The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Numerical Aspects regarding the A-priori Fisher Information of Nonlinear Models for Hydraulic Servo-Systems

Franz Dietrich, Annika Raatz and Jürgen Hesselbach

Abstract— **This article adresses aspects of the a-priori computation of the Fisher information matrix. The aim of the type of Fisher matrix discussed is the experiment design for parameter estimation of nonlinear state space models. A definition for the sampling of experiments is stated, which is convenient for software implementation. Both, measurement samples that belong to a coherent sequence *and* samples from independent experiments, may be stored together in one single data structure. Since the covariance matrix of the prediction error is required in the Fisher matrix, a method for its a-priori prediction is proposed.**

A servo-hydraulic positioning system serves as a validation example. The comparison of the a-priori and a-posteriori Fisher matrices of this example show that the method proposed is suitable to provide an approximation of an optimality criterion for experiment design.

I. Introduction

This work focuses on hydraulic servo-systems, where non-linear control has become quite common (e.g. [1]). There, the model structures are well established [2], [3]. Unfortunately, the most accurate models are highly nonlinear, which is due to the physics of hydraulics [4], [5]. Hence, 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 nonlinear models are used in practice, but only to a certain level. The models are simplified which allows the experiments to be designed by heuristic methods. 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].

Especially hydraulic systems with large valve cross sections compared to the cylinder piston diameter are vulnerable to simplifications. If the dynamics of the system are also influenced by the stiffness of the fluid transmission lines (e.g. when hoses are used), such systems must be identified in assembled configuration. Unfortunately, there is no method to design experiments for such systems adequately. This article contributes numerical considerations, which are required for the a-priori evaluation of experiments. From the results obtained here, a quality measure of the set of experiments can be computed. An optimization of this set may then take place to find the experiments wich shall finally be conducted for parameter estimation.

In the following section this article proposes a definition for experiment sampling. Based on this definition the a-priori calculation of the Fisher matrix is elaborated. The a-priori computation of the covariance matrix of the prediction error is studied, and issues of its implementation are treated in detail. Finally, the theoretical propositions and results are validated for a servo-hydraulic positioning system in order to identify practical recommendations.

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`

II. Model Structure, Parameter Estimation and Estimation accuracy

A. Model Structure

Consider a continuous state space model structure:

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

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

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 $\boldsymbol{x}$. It is assumed 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 chosen and that there exists an a-priori estimate $\tilde{\boldsymbol{\theta}}$ of the parameter vector.

B. Experiment Sampling

The task is to find the best estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$. It is assumed that L samples $\hat{\boldsymbol{y}}$ are taken from experiments which form the matrix of measured outputs $\hat{\boldsymbol{Y}} = [\hat{\boldsymbol{y}}[1] \dots \hat{\boldsymbol{y}}[L]]^T$. An estimator that uses the prediction error matrix $\boldsymbol{W} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}$ may determine the parameter vector then.

1) Definition of the Set of Experiments: The estimation problem can be considered as a model following problem. There, classically, the samples in $\hat{\boldsymbol{Y}}$ are taken from a coherent set of samples. This means that the model output $\boldsymbol{y}[l]$ is computed in discrete steps in the following way, given the initial condition $\boldsymbol{x}_0$ and the input sequence $\boldsymbol{u}[l]$:

$$\boldsymbol{x}[1] = \boldsymbol{x}[0] + \int_0^{\Delta T} \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}[1], \boldsymbol{\theta})\, dt, \tag{2}$$

$$\boldsymbol{y}[1] = \boldsymbol{g}(\boldsymbol{x}[1], \boldsymbol{u}[1], \boldsymbol{\theta})$$

$$\vdots$$

$$\boldsymbol{x}[l+1] = \boldsymbol{x}[l] + \int_{(l)\Delta T}^{(l+1)\Delta T} \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}[l]\boldsymbol{\theta})\, dt, \tag{3}$$

$$\boldsymbol{y}[l] = \boldsymbol{g}(\boldsymbol{x}[l], \boldsymbol{u}[l], \boldsymbol{\theta})$$

This means that subsequent samples are tied to the constraints of the dynamics of the system. For non-linear systems this restriction might not be favorable: Finding a single

978-1-4244-6676-4/10/$25.00 ©2010 IEEE

coherent input sequence from which all parameters may be estimated with acceptable precision is rather complicated or even impossible in this case. It is hence proposed to release the prerequisite that the samples in $\hat{\boldsymbol{Y}}$ must belong to a coherent set of experiments. For the realization of this relaxation an experiment is described by the tupel

$$\Xi = \{\boldsymbol{x}_0, \boldsymbol{u}, \Delta t\}. \tag{4}$$

Consequently, the l-th measurement $\hat{\boldsymbol{y}}[l]$ is obtained from the l-th experiment, which begins at the start condition $\boldsymbol{x}_0[l]$. $\hat{\boldsymbol{y}}[l]$ is sampled at the time instant $\Delta t[l]$. During $\Delta t[l]$ the system is fed by the constant input $\boldsymbol{u}[l]$. $\hat{\boldsymbol{Y}}$ is assembled from the set of experiments $\boldsymbol{\Xi}$ consequently:

$$\boldsymbol{\Xi} = \left\{\Xi[1], \ldots, \Xi[L]\right\}. \tag{5}$$

In this way subsequent samples in $\hat{\boldsymbol{Y}}$ are no longer tied to obey the system's equations of dynamics.

For practical reasons the definition of Ξ is extended so that coherent sets of experiments are treatable. Each experiment is augmented by a *coherency flag* $c \in 0, 1$. Its value is 1, when the experiment shall be treated coherent with the previous one, or 0 if not. If $c[l] = 1$, the initial condition of $\Xi[l]$ is equivalent to the final state of the previous experiment:

$$\Xi[l] = \{\boldsymbol{x}_0[l] = (\boldsymbol{x}(\Delta t[l-1]))[l-1], \boldsymbol{u}[l], \Delta t[1], c[l]\}. \tag{6}$$

This means that in a coherent sequence the initial conditions are defined iteratively along the constraints of $\boldsymbol{f}$, starting from the initial condition of the first experiment that is marked coherent. The definition in eq. (6) allows $\boldsymbol{\Xi}$ to contain both, pieces of coherent and pieces of non-coherent experiments. This definition helps to simplify the implementation of experiment design and parameter identification. Data manipulations may be formulated compactly. For example, $\boldsymbol{\Xi}$ may be concatenated conveniently from subsets $\boldsymbol{\Xi}_{S1}$, $\boldsymbol{\Xi}_{S2} \ldots$ of experiments where each subset is either exactly one coherent sequence or one single experiment. This is a convenient piece of infrastructure for the choice of experiments and for parameter identification.

C. A-priori Estimation of the Fisher Information Matrix

The l-th measurement point $\hat{\boldsymbol{y}}[l]$ contains some noise $\boldsymbol{v}[l]$

$$\hat{\boldsymbol{y}}[l] = \boldsymbol{y}[l] + \boldsymbol{v}[l], \quad l = 1, 2 \ldots L. \tag{7}$$

A maximum likelihood estimator is considered, which optimizes the conditional probability density

$$p\left(\hat{\boldsymbol{Y}}|\boldsymbol{U}, \hat{\boldsymbol{\theta}}\right) = \max_{\theta} p\left(\hat{\boldsymbol{Y}}|\boldsymbol{U}, \boldsymbol{\theta}\right). \tag{8}$$

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

$$\operatorname{cov}(\Delta\boldsymbol{\theta}) = \mathrm{E}\left\{\left[\hat{\boldsymbol{\theta}} - \mathrm{E}\{\hat{\boldsymbol{\theta}}\}\right]\left[\hat{\boldsymbol{\theta}} - \mathrm{E}\{\hat{\boldsymbol{\theta}}\}\right]^T\right\} \geq \boldsymbol{F}^{-1}, \tag{9}$$

where the parameter estimate is denoted by $\hat{\boldsymbol{\theta}}$ and $\mathrm{E}\{\cdot\}$ is the 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]. Goodwin [7] states an approximation of the Fisher matrix:

$$\boldsymbol{F} \approx \frac{2}{L}\sum_{l=1}^{L}\left(\frac{d\boldsymbol{w}[l]}{d\boldsymbol{\theta}}\right)^T \boldsymbol{D}(\hat{\boldsymbol{\theta}})^{-1}\frac{d\boldsymbol{w}[l]}{d\boldsymbol{\theta}}. \tag{10}$$

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

$$\frac{d\boldsymbol{w}[l]}{d\boldsymbol{\theta}} = -\frac{d\boldsymbol{y}[l]}{d\boldsymbol{\theta}} = -\left(\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}^T}\frac{d\boldsymbol{x}}{d\boldsymbol{\theta}} + \frac{\partial\boldsymbol{g}}{\partial\boldsymbol{\theta}}\right)[l]. \tag{11}$$

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

$$\frac{d\dot{\boldsymbol{x}}}{d\boldsymbol{\theta}} = \frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}^T}\frac{d\boldsymbol{x}}{d\boldsymbol{\theta}} + \frac{d\boldsymbol{f}}{d\boldsymbol{\theta}}. \tag{12}$$

The initial condition $\boldsymbol{x}_0$ is given by the experiment point $\Xi[l]$, the integration interval is given by $\Delta t[l]$.

D. A-priori Estimation of Prediction Error Covariance

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

$$\begin{aligned}\boldsymbol{D}(\hat{\boldsymbol{Y}}, \hat{\boldsymbol{\theta}}) &= \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{w}[l]\boldsymbol{w}[l]^T \\ &= \frac{1}{L}\sum_{l=1}^{L}\left(\hat{\boldsymbol{y}}[l] - \boldsymbol{g}\left(\boldsymbol{x}[l], \boldsymbol{u}[l], \hat{\boldsymbol{\theta}}\right)\right)\left(\hat{\boldsymbol{y}}[l] - \boldsymbol{g}\left(\boldsymbol{x}[l], \boldsymbol{u}[l], \hat{\boldsymbol{\theta}}\right)\right)^T.\end{aligned} \tag{13}$$

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 a method which yields an a-priori estimate of the covariance matrix $\boldsymbol{D}$ is proposed here.

A rough estimate $\tilde{\boldsymbol{\theta}}$ of the parameter vector is required a-priori (in case nothing is known about $\boldsymbol{\theta}$, heuristic choices from any available experiment data must be made) [11]. Also, the assumption is made that uncertainty in $\tilde{\boldsymbol{\theta}}$ may be approximated by a Gaussian distribution with covariance $\boldsymbol{\sigma}_{\tilde{\boldsymbol{\theta}}}$. Initially a simple choice is made, where $\boldsymbol{\sigma}_{\tilde{\theta}}$ may be reduced to a diagonal matrix:

$$\boldsymbol{\sigma}^2_{\tilde{\theta}} = \begin{bmatrix}\sigma^2(\tilde{\theta}_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2(\tilde{\theta}_N)\end{bmatrix}. \tag{14}$$

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

$$p(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) = \left(\sqrt{(2\pi)^N}\prod_{n=1}^{N}\sigma_{\tilde{\theta}_n}\right)^{-1}\exp\left(\sum_{n=1}^{N}\frac{(\theta_n - \tilde{\theta}_n)^2}{2\sigma^2_{\tilde{\theta}_n}}\right). \tag{15}$$

Based on this knowledge the parameter space in the proximity of $\tilde{\boldsymbol{\theta}}$ may be discretized. In this article, the discretization in equidistant steps of $\Delta\theta_1, ..\Delta\theta_N$ is considered. 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:

$$R = \prod_{n=1}^{N} R_n. \tag{16}$$

Then the r-th discretization point ${}^r\boldsymbol{\theta}$ is assigned to a corresponding probability ${}^rP$:

$$P({}^r\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) = {}^rP = \int\limits_{{}^r\theta_1 - \Delta\theta_1/2}^{{}^r\theta_1 + \Delta\theta_1/2} \cdots \int\limits_{{}^r\theta_N - \Delta\theta_N/2}^{{}^r\theta_N + \Delta\theta_N/2} p(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}, \sigma_{\boldsymbol{\theta}}) d\theta_1 \cdots d\theta_N \tag{17}$$

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, this article will concentrate on the equidistant case.

For each discretization point ${}^r\boldsymbol{\theta}$ the corresponding output of the model is computed (for convenience, this model will be called "detuned". The vector $({}^r\boldsymbol{\theta}_n - \tilde{\boldsymbol{\theta}}_n)/\tilde{\boldsymbol{\theta}}_n$ will be referred to as the *detuning vector*). 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 $\Xi[l]$ yields the initial condition $\boldsymbol{x}_0[l]$, and $\boldsymbol{u}[l]$ is used as constant input:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), (\boldsymbol{u})[l], {}^r\boldsymbol{\theta}\right), \quad \boldsymbol{x}(t=0) = \boldsymbol{x}_0[l], \quad t = 0 \ldots \Delta t[l] \tag{18}$$

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

$${}^r\boldsymbol{y}(\Delta t) = \boldsymbol{g}\left(\boldsymbol{x}(\Delta t), \boldsymbol{u}[l], {}^r\boldsymbol{\theta}\right) \tag{19}$$

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 $\tilde{\boldsymbol{y}}(\Delta t) = \boldsymbol{g}\left(\boldsymbol{x}(\Delta t), \boldsymbol{u}[l], \tilde{\boldsymbol{\theta}}\right)$

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

$$\tilde{\boldsymbol{\sigma}}^2[l]\left((\boldsymbol{x}[l], \boldsymbol{u}[l]), \tilde{\boldsymbol{\theta}}, \Delta t\right) = \frac{1}{\sum_{r=1}^{R} {}^rP} \sum_{r=1}^{R} {}^rP \cdot \left[({}^r\boldsymbol{y} - \tilde{\boldsymbol{y}})({}^r\boldsymbol{y} - \tilde{\boldsymbol{y}})^T\right]_{\Delta t} \tag{20}$$

This expression gives some insight into the prediction error that will be achieved by $\Xi[l]$. Similar to the sum of squares of the prediction errors which are calculated in eq. (13), the *modelled measurement covariances* $\tilde{\boldsymbol{\sigma}}^2[l]$ of the L experiment points can be summed up to retrieve the *a-priori sample covariance matrix of prediction error*. This matrix may be treated as an approximate of the sample covariance matrix in eq. (13):

Algorithm 1 Lazy calculation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ (non-coherent sequence)

Require: Ξ and LazyThreshold $= 0 \ldots 1$
for $l = 1$ to L **do**
  $\boldsymbol{D}[l] = \boldsymbol{0}; \sum P = 0$
  $\tilde{\boldsymbol{y}}[l] = \text{solve}(\boldsymbol{f}, \tilde{\boldsymbol{\theta}}, \Xi[l])$ {Nominal model}
  for $r = 1$ to R **do**
    if rand() < LazyThreshold **then**
      ${}^r\boldsymbol{y} = \text{solve}(\boldsymbol{f}, {}^r\boldsymbol{\theta}, \Xi[l])$ {Detuned model}
      $\boldsymbol{D}[l] = \boldsymbol{D}[l] + (\tilde{\boldsymbol{y}} - {}^r\boldsymbol{y}) \cdot (\tilde{\boldsymbol{y}} - {}^r\boldsymbol{y})^T \cdot {}^rP$
      $\sum P = \sum P + {}^rP$
  $\tilde{\boldsymbol{D}}_{\text{a-priori}} = \boldsymbol{D}[l] \cdot (\sum P)^{-1}$
return $\tilde{\boldsymbol{D}}_{\text{a-priori}}$

$$\tilde{\boldsymbol{D}}_{\text{a-priori}} = \frac{1}{L} \sum_{l=1}^{L} \tilde{\boldsymbol{\sigma}}^2[l]\left((\boldsymbol{x}[l], \boldsymbol{u}[l]), \tilde{\boldsymbol{\theta}}\right) \approx \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}) \tag{21}$$

$\tilde{\boldsymbol{D}}_{\text{a-priori}}$ is in the unit of measurement of $\boldsymbol{y}\boldsymbol{y}^T$. Equation (21) contains the model prediction error only, no sensor noise. If desired, it can be augmented by a covariance matrix of sensor noise $\boldsymbol{\nu}_s$. $\boldsymbol{\nu}_s[l]$ contributes to $\boldsymbol{\nu}[l]$ in eq. (1a) and acts additively on the model output, which is considered in the following way:

$$\left(\tilde{\boldsymbol{D}}_{\text{a-priori}}\right)_{+\boldsymbol{\nu}_s} = \tilde{\boldsymbol{D}}_{\text{a-priori}} + \text{cov}(\boldsymbol{\nu}_s, \boldsymbol{\nu}_s) \tag{22}$$

III. Computational Considerations regarding $\tilde{\boldsymbol{D}}_{\text{a-priori}}$

The implementation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ must maximize code efficiency: $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ is intended to be a part of the optimization of experiments. Then, $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ needs to be recalculated whenever a change is made to $\boldsymbol{\Xi}$. The major amount of computational effort is spent on solving the model equation $\boldsymbol{f}$: One calculation run of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ requires $(R+1) \cdot L$ calls of the solving function of $\boldsymbol{f}$. If an exchange algorithm shall be used, it is recommendable to buffer the $\tilde{\boldsymbol{\sigma}}^2[l]$. In that way replacing an experiment means the subtraction of $\tilde{\boldsymbol{\sigma}}^2$ of the experiment to be replaced, and calculation of $\tilde{\boldsymbol{\sigma}}^2$ of the new experiment (which requires $\boldsymbol{f}$ to be solved R times).

A. Lazy Calculation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$

In order to reduce the computational effort required for $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ *lazy computation* is proposed. This method relies on the *law of large numbers* [12]. This law states that the average of the results obtained from a large number of trials should be close to the expected value. The application to $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ means that the solutions of the detuned models $({}^r\boldsymbol{\theta})$ can be considered as trials. The expected value mentioned in the law corresponds to $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ then. Since R may be big enough to represent a large number[1], a large number of detuned models must be solved. It is proposed here to skip some of these trials, what means that some calls to

[1]The parameter space may be of high dimension, and hence $R \gg 1000$ is not uncommon.

Algorithm 2 Lazy calculation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ (coherent sequence)

```
Require: Ξ and LazyThreshold = 0 ... 1
D̃_a-priori = 0;
for r = 1 to R do
  if rand() < LazyThreshold then
    for l = 1 to L do
      if l = 1 then
        Ξ = {(u)[l], x_0[l], T[l]}
      else
        Ξ = {(u)[l], x̃[l − 1], T[l]}
      ỹ[l] = solve(f, θ̃, Ξ)                 {Nominal model}
      ʳy = solve(f, ʳθ, Ξ[l])               {Detuned model}
      Σ(w[l]w[l]^T) = Σ(w[l]w[l]^T) + (ỹ − ʳy) · (ỹ − ʳy)^T
    D̃_a-priori = D̃_a-priori + Σ(w[l]w[l]^T) · ʳP
    ΣP = ΣP + ʳP
  D̃_a-priori = D̃_a-priori · (ΣP)^−1
return D̃_a-priori
```

the solving procedure of $\boldsymbol{f}$ are skipped. Then $\sum\tilde{\boldsymbol{\sigma}}^2$ must be normalized by $\sum{}^rP$, because $\sum{}^rP \neq 1$ (cf. eq. (21)). In this way the lazy computation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ will tend towards the full $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ matrix.

Algorithm 1 shows a proposal for non-coherent $\boldsymbol{\Xi}$. There, some experiments of $\boldsymbol{\Xi}$ are skipped, but all detuned models are involved. The ratio of skipped over not skipped trials is determined by *LazyThreshold*. The normalization by $\sum{}^rP$ can be seen in step 14.

If $\boldsymbol{\Xi}$ is coherent, alg. 1 does not hold: Because coherent experiments are solved iteratively, it is not possible to skip some of the intermediate experiments. Hence, we propose to implement "lazy" computation in this case by skipping some ${}^r\boldsymbol{\theta}$. In this way, the coherent sequence of all Ξ in $\boldsymbol{\Xi}$ is solved completely, but some ${}^r\boldsymbol{\theta}$ are dropped. This variant is shown in alg. 2.

Both, algs. 1 and 2, have the same effect on the computational effort of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$. Algorithm 1 preserves the property that an individual $\sum\tilde{\boldsymbol{\sigma}}^2[l]$ for each $\Xi[l]$ is calculated, while alg. 2 obeys the coherency in $\boldsymbol{\Xi}$.

IV. Numerical Example

The equations presented above are applied to a servo-hydraulic positioning system, which actuates a Stewart-Gough-Platform. The following model is chosen:

$$\dot{\boldsymbol{x}} = \begin{bmatrix} \dot{q} \\ \ddot{q} \\ \dot{p}_A \\ \dot{p}_B \end{bmatrix} = \begin{bmatrix} \dot{q} \\ \frac{1}{m_q} \cdot (A_A \cdot p_A - A_B \cdot p_B - \tau_f) \\ \frac{E_{Oil}}{V_A(q)} (-\dot{q} \cdot A_A + Q_A + Q_{L,A}) \\ \frac{E_{Oil}}{V_B(q)} (\dot{q} \cdot A_B + Q_B + Q_{L,B}) \end{bmatrix} \tag{23a}$$

$$\boldsymbol{y} = \begin{bmatrix} q & \dot{q} p_A & p_B \end{bmatrix}^T \tag{23b}$$

$$\tau_f = a_\tau \cdot \text{sgn}(\dot{q}) + b_\tau \cdot \dot{q} \tag{23c}$$

$$Q_A = x_v c_d \sqrt{\frac{2}{\rho}} \sqrt{\Delta p_A} \tag{23d}$$

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

$$Q_B = -x_v c_d \sqrt{\frac{2}{\rho}} \sqrt{\Delta p_B} \tag{23e}$$

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

$$Q_{L,A} = c_L \cdot \left(\sqrt{p_s - p_A} - \sqrt{p_A - p_t}\right) \tag{23f}$$

$$Q_{L,B} = c_L \cdot \left(\sqrt{p_s - p_B} - \sqrt{p_B - p_t}\right) \tag{23g}$$

$$V_A(q) = A_A \cdot (q - q_{min}) + V_{0,A} \tag{23h}$$

$$V_B(q) = A_B \cdot (q_{max} - q) + V_{0,B} \tag{23i}$$

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
$\text{H}(\cdot)$: Heaviside Function
τ_f : Friction Force
m : Mass attached to Cylinder
A_A, A_B : Piston Areas
E_{Oil} : Oil Modulus
c_d,c_L, ρ: Valve Coeff., Leakage Coeff., Oil Density

The a-priori parameter vector has been chosen as

$$\tilde{\boldsymbol{\theta}} = \begin{bmatrix} E_{Oil} \\ p_t \\ B_v \\ c_L \\ V_{0,A} \\ V_{0,B} \\ a_\tau \\ b_\tau \end{bmatrix}^T = \begin{bmatrix} 960 \cdot 10^{\,6} & \frac{N}{mm^2} \\ 0.3 \cdot 10^{\,6} & \frac{N}{mm^2} \\ 6.0816 \cdot 10^{-7} & \frac{m^4}{\sqrt{N}s} \\ 1 \cdot 10^{-8} & \frac{m^4}{\sqrt{N}s} \\ 1.0254 \cdot 10^{-4} & m^3 \\ 1.5281 \cdot 10^{-4} & m^3 \\ 2 \cdot 10^{\,3} & N \\ 1 \cdot 10^{\,3} & \frac{Ns}{m} \end{bmatrix}^T . \tag{24}$$

The sequence shown in fig. 1 serves as a numerical example here. It contains 3000 equidistant samples within 0.6 s experiment duration. The input signal consists of a step from 10 % to 0 % of the full range of values. The samples are considered as a coherent sequence.

Table I shows numerical results of $\boldsymbol{D}$ under the assumption $\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{\theta}}} = \text{diag}(0.12 \cdot \tilde{\boldsymbol{\theta}}^2)$ The comparison of $\tilde{\boldsymbol{D}}_{\text{a-priori}}(\textit{Lazyness} = 1)$ versus $\tilde{\boldsymbol{D}}_{\text{a-priori}}(\textit{Lazyness} = 0.5)$ shows that *lazy calculation* yields an approximation with small errors. If the *LazynessThreshold* is decreased further, the approximation error becomes considerable, compare $\tilde{\boldsymbol{D}}_{\text{a-priori}}\big|_{Lazyness=1}$ versus $\tilde{\boldsymbol{D}}_{\text{a-priori}}(\textit{Lazyness} = 0.05)$. This validates the law of large numbers for this application.

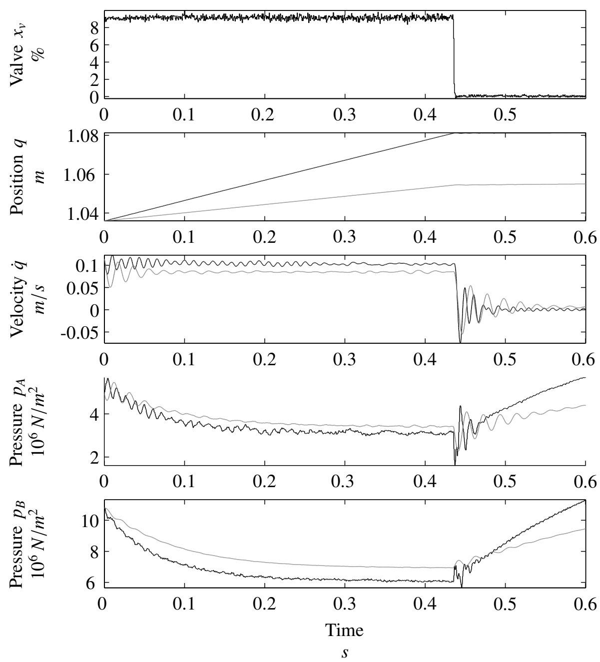

Fig. 1. Coherent experiment sequence (black: measurement, grey: simulation using $\tilde{\boldsymbol{\theta}}$).

The validity of the approximation in eq. (21) can be evaluated upon the relative error $\boldsymbol{e}$ of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$ and $\boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}})$. $\boldsymbol{e}$ is calculated element-wise:

$$e_{i,j}(\tilde{\boldsymbol{D}}_{\text{a-priori}}, \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}})) = \frac{(\tilde{\boldsymbol{D}}_{\text{a-priori}})_{i,j} - (\boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))_{i,j}}{(\boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))_{i,j}}. \quad (25)$$

Table I shows that $\boldsymbol{e}(\tilde{\boldsymbol{D}}_{\text{a-priori}}, \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))$ is not neglectable. Three reasons for this deviation are stated: At first, the detuned models are different to the nominal model only in their parameters, not in their structure. Structural insufficiencies can hence not be represented adequately. Secondly, the assumption of $\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{\theta}}}$ as a diagonal matrix is a poor choice. However, it is the most promising choice at this stage which yields a comparably simple implementation of $\tilde{\boldsymbol{D}}_{\text{a-priori}}$. Third, it must be kept in mind that $\boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}})$ is a sample covariance of $\boldsymbol{D}$, where considerable deviations may be present.

The results obtained for $\boldsymbol{D}$ form the basis for the Fisher matrix. Based on table I the a-priori Fisher matrix and the a-posteriori Fisher matrix of the sequence depicted in fig. 1 were calculated. Both Fisher matrices and their relative errors are shown in table II. It can be seen that $\boldsymbol{e}(\tilde{\boldsymbol{F}}_{\text{a-priori}}, \boldsymbol{F}_{\text{a-posteriori}})$ is not neglectable. The following conclusion is derived from this insight: Even though there is considerable prediction error in $\tilde{\boldsymbol{F}}_{\text{a-priori}}$ the differences across the elements $F_{i,j}$ are by far higher. Hence, the information content of the Fisher matrix is of much more influence than the prediction error. This stresses the hypothesis that $\tilde{\boldsymbol{F}}_{\text{a-priori}}$ provides useful information about the parameter uncertainties, at least up to a certain level. Later, when $\tilde{\boldsymbol{F}}_{\text{a-priori}}$ is used for experiment optimization, the prediction error must be considered when formulating the abort criterion of the optimization procedure. Since $\boldsymbol{F}_{\text{a-posteriori}}$ is approximated by $\tilde{\boldsymbol{F}}_{\text{a-priori}}$, it is sufficient

TABLE I

$\tilde{\boldsymbol{D}}_{\text{A-PRIORI}}$ FOR COHERENT SEQUENCE SHOWED IN FIG. 1

$\tilde{\boldsymbol{D}}_{\text{a-priori}}$, Lazyness=1:

$1.06 \cdot 10^{-05}$	$1.72 \cdot 10^{-05}$	$-1.96 \cdot 10^{+02}$	$-4.99 \cdot 10^{+02}$
$1.72 \cdot 10^{-05}$	$1.03 \cdot 10^{-04}$	$-5.77 \cdot 10^{+02}$	$-1.74 \cdot 10^{+03}$
$-1.96 \cdot 10^{+02}$	$-5.77 \cdot 10^{+02}$	$3.80 \cdot 10^{+10}$	$5.88 \cdot 10^{+10}$
$-4.99 \cdot 10^{+02}$	$-1.74 \cdot 10^{+03}$	$5.88 \cdot 10^{+10}$	$1.39 \cdot 10^{+11}$

$\tilde{\boldsymbol{D}}_{\text{a-priori}}$, Lazyness=0.5:

$1.21 \cdot 10^{-05}$	$1.96 \cdot 10^{-05}$	$-2.31 \cdot 10^{+02}$	$-5.86 \cdot 10^{+02}$
$1.96 \cdot 10^{-05}$	$1.19 \cdot 10^{-04}$	$-6.59 \cdot 10^{+02}$	$-2.00 \cdot 10^{+03}$
$-2.31 \cdot 10^{+02}$	$-6.59 \cdot 10^{+02}$	$4.48 \cdot 10^{+10}$	$7.08 \cdot 10^{+10}$
$-5.86 \cdot 10^{+02}$	$-2.00 \cdot 10^{+03}$	$7.08 \cdot 10^{+10}$	$1.67 \cdot 10^{+11}$

$\tilde{\boldsymbol{D}}_{\text{a-priori}}$, Lazyness=0.05:

$2.90 \cdot 10^{-05}$	$4.67 \cdot 10^{-05}$	$-4.21 \cdot 10^{+02}$	$-1.41 \cdot 10^{+03}$
$4.67 \cdot 10^{-05}$	$2.49 \cdot 10^{-04}$	$-1.12 \cdot 10^{+03}$	$-5.01 \cdot 10^{+03}$
$-4.21 \cdot 10^{+02}$	$-1.12 \cdot 10^{+03}$	$9.03 \cdot 10^{+10}$	$1.79 \cdot 10^{+11}$
$-1.41 \cdot 10^{+03}$	$-5.01 \cdot 10^{+03}$	$1.79 \cdot 10^{+11}$	$4.50 \cdot 10^{+11}$

$\boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}})$:

$3.35 \cdot 10^{-05}$	$4.65 \cdot 10^{-05}$	$-4.17 \cdot 10^{+02}$	$-2.87 \cdot 10^{+03}$
$4.65 \cdot 10^{-05}$	$3.83 \cdot 10^{-04}$	$-2.39 \cdot 10^{+03}$	$-6.15 \cdot 10^{+03}$
$-4.17 \cdot 10^{+02}$	$-2.39 \cdot 10^{+03}$	$7.42 \cdot 10^{+10}$	$9.51 \cdot 10^{+10}$
$-2.87 \cdot 10^{+03}$	$-6.15 \cdot 10^{+03}$	$9.51 \cdot 10^{+10}$	$3.38 \cdot 10^{+11}$

Relative error $\boldsymbol{e}(\tilde{\boldsymbol{D}}_{\text{a-priori}}(Lazyness = 1.0), \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))$:

$6.83 \cdot 10^{-01}$	$6.29 \cdot 10^{-01}$	$5.29 \cdot 10^{-01}$	$8.26 \cdot 10^{-01}$
$6.29 \cdot 10^{-01}$	$7.32 \cdot 10^{-01}$	$7.59 \cdot 10^{-01}$	$7.18 \cdot 10^{-01}$
$5.29 \cdot 10^{-01}$	$7.59 \cdot 10^{-01}$	$4.89 \cdot 10^{-01}$	$3.81 \cdot 10^{-01}$
$8.26 \cdot 10^{-01}$	$7.18 \cdot 10^{-01}$	$3.81 \cdot 10^{-01}$	$5.88 \cdot 10^{-01}$

Relative error $\boldsymbol{e}(\tilde{\boldsymbol{D}}_{\text{a-priori}}(Lazyness = 0.5), \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))$:

$6.39 \cdot 10^{-01}$	$5.78 \cdot 10^{-01}$	$4.46 \cdot 10^{-01}$	$7.96 \cdot 10^{-01}$
$5.78 \cdot 10^{-01}$	$6.89 \cdot 10^{-01}$	$7.25 \cdot 10^{-01}$	$6.75 \cdot 10^{-01}$
$4.46 \cdot 10^{-01}$	$7.25 \cdot 10^{-01}$	$3.97 \cdot 10^{-01}$	$2.55 \cdot 10^{-01}$
$7.96 \cdot 10^{-01}$	$6.75 \cdot 10^{-01}$	$2.55 \cdot 10^{-01}$	$5.07 \cdot 10^{-01}$

Relative error $\boldsymbol{e}(\tilde{\boldsymbol{D}}_{\text{a-priori}}(Lazyness = 0.05), \boldsymbol{D}(\hat{\boldsymbol{Y}}, \tilde{\boldsymbol{\theta}}))$:

$1.34 \cdot 10^{-01}$	$-5.28 \cdot 10^{-03}$	$-1.09 \cdot 10^{-02}$	$5.08 \cdot 10^{-01}$
$-5.28 \cdot 10^{-03}$	$3.50 \cdot 10^{-01}$	$5.31 \cdot 10^{-01}$	$1.85 \cdot 10^{-01}$
$-1.09 \cdot 10^{-02}$	$5.31 \cdot 10^{-01}$	$-2.16 \cdot 10^{-01}$	$-8.84 \cdot 10^{-01}$
$5.08 \cdot 10^{-01}$	$1.85 \cdot 10^{-01}$	$-8.84 \cdot 10^{-01}$	$-3.33 \cdot 10^{-01}$

for the criterion to enter a certain uncertainty region around the optimum. Note, it is reasonable for experiment optimization to balance the eigenvalues of $\boldsymbol{F}$ or of a normalized form of $\boldsymbol{F}$. As soon as the elements in $\boldsymbol{F}$ are balanced in the order of magnitude of the prediction error, much improvement in comparison to the actual situation will be achieved already.

V. CONCLUSIONS

This article adresses aspects of the a-priori computation of the Fisher information matrix. The aim of the type of Fisher matrix discussed is the experiment design for parameter estimation of nonlinear state space models. A definition of experiments is stated, which is convenient for software implementation. Both, coherent *and* non-coherent measurement samples may be stored in one data structure.

A method for the a-priori computation of the Fisher matrix is proposed. The a-priori approximation of the covariance matrix of the prediction error is elaborated. A method

TABLE II

Fisher Matrix for coherent sequence shown in fig. 1

$\tilde{\boldsymbol{F}}_{\text{a-priori}}(Lazyness = 1)$:

$2.54 \cdot 10^{-14}$	$2.63 \cdot 10^{-14}$	$-5.45 \cdot 10^{+01}$	$-2.42 \cdot 10^{+02}$	$-8.10 \cdot 10^{-02}$	$-1.78 \cdot 10^{-03}$	$1.84 \cdot 10^{-09}$	$5.42 \cdot 10^{-10}$
$2.63 \cdot 10^{-14}$	$1.42 \cdot 10^{-11}$	$-1.47 \cdot 10^{+02}$	$2.55 \cdot 10^{+03}$	$-8.97 \cdot 10^{-02}$	$-1.69 \cdot 10^{-03}$	$-2.49 \cdot 10^{-08}$	$-1.60 \cdot 10^{-09}$
$-5.45 \cdot 10^{+01}$	$-1.47 \cdot 10^{+02}$	$1.84 \cdot 10^{+17}$	$8.09 \cdot 10^{+17}$	$1.84 \cdot 10^{+14}$	$5.80 \cdot 10^{+11}$	$1.77 \cdot 10^{+07}$	$5.82 \cdot 10^{+05}$
$-2.42 \cdot 10^{+02}$	$2.55 \cdot 10^{+03}$	$8.09 \cdot 10^{+17}$	$4.36 \cdot 10^{+18}$	$8.23 \cdot 10^{+14}$	$-9.28 \cdot 10^{+10}$	$7.46 \cdot 10^{+07}$	$2.25 \cdot 10^{+06}$
$-8.10 \cdot 10^{-02}$	$-8.97 \cdot 10^{-02}$	$1.84 \cdot 10^{+14}$	$8.23 \cdot 10^{+14}$	$2.60 \cdot 10^{+11}$	$5.04 \cdot 10^{+09}$	$-2.56 \cdot 10^{+03}$	$-1.46 \cdot 10^{+03}$
$-1.78 \cdot 10^{-03}$	$-1.69 \cdot 10^{-03}$	$5.80 \cdot 10^{+11}$	$-9.28 \cdot 10^{+10}$	$5.04 \cdot 10^{+09}$	$3.50 \cdot 10^{+08}$	$-1.13 \cdot 10^{+03}$	$-1.17 \cdot 10^{+02}$
$1.84 \cdot 10^{-09}$	$-2.49 \cdot 10^{-08}$	$1.77 \cdot 10^{+07}$	$7.46 \cdot 10^{+07}$	$-2.56 \cdot 10^{+03}$	$-1.13 \cdot 10^{+03}$	$7.33 \cdot 10^{-03}$	$6.14 \cdot 10^{-04}$
$5.42 \cdot 10^{-10}$	$-1.60 \cdot 10^{-09}$	$5.82 \cdot 10^{+05}$	$2.25 \cdot 10^{+06}$	$-1.46 \cdot 10^{+03}$	$-1.17 \cdot 10^{+02}$	$6.14 \cdot 10^{-04}$	$5.85 \cdot 10^{-05}$

$\tilde{\boldsymbol{F}}_{\text{a-priori}}(Lazyness = 0.5)$:

$1.79 \cdot 10^{-14}$	$1.74 \cdot 10^{-14}$	$-3.86 \cdot 10^{+01}$	$-1.72 \cdot 10^{+02}$	$-5.70 \cdot 10^{-02}$	$-1.22 \cdot 10^{-03}$	$1.23 \cdot 10^{-09}$	$3.76 \cdot 10^{-10}$
$1.74 \cdot 10^{-14}$	$4.53 \cdot 10^{-12}$	$-9.81 \cdot 10^{+01}$	$5.78 \cdot 10^{+02}$	$-6.22 \cdot 10^{-02}$	$3.16 \cdot 10^{-04}$	$-1.65 \cdot 10^{-08}$	$-1.06 \cdot 10^{-09}$
$-3.86 \cdot 10^{+01}$	$-9.81 \cdot 10^{+01}$	$1.31 \cdot 10^{+17}$	$5.74 \cdot 10^{+17}$	$1.30 \cdot 10^{+14}$	$4.00 \cdot 10^{+11}$	$1.27 \cdot 10^{+07}$	$4.21 \cdot 10^{+05}$
$-1.72 \cdot 10^{+02}$	$5.78 \cdot 10^{+02}$	$5.74 \cdot 10^{+17}$	$2.78 \cdot 10^{+18}$	$5.83 \cdot 10^{+14}$	$1.22 \cdot 10^{+12}$	$5.36 \cdot 10^{+07}$	$1.65 \cdot 10^{+06}$
$-5.70 \cdot 10^{-02}$	$-6.22 \cdot 10^{-02}$	$1.30 \cdot 10^{+14}$	$5.83 \cdot 10^{+14}$	$1.83 \cdot 10^{+11}$	$3.50 \cdot 10^{+09}$	$-1.58 \cdot 10^{+03}$	$-1.01 \cdot 10^{+03}$
$-1.22 \cdot 10^{-03}$	$3.16 \cdot 10^{-04}$	$4.00 \cdot 10^{+11}$	$1.22 \cdot 10^{+12}$	$3.50 \cdot 10^{+09}$	$2.11 \cdot 10^{+08}$	$-7.96 \cdot 10^{+02}$	$-8.24 \cdot 10^{+01}$
$1.23 \cdot 10^{-09}$	$-1.65 \cdot 10^{-08}$	$1.27 \cdot 10^{+07}$	$5.36 \cdot 10^{+07}$	$-1.58 \cdot 10^{+03}$	$-7.96 \cdot 10^{+02}$	$5.18 \cdot 10^{-03}$	$4.33 \cdot 10^{-04}$
$3.76 \cdot 10^{-10}$	$-1.06 \cdot 10^{-09}$	$4.21 \cdot 10^{+05}$	$1.65 \cdot 10^{+06}$	$-1.01 \cdot 10^{+03}$	$-8.24 \cdot 10^{+01}$	$4.33 \cdot 10^{-04}$	$4.11 \cdot 10^{-05}$

$\boldsymbol{F}_{\text{a-posteriori}}$:

$9.79 \cdot 10^{-15}$	$1.45 \cdot 10^{-14}$	$-2.09 \cdot 10^{+01}$	$-8.95 \cdot 10^{+010}$	$-3.11 \cdot 10^{-02}$	$-7.53 \cdot 10^{-04}$	$6.88 \cdot 10^{-10}$	$2.07 \cdot 10^{-10}$
$1.45 \cdot 10^{-14}$	$1.92 \cdot 10^{-11}$	$-3.08 \cdot 10^{+01}$	$4.22 \cdot 10^{+030}$	$-4.05 \cdot 10^{-02}$	$-5.28 \cdot 10^{-03}$	$-1.37 \cdot 10^{-08}$	$-8.79 \cdot 10^{-10}$
$-2.09 \cdot 10^{+01}$	$-3.08 \cdot 10^{+01}$	$7.05 \cdot 10^{+16}$	$3.17 \cdot 10^{+17}$	$7.06 \cdot 10^{+13}$	$1.55 \cdot 10^{+11}$	$6.69 \cdot 10^{+06}$	$2.15 \cdot 10^{+05}$
$-8.95 \cdot 10^{+01}$	$4.22 \cdot 10^{+03}$	$3.17 \cdot 10^{+17}$	$2.53 \cdot 10^{+18}$	$3.12 \cdot 10^{+14}$	$-4.01 \cdot 10^{+12}$	$2.72 \cdot 10^{+07}$	$7.82 \cdot 10^{+05}$
$-3.11 \cdot 10^{-02}$	$-4.05 \cdot 10^{-02}$	$7.06 \cdot 10^{+13}$	$3.12 \cdot 10^{+14}$	$9.96 \cdot 10^{+10}$	$1.98 \cdot 10^{+09}$	$-9.38 \cdot 10^{+02}$	$-5.58 \cdot 10^{+02}$
$-7.53 \cdot 10^{-04}$	$-5.28 \cdot 10^{-03}$	$1.55 \cdot 10^{+11}$	$-4.01 \cdot 10^{+12}$	$1.98 \cdot 10^{+09}$	$2.26 \cdot 10^{+08}$	$-4.26 \cdot 10^{+02}$	$-4.49 \cdot 10^{+01}$
$6.88 \cdot 10^{-10}$	$-1.37 \cdot 10^{-08}$	$6.69 \cdot 10^{+06}$	$2.72 \cdot 10^{+07}$	$-9.38 \cdot 10^{+02}$	$-4.26 \cdot 10^{+02}$	$2.76 \cdot 10^{-03}$	$2.31 \cdot 10^{-04}$
$2.07 \cdot 10^{-10}$	$-8.79 \cdot 10^{-10}$	$2.15 \cdot 10^{+05}$	$7.82 \cdot 10^{+05}$	$-5.58 \cdot 10^{+02}$	$-4.49 \cdot 10^{+01}$	$2.31 \cdot 10^{-04}$	$2.21 \cdot 10^{-05}$

$\boldsymbol{e}(\tilde{\boldsymbol{F}}_{\text{a-priori}}, \boldsymbol{F}_{\text{a-posteriori}}(Lazyness = 1))$:

$-1.60 \cdot 10^{+00}$	$-8.12 \cdot 10^{-01}$	$-1.61 \cdot 10^{+00}$	$-1.71 \cdot 10^{+00}$	$-1.61 \cdot 10^{+00}$	$-1.37 \cdot 10^{+00}$	$-1.68 \cdot 10^{+00}$	$-1.62 \cdot 10^{+00}$
$-8.12 \cdot 10^{-01}$	$2.60 \cdot 10^{-01}$	$-3.76 \cdot 10^{+00}$	$3.96 \cdot 10^{-01}$	$-1.21 \cdot 10^{+00}$	$6.80 \cdot 10^{-01}$	$-8.14 \cdot 10^{-01}$	$-8.25 \cdot 10^{-01}$
$-1.61 \cdot 10^{+00}$	$-3.76 \cdot 10^{+00}$	$-1.61 \cdot 10^{+00}$	$-1.55 \cdot 10^{+00}$	$-1.61 \cdot 10^{+00}$	$-2.75 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$	$-1.71 \cdot 10^{+00}$
$-1.71 \cdot 10^{+00}$	$3.96 \cdot 10^{-01}$	$-1.55 \cdot 10^{+00}$	$-7.21 \cdot 10^{-01}$	$-1.64 \cdot 10^{+00}$	$9.77 \cdot 10^{-01}$	$-1.74 \cdot 10^{+00}$	$-1.87 \cdot 10^{+00}$
$-1.61 \cdot 10^{+00}$	$-1.21 \cdot 10^{+00}$	$-1.61 \cdot 10^{+00}$	$-1.64 \cdot 10^{+00}$	$-1.61 \cdot 10^{+00}$	$-1.54 \cdot 10^{+00}$	$-1.73 \cdot 10^{+00}$	$-1.62 \cdot 10^{+00}$
$-1.37 \cdot 10^{+00}$	$6.80 \cdot 10^{-01}$	$-2.75 \cdot 10^{+00}$	$9.77 \cdot 10^{-01}$	$-1.54 \cdot 10^{+00}$	$-5.48 \cdot 10^{-01}$	$-1.65 \cdot 10^{+00}$	$-1.62 \cdot 10^{+00}$
$-1.68 \cdot 10^{+00}$	$-8.14 \cdot 10^{-01}$	$-1.65 \cdot 10^{+00}$	$-1.74 \cdot 10^{+00}$	$-1.73 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$
$-1.62 \cdot 10^{+00}$	$-8.25 \cdot 10^{-01}$	$-1.71 \cdot 10^{+00}$	$-1.87 \cdot 10^{+00}$	$-1.62 \cdot 10^{+00}$	$-1.62 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$	$-1.65 \cdot 10^{+00}$

that relies on pseudo-data from detuned models is stated. *Lazy computation* is proposed to increase efficiency while accepting a small computational error. A numerical example of a servo-hydraulic positioning system illustrates the ideas. From this example it can be concluded that the Fisher matrix may be approximated a-priori using the pseudo-data method presented. Hence, the a-priori Fisher matrix is recommendable as a criterion for experiment design.

References

[1] 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.

[2] P. Chatzakos and E. Papadopoulus, "On Model-Based Control of Hydraulic Actuators," in *RAAD 2003 : Int. Workshop on Robotics in Alpe-Adria-Danube Region Cassino,*, M. Ceccarelli, Ed., 2003.

[3] D. Hisseine, "Robust tracking control for a hydraulic actuation system," in *IEEE Conf. on Control Applications*, 2005, pp. 422–427.

[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] F. Bu and B. Yao, "Nonlinear adaptive robust control of hydraulic actuators regulated by proportional directional control valves with deadband and nonlinearflow gains," in *Proc. of the American Control Conference*, ASME, Ed., vol. 6, 2000, pp. 4129–4133.

[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] J. Bernoulli and O. Sheynin, *On the law of large numbers: Part four of Ars Conjectandi*. Berlin: NG Verlag, 2005.