2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

A-priori Fisher Information of Nonlinear State Space Models for Experiment Design

Franz Dietrich, Annika Raatz and Jürgen Hesselbach

Abstract— **This article presents advances in optimal experiment design, which are intended to improve the parameter identification of nonlinear state space models. Instead of using a sequence of samples from one or just a few coherent sequences, the idea of identifying nonlinear dynamic models at distinct points in the state space is considered. In this way, the placement of the experiment points is fully flexible with respect to the set of reachable points. Also, a method for model-based generation of prediction errors is proposed, which is used to compute an a-priori estimate of the sample covariance of the prediction error. This covariance matrix may be used to approximate the Fisher information matrix a-priori. The availability of the Fisher matrix a-priori is a prerequisite for experiment optimization with respect to covariance in the parameter estimates.**

This work is driven by the problem of parameter identification of hydraulic models. There are methods for hydraulic systems regarding the estimation of parameters from experimental data, but the choice of experiments has not been treated adequately yet. A hydraulic servo system actuating a stewart platform serves as an illustrative example to which the methods above are applied.

I. Introduction

There is consensus that model-based control of nonlinear actuators improves the control performance of complex systems considerably [1], [2], [3]. Inherently the successful application of such control laws relies heavily on the quality of the models involved. This article adresses the problem of experiment design for parameter identification of nonlinear systems.

In this work hydraulic servo-systems are considered as an exemplary realization for this class of problems. Here, the model structures are well established in the literature. Unfortunately, due to the physics of hydraulics, the most accurate model types are highly nonlinear [4], [5], [6]. Hence, when modelling such a system, one has to choose a model complexity that describes the system as accurate as possible on one hand and that is simple enough so that the parameters can be identified on the other hand. In related research publications that are concerned with the determination of model parameters in practice nonlinear models are used, but only to a certain degree. The models are simplified in such a way that parameters may be estimated with acceptable uncertainty in heuristically designed experiments. For example, model structures may be obtained which are linear w.r.t. their parameters. In such cases the identification problem may be tackled by least squares analysis [6].

The authors are with the Institute of Machine Tools and Production Technology, Technische Universität Braunschweig, 38106 Braunschweig, Germany http://www.iwf.tu-bs.de {F.Dietrich, A.Raatz, J.Hesselbach}@tu-bs.de

The simplifications mainly consist of neglections of nonlinear terms or on experiment strategies that require decoupled experiments in specific test beds. Unfortunately these simplifications decrease the prediction quality of the model considerably, as will be shown in a later section. This effect becomes most important when the system to be controlled is designed for highly dynamic applications what means that it has powerful hydraulic drive systems that move low inertias. Consequently the control performance will suffer from the decreased model quality.

In the following section the class of models that is studied and a suitable estimator is introduced. Also a proposal for an experiment strategy for nonlinear state space models is stated. The design procedure that determines the experiments relies on the optimization of the Fisher matrix and therefore a novel method to obtain this matrix a-priori is explained here in detail. A procedure that is based on the Fisher matrix is chosen and adapted. It may be used to find optimum experiments. In order to illustrate the ideas a model of a hydraulic servo system is presented and a procedure for experiment optimization is proposed.

II. Model Structure and Estimator

A. Model Structure

We consider a continouos state space model structure in the multi-variate case:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\theta}), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{1a}$$

$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}) \tag{1b}$$

The functions $\boldsymbol{f}$ and $\boldsymbol{g}$ may be nonlinear but must be piecewise differentiable with respect to the parameter vector $\boldsymbol{\theta}$ and the state space vector $\boldsymbol{x}$. We assume that a model equation for $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\theta})$ and $\boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta})$ is already set up and that there exists an a-priori estimation of $\boldsymbol{\theta}$. Also a rough estimate of the variance of the parameter estimation error can be made.

B. Experiment Procedure

The task is to find the best estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$. Classically one would carry out an experiment where an input sequence $\boldsymbol{u} = [u(0), \ldots u(kT)]^T$ is given to the system and the reaction of the system is observed via measurement of the output. Out of this output measurements sequences of samples are picked to form the matrix of measured outputs $\hat{\boldsymbol{Y}} = [\hat{\boldsymbol{y}}_1 \ldots \hat{\boldsymbol{y}}_L]^T$. Then an estimator that uses the prediction error matrix $\boldsymbol{W} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}$ may determine the parameter vector.

Usually optimal experiment design for dynamic models is concerned with optimal parametrization of input signals in the time (e.g. pseudo-random binary signal, chirp, inference of multiple sine-functions) or the frequency domain (optimum shaping of random signals in frequency domain) [7]. For nonlinear systems this strategy may not be suitable because establishing criteria and optimizing them to receive an optimum input sequence is likely to fail due to the nonlinearities in the model. Practically that means that finding a single coherent input sequence from which all parameters may be estimated with acceptable precision is rather complicated or even impossible. This problem might be tackled by releasing the prerequisite that the samples in $\hat{\boldsymbol{Y}}$ belong to a coherent experiment. For the realization of this relaxation an experiment is described by the tupel

$$\mathcal{E} = \{\boldsymbol{x}_0, \boldsymbol{u}, \Delta t\}. \tag{2}$$

Consequently, the k-th measurement $\hat{\boldsymbol{y}}_k$ is taken from the k-th experiment, which begins at the start condition $(\boldsymbol{x}_0)_k$. $\hat{\boldsymbol{y}}_k$ is sampled at the time instant $(\Delta t)_k$. During $(\Delta t)_k$ the system is fed by the constant input $(\boldsymbol{u})_k$[1]. This means that each sample $\hat{\boldsymbol{y}}_k$ may be obtained from an individual experiment situation. Consequently $\hat{\boldsymbol{Y}}$ is assembled from the set of experiments $\boldsymbol{\mathcal{E}}$:

$$\boldsymbol{\mathcal{E}} = \{[\mathcal{E}_1, \ldots, \mathcal{E}_L\}. \tag{3}$$

In this way flexibility is added to the choice of the experiments, which means that subsequent samples are no longer tied to obey the system's equations of dynamics. The consequent application of this scheme to a state space model means that L distinct points in the state space of the system (*experiment points*) are picked for parameter identification. Each *experiment point* $\mathcal{E}_k$ represents a short piece of trajectory on which the system travels in a short period of time Δt.

C. Estimation Accuracy

The k-th measurement point $\hat{\boldsymbol{y}}_k$ contains some noise $\boldsymbol{\nu}_k$

$$\hat{\boldsymbol{y}}_k = \boldsymbol{y}_k + \boldsymbol{\nu}_k, \quad k = 1, 2 \ldots L. \tag{4}$$

We consider a maximum likelihood estimator which optimizes the conditional probability density

$$p\left(\hat{\boldsymbol{Y}}|\boldsymbol{U}, \hat{\boldsymbol{\theta}}\right) = \max_{\theta} p\left(\hat{\boldsymbol{Y}}|\boldsymbol{U}, \boldsymbol{\theta}\right). \tag{5}$$

The lower bound of the covariance of an unbiased estimator $\boldsymbol{\theta}$ is given by the Cramér-Rao inequality [7]:

$$\operatorname{cov}(\Delta\boldsymbol{\theta}) = \mathrm{E}\left\{\left[\hat{\boldsymbol{\theta}} - \mathrm{E}\{\hat{\boldsymbol{\theta}}\}\right]\left[\hat{\boldsymbol{\theta}} - \mathrm{E}\{\hat{\boldsymbol{\theta}}\}\right]^T\right\} \geq \boldsymbol{F}^{-1}, \tag{6}$$

where the parameter estimate is denoted by $\hat{\boldsymbol{\theta}}$ and $\mathrm{E}\{\cdot\}$ is the corresponding expected value. Therefore, the Fisher matrix $\boldsymbol{F}$ or its inverse are studied and used to indicate the information which is contained in a given set of experiments [8], [9], [10]. The approximation of the Fisher matrix may be computed upon the equation given in [7] (due to numerical reasons it is not adviseable to compute the exact of the Fisher matrix):

$$\boldsymbol{F} \approx \frac{2}{L}\sum_{k=1}^{L}\left(\frac{d\boldsymbol{w}_k}{d\boldsymbol{\theta}}\right)^T \boldsymbol{D}(\hat{\boldsymbol{\theta}})^{-1}\frac{d\boldsymbol{w}_k}{d\boldsymbol{\theta}}. \tag{7}$$

Based on the relation $\boldsymbol{w}_k = \hat{\boldsymbol{y}}_k - \boldsymbol{y}_k$ the differentiation of $\boldsymbol{w}_k$ with respect to $\boldsymbol{\theta}$ yields

$$\frac{d\boldsymbol{w}_k}{d\boldsymbol{\theta}} = -\frac{d\boldsymbol{y}_k}{d\boldsymbol{\theta}} = -\left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}^T}\frac{d\boldsymbol{x}}{d\boldsymbol{\theta}} + \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\theta}}\right)_k. \tag{8}$$

The state sensitivity $d\boldsymbol{x}/d\boldsymbol{\theta}$ may be computed via integration of the differential of $\dot{\boldsymbol{x}}$

$$\frac{d\dot{\boldsymbol{x}}}{d\boldsymbol{\theta}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}^T}\frac{d\boldsymbol{x}}{d\boldsymbol{\theta}} + \frac{d\boldsymbol{f}}{d\boldsymbol{\theta}}. \tag{9}$$

The initial condition $\boldsymbol{x}_0$ is given by the experiment point $\mathcal{E}_k$, the integration interval is given by $(\Delta t)_k$.

D. A-priori Estimation of Measurement Covariances

The covariance matrix $\boldsymbol{D}(\hat{\boldsymbol{\theta}})$ may be approximated a-posteriori by the sample covariance matrix:

$$\begin{aligned}\boldsymbol{D}(\hat{\boldsymbol{Y}}, \hat{\boldsymbol{\theta}}) &= \frac{1}{L}\sum_{k=1}^{L}\boldsymbol{w}_k\boldsymbol{w}_k^T \\ &= \frac{1}{L}\sum_{k=1}^{L}\left(\hat{\boldsymbol{y}}_k - \boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{u}_k, \hat{\boldsymbol{\theta}}\right)\right)\left(\hat{\boldsymbol{y}}_k - \boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{u}_k, \hat{\boldsymbol{\theta}}\right)\right)^T\end{aligned} \tag{10}$$

There is a dilemma which the experiment designer faces: The sample covariance matrix $\boldsymbol{D}(\hat{\boldsymbol{Y}}, \hat{\boldsymbol{\theta}})$ is not computable until the measurement data from experiments is available. Hence the information that is provided by a given sample may only be computed a-posteriori, which is in contrast to the interest of the experiment designer who requires the information content a-priori. In order to solve this dilemma we propose a method which yields an a-priori estimate of the seeked covariance matrix $\boldsymbol{D}$.

We assume that a rough estimate $\tilde{\boldsymbol{\theta}}$ of the parameter vector is known a-priori (in case nothing is known about $\boldsymbol{\theta}$, only heuristic choices for the experiment design can be made) [11]. Also we assume that the uncertainty in $\tilde{\boldsymbol{\theta}}$ may be approximated by a Gaussian distribution with covariance $\boldsymbol{\sigma}_{\tilde{\boldsymbol{\theta}}}$. For many practical problems the most simple case where $\boldsymbol{\sigma}_{\tilde{\boldsymbol{\theta}}}$ may be reduced to a diagonal matrix is a promising assumption:

$$\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{\theta}}} = \begin{bmatrix}\sigma^2(\tilde{\theta}_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2(\tilde{\theta}_N)\end{bmatrix}. \tag{11}$$

The multivariate probability density that an estimate of the parameter vector is equal to the best possible estimate is then given by

$$p(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) = \left(\sqrt{(2\pi)^N}\prod_{n=1}^{N}\sigma_{\tilde{\theta}_n}\right)^{-1}\exp\left(\sum_{n=1}^{N}\frac{(\theta_n - \tilde{\theta}_n)^2}{2\sigma^2_{\tilde{\theta}_n}}\right). \tag{12}$$

[1]Note: One could also think of the generalization where $\boldsymbol{u}$ is an arbitrary function of time. We choose to restrict the input to a constant value for the sake of simplicity here.

Based on this knowledge the parameter space in the proximity of $\tilde{\boldsymbol{\theta}}$ may be discretized. We propose the discretization in equidistant steps of $\Delta\theta_1, ..\Delta\theta_N$. The number of points wich is distributed over the parameter space of interest is the product of the number of discretization points of each parameter: $R = R_1 \cdot .. \cdot R_N$. Then the r-th discretization point ${}^r\boldsymbol{\theta}$ is assigned to a corresponding probability ${}^rP$.

$$P({}^r\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) = {}^rP = \int_{{}^r\theta_1 - \Delta\theta_1/2}^{{}^r\theta_1 + \Delta\theta_1/2} \cdots \int_{{}^r\theta_N - \Delta\theta_N/2}^{{}^r\theta_N + \Delta\theta_N/2} p(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) d\theta_1 \cdots d\theta_N \tag{13}$$

Alternatively one might prefer to choose discretization in non-equidistant steps, e.g. to receive a more dense coverage of the area near the center of the Gaussian function. Then the integral limits need to be adjusted individually for each discretization point. In order to simplify the notation we will concentrate on the equidistant case.

For each discretization point ${}^r\boldsymbol{\theta}$ the corresponding model output is computed. In order to obtain such meaningful model output, a short piece of the trajectory of $\boldsymbol{x}(t)$ is computed by solving the model eq. (1a) numerically. For this simulation the desired *experiment point* $\mathcal{E}_k$ yields the initial condition $(\boldsymbol{x}_0)_k$, and $(\boldsymbol{u})_k$ is used as constant input vector:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), (\boldsymbol{u})_k, {}^r\boldsymbol{\theta}\right), \quad \boldsymbol{x}(t=0) = (\boldsymbol{x}_0)_k, \quad t = 0 \ldots (\Delta t)_k \tag{14}$$

The model output of this simulation at $t = \Delta t$ is denoted by

$${}^r\boldsymbol{y}(\Delta t) = \boldsymbol{g}\left(\boldsymbol{x}(\Delta t), (\boldsymbol{u})_k, {}^r\boldsymbol{\theta}\right) \tag{15}$$

In the same way the trajectory of $\boldsymbol{x}(t)$ for the parameter estimate $\tilde{\boldsymbol{\theta}}$ is calculated. Here the model output at $t = \Delta t$ is denoted by $\tilde{\boldsymbol{y}}(\Delta t) = \boldsymbol{g}\left(\boldsymbol{x}(\Delta t), \boldsymbol{u}_k, \tilde{\boldsymbol{\theta}}\right)$

Then the sample covariance matrix of the model output under the assumptions stated above may be summed up over the weighed output deviations $({}^r\boldsymbol{y} - \tilde{\boldsymbol{y}})$:

$$\tilde{\boldsymbol{\sigma}}_k^2\left((\boldsymbol{x}_k, \boldsymbol{u}_k), \tilde{\boldsymbol{\theta}}, \Delta t\right) = \frac{1}{\sum_{r=1}^{R} {}^rP} \sum_{r=1}^{R} {}^rP \cdot \left[({}^r\boldsymbol{y} - \tilde{\boldsymbol{y}})({}^r\boldsymbol{y} - \tilde{\boldsymbol{y}})^T\right]_{\Delta t} \tag{16}$$

Similar to the sum of squares of the prediction errors in eq. (10) the *modelled measurement covariances* $\tilde{\boldsymbol{\sigma}}_k^2$ of the L *experiment points* can be summed up to retrieve the *a-priori sample covariance matrix*. This matrix may be treated as an approximate of the sample covariance matrix in eq. (10):

$$\tilde{\boldsymbol{D}}_{\text{a-priori}} = \frac{1}{L} \sum_{k=1}^{L} \tilde{\boldsymbol{\sigma}}_k^2\left((\boldsymbol{x}_k, \boldsymbol{u}_k), \tilde{\boldsymbol{\theta}}\right) \approx \tilde{\boldsymbol{D}}_{\text{a-posteriori}} \tag{17}$$

III. Optimization of *Experiment Points*

The inverse Fisher matrix, which is based on the *experiment points* and the *modelled measurement errors*, expresses the accuracy that a parameter identification based on the corresponding data would deliver. Hence, in order to maximize the precision of the parameter estimates, the inverse Fisher matrix should be minimized. For this purpose multiple criteria are available that derive a quality index of the Fisher matrix. It is quite common to minimize $\det(\boldsymbol{F}^{-1})$ [11], [12], which is in the considered case equivalent to the maximization of

$$\Phi(\boldsymbol{F}) = \det\left(\boldsymbol{F}\left((\boldsymbol{x}_k, \boldsymbol{u}_k), \tilde{\boldsymbol{\theta}}, \Delta t\right)\right). \tag{18}$$

We propose to employ an exchange strategy that contains randomizing elements. Such algorithms are recommendable because they can easily be established although the quality criterion may be nonlinear, discontinuous and may have many local minima. A practical implementation is shown in algorithm 1.

The number of experiment points should be chosen carefully. Each additional experiment point requires considerable extra experimental effort but might not decrease the achieveable parameter uncertainty. In a practical context it is adviseable to perform the optimization with different numbers of experiment points in order to choose the most appropriate number.

IV. Example: Hydraulic Actuator Model

A. Actuator Model

The theoretical ideas presented above are illustrated by the parameter identification problem of a servo-hydraulic positioning cylinder. The experimental platform for our investigations is a hydraulic servo system which is part of a 6-DOF *Stewart Platform* for metal forming purposes (fig. 1)[2]. This system contains six servo-hydraulic cylinders; we pick a generic model of one cylinder as an example. Each cylinder is controlled by a double-stage servo valve. A suitable state space model is given in eq. (19). The model is single-input, multi-output and highly nonlinear in both its input and its parameters[3].

$$\dot{\boldsymbol{x}} = \begin{bmatrix} \dot{q} \\ \ddot{q} \\ \dot{p}_A \\ \dot{p}_B \\ \dot{x}_v \end{bmatrix} = \begin{bmatrix} \dot{q} \\ \frac{1}{m_q} \cdot (A_A \cdot p_A - A_B \cdot p_B - \tau_f) \\ \frac{E_{Oil}}{V_A(q)} \left(-\dot{q} \cdot A_A + Q_A + Q_{L,A}\right) \\ \frac{E_{Oil}}{V_B(q)} \left(\dot{q} \cdot A_B + Q_B + Q_{L,B}\right) \\ -\frac{1}{T_1} \cdot x_v \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{1}{T_1} \end{bmatrix} \cdot u \tag{19a}$$

$$\boldsymbol{y} = \begin{bmatrix} q & p_A & p_B & x_v \end{bmatrix}^T \tag{19b}$$

$$\tau_f = a_\tau \cdot \sigma(\dot{q}) + b_\tau \cdot \dot{q} \tag{19c}$$

[2]For more details we refer to our earlier work [13].

[3]The output does neither satisfy the superposition principle w.r.t. the input nor w.r.t the parameters, cf. [11]

$$Q_A = x_v c_d \sqrt{\frac{2}{\rho}} \sqrt{\Delta p_A} \tag{19d}$$

$$= \underbrace{x_v\, c_d \sqrt{\frac{2}{\rho}}}_{B_v} \sqrt{(p_s - p_A)\,\mathrm{H}(x_v) + (p_A - p_t)\,\mathrm{H}(-x_v)}$$

$$Q_B = -x_v c_d \sqrt{\frac{2}{\rho}} \sqrt{\Delta p_B} \tag{19e}$$

$$= \underbrace{-x_v\, c_d \sqrt{\frac{2}{\rho}}}_{B_v} \sqrt{(p_s - p_B)\,\mathrm{H}(-x_v) + (p_B - p_t)\,\mathrm{H}(x_v)}$$

$$Q_{L,A} = c_L \cdot \left(\sqrt{p_s - p_A} - \sqrt{p_A - p_t} \right) \tag{19f}$$

$$Q_{L,B} = c_L \cdot \left(\sqrt{p_s - p_B} - \sqrt{p_B - p_t} \right) \tag{19g}$$

$$V_A(q) = A_A \cdot (q - q_{min}) + V_{0,A} \tag{19h}$$

$$V_B(q) = A_B \cdot (q_{max} - q) + V_{0,B} \tag{19i}$$

q : Cylinder Position
p_A, p_B : Pressures inside Chambers
x_v : Valve Spool Position
V_A, V_B : Chamber Volume
Q_A, Q_B: Oil Flow into Chambers
u : Valve Control Value
$\mathrm{H}(\cdot)$: Heaviside Function
τ_f : Friction Force
m : Mass attached to Cylinder
A_A, A_B : Piston Areas
E_{Oil} : Oil Modulus
T_1 : Time Constant of Valve
c_d : Valve Coefficient
ρ : Oil Density

In the related research publications it is quite common to simplify the the model in order to reduce the parameter space and the nonlinearities, e.g. [14], [15]. For example, in [16] the change of volume and the tank pressure are neglected. In other contributions the leakage term is neglected, e.g. [4]. Additionally the Heaviside function may be replaced by the tanh(·)-function in order to obtain differentiability in the whole state space. With these simplifications suitable experiments that provide acceptable parameter accuracy may be found heuristically [15], [4], [3].

The motivation for the work presented here is that these nonlinearities shall explicitly be taken into account. In simulations that were compared to experimental data it was found that they increase prediction quality considerably. Figure 2 shows such a comparison. There, the response of the pressure dynamics to a step of u is displayed. The upper plot shows the valve position x_v. In both, the middle and the lower plot, the measurements of the pressures p_A and p_B in the cylinder chambers are displayed (in black, see the vibrations at $t \approx 0.45s$). Additionally the plot in the middle shows the pressures that were simulated with the model of full complexity which is presented in eq. (19). In the lower plot

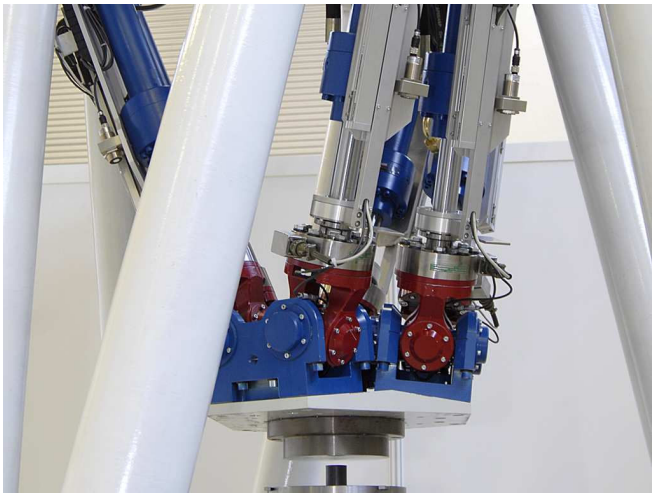

Fig. 1. 6-DOF stewart platform for flexible forming processes.

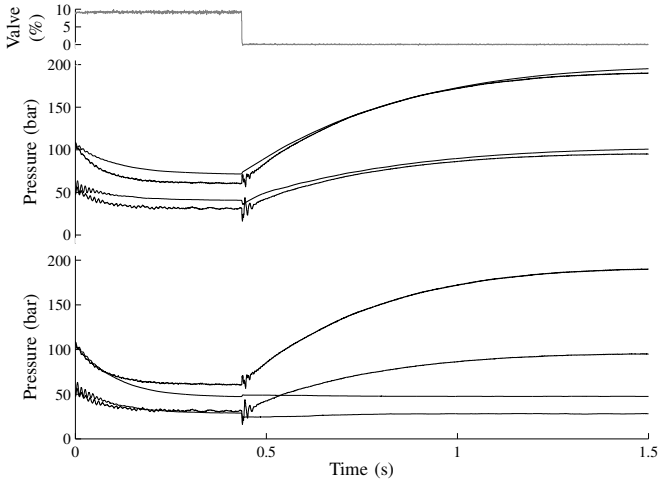

Fig. 2. Measurement of pressures and prediction of complex (center plot) and simplified model (lower plot).

the same simulation is shown, except for the difference that the leakage term was neglected ($c_L = 0$). We are especially interested in the section $t > 0.45s$, where the valve is closed again. It can be seen from the measurements that the pressures of the real system tend from a very low level towards a certain rest state. This tendency is mainly driven by the leakage inside the valve. The model that includes leakage is able to reproduce this effect properly (middle plot), whereas the simplified model (lower plot) does not reproduce this effect: As soon as the valve is closed the simplified model remains in its actual combination of pressures p_A and p_B. This shows that especially in regions where the valve piston position x_v is nearly closed the leakage term becomes important. For model based control this means that in situations where the actuator velocity is low the control values will be distorted by errors of considerable magnitude.

Algorithm 1 Optimization of Experiment Points

1: **for** $k = 1$ to L **do**
2: Create experiment point $\mathcal{E}_k = \{\boldsymbol{x}_0, \boldsymbol{u}\}_k$
3: Calculate $\frac{d\boldsymbol{w}(\mathcal{E}_k, \tilde{\boldsymbol{\theta}})}{d\boldsymbol{\theta}}$
4: Calculate $\tilde{\boldsymbol{\sigma}}^2\left(\mathcal{E}_k, \tilde{\boldsymbol{\theta}}\right)$
5: **repeat**
6: **repeat**
7: **for** $m = 1$ to M **do**
8: Create candidate $\mathcal{E}_m^{(C)} = \{\boldsymbol{x}_{0,}, \boldsymbol{u}, \Delta t\}_{\mathcal{E}_m^{(C)}}$
9: Calculate $\frac{d\boldsymbol{w}(\mathcal{E}_m^{(C)}, \tilde{\boldsymbol{\theta}})}{d\boldsymbol{\theta}}$
10: Calculate $\tilde{\boldsymbol{\sigma}}^2\left(\mathcal{E}_m^{(C)}, \tilde{\boldsymbol{\theta}}\right)$
11: $\mathcal{E}_{L+1} = \arg \min_{\mathcal{E}_1^{(C)}, .. \mathcal{E}_M^{(C)}} \begin{bmatrix} \Phi\left(\boldsymbol{F}\left(\{\mathcal{E}_1, ..\mathcal{E}_L, \mathcal{E}_1^{(C)}\}\right)\right), \dots \\ \Phi(\boldsymbol{F}\left(\{\mathcal{E}_1, ..\mathcal{E}_L, \mathcal{E}_M^{(C)}]\}\right) \end{bmatrix}$
12: **until** $\Phi\left(\boldsymbol{F}\left(\{\mathcal{E}_1, ..\mathcal{E}_L, \mathcal{E}_{L+1}\}\right)\right) > \Phi\left(\boldsymbol{F}\left(\{\mathcal{E}_1, ..\mathcal{E}_L\}\right)\right)$
13: $k = \arg \min_{n=1..L+1} \left(\Phi\left(\boldsymbol{F}(\{\mathcal{E}_1, ..\mathcal{E}_{L+1}\} \setminus \mathcal{E}_n)\right)\right)$
14: $\mathcal{E}_k = \mathcal{E}_{L+1}$
15: **until** exitcondition== **true**

B. Implementation of the Fisher Matrix and Proposal for Experiment Optimization

The vector of parameters to be estimated is

$$\boldsymbol{\theta} = \begin{bmatrix} a_\tau & b_\tau & m_q & E_{Oil} & p_t & B_v & c_L & V_{0,A} & V_{0,B} \end{bmatrix}. \quad (20)$$

It is assumed that the remaining parameters in eq. (19) are known due to their physical nature, e.g. piston areas that are known with low uncertainty from data sheets. It can be deduced from the model eqs. (19a) and (19b) that the identification problem may be decoupled into two identification problems ($\boldsymbol{\theta}_1 = [a_\tau, b_\tau, m_q]$, $\boldsymbol{\theta}_2 = [E_{Oil}, p_t, B_v, c_L, V_{0,A}, V_{0,B}]$). This is due to the presence of p_A, p_B in $\boldsymbol{y}$, which allows the pressure dynamics to be decoupled from the inertial dynamics. In order to calculate the Fisher matrix, eq. (19) requires differentiation with respect to $\boldsymbol{\theta}$. Points of the differentials that are not defined or indefinite, e.g.

$$\left.\frac{\partial \mathrm{H}(x)}{\partial x}\right|_{x=0}, \quad (21)$$

are removed from the set of possible *experiment points*.

For the optimization of $\boldsymbol{\mathcal{E}}$ we propose an exchange algorithm, which is similar to the DETMAX algorithm described in [11], see algorithm 1. Note, the computational capacity limits the choice of L, M and R (e.g. eq. (19a) is evaluated $2R \cdot M$ times per iteration, see step 4 and 10).

V. Conclusion and Future Perspectives

In this article we present advances in optimal experiment design for parameter identification of nonlinear state space models. Current design methods for the optimization of experiments with respect to parameter accuracy are not sufficiently capable of treating complex nonlinear models of dynamic systems. These circumstances are illustrated by a model for servo-hydraulic positioning cylinders for precise tracking control. We consider the idea of identifying nonlinear dynamic models at distinct points in the state space instead of using subsequent samples in one coherent experiment. Once these points are determined one experiment is carried out per sample. This way the placement of the experiment points is fully flexibile within the set of reachable points.

Based on the samples taken at these experiment points, the Fisher matrix gives an a-posteriori measure of the accuracy of the estimate. We present a method for the a-priori modelling of error covariances at the experiment points of interest. With this method the Fisher matrix becomes available a-priori what provides the opportunity of optimization of the experiments involved. A hydraulic servo system serves as an illustrative example. We propose a simple exchange algorithm with randomizing elements. Though being able to cope with nonlinearities, the effort for the implementation is rather small. Ongoing work is occupied with improvements of the convergence and a comparison against more sophisticated optimization methods.

References

[1] M. R. Sirouspour and S. E. Salcudean, "Nonlinear control of a hydraulic parallel manipulator," in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 4, 2001, pp. 3760–3765.

[2] J. Moreno, R. Kelly, and R. Campa, "Manipulator velocity control using friction compensation," *IEE Proceedings D Control Theory & Application*, no. 150(2), pp. 119–126, 2003.

[3] I. Davliakos and E. Papadopoulos, "Model-based control of a 6-dof electrohydraulic Stewart-Gough platform," *Mechanism and Machine Theory*, vol. 43, no. 11, pp. 1385–1400, 2008.

[4] G. A. Sohl and J. E. Bobrow, "Experiments and Simulations of the Nonlinear Control of a Hydraulic Servosystem," *IEEE Transactions on Control System Technology*, vol. 7, no. 2, pp. 238–247, 1999.

[5] O. Becker, I. Pietsch, and J. Hesselbach, "Robust Task-Space Control of Hydraulic Robots," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2003, pp. 4360–4365.

[6] M. K. A. Jelali, *Hydraulic servo-systems: Modelling, identification and control*. London: Springer, 2003.

[7] G. C. Goodwin and R. L. Payne, *Dynamic system identification: Experiment design and data analysis*, ser. Mathematics in science and engineering. New York, NY: Academic Press, 1977, vol. 136.

[8] L. Ljung, *System identification: Theory for the user*, 2nd ed., ser. Prentice Hall information and system sciences series. Upper Saddle River, NJ: Prentice Hall PTR, 2007.

[9] Y. Chu and J. Hahn, "Integrating parameter selection with experimental design under uncertainty for nonlinear dynamic systems," *AIChE J.*, vol. 54, no. 9, pp. 2310–2320, 2008.

[10] F. Galvanin, M. Barolo, and F. Bezzo, "Online model-based redesign of experiments for parameter estimation in dynamic systems," *Ind. Eng. Chem. Res.*, vol. 48, no. 9, pp. 4415–4427, 2009.

[11] E. Walter and L. Pronzato, *Identification of parametric models from experimental data*, ser. Communications and control engineering series. London: Springer, 1997.

[12] D. Faller, U. Klingmuller, and J. Timmer, "Simulation Methods for Optimal Experimental Design in Systems Biology," *SIMULATION - Trans. The Soc. for Modeling and Simulation*, vol. 79, no. 12, pp. 717–725, 2003.

[13] J. Hesselbach, B.-A. Behrens, F. Dietrich, S. Rathmann, and J. Poelmeyer, "Flexible forming with hexapods," *Production Engineering*, vol. 1, no. 4, pp. 429–436, 2007.

[14] D. Li and S. E. Salcudean, "Modeling, simulation, and control of a hydraulic Stewart platform," in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 4, 1997, pp. 3360–3366.

[15] T. Wey, "Modeling and observer design for hydraulic cylinders," in *6th IEEE Mediterranean Conference on Control and Systems*, 1998.

[16] W. Bernzen, T. Wey, and B. Riege, "Nonlinear control of hydraulic differential cylinders actuating a flexible robot," in *Proc. IEEE Conf. on Decision and Control*, vol. 2, 1997, pp. 1333–1334.