An application-oriented approach to dual control with excitation for closed-loop identification

Christian A. Larsson, Afrooz Ebadat, Cristian R. Rojas, Xavier Bombois, Håkan Hjalmarsson

▶ **To cite this version:**

Christian A. Larsson, Afrooz Ebadat, Cristian R. Rojas, Xavier Bombois, Håkan Hjalmarsson. An application-oriented approach to dual control with excitation for closed-loop identification. European Journal of Control, 2016, 29, pp.1-16. 10.1016/j.ejcon.2016.03.001 . hal-01355020

HAL Id: hal-01355020
https://hal.science/hal-01355020

Submitted on 25 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

An application-oriented approach to dual control with excitation for closed-loop identification

Christian A. Larsson[a,*], Afrooz Ebadat[a], Cristian R. Rojas[a], Xavier Bombois[b], Håkan Hjalmarsson[a]

[a] *ACCESS Linnaeus Center and Department of Automatic Control, Osquldas väg 10, 100 44 Stockholm, KTH Royal Institute of Technology, Sweden*
[b] *Groupe Automatique, Commande et Mécatronique, Département Méthodes pour l'Ingénierie des Systèmes, Laboratoire Ampère UMR CNRS 5005, Ecole Centrale de Lyon, France*

Abstract

System identification of systems operating in closed loop is an important problem in industrial applications, where model-based control is used to an increasing extent. For model-based controllers, plant changes over time eventually result in a mismatch between the dynamics of any initial model in the controller and the actual plant dynamics. When the mismatch becomes too large, control performance suffers and it becomes necessary to re-identify the plant to restore performance. Often the available data are not informative enough when the identification is performed in closed loop and extra excitation needs to be injected. This article considers the problem of generating such excitation with the least possible disruption to the normal operations of the plant. The methods explicitly take time domain constraints into account. The formulation leads to optimal control problems which are in general very difficult optimization problems. Computationally tractable solutions based on Markov decision processes and model predictive control are presented. The performance of the suggested algorithms is illustrated in two simulation examples comparing the novel methods and algorithms available in the literature.

Keywords: Closed-loop identification, Input design, System identification, Constrained systems, Markov decision process, Model predictive control

1. Introduction

Most modern control design approaches are model based. For example, in process industry, model predictive control (MPC) has more or less become the industry standard for control of constrained MIMO systems. Any control system has performance requirements and whenever controllers are model based, the quality of the model influences the achievable performance. This means that the implementation of model-based controllers often requires significant modeling efforts for the control to be successful. This modeling is often done using system identification and a lot of time and resources are spent on the initial, commissioning identification. However, even if the initial model gives satisfactory control, changes in the process dynamics over time can result in reduced performance and the need for re-identification to restore performance.

It is well known that identification in closed loop can cause problems due to lack of excitation in the input. The reason for this is that often the regulating properties that one desires from the controller are in conflict with the excitation properties of the signals needed for identification. The compromise between these properties has lead to the study of dual control introduced by Feldbaum (1960–61). This has been recognized in the MPC community and several MPC formulations where a dual effect is included in the input have appeared in the literature. One of the earliest seems to have been proposed by Genceli and

*Corresponding author. Tel: +46 8 790 73 22
Email addresses: christian.larsson@ee.kth.se (Christian A. Larsson), ebadat@kth.se (Afrooz Ebadat), crro@kth.se (Cristian R. Rojas), xavier.bombois@ec-lyon.fr (Xavier Bombois), hjalmars@kth.se (Håkan Hjalmarsson)

Nikolaou (1996). Similar ideas have later been proposed by Aggelogiannaki and Sarimveis (2006), Marafioti (2010) and Rathouský and Havlena (2011). They all propose amending the MPC with a constraint that renders the input persistently exciting over some horizon. This ensures that the closed-loop data can be used to estimate models consistently.

The choice of the input signal used in the identification is a very important one. A badly designed input signal could potentially ruin the (in other aspects) most well prepared identification experiment. Conversely, a carefully chosen input signal could reduce the experimental effort needed to get a certain accuracy of the identified model and simplify the system identification problem per se. This understanding has led to the growth of the branch of input or experiment design in system identification. Early contributions were thanks to, for example, Fedorov (1972), Mehra (1974) and Goodwin and Payne (1977). Later ideas, where the intended model use is taken into account, have been developed by Gevers and Ljung (1986), Hjalmarsson et al. (1996) and Bombois et al. (2006), to name a few.

This article considers the problem of optimal experiment design for system identification of constrained systems operating in closed loop. The central idea is that the quality of the identified model should be high enough to give good performance when the model is used in a controller. It is also desirable that the cost of the system identification experiment is as low as possible. The cost of an experiment depends on the application but can, for example, be specified in terms of disruption of normal operations, or the time of the experiment.

The goal of the article is to present the optimal experiment design problem and the theoretical and practical challenges that this problem entails. The article also presents suitable approximations that can be made to arrive at computationally tractable and practically implementable formulations. The problem is initially formulated as a general optimal control problem. This problem has a nice solution for linear systems without time domain signal constraints, but is computationally intractable in general. Using the framework of constrained Markov decision processes, the problem is then formulated for systems with finite state and action spaces. Although this formulation is very general, the size of the resulting optimization problem makes it computationally demanding and practical applicability is limited. Therefore, the problem is further simplified using a receding horizon formulation. This has been a successful strategy to approximate optimal control problems with constraints in many applications. For completeness, existing receding horizon formulations are also presented. In going from the most general formulation, through a series of simplifying assumptions and approximations, the challenges of the optimal experiment design problem are highlighted.

The presented ideas and algorithms explicitly take the intended model use into account by using ideas from the *application-oriented input design* framework (Hjalmarsson, 2009), which in turn is part of the least costly identification paradigm (Bombois et al., 2006). These frameworks typically result in less disruption from normal operations compared to imposing persistence of excitation.

1.1. Organization of the article

The remaining article is organized as follows. Section 2 presents the necessary mathematical background. In Section 3, the general control formulation with excitation for closed loop re-identification is introduced. Section 4 presents an MDP formulation of the problem in a general setting. In Section 5, two MPC based controllers for systems of output error type are introduced. Section 6 discusses some of the related, earlier approaches based on the idea of adding persistence of excitation to the input. Section 7 illustrates the performance of the algorithms in simulations. Finally, Section 8 concludes the article and points to some future research directions.

1.2. Notation

The symbol $\mathbb{P}\{\cdot\}$ denotes the probability of an event and $\mathbb{E}\{\cdot\}$ denotes the expectation operator for the probability spaces generated by the relevant stochastic processes. The real numbers are denoted $\mathbf{R}$. Matrices are capital letters, e.g. X, Y, and vectors are small letters, e.g. x, y. For symmetric matrices, $X \succeq 0$ means that X is positive semidefinite and for two symmetric matrices X and Y, $X \succeq Y$ means $X - Y \succeq 0$. For a vector x and matrix Q, the notation $\|x\|_Q \triangleq \sqrt{x^T Q x}$ is used.

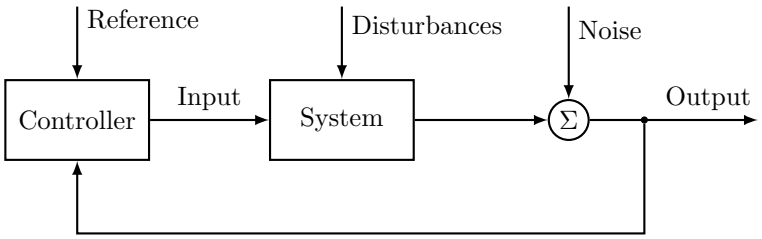

Figure 1: A general closed loop system. The system is driven by the input and disturbance processes and the controller choses the input based on feedback measurements and reference signals. Typically, the feedback is corrupted by measurement noise.

2. Preliminaries

This article deals with methods for input design for closed-loop system identification. The system, model and input design theory and assumptions used in the development of the methods are introduced here. The general closed-loop setup is shown in Figure 1.

2.1. System and model

Consider a linear, time invariant, discrete time multiple-input, multiple-output systems modeled by

$$\mathcal{M}(\theta): \begin{cases} x_{t+1} = A(\theta)x_t + B(\theta)u_t + K(\theta)v_t, \\ y_t = C(\theta)x_t + v_t, \end{cases} \tag{1}$$

where $x_t \in \mathbf{R}^n$ is the state, $u_t \in \mathbf{R}^m$ is the input, $v_t \in \mathbf{R}^p$ represents the effects of disturbances and noise and is a zero mean white sequence with covariance Λ_v, and $y_t \in \mathbf{R}^p$ is the measured output. This model class is known as innovations models and covers, e.g., ARMAX and Box-Jenkins transfer function models (Ljung, 1999). The model is parameterized by the unknown vector $\theta \in \mathbf{R}^{n_\theta}$. It is assumed that there exists a vector θ_o such that the true system $\mathcal{S}$ is given by $\mathcal{S} = \mathcal{M}(\theta_o)$.

The state and output of the system $\mathcal{S}$ can be predicted using the standard Kalman filter, given by

$$\begin{cases} \hat{x}_{t+1|t} = A(\theta_o)\hat{x}_{t|t-1} + B(\theta_o)u_t + K(\theta_o)(y_t - C(\theta_o)\hat{x}_{t|t-1}), \\ \hat{y}_{t|t-1} = C(\theta_o)\hat{x}_{t|t-1}. \end{cases} \tag{2}$$

It is well known that for Gaussian noise, this estimate is the mean square optimal predictor. For the general case, the Kalman filter gives the linear least squares estimate (see Söderström, 2002, for example).

2.2. Controller

The input to the system, u_t, is generated by the controller. The controller decides on the choice of the input at a given time instant, u_t, according to a control rule, denoted π_t. A sequence of such control rules is called a *policy* and is denoted $\pi = (\pi_1, \pi_2, \ldots)$.

There is an instantaneous cost, $c_t(x, u)$, related to the system being in a state x and using a certain control u. Based on the instantaneous costs c_t, the expected average costs can be defined, for a finite control horizon T as

$$C^\pi = \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}^\pi \left\{c_t(x_t, u_t)\right\}, \tag{3}$$

and for an infinite control horizon as

$$C^\pi = \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}^\pi \left\{c_t(x_t, u_t)\right\}, \tag{4}$$

where $\mathbb{E}^\pi$ is the expected value under policy π.

In many situations, it is desired to control the system such that the state and/or the output remain in some convex constraint set, i.e., $x_t \in \mathcal{X} \subseteq \mathbf{R}^n$ and $y_t \in \mathcal{Y} \subseteq R^p$. Furthermore, the input is also often bounded, i.e., $u_t \in \mathcal{U} \subset \mathbf{R}^m$, where $\mathcal{U}$ is also assumed to be convex. In such cases, the controller needs to ensure that these constraints are satisfied.

From these considerations, a general optimal control problem for the system in Figure 1 can be formulated as

$$\min_\pi \left\{ C^\pi \;\middle|\; \begin{array}{l} x_{t+1} = A(\theta_o)x_t + B(\theta_o)u_t + K(\theta_o)v_t, \\ \quad y_t = C(\theta_o)x_t + v_t \\ \quad u_t \in \mathcal{U}, x_t \in \mathcal{X}, y_t \in \mathcal{Y} \end{array} \right\}. \tag{5}$$

Åström (1970) gives a good introduction to optimal control. The problem (5) is often a very difficult one and analytic solutions can be obtained only in exceptional cases. The most important cases are arguably state or output feedback of unconstrained, linear system with Gaussian noise, which result in the well known LQ and LQG regulators. For the constrained case, model predictive control has become a widely used approximate solution to (5); Maciejowski (2002) gives a treatment of the basic theory. The same type of problems are considered in the Markov decision process framework. Classically, for discrete state and action spaces, while extensions to more general spaces exist. The book by Puterman (1994) introduces the subject of MDPs well.

2.3. System identification

The parameters of the model (1) are estimated using data collected from the closed-loop system in Figure 1. The fundamental difference between open- and closed-loop identification is the unavoidable correlation between the input u_t and the noise, v_t, in closed loop. Closed-loop data also often contain less information due to the robustness introduced by feedback. There are, however, situations where closed loop identification is desirable. Agüero and Goodwin (2007) and Bombois et al. (2005) discuss situations where open- or closed-loop identification is advantageous. When there are input constraints, open-loop identification is as good as closed-loop identification. On the other hand, when there are output constraints, closed-loop experiments are typically favorable. The theoretical considerations aside, there may be practical reasons why closed-loop identification is necessary. For example, the system may be unstable and therefore need to be stabilized during the identification. There may also be economic reasons for keeping the loop closed. Finally, the controller can be used to maintain signals within constraints, when necessary.

The approaches for closed-loop identification can be categorized depending on the assumptions on the nature of the feedback in the system. The *direct approach* uses the input and the measured output to directly identify the open-loop system assuming no knowledge of the feedback. The *indirect method* uses the reference and output to identify the closed-loop system. The open-loop model can be calculated from the closed-loop model assuming that the controller is known. The *joint input–output approach* considers the input and output as output from a system driven by the reference and noise. The open-loop model can be calculated assuming that the structure of the controller is known. In this article, all identification is assumed to be performed using the direct approach.

2.3.1. Prediction error identification

The prediction error method (PEM) can be used in the direct approach to estimate the parameters using the data $Z_N = \{u_t, y_t\}_{t=1}^N$. For example, Forssell and Ljung (1999) discuss closed-loop identification using PEM, and Ljung (1999) has a detailed treatment of PEM in general. The resulting parameter estimates are denoted $\hat{\theta}_N$ and are determined as

$$\hat{\theta}_N = \arg\min_\theta \frac{1}{N} \sum_{t=1}^N \ell(\hat{y}_{t|t-1}(\theta) - y_t), \tag{6}$$

where $\ell(\cdot)$ is a scalar valued function. For the choice

$$\ell\left(\hat{y}_t(\theta) - y_t\right) = \left(\hat{y}_t(\theta) - y_t\right)^T \Lambda_v^{-1} \left(\hat{y}_t(\theta) - y_t\right), \tag{7}$$

it has been proven by Ljung and Caines (1979) that, under very general assumptions, the PEM estimate is, asymptotically in the data length, distributed as

$$\left(\mathcal{I}_1^N(\theta_o)\right)^{1/2} \left(\hat{\theta}_N - \theta_o\right) \sim \mathcal{N}(0, I), \tag{8}$$

$$\mathcal{I}_1^N(\theta_o) \triangleq \sum_{t=1}^{N} \mathbb{E}\left\{\psi_t(\theta_o)\Lambda_v^{-1}\psi_t(\theta_o)^T\right\}, \tag{9}$$

$$\psi_t^i(\theta) \triangleq \frac{\partial \hat{y}_{t|t-1}(\theta)}{\partial \theta_i}, \tag{10}$$

$$\psi_t(\theta) \triangleq \begin{bmatrix}\psi_t^1(\theta) & \cdots & \psi_t^{n_\theta}(\theta)\end{bmatrix}^T \tag{11}$$

where $\mathcal{I}_1^N$ is the information matrix. The fact that the data are obtained in closed loop are of no importance for the consistency and accuracy of the estimates as long as the true system $\mathcal{S}$ can be captured by $\mathcal{M}(\theta)$. In particular, capturing the true noise properties is central for unbiased estimation of the open loop system.

2.4. Experiment design

From (2) and (9)–(11), it can be seen that the information matrix and hence also the covariance of the estimate can be influenced by the input used in during the identification experiment. This is exploited in optimal experiment design. The early experiment design for dynamic systems focused on optimality criteria related to the covariance matrix of the estimates (Fedorov, 1972; Mehra, 1974; Goodwin and Payne, 1977). The problem was either formulated in time domain by optimizing an input sequence, which leads to a nonlinear optimal control problem, or in frequency domain by optimizing the input spectrum. Since then, the focus has shifted to considering model quality in terms of the application of the estimated model. First in the *identification for control* ideas (Gevers and Ljung, 1986; Gevers, 1991; Hjalmarsson et al., 1996), and later in the *least costly experiment design* paradigm (Bombois et al., 2006).

Contemporary experiment design for dynamic systems typically includes the minimization of an *identification cost* subject to constraints on the quality of the estimated model. The identification cost can be, for example, the experiment length or extra excitation power needed. The problem is most often formulated in the frequency domain for open- or closed-loop problems. Much of the work is focused on linear systems and controllers assuming (quasi-)stationary signals. In this setting, the input is given by

$$u_t = r_t - C(q)y_t, \tag{12}$$

where $C(q)$ is the transfer function of the controller and r_t is the reference, *cf.* Figure 1. The available experiment design variables are the spectrum of the reference signal and the controller. Research has focused on formulations of the optimization problem as semidefinite programs. In principle, two approaches exist, depending on the parameterization of the input spectrum. One possibility is a *finite-dimensional spectrum parameterization*, which restricts the input spectrum to lie in a finite-dimensional subspace of the space of all spectra and therefore may result in a suboptimal solution. On the other hand, frequency-by-frequency constraints can be included in the design. The other is a *partial correlation parameterization*, which parameterizes the entire space of spectra in terms of a finite number of generalized moments and therefore does not result in suboptimal solutions. Jansson and Hjalmarsson (2005) present a unified framework for convex optimization formulations of input design problems based on finite parameterizations. They also show how a number of quality constraints can be included in the design using a finite-dimensional spectrum parameterization of the input spectrum. The joint reference–controller input design problem, when the input is given by (12), is solved for finite-dimensional parameterization by Hjalmarsson and Jansson (2008) and for the partial correlation parameterization by Hildebrand et al. (2010). Hildebrand and Gevers (2013) further extend the partial correlation parameterization to MIMO systems and simplify the solution procedure by considering a central extension of the optimal generalized moments.

2.5. Application-oriented input design

The application-oriented input design idea (Bombois et al., 2006; Hjalmarsson, 2009), is used as a basis for the development of the new control algorithms in this article. The central idea is that the quality of the model influences the performance of a control application where the model is used.

The application-oriented experiment design formulation relies on an application cost as a measure of performance degradation due to mismatch between model and system. Denote the application cost by V_{app}. For parameterized models such as (1), V_{app} can be chosen as a function of the parameter vector θ. If the true system parameters were known, the control design should result in the best possible performance, while using other parameter values may degrade the performance. From these observations, it is assumed that $V_{\text{app}}(\theta) \geq 0$ and that $V_{\text{app}}(\theta_o) = 0$ is the global minimum. A model is considered acceptable if the degradation is sufficiently small. This can be specified as an upper bound on the application cost. The model corresponding to parameter vector θ is acceptable if

$$V_{\text{app}}(\theta) \leq \frac{1}{\gamma}, \tag{13}$$

where γ is an application specific constant that relates to the required accuracy of the model. The choice of the application cost and the corresponding γ is highly application specific. By approximating $V_{\text{app}}(\theta)$ using a second order Taylor expansion around the true parameter values, and the assumptions that $V_{\text{app}}(\theta_o) = V'_{\text{app}}(\theta_o) = 0$, a set of acceptable parameters can be specified as

$$\mathcal{E}_{\text{app}}(\gamma) \triangleq \left\{ \theta \mid [\theta - \theta_o]^T V''_{\text{app}}(\theta_o)[\theta - \theta_o] \leq \frac{2}{\gamma} \right\}. \tag{14}$$

From the distribution of the estimates (8) it holds, asymptotically in N, that

$$\hat{\theta}_N \in U(\alpha) \triangleq \left\{ \theta \mid [\theta - \theta_o]^T \mathcal{I}_1^N(\theta_o) [\theta - \theta_o] \leq \chi^2_\alpha(n_\theta) \right\}, \tag{15}$$

with probability α. Here $\chi^2_\alpha(n_\theta)$ is the α-percentile of the χ^2-distribution with n_θ degrees of freedom. $U(\alpha)$ is a standard confidence ellipsoid for a Gaussian distributed random variable (Ljung, 1999).

The aim of the application-oriented input design is to find an input that with high probability, α, results in acceptable parameter estimates while at the same time it minimizes the cost of the identification experiment, i.e.,

$$\begin{aligned} &\underset{\text{input sequence}}{\text{minimize}} && \text{Experimental cost} \\ &\text{subject to} && P\left\{ V_{\text{app}}(\hat{\theta}_N) \leq \frac{1}{\gamma} \right\} \geq \alpha. \end{aligned} \tag{16}$$

The problem (16) is, however, in general nonconvex and not computationally tractable. Hjalmarsson (2009) suggests approximating the chance constraint in (16) by

$$U(\alpha) \subseteq \mathcal{E}_{\text{app}}(\gamma). \tag{17}$$

This approximation can in turn be approximated by (Hjalmarsson, 2009)

$$\mathcal{I}_1^N(\theta_o) \succeq \frac{\gamma \chi^2_\alpha(n_\theta)}{2} V''_{\text{app}}(\theta_o), \tag{18}$$

which will be refer to as the *experiment design constraint*. The expression (18) requires calculation of the Hessian of the application cost. This often requires numerical calculations and a simulations based scheme for this is given by Ebadat et al. (2014a).

The application-oriented input design above requires knowledge of the true system parameters, θ_o. This is a problem in all experiment design for system identification: the optimal design depends on the true system. In practice, one does of course not know θ_o; if the parameters were known, no identification would be needed. Therefore, one has to rely on what is known about the system before the identification, for example, an initial estimate of the parameter values. This approach will be taken later in the article. An alternative is adaptive input design (Gerencsér et al., 2009).

2.6. Information matrix

Since the experiment design constraint is used to design the excitation for the identification, it needs to be related to the input signal which is generated by the controller. The information matrix can be calculated from an extended state space description based on the sensitivites of the predictor (2) with respect to the parameters θ (Ljung and Söderström, 1983, for example). Denote this extended state space model

$$\begin{cases} \xi_{t+1} = \mathcal{A}(\theta)\xi_t + \mathcal{B}(\theta)u_t + \mathcal{K}(\theta)v_t, \\ \bar{\psi}_t = \mathcal{C}(\theta)\xi_t, \end{cases} \tag{19}$$

where

$$\begin{aligned} \xi_t &\triangleq \begin{bmatrix} x_t & \hat{x}_{t|t-1} & \frac{\partial \hat{x}_{t|t-1}}{\partial \theta_1} & \cdots & \frac{\partial \hat{x}_{t|t-1}}{\partial \theta_{n_\theta}} \end{bmatrix}^T, \\ \bar{\psi}_t &\triangleq \begin{bmatrix} \psi_t^1(\theta)^T & \cdots & \psi_t^{n_\theta}(\theta)^T \end{bmatrix}^T. \end{aligned}$$

The expressions for the matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{K}$ are given in Appendix B. Using (19), the terms of $\mathcal{I}_1^N$ can be found from

$$\mathbb{E}\left\{\psi_t(\theta)\Lambda_v^{-1}\psi_t(\theta)^T\right\}_{(i,j)} = \mathbb{E}\left\{\psi_t^i(\theta)^T\Lambda_v^{-1}\psi_t^j(\theta)\right\} = \operatorname{trace}\left[\mathbb{E}\left\{\psi_t^j(\theta)\psi_t^i(\theta)^T\right\}\Lambda_v^{-1}\right]. \tag{20}$$

For a given data record, Z_N, the Fisher information matrix can be estimated by

$$I_1^N(\theta) = \sum_{t=1}^{N} \psi_t(\theta)\Lambda_v^{-1}\psi_t(\theta)^T. \tag{21}$$

In fact, it holds under fairly mild conditions that

$$\lim_{N\to\infty} \frac{1}{N} I_1^N(\theta) = \lim_{N\to\infty} \frac{1}{N}\mathcal{I}_1^N(\theta), \text{ almost surely.} \tag{22}$$

When the limit exists,

$$\mathcal{I}(\theta) \triangleq \lim_{N\to\infty} \frac{1}{N}\mathcal{I}_1^N(\theta), \tag{23}$$

is the average per-sample information matrix.

3. Problem formulation

The problem considered in this article is the following. Consider that the system $\mathcal{S}$ is operating in closed loop with a controller based on the estimated model $\mathcal{M}(\hat{\theta})$ given by (1). Initially, the controller is such that the closed-loop performance is satisfactory, that is $V_{\text{app}}(\hat{\theta}) \leq \frac{1}{\gamma}$. At some point it is determined that due to a mismatch between the model $\mathcal{M}(\hat{\theta})$ and the plant, the closed-loop performance is no longer satisfactory. Therefore, a re-identification experiment is performed to restore the closed-loop performance. The N samples long experiment should be such that:

1. Closed-loop performance is restored (with high probability) when the new model is used in the controller, that is $V_{\text{app}}(\hat{\theta}_N) \leq \frac{1}{\gamma}$ with high probability.
2. The extra cost incurred due to the identification experiment is as small as possible.
3. Any constraints are respected during the identification.

Assuming that the direct method for identification is used together with PEM, the first two requirements on the identification experiment can be handled using the application-oriented approach in Section 2.5. The third requirement can be handled by explicitly including constraints in the input design formulation.

Measuring performance can be done in many different ways. Performance of the controller may be related to, for example, closed-loop properties such as a desired sensitivity function. For performance in terms of product quality, variance related criteria such as the index introduced by Harris (1989) can be considered. Performance may also related to the economic gain of using a certain controller (Modén and Lundh, 2013). Zagrobelny et al. (2013) give a good overview of MPC performance monitoring, but it should already be clear that there is no clear cut answer to what constitutes *good performance* and that it is highly application specific. Here it is assumed that the performance can be quantified by an application cost with the properties discussed in Section 2.5.

Two principal situations are considered for the cost of the identification experiment (Bombois et al., 2006). The first is the situation where the cost of the experiment relates to the degradation of product quality. Therefore, it is desired that the control cost is kept small during the identification. The second is the situation where the time of the experiment is the crucial cost. It is acceptable to sacrifice some product quality, as long as this results in a shorter experiment. The experiment designs for the two situations can be formulated as:

1. With the least possible effect on the control cost C^π, during an experiment of fixed length N, excite the system enough to accurately model the important system dynamics needed for acceptable control performance.
2. During the shortest possible experiment time N, excite the system as much as possible, under given constraints, to accurately model the important system dynamics needed for acceptable control performance.

Consider the system $\mathcal{S}$ modeled by $\mathcal{M}(\theta)$ in (1) and operating under the constraints

$$x \in \mathcal{X}, \qquad u \in \mathcal{U}, \qquad y \in \mathcal{Y}. \tag{24}$$

The application-oriented experiment design problem can then be formulated as the finite-horizon optimal control problem

$$\min_u \left\{ J \;\middle|\; \begin{array}{l} x_{t+1} = A(\theta_o)x_t + B(\theta_o)u_t + K(\theta_o)v_t, \\ y_t = C(\theta_o)x_t + v_t, \\ u \in \mathcal{U}, x \in \mathcal{X}, y \in \mathcal{Y}, \\ \mathcal{I}_1^N(\theta_o) \succeq \dfrac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\theta_o) \end{array} \right\}. \tag{25}$$

The objective J is the control cost, C^π, in Situation 1, and the experiment length, N, in Situation 2. The horizon considered by the controller (25) is the experiment length. Hence, the controller should find an input of length N that satisfies the experiment design constraint at the lowest cost. This results in that, after the system identification experiment, the resulting estimates fulfill the application requirement with probability α.

Alternatively, an infinite-horizon controller can be formulated for the application-oriented experiment design problem. This means that the control cost is calculated for a control horizon that is infinite. The experiment length N, on the other hand, remains finite.

In general it is intractable to use the information matrix $\mathcal{I}_1^N(\theta)$ in (25), e.g. due to that transients have to be accounted for. The average information matrix $\mathcal{I}(\theta)$ is often easier to compute (see Section 4.2.4 below for an example). In view of (23) it is common to use the approximation

$$\mathcal{I}_1^N(\theta_o) \approx N\mathcal{I}(\theta_o) \tag{26}$$

resulting in that

$$\mathcal{I}(\theta_o) \succeq \frac{\gamma\chi_\alpha^2(n_\theta)}{2N} V''_{\text{app}}(\theta_o) \tag{27}$$

replaces the last constraint in (25).

3.1. Dual properties

Explicitly including the experiment design constraint in the controller gives the controller dual properties. The minimization of the cost, ensures good control performance while the experiment design constraint ensures that the data contains enough information for good estimation. The formulation in (25) relates to the existing MPC with excitation and dual MPC formulations presented in Section 6. The significant difference is that the new formulations in this article *explicitly* take the intended application into account when the excitation is designed. The algorithms presented in Section 6, on the other hand, are primarily concerned with persistence of excitation.

3.2. Challenges

The input design problems formulated in (25) are in general difficult optimization problems. There are a number of challenges that need to be addressed. The solution may be closed-loop control and therefore the correlation between the input and any noise sources present must be accounted for. This is further complicated by the possibility that the controller is nonlinear or gives non-stationary signal distributions. When the system operates under time domain constraints on the signals, it is often observed that the resulting signal distributions are non-Gaussian even when the system is driven by Gaussian sources.

A notable exception is the quasi-stationary case with a linear output controller such that the input is given by (12). When no time domain constraints are present, the problems have essentially been solved for a number of frequency domain constraints (Hjalmarsson and Jansson, 2008; Hildebrand et al., 2010). Ways of approaching the input design problem with time domain constraints present are investigated in the following two sections. A very general Markov Decision Process approach is outlined in Section 4 while generality is traded for reduced computational complexity in Section 5 where a receding horizon approach is considered.

4. MDP approach

Markov decision processes offer a general methodology for decision making under uncertainty, such as the control problem in (5) where the noise v_t is unknown. Hordijk and Kallenberg (1984) and Altman (1999) study the problem of finding the optimal policy subject to constraints on costs of the same form as (4) and show that the optimal policy can be found by solving a linear program. In this section, the experiment design problem (25) is approached using the theory of finite, constrained MDPs. The approach works for general noise and system assumptions but is typically computationally demanding.

The section starts by defining the constrained Markov decision process needed to apply the MDP theory to the input design problem (25). The defined process is a minor extension of the constrained MDPs in Hordijk and Kallenberg (1984), by which matrix valued reward functions are introduced. The model (1) is then connected to the MDP description and the necessary components for formulating the optimal experiment design problem (5) as a constrained MPD are presented.

4.1. Matrix constrained Markov decision processes

The problem is here formulated for systems with finite state and action spaces assuming that the state of the system is available to the controller. The system is observed at discrete times $t = 1, 2, \ldots$, over an infinite horizon. The evolution of the system from one time instant to the other depends on the state of the system, the action taken by the controller and the process noise. Depending on the chosen action, the system transitions randomly from one state to another with different probabilities. It is assumed that the system has the Markov property, which means that the probability of transitioning from one state to another depends only on the current state and not on how the current state was attained.

A finite, matrix constrained Markov decision process is the tuple $\{\mathbf{X}, \mathbf{U}, \mathcal{P}, c, R\}$ where:

- $\mathbf{X}$ is a discrete state space containing a finite number of states. Generic notation for states is x, y.
- $\mathbf{U}$ is a discrete action space containing a finite number of actions. An individual action is denoted u and the set $\mathbf{U}(x) \subseteq \mathbf{U}$ are the actions available in state x.

– $\mathcal{P} = \{p_{xy}(u)\}$ are the transition probabilities from state x to state y when action u is taken, that is $p_{xy}(u) = \mathbb{P}\{x_{t+1} = y \,|x_t = x, u_t = u\}$.

– $c\colon \mathbf{X} \times \mathbf{U} \to \mathbf{R}$ is an immediate cost for state–action pair (x, u) related to the control objective, C^π.

– $R = \left\{R^k\colon \mathbf{X} \times \mathbf{U} \to \mathbf{R}^{n_k \times n_k}, k = 1, 2, \ldots, \bar{m}\right\}$ are $\bar{m}$ symmetric, matrix valued immediate rewards for state–action pair (x, u) related to the constraints. The dimensions n_k can be chosen according to the problem at hand.

It is assumed that both the immediate costs and rewards are stationary, i.e., they do not depend on time. The control u depends on the chosen policy, which can be classified according to the classes:

– Markov policies, Π_M: $\pi \in \Pi_M$ if for any time t, the rule π_t depends only on the state x_t and not on how the state was attained. Hence, the rule π_t is a mapping $\pi_t\colon \mathbf{X} \times \mathbf{U}(x) \to [0, 1]$, which gives the conditional probability of taking control action u in state x at time t, that is $\pi_t(x, u) = \mathbb{P}\{u_t = u \,|x_t = x\}$.

– Stationary policies, Π_S: $\pi \in \Pi_S$ if $\pi_1 = \pi_2 = \cdots$, that is, the rule does not depend on time. The stationary policies are a subset of Π_M.

– Stationary deterministic policies, Π_D: $\pi \in \Pi_D$ if the action u_t is a function of the state x_t. Hence, the rule π_t is a mapping $\pi_t\colon \mathbf{X} \to \mathbf{U}(x)$. The stationary deterministic policies are a subset of Π_S.

The initial distribution of the states is denoted β, that is, $\beta(x) = \mathbb{P}\{x_1 = x\}$. A given policy π, and an initial distribution β determine a unique probability measure $\mathbb{P}^\pi_\beta$ for the state and action trajectories. The corresponding expectation operator is denoted $\mathbb{E}^\pi_\beta$. Based on the immediate rewards R^k, the expected average rewards $R^{k,\pi}$ can be formulated analogously to the expected average cost C^π.

The controller has two objectives, to minimize the control cost C^π and to keep the rewards $R^{k,\pi}$ above given constraints B^k. Consequently, the control problem can be formulated as finding the optimal policy in the sense

$$\pi^\star = \arg\min_\pi \left\{C^\pi \,\middle|\, R^{k,\pi} \succeq B^k,\ k = 1, 2, \ldots, \bar{m}\right\}. \tag{28}$$

For constrained MDPs, the optimal policy depends on the initial distribution β, this dependence is however not indicated to somewhat simplify the notation.

It is seen that the experiment design problem (25) has similarities to the matrix constrained MDP (28). Firstly, the state and control constraints $\mathcal{X}$ and $\mathcal{U}$ relate to the state and action spaces $\mathbf{X}$ and $\mathbf{U}$. Secondly, the dynamics given by the state space model in (25) together with the distribution of the disturbance process v_t is an alternative description of the transition probabilities $p_{xy}(u)$. Thirdly, the excitation constraint is a matrix valued reward subject to a lower constraint.

The problem (28) can be solved by solving a semidefinite program, where the so called *state–action frequencies* (or *occupation measures*) are decision variables. The experiment design problem (25) can in turn be formulated as a matrix constrained MDP and therefore also solved using semidefinite programming. The details of arriving at the semidefinite programming formulation are given in Appendix A.

4.2. Experiment design as an MDP

Using the ideas from the previous sections and the approximation (26), an MDP solution to the experiment design problem (25) can be formulated. Find an optimal policy $\pi^\star$ such that the experiment design constraint (27) is satisfied, i.e., find

$$\pi^\star = \arg\min_\pi \left\{C^\pi \;\middle|\; \mathcal{I}(\theta_o) \succeq \frac{\gamma\chi^2_\alpha(n_\theta)}{2N} V''_{\text{app}}\right\}. \tag{29}$$

A feasible policy for (29) can be found by the following steps. To simplify notation for the constraints on the information matrix, the single output case is considered. However, extension to multiple outputs is possible.

The immediate cost and rewards are defined for the current state and actions. Since the excitation constraint is a constraint on the information matrix, the extended state space (19) is used. Therefore, the MDP is defined for the state space system

$$\xi^c_{t+1} = \mathcal{A}(\theta_o)\xi^c_t + \mathcal{B}(\theta_o)u^c_t + \mathcal{K}(\theta_o)v_t. \tag{30}$$

The superscript c is used to distinguish variables in a continuous state or action space from variables in the discrete spaces used in the MDP.

4.2.1. Discrete states and actions

The considered constrained MDP and the control problem (28) are formulated for *finite* state and action spaces while the experiment design problem in (25) is defined for more general spaces. Therefore, some discretization of the state and action spaces may be needed. This is a much studied problem where a lot of attention has been paid to the curse-of-dimensionality. Here a simple gridding technique is used.

Partition the sets $\mathcal{X}$ and $\mathcal{U}$ into discrete regions, $\mathcal{X}_x$ and $\mathcal{U}_u$, indexed by $\mathbf{X} = \{1, 2, \ldots, n_x\}$ and $\mathbf{U} = \{1, 2, \ldots, n_u\}$, such that $\mathcal{X} = \bigcup_{x=1}^{n_x} \mathcal{X}_x$ and $\mathcal{U} = \bigcup_{x=1}^{n_u} \mathcal{U}_u$. A continuous state ξ^c is classified as belonging to a certain discrete state x, for example, by

$$x = \arg\min_{x \in \mathbf{X}} \|\xi^c - \xi_x\|, \tag{31}$$

where the ξ_x are the centers of the discretization regions, and $\|\cdot\|$ is a suitable norm. The input u is classified in the same way. This allows the system to be approximated by a finite Markov chain on the discrete state space $\mathbf{X}$ and the discrete action space $\mathbf{U}$, see, for example, the work by Lunze (1998) for a discussion on the properties of the approximation.

Remark 1. In modern industrial applications, control is most often implemented digitally which means that the action space is already discrete.

4.2.2. Transition probabilities and initial distribution

In the constrained MDP description, the transition probabilities are given explicitly, while in the model (1) these probabilities depend implicitly on the properties of the stochastic innovations process, v_t. Therefore, the transition probabilities for the discretized state and action spaces need to be calculated, which can be done in many ways. One possibility is to consider the conditional probability

$$p_{xy}(u) = \mathbb{P}\left\{\xi^c_{t+1} \in \mathcal{X}_y | \xi^c_t = \xi_x, u^c_t = u_u\right\},$$

where the starting point for the transitions is considered to always be at the center of a discretization region. Assuming that the noise v_t is Gaussian distributed according to

$$v_t \sim \mathcal{N}(0, \lambda_v),$$

the state ξ^c_{t+1} is, conditionally on $\xi^c_t = \xi_x, u^c_t = u_u$, distributed as

$$\xi^c_{t+1} \sim \mathcal{N}\left(\mathcal{A}(\theta_o)\xi_x + \mathcal{B}(\theta_o)u_u, \mathcal{K}(\theta_o)\lambda_v\mathcal{K}(\theta_o)^T\right),$$

which can be used to calculate the necessary transition probabilities. This and other schemes for discretization of MDPs and calculating the corresponding transition probabilities are found in Bertsekas (1975), Chow and Tsitsiklis (1991) and Munos and Moore (2002), to name a few.

Depending on the assumptions made on the distribution of the initial state in (1), the initial distribution of the initial discrete states need to be calculated as well. This can be calculated as

$$\beta(x) = \mathbb{P}\left\{\xi^c_0 \in \mathcal{X}_x\right\}. \tag{32}$$

4.2.3. Control cost

A quadratic immediate cost based on the state of the system and the control action is used. The cost is considered to be constant over the discretization regions and given by the discretization centers, therefore the cost is chosen as

$$\begin{aligned} c(x,u) &= \mathbb{E}\left\{ x_{t+1}^T Q x_{t+1} + \Delta u_t^T R \Delta u_t + u_t^T S u_t \,|x_t = x, u_t = u \right\} \\ &= \sum_{y \in \mathbf{X}} p_{xy}(u) y^T Q y + \Delta u_t^T R \Delta u_t + u_t^T S u_t, \end{aligned} \tag{33}$$

where Q, R and S are positive semidefinite matrix weights. If the matrix $R \neq 0$, the cost (33) requires memory of the previous input, u_{t-1}. This can always be included in the MDP by adding an extra state for the past input.

4.2.4. Experiment design constraint

The experiment design constraint can be formulated using suitable immediate rewards. For the single output case, consider

$$I(x,u) = \frac{1}{\lambda_v} \mathbb{E}\left\{ \mathcal{C}(\theta_o) \xi_{t+1} \xi_{t+1}^T \mathcal{C}(\theta_o)^T \,\middle|\, x_t = x, u_t = u \right\} \tag{34}$$

$$= \frac{1}{\lambda_v} \sum_{y \in \mathbf{X}} p_{xy}(u) \mathcal{C}(\theta_o) \xi_y \xi_y^T \mathcal{C}(\theta_o)^T, \tag{35}$$

where $\mathcal{C}(\theta)$ comes from the extended state space formulation (19). The matrix $I(x,u)$ can can be interpreted as a matrix valued, immediate reward for the MDP. The average, per sample information matrix can be written as

$$\mathcal{I}(\theta_o) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\beta^\pi \left\{ I(x_t, u_t) \right\} = \sum_{\substack{x \in \mathbf{X} \\ u \in \mathbf{U}(x)}} z_{xu} I(x,u).$$

In the expression above, the variables z_{xu} are the aforementioned state–action frequencies, see Appendix Appendix A for a formal definition. It is seen that $\mathcal{I}(\theta_o)$ is in fact linear in z_{xu}, which is what allows to formulate the optimal experiment design problem as a semidefinite program.

4.2.5. MDP with experiment design

The state and action spaces, cost and rewards introduced in Sections 4.2.1–4.2.3 can be used to formulate the semidefinite program corresponding to the matrix constrained MDP (29). The solution to that SDP can then be used to find a corresponding (possibly suboptimal) stationary policy π^∞. Again, the details on how to formulate the semidefinite program and constructing the policy are given in Appendix A. The MDP with excitation is summarized in Algorithm 1.

Remark 2. Using the policy π^∞ requires knowledge of the state of the system. Often, the true state is not available and must therefore be estimated. One approach is to use the filter (2) and find the corresponding discrete state using (31) . This is simple, however, it is certainly not optimal in many cases. An alternative could be to introduce belief states for the distribution of the state estimate and based on these, calculate a distribution over the discrete states.

4.3. Discussion

The proposed MDP scheme is very general and applies to a broad class of system and disturbances as long as the predictor sensitivities and necessary transition probabilities can be found. If these are not available analytically, they can often be found using Monte Carlo techniques. On the other hand, the method suffers from the curse of dimensionality and for systems with continuous state and action spaces the

Algorithm 1 MDP with excitation

Require: $V''_{\text{app}}(\theta_o), \gamma, \alpha, N$
1: solve (29) to obtain π^∞.
2: **while** $t < N$ **do**
3: measure y_t
4: update $\hat{\xi}^c_t$ using a state estimator based on (19)
5: find discrete state x using (31)
6: find u_t from π^∞
7: apply u_t to the system
8: $t \leftarrow t + 1$
9: **end while**

discretization often leads to extremely large problem spaces. For the simple gridding technique, the problem grows exponentially in the number of states. Furthermore, since the optimization problem is a semidefinite program, smaller problems than for the linear programs in regular constrained MDPs can be handled. On the other hand, once the policy π^∞ has been found, finding the control action at a sampling instant becomes a relatively easy problem. For example, In the following section, a receding horizon approach is taken to simplify the problem and find an approximate solution to the optimal experiment design problem.

5. MPC approaches

Model predictive control (MPC) is a commonly used technique to simplify optimal control problems by using the receding horizon principle. This means that the optimization is simplified by considering a shorter horizon which reduces the number of decision variables. The drawback is that many more optimization problems need to be solved. Typically, only the first input from the solution is applied to the actual plant and the optimization is solved again in the next sampling instant. Inspired by this, MPC type formulations of the experiment design problems are proposed. The formulation easily handles much larger problems than the MDP formulation in Section 4 but is restricted to simpler disturbance characteristics. Before formulating the controller with input design, a basic MPC formulation is presented.

5.1. Model predictive control

As a starting point, the MPC formulation from Maciejowski (2002) is used. It is a receding horizon scheme with on-line optimization in each time step. The immediate cost at time t is chosen as

$$c_t(x_t, u_t) = \|y_{t+1} - r_{t+1}\|^2_{Q_t} + \|\Delta u_t\|^2_{R_t} + \|u_t\|^2_{S_t},$$

where r_{t+1} is a reference signal, $\Delta u_t \triangleq u_t - u_{t-1}$ is the control update and Q_t, R_t and S_t are tunable matrix weights. The controller minimizes the average cost over a prediction horizon N_y. For computational purposes, a control horizon $N_u \leq N_y$ is often used to reduce the number of decision variables further. This can be obtained by setting the corresponding $R_t = 0$ and $S_t = 0$ and some suitable assumption on the input for $t > N_u$. For simplicity, it is assumed that the matrix weights are constant over their respective horizons and that the inputs are constants beyond the control horizon, which are common choices in practice. This leads to the cost

$$C_t = \sum_{k=t}^{N_y-1} c_t(x_k, u_k) = \sum_{k=t}^{N_y-1} \|y_{k+1} - r_{t+k+1}\|^2_Q + \sum_{k=t}^{N_u-1} \|\Delta u_k\|^2_R + \sum_{=t}^{N_u-1} \|u_k\|^2_S.$$

If the state is not available for measurement, a common practice is to estimate the state using an observer and replace x_t by the estimate $\hat{x}_t$ and use a certainty equivalence type MPC (Maciejowski, 2002, for example).

In this case, typically the noise v_t is assumed to be $v_t = 0$ over the prediction horizon. Consequently, at time instant t, a sequence of inputs is found by solving the optimization problem

$$\begin{aligned}
\underset{\{u_k\}_{k=t}^{t+N_y-1}}{\text{minimize}} \quad & C_t \\
\text{subject to} \quad & x_{k+1} = A(\theta)x_k + B(\theta)u_k, && k = t, \ldots, t+N_y-1, \\
& y_k = C(\theta)x_k, && k = t+1, \ldots, t+N_y, \\
& u_{k+1} = u_k, && k = t+N_u-1, \ldots, t+N_y-2, \\
& u_k \in \mathcal{U}, && k = t, \ldots, t+N_y-1, \\
& x_k \in \mathcal{X}, && k = t+1, \ldots, t+N_y, \\
& y_k \in \mathcal{Y}, && k = t+1, \ldots, t+N_y.
\end{aligned} \tag{36}$$

If the state x_t is not measured, an estimate $\hat{x}_t$ is used in the state equation above. The input is constrained to be constant past the control horizon N_u. In the next time step, a measurement is taken and the procedure repeats itself at time $t+1$, in a receding horizon fashion.

The most common choice for the constraint sets are arguably box constraints for the inputs and outputs (or states). In this case, the constraints are given by

$$\mathcal{U} = \{u \,|u_{\min} \leq u \leq u_{\max}\}, \qquad \mathcal{Y} = \{y \,|y_{\min} \leq y \leq y_{\max}\}.$$

MPC has become a fairly well established control theory with formal results on stability and robustness, which are well summarized by Mayne et al. (2000).

Assumption 1. *The MPC* (36) *is feasible for each time instant and the closed loop is stable for all initial states* $x_1 \in \mathcal{X}$.

Remark 3. The recursive feasibility and stability in Assumption 1 is typically enforced using terminal constraints on the states or a terminal cost (Mayne et al., 2000). Although not explicitly included in (36), there is nothing in the formulation that restricts one from including such terms when necessary.

5.2. MPC with excitation

An MPC based approximation for the experiment design problem (25) for Situation 1, where the control cost is the main cost of the experiment, is considered here. The resulting controller will be called Model Predictive Control with eXcitation, or MPC-X (Larsson et al., 2013). First, an assumption on the noise characteristics of the system is needed.

Assumption 2. *The system is of output error type, i.e.,* $K(\theta) = 0$.

The above assumption makes the information matrix a function of the input and the noise variance only. Following the developments in Section 2.5, the experiment design constraint (18) can be added to the MPC to ensure that the resulting information matrix eventually fulfills the application specifications. A horizon, N_I for the predicted information matrix is added in the spirit of receding horizon control subject to $N_I \leq N_y$. Furthermore, a positive scaling factor, κ_t, is introduced to be able to control the excitation level at a given instant. It should hold that κ_N is at least one in order to ensure that the desire information matrix is obtained at the final time N. All in all, this gives a controller which, at time instant t, determines the input by solving the optimization problem

$$
\begin{aligned}
\underset{\{u_k\}_{k=t}^{N_y-1}}{\text{minimize}} \quad & C_t \\
\text{subject to} \quad & x_{k+1} = A(\theta_o)x_k + B(\theta_o)u_k, && k = t, \dots, N_y - 1, \\
& y_k = C(\theta_o)x_k, && k = t+1, \dots, t+N_y \\
& u_{k+1} = u_k, && t = t+N_u-1, \dots, t+N_y-2, \\
& u_k \in \mathcal{U}, && k = t, \dots, t+N_y-1, \\
& x_k \in \mathcal{X}, && k = t+1, \dots, t+N_y, \\
& y_k \in \mathcal{Y}, && k = t+1, \dots, t+N_y, \\
& \mathcal{I}_1^{t+N_I}(\theta_o) \succeq \kappa_t \frac{\gamma \chi_\alpha^2(n)}{2} V''_{\text{app}}(\theta_0).
\end{aligned}
\tag{37}
$$

In practice, the MPC-X controller in (37) cannot be implemented since it requires knowledge of the true true parameter values. Instead, one has to rely on the best available estimate of the parameters. This means that in (37), the parameters θ_o are replaced by an estimate, which will be denoted $\hat{\theta}$.

Under Assumption 2 and using the estimated parameters $\hat{\theta}$ in lieu of the true (unknown) parameters, the information matrix can be estimated by

$$
\mathcal{I}_1^{N_I}(\theta_o) \approx I_1^t(\hat{\theta}) + I_{t+1}^{t+N_I}(\hat{\theta}), \tag{38}
$$

where also the dependence of past and future data has been made explicit. We will use this approximation in (37).

Proposition 1. *The MPC with excitation in* (37) *is nonconvex. This is due to the experiment design constraint, which is quadratic in the decision variables.*

Although (37) is nonconvex, several options for convex relaxations exist. For example, the LMI relaxations of Lasserre (2000) for solving polynomial optimization problems can be used. The well-known relaxation by Shor (1987) belongs to this group and is used for input design by Manchester (2010) and by Larsson et al. (2015) for the implementation of MPC-X. Another possibility is the cyclic method by, for example, Stoica et al. (2008), which has been used in a similar setting by Ebadat et al. (2014b).

A complication due to the receding horizon principle used in MPC-X is that the addition to the information matrix is only a predicted contribution. Since only the first input from the optimal solution of (37) is implemented on the system, only the corresponding part is added to the information matrix. Therefore, it is possible that a badly tuned MPC-X controller postpones part of the information needed indefinitely. However, under certain conditions, the experiment design constraint is guaranteed to be fulfilled. This, in turn, means that the information matrix for the time segment used for identification has the desired properties. Sufficient conditions for this are given in Theorem 1, which requires the following assumption on the recursive feasibility of the optimization.

Assumption 3. *The optimization problem* (37) *is feasible at each time instant.*

Proposition 2. *A necessary condition for feasibility of* (37) *for every time instant* t *is that* $N_I \geq \operatorname{rank} V''_{app}$.

Proof. Feasibility of (37) requires that

$$
I_{t+1}^{t+N_I}(\hat{\theta}) \succeq \kappa_t \frac{\gamma \chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta}) - I_1^t(\hat{\theta}),
$$

where the right hand side has up to $\operatorname{rank} V''_{\text{app}}(\hat{\theta})$ positive eigenvalues. Since $I_{t+1}^{t+N_I}$ is the sum of N_I rank-1 matrices, it has at most N_I positive eigenvalues. Hence, $N_I \geq \operatorname{rank} V''_{\text{app}}(\hat{\theta})$ is required for feasibility for every time instance. □

Remark 4. As soon $I_1^t(\hat{\theta}) - cV''_{\text{app}}(\hat{\theta}) \geq 0$ for some $c > 0$, the condition in Proposition 2 can be relaxed.

Remark 5. The optimization problem (37) may be infeasible due to the conflicting nature of the constraints. In such a case κ_t can always be adjusted to give a feasible problem.

If κ_t is chosen such that (37) is feasible in each time instant, MPC-X renders the closed loop system stable, assuming that the MPC (36) is used for $t > N$. This follows from the fact that recursive feasibility of MPC-X ensures that $x_N \in \mathcal{X}$ and stability follows from Assumption 1. Using the above assumptions, the following results on the experiment design constraint are obtained.

Theorem 1. *Suppose that Assumptions 2 and 3 hold. If V''_{app} has full rank, κ_t is deterministic and increasing such that $\kappa_t \to \infty$, and the input is bounded, then there exists an N such that, for a fixed $\hat{\theta}$,*

$$I_1^N(\hat{\theta}) \succeq \frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta}).$$

If Assumption 3 is strengthened so that (37) is feasible for all $x \in \mathcal{X}$ at each time, then also

$$\mathcal{I}_1^N(\hat{\theta}) \succeq \frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta}).$$

PROOF. Let $\tau = \frac{\gamma\chi_\alpha^2(n_\theta)}{2}$. Assumption 3 gives

$$I_1^{t-1}(\hat{\theta}) + I_t^t(\hat{\theta}) \succeq \kappa_t \tau V''_{\text{app}}(\hat{\theta}) - I_{t+1}^{t+N_u}(\hat{\theta}), \tag{39}$$

where $I_t^t(\hat{\theta})$ is known to be added to the information matrix by Assumption 2. Take the eigenvector z corresponding to the smallest eigenvalue of the right hand side of (39) such that $\|z\| = 1$, then

$$\begin{aligned}
\lambda_{\min}\left(I_1^{t-1}(\hat{\theta}) + I_t^t(\hat{\theta})\right) &\geq z^T\left[\kappa_t \tau V''_{\text{app}}(\hat{\theta}) - I_{t+1}^{t+N_u}(\hat{\theta})\right] z \\
&\geq \kappa_t \min_{\|z\|=1} z^T \tau V''_{\text{app}}(\hat{\theta}) z - \max_{\|z\|=1} z^T I_{t+1}^{t+N_u}(\hat{\theta}) z \\
&= \kappa_t \lambda_{\min}\left(\tau V''_{\text{app}}(\hat{\theta})\right) - C,
\end{aligned}$$

where $C < \infty$ is an upper bound on the largest eigenvalue of $I_{t+1}^{t+N_u}(\hat{\theta})$, which is bounded since u_t and x_t are bounded. Hence, for N such that $\kappa_N \geq \left(\lambda_{\max}\left(\tau V''_{\text{app}}(\hat{\theta})\right) + C\right) / \lambda_{\min}\left(\tau V''_{\text{app}}(\hat{\theta})\right) < \infty$ (since $V''_{\text{app}}(\hat{\theta})$ is invertible), $\lambda_{\min}\left(I_1^N(\hat{\theta})\right) \geq \lambda_{\max}\left(\tau V''_{\text{app}}\right)$ which implies that $I_1^N(\hat{\theta}) \succeq \frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta})$.

Under the assumptions, N is deterministic. Furthermore, under the stronger assumption that (37) is feasible for all $x \in \mathcal{X}$ at each time, $I_1^N(\hat{\theta}) \succeq \frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta})$ holds for every possible trajectory. Hence, $\mathcal{I}_1^N(\hat{\theta}) = \mathbb{E}\left\{I_1^N(\hat{\theta})\right\} \succeq \mathbb{E}\left\{\frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta})\right\} = \frac{\gamma\chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}(\hat{\theta})$. □

Remark 6. The only stochastic part of the right-hand side of the excitation constraint is the parameter estimate. For a fixed parameter vector, the right-hand side is completely deterministic.

Remark 7. The bound on κ_t and the resulting N in Theorem 1 is very loose because of the hard requirement on the eigenvalues of the information matrix. In practice, the constraint is typically fulfilled for much smaller values of N.

Remark 8. The assumption that (37) is feasible for all $x \in \mathcal{X}$ at every time is fairly strong since κ_t needs to be chosen such that the problem remains feasible for every possible x for every time. The possible choices of κ_t depend on the assumptions made on the disturbances.

5.3. Minimum time MPC-X

An MPC approximation of the experiment design problem for minimum experiment time is formulated here. The controller avoids infeasibilities by optimizing the scaling factor κ_t in each sample. This can also be interpreted as the controller backing off on the excitation when required (Hjalmarsson et al., 2013).

Since the true parameters are not know, the controller is formulated using the best available estimate, as was done for MPC-X.

5.3.1. Adapting κ_t.

The algorithm aims at finding a trade-off between the identification and control performance by adapting the scaling factor, κ_t. Let $\tilde{V} = \frac{\gamma\chi^2_\alpha(n_\theta)}{2} V''_{\text{app}}(\hat{\theta})$, which can be written as $\tilde{V} = \tilde{V}^{1/2}\tilde{V}^{1/2}$ since the information matrix is symmetric and positive semidefinite. Then the experiment design constraint implies

$$I_1^t(\hat{\theta}) + I_{t+1}^{t+N_I}(\hat{\theta}) \succeq \kappa_t \frac{\gamma\chi^2_\alpha(n_\theta)}{2} V''_{\text{app}}(\hat{\theta}) \tag{40}$$

$$\Leftrightarrow \tilde{V}^{-1/2}\left[I_1^t(\hat{\theta}) + I_{t+1}^{t+N_I}(\hat{\theta})\right]\tilde{V}^{-1/2} \succeq \kappa_t I \tag{41}$$

$$\Leftrightarrow \lambda_{\min}\left\{\tilde{V}^{-1/2}\left[I_1^t(\hat{\theta}) + I_{t+1}^{t+N_I}(\hat{\theta})\right]\tilde{V}^{-1/2}\right\} \geq \kappa_t. \tag{42}$$

The inequality (42) gives an upper bound on the value of κ_t. Now κ_t can be seen as the fraction of the desired information matrix that has been achieved by the algorithm at time t. The idea of the algorithm that follows below is to maximize this fraction, i.e. the upper bound in (42).

5.3.2. Control performance requirements

Maximizing the upper bound (42) gives the highest possible excitation within the hard signal constraints. This may not be desirable since a high excitation level may result in severe degradation of the control performance as compared to using C_t in the objective function. Satisfactory control performance can be guaranteed adding the constraint

$$C_t \leq C_t^\star + \Delta C \tag{43}$$

to the problem. Here $C_t^\star$ is the the optimal value of the ordinary MPC (36) optimization at time t, and ΔC is the maximum allowed degradation of the cost.

5.3.3. Minimum time MPC-X algorithm

The above considerations result in a minimum time MPC-X method. First, the regular MPC problem (36) is solved. Second the scaling factor is maximized subject to (43), which gives the optimization

$$\begin{aligned}
\underset{\{u_k\}_{k=t}^{N_y-1}}{\text{maximize}} \quad & \lambda_{\min}\left\{\tilde{V}^{-1/2}\left[I_1^t(\hat{\theta}) + I_{t+1}^{t+N_u}(\hat{\theta})\right]\tilde{V}^{-1/2}\right\} \\
\text{subject to} \quad & x_{k+1} = A(\hat{\theta})x_k + B(\hat{\theta})u_k, && k = t, \ldots, N_y - 1, \\
& y_k = C(\hat{\theta})x_k, && k = t+1, \ldots, t+N_y, \\
& u_{k+1} = u_k, && k = t+N_u-1, \ldots, t+N_y-1, \\
& u_k \in \mathcal{U}, && k = t, \ldots, t+N_y-2, \\
& x_k \in \mathcal{X}, && k = t+1, \ldots, t+N_y, \\
& y_k \in \mathcal{Y}, && k = t+1, \ldots, t+N_y, \\
& C_t \leq C_t^\star + \Delta C.
\end{aligned} \tag{44}$$

The optimization in (44) is also nonconvex but using the same technique as for MPC-X in Section 5, a convex relaxation can be found.

The algorithm searches for the most exciting input signal that can satisfy the input and output constraints while at the same time guaranteeing that the performance of the controller does not deviate too much from the standard MPC setup. When the process is close to violating any of the constraints the algorithm reduces the scaling factor, which in turns results in less excitation in the input signal.

The minimum required time in order to satisfy the experiment design constraint, while at the same time staying in a prespecified interval around the optimal control performance, is the first time that the scaling factor becomes at least 1. This means that the experiment design constraint has been satisfied completely.

Proposition 3. *Under Assumption 1, the minimum time MPC-X algorithm is feasible in each time instant and the resulting closed loop is stable.*

PROOF. The MPC is feasible by Assumption 1 and any feasible solution to (36) is also feasible for (44), hence the minimum time MPC-X is always feasible. Feasibility of (44) ensures $x \in \mathcal{X}$ which gives stability by Assumption 1. □

6. Review of formulations with persistence of excitation

Several formulations of MPC with added excitation have been proposed earlier, some of these approaches are discussed here. All of these earlier methods differ from the solutions proposed in Sections 4 and 5 in the sense that the added excitation ensures persistence of excitation. This means that the intended use of the model is not taken into account in the same explicit way is with the methods based on the application-oriented framework. To formulate the algorithms, a definition of persistence of excitation is needed.

Definition 1. *A sequence of vectors* $\{\phi_k, k = 1, 2, \dots, \}$ *is* persistently exciting *if there exist real numbers* $\rho_1 > 0$ *and* $\rho_2 > 0$ *and integer* P *such that*

$$\rho_1 I \preceq \sum_{j=k}^{P+k-1} \phi_j \phi_j^T \preceq \rho_2 I, \quad \forall k. \tag{45}$$

For systems of the form

$$y_t = \sum_{k=1}^{n_b} b_k u_{t-k} + e_t = \theta^T \phi_t + e_t, \tag{46}$$

$$\theta = \begin{bmatrix} b_1 & \cdots & b_{n_b} \end{bmatrix}^T, \tag{47}$$

$$\phi_t = \begin{bmatrix} u_{t-1} & \cdots & u_{t-n_b} \end{bmatrix}, \tag{48}$$

persistence of excitation guarantees consistent estimation of θ (Bitmead, 1984). For systems where past outputs y_t are included in the regressors ϕ_t, sufficiently rich inputs can often guarantee persistence of excitation (see Green and Moore, 1986, for example). A very general result on the requirements on stochastic regressors (48) for consistent estimates of the parameters is thanks to Lai and Wei (1982).

Based on Definition 1, different constraints on the input can be included in MPC. The upper bound is guaranteed by the amplitude constraints of the input and is therefore removed in the following formulations. The lower bound, on the other hand, needs to be considered by the MPC and the exact formulation differs in the different approaches. The choice of the lower bound ρ_1 becomes an application dependent design choice (like the choice of γ in the application-oriented schemes) and is often a difficult one.

6.1. MPCI

Genceli and Nikolaou (1996) have proposed model predictive control with simultaneous model identification (MPCI). The MPC is modified to include the constraint

$$\sum_{k=t}^{t+N_u} \phi_k \phi_k^T \succeq (\rho_1 - \mu) I, \tag{49}$$

which forces the input to be persistently exciting over the control horizon N_u. Beyond the control horizon, MPCI constrains the input sequence to be periodic. The extra tuning parameters are ρ_1, which set the level of the excitation, and μ which is used for constraint softening, and the ideal value of $\mu = 0$. The resulting optimization problem is nonconvex, but numerical schemes for finding local solutions have been proposed.

6.2. Multiobjective MPC with identification

Aggelogiannaki and Sarimveis (2006) have formulated a multi objective MPC where the excitation is included as the top priority objective. The constraint (49) from MPCI is used but in a first step an optimal μ is found as

$$\mu^\star = \arg\min_{u,\mu} \mu \tag{50}$$

subject to the regular MPC constraints and the constraint (49). The other MPC objectives are then optimized sequentially with $\mu^\star$ imposed as a constraint. The idea is to remove the possibility of infeasible MPC and at the same time get the highest possible excitation. The solution is found using an evolutionary algorithm for the nonconvex optimization.

6.3. PE-MPC

Marafioti (2010) introduced persistently exciting model predictive control (PE-MPC) which in a way reverses the formulation of MPCI. Instead of ensuring persistence of excitation over the future control horizon, PE-MPC uses a *backward* window of length P. More formally, the constraint

$$\sum_{k=0}^{P-1} \phi_{t-k+1}\phi_{t-k+1}^T \succeq \rho_1 I, \tag{51}$$

is added to the MPC formulation at time instant t. Since

$$\sum_{k=0}^{P-1} \phi_{t-k+1}\phi_{t-k+1}^T = \sum_{k=1}^{P-1} \begin{bmatrix} u_{t-k} \\ u_{t-1-k} \\ \vdots \\ u_{t-n_b-k} \end{bmatrix} \begin{bmatrix} u_{t-k} \\ u_{t-1-k} \\ \vdots \\ u_{t-n_b-k} \end{bmatrix}^T + \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-n_b} \end{bmatrix} \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-n_b} \end{bmatrix}^T,$$

u_t is the only decision variable in the constraint (51), which therefore applies only to the next input sample. PE-MPC results in a nonconvex optimization problem but can be solved using two convex quadratic programs for certain systems. The extra tuning parameters of PE-MPC are ρ_1 and the horizon length P. Marafioti (2010) shows that PE-MPC, if feasible, admits a (possibly sub-optimal) P-periodic solution.

6.4. Dual Control by Information Maximization

Rathouský and Havlena (2011) propose adding the constraint

$$\sum_{k=t}^{t+M} \phi_k\phi_k^T \succeq \rho_1 I, \quad n_b \leq M \ll N_u, \tag{52}$$

and consider several numerical schemes for finding approximate solutions to the resulting optimization problem. Choosing ρ_1 and M is the main concern. An alternative solution, where the maximum ρ_1 is found subject to constraints on the controller performance loss, is proposed.

7. Simulation examples

In this section, simulation examples are presented to illustrate the differences between the presented methods. Simple examples are chosen to show how the shortcomings and merits of the methods. For an illustration of the practical applicability and potential of MPC-X for industrial process control, the reader is referred to Larsson et al. (2015), where the algorithm is evaluated on an industrial distillation column during normal production in the plant.

7.1. Example 1: Comparison of MDP and MPC-X

This example serves as a simple proof of concept for the presented matrix constrained MDP approach to solving the application-oriented experiment design problem. It is also shown by means of simulation how including the noise model in the design affects the expression for the information matrix and the resulting estimates.

Consider the ARX system

$$\begin{cases} x_{t+1} = -\theta_1 x_t + \theta_2 u_t - \theta_1 v_t, \\ y_t = x_t + v_t, \end{cases}$$

where v_t is Gaussian white noise with variance $\lambda_v = 1 \times 10^{-3}$, and the parameter vector $\theta = [\theta_1, \theta_2]$ is unknown and needs to be estimated. The true system is given by $\theta_0 = [0.5, 0.5]^T$, and the stationary optimal one-step ahead predictor is

$$\hat{y}_{t+1|t} = -\theta_1 y_t + \theta_2 u_t. \tag{53}$$

The state space is extended with states for the predictor sensitivities according to (19), which gives

$$\begin{aligned} \xi_{t+1} &= \begin{bmatrix} \theta_1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xi_t + \begin{bmatrix} \theta_2 \\ 0 \\ 1 \end{bmatrix} u_t + \begin{bmatrix} -\theta_1 \\ -1 \\ 0 \end{bmatrix} v_t, \\ \xi_t &= \begin{bmatrix} x_t & \frac{\partial \hat{y}_t}{\partial \theta_1} & \frac{\partial \hat{y}_t}{\partial \theta_2} \end{bmatrix}^T. \end{aligned}$$

The terms (34) needed for the excitation constraint are given by

$$\begin{aligned} I(x,u) &= \frac{1}{\lambda_v} \mathbb{E} \left\{ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xi_{t+1} \xi_{t+1}^T \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \middle| x_t = x, u_t = u \right\} \\ &= \frac{1}{\lambda_v} \mathbb{E} \left\{ \begin{bmatrix} (-x_t - v_t)^2 & (-x_t - v_t)u_t \\ (-x_t - v_t)u_t & u_t^2 \end{bmatrix} \middle| x_t = x, u_t = u \right\} \\ &= \frac{1}{\lambda_v} \begin{bmatrix} x^2 + \lambda_v & -xu \\ -xu & u^2 \end{bmatrix}. \end{aligned}$$

Therefore, no additional states are actually needed to include the excitation constraint in the controller.

The state space is discretized into 51 regions, uniformly spaced on the interval $[-1, 1]$ and the input into 21 regions, uniformly spaced on the interval $[-2, 2]$. The transition probabilities are calculated numerically by evaluating the transitions for each possible pair (x, u) in simulations. A quadratic control cost based on the immediate costs (33) is used with $Q = 2$, $R = 0$ and $S = 1$.

The identification objective is setup such that, after an experiment length of $N = 500$ samples, the average per-sample information matrix should satisfy

$$\mathcal{I}(\theta_o) \succeq \begin{bmatrix} 400 & -300 \\ -300 & 800 \end{bmatrix}. \tag{54}$$

The resulting semidefinite program is solved and a stationary policy π_X is found. For comparison, an MPC-X controller that assumes that the system is of output error type is used. MPC-X is configured with the same weights, and the horizons $N_y = N_u = 5$. ARX models are estimated using the predictor (53) for the two methods in a Monte Carlo study with 100 trials to evaluate the performance of the two methods. The 99 % confidence ellipses corresponding to the information matrices calculated from the two methods and the estimates are plotted in Figure 2. The results indicate that the performance specifications are met for both methods since the calculated confidence ellipses are inside the performance specification ellipsoid. However, with the output error model, the calculated ellipse is not the true one. This results in 15 of the estimates not fulfilling the requirements when the output error assumption is used in the experiment design, while only 1 of the estimates from the matrix constrained MDP approach does not satisfy the requirements.

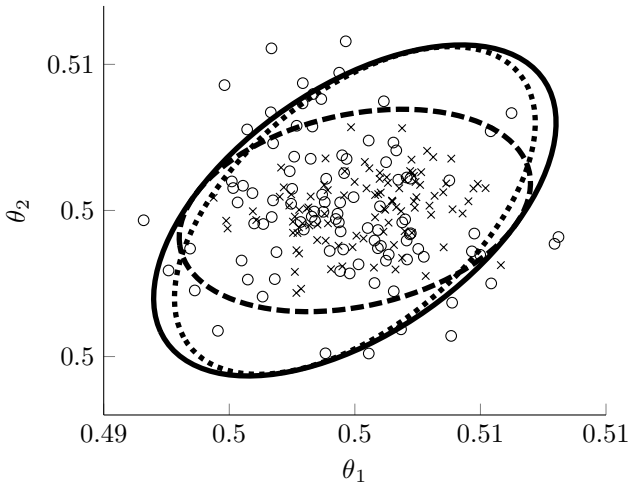

Figure 2: The confidence ellipses and estimates using the MDP approach and MPC-X for the system in Example 1. The performance specification is given by the solid ellipse. Using MDP results in the dashed ellipse and the estimates shown as crosses. MPC-X calculates the dotted confidence ellipse, which results in the estimates shown as circles. The ellipses are scaled to contain 99 % of the estimated parameters.

7.2. Example 2: Comparison of MPC based schemes

This example illustrates the differences between the two MPC based schemes in Section 5 and the persistence of excitation based schemes in Section 6 on the simulation example used by Larsson et al. (2013). MPC-X and MPC-X with back-off are implemented in Matlab using `cvx` by Grant and Boyd (2011) and the resulting SDP is solved using SeDuMi by Sturm (1999). Of the other algorithms, only PE-MPC by Marafioti (2010) is implemented. The two resulting QPs are solved using `quadprog` from the Matlab optimization toolbox.

Consider a system consisting of two interconnected tanks. An upper tank is connected to a pump with input u_t. The tank has a hole in the bottom with free flow into a lower tank, which also has a hole with free flow out of the tank. The level in the lower tank is the output, y_t, which can be measured. The system is modeled using the discrete-time, output-error model

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} \theta_3 & \theta_4 \\ 1 & 0 \end{bmatrix} x_t + \begin{bmatrix} 4.5 \\ 0 \end{bmatrix} u_t \\ y_t &= \begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix} x_t + v_t. \end{aligned} \tag{55}$$

The true system is given by the parameter values $\theta_0 = \begin{bmatrix} 0.12 & 0.059 & 0.74 & -0.14 \end{bmatrix}^T$ and the noise variance 0.01. The goal is to control the level in the lower tank using MPC with the following settings: $N_y = N_u = 5$, $Q_y = 10$, $Q_u = 1$. The considered scenario is such that the identification is started at steady state conditions of the plant. The input is constrained to be between -1 and 1.

The true parameter values are used for the initial model. This choice is motivated by the fact that the example serves to illustrate the difference between different methods, and not the effects of using estimates instead of the true parameter values. In a practical application, one would instead have to use estimated parameters, e.g. available from a commissioning identification. The application cost is chosen as

$$V_{\text{app}}(\theta) = \sum_{t=1}^{T} \|y_t(\theta_o) - y_t(\theta)\|_2^2, \tag{56}$$

over a step response of the system with MPC running. Hence, it is desired that, when the identified model is used in MPC, the step response is close to what it had been if the true system parameters were available. The desired accuracy is set to $\gamma = 100$ and should be achieved with 99 % probability. The experiments are

Table 1: Results for the algorithms used in Example 1. The value c_t is the average control cost per sample and $\sum c_t$ is the total cost over the whole experiment length.

Algorithm	var u	var y	N	c_t	$\sum c_t$
MPC	0.036	0.016	—	0.22	—
MPC-X	0.164	0.074	200	0.90	180
PE-MPC, $\rho_1 = 0.5$	0.175	0.120	211	1.36	287
MPC-X min time	0.203	0.146	82	1.66	138
PE-MPC, $\rho_1 = 1$	0.246	0.184	162	2.07	335

set up so that they run until the achieved information matrix fulfills the experiment design constraint. Two different comparisons are made.

First, the benefit of the experiment design constraint (18) over the persistence of excitation constraint is investigated in experiments of the same length. MPC-X is configured with experiment length $N = 200$ and linear scaling, $\kappa_t = \frac{t}{N}$. The PE-MPC is configured with $\rho_1 = 0.5$ and backward horizon $P = 5$, which gives a similar evolution of the smallest eigenvalue of the information matrix as with MPC-X and therefore similar experiment length.

Second, the benefit of the experiment design constraint (18) in a minimum time experiment is investigated. The PE-MPC is configured with $\rho_1 = 1$, which is the highest value that gives a feasible solution to the optimization problem. The minimum time MPC-X is configured with $\Delta C = 5$, which gives a similar evolution of the information matrix. For comparison, regular MPC is also used and the resulting information matrix is calculated.

The results of the simulations are summarized in Table 1. First, it is seen that PE-MPC with $\rho_1 = 0.5$ satisfies the experiment design constraint (18) after 211 samples which is approximately the same as the 200 samples required for MPC-X. The input variances are comparable but the resulting output variance using PE-MPC is 0.12 compared to 0.074 for MPC-X, which is also reflected in the higher control cost for PE-MPC, 287, compared to 180 for MPC-X. Second, PE-MPC with $\rho_1 = 1$ requires 162 samples to fulfill the experiment design constraint (18) compared to 82 samples for the minimum time MPC-X. This is despite the fact the PE-MPC results in higher signal variances and control cost (per sample and total). For the chosen application, MPC-X results in cheaper experiments in both cases. On the other hand, a clear benefit of PE-MPC is a much lower computation demand. The two convex QPs typically require much less time and computations to solve than the SDPs of the two MPC-X formulations.

The evolution of the smallest eigenvalue of $I_1^t - \frac{\gamma \chi_\alpha^2(n_\theta)}{2} V''_{\text{app}}$ is shown in Figure 3. The algorithms terminate when the eigenvalue becomes positive. It is clearly seen that the information available in the signals from the regular MPC is not sufficient for fulfilling the application requirements since the smallest eigenvalue is far from being positive. Typical input and output signals are shown in Figure 5. It is seen that all the formulations that add excitation in the closed loop give signals with larger variance than regular MPC.

8. Conclusions

Optimal control algorithms with dual properties have been presented. The added excitation needed for successful identification is generated in such a way that model properties that are important for the intended application are modeled well. This goal differs from the goal of previously presented algorithms where persistence of excitation is imposed on the input.

Computationally tractable approximations of the optimal controllers have been introduced. An MDP based solution to the optimal control problem is formulated. This controller can handle very general noise structures as long as the necessary expressions for the information matrix calculations can be included in an extended state space of the system to be controlled. However, the formulation relies on the finite state

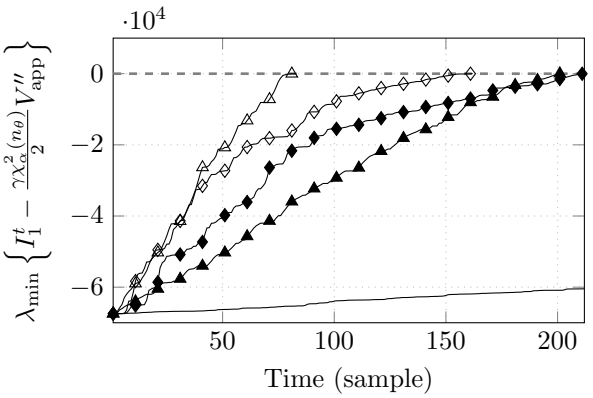

Figure 3: Evolution of the smallest eigenvalue of the difference between the information matrix and the application cost Hessian from simulation of regular MPC (——), MPC-X (—▲—), PE-MPC with $\rho_1 = 0.5$ (—♦—) and $\rho_1 = 1$ (—◇—), and minimum time MPC-X (—△—) on the system in Example 1. The algorithms terminate when the eigenvalue becomes positive.

and action space MDP framework and therefore suffers from the curse of dimensionality when continuous spaces are discretized. Better discretization schemes may alleviate some of these issues. Furthermore, it may be possible to use the techniques from MDPs over infinite spaces at the price of increased theoretical complexity.

Two algorithms based on standard model predictive control have been developed. These algorithms rely on the assumption that the system is of output error type which can be problematic in a closed loop setting. Future research should be focused on extending these formulations to more general noise models. It is believed that this can be done by leveraging on the recent results in stochastic MPC.

Finally, the presented formulations have shown in simulation to give the desired excitation properties. The resulting signals give more relevant excitation for the intended model application and disrupt normal operations to a lesser degree than if persistence of excitation is imposed. However, formal proofs of this are lacking. This should be investigated further.

9. Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 257059, The 'Autoprofit' project (www.fp7-autoprofit.eu) and the Linnaeus Center ACCESS at KTH.

References

Aggelogiannaki, E., Sarimveis, H., 2006. Multiobjective constrained MPC with simultaneous closed-loop identification. International Journal of Adaptive Control and Signal Processing 20 (4), 145–173.

Agüero, J. C., Goodwin, G. C., 2007. Choosing between open- and closed-loop experiments in linear system identification. IEEE Transactions on Automatic Control 52 (8), 1475–1480.
URL `http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4287129`

Altman, E., 1999. Constrained Markov Decision Processes. Chapman and Hall/CRC.

Åström, K. J., 1970. Introduction to Stochastic Control Theory. Academic Press, New York.

Bertsekas, D., Jun 1975. Convergence of discretization procedures in dynamic programming. IEEE Transactions on Automatic Control 20 (3), 415–419.

Bitmead, R., Mar 1984. Persistence of excitation conditions and the convergence of adaptive schemes. IEEE Transactions on Information Theory 30 (2), 183–191.

Bombois, X., Gevers, M., Scorletti, G., 2005. Open-loop versus closed-loop identification of Box–Jenkins models: a new variance analysis. In: Proceedings of the 44th IEEE Conference on Decision and Control. Sevilla, Spain, pp. 3117–3122.

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

Chow, C., Tsitsiklis, J., Aug 1991. An optimal one-way multigrid algorithm for discrete-time stochastic control. IEEE Transactions on Automatic Control 36 (8), 898–914.

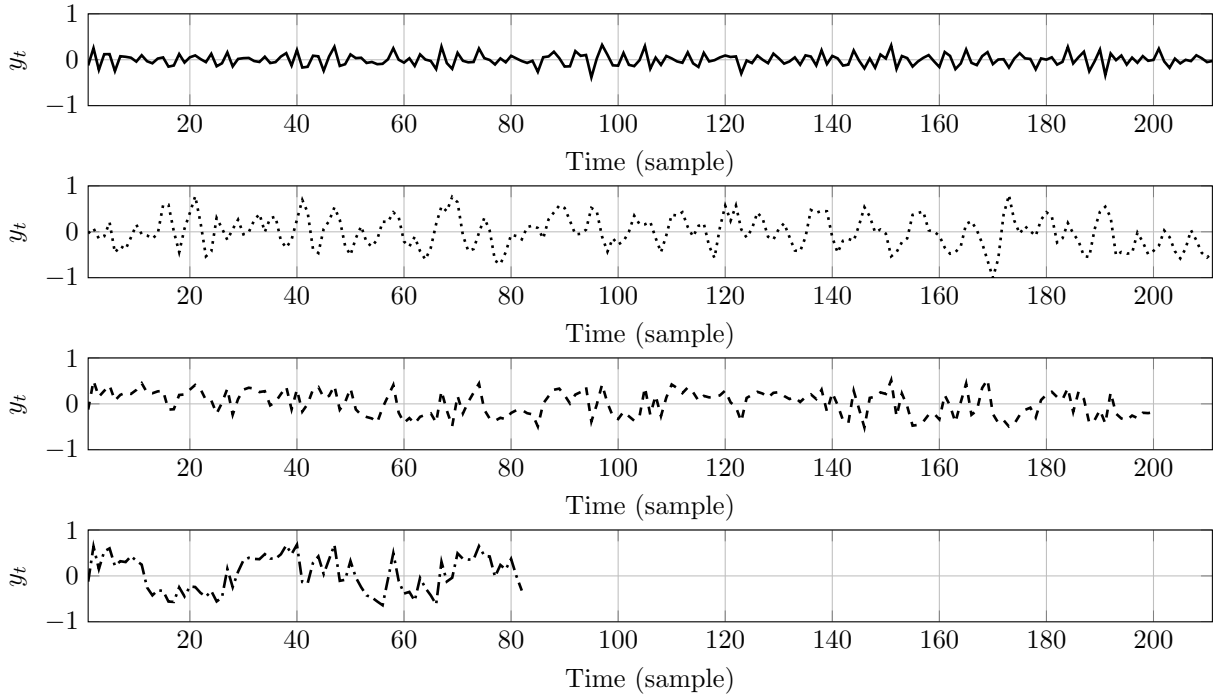

Figure 4: Typical output signals from simulation of regular MPC (——),PE-MPC (·····) with $\rho_1 = 1$, MPC-X (- - -), and minimum time MPC-X (–·–) algorithms on the system in Example 1.

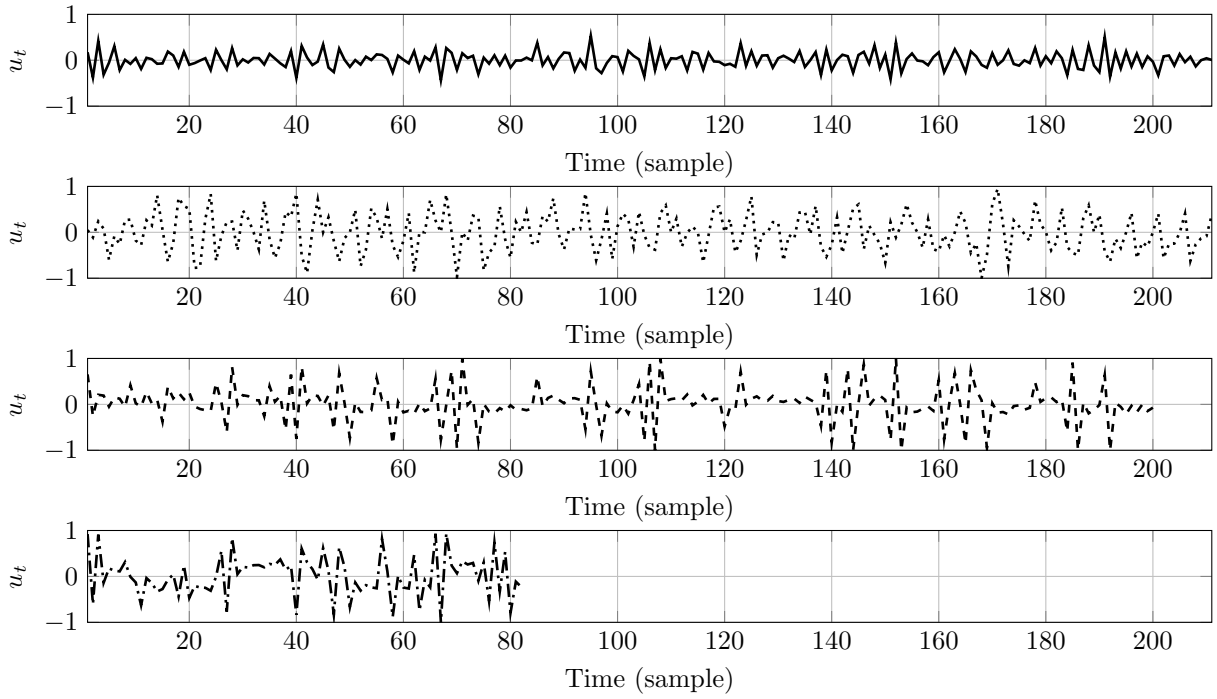

Figure 5: Typical input signals from simulation of regular MPC (——),PE-MPC (·····) with $\rho_1 = 1$, MPC-X (- - -), and minimum time MPC-X (–·–) algorithms on the system in Example 1.

Ebadat, A., Annergren, M., Larsson, C. A., Rojas, C. R., Wahlberg, B., Hjalmarsson, H., Molander, M., Sjöberg, J., 2014a. Application set approximation in optimal input design for model predictive control. In: Proceedings of the 2013 European Control Conference. Strasbourg, France.

Ebadat, A., Wahlberg, B., Hjalmarsson, H., Rojas, C. R., Hägg, P., Larsson, C. A., 2014b. Applications oriented input design in time-domain through cyclic methods. In: Proceedings of the 19th IFAC World Congress. Cape Town, South Africa.

Fedorov, V. V., 1972. Theory of Optimal Experiments. Academic Press, New York.

Feldbaum, A. A., 1960–61. Dual control theory. I–IV. Automation and Remote Control 21,22.

Forssell, U., Ljung, L., 1999. Closed-loop identification revisited. Automatica 35 (7), 1215–1241.

Genceli, H., Nikolaou, M., 1996. New approach to constrained predictive control with simultaneous model identification. AIChE Journal 42 (10), 2857–2868.

Gerencsér, L., Hjalmarsson, H., Mårtensson, J., March 2009. Identification of ARX systems with non-stationary inputs - asymptotic analysis with application to adaptive input design. Automatica 45 (3), 623–633.

Gevers, M., 1991. Connecting Identification and Robust Control: A New Challenge. In: Proc. IFAC/IFORS Symposium on Identification and System Parameter Estimation. Budapest, Hungary.

Gevers, M., Ljung, L., 1986. Optimal experiment designs with respect to the intended model application. Automatica 22 (5), 543–554.

Goodwin, G. C., Payne, R. L., 1977. Dynamic System Identification: Experiment Design and Data Analysis. Academic Press, New York.

Grant, M., Boyd, S., Feb. 2011. CVX: Matlab Software for Disciplined Convex Programming, version 1.21. http://cvxr.com/cvx.

Green, M., Moore, J. B., 1986. Persistence of excitation in linear systems. Systems & Control Letters 7 (5), 351–360.

Harris, T. J., 1989. Assessment of control loop performance. The Canadian Journal of Chemical Engineering 67 (5), 856–861.
URL http://dx.doi.org/10.1002/cjce.5450670519

Hildebrand, R., Gevers, M., Dec 2013. Central extensions in closed-loop optimal experiment design. In: Proceedings of the 52nd IEEE Conference on Decision and Control. IEEE, Florence, Italy.
URL http://dx.doi.org/10.1109/CDC.2013.6760807

Hildebrand, R., Gevers, M., Solari, G., Dec 2010. Closed-loop optimal experiment design: The partial correlation approach. In: Proceedings of the 49th IEEE Conference on Decision and Control. Atlanta, GA.

Hjalmarsson, H., 2009. System identification of complex and structured systems. European Journal of Control 15 (4), 275–310.

Hjalmarsson, H., Gevers, M., De Bruyne, F., December 1996. For model based control design criteria, closed loop identification gives better performance. Automatica 32, 1659–1673.

Hjalmarsson, H., Jansson, H., Mar 2008. Closed loop experiment design for linear time invariant dynamical systems via LMIs. Automatica 44 (3), 623–636.
URL http://dx.doi.org/10.1016/j.automatica.2007.06.022

Hjalmarsson, H., Larsson, C. A., Hägg, P., Ebadat, A., August 2013. Deliverable 3.3 — Novel algorithms for productivity preserving testing. Autoprofit project.

Hordijk, A., Kallenberg, L. C. M., 1984. Constrained undiscounted stochastic dynamic programming. Mathematics of Operations Research 9 (2), 276–289.

Jansson, H., Hjalmarsson, H., 2005. Input Design via LMIs Admitting Frequency-wise Model Specifications in Confidence Regions. IEEE Transactions on Automatic Control 50 (10), 1534–1549.

Lai, T., Wei, C., 1982. Least-squares estimates in stochastic regression-models with applications to identification and control of dynamic systems. The Annals of Statistics 10 (1), 154–166.

Larsson, C. A., Annergren, M., Hjalmarsson, H., Rojas, C. R., Bombois, X., Mesbah, A., Modén, P. E., 2013. Model predictive control with integrated experiment design for output error systems. In: Proceedings of the 2013 European Control Conference. Zürich, Switzerland.

Larsson, C. A., Rojas, C. R., Bombois, X., Hjalmarsson, H., 2015. Experimental evaluation of model predictive control with excitation (MPC-X) on an industrial depropanizer. Journal of Process Control 31 (0), 1 – 16.
URL http://www.sciencedirect.com/science/article/pii/S0959152415000578

Lasserre, J. B., Mar. 2000. Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization 11 (3), 796–817.

Ljung, L., 1999. System Identification: Theory for the User, 2nd Edition. Prentice Hall, Upper Saddle River, New Jersey.

Ljung, L., Caines, P., 1979. Asymptotic normality of prediction error estimation for approximate system models. Stochastics 3 (1), 29–46.

Ljung, L., Söderström, T., 1983. Theory and Practice of Recursive Identification. Signal Processing, Optimization and Control. The MIT Press, Cambridge, MA.

Lunze, J., 1998. On the markov property of quantised state measuirement sequences. Automatica 34 (11), 1439–1444.
URL http://www.sciencedirect.com/science/article/pii/S0005109898000995

Maciejowski, J. M., 2002. Predictive Control with Constraints. Prentice Hall, Edinburgh Gate, Harlow, Essex, England.

Manchester, I. R., Sep. 2010. Input Design for System Identification via Convex Relaxation. ArXiv e-prints.

Marafioti, G., 2010. Enhanced model predictive control: Dual control approach and state estimation issues. Ph.D. thesis, Norwegian University of Science and Technology, Department of Engineering Cybernetics.

Mayne, D., Rawlings, J., Rao, C., Scokaert, P., 2000. Constrained model predictive control: Stability and optimality. Automatica 36 (6), 789–814.

Mehra, R. K., December 1974. Optimal Input Signals for Parameter Estimation in Dynamic Systems — Survey and New Results. IEEE Transactions on Automatic Control 19(6), 753–768.

Modén, P. E., Lundh, M., 2013. Performance monitoring for model predictive control maintenance. In: Proceedings of the 2013 European Control Conference.
Munos, R., Moore, A., Nov. 2002. Variable resolution discretization in optimal control. Machine Learning 49 (2-3), 291–323. URL http://dx.doi.org/10.1023/A:1017992615625
Puterman, M. L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming, 1st Edition. John Wiley & Sons, Inc., New York, NY, USA.
Rathouský, J., Havlena, V., 2011. MPC-based approximation of dual control by information maximization. In: Proceedings of the International Conference on Process Control. Tatranská Lomnica, Slovakia, pp. 247–252.
Shor, N. Z., 1987. Quadratic optimization problems. Soviet Journal of Circuits and Systems Sciences 25 (6), 1–11.
Söderström, T., 2002. Discrete-Time Stochastic Systems: Estimation and Control. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
Stoica, P., Li, J., Zhu, X., Jun 2008. Waveform synthesis for diversity-based transmit beampattern design. IEEE Transactions on Signal Processing 56 (6), 2593–2598. URL http://dx.doi.org/10.1109/TSP.2007.916139
Sturm, J. F., 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones.
Zagrobelny, M., Ji, L., Rawlings, J. B., Dec 2013. Quis custodiet ipsos custodes? Annual Reviews in Control 37 (2), 260–270.

Appendix A. Constrained Markov decision processes

The connection between the constrained control problem

$$\min_{\pi} \left\{ C^{\pi} \,\middle|\, R^{k,\pi} \succeq B^k,\ k = 1, 2, \ldots, \bar{m} \right\}. \tag{A.1}$$

for a constrained MDP, $\{\mathbf{X}, \mathbf{U}, \mathcal{P}, c, R\}$, introduced in Section 4, and semidefinite programs is presented here. The development parallels the development of the linear programs obtained for constrained MPDs with scalar rewards as presented in, for example Hordijk and Kallenberg (1984). The difference is that the matrix valued rewards lead to semidefinite programs instead. The key to obtaining the semidefinite program are the state–action frequencies.

First, the control cost, C^{π}, and the rewards, $R^{k,\pi}$, are expressed explicitly in terms of the probabilities of state–action pairs and the initial state distribution. Let $\mathbb{P}^{\pi}_{x_1}\{x_t = x, u_t = u\}$ be the probability of the state–action pair (x, u) given initial state x_1 and policy π. Then the cost (4) is

$$C^{\pi} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{x_1 \in \mathbf{X}} \sum_{\substack{x \in \mathbf{X} \\ u \in \mathbf{U}(x)}} \beta(x_1) \mathbb{P}^{\pi}_{x_1}\{x_t = x, u_t = u\}\, c(x, u).$$

The matrix valued expected average rewards are defined analogously to the cost C^{π} as

$$R^{k,\pi} \triangleq \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{x_1 \in \mathbf{X}} \sum_{\substack{x \in \mathbf{X} \\ u \in \mathbf{U}(x)}} \beta(x_1) \mathbb{P}^{\pi}_{x_1}\{x_t = x, u_t = u\}\, R^k(x, u),$$

Now, for a given policy π, introduce the expected state–action frequencies

$$z_{xu,T}(\pi) \triangleq \frac{1}{T} \sum_{t=1}^{T} \sum_{x_1 \in \mathbf{X}} \beta(x_1) \mathbb{P}^{\pi}_{x_1}\{x_t = x, u_t = u\}.$$

The sets $\{z_{xu,T}(\pi)\}_{(x\in\mathbf{X}, u\in\mathbf{U}(x))}$ can be seen as probability measures (sometimes denoted occupation measures), which give the probability $z_{xu,T}(\pi)$ to the state–action pair (x, u).

Let Z_{π} denote the limit points, $z(\pi) = \{z_{xu}(\pi)\}$, of the vectors $z_T(\pi) = \{z_{xu,T}(\pi),\ T = 1, 2, \ldots\}$ as $T \to \infty$. The elements of Z_{π} are the infinite-horizon probability measures assigning probabilities for the state–action pairs. In general, for a given initial distribution and policy, there may be an infinite set of limit points. Therefore, the notion of a convergent policy is useful.

Definition 2 (Convergent policy). *A policy π is* convergent *if Z_{π} consists of a single element. The class of convergent policies is denoted Π_C.*

The set Z_π depends on which class the policy belongs to. Introduce the sets L, L_M, L_C, L_S, and L_D defined by

$$\begin{aligned} L &\triangleq \{z(\pi) \in Z_\pi \mid \pi \text{ is an arbitrary policy.}\}, \\ L_M &\triangleq \{z(\pi) \in Z_\pi \mid \pi \in \Pi_M\}, \\ L_C &\triangleq \{z(\pi) \in Z_\pi \mid \pi \in \Pi_C\}, \\ L_S &\triangleq \{z(\pi) \in Z_\pi \mid \pi \in \Pi_S\}, \\ L_D &\triangleq \{z(\pi) \in Z_\pi \mid \pi \in \Pi_D\}. \end{aligned}$$

Theorem 2. *$L = L(M) = L(C) = \overline{L(S)} = \overline{L(D)}$, where $\overline{X}$ is the convex hull of X.*

PROOF. See, for example, Hordijk and Kallenberg (1984).

Theorem 2 shows that, since the limit points for arbitrary policies have a corresponding convergent policy, there is no restriction in considering only convergent policies. Furthermore, the theorem shows that, in general, the optimal policy for the infinite-horizon cost situation may be non-stationary.

For convergent policies, the cost C^π and the constraints $R^{k,\pi}$ can be written in terms of the state–action frequencies as

$$C^\pi = \sum_{\substack{x \in \mathbf{X} \\ u \in \mathbf{U}(x)}} c(x,u) z_{xu}, \qquad\qquad R^{k,\pi} = \sum_{\substack{x \in \mathbf{X} \\ u \in \mathbf{U}(x)}} R^k(x,u) z_{xu},$$

which then have the natural interpretation of long run average costs or rewards. Note that the cost and rewards are linear in the state–action frequencies.

Finally, the optimization problem (28) can be solved using a semidefinite program. To this end, introduce the set Q of vectors $z = [z_{xu}^T, \tilde{z}_{xu}^T]^T$ defined by

$$Q = \left\{ z \;\middle|\; \begin{array}{c} \displaystyle\sum_{(x,u)} \{\delta_{xy} - p_{xy}(u)\} z_{xu} = 0, \; y \in \mathbf{X} \\ \displaystyle\sum_u z_{yu} + \sum_{(x,u)} \{\delta_{xy} - p_{xy}(u)\} \tilde{z}_{xu} = \beta(y), y \in \mathbf{X} \\ z_{xu}, \tilde{z}_{xu} \geq 0, \; (x,u) \in \mathbf{X} \times \mathbf{U}(x) \end{array} \right\},$$

where $\delta_{xy} \triangleq \{1, \text{if } x = y, \text{ 0 otherwise}\}$. The set Q is a polytope and can be related to the set of limit points of the state–action frequencies L by the following theorem.

Theorem 3. *L=Q.*

PROOF. See, for example, Hordijk and Kallenberg (1984). □

Theorem 3 shows that Q corresponds to the probability measures of the state–action pairs for the MDP. Therefore, it is possible to optimize over the state–action frequencies by formulating a semidefinite program with decision variables constrained to lie in Q.

Consider the semidefinite program

$$\min_{(z,\tilde{z})} \left\{ \sum_{(x,u)} c(x,u) z_{xu} \;\middle|\; \begin{array}{c} z \in Q \\ (x,u) \in \mathbf{X} \times \mathbf{U}(x) \\ \displaystyle\sum_{(x,u)} R^k(x,u) z_{xu} \succeq B^k, \\ k = 1, \ldots, \bar{m} \end{array} \right\} \tag{A.2}$$

The relationship between the matrix constrained MDP (A.1) and the semidefinite program (A.2) is formalized in the following theorem.

Theorem 4. *The following facts hold:*

1. *The semidefinite program* (A.2) *is feasible if and only if the problem* (A.1) *is feasible.*
2. *The optimal values of* (A.2) *and* (A.1) *are the same.*
3. *If π is an optimal policy for* (A.1)*, then $z(\pi)$ is optimal for* (A.2)*.*

PROOF. The proof is analogous to the proof for the linear programming case (see for example Hordijk and Kallenberg (1984)) and is outlined here. By Theorems 2 and 3, $Q = L_C$. For any convergent policy, $C^\pi = \sum_{(x,u)} c(x,u) z_{xu}$ and $R^{k,\pi} = \sum_{(x,u)} R^k(x,u) z_{xu}$. From this, (1), (2) and (3) follow. □

From the solution of (A.2), define a stationary policy π^∞ by

$$\pi^\infty_{xu} = \begin{cases} z_{xu}/z_x, & \text{if } z_x > 0, \\ \tilde{z}_{xu}/\tilde{z}_x, & \text{if } z_x = 0, \tilde{z}_x > 0, \\ \text{arbitrary}, & \text{otherwise} \end{cases} \tag{A.3}$$

$$z_x = \sum_{u \in \mathbf{U}(x)} z_{xu}, \qquad \tilde{z}_x = \sum_{u \in \mathbf{U}(x)} \tilde{z}_{xu}. \tag{A.4}$$

The stationary policy π^∞ may not be optimal for (A.1) but optimality can be checked by the following lemma.

Lemma 5. *If $z_{xu}/z_x = \tilde{z}_{xu}/\tilde{z}_x$ for all u and $x \in \{x \,:\, z_x > 0, \tilde{z}_x > 0\}$, then the stationary policy π^∞ is an optimal policy for* (28)*.*

PROOF. See, for example, Hordijk and Kallenberg (1984). □

Remark 9. It is possible to prove that if $(z, \tilde{z})$ is an optimal solution of (A.2), a corresponding optimal Markov (but not necessarily stationary) policy for (A.1) exists. However, in practice, stationary policies are simpler to work with and therefore only such policies are considered.

Appendix B. Extended state space matrices

Consider the derivatives of the predictor (2) given by

$$\frac{\partial \hat{x}_{t+1|t}}{\partial \theta_i} = \frac{\partial (A - KC)}{\partial \theta_i} \hat{x}_{t|t-1} + (A - KC) \frac{\partial \hat{x}_{t|t-1}}{\partial \theta_i} \tag{B.1}$$

$$+ \frac{\partial B}{\partial \theta_i} u_t + \frac{\partial K}{\partial \theta_i} C x_t + \frac{\partial K}{\partial \theta_i} e_t, \tag{B.2}$$

$$\frac{\partial \hat{y}_{t|t-1}}{\partial \theta_i} = \frac{\partial C}{\partial \theta_i} \hat{x}_{t|t-1} + C \frac{\partial \hat{x}_{t|t-1}}{\partial \theta_i}. \tag{B.3}$$

The matrices needed for (19) are given by

$$\mathcal{A} \triangleq \begin{bmatrix} A & 0 & 0 & 0 & 0 \\ KC & A-KC & 0 & 0 & 0 \\ \frac{\partial K}{\partial \theta_1} C & \frac{\partial (A-KC)}{\partial \theta_1} & A-KC & 0 & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ \frac{\partial K}{\partial \theta_{n_\theta}} C & \frac{\partial (A-KC)}{\partial \theta_{n_\theta}} & 0 & 0 & A-KC \end{bmatrix},$$

$$\mathcal{B} \triangleq \begin{bmatrix} B \\ B \\ \frac{\partial B}{\partial \theta_1} \\ \vdots \\ \frac{\partial B}{\partial \theta_{n_\theta}} \end{bmatrix}, \qquad \mathcal{K} \triangleq \begin{bmatrix} K \\ K \\ \frac{\partial K}{\partial \theta_1} \\ \vdots \\ \frac{\partial K}{\partial \theta_{n_\theta}} \end{bmatrix},$$

$$\mathcal{C} \triangleq \begin{bmatrix} 0 & \frac{\partial C}{\partial \theta_1} & C & 0 & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \frac{\partial C}{\partial \theta_{n_\theta}} & 0 & 0 & C \end{bmatrix}.$$