–max setting takes the following form:

Problem 2 (Experimental design): Plan an experiment $E(x_0^*, \boldsymbol{u}^*)$ with initial condition $x_0 \in \mathcal{X}_0$ and N–step input sequence $\boldsymbol{u} \doteq \{\boldsymbol{u}_k \in \mathcal{U}\}_{k=0}^{N-1}$, which minimizes, for worst possible measurements $\boldsymbol{z}^*$, parameters $\boldsymbol{\lambda}^*$, and disturbances $\boldsymbol{w}^*$, the volume of the consistent parameter set Θ_N (8), i.e. find

$$(x_0^*, \boldsymbol{u}^*, \boldsymbol{z}^*, \lambda^*, \boldsymbol{w}^*) = \arg\min_{x_0, \boldsymbol{u}} \max_{\boldsymbol{z}, \lambda, \boldsymbol{w}} \{Vol\left(\Theta_N\right)\}, \qquad (14)$$

where $Vol(\cdot) : \; Com(\mathbb{R}^{n_\lambda}) \rightarrow \mathbb{R}_+$ defines the selection criterion, and the consistent parameter set $\Theta_N = F(\Lambda, \boldsymbol{z}, x_0, \boldsymbol{u})$ as in (8)–(9) a family of polytopic sets.

Problem 2 is in general hard to solve. To obtain the desired guaranteed results we propose the following two relaxations. First, determining the exact volume of polytopic sets is very difficult for the general case $n_\lambda \geq 3$. Therefore, we consider instead the volume of the bounding orthotope $Vol\,(\mathcal{O}(\Theta_N))$. This provides a (outer) bound of the actual volume, and hence guarantees can still be provided. Second, we consider a discrete domain of initial conditions and the control set, e.g. $x_0 \in X_d = \{x_j \in \mathbb{R}^{n_x},\ j \in \{1, 2, \ldots, n_{x_d}\}\}$, and $u_k \in U_d = \{u_j \in \mathbb{R}^{n_u},\ j \in \{1, 2, \ldots, n_{u_d}\}\}$ respectively.

Problem 2 then consist in *selecting* the experiment $E(x_0^*, \boldsymbol{u}^*)$, for which the volume of $\mathcal{O}(\Theta_N)$ is minimized in worst–case.

Proposition 4.1 (Experimental selection): The experiment $E(x_0^*, \boldsymbol{u}^*)$ (15) minimizes the volume of the consistent parameter set $\mathcal{O}(\Theta_N)$ (16) for worst–case disturbances, where

$$(x_0^*, \boldsymbol{u}^*) = \arg\min_{x_0, \boldsymbol{u}} \{Vol\,(\mathcal{O}(\Theta_N))^*\}, \qquad (15)$$

$$Vol\,(\mathcal{O}(\Theta_N))^* = \max_{\boldsymbol{z}, \underline{\lambda}, \overline{\lambda}, \boldsymbol{w}} \{ \prod_{i=1}^{i=n_\lambda} (\overline{\lambda}_i^{(i)} - \underline{\lambda}_i^{(i)})\} \qquad (16)$$

$$\text{s.t.} \quad \forall i \in \mathbb{N}_{[1:n_\lambda]},$$

$$\overline{\lambda}_i^{(i)} \geq \underline{\lambda}_i^{(i)},\ \overline{\lambda}^{(i)} \in \Theta_N, \underline{\lambda}^{(i)} \in \Theta_N.$$

Hereby, $Vol(\cdot):\ Com(\mathbb{R}^{n_\lambda}) \to \mathbb{R}_+$, $\Theta_N = F(\Lambda, \boldsymbol{z}, x_0, \boldsymbol{u})$ as in (8)–(9). Proof immediately follows from construction (Prop. 3.2 and Prop. 3.3).

Analogously to (12), $2n_\lambda$ independent variables are introduced. It is important to note that for one–step ahead ($N = 1$), Problem (16) is log–max concave, i.e. a geometric program (see e.g. [8]). We discuss this important case in more detail in Section VI; for the general case, problem (16) is non–convex due to bilinear constraints, and hence requires to solve a polynomial programs. To this end, global optimization approaches can be considered, for example the method of moments [13], branch and bound procedures [24], or using a relaxation approach as in [7].

The computational complexity of the proposed approach depends in general on the number of considerable experiments. When considering a discrete input and initial domain as in Prop. 4.1, the proposed experimental selection approach requires solving $n_{x_d} n_{u_d}^N$ programs (16).

Remark 4: Note that for the trivial case $n_\lambda = 1$, $Vol(\mathcal{O}(\lambda))^* = Vol(\lambda)^*$, i.e. the input design problem is solved *exactly*. Also for the case $n_\lambda = 2$, where the consistent parameter set is an area whose measure can be *explicitly* described using vertex enumeration (e.g. following [3]), outer–bounding is not required.

Remark 5: The proposed N–step experiment selection approach can also be scheduled in closed loop, i.e. by updating the "initial" parameter set when a new measurement is available.

TABLE I

N–STEP EXPERIMENTAL DESIGN APPROACH. WORST–CASE VOLUME $Vol(\mathcal{O}(\Theta_N))^*$ AND ASSOCIATED GUARANTEED BOUNDING INTERVALS $\ell_1, \ldots, \ell_4$ FOR EXPERIMENTS $E(x_0 \equiv 0, \{u_k\}_0^{N-1})$.

N	**input**			**Volume**	**bounding intervals**			
	u_0	u_1	u_2	$\mathcal{O}(\Theta_N)$	ℓ_1	ℓ_2	ℓ_3	ℓ_4
1	0	-	-	1e4	10	10	10	10
	1	-	-	5e2	10	10	10	0.5
2	0	0	-	1e4	10	10	10	10
		1	-	5e2	10	10	10	0.5
	1	0	-	0	10	0	10	0.5
		1	-	0	10	0	10	0.5
3	0	0	0	1e4	10	10	10	10
			1	5e2	10	10	10	0.5
		1	0	0	10	0	10	0.5
			1	0	10	0	10	0.5
	1	0	0	0	0	0	0.5	0.5
			1	0	0	0	0.66	0.5
		1	0	0	0	0	0.5	0.5
			1	0	0	0	0.66	0.5

Illustrative Example 2

Consider the system

$$x_{k+1} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & 0 \end{pmatrix} \cdot x_k + \begin{pmatrix} 0 \\ \lambda_4 \end{pmatrix} \cdot u_k + \begin{pmatrix} 0 \\ w_k \end{pmatrix}, \qquad (17)$$

where the disturbance w_k can take any values in $0 \leq w_k \leq 0.5$, and the four parameters are unknown to the extend

$$\Lambda = \{\lambda_i \in \mathbb{R} : 1 \leq \lambda_i \leq 11, i = \{1, \ldots, 4\}\}.$$

We now consider the case where the initial condition $x_0 = (0, 0)^T$ is fixed, and aim to design worst case optimal inputs, considering binary input signals $U = \{0, 1\}$. Remind that future states $\boldsymbol{z} = \{x_k\}_1^N$ are unknown. Tab. I shows the results.

As a conclusion, already a three–step experiment provides improvement of the outer–bounds for the parameters in worst–case, to the extend provided in Tab. I. Here, the sequences $\boldsymbol{u} = \{1, 0, 0\}$ and $\boldsymbol{u} = \{1, 1, 0\}$ are distinguished as optimal inputs, minimizing the volume of the (anticipated) consistent parameter set and the associated bounding intervals.

V. PRACTICAL IDENTIFIABILITY

An important conclusion can be drawn from the case $Vol\,(\mathcal{O}(\Theta_N))^* = 0$ for a feasible experiment $E(x_0, \boldsymbol{u})$. Then, by construction, at least one parameter bounding interval is a singleton set, i.e. $\ell_i^N = 0$ for some i (in worst case). Hence, the respective parameter can be uniquely *identified in N–steps by* $E(x_0, \boldsymbol{u})$, in worst case.

Note that identifiability, in classical notion, is concerned with the theoretical existence of unique solutions [5], and hence strictly a mathematical problem. The identifiability problem in the worst–case membership setting, as considered here, is rather motivated from a practical point of view, namely whether point estimates of parameters can be actually obtained. To this end, we have:

Proposition 5.1 (N–step identifiability): Given a system as in (1), with unknown parameters $\lambda \in \Lambda$, bounded disturbance as in (3), and a feasible N–step experiment $E(x_0, \boldsymbol{u})$. If $\ell_i^N = 0$ with

$$\begin{aligned} \ell_i^N \quad &= \max_{\boldsymbol{z},\underline{\lambda},\overline{\lambda},\boldsymbol{w}} \{(\overline{\lambda}_i - \underline{\lambda}_i)\} \qquad (18) \\ \text{s.t.} \quad & \overline{\lambda}_i \geq \underline{\lambda}_i,\ \overline{\lambda} \in \Theta_N, \underline{\lambda} \in \Theta_N, \end{aligned}$$

$\Theta_N = F(\Lambda, \boldsymbol{z}, x_0, \boldsymbol{u})$ (8)–(9), then λ_i is *identifiable in* N *steps* by $E(x_0, \boldsymbol{u})$. If for all $i \in \{1, \ldots, n_\lambda\}$ we have $\ell_i^N = 0$, then model (1) is said identifiable (in N steps) by the experiment $E(x_0, \boldsymbol{u})$.

Prop. 5.1 provides a sufficient criterion for parameter/model identifiability. This notion of identifiability directly extends to the robust case; we say a parameter is (μ–) estimable if $\ell_i^N \leq \mu < \ell_i^0$ with $\mu \in \mathbb{R}_+$ a desired threshold and ℓ_i^0 the initial bounding interval of parameter λ_i (possibly unbounded).

As a consequence, the experimental selection approach Prop. 2 necessitates a prior selection criterion such as identifiability (Prop. 5.1), or more generally the dimension (i.e. box–counting dimension [11]) of the consistent parameter set Θ_N. To this end, the objective of (16) can be tailored to a reduced orthotope excluding identifiable parameter(s) λ_j. (15).

As an example, reconsider Ex. 2, Tab. I; the experiment with the input sequence $\boldsymbol{u} = \{1, 0, 0\}$ allows to *identify* λ_1, λ_2, and to *estimate* λ_3, λ_4 with $\mu = 0.5 \leq 10$ in worst case, i.e. for any admissible disturbances.

VI. One–step Design

In this section, we explore the possibility to *estimate* all the model parameters with one–step experiments. We consider systems as in (1) with single–entries

$$A(\lambda) = \begin{pmatrix} \lambda_{11}^A & \cdots & \lambda_{1n}^A \\ \lambda_{21}^A & \cdots & \lambda_{2n}^A \\ & \vdots & \\ \lambda_{n1}^A & \cdots & \lambda_{nn}^A \end{pmatrix}, B(\lambda) = \begin{pmatrix} \lambda_{11}^B & \cdots & \lambda_{1m}^B \\ \lambda_{21}^B & \cdots & \lambda_{2m}^B \\ & \vdots & \\ \lambda_{n1}^B & \cdots & \lambda_{nm}^B \end{pmatrix},$$

where λ_{ij}^A and λ_{il}^B ($\forall i, j \in \mathbb{N}_{[1:n_x]},\ l \in \mathbb{N}_{[1:n_u]}$) denoting the unknown parameters. Without loss of generality, we furthermore focus on the case where the disturbances are unknown with

$$W = \{w \in \mathbb{R}^{n_x} : \forall i \in \mathbb{N}_{[1:n_x]}, \underline{w}_i \leq w_i \leq \overline{w}_i\}.$$

For this simplified setup, the following "learning approach" based on one–step experimental design can be considered:

First, choose for all $j \in \{1, \ldots, n_x\}$ one–step experiments of the form

$$E^{(j)}(c_j e_j, u \equiv 0), \qquad (19)$$

where $e_j \in \mathbb{R}^{n_x}$ denote the unit–vector of the j–th coordinate and $c_j \in \mathbb{R}$ the respective amplitude scalar. Second, we choose for all $l \in \{1, \ldots, n_u\}$ one–step experiments of the form

$$E^{(l)}(x_0 \equiv 0, d_l e_l), \qquad (20)$$

where $e_l \in \mathbb{R}^{n_u}$ denote the l–th unit vector and $d_l \in \mathbb{R}$ the input's amplitude. Then:

Proposition 6.1 (One–step design): $n_x + n_u$ one–step experiments (19)–(20) are sufficient to determine all unknown parameters λ_{ij}^A and λ_{il}^B ($\forall i, j \in \mathbb{N}_{[1:n_x]},\ l \in \mathbb{N}_{[1:n_u]}$) where

$$\frac{1}{c_j}(z^{(j)} - \overline{w}) \leq \begin{pmatrix} \lambda_{1j}^A \\ \vdots \\ \lambda_{n_x j}^A \end{pmatrix} \leq \frac{1}{c_j}(z^{(j)} - \underline{w}),$$

$$\frac{1}{d_l}(z^{(l)} - \overline{w}) \leq \begin{pmatrix} \lambda_{1l}^B \\ \vdots \\ \lambda_{n_x l}^B \end{pmatrix} \leq \frac{1}{d_l}(z^{(l)} - \underline{w}),$$

with $z^{(j)}, z^{(l)} \in \mathbb{R}^{n_x}$ denoting the state measurements.

The proposed approach offers two important insights. First, the role of the initial conditions for identifiability, i.e. by choosing experimental initial conditions from a Cartesian basis of n_x linearly independent initial states, the components of the system matrix A can be inferred. This is general not possible when fixing the initial condition. In this case, the N–step experimental design approach can be used. And second, since the length of the parameter bounding intervals depend on the amplitudes c_j and d_l

$$\ell(\lambda_{ij}^A) = \frac{1}{|c_j|}(\overline{w}_i - \underline{w}_i),\ \ell(\lambda_{il}^B) = \frac{1}{|d_l|}(\overline{w}_i - \underline{w}_i),$$

the influence of the disturbance on parameter bounding intervals decreases with increasing amplitudes.

Illustrative Example 4

Reconsider the system given in Ex. 1; to infer the six model parameters, we apply Prop. 6.1, i.e. the proposed $n_x + n_u = 3$ one–step experiments, considering low E and high amplitudes $\overline{E}$. The results are provided in Tab. II.

TABLE II

One–step "learning approach". Parameter bounding intervals $\ell_1, \ldots, \ell_6$ for low (E) and high ($\overline{E}$) intense one–step experiments.

experiment			**bounding intervals**					
#	x_0	u_0	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5	ℓ_6
$E^{(1)}$	$(1, 0)$	0	0.2	1	0.2	1	1	1
$E^{(2)}$	$(0, 1)$	0	1	0.2	1	0.2	1	1
$E^{(3)}$	$(0, 0)$	1	1	1	1	1	0.2	0.2
$\overline{E}^{(1)}$	$(10, 0)$	0	0.02	1	0.02	1	1	1
$\overline{E}^{(2)}$	$(0, 10)$	0	1	0.02	1	0.02	1	1
$\overline{E}^{(3)}$	$(0, 0)$	10	1	1	1	1	0.02	0.02

Evidently, the six unknown parameters can be deduced from the proposed three (independent) one–step experiments as shown in Tab. II. Furthermore, considering experiment $E^{(3)}$ as considered also in Ex. 1, we have $\ell_5^1 \leq 0.2$ and $\ell_6^1 \leq 0.2$; this provides proof that for all possible realizations (of experiments as in Ex. 1), λ_5 and λ_6 (compare Fig. 1) are estimated to the extend $\mu \leq 0.2$ (after one–step).

Comparing the two sets of experiments $E^{(i)}$ and $\overline{E}^{(i)}$, $i=\{1,2,3\}$, it is demonstrated that high intense experiments countervail the effects of disturbances, i.e. high ample stimuli provide better estimates.

VII. CONCLUSIONS

In this contribution, we proposed a guaranteed approach for a priori experimental design and identifiability analysis of linear discrete time systems subject to bounded disturbances. Assuming bounded disturbances is, in many practical cases, more realistic and less demanding than a statistical error distribution [17]. It enables to derive the set of consistent parameters and bounding intervals, analog to confidence intervals in statistical error setting. This set is constructed using available prior information and posterior measurements in case of parameter estimation; then, the consistent parameter set is a convex polytope. For experimental design, it defines a family of polytopes, where the worst case volume provides a guaranteed upper bound, used as selection criterion.

When investigating the insightful one–step ahead case, the role of initial conditions, inputs and their scaling is outlined. As shown, $n_x + n_u$ one–step experiments are sufficient to identify all parameter of a fully parametrized system, where the influence of the disturbance on the estimates can be decreased by increasing the respective amplitudes. This is provided when the initial conditions can be manipulated freely, i.e. a basis of linear independent initial state vectors can be considered for design of experiments.

Future work will address extension of the guaranteed approach to input–output systems and to polynomial systems using relaxations [7], and experimental design for purpose of model selection.

REFERENCES

[1] Z. Artstein and S. Raković. Set invariance under output feedback: a set-dynamics approach. *International Journal of Systems Science*, 42(4):539–555, 2011.

[2] Z. Artstein and S.V. Raković. Feedback and invariance under uncertainty via set–iterates. *Automatica*, 44(2):520–525, 2008.

[3] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete and computational geometry*, 8(1):295–313, 1992.

[4] E.W. Bai, Y. Ye, and R. Tempo. Bounded error parameter estimation: A sequential analytic center approach. *IEEE Trans. Aut. Cont.*, 44:1107–1117, 1999.

[5] R. Bellman and KJ Åström. On structural identifiability. *Mathematical Biosciences*, 7(3-4):329–339, 1970.

[6] S. Borchers, S.V. Raković, and R. Findeisen. Set membership parameter estimation and design of experiments using homothety. In *18th IFAC World Congress*, 2011.

[7] S. Borchers, P. Rumschinski, S. Bosio, R. Weismantel, and R. Findeisen. A set-based framework for coherent model invalidation and parameter estimation of discrete time nonlinear systems. In *Proc. 48th Conf. on Dec. and Cont. (CDC 2009), Shanghai, China*, pages 6786–92, 2009.

[8] S. Boyd, S.J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, 2007.

[9] V. Cerone and D. Regruto. Parameter bounds for discrete-time hammerstein models with bounded output errors. *Automatic Control, IEEE Transactions on*, 48(10):1855–1860, 2003.

[10] V. Cerone and D. Regruto. Parameter bounds evaluation of wiener models with noninvertible polynomial nonlinearities. *Automatica*, 42(10):1775–1781, 2006.

[11] K.J. Falconer and J. Wiley. *Fractal geometry: mathematical foundations and applications*. Wiley New York, 2003.

[12] E. Fogel and Y.F. Huang. On the value of information in system identification–Bounded noise case 1. *Automatica*, 18(2):229–238, 1982.

[13] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[14] L. Ljung. *System identification. Theory for the user*. Prentice Hall, 2nd edition, 1998.

[15] T.H. Mattheiss. An algorithm for determining irrelevant constraints and all vertices in systems of linear inequalities. *Operations Research*, 21(1):247–260, 1973.

[16] M. Milanese and G. Belforte. Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators. *IEEE Trans. Aut. Cont.*, 27:408–414, 1982.

[17] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27(6):997–1009, 1991.

[18] S.H. Mo and J.P. Norton. Fast and robust algorithm to compute exact polytope parameter bounds. *Mathematics and Computers in Simulation*, 32(5-6):481–493, 1990.

[19] C.R. Rojas, J.S. Welsh, G.C. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43(6):993–1008, 2007.

[20] P. Rumschinski, S. Borchers, S. Bosio, R. Weismantel, and R. Findeisen. Set-based dynamical parameter estimation and model invalidation for biochemical reaction networks. *BMC Sys. Biol.*, 4:69, 2010.

[21] R. Schneider. *Convex bodies: the Brunn–Minkowski theory*. Cambridge University Press, 1993.

[22] F. Schweppe. Recursive state estimation: unknown but bounded errors and system inputs. *IEEE Trans. Aut. Cont.*, 13(1):22–28, 1968.

[23] F.C. Schweppe. *Uncertain Dynamic Systems*. Prentice Hall, Engelwood Cliffs, NJ, 1973.

[24] L.N. Vicente and P.H. Calamai. Bilevel and multilevel programming: A bibliography review. *Journal of Global optimization*, 5(3):291–306, 1994.

[25] E. Walter and H. Piet-Lahanier. Exact recursive polyhedral description of the feasible parameter set for bounded-error models. *IEEE Trans. Aut. Cont.*, 34:911–915, 1989.

[26] E. Walter and H. Piet-Lahanier. Estimation of parameter bounds from bounded-error data: a survey. *Mathematics and Computers in Simulation*, 32(5-6):481–493, 1990.

[27] H. Witsenhausen. Sets of possible states of linear systems given perturbed observations. *IEEE Trans. Aut. Cont.*, 13(5):556–558, 1968.

APPENDIX

Proof 1 (Prop. 3.1): Let $\Theta_k, x_{k+1}, x_k, u_k$ be given. By (6)–(7) we have $\Theta_{k+1} = F(\Theta_k, x_{k+1}, x_k, u_k)$, with

$$\begin{aligned} F(\Theta_k, x_{k+1}, x_k, u_k) &= \{\lambda \in \Theta : x_{k+1} - Y(x_k, u_k)\lambda \in W\} \\ &= \{\lambda \in \Theta_k : x_{k+1} - Y(x_k, u_k)\lambda = w_k,\ w_k \in W\} \\ &= \{\lambda \in \Theta_k : x_{k+1} = A(\lambda)x_k + B(\lambda)u_k + w_k,\ w_k \in W\}. \end{aligned}$$

Since $\Theta_0 \doteq \Lambda$, it follows that $\Theta_{k+1} = F(\Theta_k, x_{k+1}, x_k, u_k)$ generates the desired sequence $\{\Theta_k\}_{k=1}^{N}$ of the consistent parameter sets. □

Proof 2 (Prop. 3.2): Pick a $j \in \{0, 1, \dots, j, \dots, N-1\}$ and assume that $\Theta_j = \{\lambda \in \Lambda : M_j\lambda \leq l_j\}$. Then, by Prop. 3.1,

$$\begin{aligned} \Theta_{j+1} &= F(\Theta_j, x_{j+1}, x_j, u_j) \\ &= \{\lambda \in \Theta_j : x_{j+1} - A(\lambda)x_j - B(\lambda)u_j \in W\}. \end{aligned}$$

Hence, from the description of W (3) and Θ_j, we have:

$$\Theta_{j+1} = \{\lambda \in \Theta_j : M_{j+1}\lambda \leq l_{j+1}\}$$

with M_{j+1}, l_{j+1} as in (9). Since $\Theta_0 \doteq \Lambda$, the claim follows by induction. □