Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Applications Oriented Input Design in Time-Domain Through Cyclic Methods ⋆

A. Ebadat * **B. Wahlberg** * **H. Hjalmarsson** * **C. R. Rojas** * **P. Hägg** * **C. A. Larsson** *

* *ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (e-mail: {ebadat, bo, hjalmars, crro, pehagg, chrisla}@kth.se).*

Abstract: In this paper we propose a method for applications oriented input design for linear systems in open-loop under time-domain constraints on the amplitude of input and output signals. The method guarantees a desired control performance for the estimated model in minimum time, by imposing some lower bound on the information matrix. The problem is formulated as a time-domain optimization problem, which is non-convex. This is addressed through an alternating method, where we separate the problem into two steps and at each step we optimize the cost function with respect to one of two variables. We alternate between these two steps until convergence. A time recursive input design algorithm is performed, which enables us to use the algorithm with control. Therefore, a receding horizon framework is used to solve each optimization problem. Finally, we illustrate the method with a numerical example which shows the good ability of the proposed approach in generating an optimal input signal.

Keywords: System identification, Applications oriented input design, Alternating methods.

1. INTRODUCTION

System identification concerns the problem of data-based plant modeling and plays an important role in industry. One of the key enabling issues in any system identification problem is the choice of input signal. An appropriate input signal should be able to extract as much useful information as possible from the system. Therefore, a properly designed input signal can improve the quality of the identified model, significantly. This problem has led to formation of the topic *optimal input design.*

Optimal input design has been extensively investigated in the literature, see e.g. Ljung (1999), Goodwin and Payne (1977), and Bombois et al. (2006). The problem has been formulated in many different forms, however, one common idea is to design an input signal such that a certain accuracy is obtained during the identification while the experimental effort to obtain such an accuracy is minimized. This accuracy is often defined in terms of the application of the model and thus the identification objective is to guarantee that the estimated model belongs to the set of models that satisfies the desired control specifications, with a given probability. This induces growth of the ideas of identification for control, least costly identification and applications oriented input design, see Hjalmarsson (2005), Gevers and Ljung (1986), Bombois et al. (2006), and Hjalmarsson (2009). The problem is usually defined as an optimization problem where one tries to satisfy the requirements on the quality of the model by using the minimum experimental effort. The quality of the model can be measured by the Fisher matrix, which determines the amount of information regarding the model and, the inverse of this matrix is a lower bound on the covariance matrix for any unbiased estimator (Ljung (1999)).
For model structures linear in input, the information matrix is asymptotically an affine function of the input power spectrum. Therefore, the input design problem is usually formulated in the frequency domain and the outcome is an optimal input spectrum or an autocorrelation sequence. The optimal input values are obtained from the given optimal spectrum, see Fedorov (1972). In practice there are some constant bounds on the input signals and the resulting output signals, which should be taken into account during the experiment design. These constraints are typically expressed in the time-domain and how to handle this in frequency-domain is not evident. One way to get around this problem is to impose these constraints during the generation of a time realization of the desired input spectrum, see Gujar and Kavanagh (1968), Liu and Munson (1982), Schroeder (1970) and Larsson et al. (2013).
There are, however, some approaches that try to solve the optimal input design problem in the time domain directly, see e.g Manchester (2010) for linear systems. The main advantage is that in the time domain the constraints on the amplitude of the input and the system dynamics appears naturally and are easier to handle. However, the main difficulty that arises is that the problem is non-convex. In Manchester (2010), this problem is addressed through a semidefinite relaxation of quadratic programs and the Fisher information matrix is maximized under some constraints on the input signal.

In this paper we propose a novel method for applications oriented input design in open-loop for linear systems in the time domain, where it is straightforward to formulate the constraints on both the input and output signals. The aim is to satisfy some lower bound on the information matrix in minimum time, which can guarantee a desired control performance for the estimated model. The problem is formulated as an optimization problem by adding a positive slack variable to the lower bound. The problem, however, is non-convex, which imposes a high computational burden.

⋆ The research leading to these results has received funding from the European Unions Seventh Framework Programme (FP7/2007-2013) under grant agreement no 257059, the Autoprofit project (www.fp7-autoprofit.eu). This work was partially supported by the Swedish Research Council under contract 621-2009-4017.

We try to get around the non-convexity through alternating methods. More precisely, we solve the problem for one variable, when the other is fixed. Thus, in this paper we separate the problem into two steps where at each step we optimize the cost function with respect to one set of the variables and we alternate between these two steps until convergence. In order to detect the minimum required time the problem is formulated in a receding horizon manner and we perform a time recursive algorithm. This enables us to overcome the computational burden and makes the formulation suitable to be used with controllers such as Model Predictive Control (MPC).

The outline of the paper is as follows. In Section 2, we go through the problem formulation and mathematical backgrounds. We describe the input design problem in Section 3, followed by a description of the proposed method in Section 4. Section 5 describes the method for Finite Impulse Response models. In Section 6, we illustrate the method in a numerical example and in Section 7, some conclusions are stated.

Notations: $\mathbb{E}\{.\}$ denotes expected value. We use $(.)^{\frac{1}{2}}$ to denote a Hermitian square root of a positive definite matrix. Define $\mathcal{S}_n^+$ to be the set of positive $n \times n$ semi-definite matrices. A matrix $A \in \mathbb{R}^{m \times n}$ $(m > n)$, is said to be semi-unitary if $A^T A = I_{n\times n}$, and $I_{n\times n}$ is an $n \times n$ identity matrix.

2. PROBLEM FORMULATION

Consider the identification of discrete-time multivariate systems that are causal linear time-invariant (LTI)

$$y(t) = G_0(q)u(t) + H_0(q)e_0(t), \tag{1}$$

where $u(t) \in \mathbb{R}^{n_u}$ and $y(t) \in \mathbb{R}^{n_y}$ are the input and output vectors and $e_0(t) \in \mathbb{R}^{n_e}$ is white Gaussian noise with zero mean and covariance matrix Λ. $G_0(q)$ and $H_0(q)$ are the transfer function matrices of the system. Let q^{-1} denote the backward shift operator, e.g., $q^{-1}u(t) = u(t-1)$.
In system identification, we want to find a model of the system (1). We assume that the model is parametrized by an unknown parameter vector $\theta \in \mathbb{R}^{n_\theta}$, that is,

$$\mathcal{M}(\theta): \quad y(t) = G(q,\theta)u(t) + H(q,\theta)e(t). \tag{2}$$

In addition, we assume that the model (2) matches system (1) exactly when $\theta = \theta_o$. We call θ_o the true parameter vector. The objective of system identification is to estimate the value of θ that best describes the system, according to some quality measure. The estimated parameter vector, given N measurements in the experiment, is denoted $\hat{\theta}_N$.

3. OPTIMAL INPUT DESIGN

The idea in optimal input design in the least costly framework is to minimize an experimental effort, such as input power, while satisfying some requirements on the accuracy of the identified model. These requirements can be expressed in terms of control performance. The problem is then to design an input signal to be used in the identification experiment such that the estimated model guarantees acceptable control performance when used in the control design, namely *applications oriented experiment design* (see Bombois et al. (2006) and references therein).

3.1 Application Cost

We use the concept of application cost function to relate the plant-model mismatch to the performance degradation. We use a scalar function of θ as the application cost and denote it $V_{app}(\theta)$. The cost function is chosen such that its minimum value occurs at $\theta = \theta_o$. In particular, we assume without loss of generality that $V_{app}(\theta_o) = 0$. Note that if $V_{app}(\theta)$ is twice differentiable in a neighborhood of θ_o, this implies that $V_{app}(\theta_o) = 0$, $V'_{app}(\theta_o) = 0$ and $V''_{app}(\theta_o) \geq 0$. There are many possible choices of application functions, see e.g. Larsson et al. (2011).

The set of all acceptable parameters from an application's point of view, namely the *application set*, is defined as

$$\Theta(\gamma) = \left\{\theta : V_{app}(\theta) \leq \frac{1}{\gamma}\right\}, \tag{3}$$

where γ is a user-defined positive constant that determines the maximum allowed performance degradation. We can make a local convex approximation of $\Theta(\gamma)$ by invoking the Taylor expansion of $V_{app}(\theta)$ around θ_o and considering its mentioned properties around θ_o (see Hjalmarsson (2009)):

$$\Theta(\gamma) \approx \mathcal{E}_{app}(\gamma) = \left\{\theta : [\theta-\theta_o]^T V''_{app}(\theta_o)[\theta-\theta_o] \leq \frac{2}{\gamma}\right\}. \tag{4}$$

3.2 Applications Oriented Experiment Design

We use the prediction error method (PEM) with quadratic cost to estimate the unknown parameters of the considered system, $\theta \in \mathbb{R}^n$, from N available samples of input-output data, see Ljung (1999). A key asymptotic ($N \to \infty$) property of PEM, is that the estimated parameters lie in an *identification set* with a certain probability say α, (Ljung and Wahlberg (1992)). This set is defined as

$$\mathcal{E}_{SI}(\alpha) = \left\{\theta : [\theta-\theta_o]^T I_F(\theta_o)[\theta-\theta_o] \leq \chi^2_\alpha(n_\theta)\right\}, \tag{5}$$

where $\chi^2_\alpha(n)$ is the α-percentile of the χ^2-distribution with n degrees of freedom and I_F is the Fisher information matrix. We thus have that $\hat{\theta}_N \in \mathcal{E}_{SI}(\alpha)$ with probability α when $N \to \infty$. In this paper we assume that N is finite but sufficiently large such that asymptotic properties hold. For more details, we refer the reader to Ljung (1999).

In applications oriented input design, the input signal used in the identification experiment is designed such that the estimated model guarantees acceptable control performance when used in the control design, that is, it requires that $\hat{\theta}_N \in \Theta(\gamma)$ with high probability. One way to ensure this is to require

$$\mathcal{E}_{SI}(\alpha) \subseteq \Theta(\gamma). \tag{6}$$

Using this set constraint, the input design problem can be formulated as an optimization problem, where (6) plays the role of a constraint. In order to make the problem convex, the ellipsoidal approximation of the application set, (4), can be used in (6). Thus, both sets are ellipsoids and the problem becomes the following linear matrix inequality (LMI) in the elements of I_F:

$$\frac{1}{\chi^2_\alpha(n_\theta)} I_F(\theta_o) \geq \frac{\gamma}{2} V''_{app}(\theta_o). \tag{7}$$

Finally, a natural objective in the input design is to minimize an experiment cost, such as input power or energy or experimental time, while (7) is fulfilled, i.e.

$$\begin{aligned} \min_{\text{input}} \quad & \text{Experimental Cost} \\ \text{s.t.} \quad & \frac{1}{\chi^2_\alpha(n_\theta)} I_F(\theta_o) \geq \frac{\gamma}{2} V''_{app}(\theta_o). \end{aligned} \tag{8}$$

Since I_F is an affine function of the input spectrum in open loop identification (Ljung (1999)), the constraint (7) can be formulated as *LMIs* by linear parameterization of the input spectrum. Therefore, the optimization problem (8) is usually solved for the input spectrum and thus the outcome of the optimization is often not given as a sequence of values but rather it is given as an optimal input spectrum or an autocorrelation sequence.

4. TIME-DOMAIN OPTIMAL INPUT DESIGN

In this paper we introduce a new solution to the applications oriented input design. The objective is to satisfy the constraint (7) in *minimum time* while we are forcing the input and output signals to lie in certain convex sets. In the context of the problem (8), the experimental cost here is the minimum required time to satisfy the accuracy and input/output constraints. The problem is formulated as a time-domain optimization problem, where it is straightforward to handle constraints on input and output signals.

To be able to formulate the problem, we define a new slack variable. The constraint (7) is then satisfied if there exists a positive semidefinite matrix S such that:

$$I_F(\theta_o) - \frac{\chi^2_\alpha(n_\theta)\gamma}{2} V''_{app}(\theta_o) - S = 0, \quad S \geq 0, \tag{9}$$

where S is a positive semi-definite slack variable. We then try to minimize $J = \left\| I_F(\theta_o) - \frac{\chi^2_\alpha(n_\theta)\gamma}{2} V''_{app}(\theta_o) - S \right\|^2_F$ for $S \geq 0$, under input and output constraints. Here, $\|.\|_F$ denotes the Frobenius norm. The experiment design constraint is satisfied if we obtain $J = 0$.

In order to find the minimum required time, we perform a time recursive input design algorithm. This also makes the algorithm compatible to be used with controllers such as MPC since they are using the same context. Hence, we formulate the input design problem as the following receding horizon problem, where at each time t, we solve

$$\begin{aligned} \min_{\{u(k)\}_{k=t}^{t+N_u}, S} \quad & J_t = \left\| I_F^{t+N_u}(\theta_o) - \frac{\chi^2_\alpha(n_\theta)\gamma}{2} V''_{app}(\theta_o) - S \right\|^2_F \\ \text{s.t.} \quad & S \geq 0, \; u(k) \in \mathcal{U}, \quad k = t, \ldots, t+N_u, \\ & y(k) \in \mathcal{Y}, \quad k = t, \ldots, t+N_y, \\ & y(k) = G(q, \theta_0)u(k), \quad k = t, \ldots, t+N_y. \end{aligned} \tag{10}$$

Here, $\mathcal{U}$ and $\mathcal{Y}$ are convex constraint sets on the input and the output. These could for example correspond to system amplitude constraints. N_u and N_y are input and output horizons. If N_y is longer than N_u the input is considered zero over the rest of the output horizon. In this paper, we assume $N_u = N_y$. $I_F^{t+N_u}(\theta_o)$ is the Fisher information matrix up to time $t + N_u$, see Section 4.1. Although the solution to the problem (10) is a sequence of input values, we only apply the first value to the system and the optimization is performed again in the next time step, according to the receding horizon principle. It is also worth to note that the output constraints are only applied on the noiseless output signal (deterministic constraints), for noisy output one can impose stochastic bounds, instead.

At each time sample t, if the lower bound on the information matrix is fulfilled, i.e. $I_F^{t+N_u}(\theta_o) \geq \frac{\chi^2_\alpha(n_\theta)\gamma}{2} V''_{app}(\theta_o)$, then $J_t = 0$ holds and vice versa. We can then stop running the receding horizon (10) when $J_t = 0$ holds for the first time and consider this time to be the minimum time required to satisfy the application requirements.

To iteratively solve (10), we first need to rewrite the information matrix $I_F^{t+N_u}(\theta_o)$ in a recursive form and relate it to the input $u(t)$. Then a cyclic algorithm is proposed to address the input design problem (10).

4.1 Fisher Information Matrix

For an unbiased estimator, the inverse of the Fisher matrix is a lower bound on the covariance of the parameter estimation error, according to Cramér-Rao bound. The information matrix is (Goodwin and Payne (1977)):

$$I_F(\theta) := \mathbb{E}\Big\{ \frac{\partial \log p(y|\theta)}{\partial \theta} \frac{\partial \log p(y|\theta)}{\partial \theta}^T \Big\} \in \mathbb{R}^{n_\theta \times n_\theta}. \tag{11}$$

Considering the model (2) and assuming $e(t)$ to be a Gaussian white noise, the log likelihood function is:

$$\log p(y|\theta) = \text{constant} - \frac{1}{2} \sum_{t=1}^{N} \epsilon^T(t,\theta) \Lambda^{-1} \epsilon(t,\theta) \tag{12}$$

where N is the number of samples that are being used in the computation of the Information matrix and $\epsilon(t,\theta) \in \mathbb{R}^{n_y}$ is the prediction error given by

$$\epsilon(t,\theta) := H^{-1}(q,\theta)[y(t) - G(q,\theta)u(t)].$$

Assume that the plant and noise models are parameterized independently and let $\theta_G \in \mathbb{R}^{n_{\theta G}}$ denote the parameters of the system model while $\theta_H \in \mathbb{R}^{n_{\theta H}}$ contains the parameters of the noise model. The Fisher information matrix for data up to time $t + N_u$ is

$$I_F^{t+N_u}(\theta) := \sum_{k=1}^{k=t+N_u} \mathbb{E}\left\{ \begin{bmatrix} \frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \\ \frac{\partial \epsilon^T(t,\theta)}{\partial \theta_H} \end{bmatrix} \Lambda^{-1} \begin{bmatrix} \frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \\ \frac{\partial \epsilon^T(t,\theta)}{\partial \theta_H} \end{bmatrix}^T \right\} \tag{13}$$

where, $\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \in \mathbb{R}^{n_{\theta G} \times n_y}$ and $\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_H} \in \mathbb{R}^{n_{\theta H} \times n_y}$. Now if we assume that $\{u(t)\}$ and $\{e(t)\}$ are uncorrelated (i.e. the system is operating in open loop), we obtain

$$I_F^{t+N_u}(\theta) = \mathbb{E}\{ \begin{bmatrix} \bar{I}_F^{t+N_u}(\theta_G) & 0 \\ 0 & \bar{I}_F^{t+N_u}(\theta_H) \end{bmatrix} \}. \tag{14}$$

Since $I_F(\theta_H)$ only depends on the noise $e(t)$, the only part of information matrix that can be optimized by the choice of input signal is

$$\bar{I}_F^{t+N_u}(\theta_G) = \sum_{k=1}^{k=t+N_u} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right) \Lambda^{-1} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right)^T, \tag{15}$$

considering that $\mathbb{E}\{\bar{I}_F^{t+N_u}(\theta_G)\} = \bar{I}_F^{t+N_u}(\theta_G)$, since $\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G}$ is deterministic. On the other hand, one can write

$$\begin{aligned} \bar{I}_F^{t+N_u}(\theta_G) := & \sum_{k=1}^{k=t-1} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right) \Lambda^{-1} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right)^T \\ & + \sum_{k=t}^{k=t+N_u} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right) \Lambda^{-1} \left(\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} \right)^T. \end{aligned} \tag{16}$$

The first term in (16) depends on the values of the input signal up to time $t-1$, which are assumed to be known at time t. Therefore, we focus on the second term, which contains the inputs in the horizon in the optimization problem (10). Based on the definition of $\epsilon(t,\theta)$

$$\frac{\partial \epsilon^T(t,\theta)}{\partial \theta_G} = \left[\mathcal{F}_1(q)u(t), \mathcal{F}_2(q)u(t), \ldots, \mathcal{F}_{n_{\theta_G}}(q)u(t), \right]^T,$$

where n_{θ_G} is the number of parameters in the model and

$$\mathcal{F}_i(q)u(t) = -\left[H^{-1}(q,\theta_H) \frac{\partial G(q,\theta_G)}{\partial(\theta_G(i))} u(t) \right]^T. \tag{17}$$

Building on Manchester (2010), the elements of the reduced information matrix can be written as:

$$(\bar{I}_F^{t+N_u})_{i,j}(\theta_G) = (\bar{I}_F^{t-1})_{i,j}(\theta_G) + \sum_{k=t}^{k=t+N_u} (\mathcal{F}_i(q)u(k))\Lambda^{-1}(\mathcal{F}_j(q)u(k)), \tag{18}$$

where $i,j = 1,\ldots,n_{\theta_G}$ and $(\bar{I}_F^{t-1})_{i,j}(\theta_G)$ is obtained using available data at time instant t. Denote the impulse response of $\mathcal{F}_i(q)$ by $f_i(t)$, the maximum length of the truncated impulse responses of $\mathcal{F}_i(q)$ for $i = 1,\ldots,n_{\theta_G}$ by n, and define $(\tilde{I}_F^{t-1})_{i,j}(\theta_G)$ as the part of the information matrix depending on the future values of $u(t)$. Define

$$F_i := \begin{bmatrix} f_i(n) & f_i(n-1) & \ldots & f_i(1) & \ldots & 0 \\ 0 & f_i(n) & \ldots & f_i(2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f_i(N_u) & \ldots & f_i(1) \end{bmatrix}, \tag{19}$$

where $F_i \in \mathbb{R}^{(N_u+1)n_y \times (N_u+n-1)n_u}$, and

$$\begin{aligned} \bar{u}^*(t) &:= [u^*(t-n+1),\ldots,u^*(t-1)] \quad \in \mathbb{R}^{(n-1)n_u}, \\ \bar{u}(t) &:= [u(t),\ldots,u(t+N_u)] \quad \in \mathbb{R}^{(N_u)n_u}. \end{aligned} \tag{20}$$

The former is already known at time t while we are going to optimize the latter. We can then rewrite (18) as

$$\begin{aligned} (\tilde{I}_F^{t+N_u})_{i,j}(\theta_G) &= \left[(\bar{u}^*(t))^T \ \bar{u}(t)^T\right] F_i^T \Lambda_e^{-1} F_j \begin{bmatrix} \bar{u}^*(t) \\ \bar{u}(t) \end{bmatrix}, \\ \Lambda_e^{-1} &= I_{(N_u+1)\times(N_u+1)} \otimes \Lambda^{-1}, \end{aligned} \tag{21}$$

see Manchester (2010). Eventually we have

$$\tilde{I}_F^{t+N_u}(\theta_G) = \begin{bmatrix} \mathbf{u}^T F_1^T \Lambda_e^{-1} F_1 \mathbf{u} & \ldots & \mathbf{u}^T F_1^T \Lambda_e^{-1} F_{n_{\theta_G}} \mathbf{u} \\ \vdots & \ddots & \vdots \\ \mathbf{u}^T F_{n_{\theta_G}}^T \Lambda_e^{-1} F_1 \mathbf{u} & \ldots & \mathbf{u}^T F_{n_{\theta_G}}^T \Lambda_e^{-1} F_{n_{\theta_G}} \mathbf{u} \end{bmatrix}, \tag{22}$$

where

$$\mathbf{u} = \left[(\bar{u}^*(t))^T \ \bar{u}(t)^T\right]^T. \tag{23}$$

Therefore, $\bar{I}_F^{t+N_u}(\theta_G) = \tilde{I}_F^{t+N_u}(\theta_G) + \bar{I}_F^{t-1}(\theta_G)$. Defining

$$\Phi(\mathbf{u}) = [\Lambda_e^{-\frac{1}{2}} F_1 \mathbf{u}, \ \ldots, \ \Lambda_e^{-\frac{1}{2}} F_{n_\theta} \mathbf{u}] \in \mathbb{R}^{(N_u+1)n_y \times n_{\theta_G}}, \tag{24}$$

the Fisher information matrix can be written as:

$$\bar{I}_F^{t+N_u}(\theta_G) = \Phi(\mathbf{u})^T \Phi(\mathbf{u}) + \bar{I}_F^{t-1}(\theta_G). \tag{25}$$

Since $\Phi(\mathbf{u})$ is linear in $\mathbf{u}$, one can see that the information matrix is a quadratic function of the input sequence.

4.2 A Cyclic Algorithm

For simplicity we assume that the application cost function depends only on the plant model. Thus we can use the reduced information matrix in (10). Substituting (25) into the cost function in (10) and with some abuse of notation

$$J_t = \left\|\Phi(\mathbf{u})^T \Phi(\mathbf{u}) + \mathbf{C}(t-1) - S\right\|_F^2, \tag{26}$$

where $\mathbf{C}(t-1) = \bar{I}_F^{t-1}(\theta_0) - \frac{\chi_\alpha^2(n_\theta)\gamma}{2} V''_{app}(\theta_o)$ is a known matrix at time t which can be computed using data available at time t. The optimization problem (10) is non-convex and is in general hard to solve. However, the cost function is separable in terms of the variables, which makes it possible to find a solution of the problem through alternating algorithms (see e.g. Tropp et al. (2005) and Stoica et al. (2008)). To put it another way, we can break the problem into two smaller problems by considering only one of the variables, $\mathbf{u}$ and S, at each time. The resulting problems are easier to solve. This motivates us to propose a cyclic algorithm for this problem. The method alternates between optimizing the cost function using one of the variables while the other is fixed. Therefore, two main steps are allocated for the proposed algorithm.

Step1 Assuming S is fixed to its most recent optimal value, S_{opt}, we aim to solve the following optimization problem at time instant t:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \left\|\Phi(\mathbf{u})^T \Phi(\mathbf{u}) + \mathbf{C}(t-1) - S_{opt}\right\|_F^2 \\ \text{s.t.} \quad & \mathbf{u} = \left[(\bar{u}^*(t))^T \ \bar{u}(t)^T\right]^T \in \mathcal{U}, \\ & \mathbf{u}(k) = \bar{u}^*(t), \quad k = 1,\ldots,n-1, \\ & y(k) = G(q,\theta_0)u(k), \ y(k) \in \mathcal{Y}, \ k = t,\ldots,t+N_u. \end{aligned} \tag{27}$$

where $\mathbf{u}$ is defined in (23) and $\mathbf{u}(k)$ is the k^{th} element of $\mathbf{u}$. The optimization problem (27) is still not convex. However, the class of unconstrained signals, $\Phi(\mathbf{u})$, for which the cost function is zero, is (see Stoica et al. (2008)), $\Phi(\mathbf{u}) = U\left(S_{opt} - \mathbf{C}(t-1)\right)^{\frac{1}{2}}$, if

$$S_{opt} - \mathbf{C}(t-1) \geq 0, \tag{28}$$

where $U \in \mathbb{R}^{(N_u+1)n_y \times n_{\theta_G}}$ is a semi-unitary matrix. We will later show that the property (28) holds at time instant $t-1$. Hence, the problem (27) can be relaxed to

$$\begin{aligned} \min_{\mathbf{u},U} \quad & \left\|\Phi(\mathbf{u}) - U\left(S_{opt} - \mathbf{C}(t-1)\right)^{\frac{1}{2}}\right\|_F^2 \\ \text{s.t.} \quad & U^T U = I, \ \mathbf{u} = \left[(\bar{u}^*(t))^T \ \bar{u}(t)^T\right]^T \in \mathcal{U}, \\ & \mathbf{u}(k) = \bar{u}^*(t), \quad k = 1,\ldots,n-1, \\ & y(k) = G(q,\theta_0)u(k), \ y(k) \in \mathcal{Y}, \ k = t,\ldots,t+N_u. \end{aligned} \tag{29}$$

The cost function is still non-convex. However, this problem, in turn can be broken into two problems by considering only one of the variables and fixing the other one. Since $\Phi(\mathbf{u})$ is linear in $\mathbf{u}$ we will come up with two convex problems in terms of $\mathbf{u}$ and U. Therefore, we can again use a cyclic optimization algorithm in order to solve the problem. Here, we will use the minimization algorithm suggested in Stoica et al. (2008). The algorithm is alternating between the following two steps until convergence:

Step 1.1: Assuming U is fixed to its most recent optimal value, solve the problem (29) for $\mathbf{u}$, which is a constrained quadratic programming problem

$$\begin{aligned} \mathbf{u}_{opt} = \quad & \arg\min_{\mathbf{u}} \left\|\Phi(\mathbf{u}) - U_{opt}\left(S_{opt} - \mathbf{C}(t-1)\right)^{\frac{1}{2}}\right\|_F^2 \\ \text{s.t.} \quad & \mathbf{u} = \left[(\bar{u}^*(t))^T \ \bar{u}(t)^T\right]^T \in \mathcal{U}, \\ & \mathbf{u}(k) = \bar{u}^*(t), \quad k = 1,\ldots,n-1, \\ & y(k) \in \mathcal{Y} \quad k = t,\ldots,t+N_u, \\ & y(k) = G(q,\theta_0)u(k), \quad k = t,\ldots,t+N_u \end{aligned} \tag{30}$$

Step 1.2: Having the optimal input sequence, $\mathbf{u}_{opt}(t)$, find optimal U for (29) through Singular Value Decomposition (SVD), i.e,

$$\left(S_{opt} - \mathbf{C}(t-1)\right)^{\frac{1}{2}} \Phi(\mathbf{u}_{opt})^T = \bar{U}\Sigma\tilde{U}^T, \ U_{opt} = \tilde{U}\bar{U}^T \tag{31}$$

See Stoica et al. (2008) for more details.

Step2 Having obtained the optimal solution, $\mathbf{u}_{opt}(t)$, from the first step, we need to solve

$$\begin{aligned} \min_{S} \quad & \left\|\Phi(\mathbf{u}_{opt})^T \Phi(\mathbf{u}_{opt}) + \mathbf{C}(t-1) - S\right\|_F^2 \\ \text{s.t.} \quad & S \geq 0. \end{aligned} \tag{32}$$

An important advantage of the proposed algorithm is that we can find a closed-form solution for this step. The optimal solution of (32) is the projection of $\Phi(\mathbf{u}_{opt})^T\Phi(\mathbf{u}_{opt}) + \mathbf{C}(t-1)$ onto $\mathcal{S}^+_{n_\theta}$ (Henrion and Malick (2012)). To determine this projection note that since $\Phi(\mathbf{u}_{opt})^T\Phi(\mathbf{u}_{opt}) + \mathbf{C}(t-1)$ is symmetric, we can write

$$\Phi(\mathbf{u}_{opt})^T\Phi(\mathbf{u}_{opt}) + \mathbf{C}(t-1) = V \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_\theta}) V^T, \quad (33)$$

where λ_i are the eigenvalues and V is the corresponding orthonormal matrix of eigenvectors. Thus

$$S_{opt} = V \operatorname{diag}\big(\max(0,\lambda_1), \ldots, \max(0,\lambda_{n_\theta})\big) V^T. \quad (34)$$

See Henrion and Malick (2012) for further information. Note that $S \geq \Phi(\mathbf{u}_{opt})^T\Phi(\mathbf{u}_{opt}) + \mathbf{C}(t-1)$ according to (34), which confirms that the property (28) holds.

As mentioned before, the proposed alternating method in this paper cycles between Step 1 and Step 2. The resulting problem only involves solving a quadratic optimization problem, an SVD of a matrix with size n_θ and a projection and thus it is fast enough to address large problems. The method is summarized in the table below [1].

Algorithm Proposed Alternating Method

Initialization: choose N_u and n, $S_{opt} \leftarrow 0$,
$U_{opt} \leftarrow U_{init}$, $t \leftarrow 1$ and $J_0 \neq 0$
while $J_{t-1} \neq 0$ **do**
  i = 1
  while {Stopping criteria is not true} **do**
    Start Step 1:
      Solve (30) and $\mathbf{u}^i_{opt} \leftarrow \mathbf{u}_{opt}$
      Use $\mathbf{u}^i_{opt}$ and (31) to compute U_{opt}
      $U^i_{opt} \leftarrow U_{opt}$
    Start Step 2:
      Use $\mathbf{u}^i_{opt}$, U^i_{opt} and (33)-(34) to obtain S_{opt}
      $S^i_{opt} \leftarrow S_{opt}$
    $i \leftarrow i + 1$
  end while
  $u^*(t) \leftarrow$First sample of the optimal input signal
  Calculate $\bar{I}^t_F$, $\mathbf{C}(t)$ and J_t
  $t \leftarrow t + 1$
end while
return optimal input sequence $\{u^*(k)\}_{k=1}^{t+N_u}$

5. FIR EXAMPLE

To get a better insight into the proposed approach, we study it for a simple Finite Impulse Response (FIR) model

$$\begin{aligned} y(t,\theta) &= \theta_1 u(t-1) + \theta_2 u(t-2) + e(t), \\ \mathbb{E}\{e(t)\} &= 0 \;,\; \mathbb{E}\{e(t)^2\} = \lambda, \end{aligned} \quad (35)$$

where $\theta = [\theta_1 \;\; \theta_2]$. We aim to design an optimal input sequence with minimum length such that the identified model based on the obtained input signal can guarantee a desired control performance when it is being used in a controller. Moreover, assume that according to some physical restrictions we need $|u(t)| \leq u_{max}$, $|y(t)| \leq y_{max}$. We use the following steps.

Define desired control performance: One reasonable choice of a desired performance for the controller is the difference between the measured output when the controller is working based on the true parameters, θ_o, and when it is working based on the estimated parameters, $\hat{\theta}$, i.e.,

$$V_{app}(\hat{\theta}) = \frac{1}{N}\sum_{t=1}^{N} \left\| y(t,\theta_o) - y(t,\hat{\theta}) \right\|^2, \quad (36)$$

over a step response of the system with the controller running. Since for (36), we have $V_{app}(\theta_o) = 0$ and $V'_{app}(\theta_o) = 0$, we can approximate the set (3) by (4). The Hessian matrix can be calculated through either numerical or analytical methods, depending on the type of controller. Having defined $V_{app}(\theta_o)$, we aim to design an input sequence such that (7) is fulfilled for a given γ.

Input design: The signal generation is done through the optimization problem (10). We first need to find the Fisher information matrix. Considering (35), we attain

$$\begin{aligned} \epsilon(t) &= y(t) - \theta_1 u(t-1) - \theta_2 u(t-2), \\ \frac{\partial \epsilon^T(t,\theta)}{\partial \theta} &= -\begin{bmatrix} q^{-1}u(t) \\ q^{-2}u(t) \end{bmatrix}. \end{aligned} \quad (37)$$

Assume we are at time instant t and we aim to optimize the input signal in the prediction horizon of length N_u, putting $n = 3$, we can write (20) and (24) as

$$\bar{u}^*(t) = [u^*(t-2), u^*(t-1)], \;\; \bar{u}(t) = [u(t), \ldots, u(t+N_u)],$$
$$\mathbf{u} = [u^*(t-2), u^*(t-1), u(t), \ldots, u(t+N_u)],$$
$$\Phi(\mathbf{u}) = \frac{1}{\sqrt{\lambda}}[F_1\mathbf{u}, \;\; F_2\mathbf{u}],$$

where F_1 and F_2 are obtained using (19). For example choosing $N_u = 4$, we have

$$F_1 = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix},$$

and thus

$$\tilde{I}_F^{t+N_u}(\theta) = \frac{1}{\lambda}\sum_{k=t}^{t+N_u} \begin{bmatrix} u(k-1)u(k-1) & u(k-1)u(k-2) \\ u(k-2)u(k-1) & u(k-2)u(k-2) \end{bmatrix}. \quad (38)$$

The information matrix for the FIR system is determined by the covariances of input sequences (See Stoica and Söderström (1982)). We are now ready to find the optimal input signal, $\bar{u}(t)$, using the proposed alternating method.

6. NUMERICAL RESULTS

Consider the FIR example in Section 5, with $\theta_o = [10, \; -9]$, $u_{max} = 0.5$, $y_{max} = 5$, $N_u = 5$. Assume that we want to generate an input sequence of length $N = 100$ that when used in an system identification experiment satisfies both the application requirements and the input and output constraints. The identified model will be used in MPC, with cost function $J = \sum_{k=0}^{N_y} \|y(k+1) - r(k+1)\|^2$, the same input and output constraints as during the experiment and $r = 0$. We approximate the application set employing the method in Ebadat et al. (2014). The required accuracy is $\gamma = 100$ and we want that the estimated parameters lie in the identification set with probability $\alpha = 0.95$. The suggested method is used to obtain an optimal input sequence. For the obtained input the slack variable S, is strictly positive definite, and thus the experiment design constraint is satisfied. The application and identification ellipsoids for the obtained input are shown in Figure 1. The generated input signal has been used in

[1] One possibility for stopping criteria is to stop the iterations when the tolerance of changes in the variables is small enough.

the system identification experiment with zero mean white Gaussian noise $e(t)$ with variance $\lambda = 1$. One hundred $\hat{\theta}_N$ are estimated based on the measurements of $y(t)$, when the obtained input signal is applied to the system. To this aim the system identification toolbox in Matlab is used. In total 95% of the estimated parameters are inside the identification ellipsoid. The results are shown in Figure 1. It can be seen that $\mathcal{E}_{SI}$ is inside $\mathcal{E}_{app}$, thus, the performance requirement is fulfilled by the estimated parameters with probability 95%. The generated input has been shown in

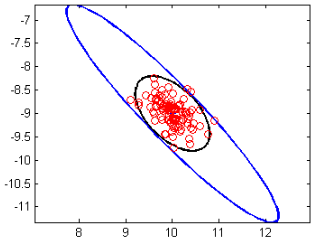

Fig. 1. $\mathcal{E}_{app}$ is the outer ellipse, $\mathcal{E}_{SI}$ is the inner ellipse and $\hat{\theta}_N$ are the small circles.

Figure 2. It can be seen that it satisfies the constraint.

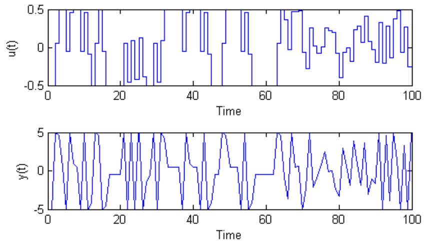

Fig. 2. Generated optimal input (top) and output (bottom) signals.

We also formulate the problem in the frequency domain using (8), where the input power is chosen to be the experimental cost. We use MOOSE, a toolbox for optimal input design implemented in MATLAB (Larsson and Annergren (2011)), in order to solve problem (8). Result is shown in Figure 3. However, as mentioned before what we are getting out of solving (8) is an input spectrum and we need to find the corresponding time realization by another optimization problem which is not an easy problem under input and output constraints (see Larsson et al. (2013)).

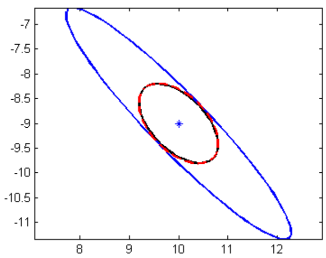

Fig. 3. $\mathcal{E}_{app}$ is the outer ellipse. The identification ellipse obtained by MOOSE is shown in red ('- -') and the one obtained by the proposed method in black ('–').

7. CONCLUSION

In this paper we introduced a new approach to generate input signals such that the estimated model based on the generated signal can guarantee a desired control performance. The method is based on satisfying a lower bound on the Fisher information matrix. The experimental cost is considered to be the minimum required time for satisfying the lower bound. One significant feature of the proposed approach is that the problem is formulated in time-domain and thus it is straightforward to handle constraint on the amplitude of the input and output signals. The problem is, however, highly non-convex. This is addressed through alternating optimization methods, where we are alternating between optimizing cost function for each variable while the others are fixed. We perform a time recursive algorithm and each optimization problem is solved in a receding horizon framework.

Future research directions include extending the method to the closed loop system identification and integrate it to MPC and we aim to design optimal input while at the same time we are concerning about the control performance.

REFERENCES

Bombois, X., Scorletti, G., Gevers, M., Van den Hof, P., and Hildebrand, R. (2006). Least costly identification experiment for control. *Automatica*, 42(10), 1651–1662.

Ebadat, A., Annergren, M., Larsson, C.A., Rojas, C.R., Wahlberg, B., Hjalmarsson, H., Molander, M., and Sjberg, J. (2014). Application set approximation in optimal input design for model predictive control. In *13th European Control Conference (ECC'14)*. Strasbourg, France, arXiv:1303.5199.

Fedorov, V. (1972). Theory of optimal experiments. *Academic Press*.

Gevers, M. and Ljung, L. (1986). Optimal experiment design with respect to intended model application. *Automatica*, 22, 543–554.

Goodwin, G. and Payne, R. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press.

Gujar, U. and Kavanagh, R. (1968). Generation of random signals with specified probability density functions and power density spectra. *IEEE Transactions on Automatic Control*, 13, 716–719.

Henrion, D. and Malick, J. (2012). Projection methods in conic optimization. *Handbook on Semidefinite, Conic and Polynomial Optimization*, 166, 565–600.

Hjalmarsson, H. (2005). From experiment design to closed-loop control. *Automatica*, 41(3), 393–438.

Hjalmarsson, H. (2009). System Identification of Complex and structured systems. *European Journal of Control*, 15, 275–310.

Larsson, C.A., Annergren, M., and Hjalmarsson, H. (2011). On Optimal Input Design in System Identification for Model Predictive Control. In *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 805–810.

Larsson, C.A., Hägg, P., and Hjalmarsson, H. (2013). Generation of excitation signals with prescribed autocorrelation for input and output constrained systems. In *European Control Conference*.

Larsson, C. and Annergren, M. (2011). Moose: Model based optimal input signal design toolbox for matlab. Technical report, Royal Institute of Technology (KTH).

Liu, B. and Munson, D.J. (1982). Generation of a random sequence having a jointly specified marginal distribution and autocovariance. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 30(6), 973–983.

Ljung, L. and Wahlberg, B. (1992). Asymptotic properties of least-squares method for estimating transfer functions and disturbance spectra. *Advances in Applied Probability*, 24(2), 412–440.

Ljung, L. (1999). *System Identification - Theory for the user*. Prentice Hall, New Jersey, 2nd edition.

Manchester, I.R. (2010). Input Design for System Identication via Convex Relaxation.pdf. In *49th IEEE Conference on Decision and Control (CDC 2010)*. USA.

Schroeder, M. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation. *Information Theory, IEEE transactions on*, 16, 85–89.

Stoica, P. and Söderström, T. (1982). A useful input parameterization for optimal experiment design. *IEEE Transactions on Automatic Control*, 27(4), 986–989.

Stoica, P., Li, J., Zhu, X., and Bin Guo (2008). Waveform Synthesis for Diversity-Based Transmit Beampattern Design. *IEEE Transactions on Signal Processing*, 56(6), 2593–2598.

Tropp, J.A., Dhillon, R.W., and Strohmer, T. (2005). structured tight frames via an alternating projection method. *IEEE Transactions on Information Theory*, 51, 188 – 209.